

ARTICLE

<https://doi.org/10.1038/s41467-019-08951-2>

OPEN

# Interfering trajectories in experimental quantum-enhanced stochastic simulation

Farzad Ghafari<sup>1</sup>, Nora Tischler<sup>1</sup>, Carlo Di Franco<sup>2,3</sup>, Jayne Thompson<sup>4</sup>, Mile Gu<sup>2,3,4</sup> & Geoff J. Pryde<sup>1</sup>

Simulations of stochastic processes play an important role in the quantitative sciences, enabling the characterisation of complex systems. Recent work has established a quantum advantage in stochastic simulation, leading to quantum devices that execute a simulation using less memory than possible by classical means. To realise this advantage it is essential that the memory register remains coherent, and coherently interacts with the processor, allowing the simulator to operate over many time steps. Here we report a multi-time-step experimental simulation of a stochastic process using less memory than the classical limit. A key feature of the photonic quantum information processor is that it creates a quantum superposition of all possible future trajectories that the system can evolve into. This superposition allows us to introduce, and demonstrate, the idea of comparing statistical futures of two classical processes via quantum interference. We demonstrate interference of two 16-dimensional quantum states, representing statistical futures of our process, with a visibility of  $0.96 \pm 0.02$ .

---

<sup>1</sup>Centre for Quantum Dynamics, Griffith University, Brisbane, QLD 4111, Australia. <sup>2</sup>School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore 639673, Singapore. <sup>3</sup>Complexity Institute, Nanyang Technological University, Singapore 639673, Singapore. <sup>4</sup>Centre for Quantum Technologies, National University of Singapore, Singapore 117543, Singapore. Correspondence and requests for materials should be addressed to F.G. (email: [farzad.ghafari@griffithuni.edu.au](mailto:farzad.ghafari@griffithuni.edu.au)) or to M.G. (email: [mgu@quantumcomplexity.org](mailto:mgu@quantumcomplexity.org)) or to G.J.P. (email: [g.pryde@griffith.edu.au](mailto:g.pryde@griffith.edu.au))

Many of the most interesting phenomena are complex—whether in urban design, meteorology or financial prediction, the systems involved feature a vast array of interacting components. Predicting and simulating such systems often requires the use of a prohibitive amount of data, evincing a pressing need for more efficient tools in algorithmic modelling and simulation.

Quantum technologies have shown the potential to dramatically reduce the amount of working memory required to simulate stochastic processes<sup>1,2</sup>. By tracking information about past observations directly within quantum states, a quantum device can replicate the system's conditional future behaviour, using less memory than the provably optimal classical limits. The key to achieving a quantum memory advantage is maintaining coherence of the quantum memory during the simulation process, enabling the encoding of relevant past information into non-orthogonal quantum states. This memory reduction comprises a new application of quantum processing, complementary to computational speedup<sup>3</sup>, cryptography<sup>4</sup>, sensing<sup>5,6</sup> and phase estimation<sup>7</sup>.

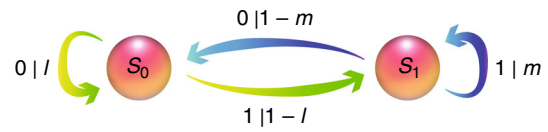
This advantage was first illustrated for simulating a particular stochastic process, where past information was encoded within non-orthogonal polarisation states of a single photon<sup>8</sup>. The scheme, however, maintained quantum coherence over only a single simulation cycle. This limitation meant that the resulting simulator exhibited a memory advantage only when simulating a single time step. To simulate multiple time steps, such a device required relevant information to be transferred to classical memory between time steps, negating any quantum advantage.

Here we develop a quantum simulator that overcomes this limitation, such that it exhibits a memory advantage when simulating multiple time steps. As an important additional benefit, our device enables us to create a quantum superposition over all potential future outcomes of a process. We illustrate that such an output lets us estimate the distinguishability in the statistical futures of two stochastic systems via quantum interference. Our experimental approach makes use of temporal (time-bin) encoding in an optical system to experimentally realise a quantum simulation over three consecutive steps, generating a coherent superposition over the process's potential future trajectories. We then implement two such quantum simulations in parallel, simultaneously generating superpositions over the trajectories for each of two independent systems. Experimentally, this corresponds to using our quantum simulators to produce and control high-dimensional quantum states. These are interfered, allowing estimation of how well the corresponding statistical futures coincide.

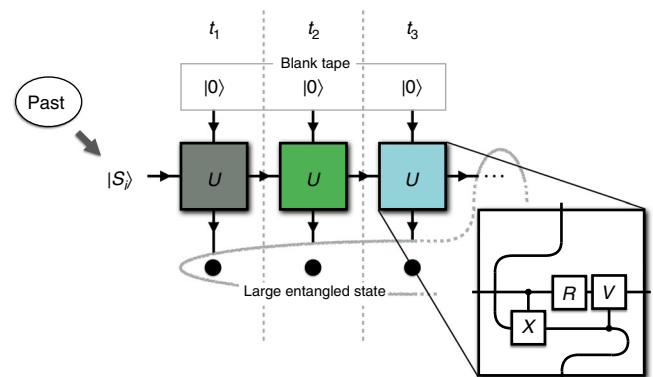
## Results

**Framework and tools.** In this work, we study a simple stochastic process known as the perturbed coin<sup>1</sup>. It consists of a binary random variable that represents the state of a coin (0 corresponds to heads and 1 to tails) inside a box. At each time step, the box is perturbed, causing the coin to flip with some probability. Afterwards, the state of the coin is emitted. In general the coin may be biased, so the probability of remaining in heads,  $l$ , can differ from the probability of remaining in tails,  $m$ , as presented in Fig. 1. Repetition of this procedure generates a string of 0s and 1s, whose statistics define the perturbed coin process.

