Drazin Inverse per Spectral perSpective

with Applications to Complex Systems

Paul M. Riechers

Joint work with James P. Crutchfield

Complexity Sciences Center
Department of Physics
University of California, Davis

October 23$^{\text{rd}}$ 2015

Support: ARO contracts W911NF-12-1-0234 and W911NF-13-1-0390
and a grant from the John Templeton Foundation

Discussion based on:

J. P. Crutchfield, C. J. Ellison and P. M. Riechers.
"Exact complexity: the spectral decomposition of intrinsic computation." arXiv:1309.3792 [cond-mat.stat-mech].

and

P. M. Riechers and J. P. Crutchfield.
"Spectral decomposition of structural complexity: the meromorphic functional calculus of nondiagonable dynamics." (In preparation.)

How to Drazin:

Invert what's invertible;
leave the rest alone.

Many complex stochastic processes of interest can be modeled by a hidden Markov model (HMM).

Any HMM will have:

- some set of states $\boldsymbol{\mathcal{S}}$,
- an alphabet $\mathcal{A}$ of observables,
- a set of output-labeled transition matrices $T^{\mathcal{A}} = \left\{ T^{(x)} : T^{(x)}_{i,j} = \Pr(\mathcal{S}_t = \sigma^j | \mathcal{S}_{t-1} = \sigma^i) \right\}_{x \in \mathcal{A}}$ constituting the row-stochastic state-to-state transition matrix $T = \sum_{x \in \mathcal{A}} T^{(x)}$.

Any HMM will have:

- some set of states $\boldsymbol{\mathcal{S}}$,
- an alphabet $\mathcal{A}$ of observables,
- a set of output-labeled transition matrices $T^{\mathcal{A}} = \left\{ T^{(x)} : T^{(x)}_{i,j} = \Pr(\mathcal{S}_t = \sigma^j | \mathcal{S}_{t-1} = \sigma^i) \right\}_{x \in \mathcal{A}}$ constituting the row-stochastic state-to-state transition matrix $T = \sum_{x \in \mathcal{A}} T^{(x)}$.

In principle, can calculate everything about a process directly from HMM, but *different questions require different HMM representations.*

Question about process reduces to function of appropriate transition dynamic.

The Zero Eigenvalue, (di)Graphically

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

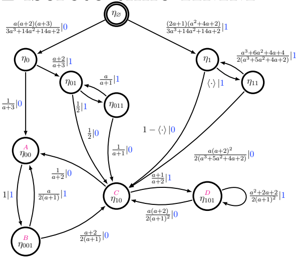
The Zero Eigenvalue, (di)Graphically

$$\Lambda_A = \left\{ \lambda \in \mathbb{C} : \det(A - \lambda I) = 0 \right\}$$
$$\text{So, } 0 \in \Lambda_A \implies \det(A) = 0 \implies A^{-1} \neq \mathrm{inv}(A)$$

The Zero Eigenvalue, (di)Graphically

$$\Lambda_A = \left\{ \lambda \in \mathbb{C} : \det(A - \lambda I) = 0 \right\}$$
$$\text{So, } 0 \in \Lambda_A \implies \det(A) = 0 \implies A^{-1} \neq \operatorname{inv}(A)$$

Discrete time HMM

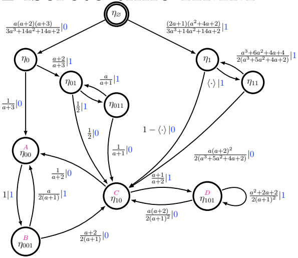


Time evolution: T^n

The Zero Eigenvalue, (di)Graphically

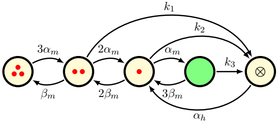
$$\Lambda_A = \left\{\lambda \in \mathbb{C} : \det(A - \lambda I) = 0\right\}$$
$$\text{So, } 0 \in \Lambda_A \implies \det(A) = 0 \implies A^{-1} \neq \operatorname{inv}(A)$$

Discrete time HMM



Time evolution: T^n

Continuous time HMM



Time evolution: e^{tG}

An Operator and its Spectrum

Spectrum

The *spectrum* of an operator A consists of the set of points $\lambda \in \mathbb{C}$ such that $\lambda I - A$ is not invertible.

An Operator and its Spectrum

Spectrum

The *spectrum* of an operator A consists of the set of points $\lambda \in \mathbb{C}$ such that $\lambda I - A$ is not invertible.

