

Exact Results Regarding the Physics of Complex Systems
via
Linear Algebra, Hidden Markov Models, and Information Theory

By

PAUL MICHAEL RIECHERS
B.S. Applied Physics (UC Davis) 2009
M.S. Electrical and Computer Engineering (UC Davis) 2012

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Physics

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

James P. Crutchfield, Chair

John B. Rundle

Daniel L. Cox

Committee in Charge

2016

Copyright © 2016 by
Paul Michael Riechers
All rights reserved.

*To my mother, Carla, for introducing me to nature;
and to my father, Peter, for introducing me to science.*

CONTENTS

Abstract	xi
Acknowledgments	xiii
1 Complex Systems, Impoverished Observables, and HMMs	1
1.1 Probability, Stochastic Dynamics, and HMMs	3
1.1.1 Probability	3
1.1.2 Shannon Entropy	4
1.1.3 Stochastic Processes	6
1.1.4 Markovity and Markov order	8
1.1.5 HMMs in general	10
1.1.6 Welcome to the jungle: Zoology of HMMs	10
1.1.7 Probabilities, conditional probabilities, and linear algebra	11
1.1.8 Beyond normal: Probabilities invoke non-normal linear operators	13
2 Beyond the Spectral Theorem	14
2.1 Chapter Overview	14
2.2 Introduction	16
2.2.1 Spectral Decomposition	17
2.2.2 Synopsis	19
2.3 Spectral Primer	20
2.4 Functional Calculi	23
2.4.1 Taylor series	23
2.4.2 Holomorphic functional calculus	24
2.4.3 Meromorphic functional calculus	25
2.5 Meromorphic Spectral Decomposition	26
2.5.1 Partial fraction expansion of the resolvent	26
2.5.2 Decomposing the identity	28
2.5.3 Dunford decomposition, decomposed	28

2.5.4	The resolvent, resolved	29
2.5.5	Meromorphic functional calculus	30
2.5.6	Evaluating the residues	31
2.5.7	Decomposing A^L	31
2.5.8	Drazin inverse	33
2.5.9	Consequences and generalizations	36
2.6	Constructing Decompositions	37
2.6.1	Projection operators of index-one eigenvalues	38
2.6.2	Eigenvectors, generalized eigenvectors, and projection operators	39
2.7	Examples and Applications	50
2.7.1	Spectra of stochastic transition operators	51
2.7.2	Randomness and memory in correlated processes	52
2.7.3	Poisson point processes	53
2.7.4	Stochastic thermodynamics	57
2.7.5	Power spectra	61
2.7.6	Operators for chaotic dynamics	69
2.8	Conclusion	71
3	Cracking the Nondiagonalizable Metadynamics of Prediction	74
3.1	Chapter Overview	74
3.2	Introduction	75
3.3	Structured Processes and their Complexities	79
3.3.1	Directly observable organization	80
3.3.2	Intrinsic predictability	81
3.3.3	The overhead of prediction	82
3.3.4	Generative complexities	83
3.4	Mealy Hidden Markov Models	85
3.4.1	Unifilar HMMs	87
3.4.2	Minimal Unifilar HMMs	87
3.4.3	The Finitary Stochastic Process Hierarchy	87

3.4.4	Continuous-time HMMs	88
3.5	Mixed State Presentations	88
3.6	Identifying the Hidden Linear Dynamic	91
3.6.1	Simple Complexity from any Presentation	91
3.6.2	Predictability from MSP of any presentation	94
3.6.3	Synchronization from MSP of generating model	98
3.6.4	Optimal prediction from MSP of ϵ -machine	100
3.6.5	Beyond the MSP	101
3.6.6	The end?	102
3.7	Spectral Theory beyond the Spectral Theorem	103
3.7.1	Spectral Primer	103
3.7.2	Eigenprojectors and right and left eigenvectors and generalized eigenvectors	105
3.7.3	Companion operators and decomposition of the resolvent	107
3.7.4	Arbitrary functions of nondiagonalizable operators	108
3.7.5	Evaluating the residues	109
3.7.6	Decomposing A^L	109
3.7.7	Drazin Inverse	110
3.8	Projection Operators for Stochastic Transition Dynamics	111
3.8.1	Row Sums: $T_\lambda \mathbf{1}\rangle = \delta_{\lambda,1} \mathbf{1}\rangle$	112
3.8.2	Expected Stationary Distribution from T_1	112
3.9	Spectra by inSppection: Useful rules for weighted digraphs	113
3.9.1	Eigenvalues by inspection	113
3.9.2	Eigenprojectors from graph structure	114
3.10	The Breadth of Possible Behaviors for Correlation and Myopic Uncertainty	116
3.10.1	Asymptotic correlation	118
3.10.2	Asymptotic entropy rate	119
3.11	Accumulated Transients for Diagonalizable Dynamics	121
3.12	Exact Complexities and Complexity Spectra for Nondiagonalizable Dynamics	123

3.12.1	Excess Entropy	125
3.12.2	The Excess Entropy Spectrum	126
3.12.3	Synchronization Information	129
3.12.4	Power Spectra	131
3.12.5	Almost Diagonalizable Dynamics	131
3.12.6	Processes of Finite Markov Order vs. Processes of Finite Symmetry Collapse	133
3.13	Coronal Spectrograms	134
3.14	Examples	135
3.14.1	$(R-k)$ -Golden Mean Process Family	135
3.14.2	Even Process	142
3.14.3	The $(\nu_0(\mathcal{W})-k)$ -Golden—Parity- (P) Process Family	142
3.15	The Importance of Predicting Beyond-Pairwise Structure: The RRX Pro- cess as a Paradigm for Processes that Compute	145
3.16	Example analytical calculation	150
3.16.1	Observed Correlation	154
3.16.2	Predictability	157
3.16.3	Synchronization to Optimal Prediction	162
3.17	Conclusion	164
4	Quantum Simplicity of Classical Complexity	167
4.1	Chapter Overview	167
4.2	Introduction	168
4.3	Two Representations of a Process	170
4.3.1	ϵ -Machine: Optimal, Minimal Predictor	171
4.3.2	q-Machine	173
4.4	Quantum Overlaps	175
4.5	Quantum Pairwise-Merger Machine	177
4.6	Overlaps from the QPMM	179
4.6.1	Finite Horizon: $(R-k)$ -Golden Mean Process	180

4.6.2	Infinite Horizon: Lollipop Process	186
4.7	Quantum-Reduced State-Complexity	191
4.8	Costs using Long Codewords	194
4.9	Conclusion	203
4.10	Appendices	205
4.10.1	Appendix A: Mealy HMMs	205
4.10.2	Appendix B: Quantum Overlaps and Cryptic Order	206
4.10.3	Appendix C: Matrices and Their Entropy	208
4.10.4	Appendix D: Examples	212
5	Correlation in Layered Structures	219
5.1	Chapter Overview	219
5.2	Introduction	219
5.3	Definitions and Notations	223
5.3.1	Correlation functions	223
5.3.2	Representing layer stacking as a hidden process	224
5.4	Expanding the Hägg-machine	
	to the <i>ABC</i> -machine	226
5.4.1	Graphical expansion method	227
5.4.2	Mixing and Nonmixing State Cycles	230
5.4.3	Rote expansion algorithm	231
5.5	Correlation Functions from HMMs	233
5.5.1	CFs from TMs	234
5.5.2	CFs from Spectral Decomposition	236
5.5.3	Asymptotic behavior of the CFs	239
5.5.4	Modes of Decay	240
5.6	Examples	241
5.6.1	3C Polytypes and Random ML Stacking: IID Processes	241
5.6.2	Random Growth and Deformation Faults in Layered 3C and 2H CPSs: The RGDF Process	251

5.6.3	Shockley–Frank Stacking Faults in 6H-SiC: The SFSF Process . . .	261
5.7	Conclusion	265
5.8	Appendix A: Transition matrices, projection operators, and bra–ket notation	270
6	Diffraction Patterns from HMMs	271
6.1	Chapter Overview	271
6.2	Introduction	272
6.3	Background	275
6.3.1	Brief History of Planar Disorder and Diffraction Patterns	275
6.3.2	Markov Models and Hidden Markov Models	277
6.4	Definitions and Notations	280
6.4.1	Correlation Functions and Stacking Notation	281
6.4.2	The Stacking Process as a Hidden Markov Model	282
6.4.3	Mixing and Nonmixing Machines	283
6.4.4	Power Spectra	284
6.5	Diffraction Patterns from Hidden Markov Models	285
6.5.1	Diffuse Scattering	286
6.5.2	Bragg Reflections	289
6.5.3	Full Spectral Treatment of the Diffuse Spectrum	292
6.6	Examples	293
6.6.1	3C Polytypes and Random ML Stacking: IID Process	293
6.6.2	Random Growth and Deformation Faults in Layered 3C and 2H CPSs: The RGDF Process	299
6.6.3	Shockley–Frank Stacking Faults in 6H-SiC: The SFSF Process . .	304
6.7	Conclusions	308
6.8	Appendix A: Hägg-to- <i>ABC</i> Machine Translation	310
7	Fluctuations When Driving Between NESSs	312
7.1	Chapter Overview	312
7.2	Introduction	313

7.2.1	Results	314
7.2.2	Synopsis	315
7.3	Driven Stochastic Dynamics	316
7.3.1	Stochastic mesoscopic dynamics and induced state-distributions	317
7.4	Energetics and Entropies	318
7.4.1	Work, heat, and their excesses	319
7.4.2	Excess environmental entropy production	322
7.4.3	Path entropies	323
7.5	Irreversibility	326
7.5.1	Detailed balance	327
7.5.2	Microscopic reversibility	327
7.5.3	Path dependence and irreversibility	328
7.6	Generalized fluctuation theorems for nonequilibrium systems	330
7.6.1	Generalized Detailed Fluctuation Theorem	330
7.6.2	Simplifications	332
7.6.3	Generalized Crooks Fluctuation Theorem	333
7.6.4	Interpretation	334
7.6.5	Translation to Steady-State Thermodynamics	336
7.6.6	Integral fluctuation theorems	338
7.6.7	Fluctuation theorems with an auxiliary variable	340
7.7	NESS Transitions in Neuronal Dynamics	343
7.7.1	Ion Channel Dynamics	343
7.7.2	Ion channel (ir)reversibility	346
7.7.3	Step Function Drive	349
7.7.4	Intrinsic feedback	354
7.8	Discussion	355
7.9	Conclusion	357
7.10	Appendices	358

7.10.1	Appendix A: Extension to Non-Markovian Instantaneous Dynamics	358
7.10.2	Appendix B: Integral fluctuation theorems with auxiliary variables	361
8	Strong Coupling, Exact Excess, and Renormalized Housekeeping	363
8.1	Chapter Overview	363
8.2	Setup and Results	364
8.2.1	Housekeeping renormalization of excess heat in hierarchical nonequilibrium steady states	371
9	Statistical–Computational Mechanics	373
9.1	Chapter Overview	373
9.2	Thermodynamic entropy	374
9.3	Non-extensivity of entropy	376
9.4	Instantaneously stationary processes of configuration	379
9.4.1	Demons and information reservoirs	380
9.5	Particles	383
9.5.1	Hamiltonians and ‘particles’	383
9.5.2	Instantaneously conditionally-stationary configurational process	384
9.6	Complexity of the Mesoscale	385
9.7	Conclusion	385

ABSTRACT

Exact Results Regarding the Physics of Complex Systems

via

Linear Algebra, Hidden Markov Models, and Information Theory

How can we ever make sense of what we observe? As a practical matter, most complex systems—that is, many-bodied systems with strongly interacting degrees of freedom—can only be observed through a time-series of relatively few functionals of their microstate. Nevertheless, because of the strong coupling, the many instantaneously-hidden degrees of freedom inject themselves over time into the observable time-series—giving us hope for inference.

This dissertation delivers new broadly-applicable results regarding the generation, prediction, and physical implication of such stochastic time-series with hidden structure. After the development of the general mathematical theory, the remainder of the dissertation can be subdivided into three parts. The first part addresses the fundamental limits of generation and predictability of structured stochastic processes. The second part identifies the possible correlation in and diffraction patterns of chaotic crystals. The third part establishes new relationships that constrain and elucidate the fluctuations and thermodynamics of nonequilibrium systems.

One of the predominant themes in this dissertation is the use of rather flexible mathematical structures called ‘hidden Markov models’. Indeed, much of this dissertation grew out of the recognition that—beyond their ability to simulate many sophisticated nonlinear and non-Markovian processes of interest—hidden Markov models enable an exact linear algebraic analysis of processes they represent. However, to proceed required the development of a generalized spectral theory for arbitrary functions of potentially nondiagonalizable operators, which is developed and utilized herein. Despite its long history, it appears (somewhat surprisingly) that not all of linear algebra had been worked out to the extent necessary to address the physics of complex systems. This extension of

the more familiar spectral theory is of interest in its own right, and has created several new and rather independent directions of inquiry.

ACKNOWLEDGMENTS

This dissertation has been long in the making and was by no means a solo journey.

Foremost, I am happy to thank my Ph.D. advisor, Jim Crutchfield. I am grateful for his mentorship over the years and for our fruitful collaboration. With his wit always about him, Jim has a characteristic way of making consequential connections among ideas from disparate realms of inquiry. I like to believe that some of this has rubbed off on me. Beyond the details of what I have learned, Jim has helped me to understand the importance of precise language—not only for communicating, but also for probing and advancing one’s own understanding of a subject. He has truly helped me to become a better thinker.

It is noteworthy that Jim has cultivated a special community and culture for research at UC Davis and beyond that resonates with both the mind and soul; I gratefully acknowledge the powerful role of this positive and stimulating atmosphere. Many among this community have been good friends and an important part of my enlightenment thus far. There are a special few that I wish to thank explicitly:

Chris Ellison: friend, mentor, and the king of mixed states.

John Mahoney, who has proven to always have a refreshingly unique way to think about a problem.

Dowman Varn, whose enthusiasm for chaotic crystals is contagious. I appreciate our many late-night discussions.

Ryan James, the master of information metrics, coding, and computers.

Sarah Marzen, for friendship, calculational mastery, skepticism, and brilliance.

Alec Boyd, for friendship, compassion, helping me to understand that detailed balance can be an emergent property, and letting loose.

Dibyendu Mandal for his thoughtful approach to science, and helping me understand what others perceive as ‘physical’ in the context of nonequilibrium thermodynamics.

Mark Goldman for helping me to assimilate some knowledge of theoretical, computational, and experimental neuroscience through his first-rate journal clubs and as his

teaching assistant.

Although it has been a while, I must thank Seth Weil and David Barlev for enlightening discussions and brain training when it was critical.

Nix Barnett, Dmitry Shemetov, Cina Aghamohammadi, Korana Burke, Gavin Crooks, Greg Wimsatt, Xincheng Lei, Pooneh Mohammadiara, Benny Brown, Adam Rupe, and all of the vibrant community of the Complexity Sciences Center, and indeed many others too numerous to gather here, have played an important and memorable role leading to this dissertation, for which I am thankful.

UC Davis, and especially the UC Davis Physics department, has been a warm and welcoming home for me for oh-so long. I am grateful especially for the wonderful physics faculty that have inspired and encouraged me, and who have worked so hard to make teaching, mentorship, and community a part of their legacy.

To Christine, my loving partner throughout my Ph.D. journey, who was most directly affected by my absentee mind-state and loved me anyway: for your encouragement, for reminding me of the grander scheme of things, for your smiles and laughter and positivity and rebelliousness, and for bringing so much joy and possibility to my life: thank you.

To my sisters, Gloria, Danielle, and Christina, who have been my best friends, enduring inspiration, and companions in creativity: for helping me to see the possibilities and for helping me to know no bounds: thank you.

To my parents, Peter and Carla: for instilling in me a strong self-confidence, for defying convention, for so clearly conveying the importance of working towards one's passions, and for endless love and support: thank you.

I am grateful and indebted to all of my family and friends who have not only tolerated my abstract pursuits, but have long encouraged and expanded my passions while sharing their own. You have brought outside light, amusement, and deeper meaning to my world. I will do my best to return the favor.

Chapter 1

Complex Systems, Impoverished Observables, and Hidden Markov Models

Hungry for understanding, we open our senses—eyes, probes, and colossal detectors—to both the natural and designed world around us in search of structure. The data deluge ensues. How do we make sense of it all? Is it noise or is there hidden structure to be inferred? How much structure exists in the data? Supposing there is predictable structure, what resources must be allocated to predict what is predictable?

Complex systems—that is, many-bodied systems with strong interactions—are usually observed through a time-series of low-bandwidth feature detectors, so that their hidden structure is only revealed over time. Because the instantaneous observations cannot capture the full resolution of each degree of freedom, the observed time-series is often stochastic and non-Markovian (i.e., memoryfull), with long-range correlations.

In general, dealing with non-Markovian dynamics is notoriously difficult [283]. Memory kernels serve some utility, but the analysis is generally not as simple as the theory of impulse response functions since the system need not be linear nor time invariant. Despite these generic difficulties, hidden Markov models (adapted from the field of machine learning) have been used with considerable success to model and analyze non-Markovian dynamics in the fields of complexity [55] and biomolecular dynamics [141].

This dissertation grew out of the recognition that—beyond their ability to *simulate* many sophisticated non-Markovian processes of interest—hidden Markov models (HMMs) can be treated as exact *mathematical objects* in the analysis of processes they represent. However, to proceed required the development of a generalized spectral theory for arbitrary functions of potentially nondiagonalizable operators. Despite its long history, it appears (somewhat surprisingly) that not all of linear algebra had been worked out to the extent necessary to address the physics of complex systems. This extension of the more familiar spectral theory is of interest in its own right, and has created several new and rather independent directions of inquiry.

The first two chapters of this dissertation set out the conceptual and mathematical foundation for utilizing HMMs in the exact analysis of complex systems. **This chapter**—an attempt at relative self-containment and an opportunity for the reader to more fully synchronize with the language of the writer—will review fundamental notions in probability, information theory, stochastic processes, and hidden Markov models. A fluency with these basics will undoubtedly aid in the appreciation of the more nuanced concepts in the later chapters. **Chapter 2** develops the generalized spectral theory necessary to fully leverage HMMs as objects of linear algebra. The subsequent chapters are somewhat modular, according to applications of the theory. After the initial development, the three domains of application I focus on are:

- The implicit structure and physical limits of the generation and prediction of a time-series (**Chapters 3 and 4**);
- Correlation and diffraction patterns of chaotic crystals (**Chapters 5 and 6**); and
- Nonequilibrium thermodynamics (**Chapters 7, 8, and 9**).

Several of these chapters have appeared in print with only minor modifications in how they appear here, and have benefited from the contributions of co-authors. I make these acknowledgements at the beginning of each of those chapters.

1.1 Probability, Stochastic Dynamics, and HMMs

1.1.1 Probability

I assume the reader has at least a basic working knowledge of probability theory, at the level of Ref. [187]. However, since we¹ will be using probabilities extensively throughout this dissertation, let us review our basic notation for probability, probability distributions, joint probability, and conditional probability:

$\Pr(A = a)$ is the *probability* that the random variable A will take on the value a (1.1)

$\Pr(A)$ is the *probability distribution* of the random variable A (1.2)

$\Pr(A, B)$ is the *joint probability distribution* of A and B (1.3)

$\Pr(A|B = b) = \frac{\Pr(A, B=b)}{\Pr(B=b)}$ is the *conditional probability distribution* for A given $B = b$ (1.4)

$\Pr(A|B)$ is the *conditional probability distribution* of A given B , as a function of B (1.5)

$\Pr(A = a|B)$ is the *conditional probability* that $A = a$ given B , as a function of B (1.6)

Note that we typically use capital letters for random variables and lower case letters for a possible outcome. Often, to be concise, we will use the shorthand that $\Pr(A|B = b) = \Pr(A|b)$, $\Pr(A = a|B = b) = \Pr(a|b)$, and so on, so long as the implicated random variables are evident by context.

Note that the conditional probability of Eq. (1.6) $\Pr(a|B)$, as a function of B , is often called the *likelihood* of $B = b$ given $A = a$. It is important to recognize that the likelihood function $\Pr(a|B)$, plotted as a function of B , is *not* a probability distribution of B —notice that, as a function of B , $\Pr(a|B)$ is not normalized in probability. Sometimes it is preferable to think of B as a parametrization of a probability distribution rather than as a random variable; in such cases, one may write $\Pr_b(A)$ rather than $\Pr(A|B = b)$.

Similar statements to the above can be restated with minor modification for probability

¹By ‘we’, I mean *you*, dear reader, and me (and co-authors wherever applicable). We are in this journey together. Indeed my hope is that you, with any luck and a bit of work, can extend what I present here.

densities, which we denote with $p(\cdot)$ rather than $\Pr(\cdot)$. Probability densities $p(A)$ are necessary when the state-space of A is uncountably infinite. In the following, wherever we can discuss either case on an equal footing, we will use $\mathbb{P}(\cdot)$ to denote either probability $\Pr(\cdot)$ or probability density $p(\cdot)$, depending on the cardinality of the random-variable's state-space.

We denote the expectation value of a function $f(A)$ of a random variable A as: $\langle f(A) \rangle$, where it is implied that we are averaging with respect to $\mathbb{P}(A)$. Sometimes it is useful to be explicit about the probability distribution that we are averaging with respect to, in which case we may write $\langle f(A) \rangle_{\mathbb{P}(A)}$ or $\langle f(A) \rangle_{\mathbb{P}(A|B=b)}$.

Allowing probability density functions to contain Dirac delta functions in addition to normal functions makes probability density functions more versatile than discrete probability distributions. Ultimately, partitioning the random-variable's state-space via a σ -algebra, à la measure theory, yields probabilities from measurable subsets of the state-space.

1.1.2 Shannon Entropy

Since we will be dealing largely with probabilistic systems, it will be important to have a tool to quantify the randomness of collections of random variables. Shannon entropy, broadly interpreted, gives us just such a tool [226, 46, 62]. Recall that collections of random variables are themselves random variables; therefore, a few simple definitions will take us quite far through implication.

For any random variable X with possible outcomes $x \in \mathcal{A}$ distributed as $X \sim \Pr(X)$, the *Shannon entropy* of the distribution is:

$$H[\Pr(X)] \equiv - \sum_{x \in \mathcal{A}} \Pr(x) \log \Pr(x) . \quad (1.7)$$

Notice that this is the expectation value $\langle -\log \Pr(X) \rangle_{\Pr(X)}$ of the *surprisal* $-\log \Pr(x)$ of a particular outcome x . In information theory, it is typical to use the base-2 logarithm $\log_2(\cdot)$, in which case entropy is quantified in units of *bits*. The entropy in bits corresponds roughly to how many yes/no questions a strategic questioner would need to ask on average to determine the outcome of the random variable. In thermodynamics, we will use the

base- e natural logarithm $\log_e(\cdot) = \ln(\cdot)$ to quantify Shannon entropy, in which case the Shannon entropy is quantified in units of *nats*—some play on ‘natural bits’. It turns out that *thermodynamic entropy* corresponds to Boltzmann’s constant k_B times a *particular* Shannon entropy—even in nonequilibrium. More on that later. Occasionally, we will use the shorthand $H(A)$ for the Shannon entropy of the distribution $\Pr(A)$.

The *conditional entropy* $H(X|Y)$ is $\langle H[\Pr(X|Y = y)] \rangle_{\mathbb{P}(Y)}$. When Y is a discrete random variable, we can write:

$$H(X|Y) = \sum_y \sum_x -\Pr(Y = y) \Pr(X = x|Y = y) \log(\Pr(X = x|Y = y)) \quad (1.8)$$

$$= \sum_y \sum_x -\Pr(X = x, Y = y) \log(\Pr(X = x|Y = y)) . \quad (1.9)$$

Because of the logarithm in the definition of entropy, many rules where probabilities would multiply (divide) have corresponding rules where entropies add (subtract). For example, for cases where Y is a discrete random variable, it is useful to keep the following easily-derivable property in mind:

$$H(X, Y) = H(X|Y) + H(Y) . \quad (1.10)$$

More generally (and usefully), this leads to the chain rule for conditional entropies:

$$H(X_1, X_2, \dots, X_N) = \sum_{n=1}^N H(X_n|X_1, \dots, X_{n-1}) . \quad (1.11)$$

We finish our brief introduction to information theory by defining the *Kullback–Leibler divergence* D_{KL} between two distributions Q and R with the same support \mathcal{A} :

$$D_{KL}(Q \| R) \equiv \sum_{x \in \mathcal{A}} Q(x) \log \frac{Q(x)}{R(x)} , \quad (1.12)$$

which is always non-negative, and is positive whenever $Q \neq R$. The *mutual information* between two variables $I(X; Y)$ can be defined as:

$$I(X; Y) \equiv D_{KL}(\Pr(X, Y) \| \Pr(X) \Pr(Y)) . \quad (1.13)$$

Hence, the mutual information between two variables is only zero if the two variables are statistically independent; i.e., if $\Pr(X, Y) = \Pr(X) \Pr(Y)$. The mutual information

therefore captures statistical dependency among random variables beyond what can be captured by linear standards of correlation. The mutual information can also be shown to be:

$$I(X; Y) = H(X) + H(Y) - H(X, Y) \quad (1.14)$$

$$= H(X) - H(X|Y) \quad (1.15)$$

$$= H(Y) - H(Y|X) , \quad (1.16)$$

which is the amount of uncertainty vanquished about one variable given knowledge of another. The entropy of a variable is the information contained in itself:

$$I(X; X) = H(X) . \quad (1.17)$$

1.1.3 Stochastic Processes

Deterministic dynamics is a special case of stochastic dynamics. So, by considering stochastic dynamics, we can greatly expand the repertoire of dynamical systems in physics.

Suppose a physical system generates an observable stochastic process $\{X_t\}_{t \in D}$, where X_t is the random variable for the observation at time t , which takes on the value $x_t \in \mathcal{A}$. The alphabet of observables \mathcal{A} can be a countable or uncountable set. The time domain D can either be discrete, in which case we typically choose the domain to be the set of integers $D = \mathbb{Z}$, or can be continuous, in which case the domain is the set of real numbers $D = \mathbb{R}$.

More generally, the domain D can be a multi-dimensional set, and does not need to represent time. For example, to complement the notion of patterns in time, we will later consider patterns in space. As a simple example, consider orientations of spin along a lattice. Later, we consider spatial correlation among types of modular layers along the stacking direction of a chaotic crystal.

The random variables $\{X_t\}_t$ can have arbitrarily sophisticated interdependencies. Observable realizations are generated probabilistically according to the constraints of these interdependencies. Critical to the analysis of general stochastic processes, these constraints yield a consistent family of measures—with a probability density over outcomes for each finite subset of observable random variables.

Often, in unraveling the consequences of a stochastic process, it is sufficient to consider families of finite-dimensional distributions. For example, in the discrete-time setting, each (t, L) -pair with $t \in \mathbb{Z}$ and $L \in \{1, 2, \dots\}$ invokes a probability distribution $\mathbb{P}(X_{t:t+L})$ over possible outcomes $x_{t:t+L} \in \mathcal{A}^L$. To be concise, we have denoted the contiguous block of random variables from time t to time $t + L - 1$ as $X_{t:t+L} \equiv X_t X_{t+1} \dots X_{t+L-1}$ —notice left inclusiveness, but right exclusiveness of the time-index. Similarly, $x_{t:t+L}$ denotes the length- L realization $x_t \dots x_{t+L-1}$. Joint probability distributions for non-contiguous random variables can be obtained by marginalizing over the intervening random variables.

A stochastic process is said to be a *stationary process* if the probability measure is invariant under a global time-shift:

$$\mathbb{P}(X_{t_1} \dots X_{t_L}) = \mathbb{P}(X_{t_1+t'} \dots X_{t_L+t'}) \text{ for all } L \in \mathbb{N} \text{ and all } t', t_1, \dots, t_L \in D. \quad (1.18)$$

Assuming stationarity eases analysis; however, stationarity is a rather confining assumption. Fortunately, we can retain similar tractability with the less stringent constraint of wide-sense stationarity (also called ‘weak-sense stationarity’). A stochastic process is said to be a *wide-sense stationary process* if:

$$\langle X_t \rangle = \langle X_0 \rangle \text{ and } \langle X_t X_{t+\tau} \rangle = \langle X_0 X_\tau \rangle \text{ for all } t, \tau \in D. \quad (1.19)$$

Not all wide-sense stationary processes are stationary processes. Later we will show that HMMs are capable of producing various types of nonstationary one-sided processes that are also not wide-sense stationary.

The simplest class of stochastic processes are *independent and identically distributed* (IID) processes: At each time, the observable is independently sampled from the same distribution as ever, oblivious of and unaffected by all past and future outcomes.

Because IID processes are easy to deal with, they are commonly assumed for many applications. For example, IID processes, like Gaussian white noise, serve as the predominant model of noise in physics—completely uncorrelated in time. See also the elementary development of information theory [46].

We will want to study processes with memory and temporal correlation. The simplest among these are the Markov processes.

1.1.4 Markovity and Markov order

Markovian processes are those for which the future is conditionally independent of the past, given the present observable state. Mathematically, a Markov process can be identified by the *Markov property*:

$$\mathbb{P}(X_{t+t'} | \{X_\tau\}_{\tau \in D \setminus (t, \infty)}) = \mathbb{P}(X_{t+t'} | X_t) \text{ for all } t' > 0, \quad (1.20)$$

where $t' \in \mathbb{N}$ for discrete-time processes, and $t' \in \mathbb{R}_{>0}$ for continuous-time processes.

Our above definition of Markovity allows us to treat discrete-time and continuous-time processes on an equal footing. However, an immediate implication for the discrete-time processes follows from the above definition of Markovity:

$$\{X_t\}_{t \in \mathbb{Z}} \text{ is a Markov process} \leftrightarrow \mathbb{P}(X_t | X_{-\infty:t}) = \mathbb{P}(X_t | X_{t-1}), \quad (1.21)$$

which, in fact, sometimes serves as the definition of Markovity for discrete-time processes [62]. Continuous-time processes do not quite enjoy the same simplification since there is no smallest time step. However, in either case the intuition is similar: For Markov processes, where you were most recently encapsulates all that is relevant about the past regarding where you will go next.

Most common stochastic models in physics assume Markovity, either explicitly or implicitly. Indeed, this is a familiar assumption in the study of dynamical systems, related to the concept that trajectories don't cross in phase space.

However, *non-Markovian processes* are typical when there are hidden degrees of freedom.

Sometimes, extended observations can make the process Markovian over an extended alphabet. These are the processes of finite Markov order. For any process, the *Markov order* R is:

$$R \equiv \inf \left\{ r \in \{0, 1, 2, \dots\} : \mathbb{P}(X_{t+t'} | \{X_\tau\}_{\tau \in D \setminus (t, \infty)}) = \mathbb{P}(X_{t+t'} | X_{t-t_1}, X_{t-t_2}, \dots, X_{t-t_r} \in \{X_\tau\}_{\tau \in D \setminus (t, \infty)}) \right. \\ \left. \text{for all } t' > 0 \text{ and for some } t_r > \dots > t_1 \geq 0 \right\}. \quad (1.22)$$

It should be clear that the Markov order depends on what parts of the system can be observed—what probes, sensory modalities, and measurement bandwidth one has. Whether

		discrete state	continuous state
		e.g., $x = (\uparrow, \downarrow)$	e.g., $x = (\vec{x}, \vec{p})$
discrete time		Markov model $\rightarrow \Pr(x_{t+1} x_t)$	Markov chain Monte Carlo (MCMC) $\rightarrow p(x_{t+1} x_t)$
continuous time		Markov jump process $\rightarrow \Pr(x_{t+t'} x_t)$	Langevin equations $\rightarrow p(x_{t+t'} x_t)$

Table 1.1: Markov processes can occur in discrete or continuous time, and with a discrete or continuous state-space. Each of the resulting four types of Markov processes have found ample utility in physics; examples are indicated.

in continuous or discrete time, Markov order answers the question: Given my apparatus, what is the minimal number of observations that I must make to be as certain as possible about the future? For discrete time processes, this simplifies to:

$$R = \inf \{ r \in \{0, 1, 2, \dots\} : \mathbb{P}(X_t | X_{-\infty:t}) = \mathbb{P}(X_t | X_{t-r:t}) \} . \quad (1.23)$$

Intuitively, Markov order is also the answer to a related question: How far into the past do I need to look to do optimal prediction? So, for example, a biased coin (or any IID process) has Markov order $R = 0$, Langevin dynamics are of Markov order $R = 1$ in the dimension of the state space, a sinusoid of unknown phase has Markov order $R = 2$, and spoken language has Markov order $R = \infty$. It is notable that our definition of Markov order implies that the Markov order of an observed scalar-valued time-series generated via projection of a linear or nonlinear deterministic differential equation is equal to its embedding dimension [184, 247], which is a welcome extension of the typical notion of Markov order. Markovian processes are those processes with Markov order 1 or Markov order 0. Many processes have infinite Markov order.

Hidden Markov models are capable of generating non-Markovian observed dynamics as a byproduct of a Markovian dynamic over latent states. Even processes with infinite Markov order can be generated by finite-state HMMs.

1.1.5 HMMs in general

HMMs are capable of generating non-Markovian observed dynamics as a byproduct of a Markovian dynamic over latent states. There are many possible types of HMMs, especially since we interpret the notion rather broadly. Nevertheless, all HMMs are unified by several characteristics that make them both powerful and tractable tools for the analysis of complex stochastic processes.

Regardless of model class, each HMM has: a *set of latent states* \mathcal{S} , a net latent-state-to-state transition operator T , and an *observable alphabet* \mathcal{A} . Let \mathcal{S}_t be the random variable for the latent state at time t , which takes on the value $s_t \in \mathcal{S}$. Let X_t be the random variable for the observation at time t , which takes on the value $x_t \in \mathcal{A}$. The following are true for all classes of hidden Markov models:

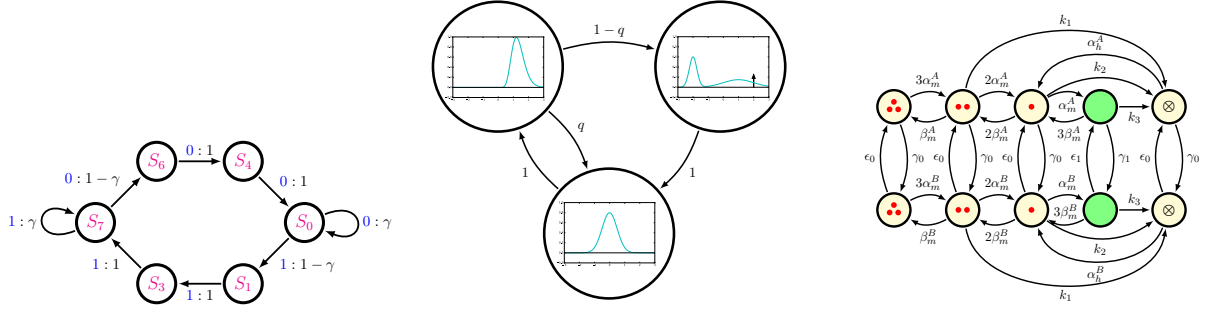
- The *observable process* $\{X_t\}_t$ can be *non-Markovian*
- The *hidden process* $\{\mathcal{S}_t\}_t$ must be *Markovian*
- The next observation is conditionally independent of all past observations, given the current latent state:

$$\mathbb{P}(X_t | \mathcal{S}_t, X_{t-1}, X_{t-2}, \dots) = \mathbb{P}(X_t | \mathcal{S}_t) \quad .$$

1.1.6 Welcome to the jungle: Zoology of HMMs

We just discussed general characteristics of HMMs. Fig. 1.1 introduces several more specific classes of HMMs by illustration. Each class entails a particular assemblage of linear operators that can be used to calculate probabilities for the observable stochastic process that it generates. These will be introduced as appropriate throughout the dissertation.

In application, much work often goes into inferring a HMM, either from observed data or first principles. Once inferred, the HMM most often serves as a numerical workhorse; e.g., for pattern matching (as in voice recognition) or to numerically simulate more data. However, we wish to emphasize that these HMMs can instead be used as exact mathematical objects, yielding new opportunity for physical *theory*.



(a) Discrete-domain edge-emitting HMM with alphabet $\mathcal{A} = \{0, 1\}$ representing either acyclic or cyclic rotations of subsequent modular layers of a close-packed SiC crystal. Each number on the right side of the colon in each of the edge labels is a transition probability.

(b) Discrete-time state-emitting HMM with real-valued observable $x \in \mathcal{A} = \mathbb{R}$. This HMM generates a stochastic process according to the state-to-state transition dynamic T and the probability density functions (pdfs) $\{p(x|s)\}_{s \in \mathcal{S}}$ associated with each state.

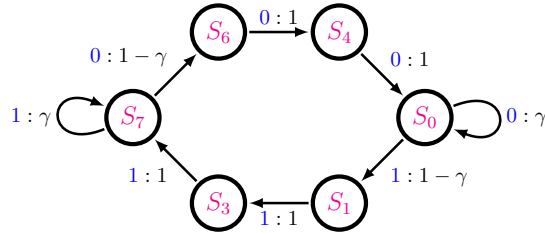
(c) Continuous-time function of a Markov chain, representing a joint molecule-environment dynamic, with symbolic observables representing conformational states of the Na^+ channel macromolecule. Edge weights here are transition rates, rather than transition probabilities.

Figure 1.1: Simple finite-state example HMMs meant to introduce several different classes of hidden Markov models, with discrete or continuous time, countable or uncountable cardinality of the observable alphabet, and observables emitted either within states or on edge transitions. Regardless of model class, each HMM has a set of latent states \mathcal{S} , net latent-state-to-state transition operator T , and observable alphabet \mathcal{A} .

1.1.7 Probabilities, conditional probabilities, and linear algebra

Despite their differences, each class of HMM allows a method to calculate probabilities by the concatenation of linear operators. Bayes rule, marginalization, and other rules of probability all find natural analogs in the rules of linear algebra. Let us explore an example calculation.

As a first example, let us consider an edge-emitting HMM, as in Fig. 1.1a:



Let $T_{i,j}^{(x)} \equiv \Pr(\mathcal{S}_{t+1} = s^j, X_t = x | \mathcal{S}_t = s^i)$ define the set of sub-stochastic observable-labeled

transition matrices relevant to this class of HMMs. Then $T = \sum_{x \in \mathcal{A}} T^{(x)}$ is a row-stochastic net latent-state-to-state transition matrix with $T_{i,j} = \Pr(\mathcal{S}_{t+1} = s^j | \mathcal{S}_t = s^i)$ and stationary distribution $\pi T = \pi$.

Assuming $0, 1 \in \mathcal{A}$, let us proceed with what should be a very simple calculation: that of $\Pr(X_0 = 1, X_2 = 0)$. We calculate:

$$\Pr(X_0 = 1, X_2 = 0) = \sum_{s, s', s'', s''' \in \mathcal{S}} \Pr(\mathcal{S}_0 = s, \mathcal{S}_1 = s', \mathcal{S}_2 = s'', \mathcal{S}_3 = s''', X_0 = 1, X_2 = 0) \quad (1.24)$$

$$= \sum_{s, s', s'', s''' \in \mathcal{S}} \Pr(\mathcal{S}_0 = s) \Pr(\mathcal{S}_1 = s', X_0 = 1 | \mathcal{S}_0 = s) \Pr(\mathcal{S}_2 = s'', \mathcal{S}_3 = s''', X_2 = 0 | \mathcal{S}_1 = s', X_0 = 1, \mathcal{S}_0 = s) \quad (1.25)$$

$$= \sum_{s, s', s'', s''' \in \mathcal{S}} \Pr(\mathcal{S}_0 = s) \Pr(\mathcal{S}_1 = s', X_0 = 1 | \mathcal{S}_0 = s) \Pr(\mathcal{S}_2 = s'', \mathcal{S}_3 = s''', X_2 = 0 | \mathcal{S}_1 = s') \quad (1.26)$$

$$= \sum_{s, s', s'', s''' \in \mathcal{S}} \Pr(\mathcal{S}_0 = s) \Pr(\mathcal{S}_1 = s', X_0 = 1 | \mathcal{S}_0 = s) \Pr(\mathcal{S}_2 = s'' | \mathcal{S}_1 = s') \Pr(\mathcal{S}_3 = s''', X_2 = 0 | \mathcal{S}_1 = s', \mathcal{S}_2 = s'') \quad (1.27)$$

$$= \sum_{s, s', s'', s''' \in \mathcal{S}} \Pr(\mathcal{S}_0 = s) \Pr(\mathcal{S}_1 = s', X_0 = 1 | \mathcal{S}_0 = s) \Pr(\mathcal{S}_2 = s'' | \mathcal{S}_1 = s') \Pr(\mathcal{S}_3 = s''', X_2 = 0 | \mathcal{S}_2 = s'') \quad (1.28)$$

$$= \sum_{s, s', s'', s''' \in \mathcal{S}} \pi_s T_{s, s'}^{(1)} T_{s', s''} T_{s'', s'''}^{(0)} \quad (1.29)$$

$$= \langle \pi | T^{(1)} T T^{(0)} | \mathbf{1} \rangle, \quad (1.30)$$

where $\langle \pi |$ is the stationary distribution over latent states, cast as a row vector, and $|\mathbf{1}\rangle$ is a column-vector of all ones. In the first line, Eq. (1.24), we expanded the probability in terms of a marginalization over the joint distribution including the latent states. With this joint distribution, we were able to decompose the probability in Eq. (1.25) according to repeated use of the chain rule of probability. Then, in going to Eq. (1.26), we invoked conditional independence: that—by the explicit assumptions built in to any HMM— \mathcal{S}_1 shields (in the sense of conditional probability) \mathcal{S}_2 , X_2 , and \mathcal{S}_3 from \mathcal{S}_0 and X_0 . Again using the chain rule and invoking conditional independence, and then identifying matrix elements and implied matrix products, we finally arrive at Eq. (1.30).

Much detailed work of marginalization, decomposing probabilities according to the chain rule, and invoking conditional independence has reduced the result in the end to simple elegant structure, alluding to deeper principles and opportunity for intuition and heuristics.

Indeed, for this type of HMM, which produces observable symbols $x \in \mathcal{A}$ as it transitions from one latent state to another, the probability of any contiguous length- L string of observables $w = x_{0:L} = x_0 x_1 \dots x_{L-1}$ can be calculated via: $\Pr(w) = \langle \pi | T^{(w)} | \mathbf{1} \rangle = \langle \pi | T^{(x_0)} T^{(x_1)} \dots T^{(x_{L-1})} | \mathbf{1} \rangle$.

A similar, albeit much more arduous, calculation would yield:

$$\Pr(X_0 = 1, X_5 = 1, X_{52} = 0) = \langle \boldsymbol{\pi} | T^{(1)} T^4 T^{(1)} T^{46} T^{(0)} | \mathbf{1} \rangle . \quad (1.31)$$

However, we start to see a trend: that marginalizing over observables corresponds to simply inserting a symbol-agnostic net state-to-state transition operator T . Many such heuristics can be assimilated into one's repertoire quite quickly, and allow a physicist to make some serious progress with serious calculations.

1.1.8 Beyond normal: Probabilities invoke non-normal linear operators

Already, from the sight of T^{46} , we know that we will want to perform some serious linear algebra. Moreover, we want to do this analytically to obtain general theoretical insights. However, the linear operators associated with hidden Markov models are nonnormal and often nondiagonalizable. The familiar spectral theorem for normal operators (which serves as the mathematical backbone for much work in physics) is therefore not applicable. The necessary extension of spectral theory is developed next.

Chapter 2

Beyond the Spectral Theorem: Spectrally Decomposing Arbitrary Functions of Nondiagonalizable Operators

This chapter has appeared online as a preprint:

P. M. Riechers & J. P. Crutchfield, (2016). “Beyond the Spectral Theorem: Spectrally Decomposing Arbitrary Functions of Nondiagonalizable Operators”. arXiv preprint arXiv:1607.06526.

2.1 Chapter Overview

Nonlinearities in finite dimensions can be linearized by projecting them into infinite dimensions. Unfortunately, often the linear operator techniques that one would then use simply fail since the operators cannot be diagonalized. This curse is well known. It also occurs for finite-dimensional linear operators. We circumvent it by developing a *meromorphic functional calculus* that can decompose arbitrary functions of nondiagonalizable linear operators in terms of their eigenvalues and projection operators. It extends the spectral theorem of normal operators to a much wider class, including circumstances in which poles and zeros of the function coincide with the operator spectrum. By allowing the direct manipulation of individual eigenspaces of nonnormal and nondiagonalizable operators, the

new theory avoids spurious divergences. As such, it yields novel insights and closed-form expressions across several areas of physics in which nondiagonalizable dynamics are relevant, including memoryful stochastic processes, open nonunitary quantum systems, and far-from-equilibrium thermodynamics.

The technical contributions include the first full treatment of arbitrary powers of an operator. In particular, we show that the Drazin inverse, previously only defined axiomatically, can be derived as the negative-one power of singular operators within the meromorphic functional calculus and we give a general method to construct it. We provide new formulae for constructing projection operators and delineate the relations between projection operators, eigenvectors, and generalized eigenvectors.

By way of illustrating its application, we explore several, rather distinct examples. First, we analyze stochastic transition operators in discrete and continuous time. Second, we show that nondiagonalizability can be a robust feature of a stochastic process, induced even by simple counting. As a result, we directly derive distributions of the Poisson process and point out that nondiagonalizability is intrinsic to it and the broad class of hidden semi-Markov processes. Third, we show that the Drazin inverse arises naturally in stochastic thermodynamics and that applying the meromorphic functional calculus provides closed-form solutions for the dynamics of key thermodynamic observables. Fourth, we show that many memoryful processes have power spectra indistinguishable from white noise, despite being highly organized. Nevertheless, whenever the power spectrum is nontrivial, it is a direct signature of the spectrum and projection operators of the process' hidden linear dynamic, with nondiagonalizable subspaces yielding qualitatively distinct line profiles. Finally, we draw connections to the Ruelle–Frobenius–Perron and Koopman operators for chaotic dynamical systems.

... the supreme goal of all theory is to make the irreducible basic elements as simple and as few as possible without having to surrender the adequate representation of a single datum of experience. A. Einstein [73, p. 165]

2.2 Introduction

Decomposing a complicated system into its constituent parts—reductionism—is one of science’s most powerful strategies for analysis and understanding. Large-scale systems with linearly coupled components give one paradigm of this success. Each can be decomposed into an equivalent system of independent elements using a similarity transformation calculated by the linear algebra of the system’s eigenvalues and eigenvectors. The physics of linear wave phenomena, whether of classical light or quantum mechanical amplitudes, sets the standard of complete reduction rather high. The dynamics is captured by an “operator” whose allowed or exhibited “modes” are the elementary behaviors out of which composite behaviors are constructed by simply weighing each mode’s contribution and adding them up.

However, one should not reduce a composite system more than is necessary nor, as is increasingly appreciated these days, more than one, in fact, can. Indeed, we live in a complex, nonlinear world whose constituents are strongly interacting. Often their key structures and memoryful behaviors emerge only over space and time. These are the complex systems. Yet, perhaps surprisingly, many complex systems with nonlinear dynamics correspond to linear operators in abstract high-dimensional spaces [131, 93, 34]. And so, there is a sense in which even these complex systems can be reduced to the study of independent nonlocal collective modes.

Reductionism, however, faces its own challenges even within its paradigmatic setting of linear systems: linear operators may have interdependent modes with irreducibly entwined behaviors. These irreducible components correspond to so-called nondiagonalizable subspaces. No similarity transformation can reduce them.

In this view, reductionism can only ever be a guide. The actual goal is to achieve a happy medium, as Einstein reminds us, of decomposing a system only to that level at which the parts are irreducible. To proceed, though, begs the original question, What happens when reductionism fails? To answer this requires revisiting one of its more successful implementations, spectral decomposition of completely reducible operators.

2.2.1 Spectral Decomposition

Spectral decomposition—splitting a linear operator into independent modes of simple behavior—has greatly accelerated progress in the physical sciences. The impact stems from the fact that spectral decomposition is not only a powerful mathematical tool for expressing the organization of large-scale systems, but also yields predictive theories with directly observable physical consequences [255]. Quantum mechanics and statistical mechanics identify the energy eigenvalues of Hamiltonians as the basic objects in thermodynamics: transitions among the energy eigenstates yield heat and work. The spectrum of eigenvalues reveals itself most directly in other kinds of spectra, such as the frequency spectra of light emitted by the gases that permeate the galactic filaments of our universe [211]. Quantized transitions, an initially mystifying feature of atomic-scale systems, correspond to distinct eigenvectors and discrete spacing between eigenvalues. The corresponding theory of spectral decomposition established the quantitative foundation of quantum mechanics.

The applications and discoveries enabled by spectral decomposition and the corresponding spectral theory fill a long list. In application, direct-bandgap semiconducting materials can be turned into light-emitting diodes (LEDs) or lasers by engineering the spatially-inhomogeneous distribution of energy eigenvalues and the occupation of their corresponding states [171]. Before their experimental discovery, anti-particles were anticipated as the nonoccupancy of negative-energy eigenstates of the Dirac Hamiltonian [67].

The spectral theory, though, extends far beyond physical science disciplines. In large measure, this arises since the evolution of any object corresponds to a linear dynamic in a sufficiently high-dimensional state space. Even nominally nonlinear dynamics over several variables, the canonical mechanism of deterministic chaos, appear as linear dynamics in appropriate infinite-dimensional shift-spaces [34]. A nondynamic version of rendering nonlinearities into linearities in a higher-dimensional feature space is exploited with much success today in machine learning by support vector machines, for example [44]. Spectral decomposition often allows a problem to be simplified by approximations that use only the dominant contributing modes. Indeed, human-face recognition can be efficiently

accomplished using a small basis of “eigenfaces” [233].

Certainly, there are many applications that highlight the importance of decomposition and the spectral theory of operators. However, a brief reflection on the mathematical history will give better context to its precise results, associated assumptions, and, more to the point, the generalizations we develop here in hopes of advancing the analysis and understanding of complex systems.

Following on early developments of operator theory by Hilbert and co-workers [45], the *spectral theorem for normal operators* reached maturity under von Neumann by the early 1930s [177, 271]. It became the mathematical backbone of much progress in physics since then, from classical partial differential equations to quantum physics. Normal operators, by definition, commute with their Hermitian conjugate: $A^\dagger A = AA^\dagger$. Examples include symmetric and orthogonal matrices in classical mechanics and Hermitian, skew-Hermitian, and unitary operators in quantum mechanics.

The spectral theorem itself is often identified as a collection of related results about normal operators; see, e.g., Ref. [107]. In the case of finite-dimensional vector spaces [105], the spectral theorem asserts that normal operators are diagonalizable and can always be diagonalized by a unitary transformation; that left and right eigenvectors (or eigenfunctions) are simply related by complex-conjugate transpose; that these eigenvectors form a complete basis; and that functions of a normal operator reduce to the action of the function on each eigenvalue. Most of these qualities survive with only moderate provisos in the infinite-dimensional case. In short, the spectral theorem makes physics governed by normal operators tractable.

The spectral theorem, though, appears powerless when faced with nonnormal and nondiagonalizable operators. What then are we to do when confronted, say, by complex interconnected systems with nonunitary time evolution, by open systems, by structures that emerge on space and time scales different from the equations of motion, or by other frontiers of physics governed by nonnormal and not-necessarily-diagonalizable operators? Where is the comparably constructive framework for calculations beyond the standard spectral theorem? Fortunately, portions of the necessary generalization have been made

within pure mathematics [70], some finding applications in engineering and control [169, 3]. However, what is available is incomplete. And, even that which is available is often not in a form adapted to perform calculations that lead to quantitative predictions.

2.2.2 Synopsis

Here, we build on previous work in functional analysis and operator theory to provide both a rigorous and constructive foundation for physically relevant calculations involving not-necessarily-diagonalizable operators. In effect, we extend the spectral theorem for normal operators to a broader setting, allowing generalized “modes” of nondiagonalizable systems to be identified and manipulated. The meromorphic functional calculus we develop extends Taylor series expansion and standard holomorphic functional calculus to analyze arbitrary functions of not-necessarily-diagonalizable operators. It readily handles singularities arising when poles (or zeros) of the function coincide with poles of the operator’s resolvent—poles that appear precisely at the eigenvalues of the operator. Pole–pole and pole–zero interactions substantially modify the complex-analytic residues within the functional calculus. A key result is that the negative-one power of a singular operator exists in the meromorphic functional calculus. It is the *Drazin inverse*, a powerful tool that is receiving increased attention in stochastic thermodynamics.

Taken altogether, the functional calculus, Drazin inverse, and methods to manipulate particular eigenspaces, are key to a thorough-going analysis of many complex systems, many now accessible for the first time. Indeed, the framework has already been fruitfully employed by the authors in several specific applications, including closed-form expressions for signal processing and information measures of hidden Markov processes [60, 204, 205] and for compressing stochastic processes over a quantum channel [203]. However, the techniques are sufficiently general they will be much more widely useful. We envision new opportunities for similar detailed analyses, ranging from biophysics to quantum field theory, wherever restrictions to normal operators and diagonalizability have been roadblocks.

With this broad scope in mind, we develop the mathematical theory first without reference to specific applications and disciplinary terminology. We later give pedagogical (yet, we hope, interesting) examples, exploring several niche, but important applications to

finite hidden Markov processes, basic stochastic process theory, nonequilibrium thermodynamics, signal processing, and nonlinear dynamical systems. At a minimum, the examples and their breadth serve to better acquaint readers with the basic methods required to employ the theory.

We introduce the meromorphic functional calculus in §2.4 through §2.5, after necessary preparation in §2.3. §2.6.1 further explores eigenprojectors, which we refer to here simply as *projection operators*. §2.6.2 makes explicit their relationship with eigenvectors and generalized eigenvectors. §2.6.2.4 then discusses simplifications of the functional calculus for special cases, while §2.7.1 takes up the spectral properties of transition operators. The examples are discussed at length in §2.7 before we close in §2.8 with suggestions on future applications and research directions.

2.3 Spectral Primer

The following is relatively self-contained, assuming basic familiarity with linear algebra at the level of Refs. [105, 169]—including eigen-decomposition and knowledge of the Jordan canonical form, partial fraction expansion (see Ref. [140]), and series expansion—and basic knowledge of complex analysis—including the residue theorem and calculation of residues at the level of Ref. [28]. For those lacking a working facility with these concepts, a quick review of §2.7’s applications may motivate reviewing them. In this section, we introduce our notation and, in doing so, remind the reader of certain basic concepts in linear algebra and complex analysis that will be used extensively in the following.

To begin, we restrict attention to operators with finite representations and only sometimes do we take the limit of dimension going to infinity. That is, we do not consider infinite-rank operators outright. While this runs counter to previous presentations in mathematical physics that consider only infinite-dimensional operators, the upshot is that they—as limiting operators—can be fully treated with a countable point spectrum. We present examples of this later on. Accordingly, we restrict our attention to operators with at most a countably infinite spectrum. Such operators share many features with finite-dimensional square matrices, and so we recall several elementary but essential facts

from matrix theory used extensively in the main development.

If A is a finite-dimensional square matrix, then its *spectrum* is simply the set Λ_A of its eigenvalues:

$$\Lambda_A = \{ \lambda \in \mathbb{C} : \det(\lambda I - A) = 0 \} ,$$

where $\det(\cdot)$ is the determinant of its argument and I is the identity matrix. The *algebraic multiplicity* a_λ of eigenvalue λ is the power of the term $(z - \lambda)$ in the characteristic polynomial $\det(zI - A)$. In contrast, the *geometric multiplicity* g_λ is the dimension of the kernel of the transformation $A - \lambda I$ or, equivalently, the number of linearly independent eigenvectors associated with the eigenvalue. The algebraic and geometric multiplicities are all equal when the matrix is diagonalizable.

Since there can be multiple subspaces associated with a single eigenvalue, corresponding to different Jordan blocks in the Jordan canonical form, it is structurally important to distinguish the *index* of the eigenvalue associated with the largest of these subspaces [69].

Definition 1. *Eigenvalue λ 's index ν_λ is the size of the largest Jordan block associated with λ .*

If $z \notin \Lambda_A$, then $\nu_z = 0$. Note that the index of the operator A itself is sometimes discussed [6]. In such contexts, the index of A is ν_0 . Hence, ν_λ corresponds to the index of $A - \lambda I$.

The index of an eigenvalue gives information beyond what the algebraic and geometric multiplicities themselves yield. Nevertheless, for $\lambda \in \Lambda_A$, it is always true that $\nu_\lambda - 1 \leq a_\lambda - g_\lambda \leq a_\lambda - 1$. In the diagonalizable case, $a_\lambda = g_\lambda$ and $\nu_\lambda = 1$ for all $\lambda \in \Lambda_A$.

The following employs basic features of complex analysis extensively in conjunction with linear algebra. Let us therefore review several elementary notions in complex analysis. Recall that a *holomorphic function* is one that is complex differentiable throughout the domain under consideration. A *pole* of order n at z_0 is a singularity that behaves as $h(z)/(z - z_0)^n$ as $z \rightarrow z_0$, where $h(z)$ is holomorphic within a neighborhood of z_0 and $h(z_0) \neq 0$. We say that $h(z)$ has a *zero* of order m at z_1 if $1/h(z)$ has a pole of order m at

z_1 . A *meromorphic function* is one that is holomorphic except possibly at a set of isolated poles within the domain under consideration.

Defined over the continuous complex variable $z \in \mathbb{C}$, A 's *resolvent*:

$$R(z; A) \equiv (zI - A)^{-1} ,$$

captures all of A 's spectral information through the poles of $R(z; A)$'s matrix elements. In fact, the resolvent contains more than just A 's spectrum: we later show that the order of each pole gives the index ν of the corresponding eigenvalue.

The spectrum Λ_A can be expressed in terms of the resolvent. Explicitly, the *point spectrum* (i.e., the set of eigenvalues) is the set of complex values z at which $zI - A$ is not a one-to-one mapping, with the implication that the inverse of $zI - A$ does not exist:

$$\Lambda_A = \{ \lambda \in \mathbb{C} : R(\lambda; A) \neq \text{inv}(\lambda I - A) \} ,$$

where $\text{inv}(\cdot)$ is the inverse of its argument. Later, via our investigation of the Drazin inverse, it should become clear that the resolvent operator can be self-consistently defined at the spectrum, despite the lack of inverse.

For infinite-rank operators, the spectrum becomes more complicated. In that case, the right point spectrum (the point spectrum of A) need not be the same as the left point spectrum (the point spectrum of A 's dual A^\top). Moreover, the spectrum may grow to include non-eigenvalues z for which the range of $zI - A$ is not dense in the vector space it transforms or for which $zI - A$ has dense range but the inverse of $zI - A$ is not bounded. These two settings give rise to the so-called residual spectrum and continuous spectrum, respectively [132]. To mitigate confusion, it should be noted that the point spectrum can be continuous, yet never coincides with the continuous spectrum just described. Moreover, understanding only countable point spectra is necessary to follow the developments here.

Each of A 's eigenvalues λ has an associated *projection operator* A_λ , which is the *residue* of the resolvent as $z \rightarrow \lambda$ [107]. Explicitly:

$$A_\lambda = \text{Res}((zI - A)^{-1}, z \rightarrow \lambda) ,$$

where $\text{Res}(\cdot, z \rightarrow \lambda)$ is the element-wise residue of its first argument as $z \rightarrow \lambda$. The projection operators are orthonormal:

$$A_\lambda A_\zeta = \delta_{\lambda,\zeta} A_\lambda . \quad (2.1)$$

and sum to the identity:

$$I = \sum_{\lambda \in \Lambda_A} A_\lambda . \quad (2.2)$$

The following discusses in detail and then derives several new properties of projection operators.

2.4 Functional Calculi

In the following, we develop an extended *functional calculus* that makes sense of arbitrary functions $f(\cdot)$ of a linear operator A . Within any functional calculus, one considers how A 's eigenvalues map to the eigenvalues of $f(A)$; which we call a *spectral mapping*. For example, it is known that holomorphic functions of bounded linear operators enjoy an especially simple spectral mapping theorem [104]:

$$\Lambda_{f(A)} = f(\Lambda_A) .$$

To fully appreciate the meromorphic functional calculus, we first state and compare the main features and limitations of alternative functional calculi.

2.4.1 Taylor series

Inspired by the Taylor expansion of scalar functions:

$$f(a) = \sum_{n=0}^{\infty} \frac{f^{(n)}(\xi)}{n!} (a - \xi)^n ,$$

a calculus for functions of an operator A can be based on the series:

$$f(A) = \sum_{n=0}^{\infty} \frac{f^{(n)}(\xi)}{n!} (A - \xi I)^n , \quad (2.3)$$

where $f^{(n)}(\xi)$ is the n^{th} derivative of $f(z)$ evaluated at $z = \xi$.

This is often used, for example, to express the exponential of A as:

$$e^A = \sum_{n=0}^{\infty} \frac{A^n}{n!} .$$

This particular series-expansion is convergent for any A since e^z is *entire*, in the sense of complex analysis. Unfortunately, even if it exists there is a limited domain of convergence for most functions. For example, suppose $f(z)$ has poles and choose a Maclaurin series; i.e., $\xi = 0$ in Eq. (2.3). Then the series only converges when A 's spectral radius is less than the radius of the innermost pole of $f(z)$. Addressing this and related issues leads directly to alternative functional calculi.

2.4.2 Holomorphic functional calculus

Holomorphic functions are well behaved, smooth functions that are complex differentiable. Given a function $f(\cdot)$ that is holomorphic within a disk enclosed by a counterclockwise contour C , its Cauchy integral formula is given by:

$$f(a) = \frac{1}{2\pi i} \oint_C f(z) (z - a)^{-1} dz , \quad (2.4)$$

Taking this as inspiration, the holomorphic functional calculus performs a contour integration of the resolvent to extend $f(\cdot)$ to operators:

$$f(A) = \frac{1}{2\pi i} \oint_{C_{\Lambda_A}} f(z) (zI - A)^{-1} dz , \quad (2.5)$$

where C_{Λ_A} is a closed counterclockwise contour that encompasses Λ_A . Assuming that $f(z)$ is holomorphic at $z = \lambda$ for all $\lambda \in \Lambda_A$, a nontrivial calculation [69] shows that Eq. (2.5) is equivalent to the holomorphic calculus defined by:

$$f(A) = \sum_{\lambda \in \Lambda_A} \sum_{m=0}^{\nu_{\lambda}-1} \frac{f^{(m)}(\lambda)}{m!} (A - \lambda I)^m A_{\lambda} . \quad (2.6)$$

After some necessary development, we will later derive Eq. (2.6) as a special case of our meromorphic functional calculus, such that Eq. (2.6) is valid whenever $f(z)$ is holomorphic at $z = \lambda$ for all $\lambda \in \Lambda_A$.

The holomorphic functional calculus was first proposed in Ref. [69] and is now in wide use; e.g., see Ref. [169, p. 603]. It agrees with the Taylor-series approach whenever

the infinite series converges, but gives an functional calculus when the series approach fails. For example, using the principal branch of the complex logarithm, the holomorphic functional calculus admits $\log(A)$ for any nonsingular matrix, with the satisfying result that $e^{\log(A)} = A$. Whereas, the Taylor series approach fails to converge for the logarithm of most matrices even if the expansion for, say, $\log(1 - z)$ is used.

The major shortcoming of the holomorphic functional calculus is that it assumes $f(z)$ is holomorphic at Λ_A . Clearly, if $f(z)$ has a pole at some $z \in \Lambda_A$, then Eq. (2.6) fails. An example of such a failure is the negative-one power of a singular operator, which we take up later on.

Several efforts have been made to extend the holomorphic functional calculus. For example, Refs. [98] and [175] define a functional calculus that extends the standard holomorphic functional calculus to include a certain class of meromorphic functions that are nevertheless still required to be *holomorphic on the point spectrum* (i.e., on the eigenvalues) of the operator. However, we are not aware of any previous work that introduces and develops the consequences of a functional calculus for functions that are meromorphic on the point spectrum—which we take up in the next few sections.

2.4.3 Meromorphic functional calculus

Meromorphic functions are holomorphic except at a set of isolated poles of the function. The resolvent of a finite-dimensional operator is meromorphic, since it is holomorphic everywhere except for poles at the eigenvalues of the operator. We will now also allow our function $f(z)$ to be meromorphic with possible poles that coincide with the poles of the resolvent.

Inspired again by the Cauchy integral formula of Eq. (2.4), but removing the restriction to holomorphic functions, our meromorphic functional calculus instead employs a partitioned contour integration of the resolvent:

$$f(A) = \sum_{\lambda \in \Lambda_A} \frac{1}{2\pi i} \oint_{C_\lambda} f(z) R(z; A) dz ,$$

where C_λ is a small counterclockwise contour around the eigenvalue λ . This and a spectral decomposition of the resolvent (to be derived later) extends the holomorphic calculus to a

much wider domain, defining:

$$f(A) = \sum_{\lambda \in \Lambda_A} \sum_{m=0}^{\nu_\lambda-1} A_\lambda (A - \lambda I)^m \frac{1}{2\pi i} \oint_{C_\lambda} \frac{f(z)}{(z - \lambda)^{m+1}} dz . \quad (2.7)$$

The contour is integrated using knowledge of $f(z)$ since meromorphic $f(z)$ can introduce poles and zeros at Λ_A that interact with the resolvent's poles.

The meromorphic functional calculus agrees with the Taylor-series approach whenever the series converges and agrees with the holomorphic functional calculus whenever $f(z)$ is holomorphic at Λ_A . However, when both the previous functional calculi fail, the meromorphic calculus extends the domain of $f(A)$ to yield surprising, yet sensible answers. For example, we show that within it, the negative-one power of a singular operator is the Drazin inverse—an operator that effectively inverts everything that is invertible.

The major assumption of our meromorphic functional calculus is that the domain of operators must have a spectrum that is at most countably infinite—e.g., A can be any compact operator. A related limitation is that singularities of $f(z)$ that coincide with Λ_A must be isolated singularities. Nevertheless, we expect that these restrictions can be lifted with proper treatment, as discussed in fuller context later.

2.5 Meromorphic Spectral Decomposition

The preceding gave an overview of the relationship between alternative functional calculi and their trade-offs, highlighting the advantages of the meromorphic functional calculus. This section leverages these advantages and employs a partial fraction expansion of the resolvent to give a general spectral decomposition of almost any function of any operator. Then, since it plays a key role in applications, we apply the functional calculus to investigate the negative-one power of singular operators—thus *deriving*, what is otherwise an operator defined axiomatically, the Drazin inverse from first principles.

2.5.1 Partial fraction expansion of the resolvent

The elements of A 's resolvent are proper rational functions that contain all of A 's spectral information. (Recall that a *proper rational function* $r(z)$ is a ratio of polynomials in z whose numerator has degree strictly less than the degree of the denominator.) In particular,

the resolvent's poles coincide with A 's eigenvalues since, for $z \notin \Lambda_A$:

$$\begin{aligned} R(z; A) &= (zI - A)^{-1} \\ &= \frac{\mathcal{C}^\top}{\det(zI - A)} \\ &= \frac{\mathcal{C}^\top}{\prod_{\lambda \in \Lambda_A} (z - \lambda)^{a_\lambda}} , \end{aligned} \tag{2.8}$$

where a_λ is the algebraic multiplicity of eigenvalue λ and \mathcal{C} is the matrix of *cofactors* of $zI - A$. That is, \mathcal{C} 's transpose \mathcal{C}^\top is the *adjugate* of $zI - A$:

$$\mathcal{C}^\top = \text{adj}(zI - A) ,$$

whose elements will be polynomial functions of z of degree less than $\sum_{\lambda \in \Lambda_A} a_\lambda$.

Recall that the partial fraction expansion of a proper rational function $r(z)$ with poles in Λ allows a unique decomposition into a sum of constant numerators divided by monomials in $z - \lambda$ up to degree a_λ , when a_λ is the order of the pole of $r(z)$ at $\lambda \in \Lambda$ [140]. Equation (2.8) thus makes it clear that the resolvent has the unique partial fraction expansion:

$$R(z; A) = \sum_{\lambda \in \Lambda_A} \sum_{m=0}^{a_\lambda-1} \frac{1}{(z - \lambda)^{m+1}} A_{\lambda,m} , \tag{2.9}$$

where $\{A_{\lambda,m}\}$ is the set of matrices with constant entries (*not* functions of z) uniquely determined elementwise by the partial fraction expansion. However, $R(z; A)$'s poles are *not* necessarily of the same order as the algebraic multiplicity of the corresponding eigenvalues since the entries of \mathcal{C} , and thus of \mathcal{C}^\top , may have zeros at A 's eigenvalues. This has the potential to render $A_{\lambda,m}$ equal to the zero matrix $\mathbf{0}$.

The Cauchy integral formula indicates that the constant matrices $\{A_{\lambda,m}\}$ of Eq. (2.9) can be obtained by the residues:

$$A_{\lambda,m} = \frac{1}{2\pi i} \oint_{C_\lambda} (z - \lambda)^m R(z; A) dz , \tag{2.10}$$

where the residues are calculated elementwise. The projection operators A_λ associated with each eigenvalue λ were already referenced in §2.3, but can now be properly introduced

as the $A_{\lambda,0}$ matrices:

$$A_\lambda = A_{\lambda,0} \quad (2.11)$$

$$= \frac{1}{2\pi i} \oint_{C_\lambda} R(z; A) dz . \quad (2.12)$$

Since $R(z; A)$'s elements are rational functions, as we just showed, it is analytic except at a finite number of isolated singularities—at A 's eigenvalues. In light of the residue theorem, this motivates the Cauchy-integral-like formula that serves as the starting point for the meromorphic functional calculus:

$$f(A) = \sum_{\lambda \in \Lambda_A} \frac{1}{2\pi i} \oint_{C_\lambda} f(z) R(z; A) dz . \quad (2.13)$$

Let's now consider several immediate consequences.

2.5.2 Decomposing the identity

Even the simplest applications of Eq. (2.13) yield insight. Consider the identity as the operator function $f(A) = A^0 = I$ that corresponds to the scalar function $f(z) = z^0 = 1$. Then, Eq. (2.13) implies:

$$\begin{aligned} I &= \sum_{\lambda \in \Lambda_A} \frac{1}{2\pi i} \oint_{C_\lambda} R(z; A) dz \\ &= \sum_{\lambda \in \Lambda_A} A_\lambda . \end{aligned}$$

This shows that the projection operators are, in fact, a decomposition of the identity, as anticipated in Eq. (2.2).

2.5.3 Dunford decomposition, decomposed

For $f(A) = A$, Eqs. (2.13) and (2.10) imply that:

$$\begin{aligned} A &= \sum_{\lambda \in \Lambda_A} \frac{1}{2\pi i} \oint_{C_\lambda} z R(z; A) dz \\ &= \sum_{\lambda \in \Lambda_A} \left[\lambda \frac{1}{2\pi i} \oint_{C_\lambda} R(z; A) dz + \frac{1}{2\pi i} \oint_{C_\lambda} (z - \lambda) R(z; A) dz \right] \\ &= \sum_{\lambda \in \Lambda_A} (\lambda A_{\lambda,0} + A_{\lambda,1}) . \end{aligned} \quad (2.14)$$

We denote the important set of nilpotent matrices $A_{\lambda,1}$ that project onto the generalized eigenspaces by relabeling them:

$$N_\lambda \equiv A_{\lambda,1} \quad (2.15)$$

$$= \frac{1}{2\pi i} \oint_{C_\lambda} (z - \lambda) R(z; A) dz . \quad (2.16)$$

Equation (2.14) is the unique *Dunford decomposition* [70]: $A = D + N$, where $D \equiv \sum_{\lambda \in \Lambda_A} \lambda A_\lambda$ is diagonalizable, $N \equiv \sum_{\lambda \in \Lambda_A} N_\lambda$ is nilpotent, and D and N commute: $[D, N] = \mathbf{0}$. This is also known as the *Jordan–Chevalley decomposition*.

The special case where A is diagonalizable implies that $N = \mathbf{0}$. And so, Eq. (2.14) simplifies to:

$$A = \sum_{\lambda \in \Lambda_A} \lambda A_\lambda .$$

2.5.4 The resolvent, resolved

As shown in Ref. [107] and can be derived from Eqs. (2.12) and (2.16):

$$A_\lambda A_\zeta = \delta_{\lambda,\zeta} A_\lambda \text{ and}$$

$$A_\lambda N_\zeta = \delta_{\lambda,\zeta} N_\lambda .$$

Due to these, our spectral decomposition of the Dunford decomposition implies that:

$$\begin{aligned} N_\lambda &= A_\lambda \left(A - \sum_{\zeta \in \Lambda_A} \zeta A_\zeta \right) \\ &= A_\lambda (A - \lambda A_\lambda) \\ &= A_\lambda (A - \lambda I) . \end{aligned} \quad (2.17)$$

Moreover:

$$A_{\lambda,m} = A_\lambda (A - \lambda I)^m . \quad (2.18)$$

It turns out that for $m > 0$: $A_{\lambda,m} = N_\lambda^m$. (See also Ref. [107, p. 483].) This leads to a generalization of the projection operator orthonormality relations of Eq. (2.1). Most generally, the operators of $\{A_{\lambda,m}\}$ are mutually related by:

$$A_{\lambda,m} A_{\zeta,n} = \delta_{\lambda,\zeta} A_{\lambda,m+n} . \quad (2.19)$$

Finally, if we recall that the index ν_λ is the dimension of the largest associated subspace, we find that the index of λ characterizes the nilpotency of N_λ : $N_\lambda^m = \mathbf{0}$ for $m \geq \nu_\lambda$. That is:

$$A_{\lambda,m} = \mathbf{0} \quad \text{for } m \geq \nu_\lambda . \quad (2.20)$$

Returning to Eq. (2.9), we see that all $A_{\lambda,m}$ with $m \geq \nu_\lambda$ are zero-matrices and so do not contribute to the sum. Thus, we can rewrite Eq. (2.9) as:

$$R(z; A) = \sum_{\lambda \in \Lambda_A} \sum_{m=0}^{\nu_\lambda-1} \frac{1}{(z - \lambda)^{m+1}} A_{\lambda,m} \quad (2.21)$$

or:

$$R(z; A) = \sum_{\lambda \in \Lambda_A} \sum_{m=0}^{\nu_\lambda-1} \frac{1}{(z - \lambda)^{m+1}} A_\lambda (A - \lambda I)^m , \quad (2.22)$$

for $z \notin \Lambda_A$.

The following sections sometimes use $A_{\lambda,m}$ in place of $A_\lambda (A - \lambda I)^m$. This is helpful both for conciseness and when applying Eq. (2.19). Nonetheless, the equality in Eq. (2.18) is a useful one to keep in mind.

2.5.5 Meromorphic functional calculus

In light of Eq. (2.13), Eq. (2.21) together with Eq. (2.18) allow us to express any function of an operator simply and solely in terms of its spectrum (i.e., its eigenvalues for the finite dimensional case), its projection operators, and itself:

$$f(A) = \sum_{\lambda \in \Lambda_A} \sum_{m=0}^{\nu_\lambda-1} A_{\lambda,m} \frac{1}{2\pi i} \oint_{C_\lambda} \frac{f(z)}{(z - \lambda)^{m+1}} dz . \quad (2.23)$$

In obtaining Eq. (2.23) we finally derived Eq. (2.7), as promised earlier in § 2.4.3. Effectively, by modulating the modes associated with the resolvent's singularities, the scalar function $f(\cdot)$ is mapped to the operator domain, where its action is expressed in each of A 's independent subspaces.

2.5.6 Evaluating the residues

Interpretation aside, how does one use this result? Equation (2.23) says that the spectral decomposition of $f(A)$ reduces to the evaluation of several residues, where:

$$\text{Res}(g(z), z \rightarrow \lambda) = \frac{1}{2\pi i} \oint_{C_\lambda} g(z) dz .$$

So, to make progress with Eq. (2.23), we must evaluate function-dependent residues of the form:

$$\text{Res}(f(z)/(z - \lambda)^{m+1}, z \rightarrow \lambda) .$$

If $f(z)$ were holomorphic at each λ , then the order of the pole would simply be the power of the denominator. We could then use Cauchy's differential formula for holomorphic functions:

$$f^{(n)}(a) = \frac{n!}{2\pi i} \oint_{C_a} \frac{f(z)}{(z - a)^{n+1}} dz , \quad (2.24)$$

for $f(z)$ holomorphic at a . And, the meromorphic calculus would reduce to the holomorphic calculus. Often, $f(z)$ will be holomorphic at least at *some* of A 's eigenvalues. And so, Eq. (2.24) is still locally a useful simplification in those special cases.

In general, though, $f(z)$ introduces poles and zeros at $\lambda \in \Lambda_A$ that change their orders. This is exactly the impetus for the generalized functional calculus. The residue of a complex-valued function $g(z)$ around its isolated pole λ of order $n + 1$ can be calculated from:

$$\text{Res}(g(z), z \rightarrow \lambda) = \frac{1}{n!} \lim_{z \rightarrow \lambda} \frac{d^n}{dz^n} [(z - \lambda)^{n+1} g(z)] .$$

2.5.7 Decomposing A^L

Equation (2.23) says that we can explicitly derive the spectral decomposition of powers of the operator A . Of course, we already did this for the special cases of A^0 and A^1 . The goal, though, is to do this in general.

For $f(A) = A^L \rightarrow f(z) = z^L$, $z = 0$ can be either a zero or a pole of $f(z)$, depending on the value of L . In either case, an eigenvalue of $\lambda = 0$ will distinguish itself in the residue

calculation of A^L via its unique ability to change the order of the pole (or zero) at $z = 0$. For example, at this special value of λ and for integer $L > 0$, $\lambda = 0$ induces poles that *cancel* with the zeros of $f(z) = z^L$, since z^L has a zero at $z = 0$ of order L . For integer $L < 0$, an eigenvalue of $\lambda = 0$ *increases* the order of the $z = 0$ pole of $f(z) = z^L$. For all other eigenvalues, the residues will be as expected. Hence, from Eq. (2.23) and inserting $f(z) = z^L$, for any $L \in \mathbb{C}$:

$$\begin{aligned}
A^L &= \left[\sum_{\substack{\lambda \in \Lambda_A \\ \lambda \neq 0}} \sum_{m=0}^{\nu_\lambda-1} A_\lambda (A - \lambda I)^m \overbrace{\left(\frac{1}{2\pi i} \oint_{C_\lambda} \frac{z^L}{(z - \lambda)^{m+1}} dz \right)}^{= \frac{1}{m!} \lim_{z \rightarrow \lambda} \frac{d^m}{dz^m} z^L = \frac{\lambda^{L-m}}{m!} \prod_{n=1}^m (L - n + 1)} \right] \\
&\quad + [0 \in \Lambda_A] \sum_{m=0}^{\nu_0-1} A_0 A^m \underbrace{\left(\frac{1}{2\pi i} \oint_{C_0} z^{L-m-1} dz \right)}_{= \delta_{L,m}} \\
&= \left[\sum_{\substack{\lambda \in \Lambda_A \\ \lambda \neq 0}} \sum_{m=0}^{\nu_\lambda-1} \binom{L}{m} \lambda^{L-m} A_\lambda (A - \lambda I)^m \right] + [0 \in \Lambda_A] \sum_{m=0}^{\nu_0-1} \delta_{L,m} A_0 A^m, \quad (2.25)
\end{aligned}$$

where $\binom{L}{m}$ is the generalized binomial coefficient:

$$\binom{L}{m} = \frac{1}{m!} \prod_{n=1}^m (L - n + 1) \quad \text{with} \quad \binom{L}{0} = 1, \quad (2.26)$$

and $[0 \in \Lambda_A]$ is the Iverson bracket which takes on value 1 if zero is an eigenvalue of A and 0 if not. $A_{\lambda,m}$ was replaced by $A_\lambda (A - \lambda I)^m$ to suggest the more explicit calculations involved with evaluating any A^L . Equation (2.25) applies to any linear operator with only isolated singularities in its resolvent.

If L is a nonnegative integer such that $L \geq \nu_\lambda - 1$ for all $\lambda \in \Lambda_A$, then:

$$A^L = \sum_{\substack{\lambda \in \Lambda_A \\ \lambda \neq 0}} \sum_{m=0}^{\nu_\lambda-1} \binom{L}{m} \lambda^{L-m} A_{\lambda,m}, \quad (2.27)$$

where $\binom{L}{m}$ is now reduced to the traditional binomial coefficient $L!/(m!(L-m)!)$.

2.5.8 Drazin inverse

If L is any negative integer, then $\binom{-|L|}{m}$ can be written as a traditional binomial coefficient $(-1)^m \binom{|L|+m-1}{m}$, yielding:

$$A^{-|L|} = \sum_{\substack{\lambda \in \Lambda_A \\ \lambda \neq 0}} \sum_{m=0}^{\nu_\lambda-1} (-1)^m \binom{|L|+m-1}{m} \lambda^{-|L|-m} A_{\lambda,m} , \quad (2.28)$$

for $-|L| \in \{-1, -2, -3, \dots\}$.

Thus, negative powers of an operator can be consistently defined even for noninvertible operators. In light of Eqs. (2.25) and (2.28), it appears that the zero eigenvalue does not even contribute to the function. It is well known, in contrast, that it wrecks havoc on the naive, oft-quoted definition of a matrix's negative power:

$$A^{-1} \stackrel{?}{=} \frac{\text{adj}(A)}{\det(A)} = \frac{\text{adj}(A)}{\prod_{\lambda \in \Lambda_A} \lambda^{a_\lambda}} ,$$

since this would imply dividing by zero. If we can accept large positive powers of singular matrices—for which the zero eigenvalue does not contribute—it seems fair to also accept negative powers that likewise involve no contribution from the zero eigenvalue.

Editorializing aside, we note that extending the definition of A^{-1} to the domain including singular operators via Eqs. (2.25) and (2.28) implies that:

$$\begin{aligned} A^{|L|} A^{-|\ell|} &= A^{-|\ell|} A^{|L|} \\ &= A^{|L|-|\ell|} \quad \text{for } |L| \geq |\ell| + \nu_0 , \end{aligned}$$

which is a very sensible and desirable condition. Moreover, we find that $AA^{-1} = I - A_0$.

Specifically, the negative-one power of any square matrix is in general *not* the same as the matrix inverse since $\text{inv}(A)$ need not exist. However, it is consistently defined via Eq. (2.28) to be:

$$A^{-1} = \sum_{\lambda \in \Lambda_A \setminus \{0\}} \sum_{m=0}^{\nu_\lambda-1} (-1)^m \lambda^{-1-m} A_{\lambda,m} . \quad (2.29)$$

This is the *Drazin inverse* $A^{\mathcal{D}}$ of A . Note that it is *not* the same as the Moore–Penrose pseudo-inverse [172, 191].

Although the Drazin inverse is usually defined axiomatically to satisfy certain criteria [19], it is naturally *derived* as the negative one power of a singular operator in the meromorphic functional calculus. We can check that it indeed satisfies the axiomatic criteria for the Drazin inverse, enumerated according to historical precedent:

$$\begin{aligned} (1^{\nu_0}) \quad & A^{\nu_0} A^{\mathcal{D}} A = A^{\nu_0} \\ (2) \quad & A^{\mathcal{D}} A A^{\mathcal{D}} = A^{\mathcal{D}} \\ (5) \quad & [A, A^{\mathcal{D}}] = 0, \end{aligned}$$

which gives rise to the Drazin inverse's moniker as the $\{1^{\nu_0}, 2, 5\}$ -inverse [19].

While A^{-1} always exists, the resolvent is nonanalytic at $z = 0$ for a singular matrix. Effectively, the meromorphic functional calculus removes the nonanalyticity of the resolvent in evaluating A^{-1} . As a result, as we can see from Eq. (2.29), the Drazin inverse inverts what is invertible; the remainder is zeroed out.

Of course, whenever A is invertible, A^{-1} is equal to $\text{inv}(A)$. However, we should not confuse this coincidence with equivalence. Moreover, despite historic notation there is no reason that the negative-one power should in general be equivalent to the inverse. Especially, if an operator is not invertible! To avoid confusing A^{-1} with $\text{inv}(A)$, we use the notation $A^{\mathcal{D}}$ for the Drazin inverse of A . Still, $A^{\mathcal{D}} = \text{inv}(A)$, whenever $0 \notin \Lambda_A$.

Amusingly, this extension of previous calculi lets us resolve an elementary but fundamental question: What is 0^{-1} ? It is certainly not infinity. Indeed, it is just as close to negative infinity! Rather: $0^{-1} = 0 \neq \text{inv}(0)$.

Although Eq. (2.29) is a constructive way to build the Drazin inverse, it imposes more work than is actually necessary. Using the meromorphic functional calculus, we can derive a new, simple construction of the Drazin inverse that requires only the original operator and the eigenvalue-0 projector.

First, assume that λ is an isolated singularity of $R(z; A)$ with finite separation at least ϵ distance from the nearest neighboring singularity. And, consider the operator-valued

function f_λ^ϵ defined via the RHS of:

$$\begin{aligned} A_\lambda &= f_\lambda^\epsilon(A) \\ &= \frac{1}{2\pi i} \oint_{\lambda + \epsilon e^{i\phi}} (\zeta I - A)^{-1} d\zeta , \end{aligned}$$

with $\lambda + \epsilon e^{i\phi}$ defining an ϵ -radius circular contour around λ . Then we see that:

$$\begin{aligned} f_\lambda^\epsilon(z) &= \frac{1}{2\pi i} \oint_{\lambda + \epsilon e^{i\phi}} (\zeta - z)^{-1} d\zeta \\ &= [z \in \mathbb{C} : |z - \lambda| < \epsilon] , \end{aligned} \tag{2.30}$$

where $[z \in \mathbb{C} : |z - \lambda| < \epsilon]$ is the Iverson bracket that takes on value 1 if z is within ϵ -distance of λ and 0 if not.

Second, we use this to find that, for any $c \in \mathbb{C} \setminus \{0\}$:

$$\begin{aligned} (A + cA_0)^{-1} &= \sum_{\lambda \in \Lambda_A} \sum_{m=0}^{\nu_\lambda-1} A_{\lambda,m} \frac{1}{2\pi i} \oint_{C_\lambda} \frac{(z + cf_0^\epsilon(z))^{-1}}{(z - \lambda)^{m+1}} dz \\ &= A^\mathcal{D} + \sum_{m=0}^{\nu_0-1} A_0 A^m \frac{1}{2\pi i} \oint_{C_0} \frac{(z + c)^{-1}}{z^{m+1}} \\ &= A^\mathcal{D} + \sum_{m=0}^{\nu_0-1} A_0 A^m (-1)^m / c^{m+1} , \end{aligned} \tag{2.31}$$

where we asserted that the contour C_0 exists within the finite ϵ -ball about the origin.

Third, we note that $A + cA_0$ is invertible for all $c \neq 0$; this can be proven by multiplying each side of Eq. (2.31) by $A + cA_0$. Hence, $(A + cA_0)^{-1} = \text{inv}(A + cA_0)$ for all $c \neq 0$.

Finally, multiplying each side of Eq. (2.31) by $I - A_0$, and recalling that $A_{0,0}A_{0,m} = A_{0,m}$, we find a useful expression for calculating the Drazin inverse of any linear operator A , given only A and A_0 . Specifically:

$$A^\mathcal{D} = (I - A_0)(A + cA_0)^{-1} . \tag{2.32}$$

which is valid for any $c \in \mathbb{C} \setminus \{0\}$. Eq. (2.32) generalizes the result found specifically for $c = -1$ in Ref. [207].

For the special case of $c = -1$, it is worthwhile to also consider the alternative construction of the Drazin inverse implied by Eq. (2.31):

$$A^\mathcal{D} = (A - A_0)^{-1} + A_0 \left(\sum_{m=0}^{\nu_0-1} A^m \right) . \tag{2.33}$$

By a spectral mapping ($\lambda \rightarrow 1 - \lambda$, for $\lambda \in \Lambda_T$), the Perron–Frobenius theorem and Eq. (2.31) yield an important consequence for any stochastic matrix T . The Perron–Frobenius theorem guarantees that T ’s eigenvalues along the unit circle are associated with a diagonalizable subspace. In particular, $\nu_1 = 1$. Spectral mapping of this result means that T ’s eigenvalue 1 maps to the eigenvalue 0 of $I - T$ and $T_1 = (I - T)_0$. Moreover:

$$[(I - T) + T_1]^{-1} = (I - T)^{\mathcal{D}} + T_1 ,$$

since $\nu_0 = 1$. This corollary of Eq. (2.31) (with $c = 1$) corresponds to a number of important and well known results in the theory of Markov processes. Indeed, $Z \equiv (I - T + T_1)^{-1}$ is called the *fundamental matrix* in that setting [129].

2.5.9 Consequences and generalizations

For an infinite-rank operator A with a continuous spectrum, the meromorphic functional calculus has the natural generalization:

$$f(A) = \frac{1}{2\pi i} \oint_{C_{\Lambda_A}} f(z)(zI - A)^{-1} dz , \quad (2.34)$$

where the contour C_{Λ_A} encloses the (possibly continuous) spectrum of A without including any unbounded contributions from $f(z)$ outside of C_{Λ_A} . The function $f(z)$ is expected to be meromorphic within C_{Λ_A} . This again deviates from the holomorphic approach, since the holomorphic functional calculus requires that $f(z)$ is analytic in a neighborhood around the spectrum; see § VII of Ref. [71]. Moreover, Eq. (2.34) allows an extension of the functional calculus of Refs. [98, 175, 22], since the function can be meromorphic at the point spectrum in addition being meromorphic on the residual and continuous spectra.

In either the finite- or infinite-rank case, whenever $f(z)$ is analytic in a neighborhood around the spectrum, the meromorphic functional calculus agrees with the holomorphic. Whenever $f(z)$ is *not* analytic in a neighborhood around the spectrum, the function is undefined in the holomorphic approach. In contrast, the meromorphic approach extends the function to the operator-valued domain and does so with novel consequences.

In particular, when $f(z)$ is *not analytic in a neighborhood around the spectrum*—say $f(z)$ is nonanalytic within A ’s spectrum at $\Xi_f \subset \Lambda_A$ —then we expect to lose both homomorphism and spectral mapping properties:

- Loss of homomorphism: $f_1(A)f_2(A) \neq (f_1 \cdot f_2)(A)$;
- Loss of naive spectral mapping: $f(\Lambda_A \setminus \Xi_f) \subset \Lambda_{f(A)}$.

A simple example of both losses arises with the Drazin inverse, above. There, $f_1(z) = z^{-1}$. Taking this and $f_2(z) = z$ combined with singular operator A leads to the loss of homomorphism: $A^{\mathcal{D}}A \neq I$. As for the second property, the spectral mapping can be altered for the candidate spectra at Ξ_f via pole–pole or pole–zero interactions in the complex contour integral. For $f(A) = A^{-1}$, how does A ’s eigenvalue of 0 get mapped into the new spectrum of $A^{\mathcal{D}}$? A naive application of the spectral mapping theorem might seem to yield an undefined quantity. But, using the meromorphic functional calculus self-consistently maps the eigenvalue as $0^{-1} = 0$. It remains to be explored whether the full spectral mapping is preserved for any function $f(A)$ under the meromorphic interpretation of $f(\lambda)$.

It should now be apparent that extending functions via the meromorphic functional calculus allows one to express novel mathematical properties, some likely capable of describing new physical phenomena. At the same time, extra care is necessary. The situation is reminiscent of the loss of commutativity in non-Abelian operator algebra: not all of the old rules apply, but the gain in nuance allows for mathematical description of important phenomena.

We chose to focus primarily on the finite-rank case here since it is sufficient to demonstrate the utility of the general projection-operator formalism. Indeed, there are ample nontrivial applications in the finite-rank setting that deserve attention. To appreciate these, we now turn to address the construction and properties of general eigenprojectors.

2.6 Constructing Decompositions

At this point, we see that projection operators are fundamental to functions of an operator. This prompts the practical question of how to actually calculate them. The next several sections address this by deriving expressions with both theoretical and applied use. We first address the projection operators associated with index-one eigenvalues. We then explicate the relationship between eigenvectors, generalized eigenvectors, and projection

operators for normal, diagonalizable, and general matrices. Finally, we address how the general results specialize in several common cases of interest. After these, we turn to examples and applications.

2.6.1 Projection operators of index-one eigenvalues

To obtain the projection operators associated with each index-one eigenvalue $\lambda \in \{\zeta \in \Lambda_A : \nu_\zeta = 1\}$, we apply the meromorphic calculus to an appropriately chosen function of A , finding:

$$\begin{aligned} \prod_{\substack{\zeta \in \Lambda_A \\ \zeta \neq \lambda}} (A - \zeta I)^{\nu_\zeta} &= \sum_{\xi \in \Lambda_A} \sum_{m=0}^{\nu_\xi-1} \frac{A_{\xi,m}}{2\pi i} \oint_{C_\xi} \frac{\prod_{\substack{\zeta \in \Lambda_A \\ \zeta \neq \lambda}} (z - \zeta)^{\nu_\zeta}}{(z - \xi)^{m+1}} dz \\ &= A_\lambda \frac{1}{2\pi i} \oint_{C_\lambda} \frac{\prod_{\substack{\zeta \in \Lambda_A \\ \zeta \neq \lambda}} (z - \zeta)^{\nu_\zeta}}{z - \lambda} dz \\ &= A_\lambda \prod_{\substack{\zeta \in \Lambda_A \\ \zeta \neq \lambda}} (\lambda - \zeta)^{\nu_\zeta} . \end{aligned}$$

Therefore, if $\nu_\lambda = 1$:

$$A_\lambda = \prod_{\substack{\zeta \in \Lambda_A \\ \zeta \neq \lambda}} \left(\frac{A - \zeta I}{\lambda - \zeta} \right)^{\nu_\zeta} . \quad (2.35)$$

As convenience dictates in our computations, we let $\nu_\zeta \rightarrow a_\zeta - g_\zeta + 1$ or even $\nu_\zeta \rightarrow a_\zeta$ in Eq. (2.35), since multiplying A_λ by $(A - \zeta I)/(\lambda - \zeta)$ has no effect for $\zeta \in \Lambda_A \setminus \{\lambda\}$ if $\nu_\lambda = 1$.

Equation (2.35) generalizes a well known result that applies when the index of *all* eigenvalues is one. That is, when the operator is diagonalizable, we have:

$$A_\lambda = \prod_{\substack{\zeta \in \Lambda_A \\ \zeta \neq \lambda}} \frac{A - \zeta I}{\lambda - \zeta} .$$

To the best of our knowledge, Eq. (2.35) is original.

Since eigenvalues can have index larger than one, not all projection operators of a nondiagonalizable operator can be found directly from Eq. (2.35). Even so, it serves three useful purposes. First, it gives a practical reduction of the eigen-analysis by finding all projection operators of index-one eigenvalues. Second, if there is only one eigenvalue that

has index larger than one—what we call the *almost diagonalizable case*—then Eq. (2.35), together with the fact that the projection operators must sum to the identity, *does* give a full solution to the set of projection operators. Third, Eq. (2.35) is a powerful theoretical tool that we can use directly to spectrally decompose functions, for example, of a stochastic matrix whose eigenvalues on the unit circle are guaranteed to be index-one by the Perron–Frobenius theorem.

Although index-one expressions have some utility, we need a more general procedure to obtain all projection operators of any linear operator. Recall that, with full generality, projection operators can also be calculated directly via residues, as in Eq. (2.12).

An alternative procedure—one that extends a method familiar at least in quantum mechanics—is to obtain the projection operators via eigenvectors. However, quantum mechanics always concerns itself with a subset of diagonalizable operators. What is the necessary generalization? For one, left and right eigenvectors are no longer simply conjugate transposes of each other. More severely, a full set of spanning eigenvectors is no longer guaranteed and we must resort to *generalized* eigenvectors. Since the relationships among eigenvectors, generalized eigenvectors, and projection operators are critical to the practical calculation of many physical observables of complex systems, we collect these results in the next section.

2.6.2 Eigenvectors, generalized eigenvectors, and projection operators

Two common questions regarding projection operators are: Why not just use eigenvectors? And, why not use the Jordan canonical form? First, the eigenvectors of a defective matrix do not form a complete basis with which to expand an arbitrary vector. One needs generalized eigenvectors for this. Second, some functions of an operator require removing, or otherwise altering, the contribution from select eigenspaces. This is most adroitly handled with the projection operator formalism where different eigenspaces (correlates of Jordan blocks) can effectively be treated separately. Moreover, even for simple cases where eigenvectors suffice, the projection operator formalism simply can be more calculationally or mathematically convenient.

That said, it is useful to understand the relationship between projection operators and generalized eigenvectors. For example, it is often useful to create projection operators from generalized eigenvectors. This section clarifies their connection using the language of matrices. In the most general case, we show that the projection operator formalism is usefully concise.

2.6.2.1 Normal matrices

Unitary, Hermitian, skew-Hermitian, orthogonal, symmetric, and skew-symmetric matrices are all special cases of normal matrices. As noted, normal matrices are those that commute with their Hermitian adjoint (complex-conjugate transpose): $AA^\dagger = A^\dagger A$. Moreover, a matrix is normal if and only if it can be diagonalized by a unitary transformation: $A = U\Lambda U^\dagger$, where the columns of the unitary matrix U are the orthonormal right eigenvectors of A corresponding to the eigenvalues ordered along the diagonal matrix Λ . For an M -by- M matrix A , the eigenvalues in Λ_A are ordered and enumerated according to the possibly degenerate M -tuple $(\Lambda_A) = (\lambda_1, \dots, \lambda_M)$. Since an eigenvalue $\lambda \in \Lambda_A$ has algebraic multiplicity $a_\lambda \geq 1$, λ appears a_λ times in the ordered tuple.

Assuming A is normal, each projection operator A_λ can be constructed as the sum of all ket–bra pairs of right-eigenvectors corresponding to λ composed with their conjugate transpose. We later introduce bras and kets more generally via generalized eigenvectors of the operator A and its dual A^\top . However, since the complex-conjugate transposition rule between dual spaces is only applicable to a ket basis derived from a normal operator, we put off using the bra-ket notation for now so as not to confuse the more familiar “normal” case with the general case.

To explicitly demonstrate this relationship between projection operators, eigenvectors,

and their Hermitian adjoints in the case of normality, observe that:

$$\begin{aligned}
A &= U \Lambda U^\dagger \\
&= \begin{bmatrix} \vec{u}_1 & \vec{u}_2 & \cdots & \vec{u}_M \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_M \end{bmatrix} \begin{bmatrix} \vec{u}_1^\dagger \\ \vec{u}_2^\dagger \\ \vdots \\ \vec{u}_M^\dagger \end{bmatrix} \\
&= \begin{bmatrix} \lambda_1 \vec{u}_1 & \lambda_2 \vec{u}_2 & \cdots & \lambda_M \vec{u}_M \end{bmatrix} \begin{bmatrix} \vec{u}_1^\dagger \\ \vec{u}_2^\dagger \\ \vdots \\ \vec{u}_M^\dagger \end{bmatrix} \\
&= \sum_{j=1}^M \lambda_j \vec{u}_j \vec{u}_j^\dagger \\
&= \sum_{\lambda \in \Lambda_A} \lambda A_\lambda .
\end{aligned}$$

Evidently, for normal matrices A :

$$A_\lambda = \sum_{j=1}^M \delta_{\lambda, \lambda_j} \vec{u}_j \vec{u}_j^\dagger .$$

And, since $\vec{u}_i^\dagger \vec{u}_j = \delta_{i,j}$, we have an orthogonal set $\{A_\lambda\}_{\lambda \in \Lambda_A}$ with the property that:

$$\begin{aligned}
A_\zeta A_\lambda &= \sum_{i=1}^M \sum_{j=1}^M \delta_{\zeta, \lambda_i} \delta_{\lambda, \lambda_j} \vec{u}_i \vec{u}_i^\dagger \vec{u}_j \vec{u}_j^\dagger \\
&= \sum_{i=1}^M \sum_{j=1}^M \delta_{\zeta, \lambda_i} \delta_{\lambda, \lambda_j} \vec{u}_i \delta_{i,j} \vec{u}_j^\dagger \\
&= \sum_{i=1}^M \delta_{\zeta, \lambda_i} \delta_{\lambda, \lambda_i} \vec{u}_i \vec{u}_i^\dagger \\
&= \delta_{\zeta, \lambda} A_\lambda .
\end{aligned}$$

Moreover:

$$\begin{aligned}
\sum_{\lambda \in \Lambda_A} A_\lambda &= \sum_{j=1}^M \vec{u}_j \vec{u}_j^\dagger \\
&= UU^\dagger \\
&= I,
\end{aligned}$$

and so on. All of the expected properties of projection operators can be established again in this restricted setting.

The rows of $U^{-1} = U^\dagger$ are A 's left-eigenvectors. In this case, they are simply the conjugate transpose of the right-eigenvector. Note that conjugate transposition is the familiar transformation rule between ket and bra spaces in quantum mechanics (see e.g., Ref. [210])—a consequence of the restriction to normal operators, as we will show. Importantly, a more general formulation of quantum mechanics would *not* have this same restricted correspondence between the dual ket and bra spaces.

To elaborate on this point, recall that vector spaces admit dual spaces and dual bases. However, there is no sense of a dual correspondence of a single ket or bra without reference to a full basis [105]. Implicitly in quantum mechanics, the basis is taken to be the basis of eigenstates of any Hermitian operator, nominally since observables are self-adjoint.

To allude to an alternative, we note that $\vec{u}_j^\dagger \vec{u}_j$ is not only the Hermitian form of inner product $\langle \vec{u}_j, \vec{u}_j \rangle$ (where $\langle \cdot, \cdot \rangle$ denotes the inner product) of the right eigenvector \vec{u}_j with itself, but importantly also the simple dot-product of the left eigenvector \vec{u}_j^\dagger and the right eigenvector \vec{u}_j , where \vec{u}_j^\dagger acts as a linear functional on \vec{u}_j . Contrary to the substantial effort devoted to the inner-product-centric theory of Hilbert spaces, this latter interpretation of $\vec{u}_j^\dagger \vec{u}_j$ —in terms of linear functionals and a left-eigenvector basis for linear functionals—is what generalizes to a consistent and constructive framework for the spectral theory beyond normal operators, as we will see shortly.

2.6.2.2 Diagonalizable matrices

By definition, diagonalizable matrices can be diagonalized, but not necessarily via a unitary transformation. All diagonalizable matrices can nevertheless be diagonalized via the transformation: $A = P\Lambda P^{-1}$, where the columns of the square matrix P are

the not-necessarily-orthogonal right eigenvectors of A corresponding to the eigenvalues ordered along the diagonal matrix Λ and where the rows of P^{-1} are A 's left eigenvectors. Importantly, the left eigenvectors need not be the Hermitian adjoint of the right eigenvectors. As a particular example, this more general setting is required for almost any transition dynamic of a Markov chain. In other words, the transition dynamic of any interesting complex network with irreversible processes serves as an example of a nonnormal operator.

Given the M -tuple of possibly-degenerate eigenvalues $(\Lambda_A) = (\lambda_1, \lambda_2, \dots, \lambda_M)$, there is a corresponding M -tuple of linearly-independent right-eigenvectors $(|\lambda_1\rangle, |\lambda_2\rangle, \dots, |\lambda_M\rangle)$ and a corresponding M -tuple of linearly-independent left-eigenvectors $(\langle\lambda_1|, \langle\lambda_2|, \dots, \langle\lambda_M|)$ such that:

$$A |\lambda_j\rangle = \lambda_j |\lambda_j\rangle$$

and:

$$\langle\lambda_j| A = \lambda_j \langle\lambda_j|$$

with the orthonormality condition that:

$$\langle\lambda_i|\lambda_j\rangle = \delta_{i,j} .$$

To avoid misinterpretation, we stress that the bras and kets that appear above are the left and right eigenvectors, respectively, and typically do *not* correspond to complex-conjugate transposition.

With these definitions in place, the projection operators for a diagonalizable matrix can be written:

$$A_\lambda = \sum_{j=1}^M \delta_{\lambda, \lambda_j} |\lambda_j\rangle \langle\lambda_j| .$$

Then:

$$\begin{aligned}
A &= \sum_{\lambda \in \Lambda_A} \lambda A_\lambda \\
&= \sum_{j=1}^M \lambda_j |\lambda_j\rangle \langle \lambda_j| \\
&= \begin{bmatrix} \lambda_1 |\lambda_1\rangle & \lambda_2 |\lambda_2\rangle & \cdots & \lambda_M |\lambda_M\rangle \end{bmatrix} \begin{bmatrix} \langle \lambda_1| \\ \langle \lambda_2| \\ \vdots \\ \langle \lambda_M| \end{bmatrix} \\
&= \begin{bmatrix} |\lambda_1\rangle & |\lambda_2\rangle & \cdots & |\lambda_M\rangle \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_M \end{bmatrix} \begin{bmatrix} \langle \lambda_1| \\ \langle \lambda_2| \\ \vdots \\ \langle \lambda_M| \end{bmatrix} \\
&= P \Lambda P^{-1} .
\end{aligned}$$

So, we see that the projection operators introduced earlier in a coordinate-free manner have a concrete representation in terms of left and right eigenvectors when the operator is diagonalizable.

2.6.2.3 Any matrix

Not all matrices can be diagonalized, but all square matrices can be put into *Jordan canonical form* via the transformation: $A = Y J Y^{-1}$ [169]. Here, the columns of the square matrix Y are the linearly independent right eigenvectors and generalized right eigenvectors corresponding to the Jordan blocks ordered along the diagonal of the block-diagonal matrix J . And, the rows of Y^{-1} are the corresponding left eigenvectors and generalized left eigenvectors, but reverse-ordered within each block, as we will show.

Let there be n Jordan blocks forming the n -tuple (J_1, J_2, \dots, J_n) , with $1 \leq n \leq M$.

The k^{th} Jordan block J_k has dimension m_k -by- m_k :

$$J_k = \underbrace{\begin{bmatrix} \lambda_k & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & \lambda_k & 1 & & & 0 & 0 \\ & 0 & \lambda_k & & & & 0 \\ \vdots & & & \ddots & \ddots & & \vdots \\ 0 & & & & \lambda_k & 1 & 0 \\ 0 & 0 & & & 0 & \lambda_k & 1 \\ 0 & 0 & 0 & \cdots & & 0 & \lambda_k \end{bmatrix}}_{m_k \text{ columns}} \left. \vphantom{\begin{bmatrix} \lambda_k & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & \lambda_k & 1 & & & 0 & 0 \\ & 0 & \lambda_k & & & & 0 \\ \vdots & & & \ddots & \ddots & & \vdots \\ 0 & & & & \lambda_k & 1 & 0 \\ 0 & 0 & & & 0 & \lambda_k & 1 \\ 0 & 0 & 0 & \cdots & & 0 & \lambda_k \end{bmatrix}} \right\} m_k \text{ rows}$$

such that:

$$\sum_{k=1}^n m_k = M .$$

Note that eigenvalue $\lambda \in \Lambda_A$ corresponds to g_λ different Jordan blocks, where g_λ is the geometric multiplicity of the eigenvalue λ . Indeed:

$$n = \sum_{\lambda \in \Lambda_A} g_\lambda .$$

Moreover, the index ν_λ of the eigenvalue λ is defined as the size of the largest Jordan block corresponding to λ . So, we write this in the current notation as:

$$\nu_\lambda = \max\{\delta_{\lambda, \lambda_k} m_k\}_{k=1}^n .$$

If the index of any eigenvalue is greater than one, then the conventional eigenvectors do not span the M -dimensional vector space. However, the set of M generalized eigenvectors does form a basis for the vector space [8].

Given the n -tuple of possibly-degenerate eigenvalues $(\Lambda_A) = (\lambda_1, \lambda_2, \dots, \lambda_n)$, there is a corresponding n -tuple of m_k -tuples of linearly-independent generalized right-eigenvectors:

$$\left((|\lambda_1^{(m)}\rangle)_{m=1}^{m_1}, (|\lambda_2^{(m)}\rangle)_{m=1}^{m_2}, \dots, (|\lambda_n^{(m)}\rangle)_{m=1}^{m_n} \right) ,$$

where:

$$(|\lambda_k^{(m)}\rangle)_{m=1}^{m_k} \equiv \left(|\lambda_k^{(1)}\rangle, |\lambda_k^{(2)}\rangle, \dots, |\lambda_k^{(m_k)}\rangle \right)$$

and a corresponding n -tuple of m_k -tuples of linearly-independent generalized left-eigenvectors:

$$\left((\langle \lambda_1^{(m)} |)_{m=1}^{m_1}, (\langle \lambda_2^{(m)} |)_{m=1}^{m_2}, \dots, (\langle \lambda_n^{(m)} |)_{m=1}^{m_n} \right),$$

where:

$$(\langle \lambda_k^{(m)} |)_{m=1}^{m_k} \equiv \left(\langle \lambda_k^{(1)} |, \langle \lambda_k^{(2)} |, \dots, \langle \lambda_k^{(m_k)} | \right)$$

such that:

$$(A - \lambda_k I) |\lambda_k^{(m+1)}\rangle = |\lambda_k^{(m)}\rangle \quad (2.36)$$

and:

$$\langle \lambda_k^{(m+1)} | (A - \lambda_k I) = \langle \lambda_k^{(m)} |, \quad (2.37)$$

for $0 \leq m \leq m_k - 1$, where $|\lambda_j^{(0)}\rangle = \vec{0}$ and $\langle \lambda_j^{(0)}| = \vec{0}$. Specifically, $|\lambda_k^{(1)}\rangle$ and $\langle \lambda_k^{(1)}|$ are conventional right and left eigenvectors, respectively.

Most directly, the generalized right and left eigenvectors can be found as the nontrivial solutions to:

$$(A - \lambda_k I)^m |\lambda_k^{(m)}\rangle = \vec{0}$$

and:

$$\langle \lambda_k^{(m)} | (A - \lambda_k I)^m = \vec{0},$$

respectively.

It should be clear from Eq. (3.16) and Eq. (2.37) that:

$$\begin{aligned} \langle \lambda_k^{(m)} | (A - \lambda_k I)^\ell |\lambda_k^{(n)}\rangle &= \langle \lambda_k^{(m-\ell)} | \lambda_k^{(n)}\rangle \\ &= \langle \lambda_k^{(m)} | \lambda_k^{(n-\ell)}\rangle, \end{aligned}$$

for $m, n \in \{0, 1, \dots, m_k\}$ and $\ell \geq 0$. At the same time, it is then easy to show that:

$$\langle \lambda_k^{(m)} | \lambda_k^{(n)}\rangle = \langle \lambda_k^{(m+n)} | \lambda_k^{(0)}\rangle = 0, \quad \text{if } m + n \leq m_k,$$

where $m, n \in \{0, 1, \dots, m_k\}$. Imposing appropriate normalization, we find that:

$$\langle \lambda_j^{(m)} | \lambda_k^{(n)} \rangle = \delta_{j,k} \delta_{m+n, m_k+1} . \quad (2.38)$$

Hence, we see that the left eigenvectors and generalized eigenvectors are a dual basis to the right eigenvectors and generalized eigenvectors. Interestingly though, within each Jordan subspace, *the most generalized left eigenvectors are dual to the least generalized right eigenvectors*, and vice versa.

(To be clear, in this terminology “least generalized” eigenvectors are the standard eigenvectors. For example, the $\langle \lambda_k^{(1)} |$ satisfying the standard eigenvector relation $\langle \lambda_k^{(1)} | A = \lambda_k \langle \lambda_k^{(1)} |$ is the least generalized left eigenvector of subspace k . By way of comparison, the “most generalized” right eigenvector of subspace k is $|\lambda_k^{(m_k)}\rangle$ satisfying the most generalized eigenvector relation $(A - \lambda_k I) |\lambda_k^{(m_k)}\rangle = |\lambda_k^{(m_k-1)}\rangle$ for subspace k . The orthonormality relation shows that the two are dual correspondents: $\langle \lambda_k^{(1)} | \lambda_k^{(m_k)} \rangle = 1$, while all other eigen-bra–eigen-ket closures utilizing these objects are null.)

With these details worked out, we find that the projection operators for a nondiagonalizable matrix can be written as:

$$A_\lambda = \sum_{k=1}^n \sum_{m=1}^{m_k} \delta_{\lambda, \lambda_k} |\lambda_k^{(m)}\rangle \langle \lambda_k^{(m_k+1-m)}| . \quad (2.39)$$

And, we see that a projection operator includes all of its left and right eigenvectors and all of its left and right generalized eigenvectors. This implies that the identity operator must also have a decomposition in terms of both eigenvectors and generalized eigenvectors:

$$\begin{aligned} I &= \sum_{\lambda \in \Lambda_A} A_\lambda \\ &= \sum_{k=1}^n \sum_{m=1}^{m_k} |\lambda_k^{(m)}\rangle \langle \lambda_k^{(m_k+1-m)}| . \end{aligned}$$

Let $[|\lambda_k^{(m)}\rangle]_{m=1}^{m_k}$ denote the column vector:

$$[|\lambda_k^{(m)}\rangle]_{m=1}^{m_k} = \begin{bmatrix} |\lambda_k^{(1)}\rangle \\ \vdots \\ |\lambda_k^{(m_k)}\rangle \end{bmatrix} ,$$

and let $[\langle \lambda_k^{(m_k+1-m)} |]_{m=1}^{m_k}$ denote the column vector:

$$[\langle \lambda_k^{(m_k+1-m)} |]_{m=1}^{m_k} = \begin{bmatrix} \langle \lambda_k^{(m_k)} | \\ \vdots \\ \langle \lambda_k^{(1)} | \end{bmatrix}.$$

Then, using the above results, and the fact that Eq. (2.37) implies that $\langle \lambda_k^{(m+1)} | A = \lambda_k \langle \lambda_k^{(m+1)} | + \langle \lambda_k^{(m)} |$, we derive the explicit generalized-eigenvector decomposition of the nondiagonalizable operator A :

$$\begin{aligned} A &= \left(\sum_{\lambda \in \Lambda_A} A_\lambda \right) A \\ &= \sum_{k=1}^n \sum_{m=1}^{m_k} |\lambda_k^{(m)}\rangle \langle \lambda_k^{(m_k+1-m)}| A \\ &= \sum_{k=1}^n \sum_{m=1}^{m_k} |\lambda_k^{(m)}\rangle \left(\lambda_k \langle \lambda_k^{(m_k+1-m)}| + \langle \lambda_k^{(m_k-m)}| \right) \\ &= \begin{bmatrix} [|\lambda_1^{(m)}\rangle]_{m=1}^{m_1} \\ [|\lambda_2^{(m)}\rangle]_{m=1}^{m_2} \\ \vdots \\ [|\lambda_n^{(m)}\rangle]_{m=1}^{m_n} \end{bmatrix}^\top \begin{bmatrix} J_1 & 0 & \cdots & 0 \\ 0 & J_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_n \end{bmatrix} \begin{bmatrix} [\langle \lambda_1^{(m_1+1-m)} |]_{m=1}^{m_1} \\ [\langle \lambda_2^{(m_2+1-m)} |]_{m=1}^{m_2} \\ \vdots \\ [\langle \lambda_n^{(m_n+1-m)} |]_{m=1}^{m_n} \end{bmatrix} \\ &= Y J Y^{-1}, \end{aligned}$$

where, defining Y as:

$$Y = \begin{bmatrix} [|\lambda_1^{(m)}\rangle]_{m=1}^{m_1} \\ [|\lambda_2^{(m)}\rangle]_{m=1}^{m_2} \\ \vdots \\ [|\lambda_n^{(m)}\rangle]_{m=1}^{m_n} \end{bmatrix}^\top,$$

we are forced by Eq. (2.38) to recognize that:

$$Y^{-1} = \begin{bmatrix} [\langle \lambda_1^{(m_1+1-m)} |]_{m=1}^{m_1} \\ [\langle \lambda_2^{(m_2+1-m)} |]_{m=1}^{m_2} \\ \vdots \\ [\langle \lambda_n^{(m_n+1-m)} |]_{m=1}^{m_n} \end{bmatrix}$$

since then $Y^{-1}Y = I$, and we recall that the inverse is guaranteed to be unique.

The above demonstrates an explicit construction for the Jordan canonical form. One advantage we learn from this explicit decomposition is that the complete set of left eigenvectors and left generalized eigenvectors (encapsulated in Y^{-1}) can be obtained from the inverse of the matrix of the complete set of right eigenvectors and generalized right eigenvectors (encoded in Y) and vice versa. One unexpected lesson, though, is that the generalized left eigenvectors appear in reverse order within each Jordan block.

Using Eqs. (2.39) and (2.18) with Eq. (2.37), we see that the nilpotent operators $A_{\lambda,m}$ with $m > 0$ further link the various generalized eigenvectors within each subspace k . Said more suggestively, generalized modes of a nondiagonalizable subspace are necessarily cooperative.

It is worth noting that the left eigenvectors and generalized left eigenvectors form a basis for all linear functionals of the vector space spanned by the right eigenvectors and generalized right eigenvectors. Moreover, the left eigenvectors and generalized left eigenvectors are exactly the dual basis to the right eigenvectors and generalized right eigenvectors by their orthonormality properties. However, neither the left nor right eigenbasis is a priori more fundamental to the operator. Sympathetically, the right eigenvectors and generalized eigenvectors form a (dual) basis for all linear functionals of the vector space spanned by the left eigenvectors and generalized eigenvectors.

2.6.2.4 Simplified calculi for special cases

In special cases, the meromorphic functional calculus reduces the general expressions above to markedly simpler forms. And, this can greatly expedite practical calculations and provide physical intuition. Here, we show which reductions can be used under which assumptions.

For functions of operators with a countable spectrum, recall that the general form of the meromorphic functional calculus is:

$$f(A) = \sum_{\lambda \in \Lambda_A} \sum_{m=0}^{\nu_\lambda-1} A_{\lambda,m} \frac{1}{2\pi i} \oint_{C_\lambda} \frac{f(z)}{(z-\lambda)^{m+1}} dz . \quad (2.40)$$

Equations (2.18) and (2.39) gave the method to calculate $A_{\lambda,m}$ in terms of eigenvectors

and generalized eigenvectors.

When the operator is *diagonalizable* (not necessarily normal), this reduces to:

$$f(A) = \sum_{\lambda \in \Lambda_A} A_\lambda \frac{1}{2\pi i} \oint_{C_\lambda} \frac{f(z)}{(z - \lambda)} dz , \quad (2.41)$$

where A_λ can now be constructed from conventional right and left eigenvectors, although $\langle \lambda_j |$ is *not* necessarily the conjugate transpose of $|\lambda_j\rangle$.

When the function is *analytic* on the spectrum of the (not necessarily diagonalizable) operator, then our functional calculus reduces to the holomorphic functional calculus:

$$f(A) = \sum_{\lambda \in \Lambda_A} \sum_{m=0}^{\nu_\lambda-1} \frac{f^{(m)}(\lambda)}{m!} A_{\lambda,m} . \quad (2.42)$$

When the function is *analytic* on the spectrum of a *diagonalizable* (not necessarily normal) operator this reduces yet again to:

$$f(A) = \sum_{\lambda \in \Lambda_A} f(\lambda) A_\lambda . \quad (2.43)$$

When the function is *analytic* on the spectrum of a *diagonalizable* (not necessarily normal) operator with *no degeneracy* this reduces even further to:

$$f(A) = \sum_{\lambda \in \Lambda_A} f(\lambda) \frac{|\lambda\rangle \langle \lambda|}{\langle \lambda | \lambda \rangle} . \quad (2.44)$$

Finally, recall that an operator is *normal* when it commutes with its conjugate transpose. If the function is *analytic* on the spectrum of a *normal* operator, then we recover the simple form enabled by the spectral theorem of normal operators familiar in physics. That is, Eq. (2.43) is applicable, but now we have the extra simplification that $\langle \lambda_j |$ is simply the conjugate transpose of $|\lambda_j\rangle$: $\langle \lambda_j | = |\lambda_j\rangle^\dagger$.

2.7 Examples and Applications

To illustrate the use and power of the meromorphic functional calculus, we now adapt it to analyze a suite of applications from quite distinct domains. First, we point to a set of example calculations for finite-dimensional operators of stochastic processes. Second, we show that the familiar Poisson process is intrinsically nondiagonalizable, and hint that

nondiagonalizability may be common more generally in semi-Markov processes. Third, we illustrate how commonly the Drazin inverse arises in nonequilibrium thermodynamics, giving a roadmap to developing closed-form expressions for a number of key observables. Fourth, we turn to signal analysis and comment on power spectra of processes generated by nondiagonalizable operators. Finally, we round out the applications with a general discussion of Ruelle–Frobenius–Perron and Koopman operators for nonlinear dynamical systems.

2.7.1 Spectra of stochastic transition operators

The preceding employed the notation that A represents a general linear operator. In the following examples, we reserve the symbol T for the operator of a stochastic transition dynamic. If the state-space is finite and has a stationary distribution, then T has a representation that is a nonnegative row-stochastic—all rows sum to unity—transition matrix.

The transition matrix’s nonnegativity guarantees that for each $\lambda \in \Lambda_T$ its complex conjugate $\bar{\lambda}$ is also in Λ_T . Moreover, the projection operator associated with the complex conjugate of λ is the complex conjugate of T_λ : $T_{\bar{\lambda}} = \overline{T_\lambda}$.

If the dynamic induced by T has a stationary distribution over the state space, then the spectral radius of T is unity and all of T ’s eigenvalues lie on or within the unit circle in the complex plane. The maximal eigenvalues have unity magnitude and $1 \in \Lambda_T$. Moreover, an extension of the Perron–Frobenius theorem guarantees that eigenvalues on the unit circle have algebraic multiplicity equal to their geometric multiplicity. And, so, $\nu_\zeta = 1$ for all $\zeta \in \{\lambda \in \Lambda_T : |\lambda| = 1\}$.

T ’s index-one eigenvalue of $\lambda = 1$ is associated with stationarity of the associated Markov process. T ’s other eigenvalues on the unit circle are roots of unity and correspond to deterministic periodicities within the process.

All of these results carry over from discrete to continuous time. In continuous time, where $e^{tG} = T_{t_0 \rightarrow t_0+t}$, T ’s stationary eigenvalue of unity maps to G ’s stationary eigenvalue of zero. If the dynamic has a stationary distribution over the state space, then the rate matrix G is row-sum zero rather than row-stochastic. T ’s eigenvalues, on or within the

unit circle, map to G 's eigenvalues with nonpositive real part in the left-hand side of the complex plane.

To reduce ambiguity in the presence of multiple operators, functions of operators, and spectral mapping, we occasionally denote eigenvectors with subscripted operators on the eigenvalues within the bra or ket. For example, $|0_G\rangle = |1_T\rangle \neq |0_{\mathcal{G}}\rangle = |1_{\mathcal{T}}\rangle \neq |0_T\rangle$ disambiguates the identification of $|0\rangle$ when we have operators G , T , \mathcal{G} , and \mathcal{T} with $T = e^{\tau G}$, $\mathcal{T} = e^{\tau \mathcal{G}}$, and $0 \in \Lambda_G, \Lambda_{\mathcal{G}}, \Lambda_T$.

2.7.2 Randomness and memory in correlated processes

The generalized spectral theory developed here has recently been applied to give the first closed-form expressions for many measures of complexity for stochastic processes that can be generated by probabilistic finite automata [60]. Rather than belabor the Kolmogorov–Chaitin notion of complexity which is inherently uncomputable [142], the new analytic framework infuses *computational mechanics* [55] with a means to compute very practical answers about an observed system's organization and to address the challenges of prediction.

For example, we can now answer the obvious questions regarding prediction: How random is a process? How much information is shared between the past and the future? How *far* into the past must we look to predict what is predictable about the future? How *much* about the observed history must be remembered to predict what is predictable about the future? And so on. The Supplementary Materials of Ref. [60] exploit the meromorphic functional calculus to answer these (and more) questions for the symbolic dynamics of a chaotic map, the spacetime domain for an elementary cellular automata, and the chaotic crystallographic structure of a close-packed polytypic material as determined from experimental X-ray diffractograms.

In the context of the current exposition, the most notable feature of the analyses across these many domains is that our imposed questions, which entail tracking an observer's state of knowledge about a process, necessarily *induce* a nondiagonalizable metadynamic that becomes the central object of analysis in each case. (This metadynamic is the so-called *mixed-state presentation* of Refs. [58, 76].)

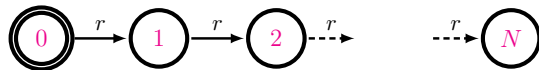


Figure 2.1: Explicit Markov-chain representation of the continuous-time truncated Poisson dynamic, giving interstate transition rates r among the first $N + 1$ counter-states. (State self-transition rates $-r$ are not depicted.) Taking the limit of $N \rightarrow \infty$ recovers the full Poisson counting distribution. It can either be time-homogeneous (transition-rate parameter r is time-independent) or time-inhomogeneous (parameter r is time-dependent).