Any device that seeks to replicate correct future statistics must retain relevant past information in a memory. This involves a prescription for configuring its memory in an appropriate state for each possible observed past, such that systematic actions on this memory recover a sequence of future outputs that are faithful to conditional future statistics. The amount of past information



**Fig. 1** Perturbed coin. The process that we study here is a coin with two outcomes, 0 and 1. The transition probabilities,  $T_{ij}$ , between different outcomes are determined by  $l$  and  $m$  for  $i, j \in \{0, 1\}$ . The optimal classical model uses the causal states ( $\{S_i\}$ ) depicted in the circles. There is a simple mapping from the past of the process to the relevant causal state: the last outcome from the coin determines the input causal state. Arrows, with the associated expressions  $j|T_{ij}$ , represent the transitions from causal states  $S_i$  to  $S_j$  with probability  $T_{ij}$ , emitting the classical outcome  $j$ . In the quantum model, the causal states become quantum states,  $\{|S_i\rangle\}$



**Fig. 2** Conceptual representation of the multiple-step quantum simulation of a perturbed coin. The memory system of the simulator is initialised in a qubit state  $|S_i\rangle$ , where  $i \in \{0, 1\}$  depends on the past of the process. The ancillary qubits are all initialised in a fixed known state  $|0\rangle$ —the logical zero state—and thus contain no information about the past of the process. At each time step  $t_k$ , the simulator interacts with the  $k$ th ancilla through the same unitary operator  $U$ . The inset shows how we implement the relevant unitary operator. The gates are a controlled- $X$ , a single-qubit rotation  $R$  such that  $R|0\rangle = |S_0\rangle$ , and a controlled- $V$  such that  $VR|1\rangle = |S_1\rangle$ . This sequence of interactions results in an entangled internal state that includes all the ancillary qubits and the memory state of the simulator. Measuring the ancillas samples the statistical distribution of the process, and at the same time the internal state of the simulator collapses into the correct memory state required for further simulation steps

stored in memory is quantified by the Shannon entropy  $C = -\sum_s d_s \log d_s$ , where  $d_s$  is the probability that the memory is in state  $s$  and the logarithm is in base 2. The minimal possible memory required,  $C_\mu$ , is known as the statistical complexity, and is an important measure of structure in complexity science<sup>9–12</sup>. For the perturbed coin (Fig. 1), the minimal information required about the past is the current state of the coin. This induces a statistical complexity of  $C_\mu = -q \log q - (1-q) \log (1-q)$ , where  $q$  represents the probability that the last outcome was heads (see Eq. (1) in Methods).

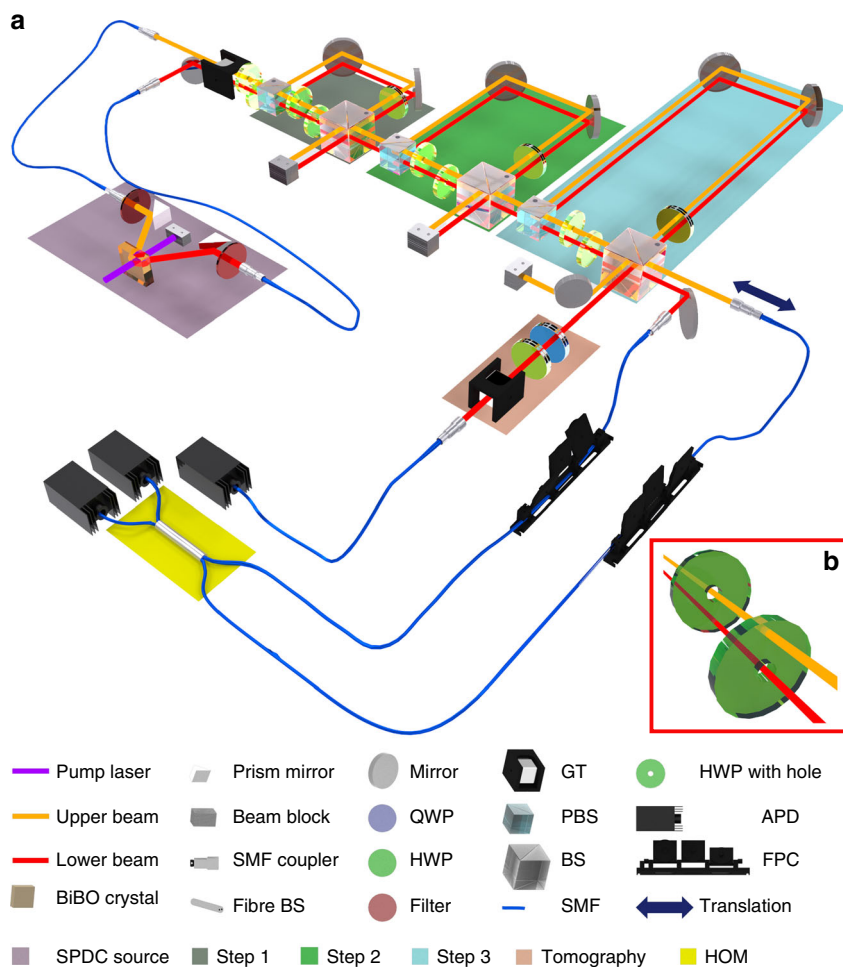
A quantum simulator can further reduce memory requirements by encoding the two possible outcomes of the process into mutually non-orthogonal states. Future statistics are then generated by a series of unitary interactions, ensuring that this entropic advantage is maintained at all times during simulation<sup>13</sup>. For the case of the perturbed coin, the quantum simulator can be implemented as shown in Fig. 2. The state of the machine encodes relevant information about past outcomes—here the state of the coin after the last step. It is represented as one of two states,  $|S_0\rangle$  or  $|S_1\rangle$ , of a quantum system that sequentially interacts

with ancillary systems. Each interaction corresponds to a time step of the stochastic process. All the ancillary systems start in a fixed state, and therefore do not contain any information. The sequence of interactions produces an entangled state. Measuring the ancillary systems after the desired number of steps provides a sample of the statistics.

**Experimental implementation.** Motivated by recent realisations of quantum walks in linear optical setups with time-bin encoding<sup>14–17</sup>, we implement the memory system and multiple ancillas—here corresponding to three time steps—by encoding on a single photon. The ancillas, which can be read to obtain the classical outcomes of the process, are encoded in the arrival time of the photon, and the memory state of the simulator is encoded in its polarisation. Thus, for a simulation of  $M$  time steps, a  $2^M$ -dimensional system corresponding to  $2^M$  different photon arrival

times replaces  $M$  distinct ancillary photons. Instead of measuring the classical outcome at each time step, our quantum information processor keeps the photon and builds up a superposition in a high-dimensional Hilbert space; in our case  $M = 3$ , and the output of the simulator is 16-dimensional (8 arrival time modes  $\times$  2 polarisation modes). The associated memory cost during this process does not increase since all operations remain unitary—and thus conserve entropy. Of course, distinct ancilla qubits could be used instead, but encoding in multiple degrees of freedom provides a convenient, effective and high-fidelity approach for small- to medium-sized photonic systems.

