Searching the minimal pure-state quantum predictive model

Jinghao Lyu

June 2021

Abstract

We investigate the non-zero phase pure-state quantum predictive model. We first set the frameworks of a quantum predictive model called a pure-state quantum model. Then we study three stochastic processes by embedding them to this quantum predictive model. We evaluate the minimal memory required to model the stochastic processes in this quantum model and compare the result to the classical $\epsilon-$machine and q-machine. The result is that by carefully choosing the parameters in our quantum model, the memory required in quantum model can be less than the memory required in $\epsilon-$machine and q-machine.

1 Introduction

Stochastic processes are ubiquitous in science. One of the interesting quantities related to a stochastic process is the minimal memory required to model this stochastic process. Classically it is solved by $\epsilon-$machine [1]. Quantumly, it is shown that q-machines require less memory. And the maximum compression is determined by the cryptic order of the stochastic process[2]. For perturbed coins[3] and Ising model[4][5] the minimal memory in quantum model is determined. In other cases, the minimal memory is unknown. In this paper we propose a new quantum model: pure-state quantum model with non-zero phases and show that by choosing appropriate phases our model even requires less memory than the q-machine. We also provide a way to search the minimal memory in pure state quantum model.

2 Frameworks

Here we consider a discrete stationary stochastic process that at each discrete time t emits an output x_i from a finite alphabet $\mathcal{X}$ with a probability distribution $\Pr(\overleftarrow{X}, \vec{X})$. For each past $\overleftarrow{x}$ we have a conditional probability $\Pr(\vec{X}|\overleftarrow{x})$. A predictive model is an algorithm which given a specific past can generate the identical conditional future behavior. Each model specifies an encoding function ϵ mapping each past to an internal state $s \in \mathcal{S}$ such that systematic actions on this

system at future time-steps will generate a string of outputs obeying $\Pr(\vec{X}|\overleftarrow{x})$. Since we consider a stationary stochastic process, at each time step we have the same internal states in our predictive model. The stochastic process can be modeled by Markovian dynamics on $\mathcal{S}$. At each time t, the predictive model in s_t generate the output x_{t+1} according to $\Pr(X_{t+1}|\overleftarrow{x}_t) = \Pr(X_{t+1}|s_t)$ and updates its internal state to $\epsilon(\overleftarrow{x}x_{t+1})$. The amount of memory needed for this model is given by Shannon entropy $H(S) = -\sum_s p_s \log_2 p_s$ where p_s is the probability $\epsilon(\overleftarrow{x}) = s$. Obviously, the stochastic process itself is a predictive model mapping each past to itself. Given a stochastic process, computational mechanics provides a framework to construct the minimal predictive model which is called $\epsilon-$ machine by mapping pasts with the same conditional future probability to a state. Each state in $\epsilon-$ machine is called a causal state. The memory that the minimal model of a stochastic process required is called the statistical complexity of the stochastic process. This is the minimal information needed to simulate the future.

A quantum predictive model is a model in which the pasts are mapped to the quantum states and for each output x there is a corresponding operator $\mathcal{T}^x$. The future conditional probabilities are encoded in the states. At each time step t the quantum state is $|\psi_t\rangle$. After the output x is observed in classical predictive model the quantum state is updated to $\mathcal{T}^x|\psi_t\rangle$. The probabilities of a given string can be given by a projective measurement. In this sense, a quantum predictive model can provide the same conditional future probabilities as in the stochastic process. Given a classical predictive model we have one corresponding quantum predictive model and we call this the quantization of a classical predictive model. We will illustrate this idea next.

2.1 Quantize classical predictive model

Here we first give a strict definition of the classical predictive models.[6]

Definition 1 (Classical predictive model) *A classical predictive model is a triplet $M = (\mathcal{R}, \mathcal{A}, \{\mathcal{T}^x, x \in \mathcal{A}\})$ of hidden states $\mathcal{R}$, an output alphabet $\mathcal{A}$, and non-negative transition matrices $\mathcal{T}^x_{\rho\rho'}$ with $x \in \mathcal{A}$ and $\rho, \rho' \in \mathcal{R}$, satisfying the properties:*

- Irreducibility : $T = \sum_{x\in\mathcal{A}} T^x$ *is stochastic*
- Unifilarity : $\mathcal{T}^x_{\rho\rho'} \sim \delta_{\rho',f(\rho,x)}$ *for a deterministic function f*

The corresponding quantized predictive model is given below.