This theme, and the inherent nondiagonalizability of prediction, is explored in greater depth elsewhere [200]. We also found that another nondiagonalizable dynamic is induced even in the context of quantum communication when determining how much memory reduction can be achieved if we generate a classical stochastic process using quantum mechanics [203].

We mention the above nondiagonalizable metadynamics primarily as a pointer to concrete worked-out examples where the meromorphic functional calculus has been employed to analyze finitary hidden Markov processes via explicitly calculated, generalized eigenvectors and projection operators. We now return to a more self-contained discussion, where we show that nondiagonalizability can be induced by the simple act of counting. Moreover, the theory developed is then applied to deliver quick and powerful results.

2.7.3 Poisson point processes

The meromorphic functional calculus leads naturally to a novel perspective on the familiar Poisson counting process—a familiar stochastic process class used widely across physics and other quantitative sciences to describe “completely random” event durations that occur over a continuous domain [13, 235, 96, 18]. The calculus shows that the basic Poisson distribution arises as the signature of a simple nondiagonalizable dynamic. More to the point, we derive the Poisson distribution directly, without requiring the limit of the discrete-time binomial distribution, as conventionally done [28].

Consider all possible counts, up to some arbitrarily large integer N . The dynamics among these first $N + 1$ counter states constitute what can be called the *truncated Poisson dynamic*. We recover the full Poisson distribution as $N \rightarrow \infty$. A Markov chain for the

truncated Poisson dynamic is shown in Fig. 2.1. The corresponding rate matrix G , for any arbitrarily large truncation N of the possible count, is:

$$G = \begin{bmatrix} -r & r & & & \\ & -r & r & & \\ & & \ddots & \ddots & \\ & & & -r & r \\ & & & & -r \end{bmatrix},$$

where G_{ij} is the rate of transitioning to state (count) j given that the system is in state (count) i . Elements not on either the main diagonal or first superdiagonal are zero. This can be rewritten succinctly as:

$$G = -rI + rD_1,$$

where I is the identity operator in N -dimensions and D_1 is the upshift-by-1 matrix in N -dimensions, with zeros everywhere, except 1s along the first superdiagonal. Let us also define the upshift-by- m matrix D_m with zeros everywhere except 1s along the m^{th} superdiagonal, such that $D_m = D_1^m$ and $D_m^n = D_{m+n}$, with $D_0 = I$. Operationally, if $\langle \delta_\ell |$ is the probability distribution over counter states that is peaked solely at state ℓ , then $\langle \delta_\ell | D_m = \langle \delta_{\ell+m} |$.

For any arbitrarily large N , G 's eigenvalues are given by $\det(G - \lambda I) = (-r - \lambda)^{N+1} = 0$, from which we see that its spectrum is the singleton: $\Lambda_G = \{-r\}$. Moreover, since it has algebraic multiplicity $a_{-r} = N + 1$ and geometric multiplicity $g_{-r} = 1$, the index of the $-r$ eigenvalue is $\nu_{-r} = N + 1$. Since $-r$ is the only eigenvalue, and all projection operators must sum to the identity, we must have the eigenprojection: $G_{-r} = I$. The lesson is that the Poisson point process is highly nondiagonalizable.

2.7.3.1 Homogeneous Poisson processes

When the transition rate r between counter states is constant in time, the net counter state-to-state transition operator from initial time 0 to later time t is given simply by:

$$T(t) = e^{tG}.$$

The functional calculus allows us to directly evaluate e^{tG} for the Poisson nondiagonalizable transition-rate operator G ; we find:

$$\begin{aligned}
T(t) &= e^{tG} \\
&= \sum_{\lambda \in \Lambda_G} \sum_{m=0}^{\nu_\lambda-1} G_\lambda (G - \lambda I)^m \left(\frac{1}{2\pi i} \oint_{C_\lambda} \frac{e^{tz}}{(z - \lambda)^m} dz \right) \\
&= \lim_{N \rightarrow \infty} \sum_{m=0}^N I(G + rI)^m \frac{1}{m!} \underbrace{\lim_{z \rightarrow -r} \frac{d^m}{dz^m} e^{tz}}_{t^m e^{-rt}} \\
&= \sum_{m=0}^{\infty} (rD_1)^m \frac{t^m e^{-rt}}{m!} \\
&= \sum_{m=0}^{\infty} D_m \frac{(rt)^m e^{-rt}}{m!} .
\end{aligned}$$

Consider the orthonormality relation $\langle \delta_i | \delta_j \rangle = \delta_{i,j}$ between counter states, where $|\delta_j\rangle$ is represented by 0s everywhere except for a 1 at counter-state j . It effectively measures the occupation probability of counter-state j . Employing the result for $T(t)$, we find the simple consequence that:

$$\begin{aligned}
\langle \delta_0 | T(t) | \delta_n \rangle &= \frac{(rt)^n e^{-rt}}{n!} \\
&= \langle \delta_n | T(t) | \delta_{m+n} \rangle .
\end{aligned}$$

That is, the probability that the counter is incremented by n in a time interval t is independent of the initial count and given by: $(rt)^n e^{-rt} / n!$.

Let us emphasize that these steps derived the Poisson distribution directly, rather than as the typical limit of the binomial distribution. Our derivation depended critically on spectral manipulations of a highly nondiagonalizable operator. Moreover, our result for the transition dynamic $T(t)$ allows a direct analysis of how *distributions* over counts evolve in time, as would be necessary, say, in a Bayesian setting with unknown prior count. This type of calculus can immediately be applied to the analysis of more sophisticated processes, for which we can generally expect nondiagonalizability to play an important functional role.

2.7.3.2 Inhomogeneous Poisson processes

Let us now generalize to time-inhomogeneous Poisson processes, where the transition rate r between count events is instantaneously uniform, but varies in time as $r(t)$. Conveniently, the associated rate matrices at different times commute with each other. Specifically, with $G_a = -aI + aD_1$ and $G_b = -bI + bD_1$, we see that:

$$[G_a, G_b] = 0 .$$

Therefore, the net counter state-to-state transition operator from time t_0 to time t_f is given by:

$$\begin{aligned} T_{t_0, t_f} &= e^{\int_{t_0}^{t_f} G(t) dt} \\ &= e^{\left(\int_{t_0}^{t_f} r(t) dt\right)(-I + D_1)} \\ &= e^{\langle r \rangle (\Delta t)(-I + D_1)} \\ &= e^{(\Delta t)G_{\langle r \rangle}} , \end{aligned} \tag{2.45}$$

where $\Delta t = t_f - t_0$ is the time elapsed and:

$$\langle r \rangle = \frac{1}{\Delta t} \int_{t_0}^{t_f} r(t) dt$$

is the average rate during that time. Given Eq. (2.45), the functional calculus proceeds just as in the time-homogeneous case to give the analogous net transition dynamic:

$$T_{t_0, t_f} = \sum_{m=0}^{\infty} D_m \frac{(\langle r \rangle \Delta t)^m e^{-\langle r \rangle \Delta t}}{m!} .$$

The probability that the count is incremented by n during the time interval Δt follows directly:

$$\langle \delta_m | T_{t_0, t_f} | \delta_{m+n} \rangle = \frac{(\langle r \rangle \Delta t)^n e^{-\langle r \rangle \Delta t}}{n!} .$$

With relative ease, our calculus allowed us to derive an important result for stochastic process theory that is nontrivial to derive by other means. Perhaps surprisingly, we see that the probability distribution over final counts induced by any rate trajectory $r(t)$ is the same as if the transition rate were held fixed at mean $\langle r \rangle$ throughout the duration.

Moreover, we can directly analyze the net evolution of distributions over counts using the derived transition operator T_{t_0, t_f} .

Note that the nondiagonalizability of the Poisson dynamic is robust in a physical sense. That is, even varying the rate parameter in time in an erratic way, the inherent structure of counting imposes a fundamental nondiagonalizable nature. That nondiagonalizability can be robust in a physical sense is significant, since one might otherwise be tempted to argue that nondiagonalizability is extremely fragile due to numerical perturbations within any matrix representation of the operator. This is simply not the case since such perturbations are physically forbidden. Rather, this simple example challenges us with the fact that some processes, even those familiar and widely used, are intrinsically nondiagonalizable. On the positive side, it appears that spectral methods can now be applied to analyze them. And, this will be particularly important in more complex, memoryful processes [164, 168, 166], including the hidden semi-Markov processes [13] that are, roughly speaking, the cross-product of hidden finite-state Markov chains and renewal processes.

2.7.4 Stochastic thermodynamics

The previous simple examples started to demonstrate the methods of the meromorphic functional calculus. Next, we show a novel application of the meromorphic functional calculus to environmentally driven mesoscopic dynamical systems, selected to give a new set of results within nonequilibrium thermodynamics. In particular, we analyze functions of singular transition-rate operators. Notably, we show that the Drazin inverse arises naturally in the general solution of Green–Kubo relations. We mention that it also arises when analyzing moments of the excess heat produced in the driven transitions atop either equilibrium steady states or nonequilibrium steady states [201].

2.7.4.1 Dynamics in independent eigenspaces

An important feature of the functional calculus is its ability to address particular eigenspaces independently when necessary. This feature is often taken for granted in the case of normal operators; say, in physical dynamical systems when analyzing stationary distributions or dominant decay modes. Consider a singular operator \mathcal{L} that is not necessarily normal and not necessarily diagonalizable and evaluate the simple yet ubiquitous integral $\int_0^\tau e^{t\mathcal{L}} dt$.

Via the meromorphic functional calculus we find:

$$\begin{aligned}
\int_0^\tau e^{t\mathcal{L}} dt &= \sum_{\lambda \in \Lambda_{\mathcal{L}}} \sum_{m=0}^{\nu_\lambda-1} \mathcal{L}_{\lambda,m} \frac{1}{2\pi i} \oint_{C_\lambda} \frac{\int_0^\tau e^{tz} dt}{(z-\lambda)^{m+1}} dz \\
&= \left(\sum_{m=0}^{\nu_0-1} \mathcal{L}_{0,m} \frac{1}{2\pi i} \oint_{C_0} \frac{z^{-1}(e^{\tau z} - 1)}{z^{m+1}} dz \right) \\
&\quad + \sum_{\lambda \in \Lambda_{\mathcal{L}} \setminus 0} \sum_{m=0}^{\nu_\lambda-1} \mathcal{L}_{\lambda,m} \frac{1}{2\pi i} \oint_{C_\lambda} \frac{z^{-1}(e^{\tau z} - 1)}{(z-\lambda)^{m+1}} dz \\
&= \left(\sum_{m=0}^{\nu_0-1} \frac{\tau^{m+1}}{(m+1)!} \mathcal{L}_{0,m} \right) + \mathcal{L}^{\mathcal{D}} (e^{\tau\mathcal{L}} - I) , \tag{2.46}
\end{aligned}$$

where $\mathcal{L}^{\mathcal{D}}$ is the Drazin inverse of \mathcal{L} , discussed earlier.

The pole–pole interaction (z^{-1} with z^{-m-1}) at $z = 0$ distinguished the 0-eigenspace in the calculations and required the meromorphic functional calculus for direct analysis. The given solution to this integral will be useful in the following.

Next, we consider the case where \mathcal{L} is the transition-rate operator among the states of a structured stochastic dynamical system. This leads to several novel consequence within stochastic thermodynamics.

2.7.4.2 Green–Kubo relations

Let us reconsider the above integral in the case when the singular operator \mathcal{L} —let us call it G —is a transition-rate operator that exhibits a single stationary distribution. By the spectral mapping $\ln \Lambda_{eG}$ of the eigenvalue $1 \in \Lambda_{eG}$ addressed in the Perron–Frobenius theorem, G ’s zero eigenmode is diagonalizable. And, by assuming a single attracting stationary distribution, the zero eigenvalue has algebraic multiplicity $a_0 = 1$. Equation (2.46) then simplifies to:

$$\int_0^\tau e^{tG} dt = \tau |0_G\rangle \langle 0_G| + G^{\mathcal{D}} (e^{\tau G} - I) . \tag{2.47}$$

Since G is a transition-rate operator, the above integral corresponds to integrated time evolution. The Drazin inverse $G^{\mathcal{D}}$ concentrates on the transient contribution beyond the persistent stationary background. In Eq. (2.47), the subscript within the left and right eigenvectors explicitly links the eigenvectors to the operator G , to reduce ambiguity. Specifically, the projector $|0_G\rangle \langle 0_G|$ maps any distribution to the stationary distribution.

Green–Kubo-type relations [100, 284] connect the out-of-steady-state transport coefficients to the time integral of steady-state autocorrelation functions. They are thus very useful for understanding out-of-steady-state dissipation due to steady-state fluctuations. (Steady state here refers to either equilibrium or nonequilibrium steady state.) Specifically, the Green–Kubo relation for a transport coefficient, κ say, is typically of the form:

$$\kappa = \int_0^\infty (\langle A(0)A(t) \rangle_{\text{s.s.}} - \langle A \rangle_{\text{s.s.}}^2) dt ,$$

where $A(0)$ and $A(t)$ are some observable of the stationary stochastic dynamical system at time 0 and time t , respectively, and the subscript $\langle \cdot \rangle_{\text{s.s.}}$ emphasizes that the expectation value is to be taken according to the steady-state distribution.

Using:

$$\begin{aligned} \langle A(0)A(t) \rangle_{\text{s.s.}} &= \text{tr}(|0_G\rangle \langle 0_G| A e^{tG} A) \\ &= \langle 0_G| A e^{tG} A |0_G\rangle , \end{aligned}$$

the transport coefficient κ can be written more explicitly in terms of the relevant transition-rate operator G for the stochastic dynamics:

$$\begin{aligned} \kappa &= \lim_{\tau \rightarrow \infty} \int_0^\tau \langle 0_G| A e^{tG} A |0_G\rangle dt - \tau \langle 0_G| A |0_G\rangle^2 \\ &= \lim_{\tau \rightarrow \infty} \langle 0_G| A \left(\int_0^\tau e^{tG} dt \right) A |0_G\rangle - \tau \langle 0_G| A |0_G\rangle^2 \\ &= \lim_{\tau \rightarrow \infty} \langle 0_G| A G^{\mathcal{D}} (e^{\tau G} - I) A |0_G\rangle \\ &= - \langle A G^{\mathcal{D}} A \rangle_{\text{s.s.}} . \end{aligned} \tag{2.48}$$

Thus, we learn that relations of Green–Kubo form are direct signatures of the Drazin inverse of the transition-rate operator for the stochastic dynamic.

The result of Eq. (2.48) holds quite generally. For example, if the steady state has some number of periodic flows, the result of Eq. (2.48) remains valid. Alternatively, in the case of nonperiodic chaotic flows—where G will be the logarithm of the Ruelle–Frobenius–Perron operator, as described later in § 2.7.6.1— $|0_G\rangle \langle 0_G|$ still induces the average over the steady-state trajectories.

In the special case where the transition-rate operator is diagonalizable, $-\langle A G^{\mathcal{D}} A \rangle_{\text{s.s.}}$ is simply the integrated contribution from a weighted sum of decaying exponentials. Transport coefficients then have a solution of the simple form:

$$\kappa = - \sum_{\lambda \in \Lambda_G \setminus 0} \frac{1}{\lambda} \langle 0_G | A G_{\lambda} A | 0_G \rangle . \quad (2.49)$$

Note that the minus sign keeps κ positive since $\text{Re}(\lambda) < 0$ for $\lambda \in \Lambda_G \setminus \{0\}$. Also, recall that G 's eigenvalues with nonzero imaginary part occur in complex-conjugate pairs and $G_{\bar{\lambda}} = \overline{G_{\lambda}}$. Moreover, if $G_{i,j}$ is the classical transition-rate from state i to state j (to disambiguate from the transposed possibility), then $\langle 0_G |$ is the stationary distribution. (The latter is sometimes denoted $\langle \pi |$ in the Markov process literature.) And, $|0_G\rangle$ is a column vector of all ones (sometimes denoted $|\mathbf{1}\rangle$) which acts to integrate contributions throughout the state space.

A relationship of the form of Eq. (2.48), between the Drazin inverse of a classical transition-rate operator and a particular Green–Kubo relation was recently found in Ref. [163] for the friction tensor for smoothly-driven transitions atop nonequilibrium steady states. Subsequently, a truncation of the eigen-expansion of the form of Eq. (2.49) was recently used in a similar context to bound a universal tradeoff between power, precision, and speed [137]. Equation (2.48) shows that a fundamental relationship between a physical property and a Drazin inverse is to be expected more generally whenever the property can be related to integrated correlation.

Notably, if a Green–Kubo-like relation integrates a cross-correlation, say between $A(t)$ and $B(t)$ rather than an autocorrelation, then we have only the slight modification:

$$\int_0^{\infty} (\langle A(0)B(t) \rangle_{\text{s.s.}} - \langle A \rangle_{\text{s.s.}} \langle B \rangle_{\text{s.s.}}) dt = - \langle A G^{\mathcal{D}} B \rangle_{\text{s.s.}} . \quad (2.50)$$

The foregoing analysis bears on both classical and quantum dynamics. G may be a so-called linear *superoperator* in the quantum regime [152]; for example, the *Lindblad superoperator* [147, 14] that operates on density operators. A Liouville-space representation [192] of the superoperator, though, exposes the superficiality of the distinction between superoperator and operator. At an abstract level, time evolution can be dis-

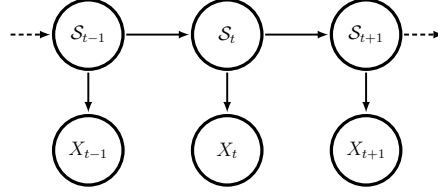


Figure 2.2: Bayes network for a state-emitting hidden Markov model graphically depicts the structure of conditional independence among random variables for the latent state $\{\mathcal{S}_n\}_{n \in \mathbb{Z}}$ at each time n and the random variables $\{X_n\}_{n \in \mathbb{Z}}$ for the observation at each time n .

cussed uniformly across subfields and reinterpretations of Eq. (2.50) will be found in each associated physical theory.

Reference [201] presents additional constructive results that emphasize the ubiquity of integrated correlation and Drazin inverses in the transitions between steady states [181], relevant to the fluctuations within any physical dynamic. Overall, these results support the broader notion that dissipation depends on the structure of correlation.

Frequency-dependent generalizations of integrated correlation have a corresponding general solution. To be slightly less abstract, later on we give novel representative formulae for a particular application: the general solution to power spectra of a process generated by any countable-state hidden Markov chain.

2.7.5 Power spectra

A signal's power spectrum quantifies how its power is distributed across frequency [242]. For a discrete-domain process it is:

$$P(\omega) = \lim_{N \rightarrow \infty} \frac{1}{N} \left\langle \left| \sum_{n=1}^N X_n e^{-i\omega n} \right|^2 \right\rangle, \quad (2.51)$$

where ω is the angular frequency and X_n is the random variable for the observation at time n . For a wide-sense stationary stochastic process, the power spectrum is also determined from the signal's autocorrelation function $\gamma(\tau)$:

$$P(\omega) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{\tau=-N}^N (N - |\tau|) \gamma(\tau) e^{-i\omega \tau}, \quad (2.52)$$

where the autocorrelation function for a wide-sense stationary stochastic process is defined:

$$\gamma(\tau) = \langle \overline{X_n} X_{n+\tau} \rangle_n .$$

The windowing function $N - |\tau|$ appearing in Eq. (2.52) is a direct consequence of Eq. (2.51). It is not imposed externally, as is common practice in signal analysis. This is important to subsequent derivations.

The question we address is how to calculate the correlation function and power spectrum given a model of the signal's generator. To this end, we briefly introduce hidden Markov models as signal generators and then use the meromorphic calculus to calculate their autocorrelation and power spectra in closed-form. This leads to several lessons. First, we see that the power spectrum is a direct fingerprint of the resolvent of the generator's time-evolution operator, analyzed along the unit circle. Second, spectrally decomposing the not-necessarily-diagonalizable time evolution operator, we derive the most general qualitative behavior of the autocorrelation function and power spectra. Third, contributions from eigenvalues on the unit circle must be extracted and dealt with separately. Contributions from eigenvalues on the unit circle correspond to Dirac delta functions—the analog of Bragg peaks in diffraction. Whereas, eigen-contributions from inside the unit circle correspond to diffuse peaks, which become sharper for eigenvalues closer to the unit circle. Finally, nondiagonalizable eigenmodes yield qualitatively different line profiles than their diagonalizable counterparts. In short, when applied to signal analysis our generalized spectral decomposition has directly measurable consequences. This has been key to analyzing low-dimensional disordered materials, for example, when adapted to X-ray diffraction spectra [239, 205, 204].

Let the 4-tuple $\mathcal{M} = (\mathcal{S}, \mathcal{A}, \mathcal{P}, T)$ be some discrete-time *state-emitting hidden Markov model* (HMM) that generates the stationary stochastic process $\dots X_{-2}X_{-1}X_0X_1X_2\dots$ according to the following. \mathcal{S} is the (finite) set of latent states of the hidden Markov chain and $\mathcal{A} \subseteq \mathbb{C}$ is the observable alphabet. \mathcal{S}_t is the random variable for the hidden state at time t that takes on values $s \in \mathcal{S}$. X_t is the random variable for the observation at time t that takes on values $x \in \mathcal{A}$. Given the latent state at time t , the possible observations are distributed according to the conditional probability density functions:

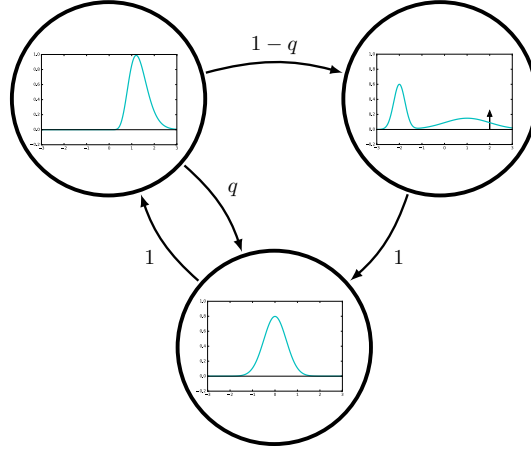


Figure 2.3: Simple 3-state state-emitting HMM that generates a stochastic process according to the state-to-state transition dynamic T and the probability density functions (pdfs) $\{p(x|s)\}_{s \in \mathcal{S}}$ associated with each state. Theorem 1 asserts that its power spectrum will be the same (with only constant offset) as the power spectrum generated from the alternative process where the pdfs in each state are solely concentrated at the Platonic average value $\langle x \rangle_{p_s(x)}$ of the former pdf associated with the state.

$\mathcal{P} = \{p(X_t = x | \mathcal{S}_t = s)\}_{s \in \mathcal{S}}$. For each $s \in \mathcal{S}$, $p(X_t = x | \mathcal{S}_t = s)$ may be abbreviated as $p(x|s)$ since the probability density function in each state is assumed not to change over t . Finally, the latent-state-to-state stochastic transition matrix T has elements $T_{i,j} = \Pr(\mathcal{S}_{t+1} = s_j | \mathcal{S}_t = s_i)$, which give the probability of transitioning from latent state s_i to s_j given that the system is in state s_i , where $s_i, s_j \in \mathcal{S}$. It is important for the subsequent derivation that we use $\Pr(\cdot)$ to denote a probability in contrast to $p(\cdot)$ which denotes a probability *density*. The Bayes network diagram of Fig. 2.2 depicts the structure of conditional independence among the random variables.

2.7.5.1 Continuous-value, discrete-state and -time processes

Figure 2.3 gives a particular HMM with continuous observable alphabet $\mathcal{A} = \mathbb{R}$ distributed according to the probability density function shown within each latent state. Processes generated as the observation of a function of a Markov chain can be of either finite or infinite Markov order. (They are, in fact, *typically* infinite Markov order in the space of processes [120].)

Directly calculating, one finds that the autocorrelation function, for $\tau > 0$, for any

such HMM is:

$$\begin{aligned}
\gamma(\tau) &= \langle \overline{X_n} X_{n+\tau} \rangle_n \\
&= \int_{x \in \mathcal{A}} \int_{x' \in \mathcal{A}} \overline{x} x' p(X_0 = x, X_\tau = x') dx dx' \\
&= \sum_{s \in \mathcal{S}} \sum_{s' \in \mathcal{S}} \int_{x \in \mathcal{A}} \int_{x' \in \mathcal{A}} \overline{x} x' p(X_0 = x, X_\tau = x', \mathcal{S}_0 = s, \mathcal{S}_\tau = s') dx dx' \\
&= \sum_{s \in \mathcal{S}} \sum_{s' \in \mathcal{S}} \int_{x \in \mathcal{A}} \int_{x' \in \mathcal{A}} \overline{x} x' \Pr(\mathcal{S}_0 = s, \mathcal{S}_\tau = s') p(X_0 = x | \mathcal{S}_0 = s) p(X_\tau = x' | \mathcal{S}_\tau = s') dx dx' \\
&= \sum_{s \in \mathcal{S}} \sum_{s' \in \mathcal{S}} \langle \boldsymbol{\pi} | \delta_s \rangle \langle \delta_s | T^\tau | \delta_{s'} \rangle \langle \delta_{s'} | \mathbf{1} \rangle \left(\int_{x \in \mathcal{A}} \overline{x} p(x | s) dx \right) \left(\int_{x' \in \mathcal{A}} x' p(x' | s') dx' \right) \\
&= \langle \boldsymbol{\pi} | \left(\sum_{s \in \mathcal{S}} \langle \overline{x} \rangle_{p(x|s)} | \delta_s \rangle \langle \delta_s | \right) T^\tau \left(\sum_{s' \in \mathcal{S}} \langle x \rangle_{p(x|s')} | \delta_{s'} \rangle \langle \delta_{s'} | \right) | \mathbf{1} \rangle ,
\end{aligned}$$

where:

$$p(X_0 = x, X_\tau = x', \mathcal{S}_0 = s, \mathcal{S}_\tau = s') = \Pr(\mathcal{S}_0 = s, \mathcal{S}_\tau = s') p(X_0 = x, X_\tau = x' | \mathcal{S}_0 = s, \mathcal{S}_\tau = s')$$

holds by definition of conditional probability. The decomposition of:

$$p(X_0 = x, X_\tau = x' | \mathcal{S}_0 = s, \mathcal{S}_\tau = s') = p(X_0 = x | \mathcal{S}_0 = s) p(X_\tau = x' | \mathcal{S}_\tau = s')$$

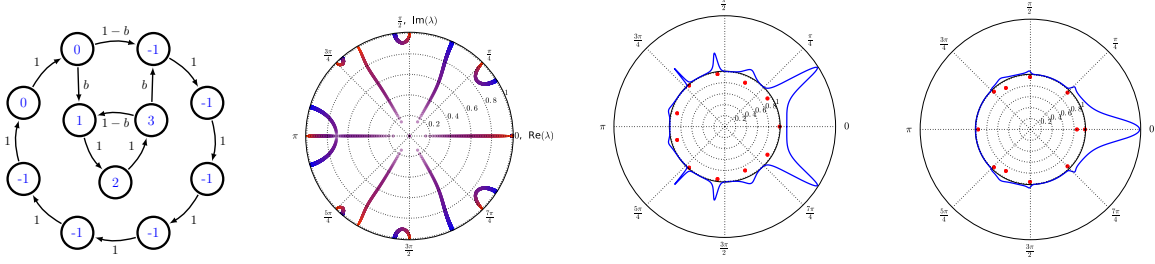
for $\tau \neq 0$ follows from the conditional independence in the relevant Bayesian network shown in Fig. 2.2. Moreover, the equality:

$$\Pr(\mathcal{S}_0 = s, \mathcal{S}_\tau = s') = \langle \boldsymbol{\pi} | \delta_s \rangle \langle \delta_s | T^\tau | \delta_{s'} \rangle \langle \delta_{s'} | \mathbf{1} \rangle$$

can be derived by marginalizing over all possible intervening state sequences. Note that $|\delta_s\rangle$ is the column vector of all 0s except for a 1 at the index corresponding to state s and $\langle \delta_s |$ is simply its transpose. Recall that $\langle \boldsymbol{\pi} | = \langle 1_T |$ is the stationary distribution induced by T over latent states and $|\mathbf{1}\rangle = |1_T\rangle$ is a column vector of all ones. Note also that $\langle \boldsymbol{\pi} | \delta_s \rangle = \Pr(s)$ and $\langle \delta_{s'} | \mathbf{1} \rangle = 1$.

Since the autocorrelation function is symmetric in τ and:

$$\begin{aligned}
\gamma(0) &= \langle |x|^2 \rangle_{p(x)} \\
&= \langle \boldsymbol{\pi} | \sum_{s \in \mathcal{S}} \langle |x|^2 \rangle_{p(x|s)} | \delta_s \rangle \rangle ,
\end{aligned}$$



(a) A b -parametrized HMM with mean values of each state's pdf $\langle x \rangle_{p(x|s)}$ indicated as the number inside each state. (b) Eigenvalue evolution for all $\lambda \in \Lambda_T$ and eigenvalues at $b = 3/4$. (c) Power spectrum and eigenvalues at $b = 3/4$. (d) Power spectrum and eigenvalues at $b = 1/4$.

Figure 2.4: Parametrized HMM generator of a stochastic process, its eigenvalue evolution, and two coronal spectrograms showing power spectra emanating from eigen-spectra.

we find the full autocorrelation function is given by:

$$\gamma(\tau) = \begin{cases} \langle |x|^2 \rangle & \text{if } \tau = 0 \\ \langle \pi | \bar{\Omega} T^{|\tau|-1} \Omega | \mathbf{1} \rangle & \text{if } |\tau| \geq 1 \end{cases},$$

where Ω is the $|\mathcal{S}|$ -by- $|\mathcal{S}|$ matrix defined by:

$$\Omega = \sum_{s \in \mathcal{S}} \langle x \rangle_{p(x|s)} |\delta_s\rangle \langle \delta_s| T. \quad (2.53)$$

The power spectrum is then calculated via Eq. (2.52) using the meromorphic calculus. In particular, the power spectrum decomposes naturally into a discrete part and a continuous part. Full details will be given elsewhere, but the derivation is similar to that given in Ref. [204] for the special case of diffraction patterns from HMMs. We note that it is important to treat individual eigenspaces separately, as our generalized calculus naturally accommodates. The end result, for the continuous part of the power spectrum, is:

$$P_c(\omega) = \langle |x|^2 \rangle + 2 \text{Re} \langle \pi | \bar{\Omega} (e^{i\omega} I - T)^{-1} \Omega | \mathbf{1} \rangle. \quad (2.54)$$

All of the ω -dependence is in the resolvent. Using the spectral expansion of the resolvent given by Eq. (2.21) allows us to better understand the qualitative possibilities for the

shape of the power spectrum:

$$P_c(\omega) = \langle |x|^2 \rangle + \sum_{\lambda \in \Lambda_T} \sum_{m=0}^{\nu_\lambda-1} 2 \operatorname{Re} \frac{\langle \boldsymbol{\pi} | \bar{\Omega} T_{\lambda,m} \Omega | \mathbf{1} \rangle}{(e^{i\omega} - \lambda)^{m+1}}. \quad (2.55)$$

Note that $\langle \boldsymbol{\pi} | \bar{\Omega} T_{\lambda,m} \Omega | \mathbf{1} \rangle$ is a complex-valued scalar and all of the frequency dependence now handily resides in the denominator.

The discrete portion (delta functions) of the power spectrum is:

$$P_d(\omega) = \sum_{k=-\infty}^{\infty} \sum_{\substack{\lambda \in \Lambda_T \\ |\lambda|=1}} 2\pi \delta(\omega - \omega_\lambda + 2\pi k) \\ \times \operatorname{Re}(\lambda^{-1} \langle \boldsymbol{\pi} | \bar{\Omega} T_\lambda \Omega | \mathbf{1} \rangle), \quad (2.56)$$

where ω_λ is related to λ by $\lambda = e^{i\omega_\lambda}$. An extension of the Perron–Frobenius theorem guarantees that the eigenvalues of T on the unit circle have index $\nu_\lambda = 1$.

When plotted as a function of the angular frequency ω around the unit circle, the power spectrum suggestively appears to emanate from the eigenvalues $\lambda \in \Lambda_T$ of the hidden linear dynamic. See Fig. 2.4 for the analysis of an example parametrized process and the last two panels for this display mode for the power spectra.

Eigenvalues of T on the unit circle yield Dirac delta functions in the power spectrum. Eigenvalues of T within the unit circle yield more diffuse line profiles, increasingly diffuse as the magnitude of the eigenvalues retreats toward the origin. Moreover, the integrated magnitude of each contribution is determined by projecting pairwise observation operators onto the eigenspace emanating the contribution. Finally, we note that nondiagonalizable eigen-modes yield qualitatively different line profiles.

Remarkably, the power spectrum generated by such process is the same as the that generated by a potentially much simpler one—a process that is a function of the same underlying Markov chain but instead emits the state-dependent *expectation value* of the observable within each state:

Theorem 1. *Let $\mathcal{P} = \{p_s(x)\}_{s \in \mathcal{S}}$ be any set of probability distribution functions over the domain $\mathcal{A} \subseteq \mathbb{C}$. Let $\mathcal{B} = \{\langle x \rangle_{p_s(x)}\}_{s \in \mathcal{S}}$ and let $\mathcal{Q} = \{\delta(x - \langle x \rangle_{p_s(x)})\}_{s \in \mathcal{S}}$. Then, the power spectrum generated by any hidden Markov model $\mathcal{M} = (\mathcal{S}, \mathcal{A}, \mathcal{P}, T)$ differs at*

most by a constant offset from the power spectrum generated by the hidden Markov model $\mathcal{M}' = (\mathcal{S}, \mathcal{B}, \mathcal{Q}, T)$ that has the same latent Markov chain but in any state $s \in \mathcal{S}$ emits, with probability one, the average value $\langle x \rangle_{p_s(x)}$ of the state-conditioned probability density function $p_s(x) \in \mathcal{P}$ of \mathcal{M} .

Proof. From Eqs. (2.54) and (2.56), we see that $P_c(\omega) + P_d(\omega) - \langle |x|^2 \rangle$ depends only on T and $\{\langle x \rangle_{p(x|s)}\}_{s \in \mathcal{S}}$. This shows that all HMMs that share the same T and $\{\langle x \rangle_{p(x|s)}\}_{s \in \mathcal{S}}$ have the same power spectrum $P(\omega) = P_c(\omega) + P_d(\omega)$ besides a constant offset determined by differences in $\langle |x|^2 \rangle$.

One immediate consequence is that *any hidden Markov chain with any arbitrary set of zero-mean distributions attached to each state, i.e.:*

$$\mathcal{P} \in \left\{ \{p(x|s)\}_{s \in \mathcal{S}} : \langle x \rangle_{p(x|s)} = 0 \text{ for all } s \in \mathcal{S} \right\},$$

generates a flat power spectrum with the appearance of white noise. On the one hand, this strongly suggests to data analysts to look beyond power spectra when attempting to extract a process' full architecture. On the other, whenever a process's power spectrum *is* structured, it is a direct fingerprint of the resolvent of the hidden linear dynamic. In short, the power spectrum is a filtered image of the resolvent along the unit circle.

The power spectrum of a particular stochastic process is shown in Fig. 2.4 and using *coronal spectrograms*, introduced in Ref. [204], it illustrates how the observed spectrum can be thought of as emanating from the spectrum of the hidden linear dynamic, as all power spectra must. Figure 2.4a shows the state-emitting HMM with state-to-state transition probabilities parametrized by b ; the mean values $\langle x \rangle_{p(x|s)}$ of each state's pdf $p(x|s)$ are indicated as the blue number inside each state. The process generated depends on the actual pdfs and the transition parameter b although, and this is our point, the power spectrum is ignorant to the details of the pdfs.

The evolution of the eigenvalues Λ_T of the transition dynamic among latent states is shown from thick blue to thin red markers in Fig. 2.4b, as we sweep the transition parameter b from 1 to 0. A subset of the eigenvalues pass continuously but very quickly through the origin of the complex plane as b passes through 1/2. The continuity of this

is not immediately apparent numerically, but can be revealed with a finer increment of b near $b \approx 1/2$. Notice the persistent eigenvalue of $\lambda_T = 1$, which is guaranteed by the Perron–Frobenius theorem.

In Fig. 2.4c and again, at another parameter setting, in Fig. 2.4d, we show the continuous part of the *power* spectrum $P_c(\omega)$ (plotted around the unit circle in solid blue) and the *eigen*-spectrum (plotted as red dots on and within the unit circle) of the state-to-state transition matrix for the 11-state hidden Markov chain (leftmost panel) that generates it. There is also a δ -function contribution to the power spectrum at $\omega = 0$ (corresponding to $\lambda_T = 1$). This is not shown. These coronal spectrograms illustrate how the power spectrum emanates from the HMM’s eigen-spectrum, with sharper peaks when the eigenvalues are closer to the unit circle. This observation is fully explained by Eq. (2.55). The integrated magnitude of each peak depends on $\langle \boldsymbol{\pi} | \bar{\Omega} | \lambda \rangle \langle \lambda | \Omega | \mathbf{1} \rangle$.

Interestingly, the apparent continuous spectrum component is the shadow of the discrete spectrum of nonunitary dynamics. This suggests that resonances in various physics domains concerned with a continuous spectrum can be modeled as simple consequences of nonunitary dynamics. Indeed, hints of this appear in the literature [176, 236, 173].

2.7.5.2 Continuous-time processes

We close this exploration of conventional signal analysis methods using the meromorphic calculus by commenting on continuous-time processes. Analogous formulae can be derived with similar methods for continuous-time hidden Markov jump processes and continuous-time deterministic (possibly chaotic) dynamics in terms of the generator G of time evolution. For example, the continuous part $P_c(\omega)$ of the power spectrum from a continuous-time deterministic dynamic has the form:

$$P_c(\omega) = 2 \operatorname{Re} \langle \boldsymbol{\pi} | \bar{\Omega} (i\omega I - G)^{-1} \Omega | \mathbf{1} \rangle .$$

Appealing to the resolvent’s spectral expansion again allows us to better understand the possible shapes of their power spectrum:

$$P_c(\omega) = \sum_{\lambda \in \Lambda_G} \sum_{m=0}^{\nu_\lambda-1} 2 \operatorname{Re} \frac{\langle \boldsymbol{\pi} | \bar{\Omega} G_{\lambda,m} \Omega | \mathbf{1} \rangle}{(i\omega - \lambda)^{m+1}} . \quad (2.57)$$

Since all of the frequency-dependence has been isolated in the denominator and $\langle \boldsymbol{\pi} | \bar{\Omega} G_{\lambda,m} \Omega | \mathbf{1} \rangle$ is a frequency-independent complex-valued constant, peaks in $P_c(\omega)$ can only arise via contributions of the form $\text{Re} \frac{c}{(i\omega - \lambda)^n}$ for $c \in \mathbb{C}$, $\omega \in \mathbb{R}$, $\lambda \in \Lambda_G$, and $n \in \mathbb{Z}_+$. This provides a rich starting point for application and further theoretical investigation. For example, Eq. (2.57) helps explain the shapes of power spectra of nonlinear dynamical systems, as have appeared, e.g., in Ref. [87]. Furthermore, it suggests an approach to the inverse problem of inferring the spectrum of the hidden linear dynamic via power spectra. In the next section, however, we develop a more general proposal for inferring eigenvalues from a time series. Further developments will appear elsewhere.

2.7.6 Operators for chaotic dynamics

Since trajectories in state-space can be generated independently of each other, any nonlinear dynamic corresponds to a linear operation on an infinite-dimensional vector-space of complex-valued distributions (in the sense of generalized functions) over the original state-space. For example, the well known Lorenz ordinary differential equations [151] are nonlinear over its three given state-space variables— x , y , and z . Nevertheless, the dynamic is linear in the infinite-dimensional vector space $D(\mathbb{R}^3)$ of distributions over \mathbb{R}^3 . Although $D(\mathbb{R}^3)$ is an unwieldy state-space, the dynamics there might be well approximated by a finite truncation of its modes.

2.7.6.1 Ruelle–Frobenius–Perron and Koopman operators

The preceding operator formalism applies, in principle at least. The question, of course, is, Is it practical and does it lead to constructive consequences? Let’s see. The right eigenvector is either $|0_G\rangle$ or $|1_T\rangle$ with $T = e^{\tau G}$ as the Ruelle–Frobenius–Perron transition operator [208, 156]. Equivalently, it is also $\boldsymbol{\pi}$, the stationary distribution, with support on attracting subsets of \mathbb{R}^3 in the case of the Lorenz dynamic. The corresponding left-eigenvector $\mathbf{1}$, either $\langle 0_G|$ or $\langle 1_T|$, is uniform over the space. Other modes of the operator’s action, according to the eigenvalues and left and right eigenvectors and generalized eigenvectors, capture the decay of arbitrary distributions on \mathbb{R}^3 .

The meromorphic spectral methods developed above give a view of the Koopman operator and Koopman modes of nominally nonlinear dynamical systems [34] that is

complementary to the Ruelle–Frobenius–Perron operator. The Koopman operator K is the adjoint—in the sense of vector spaces, not inner product spaces—of the Ruelle–Frobenius–Perron operator T : effectively the transpose $K = T^\top$. Moreover, it has the same spectrum with only right and left swapping of the eigenvectors and generalized eigenvectors.

The Ruelle–Frobenius–Perron operator T is usually associated with the evolution of probability density, while the Koopman operator K is usually associated with the evolution of linear functionals of probability density. The duality of perspectives is associative in nature: $\langle f | (T^n | \rho_0 \rangle)$ corresponds to the Ruelle–Frobenius–Perron perspective with T acting on the density ρ and $(\langle f | T^n) | \rho_0 \rangle$ corresponds to the Koopman operator $T^\top = K$ acting on the observation function f . Allowing an observation vector $\vec{f} = [f_1, f_2, \dots, f_m]$ of linear functionals, and inspecting the most general form of K^n given by Eq. (2.25) together with the generalized eigenvector decomposition of the projection operators of Eq. (2.39), yields the most general form of the dynamics in terms of Koopman modes. Each Koopman mode is a length- m vector-valued functional of a Ruelle–Frobenius–Perron right eigenvector or generalized eigenvector.

Both approaches suffer when their operators are defective. Given the meromorphic calculus’ ability to work around a wide class of such defects, adapting it the Ruelle–Frobenius–Perron and Koopman operators suggests that it may lift their decades-long restriction to only analyzing highly idealized (e.g., hyperbolic) chaotic systems.

2.7.6.2 Eigenvalues from a time series

Let’s explore an additional benefit of this view of the Ruelle–Frobenius–Perron and Koopman operators, by proposing a novel method to extract the eigenvalues of a nominally nonlinear dynamic. Let $O_N(f, z)$ be (z^{-1} times) the z -transform [32, pp. 257–262] of a length- N sequence of τ -spaced type- f observations of a dynamical system:

$$\begin{aligned} O_N(f, z) &\equiv z^{-1} \sum_{n=0}^N z^{-n} \langle f | T^n | \rho_0 \rangle \\ &\rightarrow_{N \rightarrow \infty} \langle f | (zI - T)^{-1} | \rho_0 \rangle \\ &= \sum_{\lambda \in \Lambda_T} \sum_{m=0}^{\nu_\lambda - 1} \frac{\langle f | T_{\lambda, m} | \rho_0 \rangle}{(re^{i\omega} - \lambda)^{m+1}} , \end{aligned}$$

as $N \rightarrow \infty$ for $|z| = r > 1$. Note that $\langle f|T^n|\rho_0\rangle$ is simply the f -observation of the system at time $n\tau$, when the system started in state ρ_0 . We see that this z -transform of observations automatically induces the resolvent of the hidden linear dynamic. If the process is continuous-time, then $T = e^{\tau G}$ implies $\lambda_T = e^{\tau\lambda_G}$, so that the eigenvalues should shift along the unit circle if τ changes; but the eigenvalues should be invariant to τ in the appropriate τ -dependent conformal mapping of the inside of the unit circle of the complex plane to the left half complex plane. Specifically, for any experimentally accessible choice of inter-measurement temporal spacing τ , the fundamental set of continuous time eigenvalues Λ_G can be obtained from $\lambda_G = \frac{1}{\tau} \ln \lambda_T$, where each $\lambda_T \in \Lambda_T$ is extrapolated from $c/(re^{i\omega} - \lambda_T)^n$ curves fit to $O_N(f, re^{i\omega})$ for $c \in \mathbb{C}$, large N , and fixed r .

The square magnitude of $O_N(f, z)$ is related to the power spectrum generated by f -type observations of the system. Indeed, the power spectrum generated by any type of observation of a nominally nonlinear system is a direct fingerprint of the eigenspectrum and resolvent of the hidden linear dynamic. This suggests many opportunities for inferring eigenvalues and projection operators from frequency-domain transformations of a time series.

2.8 Conclusion

The original, abstract spectral theory of normal operators rose to central importance when, in the early development of quantum mechanics, the eigenvalues of Hermitian operators were detected experimentally in the optical spectra of energetic transitions of excited electrons. We extended this powerful theory by introducing the meromorphic functional calculus, providing the spectral theory of *nonnormal* operators. Our straightforward examples suggest that the spectral properties of these general operators should also be experimentally accessible in the behavior of complex—open, strongly interacting—systems. We see a direct parallel with the success of the original spectral theory of normal operators as it made accessible the phenomena of the quantum mechanics of closed systems. This turns on nondiagonalizability and appreciating how ubiquitous it is.

Nondiagonalizability has consequences for settings as simple as counting, as shown in

§ 2.7.3. Moreover, there we found that nondiagonalizability can be robust. The Drazin inverse, the negative-one power in the meromorphic functional calculus, is quite common in the nonequilibrium thermodynamics of open systems, as we showed in § 2.7.4. Finally, we showed that the spectral character of nonnormal and nondiagonalizable operators manifests itself physically, as illustrated by Figs. 2.4c and 2.4d of § 2.7.5.

From the perspective of functional calculus, nonunitary time evolution, open systems, and non-Hermitian generators are closely related concepts since they all rely on the manipulation of nonnormal operators. Moreover, each domain is gaining traction. Nonnormal operators have recently drawn attention, from the nonequilibrium thermodynamics of nanoscale systems [92] to large-scale cosmological evolution [20]. In another arena entirely, complex directed networks [178] correspond to nonnormal and not-necessarily-diagonalizable weighted digraphs. There are even hints that nondiagonalizable network structures can be optimal for implementing certain dynamical functionality [180]. The opportunity here should be contrasted with the well established field of spectral graph theory [41] that typically considers consequences of the spectral theorem for normal operators applied to the symmetric (and thus normal) adjacency matrices and Laplacian matrices. It seems that the meromorphic calculus and its generalized spectral theory will enable a *spectral weighted digraph theory* beyond the purview of current spectral graph theory.

Even if the underlying dynamic is diagonalizable, particular questions or particular choices of observable often *induce* a nondiagonalizable hidden linear dynamic. The examples already showed this arising from the simple imposition of counting or assuming a Poissonian dynamic. In more sophisticated examples, we recently found nondiagonalizable dynamic structures in quantum memory reduction [203] and classical complexity measures [60].

Our goal has been to develop tractable, exact analytical techniques for nondiagonalizable systems. We did not discuss numerical implementation of algorithms that naturally accompany its practical application. Nevertheless, the theory does suggest new algorithms—for the Drazin inverse, projection operators, power spectra, and more. Guided by the meromorphic calculus, such algorithms can be made robust despite the common knowledge

that numerics with nondiagonalizable matrices is sensitive in certain ways.

The meromorphic calculus complements attempts to address nondiagonalizability, e.g., via pseudospectra [253, 254]. It also extends and simplifies previously known results, especially as developed by Dunford [70]. Just as the spectral theorem for normal operators enabled much theoretical progress in physics, we hope that our generalized and tractable analytic framework yields rigorous understanding for much broader classes of complex system. Importantly, the analytic framework should enable new *theory* of complex systems beyond the limited purview of numerical investigations.

While the infinite-dimensional theory is in principle readily adaptable from the present framework, special care must be taken to guarantee a similar level of tractability and generality. Nevertheless, even the finite-dimensional theory enables a new level of tractability for analyzing not-necessarily-diagonalizable systems, including nonnormal dynamics. Future work will take full advantage of the operator theory, with more emphasis on infinite-dimensional systems and continuous spectra. Another direction forward is to develop creation and annihilation operators within nondiagonalizable dynamics. In the study of complex stochastic information processing, for example, this would allow analytic study of infinite-memory processes generated by, say, stochastic pushdown and counter automata [65, 250, 166, 63]. In a physical context, such operators may aid in the study of open quantum field theories. One might finally speculate that the Drazin inverse will help to tame the divergences that arise there.

Chapter 3

Spectral Simplicity of Hidden Complexity: Cracking the Nondiagonalizable Metadynamics of Prediction

3.1 Chapter Overview

To render simple answers for complex systems, we demonstrate that almost all questions that we know how to ask about the structural complexity of a stochastic process reduce to a linear algebraic framing of time evolution if we use the appropriate hidden Markov model (HMM) generator of the process. An important lesson is that different types of questions—those of correlation, predictability, or predictive burden—invoke different classes of HMM representations. Corresponding answers are simple functions of the class-appropriate transition dynamic. Helpfully, the resulting expressions reveal unexpected similarities between complexity measures typically thought to capture distinct informational and computational properties. Along the way, the implicit structure of optimal prediction is made clear.

The theoretical advance enabling this new level of tractability is the recently introduced *meromorphic functional calculus*, which allows proper analysis of functions of nondiagonalizable linear operators even when the poles and zeros of the function coincide with the spectrum of

the operator. By avoiding spurious infinities, this results in the first closed-form expressions for many complexity measures, couched either in terms of the Drazin inverse—which is the negative-one power of a singular operator in the meromorphic functional calculus—or in terms of the eigenvalues and projection operators of the appropriate transition dynamic. We also establish special properties of the projection operators that demonstrate how they capture the organization of subprocesses within a complex system.

We apply the theoretical advances to derive closed-form expressions for correlation functions, finite-length Shannon entropy rate approximates, asymptotic entropy rate, excess entropy, transient information, transient and asymptotic state uncertainty, and synchronization information. The spectral decomposition lends intuition towards generalizing concepts of complexity. And so, we briefly address several more informative extensions of familiar complexity measures.

3.2 Introduction

Hungry for understanding, we open our senses—eyes, probes, and colossal detectors—to both the natural and designed world around us in search of structure. The data deluge ensues. How do we make sense of it all? Is it noise or is there hidden structure to be inferred [56]? How much structure exists in the data [62]? Supposing there is predictable structure, what resources must be allocated to predict what is predictable?

Herein, we investigate hidden structure and the complexities associated with prediction. Following up on our recent progress, we uncover new relationships among supposedly different challenges [60] and identify the importance of a generalized spectral theory in answering each challenge [198], yielding closed-form solutions that deliver elegance and insight. We show that different questions regarding correlation, information, or prediction induce their own structure as they sieve through the data. Resulting transitions among hidden variables summarize symmetry breaking, synchronization, and information processing. Each of these metadynamics (for example, the dynamic among belief states), though, is built up from the original observed dynamics.

Complex systems—that is, many-bodied systems with strong interactions—are usually

observed through a time-series of low-bandwidth feature detectors, so that their hidden structure is only revealed over time. Because the instantaneous observations cannot capture the full resolution of each degree of freedom, the observed time-series is often stochastic and non-Markovian, with long-range correlations. Even finite models can appear quite complex: despite admitting finite descriptions, stochastic processes with sofic support exhibit infinite-range dependencies among observable random variables. While such infinite-correlation processes are legion in complex physical and biological systems, analyzing them is generally appreciated as difficult, if not impossible. Nonetheless, we show that their complexity measures have simple closed-form expressions in both the time-asymptotic and finite-time regimes.

The major shift in perspective allowing this new level of tractability begins by recognizing that—beyond their ability to *generate* many sophisticated processes of interest—hidden Markov models (HMMs) can be treated as exact *mathematical objects* in the analysis of processes they represent. Crucially—even when dealing with nonlinear processes—most questions that we tend to ask about a process imply a linear transition dynamic over *some* hidden state space: something happens, then *something* evolves linearly in time, then we take a snapshot of our favorite characteristic about a process. Let’s call this broad type of questioning ‘cascading’, since the influence of the initial preparation cascades through state-space as time (or position) evolves, affecting our final measurement. Alternatively, some (more sophisticated?) questions involve accumulating such cascading events. The linear algebra underlying all such questions is highlighted in Table 3.1 in terms of the appropriate discrete-time transition operator T or continuous-time generator of time evolution G .

The challenge behind making our questions crumble into the simplicity of linear algebra then is relegated to identifying the appropriate hidden state space. The appropriate state space depends on the genre of the question. In this chapter, we describe closed-form solutions for complexity measures of a process. We find that many questions of complexity fall into three sub-genres and, for each of these, we identify the appropriate linear dynamic and closed-form expressions for the most important questions in the genre.

The linear algebra inside our many questions

	Discrete time	Continuous time
Cascading	$\langle \cdot T^L \cdot \rangle$	$\langle \cdot e^{tG} \cdot \rangle$
Accumulating	$\langle \cdot (\sum_L T^L) \cdot \rangle$	$\langle \cdot (\int e^{tG} dt) \cdot \rangle$

Table 3.1: Once we identify the hidden linear dynamic behind our questions, most questions we care to ask are either of the *cascading* or *accumulating* type.

This categorization of complexity measures and their implied linear transition dynamic is given in Table 3.2. This table will become more understandable as we progress, and we will return to an enhanced version of it in our conclusion.

Layering observables $x \in \mathcal{A}$ on top of the transitions between these hidden states $\sigma \in \mathcal{S}$ then turns this abstraction into a concrete hidden Markov model with observation-labeled transition matrices $\{T^{(x)} : T_{i,j}^{(x)} = \Pr(x, \sigma_j | \sigma_i)\}_{x \in \mathcal{A}}$ that sum to the net row-stochastic state-to-state transition matrix T .

Because complex processes have highly structured directional transition dynamics, we encounter the full richness of matrix algebra in the analysis of HMMs. Normal and diagonalizable restrictions, so familiar in mathematical physics, are immediately shrugged off. Notably, the analysis of complex systems *induces* a nondiagonalizable meta-dynamics, even if the observed dynamics are diagonalizable in their underlying state-space.

The new richness of nondiagonalizable dynamics presents a technical challenge in the analysis of complex processes described by HMMs. A calculus for functions of nondiagonalizable operators becomes a commonplace necessity in the exact analysis of complex processes. Moreover, complexity measures naively appear to implore us to perform illegal operations such as the inverse of a singular operator. In reality, the illegal operation is skirted because the complexity measures have us only skim off the excess transient behavior of an infinite underbelly. But what is the calculus of only accumulating the contributions of select modes of functions of a nondiagonalizable operator?

Motivated by such questions, we have recently developed a meromorphic functional calculus in Ref. [198] that enables much of the theory herein for the analysis of complex

processes described by HMMs. In addition to enabling a general spectral weighted digraph theory, these techniques are applied more specifically here to the challenges of prediction. The results developed here extend and explain those given in Ref. [60], which introduced the basic methods by focusing on closed-form expressions for various measures of intrinsic computation and applying them to several complex systems.

This meromorphic functional calculus, summarized in more detail later, deals with functions of nondiagonalizable operators when poles (or zeros) of the function coincide with poles of the operator’s resolvent—poles that appear precisely at the eigenvalues of the matrix. Pole–pole and pole–zero interactions transform the complex-analysis residues within the functional calculus. One result is that the negative-one power of a singular operator exists in the meromorphic functional calculus, and is the *Drazin inverse*.

In the following, we give the first closed-form expressions for many complexity measures in wide use—many of which turn out to be expressed most concisely in terms of a Drazin inverse. Furthermore, via spectral decomposition, we give insight into sub-processes of complex behavior in terms of the projection operators of the appropriate transition dynamic.

To get started, the following sections (§3.3 through §7.10.1) briefly review relevant background in stochastic processes and their complexities and the HMMs that generate them. Several classes of HMMs are discussed in §7.10.1. Mixed-state presentations (MSPs), which are HMM generators of a process that also track distributions induced by observation, are reviewed in §3.5 since they are relevant to complexity measures with an information-theoretic framing. We then, in §3.6, show how each complexity measure reduces to linear algebra upon the use of the appropriate HMM restriction for the genre of the question. To make progress at this point, we summarize the meromorphic functional calculus in §3.7. Several of its mathematical implications are discussed in relation to projection operators in §3.8 and a spectral weighted digraph theory in §3.9. With that all set out, in §3.10 and §3.12 we finally derive the promised closed-form complexities of a process. Common simplifications for special cases are discussed in §3.11 and §3.12.5. A suite of examples in §3.14 and §3.15 ground the theoretical developments, and are complemented with an

Question genres and their hidden linear dynamic

Genre	Example Questions	Implied linear transition dynamic
Observational	Correlations: $\gamma(L)$; Power spectra: $P(\omega)$	Transition matrix T of any HMM
Predictability	Myopic entropy: $h_\mu(L)$; Excess entropy: \mathbf{E} , $E(\omega)$	Transition matrix W of MSP of any HMM
Prediction	Causal synchrony: $\mathcal{H}^+(L)$, \mathbf{S} , $S(\omega)$	Transition matrix \mathcal{W} of MSP of ϵ -machine
\vdots	\vdots	\vdots

Table 3.2: Several genres of questions about the complexity of a process are given in the left column of the table in order of increasing sophistication. Each genre implies a different linear transition dynamic. *Observational* questions are about the superficial dynamic. *Predictability* questions are about the observation-induced dynamic over distributions (over states that generate the superficial dynamic). *Prediction* questions are about the dynamic over distributions over causally-equivalent histories of a process. Other question genres, like generation or quantum compressibility of a process, invoke other hidden linear representations not discussed here—nevertheless the tools developed here extend in obvious ways once the appropriate hidden linear dynamic is identified.

in-depth pedagogical example worked out in §3.16. Ultimately, we conclude with a brief retrospect and an eye towards future applications.

3.3 Structured Processes and their Complexities

We describe a system of interest in terms of its observed behavior, following the approach of computational mechanics, as reviewed in Ref. [55]. Again, a *process* is the collection of behaviors that the system produces and their probabilities of occurring. A process’s behaviors are described via a bi-infinite chain of random variables, denoted by capital letters $\dots X_{t-2} X_{t-1} X_t X_{t+1} X_{t+2} \dots$. A realization is indicated by lowercase letters $\dots x_{t-2} x_{t-1} x_t x_{t+1} x_{t+2} \dots$. We assume the value x_t belongs to a discrete alphabet \mathcal{A} . We work with blocks $X_{t:t'}$, where the first index is inclusive and the second exclusive: $X_{t:t'} = X_t \dots X_{t'-1}$.

A process’s probabilistic specification is a density over these chains: $\mathbb{P}(X_{-\infty:\infty})$. Practically, we work with finite blocks and their probability distributions: $\Pr(X_{t:t'})$. To simplify the development in the following, we primarily analyze stationary, ergodic processes: those

for which $\Pr(X_{t:t+L}) = \Pr(X_{0:L})$ for all $t \in \mathbb{Z}$. In such cases, we only need to consider a process's length- L *word distributions* $\Pr(X_{0:L})$.

3.3.1 Directly observable organization

A first step in understanding how processes express themselves is to analyze correlations among observables. Pairwise correlation in a sequence of observables is often summarized by the *autocorrelation function*:

$$\gamma(L) = \langle \overline{X}_t X_{t+L} \rangle_t ,$$

where the bar above X_t denotes its complex conjugate, and the angled brackets denote an average over all positions $t \in \mathbb{Z}$. Alternatively, structure in a stochastic process is often summarized by the *power spectral density*, also referred to more simply as the *power spectrum*:

$$P(\omega) = \lim_{N \rightarrow \infty} \frac{1}{N} \left\langle \left| \sum_{L=1}^N X_L e^{-i\omega L} \right|^2 \right\rangle ,$$

where $\omega \in \mathbb{R}$ is the angular frequency [242]. It is not always fully appreciated however that power spectra only capture pairwise correlation. Indeed, it is straightforward to show that the power spectrum is the windowed Fourier transform of the autocorrelation function. Thus, power spectra show how pairwise correlations are distributed among frequencies. Power spectra are common in signal processing, both in technological settings and physical experiments. As a physical example, diffraction patterns are the power spectra of a sequence of structure factors.

Green–Kubo transport coefficients are another important example measure of observable organization, but are more application-specific [100, 284]. Such coefficients epitomize the idea that dissipation depends on the structure of correlation. They usually appear in the form of integrating the autocorrelation of *derivatives* of observables, but a change of observables turns this into an integration of a standard autocorrelation function. Green–Kubo transport coefficients then involve the limit $\lim_{\omega \rightarrow 0} P(\omega)$ for the process of appropriate observables.

Unfortunately, correlation functions and power spectra give an impoverished view of a process's structural complexity, since they only consider ensemble averages over

pairwise events. Moreover, creating a list of higher-order correlations is an impractical way to summarize complexity, as seen in the connected correlation functions of statistical mechanics [25].

3.3.2 Intrinsic predictability

Information metrics can involve all orders of correlation and thus help to go beyond pairwise correlation in understanding how the information expressed thus far will affect the predictability at some later time or place. Information theory, as developed for general complex processes [62], provides a set of quantities that capture such properties using variants of Shannon's entropy $H[\cdot]$ and mutual information $I[\cdot; \cdot]$ [46]. Each quantity answers a specific question about a process's organization. For example:

- How much information is contained in the words generated? The *block entropy* [62]:

$$H(L) = - \sum_{w \in \mathcal{A}^L} \Pr(w) \log_2 \Pr(w).$$

- How random is a process? Its *entropy rate* [130]:

$$h_\mu = \lim_{L \rightarrow \infty} H(L)/L.$$

- How is this irreducible randomness approached? Via the *myopic entropy rates* [62]:

$$h_\mu(L) = H[X_0 | X_{1-L} \dots X_{-1}].$$

- How much of the future can be predicted? Its *excess entropy* [62]:

$$\mathbf{E} = I[X_{-\infty:0}; X_{0:\infty}].$$

- How much information must be extracted from the system to know how predictable it is and so see the process's intrinsic randomness? Its *transient information* [61]:

$$\mathbf{T} = \sum_{L=0}^{\infty} [\mathbf{E} + h_\mu L - H(L)].$$

We later introduce the *excess entropy spectrum* $\mathcal{E}(\omega)$, which completely, yet concisely, summarizes the structure of myopic entropy reduction, in a similar way to how the power spectrum completely describes autocorrelation. But while the power spectrum only summarizes pairwise linear correlation, the excess entropy spectrum captures all orders of nonlinear dependencies among random variables, making it a strongly qualified probe of hidden structure.

3.3.3 The overhead of prediction

The nuanced measures of a system’s predictability, as just enumerated, certainly say a lot about its intrinsic computational structure. But they leave open the question of the complexity associated with actually *doing* the prediction. To answer this challenge, a complementary set of measures address the inherent complexity of actually *predicting* what is *predictable*. How cryptic is a process?

Computational mechanics investigates optimal prediction via the structure of transitions among causal states, summarized by the process’s ϵ -machine [55]. A *causal state* $\sigma \in \mathcal{S}^+$ is an equivalence class of observable histories that all yield the same probability distribution over observable futures. Therefore, knowing a process’s current causal state, that $\mathcal{S}_0^+ = \sigma$, is sufficient for optimal prediction.

Computational mechanics provides a set of quantities that capture the overhead of synchronization and prediction, again using variants of Shannon’s entropy and mutual information. Each quantity answers a specific question about an observer’s burden of prediction. For example:

- How much historical information must be stored for optimal prediction? The *statistical complexity* [64]:

$$C_\mu = H[\mathcal{S}_0^+].$$

- How unpredictable is the causal state of the process upon observing it for a duration L ? The *myopic causal state uncertainty* [62]:

$$\mathcal{H}^+(L) = H[\mathcal{S}_0^+ | X_{-L} \dots X_{-1}].$$

- How much information must an observer extract to synchronize to the causal state?

The optimal predictor’s *synchronization information* [62]:

$$\mathbf{S} = \sum_{L=0}^{\infty} \mathcal{H}^+(L).$$

We later introduce the *optimal synchronization spectrum* $S(\omega)$, which completely, yet concisely, summarizes the structure of state-uncertainty reduction, in a similar way to how the power spectrum completely describes autocorrelation, or how the excess entropy spectrum completely describes myopic entropy reduction. All of the above quantifiers of optimal prediction can be found from the optimal synchronization spectrum.

These complexities associated with an observer’s burden in optimally predicting a process have tremendous practical relevance for artificial and biological intelligent agents taking advantage of structured stochastic environments, for prey avoiding easy prediction, and for taking advantage of predictable fluctuations in the stock market.

The prediction game has many natural generalizations. For example, since optimal prediction often requires infinite resources, sub-optimal prediction is of great practical interest. Fortunately, there are principled ways to investigate the tradeoffs between predictive accuracy and computational burden [240, 48, 167]. As another example, optimal prediction under constraints of noisy or irregular observations can be investigated with a properly generalized framework [75]. Blending the existing tools, resource-limited prediction under such observational constraints could also be investigated. In all of these cases, information metrics similar to those listed above are crucial to understanding and quantifying the tradeoffs encountered in prediction.

Now that we have highlighted the difference between prediction and predictability, we can appreciate that some processes are more cryptic than others. *Crypticity* χ , it turns out, can be quantified: It is the difference between the observer’s stored information C_μ , necessary for prediction, and the mutual information \mathbf{E} shared between past and future observables [160]. Crypticity therefore contrasts predictable information content (\mathbf{E}) with the observer’s minimal overhead of stored memory (C_μ) necessary to *do* the prediction. To actually predict what is predictable, an optimal predictor is taxed by the crypticity of the process.

3.3.4 Generative complexities

How does a physical system intrinsically compute its output process? This can depend on many particulars. Sometimes there is vast redundancy while other times Occam’s razor or resource constraints seem to have shaved the mechanism down to its minimal generative structure. And different pressures imply different senses of minimality: minimal entropy generators turn out to be distinct from minimal state generators.

Given *any* generative model [59, 61] \mathcal{M} with state-set \mathcal{S} , we can consider the corre-

sponding *myopic state uncertainty* given L sequential observations:

$$\mathcal{H}(L) = H[\mathcal{S}_0 | X_{-L:0}] ,$$

with asymptotic uncertainty $\mathcal{H} \equiv \lim_{L \rightarrow \infty} \mathcal{H}(L)$, and *excess synchronization information*:

$$\mathbf{S}' = \sum_{L=0}^{\infty} [\mathcal{H}(L) - \mathcal{H}] .$$

Such quantities are relevant even when the observer can never fully synchronize to the generative state, i.e., even when $\mathcal{H} > 0$. Each generative model has its own state complexity:

$$C(\mathcal{M}) = \mathcal{H}(0) .$$

Despite the fact that many different models can generate the same process, we can still define useful bounds on the complexity of any generator for the process. For example: The minimal *generative complexity* [150] $C_g = \min_{\mathcal{M}} C(\mathcal{M})$ is the minimal state-information a classical physical process must store to generate its future. Complexities of various genres bound each other: E.g. [150],

$$\mathbf{E} \leq C_g \leq C_{\mu} ;$$

i.e., the information that is *predictable* about the future is less than or equal to the information necessary to *produce* the future, which is less than or equal to the information necessary to *predict* the future.

A further nuance to this story has recently come to light: quantum physics can produce a stochastic process of classical observables using less memory than a classical process can. The minimal quantum memory C_q (quantified via von Neumann entropy rather than classical Shannon entropy) of generation is not known in general, but has been upper-bounded by several recent investigations [276, 102, 157, 185, 203]. In summary, it appears that:

$$\mathbf{E} \leq C_q \leq C_g \leq C_{\mu} ;$$

i.e., predictable information \leq quantum memory necessary for generation \leq classical memory necessary for generation \leq classical memory necessary for optimal prediction.

3.4 Mealy Hidden Markov Models

A unifying mathematical representation can be used to describe a wide class of process generators: the edge-labeled HMM, also known as a Mealy HMM [59]. (This is as opposed to the class-equivalent state-labeled HMMs, also known as Moore HMMs [11, 259, 261].)

Definition 2. A finite-state, edge-labeled hidden Markov model (HMM) \mathcal{M} consists of:

- A finite set of hidden states $\mathcal{S} = \{\sigma_1, \dots, \sigma_{|\mathcal{S}|}\}$.
- A finite output alphabet \mathcal{A} .
- A set of $|\mathcal{S}| \times |\mathcal{S}|$ symbol-labeled transition matrices $\{T^{(x)}\}_{x \in \mathcal{A}}$, where $T_{i,j}^{(x)} = \Pr(x, \sigma_j | \sigma_i)$ is the probability of transitioning from state σ_i to state σ_j and emitting symbol x .
The corresponding overall state-to-state transition matrix is the row-stochastic matrix $T = \sum_{x \in \mathcal{A}} T^{(x)}$.
- An initial distribution over hidden states: $\eta_0 = (\Pr(\mathcal{S}_0 = \sigma_1), \Pr(\mathcal{S}_0 = \sigma_2), \dots, \Pr(\mathcal{S}_0 = \sigma_M))$, where \mathcal{S}_t is the random variable for the hidden state at time t .

\mathcal{M} can be written as the 4-tuple: $\mathcal{M} = (\mathcal{S}, \mathcal{A}, \{T^{(x)}\}_{x \in \mathcal{A}}, \eta_0)$. In automata theory, a finite-state HMM may be called a probabilistic nondeterministic finite automaton (probabilistic NFA) [232].

Since the transition dynamics of finite-state models are represented by transition matrices amenable to linear algebra, it is useful to use a bra-ket notation where bras $\langle \cdot |$ are row vectors and kets $|\cdot\rangle$ are column vectors. One benefit is immediate recognition of object type. For example, on the one hand, any expression that forms a closed bra-ket pair—either $\langle \cdot | \cdot \rangle$ or $\langle \cdot | \cdot | \cdot \rangle$ —is a scalar quantity and commutes as a unit with anything. On the other hand, when useful, an expression of the ket-bra form $|\cdot\rangle \langle \cdot |$ can be interpreted as a matrix.

Row-stochasticity of T means that all rows sum to unity. Introducing $|\mathbf{1}\rangle$ as the column vector of all 1s, this can be restated as:

$$T |\mathbf{1}\rangle = |\mathbf{1}\rangle .$$

The reader may also recognize this as an eigen-equation: the all-ones vector $|\mathbf{1}\rangle$ is always a right eigenvector of T associated with the eigenvalue of unity.

When the internal Markov transition matrix T is irreducible, there is a unique stationary state distribution π determined by:

$$\langle \pi | T = \langle \pi | \text{ ,}$$

with the further condition that π be normalized in probability: i.e., $\langle \pi | \mathbf{1} \rangle = 1$. This will again be recognized as an eigen-equation: the stationary distribution π over the hidden states is the left eigenvector of T associated with the eigenvalue of unity.

When choosing to describe a stationary process, as we will often do in this chapter, we will choose the initial distribution over hidden states to be the stationary one: $\eta_0 = \pi$. Choosing another η_0 can be useful in many contexts, but yields non-stationary processes which we avoid for now for simplicity of the exposition.

An HMM \mathcal{M} describes process's behaviors as a formal language $\mathcal{L} \subseteq \mathcal{A}^*$ of allowed realizations. Moreover, \mathcal{M} succinctly describes a process's word distribution $\Pr(w)$ over all words $w \in \mathcal{L}$. (Appropriately, \mathcal{M} also gives zero probability to words outside of the language of the process: $\Pr(w) = 0$ for all $w \in \mathcal{L}^c$, \mathcal{L} 's complement.) Specifically, the stationary probability of observing a particular length- L word $w = x_0 x_1 \dots x_{L-1}$ is given by:

$$\Pr(w) = \langle \pi | T^{(w)} | \mathbf{1} \rangle \text{ ,} \quad (3.1)$$

where $T^{(w)} \equiv T^{(x_0)} T^{(x_1)} \dots T^{(x_{L-1})}$.

More generally, given a nonstationary distribution over hidden states, the subsequent probability of a word is:

$$\Pr(X_{t:t+L} = w | \mathcal{S}_t \sim \eta) = \langle \eta | T^{(w)} | \mathbf{1} \rangle \text{ ,} \quad (3.2)$$

where $\mathcal{S}_t \sim \eta$ means that the random variable \mathcal{S}_t is distributed as η . Such a generalized concept is actually hugely necessary: for example, most observations induce a non-stationary distribution over hidden states. Tracking these observation-induced distributions will be the role of the mixed-state presentation, introduced shortly.

3.4.1 Unifilar HMMs

A highly useful additional property that one can require of a HMM is that of *unifilarity* which guarantees that, given a start state and a sequence of observations, there is a unique path through the internal states [4]. This, in turn, allows one to directly translate properties of the internal Markov chain into properties of the observed behavior. General (non-unifilar) HMMs, in contrast, typically have an exponentially growing number of possible state paths for a linearly growing observation sequence, obviating most all quantitative connections between internal dynamics and observations.

Definition 3. A finite-state, edge-labeled, unifilar HMM (*uHMM*) is a finite-state, edge-labeled HMM with the following property:

- Unifilarity: For each state $\sigma \in \mathcal{S}$ and each symbol $x \in \mathcal{A}$ there is at most one outgoing edge from state σ that outputs symbol x .

In automata theory, a uHMM may be called a probabilistic deterministic finite automaton (probabilistic DFA) [232].

3.4.2 Minimal Unifilar HMMs

The minimal entropy uHMM of a process turns out to be the same as its minimal state uHMM which turns out to be the ϵ -machine of computational mechanics [55]. The states of the ϵ -machine, called causal states, are equivalence classes of histories that yield the same predictions for the future. More explicitly, two histories \overleftarrow{x} and $\overleftarrow{x'}$ map to the same causal state $\epsilon(\overleftarrow{x}) = \epsilon(\overleftarrow{x'}) = s$ iff $\Pr(\overrightarrow{X}|\overleftarrow{x}) = \Pr(\overrightarrow{X}|\overleftarrow{x'}) = \Pr(\overrightarrow{X}|s)$.

Converting from a uHMM to its corresponding ϵ -machine uses a variant of the familiar state-minimization algorithm in automata theory [112].

3.4.3 The Finitary Stochastic Process Hierarchy

The finite-state models in these classes form a hierarchy in terms of the processes they can generate [54]: $\text{Processes}(\epsilon\text{-machines}) = \text{Processes}(\text{uHMMs}) \subset \text{Processes}(\text{HMMs})$. That is, finite HMMs generate a strictly larger class of stochastic processes than finite uHMMs. The class of processes generated by finite uHMMs, though, is the same as generated by finite ϵ -machines.

3.4.4 Continuous-time HMMs

In continuous time, transition *rates* are more appropriate fundamental entities than transition probabilities. Continuous-time HMMs can often be obtained as the $\Delta t \rightarrow 0$ limit of an edge-labeled HMM whose edges are parameterized by Δt , but the most natural continuous-time HMM has a continuous-time generator G of time evolution over hidden states, with observables emitted as deterministic *functions of the Markov chain*, $f : \mathcal{S} \rightarrow \mathcal{A}$. Functions of a Markov chain [27, 26, 4, 251] are a sub-class of state-emitting HMMs, and continuous versions of these are yet more particular. Yet, this class of processes is as expressive as any physically realistic continuous-time HMM.

The rows of G each sum to zero, which is the continuous-time analogue of probability conservation: over a finite time interval t , marginalizing over all possible observations, the row-stochastic state-to-state transition dynamic is

$$T_{t_0 \rightarrow t_0+t} = e^{tG} .$$

The function of the Markov chain can also be given a matrix manifestation. For this purpose—with $\delta_{x,f(s)}$ as a Kronecker delta, $|\delta_s\rangle$ as the column vector of all zeros except for a one at the position for state s , and $\langle\delta_s|$ as the transpose $\langle\delta_s| = (|\delta_s\rangle)^\top$ —we introduce the continuous time observation matrices as the projectors:

$$\Gamma_x = \sum_{s \in \mathcal{S}} \delta_{x,f(s)} |\delta_s\rangle \langle\delta_s| ,$$

which sum to the identity: $\sum_{x \in \mathcal{A}} \Gamma_x = I$.

3.5 Mixed State Presentations

We have just discussed HMM representations of a process, and have indicated that several classes of HMMs (e.g., nonunifilar, unifilar, ϵ -machine) can be used to generate the same process. Moreover, within either the unifilar or nonunifilar model class, there are extensive families of HMMs that describe the same process. This flexibility suggests that we can create a HMM generator of a process that also tracks important supplementary information *in* the hidden states or through well-crafted dynamics *over* the hidden states. The HMMs

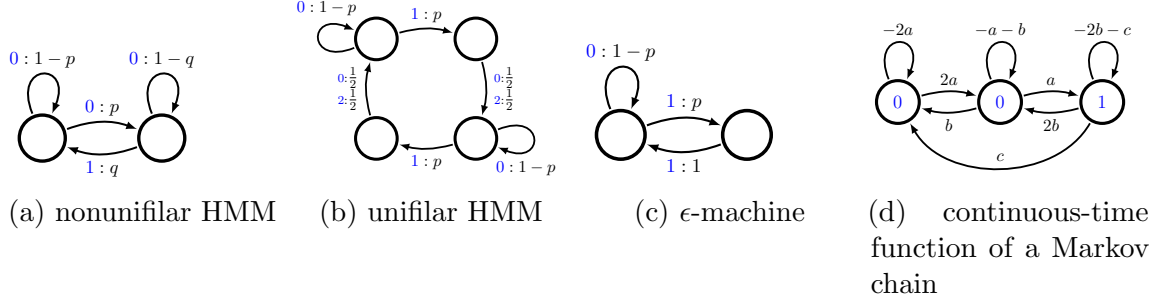


Figure 3.1: Simple examples to illustrate distinct classes of hidden Markov model, as represented by their corresponding automaton. For any setting of the transition probabilities $p, q \in (0, 1)$ and transition rates $a, b, c \in (0, \infty)$, each HMM generates an observable stochastic process over its alphabet $\mathcal{A} \subset \{0, 1, 2\}$ —the latent states themselves are not directly observable from the output process (hence, ‘hidden’ states). (a) Simple nonunifilar source: nonunifilar because two edges going out from the same state generate the same output. (b) (Non-minimal) unifilar HMM. (c) The minimal unifilar HMM for the stochastic process generated. (d) The generator of a continuous-time stochastic process.

we are about to discuss, the *mixed-state presentations* (MSPs), are a prime example of this dualistic functionality. In particular, an MSP generates a process while tracking the observation-induced distribution over the states of another generative model of the process.

As we will show in § 3.6, tracking distributions over the states of another generator makes the MSP an ideal algebraic object for closed-form complexities involving conditional entropies (which invoke conditional probabilities). We have seen in § 3.3.2 and § 3.3.3 that many complexity measures of predictability and predictive burden are indeed framed as conditional entropies, and so MSPs will be relevant to their closed-form solutions.

Mixed states were already implicit in Ref. [27], introduced in their modern form by Ref. [259], and have been used recently, e.g., in [58, 76]. Most of these efforts, however, have used mixed-states in the specific context of the *synchronizing MSP* (\mathcal{S} -MSP), which we discuss in detail later. A greatly extended development of mixed-state dynamics appears in Ref. [75]. Different information-theoretic questions will invoke different mixed-state dynamics, each of which is a unifilar presentation. Utilizing the mathematical methods developed here, we find that the desired closed-form solution are often simple functions of the transition dynamic of the appropriate MSP. The spectral character of the relevant MSP controls the behavior of information theoretic quantities. Here, we review only the

subset of the mixed-state theory required in the following.

Consider a HMM presentation $\mathcal{M} = (\mathcal{S}, \mathcal{A}, \{T^{(x)}\}_{x \in \mathcal{A}}, \pi)$ of some process in statistical equilibrium. A mixed state η could be any distribution over \mathcal{S} , but the uncountable set of points in this simplex is infinitely more than we need in the calculation of many complexity measures. For example, the set \mathcal{R}_π of mixed states that we use for the \mathcal{S} -MSP are those induced by all allowed words $w \in \mathcal{L} \subseteq \mathcal{A}^*$ from initial mixed state $\eta_0 = \pi$:

$$\mathcal{R}_\pi = \bigcup_{w \in \mathcal{L}} \frac{\langle \pi | T^{(w)} \rangle}{\langle \pi | T^{(w)} | \mathbf{1} \rangle} .$$

The cardinality of \mathcal{R}_π is finite when there are only a finite number of distinct probability distributions over the states of \mathcal{M} that can be induced by observed sequences if we start from the stationary distribution π .

If w is the first observation sequence to induce a particular distribution over \mathcal{S} , then we use the notation that this distribution is denoted by η^w . For example, if the two observation sequences 010 and 110110 both induce the same distribution over \mathcal{S} , and no word shorter than 010 induces that distribution, then the mixed state will be denoted η^{010} , corresponding to the distribution $\langle \eta^{010} | = \langle \pi | T^{(0)} T^{(1)} T^{(0)} / \langle \pi | T^{(0)} T^{(1)} T^{(0)} | \mathbf{1} \rangle$.

Since a particular observation induces a unique updated distribution from a previous distribution, the dynamic over mixed states is unifilar. Transition probabilities among mixed states can be obtained via Eq. (3.2): If $\langle \eta | T^{(x)} | \mathbf{1} \rangle > 0$ and $\langle \eta' | = \frac{\langle \eta | T^{(x)} \rangle}{\langle \eta | T^{(x)} | \mathbf{1} \rangle}$, then $\Pr(\eta', x | \eta) = \Pr(x | \eta) = \langle \eta | T^{(x)} | \mathbf{1} \rangle$. These transition probabilities over the mixed states in \mathcal{R}_π are the matrix elements for the observation-labeled transition matrices $\{W^{(x)}\}_{x \in \mathcal{A}}$ of the \mathcal{S} -MSP of \mathcal{M} :

$$\mathcal{S}\text{-MSP}(\mathcal{M}) = (\mathcal{R}_\pi, \mathcal{A}, \{W^{(x)}\}_{x \in \mathcal{A}}, \delta_\pi) ,$$

where δ_π is the distribution over \mathcal{R}_π peaked at the unique start-(mixed)-state π . The row-stochastic net mixed-state-to-state transition matrix of $\mathcal{S}\text{-MSP}(\mathcal{M})$ is: $W = \sum_{x \in \mathcal{A}} W^{(x)}$. If irreducible, then there is a unique stationary probability distribution $\langle \pi_W |$ over $\mathcal{S}\text{-MSP}(\mathcal{M})$'s states obtained by solving $\langle \pi_W | = \langle \pi_W | W$. We will use \mathcal{R}_t to denote the random variable for the state of the MSP at time t .

More generally, we will need to consider a mixed state dynamic that can start from a non-peaked distribution over the simplex. This may be counterintuitive, since a distribution over distributions should correspond to a single distribution. However: the general MSP formalism, with a non-peaked starting distribution over the simplex, allows us to consider a weighted average of behaviors originating from disparate histories—which is distinct from considering the behavior originating from a weighted average of histories. This more general MSP formalism arise in the closed form solutions for the bound information, to appear elsewhere, which has been enabled by the following mathematical developments.

Finally, we emphasize that similar linear algebraic constructions—where hidden states track relevant information—that are nevertheless *not* MSPs are just as important for answering a different set of questions about a process. Since the other constructions are not directly about predictability and prediction, we report on these findings elsewhere.

3.6 Identifying the Hidden Linear Dynamic

We are now in a position to identify the appropriate hidden linear dynamic for many questions of complexity. In part, this section addresses a very practical need for specific calculations. It should, however, also serve a broader purpose: outlining a method to identify the hidden linear dynamic behind any question from an even broader class at the frontier of complex systems science.

Identifying the linear dynamic means identifying the linear operator A such that a question of interest can be reformulated as either the cascading form $\langle \cdot | A^n | \cdot \rangle$ or as an accumulation of such cascading events. For better or worse, many well-known questions of complexity can be mapped to these simple forms. And so, we now proceed to uncover the hidden linear dynamics of the cascading questions approximately in the order they were introduced in § 3.3.

3.6.1 Simple Complexity from any Presentation

For observable correlation, any HMM transition operator will do as the linear dynamic. We simply observe, let time (or space) crank forward, and observe again. Let us make this more rigorous.

Recall the familiar autocorrelation function. For a discrete-domain process it is [202]:

$$\gamma(L) = \langle \overline{X}_t X_{t+L} \rangle_t ,$$

where $L \in \mathbb{Z}$ and the bar denotes the complex conjugate. The autocorrelation function is symmetric about $L = 0$, so we can focus on $L \geq 0$. For $L = 0$, we simply have:

$$\begin{aligned} \langle \overline{X}_t X_t \rangle_t &= \sum_{x \in \mathcal{A}} |x|^2 \Pr(X_t = x) \\ &= \sum_{x \in \mathcal{A}} |x|^2 \langle \pi | T^{(x)} | \mathbf{1} \rangle . \end{aligned}$$

For $L > 0$, we have:

$$\begin{aligned} \gamma(L) &= \langle \overline{X}_t X_{t+L} \rangle_t \\ &= \sum_{x \in \mathcal{A}} \sum_{x' \in \mathcal{A}} \overline{x} x' \Pr(X_t = x, X_{t+L} = x') \\ &= \sum_{x \in \mathcal{A}} \sum_{x' \in \mathcal{A}} \overline{x} x' \Pr(x \underbrace{* \cdots *}_{L-1 \text{ *s}} x') \\ &= \sum_{x \in \mathcal{A}} \sum_{x' \in \mathcal{A}} \overline{x} x' \sum_{w \in \mathcal{A}^{L-1}} \Pr(xwx') . \end{aligned}$$

Each ‘*’ above is a wildcard symbol denoting indifference to the particular symbol observed in its place. That is, the *s denote marginalizing over the intervening random variables.

We develop the consequence of this, explicitly calculating ¹ and finding:

$$\begin{aligned}
\gamma(L) &= \sum_{x \in \mathcal{A}} \sum_{x' \in \mathcal{A}} \bar{x}x' \sum_{w \in \mathcal{A}^{L-1}} \langle \pi | T^{(x)} T^{(w)} T^{(x')} | \mathbf{1} \rangle \\
&= \sum_{x \in \mathcal{A}} \sum_{x' \in \mathcal{A}} \bar{x}x' \langle \pi | T^{(x)} \left(\sum_{w \in \mathcal{A}^{L-1}} T^{(w)} \right) T^{(x')} | \mathbf{1} \rangle \\
&= \sum_{x \in \mathcal{A}} \sum_{x' \in \mathcal{A}} \bar{x}x' \langle \pi | T^{(x)} \underbrace{\left(\prod_{i=1}^{L-1} \sum_{x_i \in \mathcal{A}} T^{(x_i)} \right)}_{=T} T^{(x')} | \mathbf{1} \rangle \\
&= \sum_{x \in \mathcal{A}} \sum_{x' \in \mathcal{A}} \bar{x}x' \langle \pi | T^{(x)} T^{L-1} T^{(x')} | \mathbf{1} \rangle \\
&= \langle \pi | \left(\sum_{x \in \mathcal{A}} \bar{x} T^{(x)} \right) T^{L-1} \left(\sum_{x' \in \mathcal{A}} x' T^{(x')} \right) | \mathbf{1} \rangle .
\end{aligned}$$

The result is the autocorrelation in cascading form $\langle \cdot | T^t | \cdot \rangle$, which can be made extra transparent by subsuming time-independent baggage on the left and right into the bras and kets. Introducing the new row vector

$$\langle \pi \bar{\mathcal{A}} | = \langle \pi | \left(\sum_{x \in \mathcal{A}} \bar{x} T^{(x)} \right)$$

and column vector

$$| \mathcal{A} \mathbf{1} \rangle = \left(\sum_{x \in \mathcal{A}} x T^{(x)} \right) | \mathbf{1} \rangle$$

the autocorrelation function for nonzero integer τ is simply:

$$\gamma(L) = \langle \pi \bar{\mathcal{A}} | T^{|L|-1} | \mathcal{A} \mathbf{1} \rangle . \quad (3.3)$$

Clearly, the autocorrelation function is a direct (albeit filtered) signature of iterates of the transition dynamic of any presentation of a process.

This result can easily be translated to the continuous time setting. If the process is represented as a deterministic function of a continuous-time Markov chain, and we make

¹ Averaging over t invokes unconditioned word probabilities that must be calculated using the stationary probability π over the recurrent states. Effectively this ignores any transient non-stationarity that may exist in a process, since only the recurrent part of the HMM presentation plays a role in the autocorrelation function. One practical lesson is that if T has transient states, they might as well be trimmed prior to such a calculation.

the translation that:

$$\langle \pi \overline{\mathcal{A}} | = \langle \pi | \left(\sum_{x \in \mathcal{A}} \overline{x} \Gamma_x \right) \quad \text{and} \quad | \mathcal{A} \mathbf{1} \rangle = \left(\sum_{x \in \mathcal{A}} x \Gamma_x \right) | \mathbf{1} \rangle ,$$

then the autocorrelation function for any $\tau \in \mathbb{R}$ is simply:

$$\gamma(\tau) = \langle \pi \overline{\mathcal{A}} | e^{L|G} | \mathcal{A} \mathbf{1} \rangle . \quad (3.4)$$

Again, the autocorrelation function is a direct fingerprint of the transition dynamic over the hidden states.

The power spectrum is a modulated accumulation of the autocorrelation function; with some algebra, one can show that it is:

$$P(\omega) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{L=-N}^N (N - |L|) \gamma(L) e^{-i\omega L} .$$

In [202], we show that the continuous part of the power spectrum, for discrete-domain processes, is simply:

$$P_c(\omega) = \langle |x|^2 \rangle + 2 \operatorname{Re} \langle \pi \overline{\mathcal{A}} | (e^{i\omega} I - T)^{-1} | \mathcal{A} \mathbf{1} \rangle , \quad (3.5)$$

where $\operatorname{Re}(\cdot)$ denotes the real part of its argument and I is the identity matrix. Or, for continuous-domain processes:

$$P_c(\omega) = 2 \operatorname{Re} \langle \pi \overline{\mathcal{A}} | (i\omega I - G)^{-1} | \mathcal{A} \mathbf{1} \rangle . \quad (3.6)$$

Although useful, signatures of pairwise correlation are only first-order complexity measures. Common measures of complexity that include higher orders of correlation can also be written in the simple cascading form, but will require a more careful choice of representation.

3.6.2 Predictability from MSP of any presentation

Any HMM presentation allows us to calculate a process's block entropy:

$$H(L) = H[X_0 X_1 \dots X_{L-1}] ,$$

using Eq. (3.1) but, due to the exponentially growing number of words in $\mathcal{L} \cap \mathcal{A}^L$, at a computational cost that is exponential in L : $\mathcal{O}(|\mathcal{S}|^3 L |\mathcal{A}|^L)$. Consequentially, using a general HMM one cannot directly nor efficiently calculate many key complexity measures, including a process's entropy rate and excess entropy. These limitations motivate using more specialized HMM classes.

For example, it is known that the entropy rate h_μ can be calculated directly from any unifilar presentation [4]. Also, we can calculate the excess entropy directly from a process's uHMM forward and reverse states [58, 76]: $\mathbf{E} = \mathbf{I}[\overleftarrow{X}; \overrightarrow{X}] = \mathbf{I}[\mathcal{S}^+; \mathcal{S}^-]$.

However, efficient computation of *myopic* entropy rates remained elusive until recently, and we have only just found their closed-form solutions. Myopic entropy rates $h_\mu(L)$ are important because they represent the apparent entropy rate of a process if it is modeled as a finite Markov order- $(L - 1)$ process—a type of approximation made commonly in the sciences. Crucially, the difference from the true entropy rate of the process, $h_\mu(L) - h_\mu$, is the *surplus entropy rate* incurred by using an order- $L - 1$ Markov approximation. Besides an apparent loss of predictability, the excess entropy rate associated with such an approximation can lead directly to errors in inferred physical properties, such as over-estimated dissipation that would be associated with the surplus entropy rate assigned to a physical thermodynamic system [31].

Unifilarity, it turns out, is not enough to calculate $h_\mu(L)$ directly. Conditioning on histories gave a clue as to the hidden linear dynamic. In particular, the \mathcal{S} -MSP of any presentation was the missing link. Much of this subsection follows Ref. [75].

The length- L myopic entropy rate is the expected uncertainty in the L^{th} random variable X_{L-1} , given the preceding $L - 1$ random variables $X_{0:L-1}$:

$$\begin{aligned}
h_\mu(L) &\equiv \mathbf{H}(L) - \mathbf{H}(L - 1) \\
&= \mathbf{H}[X_{0:L} | \eta_0 = \pi] - \mathbf{H}[X_{0:L-1} | \eta_0 = \pi] \\
&= \mathbf{H}[X_{L-1}, X_{0:L-1} | \eta_0 = \pi] - \mathbf{H}[X_{0:L-1} | \eta_0 = \pi] \\
&= \mathbf{H}[X_{L-1} | X_{0:L-1}, \eta_0 = \pi] \quad ,
\end{aligned} \tag{3.7}$$

where we explicitly give the initial condition of ignorance: Without taking any observations

we can only assume that the initial distribution η_0 over \mathcal{M} 's states is the expected asymptotic distribution π . For a mixing ergodic process, for example, even if another distribution $\eta_{-N} = \alpha$ was known in distant past, we still have $\langle \eta_0 | = \langle \eta_{-N} | T^N \rightarrow \langle \pi |$, as $N \rightarrow \infty$.

Assuming an initial probability distribution over \mathcal{M} 's states, a given sequence of observations induces a particular sequence of updated state distributions. That is, \mathcal{S} -MSP(\mathcal{M}) is unifilar regardless of whether \mathcal{M} is unifilar or not. Or, in other words, given the \mathcal{S} -MSP's unique start state— $\mathcal{R}_0 = \pi$ —and a particular realization $X_{0:L-1} = w^{L-1}$ of the last $L - 1$ random variables, we end up at the particular mixed state $\mathcal{R}_{L-1} = \eta_{w^{L-1}} \in \mathcal{R}_\pi$. Moreover, the entropy of the next observation is uniquely determined by \mathcal{M} 's state distribution, suggesting that Eq. (3.7) becomes:

$$\mathrm{H}[X_{L-1} | X_{0:L-1}, \eta_0 = \pi] = \mathrm{H}[X_{L-1} | \mathcal{R}_{L-1}, \mathcal{R}_0 = \pi] ,$$

as proven elsewhere [75]. Intuitively, conditioning on all of the past random variables of observation is equivalent to conditioning on the random variable for the distribution induced by particular sequences of observation.

We can now recast Eq. (3.7) in terms of the \mathcal{S} -MSP, finding:

$$\begin{aligned} h_\mu(L) &= \mathrm{H}[X_{L-1} | \mathcal{R}_{L-1}, \mathcal{R}_0 = \pi] \\ &= \sum_{\eta \in \mathcal{R}_\pi} \mathrm{Pr}(\mathcal{R}_{L-1} = \eta | \mathcal{R}_0 = \pi) \mathrm{H}[X_{L-1} | \mathcal{R}_{L-1} = \eta] \\ &= \sum_{\eta \in \mathcal{R}_\pi} \langle \delta_\pi | W^{L-1} | \delta_\eta \rangle \\ &\quad \times - \sum_{x \in \mathcal{A}} \langle \delta_\eta | W^{(x)} | \mathbf{1} \rangle \log_2 \langle \delta_\eta | W^{(x)} | \mathbf{1} \rangle \\ &= \langle \delta_\pi | W^{L-1} | H(W^{\mathcal{A}}) \rangle , \end{aligned}$$

where

$$|H(W^{\mathcal{A}})\rangle \equiv - \sum_{\eta \in \mathcal{R}_\pi} |\delta_\eta\rangle \sum_{x \in \mathcal{A}} \langle \delta_\eta | W^{(x)} | \mathbf{1} \rangle \log_2 \langle \delta_\eta | W^{(x)} | \mathbf{1} \rangle$$

is simply the column vector whose i^{th} entry is the entropy of transitioning from the i^{th} state of \mathcal{S} -MSP. Importantly, $|H(W^{\mathcal{A}})\rangle$ is independent of L .

Notice that taking the logarithm of the sum of the entries of the row vector $\langle \delta_\eta | W^{(x)}$ via $\langle \delta_\eta | W^{(x)} | \mathbf{1} \rangle$ is only permissible since \mathcal{S} -MSP's unifilarity guarantees that $W^{(x)}$ has at most one nonzero entry per row. (We also use the familiar convention that $0 \log_2 0 = 0$ [46].)

The result is the particularly compact and efficient expression for the length- L myopic entropy-rates:

$$h_\mu(L) = \langle \delta_\pi | W^{L-1} | H(W^{\mathcal{A}}) \rangle , \quad (3.8)$$

where all one must do is compute the powers of the MSP transition dynamic. The computational cost is now only linear in L : $\mathcal{O}(L|\mathcal{R}_\pi|^3)$. Moreover, W is very sparse, especially with a small alphabet \mathcal{A} , and so the computational cost can be reduced even further with optimization.

With $h_\mu(L)$ in hand, the hierarchy of complexity measures that derive from it immediately follow, including h_μ , \mathbf{E} , and \mathbf{T} [62]. For example:

$$\begin{aligned} h_\mu &= \lim_{L \rightarrow \infty} h_\mu(L) , \\ \mathbf{E} &= \sum_{L=1}^{\infty} [h_\mu(L) - h_\mu] , \quad \text{and} \\ \mathbf{T} &= \sum_{L=1}^{\infty} L [h_\mu(L) - h_\mu] . \end{aligned}$$

We will discuss these quantities more and give their closed-form expressions after introducing the meromorphic functional calculus.

3.6.2.1 Continuous time?

We have seen that correlation measures are easily extended to the continuous-time domain via continuous-time HMMs. Information measures of complexity, though, are awkward in continuous time, although progress has been made recently towards understanding their structure [168, 165].

This roadblock may seem puzzling since continuous-time nonequilibrium (or equilibrium) thermodynamics can be non-problematically reformulated in information-theoretic terms. The difference is that, in thermodynamics, information quantities are over a set

of simultaneous observables rather than a temporal sequence of observables; taking the limit of continuous time doesn't change what is instantaneously observable although it wreaks havoc on information-theoretic measures of sequential expression. In nonequilibrium thermodynamics, there is still a temporal aspect, but this goes into updating the distribution over hidden states, which leads to an evolving entropy over instantaneous observables [199]. Explicitly, the thermodynamic entropy is:

$$S = k_B H((X_{0:N})_t | (Y_{0:M})_{-\infty:t}) , \quad (3.9)$$

where k_B is Boltzmann's constant, $(X_{0:N})_t$ is the random variable for the N degrees of freedom for the system under study at time t , $(Y_{0:M})_{-\infty:t}$ is the entire history of the length- M vector of environmental boundary conditions up to the present time t , and $H(\cdot)$ in this case is the Shannon entropy in units of nats [199]. Hence, the linear dynamic associated with nonequilibrium thermodynamic entropy is the dynamic over distributions (over states that produce the observable configurations of the system) induced by the environmental boundary conditions. In steady state, this reduces to the time-independent entropy of the stationary distribution over all configurations of the system, $k_B H(X_{0:N})$.

A large class of nonequilibrium thermodynamic systems can be reformulated in terms of HMMs [15, 30, 199, 201]. Many have already considered using Markovian dynamics or explicit Markov models in thermodynamics, but using *hidden* Markov models allow new analysis for transitions atop nonequilibrium steady states—for example, the first exact calculations for excess thermodynamic quantities [201].

3.6.3 Synchronization from MSP of generating model

If the physical state-space is known, then the \mathcal{S} -MSP of the generating model allows one to keep track of the observation-induced distributions over states of the generative model, naturally leading to closed-form solutions to questions about synchronizing to states of the system.

Expressing the *myopic state uncertainty*, $\mathcal{H}(L) = H[\mathcal{S}_0 | X_{-L:0}]$, in terms of the \mathcal{S} -MSP,

one finds [75]:

$$\begin{aligned}\mathcal{H}(L) &= - \sum_{w \in \mathcal{A}^L} \Pr(w) \sum_{\sigma \in \mathcal{S}} \Pr(\sigma|w) \log_2 \Pr(\sigma|w) \\ &= \sum_{\eta \in \mathcal{R}_\pi} \Pr(\mathcal{R}_L = \eta | \mathcal{R}_0 = \pi) H[\eta] .\end{aligned}$$

Here, $H[\eta]$ is the presentation-state uncertainty specified by the mixed state η :

$$H[\eta] \equiv - \sum_{\sigma \in \mathcal{S}} \langle \eta | \delta_\sigma \rangle \log_2 \langle \eta | \delta_\sigma \rangle , \quad (3.10)$$

where $|\delta_\sigma\rangle$ is the length- $|\mathcal{S}|$ column vector of all zeros except for a one at the appropriate index of the presentation-state σ .

Continuing, we re-express $\mathcal{H}(L)$ in terms of powers of the \mathcal{S} -MSP transition dynamic:

$$\begin{aligned}\mathcal{H}(L) &= \sum_{\eta \in \mathcal{R}_\pi} \langle \delta_\pi | W^L | \delta_\eta \rangle H[\eta] \\ &= \langle \delta_\pi | W^L | H[\eta] \rangle .\end{aligned} \quad (3.11)$$

Here, we defined:

$$|H[\eta]\rangle \equiv \sum_{\eta \in \mathcal{R}_\pi} |\delta_\eta\rangle H[\eta] ,$$

which is the L -independent length- $|\mathcal{R}_\pi|$ column vector whose entries are the appropriately indexed entropies of each mixed state.

The forms of Eqs. (3.8) and (3.11) highlight that $h_\mu(L+1)$ and $\mathcal{H}(L)$ differ only in the type of information being evolved: observable entropy or state entropy, as implicated by their respective kets. Each of these entropies is reduced by the distributions induced by longer observation.

Paralleling $h_\mu(L)$, there is a complementary hierarchy of complexity measures that are built from functionals of $\mathcal{H}(L)$. These include the asymptotic state uncertainty \mathcal{H} and excess synchronization information \mathbf{S}' , to mention only two:

$$\begin{aligned}\mathcal{H} &= \lim_{L \rightarrow \infty} \mathcal{H}(L) && \text{and} \\ \mathbf{S}' &= \sum_{L=0}^{\infty} [\mathcal{H}(L) - \mathcal{H}] .\end{aligned}$$

Compared to the $h_\mu(L)$ spin-offs, \mathcal{H} and \mathbf{S}' mirror the roles of h_μ and \mathbf{E} , respectively.

The state-complexity of the model, $C(\mathcal{M}) = \mathcal{H}(0) = \langle \delta_\pi | H[\eta] \rangle$, also has an analogue in the $h_\mu(L)$ hierarchy: the alphabet complexity of the process: $H[X_0] = h_\mu(1) = \langle \delta_\pi | H(W^{\mathcal{A}}) \rangle$.

3.6.4 Optimal prediction from MSP of ϵ -machine

We just reviewed the linear underpinnings of synchronizing to *any* model of a process. However, the myopic state uncertainty of the ϵ -machine has a distinguished role in telling the synchronization cost for optimal prediction of the output, regardless of the model that generated it. Using the \mathcal{S} -MSP of the ϵ -machine, the myopic state uncertainty of the ϵ -machine can be written in direct parallel to the myopic state uncertainty of any model:

$$\begin{aligned} \mathcal{H}^+(L) &= - \sum_{w \in \mathcal{A}^L} \Pr(w) \sum_{\sigma \in \mathcal{S}^+} \Pr(\sigma|w) \log_2 \Pr(\sigma|w) \\ &= \sum_{\eta \in \mathcal{R}_\pi^+} \Pr(\mathcal{R}_L = \eta | \mathcal{R}_0 = \pi) H[\eta] \\ &= \langle \delta_\pi | \mathcal{W}^L | H[\eta] \rangle . \end{aligned}$$

where we use the script \mathcal{W} to emphasize that this is now the the state-to-state transition dynamic of the MSP of the ϵ -machine specifically.

Paralleling $\mathcal{H}(L)$, an obvious hierarchy of complexity measures is built from functionals of $\mathcal{H}^+(L)$. For example, the state-complexity of the ϵ -machine is the *statistical complexity* $C_\mu = \mathcal{H}^+(0)$. The information that must be obtained to synchronize to the causal state and thus optimally predict—the causal synchronization information—is given in terms of the \mathcal{S} -MSP of the ϵ -machine by $\mathbf{S} = \sum_{L=0}^{\infty} \mathcal{H}^+(L)$.

An important difference for ϵ -machines is that they have zero asymptotic state uncertainty:

$$\mathcal{H}^+ = 0 .$$

Therefore, $\mathbf{S} = \mathbf{S}'(\epsilon\text{-machine})$. Moreover, we conjecture that $\mathbf{S} = \min_{\mathcal{M}} \sum_{L=0}^{\infty} \mathcal{H}(L)$ for any presentation \mathcal{M} that generates the process, even if $C_\mu \geq C_g$.

3.6.5 Beyond the MSP

Many of the complexity measures introduced have used a mixed-state presentation (with particular focus on the \mathcal{S} -MSP) as the appropriate linear dynamic. However, we want to emphasize that this is a testament to the questions that have become popular, and is not indicative of the general answer that one should expect in the broader approach to finding the hidden linear dynamic. Here, we hope to give a brief eye-opener for what other linear dynamics can look like for different types of questions of complexity, which we have uncovered recently, primarily in as-yet unpublished work.

For example, we have found a reverse-time mixed-functional presentation (MFP \neq MSP) of any forward-time generator. The MFP tracks the reverse-time dynamic over linear functionals $|\eta\rangle \in \mathcal{R}_1 = \left\{ \frac{T^{(w)}|\mathbf{1}\rangle}{\langle \pi | T^{(w)} | \mathbf{1} \rangle} \right\}_w$ (linear functionals of state distributions) induced by reverse-time observations. The MFP allows direct calculation of the convergence of the *preparation uncertainty* $\mathfrak{H}(L) \equiv H(\mathcal{S}_0 | X_{0:L})$ via powers of the linear MFP transition dynamic. The preparation uncertainty gives a new perspective on the transient information since $\sum_{L=0}^{\infty} (H(\mathcal{S}_0^+ | X_{0:L}) - \chi) = \mathbf{T}$ can be interpreted as the predictive advantage of hindsight. The myopic process crypticity $\chi(L)$ was previously introduced [160], where $\chi(L) = \mathfrak{H}^+(L) - \mathcal{H}^+(L)$. Since $\lim_{L \rightarrow \infty} \mathcal{H}^+(L) = \mathcal{H}^+ = 0$, the asymptotic crypticity is $\chi = \mathfrak{H}^+ + \mathcal{H}^+ = \mathfrak{H}^+$. We now glean a refined partitioning behind the sum $\sum_{L=0}^{\infty} (\chi - \chi(L)) = \mathbf{S} - \mathbf{T} \geq 0$.

Crypticity $\chi = H(\mathcal{S}_0^+ | X_{0:\infty})$ itself is only non-zero if the process' *cryptic order*,

$$k = \min \{ \ell \in \{0, 1, \dots\} : H(\mathcal{S}_0^+ | X_{-\ell:\infty}) = 0 \} ,$$

is non-zero. The cryptic order is always less than or equal to its better known cousin, the *Markov order*,

$$R = \min \{ \ell \in \{0, 1, \dots\} : H(\mathcal{S}_0^+ | X_{-\ell:0}) = 0 \} ,$$

since conditioning (on future observations $X_{0:\infty}$, in the case of cryptic order) can never increase entropy. Since the reverse causal state at time 0, \mathcal{S}_0^- , is a linear combination of forward causal states [158, 159], the forward-time *cryptic operator presentation*, which gives the forward-time observation-induced dynamic over the operators $\mathcal{O} \in \left\{ \frac{|s^-\rangle \langle \eta^w|}{\langle \eta^w | s^- \rangle} : \right.$

$s^- \in \mathcal{S}^-, \langle \eta^w | \in \mathcal{R}_\pi, \langle \eta^w | s^- \rangle > 0 \}$ allow new calculations regarding the convergence to crypticity that implicate $\Pr(\mathcal{S}_0^+ | X_{-\ell:\infty})$. However, the cryptic operator presentation is just a special case of the more general myopic bi-directional dynamic over operators $\mathcal{O} \in \left\{ \frac{|\eta^{w'}\rangle\langle\eta^w|}{\langle\eta^w|\eta^{w'}\rangle} : \langle\eta^w| \in \mathcal{R}_\pi, |\eta^{w'}\rangle \in \mathcal{R}_1, \langle\eta^w|\eta^{w'}\rangle > 0 \right\}$ induced by new observations of either the future or the past, relevant for understanding the interplay between forgetfulness and shortsightedness, $\Pr(\mathcal{S}_0 | X_{-M:0}, X_{0:N})$.

Another construction, the quantum pairwise merger machine (QPMM) gives quantum probability amplitudes on the edges of a topology derived from the ϵ -machine. The QPMM exactly quantifies how a classical process can be generated with less memory using quantum mechanics [203]. Although the QPMM is not quite a HMM, its linear transition dynamic is manipulated in much the same way as discussed here.

The list goes on. For example, detailed bounds on entropy rate convergence can be obtained from the transition dynamic of the so-called possibility machine, beyond the asymptotic result obtained in Ref. [249]. As another example, the importance of post-synchronized monitoring, as quantified by the information lost upon negligence over a duration ℓ , $b_\mu(\ell) = I(X_{0:\ell}; X_{\ell:\infty} | X_{-\infty:0})$, has been pinned down using a modified type of MSP.

These examples all find an exact solution via a parallel theory to that outlined in this chapter, applied to the linear dynamic appropriate for the corresponding question of complexity. Furthermore, these examples demonstrate the new opportunity, enabled by the associated mathematical developments, to ask and answer more nuanced and thus more directed questions about structure, predictability, and prediction.

3.6.6 The end?

In some sense, our goal is done: we have identified the appropriate transition dynamic for common questions of complexity and, by some standards, have given formulae for their exact solution. But the work so far has all been in preparation. Although we have set ourselves up nicely for linear analysis, closed-form expressions for all of the complexity measures still await the developments of the following sections. Concomitantly, at the level of qualitative understanding and scientific interpretation we have so far failed to give

answers to the simple questions:

- What is the range of possible behaviors that these complexity measures can exhibit?

and the natural follow-up question:

- What are the mechanisms that produce qualitatively different behaviors?

In the following section, we review the recently developed functional calculus that allows us to actually decompose arbitrary functions of the nondiagonalizable hidden dynamic to give conclusive answers to these fundamental questions. We expose the range of possible behaviors and identify the internal mechanisms that give rise to qualitatively different contributions to complexity.

Our investment in the next sections will allow us to express new closed-form solutions for many complexity measures beyond what we have been able to achieve so far. Besides the obvious calculational advantage, this also yields new insights into the possible behavior of the complexity measures and, moreover, their unexpected similarities with each other. In many ways, the results will shed new light upon what we were (implicitly) probing with complexity measures that we thought we knew. Constructively, the new understanding will suggest extensions from complexity magnitudes to complexity functions that succinctly capture the organization of all orders of correlation. As intuition for pairwise correlation can grow out of a power spectrum, so too can this extension unveil the viscera of both the predictability of a process and the burden of prediction.

3.7 Spectral Theory beyond the Spectral Theorem

Here we briefly review the requisite theory from Ref. [198] necessary for the subsequent development.

3.7.1 Spectral Primer

We will restrict our attention to operators that have at most a countably infinite spectrum. Such operators share many features with finite-dimensional square matrices, and so we will review some elementary but essential facts that will be used extensively in the following.

Recall that if A is a finite-dimensional square matrix, then A 's spectrum is simply its set of eigenvalues:

$$\Lambda_A = \{ \lambda \in \mathbb{C} : \det(\lambda I - A) = 0 \} ,$$

where $\det(\cdot)$ is the determinant of its argument.

For reference later, recall that the *algebraic multiplicity* a_λ of eigenvalue λ is the power of the term $(z - \lambda)$ in the characteristic polynomial $\det(zI - A)$. In contrast, the *geometric multiplicity* g_λ is the dimension of the kernel of the transformation $A - \lambda I$ or the number of linearly independent eigenvectors for the eigenvalue. The algebraic and geometric multiplicities are all equal when the matrix is diagonalizable.

Since there can be multiple subspaces associated with a single eigenvalue, corresponding to different Jordan blocks in the Jordan canonical form, it is structurally important to introduce the index of the eigenvalue to describe the depth of its deepest associated subspace.

Definition 4. The index ν_λ of eigenvalue λ is the size of the largest Jordan block associated with λ .

The index gives information beyond what the algebraic and geometric multiplicities themselves yield. Nevertheless, for $\lambda \in \Lambda_A$, it is always true that $\nu_\lambda - 1 \leq a_\lambda - g_\lambda \leq a_\lambda - 1$. In the diagonalizable case, $a_\lambda = g_\lambda$ and $\nu_\lambda = 1$ for all $\lambda \in \Lambda_A$.

The *resolvent*,

$$R(z; A) \equiv (zI - A)^{-1} ,$$

defined with the help of the continuous complex variable $z \in \mathbb{C}$, captures all of the spectral information about A through the poles of the resolvent's matrix elements. Indeed, the resolvent contains more than just the spectrum: the order of each pole gives the index of the corresponding eigenvalue.

Each eigenvalue λ of A has the associated *projection operator* A_λ , which is the residue of the resolvent as $z \rightarrow \lambda$:

$$A_\lambda \equiv \frac{1}{2\pi i} \oint_{C_\lambda} R(z; A) dz . \quad (3.12)$$

The residue of the matrix can be calculated element-wise.

The projection operators are orthonormal:

$$A_\lambda A_\zeta = \delta_{\lambda,\zeta} A_\lambda , \quad (3.13)$$

and sum to the identity:

$$I = \sum_{\lambda \in \Lambda_A} A_\lambda . \quad (3.14)$$

For cases where $\nu_\lambda = 1$, we have found that the projection operator associated with λ can be calculated as:

$$A_\lambda = \prod_{\substack{\zeta \in \Lambda_A \\ \zeta \neq \lambda}} \left(\frac{A - \zeta I}{\lambda - \zeta} \right)^{\nu_\zeta} \quad \text{if } \nu_\lambda = 1 . \quad (3.15)$$

Not all projection operators of a nondiagonalizable operator can be found directly from Eq. (3.15) (since some of them will have index larger than one). However, if there is only one eigenvalue that has index larger than one—the *almost diagonalizable* case treated in § 3.12.5—then Eq. (3.15), together with the fact that the projection operators must sum to the identity, *does* give a full solution to the set of projection operators. Next, we consider the general case, with no restriction on ν_λ .

3.7.2 Eigenprojectors and right and left eigenvectors and generalized eigenvectors

In general, the eigenprojectors can be obtained from all left and right eigenvectors and generalized eigenvectors associated with the eigenvalue. Given the n -tuple of possibly-degenerate eigenvalues $(\Lambda_A) = (\lambda_1, \lambda_2, \dots, \lambda_n)$, there is a corresponding n -tuple of m_k -tuples of linearly-independent *generalized right-eigenvectors*:

$$\left((|\lambda_1^{(m)}\rangle)_{m=1}^{m_1}, (|\lambda_2^{(m)}\rangle)_{m=1}^{m_2}, \dots, (|\lambda_n^{(m)}\rangle)_{m=1}^{m_n} \right) ,$$

where:

$$(|\lambda_k^{(m)}\rangle)_{m=1}^{m_k} \equiv \left(|\lambda_k^{(1)}\rangle, |\lambda_k^{(2)}\rangle, \dots, |\lambda_k^{(m_k)}\rangle \right)$$

and a corresponding n -tuple of m_k -tuples of linearly-independent *generalized left-eigenvectors*:

$$\left((\langle \lambda_1^{(m)} |)_{m=1}^{m_1}, (\langle \lambda_2^{(m)} |)_{m=1}^{m_2}, \dots, (\langle \lambda_n^{(m)} |)_{m=1}^{m_n} \right) ,$$

where:

$$(\langle \lambda_k^{(m)} |)_{m=1}^{m_k} \equiv \left(\langle \lambda_k^{(1)} |, \langle \lambda_k^{(2)} |, \dots, \langle \lambda_k^{(m_k)} | \right)$$

such that:

$$(A - \lambda_k I) |\lambda_k^{(m+1)}\rangle = |\lambda_k^{(m)}\rangle \quad (3.16)$$

and:

$$\langle \lambda_k^{(m+1)} | (A - \lambda_k I) = \langle \lambda_k^{(m)} | , \quad (3.17)$$

for $0 \leq m \leq m_k - 1$, where $|\lambda_j^{(0)}\rangle = \vec{0}$ and $\langle \lambda_j^{(0)}| = \vec{0}$. Specifically, $|\lambda_k^{(1)}\rangle$ and $\langle \lambda_k^{(1)}|$ are conventional right and left eigenvectors, respectively.

Recall that eigenvalue $\lambda \in \Lambda_A$ corresponds to g_λ different Jordan blocks, where g_λ is the geometric multiplicity of the eigenvalue λ . Indeed:

$$n = \sum_{\lambda \in \Lambda_A} g_\lambda .$$

Moreover, the index ν_λ of the eigenvalue λ is the size of the largest Jordan block corresponding to λ :

$$\nu_\lambda = \max \{ \delta_{\lambda, \lambda_k} m_k \}_{k=1}^n .$$

Most directly, the generalized right and left eigenvectors can be found as the nontrivial solutions to:

$$(A - \lambda_k I)^m |\lambda_k^{(m)}\rangle = \vec{0}$$

and:

$$\langle \lambda_k^{(m)} | (A - \lambda_k I)^m = \vec{0} ,$$

respectively. Imposing appropriate normalization, we find that:

$$\langle \lambda_j^{(m)} | \lambda_k^{(n)} \rangle = \delta_{j,k} \delta_{m+n, m_k+1} . \quad (3.18)$$

Crucially, right and left eigenvectors are no longer simply related by complex-conjugate transposition, and right eigenvectors are not necessarily orthogonal to each other. Rather, left eigenvectors and generalized eigenvectors form a dual basis to the right eigenvectors and generalized eigenvectors. Somewhat surprisingly, the *most* generalized left eigenvector $\langle \lambda_k^{(m_k)} |$ associated with λ_k is dual to the *least* generalized right eigenvector $|\lambda_k^{(1)}\rangle$ associated with λ_k . That is: $\langle \lambda_k^{(m_k)} | \lambda_k^{(1)} \rangle = 1$.

Explicitly, we find that the projection operators for a nondiagonalizable matrix can be written as:

$$A_\lambda = \sum_{k=1}^n \sum_{m=1}^{m_k} \delta_{\lambda, \lambda_k} |\lambda_k^{(m)}\rangle \langle \lambda_k^{(m_k+1-m)}| . \quad (3.19)$$

3.7.3 Companion operators and decomposition of the resolvent

It is useful to introduce the generalized set of *companion operators*:

$$A_{\lambda, m} = A_\lambda (A - \lambda I)^m , \quad (3.20)$$

for $\lambda \in \Lambda_A$ and $m \in \{0, 1, 2, \dots\}$. These operators satisfy the following semigroup relation:

$$A_{\lambda, m} A_{\zeta, n} = \delta_{\lambda, \zeta} A_{\lambda, m+n} . \quad (3.21)$$

$A_{\lambda, m}$ reduces to the eigenprojector for $m = 0$:

$$A_{\lambda, 0} = A_\lambda , \quad (3.22)$$

and exactly reduces to the zero-matrix for sufficiently large m :

$$A_{\lambda, m} = \mathbf{0} \quad \text{for } m \geq \nu_\lambda . \quad (3.23)$$

Crucially, we can rewrite the resolvent as a weighted sum of the companion matrices $\{A_{\lambda, m}\}$, with complex coefficients that have poles at each eigenvalue λ up to the index ν_λ of the eigenvalue. That is:

$$R(z; A) = \sum_{\lambda \in \Lambda_A} \sum_{m=0}^{\nu_\lambda-1} \frac{1}{(z - \lambda)^{m+1}} A_{\lambda, m} . \quad (3.24)$$

Ultimately this all allows us to evaluate arbitrary functions of nondiagonalizable operators, which we show next. More details can be found in Ref. [198].

3.7.4 Arbitrary functions of nondiagonalizable operators

The starting point for the *meromorphic functional calculus* [198], which gives meaning to arbitrary functions $f(\cdot)$ of any linear operator A , is the Cauchy-integral-like formula:

$$f(A) = \sum_{\lambda \in \Lambda_A} \frac{1}{2\pi i} \oint_{C_\lambda} f(z) R(z; A) dz, \quad (3.25)$$

where C_λ denotes a sufficiently small counterclockwise contour around λ in the complex plane such that no singularity of the integrand besides the possible pole at $z = \lambda$ is enclosed by the contour.

Invoking Eq. (3.24) yields the desired formulation:

$$f(A) = \sum_{\lambda \in \Lambda_A} \sum_{m=0}^{\nu_\lambda-1} A_{\lambda,m} \frac{1}{2\pi i} \oint_{C_\lambda} \frac{f(z)}{(z-\lambda)^{m+1}} dz. \quad (3.26)$$

Hence, with the eigenprojectors $\{A_\lambda\}_{\lambda \in \Lambda_A}$ in hand, evaluating an arbitrary function of the nondiagonalizable operator A comes down to the evaluation of several residues.

Typically, the evaluation of Eq. (3.26) requires less work than one might expect when looking at the equation in its full generality. For example, whenever $f(z)$ is holomorphic (i.e., well behaved) at $z = \lambda$, the residue simplifies to $\frac{1}{2\pi i} \oint_{C_\lambda} \frac{f(z)}{(z-\lambda)^{m+1}} dz = \frac{1}{m!} f^{(m)}(\lambda)$, where $f^{(m)}(\lambda)$ is the m^{th} derivative of $f(z)$ evaluated at $z = \lambda$. However, if $f(z)$ has a pole or zero at $z = \lambda$, then it crucially changes the complex contour integration. In the simplest case, when A is diagonalizable and $f(z)$ is holomorphic at Λ_A , the matrix-valued function reduces to the simple form: $f(A) = \sum_{\lambda \in \Lambda_A} f(\lambda) A_\lambda$. Moreover, if λ is non-degenerate, then $A_\lambda = \frac{|\lambda\rangle\langle\lambda|}{\langle\lambda|\lambda\rangle}$, although $\langle\lambda|$ here should be interpreted as the solution to the left eigen-equation $\langle\lambda| A = \lambda \langle\lambda|$ and in general $\langle\lambda| \neq (|\lambda\rangle)^\dagger$.

The meromorphic functional calculus agrees with the Taylor-series approach whenever the series converges and agrees with the holomorphic functional calculus of Ref. [69] whenever $f(z)$ is holomorphic at Λ_A . However, when both the previous functional calculi fail, the meromorphic functional calculus extends the domain of $f(A)$, in a way that is crucial in enabling the following analysis. For example, we will show that within the

meromorphic functional calculus, the negative-one power of a singular operator is the Drazin inverse which effectively inverts everything that is invertible and appears ubiquitous in the newfound solution to many complexity measures.

3.7.5 Evaluating the residues

How does one use this result? Eq. (3.26) says that the spectral decomposition of $f(A)$ reduces to the evaluation of several residues, where:

$$\text{Res}(g(z), z \rightarrow \lambda) = \frac{1}{2\pi i} \oint_{C_\lambda} g(z) dz .$$

So, to make progress with Eq. (3.26), we must evaluate functional-dependent residues of the form $\text{Res}(f(z)/(z - \lambda)^{m+1}, z \rightarrow \lambda)$.

Recall that the residue of a complex-valued function $g(z)$ around its isolated pole λ of order $n + 1$ can be calculated from:

$$\text{Res}(g(z), z \rightarrow \lambda) = \frac{1}{n!} \lim_{z \rightarrow \lambda} \frac{d^n}{dz^n} [(z - \lambda)^{n+1} g(z)] .$$

3.7.6 Decomposing A^L

Using Eq. (3.26), we can explicitly derive the spectral decomposition of powers of the operator A . For $f(A) = A^L \rightarrow f(z) = z^L$, $z = 0$ can be either a zero or a pole of $f(z)$, depending on the value of L . In either case, an eigenvalue of $\lambda = 0$ will distinguish itself in the residue calculation of A^L via its unique ability to change the order of the pole (or zero) at $z = 0$. For example, at this special value of λ and for integer $L > 0$, $\lambda = 0$ induces poles that *cancel* with the zeros of $f(z) = z^L$, since z^L has a zero at $z = 0$ of order L . For integer $L < 0$, an eigenvalue of $\lambda = 0$ *increases* the order of the $z = 0$ pole of $f(z) = z^L$. For all other eigenvalues, the residues will be as expected. Hence, for any $L \in \mathbb{C}$:

$$\begin{aligned} A^L = & \left[\sum_{\substack{\lambda \in \Lambda_A \\ \lambda \neq 0}} \sum_{m=0}^{\nu_\lambda-1} \binom{L}{m} \lambda^{L-m} A_{\lambda,m} \right] \\ & + [0 \in \Lambda_A] \sum_{m=0}^{\nu_0-1} \delta_{L,m} A_0 A^m , \end{aligned} \quad (3.27)$$

where $\binom{L}{m}$ is the generalized binomial coefficient:

$$\binom{L}{m} = \frac{1}{m!} \prod_{n=1}^m (L - n + 1) \quad \text{with} \quad \binom{L}{0} = 1 , \quad (3.28)$$

and $[0 \in \Lambda_A]$ is the Iverson bracket which takes on value 1 if zero is an eigenvalue of A and 0 if not. Eq. (3.27) applies to any linear operator with only isolated singularities in its resolvent.