Our experiment demonstrates that high-dimensional (here 16-dimensional) quantum states can be encoded and manipulated in photonic temporal and polarisation modes with high fidelity<sup>18,19</sup>. This complements other related works involving hybrid optical states using spatial (path and OAM) and polarisation modes<sup>20–23</sup>. It also substantiates the oft-repeated claim that combining



**Fig. 3** Experimental setup. **a** Single photons are generated from a degenerate spontaneous parametric down-conversion (SPDC) source pumped by a 410 nm continuous-wave laser. After filtering the generated photons with  $(820 \pm 1.5)$  nm bandpass filters, photons in the lower beam (red) and upper beam (orange) are separately prepared in their respective input states,  $|S_0\rangle$  or  $|S_1\rangle$ , using half-wave plates (HWP). Polarisation qubits, one from each beam, are used as memory states for the simulation of two separate and potentially different processes  $\Pi_1$  (red) and  $\Pi_2$  (orange). (For the first experiment described in the text, where only one process is required, the second SPDC output (orange) is sent straight to a heralding detector, rather than through the apparatus.) To implement the three-step simulation, three processor blocks are built—labelled Step 1, Step 2 and Step 3. In each step, path and arrival time modes are also employed to realise the relevant physical operation, as explained further in Methods. The output of one of the simulators (lower beam) is used to perform the polarisation tomography and to measure the arrival times of each photon in order to sample the statistical future. To measure the overlap of the future statistics of two processes, both photons are used, and the other outputs of the third beam splitter (BS) are interfered in a fibre BS (yellow box). An automated translation stage is used to move one of the couplers in order to vary the relative delay between the single-photon wave packets. Avalanche photodiodes (APD) and a single-photon counting module are used to count the photons. SMF stands for single-mode fibre, QWP quarter-wave plate, GT Glan-Taylor prism, PBS polarising beam splitter, and FPC fibre polarisation controller. **b** The inset shows a close-up of two vertically separated beams passing through two HWPs with holes, each of which only acts on one of the beams

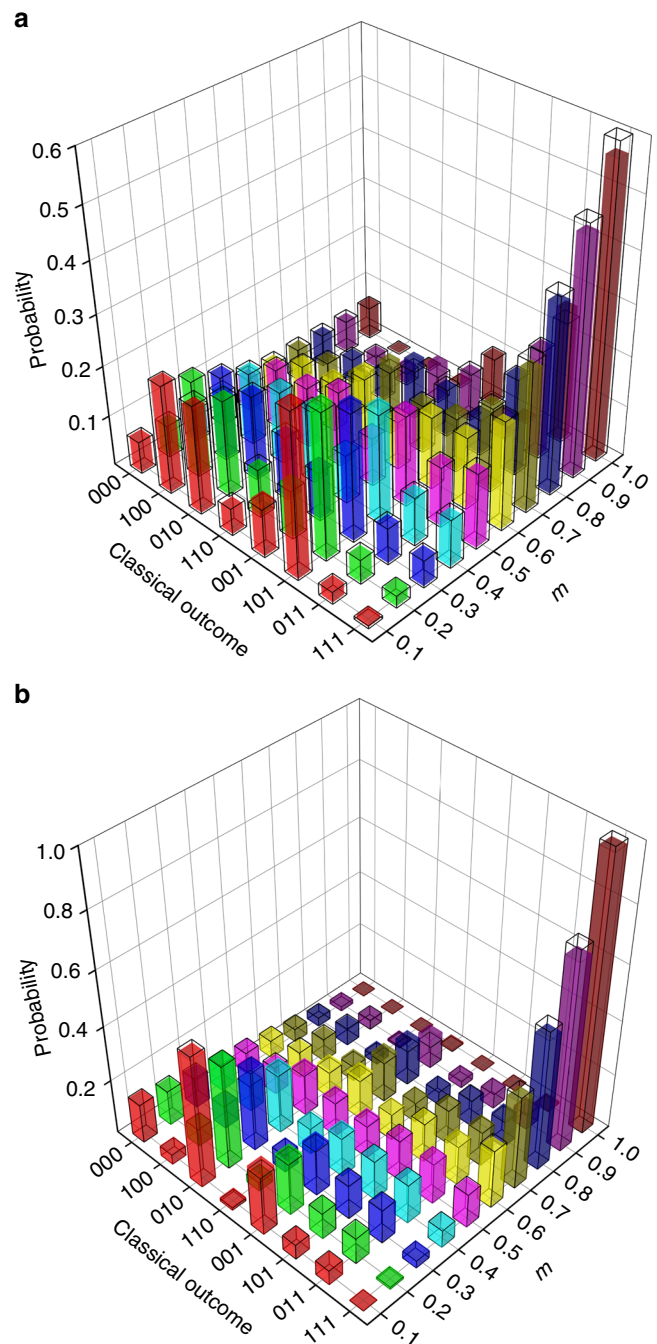
different photonic encodings<sup>24,25</sup> is a practical tool for various quantum information tasks, for example studying the remote preparation of entangled states<sup>26</sup>, complementarity<sup>27</sup>, Bell inequalities<sup>21,28,29</sup>, quantum key distribution implementations<sup>30</sup> and complete optical Bell state analysers<sup>31,32</sup>.

Our first task consists of performing the quantum simulation of the perturbed coin. In particular, we seek to verify that the simulator samples from the correct statistical distributions, and to demonstrate the memory advantage due to quantum encoding. The experimental setup is shown in Fig. 3. We generate degenerate pairs of single photons through spontaneous parametric down-conversion. One of the photons (depicted as the red, lower beam in the figure) is prepared in the state  $|S_0\rangle$  or  $|S_1\rangle$ , depending on the past of the process. It then passes through three sequential blocks, which represent the three time steps being simulated. In each block, the short and long paths correspond to outcomes 0 and 1, respectively (details in Methods). For the simulation, only one of the photons (the red beam) is used, and the other photon (orange beam in the figure) is not used except as a herald, and is measured immediately after generation (for this task, it does not go through the apparatus as shown in the figure). We then estimate the polarisation state of the red-beam photon in the tomographic reconstruction at the end of the third block, and also measure its arrival time (using the orange-beam photon as a reference). In this way, we obtain the probability distribution of the stochastic process as simulated by our quantum information processor, together with the final memory state of our simulator, which is needed for further simulation steps.

