

Dimensional Quantum Memory Advantage in the Simulation of Stochastic ProcessesFarzad Ghafari¹,[✉] Nora Tischler,^{1,*} Jayne Thompson,^{2,†} Mile Gu,^{2,3,‡} Lynden K. Shalm,⁴ Varun B. Verma,⁴ Sae Woo Nam,⁴ Raj B. Patel,^{1,5} Howard M. Wiseman,⁶ and Geoff J. Pryde¹¹*Centre for Quantum Dynamics, Griffith University, Brisbane, 4111, Australia*²*Centre for Quantum Technologies, National University of Singapore,**3 Science Drive 2, Singapore, Republic of Singapore*³*School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore 639673, Republic of Singapore*⁴*National Institute of Standards and Technology, 325 Broadway, Boulder, Colorado 80305, USA*⁵*Clarendon Laboratory, Department of Physics, Oxford University, Parks Road OX1 3PU Oxford, United Kingdom*⁶*Centre for Quantum Computation and Communication Technology (Australian Research Council), Centre for Quantum Dynamics, Griffith University, Brisbane, 4111, Australia*

(Received 15 March 2019; revised manuscript received 18 September 2019; published 17 October 2019)

Stochastic processes underlie a vast range of natural and social phenomena. Some processes such as atomic decay feature intrinsic randomness, whereas other complex processes, e.g., traffic congestion, are effectively probabilistic because we cannot track all relevant variables. To simulate a stochastic system's future behavior, information about its past must be stored, and thus memory is a key resource. Quantum information processing promises a memory advantage for stochastic simulation. Here, we report the first experimental demonstration that a quantum stochastic simulator can encode the required information in fewer dimensions than any classical simulator, thereby achieving a quantum advantage in minimal memory requirements using an individual simulator. This advantage is in contrast to recent proof-of-concept experiments, where the memory saving would only become accessible in the limit of a large number of parallel simulations. In those examples, the minimal memory registers of *individual* quantum simulators had the same dimensionality as their classical counterparts. Our photonic experiment thus establishes the potential of new, practical resource savings in the simulation of complex systems.

DOI: [10.1103/PhysRevX.9.041013](https://doi.org/10.1103/PhysRevX.9.041013)

Subject Areas: Optics, Quantum Physics, Quantum Information

I. INTRODUCTION

Stochastic processes are ubiquitous in science and technology [1,2]. Quantum information reduces the required memory storage for simulating these processes [3–16]—a newly identified advantage [3] that complements other quantum information technological enhancements. Recent first experiments confirmed the potential of this advantage [9,10,16]. However, in all past experiments, the memory enhancement was with respect to an entropic information measure. A large number of parallel simulators

would be required to exploit this entropic benefit in order to encode the required information in a memory with lower dimension. Because of the unavailability of a large number of parallel simulators, the ultimate memory advantage, a dimensional advantage, has never been achieved. Also, verifying the entropic advantage requires quantum state tomography, with state-reconstruction difficulty scaling exponentially with the problem size.

Here, we realize the first experimental demonstration of a dimensional memory advantage for simulating stochastic processes. We achieve the dimensional advantage directly with an individual quantum simulator, which uses a memory register with fewer dimensions than any classical counterpart. Our quantum encoding is achievable with any number of simulators, rather than requiring an asymptotically large array of simulators [4,17]. This encoding realizes the quantum advantage for stochastic simulation in its fullest sense. Moreover, characterizing the dimensional advantage is relatively straightforward, compared to other measures of memory, which rely on

*n.tischler@griffith.edu.au†cqttjed@nus.edu.sg‡gumile@ntu.edu.sg

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/). Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