If L is a non-negative integer such that $L \geq \nu_\lambda - 1$ for all $\lambda \in \Lambda_A$, then:

$$A^L = \sum_{\substack{\lambda \in \Lambda_A \\ \lambda \neq 0}} \sum_{m=0}^{\nu_\lambda-1} \binom{L}{m} \lambda^{L-m} A_{\lambda,m} , \quad (3.29)$$

where $\binom{L}{m}$ is now reduced to the traditional binomial coefficient $\frac{L!}{m!(L-m)!}$.

3.7.7 Drazin Inverse

The negative-one power of any linear operator—which is in general *not* the same as the inverse $\text{inv}(\cdot)$ since $\text{inv}(A)$ need not exist—is always defined via Eq. (3.27) to be:

$$A^{-1} = \sum_{\lambda \in \Lambda_A \setminus \{0\}} \sum_{m=0}^{\nu_\lambda-1} (-1)^m \lambda^{-1-m} A_{\lambda,m} . \quad (3.30)$$

Notably, when the operator is singular, we find that: $AA^{-1} = I - A_0$.

This is the *Drazin inverse* $A^{\mathcal{D}}$ of A , also known as the $\{1^{\nu_0}, 2, 5\}$ -inverse [19]. (Note that it is *not* the same as the Moore–Penrose pseudo-inverse.) Although the Drazin inverse is usually defined axiomatically to satisfy certain criteria, here it naturally falls out as the negative one power of a singular operator in the meromorphic functional calculus.

Whenever A is invertible, A^{-1} is equal to $\text{inv}(A)$. However, we should not confuse this coincidence with equivalence, and there is no reason despite historic notation that the negative-one power should in general be equivalent to the inverse—especially if an operator is not invertible! To avoid confusing A^{-1} with $\text{inv}(A)$, we will use the notation $A^{\mathcal{D}}$ for the Drazin inverse of A . Still, $A^{\mathcal{D}} = \text{inv}(A)$ whenever $0 \notin \Lambda_A$.

Although Eq. (3.30) is a constructive way to build the Drazin inverse, it suggests more work than is actually necessary. Indeed, we have derived several simple constructions of the Drazin inverse that requires only the original operator and the eigenvalue-0 projector. For example, in Ref. [198] we found that, for any $c \in \mathbb{C} \setminus \{0\}$:

$$A^{\mathcal{D}} = (I - A_0)(A + cA_0)^{-1} . \quad (3.31)$$

Later, we will also need the decomposition of $(I - W)^{\mathcal{D}}$, as it enters into many of our closed-form complexity expressions. In Ref. [198], we showed that:

$$(I - T)^{\mathcal{D}} = [I - (T - T_1)]^{-1} - T_1 \quad (3.32)$$

for any stochastic matrix T . If T is the state-transition matrix of an ergodic process, then the RHS of Eq. (3.32) becomes especially simple to evaluate since then $T_1 = |\mathbf{1}\rangle\langle\pi|$. Somewhat tangentially, the result allows a connection with the ‘fundamental matrix’ $Z = (I - T + T_1)^{-1}$ used by Kemeny and Snell [129] in their analysis of Markovian dynamics. More immediately, Eq. (3.32) will play a prominent role in upcoming derivations for excess entropy and synchronization information. It will also be useful to give the explicit spectral decomposition:

$$(I - T)^{\mathcal{D}} = \sum_{\lambda \in \Lambda_T \setminus \{1\}} \sum_{m=0}^{\nu_\lambda-1} \frac{1}{(1 - \lambda)^{m+1}} T_{\lambda,m} . \quad (3.33)$$

3.8 Projection Operators for Stochastic Transition Dynamics

The preceding employed the notation that A represents a general linear operator. In the following examples, we reserve the symbol T for the operator of a stochastic transition dynamic, as in the state-to-state transition dynamic of a HMM: $T = \sum_{x \in \mathcal{A}} T^{(x)}$. If the state-space is finite and has a stationary distribution, then T has a representation that is a nonnegative row-stochastic—i.e., all rows sum to unity—transition matrix.

We are now in a position to summarize several useful properties for the projection operators of any row-stochastic matrix T . Naturally, if one uses column-stochastic instead of row-stochastic matrices, all results can be translated by simply taking the transpose of every line in the derivations—recall that $(ABC)^{\top} = C^{\top} B^{\top} A^{\top}$.

The transition matrix’s nonnegativity guarantees that for each $\lambda \in \Lambda_T$ its complex conjugate $\bar{\lambda}$ is also in Λ_T . Moreover, the projection operator associated with the complex conjugate of λ is the complex conjugate of T_{λ} : $T_{\bar{\lambda}} = \overline{T_{\lambda}}$.

If the dynamic induced by T has a stationary distribution over the state space, then the spectral radius of T is unity and all of T ’s eigenvalues lie on or within the unit circle in

the complex plane. The maximal eigenvalues have unity magnitude and $1 \in \Lambda_T$. Moreover, an extension of the Perron–Frobenius theorem guarantees that eigenvalues on the unit circle have algebraic multiplicity equal to their geometric multiplicity. And, so, $\nu_\zeta = 1$ for all $\zeta \in \{\lambda \in \Lambda_T : |\lambda| = 1\}$.

T 's index-one eigenvalue of $\lambda = 1$ is associated with stationarity on the hidden Markov chain. T 's other eigenvalues on the unit circle are roots of unity and correspond to deterministic periodicities within the process.

3.8.1 Row Sums: $T_\lambda |\mathbf{1}\rangle = \delta_{\lambda,1} |\mathbf{1}\rangle$

If T is row-stochastic, then by definition:

$$T |\mathbf{1}\rangle = |\mathbf{1}\rangle .$$

Hence, via the general construction of the eigenprojectors Eq. (3.19) and the general orthogonality condition Eq. (3.18), we find that:

$$T_\lambda |\mathbf{1}\rangle = \delta_{\lambda,1} |\mathbf{1}\rangle . \quad (3.34)$$

This shows that T 's projection operator T_1 is row-stochastic, whereas each row of every other projection operator must sum to zero. This can also be viewed as a consequence of conservation of probability for dynamics over Markov chains.

3.8.2 Expected Stationary Distribution from T_1

If a process has no deterministic periodicities (i.e., if unity is the only eigenvalue of Λ_T on the unit circle), then every initial condition will lead to an asymptotic stationary distribution. The expected stationary distribution π_α from any initial distribution α will be:

$$\begin{aligned} \langle \pi_\alpha | &= \lim_{L \rightarrow \infty} \langle \alpha | T^L \\ &= \langle \alpha | T_1 . \end{aligned} \quad (3.35)$$

An attractive feature of Eq. (3.35) is that it holds even for nonergodic processes—those with multiple attractors.

When the stochastic process is ergodic (has only one attractor), then $a_1 = 1$ and there is only one stationary distribution π and the T_1 projection operator is:

$$T_1 = |\mathbf{1}\rangle \langle \pi| , \quad (3.36)$$

even if there are deterministic periodicities. Deterministic periodicities imply that different initial conditions may still induce different asymptotic oscillations, according to $\{T_\lambda : |\lambda| = 1\}$. In the case of ergodic processes without deterministic periodicities, every initial condition relaxes to the same steady-state distribution over the hidden states: $\langle \pi_\alpha | = \langle \alpha | T_1 = \langle \pi |$ regardless of α , so long as α is a properly normalized probability distribution.

3.9 Spectra by inSpection: Useful rules for weighted digraphs

As suggested in Ref. [198], these new results extending spectral theory to arbitrary functions of nondiagonalizable operators, presented above, allow a *spectral weighted digraph theory* beyond the purview of ‘spectral graph theory’ proper. Moreover this theory enables new analysis. In the next sections, we show how spectra and eigenprojectors can be intuited, computed, and applied in the analysis of complex systems.

3.9.1 Eigenvalues by inspection

Suppose a directed graph structure has cascading dependencies: that a cluster of nodes feeds back only among itself according to the matrix A , and feeds forward to another cluster of nodes according to the feedforward matrix B (not necessarily a square matrix); this second cluster feeds back only among itself according to the matrix C . The latter cluster of nodes might also feed forward to another cluster of nodes, but such considerations can be applied iteratively. The simple situation we have just described can be summarized (upon proper permutation of indexing) by a block matrix of the form: $W = \begin{bmatrix} A & B \\ \mathbf{0} & C \end{bmatrix}$. In

such a case, it is easy to observe that:

$$\det(W - \lambda I) = \begin{vmatrix} A - \lambda I & B \\ \mathbf{0} & C - \lambda I \end{vmatrix} = |A - \lambda I| |C - \lambda I|, \quad (3.37)$$

and so $\Lambda_W = \Lambda_A \cup \Lambda_C$. This simplification presents an opportunity to read off eigenvalues from clustered graph structures that often appear in practice, especially for transient graph structures.

Cyclic cluster structures (say, of length N and edge-weights α_1 through α_N) yield especially simple spectra:

$$\left\{ \left(\prod_{i=1}^N \alpha_i \right)^{1/N} e^{in2\pi/N} \right\}_{n=0}^{N-1}; \quad (3.38)$$

i.e., the eigenvalues are simply the N^{th} roots of the product of all of the edge-weights.

Similar ‘rules’ for reading off spectra from other cluster structures can be constructed. Although we cannot list them exhaustively here, we give another simple but useful rule in Fig. 3.2b, which also indicates the ubiquity of nondiagonalizability in weighted digraph structures. This second rule is suggestive of further generalizations where spectra can be read off from common digraph motifs.

3.9.2 Eigenprojectors from graph structure

We have just discussed how clustered directed graph structures yield simplified joint spectra. Is there any corresponding simplification of the projection operators? Indeed, there is a nice simplification that allows for iterative construction of ‘higher-level’ projectors from ‘lower-level’ clustered components. In contrast to the joint spectrum though, which completely ignored the feedforward matrix B , the emergent projectors do require B to pull the eigencontributions into the generalized setting. Fig. 3.3 summarizes the results for the simple case of non-degenerate eigenvalues. The general case can be constructed similarly.

The preceding implies a number of algorithms, both for analytic and numerical calculations. Most directly this section points to the fact that the eigen-analysis can be partitioned into a series of simpler problems that are later combined for the final solution. However, in addition to more efficient serial computation, there are opportunities for numerical parallelization of the algorithms to compute the eigenprojectors, whether they are

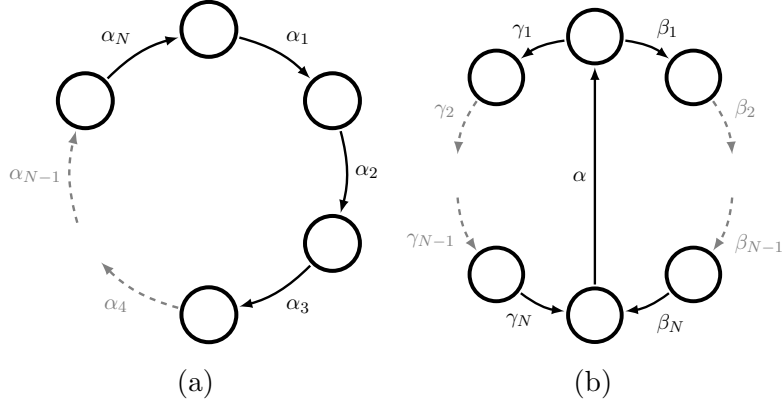


Figure 3.2: (a) Weighted digraphical representation of the feedback matrix A of a cyclic cluster structure, which would contribute eigenvalues $\Lambda_A = \left\{ \left(\prod_{i=1}^N \alpha_i \right)^{1/N} e^{in2\pi/N} \right\}_{n=0}^{N-1}$ with algebraic multiplicities $a_\lambda = 1$ for all $\lambda \in \Lambda_A$. (b) Weighted digraphical representation of the feedback matrix A of a doubly cyclic cluster structure, which would contribute eigenvalues $\Lambda_A = \{0\} \cup \left\{ \left(\alpha \left[\left(\prod_{i=1}^N \beta_i \right) + \left(\prod_{i=1}^N \gamma_i \right) \right] \right)^{\frac{1}{N+1}} e^{\frac{in2\pi}{N+1}} \right\}_{n=0}^N$ with algebraic multiplicities $a_0 = N - 1$ and $a_\lambda = 1$ for $\lambda \neq 0$. (Notice that this eigenvalue ‘rule’ depends on having the same number of β -transitions as γ -transitions.) The 0-eigenvalue only has geometric multiplicity of $g_0 = 1$, so the structure is nondiagonalable for $N > 2$; nevertheless, the generalized eigenvectors are easy to construct. The spectral analysis of the cluster structure in (b) suggests more general rules that can be constructed for reading-off eigenvalues from digraph clusters; e.g., if a *chain* of α ’s appears in the bisecting path.

computed directly, say from Eq. (3.15), or from right and left eigenvectors and generalized eigenvectors. This automation will be useful for applying our analysis to real systems with immense data produced from very high-dimensional state spaces.

With the general spectral theory now outlined, we proceed to the promised analysis of commonly used measures of complexity.

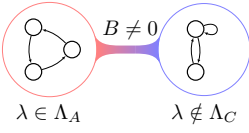
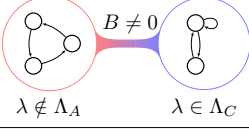
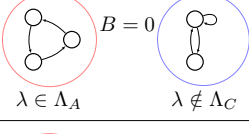
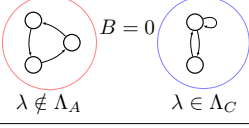
	$W_\lambda = \frac{ \lambda_W\rangle\langle\lambda_W }{\langle\lambda_W \lambda_W\rangle} = \begin{bmatrix} \lambda_A\rangle \\ \vec{0} \end{bmatrix} [\langle\lambda_A , \langle\lambda_A B(\lambda I - C)^{-1}] / \langle\lambda_A \lambda_A\rangle$
$\lambda \in \Lambda_A \quad \lambda \notin \Lambda_C$	
	$W_\lambda = \frac{ \lambda_W\rangle\langle\lambda_W }{\langle\lambda_W \lambda_W\rangle} = \begin{bmatrix} (\lambda I - A)^{-1}B \lambda_C\rangle \\ \lambda_C\rangle \end{bmatrix} [\vec{0}, \langle\lambda_C] / \langle\lambda_C \lambda_C\rangle$
$\lambda \notin \Lambda_A \quad \lambda \in \Lambda_C$	
	$W_\lambda = \frac{ \lambda_W\rangle\langle\lambda_W }{\langle\lambda_W \lambda_W\rangle} = \begin{bmatrix} \lambda_A\rangle \\ \vec{0} \end{bmatrix} [\langle\lambda_A , \vec{0}] / \langle\lambda_A \lambda_A\rangle$
$\lambda \in \Lambda_A \quad \lambda \notin \Lambda_C$	
	$W_\lambda = \frac{ \lambda_W\rangle\langle\lambda_W }{\langle\lambda_W \lambda_W\rangle} = \begin{bmatrix} \vec{0} \\ \lambda_C\rangle \end{bmatrix} [\vec{0}, \langle\lambda_C] / \langle\lambda_C \lambda_C\rangle$
$\lambda \notin \Lambda_A \quad \lambda \in \Lambda_C$	

Figure 3.3: Construction of W -eigenprojectors from ‘lower-level’ A -projectors and C -projectors, when $W = \begin{bmatrix} A & B \\ \mathbf{0} & C \end{bmatrix}$. Recall that $(\lambda I - A)^{-1}$ and $(\lambda I - C)^{-1}$ can be constructed from the lower-level projectors. For simplicity, we assume that the algebraic multiplicity $a_\lambda = 1$ in each of these cases.

3.10 The Breadth of Possible Behaviors for Correlation and Myopic Uncertainty

In light of Eq. (3.27), the autocorrelation function is simply a superposition of weighted eigen-contributions. Most generally, the spectral decomposition of $\gamma(L) = \langle \pi \bar{\mathcal{A}} | T^{|L|-1} | \mathcal{A} \mathbf{1} \rangle$ for nonzero integer L is:

$$\begin{aligned}
\gamma(L) &= \sum_{\substack{\lambda \in \Lambda_T \\ \lambda \neq 0}} \sum_{m=0}^{\nu_\lambda-1} \langle \pi \bar{\mathcal{A}} | T_{\lambda,m} | \mathcal{A} \mathbf{1} \rangle \binom{|L|-1}{m} \lambda^{|L|-1-m} \\
&\quad + [0 \in \Lambda_T] \sum_{m=0}^{\nu_0-1} \langle \pi \bar{\mathcal{A}} | T_0 T^m | \mathcal{A} \mathbf{1} \rangle \delta_{|L|-1,m}
\end{aligned} \tag{3.39}$$

$$= \gamma_{\rightsquigarrow}(L) + \gamma_{\rightarrow 0}(L) . \tag{3.40}$$

We will denote the persistent first term of Eq. (3.39) as $\gamma_{\rightsquigarrow}$, and note that it can be summarized by: $\gamma_{\rightsquigarrow}(L) = \langle \pi \bar{\mathcal{A}} | T^{\mathcal{D}} T T^{|L|-1} | \mathcal{A} \mathbf{1} \rangle = \langle \pi \bar{\mathcal{A}} | T^{\mathcal{D}} T^{|L|} | \mathcal{A} \mathbf{1} \rangle$, where $T^{\mathcal{D}}$ is the Drazin inverse of the transition matrix T . We will denote the ephemeral second term as

$\gamma_{-\circ}$, which can be written as: $\gamma_{-\circ}(L) = \langle \pi \bar{\mathcal{A}} | T_0 T^{|L|-1} | \mathcal{A} \mathbf{1} \rangle$, where T_0 is the eigenprojector associated with the eigenvalue of zero ($T_0 = \mathbf{0}$ if $0 \notin \Lambda_T$).

From Eq. (3.39), it is now apparent that the index of the zero eigenvalue of T gives a finite horizon contribution ($\gamma_{-\circ}$) to the autocorrelation function. Beyond ν_0 of T , the only L -dependence comes via a weighted sum of terms of the form $\binom{|L|-1}{m} \lambda^{|L|-1-m}$ —polynomials in L times decaying exponentials. The set $\{\langle \pi \bar{\mathcal{A}} | T_{\lambda,m} | \mathcal{A} \mathbf{1} \rangle\}$ simply weighs the amplitude of these contributions. In the diagonalizable case, the behavior of autocorrelation is simply a sum of decaying exponentials $\lambda^{|L|}$.

Similarly, in light of Eq. (3.8) and Eq. (3.27), we find the most general spectral decomposition of the myopic entropy rates $h_\mu(L) = \langle \delta_\pi | W^{L-1} | H(W^{\mathcal{A}}) \rangle$ to be:

$$\begin{aligned} h_\mu(L) &= \sum_{\substack{\lambda \in \Lambda_W \\ \lambda \neq 0}} \sum_{m=0}^{\nu_\lambda-1} \langle \delta_\pi | W_{\lambda,m} | H(W^{\mathcal{A}}) \rangle \binom{L-1}{m} \lambda^{L-1-m} \\ &\quad + [0 \in \Lambda_W] \sum_{m=0}^{\nu_0-1} \delta_{L-1,m} \langle \delta_\pi | W_0 W^m | H(W^{\mathcal{A}}) \rangle \end{aligned} \quad (3.41)$$

$$= h_{\rightsquigarrow}(L) + h_{-\circ}(L) . \quad (3.42)$$

We will denote the persistent first term of (3.41) as h_{\rightsquigarrow} , and note that it can be summarized by: $h_{\rightsquigarrow}(L) = \langle \delta_\pi | W^{\mathcal{D}} W W^{L-1} | H(W^{\mathcal{A}}) \rangle = \langle \delta_\pi | W^{\mathcal{D}} W^L | H(W^{\mathcal{A}}) \rangle$, where $W^{\mathcal{D}}$ is the Drazin inverse of the mixed-state-to-state net transition dynamic W . We will denote the ephemeral second term as $h_{-\circ}$, which can be written as: $h_{-\circ}(L) = \langle \delta_\pi | W_0 W^{L-1} | H(W^{\mathcal{A}}) \rangle$.

From Eq. (3.41), it is now apparent that the index of the zero eigenvalue of W gives a finite horizon contribution ($h_{-\circ}$) to the myopic entropy rate. Beyond ν_0 of W , the only L -dependence comes via a weighted sum of terms of the form $\binom{L-1}{m} \lambda^{L-1-m}$ —polynomials in L times decaying exponentials. The set $\{\langle \delta_\pi | W_{\lambda,m} | H(W^{\mathcal{A}}) \rangle\}$ simply weighs the amplitude of these contributions. For stationary processes we anticipate that, for all $\zeta \in \{\lambda \in \Lambda_W : |\lambda| = 1, \lambda \neq 1\}$, $\langle \delta_\pi | W_\zeta = \mathbf{0}$ and thus $\langle \delta_\pi | W_\zeta | H(W^{\mathcal{A}}) \rangle = 0$; hence we can save ourselves from superfluous calculation by excluding the non-unity eigenvalues on the unit circle when calculating the myopic entropy rate for stationary processes. In the diagonalizable case, the behavior of autocorrelation is simply a sum of decaying exponentials λ^L .

In practice $\gamma_{-\circ}$ is very often zero whereas $h_{-\circ}$ is very often non-zero. This practical difference between $\gamma_{-\circ}$ and $h_{-\circ}$ stems from the difference between typical graph structures of the respective dynamics. For a generic transition dynamic for a stationary process, zero eigenvalues (and thus $\nu_0(T)$) of T typically come from hidden symmetries in the dynamic. In contrast, the MSP of a generic transition dynamic often has tree-like ephemeral structures that are primarily responsible for the zero eigenvalues (and $\nu_0(W)$). Nevertheless, despite their practical typical differences, the same mathematical structures appear and contribute to the most general behavior of each of these cascading dynamics.

The breadth of possible behavior qualitatively shared by autocorrelation and myopic entropy rate is common to the solution of all questions that can be reformulated as a cascading hidden linear dynamic; $\mathcal{H}^+(L)$ is just one of many other examples. But, as we have already seen, different measures of a process yield the signatures of different linear operators.

Next, we expose the similarity in the qualitative behavior of asymptotics, and discuss the implications for correlation and entropy rate.

3.10.1 Asymptotic correlation

From the spectral decomposition, it is apparent that the autocorrelation converges to some constant value as $L \rightarrow \infty$, unless T has eigenvalues on the unit circle besides unity itself. This is under the assumption that ν_0 is finite, which is true for all processes generated by a finite-state HMM and also many infinite-state HMMs. If unity is the sole eigenvalue with a magnitude of one, then all other eigenvalues have magnitude less than unity and their contributions decay to negligibility for large enough L . Explicitly, if $\operatorname{argmax}_{\lambda \in \Lambda_T} |\lambda| = \{1\}$,

then:

$$\begin{aligned}
\lim_{L \rightarrow \infty} \gamma(L) &= \lim_{L \rightarrow \infty} \sum_{\substack{\lambda \in \Lambda_T \\ \lambda \neq 0}} \sum_{m=0}^{\nu_\lambda-1} \langle \pi \bar{\mathcal{A}} | T_{\lambda,m} | \mathcal{A} \mathbf{1} \rangle \binom{L-1}{m} \lambda^{L-1-m} \\
&= \langle \pi \bar{\mathcal{A}} | T_1 | \mathcal{A} \mathbf{1} \rangle \\
&= \langle \pi \bar{\mathcal{A}} | \mathbf{1} \rangle \langle \pi | \mathcal{A} \mathbf{1} \rangle \\
&= \left| \sum_{x \in \mathcal{A}} x \text{Pr}(x) \right|^2 \\
&= |\langle x \rangle|^2.
\end{aligned}$$

Above, we used the fact that $\nu_1 = 1$ and that—for an ergodic process— $T_1 = |\mathbf{1}\rangle \langle \pi|$.

If other eigenvalues in Λ_T beside unity exist on the unit circle, then the autocorrelation approaches a periodic sequence as L gets large.

3.10.2 Asymptotic entropy rate

By the Perron–Frobenius theorem, $\nu_\lambda = 1$ for all eigenvalues of W on the unit circle. Hence, In the limit of $L \rightarrow \infty$ we obtain the asymptotic entropy rate for any stationary process as:

$$h_\mu \equiv \lim_{L \rightarrow \infty} h_\mu(L) \quad (3.43)$$

$$= \lim_{L \rightarrow \infty} \sum_{\substack{\lambda \in \Lambda_W \\ |\lambda|=1}} \lambda^{L-1} \langle \delta_\pi | W_\lambda | H(W^\mathcal{A}) \rangle \quad (3.44)$$

$$= \langle \delta_\pi | W_1 | H(W^\mathcal{A}) \rangle, \quad (3.45)$$

since, for stationary processes, $\langle \delta_\pi | W_\zeta = \mathbf{0}$ for all $\zeta \in \{\lambda \in \Lambda_W : |\lambda| = 1, \lambda \neq 1\}$. For non-stationary processes, this limit may not exist, but h_μ may still be found in a suitable sense as a function of time. If the process has only one stationary distribution over mixed states, then $W_1 = |\mathbf{1}\rangle \langle \pi_W|$, and:

$$h_\mu = \langle \pi_W | H(W^\mathcal{A}) \rangle, \quad (3.46)$$

where π_W is the stationary distribution over the states of W , found either from $\langle \pi_W | = \langle \delta_\pi | W_1$ or from solving $\langle \pi_W | W = \langle \pi_W |$.

A simple but interesting example of when ergodicity is *not* the case, is for the multi-armed bandit problem, where any realization is drawn from a distribution over different biased coins (or over any other collection of IID processes) [47, 24, 63]. More generally, there could be many different memoryful structured attractors that any realization chooses amongst, with a probability distribution atop the space of attractors. In the case of many attractors (numbering the algebraic multiplicity $a_1(T) = a_1(W)$), we have the stationary mixed state eigenprojector $W_1 = \sum_{k=1}^{a_1} |1_k\rangle \langle 1_k|$ with $\langle 1_j | 1_k \rangle = \delta_{j,k}$, and the entropy becomes:

$$h_\mu = \sum_{k=1}^{a_1} \langle \delta_\pi | 1_k \rangle \langle 1_k | H(W^\mathcal{A}) \rangle \quad (3.47)$$

$$= \langle h_\mu^{(\text{attractor } k)} \rangle_k . \quad (3.48)$$

Above, $\langle \delta_\pi | 1_k \rangle$ can be interpreted as the probability of ending up in attractor k while $\langle 1_k | H(W^\mathcal{A}) \rangle$ can be interpreted as the entropy rate for sequences generated within attractor k . Thus, the entropy of the *process* may not be the same as the entropy of any particular realization if a process is non-ergodic. Rather, the entropy of the process is a weighted average of the entropies for the ensemble of sequences that constitute the process.

For unifilar \mathcal{M} , the topology, transition probabilities, and stationary distribution over the attracting set of states is the same for both \mathcal{M} and its S-MSP. Hence, we have for unifilar \mathcal{M} :

$$\begin{aligned} h_\mu &= \langle \pi_W | H(W^\mathcal{A}) \rangle \\ &= \langle \pi | H(T^\mathcal{A}) \rangle . \end{aligned} \quad (3.49)$$

One can easily show that Eq. (3.49) is equivalent to the well known closed-form expression for h_μ for unifilar presentations:

$$\langle \pi | H(T^\mathcal{A}) \rangle = - \sum_{\sigma \in \mathcal{S}} \text{Pr}(\sigma) \sum_{x \in \mathcal{A}} \sum_{\sigma' \in \mathcal{S}} T_{\sigma, \sigma'}^{(x)} \log_2(T_{\sigma, \sigma'}^{(x)}) . \quad (3.50)$$

For nonunifilar presentations, however, we must use the more general result of Eq. (3.46). This is similar to the calculation in Eq. (3.50), but must be performed over the recurrent states of a mixed-state presentation.

3.11 Accumulated Transients for Diagonalizable Dynamics

In the diagonalizable case, autocorrelation, myopic entropy rate, and myopic state uncertainty reduce to a sum of decaying exponentials. Correspondingly, we could find the power spectrum, excess entropy, and synchronization information respectively via geometric progression.

For example, if W is diagonalizable and has no zero eigenvalue, then the myopic entropy rate reduces to:

$$\begin{aligned} h_\mu(L) &= \sum_{\lambda \in \Lambda_W} \langle \delta_\pi | W_\lambda | H(W^{\mathcal{A}}) \rangle \lambda^{L-1} \\ &= \langle \delta_\pi | W_1 | H(W^{\mathcal{A}}) \rangle + \sum_{\substack{\lambda \in \Lambda_W \\ |\lambda| < 1}} \lambda^{L-1} \langle \delta_\pi | W_\lambda | H(W^{\mathcal{A}}) \rangle , \end{aligned}$$

where $\langle \delta_\pi | W_1 | H(W^{\mathcal{A}}) \rangle$ is identifiable as the entropy rate h_μ .

It then follows that the excess entropy is:

$$\begin{aligned} \mathbf{E} &\equiv \sum_{L=1}^{\infty} [h_\mu(L) - h_\mu] \\ &= \sum_{L=1}^{\infty} \sum_{\substack{\lambda \in \Lambda_W \\ |\lambda| < 1}} \lambda^{L-1} \langle \delta_\pi | W_\lambda | H(W^{\mathcal{A}}) \rangle \\ &= \sum_{\substack{\lambda \in \Lambda_W \\ |\lambda| < 1}} \langle \delta_\pi | W_\lambda | H(W^{\mathcal{A}}) \rangle \underbrace{\sum_{L=1}^{\infty} \lambda^{L-1}}_{=\sum_{L=0}^{\infty} \lambda^L = \frac{1}{1-\lambda}} \\ &= \sum_{\substack{\lambda \in \Lambda_W \\ |\lambda| < 1}} \frac{1}{1-\lambda} \langle \delta_\pi | W_\lambda | H(W^{\mathcal{A}}) \rangle . \end{aligned} \tag{3.51}$$

Note that larger eigenvalues (closer to unity magnitude) make the denominator $1 - \lambda$ closer to zero and, thus, make $\frac{1}{1-\lambda}$ larger. Hence, larger eigenvalues—modes of the mixed-state transition matrix that are slow to decay—have the potential to make the largest contributions to excess entropy. Small eigenvalues—modes of the mixed-state transition matrix that are quick to decay—cannot make much of a contribution to excess entropy. Stepping away from the language of eigenvalues, one can paraphrase the implicit lesson as:

slowly decaying transient behavior (of the distribution over distributions over the state of a process) has the most potential to make a process appear complex.

Moreover, the transient information is:

$$\begin{aligned}
\mathbf{T} &\equiv \sum_{L=1}^{\infty} L [h_{\mu}(L) - h_{\mu}] \\
&= \sum_{L=1}^{\infty} \sum_{\substack{\lambda \in \Lambda_W \\ |\lambda| < 1}} L \lambda^{L-1} \langle \delta_{\pi} | W_{\lambda} | H(W^{\mathcal{A}}) \rangle \\
&= \sum_{\substack{\lambda \in \Lambda_W \\ |\lambda| < 1}} \langle \delta_{\pi} | W_{\lambda} | H(W^{\mathcal{A}}) \rangle \underbrace{\sum_{L=1}^{\infty} L \lambda^{L-1}}_{=\sum_{L=0}^{\infty} \frac{d}{d\lambda} \lambda^L = \frac{d}{d\lambda} (\sum_{L=0}^{\infty} \lambda^L) = \frac{d}{d\lambda} (\frac{1}{1-\lambda}) = \frac{1}{(1-\lambda)^2}} \\
&= \sum_{\substack{\lambda \in \Lambda_W \\ |\lambda| < 1}} \frac{1}{(1-\lambda)^2} \langle \delta_{\pi} | W_{\lambda} | H(W^{\mathcal{A}}) \rangle .
\end{aligned}$$

We now see that the transient information is very closely related to the excess entropy, differing only via the square in the denominators. This comparison between \mathbf{E} and \mathbf{T} closed-form expressions suggests an entire hierarchy of informational quantities based on eigenvalue weighting.

Performing a similar procedure for \mathbf{S}' shows that:

$$\begin{aligned}
\mathbf{S}' &\equiv \sum_{L=0}^{\infty} [\mathcal{H}(L) - \mathcal{H}] \\
&= \sum_{L=0}^{\infty} \sum_{\substack{\lambda \in \Lambda_W \\ |\lambda| < 1}} \langle \delta_{\pi} | W_{\lambda} | H[\eta] \rangle \lambda^L \\
&= \sum_{\substack{\lambda \in \Lambda_W \\ |\lambda| < 1}} \langle \delta_{\pi} | W_{\lambda} | H[\eta] \rangle \sum_{L=0}^{\infty} \lambda^L \\
&= \sum_{\substack{\lambda \in \Lambda_W \\ |\lambda| < 1}} \frac{1}{1-\lambda} \langle \delta_{\pi} | W_{\lambda} | H[\eta] \rangle .
\end{aligned}$$

The expressions reveal a rather remarkable close relationship between \mathbf{S}' and \mathbf{E} . Define $\langle \cdot | \equiv \sum_{L=0}^{\infty} \langle \delta_{\pi} | W^L$. Then:

$$\langle \cdot | = \sum_{\substack{\lambda \in \Lambda_W \\ |\lambda| < 1}} \frac{1}{1-\lambda} \langle \delta_{\pi} | W_{\lambda} .$$

The relationship is now made plain:

$$\begin{aligned}\mathbf{E} &= \langle \cdot | H(W^{\mathcal{A}}) \rangle \quad \text{and} \\ \mathbf{S}' &= \langle \cdot | H[\eta] \rangle \quad .\end{aligned}$$

Although a bit more cumbersome, perhaps better intuition emerges if we rewrite $\langle \cdot |$ as $\langle \int \Pr(\eta, L) dL |$.

Again, large eigenvalues—slowly decaying modes of the mixed-state transition matrix—can make the largest contribution to synchronization information; small eigenvalues correspond to quickly decaying modes that do not have the opportunity to contribute. In fact, the potential of large eigenvalues to make large contributions is a recurring theme for many questions one has about a process. Simply stated, long-term behavior—what we often interpret as “complex” behavior—is dominated by a process’s largest-eigenvalue modes.

However, a word of warning must be attached. Although large-eigenvalue modes have the most *potential* to make contributions to a process’s complexity, the actual set of largest contributors also depends strongly on the amplitudes $\{\langle \delta_\pi | W_\lambda | \dots \rangle\}$, where $|\dots\rangle$ is some measure-vector of interest; e.g., $|\dots\rangle = |H[\eta]\rangle, |H(W^{\mathcal{A}})\rangle, |\mathbf{1}\rangle$, or the like.

Hence, there is startling similarity between \mathbf{E} and \mathbf{T} and another startling similarity between \mathbf{E} and \mathbf{S}' —at least this is hinted at under the assumption of diagonalizability. We would like to know the relationships among these quantities in general. However, deriving the general closed-form expressions for accumulated transients is hardly tractable by this approach. Rather, to derive the general results, we perform the meromorphic functional calculus directly at an elevated level, as we demonstrate next.

3.12 Exact Complexities and Complexity Spectra for Nondiagonalizable Dynamics

We now derive the most general closed-form solutions for several complexity measures, from which the others should become obvious. Before carrying this out, we define several

useful objects. Let $\rho(A)$ to be the spectral radius of the matrix A :

$$\rho(A) = \max_{\lambda \in \Lambda_A} |\lambda| .$$

For stochastic W , we define $\Lambda_{\rho(W)}$ to be the set of eigenvalues with unity magnitude (since $\rho(W) = 1$):

$$\Lambda_{\rho(W)} = \{\lambda \in \Lambda_W : |\lambda| = 1\} .$$

We also define

$$Q \equiv W - W_1 \tag{3.52}$$

and

$$\mathcal{Q} \equiv W - \sum_{\lambda \in \Lambda_{\rho(W)}} \lambda W_\lambda . \tag{3.53}$$

Eigenvalues with unity magnitude that are not themselves unity correspond to perfectly periodic cycles of the state-transition dynamic. By their very nature, such cycles are restricted to the recurrent states. Moreover, we expect the projection operators associated with these cycles to have no net overlap with the start-state of the MSP. So, we expect:

$$\langle \delta_\pi | W_\lambda = \vec{0} \quad \text{for all } \lambda \in \Lambda_{\rho(W)} \setminus \{1\} . \tag{3.54}$$

Hence:

$$\langle \delta_\pi | Q^L = \langle \delta_\pi | \mathcal{Q}^L . \tag{3.55}$$

We will also use the fact that, since $\rho(Q) < 1$,

$$\sum_{L=0}^{\infty} \mathcal{Q}^L = (I - \mathcal{Q})^{-1} ;$$

and furthermore:

$$\langle \delta_\pi | (I - \mathcal{Q})^{-1} = \langle \delta_\pi | (I - Q)^{-1}$$

as a consequence of Eq. (3.54) and our spectral decomposition.

We have seen that the complexity measures associated with prediction all take on a similar form in terms of the \mathcal{S} -MSP state-transition matrix, and so we expect to encounter similar forms for generically nondiagonalizable state-transition dynamics.

3.12.1 Excess Entropy

We are now ready to develop the excess entropy in full generality. Calculating, we find:

$$\begin{aligned}
\mathbf{E} &\equiv \sum_{L=1}^{\infty} [h_{\mu}(L) - h_{\mu}] \\
&= \sum_{L=1}^{\infty} [\langle \delta_{\pi} | W^{L-1} | H(W^{\mathcal{A}}) \rangle - \langle \delta_{\pi} | W_1 | H(W^{\mathcal{A}}) \rangle] \\
&= \sum_{L=0}^{\infty} [\langle \delta_{\pi} | W^L | H(W^{\mathcal{A}}) \rangle - \langle \delta_{\pi} | W_1 | H(W^{\mathcal{A}}) \rangle] \\
&= \sum_{L=0}^{\infty} \langle \delta_{\pi} | [\underbrace{(W - W_1)^L}_{\equiv Q} - \delta_{L,0} W_1] | H(W^{\mathcal{A}}) \rangle \\
&= - \underbrace{\langle \delta_{\pi} | W_1 | H(W^{\mathcal{A}}) \rangle}_{=h_{\mu}} + \sum_{L=0}^{\infty} \underbrace{\langle \delta_{\pi} | Q^L | H(W^{\mathcal{A}}) \rangle}_{=\langle \delta_{\pi} | \mathcal{Q}^L} \\
&= \langle \delta_{\pi} | \left(\sum_{L=0}^{\infty} Q^L \right) | H(W^{\mathcal{A}}) \rangle - h_{\mu} \\
&= \langle \delta_{\pi} | (I - Q)^{-1} | H(W^{\mathcal{A}}) \rangle - h_{\mu} \\
&= \langle \delta_{\pi} | (I - Q)^{-1} | H(W^{\mathcal{A}}) \rangle - h_{\mu} .
\end{aligned}$$

Note that $(I - Q)^{-1} = \text{inv}(I - Q)$ here, since unity is not an eigenvalue of Q . Indeed, the unity eigenvalue was explicitly pulled out of the former matrix to make an invertible expression. For an ergodic process, where $W_1 = |\mathbf{1}\rangle \langle \pi_W|$, this becomes:

$$\mathbf{E} = \langle \delta_{\pi} | (I - W + |\mathbf{1}\rangle \langle \pi_W|)^{-1} | H(W^{\mathcal{A}}) \rangle - h_{\mu} . \quad (3.56)$$

Computationally, Eq. (3.56) is wonderfully useful. However, the h_{μ} subtracted off the end of the expression is at first quite mysterious. Especially so, when compared to the nice result for the spectral decomposition of the excess entropy in the diagonalizable case in Eq. (3.51).

From Eq. (3.32), we see that the general solution for \mathbf{E} takes on its most elegant form

in terms of the Drazin inverse of $I - W$:

$$\begin{aligned}
\mathbf{E} &= \langle \delta_\pi | (I - Q)^{-1} | H(W^{\mathcal{A}}) \rangle - h_\mu \\
&= \langle \delta_\pi | (I - Q)^{-1} | H(W^{\mathcal{A}}) \rangle - \langle \delta_\pi | W_1 | H(W^{\mathcal{A}}) \rangle \\
&= \langle \delta_\pi | [(I - Q)^{-1} - W_1] | H(W^{\mathcal{A}}) \rangle \\
&= \langle \delta_\pi | (I - W)^{\mathcal{D}} | H(W^{\mathcal{A}}) \rangle .
\end{aligned} \tag{3.57}$$

From Eq. (3.57) and Eq. (3.33), we see that the past–future mutual information—the amount of the future that is *predictable* from the past—has the general spectral decomposition:

$$\mathbf{E} = \sum_{\lambda \in \Lambda_W \setminus \{1\}} \sum_{m=0}^{\nu_\lambda-1} \frac{1}{(1-\lambda)^{m+1}} \langle \delta_\pi | W_{\lambda,m} | H(W^{\mathcal{A}}) \rangle . \tag{3.58}$$

3.12.1.1 Persistent excess

In light of Eq. (3.42), it should be clear that there are two qualitatively distinct contributions to the excess entropy $\mathbf{E} = \mathbf{E}_{\rightsquigarrow} + \mathbf{E}_{\rightarrow\circ}$: One is the persistent leaky contributions from all L :

$$\mathbf{E}_{\rightsquigarrow} \equiv \sum_{L=1}^{\infty} [h_{\rightsquigarrow}(L) - h_\mu] \tag{3.59}$$

$$= \langle \delta_\pi | W^{\mathcal{D}} W (I - W)^{\mathcal{D}} | H(W^{\mathcal{A}}) \rangle \tag{3.60}$$

and the other is a completely ephemeral piece that contributes only up to W 's zero-eigenvalue index ν_0 :

$$\mathbf{E}_{\rightarrow\circ} \equiv \sum_{L=1}^{\infty} h_{\rightarrow\circ}(L) = \sum_{L=1}^{\nu_0} h_{\rightarrow\circ}(L) \tag{3.61}$$

$$= \langle \delta_\pi | W_0 (I - W)^{\mathcal{D}} | H(W^{\mathcal{A}}) \rangle . \tag{3.62}$$

3.12.2 The Excess Entropy Spectrum

Equation (3.57) immediately suggests that we generalize our notion of the excess entropy scalar complexity measure to a complexity *function* with continuous part defined in terms of the resolvent—something like $\langle \delta_\pi | (zI - W)^{-1} | H(W^{\mathcal{A}}) \rangle$.² Such a function would not

² However, there is ambiguity surrounding the appropriate generalization of the excess entropy. There are many functions that we could invent that approach the excess entropy as we let frequency go to zero. For

only say *how much* of the future is predictable, but would go much further to unveil the *timescales of interdependence* among the predictable features within the observations. Directly taking the z -transform of $h_\mu(L)$ comes to mind, but we would like to avoid having to track both real and imaginary parts (or, alternatively, both the magnitude and complex phase). To ameliorate this, we consider the transform of a closely related function that contains the same information.

Let $\mathfrak{d}h$ be the two-sided myopic entropy convergence function defined by:

$$\mathfrak{d}h(L) = \begin{cases} H(X_0|X_{-|L|+1:0}) & \text{for } L < 0 , \\ \log_2(\mathcal{A}) & \text{for } L = 0 , \text{ and} \\ H(X_0|X_{1:L}) & \text{for } L > 0 . \end{cases} \quad (3.63)$$

For stationary processes, it is easy to show that $H(X_0|X_{-L+1:0}) = H(X_0|X_{1:L})$, with the result that $\mathfrak{d}h$ is a symmetric function. Moreover, $\mathfrak{d}h$ then simplifies to:

$$\mathfrak{d}h(L) = h_\mu(|L|) , \quad (3.64)$$

where $h_\mu(0) \equiv \log_2(\mathcal{A})$ and, as before, $h_\mu(L) = H(X_L|X_{1:L})$ for $L \geq 1$ with $h_\mu(1) = H(X_1)$.

The symmetry of the two-sided myopic entropy convergence function $\mathfrak{d}h$ guarantees that its Fourier transform is also real and symmetric. Explicitly, the continuous part of the Fourier transform turns out to be:

$$\widetilde{\mathfrak{d}h}_c(\omega) = \mathbf{R} + 2\text{Re} \langle \delta_\pi | (e^{i\omega} I - W)^{-1} | H(W^{\mathcal{A}}) \rangle , \quad (3.65)$$

a strictly real and symmetric function of the angular frequency ω . Here, \mathbf{R} is the redundancy of the alphabet $\mathbf{R} \equiv \log_2 |\mathcal{A}| - h_\mu$, as in Ref. [62]. The transform, $\widetilde{\mathfrak{d}h}$, also has a discrete impulsive component; for stationary processes this consists solely of the

example, directly calculating from the meromorphic functional calculus, we find: $\lim_{\omega \rightarrow 0} \text{Re} \langle \delta_\pi | (e^{i\omega} I - W)^{-1} | H(W^{\mathcal{A}}) \rangle = \mathbf{E} - \frac{1}{2}h_\mu$, but we are challenged to make sense of the facts that $\text{Re} \langle \delta_\pi | (e^{i\omega} I - W)^{-1} | H(W^{\mathcal{A}}) \rangle + \frac{1}{2}h_\mu$ is not necessarily positive at all frequencies. Another direct calculation shows that $\lim_{\omega \rightarrow 0} \text{Re} \langle \delta_\pi | e^{i\omega} (e^{i\omega} I - W)^{-1} | H(W^{\mathcal{A}}) \rangle = \mathbf{E} + \frac{1}{2}h_\mu$. Enticingly, $\text{Re} \langle \delta_\pi | e^{i\omega} (e^{i\omega} I - W)^{-1} | H(W^{\mathcal{A}}) \rangle - \frac{1}{2}h_\mu$ appears to be positive over all frequencies for all examples checked. Which, if either, hints at the appropriate generalization? As we will see, the Fourier transform of the two-sided myopic entropy convergence function will make our definition of $\mathcal{E}(\omega)$ interpretable and of interest in its own right.

Dirac delta function at zero frequency:

$$\widetilde{\text{d}h}_{\text{d}}(\omega) = 2\pi h_{\mu} \sum_{k \in \mathbb{Z}} \delta(\omega + 2\pi k) ; \quad (3.66)$$

recall that the Fourier transform of a discrete-domain function is 2π -periodic in the angular frequency ω . This delta function is associated with the non-zero offset of the entropy convergence curve of entropic processes. The full transform is $\widetilde{\text{d}h}(\omega) = \widetilde{\text{d}h}_{\text{c}}(\omega) + \widetilde{\text{d}h}_{\text{d}}(\omega)$.

Direct calculation using the meromorphic functional calculus of Ref. [198] shows that:

$$\lim_{\omega \rightarrow 0} \text{Re} \langle \delta_{\pi} | (e^{i\omega} I - W)^{-1} | H(W^{\mathcal{A}}) \rangle = \mathbf{E} - \frac{1}{2} h_{\mu} . \quad (3.67)$$

This motivates the introduction of the *excess entropy spectrum* $\mathcal{E}(\omega)$:

$$\mathcal{E}(\omega) \equiv \frac{1}{2} (\widetilde{\text{d}h}(\omega) - \mathbf{R} + h_{\mu}) \quad (3.68)$$

$$= \text{Re} \langle \delta_{\pi} | (e^{i\omega} I - W)^{-1} | H(W^{\mathcal{A}}) \rangle + \frac{1}{2} h_{\mu} . \quad (3.69)$$

The excess entropy spectrum immediately displays important frequencies of apparent entropy reduction. For example, leaky period-5 processes, should have a period-5 signature in the excess entropy spectrum.

As its predecessors, the excess entropy spectrum also has the natural decomposition into it's two qualitatively distinct parts:

$$\mathcal{E}(\omega) = \mathcal{E}_{\rightsquigarrow}(\omega) + \mathcal{E}_{\rightarrow 0}(\omega) . \quad (3.70)$$

The excess entropy spectrum can serve as an intuitive and concise summary of the complexities associated with the predictability of a process. For example, given a graph of the excess entropy spectrum, the past–future mutual information can be read off as the height of the continuous part of the function as it approaches zero frequency:

$$\begin{aligned} \mathbf{E} &= \lim_{\omega \rightarrow 0} \mathcal{E}(\omega) \\ &= \mathcal{E}_{\text{c}}(\omega = 0) . \end{aligned}$$

Indeed, the *limit* of zero frequency is necessary because of the delta function in the Fourier transform at exactly zero frequency:

$$h_{\mu} = \lim_{\epsilon \rightarrow 0} \frac{1}{\pi} \int_{-\epsilon}^{\epsilon} \mathcal{E}(\omega) d\omega .$$

In some sense, the delta function represents one of the reasons why the excess entropy had been difficult to compute in the past. This can also shed light on the role of the Drazin inverse: to remove the infinite asymptotic accumulation to yield the contribution of the interesting transient structure of entropy reduction.

Notice the spectral decomposition of the excess entropy spectrum:

$$\mathcal{E}(\omega) = \sum_{\lambda \in \Lambda_W} \sum_{m=0}^{\nu_\lambda-1} \operatorname{Re} \left(\frac{\langle \delta_\pi | W_{\lambda,m} | H(W^A) \rangle}{(e^{i\omega} - \lambda)^{m+1}} \right) \quad (3.71)$$

$$\begin{aligned} &= \sum_{m=0}^{\nu_0-1} \cos((m+1)\omega) \langle \delta_\pi | W_0 W^m | H(W^A) \rangle \\ &\quad + \sum_{\lambda \in \Lambda_W \setminus 0} \sum_{m=0}^{\nu_\lambda-1} \operatorname{Re} \left(\frac{\langle \delta_\pi | W_{\lambda,m} | H(W^A) \rangle}{(e^{i\omega} - \lambda)^{m+1}} \right), \end{aligned} \quad (3.72)$$

where, in the last equality, we assume that W_0 is real. This shows that, in addition to the contribution of typical leaky modes of decay in entropy reduction, the zero-eigenvalue modes contribute uniquely to the excess entropy spectrum. In addition to Lorentzian-like curves from leaky periodicities in the MSP, the excess entropy spectrum also contains sums of cosines up to a frequencies of ν_0 that corresponds to the depth of nondiagonalizability for the MSP—the duration of ephemeral synchronization in the time domain.

3.12.3 Synchronization Information

Once expressed in terms of the \mathcal{S} -MSP transition dynamic, the derivation of the excess synchronization information \mathbf{S}' closely parallels that of the excess entropy, only with a

different $|\cdot\rangle$ appended to the end of each line. To wit:

$$\begin{aligned}
\mathbf{S}' &\equiv \sum_{L=0}^{\infty} [\mathcal{H}(L) - \mathcal{H}] \\
&= \sum_{L=0}^{\infty} [\langle \delta_{\pi} | W^L | H[\eta] \rangle - \langle \delta_{\pi} | W_1 | H[\eta] \rangle] \\
&= \sum_{L=0}^{\infty} \langle \delta_{\pi} | [\underbrace{(W - W_1)^L}_{\equiv Q} - \delta_{L,0} W_1] | H[\eta] \rangle \\
&= - \underbrace{\langle \delta_{\pi} | W_1 | H[\eta] \rangle}_{=\mathcal{H}} + \sum_{L=0}^{\infty} \underbrace{\langle \delta_{\pi} | Q^L | H[\eta] \rangle}_{=\langle \delta_{\pi} | Q^L} \\
&= \langle \delta_{\pi} | \left(\sum_{L=0}^{\infty} Q^L \right) | H[\eta] \rangle - \mathcal{H} \\
&= \langle \delta_{\pi} | (I - Q)^{-1} | H[\eta] \rangle - \mathcal{H} \\
&= \langle \delta_{\pi} | (I - Q)^{-1} | H[\eta] \rangle - \mathcal{H} .
\end{aligned}$$

For an ergodic process where $W_1 = |\mathbf{1}\rangle \langle \pi_W|$, this becomes:

$$\mathbf{S}' = \langle \delta_{\pi} | (I - W + |\mathbf{1}\rangle \langle \pi_W|)^{-1} | H[\eta] \rangle - \mathcal{H} . \quad (3.73)$$

From Eq. (3.32), we see that the general solution for \mathbf{S}' takes on its most elegant form in terms of the Drazin inverse of $I - W$:

$$\begin{aligned}
\mathbf{S}' &= \langle \delta_{\pi} | (I - Q)^{-1} | H[\eta] \rangle - \mathcal{H} \\
&= \langle \delta_{\pi} | (I - Q)^{-1} | H[\eta] \rangle - \langle \delta_{\pi} | W_1 | H[\eta] \rangle \\
&= \langle \delta_{\pi} | [(I - Q)^{-1} - W_1] | H[\eta] \rangle \\
&= \langle \delta_{\pi} | (I - W)^{\mathcal{D}} | H[\eta] \rangle .
\end{aligned} \quad (3.74)$$

From Eq. (3.74) and Eq. (3.33), we see that the excess synchronization information has the general spectral decomposition:

$$\mathbf{S}' = \sum_{\lambda \in \Lambda_W \setminus \{1\}} \sum_{m=0}^{\nu_{\lambda}-1} \frac{1}{(1 - \lambda)^{m+1}} \langle \delta_{\pi} | W_{\lambda,m} | H[\eta] \rangle . \quad (3.75)$$

Again the form of Eq. (3.74) suggests generalizing synchronization information from a complexity measure to a complexity function $\mathcal{S}(\omega)$, which, in this case, would be simply related to the Fourier-transform of the two-sided myopic state-uncertainty $\mathcal{H}(L)$.

3.12.4 Power Spectra

The extended complexity functions that we have introduced, $\mathcal{E}(\omega)$ and $\mathcal{S}(\omega)$, can give the same intuitive understanding for entropy reduction and synchronization respectively as the power spectrum $P(\omega)$ gives for pairwise correlation. Recall that the power spectrum can be written as:

$$P_c(\omega) = \langle |x|^2 \rangle + 2 \operatorname{Re} \langle \pi \bar{\mathcal{A}} | (e^{i\omega} I - T)^{-1} | \mathcal{A} \mathbf{1} \rangle .$$

We see that $(e^{i\omega} I - T)^{-1}$ is the resolvent of T evaluated along the unit circle $z = e^{i\omega}$ for $\omega \in [0, 2\pi)$. Hence by the decomposition of the resolvent, Eq. (3.24), the general spectral decomposition of the continuous part of the power spectrum will be:

$$P_c(\omega) = \langle |x|^2 \rangle + 2 \sum_{\lambda \in \Lambda_T} \sum_{m=0}^{\nu_\lambda - 1} \operatorname{Re} \frac{\langle \pi \bar{\mathcal{A}} | T_{\lambda, m} | \mathcal{A} \mathbf{1} \rangle}{(e^{i\omega} - \lambda)^{m+1}} .$$

As with $\mathcal{E}(\omega)$ and $\mathcal{S}(\omega)$, all continuous frequency dependence of the power spectrum again lies simply and entirely in the denominator of the above expression. Analogous to the discovery in Ref. [204], the delta functions of the power spectrum arise from the eigenvalues of T that lie on the unit circle:

$$P_d(\omega) = \sum_{k=-\infty}^{\infty} \sum_{\substack{\lambda \in \Lambda_T \\ |\lambda|=1}} 2\pi \delta(\omega - \omega_\lambda + 2\pi k) \operatorname{Re}(\lambda^{-1} \langle \pi \bar{\mathcal{A}} | T_\lambda | \mathcal{A} \mathbf{1} \rangle) ,$$

where ω_λ is related to λ by $\lambda = e^{i\omega_\lambda}$. An extension of the Perron–Frobenius theorem guarantees that the eigenvalues of T on the unit circle have index $\nu_\lambda = 1$.

Together, these equations yield structural constraints via constrained functional forms that can be useful for the ‘inverse problem’ of inferring the process from measured data.

3.12.5 Almost Diagonalizable Dynamics

The nondiagonalizable structures that will appear most commonly in the meta-dynamics induced by analysis will be of a special form we call ‘almost diagonalizable’: when all besides possibly one of the eigenspaces (usually the eigenspace associated with $\lambda = 0$) are diagonalizable subspaces.

In the current setting, we will say that a matrix is *almost diagonalizable* if all of its eigenvalues with magnitude greater than zero have geometric multiplicity equal to their algebraic multiplicity:

Definition 5. *W is almost diagonalizable if and only if $g_\lambda = a_\lambda$ for all $\lambda \in \Lambda_W^{\setminus 0} \equiv \Lambda_W \setminus \{0\}$.*

Fortunately, we can treat this type of non-diagonalizability relatively easily using the results of the spectral decomposition of W^L for singular matrices. First off, Eq. (3.27) simplifies to:

$$W^L = \sum_{\lambda \in \Lambda_W} \lambda^L W_\lambda + \sum_{m=1}^{\nu_0-1} \delta_{L,m} W_0 W^m . \quad (3.76)$$

To obtain the projection operators associated with each eigenvalue in $\Lambda_W^{\setminus 0}$ for an almost diagonalizable matrix W , we use Eq. (3.15) with $\nu_\lambda = 1$ for all $\lambda \in \Lambda_W^{\setminus 0}$. Finding:

$$W_\lambda = \left(\frac{W}{\lambda} \right)^{\nu_0} \prod_{\substack{\zeta \in \Lambda_W^{\setminus 0} \\ \zeta \neq \lambda}} \frac{W - \zeta I}{\lambda - \zeta} , \quad (3.77)$$

for each $\lambda \in \Lambda_W^{\setminus 0}$. Or, when more convenient in a calculation, we let $\nu_0 \rightarrow a_0 - g_0 + 1$ or even $\nu_0 \rightarrow a_0$ in Eq. (3.77), since multiplying W_λ by W/λ has no effect.

With the set of projection operators for all $\lambda \in \Lambda_W^{\setminus 0}$, we can use the fact that the projection operators sum to the identity (Eq. (3.14)) to determine the projection operator associated with the zero eigenvalue:

$$W_0 = I - \sum_{\lambda \in \Lambda_W^{\setminus 0}} W_\lambda .$$

Note that often this is sometimes simpler and easier to automate than evaluating W_0 via the methods of symbolic inversion and residues, or finding all left and right eigenvectors and generalized eigenvectors.

Almost diagonalizable metadynamics will play a prominent role for prediction of both processes of finite Markov order and the much more general class of processes with broken partial symmetries that can be detected within a finite epoch of observation—the processes of finite symmetry-collapse discussed next.

3.12.6 Processes of Finite Markov Order vs. Processes of Finite Symmetry Collapse

What if zero is the *only* eigenvalue in the transient structure of the MSP for a process? That is, what if there are no loops in the \mathcal{S} -MSP transient structure? These turn out to be finite Markov order processes.

For processes with finite Markov order R —such as, those whose support is a *subshift of finite type* [146]—the entropy-rate approximates converge for $\ell \geq R + 1$ [62]:

$$h_\mu(\ell) - h_\mu = 0 \quad \text{for } \ell \geq R + 1 , \quad (3.78)$$

or, equivalently, for $L \geq R$:

$$\langle \delta_\pi | W^L | H(W^A) \rangle - \langle \pi_W | H(W^A) \rangle = 0 \quad \text{for } L \geq R .$$

For a finite-order Markov process, all MSP transient states must have identically zero probability after R time-steps. The only way to achieve this is if the \mathcal{S} -MSP’s transient structure is an acyclic directed graph with all probability density flowing away from the unique start-state down to the recurrent structure. This means that all eigenvalues associated with the transients are zero. Moreover, the index of the zero-eigenvalue of the ϵ -machine’s \mathcal{S} -MSP is *equal* to the Markov order for finite Markov-order processes:

$$\nu_0(W) = R \quad \text{if } \Lambda_W \setminus \Lambda_T = \{0\} .$$

In contrast, for stochastic processes whose support is a *strictly sofic subshift* [146], the Markov order diverges, but ν_0 can vanish or be finite or infinite. Yet in either the finite-type or sofic case ν_0 still tracks the duration of exact state-space collapse within the transient dynamics of synchronization. This suggests that ν_0 captures the *index of broken symmetries* for strictly sofic processes, in analogy to the Markov order for subshifts of finite type; indeed the name “symmetry-collapse” captures the essence of ν_0 ’s role in both.

In the first ν_0 time-steps, symmetries are broken that synchronize an observer to the process. For the simple period-two process $\dots 010101010\dots$ the “symmetry” is the degeneracy of possible phases. After a single observation, the observer learns the phase and is completely synchronized to the process. Hence $\nu_0 = 1$ for this order-1 Markov

process. Simple periodic processes with larger periods have a longer time before the phase information is necessarily fully known; hence their larger Markov order.

For more sophisticated strictly sofic processes, there may also be some symmetries, such as phase information, that are completely broken within a finite amount of time, but this is only part of the larger transient metadynamics of synchronization. Hence the symmetries completely broken within the symmetry-collapse epoch can be in addition to lingering state uncertainties about a strictly sofic process. As a practical matter, the predictability of a process is often drastically enhanced through the finite epoch of symmetry-collapse. This should become apparent in the examples to follow.

3.13 Coronal Spectrograms

Coronal spectrograms, recently introduced in Ref. [204] (in that case, to visualize how diffraction patterns of chaotic crystals emanate from the spectrum of the hidden dynamic of the stacking process), are a broadly useful tool in visualizing complexity spectra, from power spectra to excess entropy spectra.

Coronal spectrograms display any frequency-dependent measure $f(\omega)$ of a process wrapped around the unit circle while showing the eigenvalues Λ_A of the relevant linear dynamic within the unit circle of the complex plane. An example is given in Fig. 3.4a. This setup works for discrete-domain (e.g., discrete-time or discrete-space) dynamics. For continuous-time dynamics, the coronal spectrogram unwraps—via the typical discrete-to-continuous conformal mapping of the inside of the unit circle of the complex plane to the left half of the complex plane [140]—into what we call the *coronated horizon*. We show a discrete-time version of the coronated horizon in Fig. 3.4b. Ultimately, either the coronal spectrogram or coronated horizon yields the same information and same important lesson: the eigenvalues of the hidden linear dynamic control the allowed behavior of the system.

Coronal spectrograms show undeniably that complex behavior under study behave according to the spectrum of the hidden linear dynamic. In each area of application, the relevant frequency-dependent measure $f(\omega)$ emanates from the non-zero eigenvalues of the hidden linear dynamic—the closer to the unit circle, the sharper the possible

peaks in the observation; yielding Bragg-like reflections (delta-function contributions) when the eigenvalues are actually on the unit circle. The collection of diffuse peaks is a sum of Lorentzian-like and what we might call ‘super-Lorentzian’-like line profiles. Indeed, the Lorentzian-like line profiles are the discrete-time version of a Lorentzian curve: whereas the continuous-domain Lorentzian is given by $\text{Re}(\frac{c}{\omega-\lambda})$, the continuous-to-discrete conformal mapping $\omega \rightarrow e^{i\omega}$ directly yields our discrete-domain analogs: $\text{Re}(\frac{c}{e^{i\omega}-\lambda})$. The ‘super-Lorentzian’-like line profiles have the form: $\text{Re}[(\frac{c}{e^{i\omega}-\lambda})^n]$.

Zero eigenvalues can also contribute to frequency dependence of $f(\omega)$, but only contribute sinusoidal contributions of discrete increments from $\cos(\omega)$ up to $\cos(\nu_0 \omega)$. Since such contributions are qualitatively distinct from the super-Lorentzian contributions and do not emanate radially from the eigenvalues the same way that non-zero eigenvalues emanate, we find that the coronal spectrograms are most useful for understanding the contributions of non-zero eigenvalues to the observation. Nevertheless, the two contributions can be usefully disentangled, as shown later.

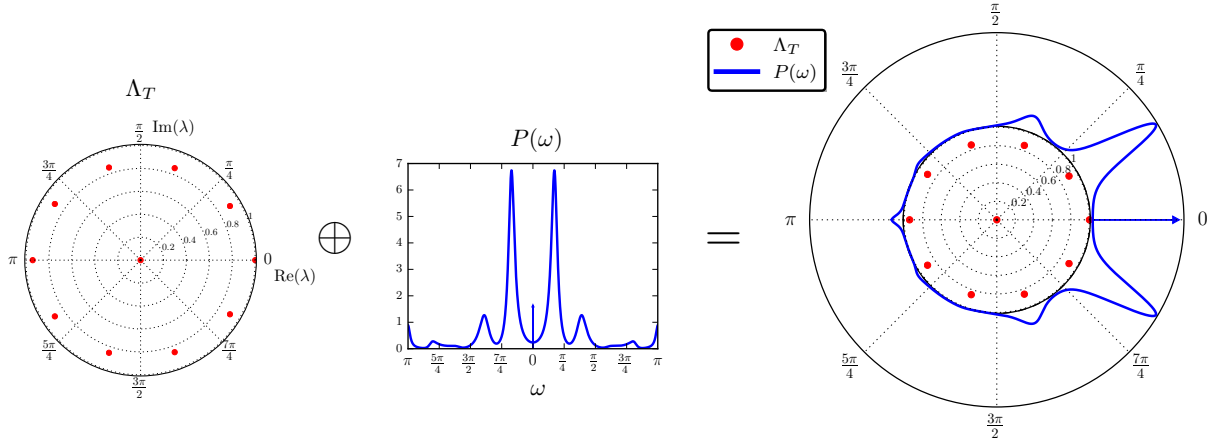
We will use both coronal spectrograms and coronated horizons to visualize various results in the following examples.

3.14 Examples

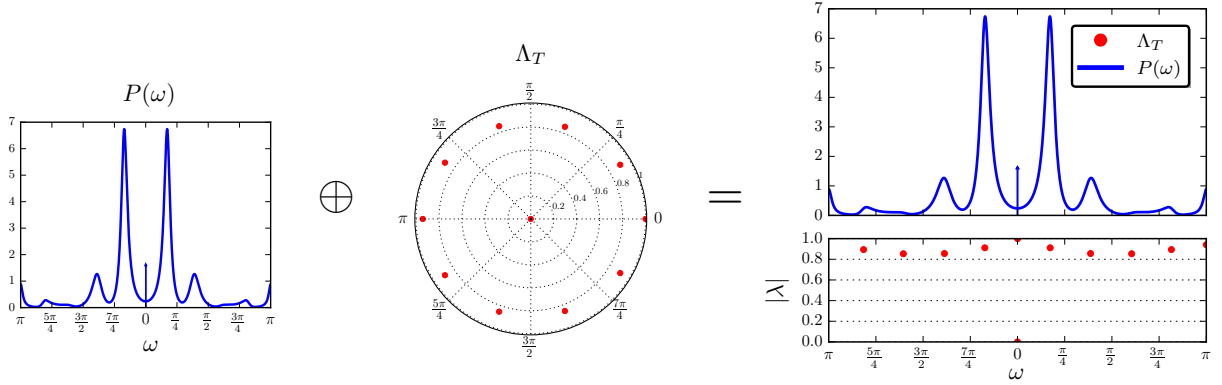
3.14.1 $(R-k)$ -Golden Mean Process Family

We will use the $(R-k)$ -Golden Mean family of processes to demonstrate the features of finite Markov order in relation to measures of complexity ³. This process family describes a unique transition-parametrized process for each Markov-order $R \in \{\nu \in \mathbb{Z} : \nu \geq 1\}$ and each cryptic-order $k \in \{\kappa \in \mathbb{Z} : 1 \leq \kappa \leq R\}$. The ϵ -machine for the (4-3)-Golden Mean Process is shown in Fig. 3.5a, from which the construction of all other $(R-k)$ -Golden Mean processes should be clear. In summary: the alphabet for any $(R-k)$ -Golden Mean process

³As a historical aside: the name ‘Golden Mean’ comes from a well-known source of the (1-1)-Golden Mean language which forbids two consecutive 1s: the symbolic dynamics of the chaotic one-dimensional shift map over the unit interval with generating partition, when the slope of the map is the golden mean $\phi = 1 + \frac{1}{\phi} = \frac{1}{2} + \frac{\sqrt{5}}{2}$ [277]. Moreover, consistency with the stationary stochastic process so generated would imply $1 - p = 1/\phi$. Generalizing to arbitrary p , and especially to arbitrary R and k takes us afar from this historical setting, into a sufficiently general setting to study the generic behavior of order- R Markov processes and the complexities of predicting them.

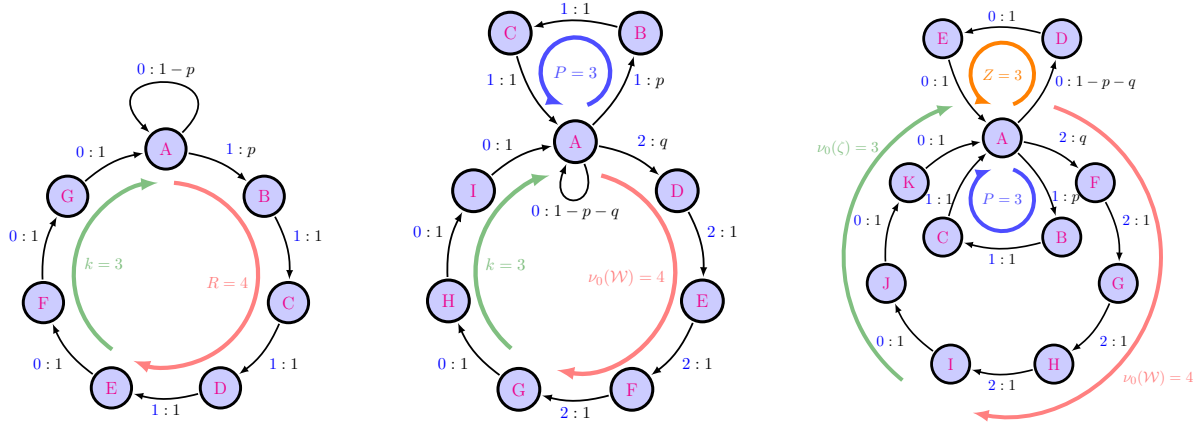


(a) A *coronal spectrogram* combines the eigenvalues Λ_A of the hidden linear dynamic together with a frequency-dependent function $f(\omega)$ of the process by wrapping $f(\omega)$ around the unit circle. It then becomes evident that $f(\omega)$ emanates from the eigenvalues.



(b) A *coronated horizon* combines the frequency-dependent function $f(\omega)$ of the process together with the eigenvalues Λ_A of the hidden linear dynamic by unwrapping the unit circle. Again, it is evident that $f(\omega)$ emanates from the eigenvalues.

Figure 3.4: Pictorial introduction to the (a) coronal spectrogram and (b) coronated horizon.



- (a) The $(4-3)$ -GM Process of the $(R-k)$ -Golden Mean family with $0 \leq k = \nu_0(\zeta) \leq R = \nu_0(\mathcal{W}) < \infty$. Members of this family have finite but tunable Markov-order R and cryptic-order k .
- (b) The $(4-3)$ -GP-(3) Process of the $(\nu_0(\mathcal{W})-k)$ -Golden Parity- (P) family, with $0 \leq k = \nu_0(\zeta) \leq \nu_0(\mathcal{W}) < R = \infty$ whenever $P > 1$. Members of this family have infinite Markov-order R , tunable finite cryptic-order k , and tunable finite symmetry-collapse index $\nu_0(\mathcal{W})$.
- (c) The $(4-3)$ -GPZ-(3-3) Process of the $(\nu_0(\mathcal{W})-\nu_0(\zeta))$ -Golden Parity- $(P-Z)$ family with $0 \leq \nu_0(\zeta) \leq \nu_0(\mathcal{W}) < k = R = \infty$ whenever $Z > 1$. Markov-order is infinite whenever either P or Z is larger than one. Cryptic-order is infinite when $Z > 1$. Members of this family have finite but tunable symmetry-collapse index $\nu_0(\mathcal{W})$ and cryptic index $\nu_0(\zeta)$.

Figure 3.5: Process families useful for understanding the roles and interplay among Markov-order R , cryptic-order k , the ‘symmetry-collapse index’ $\nu_0(\mathcal{W})$ of the zero eigenvalue of the synchronizing dynamic over mixed states, and the ‘cryptic index’ $\nu_0(\zeta)$ of the zero eigenvalue of the cryptic operator presentation. We always have $k \leq R$ and $\nu_0(\zeta) \leq \nu_0(\mathcal{W})$. Whenever $\Lambda_{\mathcal{W}} = \Lambda_T \cup \{0\}$, R is finite, $R = \nu_0(\mathcal{W})$ and $k = \nu_0(\zeta)$. Whenever $\Lambda_{\zeta} = \Lambda_T \cup \{0\}$, k is finite whether or not R is, and $k = \nu_0(\zeta)$. When k or R is infinite, the *cryptic index* and *symmetry-collapse index* pick up more nuanced features of the cryptic dynamics and synchronization dynamics.

is $\mathcal{A} = \{0, 1\}$; if the most recent history consists of at least k consecutive 0s (and no 1s since then) then there is a probability p of next observing a 1—which will necessarily entail R consecutive 1s followed by at least k consecutive 0s—and a probability $1 - p$ of simply seeing another 0.

The eigenvalues of the net state-to-state transition matrix of the recurrent component of the ϵ -machine are:

$$\Lambda_T = \left\{ \lambda \in \mathbb{C} : (\lambda - (1 - p))\lambda^{R+k-1} = p \right\}.$$

In the limit of $p \rightarrow 1$, each $(R-k)$ -Golden Mean becomes a perfectly periodic process. In this limit, the eigenvalues become evenly distributed on the unit circle: $\Lambda_T \rightarrow \{e^{in2\pi/(R+k)}\}_{n=0}^{R+k-1}$ as $p \rightarrow 1$. At the other extreme, all eigenvalues, besides the stationary eigenvalue at $z = 1$, evolve towards zero as $p \rightarrow 0$. At any setting of the transition parameter p , the non-unity eigenvalues lie approximately on a circle within the complex plane whose radius decreases nonlinearly from 1 to 0 as p is swept from 1 to 0; meanwhile, the center of this circle moves from the origin to some positive real value and back to the origin as p is swept from 1 to 0. In Fig. 3.6, we show how the eigenvalues of the (5-3)-Golden Mean Process evolve over the full range of p as it is swept from 1 to 0.

In contrast to the p -dependent spectrum of the recurrent structure just discussed, the only eigenvalue corresponding to the transient structure of the \mathcal{S} -MSP is equal to zero, regardless of the transition parameter p . Recall that this is necessarily true for any process with finite Markov order. Hence, $\Lambda_{\mathcal{W}} = \Lambda_T \cup \{0\}$, with $\nu_0(\mathcal{W}) = R$. The cryptic structure is similar: $\Lambda_{\zeta} = \Lambda_T \cup \{0\}$, with $\nu_0(\zeta) = k$, where ζ is the state-to-state transition matrix of the cryptic operator presentation.

In Table 3.3, the ϵ -machines, autocorrelation, power spectra, MSPs, myopic entropy rates, and myopic state uncertainties are compared for three p -parametrized examples of $(R-k)$ -GM processes.

The autocorrelation of each process reflects the ‘leaky periodic’ behavior, whose leakiness originates from the self-transition at state A which adds a phase-slip noise to the otherwise $(R + k)$ -periodic behavior.

For each process, the phase of the process—and hence the latent state of the ϵ -machine—

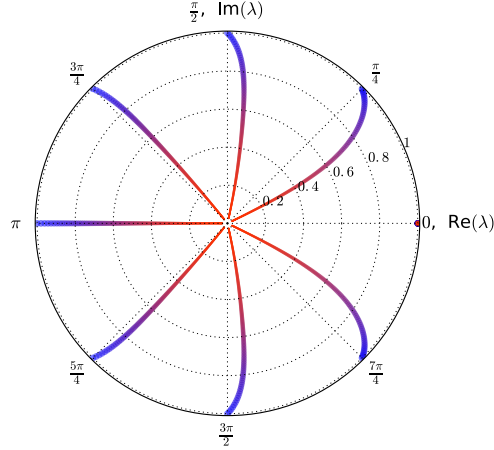


Figure 3.6: The evolution of the eigenvalues Λ_T of the recurrent part of the ϵ -machine for the (5–3)-GM Process is shown within the unit circle of the complex plane; the trajectory of each eigenvalue changes from thick blue to thin red as the transition parameter p evolves from 1 to 0. In addition to the seven eigenvalues that move from the nontrivial eighth roots of unity towards zero along nonlinear trajectories, notice the eigenvalue at $z = 1$ that does not evolve with p .

is fully resolved after R time-steps of observation. This corresponds to the depth of the S-MSP tree-like structure $\nu_0(\mathcal{W}) = R$, the complete reduction of the myopic entropy rate $h_\mu(L)$ to the true entropy rate h_μ when conditioning on observations of finite block-length $L - 1 = R$, and the complete extinction of causal state uncertainty $\mathcal{H}(L)$ after $L = R$ observations.

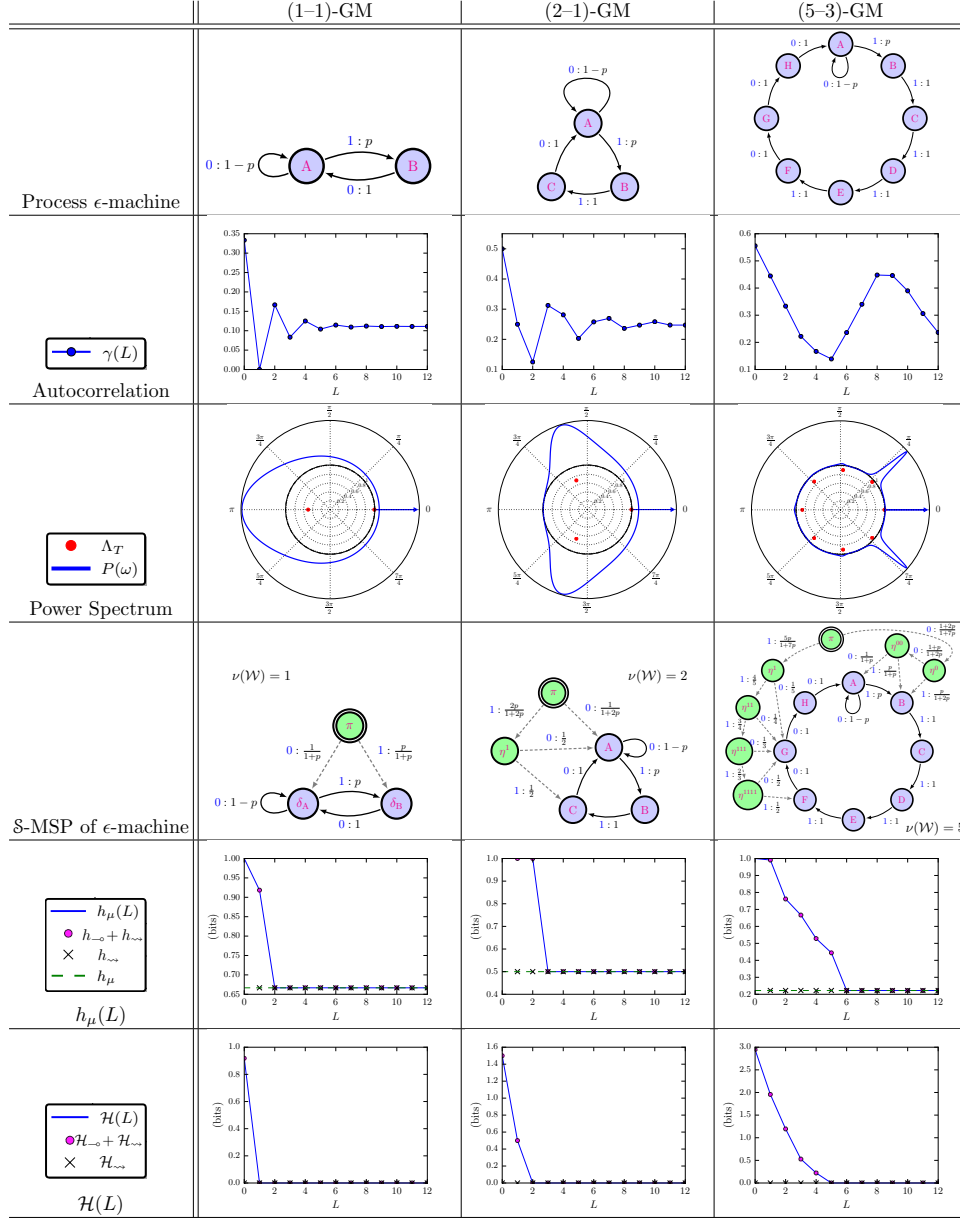


Table 3.3: Select complexity analysis for processes of finite Markov order. Quantitative data corresponds to $p = 1/2$.

As a paradigmatic example of finite Markov order, the (5-3)-Golden Mean Process has strictly tree-like structure in the transient part of its MSP, with maximum depth equal to both $\nu(W)$ and the Markov order, 5.

We have now seen the typical behavior of complexity measures for processes of finite Markov-order. We next investigate a simple process of infinite Markov-order to draw

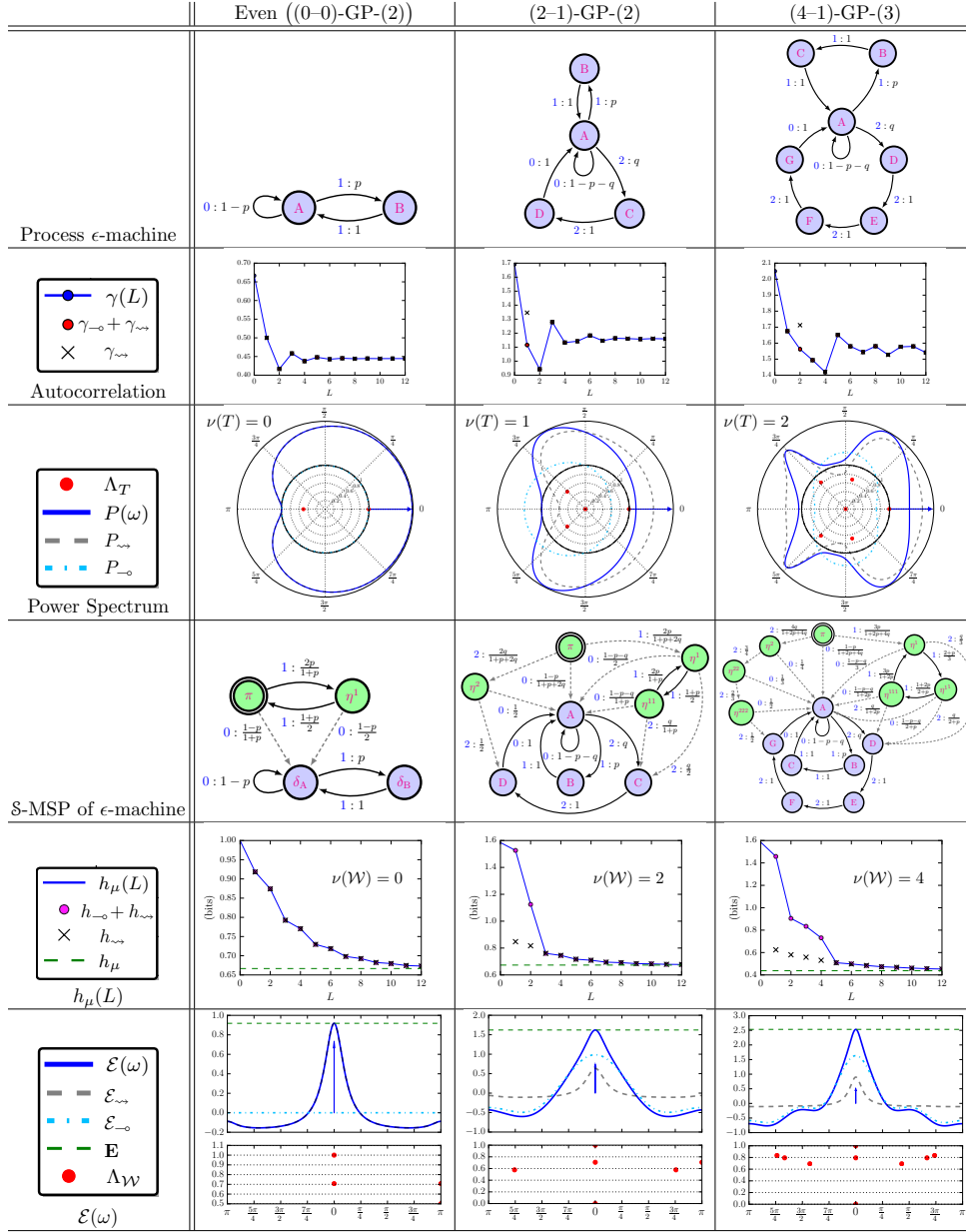


Table 3.4: Select complexity analysis for processes of infinite Markov order. Quantitative data corresponds to $p = 1/2$ and $q = 1/3$.

attention to the characteristic difference of non-zero eigenvalues in the transient structure of the MSP of infinite Markov-order processes.

3.14.2 Even Process

The Even Process, shown in the first column of Table 3.4, is a well known example of a stochastic process that cannot be fully described by any finite Markov-order approximation, yet it accommodates an apparently simple two-state hidden-Markov model description.

Infinite Markov order, in this case, stems from the fact that only even numbers of consecutive 1s are ever produced by the process. The resources necessary to track this parity induce the infinite Markov order. Moreover, the surplus entropy rate $h_\mu(L) - h_\mu$ that would be incurred upon using a finite order- $(L - 1)$ Markov approximation is seen in Table 3.4 to decay to zero only asymptotically, as the sum of decaying exponentials. This is in clear contrast to the behavior of the myopic entropy rate for the finite Markov order processes of Table 3.3, for which $h_\mu(L)$ drops to h_μ *exactly* at $L = R + 1$. Similarly, the average state uncertainty $\mathcal{H}(L)$ for infinite Markov processes converges only asymptotically—and with the same set of decay rates as $h_\mu(L)$ —to its asymptotic value of 0, although this curve is not shown in Table 3.4 for lack of space. Such everlasting decay is due to non-zero eigenvalues in the transient structure of the \mathcal{S} -MSP.

We have just considered a simple example of an infinite Markov-order process. As one generically expects for processes of infinite Markov order, the transient structure of the MSP had non-zero eigenvalues corresponding to it. In general, though, we should expect *both* types of contributions within the architecture of synchronization (as encapsulated by the transient structure of the MSP): a finite-horizon contribution to the past–future mutual information corresponding to completely ephemeral zero eigenvalue contributions to the transient structure of the MSP, and an infinite-horizon contribution to the past–future mutual information arising from non-zero eigen-contributions.

3.14.3 The $(\nu_0(\mathcal{W})-k)$ -Golden—Parity- (P) Process Family

To further explore the nature of infinite Markov order processes, we introduce the family of (ν_0-k) -Golden—Parity- (P) processes, which subsumes and extends all of the example

processes discussed so far. The role of each parameter is explained by Fig. 3.5b, which shows an automata representation of the ϵ -machine for the (4-3)-GP-(3) process.

In the case where $P = 1$, the family reduces to the (ν_0-k) -Golden Mean process family, with tunable Markov $R = \nu_0(W)$ and cryptic k orders. I.e., $(\nu_0(\mathcal{W})-k)$ -GP-(1) = $(\nu_0(\mathcal{W})-k)$ -GM. However, the Markov order becomes infinite whenever $P > 1$, in which case the index of the zero-eigenvalue of the \mathcal{S} -MSP $\nu_0(\mathcal{W})$ (which controls the finite duration necessary to resolve all broken symmetries) and the cryptic order k can still be tuned independently. The Even Process considered earlier is the (0-0)-Golden—Parity-(2) Process.

Three examples of $(\nu_0(\mathcal{W})-k)$ -Golden—Parity- (P) processes are explored in Table 3.4. The transient structure of the \mathcal{S} -MSP for the second two examples should clarify the difference between (1) the symmetry collapse associated with completely ephemeral transient states that are fully depleted of probability density after $\nu_0(\mathcal{W})$ time-steps and (2) the everlasting leaky transients whose probability density only decays asymptotically to zero as more refined ambiguity is resolved. Looking at the myopic entropy convergence $h_\mu(L)$, the affect of these distinct routes to synchronization on the predictability is clear. The process is much more predictable, on average, after $\nu_0(\mathcal{W})$ time-steps. However, the average predictability of an infinite-Markov-order process continues to increase with increasing observation duration, albeit with exponentially diminishing returns. In general, we have shown that this asymptotic convergence occurs as a sum of decaying exponentials from diagonalizable subspaces, and polynomials-times-exponentials in the case of non-diagonalizable structures associated with non-zero eigenvalues. The apparent oscillations under the exponential decays are completely described by the leaky periodicities of the eigenvalues in the transient belief states.

Finally, we note that the excess entropy spectrum shows the frequency domain picture of observation-induced predictability. $\lim_{\omega \rightarrow 0} \mathcal{E}(\omega) = \mathbf{E}$ is the total past–future mutual information, which is also the excess entropy observed before full synchronization. The $\nu_0(\mathcal{W})$ symmetry collapse contributes significantly and quickly to the total excess entropy of the last two examples, whereas the asymptotic tails of synchronization associated with

leaky periodicity of certain transient states of uncertainty accumulate their contribution to excess entropy rather slowly.

The general and broadly applicable theoretical results derived in this chapter—which yield new aid to the intuition and new answers to real-world challenges—allow novel numerical investigations and unprecedented analytical calculations for infinite-Markov-order processes. As an example of the latter, let us quote several of the exact results that can be derived—as shown in § 3.16—for the (p, q) -parametrized (2–1)-GP-(2) process that was explored in the second column of Table 3.4.

In § 3.16, we find that qualitatively distinct behavior dominates the (2–1)-GP-(2) process depending on whether the transition parameter p is larger or smaller than $2\sqrt{q} - q$. As an instantiation of a general emerging principle, we find that the behaviorally distinct regions are separated by a critical line in the (p, q) -parameter space along which the transition dynamic T becomes nondiagonalizable. For $p < 2\sqrt{q} - q$, the autocorrelation for the (2–1)-GP-(2) process for $|L| \geq 2$ has the exact solution:

$$\gamma(L) = \beta^2 + \beta q^{|L|/2} \operatorname{Re}(\zeta e^{i\omega_\xi |L|}) , \quad (3.79)$$

where $\beta \equiv \frac{2(p+2q)}{1+p+2q}$, $\zeta \equiv \frac{(\xi+1)^2(p\xi+2q)}{\xi(\xi^3+p\xi+2q)}$, $\xi \equiv -\frac{1}{2}(p+q) + i\frac{1}{2}\sqrt{4q-(p+q)^2}$, and $\omega_\xi \equiv \frac{\pi}{2} + \arctan\left(\frac{p+q}{\sqrt{4q-(p+q)^2}}\right)$. The corresponding power spectrum is:

$$\begin{aligned} P(\omega) = & \frac{8q}{1+p+2q} + \frac{2p}{1+p+2q} [1 - \cos(\omega)] \\ & + \beta \operatorname{Re}\left(\frac{\zeta\xi}{e^{i\omega} - \xi} + \frac{\zeta\xi}{e^{-i\omega} - \xi}\right) + 2\pi\beta^2 \sum_{k=-\infty}^{\infty} \delta(\omega + 2\pi k) . \end{aligned} \quad (3.80)$$

For any parameter setting, the metadynamic of observation-induced synchronization to the (2–1)-GP-(2) process is nondiagonalizable due to the index-2 eigenvalue of 0, which leads to a completely ephemeral contribution to $h_\mu(L)$ up to $L = 2$. For $L \geq 3$, we find the myopic entropy rate to asymptotically relax to the true entropy rate according to:

$$h_\mu(L) - h_\mu = \begin{cases} \frac{-p \log p + (1+p) \log(1+p) - 2p}{\sqrt{p}(1+p+2q)} p^{L/2} & \text{for odd } L \\ \frac{p \log p - (1+p) \log(1+p) + 2}{(1+p+2q)} p^{L/2} & \text{for even } L \end{cases} \quad (3.81)$$

where the true entropy rate of the process is:

$$h_\mu = \frac{-q \log q - p \log p - (1-p-q) \log(1-p-q)}{1+p+2q} . \quad (3.82)$$

Interestingly, while the autocorrelation at separation L scales as $\sim q^{L/2}$, the predictability of transitions between shifting histories of length L converges as $\sim p^{L/2}$ —two rather independent length scales. The amount of the future that can be predicted from the past is the total mutual information between the observable past and observable future:

$$\mathbf{E} = \frac{(1-p-q) \log(1-p-q) - p \log p - q \log q - (1-p) \log(1-p)}{1+p+2q} + \log(1+p+2q) . \quad (3.83)$$

However, to actually *do* this prediction requires more memory than this amount of shared information. More such quantities and many more details can be found in § 3.16.

For further exploration, we note that if one's interest lies in the investigation of structure within processes of infinite-cryptic order, a more generalized family of processes, the $(\nu_0(\mathcal{W})-\nu_0(\zeta))$ -GP- $(P-Z)$ Process Family should be studied, where $\nu_0(\zeta)$ is the index of the zero-eigenvalue of the cryptic operator presentation and the process has infinite cryptic order whenever $Z > 1$. In the preceding, we always had $Z = 1$, and $(\nu_0(\mathcal{W})-\nu_0(\zeta))$ -GP- $(P-1) = (\nu_0(\mathcal{W})-\nu_0(\zeta))$ -GP- (P) .

3.15 The Importance of Predicting Beyond-Pairwise Structure: The RRX Process as a Paradigm for Processes that Compute

The Random–Random–XOR Process (RRX Process) is a simple but non-trivial example of information processing that hints at the general features of stochastic dynamic information processing systems. The process can be interpreted as a repetitive implementation that folds out over three time-steps: a **0** or **1** is output with equal probability, another **0** or **1** is output with equal probability, and then the eXclusive-OR operation (XOR) of the last two outputs is output. The ϵ -machine for the RRX Process is shown in Fig. 3.7.

Surprisingly, there is no pairwise correlations in this process—all correlations are higher than second order. One consequence is that the power spectrum for the process is

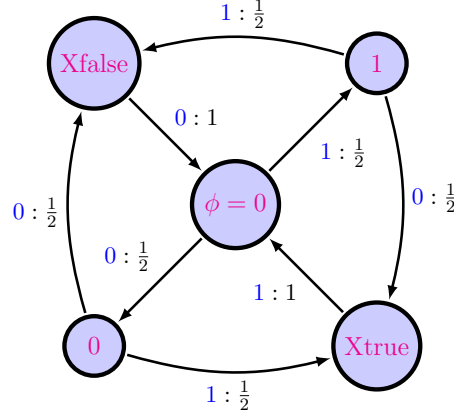


Figure 3.7: ϵ -machine of RRX process.

completely flat—yielding the apparent signature of white noise. This is shown in Fig. 3.8, and would mislead the casual observer to conclude that the observed time series has no structure when, in fact, it is only an indication that any structure must be hidden in higher-order correlations.

Nevertheless, there is clearly information being processed in a consequential way. Fortunately, the complexity measures we have discussed in this chapter will be able to resolve this higher-order structure. The structure of hidden dependencies is detected by the excess entropy spectrum, as shown in Fig. 3.12.

Let us briefly consider *why* the excess entropy spectrum was successful at detecting the nuanced structure of the RRX process. It is sometimes stated that information measures are superior to standard measures of correlation because they are able to capture nonlinear dependencies. This has been suggested as a reason to consider the *information correlation* $I(X_0; X_\tau)$ rather than autocorrelation [224]. It is natural to also consider the corresponding generalization of the power spectrum; let us define the *power-of-pairwise information spectrum* (POP information spectrum):

$$\mathcal{I}(\omega) \equiv -H(X_0) + \lim_{N \rightarrow \infty} \sum_{\tau=-N}^N e^{-i\omega\tau} I(X_0; X_\tau) . \quad (3.84)$$

However: again, the information correlation and POP information spectrum only capture *pairwise* correlation. It is easy to compute that $\mathcal{I}(\omega) = 0$ for the RRX process. Hence, as shown in Fig. 3.8, these measures are not sufficient to detect even simple computational

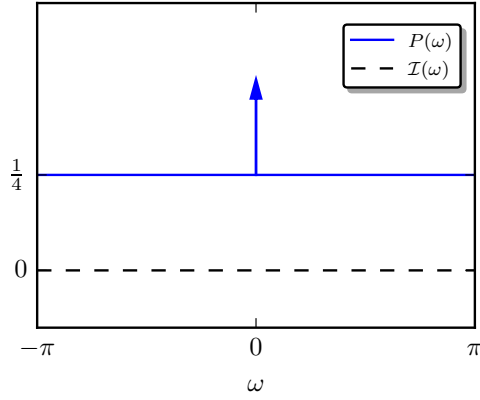


Figure 3.8: The RRX process produces a flat power spectrum, and even a null POP information spectrum, which could lead one to wrongly assume that the process is structureless white noise if only pairwise correlations are considered.

structure.

The above detour through the information correlation and POP information spectrum brings us to a deeper understanding of why the excess entropy spectrum is successful at detecting nuanced computational structure in a time series: Because the excess entropy partitions all observable random variables throughout time, the excess entropy picks up any systematic way that the past can influence the future; moreover the excess entropy *spectrum* elucidates the frequency decomposition of any such linear or nonlinear correlations. All multivariate interdependencies contribute to the excess entropy spectrum.

Let us now consider the hidden structure of the RRX process in more detail. With reference to Fig. 3.7, we observe that the expected probability density over causal states cascades through the ϵ -machine in a period-3 way, but the particular symbols drawn after each phase resetting (to $\phi = 0$) in a realization break the symmetries regarding which wings of the structure are traversed. Correspondingly, the eigenvalues of T are the three roots of unity $\{e^{in2\pi/3}\}_{n=0}^2$ and two zero eigenvalues (with $a_0(T) = g_0(T) = 2 \rightarrow \nu_0(T) = 1$).

The period-three-like structure implies a phase ambiguity in synchronizing to the process that must be resolved in the transient structure of the MSP. The \mathcal{S} -MSP of the RRX Process is shown in Fig. 3.9. There are 31 transient states of uncertainty, in addition to the five recurrent states—36 states in total. Since we have derived the \mathcal{S} -MSP of the ϵ -machine in particular, $W = \mathcal{W}$. Hence, the layout of the MSP intuitively shows the

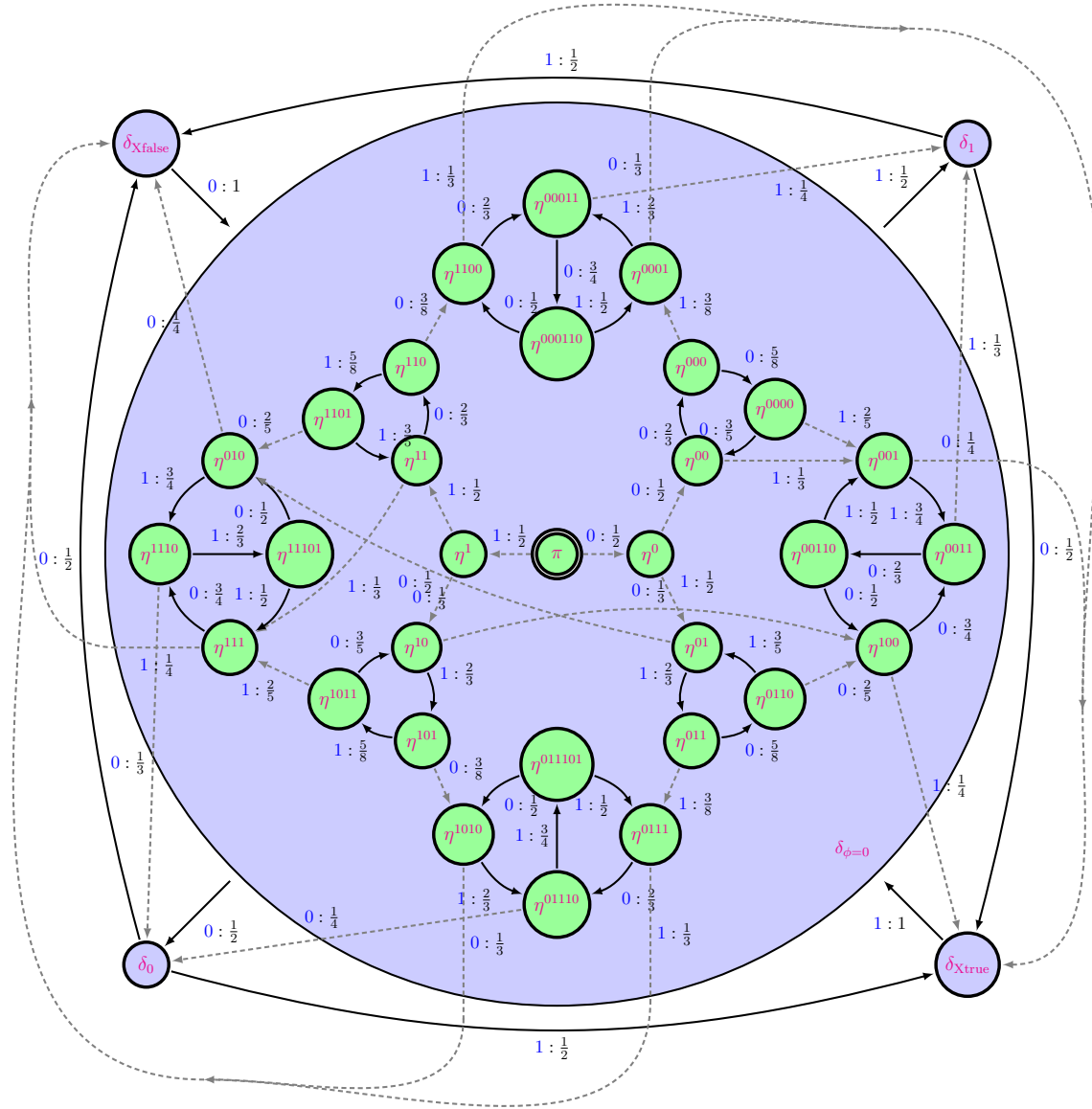


Figure 3.9: MSP of ϵ -machine of RRX process. Grayed out (and dashed) edges permanently leave the states they came from; recognizing the way these edges partition the mixed-state space allows simplified calculation of the spectrum. The graph structure is inherently non-planar: the large blue recurrent state should be visualized as being *behind* the transient states; it does not contain them.

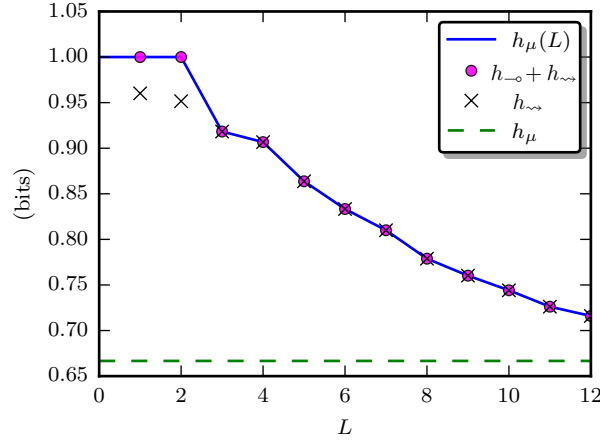


Figure 3.10: Ephemeral and persistent contributions to the myopic entropy rate. The ephemeral contribution lasts only up to $L = \nu_0(W) = 2$.

information processing involved with synchronizing to the process—the burden of an optimal predictor who will asymptotically only need to learn an average of h_μ bits per observation to fill in their knowledge of every particular about a realization.

The MSP introduces new zero eigenvalues in the transients. In particular, the first-encountered tree-like transients (starting with mixed-state π) introduce new Jordan blocks up to dimension 2. Overall, the 0-eigenspace of W has index 2, so that $\nu_0(W) = 2$.

Two different sets of leaky-period-3 structures appear in the MSP transients. There are four leaky three-state cycles, each with the same leaky-period-3 contributions to the spectrum: $\{(\frac{1}{4})^{1/3}e^{in2\pi/3}\}_{n=0}^2$. There are also four leaky four-state cycles, each with a leaky-period-3 contribution *and* symmetry-breaking 0-eigenvalue contribution to the spectrum: $\{(\frac{1}{2})^{1/3}e^{in2\pi/3}\}_{n=0}^2 \cup \{0\}$. The difference in eigenvalue magnitude $((\frac{1}{4})^{1/3}$ vs. $(\frac{1}{2})^{1/3}$) implies different timescales of synchronization associated with distinct learning tasks. An immediate lesson is that it takes longer (on average) to escape the 4-state leaky-period-3 components (from the time of arrival) than to escape the preceding 3-state leaky-period-3 components of the synchronizing metadynamic.

Figs. 3.10 and 3.11 reveal the sophisticated predictable structure that could not have been surmised from the bleakness of Fig. 3.8. Fig. 3.11 emphasizes the dominance of the slowest-decaying eigen-modes for large L . Such oscillations under the exponential

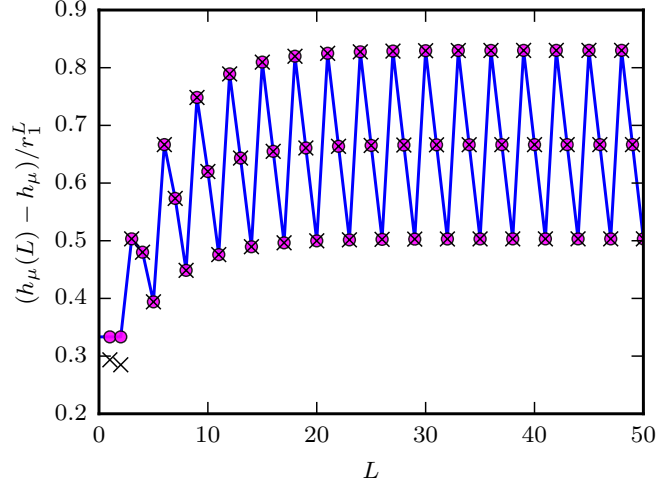


Figure 3.11: The tails of the myopic entropy convergence shown in Fig. 3.10 decay according to two different leaky period-three envelopes, corresponding to the two qualitatively different types of transient synchronization cycles in the MSP of Fig. 3.9. One of the transient cycles has a relatively fast decay rate of $r_2 = (1/4)^{1/3}$, while the slower decay rate of $r_1 = (1/2)^{1/3}$ dominates $h_\mu(L)$'s deviation from h_μ at large L .

convergence to synchronization are typical among processes. However, as seen in comparison with Fig. 3.10 much of the uncertainty may be reduced before this asymptotic mode becomes dominant. Ultimately, synchronization to prediction may involve important contributions from all modes of the mixed-state-to-state metadynamic.

As information processing becomes more complex, we anticipate that a vanishing amount of the structure in the output will be observable at low-orders of correlation. Indeed, logical computation, as implemented by universal logic gates, primarily operates above pairwise correlation. Hence, there is great motivation to move beyond measures of pairwise correlation to recognize hidden structure and to utilize higher-order structural investigations to better understand the information processing critical to the functionality of biological and engineered processes.

3.16 Example analytical calculation

Here we explicitly demonstrate the analytic calculations necessary to obtain the closed form complexity measures for the (p, q) -parametrized (2-1)-GP-(2) process, which we have

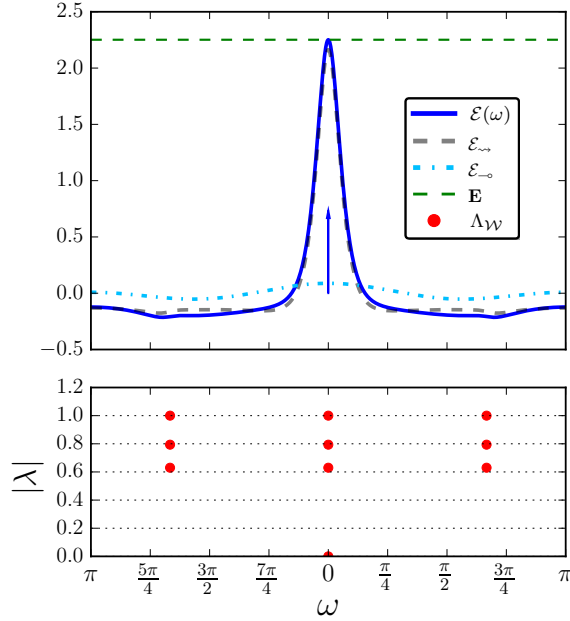
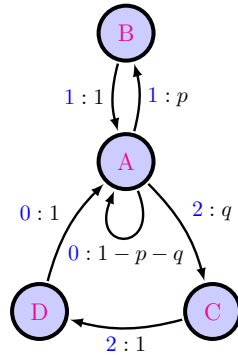


Figure 3.12: Excess entropy spectrum for RRX, together with the eigenvalues of the \mathcal{S} -MSP transition matrix \mathcal{W} . Among the power spectrum, POP information spectrum, and excess entropy spectrum, only the excess entropy spectrum is able to detect structure in the RRX process since the structure is beyond pairwise. The eigen-spectrum of the MSP of the RRX ϵ -machine and the excess entropy spectrum both indicate that the RRX process is indeed structured, with both ephemeral symmetry breaking and leaky periodicities in the convergence to optimal predictability.

already explored visually in the second column of Table 3.4. This appendix is primarily pedagogical, to promote insight and to better render particular calculational steps, and so will explicitly demonstrate a wide variety of techniques in the spirit of a tutorial.



The (p, q) -parametrized (2-1)-GP-(2) process is fully described by its ϵ -machine, whose automaton was shown in the first row and second column of Table 3.4, and is reproduced

above. More explicitly, the (2-1)-GP-(2) stationary stochastic process is generated by the HMM: $\mathcal{M}_{\text{eM}} = (\mathcal{S}, \mathcal{A}, \{T^{(x)}\}_{x \in \mathcal{A}}, \eta_0 = \pi)$, where $\mathcal{S} = \{\text{A}, \text{B}, \text{C}, \text{D}\}$ is the hidden set of causal states, $\mathcal{A} = \{0, 1, 2\}$ is the observable alphabet, $\{T^{(x)} : T_{s,s'}^{(x)} = \Pr(X_t = x, \mathcal{S}_{t+1} = s' | \mathcal{S}_t = s)\}_{x \in \mathcal{A}}$ is the set of symbol-labeled transition matrices consisting of:

$$T^{(0)} = \begin{bmatrix} 1-p-q & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}, \quad T^{(1)} = \begin{bmatrix} 0 & p & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \text{and} \quad T^{(2)} = \begin{bmatrix} 0 & 0 & q & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Together, the symbol-labeled transition matrices sum to the row-stochastic net state-to-state transition matrix:

$$T = \sum_{x \in \mathcal{A}} T^{(x)} = \begin{bmatrix} 1-p-q & p & q & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix},$$

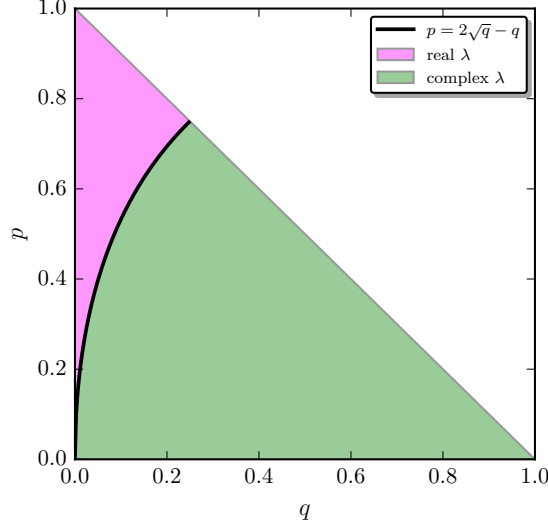
from which we find the stationary distribution over states $\langle \pi | = \langle \pi | T$ to be:

$$\langle \pi | = \frac{1}{1+p+2q} \begin{bmatrix} 1 & p & q & q \end{bmatrix}.$$

The first task of the analysis is to find the eigenvalues and associated projection operators for the net state-to-state transition matrix T . From $\det(T - \lambda I) = \lambda(\lambda - 1)(\lambda^2 + (p+q)\lambda + q) = 0$, we find the four eigenvalues to be:

$$\Lambda_T = \left\{ 1, 0, -\frac{1}{2}(p+q) \pm \frac{1}{2}\sqrt{(p+q)^2 - 4q} \right\}.$$

All eigenvalues are real for $p \geq 2\sqrt{q} - q$; two of the eigenvalues are complex with non-zero imaginary part when $p < 2\sqrt{q} - q$. Putting this together with the transition-probability consistency constraint that $p + q < 1$ yields the following mapping of the transition-parameter space:



For generic choice of the parameter setting (p, q) , all eigenvalues of T are unique, and so T is diagonalizable. However, two of the eigenvalues become degenerate along the parametric subspace $p = 2\sqrt{q} - q$, in which case the eigenvalues become $\Lambda_T = \{1, 0, -\sqrt{q}\}$. We find that the algebraic multiplicity $a_{-\sqrt{q}} = 2$ is larger than the geometric multiplicity $g_{-\sqrt{q}} = 1$, yielding nondiagonalizability ($\nu_{-\sqrt{q}} = 2$) along this $(p = 2\sqrt{q} - q)$ -submanifold. Indeed, we have found generically that eigenvalues become nondiagonalizable when they collide and scatter in the complex plane; e.g., when a pair of eigenvalues first ‘entangle’ to become complex conjugate pairs.

For any parameter setting (p, q) , we find the right eigenvectors of T from $(T - \lambda I) |\lambda\rangle = \vec{0}$ to be:

$$|0\rangle = \begin{bmatrix} 0 & 1 & -p/q & 0 \end{bmatrix}^\top$$

and

$$|\lambda\rangle = \begin{bmatrix} \lambda^2 & \lambda & 1 & \lambda \end{bmatrix}^\top \quad \text{for all } \lambda \in \Lambda_T \setminus \{0\} .$$

We find the left eigenvectors of T from $\langle\lambda| (T - \lambda I) = \vec{0}$ to be:

$$\langle 0| = \begin{bmatrix} 0 & 1 & 0 & -1 \end{bmatrix}$$

and

$$\langle \lambda | = \begin{bmatrix} \lambda^2 & p\lambda & q\lambda & q \end{bmatrix} \quad \text{for all } \lambda \in \Lambda_T \setminus \{0\} .$$

Clearly, the left eigenvectors are *not* simply the complex-conjugate transpose of the right eigenvectors.

Since T is generically diagonalizable, all of the projection operators are simply the normalized ket–bra outer products:

$$T_\lambda = \frac{|\lambda\rangle \langle \lambda|}{\langle \lambda | \lambda \rangle} ,$$

so long as $p \neq 2\sqrt{q} - q$. E.g., $T_1 = |1\rangle \langle 1| / \langle 1 | 1 \rangle = |\mathbf{1}\rangle \langle \pi|$.

Along the nondiagonalizable ($p = 2\sqrt{q} - q$)-subspace of parameter settings:

$\langle -\sqrt{q} | -\sqrt{q} \rangle = 0$, which corresponds to the fact that the right and left eigenvectors are now dual to the left and right *generalized* eigenvectors, rather than being dual to each other.

We find the projection operator for the nondiagonalizable eigenspace to be:

$$T_{-\sqrt{q}} = \frac{|-\sqrt{q}^{(1)}\rangle \langle -\sqrt{q}^{(2)}|}{\langle -\sqrt{q}^{(2)} | -\sqrt{q}^{(1)} \rangle} + \frac{|-\sqrt{q}^{(2)}\rangle \langle -\sqrt{q}^{(1)}|}{\langle -\sqrt{q}^{(1)} | -\sqrt{q}^{(2)} \rangle} ,$$

where:

$$\begin{aligned} |-\sqrt{q}^{(1)}\rangle &= \begin{bmatrix} q & -\sqrt{q} & 1 & -\sqrt{q} \end{bmatrix}^\top , \\ |-\sqrt{q}^{(2)}\rangle &= \begin{bmatrix} -\sqrt{q} & 0 & 1/\sqrt{q} & 0 \end{bmatrix}^\top , \\ \langle -\sqrt{q}^{(1)}| &= \begin{bmatrix} q & q(\sqrt{q} - 2) & -q\sqrt{q} & q \end{bmatrix} , \text{ and} \\ \langle -\sqrt{q}^{(2)}| &= \begin{bmatrix} -\sqrt{q} & 0 & 0 & \sqrt{q} \end{bmatrix} . \end{aligned}$$

Equivalently, $T_{-\sqrt{q}} = I - T_1 - T_0$. The projection operators for the remaining ($\lambda = 0, 1$) eigenspaces retains the same form as before: $T_\lambda = |\lambda\rangle \langle \lambda| / \langle \lambda | \lambda \rangle$.

3.16.1 Observed Correlation

We now have everything we need to calculate the observable correlation and power spectrum.

Recall that we have derived the general spectral decomposition of the autocorrelation

function $\gamma(L) = \langle \bar{X}_t X_{t+L} \rangle_t$ to be:

$$\begin{aligned} \gamma(L) &= \sum_{\substack{\lambda \in \Lambda_T \\ \lambda \neq 0}} \sum_{m=0}^{\nu_\lambda-1} \langle \pi \bar{\mathcal{A}} | T_{\lambda,m} | \mathcal{A} \mathbf{1} \rangle \binom{|L|-1}{m} \lambda^{|L|-1-m} \\ &\quad + [0 \in \Lambda_T] \sum_{m=0}^{\nu_0-1} \langle \pi \bar{\mathcal{A}} | T_0 T^m | \mathcal{A} \mathbf{1} \rangle \delta_{|L|-1,m} \end{aligned}$$

for nonzero integer L , where:

$$\langle \pi \bar{\mathcal{A}} | = \langle \pi | \sum_{x \in \mathcal{A}} \bar{x} T^{(x)} = \frac{1}{1+p+2q} \begin{bmatrix} p & p & 2q & 2q \end{bmatrix},$$

and:

$$| \mathcal{A} \mathbf{1} \rangle = \sum_{x \in \mathcal{A}} x T^{(x)} | \mathbf{1} \rangle = \begin{bmatrix} p+2q & 1 & 2 & 0 \end{bmatrix}^\top.$$

For generic parameter settings, this reduces to:

$$\begin{aligned} \gamma(L) &= \frac{\langle \pi \bar{\mathcal{A}} | 0 \rangle \langle 0 | \mathcal{A} \mathbf{1} \rangle}{\langle 0 | 0 \rangle} \delta_{|L|,1} + \sum_{\substack{\lambda \in \Lambda_T \\ \lambda \neq 0}} \frac{\langle \pi \bar{\mathcal{A}} | \lambda \rangle \langle \lambda | \mathcal{A} \mathbf{1} \rangle}{\langle \lambda | \lambda \rangle} \lambda^{|L|-1} \\ &= 4 \left(\frac{p+2q}{1+p+2q} \right)^2 + \frac{-p}{1+p+2q} \delta_{|L|,1} + \frac{p+2q}{1+p+2q} \sum_{\substack{\lambda \in \Lambda_T \\ \lambda \neq 0,1}} \frac{(\lambda+1)^2 (p\lambda+2q)}{\lambda^3 + p\lambda + 2q} \lambda^{|L|-1}. \end{aligned}$$

Notably, the autocorrelation splits into an ephemeral part (via the Kronecker delta) due to the zero eigenvalue, an exponentially decaying oscillatory part due to the two eigenvalues with magnitude between 0 and 1, and an asymptotic part that survives even as $L \rightarrow \infty$ due to the eigenvalue of unity. Moreover, for $p < 2\sqrt{q} - q$, the two non-trivial eigenvalues become complex conjugate pairs, which allows us to rewrite the autocorrelation for the (2-1)-GP-(2) process for $|L| \geq 2$ concisely as the exact solution:

$$\gamma(L) = \beta^2 + \beta q^{|L|/2} \operatorname{Re}(\zeta e^{i\omega_\xi |L|}),$$

where $\beta \equiv \frac{2(p+2q)}{1+p+2q}$, $\zeta \equiv \frac{(\xi+1)^2(p\xi+2q)}{\xi(\xi^3+p\xi+2q)}$, $\xi \equiv -\frac{1}{2}(p+q) + i\frac{1}{2}\sqrt{4q-(p+q)^2}$, and $\omega_\xi \equiv \frac{\pi}{2} + \arctan\left(\frac{p+q}{\sqrt{4q-(p+q)^2}}\right)$. This latter form clarifies that the magnitude of the largest non-unity eigenvalue $|\xi| = \sqrt{q}$ controls the slowest rate of decay of observed correlation $|\gamma(L)| \sim \sqrt{q}^L$, and that the complex phase of the eigenvalues determine the oscillations within this exponentially decaying envelope.

We have found the general spectral decomposition of the continuous part of the power spectrum to be:

$$P_c(\omega) = \langle |x|^2 \rangle + 2 \sum_{\lambda \in \Lambda_T} \sum_{m=0}^{\nu_\lambda-1} \operatorname{Re} \frac{\langle \pi \bar{\mathcal{A}} | T_{\lambda,m} | \mathcal{A} \mathbf{1} \rangle}{(e^{i\omega} - \lambda)^{m+1}} .$$

Also, recall from earlier that the delta functions of the power spectrum arise from the eigenvalues of T that lie on the unit circle:

$$P_d(\omega) = \sum_{k=-\infty}^{\infty} \sum_{\substack{\lambda \in \Lambda_T \\ |\lambda|=1}} 2\pi \delta(\omega - \omega_\lambda + 2\pi k) \operatorname{Re}(\lambda^{-1} \langle \pi \bar{\mathcal{A}} | T_\lambda | \mathcal{A} \mathbf{1} \rangle) ,$$

where ω_λ is related to λ by $\lambda = e^{i\omega_\lambda}$. As long as $p + q < 1$, $\lambda = 1$ is the only eigenvalue that lies on the unit circle, so that $P_d(\omega) = \sum_{k=-\infty}^{\infty} 2\pi \delta(\omega + 2\pi k) \operatorname{Re} \langle \pi \bar{\mathcal{A}} | T_1 | \mathcal{A} \mathbf{1} \rangle$, with $\operatorname{Re} \langle \pi \bar{\mathcal{A}} | T_1 | \mathcal{A} \mathbf{1} \rangle = \beta^2 = 4 \left(\frac{p+2q}{1+p+2q} \right)^2$.

Putting this all together we find the full power spectrum for the (2-1)-GP-(2) process:

$$P(\omega) = \frac{8q}{1+p+2q} + \frac{2p}{1+p+2q} [1 - \cos(\omega)] \\ + \beta \sum_{\substack{\lambda \in \Lambda_T \\ \lambda \neq 0,1}} \operatorname{Re} \left(\frac{(\lambda+1)^2(p\lambda+2q)}{\lambda^3+p\lambda+2q} \cdot \frac{1}{e^{i\omega} - \lambda} \right) + 2\pi\beta^2 \sum_{k=-\infty}^{\infty} \delta(\omega + 2\pi k) ,$$

which exhibits delta functions from the eigenvalue on the unit circle, continuous Lorentzian-like line profiles emanating from finite eigenvalues inside the unit circle and expressive of the chaotic nature of the time-series, and the unique sinuisoidal contribution that can only come from zero eigenvalues.

Whenever $p < 2\sqrt{q} - q$, the two non-trivial eigenvalues become complex conjugate pairs, which allows us to rewrite the power spectrum for the (2-1)-GP-(2) process in a more transparent manner:

$$P(\omega) = \frac{8q}{1+p+2q} + \frac{2p}{1+p+2q} [1 - \cos(\omega)] \\ + \beta \operatorname{Re} \left(\frac{\zeta \xi}{e^{i\omega} - \xi} + \frac{\zeta \xi}{e^{-i\omega} - \xi} \right) + 2\pi\beta^2 \sum_{k=-\infty}^{\infty} \delta(\omega + 2\pi k) ,$$

which is clearly symmetric about $\omega = 0$.

The level of completeness we have achieved is gratifying: we have calculated these properties exactly in closed form for this infinite-Markov order stochastic process over

the full range of possible parameter settings. Moreover, we will produce the same level of completeness in our analysis of the predictability of this process in the next section after constructing its \mathcal{S} -MSP.