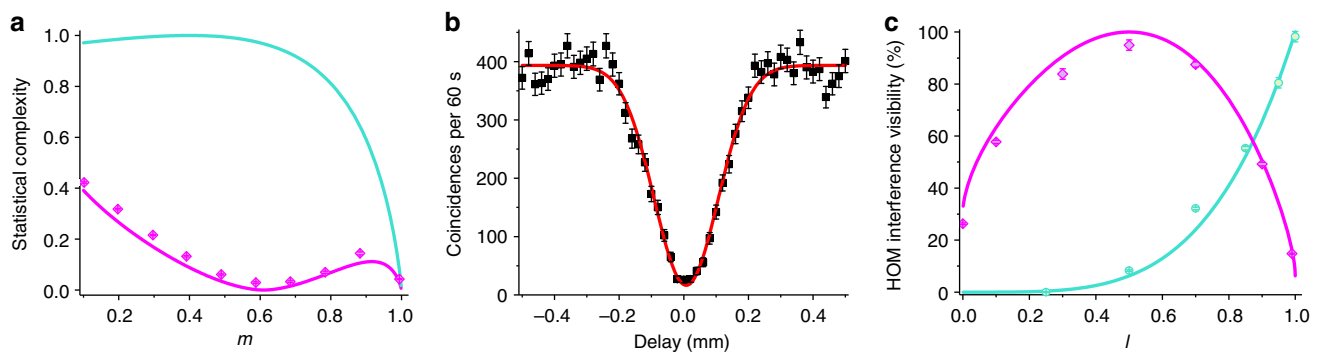
**Experimental results.** The experimentally determined outcome probabilities are shown in Fig. 4, and are close to the expected theoretical values. The main discrepancies with theory are due to small differences between nominally identical polarisation elements, and the non-identical single-mode-fibre coupling efficiency of photons taking different paths through the simulator. In order to evaluate how well they agree, we calculate the (classical) fidelity<sup>33</sup> for each set of parameters and initial conditions that we have simulated in our experiment. All the values obtained for this (classical) fidelity are larger than 0.991. Typical uncertainties are around 0.001.

To compare the use of quantum and classical resources, we use  $C_q$ , the quantum counterpart to the classical statistical complexity (the entropy of the memory register of the quantum simulator), which quantifies the memory requirement of the quantum simulator. We thus calculate  $C_q$  for this process (details in Methods). The experimental results are shown in Fig. 5a. The corresponding classical statistical complexity is also shown for the sake of comparison, demonstrating that quantum resources dramatically reduce the amount of memory needed for simulating a multi-step stochastic process.

To guarantee that the quantum memory advantage is maintained at all stages of the simulation process, we require the internal dynamics to be close to (ideally, completely) unitary. We can verify this by demonstrating the coherence of the output state that includes all the ancillary qubits and the memory state of the simulator. We observe this coherence via two-photon quantum interference. We use the complete setup of Fig. 3, where the photon depicted by the orange path is no longer measured after generation (as done previously), but also goes through the apparatus. Both the photons pass independently through the three sequential blocks, with each experiencing nominally the same optical elements (although different settings are possible). If the coherence between the different time bins



**Fig. 4** Sampling of statistical futures. **a** The coloured bars show the measured probability of different outcomes for the three simulation steps, when the initial state is  $|S_0\rangle$  with  $l = 0.4$ , and for discrete values of  $m$  in the range 0.1–1.0. The transparent bars show the theoretically calculated probabilities for the corresponding process. **b** The sampled future of the same process, when the initial state is  $|S_1\rangle$ . Uncertainties, due to the Poissonian distribution of photon counts, are so small that they are not visible in the graphs—therefore, they are not depicted. Note that the classical probability distribution is determined by the process parameters  $l$  and  $m$ , as well as the initial causal state. For example, if the last outcome of the coin is 1, the quantum simulator is initialised in state  $|S_1\rangle$ . The conditional probability of subsequently observing 111 is then  $m^3$ . For a fixed  $l = 0.4$  and increasing  $m$ , the average probability of getting 1 in the simulation thus rises accordingly. This can be seen by the higher columns in the right corners of both graphs



**Fig. 5** Experimentally determined complexity and interference visibility. **a** The quantum statistical complexity  $C_q$  of our quantum simulator, as the probability of remaining in tails during perturbation,  $m$ , is varied. The probability of remaining in heads,  $l$ , is fixed at 0.4. Due to small experimental imperfections the actual implemented values of  $l$  and  $m$  deviate slightly from the nominal round values (see Methods for more details). Data points are experimental measurements of  $C_q$ , and the magenta and turquoise curves are theoretical estimations for the quantum and classical complexities  $C_q$  and  $C_{cl}$ , respectively. **b** Two-photon interference of the superpositions of future trajectories in two implemented stochastic processes,  $\Pi_1$  and  $\Pi_2$ , with  $\Pi_1 = \Pi_2$  such that  $l = 0.5$  and  $m = 0.5$  and thus  $|S_0\rangle = |S_1\rangle$ . Since  $\Pi_1 = \Pi_2$ , an interference visibility of 100% is theoretically expected, while fitting the experimental data yields a visibility of  $0.96 \pm 0.02$ . In the graph, the number of measured twofold coincidences is depicted versus the relative delay between single-photon wave packets. **c** Magenta and turquoise elements (points—experiment; curves—theory) show the comparison of the statistical futures from two stochastic processes by two-photon interference visibility. In each case, one process ( $\Pi_1$ ) is fixed, and the other process ( $\Pi_2$ ) has fixed  $m$  but varying  $l$ . Magenta represents interference of the output states of the simulators for  $\Pi_1$  ( $|S_0\rangle$  is the input memory state,  $l = 0.5$ , and  $m = 0.5$ ) with  $\Pi_2$  ( $|S_0\rangle$  is the input memory state,  $m = 0.5$ , and  $l$  varying). Turquoise represents interference of the output states of the simulators for  $\Pi_1$  ( $|S_0\rangle$  is the input memory state,  $l = 1.0$ , and  $m = 1.0$ ) with  $\Pi_2$  ( $|S_0\rangle$  is the input memory state,  $m = 0.5$ , and  $l$  varying). Uncertainties, from the Poissonian distribution of photon counts, are so small that they are barely visible in some of the graphs

and polarisations exploited in our simulation is maintained, we expect a complete interference, which means that the visibility ideally should be unity. The result in Fig. 5b shows a visibility of  $0.96 \pm 0.02$  for the case where the theoretical output states of the apparatus are uniform superpositions of all time bins and polarisations (which is the scenario where the highest discrepancy from the ideal visibility would be expected as it is most susceptible to imperfections). The high value obtained here indicates that our simulator is (almost) implementing a unitary operator, and the entropy of our system does not significantly increase throughout the simulation process. This requirement is essential for preserving the quantum memory advantage. Moreover, apart from the specific application of this apparatus to simulate classical stochastic processes, this result is also significant in a more general context, since it demonstrates the interference of two discrete high-dimensional states with an extremely high visibility<sup>23</sup>.

