

# Extreme Quantum Advantage when Simulating Classical Systems with Long-Range Interaction

Cina Aghamohammadi,<sup>\*</sup> John R. Mahoney,<sup>†</sup> and James P. Crutchfield<sup>‡</sup>

*Complexity Sciences Center and Physics Department,  
University of California at Davis, One Shields Avenue, Davis, CA 95616*

(Dated: April 29, 2017)

## Abstract

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 has recently come to the fore: they generally require less memory than their classical counterparts. Here, we examine this quantum advantage for strongly coupled spin systems—in particular, the Dyson 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 range increases. In particular, it is impossible to simulate Dyson’s original spin chain with the most memory-efficient known classical algorithm since it requires infinite memory, while quantum simulators can do so since they use only finite memory. Thus, quantum systems can very efficiently simulate strongly coupled one-dimensional classical systems.

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

Keywords: stochastic process, Ising model, Dyson model, long-range interaction, strongly correlated system, quantum information, quantum communication, quantum complexity, quantum machine, statistical complexity

---

<sup>\*</sup> caghamohammadi@ucdavis.edu

<sup>†</sup> jrmahoney@ucdavis.edu

<sup>‡</sup> chaos@ucdavis.edu

The idea of a quantum computer, often attributed to Feynman [1], recognizes that while simulating quantum many-body systems is difficult, it is apparently something that the physical quantum system to be simulated itself accomplishes with ease. For this reason, it was conjectured that a “quantum computer”—one that operates on a quantum instead of classical substrate—might have a significant advantage in such a simulation. As modern computational technology approaches its quantum limits, the potential for a quantum advantage is becoming increasingly appealing. This has motivated diverse implementations of quantum hardware from trapped ions [2, 3], cold atoms in optical lattices [4, 5], superconducting circuits [6, 7], photons [8, 9] to liquid and solid-state NMR [10, 11] and quantum dots [12].

The phrase “quantum simulation” often refers to (as originally conceived) the simulation of a *quantum* system [13]. However, this is not the only avenue in which we find quantum advantages. For instance, there is a variety of *classical* systems that can be simulated quantumly with advantage [14] including simulating thermal states [15], fluid flows [16, 17], electromagnetic fields [18], diffusion processes [19, 20], Burger’s equation [21], and molecular dynamics [22].

Quantum advantages can also be found outside of the realm of simulation. Some mathematical problems can be solved more efficiently using a quantum computer. The most well-known of these include Shor’s factorization algorithm [23], Grover’s quantum search algorithm [24], quantum eigen-decomposition algorithm [25], and quantum linear systems algorithm [26]. For factorization, Shor’s algorithm scales polynomially [23] while the best classical algorithm currently known scales exponentially [27]. While neither algorithm has been proven optimal, many believe that the separation in scaling is real [28].

Quantum advantages also exist in the context of stochastic process generation. Sequential generation and simultaneous generation are two important problems in this field [29]. In 1988, Crutchfield and Young [30] introduced memory efficient classical algorithms for both of these problems. While there are a small number of known cases (processes) for which this algorithm can be surpassed [31–33], there remains no better general classical algorithm. Our focus here is the problem of *simultaneous generation*, the potential quantum advantage therein, and the separation in classical-quantum scaling. [Quantum algorithms for sequential generation have been studied recently [34–36].]

Reference [37] provided a quantum algorithm that can generally perform simultaneous generation using less memory than the best known classical algorithms. Recently, we introduced a new quantum algorithm—the *q-machine*—that improved this efficiency. The latter demonstrated constructively how attention to higher-order correlations in the stochastic process can lead to an improved quantum algorithm for generation [38]. A sequel provided more detailed analysis and derived the quantum advantage of the q-machine in closed form [39]. This quantum advantage has also been verified experimentally for a simple case [40]. Just as for integer factorization, proof of optimality of a simultaneous-generation algorithm is challenging in both classical and quantum settings. However, with minor restrictions, one can show that the current quantum algorithm is almost always more efficient than the classical [38].

While the existing results demonstrate a quantum advantage for generic processes, a significant question remains: *What is the scaling behavior of this advantage?* That is, to truly understand the nature of the advantage, it is critical to know how it depends on problem size. The strong separation in scaling between the classical and quantum integer factorization algorithms led many to expect that the separation will persist even as new algorithms are developed. We wish to demonstrate an analogous separation in scaling, thus solidifying the importance of the current quantum construction—the q-machine.

We choose as our testing ground the generation of equilibrium states for the one-dimensional Ising system with  $N$ -nearest neighbor interaction. Here, the coupling range  $N$  is our problem-size parameter. We choose a spin-spin coupling that decays as a power law in  $N$ . This is a natural choice, both since this model has been studied in detail over four decades [41–44] and since it provides a physically grounded benchmark.

To understand the use of such a system in this problem context, consider a one-dimensional chain of spins (with arbitrary classical Hamiltonian) in contact with a thermal reservoir. After thermalizing, the resulting bi-infinite chain of spins (considered together at some instant  $t = t_0$ ) can be regarded as a (spatial) stochastic process. Successful generation of this stochastic process is then equivalent to generating its equilibrium states.

We quantitatively define the quantum advantage as the ratio of necessary memories for classical and quantum algorithms. Our main result is that the quantum advantage scales as  $NT^2/\log T$ . We also show that classically simulating Dyson’s original model requires infinite memory. In other words, exact classical simulation of the Dyson spin chain is impossible.

**Dyson-Ising Spin Chain** We begin with a general one-dimensional ferromagnetic Ising spin chain [50, 51] defined by the Hamiltonian:

$$\mathcal{H} = - \sum_{\langle i,j \rangle} J(i,j) s_i s_j , \quad (1)$$

in contact with thermal bath at temperature  $T$ , where  $s_i$ , the spin at site  $i$ , takes on values  $\{+1, -1\}$ , and  $J(i,j) \geq 0$  is the spin coupling constant between sites  $i$  and  $j$ . [Throughout,  $T$  denotes the effective temperature  $k_B T$ .] Assuming translational symmetry, we may replace  $J(i,j)$  by  $J(k)$  with  $k \equiv |i - j|$ . 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 the following, we consider couplings that decay by a power law:

$$J(k) = \frac{J_0}{k^\delta} , \quad (2)$$

where  $\delta > 0$ . The spin chain resulting from these assumptions is called the *Dyson* model [41]. To approximate such an infinite-range system we consider one with finite-range interactions.

For every interaction range  $N$ , we define the approximating Hamiltonian:

$$\mathcal{H}_N = - \sum_i \sum_{k=1}^N \frac{J_0}{k^\delta} s_i s_{i+k} . \quad (3)$$

[ $N$  is the interaction range and should not be mistaken with the size of the lattice which is infinite here.] This class of Hamiltonians can certainly be studied in its own right, not simply as an approximation. But why is the Dyson model interesting? The ferromagnetic Ising linear spin chain with finite-range interaction cannot undergo a phase transition at any positive temperature [52]. In contrast, the Dyson model has a standard second-order phase transition for a range of  $\delta$ . Dyson analytically proved [41] that a phase transition exists for  $1 < \delta < 2$ . The existence of a transition at  $\delta = 2$  was established much later [43]. It is also known that there exists no phase transition for  $\delta > 3$  [42], 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 [44], however, this fact has resisted analytical proof. For  $\delta \leq 1$ , the model is considered nonphysical since the energy becomes non-extensive.