Resolvent

The *resolvent* of A, $\mathcal{R}(z; A) \equiv (zI - A)^{-1}$, where z is a continuous complex variable, thus contains all of the spectral information about A (and more).

A finite square matrix and its eigenvalues

- If an operator A can be represented as a finite square matrix, then its spectrum is just the set of A's *eigenvalues*:

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

A finite square matrix and its eigenvalues

- If an operator A can be represented as a finite square matrix, then its spectrum is just the set of A's *eigenvalues*:

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

- Algebraic multiplicity a_λ

A finite square matrix and its eigenvalues

- If an operator A can be represented as a finite square matrix, then its spectrum is just the set of A's *eigenvalues*:

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

- Algebraic multiplicity a_λ
- Geometric multiplicity g_λ

A finite square matrix and its eigenvalues

- If an operator A can be represented as a finite square matrix, then its spectrum is just the set of A's *eigenvalues*:

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

- Algebraic multiplicity a_λ
- Geometric multiplicity g_λ
- The *index* ν_λ of the eigenvalue λ is the size of the largest Jordan block associated with λ.

Definition

Projection Operator

The *projection operator* of A associated with the eigenvalue λ is:

$$\begin{aligned} A_\lambda &\equiv \frac{1}{2\pi i} \oint_{C_\lambda} \mathcal{R}(z; A)\, dz \\ &= \operatorname{Res}\left[(zI - A)^{-1}, z \to \lambda\right] \end{aligned}$$

Definition

Projection Operator

The *projection operator* of A associated with the eigenvalue λ is:

$$\begin{aligned} A_\lambda &\equiv \frac{1}{2\pi i} \oint_{C_\lambda} \mathcal{R}(z; A)\, dz \\ &= \text{Res}\left[(zI - A)^{-1}, z \to \lambda\right] \end{aligned}$$

If A is diagonalizable, then the projection operator can be simply expressed as:

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

Definition

Projection Operator

The *projection operator* of A associated with the eigenvalue λ is:

$$A_\lambda \equiv \frac{1}{2\pi i} \oint_{C_\lambda} \mathcal{R}(z; A)\, dz$$
$$= \text{Res}\left[(zI - A)^{-1}, z \to \lambda\right]$$

If A is diagonalizable, then the projection operator can be simply expressed as: $A_\lambda = \prod_{\zeta \in \Lambda_A \setminus \{\lambda\}} \frac{A - \zeta I}{\lambda - \zeta}$.

If $a_\lambda = 1$, then the projection operator can be simply expressed as:

$$A_\lambda = \frac{1}{\langle \boldsymbol{\lambda} | \boldsymbol{\lambda} \rangle} \left| \boldsymbol{\lambda} \right\rangle \left\langle \boldsymbol{\lambda} \right| \ ,$$

where $\langle \boldsymbol{\lambda}|$ is the left eigenvector of A associated with λ and $|\boldsymbol{\lambda}\rangle$ is the right eigenvector of A associated with λ. (Note: $\langle \boldsymbol{\lambda}| \neq |\boldsymbol{\lambda}\rangle^\dagger$!)

Some General Properties of Projection Operators

- $\{A_\lambda\}$ is a mutually orthogonal set:

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

Some General Properties of Projection Operators

- $\{A_\lambda\}$ is a mutually orthogonal set:

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

- The projection operators are a resolution of the identity:

$$I = \sum_{\lambda \in \Lambda_A} A_\lambda$$

Choosing a functional calculus

Taylor Series

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

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

Choosing a functional calculus

Taylor Series

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

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

shortcomings

- Limited domain of convergence

Choosing a functional calculus

Taylor Series

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

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

shortcomings

- Limited domain of convergence

Holomorphic Functional Calculus

Inspired by
$f(a) = \frac{1}{2\pi i} \oint_C \frac{f(z)}{(z-a)} dz.$

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

Extends $f(A)$ beyond Taylor domain.

