Extending Quantum Advantage when Simulating Strongly Coupled Classical Systems

Cina Aghamohammadi,∗ John R. Mahoney,† and James P. Crutchfield‡
Complexity Sciences Center and Physics Department,
University of California at Davis, One Shields Avenue, Davis, CA 95616
(Dated: September 19, 2016)

Classical stochastic processes can be generated by quantum simulators instead of the more standard classical ones, such as hidden Markov models. One reason for using quantum simulators is that they generally require less memory than their classical counterparts. Here, we examine this quantum advantage for strongly coupled spin systems—the Dyson-like one-dimensional Ising spin chain with variable interaction length. We find that the advantage scales with both interaction range and temperature, growing without bound as interaction increases. Thus, quantum systems can very efficiently simulate strongly coupled classical systems.

PACS numbers: 03.67.Lx 82.20.Wt 05.20.-y 05.50.+q

I. INTRODUCTION

We illustrate, what seems to be, an emerging principle relating the classical and quantum worlds: strongly correlated classical systems can be efficiently simulated by quantum systems. Moreover, we show that this quantum advantage is substantial and increases with the classical system’s degree of interaction.

Our results suggest that this principle may well lead to broad consequences. The world is nothing, if not its constituent interactions. Some go so far as to argue that interaction, not object, is the basic unit of reality [1, 2]. At a minimum, though, interactions make our world interesting and structured. They also complicate it, and rapidly as the web of interaction widens. The resulting complicatedness often confounds theory and drives us to resort to simulation.

Statistical mechanical spin systems are a prime example: A broadly used class of models that can be arbitrarily hard to analyze and, somewhat soberingly, even hard to simulate [3, 4]. While there is a plurality of complicated spin systems, we choose one that is sufficiently interesting, but also very familiar territory: the one-dimensional Ising spin chain with variable interaction range. The one-dimensionality provides tractability, while the long-range coupling yields the desired complicated configuration structure.

Selecting a system is only the start, of course. There are many different simulation methods for any given class. The differences range from the purely algorithmic to the physical substrate employed. For instance, a computer can use machine code or high-level programming languages to yield the desired result. Additionally, computers come in different varieties, in particular the substrate may be classical or quantum [5]. Just as the choice of code influences the computational resources required, so does the substrate.

Here, we report on the memory resources required by classical and quantum simulators. It is known that a quantum simulator typically requires less memory than its classical counterpart. We refer to this as the quantum advantage. Despite exploring several particular cases, very little is known about how this quantum advantage scales. Addressing this, the following shows that not only is the advantage substantial, it also increases systematically and without bound.

To establish the scalings we compare classical and quantum simulators that generate spin configurations of the one-dimensional Ising model with $N$-nearest neighbor interactions. To compute the quantum advantage, we adapt the transfer-matrix formalism to construct the two simulators. This technique allows us to numerically (but accurately) determine the scaling behavior. We find that not only is the quantum advantage generic, its growth scales with temperature and interaction length.

II. DYSON-ISING SPIN CHAIN

We begin with a general one-dimensional ferromagnetic Ising spin chain [7, 8] with Hamiltonian:

\[ \mathcal{H} = - \sum_{\langle i,j \rangle} J(i,j) s_i s_j , \]

† jrmahoney@ucdavis.edu
‡ chaos@ucdavis.edu

The memory resources required are equal if and only if the classical stochastic process is that with no crypticity [6].
in contact with thermal bath at temperature $T$, where spin $s_i$ at site $i$ takes on values $\{+1, -1\}$ and $J(i,j) \geq 0$ is the spin coupling constant between sites $i$ and $j$. Assuming translational symmetry, $J(i,j) \rightarrow J(k)$, $k \equiv |i-j|$. Commonly, $J(k)$ is a positive and monotone-decreasing function. An interaction is said to be long-range if $J(k)$ decays more slowly than exponential. In our studies, we consider couplings that decay by a power law:

$$J(k) = \frac{J_0}{k^\delta} ,$$

where $\delta > 0$. The spin chain resulting from these assumptions is called the Dyson model [9].

To approximate such a long-range system one can consider a similar system with finite-range interaction. For every interaction range $N$, we define the approximate Hamiltonian:

$$H_N = - \sum_i \sum_{k=1}^{N} \frac{J_0}{k^\delta} s_i s_{i+k} .$$

This class of Hamiltonians can certainly be studied in its own right, not simply as an approximation. Let’s explore its set of equilibrium configurations as a stochastic process.

