Photonic quantum information science for stochastic simulation and non-locality

Author:
Farzad Ghafari Jouneghani
M.Sc. Physics

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

Quantum Optics and Information Laboratory
Centre for Quantum Dynamics
School of Environment and Science

February 11, 2019
Abstract

Since its inception, quantum information processing (QIP) has been branched into many new directions. As well as well-known applications such as teleportation and metrology, researchers are beginning to investigate interdisciplinary areas such as quantum machine learning. One of the interesting areas is where quantum information science meets stochastic modelling, from the field of complexity science. Recently, it has been shown theoretically that, for simulating classical systems, quantum-assisted models and simulators are more efficient in terms of memory storage they require to do simulations, compared with classical computers. That is, for most systems, classical simulators demand an excessive amount of memory storage, while quantum simulators can do the same simulation with less memory.

The main focus of this thesis is on experimental realisation of quantum simulators that are capable of simulating stochastic processes with a reduced amount of memory. To implement the (nearly) exact simulation of stochastic processes using quantum simulators, it is essential to have low-noise state preparation, robust unitary operations, and high-precision read-out. These requirements, and the flexibility and precision of photonic quantum optics, make photonics the ideal system for developing the science of this new quantum advantage and for making strides towards its technological realisation.

In the context of stochastic simulation, I have experimentally studied three key problems that serve as stepping stones in advancing this new field. In the first experiment, an error-tolerant quantum simulator was designed, and realised to simulate a 1D Ising spin chain, using internal states that store less information than the corresponding classical approach. Furthermore, an interesting and fundamental phenomenon named the ambiguity of simplicity is witnessed. This is the inconsistency that we observe in the order of relative complexity of two systems when we change the simulators from classical to quantum. In the second experiment, a quantum simulator was built that can simulate classical processes for more than one step of the simulation at a time, storing the information from multiple steps coherently. In the third experiment, a new type of quantum memory advantage, which is based on the dimensionality of the memory register, rather than on information entropy measure, is demonstrated. This advantage is realisable in a single simulator, in contrast to previous works. Realising the dimensionality memory advantage in practical applications does not rely on running multiple simulators in parallel.
In the field of optical QIP, realising an ideal single-photon source has been a long-standing challenge. In another part of my PhD, I have worked on a source that aims to tackle some of the existing issues in this context. We built a source of high-quality entangled photons with a high heralding efficiency, which has the potential to be used in multi-photon experiments. This source was also essential to demonstrate a fundamental task in quantum non-locality, one-way steering, which was performed conclusively for the first time.
Statement of Originality

This work has not previously been submitted for a degree or diploma in any university. To the best of my knowledge and belief, the thesis contains no material previously published or written by another person except where due reference is made in the thesis itself.

Signed: Farzad Ghafari Jouneghani

Date: 11/02/2019
## Contents

Abstract

Statement of Originality

Acknowledgements

Acknowledgement of Papers included in this Thesis

1 Introduction

1.1 Theoretical background for stochastic simulation

1.1.1 System, model, and simulator

1.1.2 Computational mechanics

1.1.3 Markov machines

1.1.4 Statistical complexity

1.1.5 \( \epsilon \)-machine

1.1.6 Excess entropy

1.1.7 Quantum computational mechanics and quantum models

1.1.8 Why do quantum machines outperform classical machines?

1.1.9 Dimensional memory advantage

1.1.10 Quantum computational mechanics is a growing field

1.1.11 Quantum simulators

1.2 General methodology

1.3 High-performance source

1.3.1 Spontaneous parametric down-conversion

1.3.2 Characterising SPDC sources

1.4 Einstein-Podolsky-Rosen steering

1.4.1 Asymmetry in quantum steering; one-way steering

2 Witnessing the ambiguity of simplicity via quantum simulations of an Ising spin chain

2.1 Statement of contribution to co-authored published paper

2.2 Quantum simulations of an Ising spin chain

2.2.1 Witnessing the ambiguity of simplicity via quantum simulations of an Ising spin chain
## 3 Interfering trajectories in experimental quantum-enhanced stochastic simulation

3.1 Statement of contribution to co-authored published paper ........................................ 43

3.2 Quantum simulation over multiple time steps .............................................................. 45

3.2.1 Interfering trajectories in experimental quantum-enhanced stochastic simulation ........................................ 46

## 4 Conclusive experimental demonstration of one-way Einstein-Podolsky-Rosen steering

4.1 Statement of contribution to co-authored published paper ........................................ 57

4.2 One-way Einstein-Podolsky-Rosen steering ............................................................... 59

4.2.1 Conclusive experimental demonstration of one-way Einstein-Podolsky-Rosen steering ........................................ 59

## 5 Dimensional advantage in the simulation of stochastic processes

5.1 Statement of contribution to co-authored published paper ........................................ 71

5.2 Quantum memory advantage using a single simulator .............................................. 72

5.2.1 Dimensional advantage in the simulation of stochastic processes ........................... 72

## 6 Conclusion

80

## A List of Co-Authored Papers published during PhD not included elsewhere

84

## Bibliography

85
List of Figures

1.1 System, model, and simulator ................................. 4
1.2 Markov machine ................................................. 9
1.3 SPDC process ................................................... 20
1.4 Schematic representation of a collinear SPDC source ........ 22
1.5 HOM interference .............................................. 24
Acknowledgements

First, I would like to express my sincere gratitude to my supervisor, Geoff Pryde, for providing me with the opportunity to join his group, supporting me continuously during my PhD study and related research, sharing his immense knowledge, and for his patience and motivation. His guidance helped me through this research project and the writing of this thesis. I could not have imagined having a better advisor and mentor for my PhD study. Besides my supervisor, my sincere thanks go to my co-supervisor, Nora Tischler, for her daily support, guidance, comments, and encouragement, and for being extremely patient. Nora, thank you for answering almost all of my questions. Your broad knowledge always amazed me. And thank you for being such a sincere friend.

I thank my fellow lab mates: Joseph Ho, Morgan Weston, Kok Wei Bong, Shakib Daryanoosh, Alex Pepper, Raj Patel, and Sergei Slussarenko. I thank them for teaching me a lot of new concepts, providing stimulating discussions, for working together, and for all the enjoyable times we have shared during the last four years. They tried to turn a mathematician (as Sergei calls me) into an experimental physicist, well done! Joseph was a great friend and teacher during my early days in Australia, when things were not going very smoothly. Also, I thank my colleagues and friends in the Centre for Quantum Dynamics, as well as administration staff, Gabrielle and Nisha, for always being there to assist and support me.

