Error-tolerant witnessing of divergences in classical and quantum statistical complexity

Farzad Ghafari,1,* Mile Gu,2,3,† Joseph Ho,1 Jayne Thompson,4,3 Whei Yeap Suen,3 Howard M. Wiseman,1,5,‡ and Geoff J. Pryde1,§

1Centre for Quantum Dynamics, Griffith University, Yuggera Country, Brisbane, Queensland 4111, Australia
2Nanyang Quantum Hub, School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore 639773, Republic of Singapore
3Centre for Quantum Technologies, National University of Singapore, 3 Science Drive 2, Singapore, Republic of Singapore
4Horizon Quantum Computing, 29 Media Cir., Singapore 138565
5Centre for Quantum Computation and Communication Technology (Australian Research Council), Griffith University, Brisbane, 4111, Australia

(Dated: April 29, 2022)

How much information do we need about a process’ past to faithfully simulate its future? The statistical complexity is a prominent quantifier of structure for stochastic processes. Quantum machines, however, can simulate classical stochastic processes while storing significantly less information than their optimal classical counterparts. This implies qualitative divergences between classical and quantum statistical complexity. Here, we develop error-tolerant techniques to witness such divergences, enabling us to account for the inevitable imperfections in realising quantum stochastic simulators with present-day quantum technology. We apply these tools to experimentally verify the quantum memory advantage in simulating an Ising spin chain, even when accounting for experimental distortion. This then leads us to observe a recently conjectured effect, the ambiguity of simplicity—the notion that the relative complexity of two different processes can depend on whether we model the process using classical or quantum means of information processing.

I. INTRODUCTION

A key measure of complexity for stochastic processes in time is the minimal amount of past information a model needs to generate a statistically faithful simulation of its conditional future [1, 2]. Thus a process that changes completely randomly from one moment to the next, or one that remains unchanged over all times, are both as simple as possible in this sense. No information about their past is needed to reproduce, statistically, their future, since the future does not depend on the past. By contrast, a simple alternating sequence, ...101010..., is more complex, since one needs to remember the value of the last bit to know if the next bit will be 0 or 1. This measure, formalised as the statistical complexity, has seen use in diverse settings ranging from financial analytics to many-body physics [1, 3, 4].

The concept of statistical complexity can be generalised by allow quantum machines to store the information about the past process, and to generate the future process [5]. The so-called quantum statistical complexity can exhibit drastic qualitatives from its classical counterpart [5–7]. Specifically, the memory requirements of quantum models can exhibit more favourable scaling [8, 9], such that processes that require ever growing classical memory to simulate may only require a quantum memory of bounded dimension [10]. In addition, the relative order of what is complex can differ — a parametrised family of processes can become more complex to model classically, but less complex to model quantum mechanically — a phenomena termed the ambiguity of simplicity [6]. The reduced complexity of quantum machines can have significant operational benefit, enabling quantum-enhanced rare-event sampling [11], time-series analysis [12] and more thermally efficient means of stochastic simulation [13]. More fundamentally, they illustrate the striking feature of quantum information’s capacity to reshape our perceptions of what is complex.

However, experimental verification of these quantum–classical divergences remains challenging. One of the issues is to ensure that the demonstration is robust to experimental imperfections. Notably, the execution of quantum models assumes the availability of perfect unitary operations. Any environmental noise would result in a distorted model — one whose memory costs and predictive statistics would differ from the ideal. In previous proof-of-principle realizations of quantum models, these effects were neglected when comparing the resulting quantum memory requirements of such models with optimal classical limits [14]. Thus, two issues remained: (1) the quantum models may have simulated a slightly different process that might have significantly reduced classical statistical complexity; and (2) the noise accumulation in the quantum model may induce growing memory requirements if the simulation continued to operate ad infinitum.

Our goal here is to address these issues by develop-
ing a method to explicitly characterise and account for such imperfections. We demonstrate these techniques by building an optical quantum simulator for an iconic system in statistical physics — the one-dimensional (1D) Ising spin chain [15]. We conclusively demonstrate its reduced quantum statistical complexity by simulating the Ising system for a range of temperatures. In addition, we identify regions where there are pairs of temperatures, \( T_A \) and \( T_B \), for which the classical and quantum notions of relative complexity are reversed — providing the first experimental demonstration of the ambiguity of simplicity.