The Resolvent Resolved: Partial Fraction Decomposition of the Resolvent:

$$
\begin{aligned}
\mathcal{R}(z;A) &= (zI - A)^{-1} \\
&= \frac{\mathcal{C}^\top}{\det(zI - A)} \\
&= \frac{\mathcal{C}^\top}{\prod_{\lambda \in \Lambda_A} (z-\lambda)^{a_\lambda}} \\
&= \sum_{\lambda \in \Lambda_A} \sum_{m=0}^{a_\lambda - 1} \frac{1}{(z-\lambda)^{m+1}} A_{\lambda,m} \\
&= \sum_{\lambda \in \Lambda_A} \sum_{m=0}^{\nu_\lambda - 1} \frac{1}{(z-\lambda)^{m+1}} A_\lambda \big(A - \lambda I\big)^m
\end{aligned}
$$

for $z \notin \Lambda_A$, where $\mathcal{C}$ is the matrix of cofactors of $zI - A$; A_λ is the *projection operator* associated with λ.

Choosing a functional calculus

Taylor Series

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

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

shortcomings

- Limited domain of convergence

Holomorphic Functional Calculus

Inspired by
$f(a) = \frac{1}{2\pi i}\oint_C \frac{f(z)}{(z-a)}\, dz$.

$$\begin{aligned} f(A) &= \frac{1}{2\pi i}\oint_{C_{\Lambda_A}} f(z)(zI-A)^{-1}\, dz \\ &= \sum_{\lambda\in\Lambda_A}\sum_{m=0}^{\nu_\lambda-1} \frac{f^{(m)}(\lambda)}{m!}(A-\lambda I)^m A_\lambda \, . \end{aligned}$$

Extends $f(A)$ beyond Taylor domain.

Choosing a functional calculus

Taylor Series

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

$$f(A) = \sum_{n=0}^{\infty} \frac{f^{(n)}(\xi)}{n!} (A - \xi I)^n \; .$$

shortcomings

- Limited domain of convergence

Holomorphic Functional Calculus

Inspired by
$f(a) = \frac{1}{2\pi i} \oint_C \frac{f(z)}{(z-a)} \, dz$.

$$\begin{aligned} f(A) &= \frac{1}{2\pi i} \oint_{C_{\Lambda_A}} f(z)(zI - A)^{-1} \, dz \\ &= \sum_{\lambda \in \Lambda_A} \sum_{m=0}^{\nu_\lambda - 1} \frac{f^{(m)}(\lambda)}{m!} (A - \lambda I)^m A_\lambda \; . \end{aligned}$$

Extends $f(A)$ beyond Taylor domain.

shortcomings

- $f(z)$ must be holomorphic at Λ_A

Choosing a functional calculus

Holomorphic Functional Calculus

Inspired by
$f(a) = \frac{1}{2\pi i} \oint_C \frac{f(z)}{(z-a)} \, dz.$

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

Extends $f(A)$ beyond Taylor domain.

shortcomings

- $f(z)$ must be holomorphic at Λ_A

Choosing a functional calculus

Holomorphic Functional Calculus

Inspired by
$f(a) = \frac{1}{2\pi i} \oint_C \frac{f(z)}{(z-a)}\, dz.$

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

Extends $f(A)$ beyond Taylor domain.

shortcomings

- $f(z)$ must be holomorphic at Λ_A

Choosing a functional calculus

Meromorphic Functional Calculus

Inspired by
$f(a) = \frac{1}{2\pi i} \oint_C \frac{f(z)}{(z-a)} \, dz$.

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

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

Extends $f(A)$ beyond both Taylor and holomorphic domain.

Meromorphic functional calculus

Meromorphic functional calculus

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

- Poles and zeros of $f(z)$ can interact with poles of resolvent.
- Extends $f(A)$ beyond holomorphic domain.

Powers of Matrices

$$A^L = \sum_{\lambda \in \Lambda_A \setminus \{0\}} \sum_{m=0}^{\nu_\lambda - 1} \binom{L}{m} \lambda^{L-m} A_\lambda (A - \lambda I)^m$$
$$+ [0 \in \Lambda_A] \sum_{m=0}^{\nu_0 - 1} \delta_{L,m} A_0 A^m$$

for any $L \in \mathbb{C}$, where $\binom{L}{m}$ is the generalized binomial coefficient:

$$\binom{L}{m} = \frac{1}{m!} \prod_{n=1}^{m} (L - n + 1)$$

with $\binom{L}{0} = 1$.

Powers of Matrices

$$A^L = \sum_{\lambda \in \Lambda_A \setminus \{0\}} \sum_{m=0}^{\nu_\lambda - 1} \binom{L}{m} \lambda^{L-m} A_\lambda (A - \lambda I)^m + [0 \in \Lambda_A] \sum_{m=0}^{\nu_0 - 1} \delta_{L,m} A_0 A^m$$

for any $L \in \mathbb{C}$, $\binom{L}{m} = \frac{1}{m!} \prod_{n=1}^{m} (L - n + 1)$ with $\binom{L}{0} = 1$.