I would like to thank my collaborators in Australia and Singapore: Howard Wiseman, Carlo Di Franco, Whei Yeap Suen, Jayne Thompson, and Mile Gu for their insightful ideas and proposals, which were the basis for most of my experimental work. Mile provided me with some of his unpublished works, which was very helpful in writing my thesis.

I would also like to thank two special teachers in my life, Mr. Salehi and Mr. Opera, my math and science teachers during middle school. My family and friends have always been great encouragement, I thank them for that. Last but not the least, I would like to thank Marzi, for sharing the most valuable asset of hers, life, with me during my PhD and my life in general. I apologise for having to sacrifice so many weekends and so much family time.

I acknowledge the support by the Australian Government Research Training Program (RTP) scholarship. I would like to acknowledge the traditional owners of the land on which this PhD was undertaken at Griffith University, Nathan campus, the
Yuggera people.
Acknowledgement of Papers included in this Thesis

ALL PAPERS INCLUDED ARE CO-AUTHORED.

Section 9.1 of the Griffith University Code for the Responsible Conduct of Research ("Criteria for Authorship"), in accordance with Section 5 of the Australian Code for the Responsible Conduct of Research, states:

To be named as an author, a researcher must have made a substantial scholarly contribution to the creative or scholarly work that constitutes the research output, and be able to take public responsibility for at least that part of the work they contributed. Attribution of authorship depends to some extent on the discipline and publisher policies, but in all cases, authorship must be based on substantial contributions in a combination of one or more of:

- conception and design of the research project
- analysis and interpretation of research data
- drafting or making significant parts of the creative or scholarly work or critically revising it so as to contribute significantly to the final output.

Section 9.3 of the Griffith University Code ("Responsibilities of Researchers"), in accordance with Section 5 of the Australian Code, states:

Researchers are expected to:

- Offer authorship to all people, including research trainees, who meet the criteria for authorship listed above, but only those people.
- accept or decline offers of authorship promptly in writing.
- Include in the list of authors only those who have accepted authorship
- Appoint one author to be the executive author to record authorship and manage correspondence about the work with the publisher and other interested parties.
- Acknowledge all those who have contributed to the research, facilities or materials but who do not qualify as authors, such as research assistants, technical staff, and advisors on cultural or community knowledge. Obtain written consent to name individuals.

Included in this thesis are papers in Chapters 2, 3, 4, and 5 which are co-authored
with other researchers. My contribution to each co-authored paper is outlined at the front of the relevant chapter. The bibliographic details and publication status for these papers including all authors are:


(Signed)  
Name of Student: Farzad Ghafari Jouneghani  
(Signed)  
Supervisor: Geoff Pryde  
(Signed)  
Supervisor: Nora Tischler


(Signed)  
Name of Student: Farzad Ghafari Jouneghani  
(Signed)  
Supervisor: Geoff Pryde  
(Signed)  
Supervisor: Nora Tischler


This paper is reproduced here under a copyright license agreement, with license number RNP/18/DEC/010576, between the American Physical Society (“APS”) and Farzad Ghafari under the terms and conditions provided by the American Physical Society.
and SciPris.


Appropriate acknowledgements of those who contributed to the research but did not qualify as authors are included in each paper.
Chapter 1

Introduction

In a world with limited memory, we simply cannot distinguish between \( \pi \) and a terribly close neighbour.  

---

Rolf Landauer

Quantum information processing has attracted immense attention during the last three decades. Historically, the basic idea of quantum computing was highlighted by Feynman who doubted classical computers’ capability to simulate a physical world that is governed by the laws of quantum mechanics (1, 2). He pointed out that there was no classical simulator that could repeat the results from some specific experiments, such as the Einstein-Podolsky-Rosen paradox (3). Since then, there has been an enormous effort to theoretically and experimentally develop a universal quantum computer that can solve problems, which are not easily solvable by classical computation. Moreover, the transistors that make up conventional computers are progressively shrinking in size such that they can show the effects of quantum mechanics. Therefore, if we have to deal with quantum effects, it is sensible to use them for a beneficial purpose and invest into a computer that can operate in the quantum regime (4).

However, quantum information processing (QIP) is not just about the originally conceived notion of simulating quantum systems. After Feynman, Deutsch developed theoretical foundations of a programmable computer that can simulate any physical process—a universal quantum computer (5). That general recipe led to some interesting and potentially useful quantum algorithms with a significant advantage over their classical counterparts. The most well-known are Shor’s algorithm (6) and Grover’s search algorithm (7). The former work provided an algorithm with which factoring the product of prime numbers can be performed in polynomial time, compared to the exponential cost of known classical algorithms (6). Grover’s algorithm, for finding an element in an unsorted database of elements, offered a reduction of required operations by a square root (7). Today, there are many areas of research and technology that are inspired to use the powerful tools of QIP. Some of the most important ones are quantum communication (8, 9), quantum cryptography (10, 11), quantum metrology
and sensing (12, 13), quantum annealing (14), quantum machine learning (15), linear systems (16), and so forth. Recently, a quantum advantage was found in the context of simulating classical stochastic processes (17), where a memory advantage in simulation is offered, rather than a computational speedup. This is the topic that I will pick up in the next section and throughout this thesis.

Along with pioneering research to expand the theory of QIP, quantum hardware was essential to build a quantum computer. The majority of efforts have been focused on implementing and manipulating the basic unit of quantum information, a qubit (18). To date, several technologies have been used to make quantum systems, such as photons (19–21), trapped ions (22–25), cold atoms (26, 27), superconducting circuits (28, 29), liquid and solid-state nuclear magnetic resonance (NMR) (30), and quantum dots (31, 32).

Among the mentioned platforms, single photons have some special features. They are fairly easy to produce and manipulate and, at the same time, they are very robust to noise. However, one of the biggest drawbacks of working with photons is that they do not directly interact with one another. Nonetheless, some promising proposals are expected to overcome this issue. For instance, there are proposals to perform measurement-based quantum computing (33) where fixed-depth quantum cluster states are constructed probabilistically and efficiently (34, 35). Another unique feature of photons is that they propagate fast, positioning them as the ultimate candidate for quantum communication, able to exploit existing networks of optical fibres. These features of single photons are why they are widely used to test both the fundamental aspects of quantum mechanics (36) and implement QIP protocols (37).

The experiments in this thesis explore the realisation of quantum complexity advantages (e.g. for memory in classical stochastic simulation) using photons as the quantum information carriers. Bulk optics is employed to implement optical QIP tools and techniques. We achieve a quantum complexity advantage by experimentally realising simulators that require less memory storage compared to the best classical simulators.