Modifying this experimental setup allows us to compare two different processes,  $\Pi_1$  and  $\Pi_2$ . Clearly, one way to perform such a statistical comparison is to consider each process individually, and sample its outcomes to reconstruct the corresponding distribution. These two reconstructed distributions can then be compared. However, we notice that in our quantum simulation, all the information about the future statistics is already encoded in the state that exists in our apparatus. Thus, we do not need to collapse the superposition of possible outcomes by sampling, instead we can exploit this superposition for our task of comparing the future of processes. In particular, by simultaneously running quantum simulations of processes  $\Pi_1$  and  $\Pi_2$  in parallel and interfering the resulting output states, we can estimate the overlap of their future statistics.

In our experiment, we realise different processes by applying different operations to the two photons (red beam and orange beam) in the three blocks of the setup in Fig. 3. To implement the parameters of each process separately, we use half-wave plates with holes, which allow us to change the polarisation of one beam without affecting the other. We fix one of the processes and change the other process gradually. As the parameters defining the processes become increasingly similar, the two output

probability distributions overlap more. This is reflected in the experiment by a higher visibility value, showing how the comparison between two sets of future statistics can be evaluated via interference visibility. Results are shown in Fig. 5c, where the experimental values are close to theoretical predictions. However, there remain slight discrepancies because of experimental imperfections such as small spatial and polarisation mode mismatches. These techniques could be adapted to attain a quantum advantage in estimating the distance between two normalised vectors<sup>34</sup>, which plays an essential role in machine learning tasks such as image recognition<sup>35</sup>.

## Discussion

Our multi-step photonic implementation of a stochastic simulation has verified the memory advantage available with quantum resources. We have demonstrated that it is possible to maintain this advantage at all stages of the simulation by preserving quantum coherence, as opposed to previous experiments<sup>8,36</sup>. Further, we have shown that superpositions of process outcomes can be interfered. These techniques have the potential to reduce memory requirements in simulations of stochastic processes and to provide tools for advances in quantum machine learning and communication complexity.

The time-bin-encoding techniques in our experiment can be extended to other small- and medium-scale simulations by expanding the number of time bins. For example,  $10^8$  time-bin modes have been realised in the context of communication complexity<sup>37</sup>. However, the number of bins does not scale efficiently with the number of qubits, and thus very-large-scale simulations are not possible with this encoding. This is not a fundamental problem, as the concepts that we demonstrate can be equivalently implemented in other photonic encodings or in other qubit systems. Our current demonstration also uses non-deterministic (post-selected) mode recombination at certain beam splitters within the circuit. This implementation is convenient, but not necessary and thus not a fundamental limitation: a deterministic multi-step simulator could be realised with a step-dependent delay mechanism—for instance, a

controlled fast switch connected to fibre paths of different lengths.

The comparison of future statistics has direct relation to other protocols, such as quantum fingerprinting and state comparison in communication complexity<sup>34,37</sup>. Fingerprinting involves estimating the distance between two vectors, where the resource to be minimised is the amount of communication. For the comparison of two vectors, quantum mechanics can reduce the amount of communication required beyond classical limits. In the quantum protocol, Alice and Bob perform a SWAP test—a quantum information primitive, which compares two arbitrary states. Two-photon interference is known to be equivalent to a SWAP test<sup>38</sup>. Our comparison of futures can be cast as a similar problem. In this case, the task would be for Alice and Bob, who each have their future statistics from potentially different processes, to compare the two statistical futures<sup>34</sup>. In principle, for very high-dimensional Hilbert spaces, a comparison of statistical futures via two-photon interference can achieve a quantum advantage in communication complexity. The comparison of two vectors is also an important component of many machine learning tasks, and thus a similar advantage could extend to more general settings like speech recognition<sup>35</sup>.

**Methods**

**Theoretical background.** A discrete-time stochastic process is generally described by a joint probability distribution,  $p(\vec{X}, \vec{X})$ , where  $\vec{X} = \dots, X_{-1}, X_0$  ( $\vec{X} = X_1, X_2, \dots$ ) denotes the random variables that govern the statistics of past (future) observations. Each past (future) configuration of the random process is denoted by  $\vec{x}$  ( $\vec{x}$ ). For an observed past configuration  $\vec{x}$ , the future statistics are dictated by the conditional probability  $p(\vec{X} = \vec{x} | \vec{X} = \vec{x})$ , which we abbreviate as  $p(\vec{x} | \vec{x})$ .

By categorising all sets of past events with the same future statistics into equivalence classes (called causal states, which are encoded as memory states of the simulator), the optimal classical model (called the  $\epsilon$ -machine<sup>10,39</sup>) only needs to store the class  $\epsilon(\vec{x})$  that  $\vec{x}$  belongs to. That is, given only  $\epsilon(\vec{x})$ , the  $\epsilon$ -machine is able to make a statistically accurate inference of the process' conditional future. By observing the outcome of the stochastic process over a long time, one can infer the probability of each causal state and transition probabilities between them. For a stochastic process, the  $N$  causal states  $S = \{S_i\}_{i=1}^N$  and their relevant transition probabilities are enough to realise the  $\epsilon$ -machine model. The resulting  $\epsilon$ -machine requires<sup>33</sup>

$$C_\mu = - \sum_{i=1}^N d_i \log d_i, \tag{1}$$

bits of information about the past, where  $d_i$  is the probability that the past is in causal state  $S_i$ . No other predictive model can simulate the future while storing less information about the past. Thus,  $C_\mu$  has been termed the statistical complexity<sup>10,40,41</sup>, and is considered a fundamental measure of complexity that captures how resource-intensive it is to predict the future of a given process.

It has been theoretically proven that for many processes, including the one studied here, there exists a quantum  $\epsilon$ -machine with entropy  $C_q$ , such that  $C_q < C_\mu$ <sup>1</sup>. Similar to its classical counterpart, this quantum model is defined by its causal states  $\{|S_i\rangle\}$  and the corresponding transition probabilities. On average, the entropy of such a quantum  $\epsilon$ -machine is given by

$$C_q = -\text{Tr}(\rho \log \rho), \tag{2}$$

where  $\rho = \sum_i d_i |S_i\rangle\langle S_i|$ .