Definition 2 (Pure state quantum predictive model) *A pure state quantum predictive model is a triplet $Q = (\Sigma, \mathcal{A}, \{\mathcal{T}^x, x \in \mathcal{A}\})$ of pure quantum states Σ, an output alphabet $\mathcal{A}$, and operators $\mathcal{T}^x$ with $x \in \mathcal{A}$. For each hidden state $s \in \mathcal{R}$ in classical predictive model there is a corresponding pure state $|s\rangle \in \Sigma$. They must satisfy the following properties:*

- $\sum_{x\in\mathcal{A}} \mathcal{T}^{x\dagger}\mathcal{T}^x = 1$

- $\mathcal{T}^x|s_i\rangle = \sqrt{\Pr(f(i,x),x|i)}e^{i\theta_{i\to f(i,x)}}|s_{f(i,x)}\rangle$

where $\theta_{i\to f(i,x)}$ are tunable parameters.

Those properties come directly from the transition probabilities. In this quantum the probabilities are encoded in this way: suppose the current state is $|i\rangle$ the probability of observing output x accroding to Born rules is: $\sum_k |\langle i|\mathcal{T}^x|e_k\rangle|^2 = \mathrm{Tr}(|i\rangle\langle i|\mathcal{T}^x\mathcal{T}^{x\dagger}) = \Pr(f(i,x),x|i)$ where $|e_k\rangle$ are basis in Hilbert space. It is not hard to prove those properties. The inner products of two quantum states are given by:

$$\begin{aligned}\langle i|j\rangle &= \langle i|\mathbf{1}|j\rangle = \langle i|\sum_a \mathcal{T}^a(\mathcal{T}^a)^\dagger|j\rangle \\ &= \sum_a \sqrt{\Pr(l,a|j)\Pr(k,a|i)}(i\theta^a_{jl} - i\theta^a_{ik})\langle k|l\rangle\end{aligned} \tag{1}$$

In a quantum model $\mathcal{M}$ the statistical complexity is replaced by von Neumann entropy $\mathcal{C}(\mathcal{M})$ of the initial density matrix $\rho = \sum_s p_s|s\rangle\langle s|$. And we call a quantum model $\mathcal{M}$ of a stochastic process optimal if and only if for any other quantum model $\mathcal{M}'$, $\mathcal{C}(\mathcal{M}) \leq \mathcal{C}(\mathcal{M}')$. We can prove that in optimal quantum model there is a one-to-one correspondence between classical causal states and quantum pure states [5]. There is no benefit in differentiating two different pasts with the same conditional futures. Hence, in order to find out the optimal quantum model, we only need to consider the quantization of the $\epsilon-$machine.

In pure state quantum model the number of phases is equal to the number of transitions in the classical model. One of goals of this project is to search the minimum von Neumann entropy of a quantum model given a classical process. Luckily with the inner products of states the von Neumann entropy can be evaluated by the Gram matrix [7] [8] as the function of free phases. Then the question turns out to be finding the global minimum of a function in the entire phases space. In the next section we will show three examples how to find out the minimum von Neumann entropy given the classical model.

3 Examples

3.1 Perturbed coins

We consider a perturbed coin of which possible outcomes are A and B. At each time step, the coin is perturbed such that the coin flips with the probability p and states of coins are observed. For any $p \neq 0.5$, we have two causal states; the set of pasts ending in A, and the set of pasts ending in B.