In our photonic implementation, some of the photons are lost because of imperfect optical devices and also the probabilistic nature of our implementations. In a practical implementation, one might argue that the lost photons are lost resources which should be considered when we discuss the memory advantage. In our experimental demonstration, we do not consider the lost resources for two reasons. First of all, in principle, there is no fundamental reason why we would be limited to photonic systems or to probabilistic gates. Second, our experiments are proof-of-principle implementations which show that the quantum memory advantage exists in principle. However, for final implementations, the total resources used could be carefully accounted for, in a similar way to observing an advantage in quantum metrology (38).
Chapter 1. Introduction

After introducing principles behind the quantum advantage in Chapter 1, in Chapters 2 and 3, I will present two experiments in the context of simulating classical stochastic processes. In Chapter 4, I will report a new single-photon source and a fundamental test on quantum steering. In Chapter 5, I will use the source to implement another quantum simulator with a memory advantage. Chapter 6 is designated to final remarks and the conclusion. I start with a background on simulating stochastic processes and its relationship to QIP.

1.1 Theoretical background for stochastic simulation

1.1.1 System, model, and simulator

To understand nature, we observe it over and over. That is how we form an explanation, a theory, based on the repeated pattern of the studied case. A successful theory is one that can predict the future of the studied system reliably, matching what happens in the real case. In principle, in the realm of classical physics, most systems can be deterministically predicted, as long as we have enough knowledge about them. For instance, using Newton’s laws, we can predict the ideal behaviour—velocity and position—of a massive object, starting with knowledge of its current position and velocity. Cases such as this admit a simple deterministic explanation. Here, we are concerned with stochastic processes. These are processes that are (partially) probabilistic—their deterministic explanation is so complex that we do not have access to the internal variables and interactions required to describe them. For example, to describe the behaviour of an object in the real world with friction, inhomogeneous particles etc., many more variables, such as airflow, must also be considered. Because we do not have detailed microscopic knowledge of the relevant variables, the system may appear to behave probabilistically in some regimes. A fair coin is another interesting example where, because of our lack of knowledge, it is almost impossible to predict whether heads or tails will show when it lands. In this thesis, we study stochastic processes.

To understand a complicated system, one approach is to ignore its complicated internal interactions and treat the system as a black box, where only its output is observed. Mathematical models and physical simulators are employed to replicate the system’s behaviour. In a partially probabilistic system, there are some patterns that happen over and over. Using these repeated patterns, one can predict the future of the system based on the past observation. In this thesis, I will refer to a model as being the theoretical description of the system. Given an encoded form of the system’s output information (or at least part of it), a mathematical model predicts the probability of all possible futures conditioned on this information. Moreover, I will refer to a simulator as being the experimental realisation of the model. A simulator, which implements the mathematical model, produces a sample that statistically matches the
observed data from the system in the black box. Accordingly, a faithful model (and its corresponding simulator) is defined as one that can exactly replicate the actual system’s statistics—the results from the system and simulator are samples from the same probability distribution (see Figure 1.1).

Given a system, there might be many models and their corresponding simulators that can replicate it faithfully. An obvious question is: which one of these is favoured? Practically, we prefer a model whose corresponding simulator requires the least amount of memory storage. That is, at a given time step of the simulation, we want to have to store as little of the previously generated information as possible, in order to generate the next internal state and output of the simulator. In other words, we want to store as little as possible of the past to generate the future of the process’ statistics (39–43).

Why are we looking for models with minimum memory requirements? First, because memory is a limited resource. This limitation is more noticeable when we are dealing with complex systems, where even the best models require a huge amount of memory for simulation, to the extent that the required memory is not even available. From a fundamental point of view, we can argue that the optimal model isolates the relevant information better. In a brute-force approach, all possible past observations are memorised so as to not miss any information that might be relevant to the future prediction. However, a more efficient model has a better understanding of the system, which leads to cutting some of the irrelevant information that does not affect future prediction; therefore, this model needs less information from the past. This sentiment is also echoed in the principle of Occam’s Razor (44), and Newton’s rules of reasoning: “We are to admit no more causes of natural things, than such as are both true and sufficient to explain their appearances.” (45).

In this thesis, the studied systems are examples of stochastic processes, and will be introduced in each relevant chapter. Following the existing terminology in Reference (17), I use the words ‘model’ and ‘simulator’ interchangeably, because simulator conveys the physical implementation of the model.

1.1.2 Computational mechanics

The ideas of defining a meaningful measure of memory requirements and finding the minimum memory for simulating a process were first introduced in the context of complexity science. This field of research, which is called computational mechanics (39,

---

1It should be noted that the definition of faithful simulators only considers the ideal case. However, imperfections exist in the real world. Even the best simulators (classical or quantum) are vulnerable to error, which means that they approximately implement the desired model of the stochastic process, and therefore approximately simulate the system. In the following chapters, we discuss some of these imperfections in more depth.
Chapter 1. Introduction

Figure 1.1: **System, model, and simulator.** a. We can consider a complicated system as a black box, ignoring its internal variables and interactions. The system produces some data that we observe. b. An abstract mathematical model defines a conditional probability, which gives the probability of all possible outputs, conditioned on each possible input. An input is what we refer to as a *past*—it is an encoding of the whole or part of what has been observed in the past. An output is what we refer to as the *future* of a system. It can be one of the possible configurations (data sets) that the system generates, and we wish to predict its probability. c. A physical simulator, which implements the mathematical model, encodes the input into a physical state (1). Then, by physically implementing the conditional probability, it samples a set of data. For a faithful simulator of the system, the sample and the observation are statistically indistinguishable (2). The model and simulator convey the same concept, where simulator emphasises the physical realisation.
Chapter 1. Introduction

41, 46, 47), quantitatively defined statistical complexity as the minimum required memory for modelling a process. In Subsection 1.1.4, we will see that the statistical complexity also quantifies structure in a stochastic process; the more complex the process is, the more memory of the past is needed to simulate it. The ideas of computational mechanics have been adopted in a variety of fields, such as statistical physics (48), self-organising processes in which systems become more highly organised over time (49), hidden Markov models (50), stochastic resonance (51), chaotic dynamics (41), anomaly detection in complex dynamical systems (52), neural spike trains (53), input-output processes (54), and thermodynamics (55, 56).

Computational mechanics has been a successful theory in the realm of classical information science. That is the field in which, in order to model a stochastic process, the memory and the processor are both classical. However, even the optimal classical models are wasteful—the optimal models require storing more information from the past than they actually use. Gu et al. (17) showed that for modelling classical stochastic processes, there are more efficient models if we use quantum information carriers and processors. The term ‘efficient’ relates to the amount of information being stored, not to the simulation time. It is worth mentioning again that, besides storing less memory from the past, memory-efficient models might help us to understand the workings of complicated stochastic processes, because they tend to isolate the relevant information more effectively.