E.g.:
$\binom{0}{m} = 0$ unless $m = 0$, so:

$$A^0 = \sum_{\lambda \in \Lambda_A} A_\lambda = I \ .$$

Powers of Matrices

$$A^L = \sum_{\lambda \in \Lambda_A \setminus \{0\}} \sum_{m=0}^{\nu_\lambda - 1} \binom{L}{m} \lambda^{L-m} A_\lambda (A - \lambda I)^m + [0 \in \Lambda_A] \sum_{m=0}^{\nu_0 - 1} \delta_{L,m} A_0 A^m$$

for any $L \in \mathbb{C}$, $\binom{L}{m} = \frac{1}{m!} \prod_{n=1}^{m} (L - n + 1)$ with $\binom{L}{0} = 1$.

E.g.:
$\binom{1}{m} = 0$ unless $m = 0$ or $m = 1$, so:

$$\begin{aligned} A^1 &= \sum_{\lambda \in \Lambda_A \setminus \{0\}} \sum_{m=0}^{1} \lambda^{L-m} A_\lambda (A - \lambda I)^m + [0 \in \Lambda_A] \, A_0 A \\ &= \sum_{\lambda \in \Lambda_A} \lambda A_\lambda + \sum_{\lambda \in \Lambda_A} A_\lambda (A - \lambda I) \qquad . \end{aligned}$$

Powers of Matrices

$$A^L = \sum_{\lambda \in \Lambda_A \setminus \{0\}} \sum_{m=0}^{\nu_\lambda - 1} \binom{L}{m} \lambda^{L-m} A_\lambda (A - \lambda I)^m + [0 \in \Lambda_A] \sum_{m=0}^{\nu_0 - 1} \delta_{L,m} A_0 A^m$$

for any $L \in \mathbb{C}$, $\binom{L}{m} = \frac{1}{m!} \prod_{n=1}^{m} (L - n + 1)$ with $\binom{L}{0} = 1$.

E.g.:
$\binom{1}{m} = 0$ unless $m = 0$ or $m = 1$, so:

$$A^1 = \sum_{\lambda \in \Lambda_A \setminus \{0\}} \sum_{m=0}^{1} \lambda^{L-m} A_\lambda (A - \lambda I)^m + [0 \in \Lambda_A] \, A_0 A$$

$$= \sum_{\lambda \in \Lambda_A} \lambda A_\lambda + \sum_{\lambda \in \Lambda_A} A_\lambda (A - \lambda I) = D + N = A \; .$$

where D is diagonable, N is nilpotent, and $[D, N] = 0$.

Powers of Matrices

$$A^L = \sum_{\lambda \in \Lambda_A \setminus \{0\}} \sum_{m=0}^{\nu_\lambda - 1} \binom{L}{m} \lambda^{L-m} A_\lambda (A - \lambda I)^m + [0 \in \Lambda_A] \sum_{m=0}^{\nu_0 - 1} \delta_{L,m} A_0 A^m$$

for any $L \in \mathbb{C}$, $\binom{L}{m} = \frac{1}{m!} \prod_{n=1}^{m} (L - n + 1)$ with $\binom{L}{0} = 1$.

E.g.:
$\binom{-|L|}{m} = (-1)^m \binom{|L|+m-1}{m}$ and $\binom{m}{m} = 1$, so:

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

Powers of Matrices

$$A^L = \sum_{\lambda \in \Lambda_A \setminus \{0\}} \sum_{m=0}^{\nu_\lambda - 1} \binom{L}{m} \lambda^{L-m} A_\lambda (A - \lambda I)^m + [0 \in \Lambda_A] \sum_{m=0}^{\nu_0 - 1} \delta_{L,m} A_0 A^m$$

for any $L \in \mathbb{C}$, $\binom{L}{m} = \frac{1}{m!} \prod_{n=1}^{m} (L - n + 1)$ with $\binom{L}{0} = 1$.

E.g.:
$\binom{-|L|}{m} = (-1)^m \binom{|L|+m-1}{m}$ and $\binom{m}{m} = 1$, so:

$$A^{-1} = \sum_{\lambda \in \Lambda_A \setminus \{0\}} \sum_{m=0}^{\nu_\lambda - 1} (-1)^m \lambda^{-1-m} A_\lambda (A - \lambda I)^m \quad = A^{\mathcal{D}} \neq \mathrm{inv}(A) \, (!)$$

where $A^{\mathcal{D}}$ is the *Drazin inverse* of A.