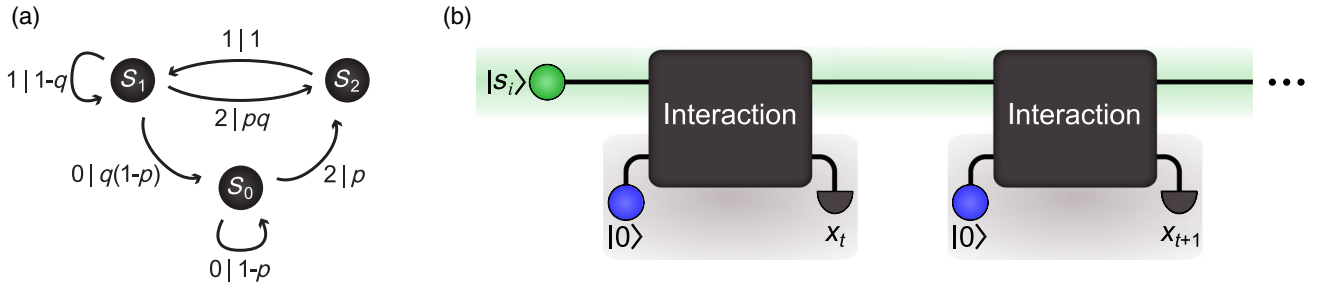


FIG. 1. The stochastic process and its simulation. (a) The simplest classical simulator for the process can be represented by a directed graph. Here, the three nodes S_1 , S_2 , and S_3 represent internal states of the machine. Each edge between two nodes, S_j and S_k , labeled $r|T_{jk}$, represents the probability T_{jk} that the machine emits output r and transitions from state S_j to state S_k . To generate correct conditional futures, one first initializes the machine in state S_i depending on the value of the last output (0, 1, or 2). Iterating the machine can then generate correct conditional future statistics. Mathematically, the probabilities T_{jk} make up the transition matrix, whose eigenvalues form the probability distribution of the causal states, called the stationary distribution $\{p_i\}_{i=0,1,2}$. (b) Dynamically, a simulator can be considered to consist of a memory (upper arm) that transfers information between time steps and sequentially interacts with a blank tape (lower arm) at each time step.

determining the information entropy of a register. Verifying information entropy requires tomographic reconstruction, which is known to be an extremely hard task for high-dimensional states.

We investigate a specific stochastic process, while noting that the advantage is theoretically known to hold for a range of other simulation tasks [14]. The process we simulate here can be understood as the output of a biased oracular coin [15] [see Fig. 1(a)]. The process involves a (possibly biased) coin in a box. At each step, the box is perturbed, such that the coin flips with some probability (p if it is heads, q if it is tails). The box outputs a 1 whenever the coin is tails. However, whenever the coin is heads, the box is able to look up its output at the subsequent future time step. If the next time step is 1, it outputs a 2; otherwise, it outputs a 0. Because of this dependence on future time steps, the classical memory for simulation is markedly increased: Instead of needing a single coin (1 bit), a three-level system is provably required [15]. This requirement does not apply to quantum simulators, which we implement using only a single qubit as the memory.

II. FRAMEWORK

Generally, as illustrated in Fig. 1(b), a quantum simulator of a stochastic process, henceforth simply referred to as a quantum simulator, accepts a memory system and an ancilla system as inputs to an interaction [3,4,14] for each simulation step. Of the two, only the memory system contains information about the past, while the ancilla system carries no information. The interaction produces an entangled state of the output memory system and a second system. Measurement of the latter provides the output of the stochastic process and collapses the memory system to the appropriate quantum state for the next

simulation step. The memory cost of such a simulator is then characterized by the amount of resources necessary to carry information between time steps in order to generate correct future statistics.

For the stochastic process of Fig. 1(a), it is known that the provably optimal simulator classifies the set of all possible pasts into three different states called causal states [3,18]. To this end, the classical processor must have three distinguishable states, $\{S_i\}_{i=0,1,2}$, as its memory. By contrast, a quantum simulator can encode the three causal states as three pure quantum states $\{|S_i\rangle\}_{i=0,1,2}$ that are not mutually orthogonal and fit within a single qubit (as described in the Appendix A).

We implement our simulator in a photonic quantum information processor, while noting that it could also be realized in other platforms in which quantum logic gates can be implemented. The memory qubit containing relevant past information is encoded in the polarization degree of freedom of a single photon, as depicted in Fig. 2(a). To simulate a time step, it is then required to take this memory as input, deliver one of three possible outputs 0,1,2 according to expected conditional statistics, and subsequently leave the memory in the correct state for the next time step. In our experiment, this process involves interacting the memory qubit with a qutrit space of three spatial modes (playing the role of the tape), followed by a controlled-NOT (C-NOT) [19,20] and a controlled-rotation (C-rotation) gate, as detailed in Fig. 2(c). The path measurement of this photon corresponds to measuring the qutrit in the logical basis, which provides the classical output (0, 1, or 2) of that step of the stochastic process. This collapses the output memory qubit, encoded again in the polarization state of a photon, to the correct conditional state, which can be characterized by quantum state tomography.

III. EXPERIMENT

We overcome constraints in the nondeterministic photonic implementation of consecutive quantum gates by introducing a nondestructive measurement realized by an additional C-NOT gate [21,22] and a corresponding ancilla photon. The photons are generated via spontaneous parametric downconversion (SPDC), and fourfold coincidences (three photons for the experiment and one “spare” photon to herald the presence of its pair) are detected using superconducting nanowire single-photon detectors (SNSPDs [23]) and coincidence logic modules. The detailed experimental setup is shown in Fig. 3, and additional details are in Appendix B.

The first goal of the experiment is to verify that the quantum simulator is performing the intended simulation. For this to be true, two criteria must be fulfilled: (i) After initialization in each of the three possible causal states, the conditional output statistics, obtained through the qutrit measurement, should match the transition probabilities that determine the stochastic process [see Fig. 1(a)].

(ii) Conditioned on the qutrit measurement outcome, the correct memory state should be produced to allow the possibility of further simulation steps.