II. MATERIALS AND METHODS

A. Framework

Statistical complexity quantifies the structure of a one-dimensional classical stochastic process. For ease of discussion, we will take this one dimension as a time dimension. Then statistical complexity measures the minimum amount of information a model must carry about the process’s past in order to or generate, its possible futures faithfully (i.e. with the correct statistics).

To be more precise, we imagine a dynamical system that emits discrete-valued outputs \( x_t \)—instances of random variables \( X_t \)—at discrete times \( t \in \mathbb{Z} \). The output string, \( \cdots, X_{-1}, X_0, X_1, X_2, \cdots \), is a stochastic process, described by a joint probability distribution, \( \mathcal{P} = P(\vec{X}, \vec{X}) \). Here \( \vec{X} = \cdots, X_{-1}, X_0 \) and \( \vec{X} = X_1, X_2, \cdots \), respectively represent the past and future strings at time \( t = 0 \). Any faithful model of the process must replicate this behaviour. That is, for each observed past \( \vec{x} \), the model must provide a systematic means of initializing a suitable machine \( \Xi \) in some state \( \epsilon(\vec{x}) \), such that repeated application of a systematic action \( M \) on \( \Xi \) sequentially generates \( x_1, x_2 \ldots \) governed by the conditional future \( P(\vec{x} | \vec{X} = \vec{x}) \). Here \( \epsilon(\cdot) \) is referred to as the encoding function, capturing precisely how the model encodes the past within its memory.

Given a faithful model of a stochastic process \( \mathcal{P} \), the amount of past information the model stores is given by \( S(\Xi) \), the Shannon entropy of \( \Xi \). The statistical complexity of this process then corresponds to the value of \( S(\Xi) \) when minimized over all faithful models [1, 3]. Traditionally, this minimization assumed that \( \Xi \) is classical, whereby the minimum is achieved by \( \epsilon \)-machines [1, 16]. This involves assigning each possible past \( \vec{x} \) to an appropriate memory state \( S_i = \epsilon(\vec{x}) \) for some \( i \in \{1, \ldots, N\} \), known as a causal state, such that two pasts are assigned to the same state if and only if their conditional futures statistics coincide. The \( \epsilon \)-machine then uses a classical memory \( \Xi \) that associates each memory state \( S_i \) with a different physical configuration. The statistical complexity is thus defined by \( C_{\mu} = -\sum_{i=1}^{N} p_i \log_2 p_i \), where

\[
p_i = \sum_{\vec{x}} P(\vec{X} = \vec{x}) \delta_{S_i, \epsilon(\vec{x})} \tag{1}
\]

is the steady-state probability that the model is in causal state \( S_i \). The identification of such causal states and the associated memory cost to store them have been used for stochastic analysis in diverse contexts, including many-body physics, neural spike trains, financial analytics and chaotic dynamics [1, 4, 17, 18].

Despite the fact that the stochastic process here is completely classical, there are benefits to be gained by considering quantum modelling, in which the machine \( \Xi \) is a quantum physical system. In particular, this has the potential of further reducing memory cost [5, 7, 19]. Such q-machines can encode relevant past information coherently—replacing each \( S_i \) with a quantum state \( |S_i\rangle \). The resulting memory then has entropy

\[
C_q = -\text{Tr}(\rho \log_2(\rho)), \tag{2}
\]

where \( \rho = \sum_{i}^{N} p_i |S_i\rangle \langle S_i| \). In this case, repeated application of a suitable quantum instrument on a system \( \Xi \) in state \( |S_i\rangle \) generates correct conditional future statistics, i.e. samples from \( P(\vec{x} | \vec{X} \in S_i) \), while leaving \( \Xi \) in the state \( |S_j\rangle \), for some \( j \in \{1, \ldots, N\} \), appropriate to the new past, \( \vec{x} \). In general, \( C_q \leq C_{\mu} \), and these two quantities can exhibit drastically different qualitative and scaling behaviour [6, 8, 10].

B. Simulating the Ising system

Determination of the true quantum statistical complexity requires minimization of \( C_q \) over all valid choices of \( |S_i\rangle \). This optimization is generally non-trivial, but has been achieved for the 1D Ising system [20]. This, coupled with the clear physical relevance of Ising systems, make them an ideal candidate for witnessing the differences between quantum and classical statistical complexity.

Specifically, we consider a 1D Ising system describing an infinite spatial chain of spins with nearest-neighbour interactions. At any non-zero temperature, this string must be described statistically. In this case we can replace the method for simulation over a series of discrete times (described above) with a simulation over discrete spatial sites, corresponding to scanning the system spatially (e.g. left to right) through spin locations \( n \). In this way, the “past” corresponds to all spins to the left of the current position and the “future” corresponds to all spin sites to the right [21], and we can use our technique to examine the spatial statistics as desired. This is illustrated in Fig. 1.