Notably, the driven quantum Dyson model has been studied experimentally of late, since it exhibits many interesting nonequilibrium phases, such as the recently introduced *discrete time crystal* (DTC) [45]. The experimental system consists of a lattice of hundreds of spin-half particles stored in a Penning trap. Particles have been chosen to be  $^9\text{Be}^+$  [46],  $^{40}\text{Ca}^+$  [47] or  $^{171}\text{Yb}^+$  [48, 49] ions. Using a combination of static electric and magnetic fields, the Penning trap confines ions. A general spin-spin coupling is implemented with an optical dipole force (ODF) induced by a pair of off-resonance laser beams. The ODF then produces Dyson-type interactions, where  $\delta$  is tunable over  $0 \leq \delta \leq 3$ . Physically,  $\delta = 1, 2, 3$  corresponds to Coulomb-like, monopole-dipole, and dipole-dipole couplings, respectively.

For these reasons this family of Hamiltonians, derived from the Dyson spin system, offer a controlled way to investigate the consequences of nontrivial correlations.

**Simulators** The concept of a stochastic process is very general. Any physical system that exhibits stochastic dynamics in time or space may be thought of as *generating* a stochastic process. Here, we consider not time evolution, but rather the spatial “dynamic”. For example, consider a one-dimensional spin chain with arbitrary classical Hamiltonian in contact with thermal reservoir. After thermalizing, 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 spatial-translation invariant configurations defines a stationary stochastic process.

We focus, in particular, on *stationary, discrete-time, discrete-valued stationary stochastic processes*. Informally, such a process can be seen as a joint probability distribution  $\mathbb{P}(\cdot)$  over the bi-infinite chain of random variables  $\dots X_{-1} X_0 X_1 \dots$ . Formally, the process denoted by  $\mathcal{P} = \{\mathcal{A}, \Sigma, \mathbb{P}(\cdot)\}$  is a probability space [53, 54]. Each spin random variable  $X_i$ ,  $i \in \mathbb{Z}$ , takes values in the set  $\mathcal{A}$ . For specificity, the observed symbols come from an alphabet  $\mathcal{A} = \{\downarrow, \uparrow\}$  of local spin states, but our results easily extend to any finite alphabet.  $\mathbb{P}(\cdot)$  is the probability

measure over the bi-infinite chain of random variables  $X_{-\infty:\infty} = \dots X_{-2}X_{-1}X_0X_1X_2\dots$  and  $\Sigma$  is the  $\sigma$ -algebra generated by the cylinder sets in  $\mathcal{A}^\infty$ . Stationarity means that  $\mathbb{P}(\cdot)$  is invariant under index translation. That is,  $\mathbb{P}(X_iX_{i+1}\dots X_{i+m}) = \mathbb{P}(X_{i+n}X_{i+1+n}\dots X_{i+m+n})$ , for all  $m \in \mathbb{Z}^+$  and  $n \in \mathbb{Z}$ . For more information on stochastic processes generated by spin system we refer to Refs. [55, 56].

Consider a device that generates a stochastic process. We call this device a *simulator* of the process if and only if there is no way to distinguish the process outputs from those of the simulator. Given a physical system that yields a stochastic process, a device that generates this process is then said to simulate the physical system. In some contexts, the word “simulation” implies an approximation. In contrast, we require our simulators to be exact.

How do these simulators work? Generally, we implement the algorithms by writing computer programs. Two common formalisms used as algorithms for generating stochastic processes are *Markov Chains* (MC) [57, 58] and *Hidden Markov Models* (HMM) [53, 59, 60]. The latter can be significantly more compact in their representations (more efficient algorithms) and, for this reason, are sometimes the preferred implementation choice.

HMMs represent the generating mechanism for a given process by a tuple  $\{\mathcal{S}, \mathcal{A}, \{T^{(x)} : x \in \mathcal{A}\}\}$ , where  $\mathcal{S}$  is a finite set of states,  $\mathcal{A}$  is a finite alphabet (set of symbols), and  $\{T^{(x)} : x \in \mathcal{A}\}$  is a set of  $|\mathcal{S}| \times |\mathcal{S}|$  substochastic symbol-labeled transition matrices. The latter’s sum  $\mathbf{T} = \sum_{x \in \mathcal{A}} T^{(x)}$  is a stochastic matrix.

As an example, consider the Even Process [61, 62]. The process can be explained by a simple procedure. Consider a biased coin that with probability  $p$  generates heads and with  $1 - p$  generates tails. To generate the Even Process we use the algorithm:

$\left\{ \begin{array}{l} \text{Step A: Flip the coin.} \\ \text{Step B: If the result is heads, output 0 and go to Step A. Else output 1 and go to Step C.} \\ \text{Step C: Output 1 and go to Step A.} \end{array} \right.$

This algorithm is depicted by the HMM shown on the left of Fig. 1a. For this HMM,  $\mathcal{S} = \{A, B\}$ ,  $\mathcal{A} = \{0, 1\}$ ,  $T^{(0)} = \begin{pmatrix} p & 0 \\ 0 & 0 \end{pmatrix}$ , and  $T^{(1)} = \begin{pmatrix} 0 & 1 - p \\ 1 & 0 \end{pmatrix}$ . The HMM, as an algorithm, simply tells the computer that: if we are in state  $A$  then, with probability  $p$ , output 0 and stay at state  $A$  and, with probability  $1 - p$ , output 1 and go to state  $B$ . If we are in state  $B$ , output 1 and go to state  $A$ .

The goal of sequential generation is to produce a very long realization of the process. For this, we use one computer with a code that runs the algorithm. At each step, the computer must memorize the current HMM state. Since the HMM has 2 states, we require 1 bit of memory for this process, independent of its bias  $p$ .

Here, though, we are interested in simultaneous generation where the goal is to generate  $M$  realizations of a process simultaneously, each of which is statistically independent of the others. The net result is  $M$  computers each with the above code, as on the right side of