To check the first criterion, we prepare each of the three causal states, whose definitions in terms of p and q are provided in Appendix A. For each input causal state, there is a probability distribution over the three possible outputs of the stochastic process. Comparing the measured distributions with the theoretical ones, we consistently obtain (classical) fidelities [24] above 0.993. For the second criterion, the collapsed output memory state is reconstructed by quantum state tomography, given each of the input causal states. The (quantum) fidelities of our experimental stationary states (see Appendix C) with the ideal stationary states are all above 0.991.

The second goal of the experiment is to demonstrate the quantum advantage in memory requirements. Most pertinent to this work is the dimensional memory advantage, which has not been demonstrated before. In this scenario, the memory size, in bits or qubits, is

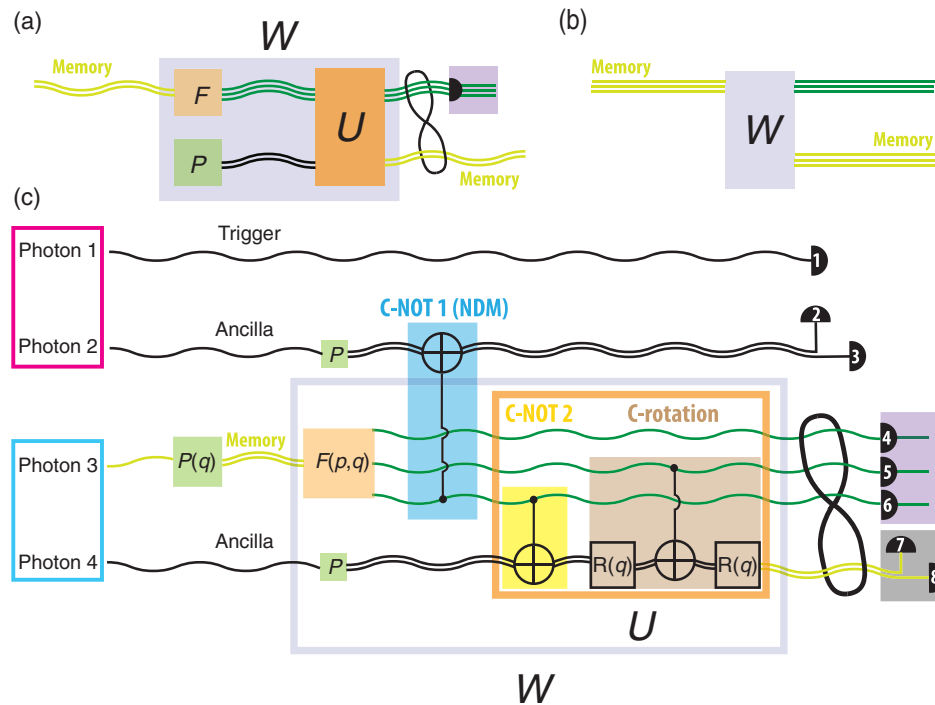


FIG. 2. Conceptual diagram of a simulation step. (a) The quantum processor (W) accepts a memory qubit and uses an ancilla. (We use wavy lines to denote quantum objects, with the number of lines in parallel indicating the dimensionality.) In the processor, the qubit undergoes a fan-out operation F to a qutrit space. The ancilla contains no information, and its preparation P is fixed. Then, a unitary operation U acts on the qutrit and ancilla, outputting an entangled state of the memory qubit and a qutrit. A projective measurement of the qutrit provides the output of the simulation step and collapses the memory qubit to the appropriate state for the next step. (b) The classical simulator requires a three-dimensional memory system. The irreversible operation W acts on the memory system to generate the classical output and the next memory state. (c) The experimental realization of the circuit in diagram (a) using linear optics gates requires an ancilla qubit (Photon 2) and its herald (Photon 1). Following the fan-out operation $F(p, q)$ on the memory qubit, we implement a gate, C-NOT 1, which performs a nondestructive measurement (NDM). Then, the unitary operation U is performed by an additional two gates, C-NOT 2 and C-rotation. The preparation $P(q)$ of the memory system (Photon 3), the fan-out operation $F(p, q)$, and the single qubit rotation $R(q)$ depend on the stochastic process parameters p and q as indicated.