For the Ising system, the energy function is

\[
H(\vec{x}, \vec{x}) = -\sum_{n} (Jx_n x_{n+1} + Bx_n), \tag{3}
\]

where \( J \) is the coupling parameter, \( B \) the magnetic field, and \( x_n \in \{-1, 1\} \) is the spin at site \( n \). For each configuration, at temperature \( T \), the joint probability distribution is given by the Boltzmann distribution [15]. We use natural units for temperature (\( k_B = 1 \)) and take the coupling
for the subsequent spin can then be generated by interacting this memory with an ancilla initial set in state \( |0\rangle \) via the controlled-unitary \( U_C \), satisfying

\[
U_C|S_0\rangle|0\rangle = \sqrt{\Gamma_{00}}|0\rangle S_0 + \sqrt{\Gamma_{01}}|1\rangle S_1,
\]

\[
U_C|S_1\rangle|0\rangle = \sqrt{\Gamma_{10}}|0\rangle S_0 + \sqrt{\Gamma_{11}}|1\rangle S_1,
\]

followed by a measurement of the ancilla qubit. In general, \( |S_0\rangle \) and \( |S_1\rangle \) are mutually non-orthogonal. Thus the quantum statistical complexity \( C_q \) is less than \( C_\mu \).

C. Accounting for Imperfection

To witness the reduced quantum statistical complexity \( C_q \) experimentally, the direct method would be to synthesize the quantum circuit representing the quantum model. However, small experimental imperfections [22] mean that instead of an ideal controlled-unitary gate, a more general transformation \( \mathcal{E} \) is implemented. The situation is depicted in Fig. 2. For the ideal process, if causal state \( \rho_i = |S_i\rangle\langle S_i| \) is the input, then the output of the circuit is \( \rho_j = |S_j\rangle\langle S_j| \) when outcome \( j \in \{0,1\} \) is obtained at the measurement. To allow for non-idealities we describe this by a map \( \mathcal{M} \), that does not necessarily preserve purity.

Because one of the input qubits (the ancilla) is always prepared in a fixed state (ideally \( |0\rangle\langle 0| \)), in place of the two-qubit completely positive (CP) trace-preserving map \( \mathcal{M} \) it is more convenient to use a one-qubit instrument. This comprises two one-qubit CP trace-decreasing maps \( \mathcal{E}_0 \) and \( \mathcal{E}_1 \), acting on the memory qubit state \( \rho_i \), depending on whether the measurement outcome was \( 0 \) or \( 1 \). The result is the output state \( \rho^o(j|i) \) multiplied by the probability \( \phi^o(j|i) \) with which it occurs:

\[
\phi^o(j|i)\rho^o(j|i) = \mathcal{E}_j(\rho_i) \tag{6}
\]

For an ideal experiment, \( \phi^o(j|i) = \Gamma_{ij} \) and \( \rho^o(j|i) = \rho_j \). But experimental imperfections can lead to deviations in both of those equations. This creates two complications: (1) the quantum device will simulate a process different from the one it is aimed to simulate (the Ising chain, in our case) and (2) the output states from one step of the simulator, \( \rho^o(j|i) \), is now not, in general, equal to either of the possible input causal states. Thus, its memory cost will generally not be \( S(\rho_i) \).

Previous experiments were proof-of-principle quantum simulators, and thus mostly ignored these imperfections for purposes of characterizing \( C_q \) [23, 24]. The entropy was taken to be the entropy of the output states after one step of the process, and there was no consideration of how small errors in this might cause the simulator to ultimately diverge in future predictions. Here, we aim to build quantum simulators that illustrate that \( C_q \) and \( C_\mu \) behave very differently even when we account for such imperfections.