### III. PROCESSES

The concept of a stochastic process is very general. Any physical system that exhibits stochastic dynamics in time may be thought of as generating a stochastic process. We focus on discrete-time, discrete-valued stationary stochastic processes. Such a process, denoted $P = \{A^\infty, \Sigma, \mathbb{P}(\cdot)\}$, is a probability space [10, 11]. Here, the observed symbols come from an alphabet $A = \{\downarrow, \uparrow\}$ of local spin states; though our results easily extend to any finite alphabet. Each random spin variable $X_i$, $i \in \mathbb{Z}$, takes values in $A$. $\mathbb{P}(\cdot)$ is the probability measure over the bi-infinite chain of random variables $X_{-\infty:0} = \ldots X_{-2}X_{-1}X_0X_1X_2 \ldots$. $\Sigma$ is the $\sigma$-algebra generated by the cylinder sets in $A^\infty$. Stationarity means that $\mathbb{P}(\cdot)$ is invariant under time translation. That is, $\mathbb{P}(X_{i_1}X_{i_2} \cdots X_{i_m}) = \mathbb{P}(X_{i_1+n}X_{i_2+n} \cdots X_{i_m+n})$, for all $m \in \mathbb{Z}^+$ and $n \in \mathbb{Z}$.

To interpret our Ising system as a stochastic process, we consider not its time evolution, but rather the spatial “dynamic”. A spin configuration at one instant of time may be thought of as having been generated left-to-right (or equivalently right-to-left). The probability distribution over these configurations defines a stochastic process $P(N,T)$ that inherits its stationarity from spin-configuration spatial translation invariance. In this way, we build on earlier work that used computational mechanics to analyze statistical structure in spatial configurations generated by spin systems [12, 13].

Now that we have defined the process of interest, let us introduce two of its simulators.

### IV. CLASSICAL AND QUANTUM SIMULATORS

What is a simulator for a stochastic process? Often, “simulation” refers to an approximation. In contrast, we require our simulators to be perfect, to generate $P$’s configurations and their probabilities exactly.

Our simulator, though, does more than correctly reproduce a probability distribution over bi-infinite configurations. Specifically, a simulator $S$ of process $P$ is an object where, given an instance of a semi-infinite “past” $x_0 = \ldots x_{-3}, x_{-2}, x_{-1}$, a query of $S$ yields a sample of the “future” $X_0 = x_0, x_1, \ldots$ from the conditional probability distribution $Pr(X_0|X_0 = x_0)$. See Fig. 1.

Physical systems, under certain assumptions such as thermal equilibrium, manifest stationary stochastic processes. When we refer to the simulation of a physical system, what we mean is the simulation of these processes.

How are these simulators implemented? Two common formalisms are Markov Chains [14, 15] and Hidden Markov Models (HMM) [10, 16, 17]. The latter can be significantly more compact in their representation and,

---

2 Throughout, $T$ denotes the effective temperature $k_B T$.  
3 For example, the Simple Nonunifilar Source (SNS) process requires an infinite-state Markov chain and consequently requires infinite memory for simulation, while its HMM requires only two states and so a single bit of memory [18].
for this reason, are sometimes the preferred implementation choice. Here, we employ a particularly useful form of HMM generators.

These HMMs represent the generating mechanism for a given process by a tuple \( \{ S, A, \{ T(x) : x \in A \} \} \) where \( S \) is a finite set of states called causal states, \( A \) is a finite discrete alphabet and \( \{ T(x) : x \in A \} \) are \( |S| \times |S| \) sub-stochastic symbol-labeled transition matrices. The latter’s sum \( T = \sum_{x \in A} T(x) \) is a stochastic matrix. A unifilar HMM is one in which each row of each substochastic matrix has at most one nonzero element.\(^4\)

**\( \epsilon \)-Machine** A given stochastic process can be correctly generated by any number of unifilar HMMs. The one requiring the minimum amount of memory for implementation is called the \( \epsilon \)-machine \(^{[23]}\) and was first introduced in Ref. \(^{[24]}\). A process’ statistical complexity \( C_\mu \) \(^{[23]}\) is the the Shannon entropy of the \( \epsilon \)-machine’s stationary state distribution: \( C_\mu = H(S) = -\sum_{x \in S} \Pr(x) \log_2 \Pr(x) \). Key to our analysis of classical simulator resources, it measures the minimal memory for a unifilar simulator of a process. \( C_\mu \) has been determined for a wide range of physical systems \(^{[25–31]}\). Helpfully, it and companion measures are directly calculable from the \( \epsilon \)-machine, many in closed-form \(^{[32]}\).