Classically the memory in order to simulate this process is 1 bits. Quantized perturbed coin is given by two quantum states $|A\rangle$ and $|B\rangle$ and two operators $\mathcal{T}^0$ and $\mathcal{T}^1$. The effects of two operators on two states are:

$$\begin{aligned}\mathcal{T}^0\ |A\rangle &= \sqrt{p}\ \exp(i\theta_1)\ |A\rangle \qquad \mathcal{T}^0\ |B\rangle = \sqrt{1-p}\ \exp(i\theta_2)\ |A\rangle \\ \mathcal{T}^1\ |B\rangle &= \sqrt{p}\ \exp(i\theta_3)\ |A\rangle \qquad \mathcal{T}^1\ |A\rangle = \sqrt{1-p}\ \exp(i\theta_4)\ |B\rangle\end{aligned} \tag{2}$$

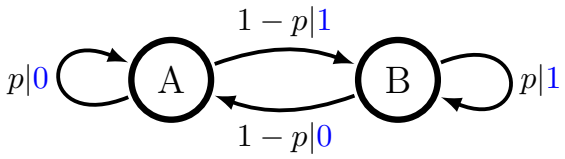

Figure 1: $\epsilon-$Machine of Perturbed coins

And the inner product of those two states is:

$$\begin{aligned} \langle B|A\rangle &= \sqrt{p}\sqrt{1-p}\Big(\exp(i\theta_1 - i\theta_2) + \exp(i\theta_4 - i\theta_3)\Big) \\ &= \sqrt{p}\sqrt{1-p}\; r\exp(i\theta) \end{aligned} \tag{3}$$

where $r \in [0,2]$ and $\theta \in [0, 2\pi]$. In order to generate the same probabilities as in the classical model the initial state we prepare is

$$\rho = \frac{1}{2}|A\rangle\langle A| + \frac{1}{2}|B\rangle\langle B| \tag{4}$$

With the inner product of two states, the von Neumann entropy of the initial mixed state can be calculated by using the Gram matrix method. After diagonalizing the density matrix two eigenvalues are

$$\lambda_1 = \frac{1}{2}\left(1 - r\sqrt{p-p^2}\right),\ \lambda_2 = \frac{1}{2}\left(1 + r\sqrt{p-p^2}\right). \tag{5}$$

The von-Neumann entropy of this system is $S = -\lambda_1 \log_2 \lambda_1 - \lambda_2 \log_2 \lambda_2$. It is obvious that when $r = 2$ the von-Neumann entropy obtains the minimum. We call it minimal pure state quantum $\epsilon-$ machine. And von Neumann entropy is independent of phase θ. When $r = 2, \theta = 0$ we have:

$$\langle A|B\rangle = 2\sqrt{p}\sqrt{1-p} \tag{6}$$

This is the same as L=1 q machine:

$$\begin{aligned} |B\rangle &= \sqrt{1-p}|0\rangle + \sqrt{p}|1\rangle \\ |A\rangle &= \sqrt{p}|0\rangle + \sqrt{1-p}|1\rangle \end{aligned} \tag{7}$$

When $p = 0.5$ and $r = 2$ one of eigenvalue is 0 and von Neumann entropy is 0. This is because when we simulate a fair coin no memory is needed. The upcoming output is independent of the past. If we set $r = 0$ the two states are orthogonal. We recover the classical model.

3.2 Nemo process

The next example is three-state Nemo Process. [2] shows that von Neumann entropy $C_q(L)$ decreases as L increases and $C_q(\infty) \sim 1.0332$. Here we first

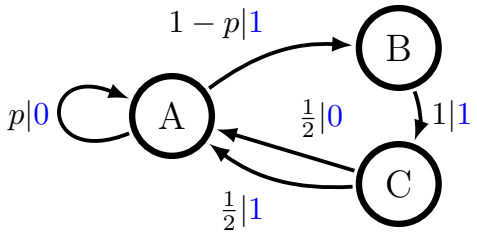

Figure 2: $\epsilon-$Machine of Nemo process

evaluate the inner products and recover the previous result. The effects of two operators on three states are:

$$\begin{aligned}
\mathcal{T}^0\ |A\rangle &= \sqrt{p}\ \exp(i\theta_1)\ |A\rangle \\
\mathcal{T}^0\ |B\rangle &= 0 \\
\mathcal{T}^0\ |C\rangle &= \sqrt{1/2}\exp(i\theta_2)|A\rangle \\
\mathcal{T}^1\ |A\rangle &= \sqrt{1-p}\ \exp(i\theta_3)\ |B\rangle \\
\mathcal{T}^1\ |B\rangle &= \exp(i\theta_4)\ |C\rangle \\
\mathcal{T}^1\ |C\rangle &= \sqrt{1/2}\exp(i\theta_5)|A\rangle.
\end{aligned} \tag{8}$$