3.16.2 Predictability

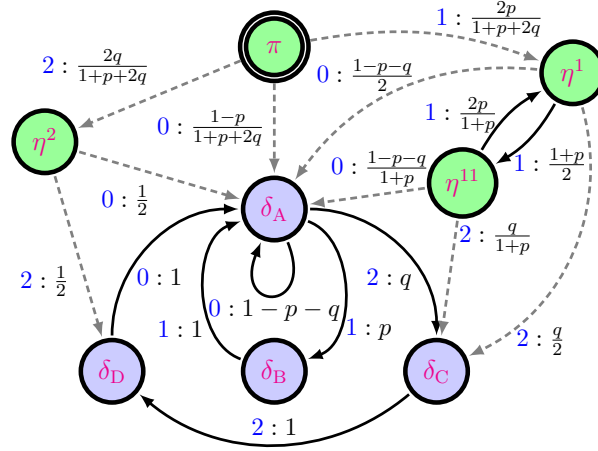
To analyze the predictability of the process, we need the \mathcal{S} -MSP of any generator of the process. Since we started already with the ϵ -machine, we will directly generate the \mathcal{S} -MSP of the ϵ -machine, and will have the mixed-state transition matrix $W = \mathcal{W}$ which will suffice for calculations of both *predictability* in this section and the synchronization necessary for *prediction* in the next.

We construct the \mathcal{S} -MSP by constructing all mixed states that can be induced by observation from the start distribution π , and calculating the transition probabilities among them. There are eight such mixed-state distributions over \mathcal{S} , although four of them (δ_A , δ_B , δ_C , and δ_D) correspond to completely synchronized peaked distributions. The other four states are new (when compared to the recurrent ϵ -machine) transient states, corresponding to transient states of uncertainty during synchronization. Calculating, we find the eight unique mixed-state distributions iteratively from $\langle \eta^{wx} | = \frac{\langle \eta^w | T^{(x)} | \mathbf{1} \rangle}{\langle \eta^w | T^{(x)} | \mathbf{1} \rangle}$, starting with $\langle \eta^x | = \frac{\langle \pi | T^{(x)} | \mathbf{1} \rangle}{\langle \pi | T^{(x)} | \mathbf{1} \rangle}$, to be:

$$\begin{aligned} \langle \pi | &= \frac{1}{1+p+2q} \begin{bmatrix} 1 & p & q & q \end{bmatrix}, \\ \langle \eta^1 | &= \frac{1}{2} \begin{bmatrix} 1 & 1 & 0 & 0 \end{bmatrix}, \\ \langle \eta^{11} | &= \frac{1}{1+p} \begin{bmatrix} 1 & p & 0 & 0 \end{bmatrix}, \\ \langle \eta^2 | &= \frac{1}{2} \begin{bmatrix} 0 & 0 & 1 & 1 \end{bmatrix}, \\ \langle \delta_A | &= \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}, \\ \langle \delta_B | &= \begin{bmatrix} 0 & 1 & 0 & 0 \end{bmatrix}, \\ \langle \delta_C | &= \begin{bmatrix} 0 & 0 & 1 & 0 \end{bmatrix}, \text{ and} \\ \langle \delta_D | &= \begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix}, \end{aligned}$$

where the transient mixed states are labeled according to the first word that induces them. These distributions constitute the set of mixed-states \mathcal{R}_π of the \mathcal{S} -MSP. Moreover, each transition probability from mixed state $\langle \eta^w |$ to mixed state $\langle \eta^{wx} |$ is calculated as: $\langle \eta^w | T^{(x)} | \mathbf{1} \rangle$. Altogether, these calculations yield the \mathcal{S} -MSP of the (2-1)-GP-(2) process,

reproduced here from Table 3.4 for convenience:



As a HMM, the S-MSP of the ϵ -machine is fully described by the 4-tuple: $\mathcal{M}_{\text{S-MSP}} = (\mathcal{R}_\pi, \mathcal{A}, \{\mathcal{W}^{(x)}\}_{x \in \mathcal{A}}, \mu_0 = \delta_\pi)$, where \mathcal{R}_π is the set of mixed states just quoted, \mathcal{A} is the same observable alphabet as before, $\{\mathcal{W}^{(x)}\}_{x \in \mathcal{A}}$ is the set of symbol-labeled transition matrices among the mixed states and $\langle \delta_\pi | = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$ is the start distribution over the mixed states.

With this new linear meta-dynamic in hand, our next step is to calculate the eigenvalues and projection operators of the net mixed-state-to-state transition dynamic $\mathcal{W} = \sum_{x \in \mathcal{A}} \mathcal{W}^{(x)}$. Explicitly, \mathcal{W} can be represented in the block-matrix form:

$$\mathcal{W} = \begin{bmatrix} A & B \\ \mathbf{0} & T \end{bmatrix} \text{ where } A = \begin{bmatrix} 0 & \frac{2p}{1+p+2q} & 0 & \frac{2q}{1+p+2q} \\ 0 & 0 & \frac{1+p}{2} & 0 \\ 0 & \frac{2p}{1+p} & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, B = \begin{bmatrix} \frac{1-p}{1+p+2q} & 0 & 0 & 0 \\ \frac{1-p-q}{2} & 0 & \frac{q}{2} & 0 \\ \frac{1-p-q}{1+p} & 0 & \frac{q}{1+p} & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \end{bmatrix}, \text{ and } T \text{ is}$$

the same as the state-to-state net transition matrix of the ϵ -machine from earlier. The eigenvalues of \mathcal{W} are thus relatively to calculate since $\det(\mathcal{W} - \lambda I) = \det(A - \lambda I)\det(T - \lambda I)$, implies that:

$$\Lambda_{\mathcal{W}} = \Lambda_A \cup \Lambda_T .$$

The new eigenvalues introduced by the feedback matrix A are: $\Lambda_A = \{0, \pm\sqrt{p}\}$ with $\nu_0 = 2$ while the other eigenvalues ($\pm\sqrt{p}$), which are found most easily from Eq. (3.38), are associated with diagonalizable subspaces. It is important to note that, while T was only

nondiagonalizable along a very special submanifold of parameter space, the mixed-state-to-state metadynamic is *generically* nondiagonalizable over all parameter settings. This nondiagonalizability corresponds to a special kind of symmetry breaking of uncertainty during synchronization.

The eigenvectors of \mathcal{W} are most easily found through a two-step process. E.g., $|\pm\sqrt{p_A}\rangle$ and $\langle\pm\sqrt{p_A}|$ (the solutions of $A|\pm\sqrt{p_A}\rangle = \pm\sqrt{p}|\pm\sqrt{p_A}\rangle$ and $\langle\pm\sqrt{p_A}|A = \pm\sqrt{p}\langle\pm\sqrt{p_A}|$) are found first, and the result is used to reduce the number of unknowns when solving the full eigen-equations (e.g., $\mathcal{W}|\pm\sqrt{p_W}\rangle = \pm\sqrt{p}|\pm\sqrt{p_W}\rangle$) for $|\pm\sqrt{p_W}\rangle$ and $\langle\pm\sqrt{p_W}|$. Similarly, we can recycle the restricted eigenvectors $|\lambda_T\rangle$ and $\langle\lambda_T|$ found earlier for the ϵ -machine to reduce the number of unknowns when solving the more general eigenvector problems for $|\lambda_W\rangle$ and $\langle\lambda_W|$ in cases where $\lambda \in \Lambda_T$. Performing such a calculation, we find:

$$\begin{aligned} |\pm\sqrt{p_W}\rangle &= \begin{bmatrix} \frac{1}{1+p+2q} & \frac{\pm\sqrt{p}}{2p} & \frac{1}{1+p} & 0 & 0 & 0 & 0 & 0 \end{bmatrix}^\top, \\ \langle\pm\sqrt{p_W}| &= \begin{bmatrix} 0 & \pm\sqrt{p} & \frac{1+p}{2} & 0 & -\frac{1\pm\sqrt{p}}{2} & -\frac{p\pm\sqrt{p}}{2} & 0 & 0 \end{bmatrix}, \\ \langle\lambda_W| &= \begin{bmatrix} 0 & 0 & 0 & 0 & \lambda^2 & p\lambda & q\lambda & q \end{bmatrix} \text{ for all } \lambda \in \Lambda_T \setminus \{0\}, \\ |1_W\rangle &= \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}^\top, \text{ and} \\ |\lambda_W\rangle &= \begin{bmatrix} 0 & \frac{\lambda(1+\lambda)}{2} & \frac{\lambda(1+\lambda)}{1+p} & \frac{1+\lambda}{2} & \lambda^2 & \lambda & 1 & \lambda \end{bmatrix}^\top \text{ for } \lambda \in \Lambda_T \setminus \{0, 1\}. \end{aligned}$$

Moreover, the eigenvectors and generalized eigenvector for \mathcal{W} corresponding to the eigenvalue of 0 are:

$$\begin{aligned} |0_1^{(1)}\rangle &= \langle\delta_\pi|^\top = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}^\top, \\ |0_1^{(2)}\rangle &= \begin{bmatrix} 0 & 0 & 0 & \frac{1+p+2q}{2q} & 0 & 0 & 0 & 0 \end{bmatrix}^\top, \\ |0_2^{(1)}\rangle &= \begin{bmatrix} 0 & \frac{1}{2} & \frac{p}{1+p} & \frac{-p}{2q} & 0 & 1 & \frac{-p}{q} & 0 \end{bmatrix}^\top, \\ \langle 0_1^{(1)}| &= \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 & \frac{-1}{2} & \frac{-1}{2} \end{bmatrix}, \\ \langle 0_1^{(2)}| &= \begin{bmatrix} \frac{1+p+2q}{2q} & 0 & \frac{-(1+p)}{2q} & 0 & 0 & 0 & \frac{-1}{2} & \frac{-1}{2} \end{bmatrix}, \text{ and} \\ \langle 0_2^{(1)}| &= \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 \end{bmatrix}, \end{aligned}$$

where we have used the notation $|0_k^{(m)}\rangle$ for indexing generalized eigenvectors introduced in § 3.7.2.

All non-degenerate eigenvalues have projection operators of the form: $\mathcal{W}_\lambda = \frac{|\lambda\rangle\langle\lambda|}{\langle\lambda|\lambda\rangle}$. However, the degenerate and nondiagonalizable subspace associated with the zero eigenvalue

has the composite projection operator:

$$\mathcal{W}_0 = \frac{|0_1^{(1)}\rangle\langle 0_1^{(2)}|}{\langle 0_1^{(2)}|0_1^{(1)}\rangle} + \frac{|0_1^{(2)}\rangle\langle 0_1^{(1)}|}{\langle 0_1^{(1)}|0_1^{(2)}\rangle} + \frac{|0_2^{(1)}\rangle\langle 0_2^{(1)}|}{\langle 0_2^{(1)}|0_2^{(1)}\rangle}.$$

The fact that $\langle \delta_\pi | \lambda_W \rangle = 0$ for all $\lambda \in \Lambda_T \setminus (\{1\} \cup \Lambda_A)$ is an instantiation of a general result that greatly simplifies the calculations relating to predictability and prediction.

The final puzzle piece for analyzing predictability is the vector of transition-entropies $|H(W^{\mathcal{A}})\rangle$. A simple calculation, utilizing the fact that $-\sum_i \frac{n_i}{d} \log\left(\frac{n_i}{d}\right) = \log(d) - \frac{1}{d} \sum_i n_i \log(n_i)$ when $\sum_i n_i = d$, yields:

$$|H(W^{\mathcal{A}})\rangle = \begin{bmatrix} \log(1+p+2q) - \frac{1}{1+p+2q} [2q \log(2q) + 2p \log(2p) + (1-p) \log(1-p)] \\ 1 - \frac{1}{2} [q \log(q) + (1+p) \log(1+p) + (1-p-q) \log(1-p-q)] \\ \log(1+p) - \frac{1}{1+p} [q \log(q) + 2p \log(2p) + (1-p-q) \log(1-p-q)] \\ 1 \\ -[q \log(q) + p \log(p) + (1-p-q) \log(1-p-q)] \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

where \log is understood to be the base-2 logarithm \log_2 .

Putting this all together, we can now calculate in full detail the myopic entropy rate $h_\mu(L)$ that would result from modeling the infinite-order (2-1)-GP-(2) process as an order- $(L-1)$ Markov process.

$$\begin{aligned} h_\mu(L) &= \sum_{\substack{\lambda \in \Lambda_W \\ \lambda \neq 0}} \sum_{m=0}^{\nu_\lambda-1} \langle \delta_\pi | W_{\lambda,m} | H(W^{\mathcal{A}}) \rangle \binom{L-1}{m} \lambda^{L-1-m} \\ &\quad + [0 \in \Lambda_W] \sum_{m=0}^{\nu_0-1} \delta_{L-1,m} \langle \delta_\pi | W_0 W^m | H(W^{\mathcal{A}}) \rangle \\ &= \delta_{L,1} \langle \delta_\pi | W_0 | H(W^{\mathcal{A}}) \rangle + \delta_{L,2} \langle \delta_\pi | W W_0 | H(W^{\mathcal{A}}) \rangle \\ &\quad + \langle \delta_\pi | W_1 | H(W^{\mathcal{A}}) \rangle + \sum_{\lambda \in \Lambda_W \setminus \Lambda_T} \langle \delta_\pi | W_\lambda | H(W^{\mathcal{A}}) \rangle \lambda^{L-1}, \end{aligned}$$

which oscillates under an exponentially decaying envelope as it approaches its asymptotic

value of:

$$\begin{aligned}
h_\mu &= \langle \delta_\pi | W_1 | H(W^{\mathcal{A}}) \rangle = \langle \delta_\pi | 1_{\mathcal{W}} \rangle \langle \pi_{\mathcal{W}} | H(W^{\mathcal{A}}) \rangle \\
&= \langle \pi_{\mathcal{W}} | H(W^{\mathcal{A}}) \rangle = \langle \pi_T | H(T^{\mathcal{A}}) \rangle \\
&= \frac{-q \log q - p \log p - (1-p-q) \log(1-p-q)}{1+p+2q} .
\end{aligned}$$

Simplifying the terms in the myopic entropy rate yields:

$$\langle \delta_\pi | W_0 | H(W^{\mathcal{A}}) \rangle = \langle \delta_\pi | H(W^{\mathcal{A}}) \rangle - \frac{1+p}{1+p+2q} \langle \delta_{\eta^{11}} | H(W^{\mathcal{A}}) \rangle$$

and

$$\langle \delta_\pi | WW_0 | H(W^{\mathcal{A}}) \rangle = \frac{2q}{1+p+2q}$$

for the two ephemeral contributions.

For $L \geq 3$, we find:

$$h_\mu(L) - h_\mu = \begin{cases} \frac{-p \log p + (1+p) \log(1+p) - 2p}{\sqrt{p}(1+p+2q)} p^{L/2} & \text{for odd } L \\ \frac{p \log p - (1+p) \log(1+p) + 2}{1+p+2q} p^{L/2} & \text{for even } L \end{cases}$$

which highlights the period-2 nature of the asymptotic decay.

The total mutual information between the observable past and observable future is the so-called excess entropy:

$$\begin{aligned}
\mathbf{E} &= \sum_{\lambda \in \Lambda_W \setminus \{1\}} \sum_{m=0}^{\nu_\lambda-1} \frac{1}{(1-\lambda)^{m+1}} \langle \delta_\pi | W_{\lambda,m} | H(W^{\mathcal{A}}) \rangle \\
&= \log(1+p+2q) + \frac{(1-p-q) \log(1-p-q) - p \log p - q \log q - (1-p) \log(1-p)}{1+p+2q} .
\end{aligned}$$

This is the total amount of the future that could possibly be predicted using past observations. The structure of *how* this information is unraveled over time is revealed in the

excess entropy spectrum:

$$\begin{aligned}
\mathcal{E}(\omega) &= \pi h_\mu \delta(\omega) + \sum_{m=0}^{\nu_0-1} \cos((m+1)\omega) \langle \delta_\pi | W_0 W^m | H(W^A) \rangle \\
&\quad + \sum_{\lambda \in \Lambda_W \setminus 0} \sum_{m=0}^{\nu_\lambda-1} \operatorname{Re} \left(\frac{\langle \delta_\pi | W_{\lambda,m} | H(W^A) \rangle}{(e^{i\omega} - \lambda)^{m+1}} \right) \\
&= \pi h_\mu \delta(\omega) + \left(\mathbf{E} + \frac{p \log p - (1+p) \log(1+p) - 2q}{1+p+2q} \right) \cos(\omega) + \frac{2q}{1+p+2q} \cos(2\omega) \\
&\quad + \frac{-p \log p + (1+p) \log(1+p)}{1+p+2q} \operatorname{Re} \left(\frac{e^{i\omega} - p}{e^{i2\omega} - p} \right) + \frac{2p}{1+p+2q} \operatorname{Re} \left(\frac{1 - e^{i\omega}}{e^{i2\omega} - p} \right),
\end{aligned}$$

from which we can observe that $\mathbf{E} = \lim_{\omega \rightarrow 0} \mathcal{E}(\omega)$.

3.16.3 Synchronization to Optimal Prediction

To analyze the burden of optimally synchronizing to the process, we need the \mathcal{S} -MSP of the ϵ -machine, which we already constructed in the last section. Hence, we can immediately evaluate the resources necessary for synchronization to and prediction of the (2-1)-GP-(2) process.

The only new puzzle piece needed for analyzing the synchronization associated with optimal prediction is the vector of mixed-state entropies $|H[\eta]\rangle$. A simple calculation yields:

$$|H[\eta]\rangle = \begin{bmatrix} \log(1+p+2q) - \frac{p \log p + 2q \log q}{1+p+2q} \\ 1 \\ \log(1+p) - \frac{p \log p}{1+p} \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

where \log is again understood to be the base-2 logarithm \log_2 .

An observer, bent on predicting the future as well as possible, must synchronize to the causal state of the dynamic. During the meta-dynamic of synchronization, the observer will on average pick up synchronization information over time according to the remaining

causal-state uncertainty $\mathcal{H}^+(L)$ after an observation duration of L :

$$\begin{aligned}
\mathcal{H}^+(L) &= \sum_{\substack{\lambda \in \Lambda_{\mathcal{W}} \\ \lambda \neq 0}} \sum_{m=0}^{\nu_{\lambda}-1} \langle \delta_{\pi} | \mathcal{W}_{\lambda,m} | H[\eta] \rangle \binom{L}{m} \lambda^{L-m} \\
&\quad + [0 \in \Lambda_{\mathcal{W}}] \sum_{m=0}^{\nu_0-1} \delta_{L,m} \langle \delta_{\pi} | \mathcal{W}_0 \mathcal{W}^m | H[\eta] \rangle \\
&= \delta_{L,0} \left[\log(1+p+2q) - \frac{2q \log q + (1+p) \log(1+p)}{1+p+2q} \right] \\
&\quad + \delta_{L,1} \frac{2q}{1+p+2q} + \frac{\langle \sqrt{p} | H[\eta] \rangle + (-1)^L \langle -\sqrt{p} | H[\eta] \rangle}{1+p+2q} p^{L/2} .
\end{aligned}$$

More explicitly, for $L \geq 2$, this becomes:

$$\mathcal{H}^+(L) = \begin{cases} \frac{2\sqrt{p}}{1+p+2q} p^{L/2} & \text{for odd } L \\ \frac{(1+p) \log(1+p) - p \log p}{1+p+2q} p^{L/2} & \text{for even } L \end{cases} .$$

The net synchronization information accumulated is:

$$\begin{aligned}
\mathbf{S} &= \sum_{\lambda \in \Lambda_{\mathcal{W}} \setminus \{1\}} \sum_{m=0}^{\nu_{\lambda}-1} \frac{1}{(1-\lambda)^{m+1}} \langle \delta_{\pi} | \mathcal{W}_{\lambda,m} | H[\eta] \rangle \\
&= \log(1+p+2q) + \frac{2q(1-\log q)}{1+p+2q} + \frac{p[2-\log p + (1+p) \log(1+p)]}{(1-p)(1+p+2q)} .
\end{aligned}$$

Even after synchronization, an observer must update an average of b_{μ} of its bits of information per observation and must keep track of a net C_{μ} bits of information to stay synchronized, where:

$$\begin{aligned}
C_{\mu} &= H[\pi] = \langle \delta_{\pi} | H[\eta] \rangle \\
&= \log(1+p+2q) - \frac{p \log p + 2q \log q}{1+p+2q} .
\end{aligned}$$

An interesting feature of prediction is the crypticity of the process,

$$\chi = C_{\mu} - \mathbf{E} = \frac{3q \log q + (1-p) \log(1-p) - (1-p-q) \log(1-p-q)}{1+p+2q} ,$$

which shows, as a function of p and q , the minimal overhead of additional (classical) memory about the past—beyond the information that the future can possibly share with it—that must be stored for optimal prediction.

In summary, we can appreciate that we have come to a rather complete analysis of the fundamental limits on predictability of our example stochastic process. This example showed many of the qualitative features, both in terms of the calculation and the system’s behavior, that should be expected generically in the analysis of prediction, based on the more general results of this chapter. The demonstrated procedure can be co-opted to work with any inference algorithm that yields a model that produces the observed dynamic, whether this is classical machine learning or Bayesian Structural Inference [243], or any other favorite tool. With a generative model in hand, synchronizing to the real world data—necessary to make good predictions about the real world’s future—follows the metadynamics of the \mathcal{S} -MSP, with the consequences for prediction that there will typically be a finite epoch of symmetry collapse followed by a slower asymptotic synchronization that allows ever-better prediction as longer observations induce a refined knowledge of what lies hidden beneath.

3.17 Conclusion

Many questions that we care to ask about a structured stochastic nonlinear process surprisingly imply a linear dynamic over a preferred hidden state space. Using the appropriate dynamic for common questions of complexity, we have revealed closed-form expressions for complexity measures as simple functions of the corresponding transition dynamic of the implied HMM. We have focused on predictability and prediction.

To make predictions about the real world, it is not sufficient to have a model of the world. Additionally, the predictor must synchronize their model to the real-world data that has been observed up to the present time. This meta-dynamic of synchronization—the transition structure among belief states—is intrinsically linear, but is typically nondiagonalizable. Despite this challenge, we were able to derive closed-form formulae that reveal general features of how the world becomes more predictable as we synchronize to it.

Recall the organizational tables from the beginning of this chapter. After working through more details, it may be a good time to consider a more nuanced formulation. We have seen that once we frame our questions in terms of the hidden linear transition dynamic,

Derivatives of cascading \uparrow	Discrete time	Continuous time
Cascading	$\langle \cdot A^L \cdot \rangle$	$\langle \cdot e^{tG} \cdot \rangle$
Integrals of cascading \downarrow		
Accumulated transients	$\langle \cdot (\sum_L (A - A_1)^L) \cdot \rangle$	$\langle \cdot (\int (e^{tG} - G_0) dt) \cdot \rangle$
modulated accumulation	$\langle \cdot (\sum_L (zA)^L) \cdot \rangle$	$\langle \cdot (\int (ze^G)^t dt) \cdot \rangle$

Table 3.5: Once we identify the hidden linear dynamic behind our questions, most questions we tend to ask are either of the *cascading* or *accumulating* type. If a complexity measure accumulates transients, the Drazin inverse is likely to appear. Interspersed accumulation can be a nice theoretical tool, since all derivatives and integrals of cascading can be calculated if we know the modified accumulation with $z \in \mathbb{C}$. With $z \in \mathbb{C}$, modulated accumulation involves an operator-valued z -transform. With $z = e^{i\omega}$ and $\omega \in \mathbb{R}$, modulated accumulation involves an operator-valued Fourier-transform.

Genre	Implied linear transition dynamic	Example Questions		
		Cascading	Accumulated transients	Modulated accumulation
OVERT OBSERVATIONAL	Transition matrix T of any HMM	Correlations, $\gamma(L)$: $\langle \pi \bar{\mathcal{A}} T^{ L -1} \mathcal{A} \mathbf{1} \rangle$	Green-Kubo transport coefficients	Power spectra, $P(\omega)$: $2\Re \langle \pi \bar{\mathcal{A}} (e^{i\omega} I - T)^{-1} \mathcal{A} \mathbf{1} \rangle$
PREDICTABILITY	Transition matrix W of MSP of any HMM	Myopic entropy rate, $h_\mu(L)$: $\langle \delta_\pi W^{L-1} H(W^{\mathcal{A}}) \rangle$	Excess entropy, \mathbf{E} : $\langle \delta_\pi (I - W)^{\mathcal{D}} H(W^{\mathcal{A}}) \rangle$	$E(z)$: $\langle \delta_\pi (zI - W)^{-1} H(W^{\mathcal{A}}) \rangle$
OPTIMAL PREDICTION	Transition matrix \mathcal{W} of MSP of ϵ -machine	Causal state uncertainty, $\mathcal{H}^+(L)$: $\langle \delta_\pi \mathcal{W}^L H[\eta] \rangle$	Synchronization info, \mathbf{S} : $\langle \delta_\pi (I - \mathcal{W})^{\mathcal{D}} H[\eta] \rangle$	$S(z)$: $\langle \delta_\pi (zI - \mathcal{W})^{-1} H[\eta] \rangle$
\vdots	\vdots	\vdots	\vdots	\vdots

Table 3.6: Several genres of questions about the complexity of a process are given in the left column of the table in order of increasing sophistication. Each genre implies a different linear transition dynamic. Closed-form formulae are given for example complexity measures, showing the deep similarity among formulae of the same column, while formulae in the same row have matching bra-ket pairs. The similarity within the column corresponds to similarity in the type of time-evolution implied by the question type. The similarity within the row corresponds to similarity in the genre of the question.

complexity measures are usually either of the cascading or accumulation type. Scalar complexity measures often accumulate only the interesting transient structure that rides on top of the asymptotics. Skimming off the asymptotics led to a Drazin inverse. Modified accumulation turns complexity scalars into complexity functions. This is summarized in Table 3.5 and Table 3.6. It is notable that Table 3.6 gives closed-form formulae for many complexity measures that previously were only expressed as infinite sums over functions of probabilities.

Let us remind ourselves: Why, in this analysis, were nondiagonalizable dynamics noteworthy? They are noteworthy because the *meta-dynamics* of diagonalizable dynamics

are generically nondiagonalizable—typically due to the zero-eigenvalue subspace that is responsible for the ephemeral epoch of symmetry collapse. We saw this explicitly with the metadynamics of transitioning among belief states. However, other metadynamics in fields beyond prediction are also generically nondiagonalizable. For example, in the analysis of quantum compression, crypticity, and other aspects of hidden structure, the relevant linear dynamic is not the MSP but is nevertheless a nondiagonalizable structure that is fruitfully analyzed with the recently generalized spectral theory of non-normal operators [198].

Chapter 4

Quantum Simplicity of Classical Complexity: Reduced State-Complexity of Quantum-Encoded Cryptic Processes

This chapter has appeared in print as:

P. M. Riechers, J. R. Mahoney, C. Aghamohammadi, & J. P. Crutchfield, (2016). “Minimized state complexity of quantum-encoded cryptic processes”. *Physical Review A*, 93(5), 052317.

At the time of writing this dissertation, the status of *minimal* memory requirements for a quantum system to generate an observable classical process is still under investigation, but it is conjectured that the construction discussed here yields the minimal memory requirements given the constraint that observations are made using projective measurements. If non-projective measurements (POVMs) are allowed, it appears that less memory may be required in at least some circumstances—however much less is currently known about this.

4.1 Chapter Overview

The predictive information required for proper trajectory sampling of a stochastic process can be more efficiently transmitted via a quantum channel than a classical one. This recent discovery allows quantum information processing to drastically reduce the memory

necessary to simulate complex classical stochastic processes. It also points to a new perspective on the intrinsic complexity that nature must employ in generating the processes we observe. The quantum advantage increases with codeword length—the length of process sequences used in constructing the quantum communication scheme. In analogy with the classical complexity measure, statistical complexity, we use this reduced communication cost as an entropic measure of state-complexity in the quantum representation. Previously difficult to compute, the quantum advantage is expressed here in closed form using spectral decomposition. This allows for efficient numerical computation of the quantum-reduced state-complexity at all encoding lengths, including infinite. Additionally, it makes clear how finite-codeword reduction in state-complexity is controlled by the classical process’ cryptic order. And, it allows asymptotic analysis of infinite-cryptic-order processes.

4.2 Introduction

To efficiently synchronize predictions of a given process over a classical communication channel two observers, call them Alice and Bob, must know the process’ internal structure and communicate the relevant history. In particular, leveraging common knowledge of the process’ dynamic, what is the minimal amount of information that Alice must communicate to Bob so that he can make the same probabilistic prediction as Alice? The answer is given by the process’ internal state information or *statistical complexity* C_μ [55].

A closely related question immediately suggests itself: is it more efficient to synchronize via a quantum communication channel that transmits qubits instead of bits? Extending early answers [102, 99], a sequence of constructions (*q-machines*) was recently introduced that offers substantial message-size reduction below C_μ [157]. In these constructions, each codeword length L yields a quantum communication cost $C_q(L) \leq C_\mu$ that decreases with increasing L . Moreover, the maximum-compression complexity, $C_q(\infty) = C_q(k)$, is achieved at a codeword length called the *cryptic order* k [158, 160]—a recently discovered classical, topological property that is a cousin to the Markov order familiar from stochastic process theory.

Reference [157] pointed out that the new efficiency in synchronization comes with a

tradeoff. Bob can only make predictions that are more specialized than Alice’s: those consistent with Alice’s but also consistent with a probabilistically generated extension of the codewords Alice uses to construct the qubits she sends. These constraints lead to a seemingly odd way for Alice and Bob to synchronize, but there is no way around this. The merits of this tradeoff are more apparent if we consider the related scenario of Alice Past synchronizing to Alice Future (‘Bob’) as she generates a realization of the process and updates her state. To *generate* a process the *future* possibilities must be synchronized with the *past* in just such a way that information shared between past and future is channeled through the *present* without violating the process’ time order. One consequence is that the communication cost $C_q(L)$ demands a more refined interpretation: it is the average state information that must be remembered to *generate* the process. Another is that $C_q(L)$ decreases with increasing L since codewords merge, yielding increasingly coincident predictions. The conclusion is that a process’ correlational structure controls the degree of quantum compression.

There are both theoretical and practical implications. On the one hand, the theory of minimized quantum-state complexity greatly broadens our notions of the structural complexity inherent in processes; for example, allowing us to quantitatively compare classical- and quantum-state memories [1]. In an applied setting, on the other, it identifies significantly reduced memory requirements for simulating complex classical stochastic processes via quantum processing.

Reduced memory requirements for stochastic simulation were recognized previously for Markov order-1 processes, whose quantum advantage saturates at $C_q(1)$ [102]. For example, it was shown that the classical nearest-neighbor one-dimensional Ising model has a less complex quantum representation [245]. Recently, the quantum advantage of reduced state-complexity was experimentally demonstrated for a simple Markovian dynamic [185].

The increasing quantum advantage discovered in Ref. [157], as encapsulated by $C_q(L)$, was challenging to calculate, analytically and numerically. This was unfortunate since for most complex processes, the optimal state-complexity reduction is only achieved asymptotically as codeword length $L \rightarrow \infty$. Moreover, without a comprehensive theory,

few conclusions could be rigorously drawn about $C_q(L)$'s convergence and limits. The following removes the roadblocks. It delivers closed-form expressions, yielding both numerical efficiency and analytic insight.

Our first contribution is the introduction of the *quantum pairwise-merger machine* (QPMM). The QPMM contains, in a compact form, all of the information required for efficient calculation of the signal-state overlaps used in the q-machine encoding. In particular, we derive closed-form expressions for overlaps in terms of the QPMM's spectrum and projection operators.

This leads to our second contribution: a decomposition of the quantum state-complexity $C_q(L)$ into two qualitatively distinct parts. The first part is present for codeword lengths only up to a finite-horizon equal to the index of the QPMM dynamic which, for the case of finite cryptic order, is equal to the process' cryptic order. This provides a nearly complete understanding of $C_q(L)$ for finite-cryptic-order processes. The second part asymptotically decays with an infinite-horizon and is present only in infinite-cryptic order processes. Moreover, we show that $C_q(L)$ oscillates under an exponentially decaying envelope and explain the relevant rates and frequencies in terms of the QPMM's spectral decomposition.

Our third contribution comes in analyzing how computing $C_q(L)$ requires efficiently manipulating quantum-state overlaps. The technique for this presented in Ref. [157] required constructing a new density matrix that respects overlaps. However, it is known that overlaps may be monitored much more directly via a Gram matrix. Here, we adapt this to improve calculational efficiency and theoretical simplicity. And, we improve matters further by introducing a new form of the Gram matrix.

Our final contribution follows from casting $C_q(L)$'s calculation in its spectral form. This has the distinct advantage that the limit of the overlaps, and thus $C_q(\infty)$, can be calculated analytically. Illustrative examples are placed throughout to ground the development.

4.3 Two Representations of a Process

The objects of interest are discrete-valued, stationary, stochastic processes. A *process* consists of a bi-infinite sequence $X_{-\infty:\infty} = \dots X_{-2}X_{-1}X_0X_1X_2\dots$ of random variables X_t

that take on one or another value in a discrete alphabet: $x_t \in \mathcal{A}$. For each time t and subsequent contiguous block-length L , a process assigns a particular probability $\Pr(w)$ to each length- L word $w = x_t \dots x_{t+L-1}$. For *stationary* processes, these probabilities are independent of t . A stationary process' *language* is that set of words $w = x_0 \dots x_{L-1}$ of any length L generated with positive probability.

In particular, we consider processes that can be generated by finite hidden Markov models (HMMs). For edge-output HMMs (i.e., Mealy HMMs), introduced more formally in App. 4.10.1, the observed symbol is generated on transitions between states.

We next consider two representations of a given process, first a canonical classical representation and then a newer quantum representation. Each utilizes the concept of a process' *causal states*, which are equivalence classes of histories that yield the same conditional probability distributions over future trajectories. Specifically, causal states are the equivalence classes induced by the *predictive equivalence relation*, \sim_ϵ , applied to observable histories $x_{-\infty:0}$:

$$\overleftarrow{x} \sim_\epsilon \overleftarrow{x}' \iff \Pr(X_{0:L}|X_{-\infty:0} = \overleftarrow{x}) = \Pr(X_{0:L}|X_{-\infty:0} = \overleftarrow{x}') \quad \text{for all } L \in \{1, 2, \dots\} .$$

Said another way, causal states are the minimal sufficient statistic of the *past* $X_{-\infty:0}$ for predicting the *future* $X_{0:\infty}$. (We use indexing $X_{a:b}$ that is left inclusive, but right exclusive.)

4.3.1 ϵ -Machine: Optimal, Minimal Predictor

While a given process generally has many alternative HMM representations, there exists a unique, canonical form—the process' ϵ -machine [55], which is a process' minimal optimal predictor. Causal states, which are by definition predictive equivalence classes of histories, are the latent states of the ϵ -machine.

Definition 6. A process' ϵ -machine is the 4-tuple $\{\mathcal{S}, \mathcal{A}, \{T^{(x)}\}_{x \in \mathcal{A}}, \boldsymbol{\pi}\}$, where \mathcal{S} is the set $\{\sigma_0, \sigma_1, \dots\}$ of the process' causal states, \mathcal{A} is the set of output symbols x , the set of matrices $\{T^{(x)} : T_{i,j}^{(x)} = \Pr(\sigma_j, x | \sigma_i)\}_{x \in \mathcal{A}}$ are the labeled transition matrices, and $\boldsymbol{\pi}$ is the stationary distribution over states.

The probability that a word $w = x_0, x_1, \dots, x_{L-1}$ is generated by an ϵ -machine is given in terms of the labeled transition matrices and the initial state distribution:

$$\Pr(w) = \boldsymbol{\pi} \prod_{i=0}^{L-1} T^{(x_i)} \mathbf{1} ,$$

where $\mathbf{1} = [1, \dots, 1]^\top$. These probabilities are constructed to agree with those of the words in a given process language.

The ensemble temporal evolution of internal state probability $\boldsymbol{\mu} = (\mu_0, \dots, \mu_{|\mathcal{S}|-1})$, with $\mu_i = \Pr(\sigma_i)$, is given by:

$$\boldsymbol{\mu}(t+1) = \boldsymbol{\mu}(t)T ,$$

where the transition matrix T is the sum over all output symbols:

$$T \equiv \sum_{x \in \mathcal{A}} T^{(x)} .$$

Transition probabilities are normalized. That is, the transition matrix T is *row-stochastic*:

$$\sum_{j=1}^{|\mathcal{S}|} T_{i,j} = \sum_{j=1}^{|\mathcal{S}|} \sum_{x \in \mathcal{A}} \Pr(\sigma_j, x | \sigma_i) = 1 .$$

Its component matrices $T_{ij}^{(x)}$ are said to be *substochastic*. Under suitable conditions on the transition matrix, $\lim_{t \rightarrow \infty} \boldsymbol{\mu}(t) = \boldsymbol{\pi}$.

Unifilarity, a property derived from the ϵ -machine equivalence relation [55], means that for each state σ_i , each symbol x may lead to at most one successor state σ_j [4]. In terms of the labeled transition matrices, for each row i and each symbol x the row $T_{ij}^{(x)}$ has at most one nonzero entry. We also will have occasion to speak of a *counifilar* HMM, which is the analogous requirement of unique labeling on transitions coming *into* each state.

One of the most important informational properties of a process, directly calculable from its ϵ -machine, is its statistical complexity C_μ [55]. Used in a variety of contexts, it quantifies the size of a process' minimal description.

Definition 7. *A process' statistical complexity C_μ is the Shannon entropy of the stationary*

distribution over its causal states:

$$\begin{aligned} C_\mu &= H[\boldsymbol{\pi}] \\ &= - \sum_{i=1}^{|\mathcal{S}|} \pi_i \log \pi_i . \end{aligned}$$

The statistical complexity has several operational meanings. For example, it is the average amount of information one gains upon learning a process' current causal state. It is also the minimal amount of information about the past that must be stored to predict the future as well as could be predicted if the entire past were stored. Most pertinent to our purposes, though, it also quantifies the communication cost of synchronizing two predicting agents through a classical channel [157].

4.3.2 q-Machine

The *q-machine* is a quantum representation of a classical stochastic process. Introduced in Ref. [157], it offers the largest reduction in state-complexity known so far among quantum models capable of generating classical processes.

A process' q-machine is constructed by first selecting a *codeword length* L . The q-machine (at L) consists of a set $\{|\eta_i(L)\rangle\}_{i=1}^{|\mathcal{S}|}$ of pure quantum *signal states* that are in one-to-one correspondence with the classical causal states $\sigma_i \in \mathcal{S}$. Each signal state $|\eta_i(L)\rangle$ encodes the set of length- L words $\{w : \Pr(w|\sigma_i) > 0\}$ that may follow causal state σ_i , as well as the corresponding conditional probability:

$$|\eta_i(L)\rangle \equiv \sum_{w \in \mathcal{A}^L} \sum_{\sigma_j \in \mathcal{S}} \sqrt{\Pr(w, \sigma_j | \sigma_i)} |w\rangle |\sigma_j\rangle , \quad (4.1)$$

where $\{|w\rangle\}_{w \in \mathcal{A}^L}$ denotes an orthonormal basis in the “word” Hilbert space with one dimension for each possible word w of length L . Similarly, $\{|\sigma_j\rangle\}_{j=1}^{|\mathcal{S}|}$ denotes an orthonormal basis in the “state” Hilbert space with one dimension for each classical causal state. The ensemble of length- L quantum signal states is then described by the density matrix:

$$\rho(L) = \sum_{i=1}^{|\mathcal{S}|} \pi_i |\eta_i(L)\rangle \langle \eta_i(L)| . \quad (4.2)$$

The ensemble's von Neumann entropy is defined in terms of its density matrix: $S(\rho) = -\text{tr}[\rho \log(\rho)]$, where $\text{tr}[\cdot]$ is the trace of its argument. Paralleling the classical statistical complexity, the quantity:

$$\begin{aligned} C_q(L) &\equiv S(\rho(L)) \\ &= -\text{tr}[\rho(L) \log(\rho(L))] , \end{aligned} \tag{4.3}$$

has the analogous operational meaning of the communication cost to send signal states over a quantum channel. Von Neumann entropy decreases with increasing signal-state overlap. It is generically smaller than the classical cost [157]: $C_q(L) \leq C_\mu$. In fact, $C_\mu = C_q$ if and only if the process' ϵ -machine is counifilar—there are no states with (at least) two similarly labeled incoming edges [102]. Notably, as we increase state number, processes with counifilar ϵ -machines represent a vanishing proportion of all possible processes [120]. The consequence is that almost all classical processes can be more compactly represented using quantum mechanics. This presents an opportunity to use quantum encoding to more efficiently represent processes.

Quantifying a process' quantum-reduced state-complexity via the von Neumann entropy of Eq. (4.3) is rooted in the existence of optimal quantum compression algorithms, such as Schumacher compression [213]. The advantage of smaller state-complexity with larger L , though, is not a consequence of the well developed theory of quantum compression. Rather it derives from carefully harnessing a model's coincident predictions by constructing a process' nonorthogonal quantum signal states. This is a new kind of quantum information processing. Notably, this quantum reduction—of requisite state-memory in the simulation of a classical stochastic process—was recently experimentally verified using non-orthogonal photon polarization signal states [185], though only for codeword-length $L = 1$. Upon both technological and theoretical advancements, the significant reduction in memory requirements quantified by $C_q(L)$ should enable efficient simulation of important complex systems whose dynamics were previously prohibitively memory intensive.

Calculating a process' quantum cost function $C_q(L)$ is challenging, however. The following shows how to circumvent the difficulties. Beyond practical calculational concerns, the theory leads to a deeper appreciation of quantum structural complexity.

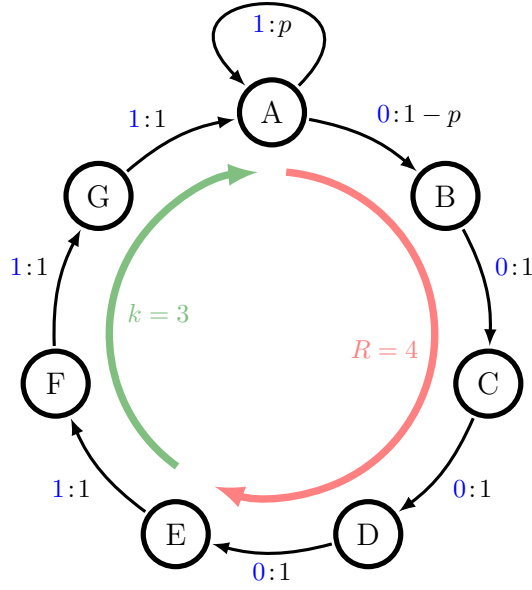


Figure 4.1: ϵ -Machine for the (4-3)-Golden Mean Process: The cycle’s red segment indicates the “Markov” portion and the green, the “cryptic” portion. The length scales R and k are tuned by changing the lengths of these two components, respectively. Edges labeled $x : p$ denote taking the state-to-state transition with probability p while emitting symbol $x \in \mathcal{A}$.

4.4 Quantum Overlaps

Reference [157] showed that the reduction $C_\mu - C_q(L)$ in state-complexity is determined by quantum overlaps between signal states in the q-machine. Accordingly, calculation of these overlaps is a primary task. Intuitively, nonorthogonal signal states correspond to causal states that yield “similar” predictions, in a sense to be explained. More rigorously, the overlap between nonorthogonal signal states is determined by words whose causal-state paths merge.

To illustrate, we compute several overlaps for the $(R-k)$ -Golden Mean Process, showing how they depend on L . (See Fig. 4.1 for its ϵ -machine state-transition diagram.) This process was designed to have a tuneable Markov order R and cryptic order k ; here, we choose $R = 4$ and $k = 3$. (Refer to Ref. [120] for more on this process and a detailed discussion of Markov and cryptic orders.)

At length $L = 0$, each signal state is simply the basis state corresponding to its causal

state: $|\eta_i(0)\rangle = |\sigma_i\rangle$. Since the ϵ -machine is minimal, there are no overlaps in the state vectors.

At length $L = 1$ codewords, we find the first nontrivial overlap. This corresponds to paths $A \xrightarrow{1} A$ and $G \xrightarrow{1} A$ merging at state A and we have:

$$\begin{aligned} |\eta_A(1)\rangle &= \sqrt{p}|1A\rangle + \sqrt{1-p}|0B\rangle \quad \text{and} \\ |\eta_G(1)\rangle &= |1A\rangle . \end{aligned}$$

This yields the overlap:

$$\langle \eta_A(1) | \eta_G(1) \rangle = \sqrt{p} .$$

Going on to length $L = 2$ codewords, more overlaps arise from mergings of more state paths. The three quantum signal states:

$$\begin{aligned} |\eta_A(2)\rangle &= p|11A\rangle + \sqrt{p(1-p)}|10B\rangle + \sqrt{(1-p)}|00C\rangle , \\ |\eta_F(2)\rangle &= |11A\rangle , \quad \text{and} \\ |\eta_G(2)\rangle &= \sqrt{p}|11A\rangle + \sqrt{1-p}|10B\rangle , \end{aligned}$$

interact to yield the overlaps:

$$\begin{aligned} \langle \eta_A(2) | \eta_F(2) \rangle &= p , \\ \langle \eta_F(2) | \eta_G(2) \rangle &= \sqrt{p} , \quad \text{and} \\ \langle \eta_A(2) | \eta_G(2) \rangle &= p\sqrt{p} + (1-p)\sqrt{p} = \sqrt{p} . \end{aligned}$$

The overlaps between (A, F) and (F, G) are new. The (A, G) overlap has the same value as that for (F, G) , however its calculation at $L = 2$ involved two terms instead of one. This is because no new merger has occurred; the $L = 1$ merger, affected by symbol 1, was simply propagated forward along two different state paths having prefix 1. There are two redundant paths: $A \xrightarrow{10} B$ overlaps $G \xrightarrow{10} B$ and $A \xrightarrow{11} A$ overlaps $G \xrightarrow{11} A$. A naive calculation of overlaps must contend with this type of redundancy.

4.5 Quantum Pairwise-Merger Machine

To calculate signal-state overlaps, we introduce the quantum pairwise-merger machine, a transient graph structure that efficiently encapsulates the organization of state paths. As we saw in the example, calculation of overlaps amounts to tracking state-path mergers. It is important that we do this in a systematic manner to avoid redundancies. The new machine does just this.

We begin by first constructing the *pairwise-merger machine* (PMM), previously introduced to compute overlaps [157]. There, probabilities were computed for each word found by scanning through the PMM. This method significantly reduced the number of words from the typically exponentially large number in a process' language and also gave a stopping criterion for PMMs with cycles. This was a vast improvement over naive constructions of the signal-state ensemble (just illustrated) and over von Neumann entropy calculation via diagonalization of an ever-growing Hilbert space.

Appropriately weighting PMM transitions yields the *quantum PMM* (QPMM), which then not only captures which states merge given which words, but also the contribution each merger makes to a quantum overlap. The QPMM has one obvious advantage over the PMM. The *particular* word that produces an overlap is ultimately unimportant; only the *amount* of overlap generated is important. Therefore, summing over symbols in the QPMM to obtain its internal state transitions removes this combinatorial factor. There are additional significant advantages to this matrix-based approach. Appreciating this requires more development.

To build the QPMM from a given process' ϵ -machine:

1. Construct the set of (unordered) pairs of (distinct) ϵ -machine states: (σ_i, σ_j) . We call these "pair-states". To this set, add a special state called SINK (short for "sink of synchronization") which is the terminal state.
2. For each pair-state (σ_i, σ_j) and each symbol $x \in \mathcal{A}$, there are three cases to address:
 - (a) If at least one of the two ϵ -machine states σ_i or σ_j has no outgoing transition on symbol x , then do nothing.

- (b) If both ϵ -machine states σ_i and σ_j have a transition on symbol x to the same state σ_m , then connect pair-state (σ_i, σ_j) to SINK with an edge labeled x . This represents a merger.
- (c) If both ϵ -machine states σ_i and σ_j have a transition on symbol x to two distinct ϵ -machine states σ_m and σ_n where $m \neq n$, then connect pair-state (σ_i, σ_j) to pair-state (σ_m, σ_n) with an edge labeled x . (There are no further restrictions on m and n .)

3. Remove all edges that are not part of a path that leads to SINK.

4. Remove all pair-states that do not have a path to SINK.

This is the PMM. Now, add information about transition probabilities to this topological structure to obtain the QPMM:

- 5. For each pair-state (σ_i, σ_j) in the PMM, add to each outgoing edge the weight $\sqrt{\Pr(x|\sigma_i) \Pr(x|\sigma_j)}$, where x is the symbol associated with that edge. Note that two states in QPMM may be connected with multiple edges (for different symbols).

Returning to our example, Fig. 4.2 gives the QPMM for the (4-3)-Golden Mean Process. Using it, we can easily determine the length at which a contribution is made to a given overlap. We consider codeword lengths $L = 1, 2, \dots$ by walking up the QPMM from SINK. For example, pair (A, G) receives a contribution of \sqrt{p} at $L = 1$. Furthermore, (A, G) receives no additional contributions at larger L . Pairs (A, F) and (F, G) , though, receive contributions $p = \sqrt{p} \times \sqrt{p}$ and $\sqrt{p} = \sqrt{p} \times 1$ at $L = 2$, respectively.

The QPMM is *not* a HMM, since the edge weights do not yield a stochastic matrix. However, like a HMM, we can consider its labeled “transition” matrices $\{\zeta^{(x)}\}$, $x \in \mathcal{A}$. Just as for their classical ϵ -machine counterparts, we index these matrices such that $\zeta_{u,v}^{(x)}$ indicates the edge going from pair-state u to pair-state v . Since the overlap contribution, and not the inducing word, is of interest, the important object is simply the resulting state-to-state substochastic matrix $\zeta = \sum_{x \in \mathcal{A}} \zeta^{(x)}$. The matrix ζ is the heart of our closed-form expressions for quantum coding costs, which follow shortly. As we noted above,

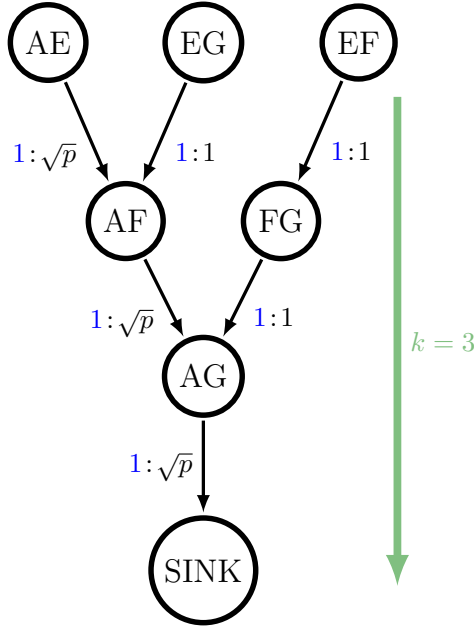


Figure 4.2: QPMM for the (4-3)-Golden Mean Process. Its depth is related to the cryptic order k .

it is this step that greatly reduces the combinatorial growth of paths that would otherwise make the calculations unwieldy.

To be explicit, our (4-3)-Golden Mean Process has:

$$\zeta = \begin{matrix} & \begin{matrix} AE & EG & EF & AF & FG & AG & SINK \end{matrix} \\ \begin{matrix} AE \\ EG \\ EF \\ AF \\ FG \\ AG \\ SINK \end{matrix} & \begin{pmatrix} 0 & 0 & 0 & \sqrt{p} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \sqrt{p} & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \sqrt{p} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \end{matrix}.$$

4.6 Overlaps from the QPMM

As we saw in the example, overlaps accumulate contributions as “probability amplitude” is pushed through the QPMM down to the SINK. Each successive overlap augmentation

can thus be expressed in terms of the next iterate of ζ :

$$\begin{aligned}\langle \eta_i(L) | \eta_j(L) \rangle &= \langle \eta_i(L-1) | \eta_j(L-1) \rangle \\ &= \langle (\sigma_i, \sigma_j) | \zeta^L | \text{SINK} \rangle .\end{aligned}$$

The general expression for quantum overlaps follows immediately:

$$\langle \eta_i(L) | \eta_j(L) \rangle = \langle (\sigma_i, \sigma_j) | \sum_{n=0}^L \zeta^n | \text{SINK} \rangle , \quad (4.4)$$

which is true for all processes by design of the QPMM. This form makes clear the cumulative nature of quantum overlaps and the fact that overlap contributions are not labeled.

Note that there are two trivial overlap types. Self-overlaps are always 1; this follows from Eq. (4.4) since $\langle (\sigma_i, \sigma_i) | = \langle \text{SINK} |$. Overlaps with no corresponding pair-state in the QPMM are defined to be zero for all L .

Now, we show that there are two behaviors that contribute to overlaps: a finite-horizon component and an infinite-horizon component. Some processes have only one type or the other, while many have both. We start with the familiar $(R-k)$ -GM, which has only finite-horizon contributions.

4.6.1 Finite Horizon: $(R-k)$ -Golden Mean Process

Overlap matrices are Hermitian, positive-semidefinite matrices and can therefore be represented as the product $A_L A_L^\dagger$. Let's use the general expression Eq. (4.4) to compute the matrix elements $(A_L A_L^\dagger)_{i,j} = \langle \eta_i(L) | \eta_j(L) \rangle$ for lengths $L = 1, 2, 3, 4$ for the $(R-k)$ -Golden Mean Process. We highlight in blue (and bold) the matrix elements that have changed from the previous length. All overlaps begin with the identity matrix, here I_7 as we have seven states in the ϵ -machine (Fig. 4.1). Then, at $L = 1$ we have one overlap.

The overlap matrix, with elements $\langle \eta_i(1) | \eta_j(1) \rangle$, is:

$$A_1 A_1^\dagger = \begin{matrix} & \begin{matrix} A & B & C & D & E & F & G \end{matrix} \\ \begin{matrix} A \\ B \\ C \\ D \\ E \\ F \\ G \end{matrix} & \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & \sqrt{p} \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ \sqrt{p} & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{matrix}.$$

Next, for $L = 2$ we find two new overlaps. The overlap matrix, with elements $\langle \eta_i(2) | \eta_j(2) \rangle$, is:

$$A_2 A_2^\dagger = \begin{matrix} & \begin{matrix} A & B & C & D & E & F & G \end{matrix} \\ \begin{matrix} A \\ B \\ C \\ D \\ E \\ F \\ G \end{matrix} & \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & p & \sqrt{p} \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ p & 0 & 0 & 0 & 0 & 1 & \sqrt{p} \\ \sqrt{p} & 0 & 0 & 0 & 0 & \sqrt{p} & 1 \end{pmatrix} \end{matrix}.$$

For $L = 3$, there are three new overlaps. The overlap matrix, with elements $\langle \eta_i(3) | \eta_j(3) \rangle$, is:

$$A_3 A_3^\dagger = \begin{matrix} & \begin{matrix} A & B & C & D & E & F & G \end{matrix} \\ \begin{matrix} A \\ B \\ C \\ D \\ E \\ F \\ G \end{matrix} & \begin{pmatrix} 1 & 0 & 0 & 0 & \sqrt{p^3} & p & \sqrt{p} \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ \sqrt{p^3} & 0 & 0 & 0 & 1 & \sqrt{p} & p \\ p & 0 & 0 & 0 & \sqrt{p} & 1 & \sqrt{p} \\ \sqrt{p} & 0 & 0 & 0 & p & \sqrt{p} & 1 \end{pmatrix} \end{matrix} .$$

Finally, for $L = 4$, we find the same matrix as $L = 3$: $\langle \eta_i(4) | \eta_j(4) \rangle = \langle \eta_i(3) | \eta_j(3) \rangle$ for all i and j . And, in fact, this is true for all $L \geq 3$. Therefore, all overlap information has been uncovered at codeword length $L = 3$.

Looking at the QPMM in Fig. 4.2, we recognize that the saturation of the overlap matrix corresponds to the finite *depth* d of the directed graph—the longest state-path through the QPMM that ends in the SINK state. Equivalently, the depth corresponds to the nilpotency of ζ :

$$d = \min\{n \in \mathbb{N} : \zeta^n = 0\} . \quad (4.5)$$

Note that the $(4 - 3)$ -Golden Mean Process QPMM is a tree of depth 4.

Whenever the QPMM is a tree or, more generally, a directed-acyclic graph (DAG), the overlaps will similarly have a finite-length horizon equal to the depth d . The nilpotency of ζ for finite-depth DAGs allows for a truncated form of the general overlap expression Eq. (4.4):

$$\langle \eta_i(L) | \eta_j(L) \rangle = \langle (\sigma_i, \sigma_j) | \sum_{n=0}^{\min(L, d-1)} \zeta^n | \text{SINK} \rangle . \quad (4.6)$$

This form is clearly advantageous for any process whose QPMM is a finite DAG. Naturally then, we are led to ask: What property of a process leads to a finite DAG? To answer this question, we reconsider how overlap is accumulated via the merging of state-paths.

Paths through the QPMM represent causal-state-path mergers. To make this more precise, we introduce the concept of an L -merge, which is most intuitively understood through Fig. 4.3:

Definition 8. An L -merge consists of a length- L word w and two state paths each of length $L + 1$ that each allow the word w ending in the same state F . We denote the word $w = (x_0, \dots, x_{L-1})$ and state paths (a_0, \dots, a_{L-1}, F) and (b_0, \dots, b_{L-1}, F) where states $a_i \neq b_i$, for all $i \in \{0, \dots, L - 1\}$ and, trivially, $F = F$, the final state in which the paths end.

Immediately, we see that every labeled path of length- L through the QPMM that ends in SINK is precisely an L -merge.

Such causal-state-path merging not only contributes to quantum overlap, but also contributes to a process' crypticity. Let \mathcal{S}_L denote the random variable for the particular causal state $\sigma \in \mathcal{S}$ at time L . Then the *crypticity* of a process—the average uncertainty about the present causal state \mathcal{S}_0 given perfect observation of the entire infinite future $x_{0:\infty}$, but not knowing the history of observations prior to the present moment—can be written as $H(\mathcal{S}_0|X_{0:\infty})$, which is accumulated at all lengths up to the *cryptic order* [59].

Definition 9. A process' cryptic order k is the minimum length L for which $H(\mathcal{S}_L|X_{0:\infty}) = 0$.

That is, given knowledge of the entire infinite future of observations, the cryptic order quantifies how far back into the past one must remember to always know the present causal state.

By way of comparison, a process' *Markov order* is:

$$R = \min\{L : H(\mathcal{S}_L|X_{0:L}) = 0\} .$$

That is, given knowledge (e.g., the ϵ -machine) of which process is being observed but without knowing future observations, the Markov order quantifies how far back into the past one must remember to always know the present causal state. A more familiar length-scale characterizing historical dependence, R depends on both path merging and path

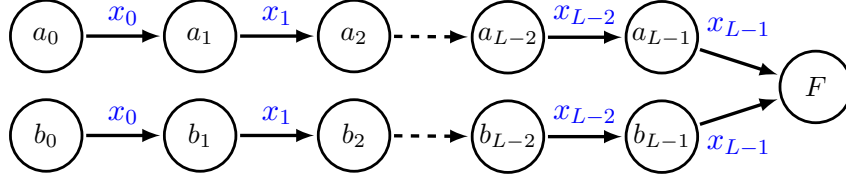


Figure 4.3: L -merge: Two causal-state paths— (a_0, \dots, a_{L-1}, F) and (b_0, \dots, b_{L-1}, F) where states $a_i \neq b_i$, for all $i \in \{0, \dots, L-1\}$ —generate the same word $w = x_0 x_1 \dots x_{L-1}$ and merge only on the last output symbol x_{L-1} into a common final state F .

termination due to disallowed transitions. The cryptic order, in contrast, effectively ignores the termination events and is therefore upper-bounded by the Markov order: $k \leq R$. This bound is also easy to see given the extra conditional variable $X_{L:\infty}$ in the definition of crypticity ($X_{0:\infty} = X_{0:L} X_{L:\infty}$) [158, 160].

The following lemma states a helpful relation between cryptic order and L -merges.

Lemma 1. *Given an ϵ -machine with cryptic order k : for $L \leq k$, there exists an L -merge; for $L > k$, there exists no L -merge.*

Proof. *See App. 4.10.2.*

Each L -merge corresponds with a real, positive contribution to some quantum overlap. By Lemma 1, for a cryptic-order k process there is at least one L -merge at each length $L \in \{1, \dots, k\}$ and none beyond k . Therefore, at least one overlap receives a real, positive contribution at each length up until k , where there are no further contributions. This leads to our result for overlap accumulation and saturation in terms of the cryptic order.

Theorem 2. *Given a process with cryptic order k , for each $L \in \{0, \dots, k\}$, each quantum overlap is a nondecreasing function of L :*

$$\langle \eta_i(L+1) | \eta_j(L+1) \rangle \geq \langle \eta_i(L) | \eta_j(L) \rangle \quad .$$

Furthermore, for each $L \in \{1, \dots, k\}$, there exists at least one overlap that is increased as a result of a corresponding L -merge. For all remaining $L \geq k$, each overlap takes the constant value $\langle \eta_i(k) | \eta_j(k) \rangle$.

Proof. *See App. 4.10.2.*

Evidently, the cryptic order is an important length scale not only for classical processes, but also when building efficient quantum encoders.

As an important corollary, this theorem also establishes the relation between a process' cryptic order and the depth of its QPMM:

$$d = k + 1 . \quad (4.7)$$

Thus, we have discovered that the process property corresponding to a finite DAG QPMM is finite cryptic order. Moreover, the cryptic order corresponds to a topological feature of the QPMM, the depth d , responsible for saturation of the overlaps.

This leads to rephrasing the truncated form of the overlaps sum in Eq. (4.4):

$$\langle \eta_i(L) | \eta_j(L) \rangle = \langle (\sigma_i, \sigma_j) | \sum_{n=0}^{\min(L,k)} \zeta^n | \text{SINK} \rangle . \quad (4.8)$$

This form is advantageous for any process that is finite cryptic order. This, of course, includes all finite Markov-order processes—processes used quite commonly in a variety of disciplines.

Since the quantum-reduced state-complexity $C_q(L)$ is a function of only π and quantum overlaps, the preceding development also gives a direct lesson about the $C_q(L)$ saturation.

Corollary 1. *$C_q(L)$ has constant value $C_q(k)$ for $L \geq k$.*

Proof. *The entropy of an ensemble of pure signal states $\{p_i, |\psi_i\rangle\}$ is a function of only probabilities p_i and overlaps $\{\langle \psi_i | \psi_j \rangle\}$. The result then follows directly from Thm. 2.*

Having established connections among depth, cryptic order, and saturation, we seem to be done analyzing quantum overlap—at least for the finite-cryptic case. To prepare for going beyond finite horizons, however, we should reflect on the spectral origin of ζ 's nilpotency.

A nilpotent matrix, such as ζ in the finite-cryptic case, has only the eigenvalue zero. This can perhaps be most easily seen if the pair-states are ordered according to their distance from SINK, so that ζ is triangular with only zeros along the diagonal.

Notably, for finite DAGs with depth $d > 1$, the standard eigenvalue–eigenvector decomposition is insufficient to form a complete basis—the corresponding ζ is necessarily

nondiagonalizable due to the geometric multiplicity of the zero eigenvalue being less than its algebraic multiplicity. *Generalized eigenvectors* must be invoked to form a complete basis [90]. Intuitively, this type of nondiagonalizability can be understood as the intrinsic interdependence among pair-states in propagating probability amplitude through a branch of the DAG. When ζ is rendered into Jordan block form via a similarity transformation, the size of the largest Jordan block associated with the zero eigenvalue is called the *index* ν_0 of the zero eigenvalue. It turns out to be equal to the depth for finite DAGs.

Summarizing, the finite-horizon case is characterized by several related features: (i) the QPMM is a DAG (of finite depth), (ii) the depth of the QPMM is one greater than the cryptic order, (iii) the matrix ζ has only the eigenvalue zero, and (iv) the depth is equal to the index of this zero-eigenvalue, meaning that ζ has at least k generalized eigenvectors. More generally, ζ can have nonzero eigenvalues and this corresponds to richer structure that we explore next.

4.6.2 Infinite Horizon: Lollipop Process

Now we ask, what happens when the QPMM is not a directed acyclic graph? That is, what happens when it contains *cycles*?

It is clear that the depth d diverges, implying that the cryptic order is infinite. Therefore, the sum in Eq. (4.4) may no longer be truncated. We also know that infinite-cryptic processes become ubiquitous as ϵ -machine state size increases [120]. Have we lost our calculational efficiencies? No, in fact, there are greater advantages yet to be gained.

We first observe that a QPMM's ζ breaks into two pieces. One has a finite horizon reminiscent of the finite cryptic order just analyzed, and the other has an infinite horizon, but is, as we now show, analytically quite tractable.

In general, a linear operator A may be decomposed using the *Dunford decomposition* [70] (also known as the Jordan–Chevalley decomposition) into:

$$A = \mathcal{D} + \mathcal{N} , \tag{4.9}$$

where \mathcal{D} is diagonalizable, \mathcal{N} is nilpotent, and \mathcal{D} and \mathcal{N} commute. In the current setting, \mathcal{N} makes the familiar finite-horizon contribution, whereas the new \mathcal{D} term has an infinite

horizon: $\mathcal{D}^n \neq 0$, for all $n < \infty$. In the context of infinite cryptic processes, the finite horizon associated with \mathcal{N} is no longer simply related to QPMM depth nor, therefore, the cryptic order which is infinite.

The systematic way to address the new diagonalizable part is via a spectral decomposition [60], where the persistent leaky features of the QPMM state probability evolution are understood as independently acting modes. It is clear that ζ always has a nilpotent component associated with a zero eigenvalue, due to the SINK state. Assuming that the remaining eigenspaces are diagonalizable, the form of the overlaps becomes:

$$\begin{aligned} \langle \eta_i(L) | \eta_j(L) \rangle = & \sum_{\xi \in \Lambda_\zeta \setminus \{0\}} \frac{1 - \xi^{L+1}}{1 - \xi} \langle (\sigma_i, \sigma_j) | \zeta_\xi | \text{SINK} \rangle \\ & + \sum_{m=0}^{\min\{L, \nu_0-1\}} \langle (\sigma_i, \sigma_j) | \zeta^m \zeta_0 | \text{SINK} \rangle, \end{aligned} \quad (4.10)$$

where Λ_ζ is the set of ζ 's eigenvalues, ζ_ξ are the projection operators corresponding to each eigenvalue, and ν_0 is the index of the zero eigenvalue, which is the size of its largest Jordan block. We refer to this as the *almost-diagonalizable* case since all eigenspaces—besides possibly the zero-eigenvalue space—are diagonalizable. This case covers all processes with generic parameters. Here, ν_0 is still responsible for the length of the finite-horizon component, but is no longer directly related to QPMM depth or process cryptic order.

Note that in the finite-cryptic order case, the only projector ζ_0 is necessarily the identity. Therefore, Eq. (4.10) reduces to the previous form in Eq. (4.8).

The spectral decomposition yields a new level of tractability for the infinite-cryptic case. The infinite-horizon piece makes contributions at all lengths, but in a regular way. This allows for direct calculation of its total contribution at any particular L , including $L \rightarrow \infty$.

To highlight this behavior, consider the (7–4)-Lollipop Process, whose ϵ -machine is shown in Fig. 4.4. It is named for the shape of its QPMM; see Fig. 4.5. This process is a simple example of one where the cryptic order is infinite and the finite-horizon length of the nilpotent contribution is tunable. Roughly speaking, the diagonalizable component comes from the “head” of the lollipop (the cycle), and the nilpotent part comes from the “stick”.

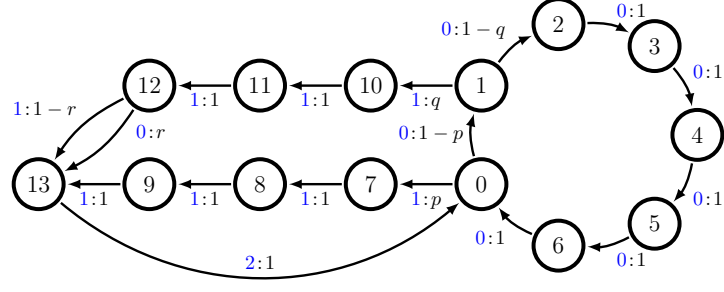


Figure 4.4: ϵ -Machine for the (7-4)-Lollipop Process. The cycle of 0s on the right leads to infinite Markov and cryptic orders.

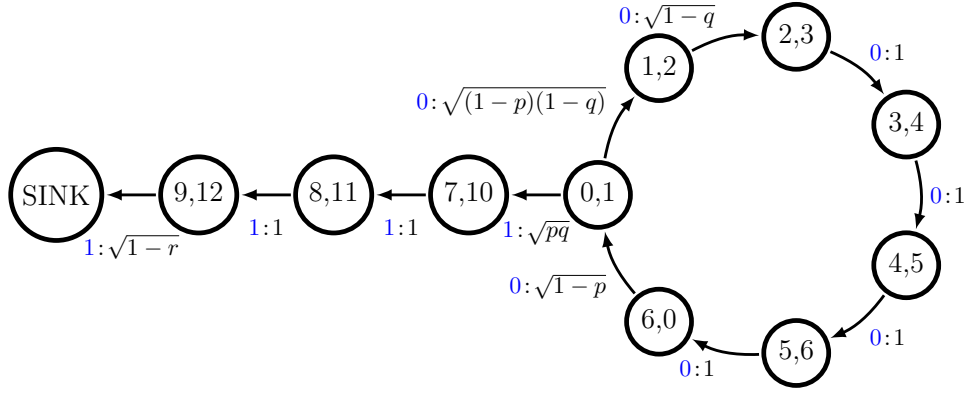


Figure 4.5: QPMM for the (7-4)-Lollipop Process.

It is straightforward to construct the general QPMM and thereby derive ζ for the $(N-M)$ -Lollipop Process. Its QPMM has N pair-states in a cyclic head. The M remaining pair-states constitute a finite-horizon ‘stick’. We find:

$$\det(\zeta - \xi I) = (-\xi)^M [(-\xi)^N - (1-p)(1-q)] ,$$

yielding:

$$\Lambda_\zeta = \left\{ 0, [(1-p)(1-q)]^{1/N} e^{in2\pi/N} \right\}_{n=0}^{N-1} , \quad (4.11)$$

with $\nu_0 = M$.

For concreteness, consider the (7-4)-Lollipop Process with transition parameters $p = q = 1/2$ and $r \in (0, 1)$. It has eigenvalues $\Lambda_\zeta = \{0, ae^{in\theta}\}$ and $\nu_0 = 4$, where $a = (1/4)^{1/7}$, $\theta = 2\pi/7$, and $n \in \{0, 1, 2, 3, 4, 5, 6\}$.

Each $\xi = ae^{in\theta}$ eigenvalue has algebraic multiplicity 1 and associated left eigenvector:

$$\begin{aligned} \langle \xi | = [2\sqrt{2}\xi^6, \sqrt{2}\xi^5, \xi^4, \xi^3, \xi^2, \xi^1, \xi^0, \\ \sqrt{2}\xi^5, \sqrt{2}\xi^4, \sqrt{2}\xi^3, \sqrt{2(1-r)}\xi^2] , \end{aligned}$$

and right eigenvector:

$$|\xi\rangle = [\frac{1}{2\xi}, 1, \sqrt{2}\xi, \sqrt{2}\xi^2, \sqrt{2}\xi^3, \sqrt{2}\xi^4, \sqrt{2}\xi^5, 0, 0, 0, 0]^\top .$$

The general relationship among left and right eigenvectors, left and right generalized eigenvectors, and projection operators, and their reduction in special cases is discussed in Ref. [198]. In the present case, notice that, since ζ is not a normal operator, the right eigenvectors are not simply the conjugate transpose of their left counterparts. (*Normal operators* by definition commute with their conjugate transpose; e.g., Hermitian operators.) The left and right eigenvectors are fundamentally different, with the differences expressing the QPMM's directed causal architecture.

Since each of these eigenvalues has algebraic multiplicity 1, the corresponding projection operators are defined in terms of right and left eigenvectors:

$$\zeta_\xi = \frac{|\xi\rangle \langle \xi|}{\langle \xi | \xi \rangle} .$$

The zero eigenvalue has algebraic multiplicity $\nu_0 = 4$ and geometric multiplicity 1, meaning that while there is only one eigenvector there are three generalized eigenvectors. The left and right eigenvectors are:

$$\begin{aligned} [0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1] \text{ and} \\ [0, 1, 0, 0, 0, 0, 0, 0, -1, 0, 0, 0]^\top . \end{aligned}$$

The three generalized left eigenvectors are:

$$\begin{aligned} [0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0] , \\ [0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0] , \text{ and} \\ [0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0] ; \end{aligned}$$

and the three generalized right eigenvectors are:

$$\begin{aligned} & [0, 0, \sqrt{2}, 0, 0, 0, 0, 0, -1, 0, 0]^\top, \\ & [0, 0, 0, \sqrt{2}, 0, 0, 0, 0, 0, -1, 0]^\top, \text{ and} \\ & [0, 0, 0, 0, \sqrt{2(1-r)}, 0, 0, 0, 0, 0, -1]^\top. \end{aligned}$$

Since the index of the zero eigenvalue is larger than 1 ($\nu_0 = 4$), the projection operator ζ_0 for the zero eigenvalue includes the contributions from both its standard and generalized eigenvectors:

$$\zeta_0 = \sum_{n=0}^3 \frac{|0_n\rangle \langle 0_{3-n}|}{\langle 0_{3-n}|0_n\rangle}, \quad (4.12)$$

where $|0_0\rangle$ and $\langle 0_0|$ are the standard right and left eigenvector respectively and $|0_n\rangle$ and $\langle 0_n|$ are the n^{th} generalized right and left eigenvector for $n \geq 1$. More generally, when the geometric multiplicity is greater than one, this sum goes over all standard and all generalized eigenvectors of the zero eigenvalue, matching pairs according to the orthonormal relations discussed in the previous chapters.

Since all projection operators must sum to the identity, the projection operator for the zero eigenvalue can be obtained alternatively from:

$$\zeta_0 = I - \sum_{\xi \in \Lambda_\zeta \setminus 0} \zeta_\xi, \quad (4.13)$$

which is often useful during calculations.

This very efficient procedure allows us to easily probe the form of quantum advantage for any process described by a finite ϵ -machine.

Finally, we jump directly to the asymptotic overlap using the following expression:

$$\begin{aligned} \langle \eta_i(\infty) | \eta_j(\infty) \rangle &= \langle (\sigma_i, \sigma_j) | \left(\sum_{n=0}^{\infty} \zeta^n \right) | \text{SINK} \rangle \\ &= \langle (\sigma_i, \sigma_j) | (I - \zeta)^{-1} | \text{SINK} \rangle. \end{aligned} \quad (4.14)$$

Note that $I - \zeta$ is invertible, since ζ is substochastic. Hence, its spectral radius is less than unity, $1 \notin \Lambda_\zeta$, and so $\det(1I - \zeta) \neq 0$. Moreover, $(I - \zeta)^{-1}$ is *equal* to the convergent Neumann series $\sum_{n=0}^{\infty} \zeta^n$ by Thm. 3 of Ref. [281, Ch. VIII § 2].

Yielding an important calculational efficiency, the form of Eq. (4.14) does not require spectral decomposition of ζ and so immediately provides the asymptotic quantum-reduction of state complexity. Finally, this form does not depend on the previous assumption of ζ being almost-diagonalizable.

4.7 Quantum-Reduced State-Complexity

The preceding development focused on computing overlaps between quantum signal states for q-machine representations of a given process. Let’s not forget that the original goal was to compute the von Neumann entropy of this ensemble—the quantum-reduced state-complexity $C_q(L)$, which is the memory that must be transferred about the state of the process to synchronize compatible predictions.

The naive approach to calculating $C_q(L)$ constructs the signal states directly and so does not make use of overlap computation. This involves working with a Hilbert space of increasing dimension, exponential in codeword length L . This quickly becomes intractable, for all but the simplest processes.

The second approach, introduced in Ref. [157], made use of the PMM to compute overlaps. These overlaps were then used to construct a density operator with those same overlaps, but in a Hilbert space of fixed size $|\mathcal{S}|$, essentially obviating the high-dimensional embedding of the naive approach. And, we just showed how to calculate overlaps in closed form. The elements of the resulting density matrix, however, are nonlinear functions of the overlaps. Besides the computational burden this entails, it makes it difficult to use the overlap matrix to theoretically infer much about the general behavior of $C_q(L)$.

Here, we present two markedly improved approaches that circumvent these barriers. We are ultimately interested in the von Neumann entropy which depends only on the spectrum of the density operator. It has been pointed out that the Gram matrix of an ensemble shares the same spectrum [122]. The *Gram matrix* for our ensemble of pure

quantum signal states is:

$$G = \begin{bmatrix} \sqrt{\pi_1 \pi_1} \langle \eta_1 | \eta_1 \rangle & \cdots & \sqrt{\pi_1 \pi_{|\mathcal{S}|}} \langle \eta_1 | \eta_{|\mathcal{S}|} \rangle \\ \vdots & \ddots & \vdots \\ \sqrt{\pi_{|\mathcal{S}|} \pi_1} \langle \eta_{|\mathcal{S}|} | \eta_1 \rangle & \cdots & \sqrt{\pi_1 \pi_{|\mathcal{S}|}} \langle \eta_{|\mathcal{S}|} | \eta_{|\mathcal{S}|} \rangle \end{bmatrix}. \quad (4.15)$$

If we define $D_{\boldsymbol{\pi}} \equiv \text{diag}(\boldsymbol{\pi})$, then $G = D_{\boldsymbol{\pi}}^{1/2} A A^\dagger D_{\boldsymbol{\pi}}^{1/2}$.

Given that it is only a small step from the overlap matrix $A A^\dagger$ to the Gram matrix G , we see the usefulness of the thoroughgoing overlap analysis above. The spectrum of G is then computed using standard methods, either symbolically or numerically.

There is another surrogate matrix that shares the spectrum but is simpler, yet again, for some calculations. We call this matrix \tilde{G} the *left-consolidated Gram matrix*:

$$\tilde{G} = \begin{bmatrix} \pi_1 \langle \eta_1 | \eta_1 \rangle & \cdots & \pi_1 \langle \eta_1 | \eta_{|\mathcal{S}|} \rangle \\ \vdots & \ddots & \vdots \\ \pi_{|\mathcal{S}|} \langle \eta_{|\mathcal{S}|} | \eta_1 \rangle & \cdots & \pi_{|\mathcal{S}|} \langle \eta_{|\mathcal{S}|} | \eta_{|\mathcal{S}|} \rangle \end{bmatrix}. \quad (4.16)$$

Note that $\tilde{G} = D_{\boldsymbol{\pi}} A A^\dagger$ —i.e., $D_{\boldsymbol{\pi}}$ has been consolidated on the left. A right-consolidated Gram matrix would work just as well for the calculation of $C_q(L)$.

Since the spectra are identical, we can calculate $C_q(L)$ directly from the density matrix $\rho(L)$, Gram matrix $G^{(L)}$, or consolidated Gram matrix $\tilde{G}^{(L)}$:

$$\begin{aligned} C_q(L) &= - \sum_{\lambda \in \Lambda_{\rho(L)}} \lambda \log \lambda \\ &= - \sum_{\lambda \in \Lambda_{G^{(L)}}} \lambda \log \lambda \\ &= - \sum_{\lambda \in \Lambda_{\tilde{G}^{(L)}}} \lambda \log \lambda. \end{aligned}$$

For further discussion, see App. 4.10.3.

Using the Gram matrix as described, we illustrate the behavior of $C_q(L)$ for the $(R-k)$ -Golden Mean (Fig. 4.6) and $(N-M)$ -Lollipop (Fig. 4.7). For each of the two process classes, we compute several instances by varying R and k and by varying N and M while holding fixed their transition parameters. Comparing the two figures, we qualitatively confirm

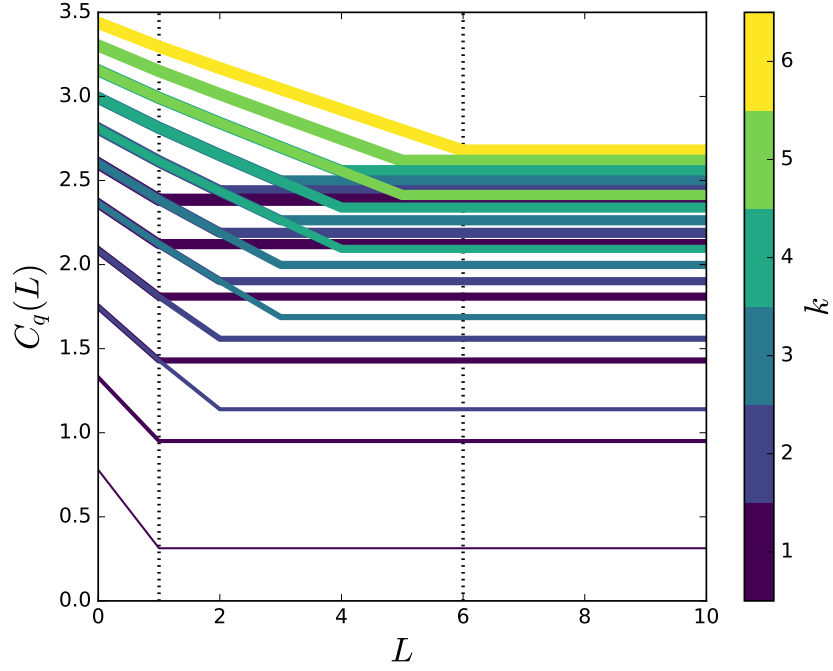


Figure 4.6: Quantum costs $C_q(L)$ for the $(R-k)$ -Golden Mean Process with $R \in \{1, \dots, 6\}$ and $k \in \{1, \dots, R\}$. R and k are indicated with line width and color, respectively. The probability of the self-loop is $p = 0.7$. $C_q(L)$ roughly linearly decreases until $L = k$ where it is then constant. Note that $(R-k)$ -GM agrees exactly with $((R+1)-(k-1))$ -GM for $L \leq k$, as explained in App. 4.10.4.

the difference between a process with only a finite-horizon contribution and one with an infinite-horizon contribution. The $(R-k)$ -Golden Mean reaches its encoding saturation at $L = k$ the cryptic order. The $(N-M)$ -Lollipop only approaches this limit asymptotically.

In contrast to the customary approach in quantum compression [213], in which an entire message is to be compressed with perfect fidelity, the compression advantage here is obtained by throwing away information that is not relevant for simulating a process—with the goal of correctly sampling from a conditional future distribution.

Recall that the quantum-reduced state-complexity $C_q(L)$ quantifies a communication cost. Specifically, it is the amount of memory about a process' state that must be queried to move the system forward in time. However, to avoid misinterpretation, we note that this cost does not have a simple relationship to the “quantum communication cost” as the phrase is sometimes used in the distributed computing setting of communication

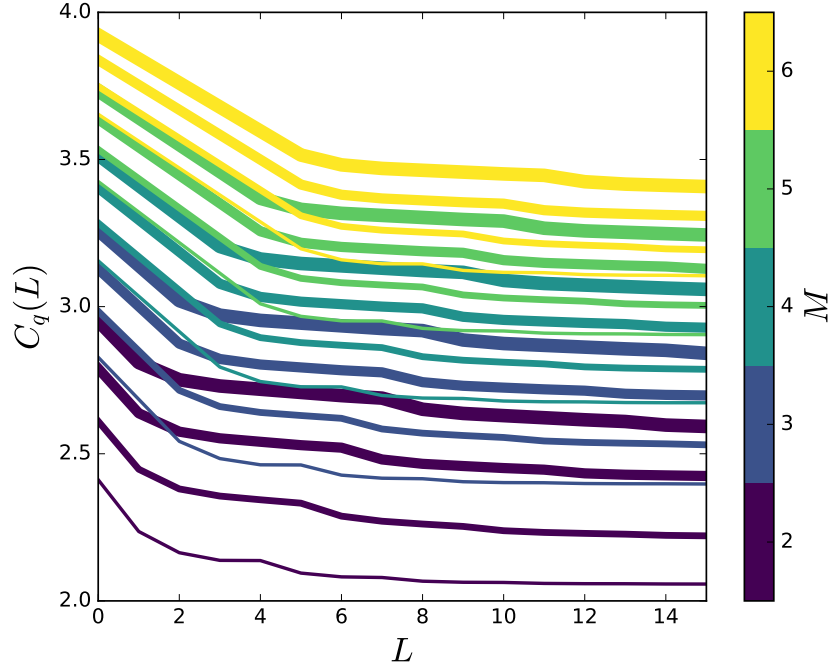


Figure 4.7: Quantum costs $C_q(L)$ for the Lollipop process for $N \in \{3, 4, 5, 6\}$, $M \in \{2, 3, 4, 5, 6\}$, $p = q = 0.5$, and $r = 0.1$. N and M are indicated with line width and color, respectively. After a fast initial decrease, these curves approach their asymptotic values more slowly.

complexity theory [33].

To supplement the details already given, annotated analytic derivations of several example processes are given in App. 4.10.4. These examples serve as a pedagogical resource, with comparison and discussion of various analytical techniques.

4.8 Costs using Long Codewords

The preceding discussed quantum-state overlaps extensively. We found that the behavior of the overlaps with L is completely described through ζ 's spectral decomposition. And, we showed that, for any L , the von Neumann entropy $C_q(L)$ can be found from the eigenvalues of the Gram matrix—a direct transformation of the overlap matrix. This is all well and good, and key progress. But, can we use this machinery to directly analyze the behavior of $C_q(L)$ as a function of L ? For infinite-cryptic processes, the answer is an especially pleasing affirmative.

This section derives $C_q(L)$'s asymptotic behavior for large L ; viz., $\nu_0 < L \leq k = \infty$. We show that a periodic pattern, exponentially decaying at the rate of the largest ζ -eigenvalue magnitude, dominates $C_q(L)$'s deviation from $C_q(\infty)$ for large L . In particular, we show two things: First, the asymptotic behavior of $C_q(L) - C_q(\infty)$ is, to first order, exponentially decreasing as r_1^L , where r_1 is ζ 's spectral radius. Second, this exponential defines an envelope for a Ψ -periodic asymptotic structure, where Ψ is the least common multiple of slowest-decaying QPMM cycle lengths.

Recall that the minimal known upper bound on state complexity is given by the asymptotic von Neumann entropy:

$$C_q(\infty) = - \sum_{\lambda^{(\infty)} \in \Lambda_{G^{(\infty)}}} \lambda^{(\infty)} \log(\lambda^{(\infty)}) .$$

We will show that when L is large, $(\delta G)^{(L)} \equiv G^{(L)} - G^{(\infty)}$ can be treated as a perturbation to $G^{(\infty)}$. From the corresponding small variations $\{(\delta\lambda)^{(L)}\}_{\lambda \in \Lambda_G}$, direct calculation of the first differential yields the approximate change in the von Neumann entropy:

$$(\delta S)^{(L)} = - \sum_{\lambda \in \Lambda_G} [\log(\lambda^{(\infty)}) + 1] (\delta\lambda)^{(L)} , \quad (4.17)$$

so long as no zero eigenvalues of $G^{(\infty)}$ prematurely vanish at finite L . Our task, therefore, is to find $(\delta\lambda)^{(L)}$ from $(\delta G)^{(L)}$ in terms of ζ 's spectral properties.

For easy reference, we first highlight our notation:

- $G^{(L)}$ is a Gram matrix at length L corresponding to $\rho(L)$.
- $\lambda^{(L)} \in \Lambda_{G^{(L)}}$ is any one of its eigenvalues.
- $|\lambda^{(L)}\rangle$ and $\langle\lambda^{(L)}|$ are the right and left eigenvectors of $G^{(L)}$ corresponding to $\lambda^{(L)}$, respectively.
- $(\delta G)^{(L)} \equiv G^{(L)} - G^{(\infty)}$ is the perturbation to $G^{(\infty)}$ investigated here.
- $\xi \in \Lambda_\zeta$ is an eigenvalue of the QPMM transition dynamic ζ .

If using G 's symmetric version, the right and left eigenvectors are simply transposes of each other: $\langle\lambda^{(L)}| = (|\lambda^{(L)}\rangle)^\top$. For simplicity of the proofs, we assume nondegeneracy of $G^{(L)}$'s eigenvalues, so that the projection operator associated with $\lambda^{(L)}$ is

$|\lambda^{(L)}\rangle \langle \lambda^{(L)}| / \langle \lambda^{(L)} | \lambda^{(L)} \rangle$, where the denominator assures normalization. Nevertheless, the eigenbasis of $G^{(L)}$ is always complete and the final result, Thm. 4, retains general validity.

Here, we show that the matrix elements of $(\delta G)^{(L)}$ are arbitrarily small for large enough L , such that first-order perturbation is appropriate for large L , and give the exact form of $(\delta G)^{(L)}$ for use in the calculation of $(\delta \lambda)^{(L)}$.

Proposition 1. *For $L \geq \nu_0$, the exact change in Gram matrix is:*

$$(\delta G)^{(L)} = - \sum_{\xi \in \Lambda_\zeta \setminus 0} \frac{\xi^{L+1}}{1-\xi} C_\xi ,$$

where C_ξ is independent of L and has matrix elements:

$$(C_\xi)_{i,j} = \sqrt{\pi_i \pi_j} \langle (\sigma_i, \sigma_j) | \zeta_\xi | \text{SINK} \rangle .$$

Proof. *We calculate:*

$$\begin{aligned} (\delta G)_{i,j}^{(L)} &= G_{i,j}^{(L)} - G_{i,j}^{(\infty)} \\ &= \sqrt{\pi_i \pi_j} \left(\langle \eta_i^{(L)} | \eta_j^{(L)} \rangle - \langle \eta_i^{(\infty)} | \eta_j^{(\infty)} \rangle \right) \\ &= -\sqrt{\pi_i \pi_j} \langle (\sigma_i, \sigma_j) | \zeta^{L+1} (1 - \zeta)^{-1} | \text{SINK} \rangle . \end{aligned}$$

If we assume that all nonzero eigenvalues of ζ correspond to diagonalizable subspaces, then for $L \geq \nu_0$, the elements of $(\delta G)^{(L)}$ have the spectral decomposition:

$$(\delta G)_{i,j}^{(L)} = - \sum_{\xi \in \Lambda_\zeta \setminus 0} \frac{\xi^{L+1}}{1-\xi} \sqrt{\pi_i \pi_j} \langle (\sigma_i, \sigma_j) | \zeta_\xi | \text{SINK} \rangle .$$

Since this decomposition is common to all matrix elements, we can factor out the $\left\{ \frac{\xi^{L+1}}{1-\xi} \right\}_\xi$, leaving the L -independent set of matrices:

$$\left\{ C_\xi : (C_\xi)_{i,j} = \sqrt{\pi_i \pi_j} \langle (\sigma_i, \sigma_j) | \zeta_\xi | \text{SINK} \rangle \right\}_{\xi \in \Lambda_\zeta} ,$$

such that:

$$(\delta G)^{(L)} = - \sum_{\xi \in \Lambda_\zeta \setminus 0} \frac{\xi^{L+1}}{1-\xi} C_\xi .$$

Proposition 2. *At large L , the first-order correction to $\lambda^{(\infty)}$ is:*

$$(\delta\lambda)^{(L)} = - \sum_{\xi \in \Lambda_\zeta \setminus 0} \frac{\xi^{L+1}}{1-\xi} \frac{\langle \lambda^{(\infty)} | C_\xi | \lambda^{(\infty)} \rangle}{\langle \lambda^{(\infty)} | \lambda^{(\infty)} \rangle} . \quad (4.18)$$

Proof. *Perturbing $G^{(\infty)}$ to $G^{(\infty)} + (\delta G)^{(L)}$, the first-order change in its eigenvalues is given by:*

$$(\delta\lambda)^{(L)} = \frac{\langle \lambda^{(\infty)} | (\delta G)^{(L)} | \lambda^{(\infty)} \rangle}{\langle \lambda^{(\infty)} | \lambda^{(\infty)} \rangle} , \quad (4.19)$$

which is standard first-order nondegenerate perturbation theory familiar in quantum mechanics, with the allowance for unnormalized bras and kets. Proposition 2 then follows directly from Eq. (4.19) and Prop. 1.

Theorem 3. *At large L , such that $\nu_0 < L \leq k = \infty$, the first-order correction to $C_q(\infty)$ is:*

$$\begin{aligned} C_q(L) - C_q(\infty) &\approx (\delta S)^{(L)} \\ &= \sum_{\xi \in \Lambda_\zeta \setminus 0} \frac{\xi^{L+1}}{1-\xi} \sum_{\lambda^{(\infty)} \in \Lambda_{G^{(\infty)}}} \langle C_\xi \rangle [\log(\lambda^{(\infty)}) + 1] , \end{aligned} \quad (4.20)$$

where:

$$\langle C_\xi \rangle \equiv \frac{\langle \lambda^{(\infty)} | C_\xi | \lambda^{(\infty)} \rangle}{\langle \lambda^{(\infty)} | \lambda^{(\infty)} \rangle} .$$

Proof. *This follows directly from Eq. (4.17) and Prop. 2.*

The large- L behavior of $C_q(L) - C_q(\infty)$ is a sum of decaying complex exponentials. And, to first order, we can even calculate the coefficient of each of these contributions.

Notice that the only L -dependence in Prop. 2 and Thm. 3 comes in the form of exponentiating eigenvalues of the QPMM transition dynamic ζ . For very large L , the dominant structure implied by Prop. 2 and Thm. 3 can be teased out by looking at the relative contributions from ζ 's first- and second-largest magnitude sets of eigenvalues.

Let r_1 be the spectral radius of ζ , shared by the largest eigenvalues $\Lambda(r_1)$: $r_1 \equiv \max\{|\xi| : \xi \in \Lambda_\zeta\}$. And, let $\Lambda(r_1) \equiv \operatorname{argmax}\{|\xi| : \xi \in \Lambda_\zeta\}$. Then, let r_2 be the second-largest

magnitude of all of ζ 's eigenvalues that differs from r_1 : $r_2 \equiv \max\{|\xi| : \xi \in \Lambda_\zeta \setminus \Lambda(r_1)\}$. And, let $\Lambda(r_2) \equiv \operatorname{argmax}\{|\xi| : \xi \in \Lambda_\zeta \setminus \Lambda(r_1)\}$. Multiple eigenvalues can belong to $\Lambda(r_1)$. Similarly, multiple eigenvalues can belong to $\Lambda(r_2)$.

Then, $0 \leq (r_2/r_1) < 1$, if ζ has at least one nonzero eigenvalue. This is the case of interest here since we are addressing those infinite-horizon processes with $k = \infty > \nu_0$. Hence, as L becomes large, $(r_2/r_1)^L$ vanishes exponentially if it is not already zero. This leads to a corollary of Prop. 2.

Corollary 2. *For $L \geq \nu_0$, the leading deviation from $\lambda^{(\infty)}$ is:*

$$(\delta\lambda)^{(L)} = -r_1^{L+1} \sum_{\xi \in \Lambda(r_1)} \frac{(\xi/|\xi|)^{L+1}}{1 - \xi} \langle C_\xi \rangle \left[1 + O\left(\left(\frac{r_2}{r_1}\right)^L\right) \right] .$$

Notice that $\xi/|\xi|$ lies on the unit circle in the complex plane. Due to their origin in cyclic graph structure, we expect each $\xi \in \Lambda(r_1)$ to have a phase in the complex plane that is a rational fraction of 2π . Hence, there is some n for which $(\xi/|\xi|)^n = 1$, for all $\xi \in \Lambda(r_1)$. The minimal such n , call it Ψ , will be of special importance:

$$\Psi \equiv \min\{n \in \mathbb{N} : (\xi/|\xi|)^n = 1 \text{ for all } \xi \in \Lambda(r_1)\} . \quad (4.21)$$

Since all $\xi \in \Lambda(r_1)$ originate from cycles in ζ 's graph, we have the result that Ψ is equal to the least common multiple of the cycle lengths implicated in $\Lambda(r_1)$.

For example, if all $\xi \in \Lambda(r_1)$ come from the same cycle in the graph of ζ , then $\Psi = |\Lambda(r_1)|$ and:

$$\Lambda(r_1) = \left\{ \xi_m = r_1 e^{im2\pi/|\Lambda(r_1)|} \right\}_{m=1}^{|\Lambda(r_1)|} .$$

That is, $\{\xi_m/|\xi_m|\}_{m=1}^{|\Lambda(r_1)|}$ are the $|\Lambda(r_1)|^{\text{th}}$ roots of unity, uniformly distributed along the unit circle. If, however, $\Lambda(r_1)$ comes from multiple cycles in ζ 's graph, then the least common multiple of the cycle lengths should be used in place of $|\Lambda(r_1)|$.

Recognizing the Ψ -periodic structure of $(\xi/|\xi|)^n$ yields a more informative corollary of Prop. 2:

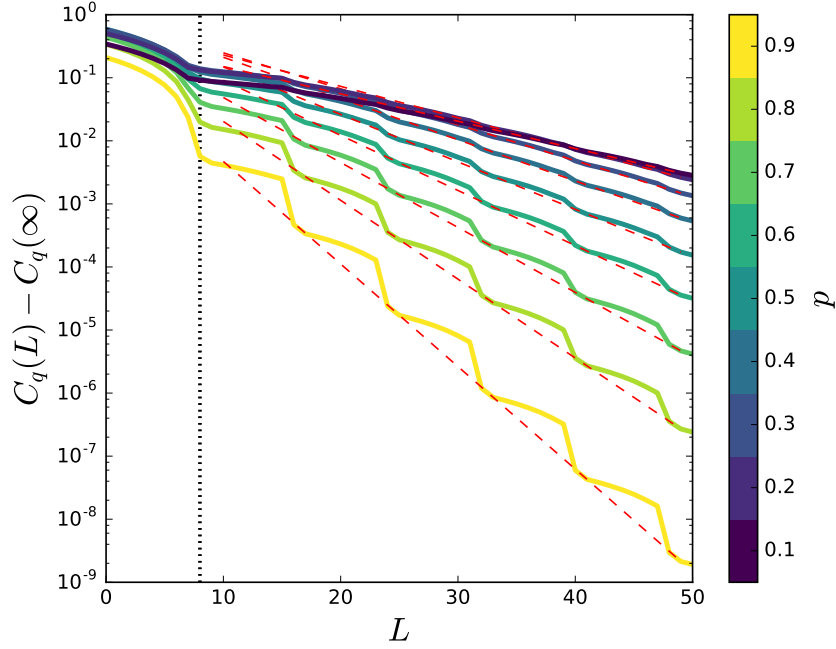


Figure 4.8: (8,8)-Lollipop with transition parameters $p \in [0.1, 0.9]$, $q = 0.5$, and $r = 0.1$. $C_q(L) - C_q(\infty)$ on semilog plot illustrates asymptotically exponential behavior. Red dashed lines, r_1^L where r_1 (no relation to r) is the spectral radius of ζ , quantify the exponential rate of decay. The height of each red dashed line is set equal to $C_q(49)$; we can see that the decay is very close to exponential even as early as $L \simeq 15$. Vertical dashed line at $L = M = 8$ shows change in behavior after the length of the “stick”.

Corollary 3. For $L \geq \nu_0$, the leading deviation from $\lambda^{(\infty)}$ is:

$$(\delta\lambda)^{(L)} = -r_1^{L+1} \sum_{\xi \in \Lambda(r_1)} \frac{(\xi/|\xi|)^{\text{mod}(L+1, \Psi)}}{1 - \xi} \langle C_\xi \rangle \times \left[1 + O\left((r_2/r_1)^L\right) \right] .$$

Hence:

$$(\delta\lambda)^{(L+\Psi)} \approx r_1^\Psi (\delta\lambda)^{(L)} . \quad (4.22)$$

We conclude that asymptotically a pattern—of changes in the density-matrix eigenvalues (with period Ψ)—decays exponentially with decay rate of r_1^Ψ per period. There are immediate implications for the pattern of asymptotic changes in $C_q(L)$ at large L .

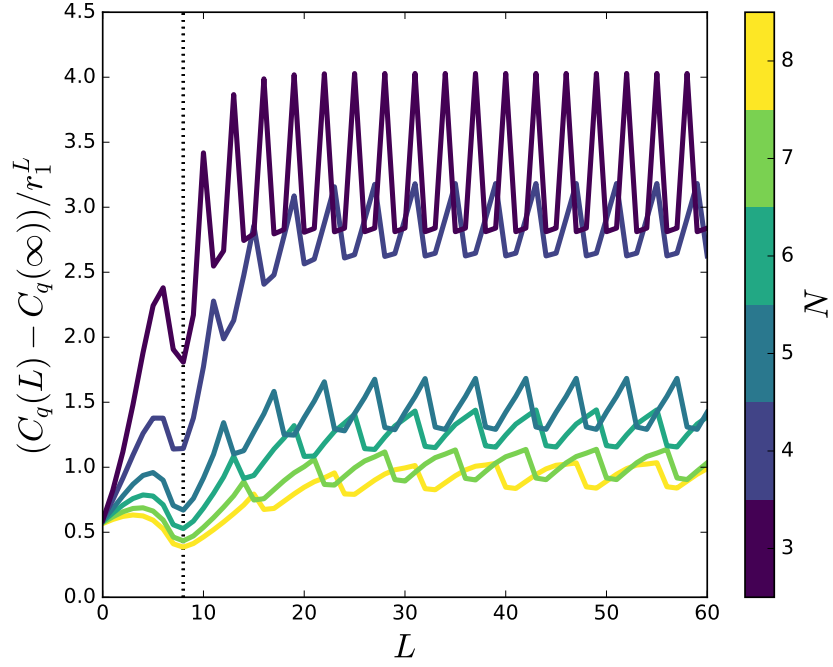


Figure 4.9: Lollipop with $N \in \{3, 4, 5, 6, 7, 8\}$ and $M = 8$, and transition parameters $p = q = 0.5$ and $r = 0.1$. $(C_q(L) - C_q(\infty))/r_1^L$ demonstrates the periodicity of asymptotic behavior. Removing the exponential envelope makes periodicity of the remaining deviation more apparent. For Lollipop, the periodicity $\Psi = |\Lambda(r_1)| = N$.

Corollary 4. *For $L \geq \nu_0$, the leading deviation from $C_q(\infty)$ is:*

$$\begin{aligned}
 C_q(L) - C_q(\infty) &\approx (\delta S)^{(L)} \\
 &= r_1^{L+1} \sum_{\xi \in \Lambda(r_1)} \frac{(\xi/|\xi|)^{\text{mod}(L+1, \Psi)}}{1 - \xi} \\
 &\quad \times \sum_{\lambda^{(\infty)} \in \Lambda_{G^{(\infty)}}} \langle C_\xi \rangle \log(\lambda^{(\infty)}) \left[1 + O\left((r_2/r_1)^L\right) \right].
 \end{aligned}$$

The most profound implication of this detailed analysis can be summarized succinctly.

Theorem 4. *For sufficiently large L :*

$$\frac{C_q(L + \Psi) - C_q(\infty)}{C_q(L) - C_q(\infty)} \approx r_1^\Psi. \quad (4.23)$$

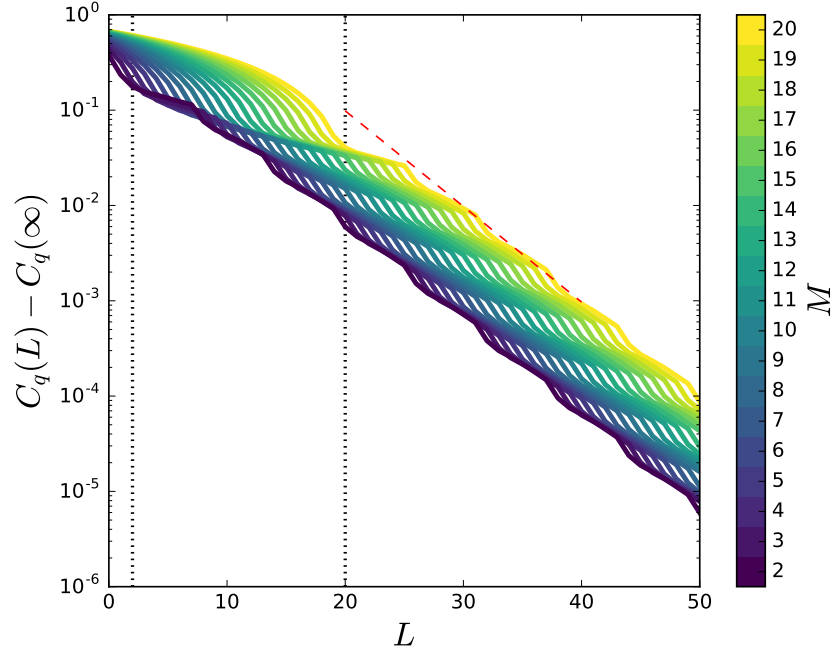


Figure 4.10: $C_q(L) - C_q(\infty)$ on a semilog plot for Lollipop with $N = 6$ and $M \in \{2, \dots, 20\}$ and transition parameters $p = q = 0.5$ and $r = 0.1$. M determines the finite-horizon length, where the nilpotent part of ζ vanishes. Vertical (dashed) lines indicate $L = 2$ and $L = 20$, the shortest and longest such length in this group.

That is, asymptotically a pattern—of changes in $C_q(L) - C_q(\infty)$ (with period Ψ)—decays exponentially with decay rate of r_1^Ψ per period ¹.

While the first-order perturbation allowed us to identify both the roles and values of r_1 and Ψ for any process and Coro. 4 would imply Thm. 4, Thm. 4 actually transcends the limitations of the first-order approximation.

Proof. Expanding $\log G^{(L)}$ in powers of $(G^{(L)} - I)$, then multiplying by $-G^{(L)}$, shows that $C_q(L) = -\text{tr}[G^{(L)} \log G^{(L)}]$ can be written as:

$$C_q(L) = - \sum_{n=0}^{\infty} a_n \text{tr}[(G^{(L)})^n] , \quad (4.24)$$

¹In principle, we need to consider two cases: the pattern decays to $C_q(\infty)$ from above or from below. In either case, the decay of the $C_q(L) - C_q(\infty)$ pattern is exponential. However, it is known that $C_q(L)$ is strictly less than $C_\mu = C_q(0)$ for any L for any noncounifilar process (and equal otherwise). Hence, we expect that $C_q(L)$ always decays from above, as corroborated by extensive numerical exploration.

for proper $a_n \in \mathbb{R}$. Using:

$$G^{(L)} = \sum_{\xi \in \Lambda_\zeta \setminus 0} \frac{1 - \xi^{L+1}}{1 - \xi} C_\xi + \sum_{m=0}^{\min\{L, \nu_0-1\}} C_{0,m} , \quad (4.25)$$

with appropriate constant matrices $C_{0,m}$, together with Eqs. (4.21) and (4.24), yields Thm. 4 with general validity.

In the simplest case, when ζ has only one largest eigenvalue, then $\Psi = |\Lambda(r_1)| = 1$ and so $C_q(L) - C_q(\infty)$ is dominated by a simple exponential decay at large L .

For the case of multiple largest eigenvalues originating from the same cycle in ζ 's graph, then $\Psi = |\Lambda(r_1)| > 1$. And so, the asymptotic behavior of $C_q(L) - C_q(\infty)$ is dominated by a decaying pattern of length $|\Lambda(r_1)|$.

For example, the Lollipop processes have an exponentially decaying pattern of length- N that dominates $C_q(L) - C_q(\infty)$ for $L > \nu_0 = M$:

$$\Psi = |\Lambda(r_1)| = N . \quad (4.26)$$

This periodic behavior is apparent in the semi-log plots of Figs. 4.8 and 4.10 and is especially emphasized in Fig. 4.9 which shows that $\Psi = N$ for various N . The figures demonstrate excellent agreement with our qualitative expectations from the above approximations.

Showing the effect of different ν_0 , Fig. 4.10 emphasizes that the initial rolloff of $C_q(L) - C_q(\infty)$ is due to $L \leq \nu_0 = M$. The dominant asymptotic behavior is reached soon after $L = \nu_0$ in this case since the remaining (i.e., nonzero) eigenvalues of the QPMM transition dynamic ζ are all in the largest-magnitude set $\Lambda(r_1)$. In other words, Thm. 3's Eq. (4.20) is not only approximated by but, in this case, also equal to the simpler expression in Coro. 4, since $r_2 = 0$.

The slope r_1 indicated in Figs. 4.8 and 4.10 corresponds to the asymptotic decay rate of the envelope for $C_q(L) - C_q(\infty)$. This asymptotic decay rate is a function of both N and p , since for Lollipop:

$$r_1 = [(1-p)(1-q)]^{1/N} . \quad (4.27)$$

Figure 4.8 shows that we have indeed identified the correct slope for different p .

The central asymptotic features of the quantum advantage $C_q(L) - C_q(\infty)$ of reduced state-complexity are all captured succinctly by Thm. 4: First, the asymptotic behavior of $C_q(L) - C_q(\infty)$ is exponentially decreasing at rate r_1 , which is the spectral radius of ζ . Second, this exponential envelope is modulated by an asymptotic Ψ -periodic structure, where Ψ is the least common multiple of slowest-decaying QPMM cycle-lengths.

These results summarize the expected behavior of the L -dependent quantum reduction of state-complexity for all classical processes that can be described by a finite-state ϵ -machine. Using codeword-length of at least the finite-horizon length ν_0 of the process' QPMM seems advisable for significant reduction of memory costs in simulations that utilize the advantage of quantum signal states discussed here. The cost-benefit analysis of further increasing encoding length for infinite-cryptic processes will be application-specific, but now has theoretical grounding in the above results.

4.9 Conclusion

We developed a detailed analytical theory of how to maximally reduce the state-complexity of a classical, stationary finite-memory stochastic process using a quantum channel. This required using the new quantum state-machine representation (q-machines) [157], carefully constructing its codewords and quantitatively monitoring their overlaps (via the quantum pairwise-merger machine), and utilizing a new matrix formulation of the overlap density matrix (consolidated Gram matrix). Applying spectral decomposition then lead directly to closed-form expressions for the quantum coding costs at any codeword length, including infinite length.

The theoretical advances give an efficient way to probe the behavior of quantum-reduced state complexity with increasing codeword length, both analytically and, when symbolic calculation become arduous, numerically. The efficient numerical algorithm (linear in L) improves on previous exponential algorithms; moreover, the infinite- L limit can now be obtained directly in finite time. Analyzing selected example processes illustrated the required calculations and also the range of phenomena that occur when compressing memoryful processes. We expect the results to be relevant for the understanding and

design of efficient simulations for complex classical stochastic systems of biological and technological importance newly enabled by the quantum reduction in memory requirements.

Particular phenomena we reported for the first time here included (i) details of how a process' cryptic order determines its quantum reduction in state complexity, (ii) transient and persistent contributions to reduced state-complexity, (iii) exponential convergence to optimum compression, and (iv) oscillations in the convergence that reveal how a process gives up its crypticity with increasing codeword length. Our results apply to both finite and infinite Markov- and cryptic-order processes.

The overall result appears as a rather complete quantitative toolkit for analyzing quantum state-compressibility of classical processes, including finite and infinite codeword closed-form expressions. That said, many issues remain, both technical and philosophical. We believe, however, that the approach's mathematical grounding and analytical and numerical efficiency will go some distance to solving them in the near future.

For example, one of the abiding questions is the meaning of process crypticity $\chi = C_\mu - \mathbf{E}$ —the difference between a process' predictable information or excess entropy \mathbf{E} and its stored state information or statistical complexity C_μ [58, 76]. Most directly, χ measures how much state information (C_μ) is hidden from observation (\mathbf{E}). Cryptic processes and even those with infinite cryptic order dominate the space of classical processes [120]. This means that generically we can compress C_μ down to $C_q(L)$. However, this begs the question of what crypticity is in the quantum domain. Now that we can work analytically in the infinite-length limit, we can explore the *quantum crypticity* $\chi_q = C_q(\infty) - \mathbf{E}$. From our studies, some not reported here, it appears that one cannot compress the state information all the way down to the excess entropy. Why? Why do not quantum models exist of “size” \mathbf{E} bits? Does this point to a future, even more parsimonious physical theory? Or, to a fundamental limitation of communication that even nature must endure, as it channels the past through the present to the future?

For another, are we really justified in comparing Shannon bits (C_μ) to qubits (C_q)? This is certainly not a new or recent puzzle. However, the results on compression bring it to the fore anew. And, whatever the outcome, the answer will change our view of

what physical pattern and structure are. Likely, the answer will have a profound effect. Assuming the comparison is valid, why is there a perceived level of classical reality that is more structurally complex when, as we demonstrated and now can calculate, processes might be more compactly represented quantum mechanically?

4.10 Appendices

4.10.1 Appendix A: Mealy HMMs

Edge-emitting HMMs are called Mealy HMMs. This should be contrasted with the state-emitting HMMs, called Moore HMMs. Mealy and Moore HMMs are different representations, but all processes that can be generated by finite-state Moore HMMs can also be generated by Mealy HMMs, and vice versa; i.e., Mealy and Moore HMMs are class equivalent. The causal equivalence relation, however, implies that the minimal classical model of a process—the ϵ -machine—is a Mealy HMM.

Definition 10. A Mealy HMM \mathcal{M} is the 4-tuple $\{\mathcal{R}, \mathcal{A}, \{T^{(x)}\}_{x \in \mathcal{A}}, \boldsymbol{\mu}_0\}$, where \mathcal{R} is the set $\{\rho_0, \rho_1, \dots\}$ of latent states, \mathcal{A} is the set of output symbols x , the set of matrices $\{T^{(x)} : T_{i,j}^{(x)} = \Pr(\rho_j, x | \rho_i)\}_{x \in \mathcal{A}}$ are the labeled transition matrices, and $\boldsymbol{\mu}_0$ is the initial distribution over latent states.

Certain non-stationary processes can be generated by Mealy HMMs when the initial distribution over latent states is not the stationary distribution. In this work, we consider only stationary processes, so $\boldsymbol{\mu}_0 = \boldsymbol{\pi}$. In such cases, time-independent word probabilities can be calculated as:

$$\Pr(w) = \boldsymbol{\pi} \prod_{i=0}^{L-1} T^{(x_i)} \mathbf{1} ,$$

where $\mathbf{1} = [1, \dots, 1]^\top$. When these probabilities are constructed to agree with those of the words in a given process language, the HMM is said to be a *presentation* of the process.

The ϵ -machine is the Mealy HMM presentation of a process, whose latent states are the causal states of the process: $\mathcal{R} = \mathcal{S}$. The ϵ -machine is provably the minimal classical unifilar generator of a process—minimal both in the number of states and the entropy over states [55].

4.10.2 Appendix B: Quantum Overlaps and Cryptic Order

Lemma 1. Given an ϵ -machine with cryptic order k : for $L \leq k$, there exists an L -merge; for $L > k$, there exists no L -merge.

Proof. By definition of cryptic order k :

$$H(\mathcal{S}_k | X_{0:\infty}) = 0 .$$

This means that for any given \vec{x} there exists a unique σ_k . Since k is the minimum such length, for $L = k - 1$ there exists some word $x_{0:\infty}$ that leaves uncertainty in causal state \mathcal{S}_{k-1} . Call two of these uncertain \mathcal{S}_{k-1} states A and B ($A \neq B$). Tracing $x_{0:\infty}$ backwards from A and B , we produce two state paths. These state paths must be distinct at each step due to ϵ -machine unifilarity. If they were not distinct at some step, they would remain so for all states going forward, particularly at \mathcal{S}_{k-1} . The next symbol x_k must take A and B to the same next state F or violate the assumption of cryptic order k . These two state paths and the word $x_{0:k}$ and the final state F make up a k -merger, meaning that cryptic order k implies the existence of a k -merger.

By removing states from the left side of this k -merger, it is easy to see that a k -merger implies the existence of all shorter L -mergers.

By unifilarity again, $H(\mathcal{S}_k | X_{0:\infty}) = 0 \rightarrow H(\mathcal{S}_L | X_{0:\infty}) = 0$, for all $L \geq k$. Assume there exists an L -merger for $L > k$ with word w . By definition of L -merger, there is then uncertainty in the state \mathcal{S}_{L-1} . This uncertainty exists for any word with w as the prefix—a set with nonzero probability. This contradicts the definition of cryptic order.

Theorem 2. Given a process with cryptic order k , for each $L \in \{0, \dots, k\}$, each quantum overlap $\langle \eta_i(L) | \eta_j(L) \rangle$ is a nondecreasing function of L . Furthermore, for each $L \in \{1, \dots, k\}$, there exists at least one overlap that is increased (as a result of a corresponding L -merge). For all remaining $L \geq k$, each overlap takes a constant value $\langle \eta_i(k) | \eta_j(k) \rangle$.

Proof. We directly calculate:

$$\begin{aligned}
\langle \eta_a(L) | \eta_b(L) \rangle &= \sum_{\substack{w, w' \in \mathcal{A}^L \\ j_L, l_L \in \{i\}_{i=1}^M}} \sqrt{T_{aj_L}^{(w)}} \sqrt{T_{bj_L}^{(w')}} \langle w | w' \rangle \langle \sigma_{l_L} | \sigma_{j_L} \rangle \\
&= \sum_{w, j_L} \sqrt{T_{aj_L}^{(w)}} \sqrt{T_{bj_L}^{(w)}} .
\end{aligned}$$

So, we have:

$$\begin{aligned}
&\langle \eta_a(L+1) | \eta_b(L+1) \rangle \\
&= \sum_{\substack{w' \in \mathcal{A}^{L+1} \\ j_{L+1}}} \sqrt{T_{aj_{L+1}}^{(w')}} \sqrt{T_{bj_{L+1}}^{(w')}} \\
&= \sum_{\substack{w \in \mathcal{A}^L, s \in \mathcal{A} \\ j_L, l_L, j_{L+1}}} \sqrt{T_{aj_n}^{(w)}} \sqrt{T_{j_n j_{L+1}}^{(s)}} \sqrt{T_{bl_L}^{(w)}} \sqrt{T_{l_L j_{L+1}}^{(s)}} \\
&= \sum_{\substack{w \in \mathcal{A}^L, s \in \mathcal{A} \\ j_L, j_{L+1}}} \sqrt{T_{aj_L}^{(w)}} \sqrt{T_{j_L j_{L+1}}^{(s)}} \sqrt{T_{bj_L}^{(w)}} \sqrt{T_{j_L j_{L+1}}^{(s)}} \\
&\quad + \sum_{\substack{w \in \mathcal{A}^L, s \in \mathcal{A} \\ j_L \neq l_L, j_{L+1}}} \sqrt{T_{aj_L}^{(w)}} \sqrt{T_{j_L j_{L+1}}^{(s)}} \sqrt{T_{bl_L}^{(w)}} \sqrt{T_{l_L j_{L+1}}^{(s)}} ,
\end{aligned}$$

The first sum represents the overlaps obtained already at length L . To see this, we split the sum to two parts, where the first contains:

$$\begin{aligned}
&\sum_{\substack{w \in \mathcal{A}^L, s \in \mathcal{A} \\ j_L, j_{L+1}}} \sqrt{T_{aj_L}^{(w)}} \sqrt{T_{j_L j_{L+1}}^{(s)}} \sqrt{T_{bj_L}^{(w)}} \sqrt{T_{j_L j_{L+1}}^{(s)}} \\
&= \sum_{\substack{w \in \mathcal{A}^L \\ j_L}} \sqrt{T_{aj_L}^{(w)}} \sqrt{T_{bj_L}^{(w)}} \left(\sum_{\substack{s \in \mathcal{A} \\ j_{L+1}}} \sqrt{T_{j_L j_{L+1}}^{(s)}} \sqrt{T_{j_L j_{L+1}}^{(s)}} \right) \\
&= \sum_{\substack{w \in \mathcal{A}^L \\ j_L}} \sqrt{T_{aj_L}^{(w)}} \sqrt{T_{bj_L}^{(w)}} \\
&= \langle \eta_a(L) | \eta_b(L) \rangle .
\end{aligned}$$

We use Lemma 1 to analyze the second sum, which represents the change in the overlaps,

finding that:

$$\sum_{\substack{w \in \mathcal{A}^L, s \in \mathcal{A} \\ j_L \neq l_L, j_{L+1}}} \sqrt{T_{aj_L}^{(w)}} \sqrt{T_{j_L j_{L+1}}^{(s)}} \sqrt{T_{bl_L}^{(w)}} \sqrt{T_{l_L j_{L+1}}^{(s)}} \geq 0 ,$$

with equality when $L \geq k$. Summarizing:

$$\langle \eta_a(L+1) | \eta_b(L+1) \rangle \geq \langle \eta_a(L) | \eta_b(L) \rangle ,$$

with equality for $L \geq k$.

Note that while the set of overlaps continues to be augmented at each length up until the cryptic order, we do not currently have a corresponding statement about the nontrivial change in $C_q(L)$ or its monotonicity. Although a proof has been elusive, it would be an important extension of our work. Nevertheless, the asymptotic analysis of Sec. 4.8 shows an overall decay of $C_q(L)$ for infinite cryptic processes. Moreover, extensive numerical exploration suggests that $C_q(L)$ is indeed monotonic at all scales for all orders of crypticity.

4.10.3 Appendix C: Matrices and Their Entropy

4.10.3.1 Density Matrix

The density matrix can now be expressed using a fixed $|\mathcal{S}|$ -by- $|\mathcal{S}|$ matrix, valid for all L .

Using the Gram-Schmidt procedure one can choose a new orthonormal basis. Let:

$$\begin{aligned} |\eta_1(L)\rangle &= |e_1^{(L)}\rangle \\ |\eta_2(L)\rangle &= a_{21}^{(L)} |e_1^{(L)}\rangle + a_{22}^{(L)} |e_2^{(L)}\rangle \\ |\eta_3(L)\rangle &= a_{31}^{(L)} |e_1^{(L)}\rangle + a_{32}^{(L)} |e_2^{(L)}\rangle + a_{33}^{(L)} |e_3^{(L)}\rangle \\ &\vdots \end{aligned}$$

and so on. Then:

$$\begin{aligned}
a_{21}^{(L)} &= \langle \eta_1(L) | \eta_2(L) \rangle \\
&= \langle (\sigma_1, \sigma_2) | \left(\sum_{n=0}^L \zeta^n \right) |\text{SINK}\rangle \ , \\
a_{22}^{(L)} &= \left(1 - | \langle \eta_1(L) | \eta_2(L) \rangle |^2 \right)^{1/2} \ , \\
a_{31}^{(L)} &= \langle \eta_1(L) | \eta_3(L) \rangle \\
&= \langle (\sigma_1, \sigma_3) | \left(\sum_{n=0}^L \zeta^n \right) |\text{SINK}\rangle \ ,
\end{aligned}$$

and so on. Now, it is useful to rewrite what we can in matrix form:

$$\begin{bmatrix} \langle \eta_1(L) | \\ \langle \eta_2(L) | \\ \langle \eta_3(L) | \\ \vdots \\ \langle \eta_{|\mathcal{S}|}(L) | \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & & & 0 \\ a_{21}^{(L)} & a_{22}^{(L)} & & \\ a_{31}^{(L)} & a_{32}^{(L)} & a_{33}^{(L)} & \\ \vdots & & & \ddots \\ a_{|\mathcal{S}|1}^{(L)} & \cdots & & a_{|\mathcal{S}||\mathcal{S}|}^{(L)} \end{bmatrix}}_{\equiv A_L} \begin{bmatrix} \langle e_1^{(L)} | \\ \langle e_2^{(L)} | \\ \langle e_3^{(L)} | \\ \vdots \\ \langle e_{|\mathcal{S}|}^{(L)} | \end{bmatrix} \ ,$$

which defines the lower-triangular matrix A_L . Note that the rightmost matrix of orthonormal basis vectors is simply the identity matrix, since we are working in that basis.

In this new basis, we construct the $|\mathcal{S}|$ -by- $|\mathcal{S}|$ density matrix as:

$$\begin{aligned}
\rho(L) &= \sum_{i=1}^{|\mathcal{S}|} \pi_i |\eta_i(L)\rangle \langle \eta_i(L)| \\
&= \begin{bmatrix} |\eta_1(L)\rangle & \cdots & |\eta_{|\mathcal{S}|}(L)\rangle \end{bmatrix} \underbrace{\begin{bmatrix} \pi_1 & & 0 \\ & \ddots & \\ 0 & & \pi_{|\mathcal{S}|} \end{bmatrix}}_{\equiv D_\pi} \begin{bmatrix} \langle \eta_1(L) | \\ \langle \eta_2(L) | \\ \langle \eta_3(L) | \\ \vdots \\ \langle \eta_{|\mathcal{S}|}(L) | \end{bmatrix} \\
&= A_L^\dagger D_\pi A_L \ .
\end{aligned}$$

Since all entries are real, the conjugate transpose is the transpose. This more general framework may be useful, however, if we want to consider the effect of adding phase to the quantum states.