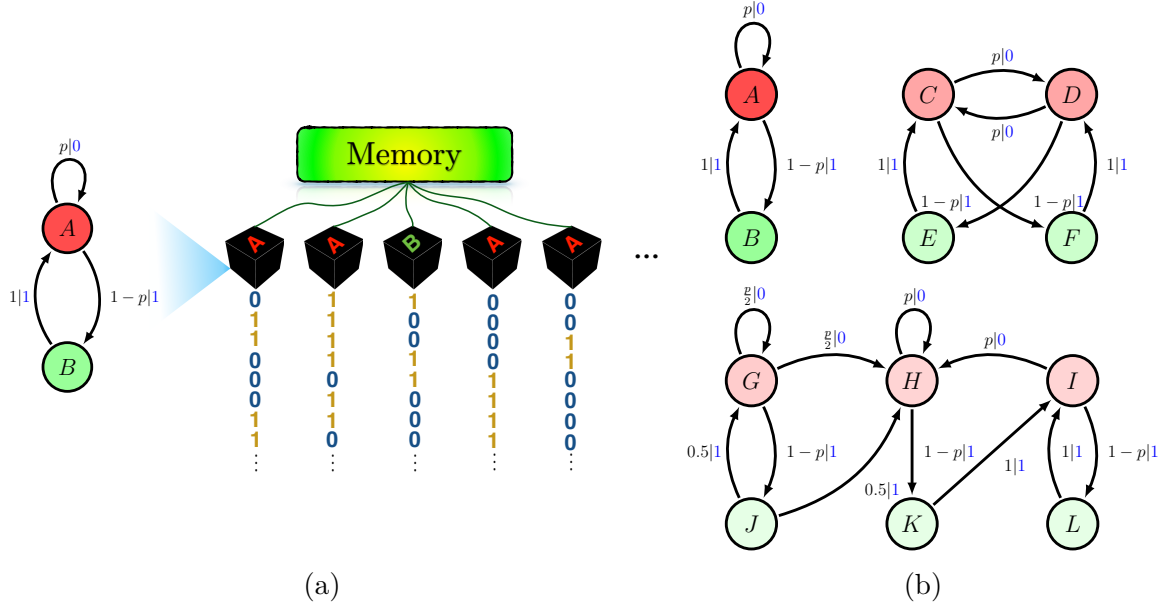


FIG. 1: (a) Left: Even Process  $\epsilon$ -machine. Right: Schematic of simultaneous generation problem. Each black box contains an Even Process generator. They all share the same memory for tracking the individual generator states. (b) Alternative HMMs: Even Process generators, each with different memory costs. Top: Unifilar HMMs, since for every state and symbol there is at most one outgoing edge from that state emitting that symbol. Below: Nonunifilar HMM, since for example state  $G$  can go to different states  $G$  or  $H$  emitting symbol  $0$ .

Fig. 1a. Similar to the sequential problem, each computer must memorize the current state of its HMM. If each computer uses its own memory, each needs 1 bit of memory as before. The total memory is then  $M$  bits.

However, we can reduce the amount of memory required by using one large shared memory among the computers. Figure 1a emphasizes this schematically. In this way, according to Shannon’s coding theorem [63], we can encode the HMM states to reduce the amount of memory down to  $MH(\mathcal{S}) \leq M$  bits, where  $H(\mathcal{S}) = -\Pr(A) \log_2 \Pr(A) - \Pr(B) \log_2 \Pr(B)$ . The memory per instance is then just  $H(\mathcal{S})$ .

Every process has an infinite number of alternative HMMs that generate it. For example, Fig. 1b shows three HMMs that each generate the Even Process, each with different  $H(\mathcal{S})$  and as a result different memory costs. Now, an important question when considering all possible generators is, which HMM needs the minimum memory or, equivalently, minimum  $H(\mathcal{S})$ ?

**$\epsilon$ -Machine** A *unifilar* HMM is one in which each row of each substochastic matrix has at most one nonzero element. Informally, this means the current state and next symbol uniquely determine the next state. Many statistical and informational quantities can be calculated in closed form from a process’s unifilar HMM; see Ref. [73] and discussion therein. For example, in Fig. 1b the top two HMMs are unifilar and the bottom one is nonunifilar. For a given process, finding the optimal HMM for simultaneous generation—an HMM with

minimum state-entropy  $H(\mathcal{S})$ —in the space of all HMMs is still an open question. Restricting to the space of unifilar HMMs, though, the optimal HMM can be found. It is the  $\epsilon$ -machine [64], first introduced in Ref. [30].  $\epsilon$ -Machine states  $\mathcal{S}$  are called *causal states*. Due to  $\epsilon$ -machine’s unifilarity property, every generated past uniquely maps to a causal state. A process’ *statistical complexity*  $C_\mu$  [64] is the the Shannon entropy of the  $\epsilon$ -machine’s stationary state distribution:  $C_\mu = H(\mathcal{S}) = -\sum_{\sigma \in \mathcal{S}} \Pr(\sigma) \log_2 \Pr(\sigma)$ . And, this is the required memory for simultaneous generation.

Attempts have been made to find smaller models among *nonunifilar* HMMs [31]. As of now, though, only a handful of examples exist [31–33, 65]. Practically speaking, the  $\epsilon$ -machine is the most memory-efficient algorithm for generating stochastic processes. Its memory  $C_\mu$  has been determined for a wide range of physical systems [66–72]. Helpfully, it and companion informational measures are directly calculable from the  $\epsilon$ -machine, many in closed-form [73]. We denote the process generated by the physical system with Hamiltonian Eq. (3) at temperature  $T$  by  $\mathcal{P}(N, T)$ . How do we construct the  $\epsilon$ -machine that simulates the process  $\mathcal{P}(N, T)$ ? First, we must define process’ *Markov order* [58]: the minimum history length  $R$  required by any simulator to correctly continue a configuration. Specifically,  $R$  is the smallest integer such that:

$$\mathbb{P}(X_t | \dots, X_{t-2}, X_{t-1}) = \mathbb{P}(X_t | X_{t-R}, \dots, X_{t-2}, X_{t-1}) . \quad (4)$$

[More precisely, an ensemble of simulators must yield an ensemble of configurations that agree (conditioned on that past) with the process’ configuration distribution.]

Reference [56, Eqs. (84) – (91)] showed that for any finite and nonzero temperature  $T$ ,  $\mathcal{P}(N, T)$  has Markov order  $N$ . One concludes that sufficient information for generation is contained in the configuration of the  $N$  previously generated spins. (Figure 2a shows this fact schematically for  $N = 2$ .) More importantly, the  $\epsilon$ -machine that simulates  $\mathcal{P}(N, T)$  has  $2^N$  causal 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* [74, 75]: the smallest integer  $K$  such that  $H[\mathcal{S}_K | X_0 X_1 \dots] = 0$ , where  $H[W | Z]$  is the conditional entropy [63] and  $\mathcal{S}_K$  is the random variable for  $K^{\text{th}}$  state of the  $\epsilon$ -machine after generating symbols  $X_0, X_1 \dots$ . Using the fact that the  $\epsilon$ -machine’s states are in one-to-one correspondence with the set of length- $N$  spin configurations [56], it is easy to see that  $\mathcal{P}(N, T)$ ’s cryptic order  $K = N$ , the Markov order. We will use this fact in the quantum algorithm construction to follow.

Figure 2b shows the  $\epsilon$ -machines of the processes  $\mathcal{P}(N, T)$  for  $N = 1, 2$ , and 3. Let’s explain. First, consider the spin process  $\mathcal{P}(1, T)$  that, as we pointed out, is a Markov-order  $R = 1$  process. This means to generate the process the simulator only need remember the last spin generated. In turn, this means the  $\epsilon$ -machine (Fig. 2b left) has two states,  $\uparrow$  and  $\downarrow$ . If the last observed spin is  $\uparrow$ , the current state is  $\uparrow$  and if it is  $\downarrow$ , the current state is  $\downarrow$ . We denote the probability of generating a  $\downarrow$  spin given a previous generated  $\uparrow$  spin by  $p_{\uparrow\downarrow}$ . The probability of an  $\uparrow$  spin following a  $\uparrow$  spin is the complement.