**Ising \( \epsilon \)-Machine** How do we construct the \( \epsilon \)-machine that simulates the process \( P(N, T) \)?

First, we must define process’ Markov order \(^{[15]}\): the minimum history length \( R \) required by any simulator to correctly continue a configuration.\(^5\) Specifically, \( R \) is the smallest integer such that:

\[
\Pr(X_1, \ldots, X_{t-2}, X_{t-1}) = \Pr(X_t|X_{t-R}, \ldots, X_{t-2}, X_{t-1})
\]

Reference \(^{[13], \text{Eqs. (84) – (91)}}\) shows that \( P(N, T) \) has Markov order \( N \) for any finite and nonzero temperature \( T \). One concludes that sufficient information for continued generation is contained in the configuration of the \( N \) previously generated spins. More importantly, the \( \epsilon \)-machine that simulates \( P(N, T) \) has \( 2^N \) states and those states are in one-to-one correspondence with the set of length-\( N \) spin configurations.

Second, another key process characteristic is its cryptic order \(^{[33, 34]}\): the smallest integer \( K \) such that \( H(S_K|X_0, X_1 \ldots) = 0 \), where \( H(W|Z) \) is the conditional entropy \(^{[35]}\) and \( S_K \) is the random variable for the \( K \)-th state of the \( \epsilon \)-machine after it generated symbols \( X_0, X_1 \ldots \). Using the fact that \( \epsilon \)-machine states are in one-to-one correspondence with the set of length-\( N \) spin configurations, it is easy to see that \( P(N, T) \)’s cryptic order \( K = N \).

Figure 2 shows the unifilar HMM generators (\( \epsilon \)-machines) of the processes \( P(N, T) \) for \( N = 1, 2, \) and \( 3 \). Let’s explain.

Consider the spin process \( P(1, T) \), a Markov-order \( R = 1 \) process. To generate the process we only need to remember the last spin generated. The \( \epsilon \)-machine (Fig. 2 top-left) has two states, \( \uparrow \) and \( \downarrow \). If the last observed spin is \( \uparrow \), then the current state is \( \uparrow \) and if it is \( \downarrow \), the current state is \( \downarrow \). We denote the probability of a \( \downarrow \) spin given a previous \( \uparrow \) spin by \( p_{\uparrow \downarrow} \). The probability of an \( \uparrow \) spin following a \( \uparrow \) spin is the complement.

Consider the process \( P(2, T) \) with Markov-order \( R = 2 \) and so longer-range interactions. Sufficient information for generation is contained in the configuration of the two previously generated spins. Thus, the \( \epsilon \)-machine (Fig. 2 top-right) has four states that we naturally label \( \uparrow \uparrow \), \( \uparrow \downarrow \), \( \downarrow \uparrow \), and \( \downarrow \downarrow \). If the last observed spin pair \( x_{t-1}x_0 \) is \( \uparrow \downarrow \), the current state is \( \uparrow \downarrow \). Given this state, the next spin will be \( \uparrow \) with probability \( p_{\uparrow \uparrow} \) and \( \downarrow \) with probability \( p_{\uparrow \downarrow} \).

Note that this scheme implies that each state has exactly two outgoing transitions. That is, not all transitions are allowed in the \( \epsilon \)-machine.

Having identified the state space, we may calculate the \( \epsilon \)-machine transition probabilities \( \{ T(x) \}_{x \in A} \). We first

\(^{\text{4}}\) A fledgling literature on minimal non-unifilar HMMs \(^{[19]}\) exists, but constructive methods are largely lacking and, as a consequence, much less is known \(^{[20–22]}\).

\(^{\text{5}}\) More precisely, we mean that an ensemble of simulators must be able to yield an ensemble of configurations that agree (conditioned on that past) with the process’ configuration distribution.
compute the transfer matrix $\mathbf{T}$ [36] and then extract conditional probabilities, following Ref. [13]. (See App. B for details.) The statistical complexity $C_\mu$ follows straightforwardly from the $\epsilon$-machine.