4.10.3.2 Von Neumann Entropy

The quantum coding cost is:

$$\begin{aligned}
C_q(L) &= -\text{tr} [\rho(L) \log \rho(L)] \\
&= -\text{tr} \left[A_L^\dagger D_\pi A_L \log(A_L^\dagger D_\pi A_L) \right] \\
&= - \sum_{\lambda \in \Lambda_{A_L^\dagger D_\pi A_L}} \lambda \log \lambda .
\end{aligned}$$

This is relatively easy to evaluate since the density matrix $\rho(L)$ is only a $|\mathcal{S}|$ -by- $|\mathcal{S}|$ function of L . Thus, we calculate $C_q(L)$ analytically from ρ 's spectrum. This, in a curious way, was already folded into ζ 's spectrum.

4.10.3.3 Gram Matrix

The A_L matrix is burdensome due to nonlinear dependence on the overlap of the quantum states. We show how to avoid this nonlinearity and instead obtain the von Neumann entropy from a transformation that yields a linear relationship with overlaps.

The *Gram matrix*, with elements $G_{mn}^{(L)} = \sqrt{\pi_m \pi_n} \langle \eta_m(L) | \eta_n(L) \rangle$, can be used instead of $\rho(L)$ to evaluate the von Neumann entropy [122]. In particular, $G^{(L)}$ has the same spectrum as $\rho(L)$, even with the same multiplicities: $\Lambda_{G^{(L)}} = \Lambda_{\rho(L)}$, while a_λ , g_λ , and ν_λ remain unchanged for all λ in the spectrum. (This is a slightly stronger statement than in Ref. [122], but is justified since $\rho(L)$ and $G^{(L)}$ are both $|\mathcal{S}|$ -by- $|\mathcal{S}|$ dimensional.)

Here, we briefly explore the relationship between $\rho(L)$ and $G^{(L)}$ and, then, focus on the closed-form expression for $G^{(L)}$. The result is more elegant than $\rho(L)$, allowing us to calculate and understand $C_q(L)$ more directly.

Earlier, we found that the density matrix can be written as:

$$\rho(L) = A_L^\dagger D_\pi A_L ,$$

which can be rewritten as:

$$\begin{aligned}
\rho(L) &= A_L^\dagger D_\pi^{1/2} D_\pi^{1/2} A_L \\
&= (D_\pi^{1/2} A_L)^\dagger D_\pi^{1/2} A_L .
\end{aligned}$$

It is easy to show that:

$$\begin{aligned}\mathrm{tr} \left[(D_\pi^{1/2} A_L)^\dagger D_\pi^{1/2} A_L \right] &= \mathrm{tr} \left[D_\pi^{1/2} A_L (D_\pi^{1/2} A_L)^\dagger \right] \\ &= \mathrm{tr} \left[D_\pi^{1/2} A_L A_L^\dagger D_\pi^{1/2} \right] .\end{aligned}$$

This means that the sum of the eigenvalues is conserved in transforming from $A_L^\dagger D_\pi A_L$ to $D_\pi^{1/2} A_L A_L^\dagger D_\pi^{1/2}$. It is less obvious that the spectrum is also conserved, but this is also true and even easy to prove. (Observe that $AB\vec{v} = \lambda\vec{v} \implies BAB\vec{v} = \lambda B\vec{v} \implies BA(B\vec{v}) = \lambda(B\vec{v})$.) Interestingly, the new object turns out to be exactly the Gram matrix, which was previously introduced, although without this explicit relationship to the density matrix. We now see that:

$$\begin{aligned}D_\pi^{1/2} A_L A_L^\dagger D_\pi^{1/2} &= D_\pi^{1/2} \begin{bmatrix} \langle \eta_1(L) | \\ \vdots \\ \langle \eta_{|\mathcal{S}|}(L) | \end{bmatrix} \begin{bmatrix} |\eta_1(L)\rangle & \cdots & |\eta_{|\mathcal{S}|}(L)\rangle \end{bmatrix} D_\pi^{1/2} \\ &= \begin{bmatrix} \sqrt{\pi_1} \langle \eta_1(L) | \\ \vdots \\ \sqrt{\pi_{|\mathcal{S}|}} \langle \eta_{|\mathcal{S}|}(L) | \end{bmatrix} \begin{bmatrix} \sqrt{\pi_1} |\eta_1(L)\rangle & \cdots & \sqrt{\pi_{|\mathcal{S}|}} |\eta_{|\mathcal{S}|}(L)\rangle \end{bmatrix} \\ &= \begin{bmatrix} \sqrt{\pi_1 \pi_1} \langle \eta_1(L) | \eta_1(L) \rangle & \cdots & \sqrt{\pi_1 \pi_{|\mathcal{S}|}} \langle \eta_1(L) | \eta_{|\mathcal{S}|}(L) \rangle \\ \vdots & \ddots & \vdots \\ \sqrt{\pi_{|\mathcal{S}|} \pi_1} \langle \eta_{|\mathcal{S}|}(L) | \eta_1(L) \rangle & \cdots & \sqrt{\pi_{|\mathcal{S}|} \pi_{|\mathcal{S}|}} \langle \eta_{|\mathcal{S}|}(L) | \eta_{|\mathcal{S}|}(L) \rangle \end{bmatrix} \\ &= G^{(L)} .\end{aligned}$$

Since the spectrum is preserved, we can use the Gram matrix directly to compute the von Neumann entropy:

$$\begin{aligned}C_q(L) &= - \sum_{\lambda \in \Lambda_{G^{(L)}}} \lambda \log \lambda \\ &= -\mathrm{tr} \left[G^{(L)} \log G^{(L)} \right] .\end{aligned}$$

4.10.3.4 Consolidated Gram Matrix

Transforming to the Gram matrix suggests a similar and even more helpful simplification that can be made while preserving the spectrum. Define the *left-consolidated Gram matrix* to be:

$$\begin{aligned}
\tilde{G}^{(L)} &\equiv D_\pi A_L A_L^\dagger \\
&= D_\pi \begin{bmatrix} \langle \eta_1(L) | \\ \vdots \\ \langle \eta_{|\mathcal{S}|}(L) | \end{bmatrix} \begin{bmatrix} |\eta_1(L)\rangle & \cdots & |\eta_{|\mathcal{S}|}(L)\rangle \end{bmatrix} \\
&= \begin{bmatrix} \pi_1 \langle \eta_1(L) | \eta_1(L) \rangle & \cdots & \pi_1 \langle \eta_1(L) | \eta_{|\mathcal{S}|}(L) \rangle \\ \vdots & \ddots & \vdots \\ \pi_{|\mathcal{S}|} \langle \eta_{|\mathcal{S}|}(L) | \eta_1(L) \rangle & \cdots & \pi_{|\mathcal{S}|} \langle \eta_{|\mathcal{S}|}(L) | \eta_{|\mathcal{S}|}(L) \rangle \end{bmatrix}.
\end{aligned}$$

Clearly, this preserves the same trace as the density matrix and previous Gram matrix. It also preserves the spectrum. And, it has the advantage of not using square-roots of two different state probabilities in each element. Rather it has a single probability attached to each element. The same is true for the right-consolidated Gram matrix $A_L A_L^\dagger D_\pi$.

Since the spectrum is preserved, we can use the consolidated Gram matrix to compute the von Neumann entropy:

$$C_q(L) = - \sum_{\lambda \in \Lambda_{\tilde{G}^{(L)}}} \lambda \log \lambda \quad (4.28)$$

$$= -\text{tr} \left[\tilde{G}^{(L)} \log \tilde{G}^{(L)} \right]. \quad (4.29)$$

4.10.4 Appendix D: Examples

Exploring several more examples will help to illustrate the methods and lead to additional observations.

4.10.4.1 Biased Coins Process

The Biased Coins Process provides a first, simple case that realizes a nontrivial quantum state entropy [102]. There are two biased coins, named A and B . The first generates 1 with probability q ; the second, 0 with probability p . A coin is picked and flipped, generating

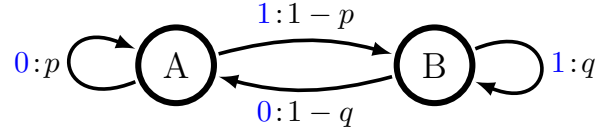


Figure 4.11: ϵ -Machine for the Biased Coins Process.

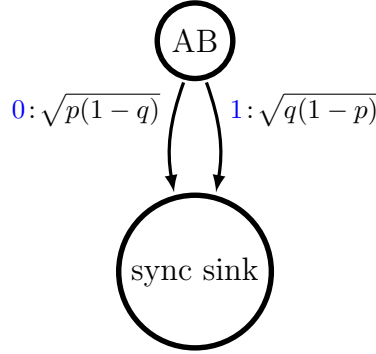


Figure 4.12: QPMM for the Biased Coins Process.

outputs 0 or 1. With probability q the other coin is used next similarly with different probability. Its two causal-state ϵ -machine is shown in Fig. 4.11.

After constructing the QPMM for the Biased Coins Process, as outlined in Figs. 4.11 and 4.12, we observe:

$$\zeta^{(0)} = \begin{bmatrix} 0 & \sqrt{p(1-q)} \\ 0 & 0 \end{bmatrix},$$

$$\zeta^{(1)} = \begin{bmatrix} 0 & \sqrt{q(1-p)} \\ 0 & 0 \end{bmatrix},$$

and so:

$$\zeta = \begin{bmatrix} 0 & \beta \\ 0 & 0 \end{bmatrix},$$

where we defined $\beta \equiv \sqrt{p(1-q)} + \sqrt{q(1-p)}$. Let's also define the suggestive quantity $\gamma \equiv (1 - \beta^2)^{-1/2}$.

The only overlap to consider is $\langle \eta_A(L) | \eta_B(L) \rangle$. For this, we note that $\langle (A, B) | = \begin{bmatrix} 1 & 0 \end{bmatrix}$. Also, $|\text{SINK}\rangle = \begin{bmatrix} 0 & 1 \end{bmatrix}^\top$.

Spectrally, ζ here is a nilpotent matrix with only a zero eigenvalue with index two: $\Lambda_\zeta = \{0\}$ and $\nu_0 = 2$. Since the projection operators must sum to the identity, we have $\zeta_0 = I$.

ζ^L is the null matrix for $L > 1$, so either by Eq. (4.6) or by Eq. (4.8), we have:

$$\langle \eta_A(L) | \eta_B(L) \rangle = \sum_{m=1}^{\min\{L, 1\}} \langle (A, B) | \zeta^m | \text{SINK} \rangle .$$

That is:

$$\langle \eta_A(L) | \eta_B(L) \rangle = \begin{cases} 0 & \text{if } L = 0 , \\ \beta & \text{if } L \geq 1 . \end{cases}$$

Entropy from the Density Matrix For the density matrix, we turn to the L -dependent orthonormal basis $\{|e_1^{(L)}\rangle, |e_2^{(L)}\rangle\}$ and use the stationary distribution over \mathcal{S} : $\pi = \begin{bmatrix} p/(p+q) & q/(p+q) \end{bmatrix}$.

Apparently, for $L = 0$ we have: $|\eta_A(0)\rangle = |e_1^{(0)}\rangle$ and $|\eta_B(0)\rangle = |e_2^{(0)}\rangle$. Hence, $\rho(0) = D_\pi$ and $C_q(0) = H_2(p/(p+q)) = C_\mu$ qubits.

For $L \geq 1$ we have: $|\eta_A(L)\rangle = |e_1^{(L)}\rangle$ and $|\eta_B(L)\rangle = a_{21}^{(L)} |e_1^{(L)}\rangle + a_{22}^{(L)} |e_2^{(L)}\rangle$, where $a_{21}^{(L)} = \langle \eta_A(L) | \eta_B(L) \rangle = \beta$ and $a_{22}^{(L)} = (1 - \beta^2)^{1/2} = \gamma^{-1}$ for $L \geq 1$. We find that:

$$A_L = \begin{bmatrix} 1 & 0 \\ \beta & \gamma^{-1} \end{bmatrix} , \text{ for } L \geq 1 .$$

Hence, for $L \geq 1$ the density matrix is:

$$\begin{aligned} \rho(L) &= A_L^\dagger D_\pi A_L \\ &= \begin{bmatrix} 1 & \beta \\ 0 & \gamma^{-1} \end{bmatrix} \begin{bmatrix} \frac{p}{p+q} & 0 \\ 0 & \frac{q}{p+q} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \beta & \gamma^{-1} \end{bmatrix} \\ &= \frac{1}{p+q} \begin{bmatrix} p & q\beta \\ 0 & q\gamma^{-1} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \beta & \gamma^{-1} \end{bmatrix} \\ &= \frac{q}{p+q} \begin{bmatrix} \frac{p}{q} + \beta^2 & \beta/\gamma \\ \beta/\gamma & 1 - \beta^2 \end{bmatrix} . \end{aligned}$$

Since:

$$\det(\rho(L) - \lambda I) = \lambda^2 - \lambda + \frac{pq}{(p+q)^2}(1 - \beta^2) ,$$

we find $\rho(L)$'s eigenvalues to be:

$$\Lambda_{\rho(L)} = \left\{ \frac{1}{2} \pm \frac{1}{2(p+q)} \sqrt{4pq\beta^2 + (p-q)^2} \right\} ,$$

which yields the von Neumann entropy for $L \geq 1$:

$$C_q(L) = - \sum_{\lambda \in \Lambda_{\rho(L)}} \lambda \log \lambda .$$

Entropy from the Consolidated Gram Matrix The left-consolidated Gram matrix for the Biased Coins Process is:

$$\tilde{G}^{(L)} = D_\pi \begin{bmatrix} \langle \eta_A(L) | \eta_A(L) \rangle & \langle \eta_A(L) | \eta_B(L) \rangle \\ \langle \eta_B(L) | \eta_A(L) \rangle & \langle \eta_B(L) | \eta_B(L) \rangle \end{bmatrix} .$$

Specifically, we have for $L = 0$:

$$\begin{aligned} \tilde{G}^{(0)} &= \frac{1}{p+q} \begin{bmatrix} p & 0 \\ 0 & q \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ &= \frac{1}{p+q} \begin{bmatrix} p & 0 \\ 0 & q \end{bmatrix} , \end{aligned}$$

and $L \geq 1$:

$$\begin{aligned} \tilde{G}^{(L)} &= \frac{1}{p+q} \begin{bmatrix} p & 0 \\ 0 & q \end{bmatrix} \begin{bmatrix} 1 & \beta \\ \beta & 1 \end{bmatrix} \\ &= \frac{1}{p+q} \begin{bmatrix} p & p\beta \\ q\beta & q \end{bmatrix} . \end{aligned}$$

$\tilde{G}^{(0)}$'s eigenvalues are simply its diagonal entries. So, $C_q(0) = H_2(p/(p+q))$ qubits.

For $L \geq 1$:

$$\det(\tilde{G}^{(L)} - \lambda I) = \lambda^2 - \lambda + \frac{pq}{(p+q)^2}(1 - \beta^2) ,$$

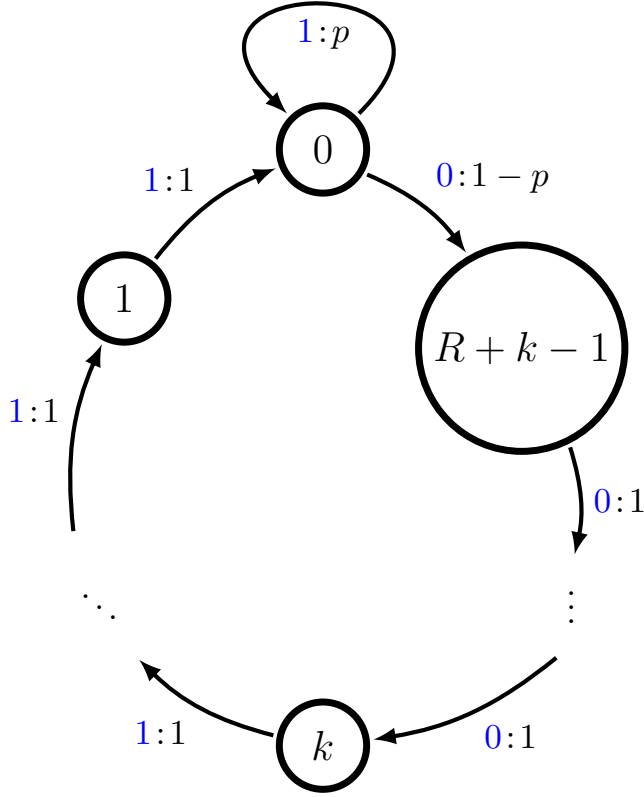


Figure 4.13: ϵ -Machine for the $(R-k)$ -Golden Mean Process.

which gives the same values for eigenvalues and entropy as we found earlier using the density matrix approach.

As the new method illustrates, there is no need to construct the density matrix. Instead, one uses the consolidated Gram matrix, which can be easily calculated from quantum overlaps. Clearly, the consolidated-Gram matrix method is more elegant for our purposes. This is evident even at $|\mathcal{S}| = 2$. This is even more critical for more complex processes since A_L grows as $|\mathcal{S}|$ grows.

4.10.4.2 $(R-k)$ -Golden Mean Process

The $(R-k)$ -Golden Mean Process is constructed to have Markov-order R and cryptic-order k . Its ϵ -machine is shown in Fig. 4.13. The 0th state σ_0 has probability $\pi_0 = 1/(R+k-p(R+k-1))$ while all other states σ_i have probability $\pi_i = (1-p)\pi_0$.

Its QPMM is strictly tree-like with depth $d = k + 1$ and maximal width k . All edges

have a unit weight except for those edges leaving A -paired states. The latter edges, numbering k in total, have an associated weight of \sqrt{p} .

The eigenvalues of the consolidated Gram matrix can be obtained from:

$$\begin{aligned} \det(\tilde{G}^{(L)} - \lambda I) &= (\pi_1 - \lambda)^{R+k-\min(L,k)-1} \\ &\times \begin{vmatrix} \pi_0 - \lambda & \pi_0 \sqrt{p} & \cdots & \pi_0 \sqrt{p}^{\min(L,k)} \\ \pi_1 \sqrt{p} & \pi_1 - \lambda & & \pi_1 \sqrt{p}^{\min(L,k)-1} \\ \vdots & & \ddots & \\ \pi_1 \sqrt{p}^{\min(L,k)} & & & \pi_1 - \lambda \end{vmatrix} \\ &= 0 , \end{aligned}$$

which directly yields the von Neumann entropy. Note that although the $C_q(L)$ is *not* actually linear in L , it appears approximately linear.

We observe that $\boldsymbol{\pi}$ is invariant under the simultaneous change of:

$$R' = R + m , \text{ while } k' = k - m , \quad (4.30)$$

for any $m \in \mathbb{Z}$. Although we insist on maintaining $R' \geq k' \geq 0$ for preservation of their functional roles. Furthermore, this transformation preserves $\det(\tilde{G}^{(L)} - \lambda I)$ for $L \leq k$ and $L \leq k'$. Hence, $C_q(L)$ is invariant to the simultaneous transformation of Eq. (4.30) for $L \leq k$ and $L \leq k'$. This explains the agreement noted in Fig. 4.6's caption—that $C_q(L)$ for $(R-k)$ -GM is the same as $C_q(L)$ for $((R+1)-(k-1))$ -GM for $L \leq k$.

To give an explicit example, let's consider the (4-3)-GM Process of Fig. 4.1. State A has probability $\pi_A = 1/(7-6p)$ while all other states have probability $\pi_i = (1-p)/(7-6p)$. Let's calculate:

- For $L = 0$:

$$\det(\tilde{G}^{(0)} - \lambda I) = (\pi_B - \lambda)^6 (\pi_A - \lambda) ,$$

yielding $\Lambda_{\tilde{G}^{(0)}} = \{\pi_B, \pi_A\}$ (with $a_{\pi_B} = 6$) and

$$C_q(0) = -6\pi_B \log \pi_B - \pi_A \log \pi_A .$$

- For $L = 1$:

$$\begin{aligned} \det(\tilde{G}^{(1)} - \lambda I) \\ = (\pi_B - \lambda)^5 [\lambda^2 - (\pi_A + \pi_B)\lambda + \pi_A\pi_B(1-p)] \ , \end{aligned}$$

yielding $\Lambda_{\tilde{G}^{(1)}} = \{\pi_B, c_+, c_-\}$ with $c_{\pm} = \frac{1}{2}(\pi_A + \pi_B) \pm \frac{1}{2}[(\pi_A + \pi_B)^2 - 4\pi_A\pi_B(1-p)]^{1/2}$
(and with $a_{\pi_B} = 5$), and:

$$C_q(1) = -5\pi_B \log \pi_B - c_+ \log c_+ - c_- \log c_- \ .$$

- For $L = 2$:

$$\begin{aligned} \det(\tilde{G}^{(2)} - \lambda I) \\ = (\pi_B - \lambda)^4 \begin{vmatrix} \pi_A - \lambda & \pi_A p^{1/2} & \pi_A p \\ \pi_B p^{1/2} & \pi_B - \lambda & \pi_B p^{1/2} \\ \pi_B p & \pi_B p^{1/2} & \pi_B - \lambda \end{vmatrix} \ . \end{aligned}$$

- For $L \geq 3$:

$$\begin{aligned} \det(\tilde{G}^{(L)} - \lambda I) \\ = \det(\tilde{G}^{(3)} - \lambda I) \\ = (\pi_B - \lambda)^3 \begin{vmatrix} \pi_A - \lambda & \pi_A p^{1/2} & \pi_A p & \pi_A p^{3/2} \\ \pi_B p^{1/2} & \pi_B - \lambda & \pi_B p^{1/2} & \pi_B p \\ \pi_B p & \pi_B p^{1/2} & \pi_B - \lambda & \pi_B p^{1/2} \\ \pi_B p^{3/2} & \pi_B p & \pi_B p^{1/2} & \pi_B - \lambda \end{vmatrix} \ . \end{aligned}$$

Chapter 5

Pairwise Correlations in Layered Close-Packed Structures

This chapter has appeared in print as:

P. M. Riechers, D. P. Varn, & J. P. Crutchfield, (2015). “Pairwise correlations in layered close-packed structures”. *Acta Crystallographica Section A: Foundations and Advances*, 71(4), 423-443.

The results carry over with minimal modification to address correlation in layered structures in general.

5.1 Chapter Overview

Given a description of the stacking statistics of layered close-packed structures in the form of a hidden Markov model, we develop analytical expressions for the pairwise correlation functions between the layers. These may be calculated analytically as explicit functions of model parameters or the expressions may be used as a fast, accurate, and efficient way to obtain numerical values. We present several examples, finding agreement with previous work as well as deriving new relations.

5.2 Introduction

There has long been an interest in planar defects or stacking faults in crystals [109, 278]. With the recent realization of the technological import of many materials prone to stacking

faults—graphene [37, 95] and SiC [282] being but two well known examples—that interest has only grown. Since stacking faults shift an entire layer of atoms, it is not unexpected that they can profoundly affect material properties. Many of these materials have more than one stable stacking configuration and additionally many metastable ones can exist [218], as well as many stacking configurations that show varying amounts of disorder. Thus, understanding stacking faults, perhaps as a prelude to controlling their kind and placement, presents a significant, but compelling challenge.

Disordered layered materials are often studied via the pairwise correlations between layers, as these correlations are experimentally accessible from the Fourier analysis of a diffraction pattern [83, 84, 264, 266] or directly from simulation studies [123, 124, 229, 228, 231, 230, 267]. Such studies yield important insights into the structural organization of materials. For example, Kabra & Pandey [123] were able to show that a model of the $2H \rightarrow 6H^1$ transformation in SiC could retain long-range order even as the short-range order was reduced. Tiwary & Pandey [248] calculated the size of domains in a model of randomly faulted close-packed structures (CPSs) by calculating the (exponential) decay rate of pairwise correlation functions between layers. Recently Estevez-Rams *et al.* [85] derived analytical expressions for the correlation functions for CPSs that contained both random growth and deformation faults, and Beyerlein *et al.* [23] demonstrated that correlation functions in finite-sized face-centered cubic crystals depend not only on the kind and amount of faulting, but additionally on their placement.

Beyond the study of layered materials, pairwise correlation information, in the form of *pair distribution functions*, has recently attracted significant attention [72]. However, as useful as the study of pairwise correlation information is, it does *not* provide a complete description of the specimen. Indeed, it has long been known that very different atomic arrangements of atoms can reproduce the same pair distribution function, although there has been recent progress in reducing this degeneracy [42]. Nor are they in general suitable for calculating material properties, such as conductivities or compressibilities.

For crystalline materials, a complete description of the specimen comes in the form

¹We will use the Ramsdell notation [183, 218] to describe well known crystalline stacking structures.

of its crystal structure, *i.e.*, the specification of the placement of all the atoms within the unit cell, coupled with the description of how the unit cells are spatially related to each other, commonly referred to as the lattice. The mathematical structure of the theory, which has been called *classical crystallography* [154], uses exact symmetries couched in the language of group theory. Determining these quantities for specimens and materials is of course the traditional purview of crystallography. For disordered materials, instead of using exact group symmetries, it has recently been suggested that the partial symmetries observed in these materials can be represented using semi-groups. This new formalism, called *chaotic crystallography* [268], employs information- and computation-theoretic measures to characterize and classify material structure. Chaotic crystallography provides a unified platform not only to calculate physical quantities of interest but also to give insight into their physical structure. For layered materials, where there is but one axis of interest, namely the organization along the stacking direction, the appropriate mathematical framework for this formalism has been identified [264, 265, 266] as *computational mechanics* [64, 55]. The mathematical entity that gives a compact, statistical description of the disordered material (along its stacking direction) is its ϵ -*machine*, a kind of hidden Markov model (HMM) [197, 74]. Chaotic crystallography also has the advantage of encompassing traditional crystal structures, so both ordered and disordered materials can be treated on the same footing in the same formalism.

It is our contention that an ϵ -machine describing a specimen's stacking includes all of the structural information necessary to calculate physical quantities that depend on the stacking statistics [268]. In the following, we demonstrate how pairwise correlation functions can be calculated either *analytically* or to a high degree of numerical certainty for an arbitrary HMM and, thus, for an arbitrary ϵ -machine. Previous researchers often calculated pairwise correlation functions for particular realizations of stacking configurations [21, 123, 230, 83, 266] or from analytic expressions constructed for particular models [248, 85, 267]. The techniques developed here, however, are the first generally applicable methods that do not rely on samples of a stacking sequence. The result delivers both an analytical solution and an efficient numerical tool. Additionally, by shifting the primary focus of

Table 5.1: A number of abbreviations are used extensively in this chapter and, as a convenience for the reader, they are tabulated here.

CF	Correlation Function	§2.1
CPS	Close-Packed Structure	§1
GM	Golden Mean (Process)	§2.2
HMM	Hidden Markov Model	§1
IID	Independent & Identically Distributed (Process)	§5.1
ML	Modular Layer	§2
RGDF	Random Growth & Deformation Fault (Process)	§5.2
SFSF	Shockley-Frank Stacking Fault (Process)	§5.3
SSC	Simple State Cycle	§3.2
TM	Transition Matrix	§2.2

study from the pair correlation functions to the HMM, new insights into the behavior of correlation functions, particularly their modes of decay to their asymptotic value, are obtained. While we will specialize to the case of CPSs for concreteness, the methods developed are extendable to other materials and stacking geometries.

Our development is organized as follows: In §5.3 we introduce nomenclature. In §5.4 we develop an algorithm to change between different representations of stacking sequences. In §5.5 we derive expressions, our main results, for the pairwise correlation functions between layers in layered CPSs. In §6.6 we consider several examples; namely, (i) a simple stacking process that represents the 3C crystal structure or a random stacking structure depending on the parameter choice, (ii) a stacking process that models random growth and deformation faults, and (iii) a stacking process inspired by recent experiments in 6H-SiC. And, in §6.7 we give our conclusions and directions for future work.

Additionally, we will find it useful to introduce a number of abbreviations that may not be familiar to the reader. As an aid, these abbreviations have been compiled in Table 5.1 along with the section of the text where they are first defined.

5.3 Definitions and Notations

We suppose the layered material is built up from identical sheets called *modular layers* (MLs) [194, 89]. The MLs are completely ordered in two dimensions and assume only one of three discrete positions, labeled A , B , or C [5, 218]. These represent the physical placement of each ML and are commonly known as the ABC -notation [183]. We define the set of possible orientations in the ABC -notation as $\mathcal{A}_P = \{A, B, C\}$. We further assume that the MLs obey the same stacking rules as CPSs, namely that two adjacent layers may not have the same orientation; *i.e.*, stacking sequences AA , BB and CC are not allowed. Exploiting this constraint, the stacking structure can be represented more compactly in the Hägg-notation: one takes the transitions between MLs as being either cyclic, ($A \rightarrow B, B \rightarrow C$, or $C \rightarrow A$), and denoted as ‘+’; or anticyclic, ($A \rightarrow C, C \rightarrow B$, or $B \rightarrow A$), and denoted as ‘−’. The Hägg-notation then gives the relative orientation of each ML to its predecessor. It is convenient to identify the usual Hägg-notation ‘+’ as ‘1’ and ‘−’ as ‘0’. Doing so, we define the set of possible relative orientations in the Hägg-notation as $\mathcal{A}_H = \{0, 1\}$. These two notations— ABC and Hägg—carry an identical message, up to an overall rotation of the specimen. Alternatively, one can say that there is freedom of choice in labeling the first ML. However, only the ABC sequences directly relate to structure factors for the specimen, making the ABC -representation of the stacking structure the natural choice for calculating pairwise correlation quantities. Indeed, it will be shown that the spectral decomposition of the ABC transition matrix directly corresponds to modes of the correlation functions relevant to diffraction patterns, whereas the same is not true for the Hägg representation.

5.3.1 Correlation functions

Let us define three statistical quantities, $Q_c(n)$, $Q_a(n)$, and $Q_s(n)$ [280]: the pairwise *correlation functions* (CFs) between MLs, where c, a, and s stand for *cyclic*, *anticyclic*, and *same*, respectively. $Q_c(n)$ is the probability that any two MLs at a separation of n are cyclically related. $Q_a(n)$ and $Q_s(n)$ are defined in a similar fashion.² Since these are

²As yet, there is no consensus on notation for these quantities. Warren [273] uses P_m^0 , P_m^+ , and P_m^- , Kabra & Pandey [123] call these $P(m)$, $Q(m)$, and $R(m)$, and Estevez *et al.* [85] use $P_0(\Delta)$, $P_1(\Delta)$, and

probabilities: $0 \leq Q_\xi(n) \leq 1$, where $\xi \in \{c, a, s\}$. Additionally, at each n it is clear that $\sum_\xi Q_\xi(n) = 1$. Notice that the CFs are defined in terms of the *ABC*-notation.

5.3.2 Representing layer stacking as a hidden process

We chose to represent a stacking sequence as the output of discrete-step, discrete-state hidden Markov model (HMM). A HMM Γ is an ordered tuple $\Gamma = (\mathcal{A}, \mathbb{S}, \mu_0, \mathbf{T})$, where \mathcal{A} is the set of symbols that one observes as the HMM's output, often called an alphabet, \mathbb{S} is a finite set of M internal states, μ_0 is an initial state probability distribution, and \mathbf{T} is a set of matrices that give the probability of making a transition between the states while outputting one of the symbols in \mathcal{A} . These transition probability matrices or more simply *transition matrices* (TMs) [190, 126] are usually written:

$$\mathcal{T}^{[s]} = \begin{bmatrix} \Pr(s, S_1|S_1) & \Pr(s, S_2|S_1) & \cdots & \Pr(s, S_M|S_1) \\ \Pr(s, S_1|S_2) & \Pr(s, S_2|S_2) & \cdots & \Pr(s, S_M|S_2) \\ \vdots & \vdots & \ddots & \vdots \\ \Pr(s, S_1|S_M) & \Pr(s, S_2|S_M) & \cdots & \Pr(s, S_M|S_M) \end{bmatrix},$$

where $s \in \mathcal{A}$ and $S_1, S_2, \dots, S_M \in \mathbb{S}$.

For a number of purposes it is convenient to work directly with the internal state TM, denote it \mathcal{T} . This is the matrix of state transition probabilities regardless of symbol, given by the sum of the symbol-labeled TMs: $\mathcal{T} = \sum_{x \in \mathcal{A}} \mathcal{T}^{[x]}$. For example, the ensemble internal state distribution evolves according to $\langle \mu_1 | = \langle \mu_0 | \mathcal{T}$. Or, more generally, $\langle \mu_L | = \langle \mu_0 | \mathcal{T}^L$. (In this notation, state distributions are row vectors.) In another use, one finds the stationary state probability distribution:

$$\langle \pi | = \left[\Pr(S_1) \quad \Pr(S_2) \quad \cdots \quad \Pr(S_M) \right],$$

as the left eigenvector of \mathcal{T} normalized in probability:

$$\langle \pi | = \langle \pi | \mathcal{T}. \quad (5.1)$$

$P_b(\Delta)$. Since we prefer to reserve the symbol ‘ P ’ for other probabilities previously established in the literature, here and elsewhere we follow the notation of Yi & Canright [280], with a slight modification of replacing ‘ $Q_r(n)$ ’ with ‘ $Q_a(n)$ ’.

The probability of any finite-length sequence of symbols can be computed exactly from these objects using linear algebra. In particular, a length- L ‘word’ $w = s_0 s_1 \dots s_{L-1} \in \mathcal{A}^L$, where \mathcal{A}^L is the set of length- L sequences, has the stationary probability:

$$\begin{aligned} \Pr(w) &= \langle \boldsymbol{\pi} | \mathcal{T}^{[w]} | \mathbf{1} \rangle \\ &= \langle \boldsymbol{\pi} | \mathcal{T}^{[s_0]} \mathcal{T}^{[s_1]} \dots \mathcal{T}^{[s_{L-1}]} | \mathbf{1} \rangle , \end{aligned}$$

where $|\mathbf{1}\rangle$ is the column-vector of all ones.

As a useful convention, we will use bras $\langle \cdot |$ to denote row vectors and kets $|\cdot\rangle$ to denote column vectors. On the one hand, any object closed by a bra on the left and ket on the right is a scalar and commutes as a unit with anything. On the other hand, a ket–bra $|\cdot\rangle \langle \cdot|$ has the dimensions of a square matrix. More explanation can be found in Appendix A.

To help familiarize the reader with TMs as they will be employed in the examples of §6.6, and to make a very useful connection to automata theory [112], we now develop these ideas using a simple pedagogical example. Let us consider a CPS stacked according to the *Golden Mean* (GM) Process [62]. The GM Process has been previously treated in the context of CPSs in Varn [270]. The GM Process is generically defined as follows: any sequence is allowed as long as there are no consecutive 0s. This is accomplished by examining the previous observed symbol: if it is 1, then the next symbol in the sequence is either 0 or 1 with equal probability; if it is 0, then the next symbol is necessarily 1.³ Thus, when scanning a sequence there are two possible states corresponding to the above two conditions. Let us call these states \mathcal{U} (next symbol is a 0 or 1 with equal probability) and \mathcal{V} (next symbol is a 1). And so, we say $\mathbb{S} = \{\mathcal{U}, \mathcal{V}\}$. The two 2-by-2 TMs for this

³The alert reader may object that there isn’t any inherent prohibition against consecutive 0s appearing in the stacking sequence of CPSs. She is quite right. While there is no general prohibition, it may be that in a *particular* specimen, whether by accident, by some natural as yet undiscovered mechanism or by the design of an experimentalist, the 00 sequence never appears. This should be reflected in the HMM. Additionally, the reader may object that rotating the specimen by 60° about the stacking direction will exchange all 0s and 1s, making such sequence prohibitions spurious. Indeed, for simple CPSs structures, $0 \Leftrightarrow 1$ is a good symmetry. (See Varn *et al.* [263] for a discussion.) While a physical rotation of the specimen doesn’t alter the physical stacking, it can alter the details of how the stacking sequence is labeled, and this *is* reflected in the HMM. In any event, the details of GM Process are not relevant to our overall theoretical development, and the prohibition of the 00 sequence just serves to create a simple HMM that is both nontrivial and easy to analyze.

process—one for each symbol in the alphabet—are given by:

$$\mathbf{T}^{[0]} = \begin{bmatrix} 0 & \frac{1}{2} \\ 0 & 0 \end{bmatrix} \text{ and } \mathbf{T}^{[1]} = \begin{bmatrix} \frac{1}{2} & 0 \\ 1 & 0 \end{bmatrix}.$$

The GM Process has internal-state TM:

$$\begin{aligned} \mathbf{T} &= \mathbf{T}^{[0]} + \mathbf{T}^{[1]} \\ &= \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ 1 & 0 \end{bmatrix}. \end{aligned}$$

The asymptotic state probabilities are found by direct application of Eq. (5.1) and are found to be $\langle \boldsymbol{\pi}_H | = \begin{bmatrix} \frac{2}{3} & \frac{1}{3} \end{bmatrix}$.

Thus a more formal definition of the GM Process is given as $\Gamma_{\text{GM}}^{(\text{H})} = (\mathcal{A}, \mathbb{S}, \mu_0, \mathbf{T}) = (\{0, 1\}, \{\mathcal{U}, \mathcal{V}\}, [\frac{2}{3} \quad \frac{1}{3}], \{\mathbf{T}^{[0]}, \mathbf{T}^{[1]}\})$.⁴ HMMs are often conveniently depicted as a *finite-state automaton* [112], a kind of labeled directed graph. For example, the labeled directed graph that represents the HMM for the GM Process is shown in Fig. 5.1. When the finite-state automaton representing the HMM is given in terms of the Hägg-notation (as is done here), we will call that the *Hägg-machine* for that process.

Thus, if the MLs of a specimen were stacked according to the GM Process, that would mean that sequences like 00 are never observed in that specimen. Alternatively, this can be restated in the *ABC*-notation: sequences of the form *ACB*, *CBA* and *BAC* are not observed in that specimen. In the next section we show how to rewrite any Hägg-machine as a finite state automaton in the *ABC*-notation, which we will call the *ABC-machine*. This is the necessary mathematical object required to calculate CFs, and will be used extensively in our theoretical development (§5.5).

5.4 Expanding the Hägg-machine to the *ABC*-machine

While simulation studies [267] and ϵ -machine spectral reconstruction [264, 265, 266, 270] express stacking structure in terms of the Hägg-machine, our theoretical methods developed

⁴Here and in the examples of §6.6, we take the stationary state probability distribution $\boldsymbol{\pi}$ as the initial probability state distribution μ_0 , as we are interested for now in the the long term behavior.

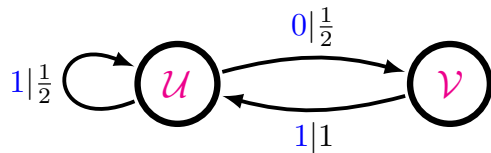


Figure 5.1: The GM Process written as a Hägg-machine. The circles indicate states, and the arcs between them are transitions, labeled by $s|p$, where s is the symbol emitted upon transition and p is the probability of making such a transition.

in §4 require that it be reexpressed in terms of the *ABC*-machine. This section is devoted to detailing techniques to do just that. We give a graphical procedure for expanding the Hägg-machine into the *ABC*-machine (§5.4.1) and then provide an algebraically equivalent algorithm (§5.4.3). We note that this expansion procedure is *not* unique and can vary up to an overall rearrangement of the columns and rows of the resulting *ABC*-machine TM. This difference, of course, does not alter the results of calculations of physical quantities.

5.4.1 Graphical expansion method

Recall that the Hägg-notation and the *ABC*-notation are equivalent representations of the stacking structure, up to an overall rotation of the crystal. Stated alternatively, in the Hägg-notation, there is an ambiguity concerning the orientation of each ML—it could be either *A*, *B* or *C*. To account for this degeneracy, when we transform to the *ABC* representation, we triple the size of the Hägg-machine. As a first step, one writes down three states for each state found in the Hägg-machine, but not the transitions between them. To distinguish among these new states of the triplet, label each with a superscript (*A*, *B* or *C*) indicating the last ML added to arrive at that state. Symbolically, this is stated:

$$\{\mathcal{S}_i\} \xrightarrow{\text{Hägg to } ABC} \left\{ \mathcal{S}_i^{[A]}, \mathcal{S}_i^{[B]}, \mathcal{S}_i^{[C]} \right\} .$$

Transitions between states on the *ABC*-machine still respect the same state labeling scheme as on the Hägg-machine (explained below), but now they store the ML information. Transitions between states on the Hägg-machine that were labeled with 1 advance the stacking sequence cyclically (*i.e.*, $A \rightarrow B \rightarrow C \rightarrow A$) and the corresponding transitions on the *ABC*-machine reflect this by taking the ML label on the initial state and advancing

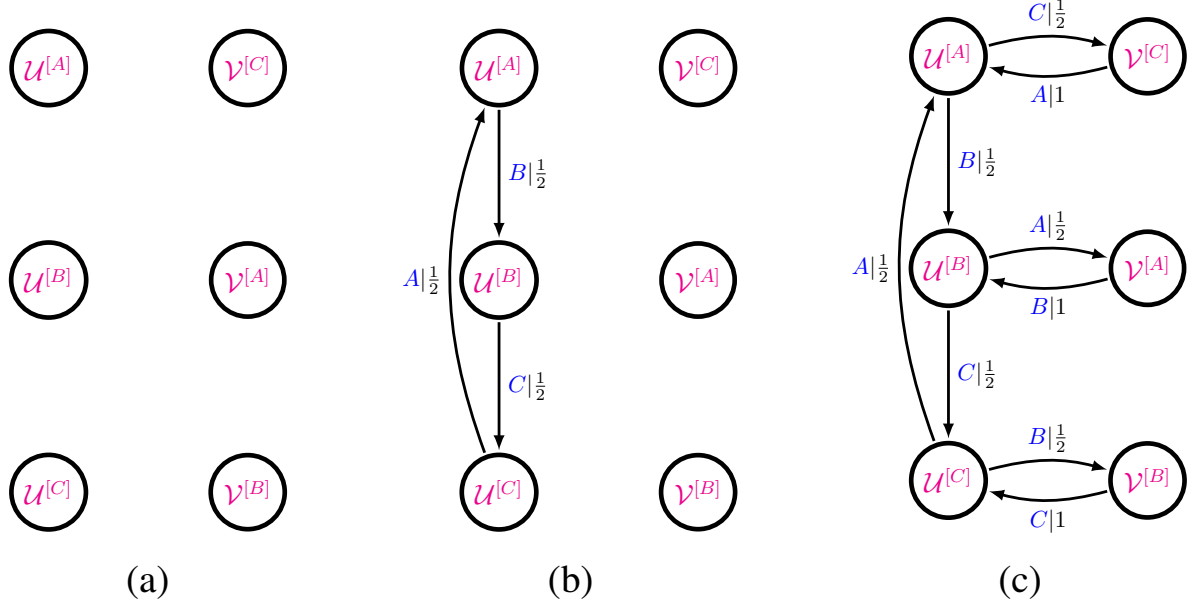


Figure 5.2: The steps to graphically expanding the Hägg-machine to the ABC -machine demonstrated using the The GM Process. (a) The two states of Hägg-machine shown in Fig. 5.1 are increased to six, as shown. (b) The self-state transition $\mathcal{U} \xrightarrow{1|\frac{1}{2}} \mathcal{U}$ of the Hägg-machine now becomes three transitions among the $\mathcal{U}^{[x]}$ of the ABC -machine. (c) The remaining transitions of the Hägg-machine are expanded in a similar fashion (see text), resulting in the ABC -machine for the GM Process.

it cyclically. In a completely analogous way, transitions labeled 0 on the Hägg-machine advance the states on the ABC -machine in an anticyclical fashion (*i.e.*, $A \rightarrow C \rightarrow B \rightarrow A$).

Continuing our GM Process example, let us write out the six ($= 3 \times 2$) states labeled with superscripts to distinguish them. This is done in Fig. 5.2(a). (It does not matter in what order these states are labeled. The scheme chosen in Fig. 5.2(a) turns out to be convenient given the state-to-state transition structure of the final ABC -machine, but any arrangement is satisfactory.) The transitions between the states on the ABC -machine preserve the labeling scheme of the original Hägg-machine. That is, if in the original Hägg-machine there is transition $\mathcal{S}_i \xrightarrow{s|p} \mathcal{S}_j$, then there must be *three* similar transitions on the ABC -machine of the form $\mathcal{S}_i^{[x]} \xrightarrow{x'|p} \mathcal{S}_j^{[x']}$, with $x, x' \in \{A, B, C\}$. Additionally, the transitions on the ABC -machine corresponding to the transitions on the Hägg-machine have the same conditional probability.

Let us consider the self-state transition on the Hägg-machine shown in Fig. 5.1: $\mathcal{U} \xrightarrow{1|\frac{1}{2}} \mathcal{U}$.

Since the corresponding transitions on the ABC -machine still respect the state labeling scheme, the self-loop on \mathcal{U} only induces transitions among the $\mathcal{U}^{[x]}$. Since a 1 advances the stacking sequence cyclically, the appropriate transitions are:

$$\begin{aligned}\mathcal{U}^{[A]} &\xrightarrow{B|\frac{1}{2}} \mathcal{U}^{[B]}, \\ \mathcal{U}^{[B]} &\xrightarrow{C|\frac{1}{2}} \mathcal{U}^{[C]}, \\ \mathcal{U}^{[C]} &\xrightarrow{A|\frac{1}{2}} \mathcal{U}^{[A]}.\end{aligned}$$

This is illustrated in Fig. 5.2(b). Applying the same procedure to the other transitions on the Hägg-machine, *i.e.*, $\mathcal{U} \xrightarrow{0|\frac{1}{2}} \mathcal{V}$ and $\mathcal{V} \xrightarrow{1|1} \mathcal{U}$, results in the completely expanded ABC -machine, and this is shown in Fig. 5.2(c).

We are now able to write down the stacking process for the GM Process from its expanded graph, Fig. 5.2(c). First, we note that the alphabet is ternary: $\mathcal{A}_P = \{A, B, C\}$. Second, there are six states on the ABC -machine, *i.e.*, $\mathcal{S} = \{\mathcal{U}^{[A]}, \mathcal{U}^{[B]}, \mathcal{U}^{[C]}, \mathcal{V}^{[A]}, \mathcal{V}^{[B]}, \mathcal{V}^{[C]}\}$. Ordering the states as above, the TMs may be directly constructed from the expanded graph, and are given by:

$$\mathcal{T}^{[A]} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad \mathcal{T}^{[B]} = \begin{bmatrix} 0 & \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

and

$$\mathcal{T}^{[C]} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

As before, the internal-state TM is simply the sum of the symbol-specific TMs, given by $\mathcal{T} = \mathcal{T}^{[A]} + \mathcal{T}^{[B]} + \mathcal{T}^{[C]}$. For the GM Process this turns out to be:

$$\mathcal{T} = \begin{bmatrix} 0 & \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

For completeness, the HMM for the GM Process *in terms of the physical stacking of MLs* is $\Gamma_{\text{GM}}^{(\text{P})} = (\mathcal{A}, \mathbb{S}, \mu_0, \mathbf{T}) = (\{A, B, C\}, \{\mathcal{U}^{[A]}, \mathcal{U}^{[B]}, \mathcal{U}^{[C]}, \mathcal{V}^{[A]}, \mathcal{V}^{[B]}, \mathcal{V}^{[C]}\}, \frac{1}{9} [2 \ 2 \ 2 \ 1 \ 1 \ 1], \{\mathcal{T}^{[A]}, \mathcal{T}^{[B]}, \mathcal{T}^{[C]}\})$. It is from such *ABC* representations that we will make direct connection to the CFs in §5.5.

5.4.2 Mixing and Nonmixing State Cycles

Observe Fig. 5.2(c)’s directed graph is *strongly connected*—any state is accessible from any other state in a finite number of transitions. It should be apparent that this need not have been the case. In fact, in this example connectivity is due to the presence of the self-state transition $\mathcal{U} \xrightarrow{1} \mathcal{U}$. The latter guarantees a strongly connected expanded graph. Had this transition been absent on the Hägg-machine, such that there were only transitions of the form $\mathcal{U} \xrightarrow{0} \mathcal{V}$ and $\mathcal{V} \xrightarrow{1} \mathcal{U}$, the expansion would have yielded a graph with three distinct, unconnected components. Only one of these graphs would be physically manifest. It is sufficient to take just one component, arbitrarily assign a *A, B* or *C* to an arbitrary state on that component, and then replace all of the $\{0, 1\}$ transitions with the appropriate $\{A, B, C\}$ transitions, as done above.

To determine whether the expansion process on a Hägg-machine results in a strongly connected graph, one can examine the set of *simple state cycles* (SSCs) and calculate the *winding number* for each. A SSC is defined analogous to a *causal state cycle* [266] on an ϵ -machine as a “finite, closed, nonself-intersecting, symbol-specific path” along the graph. The winding number W for a SSC on a Hägg-machine is similar to the parameter Δ previously defined by Yi & Canright [280] and the *cyclicity* (C) [68] for a polytype of a

CPS. The winding number differs from cyclicity as the former is not divided by the period of the cycle. We define the winding number for a SSC as:

$$W^{\text{SSC}} = n_1 - n_0 ,$$

where n_1 and n_0 are the number of 1s and the number of 0s encountered traversing the SSC, respectively. We call those SSCs *mixing* if $W^{\text{SSC}} \pmod{3} \neq 0$, and *nonmixing* if $W^{\text{SSC}} \pmod{3} = 0$. If there is at least one mixing SSC on the Hägg-machine, then the expanded *ABC*-machine will be strongly connected. For example, there are two SSCs on the Hägg-machine for the GM Process: $[\mathcal{U}]$ and $[\mathcal{UV}]$.⁵ The winding number for each is given by $W^{[\mathcal{U}]} = 1 - 0 = 1$ and $W^{[\mathcal{UV}]} = 1 - 1 = 0$. Since $W^{[\mathcal{U}]} \neq 0$ and $[\mathcal{U}]$ is thus a mixing SSC, the Hägg-machine for the GM Process will expand into a strongly connected *ABC*-machine. Let us refer to those Hägg-machines with at least one mixing SSC as *mixing Hägg-machines* and those that do not as *nonmixing Hägg-machines* and similarly for the corresponding *ABC*-machines. We find that mixing Hägg-machines, and thus mixing *ABC*-machines, are far more common than nonmixing ones and that the distinction between the two can have profound effects on the calculated quantities, such as the CFs and the diffraction pattern [204].

5.4.3 Rote expansion algorithm

To develop an algorithm for expansion, it is more convenient to change notation slightly. Let us now denote \mathcal{S} as the set of hidden recurrent states in the *ABC*-machine, indexed by integer subscripts: $\mathcal{S} = \{\mathcal{S}_i : i = 1, \dots, M_P\}$, where $M_P = |\mathcal{S}|$. Define the probability to transition from state \mathcal{S}_i to state \mathcal{S}_j on the symbol $x \in \mathcal{A}_P$ as $\mathcal{T}_{i,j}^{[x]}$. Let's gather these state-to-state transition probabilities into a $M_P \times M_P$ matrix, referring to it as the *x-transition matrix* (*x-TM*) $\mathcal{T}^{[x]}$. Thus, there will be as many *x-TMs* as there are symbols in the alphabet of the *ABC*-machine, which is always $|\mathcal{A}_P| = 3$ for CPSs.

As before, transitioning on symbol 1 has a threefold degeneracy in the *ABC* language, as it could imply any of the three transitions ($A \rightarrow B$, $B \rightarrow C$, or $C \rightarrow A$), and similarly

⁵We use the same nomenclature to denote a SSC as previously used to denote a causal state cycle: The state sequence visited traversing the cycle is given in square brackets [266]. For those cases where an ambiguity exists because the transition occurs on more than one symbol, we insert a subscript in parentheses denoting that symbol.

for 0. Thus, each labeled edge of the Hägg-machine must be split into three distinct labeled edges of the ABC -machine. Similarly, each state of the Hägg-machine maps onto three distinct states of the ABC -machine. Although we have some flexibility in indexing states in the resulting ABC -machine, we establish consistency by committing to the following construction.⁶

If M_H is the number of states in the Hägg-machine, then $M_P = 3M_H$ for mixing Hägg-machines. (The case of nonmixing Hägg-machines is treated afterward.) Let the i^{th} state of the Hägg-machine split into the $(3i-2)^{\text{th}}$ through the $(3i)^{\text{th}}$ states of the corresponding ABC -machine. Then, each labeled-edge transition from the i^{th} to the j^{th} states of the Hägg-machine maps into a 3-by-3 submatrix for each of the three labeled TMs of the ABC -machine as:

$$\left\{ T_{ij}^{[0]} \right\} \xrightarrow{\text{Hägg to } ABC} \left\{ \mathcal{T}_{3i-1,3j-2}^{[A]}, \mathcal{T}_{3i,3j-1}^{[B]}, \mathcal{T}_{3i-2,3j}^{[C]} \right\} \quad (5.2)$$

and

$$\left\{ T_{ij}^{[1]} \right\} \xrightarrow{\text{Hägg to } ABC} \left\{ \mathcal{T}_{3i,3j-2}^{[A]}, \mathcal{T}_{3i-2,3j-1}^{[B]}, \mathcal{T}_{3i-1,3j}^{[C]} \right\}. \quad (5.3)$$

We can represent the mapping of Eq. (5.2) and Eq. (5.3) more visually with the following equivalent set of statements:

$$\begin{pmatrix} \mathcal{T}_{3i-2,3j-2}^{[A]} & \mathcal{T}_{3i-2,3j-1}^{[A]} & \mathcal{T}_{3i-2,3j}^{[A]} \\ \mathcal{T}_{3i-1,3j-2}^{[A]} & \mathcal{T}_{3i-1,3j-1}^{[A]} & \mathcal{T}_{3i-1,3j}^{[A]} \\ \mathcal{T}_{3i,3j-2}^{[A]} & \mathcal{T}_{3i,3j-1}^{[A]} & \mathcal{T}_{3i,3j}^{[A]} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ T_{ij}^{[0]} & 0 & 0 \\ T_{ij}^{[1]} & 0 & 0 \end{pmatrix}, \quad (5.4)$$

$$\begin{pmatrix} \mathcal{T}_{3i-2,3j-2}^{[B]} & \mathcal{T}_{3i-2,3j-1}^{[B]} & \mathcal{T}_{3i-2,3j}^{[B]} \\ \mathcal{T}_{3i-1,3j-2}^{[B]} & \mathcal{T}_{3i-1,3j-1}^{[B]} & \mathcal{T}_{3i-1,3j}^{[B]} \\ \mathcal{T}_{3i,3j-2}^{[B]} & \mathcal{T}_{3i,3j-1}^{[B]} & \mathcal{T}_{3i,3j}^{[B]} \end{pmatrix} = \begin{pmatrix} 0 & T_{ij}^{[1]} & 0 \\ 0 & 0 & 0 \\ 0 & T_{ij}^{[0]} & 0 \end{pmatrix}, \quad (5.5)$$

and:

⁶Alternative constructions merely swap the labels of different states, but this choice of indexing affects the particular form of the TMs and how they are extracted from the Hägg-machine TMs. We choose the construction here for its intuitive and simple form.

$$\begin{pmatrix} \mathcal{T}_{3i-2,3j-2}^{[C]} & \mathcal{T}_{3i-2,3j-1}^{[C]} & \mathcal{T}_{3i-2,3j}^{[C]} \\ \mathcal{T}_{3i-1,3j-2}^{[C]} & \mathcal{T}_{3i-1,3j-1}^{[C]} & \mathcal{T}_{3i-1,3j}^{[C]} \\ \mathcal{T}_{3i,3j-2}^{[C]} & \mathcal{T}_{3i,3j-1}^{[C]} & \mathcal{T}_{3i,3j}^{[C]} \end{pmatrix} = \begin{pmatrix} 0 & 0 & \mathsf{T}_{ij}^{[0]} \\ 0 & 0 & \mathsf{T}_{ij}^{[1]} \\ 0 & 0 & 0 \end{pmatrix}, \quad (5.6)$$

which also yields the 3-by-3 submatrix for the *unlabeled ABC* TM in terms of the *labeled Hägg TMs*:

$$\begin{pmatrix} \mathcal{T}_{3i-2,3j-2} & \mathcal{T}_{3i-2,3j-1} & \mathcal{T}_{3i-2,3j} \\ \mathcal{T}_{3i-1,3j-2} & \mathcal{T}_{3i-1,3j-1} & \mathcal{T}_{3i-1,3j} \\ \mathcal{T}_{3i,3j-2} & \mathcal{T}_{3i,3j-1} & \mathcal{T}_{3i,3j} \end{pmatrix} = \begin{pmatrix} 0 & \mathsf{T}_{ij}^{[1]} & \mathsf{T}_{ij}^{[0]} \\ \mathsf{T}_{ij}^{[0]} & 0 & \mathsf{T}_{ij}^{[1]} \\ \mathsf{T}_{ij}^{[1]} & \mathsf{T}_{ij}^{[0]} & 0 \end{pmatrix}. \quad (5.7)$$

For non-mixing Hägg-machines, the above algorithm creates three disconnected *ABC*-machines, of which only one should be retained.

Furthermore, for mixing Hägg-machines, the probability from the stationary distribution over their states maps to a triplet of probabilities for the stationary distribution over the *ABC*-machine states:

$$\{p_i^{\text{H}}\} \xrightarrow{\text{Hägg to ABC}} \{3p_{3i-2}, 3p_{3i-1}, 3p_{3i}\} \quad (5.8)$$

such that:

$$\begin{aligned} \langle \boldsymbol{\pi} | &= \begin{bmatrix} p_1 & p_2 & p_3 & p_4 & \dots & p_{M_P-1} & p_{M_P} \end{bmatrix} \\ &= \frac{1}{3} \begin{bmatrix} p_1^{\text{H}} & p_1^{\text{H}} & p_1^{\text{H}} & p_2^{\text{H}} & \dots & p_{M_H}^{\text{H}} & p_{M_H}^{\text{H}} \end{bmatrix}. \end{aligned} \quad (5.9)$$

The reader should check that applying the rote expansion method given here results in the same HMM for the GM Process as we found in §5.4.1.

5.5 Correlation Functions from HMMs

In this section we derive expressions for the CFs in terms of the *ABC*-machine. We present two layers of analysis: one directly in terms of the TMs in §4.1, followed by one based on spectral analysis of the *ABC*-machine in §4.2. From the results of this latter method, we find the conditions under which the CFs decay to asymptotic values in §4.3, and additionally we discover constraints on the modes of decay of the CFs in §4.4.

We introduce the family of cyclic-relation functions $\hat{\xi}(x) \in \{\hat{c}(x), \hat{a}(x), \hat{s}(x)\}$, where, for example:

$$\hat{c}(x) = \begin{cases} B & \text{if } x = A \\ C & \text{if } x = B \\ A & \text{if } x = C \end{cases} . \quad (5.10)$$

Thus, $\hat{c}(x)$ is the cyclic permutation function. Complementarily, $\hat{a}(x)$ performs anticyclic permutation among $x \in \{A, B, C\}$; $\hat{s}(x)$ performs the identity operation among $x \in \{A, B, C\}$ and is suggestively denoted with an ‘s’ for *sameness*. In terms of the absolute position of the MLs—*i.e.*, $\mathcal{A}_P = \{A, B, C\}$ —the CFs directly relate to the products of particular sequences of TMs. This perspective suggests a way to uncover the precise relation between the CFs and the TMs. Using this, we then give a closed-form expression for $Q_\xi(n)$ for any given HMM.

5.5.1 CFs from TMs

As a prelude to developing a general method to calculate any arbitrary CF from the *ABC*-machine, let us first consider the meaning of $Q_c(3)$. In words, this is the probability that two MLs separated by two intervening MLs are cyclically related. Mathematically, we might start by writing this as:

$$Q_c(3) = \Pr(A ** B) + \Pr(B ** C) + \Pr(C ** A), \quad (5.11)$$

where $*$ is a wildcard symbol denoting an indifference for the symbol observed in its place.⁷

That is, $*$ s denote marginalizing over the intervening MLs such that, for example:

$$\Pr(A ** B) = \sum_{x_1 \in \mathcal{A}_P} \sum_{x_2 \in \mathcal{A}_P} \Pr(Ax_1x_2B). \quad (5.12)$$

⁷While it is tempting to add the stipulation that no two consecutive symbols can be the same, this will fall out naturally from $Q_s(1) = 0$ via the transition-constraints built into the *ABC*-machine construction.

Making use of the TM-formalism discussed previously, this becomes:

$$\begin{aligned}
\Pr(A * B) &= \sum_{x_1 \in \mathcal{A}_P} \sum_{x_2 \in \mathcal{A}_P} \Pr(Ax_1x_2B) \\
&= \sum_{x_1 \in \mathcal{A}_P} \sum_{x_2 \in \mathcal{A}_P} \langle \boldsymbol{\pi} | \mathcal{T}^{[A]} \mathcal{T}^{[x_1]} \mathcal{T}^{[x_2]} \mathcal{T}^{[B]} | \mathbf{1} \rangle \\
&= \langle \boldsymbol{\pi} | \mathcal{T}^{[A]} \left(\sum_{x_1 \in \mathcal{A}_P} \sum_{x_2 \in \mathcal{A}_P} \mathcal{T}^{[x_1]} \mathcal{T}^{[x_2]} \right) \mathcal{T}^{[B]} | \mathbf{1} \rangle \\
&= \langle \boldsymbol{\pi} | \mathcal{T}^{[A]} \underbrace{\left(\sum_{x_1 \in \mathcal{A}_P} \mathcal{T}^{[x_1]} \right)}_{=\mathcal{T}} \underbrace{\left(\sum_{x_2 \in \mathcal{A}_P} \mathcal{T}^{[x_2]} \right)}_{=\mathcal{T}} \mathcal{T}^{[B]} | \mathbf{1} \rangle \\
&= \langle \boldsymbol{\pi} | \mathcal{T}^{[A]} (\mathcal{T})(\mathcal{T}) \mathcal{T}^{[B]} | \mathbf{1} \rangle \\
&= \langle \boldsymbol{\pi} | \mathcal{T}^{[A]} \mathcal{T}^2 \mathcal{T}^{[B]} | \mathbf{1} \rangle ,
\end{aligned}$$

where $|\mathbf{1}\rangle$ is a column vector of 1s of length M_P . Hence, we can rewrite $Q_c(3)$ as:

$$\begin{aligned}
Q_c(3) &= \Pr(A * B) + \Pr(B * C) + \Pr(C * A) \\
&= \langle \boldsymbol{\pi} | \mathcal{T}^{[A]} \mathcal{T}^2 \mathcal{T}^{[B]} | \mathbf{1} \rangle + \langle \boldsymbol{\pi} | \mathcal{T}^{[B]} \mathcal{T}^2 \mathcal{T}^{[C]} | \mathbf{1} \rangle \\
&\quad + \langle \boldsymbol{\pi} | \mathcal{T}^{[C]} \mathcal{T}^2 \mathcal{T}^{[A]} | \mathbf{1} \rangle \\
&= \sum_{x \in \mathcal{A}_P} \langle \boldsymbol{\pi} | \mathcal{T}^{[x]} \mathcal{T}^2 \mathcal{T}^{[\hat{c}(x)]} | \mathbf{1} \rangle .
\end{aligned}$$

For *mixing ABC*-machines, $\Pr(A * B) = \Pr(B * C) = \Pr(C * A) = \frac{1}{3}Q_c(3)$, in which case the above reduces to:

$$Q_c(3) = 3 \langle \boldsymbol{\pi} | \mathcal{T}^{[x_0]} \mathcal{T}^2 \mathcal{T}^{[\hat{c}(x_0)]} | \mathbf{1} \rangle , \text{ where } x_0 \in \mathcal{A}_P.$$

The generalization to express any $Q_\xi(n)$ in terms of TMs may already be obvious by analogy. Nevertheless, we give a brief derivation for completeness, using similar concepts to those developed more explicitly above. For all $\xi \in \{c, a, s\}$ and for all $n \in \{1, 2, 3, \dots\}$,

we can write the CFs as:

$$\begin{aligned}
Q_\xi(n) &= \Pr(A \underbrace{* \cdots *}_{n-1 \text{ *s}} \hat{\xi}(A)) + \Pr(B \underbrace{* \cdots *}_{n-1 \text{ *s}} \hat{\xi}(B)) \\
&\quad + \Pr(C \underbrace{* \cdots *}_{n-1 \text{ *s}} \hat{\xi}(C)) \\
&= \sum_{x_0 \in \mathcal{A}_P} \Pr(x_0 \underbrace{* \cdots *}_{n-1 \text{ *s}} \hat{\xi}(x_0)) \\
&= \sum_{x_0 \in \mathcal{A}_P} \sum_{w \in \mathcal{A}_P^{n-1}} \Pr(x_0 w \hat{\xi}(x_0)) \\
&= \sum_{x_0 \in \mathcal{A}_P} \sum_{w \in \mathcal{A}_P^{n-1}} \langle \boldsymbol{\pi} | \mathcal{T}^{[x_0]} \mathcal{T}^{[w]} \mathcal{T}^{[\hat{\xi}(x_0)]} | \mathbf{1} \rangle \\
&= \sum_{x_0 \in \mathcal{A}_P} \langle \boldsymbol{\pi} | \mathcal{T}^{[x_0]} \left(\sum_{w \in \mathcal{A}_P^{n-1}} \mathcal{T}^{[w]} \right) \mathcal{T}^{[\hat{\xi}(x_0)]} | \mathbf{1} \rangle \\
&= \sum_{x_0 \in \mathcal{A}_P} \langle \boldsymbol{\pi} | \mathcal{T}^{[x_0]} \underbrace{\left(\prod_{i=1}^{n-1} \sum_{x_i \in \mathcal{A}_P} \mathcal{T}^{[x_i]} \right)}_{=\mathcal{T}} \mathcal{T}^{[\hat{\xi}(x_0)]} | \mathbf{1} \rangle \\
&= \sum_{x_0 \in \mathcal{A}_P} \langle \boldsymbol{\pi} | \mathcal{T}^{[x_0]} \mathcal{T}^{n-1} \mathcal{T}^{[\hat{\xi}(x_0)]} | \mathbf{1} \rangle, \tag{5.13}
\end{aligned}$$

where the stationary distribution $\langle \boldsymbol{\pi} |$ over states of the ABC -machine is found from Eq. (5.1). *The most general connection between CFs and TMs is given by Eq. (5.13) and this represents one of the main results of this paper.*

As before, we might assume on physical grounds that:

$$\Pr(A \underbrace{* \cdots *}_{n-1 \text{ *s}} \hat{\xi}(A)) = \Pr(B \underbrace{* \cdots *}_{n-1 \text{ *s}} \hat{\xi}(B)) = \Pr(C \underbrace{* \cdots *}_{n-1 \text{ *s}} \hat{\xi}(C)). \tag{5.14}$$

For example, Eq. (5.14) is always true of mixing ABC -machines. This special case yields the more constrained set of equations:

$$Q_\xi(n) = 3 \langle \boldsymbol{\pi} | \mathcal{T}^{[x_0]} \mathcal{T}^{n-1} \mathcal{T}^{[\hat{\xi}(x_0)]} | \mathbf{1} \rangle, \tag{5.15}$$

where $x_0 \in \mathcal{A}_P$.

5.5.2 CFs from Spectral Decomposition

Although Eq. (5.13) is itself an important result, we can also apply a spectral decomposition of powers of the TM to provide a closed-form that is even more useful and insightful.

Ameliorating the computational burden, this result reduces the matrix powers in the above expressions to expressions involving only powers of scalars. Also, yielding theoretical insight, the closed-forms reveal what types of behaviors can ever be expected of the CFs from stacking processes described by finite HMMs.

The most familiar case occurs when the TM is diagonalizable. Then, \mathcal{T}^{n-1} can be found via diagonalizing the TM, making use of the fact that $\mathcal{T}^L = CD^LC^{-1}$, given the eigen-decomposition $\mathcal{T} = CDC^{-1}$, where D is the diagonal matrix of eigenvalues. However, to understand the CF behavior, it is more appropriate to decompose the matrix in terms of its projection operators.

Moreover, an analytic expression for \mathcal{T}^{n-1} can be found in terms of the projection operators even when the TM is not diagonalizable. Details are given elsewhere [60, 198]. By way of summarizing, though, in the general case the L^{th} iteration of the TM follows from:

$$\mathcal{T}^L = \mathcal{Z}^{-1} \left\{ (\mathbb{I} - z^{-1}\mathcal{T})^{-1} \right\} , \quad (5.16)$$

where \mathbb{I} is the $M_P \times M_P$ identity matrix, $z \in \mathbb{C}$ is a continuous complex variable, and $\mathcal{Z}^{-1}\{\cdot\}$ denotes the inverse z -transform [182] defined to operate elementwise:

$$\mathcal{Z}^{-1}(g(z)) \equiv \frac{1}{2\pi i} \oint_C z^{L-1} g(z) dz \quad (5.17)$$

for the z -dependent matrix element $g(z)$ of $(\mathbb{I} - z^{-1}\mathcal{T})^{-1}$. Here, \oint_C indicates a counter-clockwise contour integration in the complex plane enclosing the entire unit circle.

For nonnegative integers L , and with the allowance that $0^L = \delta_{L,0}$ for the case that $0 \in \Lambda_{\mathcal{T}}$, Eq. (5.16) becomes:

$$\mathcal{T}^L = \sum_{\lambda \in \Lambda_{\mathcal{T}}} \sum_{m=0}^{\nu_{\lambda}-1} \lambda^{L-m} \binom{L}{m} \mathcal{T}_{\lambda} (\mathcal{T} - \lambda \mathbb{I})^m , \quad (5.18)$$

where $\Lambda_{\mathcal{T}} = \{\lambda \in \mathbb{C} : \det(\lambda \mathbb{I} - \mathcal{T}) = 0\}$ is the set of \mathcal{T} 's eigenvalues, \mathcal{T}_{λ} is the projection operator associated with the eigenvalue λ given by the elementwise residue of the resolvent $(z\mathbb{I} - \mathcal{T})^{-1}$ at $z \rightarrow \lambda$, the index ν_{λ} of the eigenvalue λ is the size of the largest Jordan block associated with λ , and $\binom{L}{m} = \frac{L!}{m!(L-m)!}$ is the binomial coefficient.⁸ In terms of elementwise

⁸Recall, *e.g.*, that $\binom{L}{0} = 1$, $\binom{L}{1} = L$, $\binom{L}{2} = \frac{1}{2}L(L-1)$, and $\binom{L}{L} = 1$.

contour integration, we have:

$$\mathcal{T}_\lambda = \frac{1}{2\pi i} \oint_{C_\lambda} (z\mathbb{I} - \mathcal{T})^{-1} dz, \quad (5.19)$$

where C_λ is any contour in the complex plane enclosing the point $z_0 = \lambda$ —which may or may not be a singularity depending on the particular element of the resolvent matrix—but encloses no other singularities.

As guaranteed by the Perron–Frobenius theorem, all eigenvalues of the stochastic TM \mathcal{T} lie on or within the unit circle. Moreover, the eigenvalues on the unit circle are guaranteed to have index one. The indices of all other eigenvalues must be less than or equal to one more than the difference between their algebraic a_λ and geometric g_λ multiplicities. Specifically:

$$\nu_\lambda - 1 \leq a_\lambda - g_\lambda \leq a_\lambda - 1, \text{ and } a_\lambda = g_\lambda \text{ if } |\lambda| = 1 .$$

Using Eq. (6.19) together with Eq. (5.13), the CFs can now be expressed as:

$$Q_\xi(n) = \sum_{\lambda \in \Lambda_\mathcal{T}} \sum_{m=0}^{\nu_\lambda-1} \langle \mathcal{T}_{\lambda,m}^{\xi(\mathcal{A})} \rangle \binom{n-1}{m} \lambda^{n-m-1} , \quad (5.20)$$

where $\langle \mathcal{T}_{\lambda,m}^{\xi(\mathcal{A})} \rangle$ is a complex-valued scalar:⁹

$$\langle \mathcal{T}_{\lambda,m}^{\xi(\mathcal{A})} \rangle \equiv \sum_{x_0 \in \mathcal{A}_\mathbf{P}} \langle \boldsymbol{\pi} | \mathcal{T}^{[x_0]} \mathcal{T}_\lambda (\mathcal{T} - \lambda \mathbb{I})^m \mathcal{T}^{[\hat{\xi}(x_0)]} | \mathbf{1} \rangle . \quad (5.21)$$

Evidently, the CFs’ mathematical form Eq. (5.20) is strongly constrained for any stacking process that can be described by a finite HMM. Besides the expression’s elegance, we note that its constrained form is very useful for the so-called “inverse problem” of discovering the stacking process from CFs [264, 265, 266, 270].

When \mathcal{T} is diagonalizable, $\nu_\lambda = 1$ for all λ so that Eq. (6.19) simply reduces to:

$$\mathcal{T}^L = \sum_{\lambda \in \Lambda_\mathcal{T}} \lambda^L \mathcal{T}_\lambda , \quad (5.22)$$

⁹ $\langle \mathcal{T}_{\lambda,m}^{\xi(\mathcal{A})} \rangle$ is constant with respect to the relative layer displacement n . However, $\{ \langle \mathcal{T}_{\lambda,m}^{\xi(\mathcal{A})} \rangle \}$ can be a function of a process’s parameters.

where the projection operators can be obtained more simply as:

$$\mathcal{T}_\lambda = \prod_{\substack{\zeta \in \Lambda_{\mathcal{T}} \\ \zeta \neq \lambda}} \frac{\mathcal{T} - \zeta \mathbb{I}}{\lambda - \zeta} . \quad (5.23)$$

In the diagonalizable case, Eq. (5.20) reduces to:

$$\begin{aligned} Q_\xi(n) &= \sum_{\lambda \in \Lambda_{\mathcal{T}}} \lambda^{n-1} \sum_{x_0 \in \mathcal{A}_P} \langle \boldsymbol{\pi} | \mathcal{T}^{[x_0]} \mathcal{T}_\lambda \mathcal{T}^{[\hat{\xi}(x_0)]} | \mathbf{1} \rangle \\ &= \sum_{\lambda \in \Lambda_{\mathcal{T}}} \left\langle \mathcal{T}_\lambda^{\xi(\mathcal{A})} \right\rangle \lambda^{n-1} , \end{aligned} \quad (5.24)$$

where $\left\langle \mathcal{T}_\lambda^{\xi(\mathcal{A})} \right\rangle \equiv \left\langle \mathcal{T}_{\lambda,0}^{\xi(\mathcal{A})} \right\rangle$ is again a constant:

$$\left\langle \mathcal{T}_\lambda^{\xi(\mathcal{A})} \right\rangle = \sum_{x_0 \in \mathcal{A}_P} \langle \boldsymbol{\pi} | \mathcal{T}^{[x_0]} \mathcal{T}_\lambda \mathcal{T}^{[\hat{\xi}(x_0)]} | \mathbf{1} \rangle . \quad (5.25)$$

5.5.3 Asymptotic behavior of the CFs

From the spectral decomposition, it is apparent that the CFs converge to some constant value as $n \rightarrow \infty$, unless \mathcal{T} has eigenvalues on the unit circle besides unity itself. If unity is the sole eigenvalue with a magnitude of one, then all other eigenvalues have magnitude less than unity and their contributions decay to negligibility for large enough n . Explicitly, if $\operatorname{argmax}_{\lambda \in \Lambda_{\mathcal{T}}} |\lambda| = \{1\}$, then:

$$\begin{aligned} \lim_{n \rightarrow \infty} Q_\xi(n) &= \lim_{n \rightarrow \infty} \sum_{\lambda \in \Lambda_{\mathcal{T}}} \sum_{m=0}^{\nu_\lambda-1} \left\langle \mathcal{T}_{\lambda,m}^{\xi(\mathcal{A})} \right\rangle \binom{n-1}{m} \lambda^{n-m-1} \\ &= \left\langle \mathcal{T}_1^{\xi(\mathcal{A})} \right\rangle \\ &= \sum_{x_0 \in \mathcal{A}_P} \langle \boldsymbol{\pi} | \mathcal{T}^{[x_0]} \mathcal{T}_1 \mathcal{T}^{[\hat{\xi}(x_0)]} | \mathbf{1} \rangle \\ &= \sum_{x_0 \in \mathcal{A}_P} \langle \boldsymbol{\pi} | \mathcal{T}^{[x_0]} | \mathbf{1} \rangle \langle \boldsymbol{\pi} | \mathcal{T}^{[\hat{\xi}(x_0)]} | \mathbf{1} \rangle \\ &= \sum_{x_0 \in \mathcal{A}_P} \operatorname{Pr}(x_0) \operatorname{Pr}(\hat{\xi}(x_0)) . \end{aligned}$$

Above, we used the fact that $\nu_1 = 1$ and that—for an ergodic process— $\mathcal{T}_1 = |\mathbf{1}\rangle \langle \boldsymbol{\pi}|$.

For mixing *ABC*-machines, $\operatorname{Pr}(x) = 1/3$ for all $x \in \mathcal{A}_P$. That this is so should be evident from the graphical expansion method of §5.4.1. Therefore, mixing processes with

$\operatorname{argmax}_{\lambda \in \Lambda_{\mathcal{T}}} |\lambda| = \{1\}$ have CFs that all converge to $1/3$:

$$\begin{aligned} \lim_{n \rightarrow \infty} Q_{\xi}(n) &= \sum_{x_0 \in \mathcal{A}_{\mathcal{P}}} \Pr(x_0) \Pr(\hat{\xi}(x_0)) \\ &= 3(\tfrac{1}{3} \times \tfrac{1}{3}) \\ &= \tfrac{1}{3}. \end{aligned}$$

Non-mixing processes with $\operatorname{argmax}_{\lambda \in \Lambda_{\mathcal{T}}} |\lambda| = \{1\}$ can have their CFs converging to constants other than $1/3$, depending on $\{\Pr(x) : x \in \mathcal{A}_{\mathcal{P}}\}$, although they are still constrained by $\sum_{\xi} Q_{\xi}(n) = 1$.

If other eigenvalues in $\Lambda_{\mathcal{T}}$ beside unity exist on the unit circle, then the CFs approach a periodic sequence as n gets large.

5.5.4 Modes of Decay

Since \mathcal{T} has no more eigenvalues than its dimension (*i.e.*, $|\Lambda_{\mathcal{T}}| \leq M_{\mathcal{P}}$), *Eq. (5.20) implies that the number of states in the ABC-machine for a stacking process puts an upper bound on the number of modes of decay.* Indeed, since unity is associated with stationarity, the number of modes of decay is strictly less than $M_{\mathcal{P}}$. It is important to note that these modes do not always decay strictly exponentially: They are in general the product of a decaying exponential with a polynomial in n , and the CFs are sums of these products.

Even if—due to diagonalizability of \mathcal{T} —there were only strictly exponentially decaying modes, it is simple but important to understand that there is generally more than one mode of exponential decay present in the CFs. And so, ventures to find *the* decay constant of a process are misleading unless it is explicitly acknowledged that one seeks, *e.g.*, the slowest decay mode. Even then, however, there are cases when the slowest decay mode only acts on a component of the CFs with negligible magnitude. In an extreme case, the slowest decay mode may not even be a large contributor to the CFs before the whole pattern is numerically indistinguishable from the asymptotic value.

In analyzing a broad range of CFs, nevertheless, many authors have been led to consider *correlation lengths*, also known as *characteristic lengths* [248, 270]. The form of Eq. (5.20) suggests that this perspective will often be a clumsy oversimplification for understanding

CFs. Regardless, if one wishes to assign a correlation length associated with an index-one mode of CF decay, we observe that the reciprocal of the correlation length is essentially the negative logarithm of the magnitude of the eigenvalue for that mode. We find that the typically reported correlation length ℓ_C derives from the second-largest contributing magnitude among the eigenvalues:

$$\ell_C^{-1} = -\log |\zeta| , \quad \text{for } \zeta \in \underset{\lambda \in \boldsymbol{\rho}}{\operatorname{argmax}} |\lambda| , \quad (5.26)$$

where $\boldsymbol{\rho} = \left\{ \lambda \in \Lambda_{\mathcal{T}} \setminus \{1\} : \left\langle \mathcal{T}_{\lambda}^{\xi(\mathcal{A})} \right\rangle \neq 0 \right\}$.

Guided by Eq. (5.20), we suggest that a true understanding of CF behavior involves finding $\Lambda_{\mathcal{T}}$ with the corresponding eigenvalue indices and the amplitude of each mode's contribution $\left\{ \left\langle \mathcal{T}_{\lambda,m}^{\xi(\mathcal{A})} \right\rangle \right\}$.

This now completes our theoretical development, and in the next section we apply these techniques to three examples.

5.6 Examples

5.6.1 3C Polytypes and Random ML Stacking: IID Processes

Although real materials often display much more complex behaviors, as a pedagogical exercise the random stacking of MLs in CPSs has often been treated [103] to study stacking faults. This stacking process is the simplest stacking arrangement that can be imagined,¹⁰ and there are previous analytical results that can be compared to the techniques developed here. In statistics parlance, this process is an *Independent and Identically Distributed* (IID) Process [46].

Let us assume that the placement of MLs is independent of the previous MLs scanned, except that it of course must obey the stacking constraints. The Hägg-machine that describes this process is shown in Fig. 5.3. We allow for the possibility that there might be a bias in the stacking order, and we assign a probability q that the next layer is cyclically

¹⁰This is not mere hyperbole. It is possible to quantify a process's structural organization in the form of its *statistical complexity* C_{μ} , which measures the internal information processing required to produce the pattern [64, 55, 266]. In the present case $C_{\mu} = 0$ bits, the minimum value.

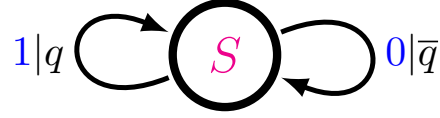


Figure 5.3: Hägg-machine for the IID Process. When $q = 1$, the IID Process generates a string of 1s, which is physically the $3C^+$ stacking structure. Conversely, when $q = 0$, the structure corresponds to the $3C^-$ structure. For $q = 1/2$, the MLs are stacked as randomly as possible. Here and elsewhere we adopt the convention that a bar over a variable means one minus that variable, *i.e.* $\bar{q} \equiv 1 - q$, with $q \in [0, 1]$.

related to its predecessor. Thus, the 1-by-1 symbol-labeled TMs for the Hägg-machine are:

$$\mathsf{T}^{[1]} = \begin{bmatrix} q \end{bmatrix} \text{ and } \mathsf{T}^{[0]} = \begin{bmatrix} \bar{q} \end{bmatrix},$$

where $\bar{q} \equiv 1 - q$, with $q \in [0, 1]$.

The physical interpretation of the IID Process is straightforward. In the case where $q = 1$, the process generates a stacking sequence of all 1s, giving a physical stacking structure of $\dots ABCABCABC \dots$. We recognize this as the $3C^+$ crystal structure. Similarly, for $q = 0$, the process generates stacking sequence of all 0s, which is the $3C^-$ crystal structure. For those cases where q is near but not quite at its extreme values, the stacking structure is 3C with randomly distributed deformation faults. When $q = \frac{1}{2}$, the MLs are stacked in a completely random fashion.

Now, we must determine whether this is a mixing or nonmixing Hägg-machine. We note that there are two SSCs, namely $[\mathsf{S}_{(0)}]$ and $[\mathsf{S}_{(1)}]$. The winding numbers for each are $W^{[\mathsf{S}_{(1)}]} = 1$ and $W^{[\mathsf{S}_{(0)}]} = 2$, respectively. Since at least one of these is not equal to zero, the Hägg-machine is mixing, and we need to expand the Hägg-machine into the *ABC*-machine. This is shown in Fig. 5.4.

The *ABC*-machine TMs can either be directly written down from inspecting Fig. 5.4 or by using the rote expansion algorithm of §5.4.3, using Eqs. (5.2) and (5.3). By either method we find the 3-by-3 TMs to be:

$$\mathcal{T}^{[A]} = \begin{bmatrix} 0 & 0 & 0 \\ \bar{q} & 0 & 0 \\ q & 0 & 0 \end{bmatrix}, \quad \mathcal{T}^{[B]} = \begin{bmatrix} 0 & q & 0 \\ 0 & 0 & 0 \\ 0 & \bar{q} & 0 \end{bmatrix} \text{ and } \mathcal{T}^{[C]} = \begin{bmatrix} 0 & 0 & \bar{q} \\ 0 & 0 & q \\ 0 & 0 & 0 \end{bmatrix}.$$

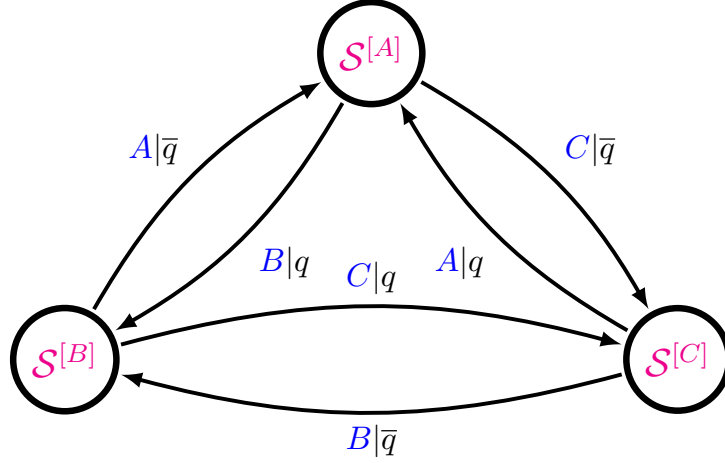


Figure 5.4: ABC -machine for the IID Process. The single state of the Hägg-machine has expanded into three.

The internal state TM then is their sum:

$$\mathcal{T} = \begin{bmatrix} 0 & q & \bar{q} \\ \bar{q} & 0 & q \\ q & \bar{q} & 0 \end{bmatrix} .$$

The eigenvalues of the ABC TM are

$$\Lambda_{\mathcal{T}} = \{1, \Omega, \Omega^*\} ,$$

where:

$$\Omega \equiv -\frac{1}{2} + i\frac{\sqrt{3}}{2}(4q^2 - 4q + 1)^{1/2}$$

and Ω^* is its complex conjugate. Already, via Eq. (5.26), we can identify what the characteristic length of the CFs will be. In particular, $\ell_C^{-1} = -\log |\Omega| = -\frac{1}{2} \log(1 - 3q + 3q^2)$ yields:

$$\ell_C = -\frac{2}{\log(1 - 3q + 3q^2)} .$$

If we identify q with the deformation faulting parameter α in the model introduced by Estevez *et al.* [85] (see the next example in §6.6.2, the RGDF Process), this is identical to the result obtained there in Eq. (35). There is much more structural information in the

CFs, however, than a single characteristic length would suggest. This fact will become especially apparent as our examples become more sophisticated.

According to Eq. (5.13), we can obtain the CFs via:

$$Q_\xi(n) = \sum_{x_0 \in \mathcal{A}_P} \langle \boldsymbol{\pi} | \mathcal{T}^{[x_0]} \mathcal{T}^{n-1} \mathcal{T}^{[\hat{\xi}(x_0)]} | \mathbf{1} \rangle .$$

The stationary distribution over the *ABC*-machine states is found from Eq. (5.1):

$$\langle \boldsymbol{\pi} | = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix} .$$

Furthermore, an analytic expression for \mathcal{T}^{n-1} follows from the z -transform as given in Eq. (5.16). As a start, we find:

$$\mathbb{I} - z^{-1} \mathcal{T} = \begin{bmatrix} 1 & -q/z & -\bar{q}/z \\ -\bar{q}/z & 1 & -q/z \\ -q/z & -\bar{q}/z & 1 \end{bmatrix}$$

and its inverse:

$$\begin{aligned} (\mathbb{I} - z^{-1} \mathcal{T})^{-1} &= \frac{1}{(1 - z^{-1})(1 - \Omega z^{-1})(1 - \Omega^* z^{-1})} \\ &\times \begin{bmatrix} 1 - q\bar{q}z^{-2} & qz^{-1} + \bar{q}^2 z^{-2} & \bar{q}z^{-1} + q^2 z^{-2} \\ \bar{q}z^{-1} + q^2 z^{-2} & 1 - q\bar{q}z^{-2} & qz^{-1} + \bar{q}^2 z^{-2} \\ qz^{-1} + \bar{q}^2 z^{-2} & \bar{q}z^{-1} + q^2 z^{-2} & 1 - q\bar{q}z^{-2} \end{bmatrix} . \end{aligned}$$

Upon partial fraction expansion, we obtain:

$$\begin{aligned}
& (\mathbb{I} - z^{-1}\mathcal{T})^{-1} \\
&= \frac{1}{3} \frac{1}{(1 - z^{-1})} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \\
&+ \frac{1}{(\Omega-1)(\Omega-\Omega^*)} \frac{1}{(1-\Omega z^{-1})} \begin{bmatrix} \Omega^2 - q\bar{q} & q\Omega + \bar{q}^2 & \bar{q}\Omega + q^2 \\ \bar{q}\Omega + q^2 & \Omega^2 - q\bar{q} & q\Omega + \bar{q}^2 \\ q\Omega + \bar{q}^2 & \bar{q}\Omega + q^2 & \Omega^2 - q\bar{q} \end{bmatrix} \\
&+ \frac{1}{(\Omega^*-1)(\Omega^*-\Omega)} \frac{1}{(1-\Omega^* z^{-1})} \begin{bmatrix} \Omega^{*2} - q\bar{q} & q\Omega^* + \bar{q}^2 & \bar{q}\Omega^* + q^2 \\ \bar{q}\Omega^* + q^2 & \Omega^{*2} - q\bar{q} & q\Omega^* + \bar{q}^2 \\ q\Omega^* + \bar{q}^2 & \bar{q}\Omega^* + q^2 & \Omega^{*2} - q\bar{q} \end{bmatrix}, \tag{5.27}
\end{aligned}$$

for $q \neq 1/2$. (The special case of $q = 1/2$ is discussed in the next subsection.) Finally, we take the inverse z -transform of Eq. (5.27) to obtain an expression for the L^{th} iterate of the TM:

$$\begin{aligned}
\mathcal{T}^L &= \mathcal{Z}^{-1} \left\{ (\mathbb{I} - z^{-1}\mathcal{T})^{-1} \right\} \\
&= \frac{1}{3} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \\
&+ \frac{\Omega^L}{(\Omega-1)(\Omega-\Omega^*)} \begin{bmatrix} \Omega^2 - q\bar{q} & q\Omega + \bar{q}^2 & \bar{q}\Omega + q^2 \\ \bar{q}\Omega + q^2 & \Omega^2 - q\bar{q} & q\Omega + \bar{q}^2 \\ q\Omega + \bar{q}^2 & \bar{q}\Omega + q^2 & \Omega^2 - q\bar{q} \end{bmatrix} \\
&+ \frac{\Omega^{*L}}{(\Omega^*-1)(\Omega^*-\Omega)} \begin{bmatrix} \Omega^{*2} - q\bar{q} & q\Omega^* + \bar{q}^2 & \bar{q}\Omega^* + q^2 \\ \bar{q}\Omega^* + q^2 & \Omega^{*2} - q\bar{q} & q\Omega^* + \bar{q}^2 \\ q\Omega^* + \bar{q}^2 & \bar{q}\Omega^* + q^2 & \Omega^{*2} - q\bar{q} \end{bmatrix}.
\end{aligned}$$

These pieces are all we need to calculate the CFs. Let's start with $Q_s(n)$. First, we find:

$$\langle \boldsymbol{\pi} | \mathcal{T}^{[A]} = \begin{bmatrix} \frac{1}{3} & 0 & 0 \end{bmatrix}$$

and:

$$\mathcal{T}^{[\hat{s}(A)]} |\mathbf{1}\rangle = \mathcal{T}^{[A]} |\mathbf{1}\rangle = \begin{bmatrix} 0 \\ \bar{q} \\ q \end{bmatrix}.$$

Then:

$$\begin{aligned} \langle \boldsymbol{\pi} | \mathcal{T}^{[A]} \mathcal{T}^{n-1} &= \frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \\ &+ \frac{1}{3} \frac{\Omega^{n-1}}{(\Omega-1)(\Omega-\Omega^*)} \begin{bmatrix} \Omega^2 - q\bar{q} & q\Omega + \bar{q}^2 & \bar{q}\Omega + q^2 \end{bmatrix} \\ &+ \frac{1}{3} \frac{\Omega^{*n-1}}{(\Omega^*-1)(\Omega^*-\Omega)} \begin{bmatrix} \Omega^{*2} - q\bar{q} & q\Omega^* + \bar{q}^2 & \bar{q}\Omega^* + q^2 \end{bmatrix} \end{aligned} \quad (5.28)$$

and:

$$\begin{aligned} \langle \boldsymbol{\pi} | \mathcal{T}^{[A]} \mathcal{T}^{n-1} \mathcal{T}^{[A]} |\mathbf{1}\rangle &= \frac{1}{9} + \frac{1}{3} \frac{\Omega^{n-1}}{(\Omega-1)(\Omega-\Omega^*)} (2q\bar{q}\Omega + \Omega\Omega^*) \\ &+ \frac{1}{3} \frac{\Omega^{*n-1}}{(\Omega^*-1)(\Omega^*-\Omega)} (2q\bar{q}\Omega^* + \Omega\Omega^*). \end{aligned} \quad (5.29)$$

One can verify that Eq. (5.15) can be applied in lieu of Eq. (5.13), which saves some effort in finding the final result, which is:

$$Q_s(n) = 1/3 + 2\text{Re} \left\{ \frac{\Omega^n}{(\Omega-1)(\Omega-\Omega^*)} (2q\bar{q} + \Omega^*) \right\}. \quad (5.30)$$

The cyclic and anticyclic CFs can also be calculated from Eq. (5.15) using the result we have already obtained in Eq. (5.28) and a quick calculation yields:

$$\mathcal{T}^{[\hat{c}(A)]} |\mathbf{1}\rangle = \mathcal{T}^{[B]} |\mathbf{1}\rangle = \begin{bmatrix} q \\ 0 \\ \bar{q} \end{bmatrix}$$

and

$$\mathcal{T}^{[\hat{a}(A)]} |\mathbf{1}\rangle = \mathcal{T}^{[C]} |\mathbf{1}\rangle = \begin{bmatrix} \bar{q} \\ q \\ 0 \end{bmatrix}.$$

Then, we have:

$$\begin{aligned} Q_c(n) &= 3 \langle \boldsymbol{\pi} | \mathcal{T}^{[A]} \mathcal{T}^{n-1} \mathcal{T}^{[B]} | \mathbf{1} \rangle \\ &= 1/3 + 2\text{Re} \left\{ \frac{\Omega^n}{(\Omega - 1)(\Omega - \Omega^*)} (\bar{q}^2 + q\Omega) \right\} \end{aligned} \quad (5.31)$$

and:

$$\begin{aligned} Q_a(n) &= 3 \langle \boldsymbol{\pi} | \mathcal{T}^{[A]} \mathcal{T}^{n-1} \mathcal{T}^{[C]} | \mathbf{1} \rangle \\ &= 1/3 + 2\text{Re} \left\{ \frac{\Omega^n}{(\Omega - 1)(\Omega - \Omega^*)} (q^2 + \bar{q}\Omega) \right\}. \end{aligned} \quad (5.32)$$

All of this subsection's results hold for the whole range of $q \in [0, \frac{1}{2}) \cup (\frac{1}{2}, 1]$, where all \mathcal{T} 's eigenvalues are distinct. However, for $q = 1/2$, the two complex conjugate eigenvalues, Ω and Ω^* , lose their imaginary components, becoming repeated eigenvalues. This requires special treatment.¹¹ We address the case of $q = 1/2$ in the next subsection, which is of interest in its own right as being the most random possible stacking sequence allowed.

5.6.1.1 A Fair Coin?

For a completely random stacking of MLs, such that $q = 1/2$, the relative orientations of the MLs are effectively assigned by a fair coin. The resulting TM is symmetric with repeated eigenvalues, implying that superficially at least, the CFs take on a special form.¹²

To obtain the CFs for the Fair Coin IID Process, we follow the procedure of the previous subsection, with all of the same results through Eq. (5.27), which with $q = 1/2$ and $\Omega|_{q=1/2} = \Omega^*|_{q=1/2} = -1/2$ can now be written as:

$$\begin{aligned} (\mathbb{I} - z^{-1}\mathcal{T})^{-1} &= \frac{1}{(1 - z^{-1})(1 + \frac{1}{2}z^{-1})^2} \\ &\times \begin{bmatrix} 1 - \frac{1}{4}z^{-2} & \frac{1}{2}z^{-1} + \frac{1}{4}z^{-2} & \frac{1}{2}z^{-1} + \frac{1}{4}z^{-2} \\ \frac{1}{2}z^{-1} + \frac{1}{4}z^{-2} & 1 - \frac{1}{4}z^{-2} & \frac{1}{2}z^{-1} + \frac{1}{4}z^{-2} \\ \frac{1}{2}z^{-1} + \frac{1}{4}z^{-2} & \frac{1}{2}z^{-1} + \frac{1}{4}z^{-2} & 1 - \frac{1}{4}z^{-2} \end{bmatrix}. \end{aligned}$$

¹¹Indeed, the straightforward z -transform approach yielding the CF equations given in this section appears to need special treatment for $q = 1/2$. However, a more direct spectral perspective as developed in §5.5.2 shows that since \mathcal{T} is diagonalizable for all q , all eigenvalues have index of one and so yield CFs of the simple form of Eq. (5.24).

¹²But since the TM remains diagonalizable, the CFs retain the simple form of Eq. (5.24).

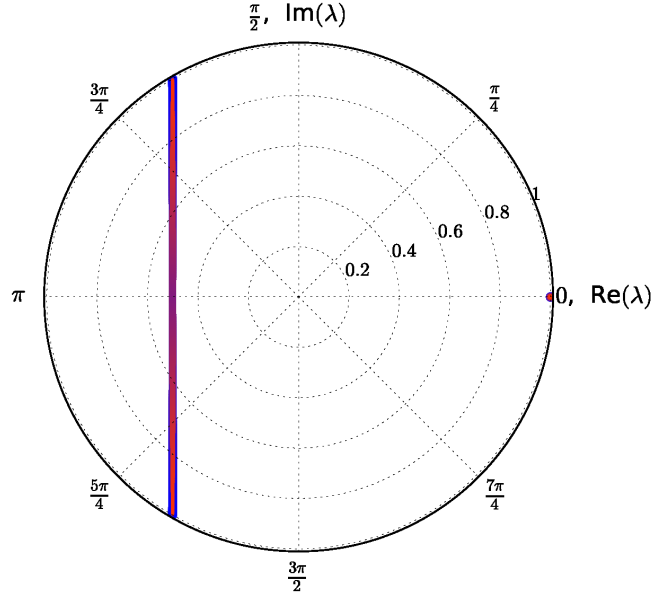


Figure 5.5: TM's eigenvalues in the complex plane for the IID Process as q is varied. Note that there is always an eigenvalue at 1.

However, the repeated factor in the denominator yields a new partial fraction expansion. Applying the inverse z -transform gives the L^{th} iterate of the TM¹³ as:

$$\begin{aligned}\mathcal{T}^L &= \mathcal{Z}^{-1} \left\{ (\mathbb{I} - z^{-1}\mathcal{T})^{-1} \right\} \\ &= \frac{1}{3} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} + \frac{1}{3} \left(-\frac{1}{2} \right)^L \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}.\end{aligned}$$

Then, we find:

$$\langle \boldsymbol{\pi} | \mathcal{T}^{[A]} \mathcal{T}^{n-1} = \frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} + \frac{1}{9} \left(-\frac{1}{2} \right)^{n-1} \begin{bmatrix} 2 & -1 & -1 \end{bmatrix}.$$

With the final result that:

$$\begin{aligned}Q_s(n) &= 3 \langle \boldsymbol{\pi} | \mathcal{T}^{[A]} \mathcal{T}^{n-1} \mathcal{T}^{[A]} | \mathbf{1} \rangle \\ &= \frac{1}{3} + \frac{2}{3} \left(-\frac{1}{2} \right)^n,\end{aligned}\tag{5.33}$$

¹³By inspection, we can verify that this decomposition still yields \mathcal{T}^0 as the identity matrix and $\mathcal{T}^1 = \mathcal{T}$, as must be the case. More interestingly, the decaying deviation from the asymptotic matrix is oscillatory.

$$\begin{aligned}
Q_c(n) &= 3 \langle \pi | \mathcal{T}^{[A]} \mathcal{T}^{n-1} \mathcal{T}^{[B]} | 1 \rangle \\
&= \frac{1}{3} - \frac{1}{3} \left(-\frac{1}{2} \right)^n,
\end{aligned} \tag{5.34}$$

and

$$\begin{aligned}
Q_a(n) &= 3 \langle \pi | \mathcal{T}^{[A]} \mathcal{T}^{n-1} \mathcal{T}^{[C]} | 1 \rangle \\
&= \frac{1}{3} - \frac{1}{3} \left(-\frac{1}{2} \right)^n.
\end{aligned} \tag{5.35}$$

For $q = 1/2$, we see that $Q_c(n)$ and $Q_a(n)$ are identical, but this is not generally the case as one can check for other values of q in Eqs. (5.31) and (5.32).

Figure 5.5 shows a graph of the TM's eigenvalues in the complex plane as q is varied. Notice that there is an eigenvalue at 1 for all values of q . This is generic feature, and we always find such an eigenvalue. The other two eigenvalues start at the other two cube roots of unity for $q \in \{0, 1\}$ and, as $q \rightarrow 1/2$, they migrate to the point $-1/2$ and become degenerate at $q = 1/2$. It is this degeneracy that requires the special treatment given in this section.

It is interesting that even the Fair Coin Hägg-machine produces structured CFs. This is because—even though the allowed transitions of the underlying ABC -machine are randomized—not all transitions are allowed. For example, if we start with an A ML, the next ML has a zero probability of being an A , a $1/2$ probability of being a B , and a $1/2$ probability of being a C . Then, the *next* ML has a rebounding $1/2$ probability of being an A while the probability of being either a B or C is each only $1/4$. So, we see that the underlying process has structure, and there is nothing we can do—given the physical constraints—to make the CFs completely random.

When we can compare our expressions for CFs at $q = 1/2$ to those derived previously by elementary means [103, 262], we find agreement. Note however that unlike in these earlier treatments, here there was no need to assume a recursion relationship.

Figures 5.6 and 5.7 show $Q_s(n)$ versus n for the IID Process with $q = 0.1$ and $q = 0.3$, respectively, as computed from Eq. (5.30). In each case the CFs decay to an asymptotic value of $1/3$, although this decay is faster for $q = 0.3$. This is not surprising, as one

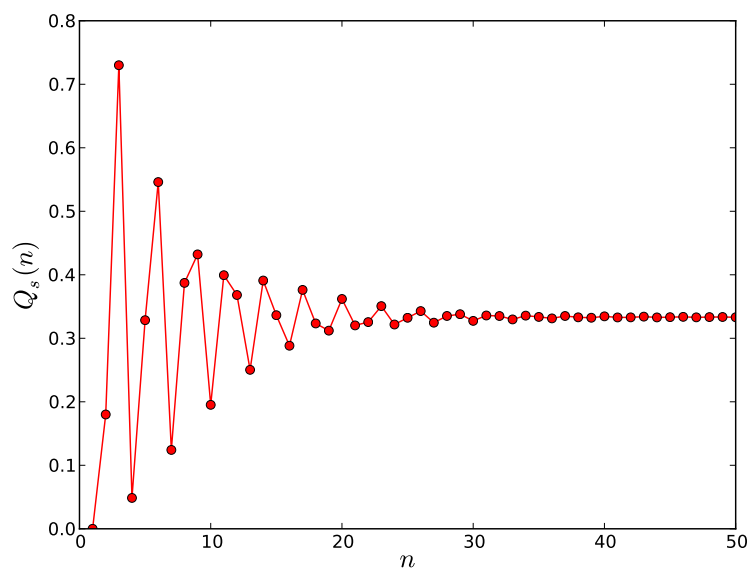


Figure 5.6: $Q_s(n)$ vs. n for $q = 0.1$ the IID Process.

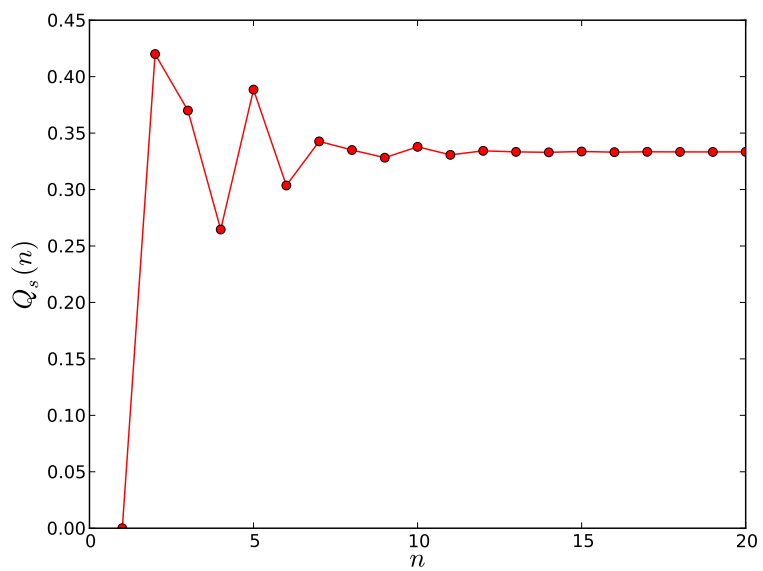


Figure 5.7: $Q_s(n)$ vs. n for $q = 0.3$ the IID Process.

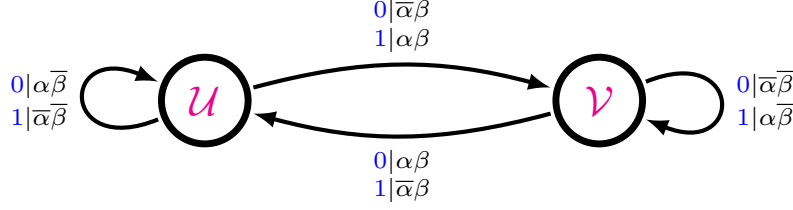


Figure 5.8: RGDF Process, first proposed by Estevez *et al.* [85] and adapted here from Panel (c) of their Fig. (2). There is a slight change in notation. We relabeled the states given as ‘f’ and ‘b’ by Estevez *et al.* [85] as ‘ \mathcal{U} ’ and ‘ \mathcal{V} ’ and, instead of drawing an arc for each of the possible eight transitions, we took advantage of the multiple transitions between the same states and labeled each arc with two transitions. There is, of course, no change in meaning; this instead provides for slightly tidier illustration. Additionally, we correct a typographical error in Estevez *et al.* [85] when we relabel the transition $b \xrightarrow{0|\alpha\beta} b$ with $\mathcal{V} \xrightarrow{0|\alpha\beta} \mathcal{V}$.

interpretation for the IID Process with $q = 0.1$ is that of a $3C^+$ crystal interspersed with 10% random deformation faults.

5.6.2 Random Growth and Deformation Faults in Layered 3C and 2H CPSs: The RGDF Process

Estevez *et al.* [85] recently showed that simultaneous random growth and deformation SFs in 2H and 3C CPSs can be modeled for all values of the fault parameters by a simple HMM, and this is shown in Fig. 6.3. We refer to this process as the *Random Growth and Deformation Faults* (RGDF) Process.¹⁴ As has become convention [273, 85], α refers to deformation faulting and β refers to growth faults.

The HMM describing the RGDF Process is unlike any of the others considered here in that on emission of a symbol from a state, the successor state is *not* uniquely specified. For example, $\mathcal{U} \xrightarrow{0} \mathcal{U}$ and $\mathcal{U} \xrightarrow{0} \mathcal{V}$; *i.e.*, being in state \mathcal{U} and emitting a 0 does not uniquely determine the next state. Such a representations were previously called *nondeterministic* [112], but to avoid a conflict in terminology we prefer the term *nonunifilar* [78, 76]. Since ϵ -machines are unifilar [64, 223], the HMM representing the RGDF Process is not an ϵ -machine. Nonetheless, the techniques we have developed are applicable: CFs do not

¹⁴ Estevez *et al.* [85] give a thorough and detailed discussion of the RGDF Process, and readers interested in a comprehensive motivation and derivation of the RGDF Process are urged to consult that reference.

require unifilar HMMs for their calculations, as do other properties such as the entropy density.

Inspecting Fig. 6.3, the RGDF Hägg-machine's TMs are seen to be (Eqs. (1) and (2) of [85]):

$$\mathsf{T}^{[0]} = \begin{bmatrix} \alpha\bar{\beta} & \bar{\alpha}\beta \\ \alpha\beta & \bar{\alpha}\bar{\beta} \end{bmatrix} \text{ and } \mathsf{T}^{[1]} = \begin{bmatrix} \bar{\alpha}\bar{\beta} & \alpha\beta \\ \bar{\alpha}\beta & \alpha\bar{\beta} \end{bmatrix},$$

where $\alpha \in [0, 1]$ and $\bar{\alpha} \equiv 1 - \alpha$, such that $\alpha + \bar{\alpha} = 1$, and $\beta \in [0, 1]$ and $\bar{\beta} \equiv 1 - \beta$, such that $\beta + \bar{\beta} = 1$. There are eight SSCs and, if at least one of them has $W^{\text{SSC}} \neq 0$, the Hägg-machine is mixing. The self-state transitions each generate a nonvanishing W^{SSC} , so for the Hägg-machine to be nonmixing, these transitions must be absent. Indeed, there are only two SSCs that have vanishing winding numbers, and these are $[\mathcal{U}_{(0)}\mathcal{V}_{(1)}]$ and $[\mathcal{U}_{(1)}\mathcal{V}_{(0)}]$. These, and only these, SSCs can exist if $\bar{\beta} = 0$ and $\alpha \in \{0, 1\}$. Thus, the Hägg-machine is nonmixing only for the parameter settings $\beta = 1$ and $\alpha \in \{0, 1\}$, which corresponds to the 2H crystal structure.

From the Hägg-machine, we obtain the corresponding TMs of the *ABC*-machine for $\alpha, \beta \in (0, 1)$ by the rote expansion method (§5.4.3):

$$\mathcal{T}^{[A]} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ \alpha\bar{\beta} & 0 & 0 & \bar{\alpha}\beta & 0 & 0 \\ \bar{\alpha}\bar{\beta} & 0 & 0 & \alpha\beta & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \alpha\beta & 0 & 0 & \bar{\alpha}\bar{\beta} & 0 & 0 \\ \bar{\alpha}\beta & 0 & 0 & \alpha\bar{\beta} & 0 & 0 \end{bmatrix}, \quad \mathcal{T}^{[B]} = \begin{bmatrix} 0 & \bar{\alpha}\bar{\beta} & 0 & 0 & \alpha\beta & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \alpha\bar{\beta} & 0 & 0 & \bar{\alpha}\beta & 0 \\ 0 & \bar{\alpha}\beta & 0 & 0 & \alpha\bar{\beta} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \alpha\beta & 0 & 0 & \bar{\alpha}\bar{\beta} & 0 \end{bmatrix},$$

and

$$\mathcal{T}^{[C]} = \begin{bmatrix} 0 & 0 & \alpha\bar{\beta} & 0 & 0 & \bar{\alpha}\beta \\ 0 & 0 & \bar{\alpha}\bar{\beta} & 0 & 0 & \alpha\beta \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \alpha\beta & 0 & 0 & \bar{\alpha}\bar{\beta} \\ 0 & 0 & \bar{\alpha}\beta & 0 & 0 & \alpha\bar{\beta} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

and the orientation-agnostic state-to-state TM:

$$\mathcal{T} = \mathcal{T}^{[A]} + \mathcal{T}^{[B]} + \mathcal{T}^{[C]}.$$

Explicitly, we have:

$$\mathcal{T} = \begin{bmatrix} 0 & \bar{\alpha}\bar{\beta} & \alpha\bar{\beta} & 0 & \alpha\beta & \bar{\alpha}\beta \\ \alpha\bar{\beta} & 0 & \bar{\alpha}\bar{\beta} & \bar{\alpha}\beta & 0 & \alpha\beta \\ \bar{\alpha}\bar{\beta} & \alpha\bar{\beta} & 0 & \alpha\beta & \bar{\alpha}\beta & 0 \\ 0 & \bar{\alpha}\beta & \alpha\beta & 0 & \alpha\bar{\beta} & \bar{\alpha}\bar{\beta} \\ \alpha\beta & 0 & \bar{\alpha}\beta & \bar{\alpha}\bar{\beta} & 0 & \alpha\bar{\beta} \\ \bar{\alpha}\beta & \alpha\beta & 0 & \alpha\bar{\beta} & \bar{\alpha}\bar{\beta} & 0 \end{bmatrix}.$$

\mathcal{T} 's eigenvalues satisfy $\det(\mathcal{T} - \lambda\mathbb{I}) = 0$. Here, with $a \equiv \alpha\beta$, $b \equiv \alpha\bar{\beta}$, $c \equiv \bar{\alpha}\beta$, and $d \equiv \bar{\alpha}\bar{\beta}$, we have:

$$\begin{aligned} \det(\mathcal{T} - \lambda\mathbb{I}) &= [(\lambda - (b + d))^2 - (a + c)^2] \\ &\quad \times [\lambda^2 + \lambda(b + d) + ac - bd - a^2 - c^2 + b^2 + d^2]^2 \\ &= 0, \end{aligned}$$

from which we obtain the eigenvalues: $\lambda = b + d \pm (a + c)$ and $\lambda = \frac{1}{2}(b + d) \pm \frac{1}{2}[4(a + c)^2 - 3(b + d)^2 + 12(bd - ac)]^{\frac{1}{2}}$. To get back to α s and β s, we note that $a + c = \beta$, $b + d = \bar{\beta}$, $ac = \beta^2\alpha\bar{\alpha}$, and $bd = \bar{\beta}^2\alpha\bar{\alpha}$. It also follows that $b + d + a + c = 1$, $b + d - (a + c) = \bar{\beta} - \beta = 1 - 2\beta$, and $bd - ac = \alpha\bar{\alpha}(\bar{\beta}^2 - \beta^2) = \alpha\bar{\alpha}(1 - 2\beta) = \alpha\bar{\alpha}(\bar{\beta} - \beta)$. Hence, after simplification, the set of \mathcal{T} 's eigenvalues can be written as:

$$\Lambda_{\mathcal{T}} = \{1, 1 - 2\beta, -\frac{1}{2}(1 - \beta) \pm \frac{1}{2}\sqrt{\sigma}\}, \quad (5.36)$$

with

$$\sigma \equiv 4\beta^2 - 3\bar{\beta}^2 + 12\alpha\bar{\alpha}(\bar{\beta} - \beta) \quad (5.37)$$

$$= -3 + 12\alpha + 6\beta - 12\alpha^2 + \beta^2 - 24\alpha\beta + 24\alpha^2\beta. \quad (5.38)$$

Except for measure-zero submanifolds along which the eigenvalues become extra degenerate, throughout the parameter range the eigenvalues' algebraic multiplicities are: $a_1 = 1$, $a_{1-2\beta} = 1$, $a_{-\frac{1}{2}(1-\beta+\sqrt{\sigma})} = 2$, and $a_{-\frac{1}{2}(1-\beta-\sqrt{\sigma})} = 2$. Moreover, the *index* of all eigenvalues is 1 except along $\sigma = 0$.

Immediately from the eigenvalues and their corresponding indices, we know all possible characteristic modes of CF decay. All that remains is to find the contributing amplitude of each characteristic mode. For comparison, note that our σ turns out to be equivalent to the all-important $-s^2$ term defined in Eq. (28) of Estevez *et al.* [85].

Eqs. (6.37) and (6.38) reveal an obvious symmetry between α and $\bar{\alpha}$ that is *not* present between β and $\bar{\beta}$. In particular, \mathcal{T} 's eigenvalues are invariant under exchange of α and $\bar{\alpha}$ —the CFs will decay in the same manner for α -values symmetric about $1/2$. There is no such symmetry between β and $\bar{\beta}$. Parameter space organization is seen nicely in Panel (c) of Fig. 6 from Estevez *et al.* [85]. Importantly, in that figure $\sigma = 0$ should be seen as the critical line organizing a phase transition in parameter space. Here, we will show that the $\sigma = 0$ line actually corresponds to nondiagonalizability of the TM and, thus, to the qualitatively different polynomial behavior in the decay of the CFs predicted by our Eq. (5.20).

Note that since \mathcal{T} is doubly-stochastic (*i.e.*, all rows sum to one *and* all columns sum to one), the all-ones vector is not only the right eigenvector associated with the eigenvalue of unity, but also the left eigenvector associated with unity. Moreover, since the stationary distribution $\langle \boldsymbol{\pi} |$ is the left eigenvector associated with unity (recall that $\langle \boldsymbol{\pi} | \mathcal{T} = \langle \boldsymbol{\pi} |$), the stationary distribution is the uniform distribution: $\langle \boldsymbol{\pi} | = \frac{1}{6} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$, *i.e.*, $\langle \boldsymbol{\pi} | = \frac{1}{6} \langle \mathbf{1} |$, for $\alpha, \beta \in (0, 1)$. Hence, throughout this range, the projection operator associated with unity is $\mathcal{T}_1 = \frac{1}{6} |\mathbf{1}\rangle \langle \mathbf{1}|$.

It is interesting to note that the eigenvalue of $1 - 2\beta$ is associated with the decay of out-of-equilibrium probability density between the Hägg states of \mathcal{U} and \mathcal{V} —or at least between the *ABC*-state clusters into which each of the Hägg states have split. Indeed, from the Hägg machine: $\Lambda_{\mathcal{T}} = \{1, 1 - 2\beta\}$. So, questions about the relative occupations of the Hägg states themselves are questions invoking the $1 - 2\beta$ projection operator. However,

due to the antisymmetry of output orientations emitted from each of these Hagg states, the $1 - 2\beta$ eigenvalue will not make any direct contribution towards answering questions about the process's output statistics. Specifically, $\langle \mathcal{T}_{1-2\beta}^{\xi(\mathcal{A})} \rangle = 0$ for all $\xi \in \{c, a, s\}$. Since $a_{1-2\beta} = 1$, the projection operator is simply the matrix product of the right and left eigenvectors associated with $1 - 2\beta$. With proper normalization, we have:

$$\mathcal{T}_{1-2\beta} = \frac{1}{6} |\mathbf{1} - 2\beta\rangle \langle \mathbf{1} - 2\beta|$$

with $|\mathbf{1} - 2\beta\rangle = \begin{bmatrix} 1 & 1 & 1 & -1 & -1 & -1 \end{bmatrix}^\top$ and $\langle \mathbf{1} - 2\beta| = \begin{bmatrix} 1 & 1 & 1 & -1 & -1 & -1 \end{bmatrix}$ where \top denotes matrix transposition. Then, one can easily check via Eq. (6.26) that indeed $\langle \mathcal{T}_{1-2\beta}^{\xi(\mathcal{A})} \rangle = 0$ for all $\xi \in \{c, a, s\}$.

To obtain an explicit expression for the CFs, we must obtain the remaining projection operators. We can always use Eq. (6.20). However, to draw attention to useful techniques, we will break the remaining analysis into two parts: one for $\sigma = 0$ and the other for $\sigma \neq 0$. In particular, for the case of $\sigma = 0$, we show that nondiagonalizability need not make the problem harder than the diagonalizable case.

5.6.2.1 $\sigma = 0$:

As mentioned earlier, the $\sigma = 0$ line is the critical line that organizes a phase transition in the ML ordering. We also find that \mathcal{T} is nondiagonalizable only along the $\sigma = 0$ submanifold. For $\sigma = 0$, the $\frac{1}{2}(1 - \beta) \pm \frac{1}{2}\sqrt{\sigma}$ eigenvalues of Eq. (6.37) collapse to a single eigenvalue so that the set of eigenvalues reduces to: $\Lambda_{\mathcal{T}|\sigma=0} = \{1, 1 - 2\beta, -\frac{1}{2}(1 - \beta)\}$ with corresponding indices: $\nu_1 = 1$, $\nu_{1-2\beta} = 1$, and $\nu_{-\beta/2} = 2$.

In this case, the projection operators are simple to obtain. As in the general case, we

have:

$$\begin{aligned}\mathcal{T}_1 &= \frac{1}{6} |\mathbf{1}\rangle \langle \mathbf{1}| \\ &= \frac{1}{6} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}\end{aligned}$$

and

$$\begin{aligned}\mathcal{T}_{1-2\beta} &= \frac{1}{6} |\mathbf{1} - 2\beta\rangle \langle \mathbf{1} - 2\beta| \\ &= \frac{1}{6} \begin{bmatrix} 1 & 1 & 1 & -1 & -1 & -1 \\ 1 & 1 & 1 & -1 & -1 & -1 \\ 1 & 1 & 1 & -1 & -1 & -1 \\ -1 & -1 & -1 & 1 & 1 & 1 \\ -1 & -1 & -1 & 1 & 1 & 1 \\ -1 & -1 & -1 & 1 & 1 & 1 \end{bmatrix}.\end{aligned}$$

Recall that the projection operators sum to the identity: $\mathbb{I} = \sum_{\lambda \in \Lambda_{\mathcal{T}}} \mathcal{T}_{\lambda} = \mathcal{T}_1 + \mathcal{T}_{1-2\beta} + \mathcal{T}_{-\bar{\beta}/2}$.