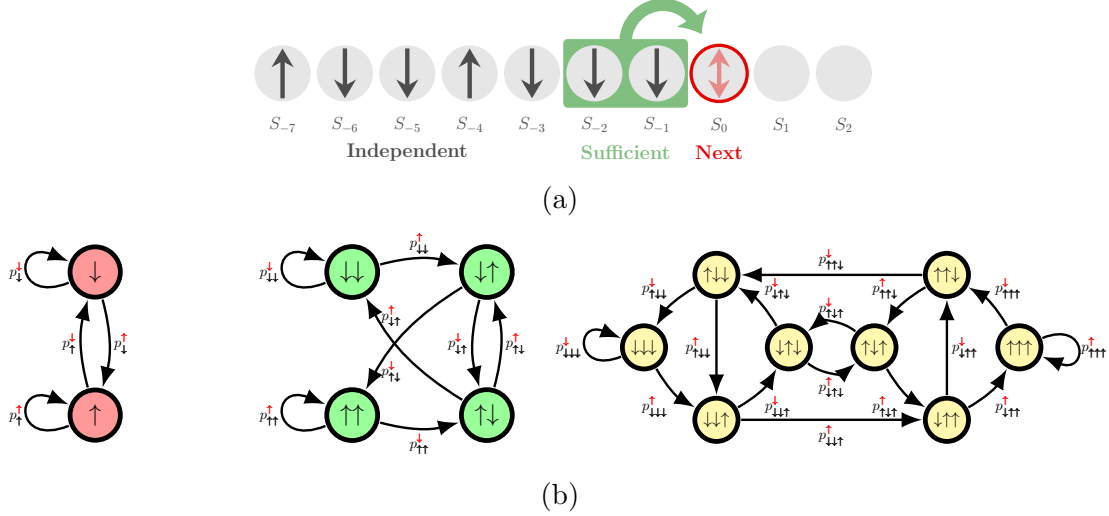


FIG. 2: (a) A Markov order- $N$  process generates a spin configuration from left-to-right. Markov order  $N = 2$  shown. The values of an isolated spin  $S_0$ , say, is undetermined. To make this (stochastic) choice consistent with the overall process and the particular instantiation on the left, it is sufficient to consider only the previous  $N$  (2) spins (highlighted in green). (b)  $\epsilon$ -Machine generators of 1D-configuration stochastic processes in Dyson-Ising systems of increasing correlational complexity ( $N = 1, 2, 3$ ):  $\mathcal{P}(1, T)$  (left),  $\mathcal{P}(2, T)$  (middle) and  $\mathcal{P}(3, T)$  (right).

Second, consider the process  $\mathcal{P}(2, T)$  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. 2b middle) has four states that we naturally label  $\uparrow\uparrow$ ,  $\uparrow\downarrow$ ,  $\downarrow\uparrow$ , and  $\downarrow\downarrow$ . If the last observed spin pair  $x_{-1}x_0$  is  $\uparrow\downarrow$ , the current state is  $\uparrow\downarrow$ . Given this state, the next spin will be  $\uparrow$  with probability  $p_{\uparrow\downarrow}^{\uparrow}$  and  $\downarrow$  with probability  $p_{\uparrow\downarrow}^{\downarrow}$ . Note that this scheme implies that each state has exactly two outgoing transitions. That is, not all state-to-state transitions are allowed in the  $\epsilon$ -machine.

Having identified the state space, to complete the  $\epsilon$ -machine construction we determine the  $\epsilon$ -machine transition probabilities  $\{T^{(x)}\}_{x \in \mathcal{A}}$ . To do this, we first compute the transfer matrix  $V$  for the Ising  $N$ -nearest neighbors with the Hamiltonian in Eq. (3) at temperature  $T$  and then extract conditional probabilities, following Ref. [56]. (See the Method section following for details.) The minimum memory for simultaneous generation or, as it is called, the statistical complexity  $C_\mu(N, T)$  of process  $\mathcal{P}(N, T)$  follows straightforwardly from the process'  $\epsilon$ -machine.

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

A process' quantum simulator is a pair  $\{f, \mathcal{M}\}$ , where  $f : \mathcal{A}^\infty \rightarrow \Omega$  is a function from the set  $\mathcal{A}^\infty$  of past sequences to a set of quantum states  $\Omega$  and  $\mathcal{M}$  is some measurement process. Given a particular past  $x_{-\infty:0}$ , applying the measurement  $\mathcal{M}$  to the quantum state  $f(x_{-\infty:0})$



leads to a correct probability distribution over future  $\mathbb{P}(x_{0:n}|x_{-\infty:0})$ . If  $f(\cdot)$  is a deterministic function, the simulator is called unifilar and if  $f$  is probabilistic function, the simulator is called nonunifilar. After generating  $x_{0:n}$ , the new past is  $x_{-\infty:n}$  and  $f$  can be used to map it to the new quantum state  $f(x_{-\infty:n})$ . By repeating the same measurement and mapping procedure, we generate a realization of the process. One can also define quantum simulator in a way that  $\mathcal{M}$  automatically maps  $f(x_{-\infty:0})$  to the correct quantum state  $f(x_{-\infty:n})$  and generates the correct probability distribution over  $x_{0:n}$  [77].

Reference [38] introduced a class of unifilar simulators, called *q-machines*, that can generate arbitrary processes. As in the classical setting, nonunifilar quantum simulators are much less well understood [34, 35, 65]. The q-machine construction depends on an encoding length  $L$ , each with its own quantum cost  $C_q(L)$ . Each of these simulators simulate the same process correctly. It is known that the cost  $C_q(L)$  is constant beyond the process' cryptic order [75]. Based on numerical evidence, it is conjectured that this is also the minimal  $C_q$  value. Thus, 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 [75]. And so,  $C_\mu/C_q$  gives us our quantitative measure of *quantum advantage*.

Reference [39] recently introduced efficient methods for calculating  $C_q$  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. The quantum advantage  $C_\mu/C_q$  has been investigated both analytically [38, 39, 76–78] and experimentally [40].

A process' q-machine is straightforward to construct from its  $\epsilon$ -machine. First, since the  $\epsilon$ -machine is unifilar, every generated past realization maps to a unique causal state. Second, every causal state  $\sigma_i$  maps to a pure quantum state  $|\eta_i\rangle$ . Using these two maps we can map every generated past realization uniquely to a quantum state. 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:

$$|\eta_i\rangle \equiv \sum_{w \in \mathcal{A}^K} \sum_{\sigma_j \in \mathcal{S}} \sqrt{\mathbb{P}(w, \sigma_j | \sigma_i)} |w\rangle |\sigma_j\rangle, \quad (5)$$

where  $w$  denotes a length- $K$  sequence and  $\mathbb{P}(w, \sigma_j | \sigma_i) = \mathbb{P}(X_0 \cdots X_{K-1} = w, \mathcal{S}_{K-1} = \sigma_j | \mathcal{S}_0 = \sigma_i)$ . 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  $|w\rangle = |x_0\rangle \cdots |x_{K-1}\rangle$ . Performing a projective measurement in the  $|w\rangle |\sigma\rangle$  basis results in a correct probability distribution. After generation of a particular realization by measurement, the next corresponding quantum state can be indicated uniquely. This means we can repeat the measurement process and continue generating the process.

Now, let us return to simultaneous generation where the goal is to generate  $M$  process realizations simultaneously where each is statistically independent of the others. As before,

we have  $M$  q-machines as in Fig. 1a. Also similar to the classical setting, we can reduce the amount of required memory by having the q-machines use a single shared memory. According to quantum coding theorem, we can encode the HMM states to reduce the amount of memory to  $MS(\rho)$  qubits where  $S(\cdot)$  is von Neumann entropy and  $\rho$  is the density matrix defined by:

$$\rho = \sum_i \pi_i |\eta_i\rangle \langle \eta_i| . \quad (6)$$

As a result, each q-machine needs  $C_q = S(\rho)$  qubits of memory for simultaneous generation.