We begin by developing a method of doing the quantum–classical comparison fairly in the presence of imperfections. We first identify target parameters \( B, J \)
and $T$ and tune our experiment to realize a quantum $\epsilon$-machine that approaches these parameters. Next, we perform quantum process tomography [25–27] of the circuit to obtain the $\mathcal{E}_0$ and $\mathcal{E}_1$ maps. From these maps, we find the states and transition probabilities $\{\rho^m_i, \Gamma^m_{ij}\}$ ($m$ is for “maps”), which describe the two-state quantum machine most closely corresponding to these maps. That is, the approximate equality $\mathcal{E}_j(\rho^m_i) \approx \Gamma^m_{ij}\rho^m_j$ is as close as possible to an equality. Here, closeness is defined in terms of trace distance (see Appendix B) We call the $\rho^m_i$ the fixed-point states. The stationary state of our machine is then $\rho^m = \rho^m_0 \rho^m_0 + \rho^m_1 \rho^m_1$, where $\rho^m_0 = \Gamma^m_{10}/(\Gamma^m_{10} + \Gamma^m_{01})$ and $\rho^m_1 = 1 - \rho^m_0$. Since the $\Gamma_{ij}$ are theoretical functions of $T$ and $B$ (equation A2 of Appendix A), we can numerically invert the equation $\Gamma_{ij}(J = 1, B^m, T^m) = \Gamma^m_{ij}$, to find the $T^m$ and $B^m$ that our real map $\mathcal{E}$ actually implement (see Appendices B and C for details).

Having determined the fixed-point states of the maps, we then go a further step by using them as input causal states. We collect statistics from our experimental simulator at various nominal values of $T$, for fixed $B$ and $J$ (The nominal values are listed in Appendix A). The binary outputs of the simulation are found by measuring the ancilla output in the logical basis, and these statistics are used to determine $\Gamma^s_{ij}$ (where $s$ denotes a value derived from the statistical output). This yields the corresponding stationary state probabilities $\rho^s_0 = \Gamma^s_{10}/(\Gamma^s_{10} + \Gamma^s_{01}) = 1 - \rho^s_1$, and inferred values of $T^s$ and $B^s$. We also define a corresponding stationary state, $\rho^s = \rho^s_0 \rho^s_0 + \rho^s_1 \rho^s_1$, that would result from repeated applications of the map with these transition probabilities. The entropy of $\rho^s$ thus corresponds to the quantum statistical complexity of the Ising process at inferred parameter values of $T^s$ and $B^s$. A comparison of this to the classical statistical complexity for the same parameter values accounts for all the effects of experimental imperfection.

FIG. 2. Process map of the physical gate. (a)) Due to experimental imperfections, a slightly non-unitary operation $\mathcal{M}$ is implemented. The $\mathcal{M}$ can be decomposed into two single-qubit completely-positive maps $\mathcal{E}_0$ and $\mathcal{E}_1$, conditional on the 0 and 1 ancilla measurement outcomes, respectively.

FIG. 3. Experimental set-up. Memory and ancilla photonic polarization qubits are generated by spontaneous parametric down conversion. The controlled-unitary operation is based on a controlled-Z gate [28] implemented by partially-polarizing beam splitters (PPBSs) and single-qubit operations implemented by wave plates. GT stands for Glan Taylor prism, and APD for avalanche photo diode. FPC are fiber polarization controllers.

III. RESULTS

A. Experimental Setup

Here, we implement one complete cycle of the $\epsilon$-machine, comprising memory state preparation, a controlled-unitary operation, and read-out (see Fig. 3). Unentangled single-photon polarization qubits are produced by degenerate spontaneous parametric down conversion (SPDC). The source was realized using a 410 nm cw pump laser and a BiO$_3$O$_5$ (BiBO) crystal cut for type-I phase matching. We use polarization to encode logical states $|0\rangle = |H\rangle$ and $|1\rangle = |V\rangle$, where $H$ and $V$ are horizontal and vertical polarization, respectively.

The optical circuit is based around single-qubit unitary rotations implemented with wave plates and a nondeterministic linear optics controlled-Z gate using three partially-polarizing beam splitters (PPBSs) [28]. Our desired $U_C$ gate can then be decomposed to these components, operations by noting

$U_C = (I \otimes V_0)(I \otimes V_1)(I \otimes H)Z_C(I \otimes H)^{-1}(I \otimes V_1)^{-1}$

, where $V_0 \langle 0 \rangle = |S_0\rangle$, $V_1$ is the rotation in the $X-Z$ plane that rotates $|0\rangle$ to the bisector of $|S_0\rangle$ and $|S_1\rangle$, and $H$ is a Hadamard operation.

The polarization qubits are measured using wave plates and avalanche photodiodes. Quantum state and process tomography are implemented using the methods in Ref. [26]. The quantum process tomography enables us to obtain the specific maps $\mathcal{E}_0$ and $\mathcal{E}_0$ and thus obtain a full description of our experimental quantum model. Following the methodology outlined above to account for these imperfections, we find the best-fit `model' parameter values $T^m$ and $B^m$, and fixed-point states $\rho^m_i$,
then the ‘statistical’ values $T^*$, $B^*$, and $\rho^s$ from the outcomes of applying the experimental map on the $\rho^m$.