**q-Machine** By studying a specific process (similar to the $\epsilon$-machine in top-left of Fig. 2), Ref. [37] recently demonstrated that quantum mechanics can simulate stochastic processes using memory capacity less than $C_\mu$. This motivates a search for more efficient quantum simulators of other stochastic processes.

A class of such simulators, called *q-machines*, applicable to arbitrary processes, was introduced in Ref. [6]. This construction depends on an encoding length $L$, each with its own q-machine and its quantum cost $C_q(L)$. The cost $C_q(L)$ saturates at a particular length, which was shown to be the process' cryptic order $K$, introduced above [34]. And so, we restrict ourselves to this choice ($L = K$) of encoding length and refer simply to the q-machine and its cost $C_q$.

The q-machine’s quantum memory $C_q$ is upper-bounded by $C_\mu$, with equality only for the special class of zero-cryptic-order processes [34]. And so, $C_\mu/C_q$ gives us our quantitative measure of *quantum advantage*. Efficient methods for calculating $C_q$ were introduced by Ref. [38] using spectral decomposition. Those results strongly suggest that the q-machine is the most memory-efficient among all unifilar quantum simulators, but as yet there is no proof.\(^6\) The quantum advantage $C_\mu/C_q$ has been investigated both analytically [6, 38, 41–43] and experimentally [44].

The q-machine is straightforward to construct from a given $\epsilon$-machine. It consists of a set $\{|\eta_i\rangle\}$ of pure quantum *signal states* in one-to-one correspondence with the classical causal states $\sigma_i \in \mathcal{S}$. Each signal state $|\eta_i\rangle$ encodes the set of length-$K$ (cryptic order) sequences that may follow $\sigma_i$, as well as each corresponding conditional probability:

$$|\eta_i\rangle \equiv \sum_{w,\sigma_j \in \mathcal{S}} \sum_{\sigma_i \in \mathcal{S}} \sqrt{\mathbf{P}(w, \sigma_j | \sigma_i)} \ | w \rangle | \sigma_j \rangle ,$$

where $w$ denotes a length-$K$ sequence and $\mathbf{P}(w, \sigma_j | \sigma_i) = \mathbf{P}(X_0 \cdots X_{K-1} = w, S_{K-1} = \sigma_j | S_0 = \sigma_i)$. The resulting Hilbert space is the product $\mathcal{H}_w \otimes \mathcal{H}_\sigma$. Factor space $\mathcal{H}_\sigma$ is of size $|\mathcal{S}|$, the number of classical causal states, with basis elements $|\sigma_i\rangle$. Factor space $\mathcal{H}_w$ is of size $|\mathcal{A}|^K$, the number of length-$K$ sequences, with basis elements $|w\rangle = |x_0\rangle \cdots |x_{K-1}\rangle$. For $\mathbf{P}(N,T)$’s $\epsilon$-machines, $|\mathcal{S}| = 2^N$ and $|\mathcal{A}|^K = 2^N$. The q-machine’s density matrix $\rho$ is defined by:

$$\rho = \sum_i \pi_i \ | \eta_i \rangle \langle \eta_i | ,$$

where $\{\pi_i\}$ is the stationary distribution over the $\epsilon$-machine’s states. This is determined from the left eigenvector of transfer matrix $\mathbf{T}$ corresponding to the eigenvalue 1. From the density matrix $\rho$, it is straightforward to calculate $C_q = S(\rho) - \rho$’s von Neumann entropy.

**V. ANALYSIS**

We begin by considering the case where couplings decay with exponent $\delta = 2$. Figure 3 displays $C_\mu(N,T)$ and $C_q(N,T)$—the $C_\mu$ and $C_q$ of processes $\mathbf{P}(N,T)$—versus $T$ for $N = 1, \ldots, 6$. The most striking feature is that the classical and quantum memory requirements exhibit qualitatively very different behaviors.