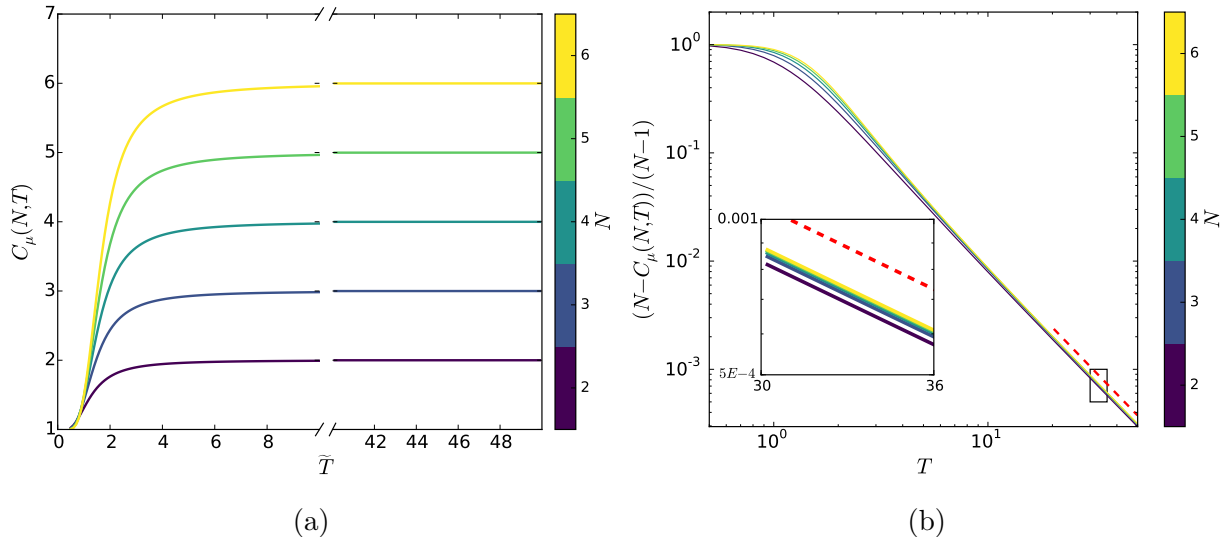


FIG. 3: (a) Classical memory  $C_\mu(N, T)$  required for simulating process  $\mathcal{P}(N, T)$  for interaction ranges  $N = 1, \dots, 6$ , a range of temperatures  $T = 1, \dots, 50$ , and  $\delta = 2$ . Note  $C_\mu(\cdot)$  is an increasing function of  $N$  and  $T$ . (b) Rescaling the classical memory requirement  $C_\mu(N, T)$  to  $(N - C_\mu)/(N - 1)$  shows a tight data collapse, which is especially strong at high temperatures ( $T > 2$ ). The asymptotic behavior is a power-law with scaling exponent  $\gamma = 2$ . The inset zooms in to show  $C_\mu$ 's convergence with increasing  $N$ . While the figure shows the case  $\delta = 2$ , the slope  $\gamma$  at high  $T$  is independent of  $\delta$ .

**Analysis** We begin by considering the case where spin couplings decay with exponent  $\delta = 2$ . Figure 3a displays  $C_\mu(N, T)$  and Fig. 4a displays  $C_q(N, T)$ —the  $C_\mu$  and  $C_q$  of processes  $\mathcal{P}(N, T)$ —versus  $T$  for interaction ranges  $N = 1, \dots, 6$ . The most striking feature is that the classical and quantum memory requirements exhibit qualitatively different behaviors.

The 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  $\mathcal{P}(1, T)$ ) high temperature makes spin- $\uparrow$  and spin- $\downarrow$ , and thus the corresponding states, equally likely. [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.]

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  $p_{\uparrow}^{\downarrow}$  and  $p_{\downarrow}^{\uparrow}$  converge to zero, while  $p_{\uparrow}^{\uparrow}$  and  $p_{\downarrow}^{\downarrow}$  converge to one. [It should be pointed out that at any finite temperature  $p_{\uparrow}^{\downarrow}$  and  $p_{\downarrow}^{\uparrow}$  are nonzero and, therefore, the  $\epsilon$ -machine states remains strongly-connected.] This is reflected in the convergence of all curves at  $C_{\mu} = 1$  bit. Equivalently, this means one needs only a single 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 ( $C_q$ ) of  $S(\rho) = 1$  qubit.

Figure 3a 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. 3b. 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 in Fig. 3b) numerical estimates from simulations indicate that this scaling is given by  $\gamma \simeq 2.000$ . The scaling determines how the classical memory saturates at high temperature. This behavior is generic for different coupling decay values  $\delta > 1$  and, more to the point, the scaling 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 4a shows that  $C_q$  is also a power-law in  $T$ . By measuring this scaling exponent in the same way as above, we determined that  $\alpha \simeq 2.000$ . Furthermore, we find analytically that for high  $T$ :

$$C_q(N, T) \propto \frac{\log_2(T)}{T^2} . \quad (7)$$

To see this, first consider nearest-neighbor coupling  $N = 1$ . Due to symmetry we have  $p \equiv \Pr(\uparrow | \uparrow) = \Pr(\downarrow | \downarrow) = F/D$ , where  $F = \exp(\beta J)$  and  $D = \exp(\beta J) + \sqrt{\exp(-2\beta J)}$  with  $\beta = 1/T$ . At high temperature  $\beta$  is small and we have  $D = 2 + \beta^2$  and  $F = 1 + \beta + \beta^2$ . Again, by symmetry we have  $\pi_1 = \pi_2 = 1/2$  and, therefore, the density matrix in Eq. (6) is:

$$\rho = \begin{pmatrix} 1/2 & \sqrt{p(1-p)} \\ \sqrt{p(1-p)} & 1/2 \end{pmatrix} , \quad (8)$$

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

$$C_q = S(\rho) \simeq - \left( \frac{\beta^2}{4} \log_2 \frac{\beta^2}{4} + \left( 1 - \frac{\beta^2}{4} \right) \log_2 \left( 1 - \frac{\beta^2}{4} \right) \right) \simeq \frac{\log_2(T)}{2T^2} . \quad (9)$$