**B. Results**

In our experiment we consider an ensemble of target Ising systems that we want to simulate, represented by various nominal values of $T$, with fixed $B = 0.30$ and $J = 1$. To clarify our procedure, we first walk through how we process the data for an example member of that ensemble. The example is the Ising systems with the nominal temperature value $T = 1.50$. We prepared the apparatus, as close as we can achieve, to simulating the Ising model with $\{J = 1, B = 0.30, T = 1.50\}$. From quantum tomography, we found the states and transition probabilities $\rho^m_i$, $\Gamma^m_{ij}$. By numerically inverting the equation $\Gamma_{ij}(J = 1, B^m, T^m) = \Gamma^m_{ij}$, the values of $T^m = 1.40 \pm 0.02$ and $B^m = 0.27 \pm 0.01$ were found. Here, the errors are obtained from Poissonian photon statistics [29]. Then, we prepared $\rho^m_i$ as the input states and collected the statistics which gave $\Gamma^s_{ij}$. The revised values of $T^* = 1.46 \pm 0.03$ and $B^* = 0.28 \pm 0.01$ were obtained by inverting equations $\Gamma_{ij}(J = 1, B^*, T^*) = \Gamma^s_{ij}$. We then repeated the same process for the other nominal $T$ values in our ensemble (see Appendix A for a list of nominal values).

As expected, there is not perfect agreement between the target parameters and the experimentally obtained parameters $T^{m(s)}$ and $B^{m(s)}$. However, for the ensemble, the mean values of $B^m$ and $B^*$ were $0.29 \pm 0.02$ and $0.28 \pm 0.05$, respectively. The errors here are standard deviation of the mean of the ensemble. The closeness to the constant nominal value of $B = 0.3$ shows that the processes implemented in the lab are not far from the target processes. Thus the variation of $B$ from its nominal value of 0.3 is negligible, and we consider the behaviour of the system as the temperature is varied (see Appendices C and D). We plot $T^m$ versus $T^*$ in Fig. 4a, demonstrating that the two methods of reconstructing the process agree well. This agreement is evidence that the simulated machines are close to the nominal ones, and have fixed-point states close to satisfying $E_j(\rho^m_i) = \Gamma^s_{ij}(\rho^m_i)$. Using equation (2), we can calculate the quantum statistical complexity both for the theoretical prediction of the steady state based on the tomographically determined model of the simulator, $C_q^m = C_q(\rho^m)$, and for the statistical results from implementing the experiment, $C_q^s = C_q(\rho^s)$. We plot these values of the quantum statistical complexity at each temperature ($T^m$ and $T^*$ respectively) in Fig. 4b for the case where we are simulating a ferromagnetic chain with nominal values of $B = 0.3$ and $J = 1$. We observe that $C_q^m$ and $C_q^s$ lie close to the es-
C. Ambiguity of Simplicity

An interesting question is whether relative simplicity is an intrinsic property of the systems being modelled, not of the models. That is, how does the notion of relative simplicity survive the transition from a classical to a quantum description [6]? Consider two Ising systems, A and B, with different temperatures $T_A$ and $T_B$. If, in the classical regime, $C^A_\mu < C^B_\mu$ (which means that A is simpler than B) and, in the quantum regime, $C^A_q < C^B_q$, then there is consistency between the two classes for processes A and B. However, if the quantum model reverses its ranking compared with the classical perspective, we have the ambiguity of simplicity [6]. The basic question, “Which process is simpler?” no longer has a well-defined answer. To mathematically describe this phenomenon, we define

$$r(T_1, T_2) = \frac{C_q(T_1) - C_q(T_2)}{C_\mu(T_1) - C_\mu(T_2)} \quad (7a)$$

$$K(T_1, T_2) = \text{Sign}(r(T_1, T_2)) \times \min\{|r(T_1, T_2)|, 1/|r(T_1, T_2)|\}. \quad (7b)$$

Here $K$ is the degree of consistency. For $-1 < K < 0$, there is ambiguity according to the definition above, and for $0 < K < 1$, the models are consistent. The magnitude $|K| \in [0, 1]$ gives an indication of the degree of consistency or discrepancy. In Fig. 5, we construct a diagram that compares all pairs of processes at different temperatures $T_1$ and $T_2$. As can be seen, the notion of relative physical simplicity, capturing which system needs less memory to simulate, depends on the models used for simulation, i.e. we observe an ambiguity.