Classical memory increases with $T$, saturating at $C_\mu = N$, since all transitions become equally likely at high temperature. As a result there are $2^N$ equally probable causal states and this means one needs $N$ bits of memory to store the system’s current state. For example, in the nearest-neighbor Ising model (process $\mathbf{P}(1,T)$) high temperature makes spin-$\uparrow$ and spin-$\downarrow$, and thus the co-

---

\(^6\) As in the classical case, nonunifilar quantum simulators are much less well understood [22, 39, 40].
responding states, equally likely.\textsuperscript{7} Also, in the low-temperature limit, this system is known to yield one of only two equally likely configurations—all spin-\(\uparrow\) or all spin-\(\downarrow\). In other words, at low temperature \(p_\uparrow\) and \(p_\downarrow\) converge to zero, while \(p_\uparrow\) and \(p_\downarrow\) converge to one.\textsuperscript{8} This is reflected in the convergence of all curves at \(C_\mu = 1\). Equivalently, this means one needs only one bit of memory to store the current state.

We can similarly understand the qualitative behavior of \(C_q(N,T)\) for a fixed \(N\). As temperature increases, all length-\(N\) signal states become equivalent. This is the same as saying that all length-\(N\) spin configurations become equally likely. As a consequence, the signal states approach one another and, thus, \(C_q(N,T)\) converges to zero.

In the low-temperature limit, the two \(N\uparrow\) and \(N\downarrow\) configurations are distinguished by the high likelihood of neighboring spins being of like type. This leads to a von Neumann entropy of \(S(\rho) = 1\).

Figure 3 reveals strong similarities in the form of \(C_\mu(T)\) at different \(N\). A simple linear scaling leads to a substantial data collapse, shown in Fig. 4. The scaled curves \((N-C_\mu)/(N-1)\) exhibit power-law behavior in \(T\) for \(T > 2\). Increasing the temperature to \(T = 300\) (beyond the scale of the Fig. 4) shows that this scaling is also especially tight at high temperature.

This behavior is generic for different coupling decay values \(\delta > 1\) and, more to the point, the scaling \((\gamma)\) is independent of \(\delta\). We do not consider \(\delta < 1\), where the system energy becomes nonextensive.

Now, we can analyze the decrease in \(C_q\) with temperature. Figure 5 shows that \(C_q\) is also a power-law in \(T\). By measuring this scaling exponent in the same way as above, we determine \(\alpha = 2.000\). Furthermore, analysis shows (App. D) that:

\[
C_q(N,T) \propto \frac{\log_2(T)}{T^2}.
\]

This verifies and adds detail to our numerical estimate.

This behavior is generic for different coupling decay values \(\delta > 1\) and, moreover, the scaling exponent \(\alpha\) is independent of \(\delta\). It is interesting to note that in this case no data-collapse rescaling is required. The exponent \(\alpha\) directly captures the extreme compactness of high temperature quantum simulations.

Taking these results together, we can now appreciate the substantial relative advantage of quantum versus classical simulators.

\textsuperscript{7} At \(T = \infty\) these processes have only a single causal state and thus \(C_\mu = 0\). This is a well known discontinuity that derives from the sudden predictive-equivalence of all of the causal states there.

\textsuperscript{8} It should be pointed out that at any finite temperature \(p_\uparrow\) and \(p_\downarrow\) are nonzero and, therefore, the \(\epsilon\)-machine states remain strongly-connected.
Define the quantum advantage $\eta$ as the ratio of the minimum required memory for the classical simulation to that for the quantum simulation:

$$\eta(N, T) \equiv C_{\mu}(N, T)/C_q(N, T).$$

For fixed temperature $T \gtrsim 2$, $C_{\mu}(N, T)$ is approximately linear in $N$ and for a fixed $N$ it is approximately independent of $T$. As a consequence, the asymptotic quantum advantage scales as:

$$\eta(N, T) \propto N \frac{T^2}{\log_2(T)},$$

which increases faster than any $T^r$ for $r < 2$. Thus, the answer to our motivating question is that the quantum advantage is amplified. Are there systems for which we find no quantum advantage? What is the distinguishing characteristic?

VI. Conclusion

It is notoriously hard to find quantum advantage and even harder to prove [45]. Here, we found just such an advantage in the realm of stochastic-process simulation. Concretely, we analyzed the $N$-nearest neighbor Ising spin system and demonstrated that its quantum advantage displays generic scaling behavior—quadratic in temperature and linear in interaction range. What does this mean? The most striking conclusion is that a highly interacting classical system can be simulated with unbounded quantum advantage. Given the simplicity of the Ising system, we conjecture that this scaling behavior may be a universal feature of quantum advantage in the simulation of extended physical systems.