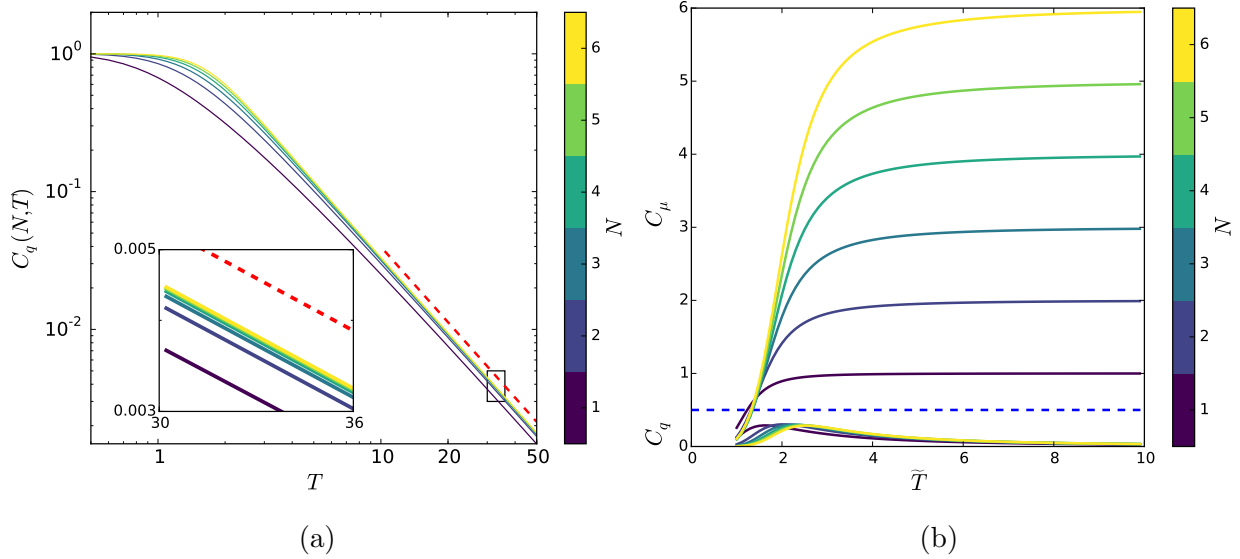


FIG. 4: (a)  $C_q(\cdot)$  is an increasing function of  $N$ , but a decreasing function of  $T$  and bounded by 1 qubit, independent of  $N$  and  $T$ . Quantum memory  $C_q(N, T)$ , similar to  $C_\mu(N, T)$ , shows a data collapse in  $N$  that is especially tight at high temperature ( $T > 2$ ). The asymptotic behavior is a power-law with numerically estimated scaling exponent  $\alpha = 2$ . (Red dashed line.) The lower inset zooms to highlight convergence with increasing  $N$ . Though the curves are for the case with  $\delta = 2$ , the slope  $\alpha$  at high  $T$  is independent of  $T$ . (b) Magnetic field effects on classical  $C_\mu(N, T)$  and quantum  $C_q(N, T)$  memory requirements for simulating the processes generated by Hamiltonian  $\widehat{\mathcal{H}}_N$  for  $N = 1, \dots, 6$  over a range of temperatures  $T = 1, \dots, 10$  at  $B = 0.3$ .  $C_q(N, T)$  curves are those under the dashed blue line.

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

$$\frac{1}{T^2} < \frac{\log_2(T)}{T^2} < \frac{1}{T^{2-r}}, \quad (10)$$

for all  $r > 0$ . This explains the fat tails of  $C_q$  for large  $T$  and establishes that for  $N = 1$  the scaling exponent is  $\alpha = 2$ .

Increasing the temperature the link between spins weakens. At high temperature the only important neighbor is the nearest. As a consequence, the high temperature behavior is similar to the case of  $N = 1$  and, in addition, it is independent of  $N$ . 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  $\delta$ . Notably, in this case no rescaling is required. The exponent 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.

Define the *quantum advantage*  $\eta$  as the ratio of the minimum required memory for the

classical simulation to the minimum required memory for the quantum simulation:

$$\eta(N, T) \equiv C_\mu(N, T)/C_q(N, T) . \quad (11)$$

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

$$\eta(N, T) \propto N \frac{T^2}{\log_2(T)} , \quad (12)$$

which scales faster than any  $T^r$  for  $r < 2$ . Thus, the answer to our motivating question is that the quantum advantage does, in fact, display scaling: it increases with interaction range  $N$  and also increases strongly with temperature  $T$ .

Up to this point we focused on finite interaction-range systems, interpreting the chosen models as a family of approximations to the Dyson model. Consider, though, Dyson’s original spin chain [41] which has infinite-range interactions. In this case, the classical memory cost of simulation diverges:  $\lim_{N \rightarrow \infty} C_\mu(N, T) \rightarrow \infty$ . That is, it is impossible to simulate the Dyson model classically. In contrast, quantum memory cost is finite— $\lim_{N \rightarrow \infty} C_q(N, T) < 1$  qubit—and so it can be simulated quantumly. There is perhaps no clearer statement of quantum advantage.

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:

$$\widehat{\mathcal{H}}_N = - \sum_i \sum_{k=1}^N \frac{J_0}{k^\delta} s_i s_{i+k} - \sum_i B s_i . \quad (13)$$

Figure 4b shows that, due to symmetry breaking at low temperature, both  $C_q(N, T)$  and  $C_\mu(N, T)$  converge to zero. (All spins at low temperature align with magnetic field and, as a consequence, no memory is needed.) The high temperature behaviors for both memory costs are the same as before, though, and the quantum advantage remains the same.

## DISCUSSION

It is notoriously hard to find quantum advantage and even harder to prove [79]. We found 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 a generic scaling behavior—quadratic in temperature and linear in interaction range. What does this mean? The most striking conclusion is that a strongly interacting classical system can be simulated with unbounded quantum advantage. One stark contrast is that it is impossible to classically simulate Dyson’s original spin chain while quantum simulators can do so and with finite memory cost.

How broadly might we expect to see this quantum advantage? Or, is it merely a feature of strongly coupled spin systems? Define a *universal* spin model as one that can simulate any other spin model. That is, by using the low-energy sector of such universal models, the physics of every classical spin model can be reproduced. Recently, Ref. [80] showed that the 2D Ising model with external fields is universal in this sense. This suggests that the quantum advantage described here may not be limited to the particular spin system we consider, but might also be universal. As a result, one should expect to see the quantum advantage for other physical systems.

The Ising model has lent great insight to condensed matter physics, however it is a classical model. Given that we are 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? And, is this their defining characteristic?

Here, we studied the cost of exact simulation of stochastic processes. Both classical and quantum costs, though, can be very different when approximation is allowed. For example, at high (but finite) temperature, we can approximate the process  $\mathcal{P}(N, T)$  as independent, identically distribution (IID). One does not require any classical or quantum memory to generate an IID process and, as a result, there would be no quantum advantage. Apparently, the difference between required classical memory for exact simulation and approximate simulation can be quite large. In contrast, the price we pay to go from approximate to exact quantum simulation is relatively small.

## METHOD

We show how to construct the  $\epsilon$ -machine simulator of the process  $\mathcal{P}(N, T)$ , following Ref. [81]. 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  $i = 1, \dots, N$ , respectively. We map the left and right block configurations each to an integer  $\eta_*$  by:

$$\eta_* = \sum_{i=1}^N \left( \frac{s_i^* + 1}{2} \right) 2^{i-1}, \quad (14)$$

where  $*$  can be either  $L$  or  $R$ . For each block we can have  $2^N$  different configurations. Consequently, the label  $\eta_*$  varies between 0 and  $2^N - 1$ . The internal energy of a given block with configuration  $\eta_*$  is given by:

$$X_{\eta_*} = -B \sum_{i=1}^N s_i^* - \sum_{i=1}^{N-1} \sum_{k=1}^{N-i} J_i s_k^* s_{k+i}^*, \quad (15)$$

and the interaction energy between two blocks is:

$$Y_{\eta_L, \eta_R} = - \sum_{i=1}^N \sum_{k=1}^i J_i s_{N-k+1}^L s_k^R . \quad (16)$$

With these we construct the transfer matrix:

$$V_{\eta_L, \eta_R} = e^{-(1/2X_{\eta_L} + Y_{\eta_L, \eta_R} + 1/2X_{\eta_R})/T} . \quad (17)$$

The right eigenvector of  $V$  corresponding to the largest eigenvalue is denoted by  $u$ . Reference [56] shows that the  $\epsilon$ -machine labeled-transition matrices can be written as:

$$T_{\eta_0, \eta_1}^{(x)} = \begin{cases} \frac{1}{\lambda} V_{\eta_0, \eta_1} \frac{u_{\eta_1}}{u_{\eta_0}}, & \eta_1 = (\lfloor \frac{\eta_0}{2} \rfloor + x(2^{N-1})) \\ 0, & \text{otherwise} \end{cases} , \quad (18)$$

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

## ACKNOWLEDGMENTS

We thank both reviewers for their insightful comments. JPC thanks the Santa Fe Institute for its hospitality during visits as an External Faculty member. 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.