Drazin Inverse, Spectrally

$$A^{\mathcal{D}} = \sum_{\lambda \in \Lambda_A \setminus \{0\}} \sum_{m=0}^{\nu_\lambda - 1} (-1)^m \lambda^{-1-m} A_\lambda (A - \lambda I)^m \ .$$

When A is diagonable, this reduces to:

$$A^{\mathcal{D}} = \sum_{\lambda \in \Lambda_A \setminus \{0\}} \lambda^{-1} A_\lambda \qquad \text{if } \nu_\lambda = 1 \text{ for all } \lambda \in \Lambda_A \ .$$

Drazin inverse, Axiomatically

A^{-1} from the meromorphic functional calculus satisfies all of the properties of the unique Drazin inverse as it is axiomatically defined for matrices with index ν_0:

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

So $A^{-1} = A^{\mathcal{D}}$,

a.k.a., the $\{1^{\nu_0}, 2, 5\}$-inverse.

Drazin inverse or Group inverse?

For $\nu_0 = 1$, $A^{\mathcal{D}} = A^{\sharp}$.
For $\nu_0 > 1$, $A^{\sharp}$ does not exist.

We are really bad at asking novel questions!

Most questions are of the form:

$$\langle \cdot | \, T^L \, | \cdot \rangle$$

or

$$\sum_{n=0}^{L} \langle \cdot | \, T^n \, | \cdot \rangle$$

We are really bad at asking novel questions!

Most questions are of the form:

$$\langle \cdot | \, T^L \, | \cdot \rangle \qquad \text{or} \quad \langle \cdot | \, e^{tG} \, | \cdot \rangle$$

or

$$\sum_{n=0}^{L} \langle \cdot | \, T^n \, | \cdot \rangle \qquad \text{or} \quad \int \langle \cdot | \, e^{tG} \, | \cdot \rangle \, dt$$

We are really bad at asking novel questions!

Most questions are of the form:

$$\langle \cdot | \, T^L \, | \cdot \rangle \qquad \text{or} \quad \langle \cdot | \, e^{tG} \, | \cdot \rangle$$

or

$$\sum_{n=0}^{L} \langle \cdot | \, T^n \, | \cdot \rangle \qquad \text{or} \quad \int \langle \cdot | \, e^{tG} \, | \cdot \rangle \; dt$$

Different representations of a process (diff. Ts or diff. Gs) needed to answer different questions (using linear algebra).

We are really bad at asking novel questions!

Most questions are of the form:

$$\langle \cdot | \, T^L \, | \cdot \rangle \qquad \text{or} \quad \langle \cdot | \, e^{tG} \, | \cdot \rangle$$

or

$$\sum_{n=0}^{L} \langle \cdot | \, T^n \, | \cdot \rangle \qquad \text{or} \quad \int \langle \cdot | \, e^{tG} \, | \cdot \rangle \, dt$$

Different representations of a process (diff. Ts or diff. Gs) needed to answer different questions (using linear algebra).

- Every time we ask a question like the bottom row, we invoke the Drazin inverse.

Correlations and Power Spectra are directly about observables:
$\longrightarrow T$ of any representative HMM will do:

$$\gamma[\tau] = \left\langle \overline{X}_n X_{n+\tau} \right\rangle_n$$

Correlations and Power Spectra are directly about observables:
$\longrightarrow T$ of any representative HMM will do:

$$\begin{aligned}\gamma[\tau] &= \left\langle \overline{X}_n X_{n+\tau} \right\rangle_n \\ &= \langle \overline{\mathcal{A}} | \, T^{|\tau|-1} \, | \mathcal{A} \rangle \ .\end{aligned}$$

$$P_{\mathrm{c}}(\omega) = 2 \, \mathrm{Re} \, \langle \overline{\mathcal{A}} | \, (e^{i\omega} I - T)^{-1} \, | \mathcal{A} \rangle + \text{const} \ .$$

Information metrics require *distributions* over states of any rep. of process:
$\longrightarrow$ Use W of mixed state presentation (MSP) of any rep. of process.

E.g.,
Myopic entropy rates:

$$H(X_L|X_{0:L}) = h_\mu(L) = \langle \cdot | \, W^n \, | \cdot \rangle$$

Information metrics require *distributions* over states of any rep. of process:
$\longrightarrow$ Use W of mixed state presentation (MSP) of any rep. of process.