**Three-step simulation of a perturbed coin.** For the perturbed coin process, the optimal quantum causal states can be written as<sup>1</sup>:

$$|S_0\rangle = \sqrt{l}|0\rangle + \sqrt{1-l}|1\rangle, \tag{3}$$

$$|S_1\rangle = \sqrt{1-m}|0\rangle + \sqrt{m}|1\rangle. \tag{4}$$

To give an example of the output state of our simulator, let us consider a perturbed coin defined by its parameters  $l$  and  $m$ , which we denote as process  $\Pi_1$ . The output of the corresponding quantum  $\epsilon$ -machine after three time steps is given

by the superposition

$$\sum_{x_n} \sqrt{p(x_1, x_2, x_3 | S_i, \Pi_1)} |x_1, x_2, x_3\rangle |S_{x_3}\rangle, \tag{5}$$

where  $n = \{1, 2, 3\}$  and  $p(x_1, x_2, x_3 | S_i, \Pi_1)$  is the probability to obtain  $x_1, x_2$  and  $x_3$  as the outcomes of three time steps of the process  $\Pi_1$  when the input causal state is  $|S_i\rangle$ . The value of  $p$  can be evaluated theoretically from the transition probabilities between causal states (Fig. 1). The variables  $x_n \in \{0, 1\}$  are the configurations of random variables  $x_1, x_2$  and  $x_3$ , respectively. To sample from the future statistics of the perturbed coin process, we perform a simultaneous measurement of all the ancillary qubits after the three time steps. By also characterising the polarisation state of the photon in each case, we can tomographically reconstruct the output state associated with each time bin, and thus experimentally determine the statistical complexity of the simulation. To calculate the statistical complexity,  $C_q$ , for this process, we need to find the state  $\rho$ :

$$\rho = d_0 \sum_{x_n} p(x_1, x_2, x_3 | S_0, \Pi_1) \rho_{\text{pol}|S_0} + d_1 \sum_{x_n} p(x_1, x_2, x_3 | S_1, \Pi_1) \rho_{\text{pol}|S_1}, \tag{6}$$

where

$$d_0 = \frac{\sum_{x_1, x_2} p(x_1, x_2, x_3 = 0 | S_i, \Pi_1)}{\sum_{x_1, x_2} p(x_1, x_2, x_3 = 1 | S_0, \Pi_1) + \sum_{x_1, x_2} p(x_1, x_2, x_3 = 0 | S_i, \Pi_1)},$$

$d_1 = 1 - d_0$ , and  $S_{\text{pol}|S_i}$  is the tomographically reconstructed polarisation state at each arrival time, conditioned on the input memory state being encoded in  $|S_i\rangle$ .

**Verifying the unitarity of the processor via two-photon quantum interference.**

To verify that the operation is unitary, which guarantees the conservation of the entropy, we need to show that the superposition of different modes, both in time and polarisation, is coherent and that this coherence is maintained throughout the whole process. Using a pure state as the input and viewing the entire simulation as a black box, the output of the unitary operations inside the box should ideally be a pure state. In order to experimentally demonstrate this, we consider the case of simultaneously implementing two setups to model two identical processes,  $\Pi_1 = \Pi_2$ . It is possible to verify that two uncorrelated single photons are in identical pure states via two-photon interference—the Hong-Ou-Mandel (HOM) effect. The visibility of the interference,  $v = \frac{P_{\text{max}} - P_{\text{min}}}{P_{\text{max}}}$ , where  $P_{\text{max}}$  ( $P_{\text{min}}$ ) is the maximum (minimum) of two-photon coincidence detections measured when varying the delay between the two beams, can only be unity if the photons are in pure and identical states.

**Comparison of future statistics.** The case of unequal processes also provides useful information. If  $\Pi_1 \neq \Pi_2$  and the output states are pure, the overlap of different future output statistics can be deduced by interfering the output photons. For two photons in states  $|\psi\rangle$  and  $|\phi\rangle$  entering two input ports of a 50:50 beam splitter, the probability of finding a coincidence is  $\frac{1 - |\langle\phi|\psi\rangle|^2}{2}$ , where  $\langle\phi|\psi\rangle$  is the overlap of the two states. Therefore, one can use the HOM interference visibility  $v$  to estimate overlaps, by noting that  $v = |\langle\phi|\psi\rangle|^2$ . For our stochastic processes, the overlaps of the photonic output states are directly related to the overlaps of the future statistics produced by the two processes. For two different processes  $\Pi_1$  and  $\Pi_2$ , let  $|S_i\rangle$  be a causal state of  $\Pi_1$ , and  $|T_j\rangle$  be a causal state of  $\Pi_2$ . Using Eq. (5), in general the overlap between the respective outputs of the quantum simulators for  $\Pi_1$  and  $\Pi_2$  will be

$$\sum_{x_n} \sqrt{p(x_1, x_2, x_3 | S_i, \Pi_1) p(x_1, x_2, x_3 | T_j, \Pi_2)} \langle S_{x_3} | T_{x_3} \rangle. \tag{7}$$

Since the perturbed coin process has Markov order one, and there is a one-to-one correspondence between the classical outcome and the causal state the machine transitions to, interfering the output states from a pair of quantum simulators for  $\Pi_1$  and  $\Pi_2$  as in Eq. (7), actually results in an overlap

$$\sum_{x_n} \sqrt{p(x_1, x_2, x_3, x_4 | S_i, \Pi_1) p(x_1, x_2, x_3, x_4 | T_j, \Pi_2)}. \tag{8}$$

That is, in this special case we are able to find the difference between conditional futures up to one additional time step. Therefore, we can use our photonic quantum information processor for two tasks: (1) to simulate the future outcomes over three time steps of the classical stochastic process; and (2) to estimate the overlap of the future output statistics over four time steps.

**Details of the experimental design.** The schematic in Fig. 3 shows how we implement the multi-step quantum-enhanced stochastic processor. Consider, for instance, the scenario where we want to sample the statistics. For one process (i.e. for one beam) a single photon is injected in the left-hand side of the circuit, from the source, with the state  $|0\rangle$  ( $|1\rangle$ ), which is encoded as  $|H\rangle = \text{horizontal}$  ( $|V\rangle = \text{vertical}$ ) polarisation. The first wave plate creates the desired initial causal state of our perturbed coin, either  $|S_0\rangle_{\text{pol}} = \sqrt{l}|H\rangle_{\text{pol}} + \sqrt{1-l}|V\rangle_{\text{pol}}$  or  $|S_1\rangle_{\text{pol}} = \sqrt{1-m}|H\rangle_{\text{pol}} + \sqrt{m}|V\rangle_{\text{pol}}$ . The purpose of the first block is to transform a photon with a causal state encoded in polarisation into an appropriately weighted superposition of the classical outcomes of the first step encoded in the arrival time (denoted here as the delay degree of freedom,  $\text{del}$ ), with the corresponding next causal state encoded in the polarisation:

$$|S_0\rangle_{\text{pol}}|0\rangle_{\text{del}} \rightarrow \sqrt{l}|S_0\rangle_{\text{pol}}|0\rangle_{\text{del}} + \sqrt{1-l}|S_1\rangle_{\text{pol}}|t_1\rangle_{\text{del}}, \quad (9)$$

$$|S_1\rangle_{\text{pol}}|0\rangle_{\text{del}} \rightarrow \sqrt{1-m}|S_0\rangle_{\text{pol}}|0\rangle_{\text{del}} + \sqrt{m}|S_1\rangle_{\text{pol}}|t_1\rangle_{\text{del}}. \quad (10)$$