## AUTHOR CONTRIBUTIONS STATEMENT

CA, JPC, and JRM conceived of the project and contributed to the writing of the final manuscript. CA and JRM developed theory and performed the calculations. JPC supervised the project.

## COMPETING FINANCIAL INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

---

[1] Feynman, R.P. Simulating physics with computers. *Intl. J. Theo. Physics* **21**, 467–488 (1982)

- [2] Kihwan, K. *et al.* Quantum simulation of frustrated Ising spins with trapped ions. *Nature* **465**, 590–593 (2010)
- [3] Blatt, R. & Roos, C.F. Quantum simulations with trapped ions. *Nature Phys* **8**, 277–284 (2012)
- [4] Jaksch, D. & Zoller, P. The cold atom Hubbard toolbox. *Ann. Phys* **315**, 52–79 (2005)
- [5] Bloch, I., Dalibard, J. & Nascimbene, S. Quantum simulations with ultracold quantum gases. *Nature Phys* **8**, 267–276 (2012)
- [6] Clarke, J. & Wilhelm, F.K. Superconducting quantum bits. *Nature* **453**, 1031–1042 (2008)
- [7] Houck, A.A., Türeci, H.E. & Koch, J. On-chip quantum simulation with superconducting circuits. *Nature Phys* **8**, 292–299 (2012)
- [8] Lanyon, B.P., *et al.* Towards quantum chemistry on a quantum computer. *Nature Chem.* **2**, 106–111 (2010)
- [9] A. Aspuru-Guzik and P. Walther. Photonic quantum simulators. *Nature Phys.* **8**, 285–291 (2012)
- [10] Du, J. *et al.* NMR implementation of a molecular hydrogen quantum simulation with adiabatic state preparation. *Phys. Rev. Let.* **104**, 030502 (2010).
- [11] Zhang, J., Yung, M.H., Laflamme, R., Aspuru-Guzik, A. & Baugh, J. Digital quantum simulation of the statistical mechanics of a frustrated magnet. *Nature Comm.* **3**, 880 (2012).
- [12] Byrnes, T., Kim, N.Y., Kusudo, K., & Yamamoto, Y. Quantum simulation of Fermi-Hubbard models in semiconductor quantum-dot arrays. *Phys. Rev. B.* **78**, 075320 (2008).
- [13] Georgescu, I.M., Ashhab, S. & Nori, F. Quantum simulation. *Rev. Mod. Phys.*, **86**, 153 (2014).
- [14] Meyer, D.A. Quantum computing classical physics. *Phil. Trans. Roy. Soc. London A.* **360**, 395–405 (2002).
- [15] Yung, M.H., Nagaj, D., Whitfield, J.D., & Aspuru-Guzik, A. Simulation of classical thermal states on a quantum computer: A transfer-matrix approach. *Phys. Rev. A.* **82**, 060302 (2010).
- [16] Yepez, J. Quantum lattice-gas model for computational fluid dynamics. *Phys. Rev. E.* **63**, 046702 (2001).
- [17] Yepez, J. Quantum computation of fluid dynamics. In *Quantum Computing and Quantum Communications*, pages 34–60. Springer, 1999.
- [18] Sinha, S. & Russer, P. Quantum computing algorithm for electromagnetic field simulation. *Quant. Info. Processing.* **9**, 385–404 (2010).
- [19] Yepez, J. Quantum lattice-gas model for the diffusion equation. *Intl. J. Mod. Phys. C.* **12**, 1285–1303 (2001).
- [20] Berman, G.P., Ezhov, A.A., Kamenev, D.I. & Yepez, J. Simulation of the diffusion equation on a type-ii quantum computer. *Phys. Rev. A.* **66**, 012310 (2002).
- [21] Yepez, J. Quantum lattice-gas model for the burgers equation. *J. Stat. Phys.* **107**, 203–224 (2002).
- [22] Harris, S.A. & Kendon, V.M. Quantum-assisted biomolecular modelling. *Phil. Trans. Roy. Soc. London A.* **368**, 3581–3592 (2010).



- [23] Shor, P.W. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM Review* **41**, 303–332 (1999).
- [24] Grover, L.K. A fast quantum mechanical algorithm for database search. In *Proceedings of the Twenty-Eighth Annual ACM Symposium on Theory of Computing*, pages 212–219 (1996).
- [25] Abrams, D.S. & Lloyd, S. Quantum algorithm providing exponential speed increase for finding eigenvalues and eigenvectors. *Phys. Rev. Lett.* **83** 5162 (1999).
- [26] Harrow, A.W., Hassidim, A., & Lloyd, S. Quantum algorithm for linear systems of equations. *Phys. Rev. Lett.* **103** 150502 (2009).
- [27] Pomerance, C. Fast, rigorous factorization and discrete logarithm algorithms. *Discrete Algo. Complexity* 119–143 (1987).
- [28] Bouland, A. Establishing quantum advantage. *XRDS: Crossroads, The ACM Magazine for Students* **23**, 40–44 (2016).
- [29] Aghamohammadi, C. & Crutchfield, J.P. Minimum memory for generating rare events. *Phys. Rev. E*. **95** 032101 (2017).
- [30] Crutchfield, J.P. & Young, K. Inferring statistical complexity. *Phys. Rev. Lett.* **63**, 105–108 (1989).
- [31] Löhr, W. & Ay, N. Non-sufficient memories that are sufficient for prediction. In *International Conference on Complex Sciences*, pages 265–276. Springer, (2009).
- [32] Löhr, W. & Ay, N. On the generative nature of prediction. *Adv. Complex Sys.* **12**, 169–194 (2009).
- [33] Löhr, W. Predictive models and generative complexity. *J. Systems Sci. Complex.* **25**, 30–45 (2012).
- [34] Monras, A., Beige, A. & Wiesner, K. Hidden quantum Markov models and non-adaptive read-out of many-body states. *arXiv:1002.2337* (2010).
- [35] Monras, A. & Winter, A. Quantum learning of classical stochastic processes: The completely positive realization problem. *J. Math. Phys.* **57**, 015219 (2016).
- [36] Stark, C.J. & Harrow, A.W. Compressibility of positive semidefinite factorizations and quantum models. *IEEE Trans. Info. Th.* **62** 2867–2880 (2016).
- [37] Gu, M., Wiesner, K., Rieper, E. & Vedral, V. Quantum mechanics can reduce the complexity of classical models. *Nature Comm.* **3** (2012).
- [38] Mahoney, J.R., Aghamohammadi, C. & Crutchfield, J.P. Occam’s quantum stop: Synchronizing and compressing classical cryptic processes via a quantum channel. *Sci. Reports.* **6**, (2016).
- [39] Riechers, P.M., Mahoney, J.R., Aghamohammadi, C. & Crutchfield, J.P. Minimized state complexity of quantum-encoded cryptic processes. *Phys. Rev. A*. **93**, 052317 (2016).
- [40] Palsson, M.S., Gu, M., Ho, J., Wiseman, H.M. & Pryde, G.J. Experimental quantum processing enhancement in modelling stochastic processes. *Science Advances* **3**, e1601302 (2017).
- [41] Dyson, F.J. Existence of a phase-transition in a one-dimensional Ising ferromagnet. *Comm. Math. Phys.* **12**, 91–107 (1969).