And so, it is easy to obtain the remaining projection operator:

$$\begin{aligned}\mathcal{T}_{-\bar{\beta}/2} &= \mathbb{I} - \mathcal{T}_1 - \mathcal{T}_{1-2\beta} \\ &= \frac{1}{3} \begin{bmatrix} 2 & -1 & -1 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 & 0 \\ -1 & -1 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & -1 & -1 \\ 0 & 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & -1 & -1 & 2 \end{bmatrix}.\end{aligned}$$

Note that $3 \langle \pi | \mathcal{T}^{[A]} = \frac{1}{2} \langle \mathbf{1} | \mathcal{T}^{[A]} = \frac{1}{2} \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}$ and that:

$$\mathcal{T}^{[A]} |\mathbf{1}\rangle = \begin{bmatrix} 0 \\ \alpha\bar{\beta} + \bar{\alpha}\beta \\ \alpha\beta + \bar{\alpha}\bar{\beta} \\ 0 \\ \alpha\beta + \bar{\alpha}\bar{\beta} \\ \alpha\bar{\beta} + \bar{\alpha}\beta \end{bmatrix}, \mathcal{T}^{[B]} |\mathbf{1}\rangle = \begin{bmatrix} \alpha\beta + \bar{\alpha}\bar{\beta} \\ 0 \\ \alpha\bar{\beta} + \bar{\alpha}\beta \\ \alpha\bar{\beta} + \bar{\alpha}\beta \\ 0 \\ \alpha\beta + \bar{\alpha}\bar{\beta} \end{bmatrix},$$

$$\text{and } \mathcal{T}^{[C]} |\mathbf{1}\rangle = \begin{bmatrix} \alpha\bar{\beta} + \bar{\alpha}\beta \\ \alpha\beta + \bar{\alpha}\bar{\beta} \\ 0 \\ \alpha\beta + \bar{\alpha}\bar{\beta} \\ \alpha\bar{\beta} + \bar{\alpha}\beta \\ 0 \end{bmatrix}.$$

Then, according to Eq. (5.20), with $\langle \mathcal{T}_1^{\xi(\mathcal{A})} \rangle = \frac{1}{3}$, $\langle \mathcal{T}_{1-2\beta}^{\xi(\mathcal{A})} \rangle = 0$, $\langle \mathcal{T}_{-\bar{\beta}/2}^{s(\mathcal{A})} \rangle = -\frac{1}{3}$, $\langle \mathcal{T}_{-\bar{\beta}/2}^{c(\mathcal{A})} \rangle = \langle \mathcal{T}_{-\bar{\beta}/2}^{a(\mathcal{A})} \rangle = \frac{1}{6}$, $\langle \mathcal{T}_{-\bar{\beta}/2,1}^{s(\mathcal{A})} \rangle = \frac{1}{6}(\sigma + \beta - \beta^2) = \frac{1}{6}\beta\bar{\beta}$, and $\langle \mathcal{T}_{-\bar{\beta}/2,1}^{c(\mathcal{A})} \rangle = \langle \mathcal{T}_{-\bar{\beta}/2,1}^{a(\mathcal{A})} \rangle = -\frac{1}{12}(\sigma + \beta - \beta^2) = -\frac{1}{12}\beta\bar{\beta}$, the CFs are:

$$\begin{aligned} Q_\xi(n) &= \sum_{\lambda \in \Lambda_{\mathcal{T}}} \sum_{m=0}^{\nu_\lambda-1} \langle \mathcal{T}_{\lambda,m}^{\xi(\mathcal{A})} \rangle \binom{n-1}{m} \lambda^{n-m-1} \\ &= \langle \mathcal{T}_1^{\xi(\mathcal{A})} \rangle + \sum_{m=0}^1 \langle \mathcal{T}_{-\bar{\beta}/2,m}^{\xi(\mathcal{A})} \rangle \binom{n-1}{m} (-\bar{\beta}/2)^{n-m-1} \\ &= \frac{1}{3} + \left[\langle \mathcal{T}_{-\bar{\beta}/2}^{\xi(\mathcal{A})} \rangle - \frac{2}{\bar{\beta}} \langle \mathcal{T}_{-\bar{\beta}/2,1}^{\xi(\mathcal{A})} \rangle (n-1) \right] (-\bar{\beta}/2)^{n-1}. \end{aligned}$$

Specifically:

$$Q_s(n) = \frac{1}{3} \left[1 + 2 \left(1 + \frac{\beta}{\bar{\beta}} n \right) (-\bar{\beta}/2)^n \right], \quad (5.39)$$

and

$$Q_c(n) = Q_a(n) = \frac{1}{3} \left[1 - \left(1 + \frac{\beta}{\bar{\beta}} n \right) (-\bar{\beta}/2)^n \right]. \quad (5.40)$$

5.6.2.2 $\sigma \neq 0$:

For any value of σ , we can obtain the projection operators via Eq. (6.20). In addition to those quoted above and, in terms of the former $\mathcal{T}_{-\bar{\beta}/2}$, the remaining projection operators turn out to be:

$$\mathcal{T}_{\frac{-\bar{\beta} \pm \sqrt{\sigma}}{2}} = \pm \frac{1}{\sqrt{\sigma}} \mathcal{T}_{-\bar{\beta}/2} \left[\mathcal{T} + \left(\frac{\bar{\beta} \pm \sqrt{\sigma}}{2} \right) \mathbb{I} \right] .$$

Since the $1 - 2\beta$ eigen-contribution is null and since:

$$\begin{aligned} \langle T_1^{\xi(\mathcal{A})} \rangle &= 1/3 , \\ \langle T_{\frac{-\bar{\beta} \pm \sqrt{\sigma}}{2}}^{s(\mathcal{A})} \rangle &= \frac{1}{6} \left[-1 \pm \left(\sqrt{\sigma} + \frac{\beta \bar{\beta}}{\sqrt{\sigma}} \right) \right] \\ &= \pm \frac{1}{6} \left(1 \mp \frac{\beta}{\sqrt{\sigma}} \right) (\sqrt{\sigma} \mp \bar{\beta}) , \text{ and} \\ \langle T_{\frac{-\bar{\beta} \pm \sqrt{\sigma}}{2}}^{c(\mathcal{A})} \rangle &= \langle T_{\frac{-\bar{\beta} \pm \sqrt{\sigma}}{2}}^{a(\mathcal{A})} \rangle \\ &= \frac{1}{12} \left[1 \mp \left(\sqrt{\sigma} + \frac{\beta \bar{\beta}}{\sqrt{\sigma}} \right) \right] \\ &= \mp \frac{1}{12} \left(1 \mp \frac{\beta}{\sqrt{\sigma}} \right) (\sqrt{\sigma} \mp \bar{\beta}) , \end{aligned}$$

the CFs for $\sigma \neq 0$ are:

$$\begin{aligned} Q_{\xi}(n) &= \sum_{\lambda \in \Lambda_{\mathcal{T}}} \lambda^{n-1} \sum_{x_0 \in \mathcal{A}_{\mathcal{P}}} \langle \pi | \mathcal{T}^{[x_0]} \mathcal{T}_{\lambda} \mathcal{T}^{[\xi(x_0)]} | \mathbf{1} \rangle \\ &= \frac{1}{3} + \sum_{\lambda \in \{\frac{-\bar{\beta} \pm \sqrt{\sigma}}{2}\}} \langle \mathcal{T}_{\lambda}^{\xi(\mathcal{A})} \rangle \lambda^{n-1} . \end{aligned} \quad (5.41)$$

Specifically, for $\xi = s$:

$$\begin{aligned} Q_s(n) &= \frac{1}{3} + \frac{1}{6} \left(1 - \frac{\beta}{\sqrt{\sigma}} \right) (\sqrt{\sigma} - \bar{\beta}) \left(\frac{-\bar{\beta} + \sqrt{\sigma}}{2} \right)^{n-1} \\ &\quad - \frac{1}{6} \left(1 + \frac{\beta}{\sqrt{\sigma}} \right) (\sqrt{\sigma} + \bar{\beta}) \left(\frac{-\bar{\beta} - \sqrt{\sigma}}{2} \right)^{n-1} \\ &= \frac{1}{3} \left[1 + \left(1 - \frac{\beta}{\sqrt{\sigma}} \right) \left(\frac{-\bar{\beta} + \sqrt{\sigma}}{2} \right)^n + \left(1 + \frac{\beta}{\sqrt{\sigma}} \right) \left(\frac{-\bar{\beta} - \sqrt{\sigma}}{2} \right)^n \right] , \end{aligned} \quad (5.42)$$

and we recover Eq. (29) of Estevez *et al.* [85].

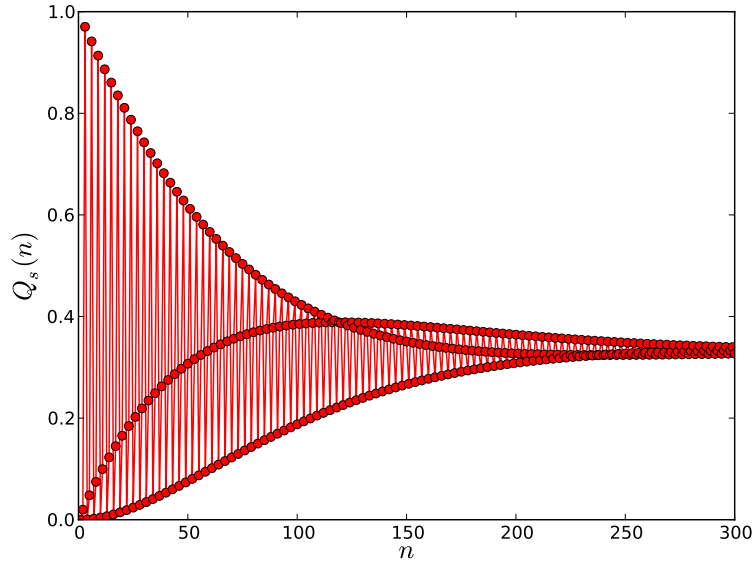


Figure 5.9: $Q_s(n)$ vs. n with $\alpha = 0.01$ and $\beta = 0$ for the RGDF Process. This should be compared to Panel (b) of Fig. 8 in Estevez *et al.* [85]. Although different means were used to make the calculations, they appear to be identical.

Estevez *et al.* [85] recount the embarrassingly long list of recent failures of previous attempts to analyze organization in RGDF-like processes. These failures resulted from not obtaining all of the terms in the CFs, which in turn stem primarily from not using a sufficiently clever ansatz in their methods, together with not knowing how many terms there should be. In contrast, even when casually observing the number of HMM states, our method gives immediate knowledge of the number of terms. Our method is generally applicable with straightforward steps to actually calculate all the terms once and for all.

Figures 5.9, 5.10, 5.11 and 5.12 show plots of $Q_s(n)$ versus n for the RGDF Process at different values of α and β . The first two graphs, Figs. 5.9 and 5.10, were previously produced by Estevez *et al.* [85] and appear to be identical to our results. The second pair of graphs for the RGDF Process, Figs. 5.11 and 5.12 show the behavior of the CFs for larger values of α and β , but with the numerical values of each exchanged ($0.1 \Leftrightarrow 0.2$). The CFs are clearly sensitive to the kind of faulting present, as one would expect. However, each does decay to $1/3$, as they must.

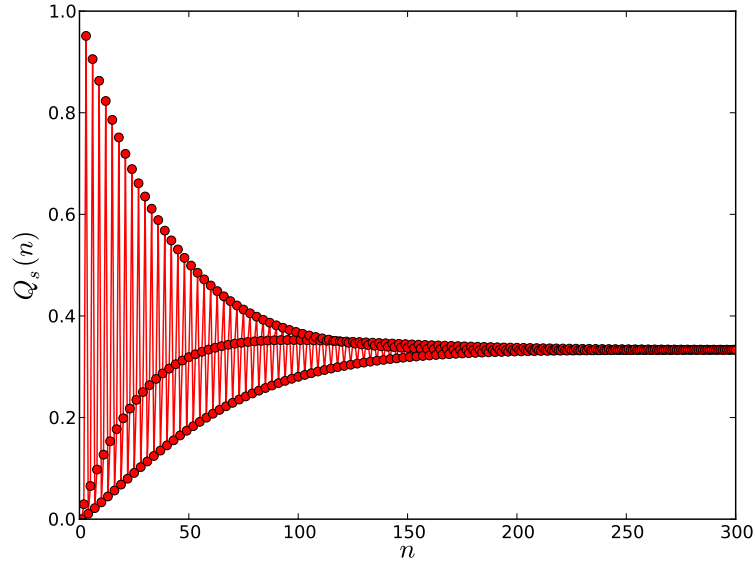


Figure 5.10: $Q_s(n)$ vs. n with $\alpha = 0.01$ and $\beta = 0.01$ for the RGDF Process. Comparison with Panel (d) of Fig. 8 in Estevez *et al.* [85] shows an identical result.

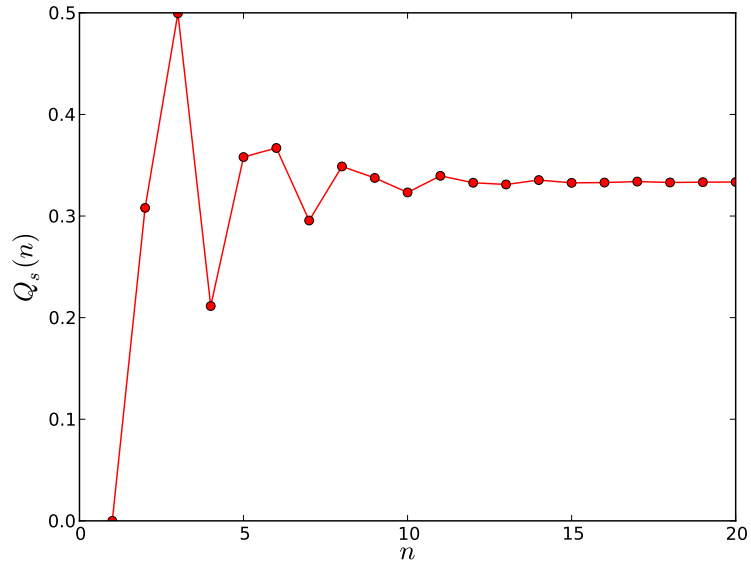


Figure 5.11: $Q_s(n)$ vs. n with $\alpha = 0.1$ and $\beta = 0.2$ for the RGDF Process.

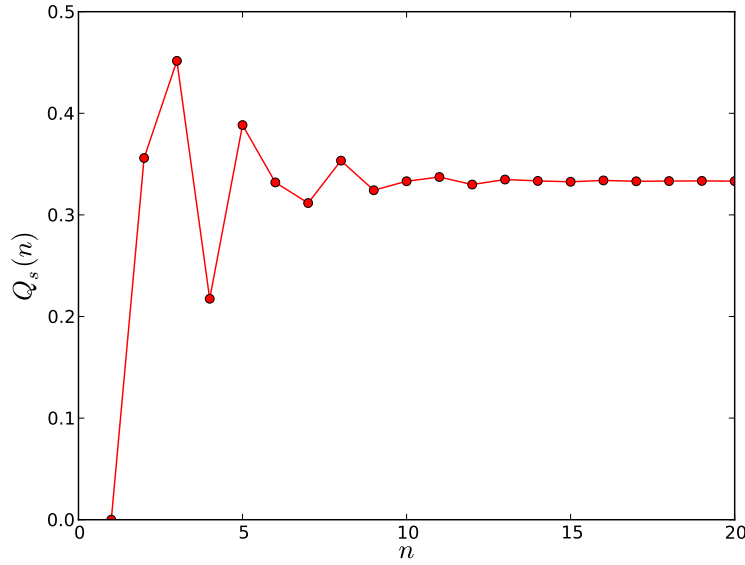


Figure 5.12: $Q_s(n)$ vs. n with $\alpha = 0.2$ and $\beta = 0.1$ for the RGDF Process.

5.6.3 Shockley–Frank Stacking Faults in 6H-SiC: The SFSF Process

While promising as a material for next generation electronic components, fabricating SiC crystals of a specified polytype remains challenging. Recently Sun *et al.* [246] reported experiments on 6H-SiC epilayers ($\sim 200 \mu\text{m}$ thick) produced by the fast sublimation growth process at 1775°C . Using high resolution transmission electron microscopy, they were able to survey the kind and amount of particular stacking faults present. In the Hägg notation 6H-SiC is specified by 000111, and this is written in the Zhdanov notation as (3,3) [183]. Thus, unfaulted 6H-SiC can be thought of as alternating blocks of size-three domains. Ab initio super-cell calculations by Iwata *et al.* [116] predicted that the Shockley defects (4,2), (5,1), (9,3), and (10,2) should be present, with the (4,2) defect having the lowest energy and, thus, it presumably should be the most common. Of these, however, Sun *et al.* [246] observed only the (9,3) defect (given there as (3,9)) and, at that, only once. Instead, the most commonly observed defects were (3,4), (3,5), (3,6), and (3,7), appearing nine, two, two, and three times respectively, with isolated instances of other stacking fault sequences. They postulated that combined Shockley–Frank defects [110] could produce these results.

The (3,4) stacking sequences could be explained as external Frank stacking faults, and the other observed faults could result from further Shockley defects merging with these (3,4) SFs. We call this process the *Shockley-Frank Stacking Fault* (SFSF) Process.

Inspired by these observations, we ask what causal-state structure could produce such stacking sequences. We suggest that the ϵ -machine shown in Fig. 6.5 is a potential candidate, with $\gamma \in [0, 1]$ as the sole faulting parameter. (Here, we must insist that only a thorough analysis, with significantly more data, such as that obtainable from high resolution transmission electron microscopy or a diffraction pattern, can properly reveal the appropriate causal-state structure. The SFSF Process is given primarily to illustrate our methods.) For weakly faulted crystals ($\gamma \approx 0$), as seems to be the case here, there must be a causal state cycle that gives the 6H structure, and we see that the causal-state sequence $[S_7 S_6 S_4 S_0 S_1 S_3]$ does that. Indeed, if the fault parameter γ were identically zero, then this ϵ -machine would give only the 6H structure. Sun *et al.* [246]’s observations suggest that deviations from 6H structure occur (almost) always as *additions* to the size-three 0 or 1 domains. The self-state transitions on S_7 and S_0 have just this effect: After seeing three consecutive 1s (0s), with probability γ the current domain will increase in size to four. And likewise, with probability γ , size-four domains will increase to size-five domains. Thus, with decreasing probability, the faults (3,4), (3,5) ... can be modeled by this ϵ -machine. Notice that the causal architecture prevents domains of any size less than three, which is consistent with the bulk of the observations by Sun *et al.* [246].¹⁵ Also, this ϵ -machine does predict (4,4) sequences, which Sun *et al.* [246] observed once. Thus, qualitatively, and approximately quantitatively, the proposed ϵ -machine largely reproduces the observations of Sun *et al.* [246].

We begin by identifying the SSCs on the HMM, the ϵ -machine shown in Fig. 6.5. We find that there are three, $[S_7]$, $[S_0]$ and $[S_7 S_6 S_4 S_0 S_1 S_3]$. We calculate the winding numbers to be $W^{[S_7]} = 1$, $W^{[S_0]} = 2$, and $W^{[S_7 S_6 S_4 S_0 S_1 S_3]} = 0$. The first two of these SSCs vanish if

¹⁵They did observe a single (3,2) sequence (see their Table I), and the SFSF Process cannot reproduce that structure. Additional causal states and/or transitions would be needed to accommodate this additional stacking structure. One obvious and simple modification that would produce domains of size-two would be to allow the transitions $S_3 \xrightarrow{0} S_6$ and $S_4 \xrightarrow{1} S_1$ with some small probability. However, in the interest of maintaining a reasonably clear example, we neglect this possibility.

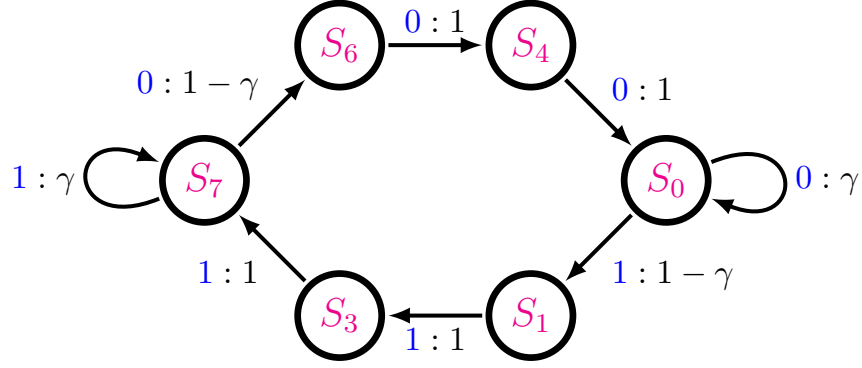


Figure 5.13: Hägg-machine for the SFSF Process, inspired by the observations of Sun *et al.* [246]. We observe that there is one faulting parameter $\gamma \in [0, 1]$ and three SSCs. Or, equivalently three causal state cycles, as this graph is also an ϵ -machine. The three SSCs are $[S_7]$, $[S_0]$ and $[S_7S_6S_4S_0S_1S_3]$. The latter we recognize as the 6H structure if $\gamma = 0$. For large values of γ , *i.e.*, as $\gamma \rightarrow 1$, this process approaches a twinned 3C structure, although the faulting is *not* random. The causal state architecture prevents the occurrence of domains of size-three or less.

$\gamma = 0$, giving a nonmixing Hägg-machine. Thus, for $\gamma \neq 0$ the Hägg-machine is mixing and we proceed with the case of $\gamma \in (0, 1]$.

By inspection we write down the two 6-by-6 TMs of the Hägg-machine as:

$$\mathbf{T}^{[0]} = \begin{bmatrix} \gamma & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \bar{\gamma} & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

and:

$$\mathbf{T}^{[1]} = \begin{bmatrix} 0 & \bar{\gamma} & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & \gamma & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

where the states are ordered S_0, S_1, S_3, S_7, S_6 , and S_4 . The internal state TM is their sum:

$$\mathbf{T} = \begin{bmatrix} \gamma & \bar{\gamma} & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & \gamma & \bar{\gamma} & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Since the six-state Hägg-machine generates an $(3 \times 6 =)$ eighteen-state *ABC*-machine, we do not explicitly write out the TMs of the *ABC*-machine. Nevertheless, it is straightforward to expand the Hägg-machine to the *ABC*-machine via the rote expansion method of §5.4.3. It is also straightforward to apply Eq. (5.15) to obtain the CFs as a function of the faulting parameter γ . To use Eq. (5.15), note that the stationary distribution over the *ABC*-machine can be obtained via Eq. (5.9) with:

$$\langle \boldsymbol{\pi}_H | = \frac{1}{6-4\gamma} \begin{bmatrix} 1 & \bar{\gamma} & \bar{\gamma} & 1 & \bar{\gamma} & \bar{\gamma} \end{bmatrix}$$

as the stationary distribution over the Hägg-machine.

The eigenvalues of the Hägg-TM can be obtained as the solutions of $\det(\mathbf{T} - \lambda \mathbb{I}) = (\lambda - \gamma)^2 \lambda^4 - \bar{\gamma}^2 = 0$. These include $1, -\frac{1}{2}\bar{\gamma} \pm \sqrt{\gamma^2 + 2\gamma - 3}$, and three other eigenvalues involving cube roots. Their values are plotted in the complex plane Fig. 5.14 as we sweep through γ .

The eigenvalues of the *ABC*-TM are similarly obtained as the solutions of $\det(\mathcal{T} - \lambda \mathbb{I}) = 0$. The real and imaginary parts of these eigenvalues are plotted in Fig. 5.15. Note that $\Lambda_{\mathcal{T}}$ inherits $\Lambda_{\mathbf{T}}$ as the backbone for its more complex structure, just as $\Lambda_{\mathbf{T}} \subseteq \Lambda_{\mathcal{T}}$ for all of our previous examples. The eigenvalues in $\Lambda_{\mathcal{T}}$ are, of course, those most directly responsible for the structure of the CFs.

The SFSF Process's CFs are shown for several example parameter values of γ in Figs. 5.16, 5.17, 5.18, and 5.19 calculated directly from numerical implementation of Eq. (5.15). As the faulting parameter is increased from $0.01 \rightarrow 0.5$, the CFs begin to decay more quickly. However, for $\gamma = 0.9$, the correlation length increases as the eigenvalues,

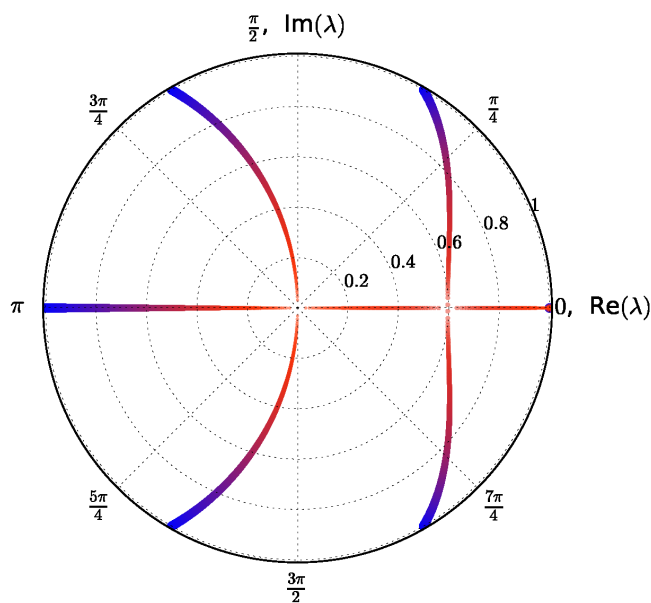


Figure 5.14: The six eigenvalues of the Hägg-machine as they evolve from $\gamma = 0$ (thickest blue markings) to $\gamma = 1$ (thinnest red markings). Note that the eigenvalues at $\gamma = 0$ are the six roots of unity. Unity is a persistent eigenvalue. Four of the eigenvalues approach 0 as $\gamma \rightarrow 1$. Another of the eigenvalues approaches unity as $\gamma \rightarrow 1$. The eigenvalues are nondegenerate throughout the parameter range except for the transformation event where the two eigenvalues on the right collide and scatter upon losing their imaginary parts.

near the nontrivial cube-roots of unity, loop back toward the unit circle. The behavior near $\gamma = 0.9$ suggests a longer ranged and more regularly structured specimen, even though there are fewer significant eigen-contributions to the specimen's structure. Indeed, the bulk of the structure is now more apparent but less sophisticated.

5.7 Conclusion

We introduced a new approach to exactly determining CFs directly from HMMs that describe layered CPSs. The calculation can be done either with high numerical accuracy and efficiency, as we have shown in the CF plots for each example, or analytically, as was done for the IID and RGDF Processes.

The mathematical object that assumes central importance here is the HMM. While we appreciate the value that studying CFs and, more generally, pair distribution functions brings to understanding material structure, pairwise correlation information is better

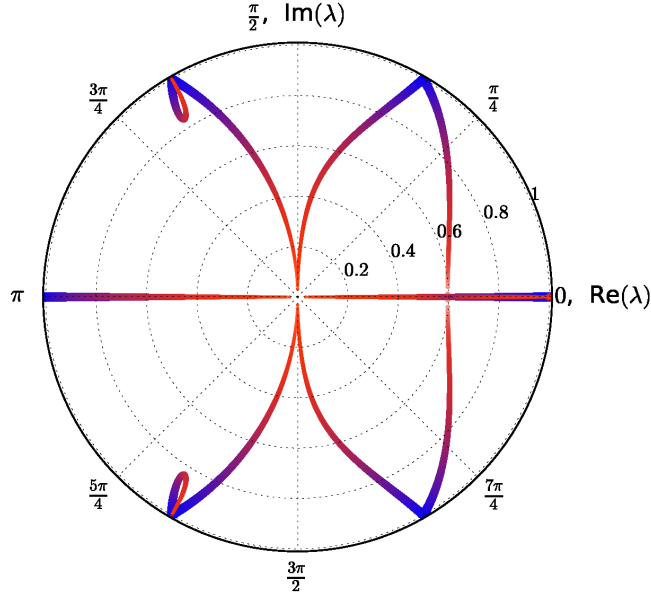


Figure 5.15: The eighteen eigenvalues of the *ABC*-machine as they evolve from $\gamma = 0$ (thickest blue markings) to $\gamma = 1$ (thinnest red markings). Note that the eigenvalues at $\gamma = 0$ are still the six roots of unity. The new eigenvalues introduced via transformation to the *ABC*-machine all appear in degenerate (but diagonalizable) pairs. In terms of increasing γ , these include eigenvalues approaching zero from ± 1 , eigenvalues taking a left branch towards zero as they lose their imaginary parts, and eigenvalues looping away and back towards the nontrivial cube-roots of unity.

thought of as a consequence of a more fundamental entity (*i.e.*, the HMM) than one of intrinsic importance. This becomes clear when we consider that the structure is completely contained in the very compact HMM representation. More to the point, all of the correlation information is directly calculable from it, as we demonstrated. In contrast, the task of inverting correlation information to specify the underlying organization of a material's structure, *i.e.*, its HMM, is highly nontrivial. Over the past century considerable effort has been expended to invert diffraction patterns, the Fourier transform of the CFs, into these compact structural models.¹⁶ The work of Warren [273], Krishna and coworkers [214, 216, 217, 215, 219], Berliner & Werner [21] and that of our own group [264, 265, 266, 270], to mention a few, all stand in testament to this effort.

¹⁶We have not explicitly made the connection here, but almost all previous models of planar disorder can generically be expressed as HMMs.

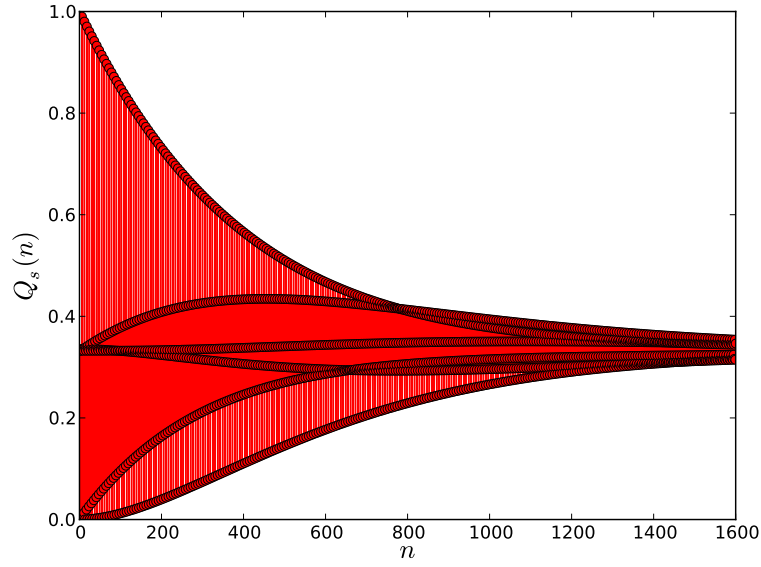


Figure 5.16: $Q_s(n)$ vs. n for the SFSF Process with $\gamma = 0.01$. This specimen is only very weakly faulted and, hence, there are small decay constants giving a slow decay to $1/3$.

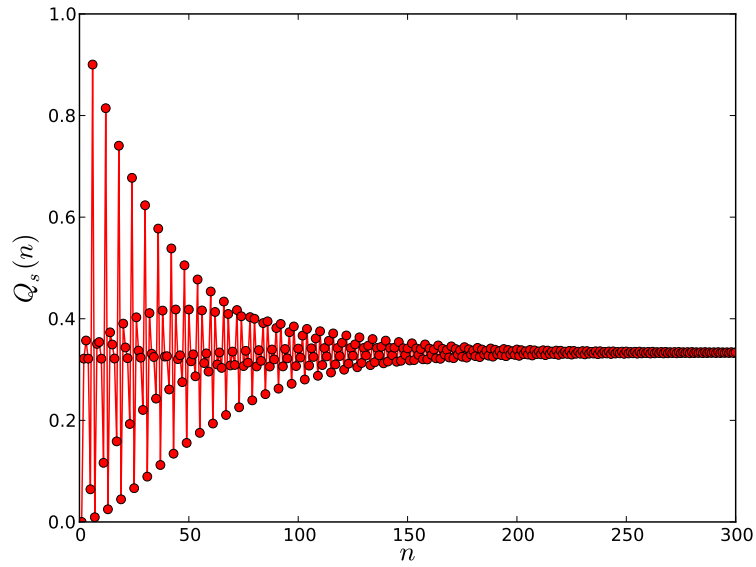


Figure 5.17: $Q_s(n)$ vs. n for the SFSF Process with $\gamma = 0.1$. With increasing γ , the CFs approach their asymptotic value of $1/3$ much more quickly.

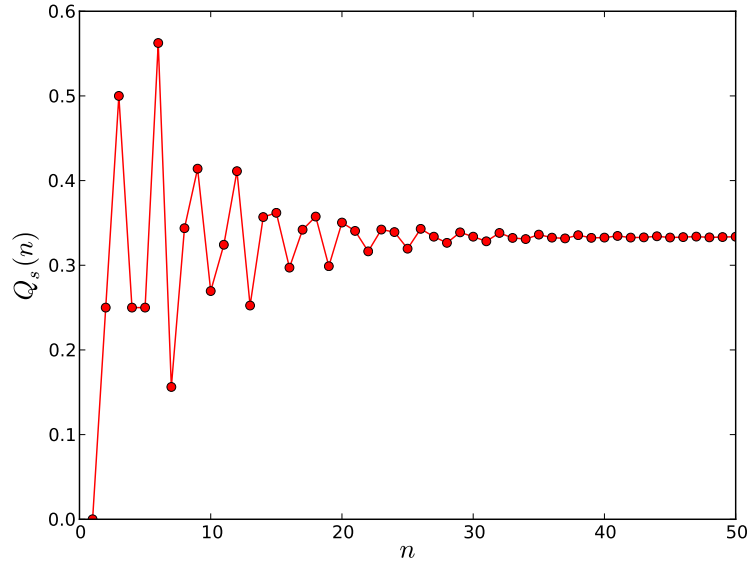


Figure 5.18: $Q_s(n)$ vs. n for the SFSF Process with $\gamma = 0.5$. Here, the specimen is quite disordered, and the CFs decay quickly.

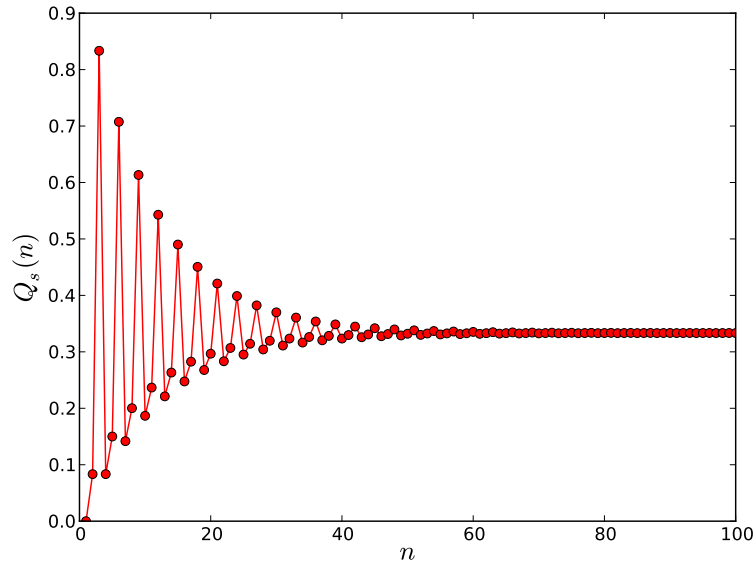


Figure 5.19: $Q_s(n)$ vs. n for the SFSF Process with $\gamma = 0.9$. The slower CF decay suggests that the process is now less disordered than the $\gamma = 0.5$ case. Notice that this CF is large for $n \bmod (3) = 0$, indicating strong correlation between MLs separated by a multiple of three MLs. This is the kind of behavior that one expects from a twinned 3C crystal.

Although the presentation concentrated on CFs in layered CPSs, the potential impact of the new approach is far wider. First, we note that it was necessary to make some assumptions about the geometry of the stacking process, *i.e.*, the number of possible orientations of each ML and how two MLs can be stacked, in order to demonstrate numerical results and make contact with previous work. These assumptions however in no way limit the applicability: Any set of stacking rules over a finite number of possible positions is amenable to this treatment. Second, it may seem that starting with a HMM is unnecessarily restrictive. It is not. Given a sample of the stacking process from (say) a simulation study, there are techniques that now have become standard for finding the ϵ -machine, a kind of HMM, that describes the process. The subtree-merging method [64] and causal-state splitting reconstruction [225] are perhaps the best known, but recently a new procedure based on Bayesian inference has been developed [243]. Finally, a HMM may be proposed on theoretical grounds, as done with the RGDF and SFSF HMMs in our second and third examples. And, for the case when a diffraction pattern is available, there is ϵ -machine spectral reconstruction theory [264, 265, 266, 270]. We anticipate that HMMs will become the standard representation for describing layered structures.

The approach presented here should also be viewed in the larger context of our recent research thrusts. While crystallography has historically struggled to settle on a formalism to describe disordered structures, we propose that such a framework has been identified—at least for layered materials. Based in computational mechanics, *chaotic crystallography* [268] employs information theory as a key component to characterize disordered materials. Although the use of information theory in crystallography has been previously proposed by Mackay and coworkers [154, 155, 36], chaotic crystallography realizes this goal. Additionally, using spectral methods in the spirit of §5.5.2, information- and computation-theoretic measures are now directly calculable from ϵ -machines [60, 198]. And importantly, a sequel will demonstrate how spectral methods can give both a fast and efficient method for calculating the diffraction pattern of layered CPSs or analytical expressions thereof [204].

5.8 Appendix A: Transition matrices, projection operators, and bra–ket notation

The sub-stochastic symbol-labeled transition matrices of a HMM sum to give our probability-conserving right-stochastic (*i.e.*, all rows sum to one) state-to-state transition matrix $\mathcal{T} = \sum_{x \in \mathcal{A}} \mathcal{T}^{[x]}$. All eigenvalues of this matrix lie on or within the unit circle in the complex plane. We use a bra–ket notation to easily identify bras $\langle \cdot |$ as row vectors and kets $|\cdot\rangle$ as column vectors. A special set of kets are right eigenvectors of \mathcal{T} . The right eigenvectors satisfy:

$$\mathcal{T} |\lambda\rangle = \lambda |\lambda\rangle$$

for $\lambda \in \Lambda_{\mathcal{T}}$. A corresponding set of bras are left eigenvectors of \mathcal{T} . The left eigenvectors satisfy:

$$\langle \lambda | \mathcal{T} = \lambda \langle \lambda |$$

for $\lambda \in \Lambda_{\mathcal{T}}$. In general—unlike the special case in quantum mechanics—the left eigenvectors are not merely the conjugate transpose of the right eigenvectors. Projection operators play a prominent role in the development of our results. For the general case, they can be obtained from Eq. (6.20). However, in the simplest case of a projection operator of a non-degenerate eigenvalue, the projection operator can simply be expressed as:

$$\mathcal{T}_{\lambda} = \frac{1}{\langle \lambda | \lambda \rangle} |\lambda\rangle \langle \lambda| ,$$

where the denominator is simply a normalizing factor. The only normalization convention we impose is that $|\mathbf{1}\rangle$ represents the right eigenvector associated with unity of all ones and $\langle \boldsymbol{\pi} |$ is the left eigenvector associated with unity normalized in probability. This normalization allows $\mathcal{T}_1 = |\mathbf{1}\rangle \langle \boldsymbol{\pi} |$ and consistent interpretation of unconditioned probabilities: $\Pr(w) = \langle \boldsymbol{\pi} | \mathcal{T}^{[w]} | \mathbf{1} \rangle$. Further useful results and discussion of the mathematical machinery can be found in [198].

Chapter 6

Diffraction Patterns of Layered Close-packed Structures from Hidden Markov Models

An earlier version of this chapter has appeared online as a preprint:

P. M. Riechers, D. P. Varn, & J. P. Crutchfield, (2014). “Diffraction Patterns of Layered Close-packed Structures from Hidden Markov Models”. arXiv preprint arXiv:1410.5028.

The results carry over with some modification to address diffraction patterns of layered structures in general. The novel insight conveyed in this chapter—that diffraction spectra are a direct signature of eigen-spectra—should be expected quite generally, even beyond layered structures.

6.1 Chapter Overview

We develop a method to calculate the diffraction pattern for layered close-packed structures stacked according to the wide class of processes expressible as hidden Markov models. We show that in the limit of large crystals, the diffraction pattern is a particularly simple function of parameters that specify the hidden Markov model. We give three elementary but important examples that demonstrate this result, deriving expressions for the diffraction pattern of close-packed structures stacked: (i) independently, (ii) as infinite-Markov-order randomly faulted 2H and 3C specimens for the entire range of growth and deformation faulting probabilities, and (iii) as a hidden Markov model that describes Shockley-Frank

stacking faults in 6H-SiC. We show that the eigenvalues of the transition matrix—as defined by the hidden Markov model—organize and structure the diffraction pattern. To illuminate this connection, we introduce a new visualization technique, the *coronal spectrogram*, that makes explicit the tight relationship between the eigenvalues and the diffraction pattern. In this way, we effectively solve the broad problem of calculating a diffraction pattern—either analytically or numerically—for any layered close-packed stacking structure—ordered or disordered—where the stacking process can be expressed as a hidden Markov model.

6.2 Introduction

Since the fundamental forces that bind materials are isotropic, it may be surprising that so many materials have such strongly anisotropic structures. That is, the bonds may not be of equal strength in all directions, instead showing relatively strong preferential bonding in two of the directions, with much weaker interactions along a third direction. We say that the resulting material is *layer-like* with strong *intralayer* bonding and weaker *interlayer* bonding. Other mechanisms may contribute to a material having a layer-like structure. For example, some form as a result of a layer-by-layer growth process; and, of course, it is now possible to engineer artificially layered materials with great precision [95, 149]. These various mechanisms conspire to create a great variety of materials that are conveniently thought of as being composed of stacked layers. Examples are numerous: micas such as muscovite and phlogopite [133]; III-V compounds such as GaS and InSe [218], which are of interest due their electronic and optical properties; SiC and ZnS [218], which are known to have hundreds of stable crystalline structures resulting from their layer-like properties; perovskites such as BaRuO₃ and CaTiO₃ [91]; graphene and hexagonal BN [95], which have attracted great interest for their electronic properties; and lastly, metal dichalcogenides such as MoS₂ and WSe₂ [127].

Although restricted to a few alternatives, often there is more than one way to stack two adjacent layers, and much of the interesting physics and crystallography occurs as one traverses a material perpendicular to these layers in what is called the *stacking*

direction [268]. One such phenomenon is *polytypism* [257], where a material built up of identical layers can coexist in two or more stable crystalline stacking configurations. SiC in particular is known to have hundreds of such periodic (and hence crystalline) stacking configurations, some with periodicities extending over one hundred layers [218]. Given the weak layer-to-layer interactions in many of these materials as well as the small energy differences between some stacking sequences, they are prone to errors in the stacking configuration called *stacking faults* [110, 115]. These deviations from periodic stacking can occur during the growth process or through some post-formation stress to the specimen, be it mechanical, thermal, or irradiative [218]. The resulting specimen is disordered, sometimes even to the point of losing any meaningful sense of an underlying periodic stacking structure [161, 269]. While these planar defects have, of course, been known and studied for some time they have often been viewed as a nuisance or feature to be minimized or eliminated [35]. Increasingly though, materials scientists are appreciating the sometimes unexpected role that disorder and crystal defects play in material properties. For example, disordered graphene nanosheets can be used to improve the performance of high-capacity Li ion batteries [186], and so-called ‘defect engineering’ in semiconductors is attracting wide attention [220].

Given the growing technological import of disordered materials, methods of detecting experimentally and then describing the resulting structures are needed. Since its discovery in the early 20th century, X-ray diffraction has proved invaluable for determining material structure, for both ordered and disordered materials [275]. Typically in a diffraction experiment, one uses a probe, most often a finely collimated beam of monochromatic X-rays, electrons, or neutrons focused on the specimen at some angle and subsequently one collects the diffracted reflections. When this is performed over a variety of angles, one generates a *diffraction pattern*.¹ While many physical effects can influence the diffraction pattern [97], most can be corrected for and the diffraction pattern is profitably viewed as the power spectrum of the real space arrangement of atoms. Thus, the diffraction pattern

¹Common geometries and techniques for diffraction experiments, particularly X-ray diffraction, can be found in any text on condensed matter physics, such as Ashcroft & Mermin [5], or any crystallographic text such as Giacovazza *et al.*[97], Warren [273], and Guinier [103].

is highly sensitive to material structure and diffraction experiments remain the workhorse of modern material structure determination.

It is also important to efficiently, compactly, and consistently describe such disordered layered materials. *Classical crystallography*, based on the geometric symmetries found in perfectly crystalline materials and couched in the language of finite group theory, has enjoyed enormous success. However, there has long been a call to expand the range of classical crystallography to treat other structures, particularly those that are not perfectly ordered [154, 36]. Indeed, classical crystallography has not proved nearly as useful when characterizing materials that show disorder, especially if that disorder is profound. To circumvent these limitations, a new perspective on the structure of materials has recently been proposed. *Chaotic crystallography* is the application of information- and computation-theoretic methods to the discovery and description of disorder in materials [268, 12]. Drawing from concepts developed in information theory [46], theoretical computer science [190, 112] and nonlinear dynamics [244], chaotic crystallography adapts and applies *computational mechanics* [55, 64] to the problem of disorder in materials. Instead of relying on exact geometric symmetries, chaotic crystallography captures structure in terms of partial and noisy symmetries, formulated in the language of semi-groups. Setting order and disorder on equal footing, chaotic crystallography provides novel platform to understand material structure.

Quasi-one-dimensional materials are particularly amenable to analysis by the methods of chaotic crystallography. One scans the specimen along the stacking direction recording the kind and orientation of the layers as they are encountered, assigning a symbol to each of the possible kinds and orientations. The resulting list of symbols is called the *stacking sequence* [264], and the effective stochastic process induced is called the *stacking process* [264]. Sequential symbolic data such as this have been well-studied in the physical sciences and are sometimes analyzed using *hidden Markov models* [197, 74]. A rich set of analytical tools has been developed to study hidden Markov models, and here we will augment that toolset by demonstrating how the diffraction pattern can be directly calculated from an arbitrary hidden Markov model, either efficiently to great numerical

accuracy or analytically.

We report two major advances. (i) Although the problem of connecting disordered stacking structures to diffraction patterns has been addressed by many previous authors, we offer a new derivation of an expression to calculate the diffracted intensity of close-packed layered materials in terms of the hidden Markov model that describes the stacking process. This is carried out by explicitly treating either ordered or disordered stacking configurations as an arbitrary hidden Markov model, which to our knowledge has not been previously done. (ii) Spectral analysis of the hidden Markov model will lead us to a new way to visualize diffraction patterns for layered materials, in what we call *coronal spectrograms*. We will introduce these versatile and informative graphs in §6.6, where we will find that they result in significant insights into stacking structures.

Our development is organized as follows: §6.3 reviews historical milestones in the study of diffraction patterns from layered materials and familiarizes the reader with modeling stochastic processes; §6.4 fixes nomenclature and definitions; §6.5 derives a general expression for the diffraction pattern of layered close-packed structures in terms of an arbitrary hidden Markov model; §6.6 introduces coronal spectrograms and considers several examples, namely (i) an independently distributed process that can model 3C or random stacking structures, (ii) an infinite-order-Markov stacking process that represents any amount of random growth and deformation faults in 3C and 2H structures, and (iii) a stacking process inspired by recent experiments in 6H-SiC; and §6.7 gives our conclusions and outlines directions for future work.

6.3 Background

6.3.1 Brief History of Planar Disorder and Diffraction Patterns

There is a long history of researchers treating the problem of calculating diffraction patterns from layered materials and the interested reader is encouraged to consult Treacy *et al.* [252] for a thorough and detailed treatment or Estevez-Rams [82] for a shorter but more recent exposition.

Landau [138] and Lifshitz [145] are usually credited as the first to treat the problem,

when in 1937 they considered the effect of stacking disorder, assuming no correlation among the stacking defects. In 1942 Hendricks & Teller [109] developed the *matrix formalism method*, where a transition matrix specifies the probability of observing a layer, possibly dependent on a previous history of observed layers. They considered four cases: (i) no interaction between layers with each layer having the same form factor, (ii) randomly distributed layers having variable spacing, (iii) the probabilities of the occurrence of layers depending on the nearest neighbors, and (iv) the probabilities of occurrence of layers depending on nearest and next nearest neighbors. Also in 1942 Wilson [278] examined the case of hexagonal closed-packed Co with randomly distributed planar faults, pioneering the *difference equation method*. Jagodzinski [118, 119] in 1949 introduced the *reichweite*, denoted s , that allows variable range of influence between layers in close-packed structures. He found solutions for small s , but noted that the calculations become onerous for $s > 3$. In 1986 Berliner & Werner [21] demonstrated that the diffraction pattern for a specific stacking sequence can be calculated directly by *Monte Carlo* or *crystal growth methods*, where one simulates a crystal by ‘growing’ it according to some rules, and then sums the contributions to the reflected wave for all layers in the sample. They applied this technique to growth and deformation faults in Li and compared the results with experimental neutron diffraction studies. Many other researchers have expanded on these methods, leaving an extensive literature [252, 82, 38].

However, most relevant to the development we pursue here is the 1991 seminal work by Treacy *et al.* [252]. They offered a *general recursion matrix method* to calculate the diffraction pattern for layered materials, implementing it in a freely available and well-used computer software package called *DIFFaX*. Key to their approach is the recognition that by a judicious grouping of terms it is possible to perform analytically the sum over layers to calculate the diffraction pattern. As part of their procedure, they assume a transition matrix that gives the probabilities of transitioning to any other layer, given the current layer.

To place our contributions in context of these previous efforts, it is necessary to discuss how these models can be classified according to their computational-theoretic capabilities.

In the next section we offer a tour of the differences between the kinds of models that have been used and in §6.46.4.2 we give a mathematically grounded treatment of the stacking process as generated by a hidden Markov model.

6.3.2 Markov Models and Hidden Markov Models

Let us consider models that are capable of probabilistically generating a sequence of symbols, where the symbols are taken from some finite set. Important distinctions between different models rest on how they condition the probability of observing a current symbol on a previously observed, perhaps even infinite, history of symbols. The simplest assumption is that the current symbol in the sequence takes its particular value independent of all the other symbols in the sequence. If, further, the probability distribution over symbols is invariant to a shift in sequence position, the model is called *independent and identically distributed* (IID). By definition, these models admit no correlation between symbols.

The next level of complication occurs when we allow the value of the current symbol to depend on the previous symbol, and *only* the previous symbol. These models are called Markov models (MMs). Since it is only the value of the previous symbol that is needed to give the best prediction of the current symbol, it is possible to label the states of the MM by the current symbol. Trivially, then, one knows the sequence of states by the observed symbol sequence, since they are identical. Concomitantly, for MMs, the total number of states of the model is strictly limited by the number of the kinds of symbols observed.

An additional complication occurs when it is necessary to know not only the previous symbol, but a finite history of previous symbols. These *finite-order Markov models* (fMMs) are often further specified by the length of the history needed, sometimes denoted r [270], and are called r th-order Markov models. For $r = 1$, fMMs degenerate to the MMs considered earlier, and if we allow the case that $r = 0$, then zeroth-order Markov models generate the IID processes. Note that for $r \geq 2$ it is no longer possible to label the states by the observed symbol. In this sense, the states of the model become *hidden*, as the sequence of states is no longer just the observed sequence of symbols.

Perhaps surprisingly, often the next symbol in a sequence depends on an infinite history of symbols. This is best explained by giving an example. Suppose that a sequence

composed of only 0s and 1s conforms to the following rule: one can observe a consecutive sequence of 1s of any length; however, between any two 1s there must be an *even* number of 0s. Thus, on observing a 0, one also needs to know the symbol preceding it. If that symbol were 1, such that the two-symbol history was 10, then to retain the evenness of 0s the next symbol would perforce be 0. If it was not, then we would have the sequence 101, which is not allowed. Indeed, it becomes apparent that when one observes a consecutive sequence of 0s, it is important to keep track of whether there has been an even or odd number of them. Since this history could extend back indefinitely, this process, called the *even process* [274, 9], entails a special kind of infinite memory. Nonetheless, it is possible to describe this process by using only two states: (a) if the history of symbols ends in 1 or an even number of consecutive 0s, then the next symbol is either 0 or 1; (b) if the history is an odd number of 0s, then the next symbol is 0. Borrowing terminology from the computation theory of formal languages, the even process is an example of a *strictly sofic process* [274, 9] and has been treated in the context of close-packed layered materials elsewhere [270]. Models such as the even process that remember an infinite past with a finite memory (using a finite number of states) are called *infinite-order Markov models* (iMMs).

For IIDs, MMs, and fMMs, if the system is in a particular state and makes a transition to another state, the symbol observed on this transition uniquely determines the successor state. This property is called *unifilarity* [76]. If this condition is relaxed so that the state and the observed symbol are not sufficient to determine the next state, these models are *nonunifilar*. This distinction becomes important for infinite-order Markov models, and those with unifilarity are called *unifilar infinite-order Markov models* (uiMMs) and those without it are called *nonunifilar infinite-order Markov models* (niMMs). From the development, it should be apparent that these models are hierarchically related, such that models lower on the hierarchy are special cases of those above it.² The model-class hierarchy is:

$$\text{IID} \subset \text{MM} \subset \text{fMM} \subset \text{uiMM} \subset \text{niMM} . \quad (6.1)$$

²For a full discussion, see Crutchfield [54] and Crutchfield & Marzen [63].

Let us revisit previous models of disorder and classify them by their place on the hierarchy. Landau [138], Lifshitz [145], and the first two cases of Hendricks & Teller [109] are IID models. Hendricks & Teller’s third case, where the probabilities only depend on the single preceding layer, are MMs. Finally, by grouping layers, Hendricks & Teller were able to effectively construct a 2nd-order MM (*i.e.*, a fMM) in their last case. Similarly, a careful treatment of the model proposed by Wilson [278] shows that it is a 2nd-order MM, as he conditions the fault probability on the values of the previous two layers.³ Jagodzinski’s *reichweite* [118, 119] is clearly related to the Markov order. If one writes the stacking sequence in the *HK*-notation [218, 183], then $s = 3$ can be expressed as a MM and $s = 4$ as a 2nd-order MM. The Monte Carlo method of Berliner & Werner [21] is not a distinct model class in the same sense that the others are, since there is no ‘compactification’ of the rules governing the stacking process. By explicit calculation from a particular stacking sequence, it can effectively calculate the diffraction pattern for any of the models in Eq. (6.1). ϵ -Machine spectral reconstruction theory [264, 265, 266, 270], an alternative sequence-based estimation method, has been implemented up to 3rd-order MMs, although in previous applications it used the Monte Carlo method to calculate diffraction patterns. Lastly, the recursion matrix method of Treacy *et al.* [252] explicitly treats MMs, although by grouping layers as proposed by Hendricks & Teller, it can treat fMMs.

It should be noted that the Monte Carlo method, though now in common usage, is not altogether satisfactory. One difficulty lies in the statistical fluctuations inherent in the finite-size samples it uses. It is known that these lead to estimation errors in the power spectrum on the order of the magnitude of the power spectrum itself [125]. This difficulty can be ameliorated by taking many samples and averaging [179] or by using a smoothing procedure [266]. A second concern arises when repeated comparison of calculated and experimental diffraction patterns is needed for inference procedures such as reverse Monte Carlo modeling [128] and differential evolution and genetic algorithms [170]. For samples of reasonable size, the Monte Carlo method is clearly much less efficient than the matrix

³There is a subtlety here. Wilson defines his model in terms of the *ABC*-notation of the close-packed layers. By rewriting the stacking sequence in a different nomenclature, it is possible to recast his model as a 1st-order MM or even simply a MM. This particular model, as well as the dependence of the Markov order on the notation used, is treated in detail elsewhere [269].

formalism or difference equation methods.

It might be thought that the description of stacked layers would not require such sophisticated models. However, especially when a layered specimen is undergoing solid-state transformation, such as by annealing, effectively infinite-range memories can be induced [267]. As the layers shift, energetic considerations can restrict the number of paths to disorder, leaving a remnant of the crystalline stacking structure imprinted in the disordered specimen. Equilibrium is often not obtained, and the structure finds itself caught in a local minimum of the free energy, lacking the necessary activation energy to explore all of phase space. In addition to their expected importance, neglecting infinite-order Markov models in favor of only finite-order Markov models severely restricts the space of possible models considered. James *et al.*[120] showed that as a function of the number of states, the number of topologically distinct finite-order MMs is dwarfed as compared to the number of distinct infinite-order MMs. For six-state processes, a full 98% of the models were infinite-order Markov. Thus, models capable of capturing infinite-range memory are needed.

To our knowledge, no method has yet been shown to calculate diffraction patterns of infinite-order Markov models (uiMMs or niMMs) either analytically or in closed form. In the following, we offer a general, analytical solution to this problem for stacking processes in close-packed structures describable as a hidden Markov model.

6.4 Definitions and Notations

We will confine our exposition to close-packed structures (CPSs). For many materials, such as SiC, CdI₂, and GeSe [218], there may be several layers of atoms that act as a unit and it is these units which obey the close-packing rules. We will refer to these units as *modular layers* (MLs) [89, 263]. Let us make the following assumptions concerning the stacking of MLs in CPSs: (i) the MLs themselves are undefected and free of any distortions; (ii) the spacing between MLs does not depend on the local stacking arrangement; (iii) each ML has the same scattering power; and (iv) the faults extend laterally completely across the crystal.

Additionally, we assume that the unconditioned probability of finding a given stacking sequence remains constant through the crystal. (In statistics parlance, we assume that the stacking process is *weak-sense stationary*. Physically, the process is spatial-translation invariant.)

6.4.1 Correlation Functions and Stacking Notation

In CPSs, each ML may assume one of three possible orientations, usually labeled A , B , and C [5]. We say that two MLs in a sequence are *cyclically* related if the ML further along in the sequence can be obtained from the earlier ML via a cyclic permutation (*i.e.*, $A \rightarrow B \rightarrow C \rightarrow A$), and *anticyclically* related if it can be obtained via an anticyclic permutation (*i.e.*, $A \rightarrow C \rightarrow B \rightarrow A$). It is convenient to introduce three statistical quantities [266], $Q_c(n)$, $Q_a(n)$, and $Q_s(n)$: the pairwise (auto)correlation functions (CFs) between MLs that are the probability any two MLs at separation n are related cyclically (c), anticyclically (a), or have the same orientation (s), respectively. It is also useful to introduce a family of cyclic-relation functions [205] $\hat{\xi}(x) \in \{\hat{c}(x), \hat{a}(x), \hat{s}(x)\}$, where, for example:

$$\hat{c}(x) = \begin{cases} B & \text{if } x = A \\ C & \text{if } x = B \\ A & \text{if } x = C \end{cases} . \quad (6.2)$$

The other two operators, $\hat{a}(x)$ and $\hat{s}(x)$, are defined in an obviously analogous fashion.

It is sometimes advantageous to exploit the constraint that no two adjacent MLs may occupy the same orientation in CPSs. Thus, we sometimes use the Hägg-notation, where cyclic transitions between adjacent MLs are denoted with ‘+’, and anticyclic ones with ‘-’. (Ortiz *et al.* [183] give an excellent treatment of the various notations used to describe CPSs.) Often it is more convenient to substitute ‘1’ for ‘+’ and ‘0’ for ‘-’ and we make this substitution throughout. The two notations, the Hägg-notation and the ABC -notation, carry an equivalent message (up to an overall rotation of the specimen about the stacking direction), albeit in different tongues.

6.4.2 The Stacking Process as a Hidden Markov Model

Previously, it was shown that the stacking process for many cases of practical interest can be written as a discrete-step, discrete-state hidden Markov model (HMM) [205], and we review notations and conventions now.

We assume that the statistics of the stacking process are known and can be expressed as a HMM in the form of an ordered tuple $\Gamma = (\mathcal{A}, \mathbb{S}, \mu_0, \mathbf{T})$, where \mathcal{A} is a set of symbols output by the process and often called an *alphabet*, \mathbb{S} is a finite set of M internal (and possibly hidden) states, μ_0 is an initial state probability distribution, and \mathbf{T} is a set of $|\mathcal{A}|$ M -by- M transition matrices (TMs) that give the transition probabilities between states on emission of one of the symbols in \mathcal{A} .

For the CPSs, the output symbols are just ML orientations and, thus, this alphabet can either be written in the Hägg-notation or the *ABC*-notation. Since the latter is more convenient for our purposes, we take $\mathcal{A} = \mathcal{A}_P \equiv \{A, B, C\}$. \mathbb{S} is the set of M states that comprise the process; *i.e.*, $\mathbb{S} = \{\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_M\}$. Lastly, there is one $M \times M$ TM for each output symbol, so that $\mathbf{T} = \{\mathcal{T}^{[A]}, \mathcal{T}^{[B]}, \mathcal{T}^{[C]}\}$. These emission-labeled transition probability matrices are of the form:

$$\mathcal{T}^{[x]} = \begin{bmatrix} \Pr(x, \mathcal{S}_1 | \mathcal{S}_1) & \Pr(x, \mathcal{S}_2 | \mathcal{S}_1) & \cdots & \Pr(x, \mathcal{S}_M | \mathcal{S}_1) \\ \Pr(x, \mathcal{S}_1 | \mathcal{S}_2) & \Pr(x, \mathcal{S}_2 | \mathcal{S}_2) & \cdots & \Pr(x, \mathcal{S}_M | \mathcal{S}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \Pr(x, \mathcal{S}_1 | \mathcal{S}_M) & \Pr(x, \mathcal{S}_2 | \mathcal{S}_M) & \cdots & \Pr(x, \mathcal{S}_M | \mathcal{S}_M) \end{bmatrix},$$

where $x \in \mathcal{A}_P$ and $\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_M \in \mathbb{S}$.

It is often useful to have the total state-to-state TM, whose components are the probability of transitions independent of the output symbol, and it is given by the row-stochastic matrix $\mathcal{T} = \mathcal{T}^{[A]} + \mathcal{T}^{[B]} + \mathcal{T}^{[C]}$. There also exists a stationary distribution $\boldsymbol{\pi} = (\Pr(\mathcal{S}_1), \dots, \Pr(\mathcal{S}_M))$ over the hidden states, such that $\langle \boldsymbol{\pi} | = \langle \boldsymbol{\pi} | \mathcal{T}$. We make limited use of a bra-ket notation throughout the following, where bras $\langle \cdot |$ represent row vectors and kets $|\cdot\rangle$ represent column vectors. Bra-ket closures, $\langle \cdot \rangle$ or $\langle \cdot | \cdot \rangle$, are scalars and commute as a unit with anything.

In the Hägg representation, the state-to-state transition matrix is $\mathbf{T} = \mathbf{T}^{[0]} + \mathbf{T}^{[1]}$. In

that case the stationary distribution π_H can be obtained from $\langle \pi_H | = \langle \pi_H | T$.

HMMs are often depicted as labeled directed graphs called probabilistic finite-state automata (FSA) [112, 190]. When written using the *ABC*-notation, we refer to such an automaton as the *ABC*-machine and, similarly, when written in terms of the Hägg-notation, such an automaton is referred to as the Hägg-machine. It is a straightforward task to translate a Hägg-machine into an *ABC*-machine [205]. For completeness, we reproduce the minimal algorithm in the appendix.

We note that while Hägg-notation and Hägg-machines are useful shorthand, the primary mathematical object for the development here is the *ABC*-machine, since this describes the stacking process in the natural language of the $\{Q_\xi(n)\}$. It is, however, often easier to give just the Hägg-machine since the expansion procedure is straightforward. Fundamentally, however, it is the *ABC* sequences that directly relate to structure factors for the specimen. And, this practical consideration is the principle reason for using the *ABC*-notation and *ABC*-machines.

6.4.3 Mixing and Nonmixing Machines

When expanding the Hägg-machine into an *ABC*-machine, two important cases emerge: *mixing* and *nonmixing* Hägg-machines. Which of these two cases we are considering has implications for the resultant DP, and so it is important to distinguish them [205].

In the expansion process, the number of states is tripled to account for the possible degeneracy of the *ABC*-notation. That is, we require that the *ABC*-machine keep track of not only the relative orientation between adjacent MLs (as the Hägg-machine does), but also the absolute *A*, *B*, or *C* orientation. In doing so, we allow a state architecture that can accommodate this increased representation requirement. For *mixing* machines, the resultant FSA is *strongly connected*, such that any state is accessible to any other state in a finite number of transitions. We find that this is by far the more common case. For *nonmixing* machines, the resultant graph is not strongly connected, but instead breaks into three unconnected graphs, each retaining the state structure of the original Hägg-machine. Only one of these graphs is physically realized in any given specimen, and we may arbitrarily choose to treat just one of them. The deciding factor on whether

a machine is mixing or nonmixing depends on its architecture: if there exists at least one closed, nonself-intersecting state path that corresponds to an overall rotation of the specimen, then the machine is mixing. The closed path is called a *simple state cycle* (SSC) on a FSA or a *causal state cycle* (CSC) if the FSA is also an ϵ -machine. All of the examples we consider here are mixing machines over most of their parameter range.

6.4.4 Power Spectra

Since we are considering only finite-state HMMs, \mathcal{T} is a finite-dimensional square matrix, and so its spectrum is just its set of eigenvalues:

$$\Lambda_{\mathcal{T}} = \{\lambda \in \mathbb{C} : \det(\lambda \mathbb{I} - \mathcal{T}) = 0\} , \quad (6.3)$$

where \mathbb{I} is the $M \times M$ identity matrix. Since \mathcal{T} is row-stochastic (*i.e.*, all rows sum to one), all of its eigenvalues live on or within the unit circle in the complex plane. The connection between the operator's spectrum and the diffraction spectrum will become clear shortly. In brief, though, eigenvalues along the unit circle lead to Bragg peaks; eigenvalues within the unit circle are responsible for diffuse peaks associated with disorder—the diffuse diffraction pattern (DP) is the shadow that these eigen-contributions cast along the unit circle.

In the limit of infinite-length sequences,⁴ power spectra generally can be thought of as having three distinct contributions; namely, *pure point* (*pp*), *absolutely continuous* (*ac*), and *singular continuous* (*sc*). Thus, a typical power spectrum $\mathcal{P}(\omega)$ can be decomposed into [9, 101]:

$$\mathcal{P}(\omega) = \mathcal{P}_{pp}(\omega) + \mathcal{P}_{ac}(\omega) + \mathcal{P}_{sc}(\omega) . \quad (6.4)$$

Pure point spectra are physically realized as Bragg reflections in DPs, and diffuse or broadband scattering is associated with the absolutely continuous part. Singular continuous spectra are not often observed in DPs from quasi-one-dimensional crystals, although specimens can be engineered to have a singular continuous portion in the DP, as for example layered GaAs-AlAs heterostructures stacked according to the Thue-Morse process [7]. Since these more exotic processes are not expressible as finite-state HMMs, we do not consider them further for now.

⁴For power spectra of finite-length sequences, there is no clear distinction among these features.

It might be thought that the more pedestrian forms of disorder in layered materials—such as growth, deformation, or layer-displacement faults—always destroy the long-range periodicity along the stacking direction and, thus, ‘true’ Bragg reflections need not be treated. (This is in contrast to those cases where there is little disorder and the integrity of the Bragg reflections is largely preserved.) In fact, there are occasions, such as solid-state transformations in materials with competing interactions between MLs [123, 279] or those with disordered and degenerate ground states [280, 263] that do maintain long-range correlations. Hence, it is not possible to exclude the existence of Bragg reflections *a priori*. Thus, we generally consider both Bragg reflections (B) and diffuse scattering (D) here and write the DP $I(\ell)$ as having two contributions:

$$I(\ell) = I_B(\ell) + I_D(\ell) , \quad (6.5)$$

where $\ell \in \mathbb{R}$ is a continuous variable that indexes the magnitude of the perpendicular component of the diffracted wave: $k = \omega/c = 2\pi\ell/c$ with c being the distance between adjacent MLs of the crystal.

Fortunately, knowledge of the HMM allows us to select beforehand those values of ℓ potentially contributing Bragg reflections. Let $\Lambda_{\rho(\mathcal{T})} \equiv \{\lambda \in \Lambda_{\mathcal{T}} : |\lambda| = 1\}$. The values of ℓ for which $e^{i2\pi\ell} \in \Lambda_{\rho(\mathcal{T})}$ are the only ones where there may possibly exist Bragg reflections. It is immediately apparent, then, that the total number of Bragg reflections within a unit interval of ℓ in the DP cannot be more than the number M of HMM states. Conversely, the total number of Bragg reflections sets a minimum on M .

6.5 Diffraction Patterns from Hidden Markov Models

With definitions and notations in place, we now derive our main results: analytical expressions for the DP in terms of the parameters that define a given HMM. We split our treatment into two steps: (i) we first treat the diffuse part of the spectrum and, then, (ii) we treat those z -values ($z \equiv e^{i\omega} = e^{i2\pi\ell}$) corresponding to eigenvalues of the TM along the unit circle.

6.5.1 Diffuse Scattering

The corrected DP⁵ for CPSs along a row defined by $h_0 - k_0 = 1 \pmod{3}$, where h_0, k_0 are components of the reciprocal lattice vectors in the plane of the MLs, can be written as [280, 81, 266]:

$$I^{(N)}(\ell) = \frac{\sin^2(N\pi\ell)}{N \sin^2(\pi\ell)} - \frac{2\sqrt{3}}{N} \sum_{n=1}^N (N-n) \left[Q_c(n) \cos\left(2\pi n\ell + \frac{\pi}{6}\right) + Q_a(n) \cos\left(2\pi n\ell - \frac{\pi}{6}\right) \right] \quad (6.6)$$

$$= \frac{\sin^2(N\pi\ell)}{N \sin^2(\pi\ell)} - \frac{2\sqrt{3}}{N} \Re \left\{ \sum_{n=1}^N (N-n) \left[Q_c(n) e^{-i2\pi n\ell} e^{-i\pi/6} + Q_a(n) e^{-i2\pi n\ell} e^{i\pi/6} \right] \right\} .$$

$Q_c(n)$ and $Q_a(n)$ are the previously defined CFs and N is the total number of MLs in the specimen.⁶ The superscript N on $I^{(N)}(\ell)$ reminds us that this expression for the diffuse DP depends on the number of MLs. The first term in Eq. (6.6) is the Fejér kernel. As the number of MLs becomes infinite, this term will tend to a δ -function at integer values of ℓ , which may be altered or eliminated by δ -function contributions from the summation: an issue we address shortly. It is only the second term, the summation, that results in diffuse scattering even as $N \rightarrow \infty$. It has previously been shown [205] that the CFs, in turn, can be written in terms of the labeled and unlabeled TMs of the underlying stacking process as:

$$Q_\xi(n) = \sum_{x \in \mathcal{A}_P} \langle \boldsymbol{\pi} | \mathcal{T}^{[x]} \mathcal{T}^{n-1} \mathcal{T}^{[\hat{\xi}(x)]} | \mathbf{1} \rangle , \quad (6.7)$$

where we denote the asymptotic probability distribution over the HMM states as the length- M row vector $\langle \boldsymbol{\pi} |$ and a length- M column vector of 1s as $|\mathbf{1}\rangle$. For mixing processes,

⁵As previously done [266], we divide out those factors associated with experimental corrections to the observed DP, as well as the total number N of MLs, so that $I(\ell)$ has only those contributions arising from the stacking structure itself. Here and elsewhere, we refer to $I(\ell)$ simply as the DP. Note however that factoring out the structure factors, etc. is not necessary. A more general expression is obtained by retaining these effects in the expression for the diffracted intensity. Note that the treatment by Treacy *et al.* [252] retains these factors.

⁶It may seem that specializing to such a specific expression for the DP at this stage limits the applicability of the approach. While the development here is restricted to the case of CPS, under mild conditions, the Wiener-Khinchin theorem [9] guarantees that power spectra can be written in terms of pair autocorrelation functions, as is done here. This makes the spectral decomposition rather generic.

Eq. (6.7) simplifies to the more restricted set of equations:

$$Q_\xi(n) = 3 \langle \boldsymbol{\pi} | \mathcal{T}^{[x]} \mathcal{T}^{n-1} \mathcal{T}^{[\hat{\xi}(x)]} | \mathbf{1} \rangle, \text{ where } x \in \mathcal{A}_P. \quad (6.8)$$

Thus, we can rewrite the DP directly in terms of the TMs of the underlying stacking process as:

$$I^{(N)}(\ell) = \frac{\sin^2(N\pi\ell)}{N \sin^2(\pi\ell)} - \frac{2\sqrt{3}}{N} \Re \left\{ \sum_{x \in \mathcal{A}_P} \langle \boldsymbol{\pi} | \mathcal{T}^{[x]} \left(\sum_{n=1}^N (N-n) z^{-n} \mathcal{T}^{n-1} \right) (e^{-i\pi/6} \mathcal{T}^{[\hat{c}(x)]} + e^{i\pi/6} \mathcal{T}^{[\hat{a}(x)]}) | \mathbf{1} \rangle \right\}, \quad (6.9)$$

where we have introduced the ℓ -dependent variable $z \equiv e^{i2\pi\ell}$. Furthermore, we can evaluate the summation over n in Eq. (6.9) analytically. First, we note that the summation can be re-indexed and split up as:

$$\sum_{n=1}^N (N-n) z^{-n} \mathcal{T}^{n-1} = z^{-1} \sum_{\eta=0}^{N-1} (N-1-\eta) (\mathcal{T}/z)^\eta \quad (6.10)$$

$$= z^{-1} \left\{ (N-1) \left[\sum_{\eta=0}^{N-1} (\mathcal{T}/z)^\eta \right] - \left[\sum_{\eta=0}^{N-1} \eta (\mathcal{T}/z)^\eta \right] \right\}. \quad (6.11)$$

For finite positive integer N , it is always true that:

$$(z\mathbb{I} - \mathcal{T}) \sum_{\eta=0}^{N-1} (\mathcal{T}/z)^\eta = z [\mathbb{I} - (\mathcal{T}/z)^N] \quad (6.12)$$

and

$$(z\mathbb{I} - \mathcal{T}) \sum_{\eta=0}^{N-1} \eta (\mathcal{T}/z)^\eta = z \left\{ \left[\sum_{\eta=0}^{N-1} (\mathcal{T}/z)^\eta \right] - N(\mathcal{T}/z)^N - [\mathbb{I} - (\mathcal{T}/z)^N] \right\}. \quad (6.13)$$

Hence, for $z \notin \Lambda_{\mathcal{T}}$, $z\mathbb{I} - \mathcal{T}$ is invertible and we have:

$$\sum_{n=1}^N (N-n) z^{-n} \mathcal{T}^{n-1} = (z\mathbb{I} - \mathcal{T})^{-1} \{ N\mathbb{I} - z(z\mathbb{I} - \mathcal{T})^{-1} [\mathbb{I} - (\mathcal{T}/z)^N] \}. \quad (6.14)$$

Putting this all together, we find the expected value of the finite- N DP for all $z = e^{i2\pi\ell} \notin \Lambda_{\mathcal{T}}$:

$$\begin{aligned} I^{(N)}(\ell) &= \frac{\sin^2(N\pi\ell)}{N \sin^2(\pi\ell)} \\ &\quad - 2\sqrt{3} \Re \left\{ \sum_{x \in \mathcal{A}_P} \langle \boldsymbol{\pi} | \mathcal{T}^{[x]} (z\mathbb{I} - \mathcal{T})^{-1} \left\{ \mathbb{I} - \frac{z}{N} (z\mathbb{I} - \mathcal{T})^{-1} [\mathbb{I} - (\mathcal{T}/z)^N] \right\} \right. \\ &\quad \times \left. (e^{-i\pi/6} \mathcal{T}^{[\hat{c}(x)]} + e^{i\pi/6} \mathcal{T}^{[\hat{a}(x)]}) | \mathbf{1} \rangle \right\}, \end{aligned} \quad (6.15)$$

with $z \equiv e^{i2\pi\ell}$. This gives the most general relationship between the DP and the TMs of the underlying stacking process. We see that the effects of finite crystal size come into the diffuse DP via a $1/N$ -decaying term containing the N^{th} power of both z^{-1} and the unlabeled TM. This powerful result directly links the stacking process rules to the observed DP and, additionally, already includes the effects of finite specimen size.

Eq. (6.15) should be compared with Eqs. (16) and (17) of Treacy *et al.*[252]. Note the similar form, especially that the *resolvent* [205], $(z\mathbb{I} - \mathcal{T})^{-1}$, finds an analog in the term $(I - T)^{-1}$ of Eq. (16) of Treacy *et al.* However our result is obtained by assuming a HMM that describes the stacking process, and not a judicious grouping of terms as has been done previously [252]. Whether the results of Treacy *et al.* are sufficiently general to include infinite-order Markov models, or can be modified to do so, remains an open question.

For many cases of practical interest, the specimen can be treated as effectively infinite along the stacking direction. (In follow-on work, we explore the effects of finite specimen size.) In this limiting case, the relationship between the diffuse DP and the TMs becomes especially simple. In particular, as $N \rightarrow \infty$ the DP's diffuse part becomes:

$$\begin{aligned} l_D(\ell) &= \lim_{N \rightarrow \infty} l^{(N)}(\ell) \\ &= -2\sqrt{3} \Re \left\{ \sum_{x \in \mathcal{A}_P} \langle \boldsymbol{\pi} | \mathcal{T}^{[x]} (z\mathbb{I} - \mathcal{T})^{-1} (e^{-i\pi/6} \mathcal{T}^{[\hat{c}(x)]} + e^{i\pi/6} \mathcal{T}^{[\hat{a}(x)]}) | \mathbf{1} \rangle \right\} , \end{aligned} \quad (6.16)$$

for all $z = e^{i2\pi\ell} \notin \Lambda_{\mathcal{T}}$. For mixing processes, this reduces to:

$$l_D(\ell) = -6\sqrt{3} \Re \left\{ \langle \boldsymbol{\pi} | \mathcal{T}^{[x]} (z\mathbb{I} - \mathcal{T})^{-1} (e^{-i\pi/6} \mathcal{T}^{[\hat{c}(x)]} + e^{i\pi/6} \mathcal{T}^{[\hat{a}(x)]}) | \mathbf{1} \rangle \right\} , \quad (6.17)$$

for any $x \in \mathcal{A}_P$. Note that there are no powers of the TM that need to be calculated in either of these cases. Rather, *the DP is a direct fingerprint of the noniterated TMs*. The simple elegance of Eq. (6.16) relating the DP and TMs suggests that there is a link of fundamental conceptual importance between them. The examples to follow draw out this connection.

The important role that \mathcal{T} 's eigenvalues $\Lambda_{\mathcal{T}}$ play in the DP should now be clear: they are the poles of the resolvent matrix $(\zeta\mathbb{I} - \mathcal{T})^{-1}$ with $\zeta \in \mathbb{C}$. Since the DP is a simple

function of the resolvent evaluated along the unit circle, $\Lambda_{\mathcal{T}}$ plays a critical organizational role in the DP's structure. Any peaks in the DP are shadows of the poles of the resolvent filtered through the appropriate row and column vectors and cast out radially onto the unit circle. Peaks in the DP become more diffuse as the corresponding eigenvalues withdraw towards the origin of the complex plane. They approach δ -functions as the corresponding eigenvalues approach the unit circle. §6.6's examples demonstrate this graphically.

6.5.2 Bragg Reflections

The eigenvalues $\Lambda_{\rho(\mathcal{T})} \subset \Lambda_{\mathcal{T}}$ along the unit circle are responsible for Bragg peaks, and we treat this case now. For finite- N , the eigenvalues along the unit circle give rise to Dirichlet kernels. As $N \rightarrow \infty$, the analysis becomes somewhat simpler since the Dirichlet kernel and Fejér kernel both tend to δ -functions.

As $N \rightarrow \infty$, the summation over n in Eq. (6.9) divided by the total number of MLs becomes:

$$\lim_{N \rightarrow \infty} \sum_{n=1}^N \frac{N-n}{N} z^{-n} \mathcal{T}^{n-1} = z^{-1} \sum_{\eta=0}^{\infty} (\mathcal{T}/z)^{\eta} . \quad (6.18)$$

At this point, it is pertinent to use the recently developed spectral decomposition of \mathcal{T}^L [205]. With the allowance that $0^{L-m} = \delta_{L-m,0}$ for the case that $0 \in \Lambda_{\mathcal{T}}$, this is:

$$\mathcal{T}^L = \sum_{\lambda \in \Lambda_{\mathcal{T}}} \sum_{m=0}^{\nu_{\lambda}-1} \lambda^{L-m} \binom{L}{m} \mathcal{T}_{\lambda} (\mathcal{T} - \lambda \mathbb{I})^m , \quad (6.19)$$

where (i) \mathcal{T}_{λ} is the projection operator associated with the eigenvalue λ given by the elementwise residue of the resolvent $(z\mathbb{I} - \mathcal{T})^{-1}$ at $z \rightarrow \lambda$, (ii) the index ν_{λ} of the eigenvalue λ is the size of the largest Jordan block associated with λ , and (iii) $\binom{L}{m} = \frac{1}{m!} \prod_{n=1}^m (L-n+1)$ is the generalized binomial coefficient. In terms of elementwise counter-clockwise contour integration, we have:

$$\mathcal{T}_{\lambda} = \frac{1}{2\pi i} \oint_{C_{\lambda}} (z\mathbb{I} - \mathcal{T})^{-1} dz , \quad (6.20)$$

where C_{λ} is any contour in the complex plane enclosing the point $z_0 = \lambda$ —which may or may not be a singularity depending on the particular element of the resolvent matrix—but

encloses no other singularities. Usefully, the projection operators are a mutually orthogonal set such that for $\zeta, \lambda \in \Lambda_{\mathcal{T}}$, we have:

$$\mathcal{T}_{\zeta} \mathcal{T}_{\lambda} = \delta_{\zeta, \lambda} \mathcal{T}_{\lambda} .$$

The Perron–Frobenius theorem guarantees that all eigenvalues of the stochastic TM \mathcal{T} lie on or within the unit circle. Moreover—and very important to our discussion on Bragg reflections—the eigenvalues on the unit circle are guaranteed to have an index of one. The indices of all other eigenvalues must be less than or equal to one more than the difference between their algebraic a_{λ} and geometric g_{λ} multiplicities. Specifically:

$$\nu_{\lambda} - 1 \leq a_{\lambda} - g_{\lambda} \leq a_{\lambda} - 1$$

and

$$\nu_{\lambda} = 1 \text{ if } |\lambda| = 1 .$$

Taking advantage of the index-one nature of the eigenvalues on the unit circle, we can define:

$$\Xi \equiv \sum_{\zeta \in \Lambda_{\rho(\mathcal{T})}} \zeta \mathcal{T}_{\zeta}$$

and

$$F \equiv \mathcal{T} - \Xi .$$

Then, the summation on the right-hand side of Eq. (6.18) becomes:

$$\sum_{\eta=0}^{\infty} (\mathcal{T}/z)^{\eta} = \left[\sum_{\eta=0}^{\infty} (F/z)^{\eta} \right] + \left[\sum_{\eta=0}^{\infty} (\Xi/z)^{\eta} \right] . \quad (6.21)$$

In the above, only the summation involving Ξ is capable of contributing δ -functions. And so, expanding this summation, yields:

$$\sum_{\eta=0}^{\infty} (\Xi/z)^{\eta} = \sum_{\lambda \in \Lambda_{\rho(\mathcal{T})}} \mathcal{T}_{\lambda} \sum_{\eta=0}^{\infty} (\lambda/z)^{\eta} \quad (6.22)$$

$$= \sum_{\lambda \in \Lambda_{\rho(\mathcal{T})}} \mathcal{T}_{\lambda} \sum_{\eta=0}^{\infty} e^{i2\pi(\ell_{\lambda}-\ell)\eta} , \quad (6.23)$$

where ℓ_λ is related to λ by $\lambda = e^{i2\pi\ell_\lambda}$ over some appropriate length-one ℓ -interval.

Using properties of the discrete-time Fourier transform [182], we can finally pull the δ -functions out of Eq. (6.23). In particular:

$$\sum_{\eta=0}^{\infty} e^{i2\pi(\ell_\lambda - \ell)\eta} = \frac{1}{1 - e^{i2\pi(\ell_\lambda - \ell)}} + \sum_{k=-\infty}^{\infty} \frac{1}{2} \delta(\ell - \ell_\lambda + k) . \quad (6.24)$$

Identifying the context of Eq. (6.24) within Eq. (6.9) shows that the potential δ -function at ℓ_λ (and at its integer-offset values) has magnitude:⁷

$$\begin{aligned} \Delta_\lambda &\equiv \lim_{\epsilon \rightarrow 0} \int_{\ell_\lambda - \epsilon}^{\ell_\lambda + \epsilon} I(\ell) d\ell \\ &= -\sqrt{3} \Re \left\{ \lambda^{-1} \left[\left\langle \mathcal{T}_\lambda^{\hat{c}(\mathcal{A})} \right\rangle e^{-i\pi/6} + \left\langle \mathcal{T}_\lambda^{\hat{a}(\mathcal{A})} \right\rangle e^{i\pi/6} \right] \right\} \end{aligned} \quad (6.25)$$

contributed via the summation of Eq. (6.9), where:

$$\left\langle \mathcal{T}_\lambda^{\xi(\mathcal{A})} \right\rangle \equiv \sum_{x_0 \in \mathcal{A}_P} \langle \boldsymbol{\pi} | \mathcal{T}^{[x_0]} \mathcal{T}_\lambda \mathcal{T}^{[\hat{\xi}(x_0)]} | \mathbf{1} \rangle . \quad (6.26)$$

Finally, considering Eq. (6.25) together with the contribution of the persistent Fejér kernel, the discrete part of the DP is given by:

$$I_B(\ell) = \sum_{k=-\infty}^{\infty} \sum_{\lambda \in \Lambda_\rho(\mathcal{T})} (\delta_{\lambda,1} + \Delta_\lambda) \delta(\ell - \ell_\lambda + k) , \quad (6.27)$$

where $\delta_{\lambda,1}$ is a Kronecker delta and $\delta(\ell - \ell_\lambda + k)$ is a Dirac δ -function.

In particular, the presence of the Bragg reflection at integer ℓ (zero frequency) depends strongly on whether the stacking process is mixing. In any case, the magnitude of these δ -functions at integer ℓ is $1 + \Delta_1$. For an ergodic process $\mathcal{T}_1 = |\mathbf{1}\rangle \langle \boldsymbol{\pi}|$, so we have:

$$\left\langle \mathcal{T}_1^{\hat{\xi}(\mathcal{A})} \right\rangle = \sum_{x_0 \in \mathcal{A}_P} \langle \boldsymbol{\pi} | \mathcal{T}^{[x_0]} | \mathbf{1} \rangle \langle \boldsymbol{\pi} | \mathcal{T}^{[\hat{\xi}(x_0)]} | \mathbf{1} \rangle . \quad (6.28)$$

⁷By magnitude, we mean the ℓ -integral over the δ -function. If integrating with respect to a related variable, then the magnitude of the δ -function changes accordingly. As a simple example, integrating over $\omega = 2\pi\ell$ changes the magnitude of the δ -function by a factor of 2π .

For *mixing* ABC-machines, $\langle \boldsymbol{\pi} | \mathcal{T}^{[x]} | \mathbf{1} \rangle = \Pr(x) = 1/3$ for all $x \in \mathcal{A}_P$, giving $\langle \mathcal{T}_1^{\xi(\mathcal{A})} \rangle = 1/3$. Hence:

$$\begin{aligned} \Delta_1 &= -\frac{\sqrt{3}}{3} \Re \{ e^{-i\pi/6} + e^{i\pi/6} \} \\ &= -\frac{2\sqrt{3}}{3} \cos(\pi/6) \\ &= -1, \end{aligned} \tag{6.29}$$

and the integer- ℓ δ -functions are extinguished for all mixing processes.

For *nonmixing* processes, the probability of each ML is *not* necessarily the same, and the magnitude of the δ -function at integer- ℓ will reflect the heterogeneity of the single-symbol statistics.

6.5.3 Full Spectral Treatment of the Diffuse Spectrum

From Eq. (6.16), it is clear that the diffuse part of the DP is directly related to the resolvent $(z\mathbb{I} - \mathcal{T})^{-1}$ of the state-to-state TM evaluated along the unit circle. According to Riechers & Crutchfield [205] the resolvent can be expressed in terms of the projection operators:

$$(z\mathbb{I} - \mathcal{T})^{-1} = \sum_{\lambda \in \Lambda_{\mathcal{T}}} \sum_{m=0}^{\nu_{\lambda}-1} \frac{1}{(z-\lambda)^{m+1}} \mathcal{T}_{\lambda} (\mathcal{T} - \lambda\mathbb{I})^m. \tag{6.30}$$

Hence, Eq. (6.16) can be expressed as:

$$\mathbf{l}_D(\ell) = -2\sqrt{3} \Re \left\{ \sum_{\lambda \in \Lambda_{\mathcal{T}}} \sum_{m=0}^{\nu_{\lambda}-1} \frac{1}{(z-\lambda)^{m+1}} \left[\langle \mathcal{T}_{\lambda,m}^{\hat{c}(\mathcal{A})} \rangle e^{-i\pi/6} + \langle \mathcal{T}_{\lambda,m}^{\hat{a}(\mathcal{A})} \rangle e^{i\pi/6} \right] \right\}, \tag{6.31}$$

where $\langle \mathcal{T}_{\lambda,m}^{\xi(\mathcal{A})} \rangle$ is a complex-valued scalar:⁸

$$\langle \mathcal{T}_{\lambda,m}^{\xi(\mathcal{A})} \rangle \equiv \sum_{x_0 \in \mathcal{A}_P} \langle \boldsymbol{\pi} | \mathcal{T}^{[x_0]} \mathcal{T}_{\lambda} (\mathcal{T} - \lambda\mathbb{I})^m \mathcal{T}^{[\hat{\xi}(x_0)]} | \mathbf{1} \rangle. \tag{6.32}$$

Moreover, if $\langle \mathcal{T}_{\lambda,m}^{\hat{c}(\mathcal{A})} \rangle = \langle \mathcal{T}_{\lambda,m}^{\hat{a}(\mathcal{A})} \rangle$ for all λ and all m , then Eq. (6.31) simplifies to:

$$\mathbf{l}_D(\ell) = -6 \Re \left\{ \sum_{\lambda \in \Lambda_{\mathcal{T}}} \sum_{m=0}^{\nu_{\lambda}-1} \frac{\langle \mathcal{T}_{\lambda,m}^{\hat{c}(\mathcal{A})} \rangle}{(z-\lambda)^{m+1}} \right\}. \tag{6.33}$$

⁸ $\langle \mathcal{T}_{\lambda,m}^{\xi(\mathcal{A})} \rangle$ is constant with respect to the relative layer displacement n . However, $\{ \langle \mathcal{T}_{\lambda,m}^{\xi(\mathcal{A})} \rangle \}$ can be a function of a process's parameters.

6.6 Examples

To illustrate the theory, we treat in some detail three examples for which we previously [205] estimated the CFs directly from the HMM. Throughout the examples, we find it particularly revealing to plot the DP and TM eigenvalues via, what we call, the *coronal spectrogram*. This takes advantage of the fact that the DP is periodic in ℓ with period one and that the TM's eigenvalues lie on or within the unit circle in the complex plane. Thus, a coronal spectrogram is any frequency-dependent graph emanating radially from the unit circle, while the unit circle and its interior are concurrently used for its portion of the complex plane to plot the poles of the resolvent of the underlying process's transition dynamic. (Here, the poles of the resolvent are simply the eigenvalues $\Lambda_{\mathcal{T}}$ of \mathcal{T} , since \mathcal{T} is finite dimensional.)

Coronal spectrograms plot the DP as a function of the polar angle $\omega = 2\pi\ell$. The radial extent of the corona is normalized to have the same maximal value for each figure here. With our particular interest in the DP of CPSs, we plot all eigenvalues in $\Lambda_{\mathcal{T}}$ as (red, online) dots and also plot all eigenvalues in $\Lambda_{\mathcal{T}}$ as (black) \times s. Note that $\Lambda_{\mathcal{T}} \subset \Lambda_{\mathcal{T}}$. In all of our examples, it appears that only the eigenvalues introduced in generating the *ABC*-machine from the Hägg-machine (dots without \times s through them) are capable of producing DP peaks. For *nonmixing* processes this is not true, since the Hägg-machine and *ABC*-machine share the same topology and the same set of eigenvalues.

6.6.1 3C Polytypes and Random ML Stacking: IID Process

The independent and identically distributed Hägg process is the simplest ML stacking process in a CPS that one can consider. Although we work out this example largely as a pedagogical exercise, in limiting cases it can be thought of as random deformation faulting in face-center cubic (FCC) (aka 3C) crystals.

We define the *independent and identically distributed* (IID) stacking process as such: when transitioning between adjacent MLs, a ML will be cyclically related to the previous ML with probability $q \in [0, 1]$. Due to stacking constraints, the ML will otherwise be anticyclically related to its predecessor with probability $\bar{q} \equiv 1 - q$.⁹ The Hägg-machine

⁹Here and in the following examples, we define a bar over a variable to mean one minus that variable:

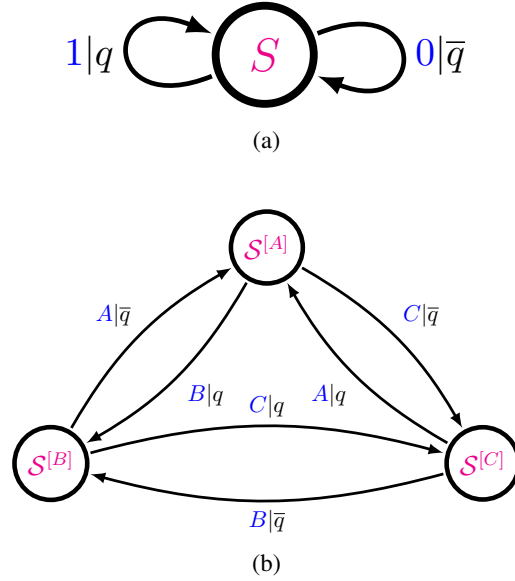


Figure 6.1: The (a) Hägg-machine and the (b) ABC -machine for the IID Process, $q \in [0, 1]$. When $q = 1$, the IID process generates a string of 1s, which is physically the $3C^+$ stacking structure. Conversely, for $q = 0$, the structure corresponds to the $3C^-$ structure. For $q = 0.5$, the MLs are stacked as randomly as possible. Notice that the single state of the Hägg-machine has split into a three-state ABC -machine. This trebling of states is a generic feature of expanding mixing Hägg-machines into ABC -machines. It should be observed when written in the ABC -notation the process is actually a MM. Changing from one stacking notation to another can affect the order of the Markov model, and is discussed elsewhere [269]. (From Riechers *et al.*[205], used with permission.)

and ABC -machine for the IID Process are given in Fig. 6.1.

It is useful to consider limiting cases for q . When $q = 0.5$, the stacking is completely random, subject only to the stacking constraints preventing two adjacent MLs from having the same orientation. As $q \rightarrow 1$, adjacent MLs are almost always cyclically related, and the specimen can be thought of as a $3C^+$ crystal with randomly distributed deformation faults [273] with probability \bar{q} . As $q \rightarrow 0$, it is also a $3C$ crystal with randomly distributed deformation faults, except that the MLs are anticyclically related, which we denote as $3C^-$. This is summarized in Table 6.1.

The TMs in ABC -notation are:

$$\bar{x} \equiv 1 - x.$$

Table 6.1: The limiting material structures for the IID Process. Key: DF - deformation fault; Ran - completely random stacking.

$q = 0$	$q \approx 0$	$q = \bar{q} = \frac{1}{2}$	$\bar{q} \approx 0$	$\bar{q} = 0$
$3C^-$	$3C^-/\text{DF}$	Ran	$3C^+/\text{DF}$	$3C^+$

$$\mathcal{T}^{[A]} = \begin{bmatrix} 0 & 0 & 0 \\ \bar{q} & 0 & 0 \\ q & 0 & 0 \end{bmatrix}, \quad \mathcal{T}^{[B]} = \begin{bmatrix} 0 & q & 0 \\ 0 & 0 & 0 \\ 0 & \bar{q} & 0 \end{bmatrix} \quad \text{and} \quad \mathcal{T}^{[C]} = \begin{bmatrix} 0 & 0 & \bar{q} \\ 0 & 0 & q \\ 0 & 0 & 0 \end{bmatrix}.$$

The internal state TM then is their sum:

$$\mathcal{T} = \begin{bmatrix} 0 & q & \bar{q} \\ \bar{q} & 0 & q \\ q & \bar{q} & 0 \end{bmatrix}.$$

The eigenvalues of the ABC TM are

$$\Lambda_{\mathcal{T}} = \{1, \Omega, \Omega^*\},$$

where:

$$\Omega \equiv -\frac{1}{2} + i\frac{\sqrt{3}}{2}(4q^2 - 4q + 1)^{1/2}$$

and Ω^* is its complex conjugate.

Furthermore, the stationary distribution over states of the ABC -machine can be found from $\langle \boldsymbol{\pi} | = \langle \boldsymbol{\pi} | \mathcal{T}$ to be:

$$\langle \boldsymbol{\pi} | = \left[\frac{1}{3} \quad \frac{1}{3} \quad \frac{1}{3} \right].$$

For $q \in (0, 1)$, none of the eigenvalues in $\Lambda_{\mathcal{T}}$ besides unity lie on the unit circle of the complex plane, and so there is no possibility of Bragg reflections at non-integer ℓ . Moreover, since the process is mixing, the Bragg peak at integer ℓ is also absent. Thus we need only find the diffuse DP. To calculate the $I_D(\ell)$ as given in Eq. (6.16), we are only

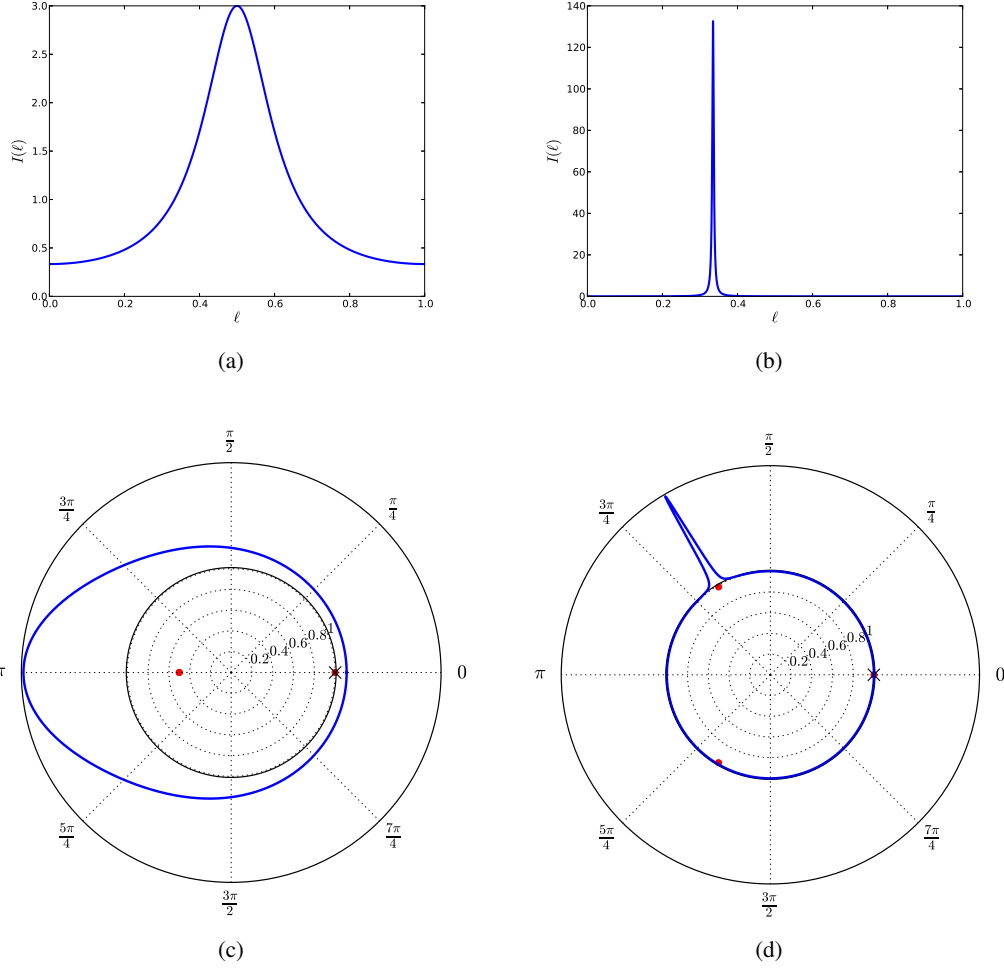


Figure 6.2: IID Process diffraction patterns for (a) $q = 0.5$ and (b) $q = 0.99$, as calculated from Eq. (6.35). Notice that as $q \rightarrow 1$, the DP approaches that of a $3C^+$ crystal. For values of q close to but less than 1, the specimen is $3C^+$ with randomly distributed deformation faults. (c) The coronal spectrogram corresponding to $q = 0.5$. The enhanced scattering at $\ell = 0.5$ in (a) is replaced with the bulge at $\omega = \pi$. There are three eigenvalues for the IID Process, one at $z = 1$ and a degenerate pair at $z = -0.5$. (d) The coronal spectrogram corresponding to $q = 0.99$. The Bragg-like peak at $\ell = 0.33$ in (b) is now represented as a Bragg-like peak at $\omega = 2\pi/3$. Notice how the degenerate eigenvalues in (c) have split and migrated away from the real axis. As they approach the boundary of the unit circle, their presence makes possible Bragg-like reflections in the DP. However, eigenvalues near the unit circle are a *necessary*, but *not sufficient* condition for Bragg-like reflections. This is seen in the eigenvalue in the third quadrant that is *not* accompanied by a Bragg-like reflection.

missing $(z\mathbb{I} - \mathcal{T})^{-1}$, which is given by:

$$(z\mathbb{I} - \mathcal{T})^{-1} = \frac{1}{(z-1)(z-\Omega)(z-\Omega^*)} \times \begin{bmatrix} z^2 - q\bar{q} & qz + \bar{q}^2 & \bar{q}z + q^2 \\ \bar{q}z + q^2 & z^2 - q\bar{q} & qz + \bar{q}^2 \\ qz + \bar{q}^2 & \bar{q}z + q^2 & z^2 - q\bar{q} \end{bmatrix}.$$

Then, with:

$$\langle \boldsymbol{\pi} | \mathcal{T}^{[A]} = \begin{bmatrix} \frac{1}{3} & 0 & 0 \end{bmatrix},$$

we can write

$$\langle \boldsymbol{\pi} | \mathcal{T}^{[A]} (z\mathbb{I} - \mathcal{T})^{-1} = \frac{1}{3} \frac{1}{(z-1)(z-\Omega)(z-\Omega^*)} \times \begin{bmatrix} z^2 - q\bar{q} & qz + \bar{q}^2 & \bar{q}z + q^2 \end{bmatrix},$$

where:

$$\mathcal{T}^{[\hat{c}(A)]} |\mathbf{1}\rangle = \mathcal{T}^{[B]} |\mathbf{1}\rangle = \begin{bmatrix} q \\ 0 \\ \bar{q} \end{bmatrix} \quad \text{and} \quad \mathcal{T}^{[\hat{a}(A)]} |\mathbf{1}\rangle = \mathcal{T}^{[C]} |\mathbf{1}\rangle = \begin{bmatrix} \bar{q} \\ q \\ 0 \end{bmatrix}.$$

From Eq. (6.17), the DP becomes:

$$I_D(\ell) = -\frac{6\sqrt{3}}{2} \Re \left\{ \frac{e^{i\pi/6} z (q^2 + \bar{q}z) + e^{-i\pi/6} z (\bar{q}^2 + qz)}{(z-1)(z-\Omega)(z-\Omega^*)} \right\} \quad (6.34)$$

$$= -2\sqrt{3} \Re \left\{ z \frac{e^{i\pi/6} (q^2 + \bar{q}z) + e^{-i\pi/6} (\bar{q}^2 + qz)}{(z-1)(z^2 + z + 1 - 3q\bar{q})} \right\}. \quad (6.35)$$

For the case of the most random possible stacking in CPSs, where $q = \bar{q} = \frac{1}{2}$, this simplifies to:

$$I_D(\ell) = \frac{3/4}{5/4 + \cos(2\pi\ell)}. \quad (6.36)$$

This result was obtained previously by more elementary means. The results are in agreement [103].

Figure 6.2 shows DPs and coronal spectrograms for $q = 0.5$ and $q = 0.99$. Figure 6.2(a) gives the DP for a maximally disordered stacking process. The spectrum is entirely diffuse with broadband enhancement near $\ell = 0.5$. In contrast, the DP for $q = 0.99$ in Fig. 6.2(b) shows a strong Bragg-like reflection at $\ell = 0.33$, which we recognize as just the $3C^+$

stacking structure, with a small amount of (as it turns out in this case) deformation faulting. The other two panels in Fig. 6.2, (c) and (d), are coronal spectrograms giving DPs for these two cases as the radially emanating curve outside the unit circle, but now the three eigenvalues of the total TM are plotted interior to the unit circle. As always, there is a single eigenvalue at $z = 1$. In panel (c), the other two degenerate eigenvalues occur at $z = -0.5$, ‘casting a shadow’ on the unit circle in the form of enhanced power at $\omega = \pi$. In panel (d), these eigenvalues split and move away from the real axis closer to the unit circle. In doing so, one casts a more focused shadow in the form of a Bragg-like reflection at $\omega = 2\pi/3$. For $q = 1$, this eigenvalue finally comes to rest on the unit circle, and the Bragg-like reflection becomes a true Bragg peak, as explored shortly. Note that the other eigenvalue does not give rise to enhanced scattering. We find that having an eigenvalue near the unit circle is necessary to produce enhanced scattering, but the presence of such an eigenvalue does not necessarily guarantee Bragg-like reflections.

6.6.1.1 Bragg Peaks from 3C

For the case of $q \in \{0, 1\}$, we recover perfect crystalline structure. Although the presence, placement, and magnitude of Bragg peaks are well known from other methods, we show the comprehensive consistency of our method via the example of $q = 1$ ($\bar{q} = 0$). In this case: $\Lambda_{\mathcal{T}} = \{1, \Omega, \Omega^*\}$ with $\Omega = -\frac{1}{2} + i\frac{\sqrt{3}}{2} = e^{i2\pi/3}$, so that $\ell_{\Omega} = 1/3$ and $\ell_{\Omega^*} = 2/3$, and the two relevant projection operators reduce to:

$$\mathcal{T}_{\Omega} = \frac{1}{(\Omega - 1)(\Omega - \Omega^*)} \begin{bmatrix} \Omega^2 & \Omega & 1 \\ 1 & \Omega^2 & \Omega \\ \Omega & 1 & \Omega^2 \end{bmatrix} \text{ and } \mathcal{T}_{\Omega^*} = \frac{1}{(\Omega^* - 1)(\Omega^* - \Omega)} \begin{bmatrix} \Omega^{*2} & \Omega^* & 1 \\ 1 & \Omega^{*2} & \Omega^* \\ \Omega^* & 1 & \Omega^{*2} \end{bmatrix}.$$

From Eq. (6.26), we have:

$$\langle \mathcal{T}_{\Omega}^{\hat{\mathcal{C}}(\mathcal{A})} \rangle = \frac{\Omega^2}{(\Omega - 1)(\Omega - \Omega^*)}, \quad \langle \mathcal{T}_{\Omega}^{\hat{\mathcal{A}}(\mathcal{A})} \rangle = \frac{\Omega}{(\Omega - 1)(\Omega - \Omega^*)},$$

and

$$\langle \mathcal{T}_{\Omega^*}^{\hat{\mathcal{C}}(\mathcal{A})} \rangle = \frac{\Omega^{*2}}{(\Omega^* - 1)(\Omega^* - \Omega)}, \quad \langle \mathcal{T}_{\Omega^*}^{\hat{\mathcal{A}}(\mathcal{A})} \rangle = \frac{\Omega^*}{(\Omega^* - 1)(\Omega^* - \Omega)},$$

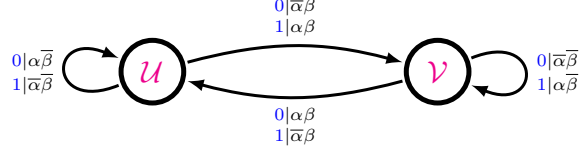


Figure 6.3: The Hägg-machine for the RGDF Process as proposed by Estevez *et al.*[85]. This two-state machine is a niMM and has two parameters, $\alpha \in [0, 1]$ and $\beta \in [0, 1]$, the probability of deformation and growth faults in CPSs, respectively. (From Riechers *et al.*[205], used with permission.)

which from Eq. (6.25) yields:

$$\Delta_{\Omega} = 1 \quad \text{and} \quad \Delta_{\Omega^*} = 0 .$$

Then, using Eq. (6.27), the DP's discrete part becomes:

$$l_B(\ell) = \sum_{k=-\infty}^{\infty} \delta(\ell - \frac{1}{3} + k) ,$$

as it ought to be for $3C^+$.

6.6.2 Random Growth and Deformation Faults in Layered 3C and 2H CPSs: The RGDF Process

As a simple model of faulting in CPSs, combined random growth and deformation faults are often assumed if the faulting probabilities are believed to be small. However, until now there has not been an analytical expression available for the DP for all values of the faulting parameters, and we derive such an expression here.

The HMM for the *Random Growth and Deformation Faults* (RGDF) process was first proposed by Estevez-Rams *et al.* [85] and the Hägg-machine is shown in Fig. 6.3. The process has two parameters, $\alpha \in [0, 1]$ and $\beta \in [0, 1]$, that (at least for small values) are interpreted as the probability of deformation and growth faults, respectively. The stacking process, however, is described best on its own terms—in terms of the HMM, which captures the causal architecture of the stacking for all parameter values.

It is instructive to consider limiting values of α and β . For $\alpha = \beta = 0$, the stacking structure is simply 3C. The machine splits into two distinct machines: each machine has one state with a single self-state transition, corresponding to the $3C^+$ stacking

Table 6.2: The limiting material structures for the RGDF Process. Key: GF - growth fault; DF - deformation fault; Ran - completely random stacking.

	$\beta = 0$	$\beta \approx 0$	$\beta = \bar{\beta} = 1/2$	$\bar{\beta} \approx 0$	$\bar{\beta} = 0$
$\alpha = 0$	3C	3C/GF	Ran	2H/GF	2H
$\alpha \approx 0$	3C/DF	3C/DF,GF	Ran	2H/DF,GF	2H/DF
$\alpha = \frac{1}{2}$	Ran	Ran	Ran	Ran	Ran

structure and the other to 3C⁻ stacking structure. The 2H stacking structure occurs when $\alpha = \bar{\beta} = 0$. Typically growth faults are introduced as β strays from these limiting values, and deformation faults appear when α becomes small but nonvanishing. When $\alpha = 1/2$, the stacking becomes completely random, regardless of the value of β . This is summarized in Table 6.2.

The RGDF Hägg-machine's TMs are:

$$\mathsf{T}^{[0]} = \begin{bmatrix} \alpha\bar{\beta} & \bar{\alpha}\beta \\ \alpha\beta & \bar{\alpha}\bar{\beta} \end{bmatrix} \text{ and } \mathsf{T}^{[1]} = \begin{bmatrix} \bar{\alpha}\bar{\beta} & \alpha\beta \\ \bar{\alpha}\beta & \alpha\bar{\beta} \end{bmatrix}.$$

The Hägg-machine is nonmixing only for the parameter settings $\beta = 1$ and $\alpha \in \{0, 1\}$, giving rise to 2H crystal structure.

From the Hägg-machine, we obtain the corresponding TMs of the *ABC*-machine for $\alpha, \beta \in (0, 1)$ [205]:

$$\mathcal{T}^{[A]} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ \alpha\bar{\beta} & 0 & 0 & \bar{\alpha}\beta & 0 & 0 \\ \bar{\alpha}\bar{\beta} & 0 & 0 & \alpha\beta & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \alpha\beta & 0 & 0 & \bar{\alpha}\bar{\beta} & 0 & 0 \\ \bar{\alpha}\beta & 0 & 0 & \alpha\bar{\beta} & 0 & 0 \end{bmatrix}, \quad \mathcal{T}^{[B]} = \begin{bmatrix} 0 & \bar{\alpha}\bar{\beta} & 0 & 0 & \alpha\beta & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \alpha\bar{\beta} & 0 & 0 & \bar{\alpha}\beta & 0 \\ 0 & \bar{\alpha}\beta & 0 & 0 & \alpha\bar{\beta} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \alpha\beta & 0 & 0 & \bar{\alpha}\bar{\beta} & 0 \end{bmatrix}$$

and

$$\mathcal{T}^{[C]} = \begin{bmatrix} 0 & 0 & \alpha\bar{\beta} & 0 & 0 & \bar{\alpha}\beta \\ 0 & 0 & \bar{\alpha}\bar{\beta} & 0 & 0 & \alpha\beta \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \alpha\beta & 0 & 0 & \bar{\alpha}\bar{\beta} \\ 0 & 0 & \bar{\alpha}\beta & 0 & 0 & \alpha\bar{\beta} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

and the orientation-agnostic state-to-state TM:

$$\mathcal{T} = \mathcal{T}^{[A]} + \mathcal{T}^{[B]} + \mathcal{T}^{[C]}.$$

Explicitly, we have:

$$\mathcal{T} = \begin{bmatrix} 0 & \bar{\alpha}\bar{\beta} & \alpha\bar{\beta} & 0 & \alpha\beta & \bar{\alpha}\beta \\ \alpha\bar{\beta} & 0 & \bar{\alpha}\bar{\beta} & \bar{\alpha}\beta & 0 & \alpha\beta \\ \bar{\alpha}\bar{\beta} & \alpha\bar{\beta} & 0 & \alpha\beta & \bar{\alpha}\beta & 0 \\ 0 & \bar{\alpha}\beta & \alpha\beta & 0 & \alpha\bar{\beta} & \bar{\alpha}\bar{\beta} \\ \alpha\beta & 0 & \bar{\alpha}\beta & \bar{\alpha}\bar{\beta} & 0 & \alpha\bar{\beta} \\ \bar{\alpha}\beta & \alpha\beta & 0 & \alpha\bar{\beta} & \bar{\alpha}\bar{\beta} & 0 \end{bmatrix}.$$

\mathcal{T} 's eigenvalues satisfy $\det(\mathcal{T} - \lambda\mathbb{I}) = 0$, from which we obtain the eigenvalues [205]:

$$\Lambda_{\mathcal{T}} = \left\{ 1, 1 - 2\beta, -\frac{1}{2}(1 - \beta) \pm \frac{1}{2}\sqrt{\sigma} \right\}, \quad (6.37)$$

with

$$\sigma \equiv 4\beta^2 - 3\bar{\beta}^2 + 12\alpha\bar{\alpha}(\bar{\beta} - \beta) \quad (6.38)$$

$$= -3 + 12\alpha + 6\beta - 12\alpha^2 + \beta^2 - 24\alpha\beta + 24\alpha^2\beta. \quad (6.39)$$

Except for measure-zero submanifolds along which the eigenvalues become extra degenerate, throughout the parameter range the eigenvalues' algebraic multiplicities are: $a_1 = 1$, $a_{1-2\beta} = 1$, $a_{-\frac{1}{2}(1-\beta+\sqrt{\sigma})} = 2$, and $a_{-\frac{1}{2}(1-\beta-\sqrt{\sigma})} = 2$. Moreover, the *index* of all eigenvalues is 1 except along $\sigma = 0$. Hence, due to their qualitative difference, we treat the cases of $\sigma = 0$ and $\sigma \neq 0$ separately.

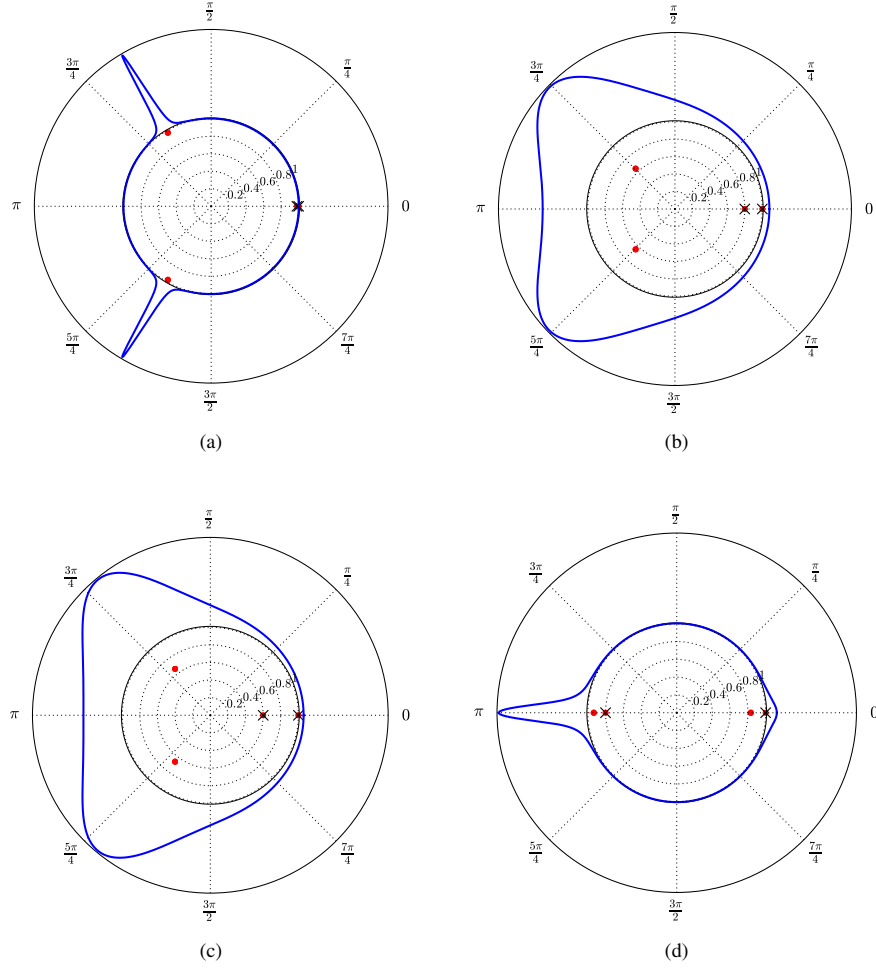


Figure 6.4: Coronal spectrograms showing the DP and eigenvalues for the RGDF Process: (a) $\alpha = 0.01$, $\beta = 0.01$; (b) $\alpha = 0.2$, $\beta = 0.1$; (c) $\alpha = 0.1$, $\beta = 0.2$; and (d) $\alpha = 0.01$, $\beta = 0.9$. Note how the eigenvalues organize the DP: as the eigenvalues approach the unit circle, the DP becomes enhanced. Also, note that nowhere is there enhanced scattering without an underlying eigenvalue of the TM driving it.

6.6.2.1 $\sigma = 0$:

Riechers *et al.* [205] found that:

$$\left\langle \mathcal{T}_1^{\hat{c}(\mathcal{A})} \right\rangle = \left\langle \mathcal{T}_1^{\hat{a}(\mathcal{A})} \right\rangle = \frac{1}{3}, \quad \left\langle \mathcal{T}_{1-2\beta}^{\hat{c}(\mathcal{A})} \right\rangle = \left\langle \mathcal{T}_{1-2\beta}^{\hat{a}(\mathcal{A})} \right\rangle = 0, \quad \left\langle \mathcal{T}_{-\bar{\beta}/2}^{\hat{c}(\mathcal{A})} \right\rangle = \left\langle \mathcal{T}_{-\bar{\beta}/2}^{\hat{a}(\mathcal{A})} \right\rangle = \frac{1}{6},$$

and

$$\left\langle \mathcal{T}_{-\bar{\beta}/2,1}^{\hat{c}(\mathcal{A})} \right\rangle = \left\langle \mathcal{T}_{-\bar{\beta}/2,1}^{\hat{a}(\mathcal{A})} \right\rangle = -\frac{1}{12}\beta\bar{\beta},$$

for the case of $\sigma = 0$. According to Eq. (6.33), the DP for $\sigma = 0$ is thus:

$$\begin{aligned} \mathbf{l}_D(\ell) &= -6 \Re \left\{ \sum_{\lambda \in \Lambda_{\mathcal{T}}} \sum_{m=0}^{\nu_{\lambda}-1} \frac{\left\langle \mathcal{T}_{\lambda,m}^{\hat{c}(\mathcal{A})} \right\rangle}{(z-\lambda)^{m+1}} \right\} \\ &= -6 \Re \left\{ \frac{\left\langle \mathcal{T}_1^{\hat{c}(\mathcal{A})} \right\rangle}{z-1} + \frac{\left\langle \mathcal{T}_{-\bar{\beta}/2}^{\hat{c}(\mathcal{A})} \right\rangle}{z+\bar{\beta}/2} + \frac{\left\langle \mathcal{T}_{-\bar{\beta}/2,1}^{\hat{c}(\mathcal{A})} \right\rangle}{(z+\bar{\beta}/2)^2} \right\} \\ &= \Re \left\{ -\frac{2}{z-1} - \frac{1}{z+\bar{\beta}/2} + \frac{\beta\bar{\beta}/2}{(z+\bar{\beta}/2)^2} \right\} \\ &= 1 - \Re \left\{ \frac{z+\bar{\beta}^2/2}{(z+\bar{\beta}/2)^2} \right\}. \end{aligned} \tag{6.40}$$

6.6.2.2 $\sigma \neq 0$:

Riechers *et al.* [205] also found that:

$$\left\langle \mathcal{T}_1^{\hat{c}(\mathcal{A})} \right\rangle = \left\langle \mathcal{T}_1^{\hat{a}(\mathcal{A})} \right\rangle = \frac{1}{3}, \quad \left\langle \mathcal{T}_{1-2\beta}^{\hat{c}(\mathcal{A})} \right\rangle = \left\langle \mathcal{T}_{1-2\beta}^{\hat{a}(\mathcal{A})} \right\rangle = 0,$$

$$\left\langle \mathcal{T}_{\frac{-\bar{\beta}+\sqrt{\sigma}}{2}}^{\hat{c}(\mathcal{A})} \right\rangle = \left\langle \mathcal{T}_{\frac{-\bar{\beta}+\sqrt{\sigma}}{2}}^{\hat{a}(\mathcal{A})} \right\rangle = -\frac{1}{12} \left(1 - \frac{\beta}{\sqrt{\sigma}} \right) (\sqrt{\sigma} - \bar{\beta}),$$

and

$$\left\langle \mathcal{T}_{\frac{-\bar{\beta}-\sqrt{\sigma}}{2}}^{\hat{c}(\mathcal{A})} \right\rangle = \left\langle \mathcal{T}_{\frac{-\bar{\beta}-\sqrt{\sigma}}{2}}^{\hat{a}(\mathcal{A})} \right\rangle = \frac{1}{12} \left(1 + \frac{\beta}{\sqrt{\sigma}} \right) (\sqrt{\sigma} + \bar{\beta}),$$

for $\sigma \neq 0$. According to Eq. (6.33), the DP for $\sigma \neq 0$ is:

$$\begin{aligned}
I_D(\ell) &= -6 \Re \left\{ \sum_{\lambda \in \Lambda_T} \frac{\langle \mathcal{T}_\lambda^{\hat{c}(\mathcal{A})} \rangle}{z - \lambda} \right\} \\
&= -6 \Re \left\{ \frac{\langle \mathcal{T}_1^{\hat{c}(\mathcal{A})} \rangle}{z - 1} + \frac{\langle \mathcal{T}_{\frac{-\bar{\beta} + \sqrt{\sigma}}{2}}^{\hat{c}(\mathcal{A})} \rangle}{z - \frac{\sqrt{\sigma} - \bar{\beta}}{2}} + \frac{\langle \mathcal{T}_{\frac{-\bar{\beta} - \sqrt{\sigma}}{2}}^{\hat{c}(\mathcal{A})} \rangle}{z + \frac{\sqrt{\sigma} + \bar{\beta}}{2}} \right\} \\
&= 1 + \frac{1}{2} \Re \left\{ \frac{1 - \beta/\sqrt{\sigma}}{\frac{z}{\sqrt{\sigma} - \beta} - \frac{1}{2}} - \frac{1 + \beta/\sqrt{\sigma}}{\frac{z}{\sqrt{\sigma} + \beta} + \frac{1}{2}} \right\}. \tag{6.41}
\end{aligned}$$

Figure 6.4 gives several coronal spectrograms for various values of the parameters α and β . It is instructive to examine the influence of the TM's eigenvalues on the placement and intensity of the Bragg-like reflections. In panel (a) there are two strong reflections, one each at $\omega = 2\pi/3$ and $4\pi/3$, signaling a twinned 3C structure, when the faulting parameters are set to $\alpha = \beta = 0.01$. Each is accompanied by an eigenvalue close to the surface of the unit circle. As the disorder is increased, see panels (b) and (c), TM eigenvalues retreat toward the center of the unit circle and the two strong reflections become diffuse. However, in the final panel (d), the faulting parameters ($\alpha = 0.01$, $\beta = 0.9$) are set such that the material has apparently undergone a phase transition from prominently 3C stacking structure to prominently 2H stacking structure. Indeed, the eigenvalues have coalesced through the critical point of $\sigma = 0$ (as σ changes from negative to positive) and emerge on the other side of the phase transition mutually scattered along the real axis and approaching the edge of the unit circle, giving rise to the 2H-like protrusions in the DP. This demonstrates again how the eigenvalues orchestrate the placement and intensity of the Bragg-like peaks.

6.6.3 Shockley–Frank Stacking Faults in 6H-SiC: The SFSF Process

SiC has been the intense focus of both experimental and theoretical investigations for some time due to its promise as a material suitable for next-generation electronic devices. However, it is known that SiC can have many different stacking configurations—some ordered and some disordered[218]—and these different stacking configurations can profoundly affect material properties. Despite considerable effort to grow commercial SiC wafers that

Table 6.3: Limiting material structures for the SFSF Process. Key: SF - Shockley–Frank fault; NGF - nonrandom growth fault.

$\gamma = 0$	$\gamma \approx 0$	$\bar{\gamma} \approx 0$	$\bar{\gamma} = 0$
6H	6H/SF	3C/NGF	3C

are purely crystalline—*i.e.*, that have no stacking defects—reliable techniques have not yet been developed. It is therefore important to better understand and characterize the nature of the defects in order to better control them.

Recently, Sun *et al.* [246] reported experiments on 6H-SiC that used a combination of low temperature photoluminescence and high resolution transmission electron microscopy (HRTEM). One of the more common crystalline forms of SiC, the 6H stacking structure is simply the sequence $\dots \text{ABCACBA} \dots$, or in terms of the Hägg-notation, $\dots 111000 \dots$. The most common stacking fault in 6H-SiC identified by HRTEM can be explained as the result of one extrinsic Frank stacking fault coupled with one Shockley stacking fault [110]. Physically, the resultant stacking structure corresponds to the insertion of an additional SiC ML so that one has instead $\dots 111000\underline{0}111000 \dots$, where the underlined spin is the inserted ML.