Quantum computational mechanics and its applications have been developing very rapidly. Theoretically, it has been applied to Markovian Ising systems (57), real-valued stochastic processes (58), systems with long-range interactions (59), relative complexity of processes (59), rare-event sampling (60), input–output processes (61), and continuous-time stochastic processes (62). Moreover, it has been used to investigate fundamental questions about quantum interpretations (63) and causal asymmetry in predictive modelling (64). Besides applications of quantum models, developing more efficient models with less memory requirements has been a central research line (65–67). I will give a more detailed review of the relevant literature in Subsection 1.1.10. A major advance in this field has been the experimental implementation of quantum-enhanced simulators of stochastic processes, demonstrating the statistical complexity advantage and driving innovation in the field. In 2017, Palsson et al. (68) realised the first quantum simulator that simulated a stochastic process with less memory compared to the optimal classical simulator. In this thesis, I report three experiments that advance the task of stochastic simulation with quantum processors in several important new directions. Before delving into each project individually, I will provide a mathematical background that is useful for understanding all projects.
1.1.3 Markov machines

In this subsection, I introduce the basic tools that are needed to understand the predictive modelling of stochastic processes. Most concepts are borrowed from classical computational mechanics, where they were initially defined \((42, 46)\). Some of the notation, definitions, structure, and concepts in this subsection and subsequent ones (to 1.1.9) are heavily based on an unpublished review paper written by my collaborator Mile Gu and his colleagues (private communication).

A random variable is denoted by \(X\), with \(x\) as one of its possible configurations from the domain \(X\). The probability of \(X\) taking \(x\) is \(p_x := P(X = x)^2\). The Shannon entropy \((69)\) of this variable is given by

\[
H(X) = -\sum_x p_x \log p_x,
\]

with all logarithms in this thesis are taken to base 2. This entropy captures the informal uncertainty embedded in the random variable.

However, information is not only an isolated abstract concept; it is attached to a physical realisation. It can be a spin in a magnetic field, or a hole in a card \((70–73)\). Analogously, a random variable \(X\) is tied to a physical system \(\Xi\). There is a one-to-one map from the domain \(X\) to \(\Xi\)'s configuration space, \(\Omega\), with the probability \(p_i = p_x\) of \(s_i\) being the physical state representing \(x\). Accordingly, the entropy of \(\Xi\), \(H(\Xi)\), is equal to the entropy of random variable \(X\). If \(\Xi\) is a quantum system, its entropy is determined by the von Neumann entropy \((18)\):

\[
H(\Xi) = -\text{Tr}(\rho \log \rho),
\]

where \(\rho = \sum_i p_i |s_i \rangle \langle s_i|\) is the density operator of physical system \(\Xi\). If \(\{|s_i\}\) are mutually orthogonal states, \(H(\Xi)\) reduces to the Shannon entropy, \(H(X)\).

A discrete stochastic process is a dynamical system with a state sampled at discrete times \(t \in \mathbb{Z}\), with an outcome \(x_t \in \mathcal{X}\) at time \(t\) governed by the random variable \(X_t\). The behaviour of the system is characterised by a joint probability distribution, \(P = P(\vec{X}_k, \vec{X}_k^t)\), where \(\vec{X}_k = \ldots, X_{k-1}, X_k\) and \(\vec{X}_k^t = X_{k+1}, X_{k+2}, \ldots\) denote the random variables that respectively govern the statistics of past and future observations with respect to the time \(t = k\), which is the present time. The systems that we study in this thesis are stationary: \(P(X_{0:k}) = P(X_{t:t+k})\) for any \(t\) and \(k \in \mathbb{Z}\), where \(P(X_{0:k}) = P(X_0, X_1, \ldots, X_k)\). This implies that the system’s behaviour is time invariant, and we are free to pick the present time, \(t\). We choose \(t = 0\) to be the present time and we write \(P(\vec{X}, \vec{X}^t)\) instead of \(P(\vec{X}_t, \vec{X}_k^t)\). A simple example of a stationary process, known as the perturbed coin, is one where at each time step, the probability of a coin flipping

\(^2\)\(P(x)\) and \(p(x)\) denote the probability of \(x\).

\(^3\)Computational mechanics may also study non-stationary processes \((49, 74)\).
from tails (binary 0) to heads (binary 1) is given by $p$, flipping from heads to tails by $q$, staying in heads by $1 - p$, and staying in tails by $1 - q$.

For each past configuration of a random process, denoted by $\vec{x}$, the future behaviour of the system is governed by $P(\vec{X} | \vec{X} = \vec{x})$. A faithful model replicates the system’s future behaviour using the information from the past—that is, it samples a future configuration, $\vec{x}'$, from $P(\vec{X} = \vec{x}' | \vec{X} = \vec{x})$. Let us assume our model uses the brute-force approach. That means the model requires the simulator to memorise all information contained in $\vec{x}$. Therefore, to model a process $\mathcal{P}$, the physical system $\Xi$ with a one-to-one mapping from its configuration space $\Omega$ to all possible past configurations is required. This is usually inefficient. Consider a simple system such as a perturbed coin with $p = q = 0.5$. For memorising the evolution of the coin after $N$ steps, a system with $|\Omega| = 2^N$ is needed ($|\ldots|$ counts the number of elements a set contains). However, in practice, because the coin is random, observation of the past does not carry any information about the future.

What are other, more efficient, models? Computational mechanics answers this question in a systematic way. To do so, the first objective is to quantify the memory requirements for modelling. Computational mechanics approaches this issue by introducing the broadest model, called the Markov representation (42). The corresponding physical realisation of this model is known as a Markov machine (75). In this approach, the relevant information about $\vec{x}$ is encoded in an internal state of a physical system $\Xi$, such that systematic actions on $\Xi$ replicate the system’s future with correct statistics $P(\vec{X} | \vec{X} = \vec{x})$. The amount of information that $\Xi$ carries represents the required information to predict the system’s future.

A general Markov machine is represented by the pair $\mathcal{M} = (\Omega, M)$. As previously stated, $\Omega$ is the configuration space of the physical system $\Xi$, and $M$ is the physical process acting on $\Xi$ and an ancillary system $A$. The physical process is defined by probabilities $\{T^k_{ij}\}$, where $T^k_{ij}$ is the probability that the internal state will transfer from $s_i$ to $s_j$, while writing $x$ on the ancilla system. It is the ancilla that will be read out to give $x$. A representation $\mathcal{M}$ is called a faithful representation of process $\mathcal{P}$ if applying $M$ on $\Xi$ for infinite time generates the stochastic process $\mathcal{P}$. Note that the ancilla system always starts, at the beginning of each time step, in a fixed state, e.g. $x = 0$; therefore, it is memoryless—$H(A) = 0$ (42). Figure 1.2 provides a conceptual diagram of the Markov machine.