And inner products are

$$\begin{aligned}
\langle A|B\rangle &= \frac{\sqrt{p(1-p)}}{1+p}\exp(i\theta_1-\theta_2-\theta_3+\theta_5) \\
\langle B|C\rangle &= \frac{\sqrt{p}}{1+p}\exp(i\theta_1-\theta_2-\theta_4+\theta_5) \\
\langle A|C\rangle &= \frac{\sqrt{2p}}{1+p}\exp(i\theta_1-\theta_2).
\end{aligned} \tag{9}$$

In order to simplify the calculation, we choose the phases of $|A\rangle$ and $|B\rangle$ such that $\langle A|B\rangle$ and $\langle B|C\rangle$ are real numbers leaving only one phase θ in $\langle A|C\rangle$:

$$\begin{aligned}
\langle A|B\rangle &= \frac{\sqrt{p(1-p)}}{1+p} \\
\langle B|C\rangle &= \frac{\sqrt{p}}{1+p} \\
\langle A|C\rangle &= \frac{\sqrt{2p}}{1+p}\exp(i\theta).
\end{aligned} \tag{10}$$

The initial mixed state we prepare is

$$\rho = \frac{1}{3-2p}|A\rangle\langle A| + \frac{1-p}{3-2p}|B\rangle\langle B| + \frac{1-p}{3-2p}|C\rangle\langle C|. \tag{11}$$

Here we choose $p = 0.666$ same as in and evaluate von Neumann entropy as a function of the phase θ. As expected when $\theta = 0$ von Neumann entropy is the

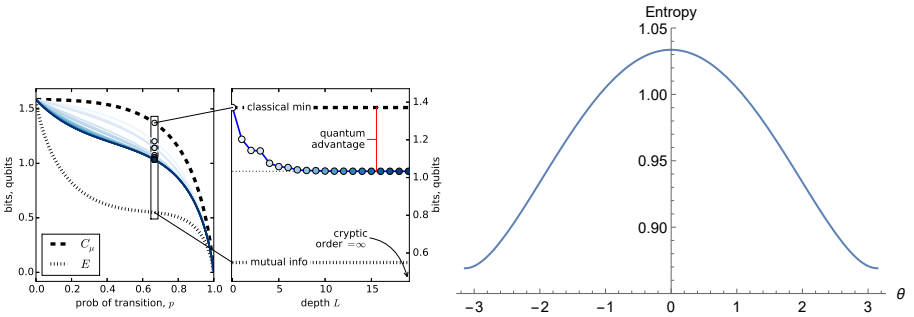

Figure 3: Entropy for p=0.666 Nemo Process. (Mahoney, Aghamohammadi, Crutchfield. 2016)

same as asymptotic $C_q(L)$ in q-machine. Since in q-machine we set all phases to be 0. And we can see the von Neumann entropy reaches minimum when $\theta = \pi$:

$$\begin{aligned} \langle A|B\rangle &= \frac{\sqrt{p(1-p)}}{1+p} \\ \langle B|C\rangle &= \frac{\sqrt{p}}{1+p} \\ \langle A|C\rangle &= -\frac{\sqrt{2p}}{1+p} \end{aligned} \tag{12}$$

We can check that one of the eigenvalues is 0 when $p = 1$ or $p = \frac{\sqrt[3]{9+\sqrt{87}}}{6^{2/3}} - \frac{1}{\sqrt[3]{6\left(9+\sqrt{87}\right)}} \sim 0.589755$. When $p = 1$ there are only two current states so we only need two states to model this process. When $p \sim 0.589755$ there is the dimensional advantage in quantum model. Classical 3-state process can be modeled by 2 states in quantum mechanics.