IV. DISCUSSION AND CONCLUSION

Here, we simulated an Ising system using a quantum processor that can drastically reduces memory require-ments about the system’s past, suggesting that complexity exhibits significantly different behaviour in the quantum regime. This involved devising a method to compare quantum and classical notions of statistical measures of complexity that accounts for for experimental imperfections. Moreover, we provide the first experimental witness for “ambiguity of simplicity” [6] — a peculiar feature where the relative order of what is more complex differs depending on whether we allow quantum simulators.

Our work opens a number of interesting future avenues. One observation is that our measure of memory cost is entropic, in line with the most convention in classical literature. Thus, its operational meaning is mostly reflected in the i.i.d regime, corresponding to most costs when simulating a large number of stochastic processes simultaneously. In near-term devices, an alternative single-shot measure using memory dimension may be more relevant. Recent theory and experiments work indicate a quantum advantage in simulation persists and can also scale in this single-shot regime [10, 14]. Our methods presented here could thus enable a more robust experimental demonstration of such a single-shot simulation advantage. Indeed, the ambiguity of simplicity has not been demonstrated, not within such single-shot regimes, providing further impetus. Meanwhile, there are several other related tasks where quantum devices have the potential at an advantage in memory-limited regimes. Examples in-

![Consistency graph](image-url)
clude quantum clocks, adaptive agents, and solutions of certain promise problems [30–34], all of which provide potential candidates for theoretical investigation and experimental realization.

V. ACKNOWLEDGMENTS

We thank Raj B. Patel for useful discussions and help in performing the experiment. The authors also thank Nora Tischler for discussions and comments on the manuscript. This work was supported by the National Research Foundation of Singapore Fellowship No. NRF-NRFF2016-02, the FQXi R-710-000-146-720 Grant “Are quantum agents more energetically efficient at making predictions?” from the Foundational Questions Institute and Fetzer Franklin Fund (a donor-advised fund of Silicon Valley Community Foundation) and the Australian Research Council (project no. DP160101911). H.M.W.’s contribution was funded under the Centre of Excellence Grant CE170100012 (Centre for Quantum Computation and Communication Technology). F.G. and J.H. acknowledge support by the Australian Government Research Training Program (RTP) scholarship.

Appendix A: Ising model

Different Ising systems may be specified by different $T, J$ and $B$. We choose the nominal values of $\{J = 1, B = 0.3\}$ as an example of the ferromagnetic regime, and simulate the chain for a range of different nominal temperatures

$$ T \in \{0.75, 1, 1.25, 1.5, 1.75, 2.25, 2.75, 3, 4, 5, 6, 8, 10, 12, 14\}. $$

These are used to calculate nominal values of $\Gamma_{ij}(J, B, T)$ and to realize the causal states defined in equation (3) in the main text. The transition probabilities are given by [35]:

$$\begin{align*}
\Gamma_{00} &= e^{\frac{B}{J} + \frac{J}{D}}, \\
\Gamma_{01} &= 1 - \Gamma_{00}, \\
\Gamma_{11} &= e^{-\frac{B}{J} + \frac{J}{D}}, \\
\Gamma_{10} &= 1 - \Gamma_{11},
\end{align*}$$

(A1)

where $D = \exp(JT) \cosh(BT) + \sqrt{\exp(-2JT) + \exp(2JT) \sinh^2(BT)}$.

Appendix B: Fixed-point states

In the ideal case defined in equation (5) in the main text, if we get measurement outcome $j$ with probability $\Gamma_{ij}$, then $E_j(\rho_i) = \Gamma_{ij}\rho_j$. (Here, the causal state $\rho_i = |S_i\rangle\langle S_i|$ is the input, $\rho_j = |S_j\rangle\langle S_j|$ is the output state of the circuit, and $E_0$ and $E_1$ are the experimentally-implemented maps which are characterized through quantum process tomography performed on the one-qubit process [26, 27].) However, in practice, a slightly different (but very close) output state $\rho'_{ij}(j|i)$ is obtained: it turns out that $\rho'_{ij}(j|i) \neq \Gamma_{ij}\rho_{ij}(j|i)$, motivating a theoretical question: “Given map $E$, can we find $\Gamma_{ij}$ and $\rho_i$ (for $i = 0, 1$) such that $E_j(\rho_i) = \Gamma_{ij}\rho_{ij}$, exactly, for $j = 0, 1$?” Experimental tests indicate that the answer to this question is generally “no”, but it can be close. Instead, we find the best solution for $\Gamma_{ij}$ and $\rho_i$ with $i = 0, 1$, as

$$\{\rho'_{ij}, \Gamma_{ij}^m\} = \arg \min_{\{\rho_i, \Gamma_{ij}\}} \left\{ \sum_{i,j=0,1} \|E_j(\rho_i) - \Gamma_{ij}\rho_j\| \right\},$$