Since a Markov machine represents a stationary process, its dynamics are also invariant under time translation. If we consider the machine as a channel from past to future, the physical state of $\Xi$ is the only thing that carries information in this channel. Analogous to the process itself, at each time step $t$, the state of $\Xi$ is governed by the random variable $S_t$, where the sequence $(\vec{S}, \vec{S}')$ is a Markov chain; the variable $S_t$ only depends on $S_{t-1}$. This motivates us to define the stationary state of $\Xi$, which is the average state of the system after an infinitely long time. For example, for
Chapter 1. Introduction

FIGURE 1.2: Markov machine. The observation about the past is encoded in the relevant state of $\Xi$ (here, $s_i$) at present time $t_0$, where from the ancilla tape, the ancilla used is shown by bold red rectangles, carrying no information. To carry on the simulation, the physical process, $M$, acts on both $\Xi$ and the ancilla. Accordingly, in the next time step $t_1$, the internal state of $\Xi$ transits to $s_j$ and $x_1$ is written on the ancilla with probability $T_{ij}^{x_1}$ (1). The transition to $s_j$ is one of the possible options, where transition to $s_k$ (2), with the relevant probability, is another one. To continue the simulation beyond $t_1$, this procedure is repeated, using a fresh ancilla at each time step.
Chapter 1. Introduction

A quantum system the stationary mixture is given by $\rho = \sum_i p_i |s_i\rangle \langle s_i|$, where $p_i$ is the stationary probability of each state $|s_i\rangle$. The classical stationary state is defined in the same way, except the set \{s_i\} is an orthogonal basis.\(^4\)

We now have enough tools to quantify the amount of information that a Markov machine needs to faithfully replicate a system’s behaviour. The amount of resources that are required to store the internal state of $\Xi$ quantifies the required memory of the machine. There are different measures, of which information entropy is the most popular. The presentation statistical complexity\(^5\) of a model $\mathcal{M}$ is the information entropy of $\Xi$ (42). If the machine is classical, then it is given by

$$C(\mathcal{M}) = - \sum p_i \log p_i,$$

where $p_i$ is the stationary probability that $\Xi$ is in the state $s_i$.

In the limit of a very large number ($N$) of parallel classical simulations, the number of required bits to run them is given by $NC(\mathcal{M})$, which means the average required memory for each of the machines, in the independent and identically distributed (i.i.d.) limit, is $C(\mathcal{M})$. Another measure that we use is the information capacity of a $d$-dimensional system, $\log(d)$, which provides the upper bound for information entropy: $C(\mathcal{M}) < \log(\text{Dim}(\Xi))$, where $\text{Dim}(\Xi)$ refers to the dimension of $\Xi$ (41).

In computational mechanics, Shannon entropy is the one used to compare different representations—$\mathcal{M}$ is considered simpler than $\mathcal{M}'$ if and only if $C(\mathcal{M}) \leq C(\mathcal{M}')$.

It is worth mentioning that, in this thesis, we only consider non-oracular models. That means all information contained in $\Xi$ exclusively comes from the past observations, $\Xi$. In other words, these models use information only from the past to sample the system’s future. Such models are also referred to as causal models (41, 46).\(^6\)

1.1.4 Statistical complexity

The minimum presentation statistical complexity between all faithful causal models is simply called statistical complexity, i.e. $C_\mu = \min_{\mathcal{M}} C(\mathcal{M})$.\(^7\) Statistical complexity, therefore, captures the minimum amount of required information to faithfully replicate a system’s behaviour, conditioned on the assumption that $\Xi$ does not have access to information about the future not originating from the past.\(^8\)

---

\(^4\)Computational mechanics may also study non-Markovian processes (59, 66).

\(^5\)This terminology, presentation statistical complexity instead of representation statistical complexity, is mostly historically motivated (42).

\(^6\)This feature, exclusively using the past information, is guaranteed in synchronising representations. If $\lim_{k \to \infty} H(S_0|X_{-k}\Xi) = 0$, the representation $\mathcal{M}$ is synchronising. It is proven that the simplest causal model of $\mathcal{P}$ is always synchronising (64); therefore, the studied models here are limited to synchronising ones.

\(^7\)We use the subscript $\mu$ for historical reasons.

\(^8\)Statistical complexity is also a good measure of structure, which identifies how much structure (or pattern) exists in a stochastic process that can be used to replicate the system’s behaviour. There is a difference between disorder and statistical complexity, although both of them are described by the
1.1.5 \textit{c-machine}

The optimal classical model of a process \( P \) is defined as its simplest classical non-oracular representation \( (42) \)—the model that uses the least amount of information exclusively from the past to faithfully predict the future. This model is denoted by the tuple \( M = (f, \Omega, M) \) with \( C(M) = C_\mu \), where \( f \) is an encoding function \( f : \bar{X} \rightarrow \Omega \) such that \( P(\bar{X} | \bar{x} = x) = P(\bar{x} | \bar{X} = \bar{x}) \). As defined, \( C_\mu \) is the statistical complexity of \( P \). Generally, the tuple \( M = (f, \Omega, M) \) is known as a predictive model.\(^9\)

As previously stated, optimality is all about using the least memory. For the optimal model, there should be a specific approach to how the model maps \( \bar{x} \) to \( \Omega \) states, in contrast to the brute-force approach for example, which maps each of the past events to a distinct state. The optimal approach is to bundle relevant pasts into an equivalence class, and then instead of memorising all of the past events, just store equivalence classes which any given \( \bar{x} \) belongs to. The equivalence relation, between two past events, is defined by \( (41, 46) \)

\[ \bar{x}_1 \sim \bar{x}_2 \iff f(\bar{x}_1) = f(\bar{x}_2). \] (1.4)

Therefore, for the optimal model, elements of \( \Omega \) are special, and they are called causal states \( (41) \). That is, if \( M = (f, \Omega, M) \) is the optimal classical model of \( P \), then the elements of \( \Omega \) are causal states of \( P \). Computational mechanics formalises this in a smart way, such that there is no need to distinguish between two previous pasts (\( \bar{x}_1 \) and \( \bar{x}_2 \)), if and only if they lead to the same future statistics. That is,

\[ \bar{x}_1 \sim \bar{x}_2 \iff P(\bar{X} | \bar{x} = \bar{x}_1) = P(\bar{X} | \bar{x} = \bar{x}_2). \] (1.5)

Then \( \bar{x}_1 \) and \( \bar{x}_2 \) are causally equivalent. With this causal equivalence relation, the past configurations are classified into a set of causal states, \( S = \{s_i\}^{10}(41) \).