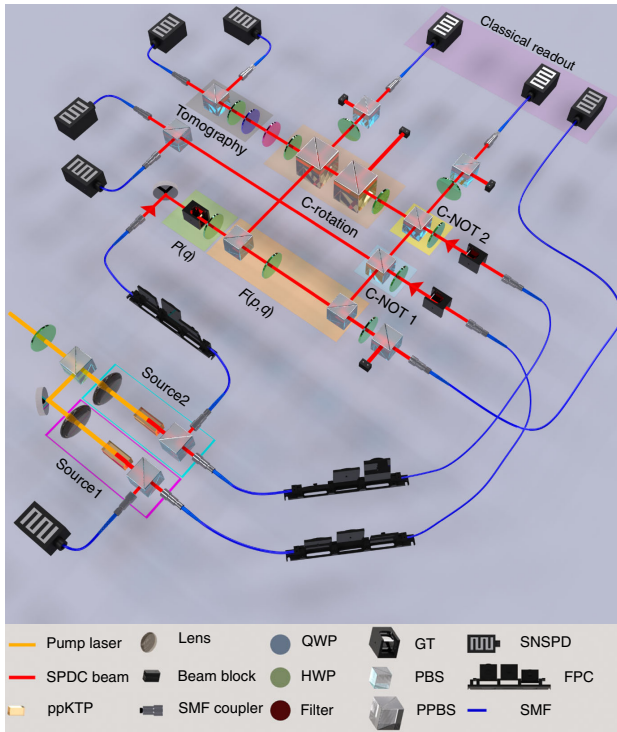


FIG. 3. Experimental setup. Single photons are generated from SPDC events. The herald photon from Source 1 is sent straight to a heralding detector. The polarization of the memory system is used to encode the relevant causal state in a qubit, using a half-wave plate (HWP). Ancillas are prepared in a fixed polarization using HWPs. To implement the fanning out from the memory qubit to a qutrit, a HWP and polarizing beam splitters (PBSs) are used. Each of the C-NOT 1 and C-NOT 2 gates is implemented using a HWP and a PBS. The C-rotation gate is realized via HWPs and partially polarizing beam splitters (PPBSs). In order to vary the relative delay between the single-photon wave packets, an automated translation stage is used to move one of the couplers. Classical readout is performed via projective measurements on the path modes of the qutrit, which collapses the memory state to the appropriate causal state. To verify the memory qubit, its state is reconstructed via quantum state tomography. A telecom bandpass filter is used in the tomography arm in order to spectrally filter the SPDC photons and maximize the visibility of the quantum interference. P stands for state preparation, SMF for single-mode fiber, QWP for quarter-wave plate, GT for Glan-Taylor prism, and FPC for fiber polarization controller. For more details, see Appendix B.

measured by the max-entropy, which is simply $\log_2 d$, where d is the dimensionality of the memory system [4,17]. We denote this max-entropy memory size as D_μ for the best classical simulator and D_Q for our quantum simulator. Since the information about the past is encoded in the polarization of a single photon, both at the beginning and at the end of the simulated step, the memory system that connects steps is obviously confined to a qubit space. In contrast to this two-level quantum system, the optimal classical simulator requires a three-level system [15].

Thus, there is a clear dimensional quantum advantage in memory. This advantage enables the first quantum simulators that use less physical resources (qubits compared to bits) when performing individual simulations. In complexity science, the max-entropy of the model, D_μ , which captures the minimal dimensions required to store the past for future simulation, is known as topological state complexity [25,26], and we provide the first experimental demonstration that the quantum analogue D_Q can be reduced.

For completeness, we also investigate the resource advantage one could obtain when multiple simulations are run in parallel. Here, the required memory is no longer determined by the dimensionality of the memory system alone. In the limit of a very large number (N) of parallel simulations [the independent and identically distributed (i.i.d.) case [4,17]], the minimum required memory to replicate the process faithfully is given by NC , where C is called the statistical complexity [25]. The classical statistical complexity [25] C_μ is the Shannon entropy of the stationary distribution over causal states, while the quantum statistical complexity [3] C_Q is the von Neumann entropy of the quantum stationary state (see Appendix C for mathematical definitions).

Figure 4(a) illustrates the theoretically expected statistical complexities C_μ and C_Q for all possible values of p and q , showing the potential for a significant asymptotic quantum advantage over a large region of the parameter space. We perform the simulation for sets of (p, q) values along several cross sections. The experimental values of C_Q , shown in Figs. 4(b)–4(e), are determined from the density matrices of the output memory system and the transition probabilities (see Appendix C). The slight deviations of the experimental data compared to the theoretical curves arise from experimental imperfections such as reduced qubit purity from imperfect nonclassical interference, small imperfections and setting errors in polarization-dependent elements, and a minor imbalance in detector efficiencies. These results nevertheless demonstrate a substantial quantum advantage in the required memory for simulation in the i.i.d. case (i.e., asymptotic advantage).

Thus, our quantum simulator has an advantage over its classical counterpart both for the individual simulator and the i.i.d. scenario. Remarkably, we even simulate processes, marked by the shaded regions in Figs. 4(b)–4(e), where the classical statistical complexity C_μ exceeds one bit. In these cases, we have a gap between both quantum measures and both classical measures: $C_Q < D_Q < C_\mu < D_\mu$. (Note that the max-entropies D_Q and D_μ always form upper bounds on their i.i.d. variants C_Q and C_μ , but small values of C_Q and C_μ do not guarantee any bounds on D_Q and D_μ .)