- [42] Fisher, M.E., Ma, S.K. & Nickel, B.G. Critical exponents for long-range interactions. *Phys. Rev. Lett.* **29**, 917 (1972).
- [43] Fröhlich, J. & Spencer, T. The phase transition in the one-dimensional Ising model with  $1/r^2$  interaction energy. *Comm. Math. Phys.* **84**, 87–101 (1982).
- [44] Blanchard, T., Picco, M. & Rajabpour, M.A. Influence of long-range interactions on the critical behavior of the Ising model. *Europhys. Lett.* **101**, 56003 (2013).
- [45] Zhang, J., Hess, P. W., Kyprianidis, A., Becker, P., Lee, A., Smith, J., Pagano, G., Potirniche, I.D., Potter, A. C., Vishwanath, A., Yao, N. Y. & Monroe, C. Observation of a Discrete Time Crystal. *Nature*. **543**, 217-220 (2017).
- [46] Britton, J.W., Sawyer, B.C., Keith, A.C., Wang, C.C.J., Freericks, J.K., Uys, H., Biercuk, M.J., & Bollinger, J.J. Engineered two-dimensional Ising interactions in a trapped-ion quantum simulator with hundreds of spins *Nature*. **484**, 489-492 (2012).
- [47] Jurcevic, P., Lanyon, B.P., Hauke, P., Hempel, C., Zoller, P., Blatt, R. & Roos C.F. Quasiparticle engineering and entanglement propagation in a quantum many-body system *Nature*. **511**, 202-205 (2014).
- [48] Islam, R., Senko, C., Campbell, W.C., Korenblit, S., Smith, J., Lee, A., Edwards, E.E., Wang, C.C.J., Freericks, J.K., & Monroe, C. Emergence and frustration of magnetism with variable-range interactions in a quantum simulator *Science*. **340**, 583-587 (2013).
- [49] Smith, J., Lee, A., Richerme, P., Neyenhuis, B., Hess, P.W., Hauke, P., Heyl, M., Huse, D.A. and Monroe, C. Many-body localization in a quantum simulator with programmable random disorder *Nature Phys.* **12**, 907-911 (2016).
- [50] Baxter, R.J. *Exactly solved models in statistical mechanics*. (Courier Corporation, 2007).
- [51] Aghamohammadi, A, Aghamohammadi, C. & Khorrami, M. Externally driven one-dimensional Ising model. *J. Stat. Mech.* **2012** P02004 (2012).
- [52] Rushbrooke, G.S. & Ursell, H.D. On one-dimensional regular assemblies. In *Mathematical Proceedings of the Cambridge Philosophical Society*, **44**, 263–271. (Cambridge University Press, 1948).
- [53] Upper, D.R. *Theory and Algorithms for Hidden Markov Models and Generalized Hidden Markov Models*. PhD thesis, University of California, Berkeley, (1997).
- [54] Travers, N.F. *Bounds on Convergence of Entropy Rate Approximations in Hidden Markov Processes*. PhD thesis, University of California, Davis, (2013).
- [55] Crutchfield, J.P. & Feldman, D.P. Statistical complexity of simple one-dimensional spin systems. *Phys. Rev. E*. **55**, 1239–1243 (1997).
- [56] Feldman, D.P. & Crutchfield, J.P. Discovering non-critical organization: Statistical mechanical, information theoretic, and computational views of patterns in simple one-dimensional spin systems. Santa Fe Institute Paper 98-04-026 (1998).
- [57] Norris, J.R. *Markov Chains*, volume 2. (Cambridge University Press, 1998).
- [58] Levin, D.A., Peres, Y. & Wilmer, E.L. *Markov Chains and Mixing Times*. (American Mathematical Society, 2009).

- [59] Rabiner, L.R. & Juang, B.H. An introduction to hidden Markov models. *IEEE ASSP Magazine*. **3**, 4–16 (1986).
- [60] Rabiner, L.R. A tutorial on hidden Markov models and selected applications in speech recognition. *Proc of the IEEE* **77**, 257–286 (1989).
- [61] Weiss, B. Subshifts of finite type and sofic systems. *Monatsh. Math.* **77**, 462 (1973).
- [62] Crutchfield, J.P. Semantics and thermodynamics. In *Santa Fe Institute Studies in the Sciences of Complexity*. **12**, 317–359 (1992).
- [63] Cover, T.M. & Thomas, J.A. *Elements of Information Theory*. (Wiley-Interscience, New York, second edition, 2006).
- [64] Crutchfield, J.P. Between order and chaos. *Nature Phys.* **8**, 17–24 (2012).
- [65] Gmeiner, P. Equality conditions for internal entropies of certain classical and quantum models. *arXiv:1108.5303* (2011).
- [66] Perry, N. & Binder, P.M. Finite statistical complexity for sofic systems. *Phys. Rev. E*. **60**, 459–463 (1999).
- [67] Delgado, J. & Solé, R.V. Collective-induced computation. *Phys. Rev. E*. **55**, 2338–2344 (1997).
- [68] Nerukh, D., Jensen, C.H. & Glen, R.C. Identifying and correcting non-Markov states in peptide conformational dynamics. *J. Chem. Phys.* **132**, 084104 (2010).
- [69] Nerukh, D. Non-Markov state model of peptide dynamics. *J. Mole. Liquids*. **176**, 65–70 (2012).
- [70] Kelly, D., Dillingham, M., Hudson, A. & Wiesner, K. A new method for inferring hidden Markov models from noisy time sequences. *PLoS One*. **7** e29703 (2012).
- [71] Li, C.B. & Komatsuzaki, T. Aggregated Markov model using time series of a single molecule dwell times with a minimum of excessive information. *Phys. Rev. Lett.* **111**, 058301 (2013).
- [72] Varn, D.P. & Crutchfield, J.P. Chaotic crystallography: How the physics of information reveals structural order in materials. *Current Opinion Chem. Engin.* **7**, 47–56 (2015).
- [73] Crutchfield, J.P., Riechers, P. & Ellison, C.J. Exact complexity: Spectral decomposition of intrinsic computation. *Phys. Lett. A*. **380** 998–1002 (2016).
- [74] Mahoney, J.R., Ellison, C.J. & Crutchfield, J.P. Information accessibility and cryptic processes. *J. Phys. A*. **42**, 362002 (2009).
- [75] Mahoney, J.R., Ellison, C.j., James, R.G. & Crutchfield J.P. How hidden are hidden processes? A primer on crypticity and entropy convergence. *Chaos* **21**, 037112 (2011).
- [76] Aghamohammadi, C., Mahoney, J.R. & Crutchfield, J.P. The ambiguity of simplicity in quantum and classical simulation. *Phys. Lett. A*. **381**, 1223–1227 (2017).
- [77] Suen, W.Y., Thompson, J., Garner, A.J.P., Vedral, V. & Gu, M. The classical-quantum divergence of complexity in the Ising spin chain. *arXiv:1511.05738* (2015).
- [78] Tan, R., Terno, D.R., Thompson, J., Vedral, V. & Gu, M. Towards quantifying complexity with quantum mechanics. *European Physical J. Plus*. **129**, 1–12 (2014).
- [79] Dale, H., Jennings, D. & Rudolph, T. Provable quantum advantage in randomness processing. *Nature Comm.* **6**, (2015).
- [80] De las Cuevas, G. & Cubitt, T.S. Simple universal models capture all classical spin physics. *Science*. **351**, 1180–1183 (2016).

[81] Dobson, J.F. Many-neighbored Ising chain. *J. Math. Phys.* **10**, 40–45 (1969).