The optimal model for a process, which is called the \( \textit{c-machine} \) \( (41) \) for which the elements of \( \Omega \) are in a one-to-one correspondence with causal states of the process. Given a stochastic process \( P \), the \( \textit{c-machine} \) is identified by a predictive model \( (\epsilon, \Omega, M) \), where \( \Omega = \{s_i\} \) is the set of causal states of \( P \). The encoding function, \( \epsilon \), is

---

\(^9\)Because the models are synchronising, there exists the encoding function \( f \). In other words, by knowing only \( \bar{x} \), one can dial \( \Xi \) into the state \( s_{\bar{x}} \), and by repeated action of \( M \) on \( \Xi \) the statistically correct future of the system is generated. The ordered tuple \( (f, \Omega, M) \) is known as a predictive model. The information entropy of \( \Xi \) is interpreted as the amount of information that the model demands exclusively from the past to predict the future.

Each predictive model \( (f, \Omega, M) \) provides a benchmark for simulating the stochastic process \( P \). That is, we need a physical system \( \Xi \). The physical system is initialised in \( f(\bar{x}) \), then a physical map \( M \) acts on the initial state and an ancillary system \( A \). As a result, the ancilla is transformed to \( x' \) with probability \( T_{f(\bar{x})}^{x_1}(X_1 = x_1 | \bar{X} = \bar{x}) \) while the internal state of \( \Xi \) transits from \( f(\bar{x}) \) to \( f(\bar{x}_1) \), where \( x_1 = \Xi_{x_1} \).

\(^{10} \)In our papers, the causal states may also be denoted by uppercase \( S_i \).
defined such that \( c(\vec{x}) = s_k \) if and only if \( \vec{x} \in s_k \). The physical process, \( M \), is defined by the transition probabilities \( T_{jk} \) and produces the correct statistics. That is, after the past sequence \( \vec{x} \in s_j \), in the next step of the simulation, the machine will emit \( x \), and transfer from past \( \vec{x} \) to \( \vec{x} x \), such that \( \vec{x} x \in s_k \). It is proven that the \( \epsilon \)-machine is the optimal and unique (up to relabelling causal states) causal model. Because it is the optimal model, its presentation statistical complexity gives the statistical complexity of the process \( P \):

\[
C_\mu = H(S) = -\sum p_i \log p_i, \tag{1.6}
\]

where \( S \) is the random variable that governs the internal states of \( \Xi \), which are the causal states. The probability of each causal state is given by \( p_i = P(\vec{x}) \). Note that \( C_\mu \) is a property of the process \( P \), capturing the minimum causal structure (from the past) required for understanding the probabilistic future of the process.\(^{11}\)

### 1.1.6 Excess entropy

\( \epsilon \)-machines are the optimal classical models, but are they the ideal models? Do they still waste any information? In fact, although they are classically optimal, most of the time they are very wasteful. Recall the assumption about predictive models: they only use the information observed in the past, not oracular information. The maximum amount of information that can be extracted from the past to predict the future is the mutual information between past and future, which is called the excess entropy \((39, 78)\).

The excess entropy of the process \( P = P(\vec{X}, \vec{X}) \) is

\[
E = I(\vec{X}; \vec{X}) = H(\vec{X}) + H(\vec{X}) - H(\vec{X}, \vec{X}), \tag{1.7}
\]

where \( I(\vec{X}; \vec{X}) \) is the mutual information between past and future. For the cases where \( C_\mu > E \), even the optimal classical model is wasteful, and this is true for most stochastic processes \((79)\). The gap between \( E \) and \( C_\mu \) is attributed to crypticity, \( \xi = H(S_0|\vec{X}) \).\(^{12}\)

\(^{11}\) Although the \( \epsilon \)-machine gives a systematic way to evaluate \( C_\mu \), other algorithms exist for finding causal states and therefore estimating statistical complexity from the observed data. To do so, one starts from the assumption that the process is simple (only needs one causal state), and then iteratively adds new causal states as required by the model \((47, 76, 77)\).

\(^{12}\) The interpretation of crypticity is rather subtle. Crutchfield et al. \((79)\) described crypticity as “the difference between observed information (\( E \)) and a process’s stored information (\( C_\mu \))”. The question is: why does a process store more information than can be observed? Non-zero crypticity implies that for an \( \epsilon \)-machine of \( P \), which is initialised in the state \( s_0 \) at \( t = 0 \), regardless of how long the output statistics of the machine are observed, it is impossible to deduce what the initial causal state of the machine was. The cryptic order, \( k \), of a stochastic process is the number of steps into the past that the observer of the \( \epsilon \)-machine needs to look, in order to deduce its current causal state. A process with non-zero cryptic order has non-zero crypticity, and vice versa. Therefore, the machine stores information that is more than what is needed to predict the future, and so some of the information is wasted. The required information for prediction is \( E \), while \( C_\mu \) is the required information for synchronised modelling. Different interpretations of crypticity are discussed in References \((80, 81)\).
To sum up, even the optimal classical models, \( \epsilon \)-machines, are wasteful. They use excessive information about the past—more than is observable in the future statistics. This also implies that some of the information is erased during the simulation, which makes them thermodynamically irreversible (56, 82).

1.1.7 Quantum computational mechanics and quantum models

It was not until very recently that quantum information processing met computational mechanics. The main motivation was to find predictive models that need less memory from the past to faithfully predict a system’s behaviour. In 2012, Gu et al. theoretically proved that for any \( \epsilon \)-machine with \( C_\mu > E \), there exists a quantum model with entropy \( C_q \), where \( C_c > C_q \geq E \) (17). Thus, quantum encodings provide a saving in information storage for stochastic simulation.

Similar to classical computational mechanics, there exists a quantum counterpart of predictive models. A quantum predictive model is defined by a tuple \( Q = (f, \Omega, M) \), where \( f : \vec{X} \rightarrow \Omega \) is an encoding function that maps the past \( \vec{x} \) to \( |s_i\rangle = |s_i\rangle \), one of the discrete internal quantum states of the physical system, \( \Xi \), from set \( \Omega \). \( M \) is a quantum process. \( Q \) is the quantum model of the process \( \mathcal{P} \), where the action of \( M \) on the initial state, \( f(\vec{X}) \), generates an output \( x \) with probability \( P(X_1 = x|\vec{X} = \vec{x}) \) and transits the state \( f(\vec{x}) \) to \( f(\vec{x} x) \). Note, in contrast to the classical definition, here \( \Omega \) is a set of quantum states, and \( M \) is a quantum process. One can infer that classical models are a special case of quantum models where all states in \( \Omega \) are mutually orthogonal.