3.3 3-state MBW process

The last example we study is 3-state MBW process which is a Markov process. The q-machines of this process are studied in [6]. Here we still first list the effect of operators on states:

$$\begin{aligned} \mathcal{T}^A|A\rangle = \sqrt{\frac{2}{3}}e^{i\phi_{AA}}, \mathcal{T}^B|A\rangle = \sqrt{\frac{1}{6}}e^{i\phi_{AB}}, \mathcal{T}^C|A\rangle = \sqrt{\frac{1}{6}}e^{i\phi_{AC}} \\ \mathcal{T}^A|B\rangle = \sqrt{\frac{1}{6}}e^{i\phi_{BA}}, \mathcal{T}^B|B\rangle = \sqrt{\frac{2}{3}}e^{i\phi_{BB}}, \mathcal{T}^C|B\rangle = \sqrt{\frac{1}{6}}e^{i\phi_{BC}} \\ \mathcal{T}^A|C\rangle = \sqrt{\frac{1}{6}}e^{i\phi_{CA}}, \mathcal{T}^B|C\rangle = \sqrt{\frac{1}{6}}e^{i\phi_{CB}}, \mathcal{T}^C|C\rangle = \sqrt{\frac{2}{3}}e^{i\phi_{CC}} \end{aligned} \tag{13}$$

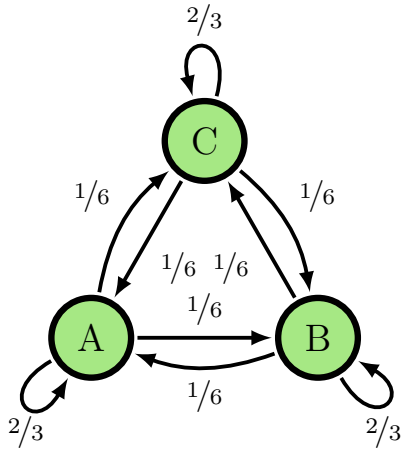

Figure 4: $\epsilon-$Machine of 3MBW process

And inner products are given by:

$$\begin{aligned}
\langle A|B\rangle &= \frac{1}{3}e^{i(\phi_{BA}-\phi_{AA})} + \frac{1}{3}e^{i(\phi_{BB}-\phi_{AB})} + \frac{1}{6}e^{i(\phi_{BC}-\phi_{AC})} = r_{AB}e^{i\theta_1} \\
\langle A|C\rangle &= \frac{1}{3}e^{i(\phi_{CA}-\phi_{AA})} + \frac{1}{3}e^{i(\phi_{CB}-\phi_{AB})} + \frac{1}{6}e^{i(\phi_{CB}-\phi_{AB})} = r_{AC}e^{i\theta_2} \\
\langle B|C\rangle &= \frac{1}{3}e^{i(\phi_{BB}-\phi_{BC})} + \frac{1}{3}e^{i(\phi_{CB}-\phi_{CC})} + \frac{1}{6}e^{i(\phi_{AB}-\phi_{AC})} = r_{BC}e^{i\theta_3}
\end{aligned} \tag{14}$$

Since we have nine free parameters in this model the overall phases $\theta_{1,2,3} \in [0, 2\pi]$ and $r_{AB,AC,BC} \in [0, \frac{5}{6}]$. We can choose the phases of B and C such that $\langle A|B\rangle$ and $\langle A|C\rangle$ are real number:

$$\begin{aligned}
\langle A|B\rangle &= r_{AB} \\
\langle A|C\rangle &= r_{AC} \\
\langle B|C\rangle &= r_{BC}e^{i\phi}
\end{aligned} \tag{15}$$

The mixed state we prepare is $\rho = \frac{1}{3}|A\rangle\langle A| + \frac{1}{3}|B\rangle\langle B| + \frac{1}{3}|C\rangle\langle C|$. The other constraint we have on this system is that we ask the three eigenvalues of the density matrix to be positive: $1+2\, r_{AB}\, r_{BC}\, r_{AC}\cos(\phi) \geq r_{AB}^2 + r_{BC}^2 + r_{AC}^2$. We scan the entire phase space and find that the minimum von Neumann entropy is obtained at $r_{AB} = r_{BC} = r_{AC} = \frac{5}{6}, \cos\theta = \frac{117}{125}$. The minimum von Neumann entropy and eigenvalues are $S = 0.515109$ and $\{0, \frac{1}{18}\left(9 - 4\sqrt{3}\right), \frac{1}{18}\left(9 + 4\sqrt{3}\right)\}$. In this process the quantum model also enjoys dimensional advantage. q-machines can be treated as special cases in our model. For 3 dimension q-machine $r_{AB} = r_{BC} = r_{AC} = \frac{5}{6}, \cos\theta = 1$ and for 2 dimension q-machine $r_{AB} = r_{BC} = r_{AC} = \frac{1}{2}, \cos\theta = -1$. The von Neumann entropies are $S_{q3} = 0.61$ and $S_{q2} = 1$ bits.