The Ising model has contributed great insights to condensed matter physics, however, it is classical. Given our examining the difference between classical and quantum simulators, it is natural to wonder about this difference in the context of a truly quantum Hamiltonian. Is the quantum advantage amplified? Are there systems for which we find no quantum advantage? What is the distinguishing characteristic?

For finite-range interaction in one dimension, there is no Ising phase transition. How might the quantum advantage change in the presence of such a transition? Given that the quantum advantage here scaled with the interaction range, we might expect that near the critical temperature, where long-range interactions are important, the quantum advantage is amplified.

Acknowledgments

We thank Mehrnaz Anvari for her suggestions and the Santa Fe Institute for its hospitality during visits. This material is based upon work supported by, or in part by, the John Templeton Foundation and the U. S. Army Research Laboratory and the U. S. Army Research Office under contracts W911NF-13-1-0390 and W911NF-13-1-0340.

[1] A. N. Whitehead. Process and Reality. The Free Press, New York, corrected edition, 1978.
[2] D. Ross, J. Ladyman, D. Spurrett, and J. Collier. Everything Must Go: Metaphysics Naturalized. Oxford University Press, 2007.
[3] P. W. Anderson. Spin glass VI: Spin glass as cornucopia. Physics Today, September:9–11, 1989.
[4] D. K. Stein and C. M. Newman. Spin Glasses and Complexity. Princeton University Press, Princeton, New Jersey, 2013.
[5] R. Feynman. Simulating physics with computers. Intl. J. Theo. Phys., 21(6/7):467–488, 1982.
[6] J. R. Mahoney, C. Aghamohammadi, and J. P. Crutchfield. Occam’s quantum strop: Synchronizing and compressing classical cryptic processes via a quantum channel. Scientific Reports, 6, 2016.
[7] R. J. Baxter. Exactly solved models in statistical mechanics. Courier Corporation, 2007.
[8] A.. Aghamohammadi, C. Aghamohammadi, and M. Khorrami. Externally driven one-dimensional ising model. J. Stat. Mechanics: Theory and Experiment, 2012(02):P02004, 2012.
[9] F. J. Dyson. Existence of a phase-transition in a one-dimensional Ising ferromagnet. Comm. Math. Physics, 12(2):91–107, 1969.
[10] 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.
[11] N. F. Travers. Bounds on Convergence of Entropy Rate Approximations in Hidden Markov Processes. PhD thesis, University of California, Davis, 2013.
[12] J. P. Crutchfield and D. P. Feldman. Statistical complexity of simple one-dimensional spin systems. Phys. Rev. E, 55(2):R1239–R1243, 1997.
[13] D. P. Feldman and J. P. Crutchfield. Discovering noncritical organization: Statistical mechanical, information theoretic, and computational views of patterns in simple one-dimensional spin systems. Santa Fe Institute, 1998. Santa Fe Institute Working Paper 98-04-026.
[14] J. R. Norris. Markov Chains, volume 2. Cambridge University Press, 1998.
Appendix A: Why the Dyson model?

The ferromagnetic Ising spin linear chain with finite-range interaction cannot undergo a phase transition at any positive temperature [46]. In contrast, the Dyson model has a standard second-order phase transition for a range of $\delta$. It was analytically proven by Dyson [9] that a phase transition exists for $1 < \delta < 2$. The existence of a transition at $\delta = 2$ was proven much later on [47]. It is also known that there exists no phase transition for $\delta > 3$ [48] where it behaves as a short-range system. Finally, it was demonstrated numerically that the parameter regime $2 < \delta \leq 3$ contains a phase transition [49], however, this fact has resisted analytical proof. For $\delta \leq 1$, the model is considered nonphysical since the energy becomes non-extensive.

For these reasons we selected the Dyson spin model as it provides the simplicity of 1D configurations, while generating nontrivially correlation spin configurations.