The presentation complexity of a quantum model is given by the information entropy of the physical system:

\[
C_q(Q) = -\text{Tr}(\rho \log \rho),
\]

where \( \rho = \sum_i p_i |S_i\rangle \langle S_i| \) is the stationary mixture state, and \( p_i \) is the stationary probability of each state; \( p_i = P(f(\vec{x})) \). The stationary state is the mixture of the pure states, \( \{|s_i\rangle\} \), which implies the average state of the memory system evolved for an infinitely long time\(^{13} \). The operational meaning of the \( C_q(Q) \) is more or less the same as \( C(M) \): in the case of an asymptotically large number \( N \) of parallel simulators (i.i.d. case), it is possible to use Schumacher compression (83) to compress the required information from the past into \( NC_q(Q) \) qubits. This means the average memory usage of each simulator is \( C_q(Q) \), which could be lower than the classical average \( C(M) \). Similarly, for the dimensionality of the system, \( \log(\text{Dim}(\Xi)) \) is an upper bound of \( C_q(Q) \).

The optimal quantum model, in analogy to the classical case, is called the quantum \( \epsilon \)-machine, \( Q = (\epsilon_q, \Omega, M) \), where \( C_q(Q') \geq C_q(Q) \) for all faithful quantum

\(^{13}\)For classical models, the stationary state is defined in the same way; however, the internal states of the memory are orthogonal.
models $Q'$ of $P$. Moreover, the **quantum statistical complexity** is defined as $C_q = \min(C_q(Q))$, where \{ $Q$ \} is the set of faithful predictive models for $P$. It is obvious that, because the quantum $\epsilon$-machine is the optimal quantum predictive model, its presentation statistical complexity is the quantum statistical complexity.

By analogy with the classical causal state $s_i \in S$, there exists a quantum causal state for each classical one. Note that the number of causal states depends on the process, not the predictive model. The quantum causal state is defined by (17):

$$|s_i\rangle = \sum_{x \in X} \sum_{k=1}^{|S|} \sqrt{T_{ik}} |x\rangle |k\rangle,$$

with the encoding function $f(x) = |s_i\rangle$ if and only if $x \in s_i$. The transition probabilities $T_{ik}$ are the same probabilities as for the classical $\epsilon$-machine. That is, if the machine is initialised in the state $f(x) = |s_i\rangle$, then with the probability $T_{ik}$ it emits $x$, and the internal state of the machine transits to $f(x x) = |s_k\rangle$. (1.9)

For Markovian processes, the knowledge of the very last outcome of the process, $x_0$, is enough to determine the causal states. The processes we study in this thesis are Markovian, where $|X| = |S|$ and \{ $|x\rangle$ \} and \{ $|k\rangle$ \} coincide, with causal states simplify to:

$$|s_i\rangle = \sqrt{\sum_{k=1}^{|S|} T_{ik}} |k\rangle.$$ (1.10)

### 1.1.8 Why do quantum machines outperform classical machines?

Why do quantum machines outperform classical ones? The overlap of the internal states of the machines explains it. Let us assume $Q = (f, \Omega, M)$ is a quantum model of the process $P$, with encoding function $f(x) = |\psi_i\rangle$. Then

$$|\langle \psi_i | \psi_j \rangle| \leq \sum_x \sqrt{P_i(x) P_j(x)}.$$ (1.11)

This indicates that the overlap of the two internal states must be at least as distinguishable as the future statistics they correspond to. In the classical machine, $|\langle \psi_i | \psi_j \rangle| = 0$ for all $i$ and $j$. However, for Markovian processes, the causal states in Equation (1.9) satisfy $|\langle s_i | s_j \rangle| = \sum_x \sqrt{P_i(x) P_j(x)}$. Intuitively, this tells us that when the Markovian process is not completely random, $\sum_x \sqrt{P_i(x) P_j(x)} \neq 0$, the quantum model uses non-orthogonal states as its internal states and still samples the correct statistics. This is in contrast to the classical model which uses orthogonal states. Consequently, the von Neumann entropy of the quantum state $\rho$, $C_q$, will be less than the Shannon

---

14 As previously stated, optimal models are limited to models with a one-to-one correspondence between the causal states and elements of $\Omega$ (57).
entropy of the classical causal states, $C_\mu$ (84). In this thesis, the case studies are Markovian\textsuperscript{15}, therefore we use the construction introduced in Reference (17)\textsuperscript{16}.

The quantum memory advantage exists because of the non-orthogonality of the causal states—they may be expressed as coherent superpositions of orthogonal states. To have a practical simulator, this coherence should be maintained throughout the entire simulation. For example, we do not want to collapse the memory system to a classical state and reprepare it at each simulation step, because then the entropy of the memory rises to the classical level at each step. To maintain the physical system’s internal entropy, which is the quantifier of the required memory, the quantum process should be unitary. That is one of the reasons why the process, $M$, is a crucial concept in predictive modelling. In classical machines, a transition matrix, formed by transition elements $T_{ik}$, is adequate to describe the classical process. However, this is not immediately obvious for quantum machines. In Reference (67), Binder and co-authors introduce a systematic approach to find a unitary, $U$, fulfilling certain criteria for quantum machines, which simulates a discrete-valued stationary process. Their work is important because it provides the means to make a quantum machine which generates the correct statistics of the stochastic process, and simultaneously does not increase the internal entropy to the classical level. (This is in contrast to Reference (66), where at each step of the simulation the quantum memory needs to be measured and collapsed to classical information, increasing the entropy to the classical level.) Another important aspect of their work is that, in their procedure, the internal state of the machine stays in a Hilbert space of size $S$, instead of growing with cryptic order of a process, $k$. This is important for practical applications, where we would like to simulate processes with a large cryptic order.

1.1.9 Dimensional memory advantage

Here, I elaborate on the dimensionality of $\Xi$ as an alternative measure of required memory for simulation. As discussed, historically, most of the research in computational mechanics focused on information entropy (Shannon and von Neumann) as the measure for comparing models. However, achieving the quantum memory advantage with these measures is only practical in the limit of $N \gg 1$ parallel simulators, which is not feasible with current quantum technology and might also not be desirable. Another important measure of memory, which is scale-independent and feasible

\textsuperscript{15}The examples studied in this thesis are Markovian with cryptic order 1.