4 Unitary realization

In the frameworks discussed above, the operators $\mathcal{T}^x$ are not unitary. In [9] they propose a unitary simulator. Based on their work we can design a unitary operator for our model. The idea is to consider our system as a subsystem of a larger system:

$$U|s_i\rangle|0\rangle = \sum_x \sqrt{\Pr(f(i,x),x|i)} e^{i\theta_{i\to f(i,x)}} |s_{f(i,x)}\rangle|x\rangle \tag{16}$$

where $|x\rangle$ are orthogonal states. At each time step we measure in the $|x\rangle$ basis. This generate the probability correctly. And operator U is unitary.

5 Conclusion and open questions

In this project, we restudy the way to quantize the classical predictive model. In the quantized model, we have free parameters of which number is equal to the number of transitions in the classical predictive model. In general von Neumann entropy of the quantum system is dependent on the choice of the phases. We provide a way to find the minimum von Neumann entropy in simple processes. Those are the minimal pure-state quantum model. However several questions remain unanswered: a) Our project is mainly focused on simple processes. The number of free parameters is equal to the number of transitions. For a general process, the task of finding the minimal model could be time-consuming. For all three processes we study in this project we find that all three minimal models are at the boundary of the phase space. It is unclear whether this is true in general or not. b) In this project we assign each classical causal state to a pure quantum state. Are there any other frameworks in quantum model? For example, assigning a causal state to a mixed quantum state. c) For pure-state quantum models we search the minimal model by quantizing the $\epsilon-$machine. If we start with a non-minimal classical model, can we reach the same minimal pure-state quantum model by quantizing the non-minimal classical model?

References

[1] Cosma Rohilla Shalizi and James P Crutchfield. Computational mechanics: Pattern and prediction, structure and simplicity. *Journal of statistical physics*, 104(3):817–879, 2001.

[2] John R Mahoney, Cina Aghamohammadi, and James P Crutchfield. Occam's quantum strop: Synchronizing and compressing classical cryptic processes via a quantum channel. *Scientific reports*, 6(1):1–11, 2016.

[3] Jayne Thompson, Andrew J. P. Garner, John R. Mahoney, James P. Crutchfield, Vlatko Vedral, and Mile Gu. Causal asymmetry in a quantum world. *Phys. Rev. X*, 8:031013, Jul 2018.

[4] Mile Gu, Karoline Wiesner, Elisabeth Rieper, and Vlatko Vedral. Quantum mechanics can reduce the complexity of classical models. *Nature communications*, 3(1):1–5, 2012.

[5] Whei Yeap Suen, Jayne Thompson, Andrew JP Garner, Vlatko Vedral, and Mile Gu. The classical-quantum divergence of complexity in modelling spin chains. *Quantum*, 1:25, 2017.

[6] Samuel P Loomis and James P Crutchfield. Strong and weak optimizations in classical and quantum models of stochastic processes. *Journal of Statistical Physics*, 176(6):1317–1342, 2019.

[7] Richard Jozsa and Jürgen Schlienz. Distinguishability of states and von neumann entropy. *Physical Review A*, 62(1):012301, 2000.

[8] Paul M Riechers, John R Mahoney, Cina Aghamohammadi, and James P Crutchfield. A closed-form shave from occam's quantum razor: Exact results for quantum compression. *arXiv preprint arXiv:1510.08186*, 2015.

[9] Felix C Binder, Jayne Thompson, and Mile Gu. Practical unitary simulator for non-markovian complex processes. *Physical review letters*, 120(24):240502, 2018.