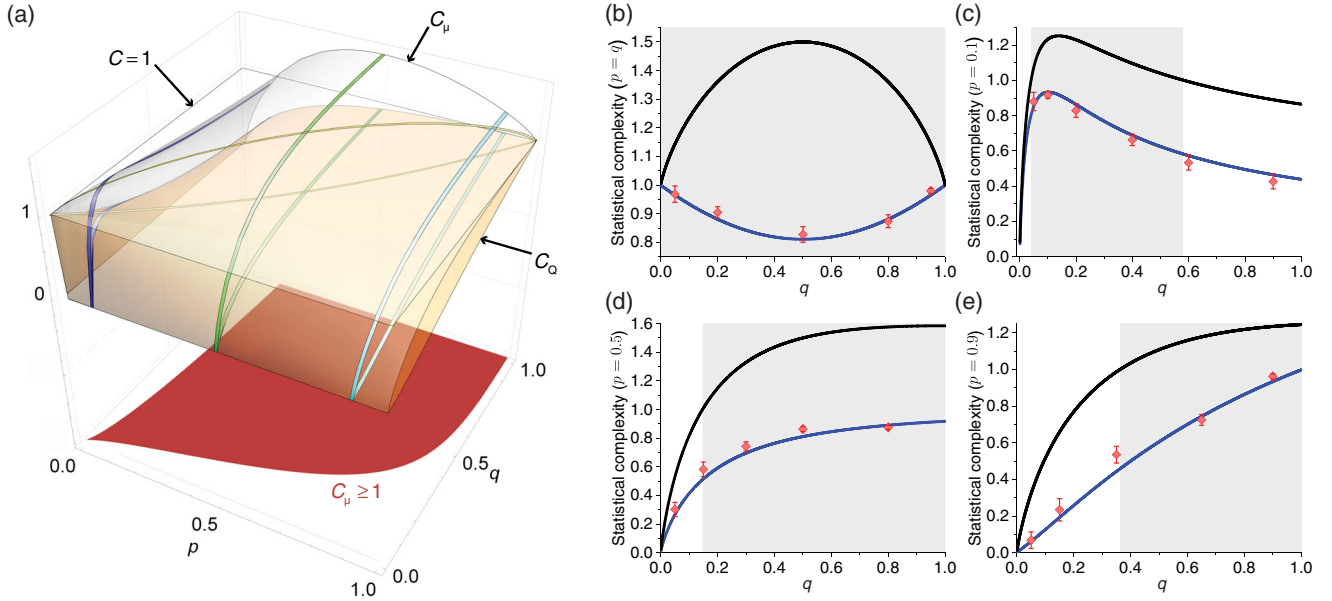


FIG. 4. Statistical complexity of the classical and quantum simulators. (a) The theoretically calculated statistical complexity. The pale grey surface depicts C_μ , while the pale orange surface shows C_Q . The transparent plane marks the value $C = 1$. The yellow, purple, green, and cyan cuts illustrate specific cross sections, which are experimentally probed and shown in panels (b), (c), (d), and (e), respectively. The red projection on the floor illustrates the (p, q) values for which $C_\mu \geq 1$. (b)–(e) The quantum simulator is used to investigate several sets of processes with different values of p and q . The entropy of the reconstructed stationary states (see Appendix C) determines the quantum statistical complexity (red dots). The black and blue curves represent the theoretical C_μ and C_Q , respectively. The plots demonstrate a considerable memory advantage for the i.i.d. case. Furthermore, the grey shaded areas mark processes where the complexity C_μ of the classical simulator exceeds one bit, while the quantum simulation runs with only one memory qubit. Uncertainties are estimated from the Poissonian distribution of photon counts.

IV. DISCUSSION

In conclusion, we have shown that quantum information processing enables the simulation of a stochastic process with a memory that is smaller in terms of its dimensionality (the number of orthogonal states it can support), as compared to any classical counterpart. This demonstrated decrease in the dimensionality of the memory system establishes a new type of memory saving—namely, a dimensional memory advantage, complementing prior works that established entropic reductions in memory [9,10,16]. Whereas the latter quantified quantum memory advantages in the asymptotic limit of simulating many stochastic processes simultaneously, the dimensionality reduction indicates an immediate reduction in the physical resources used (number of qubits compared to bits) when simulating stochastic processes at the level of an individual simulator. Here, we were able to use a single qubit to generate the same predictions that would otherwise require at least two classical bits.

One crucial point to make is that dimensional and entropic advantages are distinct. Having a dimensional memory advantage for individual simulators guarantees a memory advantage in the entropic setting but not necessarily vice versa. Indeed, there exist processes for which it is possible to design quantum models with reduced entropic memory but not reduced dimensionality. Meanwhile, recent works illustrated that the quantum

simulator optimized to reduce memory dimensionality is not necessarily the same as the one that is optimized for memory entropy—a uniquely quantum phenomenon that has no classical counterpart [27,28]. Indeed, the present experiment allows us to study both the dimensionality and the statistical complexity of the memory system and provides a means to explore this phenomenon.