Inspired by these findings, we suggest a simple HMM for the *Shockley–Frank stacking fault* (SFSF) process that replicates this structure, and this is shown in Fig. 6.5. Our motivation here is largely pedagogical, and certainly more detailed experiments are required to confidently propose a structure, but this HMM reproduces at least qualitatively the observed structure. The model has a single parameter $\gamma \in [0, 1]$. As before, it is instructive to consider limiting cases of γ . For $\gamma = 0$, we have the pure 6H structure and, for small γ , Shockley–Frank defects are introduced into this stacking structure. As $\gamma \rightarrow 1$, the structure transitions into a twinned 3C crystal. However, unlike the previous example, this twinning is not random. Instead, the architecture of the machine requires that at least three 0s or 1s must be seen before there is a possibility of reversing the chirality, *i.e.*, before there is twinning. These limiting cases are summarized in Table 6.3.

For $\gamma \in (0, 1)$ the Hägg-machine is mixing and we proceed with this case. By inspection,

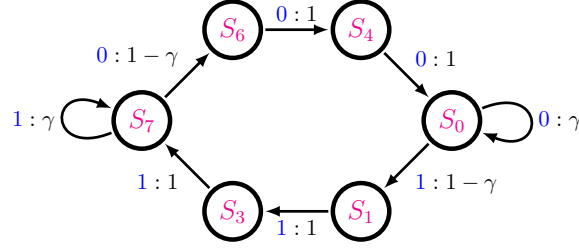


Figure 6.5: The Hägg-machine for the SFSF Process expressed as a 3rd-order MM. There is one faulting parameter $\gamma \in [0, 1]$ and three SSCs or, equivalently, three CSCs, as this machine is also an ϵ -machine. The three SSCs are $[S_7]$, $[S_0]$ and $[S_7S_6S_4S_0S_1S_3]$. The latter we recognize as the 6H structure if $\gamma = 0$. For large values of γ —*i.e.*, as $\gamma \rightarrow 1$ —this process approaches a twinned 3C structure, although the faulting is *not* random. The causal-state architecture prevents the occurrence of domains of size-three or less. (From Riechers *et al.* [205] Used with permission.)

we write down the two 6-by-6 TMs of the Hägg-machine as:

$$T^{[0]} = \begin{bmatrix} \gamma & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \bar{\gamma} & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad T^{[1]} = \begin{bmatrix} 0 & \bar{\gamma} & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & \gamma & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

where the states are ordered S_0, S_1, S_3, S_7, S_6 , and S_4 . The internal state TM is their sum:

$$T = \begin{bmatrix} \gamma & \bar{\gamma} & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & \gamma & \bar{\gamma} & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Since the six-state Hägg-machine generates an $(3 \times 6 =)$ eighteen-state *ABC*-machine, we do not explicitly write out its TMs. Nevertheless, it is straightforward to expand the Hägg-machine to the *ABC*-machine via the rote expansion method [205]. It is also straightforward to apply Eq. (6.17) to obtain the DP as a function of the faulting parameter

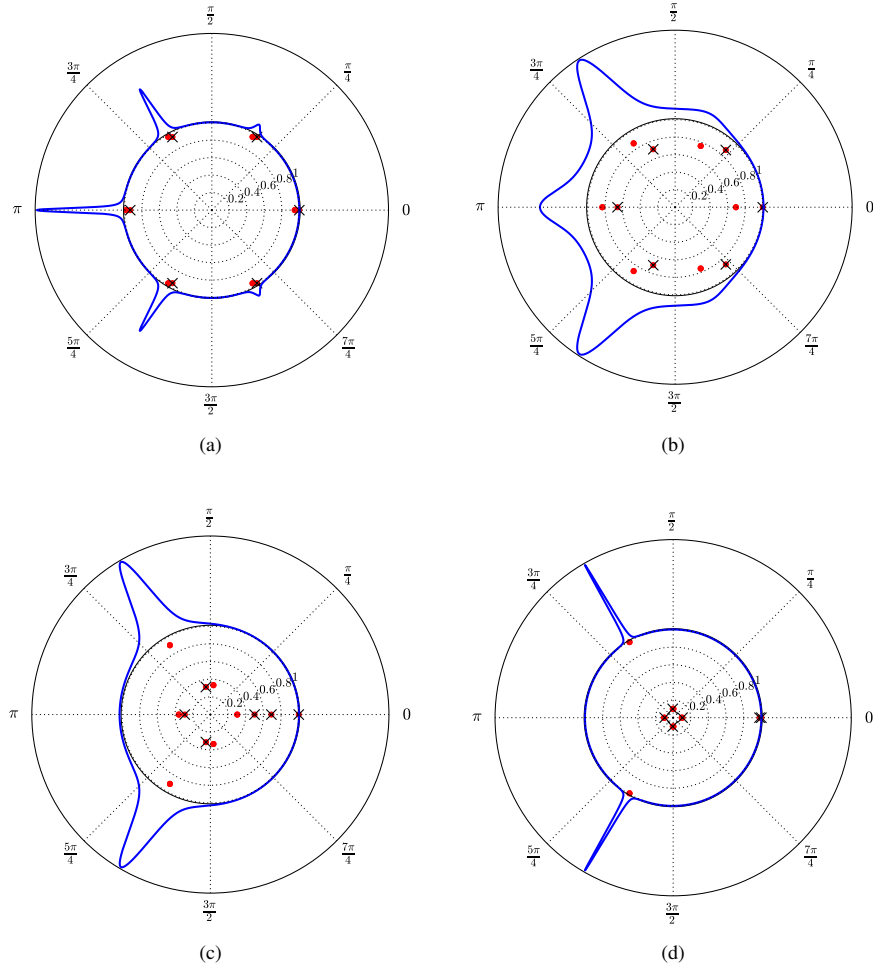


Figure 6.6: Coronal spectrograms showing the evolution of the DP and its eigenvalues for the SFSF Process. (a) $\gamma = 0.1$, (b) $\gamma = 0.5$, (c) $\gamma = 0.9$, and (d) $\gamma = 0.99$. In (a) the faulting is weak and the DP has the six degraded Bragg-like reflections characteristic of the 6H stacking structure. In (b), the faulting is more severe, with the concomitant erosion of the Bragg-like reflections, especially for $\omega = \pi$. In panel (c) the 6H character has been eliminated, and the Bragg-like peaks at $\omega = 2\pi/3$ and $4\pi/3$ are now associated with a twinned 3C stacking structure. In panel (d), the Bragg-like reflections sharpen as the probability of short 3C sequences stacking sequences decreases.

γ . To use Eq. (6.17), note that the stationary distribution over the *ABC*-machine can be obtained from Eq. (6.42) with:

$$\langle \pi_H | = \frac{1}{6-4\gamma} \begin{bmatrix} 1 & \bar{\gamma} & \bar{\gamma} & 1 & \bar{\gamma} & \bar{\gamma} \end{bmatrix}$$

as the stationary distribution over the Hägg-machine.

The eigenvalues of the Hägg TM can be obtained as the solutions of $\det(\mathbf{T} - \lambda \mathbb{I}) = (\lambda - \gamma)^2 \lambda^4 - \bar{\gamma}^2 = 0$. These include 1 , $-\frac{1}{2}\bar{\gamma} \pm \sqrt{\gamma^2 + 2\gamma - 3}$, and three other eigenvalues involving cube roots.

The eigenvalues of the *ABC* TM are obtained similarly as the solutions of $\det(\mathcal{T} - \lambda \mathbb{I}) = 0$. Note that $\Lambda_{\mathcal{T}}$ inherits $\Lambda_{\mathbf{T}}$ as the backbone for its more complex structure, just as $\Lambda_{\mathbf{T}} \subseteq \Lambda_{\mathcal{T}}$ for all of our previous examples. The eigenvalues in $\Lambda_{\mathcal{T}}$ are, of course, those most directly responsible for the structure of the CFs. Since the *ABC*-machine has eighteen states, there are eighteen eigenvalues contributing to the behavior of the DP; although several eigenvalues are degenerate. Hence, the SFSF Process is capable of a richer DP than the previous two examples.

The coronal spectrograms for the SFSF Process are shown for several example values of γ in Fig. 6.6. Over the range of γ values the stacking structure changes from a nearly perfect 6H crystal through a disordered phase finally becoming a twinned 3C structure. Most notable in Fig. 6.6 is how the eigenvalues of the total TM dictate the placement of the Bragg-like reflections. Phrased alternatively, the Bragg-like reflections appear to literally track the movement of the eigenvalues as they evolve during transformation.

6.7 Conclusions

We showed how the DP of layered CPSs, as described by an arbitrary HMM stacking process, is calculated either analytically or to a high degree of numerical certainty directly without restriction to finite-Markov-order and without needing finite samples of the stacking sequence. Our expressions for the DP are similar to those previously obtained by Treacy *et al.*[252], but since our starting point is an arbitrary HMM, ours are guaranteed to be valid for both finite- and infinite-order Markov models. Along the way, we uncovered a remarkably simple relationship between the DP and the HMM. The former is given by

straightforward, standard matrix manipulations of the latter. Critically, in the case of an infinite number of MLs, this relationship does not involve powers of the TM.

The connection yields important insights. (i) The number of Bragg and Bragg-like reflections in the DP is limited by the size of the TMs that define the HMM. Thus, knowing only the number of machine states reveals the maximum possible number of Bragg and Bragg-like reflections. (ii) As a corollary, given a DP, the number of Bragg and Bragg-like reflections puts a minimum on the number of HMM states. For the problem of inferring the HMM from experimental DPs, this gives powerful clues about the HMM (and so internal mechanism) architecture. (iii) The eigenvalues *within* the unit circle organize the diffuse Bragg-like reflections. Only TM eigenvalues *on* the unit circle correspond to those ℓ -values that potentially can result in true Bragg peaks. (iv) The expansion of the Hagg-machine into the *ABC*-machine, necessary for the appropriate matrix manipulations, showed that there are two kinds of machines and, hence, two kinds of stacking process important in CPSs: mixing and nonmixing processes. In addition to the calculational shortcuts given by the former, mixing machines ensure that there are no true Bragg reflections at integer- ℓ . (v) Conversely, the presence of Bragg peaks at integer- ℓ is an unmistakable sign that a stacking process is nonmixing. Again, this puts important constraints on the HMM state architecture, useful for the problem of inverting the DP to find the HMM. (vi) For mixing processes, the ML probabilities must all be one-third, *i.e.*, $\Pr(A) = \Pr(B) = \Pr(C) = 1/3$.

New in the theory is the introduction of coronal spectrograms, a convenient way to visualize the interplay between a frequency-domain functional of a process and the eigenvalues of the process's TMs. In our case, the frequency-domain functional was the DP: the power spectrum of the sequence of ML structure factors. In each of the examples, the movement of the eigenvalues (as the HMM parameters change) were echoed by movement of their 'shadow'—the Bragg-like peaks in the power spectrum. While this technique was explored in the context of DPs from layered materials, this visualization tool is by no means confined to DPs or layered materials. We suspect that in other areas where power spectra and HMMs are studied, this technique will become a useful analysis tool.

There are several important research directions to follow in further refining and

extending the theory and developing applications. (i) While specialized to CPSs here, the basic techniques extend to other stacking geometries and other materials, including the gamut of technologically cutting-edge heterostructures of stacked 2D materials. (ii) With the ability to analytically calculate DPs and CFs [205] from arbitrary HMMs, the number of physical and information- and computation-theoretic quantities amenable to such a treatment continues to expand. Statistical complexity, the Shannon entropy rate, and memory length have long been calculable from the ϵ -machine [64, 265, 266, 270], but recently the excess entropy, transient information, and synchronization time have also been shown to be exactly calculable from the ϵ -machine [60, 205]. This portends well that additional quantities, especially those of physical import such as band structure in chaotic crystals, may also be treatable with exact methods. (iii) Improved calculational techniques raise the possibility of improved inference methods, so that more kinds of stacking process may be discovered from DPs. An important research direction then is to incorporate these improved methods into more flexible, more sensitive inference algorithms.

Finally, the spectral methods pursued here increase the tools available to chaotic crystallography for the discovery, description, and categorization of both ordered and disordered (chaotic) crystals. With these tools in hand, we will more readily identify key features of the hidden structures responsible for novel physical properties of materials.

6.8 Appendix A: Hägg-to- ABC Machine Translation

If M_H is the number of states in the Hägg-machine and M is the number of states in the ABC -machine, then $M = 3M_H$ for mixing Hägg-machines. Let the i^{th} state of the Hägg-machine split into the $(3i - 2)^{\text{th}}$ through the $(3i)^{\text{th}}$ states of the corresponding ABC -machine. Then, each labeled-edge transition from the i^{th} to the j^{th} states of the Hägg-machine maps into a 3-by-3 submatrix for each of the three labeled TMs of the ABC -machine as:

$$\left\{ T_{ij}^{[0]} \right\} \xrightarrow{\text{Hägg to } ABC} \left\{ \mathcal{T}_{3i-1, 3j-2}^{[A]}, \mathcal{T}_{3i, 3j-1}^{[B]}, \mathcal{T}_{3i-2, 3j}^{[C]} \right\}$$

and

$$\left\{ \mathsf{T}_{ij}^{[1]} \right\} \xrightarrow{\text{Hägg to } ABC} \left\{ \mathcal{T}_{3i,3j-2}^{[A]}, \mathcal{T}_{3i-2,3j-1}^{[B]}, \mathcal{T}_{3i-1,3j}^{[C]} \right\}.$$

For nonmixing Hägg-machines, the above algorithm creates three disconnected ABC -machines, of which only one need be retained.

Furthermore, for mixing Hägg-machines, the probability from the stationary distribution over their states maps to a triplet of probabilities for the stationary distribution over the ABC -machine states:

$$\{p_i^{\text{H}}\} \xrightarrow{\text{Hägg to } ABC} \{3p_{3i-2}, 3p_{3i-1}, 3p_{3i}\} . \quad (6.42)$$

A thorough exposition of these procedures is given by Riechers *et al.* [205].

Chapter 7

Fluctuations When Driving Between Nonequilibrium Steady States

This chapter has appeared online as a preprint:

P. M. Riechers & J. P. Crutchfield, (2016). “Fluctuations When Driving Between Nonequilibrium Steady States”. arXiv preprint arXiv:1610.09444.

7.1 Chapter Overview

Maintained by environmental fluxes, biological systems are thermodynamic processes that operate far from equilibrium without detailed-balance dynamics. Yet, they often exhibit well defined nonequilibrium steady states (NESSs). More importantly, critical thermodynamic functionality arises directly from transitions among their NESSs, driven by environmental switching. Here, we identify constraints on excess thermodynamic quantities that ride above the NESS housekeeping background. We do this by extending the Crooks fluctuation theorem to transitions among NESSs, without invoking an unphysical dual dynamics. This and corresponding integral fluctuation theorems determine how much work must be expended when controlling systems maintained far from equilibrium. This generalizes feedback control theory, showing that Maxwellian Demons can leverage mesoscopic-state information to take advantage of the excess energetics in NESS transitions. Altogether, these point to universal thermodynamic laws that are immediately applicable to the accessible degrees of freedom within the effective dynamic at any emergent level

of hierarchical organization. By way of illustration, this readily allows analyzing a voltage-gated sodium ion channel whose molecular conformational dynamics play a critical functional role in propagating action potentials in mammalian neuronal membranes.

7.2 Introduction

The sun shines; ATP is abundant; power is supplied. These are the generous settings in which we find many complex biological systems, buoyed steadily out of equilibrium by energy fluxes in their environment. The resulting steady-state dynamics exhibit various types of directionality, including periodic oscillations and macroscopic thermodynamic functionality. These behaviors contrast rather sharply with the deathly isotropy of equilibrium detailed-balance dynamics—where fluxes are absent and state transition rates depend only on relative asymptotic state-occupation probabilities.

Detailed balance and its implied dynamical reversibility, though, are common tenets of equilibrium thermodynamics and statistical mechanics. They are technically necessary when applying much of the associated theory relevant to equilibrium and reversible (e.g., quasistatic) transitions between equilibrium macrostates [53]. Detailed balance is even assumed by several modern theorems that influence our understanding of the structure of fluctuations and limitations on work performed far from equilibrium [49, 209]. The natural world, though, is replete with systems that violate detailed balance, such as small biological molecules constantly driven out of equilibrium through interactions with their biochemical environment [272, 193].

Far from happenstance and disruption, the probability currents through the effective state-space of these nonequilibrium systems enable crucial thermodynamic functionality [139, 195]. Even rare fluctuations play an important functional role [114, 148, 57]. While constant environmental pressure can drive a system into a nonequilibrium steady state (NESS), complex biological systems are often driven farther—far from even any NESS. Moreover, such system–environment dynamics involve feedback between system and environment states. Although many believe these facilitate the necessary complex processes that sustain life, their very nature seems to preclude most, if not all, hope of a

universal theoretical framework for quantitative predictions. To ameliorate the roadblock, we present a consistent thermodynamics that is not only descriptive, but constructive, tractable, and predictive, even when irreversible dynamics transition between NESSs.

Beyond laying out the structure of fluctuations during NESS transitions, this thermodynamics sets the stage to understand how one level of organization gives way to another. In particular, using it a sequel renormalizes nonequilibrium housekeeping background to show how to maintain a hierarchy of steady-state dynamics. Said simply, at each level of hierarchical organization, controllable degrees of freedom are subject to universal thermodynamic laws that tie their fluctuations and functionality to dissipation at lower levels.

7.2.1 Results

Nonequilibrium thermodynamics progressed markedly over the last two decades on at least two fronts. First, by taking the ‘dynamics’ in ‘thermodynamics’ seriously, fluctuation theorems (FTs) transformed previous inequalities, such as the classical Second Law of Thermodynamics, into subsuming equalities that exactly express the distribution of thermodynamic variations. (These have been derived by many authors now in a wide range of physical settings; see, e.g., Refs. [222] and [238] for lucid reviews.) Second, steady-state thermodynamics (SST) showed that NESSs play a role in nonequilibrium analogous to that of equilibrium macrostates in equilibrium. In this view heat decomposes into the *housekeeping heat* Q_{hk} needed to sustain NESSs and the *excess heat* Q_{ex} dissipated in transitions between them [181, 108, 256]. Bolstering SST, recent efforts generalized the Clausius inequality (describing excess heat produced beyond the change in system entropy) to smoothly driven transitions between NESSs [163]. Taken together, these results established an integral fluctuation theorem for the excess work in NESS transitions and, consequently, a generalized Second Law for excess entropy produced beyond housekeeping during driven NESS transitions.

The following extends SST by introducing several new FTs, highlighting correspondences between nonequilibrium and equilibrium relations. First, we provide detailed (i.e., nonintegrated) fluctuation theorems, rather than integral fluctuation theorems for driven

NESS transitions. (Integral FTs follow directly, in any case.) This constrains distributions of excess work W_{ex} exerted when controlling nonequilibrium systems. Second, we jointly bound housekeeping and excess work distributions. For example, for time-symmetric driving we show that the joint probability of excess work and housekeeping heat respect the strong constraint:

$$\frac{\Pr(W_{\text{ex}}, Q_{\text{hk}})}{\Pr(-W_{\text{ex}}, -Q_{\text{hk}})} = e^{\beta W_{\text{ex}}} e^{\beta Q_{\text{hk}}} .$$

When the transitions are nonequilibrium excursions between *equilibrium* steady states, this reduces to the Crooks FT [50]. Third, we derive the detailed FTs for entropy production even when temperature varies in space and time. They are expressed in terms of excess environmental entropy production Ω and irreversibility Ψ , even when the irreversibility is “housekeeping” not strictly associated with heat. Finally, we quantify a system’s net *path irreversibility* Ψ with the accumulated violation of detailed balance in the effective dynamic. In the isothermal setting, for example, the irreversibility is the housekeeping heat, maintaining the system in its nonequilibrium dynamic: $\Psi = \beta Q_{\text{hk}}$. Importantly, we can determine the minimum housekeeping heat without appealing to the system’s Hamiltonian.

Extending SST in this way reveals universal constraints on excess thermodynamic quantities—effective energies accessible above the housekeeping background. Looking forward, this allows one to analyze nondetailed-balanced stochastic dynamics—and thus contributes an understanding of the role of hierarchy—in the thermodynamics of replication [77] and the thermodynamics of learning [54]. Moreover, this identifies how complex, possibly intelligent, thermodynamic systems leverage (designed or intrinsic) irreversibility in their own state-space to harness energy from structured environments.

7.2.2 Synopsis

Section 7.3 sets up our approach, introducing notation, discussing input-dependent system dynamics, and establishing fundamental relationships among nonequilibrium thermodynamic quantities. Section 7.4.1 introduces excess heat and excess work in analogy to classical heat and work. Ultimately though, the related excess environmental entropy

production Ω discussed in § 7.4.2 generalizes these to the case of temperature inhomogeneity over spacetime. Section 7.4.3 demonstrates that path entropies are the fundamental objects of nonequilibrium thermodynamics. In steady state, unaveraged path entropies reduce to the steady-state surprisal ϕ . Deviations from the asymptotic surprisal contribute to a nonsteady-state additional free energy. All of these quantities play a central role in the subsequent development.

Before delving into irreversibility, though, we first address what is meant by reversibility. Therefore, § 7.5.1 and § 7.5.2 discuss detailed balance, microscopic reversibility, and the close relationship between them. Section 7.5.3 then introduces path dependence and reverse-path dependence and explains how together they yield a system’s irreversibility Ψ .

With this laid out, § 7.6.1 and § 7.6.3 derive the detailed FTs in terms of excess environmental entropy production Ω and irreversibility Ψ . One sees that in the isothermal setting $\Psi = \beta Q_{\text{hk}}$ and the excess entropy production is directly related to the excess work. This allows § 7.6.5 to explain how these results extend SST.

Sections 7.6.6 and 7.6.7 finish our investigation of NESS FTs by deriving several integral FTs. This, in effect, extends feedback control, as developed in Refs. [209] and [113], to SST. We note that such environmental feedback is intrinsic to natural systems.

For concreteness, § 7.7 analyzes a simple but biologically important prototype system: voltage-gated sodium ion channels. These are complex macromolecules that violate detailed balance in order to perform critical biological functioning far from equilibrium. Finally, appendices discuss non-Markovian dynamics and comment on the bounds provided by integral fluctuation theorems for auxiliary variables.

7.3 Driven Stochastic Dynamics

We consider a classical system—the *system under study*—with time-dependent driving via environmentally determined parameters; e.g., time-dependent temperature, voltage, and piston position. Hence, the environmental control input X_t at time t , taking on values $x_t \in \mathcal{X}$, will typically be a vector object. The system under study is assumed to have a countable set \mathcal{S} of states. The random variable \mathcal{S}_t for the state at time t takes on

values $s_t \in \mathcal{S}$. We assume that the environment’s control value (current *input*) x and the system’s physical state (current *state*) s are sufficient to determine the system’s net effective energy—the *nonequilibrium potential* $\phi(x, s)$. Even with constant environmental input, the system dynamic need not be detailed balance.

7.3.1 Stochastic mesoscopic dynamics and induced state-distributions

We assume the current environmental input x determines the instantaneous stochastic transition dynamic over the system’s observable mesoscopic states. However, that input can itself depend arbitrarily on all previous input and state history. That is, we assume that the \mathcal{S} -to- \mathcal{S} transitions are instantaneously Markovian given the input. Over time, though, different inputs induce different Markov chains over system states.

Note that the Markov assumption is common, although often implicit, and we follow this here to isolate the novel implications of nondetailed-balance dynamics. Nevertheless, the results generalize to infinite Markov order by modeling system states as the observable output of many-to-one mappings of latent states of an input-controllable hidden Markov chain. Appendix 7.10.1 details this generalization.

We do not restrict the environment’s driving process, allowing arbitrary non-Markovity, feedback, and nonstationarity. Thus, the joint system-environment dynamic can be non-Markovian even if the instantaneous system dynamic is. Such a setup is quite general, and so the results to follow extend others known for SST. We also follow stochastic thermodynamics in the use of (arbitrarily small) discrete-time steps. Nevertheless, it is usually easy to take the continuous-time limit. As, in fact, we do in the example at the end.

Hence, the Markovian dynamic is described by a (possibly infinite) set of input-conditioned transition matrices over the state set \mathcal{S} : $\{\mathbb{T}^{(\mathcal{S} \rightarrow \mathcal{S}|x)}\}_{x \in \mathcal{X}}$, where $\mathbb{T}_{i,j}^{(\mathcal{S} \rightarrow \mathcal{S}|x)} = \Pr(\mathcal{S}_t = s^j | \mathcal{S}_{t-1} = s^i, X_t = x)$ is the probability that the system is in state s^j at time t given that the system was in state s^i at time $t - 1$ and the instantaneous environmental input controlling the system was x .

The Perron–Frobenius theorem guarantees that there is a stationary distribution $\boldsymbol{\pi}_x$

over states associated with each fixed input x . These are the state distributions associated with the system's nonequilibrium steady states (NESSs). For simplicity, and unless otherwise stated, we assume that a fixed input x eventually induces a unique NESS.

We denote distributions over the system states as bold Greek symbols; such as $\boldsymbol{\mu}$. We denote the state random variable \mathcal{S} being distributed according to $\boldsymbol{\mu}$ via $\mathcal{S} \sim \boldsymbol{\mu}$. It will often be convenient to cast $\boldsymbol{\mu}$ as a row-vector, in which case it appears as the bra $\langle \boldsymbol{\mu} |$. Putting this altogether, a sequence of driving inputs updates the state distribution as follows:

$$\begin{aligned} \langle \boldsymbol{\mu}_{t+n} | &= \langle \boldsymbol{\mu}_t | \mathsf{T}^{(\mathcal{S} \rightarrow \mathcal{S}|_{x:t:t+n})} \\ &= \langle \boldsymbol{\mu}_t | \mathsf{T}^{(\mathcal{S} \rightarrow \mathcal{S}|_x)} \mathsf{T}^{(\mathcal{S} \rightarrow \mathcal{S}|_{x_{t+1}})} \dots \mathsf{T}^{(\mathcal{S} \rightarrow \mathcal{S}|_{x_{t+n-1}})} . \end{aligned}$$

(Time indexing here and throughout is denoted by subscripts $t:t'$ that are left-inclusive and right-exclusive.) An infinite driving history $\overleftarrow{x} = \dots x_{-2}x_{-1}$ induces a distribution $\boldsymbol{\mu}(\overleftarrow{x})$ over the state space. The so-called *steady-state distribution* associated with the environmental drive value x , induced by tireless repetition of x , is:

$$\langle \boldsymbol{\pi}_x | = \lim_{n \rightarrow \infty} \langle \boldsymbol{\mu}_0 | \left(\mathsf{T}^{(\mathcal{S} \rightarrow \mathcal{S}|_x)} \right)^n .$$

Usefully, $\boldsymbol{\pi}_x$ can also be found as the left eigenvector of $\mathsf{T}^{(\mathcal{S} \rightarrow \mathcal{S}|_x)}$ associated with the eigenvalue of unity ¹:

$$\langle \boldsymbol{\pi}_x | = \langle \boldsymbol{\pi}_x | \mathsf{T}^{(\mathcal{S} \rightarrow \mathcal{S}|_x)} . \quad (7.1)$$

The assumption that observable state-to-state transitions are instantaneously Markovian allows the state distribution $\boldsymbol{\mu}$ to summarize the causal relevance of the entire driving history \overleftarrow{x} .

7.4 Energetics and Entropies

For later comparison, we recount the basics of a statistical mechanics description of the thermodynamics of a system exchanging energy with a large environment, imposing fixed

¹We ignore nonergodicity to simplify the development. The approach, though, handles nonergodicity just as well. However, distracting nuances arise that we do not wish to dwell on. For example, if the Markov chain has more than one attracting component for a particular x , then $\boldsymbol{\pi}_x$ is not unique, but can be constructed as any one of infinitely many probability-normalized linear superpositions of left eigenvectors of $\mathsf{T}^{(\mathcal{S} \rightarrow \mathcal{S}|_x)}$ associated with the eigenvalue of unity.

constraints indexed as x . The many-body Hamiltonian $\mathcal{H}(x)$ has energy eigenvalues $\{E(x, s)\}$, where s indexes the energy eigenstates. The canonical distribution is $\pi_x(s) = e^{-\beta[E(x, s) - F_{\text{eq}}(x)]}$, at fixed x . This distribution is the equilibrium steady “state” associated with x , where $\beta^{-1} \equiv k_{\text{B}}T$, T is the temperature of the macroscopic environment surrounding the system, and $F_{\text{eq}}(x)$ is the associated equilibrium free energy.

7.4.1 Work, heat, and their excesses

Work W is environmentally driven energy change. Within one time-step it is given by ²:

$$W[x_{n-1} \rightarrow x_n; s_{n-1}] = E(x_n, s_{n-1}) - E(x_{n-1}, s_{n-1}) .$$

Heat Q is the change in system energy due to its internal response to the environmental drive; e.g., a molecule’s change in conformation. Within one time-step the heat is:

$$Q[x_n; s_{n-1} \rightarrow s_n] = E(x_n, s_n) - E(x_n, s_{n-1}) .$$

Over the course of driving the system from $t = 0$ to $t = N\Delta t = \tau$, the net energy change is then:

$$\begin{aligned} \Delta E &= E(x_N, s_N) - E(x_0, s_0) \\ &= W + Q , \end{aligned}$$

where the net work and net heat are:

$$W = \sum_{n=1}^N W[x_{n-1} \rightarrow x_n; s_{n-1}]$$

and:

$$Q = \sum_{n=1}^N Q[x_n; s_{n-1} \rightarrow s_n] ,$$

respectively. Here, and later on, Δ applied to a quantity refers to its change over one time step, where the step is given by context.

When the system strongly couples to a substrate with uncontrolled energy fluxes, steady-state dynamics are often established far from equilibrium, even when environmental

²We start in a discrete-time setup, but later translate to continuous time

parameters are held fixed. That is, for fixed driving $\dots xxxxx \dots$, the system settles down to a NESS with a distribution over observable system states given by the *nonequilibrium potential* $\phi(x, s)$:

$$\pi_x(s) = e^{-\phi(x, s)} . \quad (7.2)$$

In this, $\phi(x, s)$ plays a role roughly analogous to energy eigenvalues. Thus, the thermodynamics of accessible energetics—the excess heat generated and work irretrievably performed in driving between NESSs—follows analogously to its equilibrium counterpart. This is complementary to recent SST studies [108, 256, 80, 163].

If steady-state free energies $F_{\text{ss}}(x)$ and *effective* energies $E_{\text{eff}}(x, s)$ could be uniquely (and usefully) defined, then the nonequilibrium potential would be:

$$\phi(x, s) = \beta[E_{\text{eff}}(x, s) - F_{\text{ss}}(x)] .$$

However, the assignment of steady-state free energies is problematic. Nevertheless, $\phi(x, s)$ retains meaning since it quantifies the steady-state *surprisal* of observing state s :

$$\phi(x, s) = -\ln \pi_x(s) .$$

The surprisal is Shannon’s *self-information* [46]—the unaveraged individual-event entropy measuring how surprising a specific event is. Intuitively, we must do work to make otherwise unlikely things happen.

SST’s *excess* work and heat can be defined via changes in steady-state surprisal ϕ , analogous to how equilibrium quantities are in terms of energy changes. For clarity, we temporarily restrict ourselves to the isothermal setting, but we can easily adapt to time-varying temperatures.

Excess work W_{ex} is environmentally driven change in nonequilibrium potential:

$$W_{\text{ex}}[x_{n-1} \rightarrow x_n; s_{n-1}] = \beta^{-1}[\phi(x_n, s_{n-1}) - \phi(x_{n-1}, s_{n-1})] ,$$

over one time-step. *Excess heat* Q_{ex} is the change in nonequilibrium potential due to the system’s response:

$$Q_{\text{ex}}[x_n; s_{n-1} \rightarrow s_n] = \beta^{-1}[\phi(x_n, s_n) - \phi(x_n, s_{n-1})] ,$$

over one time-step. When driving from $t = 0$ to $t = N\Delta t = \tau$, the net change in nonequilibrium potential is:

$$\begin{aligned}\Delta\phi &= \phi(x_N, s_N) - \phi(x_0, s_0) \\ &= \beta(W_{\text{ex}} + Q_{\text{ex}}) \\ &= -\ln \frac{\pi_{x_N}(s_N)}{\pi_{x_0}(s_0)} ,\end{aligned}\tag{7.3}$$

where the net excess work and net excess heat are:

$$W_{\text{ex}} = \sum_{n=1}^N W_{\text{ex}}[x_{n-1} \rightarrow x_n; s_{n-1}] \tag{7.4}$$

and:

$$Q_{\text{ex}} = \sum_{n=1}^N Q_{\text{ex}}[x_n; s_{n-1} \rightarrow s_n] , \tag{7.5}$$

respectively.

This approach to excess heat Q_{ex} coincides with SST's definition and reduces to total heat in equilibrium transitions. Importantly, it follows as closely as possible the equilibrium approach to total heat Q outlined above and deviates from the typical starting point: $Q_{\text{ex}} \equiv Q - Q_{\text{hk}}$, where Q_{hk} is the so-called *housekeeping heat*. In contrast, excess work W_{ex} does *not* reduce to the total work in equilibrium transitions. Rather, W_{ex} goes over to $W - \Delta F_{\text{eq}}$, if the steady states are near equilibrium. And, this fortuitously coincides with its previous narrower use in describing transitions atop equilibrium steady states—the work exerted beyond the change in free energy [241].

The excess heat Q_{ex} can be interpreted as the heat dissipated during transitions between NESSs. Similarly, the excess work W_{ex} can be interpreted as the work that *would* be dissipated if the system is allowed to relax back to a NESS. The difference between excess work W_{ex} and *dissipated work*, denoted W_{diss} , depends on a notion of excess nonequilibrium free energy, discussed shortly.

This framing leads us to see that heat is how small, possibly intelligent, systems store and transform energy via their *own agency*. This stance also moves us away from any unjustified biases that heat is necessarily wasteful. For example, an increase in heat may

indicate that a system has harvested energy, and the emission of heat may indicate an intrinsic computation [64] in the system’s state-space. The *efficiency* of the tradeoff—spending stored energy to achieve some utility—then comes into question. It is inefficiency in this sense that is by its nature wasteful.

7.4.2 Excess environmental entropy production

In isothermal transitions between equilibrium steady states, the environmental entropy production is [50]:

$$\begin{aligned}\Omega_{\text{eq}} &= \beta(W - \Delta F_{\text{eq}}) \\ &= -\beta Q - \ln \frac{\pi_{x_N}(s_N)}{\pi_{x_0}(s_0)} .\end{aligned}$$

This extends to SST by defining the *excess environmental entropy production*:

$$\begin{aligned}\Omega &= \beta W_{\text{ex}} \\ &= -\beta Q_{\text{ex}} - \ln \frac{\pi_{x_N}(s_N)}{\pi_{x_0}(s_0)} ,\end{aligned}\tag{7.6}$$

This has also been referred to as the “nonadiabatic component of entropy production” [79, 80, 10, 163]. Note that $-\ln(\pi_{x_N}(s_N)/\pi_{x_0}(s_0)) = \Delta\phi$, recovering Eq. (7.3)’s change in nonequilibrium potential ϕ .

Recalling the definitions of Q_{ex} in terms of steady state surprisals and $\phi(x, s) = -\ln \pi_x(s)$, we see that:

$$\begin{aligned}e^{-\beta Q_{\text{ex}}} &= e^{-\sum_{n=1}^N [\phi(x_n, s_n) - \phi(x_n, s_{n-1})]} \\ &= \prod_{n=1}^N \frac{\pi_{x_n}(s_n)}{\pi_{x_n}(s_{n-1})} .\end{aligned}\tag{7.7}$$

And so, Eq. (7.6) gives:

$$\begin{aligned}e^{\Omega(x_{0:N+1}, s_{0:N})} &= \frac{\pi_{x_0}(s_0)}{\pi_{x_N}(s_N)} \prod_{n=1}^N \frac{\pi_{x_n}(s_n)}{\pi_{x_n}(s_{n-1})} \\ &= \prod_{n=0}^{N-1} \frac{\pi_{x_n}(s_n)}{\pi_{x_{n+1}}(s_n)} .\end{aligned}\tag{7.8}$$

If temperature varies, then the above still holds if we replace the equilibrium probabilities with the temperature-dependent equilibrium probabilities. Thus, to go beyond the isothermal setting, we use Eq. (7.8) as the defining relationship for the excess environmental entropy production Ω . If temperature is spatially homogeneous, then it is equivalent to:

$$\Omega = \Delta\phi - \frac{1}{k_B} \int \frac{\delta Q_{\text{ex}}}{T} .$$

However, spatially inhomogeneous temperatures can also be addressed by folding temperature dependence into the environmental input x .

We return to these expressions and explore their role in generalized fluctuation theorems once we develop the necessary quantitative notions of irreversibility.

7.4.3 Path entropies

In steady state, the system state probability distribution has a Boltzmann exponential dependence on the effective energies. Naturally, out of steady state the distribution is something different. There is a nonsteady-state free energy associated with this out-of-steady-state distribution, since the system can do work (or computations) at the cost of relaxing the distribution.

Nonsteady-state free energies are controlled by path entropies, which come in several varieties. Here, we are especially interested in the controllable unaveraged state surprisals induced by the driving path \overleftarrow{x} :

$$h^{(s|x_{-\infty:t+1})} = -\ln \Pr(\mathcal{S}_t = s | x_{-\infty:t+1}) . \quad (7.9)$$

Since a semi-infinite history induces a particular distribution over system states, this can be usefully recast in terms of the initial distribution $\boldsymbol{\mu}_0$ induced by the path $x_{-\infty:1}$ and the driving history $x_{1:t+1}$ since then:

$$\begin{aligned} h^{(s|\boldsymbol{\mu}_0, x_{1:t+1})} &= -\ln \Pr(\mathcal{S}_t = s | \mathcal{S}_0 \sim \boldsymbol{\mu}_0, x_{1:t+1}) \\ &= -\ln \langle \boldsymbol{\mu}_0 | \mathsf{T}^{(\boldsymbol{\mathcal{S}} \rightarrow \boldsymbol{\mathcal{S}} | x_{1:t+1})} | s \rangle , \end{aligned} \quad (7.10)$$

where $\Pr(\mathcal{S}_t = s | \mathcal{S}_0 \sim \boldsymbol{\mu}_0, x_{1:t+1})$ is the probability that the state is s at time t , under the

measure induced when the initial state $\mathcal{S}_0 \sim \boldsymbol{\mu}_0$ (distributed according to $\boldsymbol{\mu}_0$)³ and given the driving history $x_{1:t+1} = x_1 \dots x_t$ since the initial time.

Alternatively, consider the distribution $\boldsymbol{\mu}$ induced from a start distribution by the driving history since the start. Then the path-induced state-surprisal can be expressed simply in terms of the *present* environmental-history-induced distribution over system states and the candidate state s :

$$\begin{aligned} h^{(s|\boldsymbol{\mu})} &= -\ln \Pr(\mathcal{S}_t = s | \mathcal{S}_t \sim \boldsymbol{\mu}) \\ &= -\ln \langle \boldsymbol{\mu} | s \rangle . \end{aligned} \quad (7.11)$$

Thermodynamic units of entropy are recovered by multiplying the Shannon-like path surprisals by Boltzmann's constant: $\mathfrak{s} = k_B h$.

Averaging the path-induced state-surprisal over states gives a genuine input-conditioned Shannon entropy:

$$\begin{aligned} \langle h^{(s_t | \overleftarrow{x}_t)} \rangle_{\Pr(s_t | \overleftarrow{x}_t)} &= -\left\langle \ln \Pr(s_t | \overleftarrow{x}_t) \right\rangle_{\Pr(s_t | \overleftarrow{x}_t)} \\ &= -\sum_{s_t} \Pr(s_t | \overleftarrow{x}_t) \ln \Pr(s_t | \overleftarrow{x}_t) \\ &= H[\mathcal{S}_t | \overleftarrow{X}_t = \overleftarrow{x}_t] , \end{aligned} \quad (7.12)$$

where $H[\cdot|\cdot]$ is the conditional Shannon entropy in units of nats.

It follows directly that the state-averaged path entropy $k_B H[\mathcal{S}_t | \overleftarrow{x}_t]$ is an extension of the system's steady-state nonequilibrium entropy S_{ss} . In steady-state, the state-averaged path entropy reduces to:

$$\begin{aligned} k_B H[\mathcal{S}_t | \overleftarrow{X}_t = \dots xxx] &= -k_B H[\mathcal{S}_t | \mathcal{S}_t \sim \boldsymbol{\pi}_x] \\ &= -k_B \sum_{s \in \mathcal{S}} \pi_x(s) \ln \pi_x(s) \\ &= S_{\text{ss}}(x) . \end{aligned} \quad (7.13)$$

³To be more precise, we write $\Pr(\mathcal{S}_t = s | \mathcal{S}_0 \sim \boldsymbol{\mu}_0, x_{1:t+1})$ as $\Pr_{\mathcal{S}_0 \sim \boldsymbol{\mu}_0}(\mathcal{S}_t = s | x_{1:t+1})$, since the probability is not *conditioned* on $\boldsymbol{\mu}_0$ —a probability measure for subsequent state sequences. Here, we simply gloss over this nuance, later adopting the shorthand: $\Pr(\mathcal{S}_t = s | \boldsymbol{\mu}_0, x_{1:t+1})$.

The system steady-state nonequilibrium entropy S_{ss} has been discussed as a fundamental entity in SST; e.g., see Refs. [108] and [163]. However, Eq. (7.12) ($\times k_B$) gives the appropriate extension for the thermodynamic entropy of a nonequilibrium system that is *not in steady state*. Rather, it is the entropy over system states given the entire history of environmental driving.

When \mathcal{S} is the set of microstates, rather than, say, observable mesoscopic states, the unaveraged nonequilibrium free energy F enjoys the familiar relationship between energy E and (path) entropy \mathfrak{s} :

$$F^{(s_t|x_{-\infty:t+1})} \equiv E(x_t, s_t) - T\mathfrak{s}^{(s_t|x_{-\infty:t+1})} \quad (7.14)$$

$$= F_{\text{eq}}(x_t) + \beta^{-1} \ln \frac{\Pr(s_t|x_{-\infty:t+1})}{\pi_{x_t}(s_t)} . \quad (7.15)$$

Or, averaging over states:

$$\mathcal{F}(t) = U(t) - \beta^{-1} \mathbb{H}[\mathcal{S}_t|x_{-\infty:t+1}] \quad (7.16)$$

$$= F_{\text{eq}}(x_t) + \beta^{-1} D_{\text{KL}}(\Pr(\mathcal{S}_t|x_{-\infty:t+1}) || \boldsymbol{\pi}_{x_t}) ,$$

where $\mathcal{F}(t)$ is the expected instantaneous nonequilibrium free energy, $U(t)$ is the expected instantaneous thermal energy, and $D_{\text{KL}}(\cdot||\cdot)$ is the Kullback–Leibler divergence [46]. Recognizing $k_B \mathbb{H}[\mathcal{S}_t|x_{-\infty:t+1}]$ as the natural extension of a system’s thermodynamic entropy, Eq. (7.16) is familiar from equilibrium thermodynamics, but it is now applicable arbitrarily far from equilibrium and at any time t using the instantaneous temperature. This is not the first statement of such a generalized relationship; compare, e.g., Refs. [94, 234]. In equilibrium, the expected value of the path entropy (using microstates) reduces to *the* equilibrium entropy of a system.

In the setting of effective states and NESS surprisals, we can no longer directly use Eq. (7.14). Nevertheless, by analogy with Eq. (7.15), we can still identify the *nonsteady-state addition* $\gamma(\cdot|\cdot)$ to free energy as:

$$\beta^{-1} \gamma(s|\boldsymbol{\mu}, x) \equiv \beta^{-1} \ln \frac{\Pr(\mathcal{S}_t = s | \mathcal{S}_{t-1} \sim \boldsymbol{\mu}, X_t = x)}{\pi_x(s)} .$$

Expressed differently, it is:

$$\begin{aligned}\gamma(s|\boldsymbol{\mu}, x) &= h^{(s|\boldsymbol{\pi}_x)} - h^{(s|\boldsymbol{\mu}, x)} \\ &= \phi(x, s) - h^{(s|\boldsymbol{\mu}, x)} .\end{aligned}$$

Averaging over states this becomes the Kullback–Leibler divergence between nonsteady state and steady state distributions:

$$\langle \gamma(s|\boldsymbol{\mu}, x) \rangle = D_{\text{KL}}[\text{Pr}(\mathcal{S}_t|\mathcal{S}_{t-1} \sim \boldsymbol{\mu}, X_t = x) || \boldsymbol{\pi}_x] ,$$

which is nonnegative.

Identifying the nonsteady-state contribution to the free energy allows us to introduce the *dissipated work*:

$$W_{\text{diss}} \equiv W_{\text{ex}} - \beta^{-1} \Delta \gamma ,$$

to account for the fact that excess work is not fully dissipated until the distribution relaxes back to steady state $\boldsymbol{\pi}_x$. An important consequence is that the excess work dissipated can be reclaimed by a subsequent “fluctuation” with $W_{\text{ex}} < 0$ in the midst of a driven nonequilibrium excursion.

The role of the nonsteady-state contribution to free energy will be apparent in the FTs to come shortly. This generalizes similar FTs that are restricted to starting and possibly ending in a steady state $\boldsymbol{\pi}_x$. The generalization here is key to analyzing complex systems, since many simply cannot be initiated in a steady state without losing their essential character.

7.5 Irreversibility

To emphasize, the preceding did not reference and does not require detailed balance. However, to ground the coming development, we need to describe the roles of reversibility, detailed balance, and their violations. At a minimum, this is due to most FTs assuming reversibility of the effective dynamic over states. Having established the necessary concepts and giving a measure of the irreversibility of the effective dynamic, we finally move on to FTs for nondetailed balanced processes.

7.5.1 Detailed balance

Transitioning from state a to state b , say, invoking detailed balance assumes that:

$$\frac{\Pr(\mathcal{S}_n = a | \mathcal{S}_{n-1} = b, X_n = x)}{\Pr(\mathcal{S}_n = b | \mathcal{S}_{n-1} = a, X_n = x)} = \frac{\pi_x(a)}{\pi_x(b)}.$$

Though, we do *not* assume detailed balance over the states considered here, we refer to it occasionally. For example, assuming detailed balance, microscopic reversibility and the standard Crooks fluctuation theorem (CFT) follow almost immediately.

In contrast, complex systems sustained out of equilibrium by an active substrate generically evolve via nondetailed-balance dynamics. To wit, many examples of nondetailed-balance dynamics are exhibited by chemical kinetics in biological systems [196, 144, 143]. We conclude with a thorough-going thermodynamic analysis of one neurobiological example.

7.5.2 Microscopic reversibility

Consider a particular realization of interleaved environmental-input sequence $\dots x^1 x^2 \dots x^{N-1} \dots$ and system-state sequence $\dots s^1 s^2 \dots s^{N-1} \dots$:

$$\begin{array}{ccccccc} & & \mathbf{x}, \mathbf{x}^R & & & & \\ & \overbrace{\quad\quad\quad} & & \overbrace{\quad\quad\quad} & & & \\ x^0 & x^1 & x^2 & \dots & x^{N-1} & x^N & \\ & \underbrace{\quad\quad\quad} & & \underbrace{\quad\quad\quad} & & & \\ & s^0 & s^1 & s^2 & s^{N-2} & s^{N-1} & s^N \\ & & \underbrace{\quad\quad\quad} & & \underbrace{\quad\quad\quad} & & \\ & & \mathbf{s} & & \mathbf{s}^R_{\leftarrow} & & \end{array}$$

There are several length- $(N-1)$ subsequences in play here, including the forward trajectory $\mathbf{x} = x^1 x^2 \dots x^{N-2} x^{N-1}$ of the environmental driving and the forward trajectory $\mathbf{s} = s^1 s^2 \dots s^{N-2} s^{N-1}$ of the state sequence. Furthermore, let $\mathbf{x}^R = x^{N-1} x^{N-2} \dots x^2 x^1$ be the time-reversal of the environmental driving \mathbf{x} and $\mathbf{s}^R_{\leftarrow} = s^{N-2} s^{N-3} \dots s^1 s^0$ the time-reversal of the time-shifted state sequence \mathbf{s} .

For example, if $\mathcal{X} = \{0, 1\}$ and $\mathcal{S} = \{a, b, c\}$, then \mathbf{x} may be the sequence 00101110...11000010 and \mathbf{s} the sequence *acaaaaba...abaccabc*. Then \mathbf{x}^R is the sequence 01000011...01110100. Taking the time reversal of the state sequence, we have *cbaccaba...abaaaaca*. However, since $\mathbf{s}^R_{\leftarrow}$ is also time-shifted by one time-step, we must drop the first c and append another symbol, say a . Then $\mathbf{s}^R_{\leftarrow}$ is the sequence *baccaba...abaaaaca*.

Let Q_F be the excess heat of the joint forward sequences \mathbf{x} and $s^0\mathbf{s}$, according to Eq. (7.5). By definition, a system–environment effective dynamic is *microscopically reversible* if:

$$\frac{\Pr(\mathcal{S}_{1:N} = \mathbf{s} | \mathcal{S}_0 = s^0, X_{1:N} = \mathbf{x})}{\Pr(\mathcal{S}_{1:N} = \mathbf{s}_{\leftarrow}^R | \mathcal{S}_0 = s^{N-1}, X_{1:N} = \mathbf{x}^R)} = e^{-\beta Q_F} ,$$

for any $s^0 \in \mathcal{S}$, $\mathbf{s} \in \mathcal{S}^{N-1}$, and $\mathbf{x} \in \mathcal{X}^{N-1}$. As a useful visual aid, we can re-express this as:

$$\frac{\Pr(s^0 \xrightarrow{x^1} s^1 \dots s^{N-2} \xrightarrow{x^{N-1}} s^{N-1} | s^0, \mathbf{x})}{\Pr(s^0 \xleftarrow{x^1} s^1 \dots s^{N-2} \xleftarrow{x^{N-1}} s^{N-1} | s^{N-1}, \mathbf{x}^R)} = e^{-\beta Q_F} .$$

Otherwise, microscopic reversibility is broken.

Although, microscopic reversibility has also been referred to as a “detailed fluctuation theorem”, it is actually an assumption appropriate only in special cases. For example, Eq. (7.7) shows that if the dynamics are Markovian over states (given input) and obey detailed balance (à la §7.5.1), then microscopic reversibility is satisfied for arbitrary non-Markovian inputs. In essence, this is the justification of microscopic reversibility suggested by Crooks [49, 50] from which his eponymous fluctuation theorem follows.

In this view, detailed balance and microscopic reversibility are effectively the same assumption since each implies the other. Section § 7.6.3 generalizes the CFT to describe fluctuation laws in the absence of microscopic reversibility.

7.5.3 Path dependence and irreversibility

The importance of state-space path dependence is captured via an informational quantity Υ we call the *path relevance* of a state sequence $s_{1:N}$ given initial state s_0 and input sequence $x_{1:N}$:

$$\Upsilon(s_{1:N} | s_0, x_{1:N}) \equiv \ln \frac{\Pr(s_{1:N} | s_0, x_{1:N})}{\prod_{n=1}^{N-1} \pi_{x_n}(s_n)} . \quad (7.17)$$

(The branching Pythagorean letter Υ recognizes its ancient symbolism—divergent consequences of choosing one path over another.) Note that the equilibrium probabilities in the denominator do not depend on the original state, whereas the numerator (even after factoring) depends on state-to-state transitions. As the environmental input drives the

system probability density through its state-space, path relevance develops in the state sequence. A joint sequence lacks path relevance, if $\Upsilon = 0$ for that sequence.

Whenever state transitions are Markovian given the input, the numerator in Eq. (7.17) simplifies to:

$$\Pr(s_{1:N}|x_{1:N}, s_0) = \prod_{n=1}^{N-1} \Pr(s_n|s_{n-1}, x_n) ,$$

and the path relevance becomes:

$$\Upsilon = \sum_{n=1}^{N-1} \ln \frac{\Pr(s_n|s_{n-1}, x_n)}{\pi_{x_n}(s_n)} .$$

Thus, there is path relevance even for Markov processes. The actual driving history matters. When a system is *non*-Markovian, there are yet additional contributions to path relevance.

Path relevance of a particular state sequence given a particular driving is a system feature, regardless of the environment in which the system finds itself. However, expectation values involving the above relationship can reflect the environment's nature.

For our development, we find it useful to consider both the forward-path dependence and the reverse-path dependence of a particular joint sequence: $x^1 \dots x^{N-1}$ and $s^0 \dots s^{N-1}$. The *forward-path dependence* is as expected:

$$\begin{aligned} \Upsilon &= \Upsilon(\mathbf{s}|s^0, \mathbf{x}) \\ &= \ln \frac{\Pr(\mathbf{s}|s^0, \mathbf{x})}{\prod_{n=1}^{N-1} \pi_{x^n}(s^n)} , \end{aligned} \tag{7.18}$$

and, similarly, the *reverse-path dependence* is:

$$\begin{aligned} \mathcal{J} &= \mathcal{J}(\mathbf{s}|s^0, \mathbf{x}) \\ &= \Upsilon(\mathbf{s}_{\leftarrow}^R|s^{N-1}, \mathbf{x}^R) \\ &= \ln \frac{\Pr(\mathbf{s}_{\leftarrow}^R|s^{N-1}, \mathbf{x}^R)}{\prod_{n=1}^{N-1} \pi_{x^n}(s^{n-1})} . \end{aligned} \tag{7.19}$$

And, finally, we have the *irreversibility*:

$$\Psi \equiv \Upsilon - \mathcal{J} , \tag{7.20}$$

the *net directional relevance*—of a particular path \mathbf{s} given s^0 and \mathbf{x} . Nonzero Ψ indicates the irrevocable consequences of path traversal. *Microscopically reversible dynamics have $\Psi = 0$ for all paths with nonzero probability*, indicating no divergence in path branching anywhere through the state-space. And so, $\Psi = 0$ for all paths with nonzero probability if and only if the dynamic satisfies detailed balance. In short, Ψ quantifies the imbalance in path reciprocity along a driven state-sequence.

Sometimes \mathcal{L} can be $-\infty$ for an allowed forward path $\Pr(\mathbf{s}|s^0, \mathbf{x}) > 0$, corresponding to a forbidden reverse path $\Pr(\mathbf{s}_{\leftarrow}^R|s^{N-1}, \mathbf{x}^R) = 0$. This is a situation that never arises with detailed balance dynamics. Such paths are infinitely irreversible: $\Psi = \infty$.

7.6 Generalized fluctuation theorems for nonequilibrium systems

Absent microscopic reversibility, the architecture of transitions over state-space matters. More concretely, we will constructively show how this architecture affects the nonequilibrium thermodynamics of complex systems.

7.6.1 Generalized Detailed Fluctuation Theorem

Assume the system under study starts from some distribution $\boldsymbol{\mu}_F$ and that the associated reverse trajectory (when starting from some other distribution $\boldsymbol{\mu}_R$) is allowed—that is, it has nonzero probability. Then the ratio of conditional probabilities of a state sequence

(given a driving sequence) to the reversed state sequence (given reversed driving) is:

$$\begin{aligned}
& \frac{\Pr(\boldsymbol{\mu}_F \xrightarrow{x^0} s^0 \xrightarrow{x^1} s^1 \dots s^{N-2} \xrightarrow{x^{N-1}} s^{N-1} | \boldsymbol{\mu}_F, x^0 \mathbf{x})}{\Pr(s^0 \xleftarrow{x^1} s^1 \dots s^{N-2} \xleftarrow{x^{N-1}} s^{N-1} \xleftarrow{x^N} \boldsymbol{\mu}_R | \boldsymbol{\mu}_R, x^N \mathbf{x}^R)} \\
&= \frac{\Pr(\mathcal{S}_{0:N} = s^0 \mathbf{s} | \mathcal{S}_{-1} \sim \boldsymbol{\mu}_F, X_{0:N} = x^0 \mathbf{x})}{\Pr(\mathcal{S}_{0:N} = s^{N-1} \mathbf{s}_{\leftarrow}^R | \mathcal{S}_{-1} \sim \boldsymbol{\mu}_R, X_{0:N} = x^N \mathbf{x}^R)} \\
&= \frac{\Pr(s^0 | \boldsymbol{\mu}_F, x^0)}{\Pr(s^{N-1} | \boldsymbol{\mu}_R, x^N)} \frac{\Pr(\mathbf{s} | s^0, \mathbf{x})}{\Pr(\mathbf{s}_{\leftarrow}^R | s^{N-1}, \mathbf{x}^R)} \\
&= \frac{\Pr(s^0 | \boldsymbol{\mu}_F, x^0)}{\pi_{x^0}(s^0)} \frac{\pi_{x^N}(s^{N-1})}{\Pr(s^{N-1} | \boldsymbol{\mu}_R, x^N)} \frac{\Pr(\mathbf{s} | s^0, \mathbf{x})}{\prod_{n=1}^{N-1} \pi_{x^n}(s^n)} \frac{\prod_{n=1}^{N-1} \pi_{x^n}(s^{n-1})}{\Pr(\mathbf{s}_{\leftarrow}^R | s^{N-1}, \mathbf{x}^R)} \\
&\quad \times \prod_{n=0}^{N-1} \frac{\pi_{x^n}(s^n)}{\pi_{x^{n+1}}(s^n)} \\
&= \frac{\Pr(s^0 | \boldsymbol{\mu}_F, x^0)}{\pi_{x^0}(s^0)} \frac{\pi_{x^N}(s^{N-1})}{\Pr(s^{N-1} | \boldsymbol{\mu}_R, x^N)} e^{\Psi_F} e^{\Omega_F} \\
&= e^{\gamma(s^0 | \boldsymbol{\mu}_F, x^0) - \gamma(s^{N-1} | \boldsymbol{\mu}_R, x^N)} e^{\Omega_F + \Psi_F}, \tag{7.21}
\end{aligned}$$

where $\Omega_F = \Omega(X_{0:N+1} = x^0 \mathbf{x}^N, \mathcal{S}_{0:N} = s^0 \mathbf{s})$ is the excess environmental entropy production in the forward trajectory, $\Psi_F = \Psi(\mathcal{S}_{0:N} = s^0 \mathbf{s} | X_{1:N} = \mathbf{x})$ is the irreversibility of the forward trajectory, and $\gamma(s | \boldsymbol{\mu}, x) = \ln(\Pr(s | \boldsymbol{\mu}, x) / \pi_x(s))$ is the nonsteady-state addition to free energy associated with being in the nonsteady-state distribution $\boldsymbol{\mu}$ with environmental drive x .

Since Ψ_F can diverge for forward paths with nonzero probability, we typically rewrite Eq. (7.21) as the “less divergent” expression:

$$e^{-\gamma_F} \Pr(s^0 \mathbf{s} | \boldsymbol{\mu}_F, x^0 \mathbf{x}) e^{-\Psi_F} = e^{-\gamma_R} \Pr(s^{N-1} \mathbf{s}_{\leftarrow}^R | \boldsymbol{\mu}_R, x^N \mathbf{x}^R) e^{\Omega_F}. \tag{7.22}$$

Eq. (7.22) is the fundamental relation for all that follows: it relates the probabilities of forward and reverse trajectories via entropy production Ω_F of the forward path, irreversibility Ψ_F of the forward path, and change $\beta^{-1}(\gamma_F - \gamma_R)$ in the nonsteady-state addition to free energy between the forward and reverse start-distributions.

In what follows it will be all too easy to write seemingly divergent expressions. Such divergences do not manifest themselves when taking expectation values for physical quantities involving them, since they come weighted with zero probability. This is similar

to the reasonable convention for Shannon entropies that $0 \log 0 = 0$. Nevertheless, caution is advised when probabilities vanish.

7.6.2 Simplifications

Before proceeding and to aid understanding, let's consider several special cases. If the forward drive or protocol begins with the system equilibrated to the static environmental drive x^0 , then $\boldsymbol{\mu}_F = \boldsymbol{\pi}_{x^0}$ and $\gamma(s^0|\boldsymbol{\mu}_F, x^0) = 0$. Similarly, if the reverse protocol begins with the system equilibrated to the static environmental drive x^N , then $\boldsymbol{\mu}_R = \boldsymbol{\pi}_{x^N}$ and $\gamma(s^N|\boldsymbol{\mu}_R, x^N) = 0$. In this case, Eq. (7.22) simplifies to:

$$\frac{\Pr(\boldsymbol{\pi}_{x^0} \xrightarrow{x^0} s^0 \cdots \xrightarrow{x^{N-1}} s^{N-1} | \boldsymbol{\pi}_{x^0}, x^0 \mathbf{x})}{\Pr(s^0 \xleftarrow{x^1} \cdots s^{N-1} \xleftarrow{x^N} \boldsymbol{\pi}_{x^N} | \boldsymbol{\pi}_{x^N}, x^N \mathbf{x}^R)} = e^{\Omega_F + \Psi_F}.$$

As a separate matter, if the dynamics are microscopically reversible, then $\Psi = 0$. Consider the very special case where (i) the dynamics are microscopically reversible, (ii) the forward driving begins with the system equilibrated with x^0 , and (iii) the reverse driving begins with the system equilibrated with x^N . Then, the ratio of probabilities of observing a forward state sequence (given forward driving) and observing the reversal of that state sequence (given the reversal of that driving) is simply e^{Ω_F} . That is, the forward sequence is exponentially more likely if it has positive entropy production.

Apparently, the more general case is more nuanced and, beyond depending on a nonsteady-state starting distribution, it depends strongly on the architecture of branching transitions among states.

Another interesting special case is if $x^N = x^{N-1}$ and $\boldsymbol{\mu}_R$ is the distribution that the forward driving induces from $\boldsymbol{\mu}_F$. Then the *dissipated work* $W_{\text{diss}} \equiv W_{\text{ex}} - \beta^{-1} \Delta \gamma$ associated with the forward trajectory comes into play. (Recall that $\beta^{-1} \Delta \gamma$ is the change in nonsteady-state contributions to free energy.) Then the ratio of forward- and reverse-path probabilities is:

$$\begin{aligned} \frac{\Pr(\boldsymbol{\mu}_F \xrightarrow{x^0} \cdots \xrightarrow{x^{N-1}} s^{N-1} | \boldsymbol{\mu}_F, \mathbf{x})}{\Pr(s^0 \xleftarrow{x^1} \cdots \xleftarrow{x^N} \boldsymbol{\mu}(\boldsymbol{\mu}_F, \mathbf{x}) | \boldsymbol{\mu}(\boldsymbol{\mu}_F, \mathbf{x}), \mathbf{x}^R)} &= e^{\Psi_F} e^{\beta[W_{\text{ex}} - \beta^{-1} \Delta \gamma]} \\ &= e^{\Psi_F} e^{\beta W_{\text{diss}}} . \end{aligned}$$

Even in the case of microscopic reversibility, this is useful since it generalizes previous FTs to nonequilibrium start and end distributions. In the case of microscopic reversibility, $\Psi_F = 0$ and so the ratio of forward- and reverse-path probabilities from any nonequilibrium start and end distribution is exponential $e^{\beta W_{\text{diss}}}$ in the dissipated work. Thus, an experimental test of this result is one with time-symmetric driving. The forward protocol corresponds to the first half of the driving while the reverse protocol is the second half. Clearly, the final nonequilibrium distribution for the forward protocol is the same as the initial nonequilibrium distribution for the reverse protocol. The dissipated work then corresponds to that dissipated in the first half of the driving. Practically, in cases where the dynamic is not microscopically reversible, this allows experimentally extracting the system's irreversibility Ψ .

7.6.3 Generalized Crooks Fluctuation Theorem

We can now turn to the irreversible analog of the Crooks Fluctuation Theorem (CFT).

First, we note that both entropy production Ω and irreversibility Ψ are odd under time reversal. Explicitly, we have:

$$\begin{aligned}
\Omega(X_{0:N+1} = x^0 \mathbf{x} x^N, S_{0:N} = s^0 \mathbf{s}) &= \ln \prod_{n=0}^{N-1} \frac{\pi_{x^n}(s^n)}{\pi_{x^{n+1}}(s^n)} \\
&= -\ln \prod_{n=0}^{N-1} \frac{\pi_{x^{n+1}}(s^n)}{\pi_{x^n}(s^n)} \\
&= -\ln \prod_{n=0}^{N-1} \frac{\pi_{x^{N-n}}(s^{N-1-n})}{\pi_{x^{N-1-n}}(s^{N-1-n})} \\
&= -\Omega(X_{0:N+1} = x^N \mathbf{x}^R x^0, S_{0:N} = s^{N-1} \mathbf{s}_{\leftarrow}^R)
\end{aligned}$$

and:

$$\begin{aligned}
\Psi(S_{0:N} = s^0 \mathbf{s} | X_{1:N} = \mathbf{x}) &= \ln \left[\frac{\Pr(\mathbf{s} | s^0, \mathbf{x})}{\Pr(\mathbf{s}_{\leftarrow}^R | s^{N-1}, \mathbf{x}^R)} \prod_{n=1}^{N-1} \frac{\pi_{x^n}(s^{n-1})}{\pi_{x^n}(s^n)} \right] \\
&= -\ln \left[\frac{\Pr(\mathbf{s}_{\leftarrow}^R | s^{N-1}, \mathbf{x}^R)}{\Pr(\mathbf{s} | s^0, \mathbf{x})} \prod_{n=1}^{N-1} \frac{\pi_{x^n}(s^n)}{\pi_{x^n}(s^{n-1})} \right] \\
&= -\Psi(S_{0:N} = s^{N-1} \mathbf{s}_{\leftarrow}^R | X_{1:N} = \mathbf{x}^R) .
\end{aligned}$$

For brevity, let

$$\Omega_F \equiv \Omega(X_{0:N+1} = x^0 \mathbf{x} x^N, \mathcal{S}_{0:N} = s^0 \mathbf{s}) \text{ and}$$

$$\Omega_R \equiv \Omega(X_{0:N+1} = x^N \mathbf{x}^R x^0, \mathcal{S}_{0:N} = s^{N-1} \mathbf{s}_{\leftarrow}^R).$$

And, similarly,

$$\Psi_F \equiv \Psi(\mathcal{S}_{0:N} = s^0 \mathbf{s} | X_{1:N} = \mathbf{x}) \text{ and}$$

$$\Psi_R \equiv \Psi(\mathcal{S}_{0:N} = s^{N-1} \mathbf{s}_{\leftarrow}^R | X_{1:N} = \mathbf{x}^R).$$

In this notation, we just established that $\Omega_F = -\Omega_R$ and $\Psi_F = -\Psi_R$.

Second, if we now choose $\boldsymbol{\mu}_F = \boldsymbol{\pi}_{\mathbf{x}^0}$ and $\boldsymbol{\mu}_R = \boldsymbol{\pi}_{\mathbf{x}^N}$ and marginalize over all possible state trajectories, we find that the joint probability of entropy production and irreversibility given the driving protocol starting from an equilibrium distribution is:

$$\begin{aligned} \Pr(\Omega, \Psi | \boldsymbol{\pi}_{\mathbf{x}^0}, x^0 \mathbf{x} x^N) &= \sum_{s^0 \mathbf{s} \in \mathcal{S}^N} \Pr(s^0 \mathbf{s} | \boldsymbol{\pi}_{\mathbf{x}^0}, x^0 \mathbf{x}) \delta_{\Omega, \Omega_F} \delta_{\Psi, \Psi_F} \\ &= \sum_{s^0 \mathbf{s} \in \mathcal{S}^N} e^{\Omega_F} e^{\Psi_F} \Pr(s^{N-1} \mathbf{s}_{\leftarrow}^R | \boldsymbol{\pi}_{\mathbf{x}^N}, x^N \mathbf{x}^R) \delta_{\Omega, \Omega_F} \delta_{\Psi, \Psi_F} \\ &= e^{\Omega} e^{\Psi} \sum_{s^0 \mathbf{s} \in \mathcal{S}^N} \Pr(s^{N-1} \mathbf{s}_{\leftarrow}^R | \boldsymbol{\pi}_{\mathbf{x}^N}, x^N \mathbf{x}^R) \delta_{\Omega, \Omega_F} \delta_{\Psi, \Psi_F} \\ &= e^{\Omega} e^{\Psi} \sum_{s^{N-1} \mathbf{s}_{\leftarrow}^R \in \mathcal{S}^N} \Pr(s^{N-1} \mathbf{s}_{\leftarrow}^R | \boldsymbol{\pi}_{\mathbf{x}^N}, x^N \mathbf{x}^R) \delta_{\Omega, -\Omega_R} \delta_{\Psi, -\Psi_R} \\ &= e^{\Omega} e^{\Psi} \Pr(-\Omega, -\Psi | \boldsymbol{\pi}_{\mathbf{x}^N}, x^N \mathbf{x}^R x^0) . \end{aligned}$$

Finally, we rewrite this to give the *extended CFT for irreversible processes*:

$$\frac{\Pr(\Omega, \Psi | \boldsymbol{\pi}_{\mathbf{x}^0}, x^0 \mathbf{x} x^N)}{\Pr(-\Omega, -\Psi | \boldsymbol{\pi}_{\mathbf{x}^N}, x^N \mathbf{x}^R x^0)} = e^{\Psi} e^{\Omega} . \quad (7.23)$$

7.6.4 Interpretation

In the special case of isothermal time-symmetric driving— $x^0 \mathbf{x} x^0 = x^0 \mathbf{x}^R x^0 = x^0 x^1 x^2 \dots x^2 x^1 x^0$ —and starting from an equilibrium distribution, Eq. (7.23) provides a useful comparison between values of excess work achieved by the single time-symmetric driving protocol:

$$\frac{\Pr(W_{\text{ex}}, \Psi)}{\Pr(-W_{\text{ex}}, -\Psi)} = e^{\Psi} e^{\beta W_{\text{ex}}} . \quad (7.24)$$

Equation (7.23) should be compared to the original CFT that, in its most general form,

can be written (with necessary interpretation) as [51]:

$$\frac{\text{Pr}_F(\Omega)}{\text{Pr}_R(-\Omega)} = e^\Omega . \quad (7.25)$$

It is tempting to write Eq. (7.25) as:

$$\frac{\text{Pr}(\Omega | \boldsymbol{\pi}_{\mathbf{x}^0}, x^0 \mathbf{x}^N)}{\text{Pr}(-\Omega | \boldsymbol{\pi}_{\mathbf{x}^N}, x^N \mathbf{x}^R x^0)} \stackrel{?}{=} e^\Omega . \quad (7.26)$$

This form presents some concerns, however. In the case of detailed balance, though, $\Psi = 0$ for all trajectories, and so our Eq. (7.23) guarantees Eq. (7.26) in the case of detailed balance. Crooks' original CFT derivation [49, 50] also assumed detailed balance, and so Eq. (7.26) was implied.

However, absent detailed balance, Eq. (7.25) has a rather different interpretation: $\text{Pr}_R(\cdot)$ then implies not only the reversed driving, but also that the distribution describes a different “reversed” system that is not of direct physical relevance [51, 39]. One consequence is that the probabilities in the numerator and denominator are not comparable in any physical sense. So, in general, we have:

$$\frac{\text{Pr}(\Omega | \boldsymbol{\pi}_{\mathbf{x}^0}, x^0 \mathbf{x}^N)}{\text{Pr}(-\Omega | \boldsymbol{\pi}_{\mathbf{x}^N}, x^N \mathbf{x}^R x^0)} \neq e^\Omega . \quad (7.27)$$

In contrast, our irreversible CFT in Eq. (7.23) compares probabilities of entropy production (and path irreversibility) for the same thermodynamic system under a control protocol and under the reversed control protocol. Equation (7.23), unlike equalities involving an unphysical dual dynamic as in Eq. (7.25), allows a clear and meaningful physical interpretation of the relationship between entropies produced and, moreover, is not limited by assuming detailed balance.

Note that our Eq. (7.23), expressed in terms of excess environmental entropy production Ω and path irreversibility Ψ , does not make explicit mention of temperature. Indeed, if temperature dependence is folded into different environmental inputs x , then Eq. (7.23) applies just as well to systems driven by environments with spatially inhomogeneous temperature distributions that change in time. Explicitly, $\boldsymbol{\pi}_{\mathbf{x}}$ and $\boldsymbol{\pi}_{\mathbf{x}'}$ could represent the distribution over effective states induced by environmental conditions associated with x and x' *including* their different spatial distributions of temperature.

7.6.5 Translation to Steady-State Thermodynamics

A better understanding of the irreversible CFT comes by comparing it to recent related work. Most directly, our results complement those on driven transitions between NESSs. Specifically, the importance of nondetailed-balanced dynamics in enabling the organization of complex nonequilibrium behavior has been considered previously. For example, Ref. [94] also introduced a path entropy which is an ensemble average of that considered here.

Another comparison is found in Ref. [108]’s nonequilibrium thermodynamics over NESSs using housekeeping Q_{hk} and excess Q_{ex} heats. While that treatment focused on Langevin dynamics, we find that in general Q_{hk} corresponds directly to our path irreversibility Ψ . Specifically, in the isothermal setting there, according to Eq. (35), we have:

$$\beta Q_{\text{hk}} \approx \Psi .$$

Indeed, for isothermal Markovian dynamics Eq. (7.7) of Ref. [106] suggests (via their Eqs. (2.11) and (7.1)) that this is in fact an equality:

$$\beta Q_{\text{hk}} = \Psi . \tag{7.28}$$

Reference [80] called the irreversibility Ψ the *adiabatic contribution* to entropy production. Several related translations from Ref. [108] to our setting can also be easily made: $\rho_{\text{ss}}(s; x) \rightarrow \pi_x(s)$, $\phi(s; x) \rightarrow -\log \pi_x(s)$, $\Delta S \rightarrow \Delta S_{\text{ss}}$, and $\beta Q_{\text{ex}} + \Delta \phi \rightarrow \Omega$. Hence, $\langle \Omega \rangle \geq 0$ (for Langevin systems) is Ref. [108]’s main result. From these connections, we see that our development not only provides new constraints on detailed fluctuations, but also extends these earlier results beyond Langevin systems.

Exposing these translations allows reformulating our detailed fluctuation theorems to steady-state thermodynamics (SST). We have:

$$\begin{aligned} e^{\gamma(s^0|\boldsymbol{\mu}_{\text{F}}, x^0) - \gamma(s^{N-1}|\boldsymbol{\mu}_{\text{R}}, x^N)} e^{\Omega_{\text{F}} + \Psi_{\text{F}}} &= e^{\beta(Q_{\text{ex}} + Q_{\text{hk}}) + \Delta S^{\text{sys}}} \\ &= e^{\Delta S_{\text{F}}^{\text{tot}}} , \end{aligned}$$

where $\Delta S^{\text{sys}} \equiv -\ln \frac{\Pr(s^{N-1}|\boldsymbol{\mu}_{\text{R}}, x^N)}{\Pr(s^0|\boldsymbol{\mu}_{\text{F}}, x^0)}$ when we choose $\langle \boldsymbol{\mu}_{\text{R}} | = \langle \boldsymbol{\mu}_{\text{F}} | \mathsf{T}^{(\boldsymbol{s} \rightarrow \boldsymbol{s} | x^0 \mathbf{x})}$ and where $S_{\text{F}}^{\text{tot}}$

is the total change in entropy in forward time. This yields:

$$\frac{\Pr(\mathcal{S}_{0:N} = s^0 \mathbf{s} | \mathcal{S}_{-1} \sim \boldsymbol{\mu}_F, X_{0:N} = x^0 \mathbf{x})}{\Pr(\mathcal{S}_{0:N} = s^{N-1} \mathbf{s}_{\leftarrow}^R | \mathcal{S}_{-1} \sim \boldsymbol{\mu}_R, X_{0:N} = x^N \mathbf{x}^R)} = e^{\Delta S_F^{\text{tot}}} .$$

And so, we immediately see that:

$$\langle e^{-\Delta S_F^{\text{tot}}} \rangle_{\Pr(\mathcal{S}_{0:N} = s^0 \mathbf{s} | \mathcal{S}_{-1} \sim \boldsymbol{\mu}_F, X_{0:N} = x^0 \mathbf{x})} = 1 .$$

This extends the validity of Ref. [221]’s general integral fluctuation theorem beyond Langevin dynamics. Since the total change in entropy is time asymmetric— $\Delta S_F^{\text{tot}} = -\Delta S_R^{\text{tot}}$ —we obtain the most direct CFT generalization valid outside of detailed balance:

$$\frac{\Pr(\Delta S^{\text{tot}} | \boldsymbol{\pi}_{\mathbf{x}^0}, x^0 \mathbf{x} x^N)}{\Pr(-\Delta S^{\text{tot}} | \boldsymbol{\pi}_{\mathbf{x}^N}, x^N \mathbf{x}^R x^0)} = e^{\Delta S^{\text{tot}}} . \quad (7.29)$$

Again, this does not invoke a dual, unphysical dynamic. Equation (7.29) has been reported previously in various settings; see, e.g., Eq. (21) of Ref. [80] and Eq. (43) of Ref. [39]. The result gives a detailed fluctuation relation for the change in total entropy production when transitioning between steady states.

The new detailed fluctuation theorem of Eq. (7.23) for joint distributions goes further in refining SST. If starting in a steady state and executing a protocol in an isothermal environment, we find that:

$$\frac{\Pr(W_{\text{ex}}, Q_{\text{hk}} | \boldsymbol{\pi}_{\mathbf{x}^0}, x^0 \mathbf{x} x^N)}{\Pr(-W_{\text{ex}}, -Q_{\text{hk}} | \boldsymbol{\pi}_{\mathbf{x}^N}, x^N \mathbf{x}^R x^0)} = e^{\beta Q_{\text{hk}}} e^{\beta W_{\text{ex}}} .$$

This novel relation gives strong constraints on the thermodynamic behavior of systems driven between NESSs, since it constrains the *joint* distribution for excess work and housekeeping heat. Moreover, nonsteady-state additions to free energy are predicted when an experiment does not start in steady state.

In the special case of time-symmetric driving— $x^0 \mathbf{x} x^0 = x^0 \mathbf{x}^R x^0 = x^0 x^1 x^2 \dots x^2 x^1 x^0$ —and starting from an equilibrium distribution, the preceding expression reduces to a useful comparison between excess work values achieved by the single time-symmetric protocol:

$$\frac{\Pr(W_{\text{ex}}, Q_{\text{hk}})}{\Pr(-W_{\text{ex}}, -Q_{\text{hk}})} = e^{\beta Q_{\text{hk}}} e^{\beta W_{\text{ex}}} .$$

Similar results were recently derived in Ref. [136] under more restrictive assumptions for underdamped Langevin systems.

This all said, one must use caution and not always identify Ψ with βQ_{hk} . Most importantly, not all sources of irreversibility are naturally characterized as “heat”. Thinking of irreversibility on its own dynamical terms is best.

7.6.6 Integral fluctuation theorems

Integral fluctuation theorems in the absence of detailed balance, starting arbitrarily far from equilibrium, also follow straightforwardly. One generalization of the integral fluctuation theorem [121] is:

$$\begin{aligned}
& \langle e^{-\beta W_{\text{diss}} - \Psi} \rangle_{\text{Pr}(s_{0:N} | \boldsymbol{\mu}_F, \mathbf{x})} \\
&= \sum_{s_{0:N} \in \mathcal{S}^N} \text{Pr}(s_{0:N} | \boldsymbol{\mu}_F, \mathbf{x}) e^{-\beta W_{\text{diss}} - \Psi} \\
&= \sum_{s_{0:N} \in \mathcal{S}^N} \text{Pr}(s^0 \xleftarrow{x^1} \dots \xleftarrow{x^N} \boldsymbol{\mu}(\boldsymbol{\mu}_F, \mathbf{x}) | \boldsymbol{\mu}(\boldsymbol{\mu}_F, \mathbf{x}), \mathbf{x}^R) \\
&= 1 .
\end{aligned} \tag{7.30}$$

If the input is stochastic, then averaging over the input also gives:

$$\langle e^{-\beta W_{\text{diss}} - \Psi} \rangle_{\text{Pr}(x_{0:N}, s_{0:N} | \boldsymbol{\mu}_F)} = 1 .$$

Note that this relation does *not* require the system to be in steady state at any time.

From the concavity of the exponential function, it is tempting to assert a corresponding generalized Second Law of SST as:

$$\langle W_{\text{diss}} \rangle \geq - \langle Q_{\text{hk}} \rangle . \tag{7.31}$$

Although Eq. (7.31) is true, notably it is neither a strong nor useful bound. Let’s address this. Note that:

$$\langle e^{-\Psi} \rangle_{\text{Pr}(s_{0:N} | \boldsymbol{\mu}_F, \mathbf{x})} = 1 \tag{7.32}$$

and:

$$\langle e^{-\beta W_{\text{diss}}} \rangle_{\text{Pr}(s_{0:N} | \boldsymbol{\mu}_F, \mathbf{x})} = 1 . \tag{7.33}$$

Both follow from the normalization of probabilities of the conjugate dynamic. Therefore, $\langle \Psi \rangle \geq 0$ and $\langle W_{\text{diss}} \rangle \geq 0$. And, hence $\langle Q_{\text{hk}} \rangle \geq 0$ as shown in Ref. [237]. So, Eq. (7.31) is devoid of utility. Nevertheless, Eq. (7.30) puts a novel constraint on the joint distributions of W_{diss} and Ψ .

Introducing an artificial conjugate dynamic following Ref. [51] and following the derivation there with ϕ/β in place of E , when starting in the steady state distribution $\pi_{\mathbf{x}_0}$, we can show that:

$$\langle e^{-\Omega} \rangle_{\text{Pr}(s_{0:N}|\pi_{\mathbf{x}_0}, \mathbf{x})} = 1 , \quad (7.34)$$

which implies the restriction $\langle \Omega \rangle \geq 0$. Despite similar appearance, this result has meaning beyond Crooks' derivation of the Jarzynski equality, as it now also applies *atop nonequilibrium steady states*. Recall that Ω has general meaning as in Eq. (7.6): $\Omega = \beta W_{\text{ex}} = -\beta Q_{\text{ex}} + \Delta\phi$. So, Eq. (7.34) becomes:

$$\langle e^{-\beta W_{\text{ex}}} \rangle_{\text{Pr}(s_{0:N}|\pi_{\mathbf{x}_0}, \mathbf{x})} = 1 . \quad (7.35)$$

Effectively, this is Ref. [108]'s relation that, with our sign convention for Q_{ex} , implies:

$$\begin{aligned} \langle \Omega \rangle &= \langle -\beta Q_{\text{ex}} + \Delta\phi \rangle \\ &\geq 0 , \end{aligned} \quad (7.36)$$

for processes that start in steady state.

However, using Eq. (7.33), we find a more precise constraint on expected excess entropy production whether or not the system starts in steady state:

$$\langle \Omega \rangle \geq \Delta \langle \gamma \rangle , \quad (7.37)$$

where the RHS can be positive or negative, but can only be negative if the system starts out of steady state. When starting in a steady state, this yields:

$$\langle \Omega \rangle \geq \langle \gamma_{\text{final}} \rangle \quad (7.38)$$

$$= D_{\text{KL}}[\text{Pr}(\mathcal{S}_t | \mathcal{S}_0 \sim \pi_{\mathbf{x}_0}, x_{1:t+1}) || \pi_{\mathbf{x}_t}] , \quad (7.39)$$

which is a stronger constraint than the previous result of Eq. (7.36), since the RHS is always positive for $\Pr(\mathcal{S}_t|\mathcal{S}_0 \sim \pi_{x_0, x_{1:t+1}}) \neq \pi_{x_t}$. Eq. (7.39) extends the validity of the main result obtained in Ref. [260] to now include the possibility of starting in a nonequilibrium steady state and allowing for non-detailed-balanced dynamics. (Note that ‘ W_{diss} ’ in Ref. [260] corresponds to our W_{ex} —it is excess work that is not necessarily yet dissipated.)

Integral fluctuation theorems for systems with controlled or intrinsic feedback also directly follow, as we now show, extending the theory of feedback control to the setting of transitions between NESSs.

7.6.7 Fluctuation theorems with an auxiliary variable

Actions made by a complex thermodynamic system can couple back from the environment to influence the system’s future input. To achieve this, the system may express an auxiliary random variable Y_t —the current “output” that takes on the values $y \in \mathcal{Y}$ and is instantaneously energetically mute, but may influence the future input and so does have energetic relevance.

The variable Y_t could be measurement, output, or any other auxiliary variable that influences the state or input sequences. To be concise, we introduce a shorthand for the time-ordered sequences of random variables: $\vec{X} \equiv X_{0:N}$, $\vec{S} \equiv \mathcal{S}_{0:N}$, and $\vec{Y} \equiv Y_{0:N}$. And, for particular realizations of the sequences: $\vec{x} \equiv x^{0:N}$, $\vec{s} \equiv s^{0:N}$, and $\vec{y} \equiv y^{0:N}$. When time reversing realizations, we let $\overleftarrow{x} = x^{N-1}x^{N-2} \dots x^1x^0$ and $\overleftarrow{s} = s^{N-1}s^{N-2} \dots s^1s^0$. To clarify further, \vec{x} appearing inside a probability implies $\vec{X} = \vec{x}$ and \overleftarrow{s} appearing inside a probability implies $\overleftarrow{S} = \overleftarrow{s}$.

We quantify how much the auxiliary variable is independently informed from the state sequence—beyond what could be known if given only the initial distribution over states and the driving history—via the unaveraged conditional mutual information:

$$\begin{aligned} i(\vec{s}; \vec{y} | \vec{x}, \mu_F) &\equiv \ln \frac{\Pr(\vec{s}, \vec{y} | \vec{x}, \mu_F)}{\Pr(\vec{y} | \vec{x}, \mu_F) \Pr(\vec{s} | \vec{x}, \mu_F)} \\ &= \ln \frac{\Pr(\vec{s}, \vec{y}, \vec{x} | \mu_F)}{\Pr(\vec{y}, \vec{x} | \mu_F) \Pr(\vec{s} | \vec{x}, \mu_F)}. \end{aligned}$$

Note that averaging over the input, state, and auxiliary sequences gives the familiar conditional mutual information: $I[\vec{S}; \vec{Y} | \vec{X}, \boldsymbol{\mu}_F] = \langle i[\vec{S}; \vec{Y} | \vec{x}, \boldsymbol{\mu}_F] \rangle_{\text{Pr}(\vec{x}, \vec{s}, \vec{y} | \boldsymbol{\mu}_F)}$.

As detailed in the App. 7.10.2:

$$e^{\beta W_{\text{diss}} + i[\vec{S}; \vec{Y} | \vec{x}, \boldsymbol{\mu}_F] + \Psi} = \frac{\text{Pr}(\vec{s}, \vec{y}, \vec{x} | \boldsymbol{\mu}_F)}{\text{Pr}(\vec{y}, \vec{x} | \boldsymbol{\mu}_F) \text{Pr}(\vec{s} | \vec{x}, \boldsymbol{\mu}_R)},$$

where $\boldsymbol{\mu}_R = \boldsymbol{\mu}(\boldsymbol{\mu}_F, \vec{x})$. This leads directly to the integral fluctuation theorem:

$$\left\langle e^{-\beta W_{\text{diss}} - i[\vec{S}; \vec{Y} | \vec{x}, \boldsymbol{\mu}_F] - \Psi} \right\rangle_{\text{Pr}(\vec{s}, \vec{y}, \vec{x} | \boldsymbol{\mu}_F)} = 1. \quad (7.40)$$

However, as before, the resulting bound on $\langle W_{\text{diss}} \rangle$ is not the tightest possible. Alternatively, we can invoke the normalization of conjugate dynamics to show:

$$\left\langle e^{-\beta W_{\text{diss}} - i[\vec{S}; \vec{Y} | \vec{x}, \boldsymbol{\mu}_F]} \right\rangle_{\text{Pr}(\vec{s}, \vec{y}, \vec{x} | \boldsymbol{\mu}_F)} = 1. \quad (7.41)$$

This implies a new lower bound for the revised Second Law of Thermodynamics:

$$\langle W_{\text{diss}} \rangle \geq -k_B T I[\vec{S}; \vec{Y} | \vec{X}, \boldsymbol{\mu}_F], \quad (7.42)$$

enabled by the conditional mutual information between state-sequence and auxiliary sequence, given input-sequence. Notably, this relation holds arbitrarily far from equilibrium and allows for the starting and ending distributions to be nonsteady-state.

We may also be interested in the unaveraged unconditioned mutual information between the auxiliary variable sequence and the joint input-state sequence. Then, using:

$$i[\vec{y}; \vec{x} \vec{s} | \boldsymbol{\mu}_F] \equiv \ln \frac{\text{Pr}(\vec{x}, \vec{s}, \vec{y} | \boldsymbol{\mu}_F)}{\text{Pr}(\vec{y} | \boldsymbol{\mu}_F) \text{Pr}(\vec{x}, \vec{s} | \boldsymbol{\mu}_F)},$$

we find that, in general:

$$\langle W_{\text{diss}} \rangle \geq -k_B T I[\vec{Y}; \vec{X} \vec{S} | \boldsymbol{\mu}_F] \quad (7.43)$$

and when starting in steady-state:

$$\langle \Omega \rangle \geq -I[\vec{Y}; \vec{X} \vec{S} | \boldsymbol{\pi}_{x^0}]. \quad (7.44)$$

One can now continue in this fashion to successively derive a seeming unending sequence of fluctuation theorems. Let's stop, though, with one more and discuss its interpretations and applications.

As a final set of example integral fluctuation theorems, we follow Ref. [209] in defining:

$$i_{\text{SU}} \equiv \ln \frac{\Pr(\vec{y}, \vec{s} | \boldsymbol{\mu}_0)}{\Pr(\vec{y} | \boldsymbol{\mu}_0) \Pr(\vec{s} | \boldsymbol{\mu}_0, \vec{x})} .$$

(This is Ref. [209]’s I_C , if $\boldsymbol{\mu}_0 \rightarrow \boldsymbol{\pi}_{x^0}$.) Technically speaking, this is not a mutual information, even upon averaging. Then, we arrive at the integral fluctuation theorems:

$$\langle e^{-W_{\text{diss}} - i_{\text{SU}} - \Psi} \rangle_{\Pr(\vec{s}, \vec{y}, \vec{x} | \boldsymbol{\mu}_0)} = 1$$

and

$$\langle e^{-W_{\text{diss}} - i_{\text{SU}}} \rangle_{\Pr(\vec{s}, \vec{y}, \vec{x} | \boldsymbol{\mu}_0)} = 1 .$$

When starting from a steady-state distribution, we have the most direct generalization of Ref. [209]’s feedback control result, but extended to not require detailed balance:

$$\langle e^{-\Omega - i_{\text{SU}}} \rangle_{\Pr(\vec{s}, \vec{y}, \vec{x} | \boldsymbol{\pi}_{\mathbf{x}^0})} = 1 .$$

When starting from a NESS, this suggests that:

$$\langle \Omega \rangle \geq -I_{\text{SU}} . \quad (7.45)$$

When the dynamics are detailed balance, this naturally reduces to the well known results of Ref. [209] and others: $\langle W \rangle \geq \Delta F_{\text{eq}} - k_B T I_{\text{SU}}$.

In the feedback control setting, Y_n is said to be the random variable for measurements at time n . This suggests that Y_n is a function of \mathcal{S}_n and the outcome of Y_n effectively induces different Markov chains over the states since X_{n+1} is a function of Y_n —i.e., $x_{n+1}(y_n(s_n))$.

With our interest in complex autonomous systems, we note that our results give new bounds on the Second Law of Thermodynamics for highly structured complex systems strongly coupled to an environment. A preliminary application of this was presented out in Ref. [212]. We offer our own in the next section. Analysis of thermodynamic systems with the agency to influence their environmental input, via some kind of coupling or feedback, say, will likely benefit from our extended theory.

How can we reconcile this with other inequalities without auxiliary Y ? The other inequalities used averages of variable occurrence already conditioned on \vec{x} . However,

if input x and states s can influence each other dynamically through auxiliary y , then averaging over their joint dynamic allows less dissipation than the traditional Second Law suggests.

If \mathcal{S} represents the random variable for one subset of a system’s degrees of freedom, and Y represents the random variable for another subset of a system’s degrees of freedom, then the intrinsic nonextensivity of the thermodynamic entropy $S(\mathcal{S}, Y|X) = S(\mathcal{S}|X) + S(Y|X) - k_B I(\mathcal{S}; Y|X)$ goes a long way towards explaining the physics of information stimulating the recent resurgence of Maxwellian demonology. This viewpoint will be further developed elsewhere.

7.7 NESS Transitions in Neuronal Dynamics

Acting in concert, voltage-gated sodium ion channels and potassium ion channels are the primary thermodynamic substrate that drives the evolution of membrane potentials in neurons [117]. Together, these voltage-gated channels are the primary generators of the action potentials or “spikes” that are the basic signals whose collective patterns support neural information processing [206]. In experiments, if the cell membrane is voltage clamped, then the channels approach a stationary distribution over their conformational states according to the effective energies of their biomolecular conformations at that voltage [111]. However, absent clamping, the channels influence their own voltage input dynamically through their current output. The result is spontaneous spiking patterns.

Although this is not the setting in which to analyze the full richness of ion channel interactions, we will use the sodium channel under different voltage-driving protocols as a relatively straightforward example to demonstrate the insights on NESS transitions gained from preceding theoretical results. That is, while potassium ion channels are somewhat structured, the sodium ion channel exhibits a more structured and so more illustrative dynamic over its coarse-grained state space of functional protein conformations.

7.7.1 Ion Channel Dynamics

For sodium ions to move through the neural membrane, a channel’s activation gates must be open and the deactivation gate must not yet plug the channel [66]. The rates of

transitions among the conformational states have a highly nontrivial dependence on voltage across the cell membrane. Beyond this voltage dependence, while the activation gates act largely independently of one another, the inactivation gate cannot plug the channel until at least some of the activation gates are open. This causal architecture was not yet captured by the relatively macroscopic differential equations introduced in the pioneering work of Hodgkin and Huxley [111]. Since then, however, it has been summarized by experimentally-motivated voltage-dependent Markov chains over the causally relevant conformational states [189]. Here, we follow the model implied in Ref. [66], whose voltage-dependent Markov chain we show in Fig. 7.1.

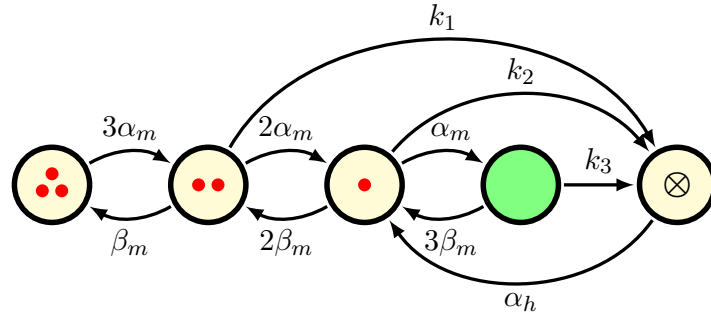


Figure 7.1: Markov chain representation of the input-conditioned state-to-state rate matrix $G^{(\mathcal{S} \rightarrow \mathcal{S}|v)}$. Self-transitions are implied but, for tidiness, are not shown explicitly. These coarse-grainings of conformational states have biologically important functional interpretations. The number of (red) dots in each effective state corresponds to the number of activation gates that close off the channel. For example, when the channel is in the leftmost state, three activation gates are still active in blocking the channel. The solid (green) state corresponds to the channel being open. This is the only one of the five states in which sodium current can flow. The last state, marked \otimes , corresponds to channel inactivation by the inactivation gate—when the channel is plugged by its “ball and chain”. Subsequent figures use the state numbering 1 through 5 for state identification, which corresponds to enumeration of the states from left to right here.

As the mesoscopic system of thermodynamic interest, the voltage-dependent Markov chain can be re-interpreted as a transducer that takes in voltage $v \in \mathbb{R}$ across the cell membrane as its input and makes transitions over its conformational state space \mathcal{S} according to an infinite set of transition matrices $\{\mathbf{T}^{(\mathcal{S} \rightarrow \mathcal{S}|v)}\}_{v \in \mathbb{R}}$. Although the set is uncountable, the voltage-conditioned transition matrices are all described succinctly via time-independent functions of the voltage appearing in the transition elements; denoted

α_m , β_m , and α_h . Time discretization of the continuous-time dynamic is straightforward and well behaved as $\Delta t \rightarrow 0$. If the voltage v is approximately constant during the infinitesimal interval Δt , then the state-to-state transition matrix is:

$$\begin{aligned} \mathbf{T}_{\Delta t}^{(\mathcal{S} \rightarrow \mathcal{S}|v)} &= e^{(\Delta t)G^{(\mathcal{S} \rightarrow \mathcal{S}|v)}} \\ &\approx I + (\Delta t)G^{(\mathcal{S} \rightarrow \mathcal{S}|v)} , \end{aligned}$$

where I is the identity matrix and $G^{(\mathcal{S} \rightarrow \mathcal{S}|v)}$ is the infinitesimal generator of time evolution:

$$G^{(\mathcal{S} \rightarrow \mathcal{S}|v)} \equiv \begin{bmatrix} -3\alpha_m & 3\alpha_m & 0 & 0 & 0 \\ \beta_m & -(2\alpha_m + \beta_m + k_1) & 2\alpha_m & 0 & k_1 \\ 0 & 2\beta_m & -(\alpha_m + 2\beta_m + k_2) & \alpha_m & k_2 \\ 0 & 0 & 3\beta_m & -(3\beta_m + k_3) & k_3 \\ 0 & 0 & \alpha_h & 0 & -\alpha_h \end{bmatrix} . \quad (7.46)$$

Specifically, α_m , β_m , and α_h are voltage-dependent variables, as found in the Hodgkin and Huxley model [111, 66]:

$$\begin{aligned} \alpha_m(v) &= \frac{(v + 40 \text{ mV})/10 \text{ mV}}{1 - \exp [-(v + 40 \text{ mV})/10 \text{ mV}] } , \\ \beta_m(v) &= 4 \exp [-(v + 65 \text{ mV})/18 \text{ mV}] , \end{aligned}$$

and:

$$\alpha_h(v) = \frac{7}{100} \exp [-(v + 65 \text{ mV})/20 \text{ mV}] .$$

See Fig. 7.2. The reaction-rate constants are $k_1 = 6/25 \text{ ms}^{-1}$, $k_2 = 2/5 \text{ ms}^{-1}$, and $k_3 = 3/2 \text{ ms}^{-1}$.

We developed new spectral decomposition methods from the meromorphic functional calculus [60, 198] to circumvent the inherent ill-conditioning in ion channel dynamics. Using these we can analytically calculate most, if not all, properties—e.g., dynamics, expected current, thermodynamics, information measures, and the like—about this model directly from the transition dynamic.

Since we are interested in thermodynamics, though, let us focus on determining the steady-state surprisals of the conformational states. For any persistent environmental input, the effective energies of the various conformational states are determined by their relative stationary occupation probability; according to Eq. (7.2), $\pi_x(s) = e^{-\phi(x,s)}$. The stationary distribution π_v induced by persistent v is the left eigenvector of $T_\tau^{(\mathcal{S} \rightarrow \mathcal{S}|v)}$ associated with the eigenvalue of unity. Equivalently, and more convenient in this case, π_v is the left eigenvector of $G^{(\mathcal{S} \rightarrow \mathcal{S}|v)}$ associated with the eigenvalue of zero. Via $\langle \pi_v | G^{(\mathcal{S} \rightarrow \mathcal{S}|v)} = \vec{0}$, we find that the steady state distribution for any persistent v is:

$$\begin{aligned} \pi_v &\propto \left(\frac{\beta_m}{3\alpha_m}, 1, \frac{2\alpha_m+k_1}{2\beta_m}, \frac{\alpha_m}{2\beta_m} \left(\frac{2\alpha_m+k_1}{3\beta_m+k_3} \right), \frac{1}{\alpha_h} \left[k_1 + \frac{2\alpha_m+k_1}{2\beta_m} \left(k_2 + \frac{k_3\alpha_m}{3\beta_m+k_3} \right) \right] \right) \\ &\propto e^{-\phi(v)}, \end{aligned}$$

which immediately yields the steady-state surprisals for conformational states at a constant environmental input v . The steady-state surprisal is shown for each conformational state in Fig. 7.3, as a function of the voltage-clamped membrane potential v .

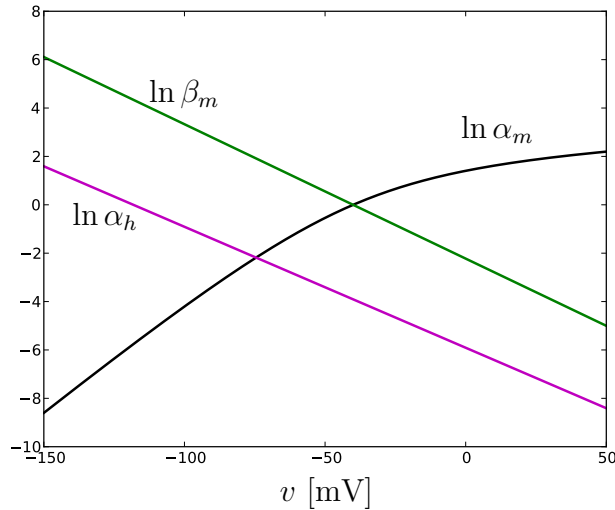


Figure 7.2: Markov transition parameter voltage dependencies: Plots of $\ln \alpha_m$, $\ln \beta_m$, and $\ln \alpha_h$. These plots show that at -100 mV, $\beta_m \gg \alpha_h \gg \alpha_m \approx 0$. At $+10$ mV, $\alpha_m \gg \beta_m \gg \alpha_h \approx 0$.

7.7.2 Ion channel (ir)reversibility

Recall that detailed balance is the condition that, for states a and b and environmental input x , $\Pr(b \xrightarrow{x} a) / \Pr(a \xrightarrow{x} b) = \pi_x(a) / \pi_x(b)$. Interestingly, in this biologically inspired

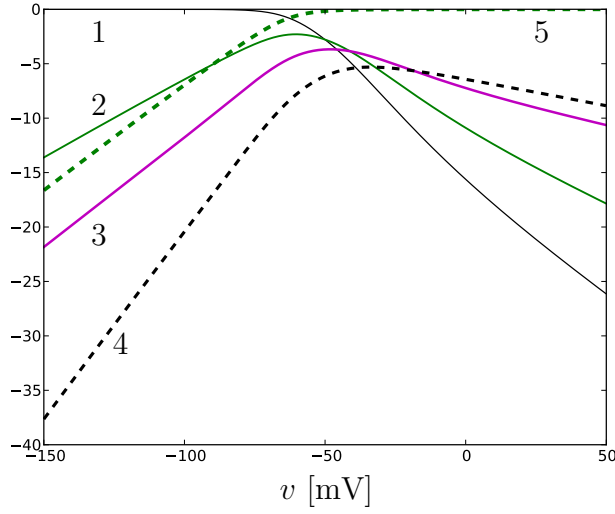


Figure 7.3: Steady-state distribution voltage dependence: Negative of the steady-state surprisals, $\ln \pi_v(s) = -\phi(v, s)$ for each conformational state. Each curve labeled by the state-number to which it corresponds. Note that -100 mV and $+10$ mV (relevant for later) are extremes in that $\pi_{v_a} \approx \delta_1$ and $\pi_{v_b} \approx \delta_5$.

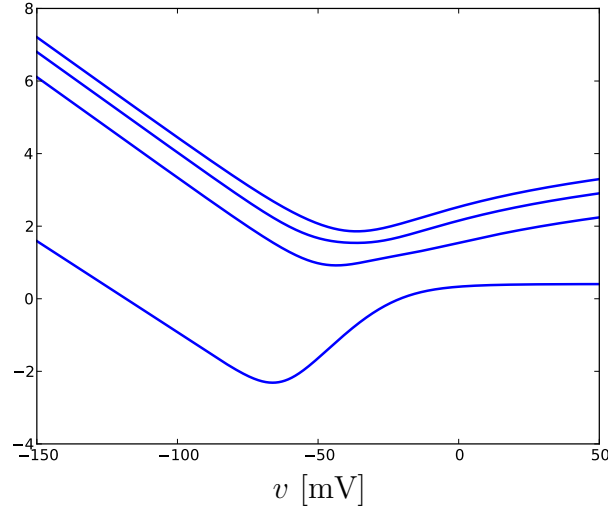


Figure 7.4: Modes of the state-to-state dynamic: All eigenvalues of G are real and nonpositive. There is a zero eigenvalue associated with stationarity and four negative eigenvalues associated with decay rates from the states. Plots of $\ln[-\lambda(v)]$ for $\lambda(v) \in \Lambda_{G(\mathbf{s} \rightarrow \mathbf{s}|v)}$. Smaller $\ln(-\lambda)$ corresponds to longer time-scales. The zero eigenvalue maps to $-\infty$.

model, detailed balance is satisfied by *several, but not all* of the state-transition pairs.

For example, for very small Δt :

$$\begin{aligned}\frac{\Pr(1 \xrightarrow{v} 2)}{\Pr(2 \xrightarrow{v} 1)} &= \frac{3\alpha_m}{\beta_m} \\ &= \frac{\pi_v(2)}{\pi_v(1)}.\end{aligned}$$

That is, this transition pair satisfies detailed balance. Hence, all transitions between these states are completely reversible: $\Psi(2|1, v) = \Psi(1|2, v) = \Psi(22121112|2, v) = 0$.

However, this does not hold for other transition pairs. Consider transitions between states 2 and 3:

$$\begin{aligned}\frac{\Pr(2 \xrightarrow{v} 3)}{\Pr(3 \xrightarrow{v} 2)} &= \frac{\alpha_m}{\beta_m} \\ &\neq \frac{\pi_v(3)}{\pi_v(2)} \\ &= \frac{\alpha_m + k_1/2}{\beta_m}.\end{aligned}$$

Most other transitions also violate detailed balance.

Since detailed balance does not hold for the effective dynamic, the theory developed above is essential to analyzing the sodium ion channel thermodynamics. Moreover, the fact that $G^{(\mathcal{S} \rightarrow \mathcal{S}|v)}$ has null entries that are nonzero for its transpose implies that paths involving these transitions will be infinitely irreversible: $\Psi = \infty$ for such paths as $\Delta t \rightarrow 0$; specifically, the transitions of $G^{(\mathcal{S} \rightarrow \mathcal{S}|v)}$ with rates k_1 and k_3 . Forbidden transitions are an extreme form of irreversibility that are nevertheless commonly observed for complex systems, as the ion channel so readily illustrates. In it, the asymmetry in allowed transitions can be traced to different *mechanisms* facilitating different paths through the state space. Whether the irreversibility is truly infinite or just practically infinite does not matter much for the excess thermodynamics, although it will of course affect the calculated distribution of Ψ . (Conventional, linear algebraic methods are inadequate to overcome these technical challenges. The spectral decomposition methods, mentioned above, are required.)

7.7.3 Step Function Drive

With this understanding of ion channel NESSs, let's now turn to the thermodynamics induced by driving between them. We first consider the particular voltage protocol of $v_a \equiv -100$ mV for all time except a $v_b \equiv 10$ mV pulse for 5 ms starting at $t = 0$. This is an example of continuous-time dynamics and deterministic driving. The system begins equilibrated with the static environmental drive $v_a = -100$ mV. The initial distribution over \mathcal{S} is thus $\mu_0 = \pi_{v_a}$, where π_{v_a} is the left eigenvector of $G^{(\mathcal{S} \rightarrow \mathcal{S}|v_a)}$ associated with the eigenvalue of zero.

During an epoch of fixed $v = V$, the net transition dynamic after τ ms becomes:

$$T_\tau^{(\mathcal{S} \rightarrow \mathcal{S}|v=V)} = e^{\tau G^{(\mathcal{S} \rightarrow \mathcal{S}|v=V)}} .$$

Therefore, the distribution over states induced by the driving protocol is:

$$\langle \mu_t | = \begin{cases} \langle \pi_{v_a} | & \text{for } t \leq 0 \\ \langle \pi_{v_a} | e^{tG_b} & \text{for } 0 < t \leq 5 \text{ ms} , \\ \langle \pi_{v_a} | e^{5G_b} e^{(t-5)G_a} & \text{for } t > 5 \text{ ms} \end{cases} , \quad (7.47)$$

where, for brevity, we defined: $G_a \equiv G^{(\mathcal{S} \rightarrow \mathcal{S}|v_a)}$ and $G_b \equiv G^{(\mathcal{S} \rightarrow \mathcal{S}|v_b)}$. These are especially useful when expressing the rate matrix via its spectral decomposition, using the methods of Refs. [60, 198]. Besides the zero eigenvalue, there are only four other eigenvalues of G that are determined via $\det(\lambda I - G) = 0$.

Figure 7.4 shows G 's eigenvalues as a function of v , which indicates the voltage-dependent timescales of probability decay for modes of occupation probability. The associated decay rates play a prominent role in Fig. 7.5, which shows the time-dependent distribution induced over states by the 5 ms voltage pulse driving protocol.

Having the distribution over states at all times is powerful knowledge. For example, since the current through a single channel is binary—either 0 or $I_0(v)$ —and since current only flows in the open conformation, the expected current through the channel is $I_0(v) = g_0[v - V_{Na}]$ times the expectation value of being in the open state:

$$\langle I(t) \rangle = g_0 [v(t) - V_{Na}] \langle \mu(t) | \delta_{\text{open}} \rangle ,$$

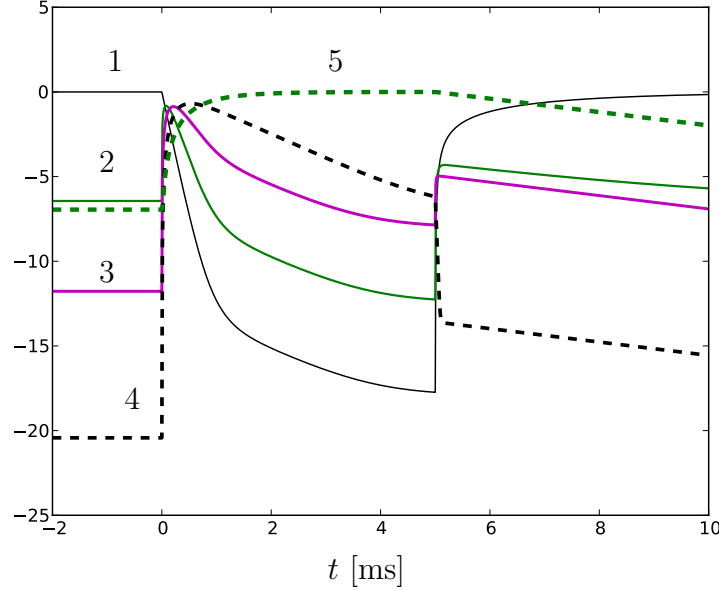


Figure 7.5: Na^+ ion channel NESS transitions: temporal evolution of the distribution μ_t of ion-channel conformational states induced by a deterministic 5 ms voltage pulse is shown via plots of $\ln \langle \mu_t | s \rangle$ for all $s \in \mathcal{S}$. Curves are labeled by the state-number to which each component corresponds.

where $\delta_{\text{open}} = (0, 0, 0, 1, 0)$, g_0 is the conductance of an open Na^+ channel, and $V_{\text{Na}} = \frac{k_B T}{e^+} \ln \frac{[\text{Na}^+]_{\text{out}}}{[\text{Na}^+]_{\text{in}}} \approx 90$ mV is the Nernst potential for sodium in a typical mammalian neuron [66]. To be clear $\langle I(t) \rangle$ is what would be observed from an ensemble of channels in a local patch of cell membrane experiencing the same driving. The current produced from the Markovian model appears to be more realistic than what would be expected from the Hodgkin–Huxley model [66]. Moreover, using our spectral-decomposition methods for functions of a Markov chain [60, 198], this current can now be obtained in closed-form.

Let us start the thermodynamic investigation by considering excess work W_{ex} . With $\tau = N\Delta t$, we take the limit of $\Delta t \rightarrow 0$ while keeping the product $N\Delta t = \tau$ constant. Then the expected excess work per $k_B T$, from time t_0 to time $t_0 + \tau$, is:

$$\beta \langle W_{\text{ex}} \rangle = \int_{t_0}^{t_0 + \tau} \langle \mu_t | d\phi_{v(t)} / dt \rangle dt .$$

However, it should be clear that, for this stepped voltage protocol, excess work is *only performed on this system at the very onset and subsequently at the end* of the step driving. Indeed, this is the only time that the driving $v(t)$ changes and, thus, the only time that

the state-dependent rate of work $d|\phi_{v(t)}\rangle/dt$ is nonzero. As we let $\Delta t \rightarrow 0$, the expected excess work (divided by $k_B T$) near the onset of driving becomes a step function with height:

$$\begin{aligned} & \lim_{\epsilon \rightarrow 0^+} \langle \Omega(t = \epsilon) - \Omega(t = -\epsilon) \rangle \\ &= \sum_{s \in \mathcal{S}} \langle \boldsymbol{\pi}_{\mathbf{v}_a} | s \rangle [\phi(10\text{mV}, s) - \phi(-100\text{mV}, s)] , \end{aligned}$$

where $\langle \boldsymbol{\pi}_{\mathbf{v}_a} | s \rangle = \pi_{-100 \text{ mV}}(s)$.

Indeed, for this singular event, the full distribution of work performed can be given according to the probabilities that the system was in a particular state when the driving was applied. For $0 < t < 5 \text{ ms}$, the probability density function for βW_{ex} is:

$$\begin{aligned} p(\Omega) &= \sum_{s \in \mathcal{S}} \pi_{-100 \text{ mV}}(s) \\ &\quad \times \delta\left(\Omega - [\phi(10\text{mV}, s) - \phi(-100\text{mV}, s)]\right) , \end{aligned}$$

where $\delta(\cdot)$ here is the Dirac delta function. For $t > 5 \text{ ms}$, the full excess environmental entropy production (EEEP) probability density function (pdf) is:

$$\begin{aligned} p(\Omega) &= \sum_{s, s' \in \mathcal{S}} \langle \boldsymbol{\pi}_{\mathbf{v}_a} | s \rangle \langle s | e^{5G_b} | s' \rangle \\ &\quad \times \delta\left(\Omega - [\phi(v_b, s) - \phi(v_a, s)] - [\phi(v_a, s') - \phi(v_b, s')]\right). \end{aligned}$$

From the Dirac delta function's argument and the sum over s and s' , it is clear that every nonzero-probability EEEP value Ω also has a nonzero probability for the negative $-\Omega$ of that EEEP value.

For the time-symmetric 5 ms voltage-pulse driving, Eq. (7.24) tells us that $\Pr(\Omega, \Psi)/\Pr(-\Omega, -\Psi) = e^\Psi e^\Omega$. Since there are infinitely many Ψ values to account for, we do not plot the joint distribution explicitly. However, we can appreciate the necessity of the relationship by comparing it to the naive CFT interpretation that, for this case, suggests $\Pr(\Omega)/\Pr(-\Omega) = e^\Omega$. Figure 7.6 compares these by plotting $e^{-\Omega} \Pr(\Omega)/\Pr(-\Omega)$.

Allowed values of the excess work that do *not* lie on $e^{-\Omega} \Pr(\Omega)/\Pr(-\Omega) = 1$ demonstrate deviations from the naive CFT interpretation. Since the constant-voltage steady

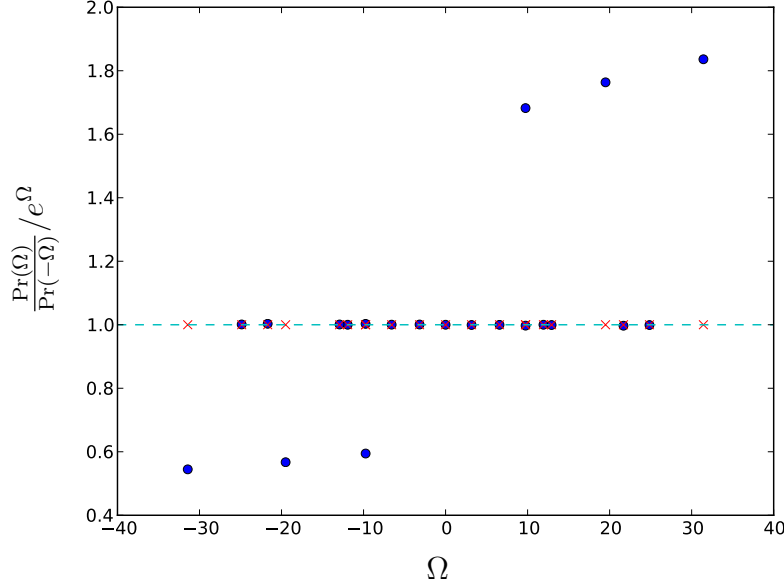


Figure 7.6: Deviations from and agreement with the Crooks Fluctuation Theorem for nondetailed balance dynamics: Exact calculation of $\frac{\text{Pr}(\Omega)}{\text{Pr}(-\Omega)}/e^\Omega$ (blue dots) for all allowed values of $\Omega = \beta W_{\text{ex}}$ during the pulse drive. Since the system starts in equilibrium and since the driving is time-symmetric, a naive CFT interpretation suggests that all values lie at unity, dashed line (blue) and marked with a \times (red) wherever an allowed excess work value appears. Interestingly, many of the allowed work values *do* still fall on or very near unity. Absent detailed balance, though, Eq. (7.23) must be used to account for the actual distribution of excess environmental entropy production and path irreversibilities that, in addition to all other values, yields the six deviant markings (blue dots) above and below unity.

states are nonequilibrium and, thus, not microscopically reversible—i.e., $\Psi \neq 0$ for some state paths—the naive CFT interpretation cannot be true despite the time-symmetric driving. Perhaps the most surprising feature in Fig. 7.6 is that many of the probability ratios still *do* (almost) fall on the naive CFT line at unity. In part, this is due to a subset of the cycles in the NESS dynamic obeying detailed balance. Another contributing factor is that longer durations τ of fixed v induces a *net* dynamic $e^{\tau G}$ that *approaches* a detailed-balanced dynamic. That the values in Fig. 7.6 are sensible can be verified by checking the ratio of the joint probabilities $\langle \pi_{v_a} | s \rangle \langle s | e^{5G_b} | s' \rangle$ to the value of the joint probability with s and s' swapped.

In stark contrast to the instantaneous work contribution, the system’s excess heat Q_{ex} unfolds over time, exhibiting a rich structure governed by the trajectories through the

conformational state-space. The expected excess heat per $k_B T$ is:

$$\beta \langle Q_{\text{ex}} \rangle = \int_{t_0}^{t_0+\tau} \langle \dot{\mu} | \phi_{v(t)} \rangle dt . \quad (7.48)$$

over a duration τ , if starting at time t_0 . Since $v(t)$ is constant except at the two instants of change, the integral is easily solved exactly using the fundamental theorem of calculus and Eq. (7.47).

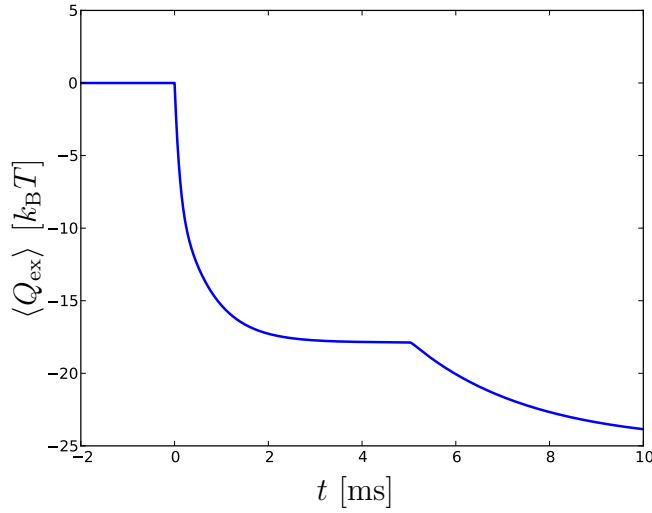


Figure 7.7: Excess heat Q_{ex} per ion channel for an ensemble of ion channels embedded in a local patch of cell membrane; in units of $k_B T \approx 26$ meV. The two bouts of relaxation correspond to the ion channel adapting to sudden changes in voltage across the cell membrane.

For $t_0 < t_0 + \tau < 0$, the system is in the initial steady-state and has a time-independent heat and so a constant excess heat rate that vanishes:

$$\frac{\langle Q_{\text{ex}} \rangle}{\tau} = 0 .$$

Figure 7.7 shows the expected excess heat $\langle Q_{\text{ex}} \rangle$ over the course of the voltage-drive protocol. The steady-state average rate of excess heat production within any steady state is necessarily zero. However, the channel macromolecule responds to changes in the environment via conformational changes and corresponding heat productions that unfold on the timescale of milliseconds⁴. Notably, the expected heat is on the order of tens of

⁴The characteristic timescale is actually the net result of a combination of timescales from the inverse eigenvalues of G . Of necessity, these are the same timescales that determine the relaxation of the state distribution.

$k_{\text{B}}T$, which for mammalian neurons is $k_{\text{B}}T = 1/\beta \approx 26$ meV.

For $0 = t_0 < t_0 + \tau < 5$ ms, the system has a time-varying heat production as it synchronizes to the new nonequilibrium steady state. From Eq. (7.47), we know that $\langle \mu_t | = \langle \pi_{v_a} | e^{tG_b}$ in this case. At the same time, the steady-state-surprisal $|\phi_{v(t)}\rangle$ is time-independent during this epoch since v is temporarily fixed at $v_b = 10$ mV.

Again, as seen in Fig. 7.7, the expected excess heat drops for several milliseconds after the final voltage switch at 5 ms, as the ion channel re-adapts to its original steady state. For this second bout of relaxation the adaptation is slower since, in accordance with Fig. 7.4, the slowest timescale at $v_a = -100$ mV is slower than the slowest timescale at $v_b = 10$ mV.

Overall, we see that the excess thermodynamic quantities are well behaved and accessible: without needing to know the background biological upkeep of the Na^+ ion channel, we can access and control coarse degrees of freedom of the channel macromolecule via modulating the voltage across the cell membrane. Moreover, for the Na^+ channel, state-measurement and feedback on the timescale of milliseconds would allow significant alterations of heat and entropy production. Fortuitously, this suggests an accessible platform for laboratory experimentation. Next, we comment briefly on an intrinsic type of measurement and feedback that happens in vivo every moment.

7.7.4 Intrinsic feedback

Having come this far, we close illustrating the thermodynamics of NESS transitions with a final application. In a biologically active (in vivo) neuron, the input membrane voltage at each time depends on integrated current—a functional of the state distribution—up to that time. Our relations for modified integral fluctuation relations describe the thermodynamic agency of Na^+ channels in vivo, whereas conventional fluctuation relations fall short. Although there is certainly feedback in vivo, it is not the “feedback control” discussed recently. Importantly, no “outsider” forces the feedback; the feedback is intrinsic—woven into the system–environment joint dynamic. We leave a thorough investigation of the thermodynamics of intrinsic feedback to elsewhere. The success here, however, already suggests investigating other natural systems with intrinsic feedback—in joint nonlinear

dynamics and complex networks—to test the new fluctuation theorems and computational methods in a broader class of interacting complex nonequilibrium systems.

7.8 Discussion

In light of our refined detailed fluctuation theorem Eq. (7.23) for nondetailed-balanced dynamics, we referred to the common belief in Eq. (7.27) as the naive CFT interpretation since it appeals to a nonphysical conjugate dynamics, as in Eq. (7.25). Similarly, we referred to failures of Eq. (7.26) as CFT violations. Nonetheless, with proper interpretation using the unphysical conjugate dynamics, Eq. (7.25) is mathematically correct even without detailed balance and can be a useful device for establishing integral FTs.

The CFT is often misinterpreted, though, despite receiving widespread attention. For example, Ref. [77] appealed to the naive CFT interpretation in the case of nondetailed balance dynamics of self-replication. However, as we showed, such applications are not valid. So, the statistical physics of self-replication either depends on an assumption of detailed balance or deserves a generalization. We believe the latter should be straightforward using our results, new spectral methods, and a derivation paralleling Ref. [77].

We hope that our nonintegral FTs—especially Eq. (7.23) that constrains the joint distribution of excess and housekeeping entropies—will provide better physical intuition for the structure of effective dynamics outside detailed balance. Path irreversibility clearly plays a prominent role. Although the preceding introduced a unifying framework, certain subclasses of path irreversibility have already been proposed recently.

In certain applications, for example, path irreversibility is governed by differences in chemical potential. In such cases, the irreversibility is quantitatively related to cycle affinities; see, for example, Ref. [134]. It essentially discovered a special case of the results developed here specifically applicable to the interesting example of a kinesin motor protein.