Appendix B: $\epsilon$-Machine Construction

We show how to construct the $\epsilon$-machine simulator of the process $P(N, T)$. Consider a block of spins of length $2N$, divided equally into two blocks. We denote spins in the left (L) and right (R) halves by: $s_i^L$ and $s_i^R$ for $i = 1, \cdots, N$, respectively. We map the left and right block configurations each to an integer $\eta_X$ by:

$$\eta_s = \sum_{i=1}^{N} \left( s_i^s + \frac{1}{2} \right) 2^{i-1},$$

$s \in \{L, R\}$. The blocks internal energies are given by:

$$X_{\eta_s} = -B \sum_{i=1}^{N} s_i^L - \sum_{i=1}^{N-1} \sum_{k=1}^{N-i} J_i s_i^L s_{i+k}^L,$$

and the correlated energy between two blocks is:

$$Y_{\eta_L, \eta_R} = -\sum_{i=1}^{N} \sum_{k=1}^{i} J_i s_i^L s_{i+k}^R.$$

With these we construct the transfer matrix:

$$V_{\eta_L, \eta_R} = \exp \left( -\frac{1}{T} (\epsilon X_{\eta_L} + Y_{\eta_L, \eta_R} + \frac{1}{2} X_{\eta_R}) \right).$$

Reference [13] shows that the $\epsilon$-machine labeled-transition matrices can be written as:

$$T^{(x)}_{\eta_0, \eta_1} = \begin{cases} \frac{1}{X} V_{\eta_0, \eta_1} & \eta_1 = (\lfloor \frac{\eta_0}{2} \rfloor + x * 2^{N-1}) \\
0 & \eta_1 \neq (\lfloor \frac{\eta_0}{2} \rfloor + x * 2^{N-1}) \end{cases}.$$

Then the $\epsilon$-machine simulator of $P(N, T)$ is $\{S, A, \{T^{(x)}\}_{x \in A}\}$ where $A = \{0, 1\}$ and $S = \{i : 1 \leq i \leq 2^N\}$.

Appendix C: Presence of Magnetic Field

Naturally, one might ask how our results are modified by the presence of an external magnetic field. Consider the one-dimensional ferromagnetic Ising spin chain with Hamiltonian:

$$\hat{H}_N = -\sum_{i=1}^{N} \sum_{k=1}^{N} \frac{J_0}{k^2} s_i s_{i+k} - \sum_{i} B s_i.$$

Figure 6 shows that, due to the symmetry breaking at low temperature, both $C_q(N, T)$ and $C_\mu(N, T)$ converge to zero. (All the spins at low temperature align with magnetic field and, as a consequence, no memory is needed.)

The high temperature behaviors for both functions are the same as before and the quantum advantage remains the same.
Appendix D: High Temperature Behavior

Consider first the case $N = 1$. Due to symmetry we have $p \equiv \Pr(\uparrow | \uparrow) = \Pr(\downarrow | \downarrow) = N/D$, where $N = \exp(\beta J)$ and $D = \exp(\beta J) + \sqrt{\exp(-2\beta J)}$ with $\beta = 1/T$. At high temperature $\beta$ is small and we have:

$$D = 2 + \beta^2,$$
$$N = 1 + \beta + \beta^2.$$  

Again, due to symmetry we have $\pi_1 = \pi_2 = 1/2$. Therefore, the density matrix in Eq. (1) is:

$$\rho = \begin{pmatrix}
\frac{1}{2} & \sqrt{p(1-p)} \\
\sqrt{p(1-p)} & \frac{1}{2}
\end{pmatrix},$$

which has two eigenvalues $\beta^2/4$ and $1 - \beta^2/4$. As a consequence, $C_q$—$\rho$’s von Neumann entropy—is:

$$C_q = S(\rho) \simeq -\left(\frac{\beta^2}{4} \log_2 \frac{\beta^2}{4} + \left(1 - \frac{\beta^2}{4}\right) \log_2 \left(1 - \frac{\beta^2}{4}\right)\right) \simeq \frac{\log_2(T)}{2T^2}.$$  

Increasing the temperature, the interaction between spins weakens. At high temperature the only important neighbor is the nearest neighbor. And so, the high-temperature behavior is similar to the case of $N = 1$ and is independent of $N$.

Examining the numerator, for any $r > 0$ we have $\log_2(T) < T^r$. So, for large $T$ and for all $r > 0$:

$$\frac{1}{T^2} < \frac{\log_2(T)}{T^2} < \frac{1}{T^{r+2}}.$$

This explains the fat tails of $C_q$ for large $T$. More to the point, it shows that for $N = 1$ the scaling exponent is $\alpha = 2$. 