(B1)

where $\arg \min_x[f(x)]$ means the value of the arguments $x$ that minimize the function $f(x)$, and $\|\ldots\|$ is the trace distance [25].

Appendix C: Deviation of implemented Ising parameters from nominal values

The implemented mean value of $B^m$, $0.29 \pm 0.02$, is consistent with the nominal value of 0.3. However, the difference between the nominal value of $T$ and the implemented value of $T$ increases when we aim for high temperatures. As an example, the implemented $T^m$ is 0.751 when the nominal value is 0.750, while $T^m$ is 3.237 when the nominal value is 4.000. The implemented values for $T$ can be seen in Fig. 4a of the main text. Also, as it can be seen in Fig. 6, for values close to 0, the implemented $T^m$ and $T^*$ versus nominal values almost follow a linear trend, while for higher temperature it starts to saturate.
It is possible to obtain an intuitive understanding of why the simulator imperfection increases slightly with increasing $T$ parameter. At low $T$, the quantum causal states are close together (hence the low statistical complexity) but also relatively close to a logical state. At higher $T$, the states are close together but closer to equal superpositions of logical states. Thus gate imperfections that affect coherence (e.g. imperfect or variable phase offsets from the apparatus) are more pronounced on average. Also, the sensitivity of the inversion to find $T^m$ varies with $T$.

Appendix D: Theoretical prediction for the statistical complexity of the real simulator

The simulator maps to an Ising system with temperature $T^m$ and magnetic field $B^m$, which vary slightly from the target values in equation (A1) and (constant) $B = 0.3$, respectively. While the temperature is the main parameter that varies between different system, the estimated $B^m$’s are not exactly constant. The mean value of $B^m$ is close to 0.30, but there is some spread and one can construct a value and an uncertainty band, $B^m \in [0.27 \pm 0.01, 0.31 \pm 0.03]$ . For each $B^m$ values, we define an upper bound and a lower bound as $B^+ = B^m + \sigma$ and $B^- = B^m - \sigma$ respectively, where $\sigma$ is the error bar calculated from Poisson distribution. The theoretically predicted quantum statistical complexity for the real simulator is given by $C_q(B \in \{B^+ , B^\} , T^m)$ for $J = 1$. The $C_q$ values corresponding to the upper and lower bound resulted in the grey bounds in Fig. 4b in the main text.