On a more practical level, dimensional reduction also has notable operational benefits. Since entropy is an i.i.d. quantity, it is not directly observable in a single run of the experiment. As such, verification of entropic advantage requires the use of state tomography. For more complex processes that entail high-dimensional memory systems, the quantum state tomography would require increased resources (such as photons, modes, and detectors) and could become prohibitively time-consuming. In contrast, verifying a dimensionality advantage remains relatively straightforward because it is based on counting dimensions of a Hilbert space rather than characterizing quantum states. This simplicity may become especially pertinent in the context of conducting experimental tests to verify scaling advantages of quantum models. Such scaling advantages have recently been predicted to also be possible at the level of individual simulators and to even admit scenarios with an unbounded dimensional advantage, where a classical simulator would require infinite memory

but a quantum simulator would only require a finite-dimensional memory [15,29].

ACKNOWLEDGMENTS

The authors would like to thank Thomas Elliott and Chengran Yang for useful discussions. F. G. acknowledges support by the Australian Government Research Training Program (RTP). N. T. was supported by the Griffith University Postdoctoral Fellowship Scheme. This research was funded by the Australian Research Council (Projects No. DP160101911 and No. CE170100012), Singapore National Research Foundation Fellowship NRF-NRFF2016-02, the NRF-ANR grant NRF2017-NRF-ANR004 VanQuTe, the John Templeton Foundation (Grant No. 53914), and the FQXi grant ‘‘Are quantum agents more energetically efficient at making predictions?’’ We acknowledge the traditional owners of the land on which this work was undertaken at Griffith University, the Yuggera people.

APPENDIX A: STOCHASTIC PROCESSES

A stochastic process [1,2] evolving in discrete time is a collection of random variables $\{\dots, X_{t-1}, X_t, X_{t+1}, X_{t+2}, \dots\}$, where the previously observed variables $\{\dots, X_{t-1}, X_t\}$ are considered the past of the process, i.e., the list of past outputs. A faithful simulator is one that correctly generates the process’s future statistical behavior based on a given configuration of its past. The memory system of the simulator must store sufficient information about the past configuration to enable this faithful simulation [18]. Then, a processor acts on the memory, generating a new classical output X_{t+1} and updating the memory to be ready for the next step.

For optimal simulation of the process that we study here [15], the most recent output X_t is sufficient for determining the memory state for step $t + 1$ [25]. The possible memory states are called causal states [18,25], and there are three of them for this process. The classical causal states are perfectly distinguishable states, $\{S_i\}_{i=0,1,2}$. The quantum causal states $\{|S'_i\rangle\}_{i=0,1,2}$ can be similarly defined as

$$\begin{aligned} |S'_0\rangle &= \sqrt{1-p}|0\rangle + \sqrt{p}|2\rangle, \\ |S'_1\rangle &= \sqrt{q(1-p)}|0\rangle + \sqrt{1-q}|1\rangle + \sqrt{pq}|2\rangle, \\ |S'_2\rangle &= |1\rangle. \end{aligned} \tag{A1}$$

However, by choosing a different basis, these states can be mapped to a single qubit space [15,30]:

$$\begin{aligned} |S_0\rangle &= |0\rangle, \\ |S_1\rangle &= \sqrt{q}|0\rangle + \sqrt{1-q}|1\rangle, \\ |S_2\rangle &= |1\rangle, \end{aligned} \tag{A2}$$

where $|0\rangle, |1\rangle$ form an orthogonal basis.

APPENDIX B: EXPERIMENTAL DETAILS

Four photons are generated via SPDC, as shown in Fig. 3. To generate the photons, two SPDC sources are realized using a 775-nm Ti-sapphire picosecond-pulse-length pump laser and ppKTP (46.20- μ m poling period) crystals cut for type-II collinear degenerate phase matching [31,32]. The photons are not entangled in polarization. The crystal temperature is controlled at 25 °C by a temperature controller. The bandpass filter is centered at 1550 nm and has a full width at half maximum of 8.8 nm.

To run the simulator, the causal states in Eq. (A2) are encoded in the polarization degree of freedom of a single photon acting as the memory system. We use polarization modes such that $|0\rangle = |H\rangle$ and $|1\rangle = |V\rangle$, where H and V are horizontal and vertical polarizations, respectively.

The fan-out transformation implements the basis change from Eq. (A2) to Eq. (A1), so the three paths correspond to orthogonal states $|0\rangle, |1\rangle$, and $|2\rangle$. The experimental setup contains nondeterministic two-qubit gates. The C-NOT gates 1 and 2 are realized with a HWP, a PBS, and postselective detection. This simplified version (compared to a universal photonic C-NOT gate [19]) provides an increased success probability and works correctly since the photons in the two input spatial modes always have a fixed polarization. The C-rotation gate is comprised of two single-qubit rotation gates $R(q)$ and a two-qubit controlled-Z gate. This controlled-Z gate is based on the scheme in Ref. [20], which uses three PPBSs. However, we only require two because of the fixed polarization in one of the input spatial modes. Fourfold coincidences are detected in a 5-ns coincidence window, using SNSPDs and fast-counting electronics.