This is achieved by temporarily using the photon path as an auxiliary degree of freedom: a polarising beam splitter maps the polarisation degree of freedom onto the path, which is then copied onto the arrival time through the use of different path lengths ( $|H\rangle \rightarrow \text{short path}$  and  $|V\rangle \rightarrow \text{long path}$ ). By using a wave plate in each of the two paths, a path-dependent (and therefore, arrival-time-dependent) transformation of the polarisation into one of the two causal states is achieved:  $|S_0\rangle_{\text{pol}}$  in the short path and  $|S_1\rangle_{\text{pol}}$  in the long path.

Next, the information on the path degree of freedom is erased, to avoid an exponential scaling of the number of paths (and optical elements in the experiment) with the number of time steps. To this end, the paths are recombined in a 50:50 beam splitter, and subsequently post-selected for the photon exiting in the right output arm at the end of the first block (Fig. 3). This means that we will lose half of our photons at the beam splitter, but in each run that we post-select, the evolution is unitary because the post-selection ensures that no photon is detected in the other output arm. By repeating the described block at each time step, we have three blocks to realise a three-step machine. The use of a sequence of interferometers has also been demonstrated in other experiments to study different topics in quantum information, such as non-Markovian dynamics and sequential state discrimination<sup>42,43</sup>.

To be able to attribute a different arrival time to each sequence of classical outcomes, we require a unique path length for every possible combination of short or long paths within the three blocks. The delays are implemented as  $t_1 = 2$  ns at the first step,  $t_2 = 4$  ns at the second and  $t_3 = 8$  ns at the third step. The arrival times are discriminated by time-resolving single-photon detectors. The coincidence window for HOM interference is long enough to include the state which is spread out in a 14 ns time interval.

After the third step, we have the measurement stage at one output arm of the third BS and the circuit continues at the other, which is exploited for the second task of our work. In order to run our simulation and estimate the memory efficiency of this scheme compared to the optimal classical one, we measure the final arrival times (encoding the three ancillary qubits of the original scheme) and reconstruct the final polarisation state of the photon. This can be done simultaneously at the tomography stage, by also measuring the arrival times of the photons, allowing a full reconstruction of the polarisation state and arrival time.

The same apparatus can be exploited for the interference part of our experiment, the only difference being that now two single photons are injected in the setup. They both pass through the three blocks described above. When we want to verify the unitarity of our simulation, the elements in the blocks are the same for both the photons, so as to have  $\Pi_1 = \Pi_2$ ; on the other hand, they are different when we want to compare the future statistics of two different processes ( $\Pi_1 \neq \Pi_2$ ). After the output of the third block, the two photons interfere in a fibre BS and the number of coincidences is measured.

**Details of the  $l$  and  $m$  parameters used in the experiment.** The simulated process, for which the  $C_q$  results are depicted in Fig. 5a, is a perturbed coin with parameters  $l = 0.4$  and  $m$  ranging from 0.1 to 1.0 in increments of 0.1. Due to experimental imperfections the actual implemented values of  $l$  and  $m$  slightly deviate from the nominal ones ( $l = \{0.397, 0.197, 0.297, 0.391, 0.490, 0.588, 0.685, 0.784, 0.882, 0.994\}$ ). In Fig. 5c, the turquoise and magenta colours both show the case of two processes. For the turquoise graph, the fixed process is a stochastic process of a perturbed coin with input causal state  $|S_0\rangle$ ,  $l = 1.0$ , and  $m = 1.0$ . The varying stochastic processes are the ones with input causal state  $|S_0\rangle$ ,  $m = 0.5$ , and nominal  $l = \{0.25, 0.50, 0.70, 0.85, 0.95, 1.0\}$  (the parameter  $l$  is used to change between different processes). For the magenta graph, the fixed stochastic process is a perturbed coin with input causal state  $|S_0\rangle$ ,  $l = 0.5$ , and  $m = 0.5$ . The varying ones are the stochastic processes with input causal state  $|S_0\rangle$ ,  $m = 0.5$ , and nominal  $l = \{0.00, 0.10, 0.30, 0.50, 0.70, 0.90, 0.99\}$ .

## Data availability

All data, relevant to the information and figures presented in this manuscript, are available upon reasonable request.