E.g.,
Myopic entropy rates:

$$H(X_L | X_{0:L}) = h_\mu(L) = \langle \cdot | \, W^n \, | \cdot \rangle$$

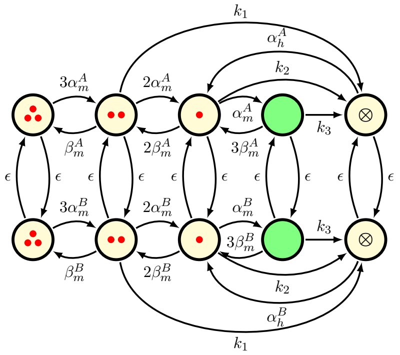
Past–future mutual info:

$$I(X_{:0}; X_{0:}) = \mathbf{E} = \sum_{L=1}^{\infty} \left[h_\mu(L) - h_\mu \right] = \langle \cdot | \, (I - W)^{\mathcal{D}} \, | \cdot \rangle$$

Measures of intrinsic computation require distributions over *causal* states:
$\longrightarrow$ Use $\mathcal{W}$ of MSP of ϵ-machine.

$\ldots$

Joint Environmental-driving–System-state Dynamic



$$Q_{\text{ex}} = \sum_{n=1}^{N} Q_{\text{ex}}[x_n; s_{n-1} \to s_n]$$

$$Q_{\mathrm{ex}} = \textstyle\sum_{n=1}^{N} Q_{\mathrm{ex}}[x_n; s_{n-1} \to s_n]$$

$$\begin{aligned}\langle Q_{\mathrm{ex}} \rangle &= \langle \boldsymbol{\mu}_0 | \Big(\sum_{n=0}^{N-1} \mathcal{T}^n \Big) Q^{(X\mathcal{S})} \, |\mathbf{1}\rangle \\ &\to \langle \boldsymbol{\mu}_0 | \Big(\int_{t=0}^{\tau} e^{t\mathcal{G}} \, dt \Big) \mathcal{Q}^{(X\mathcal{S})} \, |\mathbf{1}\rangle\end{aligned}$$

For a broad subset of HMMs, the heat matrix turns out to be:

$$\mathcal{Q}_{ij}^{(X\mathcal{S})} = \frac{\mathcal{G}_{ij}}{\beta_j} \left[\ln\left(\pi_{x^j}(s^i)\right) - \ln\left(\pi_{x^j}(s^j)\right) \right] \; .$$

$Q_{\mathrm{ex}} = \sum_{n=1}^{N} Q_{\mathrm{ex}}[x_n; s_{n-1} \to s_n]$

$$\begin{aligned}\langle Q_{\mathrm{ex}} \rangle &= \langle \boldsymbol{\mu}_0 | \Big(\sum_{n=0}^{N-1} \mathcal{T}^n \Big) Q^{(X\mathcal{S})} | \mathbf{1} \rangle \\ &\to \langle \boldsymbol{\mu}_0 | \Big(\int_{t=0}^{\tau} e^{t\mathcal{G}} \, dt \Big) \mathcal{Q}^{(X\mathcal{S})} | \mathbf{1} \rangle \\ &= \tau \, \langle \boldsymbol{\pi} | \mathcal{Q}^{(X\mathcal{S})} | \mathbf{1} \rangle + \langle \boldsymbol{\mu}_0 | \, \mathcal{G}^{\mathcal{D}} \big(e^{\tau\mathcal{G}} - I \big) \mathcal{Q}^{(X\mathcal{S})} | \mathbf{1} \rangle\end{aligned}$$

since $\int_0^{\tau} e^{t\mathcal{G}} \, dt = \tau \, | \mathbf{1} \rangle \langle \boldsymbol{\pi} | + \mathcal{G}^{\mathcal{D}} \big(e^{\tau\mathcal{G}} - I \big)$.

For a broad subset of HMMs, the heat matrix turns out to be:

$$\mathcal{Q}_{ij}^{(X\mathcal{S})} = \frac{\mathcal{G}_{ij}}{\beta_j} \left[\ln \big(\pi_{x^j}(s^i) \big) - \ln \big(\pi_{x^j}(s^j) \big) \right] \ .$$

Where did you come from?
Nobody knows.
But Drazin knows best.

Remove your asymptotics;
I want to see your transients.

$\mathcal{L}$ove,

$\mathcal{D}razin$