The detection channels have slightly different efficiencies, which may affect the probabilities determined from the various coincidence detection combinations and thus the inferred transition probabilities. The possible fourfold detection combinations are formed by coincidence detections between detectors from each of the following four sets [see Fig. 2(c)]: $\{1\}$, $\{2, 3\}$, $\{4, 5, 6\}$, and $\{7, 8\}$. Therefore, the detectors within each set should ideally have the same efficiencies. In the experiment, the detectors are installed in such a way as to match this criterion as closely as possible.

APPENDIX C: STATISTICAL COMPLEXITY

The statistical complexity [18,25,33] is the minimal memory a model needs to generate future statistics correctly using only information from past observations. The classical statistical complexity is

$$C_\mu = -\sum_i p_i \log_2 p_i, \tag{C1}$$

where p_i is the probability of each causal state in the stationary stochastic process, i.e., in the limit of a long

evolution. The quantum statistical complexity is defined as [3]

$$C_Q = -\text{Tr}(\rho \log_2(\rho)), \quad (\text{C2})$$

where $\rho = \sum_i p_i |S_i\rangle\langle S_i|$ is the quantum stationary state.

Our simulator implements the provably optimal model, the so-called quantum epsilon machine [3,15,25]. Therefore, we can measure C_Q by inputting the causal states described in Eq. (A2) for a given set of p and q values. The stationary state ρ is calculated as

$$\begin{aligned} \rho = & d_0 \sum_{i=0}^2 T_{0i} S_{\text{pol}|S_0} + d_1 \sum_{i=0}^2 T_{1i} S_{\text{pol}|S_1} \\ & + d_2 \sum_{i=0}^2 T_{2i} S_{\text{pol}|S_2}, \end{aligned} \quad (\text{C3})$$

where $\{d_i\}_{i=0,1,2}$ are the eigenvalues of the experimentally measured transition matrix:

$$T = \begin{pmatrix} T_{00} & T_{10} & T_{20} \\ T_{01} & T_{11} & T_{21} \\ T_{02} & T_{12} & T_{22} \end{pmatrix}. \quad (\text{C4})$$

Here, T_{ij} is the probability of classical output j when the input causal state is $|S_i\rangle$. Moreover, $S_{\text{pol}|S_i}$ is the reconstructed polarization state of the output memory system when the input causal state is $|S_i\rangle$.

[1] K. Jacobs, *Stochastic Processes for Physicists: Understanding Noisy Systems* (Cambridge University Press, Cambridge, England, 2010).

[2] J. Duan, *An Introduction to Stochastic Dynamics* (Cambridge University Press, Cambridge, England, 2015).

[3] M. Gu, K. Wiesner, E. Rieper, and V. Vedral, *Quantum Mechanics Can Reduce the Complexity of Classical Models*, *Nat. Commun.* **3**, 762 (2012).

[4] J. R. Mahoney, C. Aghamohammadi, and J. P. Crutchfield, *Occam's Quantum Strop: Synchronizing and Compressing Classical Cryptic Processes via a Quantum Channel*, *Sci. Rep.* **6**, 20495 (2016).

[5] P. M. Riechers, J. R. Mahoney, C. Aghamohammadi, and J. P. Crutchfield, *Minimized State Complexity of Quantum-Encoded Cryptic Processes*, *Phys. Rev. A* **93**, 052317 (2016).

[6] A. Cabello, M. Gu, O. Gühne, J.-Å. Larsson, and K. Wiesner, *Thermodynamical Cost of Some Interpretations of Quantum Theory*, *Phys. Rev. A* **94**, 052127 (2016).

[7] J. Thompson, A. J. P. Garner, V. Vedral, and M. Gu, *Using Quantum Theory to Simplify Input-Output Processes*, *npj Quantum Inf.* **3**, 6 (2017).

[8] W. Y. Suen, J. Thompson, A. J. P. Garner, V. Vedral, and M. Gu, *The Classical-Quantum Divergence of Complexity in Modelling Spin Chains*, *Quantum* **1**, 25 (2017).

[9] M. S. Palsson, M. Gu, J. Ho, H. M. Wiseman, and G. J. Pryde, *Experimentally Modeling Stochastic Processes with Less Memory by the Use of a Quantum Processor*, *Sci. Adv.* **3**, e1601302 (2017).

[10] F. Ghafari, M. Gu, J. Ho, J. Thompson, W. Y. Suen, H. M. Wiseman, and G. J. Pryde, *Observing the Ambiguity of Simplicity via Quantum Simulations of an Ising Spin Chain*, *arXiv:1711.03661*.

[11] T. J. Elliott and M. Gu, *Superior Memory Efficiency of Quantum Devices for the Simulation of Continuous-Time Stochastic Processes*, *npj Quantum Inf.* **4**, 18 (2018).

[12] C. Aghamohammadi, S. P. Loomis, J. R. Mahoney, and J. P. Crutchfield, *Extreme Quantum Memory Advantage for Rare-Event Sampling*, *Phys. Rev. X* **8**, 011025 (2018).

[13] A. Cabello, M. Gu, O. Gühne, and Z.-P. Xu, *Optimal Classical Simulation of State-Independent Quantum Contextuality*, *Phys. Rev. Lett.* **120**, 130401 (2018).