Received: 27 July 2018 Accepted: 4 February 2019

Published online: 09 April 2019

## References

- Gu, M., Wiesner, K., Rieper, E. & Vedral, V. Quantum mechanics can reduce the complexity of classical models. *Nat. Commun.* **3**, 762 (2012).
- 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).
- Lloyd, S. Universal quantum simulators. *Science* **273**, 1073–1078 (1996).
- Bennett, C. H. & Brassard, G. Quantum cryptography: Public key distribution and coin tossing. *Proc. IEEE International Conference on Computers, Systems, and Signal Processing*, 175–179 (Bangalore, 1984).
- Giovannetti, V., Lloyd, S. & Maccone, L. Quantum-enhanced measurements: Beating the standard quantum limit. *Science* **306**, 1330–1336 (2004).
- Slussarenko, S. et al. Unconditional violation of the shot-noise limit in photonic quantum metrology. *Nat. Photonics* **11**, 700–703 (2017).
- Xiang, G. Y., Higgins, B. L., Berry, D. W., Wiseman, H. M. & Pryde, G. J. Entanglement-enhanced measurement of a completely unknown optical phase. *Nat. Photonics* **5**, 43–47 (2011).
- 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**, e1601302 (2017).
- Grassberger, P. Toward a quantitative theory of self-generated complexity. *Int. J. Theor. Phys.* **25**, 907–938 (1986).
- Crutchfield, J. P. & Young, K. Inferring statistical complexity. *Phys. Rev. Lett.* **63**, 105 (1989).
- Shalizi, C. R. & Crutchfield, J. P. Computational mechanics: Pattern and prediction, structure and simplicity. *J. Stat. Phys.* **104**, 817–879 (2001).
- Crutchfield, J. P., Ellison, C. J. & Mahoney, J. R. Time's barbed arrow: Irreversibility, crypticity, and stored information. *Phys. Rev. Lett.* **103**, 094101 (2009).
- Binder, F. C., Thompson, J. & Gu, M. Practical unitary simulator for non-Markovian complex processes. *Phys. Rev. Lett.* **120**, 240502 (2018).
- Schreiber, A. et al. Photons walking the line: A quantum walk with adjustable coin operations. *Phys. Rev. Lett.* **104**, 050502 (2010).
- Schreiber, A. et al. A 2D quantum walk simulation of two-particle dynamics. *Science* **336**, 55–58 (2012).
- Jeong, Y.-C., Di Franco, C., Lim, H.-T., Kim, M. S. & Kim, Y.-H. Experimental realization of a delayed-choice quantum walk. *Nat. Commun.* **4**, 2471 (2013).
- Boutari, J. et al. Large scale quantum walks by means of optical fiber cavities. *J. Opt.* **18**, 094007 (2016).
- Franson, J. D. Bell inequality for position and time. *Phys. Rev. Lett.* **62**, 2205 (1989).
- Kwiat, P. G., Steinberg, A. M. & Chiao, R. Y. High-visibility interference in a Bell-inequality experiment for energy and time. *Phys. Rev. A* **47**, R2472 (1993).
- Takeuchi, S. Experimental demonstration of a three-qubit quantum computation algorithm using a single photon and linear optics. *Phys. Rev. A* **62**, 032301 (2000).
- Ma, X.-s., Qarry, A., Kofler, J., Jennewein, T. & Zeilinger, A. Experimental violation of a Bell inequality with two different degrees of freedom of entangled particle pairs. *Phys. Rev. A* **79**, 042101 (2009).
- Nagali, E. & Sciarrino, F. Generation of hybrid polarization-orbital angular momentum entangled states. *Opt. Express* **18**, 18243–18248 (2010).
- Zhang, Y. et al. Engineering two-photon high-dimensional states through quantum interference. *Sci. Adv.* **2**, e1501165 (2016).
- Kwiat, P. G. Hyper-entangled states. *J. Mod. Opt.* **44**, 2173–2184 (1997).
- Barreiro, J. T., Langford, N. K., Peters, N. A. & Kwiat, P. G. Generation of hyperentangled photon pairs. *Phys. Rev. Lett.* **95**, 260501 (2005).
- Barreiro, J. T., Wei, T.-C. & Kwiat, P. G. Remote preparation of single-photon "hybrid" entangled and vector-polarization states. *Phys. Rev. Lett.* **105**, 030407 (2010).
- Nogueira, W. et al. Interference and complementarity for two-photon hybrid entangled states. *Phys. Rev. A* **82**, 042104 (2010).
- Vallés, A. et al. Generation of tunable entanglement and violation of a Bell-like inequality between different degrees of freedom of a single photon. *Phys. Rev. A* **90**, 052326 (2014).

29. Dada, A. C., Leach, J., Buller, G. S., Padgett, M. J. & Andersson, E. Experimental high-dimensional two-photon entanglement and violations of generalized Bell inequalities. *Nat. Phys.* **7**, 677–680 (2011).
30. Takemoto, K. et al. Quantum key distribution over 120 km using ultrahigh purity single-photon source and superconducting single-photon detectors. *Sci. Rep.* **5**, 14383 (2015).
31. Walborn, S. P., Pádua, S. & Monken, C. H. Hyperentanglement-assisted Bell-state analysis. *Phys. Rev. A* **68**, 042313 (2003).
32. Wei, T.-C., Barreiro, J. T. & Kwiat, P. G. Hyperentangled Bell-state analysis. *Phys. Rev. A* **75**, 060305(R) (2007).
33. Nielsen, M. A. & Chuang, I. L. *Quantum Computation and Quantum Information* (Cambridge Univ. Press, New York, 2010).
34. Kumar, N., Diamanti, E. & Kerenidis, I. Efficient quantum communications with coherent state fingerprints over multiple channels. *Phys. Rev. A* **95**, 032337 (2017).
35. Shalev-Shwartz, S. & Ben-David, S. *Understanding Machine Learning: From Theory to Algorithms* (Cambridge Univ. Press, New York, 2014).
36. Ghafari, F. et al. Observing the ambiguity of simplicity via quantum simulations of an Ising spin chain. Preprint at <http://arxiv.org/abs/1711.03661> (2017).
37. Xu, F. et al. Experimental quantum fingerprinting with weak coherent pulses. *Nat. Commun.* **6**, 8735 (2015).
38. Garcia-Escartin, J. C. & Chamorro-Posada, P. SWAP test and Hong-Ou-Mandel effect are equivalent. *Phys. Rev. A* **87**, 052330 (2013).
39. Crutchfield, J. P. The calculi of emergence: Computation, dynamics and induction. *Physica D* **75**, 11–54 (1994).
40. Zambella, D. & Grassberger, P. Complexity of forecasting in a class of simple models. *Complex Syst.* **2**, 269–303 (1988).
41. Crutchfield, J. P. Between order and chaos. *Nat. Phys.* **8**, 17–24 (2012).
42. Chiuri, A., Greganti, C., Mazzola, L., Paternostro, M. & Mataloni, P. Linear optics simulation of quantum non-Markovian dynamics. *Sci. Rep.* **2**, 968 (2012).
43. Nagali, E. et al. Testing sequential quantum measurements: How can maximal knowledge be extracted? *Sci. Rep.* **2**, 443 (2012).

## Acknowledgements

We thank Raj B. Patel for helpful contributions. This research was funded, in part, by the Australian Research Council (project no. DP160101911), the Lee Kuan Yew Endowment Fund (Postdoctoral Fellowship), Singapore Ministry of Education Tier 1 grant RG190/17, Singapore National Research Foundation Fellowship NRF-NRFF2016-02 and NRF-ANR

grant NRF2017-NRF-ANR004 VanQuTe, and the FQXi Large Grant: the role of quantum effects in simplifying adaptive agents. F.G. acknowledges support by the Australian Government Research Training Program (RTP) scholarship. We acknowledge the traditional owners of the land on which this work was undertaken at Griffith University, the Yuggera people.

## Author contributions

F.G., N.T. and C.D.F. designed the experimental setup; F.G. and N.T. performed the experiment and analysed the data. C.D.F., M.G. and J.T. conducted the theory of the project, as well as contributing to the data analysis. G.J.P. played a significant role in the project conceptualisation, provided experimental assistance and oversaw all aspects of the project. All authors contributed to writing the manuscript.

## Additional information

**Competing interests:** The authors declare no competing interests.

**Reprints and permission** information is available online at <http://npg.nature.com/reprintsandpermissions/>

**Journal peer review information:** *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019