[1] Crutchfield, J. P. & Young, K. Inferring statistical complexity. Phys. Rev. Lett. 63, 105 (1989).
[2] Grassberger, P. Toward a quantitative theory of self-generated complexity. Int. J. Theor. Phys. 25, 907–938 (1986).
[3] Shalizi, C. R. & Crutchfield, J. P. Computational mechanics: Pattern and prediction, structure and simplicity. J. Stat. Phys. 104, 817–879 (2001).
[4] Crutchfield, J. P., Ellison, C. J. & Mahoney, J. R. Time’s barbed arrow: Irreversibility, crypticity, and stored information. Phys. Rev. Lett. 103, 094101 (2009).
[5] Gu, M., Wiesner, K., Rieper, E. & Vedral, V. Quantum mechanics can reduce the complexity of classical models. Nat. Commun. 3, 762 (2012).
[6] Aghamohammadi, C., Mahoney, J. R. & Crutchfield, J. P. The ambiguity of simplicity in quantum and classical simulation. Phys. Lett. A 381, 1223–1227 (2017).
[7] Binder, F. C., Thompson, J. & Gu, M. Practical unitary simulator for non-markovian complex processes. Phys. Rev. Lett. 120, 240502 (2018).
[8] Garner, A. J., Liu, Q., Thompson, J., Vedral, V. et al. Provably unbounded memory advantage in stochastic simulation using quantum mechanics. New Journal of Physics 19, 103009 (2017).
[9] Aghamohammadi, C., Mahoney, J. R. & Crutchfield, J. P. Extreme quantum advantage when simulating classical systems with long-range interaction. Sci. Rep. 7, 6735 (2017).
[10] Elliott, T. J. et al. Extreme dimensionality reduction with quantum modeling. Physical Review Letters 125, 260501 (2020).
[11] Aghamohammadi, C., Loomis, S. P., Mahoney, J. R. & Crutchfield, J. P. Extreme quantum memory advantage for rare-event sampling. Physical Review X 8, 011025 (2018).
[12] Elliott, T. J. & Gu, M. Superior memory efficiency of quantum devices for the simulation of continuous-time stochastic processes. npj Quantum Inf. 4, 18 (2018).
[13] Cabello, A., Gu, M., Gühne, O., Larsson, J.-r. & Wiesner, K. Thermodynamical cost of some interpretations of quantum theory. Phys. Rev. A 94, 052127 (2016).
[14] Ghafari, F. et al. Dimensional quantum memory advantage in the simulation of stochastic processes. Physical Review X 9, 041013 (2019).
[15] Pathria, R. Statistical Mechanics (Elsevier, 1972).
[16] Crutchfield, J. P. The calculi of emergence: computation, dynamics and induction. Physica D 75, 11–54 (1994).
[17] Crutchfield, J. P. Between order and chaos. Nat. Phys. 8, 17–24 (2012).
[18] Crutchfield, J. P. & Feldman, D. P. Regularities unseen, randomness observed: Levels of entropy convergence. Chaos 13, 25–54 (2003).
[19] Mahoney, J. R., Aghamohammadi, C. & Crutchfield, J. P. Occam’s quantum strop: Synchronizing and compressing classical cryptic processes via a quantum channel. Sci. Rep. 6, 20495 (2016).
[20] Suen, W. Y., Thompson, J., Garner, A. J. P., Vedral, V. & Gu, M. The classical-quantum divergence of complexity in modelling spin chains. Quantum 1, 25 (2017).
[21] Crutchfield, J. P. & Feldman, D. P. Statistical complexity of simple one-dimensional spin systems. Phys. Rev. E 55, R1239 (1997).
[22] Rolde, P. P., Pryde, G. J., O’Brien, J. L. & Ralph, T. C. Quantum-gate characterization in an extended hilbert space. Physical Review A 72, 032306 (2005).
[23] Palsson, M. S., Gu, M., Ho, J., Wiseman, H. M. & Pryde, G. J. Experimentally modeling stochastic processes with less memory by the use of a quantum processor. Sci. Adv. 3, 1601302 (2017).
[24] Ghafari, F. et al. Interfering trajectories in experimental quantum-enhanced stochastic simulation. Nature communications 10, 1–8 (2019).
[25] Nielsen, M. A. & Chuang, I. L. Quantum Computation And Quantum Information (Cambridge university press,
[26] White, A. G. et al. Measuring two-qubit gates. J. Opt. Soc. Am. B 24, 172–183 (2007).

[27] Bongioanni, I., Sansoni, L., Sciarrino, F., Vallone, G. & Mataloni, P. Experimental quantum process tomography of non-trace-preserving maps. Phys. Rev. A 82, 042307 (2010).

[28] Langford, N. K. et al. Demonstration of a simple entangling optical gate and its use in bell-state analysis. Phys. Rev. Lett. 95, 210504 (2005).

[29] Fox, A. M., Fox, M. et al. Quantum optics: an introduction, vol. 15 (Oxford university press, 2006).

[30] Yang, Y., Baumgärtner, L., Silva, R. & Renner, R. Accuracy enhancing protocols for quantum clocks. arXiv preprint arXiv:1905.09707 (2019).

[31] Budroni, C., Vitagliano, G. & Woods, M. P. Ticking-clock performance enhanced by nonclassical temporal correlations. Physical Review Research 3, 033051 (2021).

[32] Thompson, J., Garner, A. J., Vedral, V. & Gu, M. Using quantum theory to simplify input–output processes. npj Quantum Information 3, 1–8 (2017).

[33] Elliott, T. J., Gu, M., Garner, A. J. & Thompson, J. Quantum adaptive agents with efficient long-term memories. arXiv preprint arXiv:2108.10876 (2021).

[34] Tian, Y., Feng, T., Luo, M., Zheng, S. & Zhou, X. Experimental demonstration of quantum finite automaton. npj Quantum Information 5, 1–5 (2019).

[35] Feldman, D. P. Computational mechanics of classical spin systems. Ph.D. thesis, University of California, Davis (1998).