[14] F. C. Binder, J. Thompson, and M. Gu, *Practical Unitary Simulator for Non-Markovian Complex Processes*, *Phys. Rev. Lett.* **120**, 240502 (2018).

[15] J. Thompson, A. J. P. Garner, J. R. Mahoney, J. P. Crutchfield, V. Vedral, and M. Gu, *Causal Asymmetry in a Quantum World*, *Phys. Rev. X* **8**, 031013 (2018).

[16] F. Ghafari, N. Tischler, C. Di Franco, J. Thompson, M. Gu, and G. J. Pryde, *Interfering Trajectories in Experimental Quantum-Enhanced Stochastic Simulation*, *Nat. Commun.* **10**, 1630 (2019).

[17] B. Schumacher, *Quantum Coding*, *Phys. Rev. A* **51**, 2738 (1995).

[18] C. R. Shalizi and J. P. Crutchfield, *Computational Mechanics: Pattern and Prediction, Structure and Simplicity*, *J. Stat. Phys.* **104**, 817 (2001).

[19] J. L. O'Brien, G. J. Pryde, A. G. White, T. C. Ralph, and D. Branning, *Demonstration of an All-Optical Quantum Controlled-NOT Gate*, *Nature (London)* **426**, 264 (2003).

[20] N. K. Langford, T. J. Weinhold, R. Prevedel, K. J. Resch, A. Gilchrist, J. L. O'Brien, G. J. Pryde, and A. G. White, *Demonstration of a Simple Entangling Optical Gate and Its Use in Bell-State Analysis*, *Phys. Rev. Lett.* **95**, 210504 (2005).

[21] G. J. Pryde, J. L. O'Brien, A. G. White, S. D. Bartlett, and T. C. Ralph, *Measuring a Photonic Qubit without Destroying It*, *Phys. Rev. Lett.* **92**, 190402 (2004).

[22] T. C. Ralph, S. D. Bartlett, J. L. O'Brien, G. J. Pryde, and H. M. Wiseman, *Quantum Nondemolition Measurements for Quantum Information*, *Phys. Rev. A* **73**, 012113 (2006).

[23] B. Baek, A. E. Lita, V. Verma, and S. W. Nam, *Superconducting $a\text{-W}_x\text{Si}_{1-x}$ Nanowire Single-Photon Detector with Saturated Internal Quantum Efficiency from Visible to 1850 nm*, *Appl. Phys. Lett.* **98**, 251105 (2011).

[24] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, England, 2010).

[25] J. P. Crutchfield and K. Young, *Inferring Statistical Complexity*, *Phys. Rev. Lett.* **63**, 105 (1989).

- [26] N. Barnett and J. P. Crutchfield, *Computational Mechanics of Input–Output Processes: Structured Transformations and the ϵ -Transducer*, *J. Stat. Phys.* **161**, 404 (2015).
- [27] S. Loomis and J. P. Crutchfield, *Strong and Weak Optimizations in Classical and Quantum Models of Stochastic Processes*, [arXiv:1808.08639](https://arxiv.org/abs/1808.08639).
- [28] Q. Liu, T. J. Elliott, F. C. Binder, C. Di Franco, and M. Gu, *Optimal Stochastic Modeling with Unitary Quantum Dynamics*, *Phys. Rev. A* **99**, 062110 (2019).
- [29] T. J. Elliott, C. Yang, F. C. Binder, A. J. P. Garner, J. Thompson, and M. Gu, *Extreme Dimensional Compression with Quantum Modelling*, [arXiv:1909.02817](https://arxiv.org/abs/1909.02817).
- [30] A. Pepper, N. Tischler, and G. J. Pryde, *Experimental Realization of a Quantum Autoencoder: The Compression of Qutrits via Machine Learning*, *Phys. Rev. Lett.* **122**, 060501 (2019).
- [31] M. M. Weston, H. M. Chrzanowski, S. Wollmann, A. Boston, J. Ho, L. K. Shalm, V. B. Verma, M. S. Allman, S. W. Nam, R. B. Patel, S. Slussarenko, and G. J. Pryde, *Efficient and Pure Femtosecond-Pulse-Length Source of Polarization-Entangled Photons*, *Opt. Express* **24**, 10869 (2016).
- [32] N. Tischler, F. Ghafari, T. J. Baker, S. Slussarenko, R. B. Patel, M. M. Weston, S. Wollmann, L. K. Shalm, V. B. Verma, S. W. Nam, H. C. Nguyen, H. M. Wiseman, and G. J. Pryde, *Conclusive Experimental Demonstration of One-Way Einstein-Podolsky-Rosen Steering*, *Phys. Rev. Lett.* **121**, 100401 (2018).
- [33] J. P. Crutchfield, C. J. Ellison, and J. R. Mahoney, *Time’s Barbed Arrow: Irreversibility, Crypticity, and Stored Information*, *Phys. Rev. Lett.* **103**, 094101 (2009).