Recently, Ref. [174] elaborated on one type of irreversibility, called *absolute irreversibility*, that at first appears to constitute an extreme contribution to the total path irreversibility Ψ . This indeed is one interpretation, but not the full story. On closer examination, its result appears to coincide most directly with Eq. (7.33) which must be

used when starting in a nonsteady-state, rather than with a violation of Eq. (7.34) which is simply inapplicable when starting in a nonsteady-state. We reinterpret that work as a testament to the importance of the nonsteady-state contribution to free energy change, $\beta^{-1}\Delta(\gamma)$. Explicitly:

$$\begin{aligned} \langle e^{-\beta W_{\text{diss}}} \rangle_{\text{Pr}(s_{0:N}|\mu_F, \mathbf{x})} &= 1 \\ \rightarrow \langle \Omega \rangle &\geq \Delta \langle \gamma \rangle . \end{aligned}$$

This captures, for example, the entropy change associated with free expansion. From our viewpoint, however, any absolute irreversibility is only one extreme of the broader generalization introduced above to explore the consequences of irreversibility and nonsteady-state additions to free energy.

To frame our results in yet another way, we note that the “feedback control” imposed by an experimenter on an otherwise detailed-balanced system is a rather limited form of CFT violation. Yet it appears to be the only one having gained much recognition. This is odd. Hysteresis, to take one example, common in paradigmatically complex physical systems, provides a more physical manifestation of CFT violation. Even this is still a relatively tame deviation from detailed balance. Living systems are the true flagship of complex physical agents with intrinsic computational feedback across many levels of their organization. Our fluctuation relations describe all of these aspects, together.

In particular, they suggest how a system’s intrinsic model of its environment, together with an action policy that leverages knowledge captured in the model to control the environment, allows the system to play the survival game to its thermodynamic advantage. For example, an agent can use information about the environment to increase its nonsteady-state free energy and perform useful work—a phenomenon that is not only reminiscent of living beings, but also comes very near to defining them.

We hope that our results and methods stimulate investigating the excess thermodynamics of systems with intrinsic feedback—from designed “toy demons” to complex biological molecules affected by and simultaneously affecting their environments. Several biological examples that suggest themselves include kinesin motors [86], drug-operated channels [43], and dynamic synapses [135], just to name a few.

7.9 Conclusion

We presented generalized fluctuation theorems for complex nonequilibrium systems driven between NESSs. In addition to the detailed FTs that constrain joint distributions of excess and housekeeping quantities, we introduced integral fluctuation theorems in the presence of an auxiliary variable. The auxiliary variable need not be measurement nor any other meddling of an outsider. Due to this, it generalizes the theory of “feedback control” to the setting of arbitrary intrinsic feedback between system and environment.

A sequel to the above derives exact closed-form expressions for the moments of excess heat and excess work when the joint system–environment dynamic is governed by a (finite or countably infinite) discrete- or continuous-time hidden Markov model. A joint system can always be modeled as a joint hidden Markov model—at least as an approximation to the true joint dynamics. For this reason, our exact results should provide broadly applicable tools. The latter have particular theoretical advantage in giving access to what occurs in transient and asymptotic dynamics of excess thermodynamic quantities atop NESSs.

In summary, the traditional laws of thermodynamics are largely preserved for the renormalized “excess” thermodynamic quantities that arise naturally when considering nondetailed-balanced complex systems. However, the laws must be modified by the entropic contribution of path irreversibility. We noted that the latter turns out to be equivalent to steady-state thermodynamics’ housekeeping entropy.

Our relations still hold for excursions between equilibrium steady states, but we then have the simplification that $\Psi = \beta Q_{\text{hk}} = 0$. Consistently, equilibrium thermodynamics is a reduction of the theory of excess thermodynamic quantities with no housekeeping terms—when all paths are microscopically reversible.

Layers of emergence, typical of the biological world [227, Fig. 6], beg renormalization in terms of a hierarchy of housekeeping backgrounds [2]. The opportunity offered up by emergent levels of novel organization is a new richness in nondetailed-balanced effective dynamics—dynamics and structure that can be exploited by intelligent thermodynamic agency [29, 31]. We consider the thermodynamics of agency in a sequel, analyzing a simple

autonomous agent that harvests energy by leveraging hidden correlations in a fluctuating environment.

We leave the development for now, but with an encouraging lesson: Even in nonstationary nonequilibrium, there is excess thermodynamic structure at any level of observation that we can access, control, and harness.

7.10 Appendices

7.10.1 Appendix A: Extension to Non-Markovian Instantaneous Dynamics

Commonly, theoretical developments assume state-to-state transitions are instantaneously Markovian given the input. This assumption works well for many cases, but fails in others with strong coupling between system and environment. Fortunately, we can straightforwardly generalize the results of stochastic thermodynamics by considering a system's observable states to be functions of latent variables \mathcal{R} . The goal in the following is to highlight the necessary changes, so that it should be relatively direct to adapt our derivations to the non-Markovian dynamics.

7.10.1.1 Latent states, system states, and their many distributions

Even with constant environmental input, the dynamic over a system's states need not obey detailed balance nor exhibit any finite Markov order. We assume that the classical observed states \mathcal{S} are functions $f : \mathcal{R} \rightarrow \mathcal{S}$ of a latent Markov chain. We also assume that the stochastic transitions among latent states are determined by the current environmental input $x \in \mathcal{X}$, which can depend arbitrarily on all previous input and system-state history. The Perron–Frobenius theorem guarantees that there is a stationary distribution over latent states associated with each fixed input x ; the function of the Markov chain maps this stationary distribution over latent states into the stationary distribution over system states. These are the stationary distributions associated with system NESSs.

We assume too that the \mathcal{R} -to- \mathcal{R} transitions are Markovian given the input. However, different inputs induce different Markov chains over the latent states. This can be described by a (possibly infinite) set of input-conditioned transition matrices over the latent state

set \mathcal{R} : $\{\mathsf{T}^{(\mathcal{R} \rightarrow \mathcal{R}|x)}\}_{x \in \mathcal{X}}$, where $\mathsf{T}_{i,j}^{(\mathcal{R} \rightarrow \mathcal{R}|x)} = \Pr(\mathcal{R}_t = r^j | \mathcal{R}_{t-1} = r^i, X_t = x)$. Probabilities regarding actual state paths can be obtained from the latent-state-to-state transition dynamic together with the observable-state projectors, which we now define.

We denote distributions over the latent states as bold Greek symbols, such as $\boldsymbol{\mu}$. As in the main text, it is convenient to cast $\boldsymbol{\mu}$ as a row-vector, in which case it appears as the bra $\langle \boldsymbol{\mu} |$. The distribution over latent states \mathcal{R} implies a distinct distribution over observable states \mathcal{S} . A sequence of driving inputs updates the distribution: $\boldsymbol{\mu}_{t+n}(\boldsymbol{\mu}_t, x_{t:t+n})$. In particular:

$$\begin{aligned} \langle \boldsymbol{\mu}_{t+n} | &= \langle \boldsymbol{\mu}_t | \mathsf{T}^{(\mathcal{R} \rightarrow \mathcal{R}|x_{t:t+n})} \\ &= \langle \boldsymbol{\mu}_t | \mathsf{T}^{(\mathcal{R} \rightarrow \mathcal{R}|x_t)} \mathsf{T}^{(\mathcal{R} \rightarrow \mathcal{R}|x_{t+1})} \dots \mathsf{T}^{(\mathcal{R} \rightarrow \mathcal{R}|x_{t+n-1})} . \end{aligned}$$

(Recall that time indexing is denoted by subscript ranges $n : m$ that are left-inclusive and right-exclusive.) An infinite driving history \vec{x} induces a distribution $\boldsymbol{\mu}(\vec{x})$ over the state space, and $\boldsymbol{\pi}_x$ is the specific distribution induced by tireless repetition of the single environmental drive x . This is the so-called “equilibrium distribution” associated with equilibrating with the environmental drive x . Explicitly:

$$\langle \boldsymbol{\pi}_x | = \lim_{n \rightarrow \infty} \langle \boldsymbol{\mu}_0 | \left(\mathsf{T}^{(\mathcal{R} \rightarrow \mathcal{R}|x)} \right)^n .$$

Usefully, $\boldsymbol{\pi}_x$ can also be found as the left eigenvector of $\mathsf{T}^{(\mathcal{R} \rightarrow \mathcal{R}|x)}$ associated with the eigenvalue of unity:

$$\langle \boldsymbol{\pi}_x | = \langle \boldsymbol{\pi}_x | \mathsf{T}^{(\mathcal{R} \rightarrow \mathcal{R}|x)} . \quad (7.49)$$

The canonical equilibrium probabilities are this vector’s projection onto observable states: $\pi_x(s) = \langle \boldsymbol{\pi}_x | s \rangle$, where $|s\rangle = |\delta_{s,f(r)}\rangle$ has a vector-representation in the latent-state basis with elements of all 0s except 1s where the latent state maps to the observable state s .

Assuming latent-state-to-state transitions are Markovian allows the distribution $\boldsymbol{\mu}$ over these latent states to summarize the causal relevance of the entire driving history.

7.10.1.2 Implications

A semi-infinite history induces a particular distribution over system latent states and implies another particular distribution over its observable states. This can be usefully

recast in terms of the “start” (or initial) distribution $\boldsymbol{\mu}_0$ induced by the path $x_{-\infty:1}$ and the driving history $x_{1:t+1}$ since then, giving the entropy of the induced state distribution:

$$\begin{aligned} h^{(s|\boldsymbol{\mu}_0, x_{1:t+1})} &= -\ln \Pr(\mathcal{S}_t = s | \boldsymbol{\mu}_0, x_{1:t+1}) \\ &= -\ln \langle \boldsymbol{\mu}_0 | \mathsf{T}^{(\mathcal{R} \rightarrow \mathcal{R} | x_{1:t+1})} | s \rangle . \end{aligned}$$

Or, employing the new distribution and the driving history since then, the path entropy (functional of state and driving history) can be expressed simply in terms of the *current* distribution over latent states and the candidate observable state s :

$$\begin{aligned} h^{(s|\boldsymbol{\mu})} &= -\ln \Pr(\mathcal{S}_t = s | \mathcal{R}_t \sim \boldsymbol{\mu}) \\ &= -\ln \langle \boldsymbol{\mu} | s \rangle . \end{aligned}$$

Averaging the path-conditional state entropy over observable states again gives a genuine input-conditioned Shannon state entropy:

$$\langle h^{(s_t | \overleftarrow{x}_t)} \rangle_{\Pr(s_t | \overleftarrow{x}_t)} = \mathsf{H}[\mathcal{S}_t | \overleftarrow{X}_t = \overleftarrow{x}_t] .$$

It is again easy to show that the state-averaged path entropy $k_B \mathsf{H}[\mathcal{S}_t | \overleftarrow{x}_t]$ is an extension of the system’s steady-state nonequilibrium entropy. In steady-state, the state-averaged path entropy reduces to:

$$\begin{aligned} k_B \mathsf{H}[\mathcal{S}_t | \overleftarrow{X}_t = \dots xxx] &= -k_B \mathsf{H}[\mathcal{S}_t | \mathcal{R}_t \sim \boldsymbol{\pi}_x] \\ &= -k_B \sum_{s \in \mathcal{S}} \pi_x(s) \ln \pi_x(s) \\ &= S_{\text{ss}}(x) . \end{aligned}$$

The *nonsteady-state addition to free energy* is:

$$\beta^{-1} \gamma(s | \boldsymbol{\mu}, x) \equiv \beta^{-1} \ln \frac{\Pr(\mathcal{S}_t = s | \mathcal{R}_{t-1} \sim \boldsymbol{\mu}, X_t = x)}{\pi_x(s)} .$$

Averaging over observable states this becomes the relative entropy:

$$\langle \gamma(s | \boldsymbol{\mu}, x) \rangle = D_{\text{KL}} [\Pr(\mathcal{S}_t | \mathcal{R}_{t-1} \sim \boldsymbol{\mu}, X_t = x) || \boldsymbol{\pi}_x] ,$$

which is always nonnegative.

Using this setup and decomposing:

$$\frac{\Pr(\mathcal{S}_{0:N} = s^0 \mathbf{s} | \mathcal{R}_{-1} \sim \boldsymbol{\mu}_F, X_{0:N} = x^0 \mathbf{x})}{\Pr(\mathcal{S}_{0:N} = s^{N-1} \mathbf{s}_{\leftarrow}^R | \mathcal{R}_{-1} \sim \boldsymbol{\mu}_R, X_{0:N} = x^N \mathbf{x}^R)}$$

in analogy with Eq. (7.21), it is straightforward to extend the remaining results of the main body to the setting in which observed states are functions of a Markov chain. Notably, the path dependencies pick up new contributions from non-Markovity. Also, knowledge of distributions over latent states provides a thermodynamic advantage to Maxwellian Demons.

7.10.2 Appendix B: Integral fluctuation theorems with auxiliary variables

Recall that we quantify how much the auxiliary variable independently informs the state sequence via the nonaveraged conditional mutual information:

$$\begin{aligned} i[\vec{s}; \vec{y} | \vec{x}, \boldsymbol{\mu}_F] &\equiv \ln \frac{\Pr(\vec{s}, \vec{y} | \vec{x}, \boldsymbol{\mu}_F)}{\Pr(\vec{y} | \vec{x}, \boldsymbol{\mu}_F) \Pr(\vec{s} | \vec{x}, \boldsymbol{\mu}_F)} \\ &= \ln \frac{\Pr(\vec{s}, \vec{y}, \vec{x} | \boldsymbol{\mu}_F)}{\Pr(\vec{y}, \vec{x} | \boldsymbol{\mu}_F) \Pr(\vec{s} | \vec{x}, \boldsymbol{\mu}_F)} . \end{aligned}$$

Note that averaging over the input, state, and auxiliary sequences gives the familiar conditional mutual information:

$$I[\mathcal{S}_{0:N}; Y_{0:N} | X_{0:N}, \boldsymbol{\mu}_F] = \langle i[\vec{s}; \vec{y} | \vec{x}, \boldsymbol{\mu}_F] \rangle_{\Pr(x_{0:N}, s_{0:N}, y_{0:N} | \boldsymbol{\mu}_F)} .$$

(Averaging over distributions is the same as being given the distribution, since the distribution over distributions is assumed to be peaked at $\boldsymbol{\mu}_F$.)

Noting that:

$$\begin{aligned} e^{\beta W_{\text{diss}} + i(\vec{s}; \vec{y} | \vec{x}, \boldsymbol{\mu}_F) + \Psi} &= e^{\Omega + i(\vec{s}; \vec{y} | \vec{x}, \boldsymbol{\mu}_F) + \Psi + (\gamma_F - \gamma_R)} \\ &= \frac{\Pr(\vec{s}, \vec{y}, \vec{x} | \boldsymbol{\mu}_F)}{\Pr(\vec{y}, \vec{x} | \boldsymbol{\mu}_F) \Pr(s^{N-1} \mathbf{s}_{\leftarrow}^R | \mathbf{x}^R x^0, \boldsymbol{\mu}_R)} \\ &= \frac{\Pr(\vec{s}, \vec{y}, \vec{x} | \boldsymbol{\mu}_F)}{\Pr(\vec{y}, \vec{x} | \boldsymbol{\mu}_F) \Pr(\vec{s} | \vec{x}, \boldsymbol{\mu}_R)} , \end{aligned}$$

where $\boldsymbol{\mu}_R = \boldsymbol{\mu}(\boldsymbol{\mu}_F, \vec{x})$, we have the integral fluctuation theorem (IFT):

$$\begin{aligned}
& \left\langle e^{-\beta W_{\text{diss}} - i(\vec{s}; \vec{y} | \vec{x}, \boldsymbol{\mu}_F) - \Psi} \right\rangle_{\text{Pr}(\vec{s}, \vec{y}, \vec{x} | \boldsymbol{\mu}_F)} \\
&= \sum_{\vec{x}, \vec{s}, \vec{y}} \text{Pr}(\vec{s}, \vec{y}, \vec{x} | \boldsymbol{\mu}_F) \frac{\text{Pr}(\vec{y}, \vec{x} | \boldsymbol{\mu}_F) \text{Pr}(\overleftarrow{s} | \overleftarrow{x}, \boldsymbol{\mu}_R)}{\text{Pr}(\vec{s}, \vec{y}, \vec{x} | \boldsymbol{\mu}_F)} \\
&= \sum_{\vec{x}, \vec{s}, \vec{y}} \text{Pr}(\vec{y}, \vec{x} | \boldsymbol{\mu}_F) \text{Pr}(\overleftarrow{s} | \overleftarrow{x}, \boldsymbol{\mu}_R) \\
&= \sum_{\vec{x}, \vec{y}} \text{Pr}(\vec{y}, \vec{x} | \boldsymbol{\mu}_F) \sum_{\overleftarrow{s}} \text{Pr}(\overleftarrow{s} | \overleftarrow{x}, \boldsymbol{\mu}_R) \\
&= \sum_{\vec{x}, \vec{y}} \text{Pr}(\vec{y}, \vec{x} | \boldsymbol{\mu}_F) \\
&= 1 .
\end{aligned}$$

Notably, this relation holds arbitrarily far from equilibrium and allows for the starting and ending distributions to both be nonsteady-state. It is tempting to conclude that the revised Second Law of Thermodynamics should read:

$$\langle W_{\text{diss}} \rangle \geq -k_B T \text{I}[\vec{\mathcal{S}}; \vec{Y} | \vec{X}, \boldsymbol{\mu}_F] - \langle Q_{\text{hk}} \rangle , \quad (7.50)$$

which includes the effects of both irreversibility and conditional mutual information between state-sequence and auxiliary sequence, given input-sequence. However, we expect that $\langle Q_{\text{hk}} \rangle > 0$, so Eq. (7.50) is not the strongest bound derivable. Dropping Ψ from the IFT still yields a true equality, but the derivation runs differently since it depends on the normalization of the conjugate dynamic. Although IFTs with Ψ may be useful for other reasons, it is the non- Ψ IFTs that seems to yield the tighter bound for the revised Second Laws of information thermodynamics without detailed balance.

Chapter 8

Strongly Coupled Systems, Exact Excess, and Renormalized Housekeeping

8.1 Chapter Overview

When a small but complex thermodynamic agent (such as a biological macromolecule) strongly couples with a structured stochastic environment, the joint interaction produces an emergent dynamic of non-Markovian transitions among the agent's conformational states. With feedback and constant driving, the agent never has the chance to settle even into one of its nonequilibrium steady states that would be induced asymptotically by a particular environmental condition. On the surface, the situation seems intractable. Nevertheless, we show that moments of the excess work and excess heat distributions can be calculated exactly, directly from the spectrum and projection operators of the transition rate operator of the joint dynamic. The average excess heat at one level of description becomes the additional housekeeping heat to maintain the next level of description. This framework naturally accommodates two-way feedback, highlighting the role of information transduction in the stochastic thermodynamics of biologically relevant continuous-time systems.

8.2 Setup and Results

Let us consider a system-under-study (SUS) that transduces a dynamic environmental signal, such that a particular setting of the environmental control vector $\vec{v} \in \mathcal{V}$ induces a particular stochastic transition-rate dynamic $G^{(\mathcal{S} \rightarrow \mathcal{S}|\vec{v})}$ over the states \mathcal{S} of the SUS. The random variable for the state of the SUS at time t is \mathcal{S}_t , which can take on values $s \in \mathcal{S}$. Specifically, we will consider a representation of $G^{(\mathcal{S} \rightarrow \mathcal{S}|\vec{v})}$ where $G_{s,s'}^{(\mathcal{S} \rightarrow \mathcal{S}|\vec{v})} = \langle s | G^{(\mathcal{S} \rightarrow \mathcal{S}|\vec{v})} | s' \rangle$ gives the \vec{v} -dependent rate of transitioning from state s to state s' , given that the SUS is in state s , where $s, s' \in \mathcal{S}$.

Let μ be a probability distribution over the states \mathcal{S} of the SUS. Given the instantaneous environmental input $\vec{v} \in \mathcal{V}$, the instantaneous dynamic of the ensemble is thus given by:

$$\frac{d}{dt} \langle \mu | = \langle \mu | G^{(\mathcal{S} \rightarrow \mathcal{S}|\vec{v})}, \quad (8.1)$$

The environmental driving can be time dependent. Nevertheless if \vec{v} were constant, then the linearity of Eq. (8.1) guarantees that there is a set of instantaneous modes of decay to the control-dependent asymptotic probability distribution $\pi_{\vec{v}}$ over \mathcal{S} induced by fixing the control at \vec{v} for a sufficiently long time. I.e.:

$$\langle \pi_{\vec{v}} | G^{(\mathcal{S} \rightarrow \mathcal{S}|\vec{v})} = \langle 0_{G(\mathcal{S} \rightarrow \mathcal{S}|\vec{v})} | G^{(\mathcal{S} \rightarrow \mathcal{S}|\vec{v})} = \vec{0}. \quad (8.2)$$

In words, the stationary distribution $\pi_{\vec{v}}$ is the left eigenvector of $G^{(\mathcal{S} \rightarrow \mathcal{S}|\vec{v})}$ associated with its index-one eigenvalue of 0, and we assume that each fixed \vec{v} induces a single attractor. The steady-state distribution $\pi_{\vec{v}}$ can either be an equilibrium steady state or a nonequilibrium steady state; our formalism applies just as well in either case. Notably, the right eigenvector of $G^{(\mathcal{S} \rightarrow \mathcal{S}|\vec{v})}$ associated with the zero eigenvalue is a column vector of all ones denoted $|0_{G(\mathcal{S} \rightarrow \mathcal{S}|\vec{v})}\rangle = |\mathbf{1}\rangle$.

In the special case where \mathcal{S} is the set of energy eigenstates and $G^{(\mathcal{S} \rightarrow \mathcal{S}|\vec{v})}$ is a detail-balanced dynamic, $\pi_{\vec{v}}$ is simply a state of thermodynamic equilibrium (rather than a nonequilibrium steady state) and we expect that the equilibrium state has the familiar Boltzmann distribution over energies: $\pi_{\vec{v}}(s) = \langle \pi_{\vec{v}} | s \rangle = e^{-\beta(E_{\vec{v}}(s) - F_{\vec{v}})} \propto e^{-\beta E_{\vec{v}}(s)}$.

However, more generally, $\pi_{\vec{v}}$ will correspond to a nonequilibrium steady state, where the microscopic Hamiltonian may not be known and only effective energies can be inferred

from the dynamics. In such a case, it is useful to utilize the *steady-state surprisal* $\phi_{\vec{v}}$ of steady-state thermodynamics [108, 256, 163, 199]. The steady-state surprisal $\phi_{\vec{v}}$ associated with \vec{v} is defined via:

$$\pi_{\vec{v}} = e^{-\phi_{\vec{v}}} . \quad (8.3)$$

I.e., $\phi_{\vec{v}} = -\ln \pi_{\vec{v}}$. Or, elementwise: $\phi_{\vec{v}}(s) = -\ln \pi_{\vec{v}}(s)$ for each $s \in \mathcal{S}$. The nonequilibrium potential replaces the need for an effective Hamiltonian as it plays the role of the asymptotic surprisal of state occupation (‘surprisal’, sometimes referred to as ‘self-information’ is the precursor to entropy; Shannon entropy is obtained from averaging surprisal over states [46]) that would be induced under steady application of the effective Hamiltonian: $\phi_{\vec{v}} \sim \beta H_{\vec{v}}^{\text{eff}} + \ln Z_{\vec{v}}$. Note that in addition to the effective Hamiltonian, the normalization factor (which encodes the free energy in the case of equilibrium systems) is also subsumed by the steady-state surprisal.

Whatever the environmental control protocol is as a function of time, the change in steady-state surprisal breaks up into contributions of *excess heat* Q_{ex} and *excess work* W_{ex} [199]. In the isothermal setting, with inverse temperature $\beta = \frac{1}{k_B T}$, their ensemble averages can be found by:

$$\Delta \langle \phi_{\vec{v}} \rangle = \int \langle \dot{\phi}_{\vec{v}} \rangle dt = \underbrace{\int \langle \dot{\boldsymbol{\mu}} | \phi_{\vec{v}} \rangle dt}_{\langle \dot{\mathcal{U}}_{\text{ex}} \rangle = \beta \langle \dot{Q}_{\text{ex}} \rangle} + \underbrace{\int \langle \boldsymbol{\mu} | \dot{\phi}_{\vec{v}} \rangle dt}_{\langle \dot{\Omega}_{\text{ex}} \rangle = \beta \langle \dot{W}_{\text{ex}} \rangle} , \quad (8.4)$$

This accounting is analogous to how net change of internal energy breaks up into contributions of total heat and total work.

The excess heat is the change in $(k_B T \times)$ steady-state surprisal due to a relaxation of the system, with the effective energy levels held fixed. It is a heat in excess of the housekeeping heat (Q_{hk} , which is needed to maintain the steady state): $Q = Q_{\text{hk}} + Q_{\text{ex}}$. Over any interval τ during which the environmental input \vec{v} and temperature T are constant, the excess heat and average excess heat are simply $Q_{\text{ex}} = k_B T (\phi_{\vec{v}}(\mathcal{S}_{\tau}) - \phi_{\vec{v}}(\mathcal{S}_0))$ and $\langle Q_{\text{ex}} \rangle = k_B T (\langle \boldsymbol{\mu}_{\tau} | \phi_{\vec{v}} \rangle - \langle \boldsymbol{\mu}_0 | \phi_{\vec{v}} \rangle)$ respectively.

Excess work is the change in $(k_B T \times)$ steady-state surprisal due to a shift in the effective energy levels, while they remain occupied. Over any short interval $t \in [0, \tau]$ during which

the state \mathcal{S}_0 and temperature T are held fixed, the excess work and average excess work are simply $W_{\text{ex}} = k_{\text{B}}T(\phi_{\vec{v}(\tau)}(\mathcal{S}_0) - \phi_{\vec{v}(0)}(\mathcal{S}_0))$ and $\langle W_{\text{ex}} \rangle = k_{\text{B}}T(\langle \boldsymbol{\mu}_0 | \phi_{\vec{v}(\tau)} \rangle - \langle \boldsymbol{\mu}_0 | \phi_{\vec{v}(0)} \rangle)$ respectively. In considering processes that transition between equilibrium states, the excess work reduces to the work in excess of the free energy change: $W_{\text{ex}} \rightarrow W - (F_{\vec{v}(\tau)} - F_{\vec{v}(0)})$. The excess work is zero over any interval during which \vec{v} is constant.

When temperature is varying or spatially inhomogeneous, it is more natural to skip over the definitions of excess heat and excess work and simply account for entropy production. In any case, we can always decompose the change in non-equilibrium potential into the *excess heat-entropy* \mathcal{U}_{ex} and *excess work-entropy* Ω_{ex} . From the viewpoint of the experimenter in the environment, the useful quantity to emphasize is often the excess work-entropy produced in driving the system:

$$\Omega_{\text{ex}} = \Delta\phi_{\vec{v}} - \mathcal{U}_{\text{ex}} . \quad (8.5)$$

In general, the environmental control can be a deterministic or stochastic function of the joint history $(\overleftarrow{s}, \overleftarrow{v})$ of system-state and prior control. Here, let us only consider the case where the environmental control is an s -dependent stochastic process with a finite number of settings (i.e., $|\mathcal{V}|$ is arbitrarily large but finite): the environment may change its influence depending on the current state of the SUS. The s -dependent transition rates among the environmental states can be described as another transducing rate operator: $G^{(\mathcal{V} \rightarrow \mathcal{V}|s)}$. This allowance for feedback accommodates ‘intelligent’ environments, that take advantage of the micro-structure of the SUS. For example, the dependence on the present SUS state alone allows for most schemes of Maxwell’s demons that have become popular gedanken experiments.

Note also that feedback occurs naturally in complex physical systems. For example, the ensemble history of sodium ion channel conformations in mammalian neural membranes determines the integrated Na^+ current that changes the membrane voltage that influences the transition rates among conformations of the ion channels. In the case without s -dependence, the system is simply driven without feedback. An example is voltage-clamped driving of the membrane potential.

When the SUS and environment couple, each transduces the signal from the other to

produce a joint process over a joint state-space $\mathcal{U} = \mathcal{S} \times \mathcal{V} = \{(s, v)\}_{s \in \mathcal{S}, v \in \mathcal{V}}$ equipped with the projective functions $f_s : \mathcal{U} \rightarrow \mathcal{S}$ and $f_v : \mathcal{U} \rightarrow \mathcal{V}$. Transitions over the joint state space are summarized by the joint transition rate matrix $\mathcal{G} = \mathcal{G}^{(\mathcal{U} \rightarrow \mathcal{U})}$ over the joint state space, such that:

$$\mathcal{G}_{(s,v),(s',v')} = \begin{cases} G_{s,s}^{(\mathcal{S} \rightarrow \mathcal{S}|v)} + G_{v,v}^{(\mathcal{V} \rightarrow \mathcal{V}|s)} & \text{if } s=s' \text{ and } v=v' \\ G_{v,v'}^{(\mathcal{V} \rightarrow \mathcal{V}|s)} & \text{if } s=s' \text{ and } v \neq v' \\ G_{s,s'}^{(\mathcal{S} \rightarrow \mathcal{S}|v)} & \text{if } s \neq s' \text{ and } v=v' \\ 0 & \text{if } s \neq s' \text{ and } v \neq v' \end{cases}. \quad (8.6)$$

A simple example construction is given in Fig. 8.1. If the environment behaves independently of the state of the SUS, then the joint dynamic simplifies slightly since then $G^{(\mathcal{V} \rightarrow \mathcal{V}|s)} = G^{(\mathcal{V} \rightarrow \mathcal{V})}$ for all $s \in \mathcal{S}$.

The joint dynamic has a spectral decomposition unique from either of the transducing generators that go into building it. $\langle 0_{\mathcal{G}} |$ is the distribution over the joint state-space left invariant to the joint time-evolution operator $e^{\tau \mathcal{G}}$. Note that the restriction of $\langle 0_{\mathcal{G}} |$ to the states $\mathcal{U}_{\vec{v}} = \{u \in \mathcal{U} : f_v(u) = \vec{v}\}$ is *not* typically proportional to $\pi_{\vec{v}}$. However, since \mathcal{G} is still a transition-rate operator, $|0_{\mathcal{G}}\rangle$ is again a column vector of all ones which, with some benign abuse of notation (due to differences in dimensionality), we will again denote as $|\mathbf{1}\rangle$.

It is now useful to introduce the infinitesimal excess heat-entropy matrix for the joint system, which can be represented as a block-diagonal matrix. First, we introduce the \vec{v} -dependent infinitesimal heat-entropy matrices by the commutator:

$$\frac{\delta(\beta_{\vec{v}} \mathcal{Q}_{\vec{v}})}{dt} = [G^{(\mathcal{S} \rightarrow \mathcal{S}|\vec{v})}, \text{diag}(\phi_{\vec{v}})_{s \in \mathcal{S}}]. \quad (8.7)$$

For example, the \vec{v} -dependent infinitesimal heat-entropy matrix yields the expected infinitesimal heat via: $\langle \mu | \delta(\beta_{\vec{v}} \mathcal{Q}_{\vec{v}}) | \mathbf{1} \rangle = \langle \mu | [G^{(\mathcal{S} \rightarrow \mathcal{S}|\vec{v})}, \text{diag}(\phi_{\vec{v}})] | \mathbf{1} \rangle dt = \langle \dot{\mu} | \phi_{\vec{v}} \rangle dt$, since $G^{(\mathcal{S} \rightarrow \mathcal{S}|\vec{v})} | \mathbf{1} \rangle = \vec{0}$. However, the primary utility of the infinitesimal heat-entropy matrix is in calculations of joint system–environment dynamics, especially for higher-order moments of expected heat.

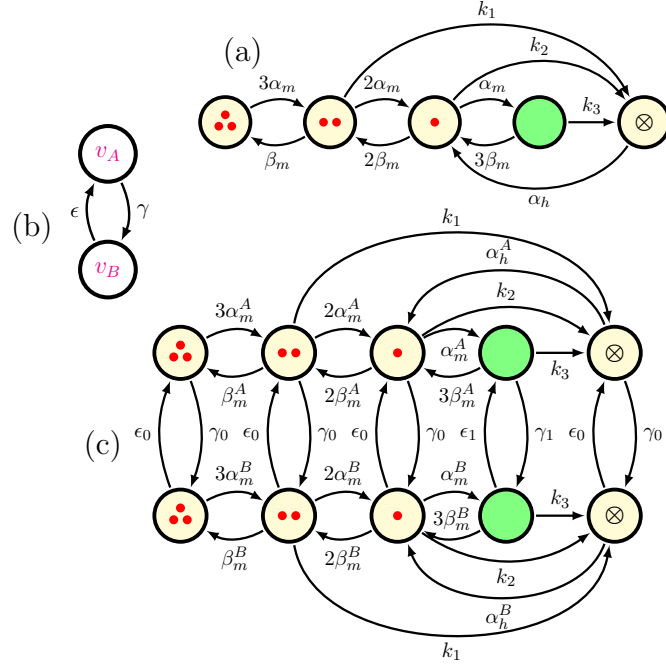


Figure 8.1: Here we give a simple example to illustrate the construction of a joint dynamic from system (Na^+ channel) and environment (experimenter with voltage control) that each feed back on the other. Note that all self-transition rates are implied, but not shown explicitly in hopes of maintaining a visibly palatable presentation. (a) An ion channel makes transitions among its relevant conformational states \mathcal{S} at rates that are dependent on the voltage $\mathbf{V} \subset \mathbb{R}$ accross the cell membrane. This simple five-state model suggested in Ref. [66] was inspired by experiments by Patlak in Ref. [189]. There is only one molecular conformation of the channel, indicated with solid green, in which sodium ions are allowed to flow. States with red dots indicate the number of activation gates that are closed. The final blocked state, indicated with the \otimes , is the conformation where the channel has been plugged up by the ball-and-chain-like inactivation gate. (b) An experimenter can clamp the membrane voltage to a desired value. In this simple scenario, an experimenter toggles between two voltage-clamped membrane potentials $\mathbf{V} = \{v_A, v_B\}$ at different rates that depend on whether the ion channel is in a conducting conformation or not. (c) The resulting joint dynamic is non-Markovian over \mathcal{S} and non-Markovian over \mathbf{V} ; nevertheless it is Markovian over the joint state-space \mathcal{U} . The joint rate operator allows exact calculation of excess thermodynamic quantities in terms of its spectral properties.

The joint infinitesimal heat-entropy matrix can be constructed as the block diagonal matrix:

$$\dot{H} = \text{diag}\left(\frac{\delta(\beta_{\vec{v}} Q_{\vec{v}})}{dt}\right)_{\vec{v} \in \mathbf{V}} \quad (8.8)$$

$$= \text{diag}\left([G^{(\mathbf{S} \rightarrow \mathbf{S}|\vec{v})}, \text{diag}(\phi_{\vec{v}})_{s \in \mathbf{S}}]\right)_{\vec{v} \in \mathbf{V}}. \quad (8.9)$$

Note that the operator \dot{H} is constant in time.

Starting from any initial distribution over \mathbf{U} , $\langle \eta_0 |$, and assuming an isothermal setting so that $\langle \mathcal{U}_{\text{ex}} \rangle = \beta \langle Q_{\text{ex}} \rangle$, the first moment of the expected excess heat (per $k_B T$) for the SUS under the joint dynamic is:

$$\beta \langle Q_{\text{ex}} \rangle = \int_0^\tau \langle \eta_t | \dot{H} | \mathbf{1} \rangle dt \quad (8.10)$$

$$= \langle \eta_0 | \left(\int_0^\tau e^{t\mathcal{G}} dt \right) \dot{H} | \mathbf{1} \rangle \quad (8.11)$$

$$= \langle \eta_0 | \left(\tau |0_{\mathcal{G}}\rangle \langle 0_{\mathcal{G}}| + \mathcal{G}^{\mathcal{D}}(e^{\tau\mathcal{G}} - I) \right) \dot{H} | \mathbf{1} \rangle \quad (8.12)$$

$$= \tau \langle 0_{\mathcal{G}} | \dot{H} | \mathbf{1} \rangle + (\langle \eta_\tau | - \langle \eta_0 |) \mathcal{G}^{\mathcal{D}} \dot{H} | \mathbf{1} \rangle, \quad (8.13)$$

where $\mathcal{G}^{\mathcal{D}}$ is the Drazin inverse of the joint transition rate matrix \mathcal{G} , and we have used the fact that $\langle \eta_0 | \mathbf{1} \rangle = 1$ for any normalized η_0 . Note that the first term on the RHS of Eq. (8.13) grows exactly linearly in time: $\langle 0_{\mathcal{G}} | \dot{H} | \mathbf{1} \rangle$ is the rate of excess heat-entropy production in the steady-state joint dynamic atop nonequilibrium steady states. The second term on the RHS of Eq. (8.13) accumulates only transient contributions to the excess heat in the relaxation of the joint system to the joint steady state. Asymptotically, as $\langle \eta_\tau | \rightarrow \langle 0_{\mathcal{G}} |$ (if the joint dynamic has no exact periodicities), the transient contribution saturates at $-\langle \eta_0 | \mathcal{G}^{\mathcal{D}} \dot{H} | \mathbf{1} \rangle$.

In the steady-state of the joint process:

$$\frac{\langle Q_{\text{ex}} \rangle}{\tau} = \langle 0_{\mathcal{G}} | \dot{H} | \mathbf{1} \rangle k_B T \quad (8.14)$$

for any finite interval τ . Excess heat is produced on average, which at first appears as a contradiction. However, the average excess heat at one level of description becomes the additional housekeeping heat to maintain the next level of description.

Meanwhile, the second moment of the expected (non-renormalized) excess heat (per $(k_B T)^2$) for the SUS under the steady-state joint dynamic is:

$$\beta^2 \langle Q_{\text{ex}}^2 \rangle = \left\langle \left(\int_0^\tau (\beta \dot{Q}_{\text{ex}})_t dt \right) \left(\int_0^\tau (\beta \dot{Q}_{\text{ex}})_{t'} dt' \right) \right\rangle \quad (8.15)$$

$$= \int_0^\tau \int_0^\tau \left\langle (\beta \dot{Q}_{\text{ex}})_t (\beta \dot{Q}_{\text{ex}})_{t'} \right\rangle dt' dt \quad (8.16)$$

$$= 2 \int_0^\tau \int_0^{\tau-t} \left\langle (\beta \dot{Q}_{\text{ex}})_0 (\beta \dot{Q}_{\text{ex}})_\ell \right\rangle d\ell dt \quad (8.17)$$

$$= 2 \int_0^\tau \int_0^{\tau-t} \langle 0_{\mathcal{G}} | \dot{H} e^{\ell \mathcal{G}} \dot{H} | \mathbf{1} \rangle d\ell dt \quad (8.18)$$

$$= \tau^2 \langle 0_{\mathcal{G}} | \dot{H} | \mathbf{1} \rangle^2 - 2\tau \langle 0_{\mathcal{G}} | \dot{H} \mathcal{G}^{\mathcal{D}} \dot{H} | \mathbf{1} \rangle + 2 \langle 0_{\mathcal{G}} | \dot{H} \mathcal{G}^{\mathcal{D}} (e^{\tau \mathcal{G}} - I) \mathcal{G}^{\mathcal{D}} \dot{H} | \mathbf{1} \rangle . \quad (8.19)$$

As the duration $\tau \rightarrow \infty$, the last term on the RHS of Eq. (8.19) saturates to:

$$-2 \langle 0_{\mathcal{G}} | \dot{H} (\mathcal{G}^{\mathcal{D}})^2 \dot{H} | \mathbf{1} \rangle .$$

The variance of excess heat in the steady state of the joint dynamic is, by $\text{var}(Q_{\text{ex}}) = \langle Q_{\text{ex}}^2 \rangle - \langle Q_{\text{ex}} \rangle^2$, therefore:

$$\begin{aligned} \text{var}(Q_{\text{ex}}) &= -2(k_B T)^2 \left(\langle 0_{\mathcal{G}} | \dot{H} \mathcal{G}^{\mathcal{D}} \dot{H} | \mathbf{1} \rangle \tau \right. \\ &\quad \left. - \langle 0_{\mathcal{G}} | \dot{H} \mathcal{G}^{\mathcal{D}} (e^{\tau \mathcal{G}} - I) \mathcal{G}^{\mathcal{D}} \dot{H} | \mathbf{1} \rangle \right) . \end{aligned} \quad (8.20)$$

This implies that the variance of excess heat per time in the steady-state joint dynamic scales as:

$$\frac{\text{var}(Q_{\text{ex}})}{\tau} \approx -2 \langle 0_{\mathcal{G}} | \dot{H} \mathcal{G}^{\mathcal{D}} \dot{H} | \mathbf{1} \rangle (k_B T)^2 \quad \text{for } \tau \gg |\frac{1}{\lambda_{\mathcal{G}}}|_{\text{max}} . \quad (8.21)$$

Above, $|1/\lambda_{\mathcal{G}}|_{\text{max}}$ is the time-constant of the slowest contributing mode of decay; specifically: the inverse magnitude of the smallest-magnitude eigenvalue of \mathcal{G} , besides 0, that contributes to $\langle 0_{\mathcal{G}} | \dot{H} \mathcal{G}^{\mathcal{D}} \mathcal{G}^{\mathcal{D}} \dot{H} | \mathbf{1} \rangle$.

We can recognize the RHS of Eq. (8.21) as the solution of a Green–Kubo form, relating the rate of growth of the variance of heat-entropy production in the steady-state of the joint system–environment dynamic to the integral of autocorrelation of heat puffs in the joint steady state:

$$\frac{\text{var}(Q_{\text{ex}})}{2(k_B T)^2 \tau} \approx \int_0^\infty (\langle \dot{H}(0) \dot{H}(t) \rangle_{\text{s.s.}} - \langle \dot{H} \rangle_{\text{s.s.}}^2) dt , \quad (8.22)$$

which is an approximate equality for $\tau \gg |1/\lambda_G|_{\max}$. Over long timescales ($\tau \gg |1/\lambda_G|_{\max}$), the standard deviation of excess heat depends on integrated correlation, is proportional to temperature, and is proportional to the square-root of the time elapsed. Analogous results can be formulated for the excess work W_{ex} via the joint infinitesimal work-entropy matrix: $\dot{\Sigma} = \text{diag}\left([G(\mathbf{v} \rightarrow \mathbf{v}^s), \text{diag}(\phi_{\vec{v}}(s))_{\vec{v} \in \mathbf{v}}]\right)_{s \in \mathcal{S}}$.

Overall, these results support the broader notion that dissipation depends on the structure of correlation.

8.2.1 Housekeeping renormalization of excess heat in hierarchical nonequilibrium steady states

In the steady-state of the joint process:

$$\frac{\langle Q_{\text{ex}} \rangle}{\tau} = \langle 0_G | \dot{H} | \mathbf{1} \rangle k_B T \quad (8.23)$$

for any finite interval τ . Excess heat is produced on average, which at first appears as a contradiction. However, the average excess heat at one level of description becomes the additional housekeeping heat to maintain the next level of description. Hence, we have found how to renormalize the housekeeping to obtain hierarchical thermodynamic descriptions of a nonequilibrium system:

$$\langle \dot{Q}_{\text{hk}}^{(\text{renorm})} \rangle = \langle \dot{Q}_{\text{hk}}^{(\text{former})} \rangle + \langle 0_G | \dot{H} | \mathbf{1} \rangle k_B T . \quad (8.24)$$

At the next higher level of excess, the new excess heat will be that heat produced in transitions among different joint dynamics. The renormalized excess heat contributions will be of the form:

$$\beta \langle Q_{\text{ex}}^{(\text{renorm})} \rangle = (\langle \eta_\tau | - \langle \eta_0 |) \mathcal{G}_j^{\mathcal{D}} \dot{H}_j | \mathbf{1} \rangle , \quad (8.25)$$

over any interval τ where the joint dynamic is held fixed, and j now indexes the different possible joint dynamics. Infinitesimally, this can be written as:

$$\beta \delta \langle Q_{\text{ex}}^{(\text{renorm})} \rangle = \beta \lim_{\tau \rightarrow 0} \frac{\langle Q_{\text{ex}}^{(\text{renorm})} \rangle}{\tau} \quad (8.26)$$

$$= \langle \dot{\eta} | \mathcal{G}_j^{\mathcal{D}} \dot{I}_j | \mathbf{1} \rangle \quad (8.27)$$

$$= \langle \eta | \mathcal{G}_j \mathcal{G}_j^{\mathcal{D}} \dot{I}_j | \mathbf{1} \rangle \quad (8.28)$$

$$= \langle \eta | \dot{I}_j | \mathbf{1} \rangle - \beta \langle \dot{Q}_{\text{hk}}^{(\text{add})} \rangle, \quad (8.29)$$

where $\langle \dot{Q}_{\text{hk}}^{(\text{add})} \rangle = \langle 0_{\mathcal{G}} | \dot{I} | \mathbf{1} \rangle k_{\text{B}} T$, and we have used the fact that $\mathcal{G}_j \mathcal{G}_j^{\mathcal{D}} = I - |0_{\mathcal{G}_j}\rangle \langle 0_{\mathcal{G}_j}|$. This suggests that the new heat-entropy matrix at the next hierarchical level of excess will be:

$$\dot{I}^{(\text{renorm})} = \text{diag} \left([\mathcal{G}_j, \mathcal{G}_j^{\mathcal{D}} \dot{I}_j] \right)_j, \quad (8.30)$$

although the details are left for further exploration.

To summarize: When systems are strongly coupled with two-way feedback, it is often desirable to change the level of description in one's model, to treat the joint dynamic as the status quo. Consequentially, the average excess heat at the former level of description becomes the additional housekeeping heat to maintain the new level of description. Here, we have described how to renormalize the housekeeping to obtain hierarchical thermodynamic descriptions of a nonequilibrium system.

It would be worthwhile to apply this procedure recursively to a subsystem 'driven' by the rest of the system it is strongly coupled with. This procedure, applicable far from equilibrium, will likely have an interesting relationship with the standard renormalization group methods of equilibrium statistical mechanics.

Chapter 9

Statistical–Computational Mechanics: Thermodynamics from Information-Theoretic Symmetries

9.1 Chapter Overview

We start from a well-supported definition of nonequilibrium thermodynamic entropy that extends the accepted definitions of equilibrium entropy and steady state entropy of a system. We point out that non-extensivity of entropy should be expected (whether in or out of equilibrium) for two subsystems with strongly coupled degrees of freedom. Moreover, we derive that this deviation from extensivity is precisely the mutual information between the two subsystems. Indeed, the difficulty of calculating entropy for complex systems stems from the fact that small subsystems are not independent.

We proceed to identify classes of information symmetries among the degrees of freedom typical of physical systems in the thermodynamic limit of large system size. In doing so, we develop a synthesis of statistical mechanics [188] (the workhorse for deriving properties of systems in thermodynamic equilibrium) and computational mechanics [55] (an information theoretic framework for discovering and quantifying pattern)—a synthesis that yields a practical mechanics for calculating thermodynamic properties of strongly interacting systems arbitrarily far from equilibrium. In recognition of its two main ingredients, we call this synthesis ‘Statistical–Computational Mechanics’.

When the physical system is configured according to a stationary stochastic process, we show that the thermodynamic entropy has contributions from both the Shannon entropy rate of the observed pattern and the excess entropy required to synchronize to this pattern. This result allows us to refine recent predictions for the synchronization cost of energy-harvesting information ratchets. It also suggests a new approach to inferring the thermodynamics of strongly interacting particles.

Even when the configurational degrees of freedom (or groupings of them) do not constitute a stochastic process, we discuss how the much more general notion of *conditional stationarity* (among system-specific groupings of degrees of freedom) can be used to infer nonequilibrium entropy from the instantaneously observed pattern. As an example, this can be adapted to show why and exactly how the Shannon entropy density of patterns in Ising-like models in any dimension contribute to the nonequilibrium entropy.

Finally, we discuss how the utility of this theory breaks down for the nonequilibrium thermodynamics of nanoscale and mesoscopic systems, where the thermodynamic limit is invalid. We discuss how the excess entropy of small configuration becomes more important and how the configurational entropy rate loses precedence to the entropy rate of the dynamics.

9.2 Thermodynamic entropy

Quite generally, we can suppose that a physical system has N degrees of freedom at any time t , with corresponding random variables $X_n(t)$ for $n \in \{0, 1, \dots, N-1\}$, where N is arbitrarily large, but finite. The collection of these random variables $X_{0:N}(t)$ is itself a random variable. The entropy of $X_{0:N}(t)$ at any time—or equivalently, the entropy of the high-dimensional joint distribution over the random variables for each degree of freedom—is the thermodynamic entropy. More precisely, if an active environment is capable of doing work on the system via some driving protocol $Y_{-\infty:t}$ —where Y_t may be a vector of deterministic controls, e.g. $(\vec{E}(t), \vec{B}(t))$, or, more generally, a random variable with M degrees of freedom $Y_{0:M}(t)$ —then it has been proposed [94, 234, 199] that the

thermodynamic entropy of the system at any time t is:

$$S \equiv k_B H \left[\Pr(X_{0:N}(t) | Y_{-\infty:t}) \right] , \quad (9.1)$$

where $H[\cdot]$ is the Shannon entropy of the distribution. Note that the subscripted index for Y is not from the same domain as the subscripted index for X .

The above definition Eq. (9.1), in terms of degrees of freedom, is an equivalent but more directly *physical* statement of the nonequilibrium entropy S than explicitly proposed in Refs. [94, 234, 199]. In the case of NESSs, this nonequilibrium entropy reduces to the steady-state entropy S_{ss} as shown in Ref. [199], which has received strong validation as a physically relevant quantity for steady states [108, 256, 163].

We use probability rather than probability density because each observable degree of freedom has a measurable probability distribution by assumption of observability. This use of discrete probabilities is in agreement with nearly all previous work in thermodynamics—recall that even apparently continuous spatial and momentum coordinates are discretized by atomic phase-volume units of Planck’s constant: $h = dx dp$.

The active environment may drive the system arbitrarily far from equilibrium, may drive it quasistatically, or may not drive it at all; all three cases are addressed in this unifying framework.

Let \mathcal{A}_n be the state-space for the single-degree-of-freedom random-variable X_n . Then $\mathcal{A}_{\text{net}} \equiv \cup_{n=1}^N \mathcal{A}_n$ will be the net state-space for the system $X_{0:N}$ at any instant. If the boundary conditions are held fixed, such that $Y_t = y$ for all $t > t_0$, then the system will relax to the y -parametrized equilibrium distribution π_y over the state-space \mathcal{A}_{net} . I.e., $X_{0:N}(t)$ converges in probability to the random variable $X_{0:N}(\infty) \sim \pi_y$ as $t \rightarrow \infty$ if Y_t is held fixed at y . The relevant timescale of relaxation depends on the system under study.

Insofar as actual probabilities are relevant to dynamical systems—which they are—we propose that the general nonequilibrium entropy S is a physically relevant extension of the equilibrium entropy S_{eq} . It is easy to see that the general nonequilibrium entropy relaxes to the equilibrium entropy when all external driving ceases to be dynamic, staying fixed

at the value of $Y(t) = y$ for all $t > 0$:

$$\lim_{t \rightarrow \infty} S(t) = k_B H[\boldsymbol{\pi}_y] \quad (9.2)$$

$$= S_{\text{eq}} . \quad (9.3)$$

Moreover, the difference in expected internal energy and (environmental temperature T times) nonequilibrium entropy at any time yields a nonequilibrium free energy that is greater than the reference equilibrium free energy at that time by the Kullback–Leibler divergence:

$$F_{\text{add}}(t) = k_B T D_{\text{KL}} \left(\text{Pr}(X_{0:N}(t) | Y_{-\infty:t}) \parallel \boldsymbol{\pi}_{Y_t} \right) , \quad (9.4)$$

which is always non-negative.

9.3 Non-extensivity of entropy

It is easy to see that two sub-systems generally have non-extensive thermodynamic entropy, whether or not they are in thermodynamic equilibrium. Let subsystem A be the collection of degrees-of-freedom $X_{0:a}$, while subsystem B will be the complementary collection of degrees-of-freedom $X_{a:N}$. The joint system is $A \cup B$. Directly calculating, we find that:

$$S(A, B) = S(A) + S(B) - I(A; B) . \quad (9.5)$$

More explicitly, if we consider the disturbances to the joint system as a random variable $Y_{-\infty:t}$ rather than as a mere parameter, we find that the joint entropy discounts a conditional mutual information between the subsystems from what would otherwise be an extensive entropy:

$$S(A, B) = k_B H(A, B | Y_{-\infty:t}) \quad (9.6)$$

$$= k_B H(A | Y_{-\infty:t}) + k_B H(B | Y_{-\infty:t}) + I(A; B | Y_{-\infty:t}) \quad (9.7)$$

$$= S(A) + S(B) - I(A; B | Y_{-\infty:t}) . \quad (9.8)$$

Hence, whether in or out of equilibrium, *the deviation from extensivity is precisely the mutual information between two subsystems.*

Despite its simplicity, Eq. (9.5) goes a long way towards explaining the basic physics behind the recent resurgence of Maxwellian demonology, including new results affirming the thermodynamic relevance of information in feedback control [113, 209], prediction [241], and energy-harvesting ratchets [153, 30]. For example, if system B is itself an autonomous ‘Maxwell’s demon’, then it can reduce the net entropy of the joint system $A \cup B$ by becoming more correlated with system A over time.¹ System B can then leverage this information to extract work from system A at the cost of de-correlating from it. As another interpretation (elaborating on Ref. [241]) if system B is a biological system while system A is its structured environment, then system B can suppress entropy (locally) and even harvest energy for itself by building a model of its environment and inferring the current state of the environment.² In each of these scenarios, Eq. (9.8) suggests that the two subsystems may become correlated through mutual experience induced by the larger universe $Y_{-\infty:t}$. In such cases, the biological system may not need to exert its own efforts to reap the thermodynamic rewards of being strongly correlated with system A . Visible light, for example, may just bounce off of structured objects (tasty morsels, say) in the surroundings, into an observer’s pupils by generosity of the beaming sun.

However, this non-extensivity of entropy will only be appreciable when the mutual information among the degrees of freedom is comparable to the entropy of either subsystem. There are three obvious scenarios that can lead to noticeable non-extensivity of the entropy:

1. Strong interaction (and thus inter-dependence) between the degrees of freedom of each subsystem
2. A large boundary between the two subsystems
3. Small system size—Or, at least, small subsystem size

¹Doing so, however, will likely induce an increased entropy in some external system C due to overall entropy balance of the universe.

²However, if B ’s predictions are poor, then the two systems will become less correlated with time, resulting in an increase in entropy that will correspond to dissipated work if the biological system is to work to return to an enlightened predictive state.

All three points are practical considerations, but condition (1) must be met to some extent as a prerequisite for the other two conditions to matter. Indeed, the entropy is exactly extensive when the two subsystems are statistically independent.

It is tempting to wonder whether the non-extensivity corresponds to Tsallis entropy [258] or some other more general super-statistics (see, e.g., Ref. [16]). However, we find that the claimed motivation (i.e., enabling non-extensive entropy) behind at least the Tsallis entropy infuses confusion into the perceived source of extensivity in standard Statistical Mechanics, as evidenced by much follow up work [40]. Extensivity does *not* follow from $S_{\text{eq}} \equiv k_B H[\boldsymbol{\pi}]$ nor from the microcanonical specialization $S_{\text{eq}}^{\text{micro}} \equiv k_B \log \Omega$, where Ω is the number of microstates accessible to a system at fixed energy (and particle number and volume). Rather, it is the assumption of *statistical independence* among degrees of freedom—the assumption that $\Pr(X_i X_j) = \Pr(X_i) \Pr(X_j)$ for any $X_i \in A$ and $X_j \in B$ —that leads to extensive entropy $S(A, B) = S(A) + S(B)$. This assumption is related to the assumption that energies are not shifted upon addition of new constituents to the system, which is patently false when there are interactions. So, what is behind the intuition that entropy is extensive? It is this: weak interactions lead to approximate statistical independence which implies approximate extensivity of the entropy. Nevertheless, the works quoted within Ref. [258] seem to have been written with a more acute awareness of this situation. Despite these criticisms, super-statistics have the opportunity to play a constructive role for understanding how non-equilibrium probability distributions project onto the equilibrium probability distributions [17, 52]. But this is distinct from enabling extensive entropy.

Non-extensivity, properly understood, is critical to understanding the complexity of a system. Here, we have shown that this deviation from extensivity follows directly from Eq. (9.1) and is precisely the mutual information between two subsystems.

Moreover, it may be recognized that the non-additivity of entropy for small portions of a system is exactly why statistical mechanics becomes difficult for complex systems! The entropy of the 2-D Ising model is clearly not just the sum of the entropies for each spin. It is the interdependencies (among degrees of freedom through the Hamiltonian) that reduce total entropy in a nontrivial way as a function of system parameters.

In the following, we exploit this utmost importance of mutual information and conditional entropies to provide a new approach for thermodynamic calculations.

9.4 Instantaneously stationary processes of configuration

If the number of degrees of freedom N for the system is very large—as expected explicitly in the definition of ‘thermodynamic limit’—then $\{X_n\}_{n \in D_{\text{d.o.f.}}}$ can profitably be treated as a stochastic process over the domain enumerating the degrees of freedom $D_{\text{d.o.f.}} = \{1, 2, \dots, N\}$. *If this process $\{X_n\}_n$ is stationary*, we can talk about the *configurational Shannon entropy rate* $h_\mu(\{X_n\}_n)$ of the instantaneous configuration of the system:

$$h_\mu(\{X_i\}_i) = \lim_{n \rightarrow \infty} H(X_n | X_{0:n}) . \quad (9.9)$$

Whether in or out of equilibrium, the total instantaneous thermodynamic entropy $S(t)$ of the system is given by:

$$\frac{1}{k_B} S(t) = H(X_{0:N}(t) | Y_{-\infty:t}) \quad (9.10)$$

$$= N h_\mu(\{X_i\}_i) + \sum_{n=0}^{N-1} \left(H(X_n | X_{0:n}) - h_\mu(\{X_i\}_i) \right) \quad (9.11)$$

$$= N h_\mu(\{X_i\}_i) + \mathbf{E}_{\text{config}}^{(N)} . \quad (9.12)$$

In the thermodynamic limit of large system size, the rightmost term converges to the *configurational excess-entropy*:

$$\mathbf{E}_{\text{config}} = \lim_{N \rightarrow \infty} \mathbf{E}_{\text{config}}^{(N)} , \quad (9.13)$$

where

$$\mathbf{E}_{\text{config}}^{(N)} \equiv \sum_{n=0}^{N-1} \left(H(X_n | X_{0:n}) - h_\mu(\{X_i\}_i) \right) . \quad (9.14)$$

Furthermore, the configurational excess entropy of the system can be shown to be how much one half of the system knows about the other half of the system:

$$\mathbf{E}_{\text{config}}^{(N)} = I(X_{0:\lfloor N/2 \rfloor}; X_{\lfloor N/2 \rfloor:N}) . \quad (9.15)$$

As its name suggests, the configurational excess entropy is the excess thermodynamic entropy incurred by not yet being synchronized to the X_n stationary process—the information that would have been predictable if sufficiently many random variables had already been analyzed in another part of the system. But alas, we must start observing somewhere. This excess entropy may also be considered the cost of amnesia: supposedly, if you do not keep track of some information about previous random variables, you may lose out on discounting future statistical correlations.

It should be apparent from Eq. (9.12) that the configurational excess entropy is the non-extensive part of the thermodynamic entropy. In the thermodynamic limit, $\mathbf{E}_{\text{config}}$ will often become a negligible fraction of the total entropy. Although, notably, the configurational excess entropy should be larger for systems with longer-range interactions. At one extreme, for systems with zero entropy rate, the configurational excess-entropy will constitute *all* of the system’s entropy.

Eq. (9.12) is the first thermodynamic derivation that makes it clear *why* configurational excess entropy can be used as a general order parameter for a system, as has been previously suggested [88]. Near equilibrium critical points, while the entropy rate is relatively insensitive to small changes of control parameters, the excess entropy peaks as the system is overwhelmed with long-range correlations. Out of equilibrium, the excess entropy is just as keen at detecting thermodynamically relevant changes in structure.

One important implication of Eq. (9.12) is a rather practical one for calculating thermodynamic quantities from data obtained from direct observation of nonequilibrium systems:

We can calculate the nonequilibrium entropy of a system at any time, upon any arbitrary history of driving, simply by calculating two computational-mechanical properties of the present configuration: the *configurational entropy rate* and the *configurational excess entropy*.

9.4.1 Demons and information reservoirs

In general, the total thermodynamic entropy of the system is the sum of both the configurational entropy rate and the configurational excess entropy. For one, this sheds light

on how the difference in Shannon entropy rates of configurations of symbols on input and output strings can serve as a thermodynamic work resource, as recently discovered [162, 30]. In Ref. [30], it was shown that the asymptotic rate of work extraction by an ‘information ratchet’ is bounded by:

$$\lim_{N \rightarrow \infty} \frac{\langle W \rangle}{N} \leq \Delta h_\mu(\{X_i\}_i) . \quad (9.16)$$

Indeed, it has been shown that the difference in entropy rates $\Delta h_\mu(\{X_i\}_i)$ is more fundamental than the differences in single-variable entropy $\Delta H(X_i)$. Here, our work gives a further testable prediction: When the total change in entropy is due to the change in entropy of the bit string and the change in entropy of the heat bath, the total work that can be extracted from a length- N bit string (not just the rate) must be bounded by:

$$\frac{1}{k_B} \left\langle \int \frac{\dot{W}}{T} dt \right\rangle \leq N \Delta h_\mu(\{X_i\}_i) + \Delta \mathbf{E}_{\text{config}}^{(N)} . \quad (9.17)$$

For most strings of appreciable length, $\Delta \mathbf{E}_{\text{config}}^{(N)} \approx \Delta \mathbf{E}_{\text{config}}$. In the isothermal setting, and after appreciable length of the bit string has been transformed, this reduces to:

$$\beta \langle W \rangle \leq N \Delta h_\mu(\{X_i\}_i) + \Delta \mathbf{E}_{\text{config}} . \quad (9.18)$$

In the thermodynamic limit, the rate of work extraction will asymptote to the constraint given by Eq. (9.16). However, this modification by the change in configurational excess entropy corresponds to the inevitable cost of synchronizing to an informational work resource. A special case of such a synchronization cost was recently investigated in [29]. Supposing that the output can be transformed to an IID string, which would have no excess entropy, our work suggests the general constraints on synchronization will cost the user at least $k_B T \mathbf{E}_{\text{config}}$ of un-usable work as a start-up cost to the asymptotic work-harvesting rate.

The above bound on work extraction considered explicitly only the change in entropy of the bit string $\Delta S(A) = N \Delta h_\mu(\{X_i\}_i) + \Delta \mathbf{E}_{\text{config}}^{(N)}$, before and after use. Implicit was also the change in entropy of the heat bath $\Delta S(C)$, before and after use, supposing that the heat bath is structureless and thus uncorrelated with A such that $\Delta S(A, C) = \Delta S(A) + \Delta S(C)$. Heat lost to the bath is assumed irrevocable.

More generally, we should also consider the change in entropy within the demon's state-space, $\Delta S(B)$ and the change in correlation between the demon and the bit-string $\Delta I(A; B)$. In the isothermal setting, the total work extractable will be:

$$\beta \langle W \rangle \leq \Delta S(A, B) \quad (9.19)$$

$$= \Delta S(A) + \Delta S(B) - \Delta I(A; B) \quad (9.20)$$

$$= N\Delta h_\mu(\{X_i\}_i) + \Delta \mathbf{E}_{\text{config}}^{(N)} + \Delta S(B) - \Delta I(X_{0:N}; B) . \quad (9.21)$$

This more thorough bound on work extraction implies that the disordering of the demon's 'mind' can also be seen as a work resource, since the mind in addition to the bit string can be disordered. This aspect of energy harvesting was recently discussed as the 'thermodynamics of memory' in Ref. [31]. On the other hand, if the mind becomes *more* ordered while observing the bit string, then this will amount to less work harvested. This latter point seems related to the cost of synchronization, which is the excess entropy of the input if the demon's mind corresponds approximately to a simplex with the stationary distribution over bit-string states as a prior. Overall, this suggests that the demon's *model* of the incoming string has thermodynamic relevance. Minimal models allow greater work extraction. Start states are better than start simplices. To summarize, here is pointwise how to *maximally* extract work:

- Change the entropy rate of the string from low entropy (ordered) to high entropy (disordered) [Gain up to $Nk_B T \Delta h_\mu$ of work]
- Change the synchronization cost of the string from easily synchronizable (low \mathbf{E}) to hardly synchronizable (high \mathbf{E}) [Gain up to $\Delta \mathbf{E}_{\text{config}}$ of work]
- Scramble the demon's mind [Gain up to $\Delta S_{\text{demon's mind}}$ of work]
- Let the demon become less correlated with the bit string [Gain up to $\Delta I(X_{0:N}; B)$ of work]

However, not each of these can be satisfied simultaneously. Moreover, each of the above items become work detractors upon flipping the time ordering of the relevant change.

Therefore, the engineering of Maxwellian demons will require optimizing the tradeoffs given design constraints.

Often, groupings of the degrees of freedom constitute a stochastic process, even when the full enumeration of them do not. We next show how symmetries of the Hamiltonian suggest which groupings of random variables can be treated as particles.

9.5 Particles

9.5.1 Hamiltonians and ‘particles’

Even if the full enumeration of the degrees of freedom does not constitute a strictly stationary process, we can often partition the random variables into degrees-of-freedom packets called ‘particles’ to gain some further tractability. Indeed, we will take an unconventional approach by defining different particle types as different equivalence classes of groupings of random variables. In this view, non-interacting particles are independent random variables. Interacting particles are dependent, but somehow symmetrized in their interaction.³ Often the the particles are indicated via symmetries of the Hamiltonian.

Suppose there are γ degrees of freedom associated with each of the \mathbf{n} particles; $\gamma = 7$ would be necessary for position, momentum, and spin-1/2 for a mobile particle in three dimensions. Let the particles be the collection of random variables $P_i = X_{i\gamma:(i+1)\gamma}$. Often, $\{P_i\}_i$ will be a strictly stationary process whereas $\{X_n\}_n$ would be a cyclostationary process. The configurational entropy rate and configurational excess entropy for the *particle* configurations can now be used to calculate the total system entropy, analogous to § 9.4.

In equilibrium, the stochastic process for X_n is specified by the equilibrium distribution π_y . For strongly interacting particles, the entropy of the system will be significantly lower than for non-interacting particles of the same makeup—this is because of the partial

³The thoughtful reader will recognize that intrinsically *identical* particles cannot be treated as individual random variables. Such a treatment would assert distinguishability from the outset. So how do we treat identical particles? As always, the rigorous approach would be with the appropriate degrees of freedom! In the case of identical particles with position, momentum, and spin in a finite volume, it is useful to choose the degrees of freedom to be choices within each point of phase space: Occupied or not? If so, what value of spin? At the surface, this looks like many additional degrees of freedom. Well, it is—but each has much lower entropy. The theoretical move is akin to the so-called second-quantization. The net result is the same in all comparable cases—i.e., when the particles are distinguishable.

redundancy of certain degrees of freedom when they strongly influence each other. For example, electrons that are close to each other have a significant bias in the probability of their joint spin orientations due to the exchange interaction. Hence, strong interaction implies $h_\mu(\{P_i\}_i) \ll H(P_i)$ and $S(t) \ll \mathbf{n}k_B H(P_i(t))$.

9.5.1.1 Hamiltonians with complete symmetry among particles

For Hamiltonians that treat each particle symmetrically—that the Hamiltonian is invariant under the exchange of P_i and P_j —we find that $\Pr(P_i) = \Pr(P_j)$, $\Pr(P_i|P_j) = \Pr(P_j|P_i)$, $\Pr(P_i|P_j, P_k) = \Pr(P_j|P_i, P_k) = \Pr(P_k|P_i, P_j)$, and so on, for all $i, j, k, \dots \in \{1, \dots, \mathbf{n}\}$; $i \neq j \neq k \dots$ so long as the particles are initially homogeneously distributed. Hence, resulting configurations, even if the Hamiltonian is time-varying, will be stationary configuration processes. Moreover, we then have:

$$\frac{1}{k_B} S = H(P_1, P_2, \dots, P_{\mathbf{n}}) \quad (9.22)$$

$$= \sum_{i=0}^{\mathbf{n}} H(P_i|P_{1:i}) \quad (9.23)$$

$$\leq (\mathbf{n} - \ell) H(P_\ell|P_{1:\ell}) + \sum_{i=0}^{\ell} H(P_i|P_{1:i}) , \quad (9.24)$$

for any ℓ . We can use Eq. (9.24) to put an upper bound on the entropy of any arbitrarily strongly-interacting collection of particles with homogeneous initial conditions and time-varying (but particle-symmetric) Hamiltonian.

A trivial (and extreme) case is for non-interacting particles, for which $\Pr(P_i|P_j) = \Pr(P_i)$ for all i, j . Then $S = \mathbf{n}k_B H(P_i)$, even for time-varying Hamiltonians as long as the particles were started homogeneously and are acted on indifferently.

9.5.2 Instantaneously conditionally-stationary configurational process

Much more general than the constraint of stationarity is the opportunity for conditional stationarity among groupings of degrees of freedom. Multi-indices can be useful for identifying conditional stationarity, especially if the model is explicitly multi-dimensional.

Conditional stationarity is a symmetry whereby conditional entropies are invariant under a specified translation of multi-indices. Such circumstances are induced by the

dynamics and indicate how the thermodynamic entropy can be obtained from observable Shannon entropies of the configurational degrees of freedom.

9.6 Complexity of the Mesoscale

For nanoscale systems, we cannot justify approximations that require the thermodynamic limit. As one consequence, if we continue to treat the joint state of the system as a stochastic process over degrees of freedom, then the excess entropy—the *predictable information* about the system—can now constitute a large fraction of the total entropy of the system. However—in part because of this dominance of excess entropy—we may also find that, without the thermodynamic limit, there is little utility in treating the different degrees of freedom as elements of a stochastic process. Rather, we may choose simply to treat the joint random variable $X_{0:N}$ on its own terms. In the mesoscopic regime—without the aid of the thermodynamic limit, but also with more degrees of freedom than one can easily handle simultaneously—we encounter theoretical justification for awkwardness in analysis. Indeed, the difficulties encountered at the scale of macromolecules and molecular machines is well known across many fields. However, the difficulty of analysis corresponds to a rich intrinsic utility of the degrees of freedom. This is a scale at which number of computations per degrees of freedom may be maximized because of the strong interaction among all constituents.

9.7 Conclusion

According to Eq. (9.12), the ultimate quest of thermodynamics (with special difficulty for nonequilibrium thermodynamics) is to identify the configurational stochastic process $\{X_n\}_{n \in D_{\text{d.o.f.}}}$ induced by the driving history $Y_{-\infty:t}$. Combining the tools of both statistical mechanics and computational mechanics, we wager optimism that we can analyze the thermodynamics of many complex physical systems, even far from equilibrium.

REFERENCES

- [1] C. Aghamohammadi, J. R. Mahoney, and J. P. Crutchfield. The ambiguity of simplicity. 2016. [arxiv.org:1602.08646](https://arxiv.org/abs/1602.08646) [quant-ph]. 169
- [2] P. W. Anderson. More is different. *Science*, 177(4047):393–396, 1972. 357
- [3] P. J. Antsaklis and A. N. Michel. *A Linear Systems Primer*. Springer Science & Business Media, New York, New York, 2007. 19
- [4] R.B. Ash. *Information Theory*. Dover books on advanced mathematics. Dover Publications, 1965. 87, 88, 95, 172
- [5] N. W. Ashcroft and N. D. Mermin. *Solid State Physics*. Saunders College Publishing, New York, 1976. 223, 273, 281
- [6] M. Atiyah, R. Bott, and V. K. Patodi. On the heat equation and the index theorem. *Inventiones Mathematicae*, 19(4):279–330, 1973. 21
- [7] Françoise Axel and Hikaru Terauchi. High-resolution x-ray-diffraction spectra of thue-morse gaas-alas heterostructures: Towards a novel description of disorder. *Physical Review Letters*, 66:2223–2226, 1991. 284
- [8] S. J. Axler. *Linear Algebra Done Right*, volume 2. Springer, New York, New York, 1997. 45
- [9] R. Badii and A. Politi. *Complexity: Hierarchical Structures and Scaling in Physics*, volume 6 of *Cambridge Nonlinear Science Series*. Cambridge University Press, 1997. 278, 284, 286
- [10] G. B. Bagci, U. Tirnakli, and J. Kurths. The second law for the transitions between the non-equilibrium steady states. 2012. [arXiv:1205.6119](https://arxiv.org/abs/1205.6119). 322
- [11] V. Balasubramanian. Equivalence and reduction of hidden markov models. Master’s thesis, MIT, Jan. 1993. 85
- [12] P. Ball. Material Witness: Beyond the crystal. *Nat. Mater.*, 13:1003, 2014. 274
- [13] V. S. Barbu and N. Limnios. *Semi-Markov Chains and Hidden semi-Markov Models toward Applications: Their Use in Reliability and DNA Analysis*, volume 191. Springer, New York, 2008. 53, 57
- [14] S. M. Barnett and S. Stenholm. Spectral decomposition of the Lindblad operator. *J. Mod. Optics*, 47(14-15):2869–2882, 2000. 60
- [15] J. Bechhoefer. Hidden markov models for stochastic thermodynamics. *New Journal of Physics*, 17(7):075003, 2015. 98

- [16] C. Beck and E. G. D. Cohen. Superstatistics. *Physica A: Statistical mechanics and its applications*, 322:267–275, 2003. 378
- [17] C. Beck, E. G. D. Cohen, and H. L. Swinney. From time series to superstatistics. *Phys. Rev. E*, 72:056133, Nov 2005. 378
- [18] F. Beichelt. *Stochastic Processes in Science, Engineering and Finance*. Chapman and Hall, New York, 2006. 53
- [19] A. Ben-Israel and T.N.E. Greville. *Generalized Inverses: Theory and Applications*. CMS Books in Mathematics. Springer, 2003. 34, 110
- [20] N. Berkovits and D. Witten. Conformal supergravity in twistor-string theory. *J. High Energy Physics*, 2004(08):009, 2004. 72
- [21] R. Berliner and S.A. Werner. Effects of stacking faults on diffraction: The structure of lithium metal. *Physical Review B*, 34:3586–3603, 1986. 221, 266, 276, 279
- [22] T. Bermúdez. Meromorphic functional calculus and local spectral theory. *Rocky Mountain J. Math.*, 29(2):437–447, 1999. 36
- [23] K. R. Beyerlein, R. L. Snyder, and P. Scardi. Faulting in finite face-centered-cubic crystallites. *Acta Crystallogr. Sec. A*, 67(3):252–263, 2011. 220
- [24] W. Bialek, I. Nemenman, and N. Tishby. Predictability, complexity, and learning. *Neural Comput.*, 13(11):2409–2463, November 2001. 120
- [25] J. J. Binney, N. J. Dowrick, A. J. Fisher, and M. E. J. Newman. *The Theory of Critical Phenomena*. Oxford University Press, Oxford, 1992. 81
- [26] J. J. Birch. Approximations for the entropy for functions of Markov chains. *Ann. Math. Statist.*, 33(2):930–938, 1962. 88
- [27] D. Blackwell. The entropy of functions of finite-state Markov chains. volume 28, pages 13–20, Publishing House of the Czechoslovak Academy of Sciences, Prague, 1957. 88, 89
- [28] M. L. Boas. *Mathematical Methods in the Physical Sciences*, volume 2. Wiley and Sons, New York, New York, 1966. 20, 53
- [29] A. B. Boyd, D. Mandal, and J. P. Crutchfield. Correlation-powered information engines and the thermodynamics of self-correction. *arXiv:1606.08506*, 2016. 357, 381
- [30] A. B. Boyd, D. Mandal, and J. P. Crutchfield. Identifying functional thermodynamics in autonomous maxwellian ratchets. *New Journal of Physics*, 18(2):023049, 2016. 98, 377, 381

- [31] A. B. Boyd, D. Mandal, and J. P. Crutchfield. Leveraging environmental correlations: The thermodynamics of requisite variety. *arXiv preprint arXiv:1609.05353*, 2016. 95, 357, 382
- [32] R. Bracewell. *The Fourier Transform and Its Applications*. McGraw-Hill, New York, third edition, 1999. 70
- [33] G. Brassard. Quantum communication complexity (a survey). *Found. Physics*, 33(11):1593–1616, 2003. 194
- [34] M. Budii, R. Mohr, and I. Mezic. Applied Koopmanism. *Chaos*, 22(4), 2012. 16, 17, 69
- [35] P. Caroff, J. Bolinsson, and J. Johansson. Crystal phases in III-V nanowires: From random toward engineered polytypism. *IEEE Journal of Selected Topics in Quantum Electronics*, 17:829–846, 2011. 273
- [36] J. H. E. Cartwright and A. L. Mackay. Beyond crystals: the dialectic of materials and information. *Phil. Trans. R. Soc. A*, 370:2807–2822, 2012. 269, 274
- [37] A. H. Castro Neto, F. Guinea, N. M. R. Peres, K. S. Novoselov, and A. K. Geim. The electronic properties of graphene. *Rev. Mod. Phys.*, 81:109–162, Jan 2009. 220
- [38] S. V. Cherepanova. X-ray scattering on one-dimensional disordered structures. *J. Struct. Chem.*, 53:S109–S132, 2012. 276
- [39] V. Y. Chernyak, M. Chertkov, and C. Jarzynski. Path-integral analysis of fluctuation theorems for general langevin processes. *Journal of Statistical Mechanics: Theory and Experiment*, 2006(08):P08001, 2006. 335, 337
- [40] A. Cho. A fresh take on disorder, or disorderly science? *Science*, 297(5585):1268–1269, 2002. 378
- [41] F. R. K. Chung. *Spectral Graph Theory*, volume 92. American Mathematical Soc., Providence, Rhode Island, 1997. 72
- [42] M. J. Cliffe, M. T. Dove, D. A. Drabold, and A. L. Goodwin. Structure determination of disordered materials from diffraction data. *Physical Review Letters*, 104:125501, 2010. 220
- [43] D. Colquhoun and A. G. Hawkes. Relaxation and fluctuations of membrane currents that flow through drug-operated channels. *Proceedings of the Royal Society of London B: Biological Sciences*, 199(1135):231–262, 1977. 356
- [44] C. Cortes and V. Vapnik. Support-vector networks. *Machine Learning*, 20(3):273–297, 1995. 17
- [45] R. Courant and D. Hilbert. *Methods of mathematical physics: first English edition*, volume 1. Interscience Publishers, 1953. 18

- [46] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. Wiley-Interscience, New York, second edition, 2006. 4, 7, 81, 97, 241, 274, 320, 325, 365
- [47] T.M. Cover and M.E. Hellman. The two-armed-bandit problem with time-invariant finite memory. *Information Theory, IEEE Transactions on*, 16(2):185–195, Mar 1970. 120
- [48] F. Creutzig, A. Globerson, and N. Tishby. Past-future information bottleneck in dynamical systems. *Physical Review E*, 79(4):041925, 2009. 83
- [49] G. E. Crooks. Nonequilibrium measurements of free energy differences for microscopically reversible markovian systems. *J. Stat. Phys.*, 90(5/6):1481–1487, 1998. 313, 328, 335
- [50] G. E. Crooks. Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences. *Physical Review E*, 60(3):2721–2726, 1999. 315, 322, 328, 335
- [51] G. E. Crooks. Path-ensemble averages in systems driven far from equilibrium. *Physical Review E*, 61(3):2361–2366, 2000. 335, 339
- [52] G. E. Crooks. Beyond boltzmann-gibbs statistics: Maximum entropy hyperensembles out of equilibrium. *Physical Review E*, 75(4):041119, 2007. 378
- [53] G. E. Crooks. On thermodynamic and microscopic reversibility. *Journal of Statistical Mechanics: Theory and Experiment*, 2011(07):P07008, 2011. 313
- [54] J. P. Crutchfield. The calculi of emergence: computation, dynamics and induction. *Physica D: Nonlinear Phenomena*, 75(13):11 – 54, 1994. 87, 278, 315
- [55] J. P. Crutchfield. Between order and chaos. *Nature Physics*, 8:17–24, 2012. 1, 52, 79, 82, 87, 168, 171, 172, 205, 221, 241, 274, 373
- [56] J. P. Crutchfield. The dreams of theory. *Wiley Interdisciplinary Reviews: Computational Statistics*, 6(2):75–79, 2014. 75
- [57] J. P. Crutchfield and C. Aghamohammadi. Not all fluctuations are created equal: Spontaneous variations in thermodynamic function. 2016. Santa Fe Institute Working Paper 16-09-018; arxiv.org:1609.02519 [math-ph]. 313
- [58] J. P. Crutchfield, C. J. Ellison, and J. R. Mahoney. Time’s barbed arrow: Irreversibility, crypticity, and stored information. *Phys. Rev. Lett.*, 103(9):094101, 2009. 52, 89, 95, 204
- [59] J. P. Crutchfield, C. J. Ellison, J. R. Mahoney, and R. G. James. Synchronization and control in intrinsic and designed computation: An information-theoretic analysis of competing models of stochastic computation. *CHAOS*, 20(3):037105, 2010. 83, 85, 183

- [60] J. P. Crutchfield, C. J. Ellison, and P. M. Riechers. Exact complexity: The spectral decomposition of intrinsic computation. *Santa Fe Institute Working Paper 2013-09-028*, 2013. 19, 52, 72, 75, 78, 187, 237, 269, 310, 345, 349, 350
- [61] J. P. Crutchfield and D. P. Feldman. Synchronizing to the environment: Information theoretic limits on agent learning. *Adv. in Complex Systems*, 4(2):251–264, 2001. 81, 83
- [62] J. P. Crutchfield and D. P. Feldman. Regularities unseen, randomness observed: Levels of entropy convergence. *CHAOS*, 13(1):25–54, 2003. 4, 8, 75, 81, 82, 97, 127, 133, 225
- [63] J. P. Crutchfield and S. E. Marzen. Signatures of infinity: Nonergodicity and resource scaling in prediction, complexity, and learning. *Phys. Rev. E*, 91:050106, May 2015. 73, 120, 278
- [64] J. P. Crutchfield and K. Young. Inferring statistical complexity. *Phys. Rev. Lett.*, 63:105–108, 1989. 82, 221, 241, 251, 269, 274, 310, 322
- [65] J. P. Crutchfield and K. Young. Computation at the onset of chaos. In W. Zurek, editor, *Entropy, Complexity, and the Physics of Information*, volume VIII of *SFI Studies in the Sciences of Complexity*, pages 223 – 269, Reading, Massachusetts, 1990. Addison-Wesley. 73
- [66] P. Dayan and L. F. Abbott. *Theoretical Neuroscience*. 2005. 343, 344, 345, 350, 368
- [67] P. A. M. Dirac. Theory of electrons and positrons. In *Nobel Lecture, Physics 1922–1941*. Elsevier Publishing Company, Amsterdam, 1965. 17
- [68] K. Dornberger-Schiff and H. Schmittler. Determination of cyclicity, hexagonality and other properties of polytypes. *Acta Crystallogr. Sec. A*, 27:216–219, 1971. 230
- [69] N. Dunford. Spectral theory i. convergence to projections. *Transactions of the American Mathematical Society*, 54(2):pp. 185–217, 1943. 21, 24, 108
- [70] N. Dunford. Spectral operators. *Pacific J. Math.*, 4(3):321–354, 1954. 19, 29, 73, 186
- [71] N. Dunford and J. T. Schwartz. *Linear Operators*. Interscience Publishers, New York, 1967. 36
- [72] T. Egami and S. J. L. Billinge. *Underneath the Bragg Peaks: Structural Analysis of Complex Materials*, volume 16 of *Pergamon Materials Series*. Pergamon, New York, 2 edition, 2013. 220
- [73] A. Einstein. On the method of theoretical physics. *Philosophy of Science*, 1(2):163–169, April 1934. The Herbert Spencer Lecture, delivered at Oxford (10 June 1953). 15

- [74] R. J. Elliot, L. Aggoun, and J. B. Moore. *Hidden Markov Models: Estimation and Control*, volume 29 of *Applications of Mathematics*. Springer, New York, 1995. 221, 274
- [75] C. J. Ellison and J. P. Crutchfield. States of states of uncertainty. in preparation. 83, 89, 95, 96, 99
- [76] C. J. Ellison, J. R. Mahoney, and J. P. Crutchfield. Prediction, retrodiction, and the amount of information stored in the present. *J. Stat. Phys.*, 136(6):1005–1034, 2009. 52, 89, 95, 204, 251, 278
- [77] J. L. England. Statistical physics of self-replication. *The Journal of Chemical Physics*, 139(12):–, 2013. 315, 355
- [78] Y. Ephraim and N. Merhav. Hidden markov processes. *IEEE Trans. Inf. Theory*, 48:1518–1569, 2002. 251
- [79] M. Esposito, U. Harbola, and S. Mukamel. Entropy fluctuation theorems in driven open systems: Application to electron counting statistics. *Phys. Rev. E*, 76:031132, Sep 2007. 322
- [80] M. Esposito and C. Van den Broeck. Three detailed fluctuation theorems. *Phys. Rev. Lett.*, 104:090601, Mar 2010. 320, 322, 336, 337
- [81] E. Estevez-Rams, B. Aragon-Fernandez, H. Fuess, and A. Penton-Madrigal. Diffraction patterns of stacked layer crystals. *Physical Review B*, 68:064111, 2003. 286
- [82] E. Estevez-Rams, A. Pentón Madrigal, P. Scardi, and M. Leoni. Powder diffraction characterization of stacking disorder. *Z. Kristallogr. Suppl.*, 26:99–104, 2007. 275, 276
- [83] E. Estevez-Rams, J. Martinez, A. Penton-Madrigal, and R. Lora-Serrano. Direct solution of the diffraction pattern of a crystal with planar faulting. *Physical Review B*, 63:054109, 2001. 220, 221
- [84] E. Estevez-Rams, A. Penton-Madrigal, R. Lora-Serrano, and J. Martinez-Garcia. Direct determination of microstructural parameters from the x-ray diffraction profile of a crystal with stacking faults. *J. Appl. Crystallogr.*, 34:730–736, 2001. 220
- [85] E. Estevez-Rams, U. Welzel, A. Pentón Madrigal, and E. J. Mittemeijer. Stacking and twin faults in close-packed crystal structures: exact description of random faulting statistics for the full range of faulting probabilities. *Acta Crystallogr. Sec. A*, 64:537–548, 2008. 220, 221, 223, 243, 251, 252, 254, 258, 259, 260, 299
- [86] B. Altaner et al. Fluctuating currents in stochastic thermodynamics ii: Energy conversion and nonequilibrium response in kinesin models. *arXiv:1504.03648 [cond-mat.stat-mech]*, 2015. 356

- [87] J. D. Farmer, J. P. Crutchfield, H. Froehling, N. H. Packard, and R. S. Shaw. Power spectra and mixing properties of strange attractors. *Ann. New York Acad. Sci.*, 357:453, 1980. 69
- [88] D. P. Feldman and J. P. Crutchfield. Structural information in two-dimensional patterns: Entropy convergence and excess entropy. *Phys. Rev. E*, 67:051104, May 2003. 380
- [89] G. Ferraris, E. Makovicky, and S. Merlino. *Crystallography of Modular Materials*, volume 15 of *International Union of Crystallography Monographs on Crystallography*. Oxford University Press, Oxford, 2008. 223, 280
- [90] J. N. Franklin. *Matrix Theory*. Dover Publications, New York, 2000. 186
- [91] F. S. Galasso. *Structure, Properties and Preparation of Perovskite-Type Compounds*, volume 5 of *International Series in Monographs in Solid State Physics*. Pergamon Press, Oxford, 1969. 272
- [92] B. Gardas, S. Deffner, and A. Saxena. Non-hermitian quantum thermodynamics. *Sci. Reports*, 6:23408, 2016. 72
- [93] P. Gaspard, G. Nicolis, A. Provata, and S. Tasaki. Spectral signature of the pitchfork bifurcation: Liouville equation approach. *Phys. Rev. E*, 51:74–94, Jan 1995. 16
- [94] B. Gaveau and L.S. Schulman. A general framework for non-equilibrium phenomena: the master equation and its formal consequences. *Physics Letters A*, 229(6):347 – 353, 1997. 325, 336, 374, 375
- [95] A. K. Geim and I. V. Grigorieva. Van der waals heterostructures. *Nature*, 499:419–425, 2013. 220, 272
- [96] W. Gerstner and W. Kistler. Statistics of spike trains. In *Spiking Neuron Models*. Cambridge University Press, Cambridge, United Kingdom, 2002. 53
- [97] C. Giacovazzo, H. L. Monaco, G. Artioli, D. Viterbo, M. Milanesio, G. Ferraris, G. Gilli, P. Gilli, G. Zanotti, and M. Catti. *Fundamentals of Crystallography*, volume 15 of *IUCr Texts on Crystallography*. Oxford University Press, 3 edition, 2011. 273
- [98] H. A. Gindler. An operational calculus for meromorphic functions. *Nagoya Math. J.*, 26:31–38, 1966. 25, 36
- [99] P. Gmeiner. Equality conditions for internal entropies of certain classical and quantum models. *arXiv:1108.5303*, 2011. 168
- [100] M. S. Green. Markoff random processes and the statistical mechanics of time-dependent phenomena. II. Irreversible processes in fluids. *J. Chem. Physics*, 22(3):398–413, 1954. 59, 80

- [101] U. Grimm. Aperiodic crystals and beyond. *Acta Crystallogr. Sec. B*, 71:258–274, 2015. 284
- [102] M. Gu, K. Wiesner, E. Rieper, and V. Vedral. Quantum mechanics can reduce the complexity of classical models. *Nature communications*, 3:762, 2012. 84, 168, 169, 174, 212
- [103] A. Guinier. *X-Ray Diffraction in Crystals, Imperfect Crystals, and Amorphous Bodies*. W. H. Freeman and Company, New York, 1963. 241, 249, 273, 297
- [104] M. Haase. Spectral mapping theorems for holomorphic functional calculi. *J. London Math. Soc.*, 71(3):723–739, 2005. 23
- [105] R. R. Halmos. *Finite-Dimensional Vector Spaces*. D. Van Nostrand Company, 1958. 18, 20, 42
- [106] R. J. Harris and G. M. Schtz. Fluctuation theorems for stochastic dynamics. *Journal of Statistical Mechanics: Theory and Experiment*, 2007(07):P07020, 2007. 336
- [107] S. Hassani. *Mathematical Physics*. Springer, New York, 1999. 18, 22, 29
- [108] T. Hatano and S. Sasa. Steady-state thermodynamics of langevin systems. *Phys. Rev. Lett.*, 86:3463–3466, Apr 2001. 314, 320, 325, 336, 339, 365, 375
- [109] S. Hendricks and E. Teller. X-ray interference in partially ordered layer lattices. *J. Chem. Phys.*, 10:147–167, 1942. 219, 276, 279
- [110] J. P. Hirth and J. Lothe. *Theory of Dislocations*. McGraw-Hill, New York, 2 edition, 1968. 261, 273, 305
- [111] A. L. Hodgkin and A. F. Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *J. Physio.*, 117(4):500, 1952. 343, 344, 345
- [112] J. E. Hopcroft and J. D. Ullman. *Introduction to Automata Theory, Languages, and Computation*. Addison-Wesley, Reading, 1979. 87, 225, 226, 251, 274, 283
- [113] J. M. Horowitz and S. Vaikuntanathan. Nonequilibrium detailed fluctuation theorem for repeated discrete feedback. *Phys. Rev. E*, 82:061120, Dec 2010. 316, 377
- [114] W. Horsthemke. Noise induced transitions. In C. Vidal and A. Pacault, editors, *Non-Equilibrium Dynamics in Chemical Systems: Proceedings of the International Symposium, Bordeaux, France, September 3–7, 1984*, pages 150–160, Berlin, Heidelberg, 1984. Springer. 313
- [115] D. Hull and D. J. Bacon. *Introduction to Dislocations*. Butterworth-Heinemann, New York, fifth edition, 2011. 273

- [116] H. P. Iwata, U. Lindefelt, S. Öberg, and P. R. Briddon. Ab initio study of 3c inclusions and stacking fault/stacking fault interactions in 6h-sic. *J. Appl. Phys.*, 94:4972–4979, 2003. 261
- [117] E. M. Izhikevich. *Dynamical Systems in Neuroscience*. Computational Neuroscience Series. MIT Press, Boston, Massachusetts, 2010. 343
- [118] H. Jagodzinski. Eindimensionale fehlordnung in kristallen und ihr einfluss auf die röntgeninterferenzen. I. Berechnung des fehlordnungsgrades aus den röntgenintensitäten. *Acta Crystallogr.*, 2:201–207, 1949. 276, 279
- [119] H. Jagodzinski. Eindimensionale fehlordnung in kristallen und ihr einfluss auf die röntgeninterferenzen. II. Berechnung der fehlgeordneten dichtesten kugelpackungen mit wechselwirkungen der reichweite 3. *Acta Crystallogr.*, 2:208–214, 1949. 276, 279
- [120] R. G. James, J. R. Mahoney, C. J. Ellison, and J. P. Crutchfield. Many roads to synchrony: Natural time scales and their algorithms. *Phys. Rev E*, 89:042135, 2014. 63, 174, 175, 186, 204, 280
- [121] C. Jarzynski. Nonequilibrium equality for free energy differences. *Phys. Rev. Lett.*, 78(14):2690–2693, 1997. 338
- [122] R. Jozsa and J. Schlienz. Distinguishability of states and von Neumann entropy. *Phys. Rev. A*, 62:012301, 2000. 191, 210
- [123] V. K. Kabra and Dhananjai Pandey. Long-range ordered phases without short-range correlations. *Physical Review Letters*, 61:1493–1496, 1988. 220, 221, 223, 285
- [124] V. K. Kabra and Dhananjai Pandey. A computer simulation approach for the calculation of diffuse intensity distributions from crystals undergoing the 2h to 6h transformation by the layer displacement mechanism. *Acta Crystallogr. Sec. A*, 51:329–335, 1995. 220
- [125] H. Kantz and T. Schreiber. *Nonlinear Time Series Analysis*. Cambridge University Press, Cambridge, 2 edition, 2004. 279
- [126] S. Karlin and H. M. Taylor. *A First Course in Stochastic Processes*. Academic Press, New York, 2 edition, 1975. 224
- [127] H. Katzke, P. Tolédano, and W. Depmeier. Phase transitions between polytypes and intralayer superstructures in transition metal dichalcogenides. *Phys. Rev. B*, 69:134111, 2004. 272
- [128] D. A. Keen and R. L. McGreevy. Structural modelling of glasses using reverse Monte Carlo simulation. *Nature*, 344:423–425, 1990. 279
- [129] J. G. Kemeny and J. L. Snell. *Finite markov chains*, volume 356. 1960. 36, 111

- [130] A. N. Kolmogorov. Entropy per unit time as a metric invariant of automorphisms. *Dokl. Akad. Nauk. SSSR*, 124:754, 1959. (Russian) Math. Rev. vol. 21, no. 2035b. 81
- [131] B. O. Koopman. Hamiltonian systems and transformation in hilbert space. *Proceedings of the National Academy of Sciences*, 17(5):315–318, 1931. 16
- [132] C. S. Kubrusly. *Spectral theory of operators on Hilbert spaces*. Springer Science & Business Media, 2012. 22
- [133] Y. Kuwahara. Comparison of the surface structure of the tetrahedral sheets of muscovite and phlogopite by AFM. *Physics and Chemistry of Minerals*, 28(1):1–8, 2001. 272
- [134] D. Lacoste, A. W.C. Lau, and K. Mallick. Fluctuation theorem and large deviation function for a solvable model of a molecular motor. *Phys. Rev. E*, 78:011915, Jul 2008. 355
- [135] S. Lahiri and S. Ganguli. A memory frontier for complex synapses. In C. J. C. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K. Q. Weinberger, editors, *Advances in Neural Information Processing Systems 26*, pages 1034–1042. Curran Associates, Inc., 2013. 356
- [136] S. Lahiri and A. M. Jayannavar. Fluctuation theorems for excess and housekeeping heat for underdamped langevin systems. *The European Physical Journal B*, 87(9), 2014. 338
- [137] S. Lahiri, J. Sohl-Dickstein, and S. Ganguli. A universal tradeoff between power, precision and speed in physical communication. *arXiv:1603.07758 [cond-mat.stat-mech]*, 2016. 60
- [138] L. Landau. The scattering of X-rays by crystals with a variable lamellar structure. *Phys. Z. Sowjetunion*, 12:579, 1937. 275, 279
- [139] R. Landauer. Statistical physics of machinery: forgotten middle-ground. *Physica A: Statistical Mechanics and its Applications*, 194(1-4):551–562, 1993. 313
- [140] B. Latni. *Signal processing and linear systems*. Oxford University Press, USA, 1998. 20, 27, 134
- [141] C.-B. Li, H. Yang, and T. Komatsuzaki. Multiscale complex network of protein conformational fluctuations in single-molecule time series. *Proceedings of the National Academy of Sciences USA*, 105:536–541, 2008. 1
- [142] M. Li and P. M. B. Vitanyi. *An Introduction to Kolmogorov Complexity and its Applications*. Springer-Verlag, New York, 1993. 52
- [143] S. Liepelt and R. Lipowsky. Kinesin’s network of chemomechanical motor cycles. *Phys. Rev. Lett.*, 98:258102, Jun 2007. 327

- [144] Liepelt, S. and Lipowsky, R. Steady-state balance conditions for molecular motor cycles and stochastic nonequilibrium processes. *EPL*, 77(5):50002, 2007. 327
- [145] I. M. Lifschitz. On the theory of scattering of X-rays by crystals of variable structure. *Phys. Z. Sowjetunion*, 12:623, 1937. 275, 279
- [146] D. Lind and B. Marcus. *An Introduction to Symbolic Dynamics and Coding*. Cambridge University Press, New York, 1995. 133
- [147] G. Lindblad. On the generators of quantum dynamical semigroups. *Comm. Math. Physics*, 48(2):119–130, 1976. 60
- [148] B. Lindner, J. Garcia-Ojalvo, A. Neiman, and L. Schimansky-Geier. Effects of noise in excitable systems. *Physics Reports*, 392(6):321 – 424, 2004. 313
- [149] J. Liu, M. Kargarian, M. Kareev, B. Gray, P. J. Ryan, A. Cruz, N. Tahir, Y. Chuang, J. Guo, J. M. Rondinelli, J. W. Freeland, G. A. Fiete, and J. Chakhalian. Heterointerface engineered electronic and magnetic phases of NdNiO₃ thin films. *Nat. Comm.*, 4:2714, 2013. 272
- [150] W. Lohr and N. Ay. Non-sufficient memories that are sufficient for prediction. In *Complex Sciences*, volume 4 of *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, pages 265–276. 2009. 84
- [151] E. N. Lorenz. Deterministic nonperiodic flow. *J. Atmos. Sci.*, 20(2):130–141, 1963. 69
- [152] P. Löwdin. On operators, superoperators, Hamiltonians, and Liouvillians. *Intl. J. Quant. Chem.*, 22(S16):485–560, 1982. 60
- [153] Z. Lu, D. Mandal, and C. Jarzynski. Engineering maxwell’s demon. *Physics Today*, 67(8):60, 2014. 377
- [154] A. L. Mackay. Generalized crystallography. *Computers & Mathematics with Applications*, B12(1/2):21–37, 1986. 221, 269, 274
- [155] A. L. Mackay. Generalized crystallography. *Structural Chemistry*, 13(3/4):215–220, 2002. 269
- [156] M. C. Mackey. *Time’s arrow: the origins of thermodynamic behavior*. Courier Corporation, 2003. 69
- [157] J. R. Mahoney, C. Aghamohammadi, and J. P. Crutchfield. Occam’s quantum stop: Synchronizing and compressing classical cryptic processes via a quantum channel. *Scientific Reports*, 6:20495, 2016. 84, 168, 169, 170, 173, 174, 175, 177, 191, 203
- [158] J. R. Mahoney, C. J. Ellison, and J. P. Crutchfield. Information accessibility and cryptic processes. *J. Phys. A: Math. Theo.*, 42:362002, 2009. 101, 168, 184

- [159] J. R. Mahoney, C. J. Ellison, and J. P. Crutchfield. Information accessibility and cryptic processes: Linear combinations of causal states. 2009. [arxiv.org:0906.5099](https://arxiv.org/abs/0906.5099) [cond-mat]. 101
- [160] J. R. Mahoney, C. J. Ellison, R. G. James, and J. P. Crutchfield. How hidden are hidden processes? a primer on crypticity and entropy convergence. *CHAOS*, 21(3):037112, 2011. 83, 101, 168, 184
- [161] T. L. Malkin, B. J. Murray, A. V. Brukhno, J. Anwar, and C. G. Salzmann. Structure of ice crystallized from supercooled water. *Proceedings of the National Academy of Sciences of the United States of America*, 109(4):1041–1045, 2012. 273
- [162] D. Mandal and C. Jarzynski. Work and information processing in a solvable model of maxwells demon. *Proceedings of the National Academy of Sciences*, 109(29):11641–11645, 2012. 381
- [163] D. Mandal and C. Jarzynski. Analysis of slow transitions between nonequilibrium steady states. *Journal of Statistical Mechanics: Theory and Experiment*, 2016(6):063204, 2016. 60, 314, 320, 322, 325, 365, 375
- [164] S. Marzen and J. P. Crutchfield. Informational and causal architecture of discrete-time renewal processes. *Entropy*, 17(7):4891–4917, 2015. 57
- [165] S. Marzen and J. P. Crutchfield. Informational and causal architecture of continuous-time renewal and hidden semi-markov processes. 2016. [arxiv.org:1611.01099](https://arxiv.org/abs/1611.01099). 97
- [166] S. Marzen and J. P. Crutchfield. Statistical signatures of structural organization: The case of long memory in renewal processes. *Phys. Lett. A*, 380(17):1517–1525, 2016. 57, 73
- [167] S. E. Marzen and J. P. Crutchfield. Predictive rate-distortion for infinite-order markov processes. *Journal of Statistical Physics*, 163(6):1312–1338, 2016. 83
- [168] S. E. Marzen, M. R. DeWeese, and J. P. Crutchfield. Time resolution dependence of information measures for spiking neurons: scaling and universality. *Frontiers in Computational Neuroscience*, 9(105), 2015. 57, 97
- [169] C. D. Meyer. *Matrix analysis and applied linear algebra*. Siam, 2000. 19, 20, 24, 44
- [170] T. M. Michels-Clark, V. E. Lynch, C. M. Hoffmann, J. Hauser, T. Weber, R. Harrison, and H. B. Bürgi. Analyzing diffuse scattering with supercomputers. *J. Appl. Crystallogr.*, 46:1616–1625, 2013. 279
- [171] A.G. Milnes. Semiconductor heterojunction topics: Introduction and overview. *Solid-State Electronics*, 29(2):99 – 121, 1986. 17
- [172] E. H. Moore. On the reciprocal of the general algebraic matrix. *Bull. Am. Math. Soc.*, 26, 1920. 33

- [173] A. Mostafazadeh. Spectral singularities of complex scattering potentials and infinite reflection and transmission coefficients at real energies. *Phys. Rev. Lett.*, 102:220402, 2009. 68
- [174] Y. Murashita, K. Funo, and M. Ueda. Nonequilibrium equalities in absolutely irreversible processes. *Phys. Rev. E*, 90:042110, Oct 2014. 355
- [175] B. Nagy. On an operational calculus for meromorphic functions. *Acta Math. Acad. Sci. Hungarica*, 33(3):379–390, 1979. 25, 36
- [176] E. Narevicius, P. Serra, and N. Moiseyev. Critical phenomena associated with self-orthogonality in non-Hermitian quantum mechanics. *EPL (Europhys. Lett.)*, 62(6):789, 2003. 68
- [177] J. von Neumann. Zur algebra der funktionaloperationen und theorie der normalen operatoren. *Math. Annalen*, 102:370–427, 1930. 18
- [178] M. Newman. *Networks: an introduction*. Oxford University Press, Oxford, United Kingdom, 2010. 72
- [179] B. I. Nikolin and A. Y. Babkevich. A Monte Carlo simulation of random stacking faults in close-packed structures. *Acta Crystallogr. Sec. A*, 45:797–801, 1989. 279
- [180] T. Nishikawa and A. E. Motter. Synchronization is optimal in nondiagonalizable networks. *Phys. Rev. E*, 73:065106, Jun 2006. 72
- [181] Y. Oono and M. Paniconi. Steady state thermodynamics. *Progress of Theoretical Physics Supplement*, 130:29–44, 1998. 61, 314
- [182] A. V. Oppenheim and R. W. Schaffer. *Digital Signal Processing*. Prentice-Hall, Englewood Cliffs, 1975. 237, 291
- [183] A. L. Ortiz, F. Sánchez-Bajo, F. L. Cumbrera, and F. Guiberteau. The prolific polytypism of silicon carbide. *J. Appl. Crystallogr.*, 46:242–247, 2013. 220, 223, 261, 279, 281
- [184] N. H. Packard, J. P. Crutchfield, J. D. Farmer, and R. S. Shaw. Geometry from a time series. *Phys. Rev. Lett.*, 45:712–716, Sep 1980. 9
- [185] M. S. Palsson, M. Gu, J. Ho, H. M. Wiseman, and Geoff J. Pryde. Experimental quantum processing enhancement in modelling stochastic processes. *arXiv:1602.05683 [quant-ph]*, 2016. 84, 169, 174
- [186] D. Pan, S. Wang, B. Zhao, M. Wu, H. Zhang, Y. Wang, and Z. Jiao. Li storage properties of disordered graphene nanosheets. *Chemistry of Materials*, 21(14):3136–3142, 2009. 273
- [187] E. Parzen. *Stochastic processes*, volume Classics in Applied Mathematics; 24. SIAM Classics in Applied Mathematics, 1999. 3

- [188] R. K. Pathria and P. D. Beale. *Statistical mechanics, 3rd Edition*. Academic Press, 2011. 373
- [189] J. Patlak. Molecular kinetics of voltage-dependent na⁺ channels. *Physiological Reviews*, 71(4):1047–1080, 1991. 344, 368
- [190] A. Paz. *Introduction to Probabilistic Automata*. Academic Press, New York, 1971. 224, 274, 283
- [191] R. Penrose. A generalized inverse for matrices. *Math. Proc. Cambridge Phil. Soc.*, 51, 1955. 33
- [192] T. Petrosky and I. Prigogine. The Liouville space extension of quantum mechanics. *Adv. Chem. Phys*, 99:1–120, 1997. 60
- [193] M. Polettini and M. Esposito. Irreversible thermodynamics of open chemical networks. i. emergent cycles and broken conservation laws. *The Journal of Chemical Physics*, 141(2), 2014. 313
- [194] G. D. Price. Polytypism and the factors determining the stability of spinelloid structures. *Phys. Chem. Minerals*, 10:77–83, 1983. 223
- [195] H. Qian. Nonequilibrium steady-state circulation and heat dissipation functional. *Phys. Rev. E*, 64:022101, Jul 2001. 313
- [196] H. Qian. Cycle kinetics, steady state thermodynamics and motors a paradigm for living matter physics. *Journal of Physics: Condensed Matter*, 17(47):S3783, 2005. 327
- [197] L. R. Rabiner. A tutorial on hidden Markov models and selected applications. *IEEE Proc.*, 77:257, 1989. 221, 274
- [198] P. M. Riechers and J. P. Crutchfield. Beyond the spectral theorem: Spectrally decomposing arbitrary functions of nondiagonalizable operators. *arXiv:1607.06526*, 2016. 75, 77, 103, 108, 110, 111, 113, 128, 166, 189, 237, 269, 270, 345, 349, 350
- [199] P. M. Riechers and J. P. Crutchfield. Fluctuations when driving between nonequilibrium steady states. *arXiv: 1610.09444*, 2016. 98, 365, 374, 375
- [200] P. M. Riechers and J. P. Crutchfield. Spectral simplicity of hidden complexity: Cracking the nondiagonalizable metadynamics of prediction. 2016. 53
- [201] P. M. Riechers and J. P. Crutchfield. Exact excess, renormalized housekeeping. in preparation. 57, 61, 98
- [202] P. M. Riechers and J. P. Crutchfield. Power spectra of stochastic processes from transition matrices of hidden markov models. in preparation. 92, 94

- [203] P. M. Riechers, J. R. Mahoney, C. Aghamohammadi, and J. P. Crutchfield. Minimized state complexity of quantum-encoded cryptic processes. *Phys. Rev. A*, 93:052317, May 2016. 19, 53, 72, 84, 102
- [204] P. M. Riechers, D. P. Varn, and J. P. Crutchfield. Diffraction patterns of layered close-packed structures from hidden markov models. 2014. Santa Fe Institute Working Paper 2014-10-038; arxiv.org:1410.5028 [cond-mat.mtrl-sci]. 19, 62, 65, 67, 131, 134, 231, 269
- [205] P. M. Riechers, D. P. Varn, and J. P. Crutchfield. Pairwise correlations in layered close-packed structures. *Acta Cryst. A*, 71:423–443, 2015. 19, 62, 281, 282, 283, 286, 288, 289, 292, 293, 294, 299, 300, 301, 303, 306, 310, 311
- [206] F. Rieke, D. Warland, R. de Ruyter van Steveninck, and W. Bialek. *Spikes: Exploring the Neural Code*. Bradford Books, New York, 1999. 343
- [207] U. G. Rothblum. A representation of the Drazin inverse and characterizations of the index. *SIAM J. App. Math.*, 31(4):646–648, 1976. 35
- [208] D. Ruelle and F. Takens. On the nature of turbulence. *Comm. Math. Physics*, 20(3):167–192, 1971. 69
- [209] T. Sagawa and M. Ueda. Nonequilibrium thermodynamics of feedback control. *Phys. Rev. E*, 85:021104, Feb 2012. 313, 316, 342, 377
- [210] J. J. Sakurai and J. J. Napolitano. *Modern Quantum Mechanics*. Addison-Wesley, San Francisco, California, 2011. 42
- [211] A. Sandage and G. A. Tammann. Steps toward the hubble constant. vii-distances to ngc 2403, m101, and the virgo cluster using 21 centimeter line widths compared with optical methods: The global value of h sub 0. *Astrophys. J.*, 210:7–24, 1976. 17
- [212] P. Sartori, L. Granger, C. F. Lee, and J. M. Horowitz. Thermodynamic costs of information processing in sensory adaptation. *PLoS Comput Biol*, 10(12):e1003974, 12 2014. 342
- [213] B. Schumacher. Quantum coding. *Phys. Rev. A*, 51:2738–2747, 1995. 174, 193
- [214] M. T. Sebastian and P. Krishna. An x-ray diffraction study of faulting in single crystal of cubic zns grown from the vapour phase. *Philos. Mag. A*, 49:809–821, 1984. 266
- [215] M. T. Sebastian and P. Krishna. Diffraction effects from hcp crystals with random growth faults undergoing transformation to the fcc phase by the deformation mechanism. *Phys. Stat. Sol. A*, 101:329–337, 1987. 266

- [216] M. T. Sebastian and P. Krishna. X-ray investigation of the mechanism of phase transformation in single crystals of ZnS , $\text{Zn}_x\text{Cd}_{1-x}$ and $\text{Zn}_x\text{Mn}_{1-x}\text{S}$ (i) calculation of diffraction effects by a three parameter model. *Crys. Res. Tech.*, 22:929–941, 1987. 266
- [217] M. T. Sebastian and P. Krishna. X-ray investigation of the mechanism of phase transformation in single crystals of ZnS , $\text{Zn}_x\text{Cd}_{1-x}$ and $\text{Zn}_x\text{Mn}_{1-x}\text{S}$ (ii) comparison of observed and calculated diffraction effects. *Crys. Res. Tech.*, 22:1063–1072, 1987. 266
- [218] M. T. Sebastian and P. Krishna. *Random, Non-Random and Periodic Faulting in Crystals*. Gordon and Breach, The Netherlands, 1994. 220, 223, 272, 273, 279, 280, 304
- [219] M. T. Sebastian, K. Narayanan, and P. Krishna. X-ray diffraction effects from 2h crystals undergoing transformation to the 3c structure by the layer displacement mechanism. *Phys. Stat. Sol. A*, 102:241–249, 1987. 266
- [220] E. G. Seebauer and K. W. Noh. Trends in semiconductor defect engineering at the nanoscale. *Mater. Sci. Eng. R*, 70:151–168, 2010. 273
- [221] U. Seifert. Entropy production along a stochastic trajectory and an integral fluctuation theorem. *Phys. Rev. Lett.*, 95:040602, Jul 2005. 337
- [222] U. Seifert. Stochastic thermodynamics, fluctuation theorems and molecular machines. *Reports on Progress in Physics*, 75(12):126001, 2012. 314
- [223] C. R. Shalizi and J. P. Crutchfield. Computational mechanics: Pattern and prediction, structure and simplicity. *J. Stat. Phys.*, 104:817–881, 2001. 251
- [224] C. R. Shalizi and A. Kontorovich. *Almost none of the theory of stochastic processes*. 2010. 146
- [225] C. R. Shalizi, K. L. Shalizi, and J. P. Crutchfield. Pattern discovery in time series, Part I: Theory, algorithm, analysis, and convergence. *Santa Fe Institute Working Paper 02-10-060*; arXiv.org/abs/cs.LG/0210025, 2002. 269
- [226] C. E. Shannon. A mathematical theory of communication. *Bell Sys. Tech. J.*, 27:379–423, 623–656, 1948. 4
- [227] Q. Shen, Q. Hao, and S. M. Gruner. Macromolecular phasing. *Physics Today*, 59(3):46–52, 2006. 357
- [228] S. P. Shrestha and D. Pandey. Application of 1d kinetic ising model for studying the evolution of diffuse scattering during hcp (2h) to fcc (3c) martensitic transition. *Europhys. Lett.*, 34(4):269–274, 1996. 220

- [229] S. P. Shrestha and D. Pandey. Monte carlo study of the evolution of diffuse scattering and coherent modulation during hcp to fcc martensitic transition – ii. domain growth kinetics by spinodal process. *Acta Mater.*, 44:4949–4960, 1996. 220
- [230] S. P. Shrestha and D. Pandey. Kinetics of martensitic type restacking transitions: Dynamic scaling, universal growth exponent and evolution of diffuse scattering. *Proc. R. Soc. Lond. A*, 453:1311–1330, 1997. 220, 221
- [231] S. P. Shrestha, V. Tripathi, V. K. Kabra, and D. Pandey. Monte carlo study of the evolution of diffuse scattering and coherent modulation during hcp to fcc martensitic transition – i. infinitely strong repulsive model. *Acta Mater.*, 44:4937–4947, 1996. 220
- [232] M. Sipser. *Introduction to the Theory of Computation*. Cengage Learning, 2012. 85, 87
- [233] L. Sirovich and M. Kirby. Low-dimensional procedure for the characterization of human faces. *J. Opt. Soc. Am. A*, 4(3):519–524, Mar 1987. 18
- [234] D. A. Sivak and G. E. Crooks. Near-equilibrium measurements of nonequilibrium free energy. *Physical Review Letters*, 108(15), 2012. 325, 374, 375
- [235] W. L. Smith. Renewal theory and its ramifications. *J. Roy. Stat. Soc. B*, 20(2):243–302, 1958. 53
- [236] A. V. Sokolov, A. A. Andrianov, and F. Cannata. Non-Hermitian quantum mechanics of non-diagonalizable Hamiltonians: Puzzles with self-orthogonal states. *J. Physics A*, 39(32):10207, 2006. 68
- [237] T. Speck and U. Seifert. Integral fluctuation theorem for the housekeeping heat. *Journal of Physics A: Mathematical and General*, 38(34):L581, 2005. 339
- [238] R. Spinney and I. Ford. *Fluctuation Relations: A Pedagogical Overview*, pages 3–56. Wiley-VCH Verlag GmbH & Co. KGaA, 2013. 314
- [239] H. Stark, W. R. Bennett, and M. Arm. Design considerations in power spectra measurements by diffraction of coherent light. *Appl. Opt.*, 8(11):2165–2172, Nov 1969. 62
- [240] S. Still, J. P. Crutchfield, and C. J. Ellison. Optimal causal inference: Estimating stored information and approximating causal architecture. *CHAOS*, 20(3):037111, 2010. 83
- [241] S. Still, D. A. Sivak, A. J. Bell, and G. E. Crooks. Thermodynamics of prediction. *Phys. Rev. Lett.*, 109:120604, Sep 2012. 321, 377
- [242] P. Stoica and R. L. Moses. *Spectral Analysis of Signals*. Pearson Prentice Hall, Upper Saddle River, New Jersey, 2005. 61, 80

- [243] C. C. Streliaoff and J. P. Crutchfield. Bayesian structural inference for hidden processes. *Phys. Rev. E*, 89:042119, Apr 2014. 164, 269
- [244] S. H. Strogatz. *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering*. Westview Press, 2001. 274
- [245] W. Y. Suen, J. Thompson, A. J. P. Garner, V. Vedral, and M. Gu. The classical-quantum divergence of complexity in the Ising spin chain. *arXiv:1511.05738*, 2015. 169
- [246] J. W. Sun, T. Robert, A. Andreadou, A. Mantzari, V. Jokubavicius, R. Yakimova, J. Camassel, S. Juillaguet, E. K. Polychroniadis, and M. Syväjärvi. Shockley-frank stacking faults in 6h-sic. *J. Appl. Phys.*, 111:113527, 2012. 261, 262, 263, 305
- [247] F. Takens. Detecting strange attractors in turbulence. In *Dynamical systems and turbulence, Warwick 1980*, pages 366–381. Springer, 1981. 9
- [248] P. Tiwary and D. Pandey. Scaling behaviour of pair correlation functions for randomly faulted hexagonal close-packed structures. *Acta Crystallogr. Sec. A*, 63:289–296, 2007. 220, 221, 240
- [249] N. Travers and J. P. Crutchfield. Asymptotic synchronization for finite-state sources. *J. Stat. Phys.*, 145(5):1202–1223, 2011. 102
- [250] N. Travers and J. P. Crutchfield. Infinite excess entropy processes with countable-state generators. *Entropy*, 16:1396–1413, 2014. 73
- [251] N. F. Travers. Exponential bounds for convergence of entropy rate approximations in hidden markov models satisfying a path-mergeability condition. *Stochastic Processes and their Applications*, 124(12):4149 – 4170, 2014. 88
- [252] M. M. J. Treacy, J. M. Newsam, and M. W. Deem. A general recursion method for calculating diffracted intensities from crystals containing planar faults. *Proc. R. Soc. London, Ser. A*, 433:499–520, 1991. 275, 276, 279, 286, 288, 308
- [253] L. N. Trefethen. Pseudospectra of linear operators. *SIAM Review*, 39(3):383–406, 1997. 73
- [254] L. N. Trefethen and M. Embree. *Spectra and pseudospectra: The behavior of nonnormal matrices and operators*. Princeton University Press, Princeton, New Jersey, 2005. 73
- [255] N. Trefethen. Favorite eigenvalue problems. *SIAM News*, 44(10), Dec 2011. 17
- [256] E. H. Trepagnier, C. Jarzynski, F. Ritort, G. E. Crooks, C. J. Bustamante, and J. Liphardt. Experimental test of hatano and sasa’s nonequilibrium steady-state equality. *Proceedings of the National Academy of Sciences of the United States of America*, 101(42):15038–15041, 2004. 314, 320, 365, 375

- [257] G. C. Trigunayat. A survey of the phenomenon of polytypism in crystals. *Solid State Ionics*, 48(1/2):3–70, 1991. 273
- [258] C. Tsallis, F. Baldovin, R. Cerbino, and P. Pierobon. Introduction to nonextensive statistical mechanics and thermodynamics. *Proceedings of the International School of Physics "Enrico Fermi"*, 155: The Physics of Complex Systems (New Advances and Perspectives:229 – 252, 2003. 378
- [259] D. R. Upper. *Theory and Algorithms for Hidden Markov Models and Generalized Hidden Markov Models*. PhD thesis, University of California, Berkeley, 1997. Published by University Microfilms Intl, Ann Arbor, Michigan. 85, 89
- [260] S. Vaikuntanathan and C. Jarzynski. Dissipation and lag in irreversible processes. *EPL (Europhysics Letters)*, 87(6):60005, 2009. 340
- [261] B. Vanluyten, J. C. Willems, and B. De Moor. Equivalence of state representations for hidden markov models. *Systems & Control Letters*, 57(5):410 – 419, 2008. 85
- [262] D. P. Varn. *Language Extraction from ZnS*. PhD thesis, University of Tennessee, Knoxville, 2001. 249
- [263] D. P. Varn and G. S. Canright. The crystal problem for polytypes. *Acta Crystallogr. Sec. A*, 57:4–19, 2001. 225, 280, 285
- [264] D. P. Varn, G. S. Canright, and J. P. Crutchfield. Discovering planar disorder in close-packed structures from x-ray diffraction: Beyond the fault model. *Physical Review B*, 66:174110, 2002. 220, 221, 226, 238, 266, 269, 274, 279
- [265] D. P. Varn, G. S. Canright, and J. P. Crutchfield. Inferring planar disorder in close-packed structures *via* ϵ -machine spectral reconstruction theory: structure and intrinsic computation in zinc sulfide. *Acta Crystallogr. Sec. B*, 63:169–182, 2007. 221, 226, 238, 266, 269, 279, 310
- [266] D. P. Varn, G. S. Canright, and J. P. Crutchfield. ϵ -machine spectral reconstruction theory: a direct method for inferring planar disorder and structure from x-ray diffraction studies. *Acta Crystallogr. Sec. A*, 69:197–206, 2013. 220, 221, 226, 230, 231, 238, 241, 266, 269, 279, 281, 286, 310
- [267] D. P. Varn and J. P. Crutchfield. From finite to infinite range order via annealing: The causal architecture of deformation faulting in annealed close-packed crystals. *Phys. Lett. A*, 324(4):299–307, 2004. 220, 221, 226, 280
- [268] D. P. Varn and J. P. Crutchfield. Chaotic crystallography: How the physics of information reveals structural order in materials. *Curr. Opin. Chem. Eng.*, 7:47–56, 2015. 221, 269, 273, 274
- [269] D. P. Varn and J. P. Crutchfield. Graphical representations of stacking structure in disordered layered materials, 2015. manuscript in preparation. 273, 279, 294

- [270] Dowman P. Varn, Geoffrey S. Canright, and James P. Crutchfield. Inferring planar disorder in close-packed structures *via* ϵ -machine spectral reconstruction theory: examples from simulated diffraction spectra. *Acta Crystallogr. Sec. A*, 69:413–426, 2013. 225, 226, 238, 240, 266, 269, 277, 278, 279, 310
- [271] J. von Neumann. *Mathematical Foundations of Quantum Mechanics*. Princeton University Press, Princeton, New Jersey, 1955. 18
- [272] H. Wang and G. Oster. Energy transduction in the f1 motor of atp synthase. *Nature*, 396(6708):279–282, 1998. 313
- [273] B. E. Warren. *X-Ray Diffraction*. Addison-Wesley, 1969. 223, 251, 266, 273, 294
- [274] B. Weiss. Subshifts of finite type and sofic systems. *Monastsh. Math.*, 77:462, 1973. 278
- [275] T. R. Welberry. One hundred years of diffuse x-ray scattering. *Metallurgical and Materials Transactions A*, 45:75–84, 2014. 273
- [276] K. Wiesner. Nature computes: Information processing in quantum dynamical systems. *Chaos*, 20(3), 2010. 84
- [277] S. G. Williams. Introduction to symbolic dynamics. In *Proceedings of Symposia in Applied Mathematics*, volume 60, pages 1–12, 2004. 135
- [278] A. J. C. Wilson. Imperfection in the structure of cobalt ii. mathematical treatment of the proposed structure. *Proc. R. Soc. Ser. A*, 180:277–285, 1942. 219, 276, 279
- [279] J. Yeomans. The theory and application of axial Ising models. *Solid State Physics*, 41:151–200, 1988. 285
- [280] J. Yi and G. S. Canright. Possible disordered ground states for layered solids and their diffraction patterns. *Physical Review B*, 53:5198–5210, 1996. 223, 224, 230, 285, 286
- [281] K. Yosida. *Functional Analysis*. Classics in Mathematics. Cambridge University Press, 1995. 190
- [282] K. Zekentes and K. Rogdakis. Sic nanowires: material and devices. *Journal of Physics D: Applied Physics*, 44(13):133001, 2011. 220
- [283] R. Zwanzig. Memory effects in irreversible thermodynamics. *Phys. Rev.*, 124:983–992, Nov 1961. 1
- [284] R. Zwanzig. Time-correlation functions and transport coefficients in statistical mechanics. *Ann. Rev. Phys. Chemistry*, 16(1):67–102, 1965. 59, 80