\textsuperscript{16}What about non-Markovian processes? Recently, Mahoney et al. (66) introduced an algorithm for implementing a general form of quantum machine, the q-machine, which goes beyond the construction in Reference (17) and Equation (1.9). The form that they introduced for quantum states saturates Equation (1.11), regardless of whether the process is Markovian or non-Markovian. However, they also argue that, for a q-machine, the extra quantum memory advantage comes at the cost of missing some predictive knowledge. The idea of the q-machine also paved the path for calculating the presentation statistical complexity for processes that have large cryptic order (85).
Chapter 1. Introduction

with any number of simulators, is referred to as topological complexity \((41)\). In classical computational mechanics, it is defined by:

\[
D_\mu := \log(\text{Dim}(S)),
\]

and the quantum version is given by

\[
D_q := \log \left( \text{rank}(\rho) \right).
\]

The topological advantage is achievable using a single simulator, because the classical information carrier is a physical system with higher dimension than the Hilbert-space dimension of the quantum memory. There exist processes where \(D_q \leq D_\mu\). One example is the process studied in Chapter 5 of this thesis, where we have experimentally demonstrated a dimensional quantum memory advantage. Our work is not an isolated example. In Reference \((64)\), some other instances with dimensional advantage are theoretically studied. Moreover, in References \((86)\), it has been shown that models having the ability to tune quantum phases, phase-enhanced models, can provide a dimensional advantage.

1.1.10 Quantum computational mechanics is a growing field

One of the unsolved challenges in the field of quantum computational mechanics is to prove the optimality of quantum models—to prove that a specific quantum model requires the lowest amount of memory for simulation, compared to all quantum models. This is in contrast to the classical \(\epsilon\)-machine which is the provably optimal classical model\(^\text{17}\). There is no systematic approach to this non-trivial problem yet. There are a few examples for which the optimality has been proven. One of them is the 1D Ising spin chain \((57)\), where the optimality of quantum \(\epsilon\)-machine for modelling this system is proven. In Reference \((64)\), Thompson and co-workers found the quantum optimal model for the perturbed coin process, and also a time-reversed version of it called the heralding coin. One of the reasons for the subtlety of proving the optimal quantum model is that the von Neumann entropy depends also on the phase of quantum states \((87)\). This phase dependence of entropy was not considered in developing most of the initial models. However, recently there has been an investigation of phase-enhanced quantum models \((86)\). The theory suggests that there is a competition between two measures of memory advantages, information entropy and topological complexity, when it comes to quantum modelling with phase enhanced models. That is, the phase can be tuned to minimise the von Neumann entropy; however, this is not always the phase that minimises topological complexity \((88)\).

\(^{17}\) Note that quantum optimality is different to the quantum advantage in memory usage, where quantum models outperform optimal classical models.
Quantum computational mechanics has been developing in several other directions. In Reference (57), the Ising spin chain was theoretically studied as an interesting physical process where the divergence between the required quantum memory and the classical entropy is demonstrated. In Reference (89), the same system was investigated to theoretically compare the relative order of complexity of two processes, using classical and quantum simulators. The experiment presented in Chapter 2 is based on these two works. Ising systems with long-range interaction are studied in Reference (59), where it is shown that there is potentially an unbounded quantum memory (information entropy) advantage for simulating these chains. The application of quantum modelling is also adopted to study more efficient approaches to sample rare events such as earthquakes and stock market crashes (60). Continuous-time processes, where \( t \) is a continuous index, also show an unbounded quantum memory advantage when simulating the future of these processes (62). Another example of quantum unbounded advantage is theoretically studied in the predictive modelling of real-valued stochastic processes, where evaluation of a real variable to high precision is modelled (58). An input-output process is a system where at each time step, \( t \), it takes an input governed by a random variable \( X_t \), and emits an output from another random variable \( Y_t \); these processes can be used for analysing systems interacting with an environment. Quantum models for discrete input-output processes are studied in Reference (61). Recently, there has also been an effort to connect many-body systems and their common formalism, such as matrix product states, to stochastic modelling (90).

1.1.11 Quantum simulators

So far, I have mostly described the theoretical ideas of classical and quantum models. One of the essential parts of simulating stochastic processes is the simulator—the physical realisation of the model. Put simply, we need to run our models on a computer.

On the experimental side, the first quantum \( \epsilon \)-machine experiment, in Reference (68), implemented a single-step quantum simulation of the perturbed coin. In that experiment, both classical and quantum simulators were realised, and it was shown that for a single step of the simulation, the quantum statistical complexity was by far less than the classical statistical complexity. The experiment was the first one of its type and opened the door to the study of quantum-enhanced stochastic simulation experiments.

In the experiments reported in this thesis, I tackled a range of important advances in quantum \( \epsilon \)-machines. The perturbed coin is a process mostly studied in computational mechanics (17) but may be less familiar to physicists. We demonstrated how to do the stochastic simulation in a physically motivated problem. After the theory work of Reference (57), we built a simulator capable of simulating a 1D Ising system
Chapter 1. Introduction

over one step, both classically and quantumly. Besides simulating a physical system with a statistical complexity advantage, we developed an approach which makes our simulator more error-tolerant. Moreover, we investigated an interesting phenomenon called *ambiguity of simplicity*, where the relative complexity of two processes is compared when the simulator is swapped from classical to quantum. We experimentally showed that the relative complexity is not fixed, and it might change if the simulator is changed. The results of this work are presented in Chapter 2.

Both of the above experimental works (68, 91) simulate the processes for only a single step. For quantum simulators to be practically applicable, it is necessary to go beyond a single-step simulation. In a multi-step simulation, realising the quantum unitary process and maintaining the quantum coherence and memory advantage are crucial. The theory proposal in References (17, 57) proposed that for each extra simulation step an extra ancilla qubit be employed. However, instead of adding extra single photons as ancillas, in our experiment we used a high-dimensional single photon encoding using different degrees of freedom. Our method needs the control, manipulation, and interference of high-dimensional quantum states, which is performed with a very high quality. The details of this work are presented in Chapter 3.

In the experiments of Chapters 2 and 3, the information entropy measure (statistical complexity) was used to compare different models and their memory requirements. One of the drawbacks of using this measure is that to achieve a practical memory advantage, there is a need for an asymptotically large number of simulators running in parallel. Currently, this is not experimentally possible, and might not even be desirable. Motivated by this issue, we implemented a quantum simulator capable of simulating a process with a dimensional quantum memory advantage. To perform the experiment, we developed a novel design that made the simulation possible in a four-photon circuit. The results of this work are presented in Chapter 5. To experimentally implement the quantum dimensional advantage, we had to overcome some experimental challenges, one of which was building a high performance four-photon source. In Sections 1.3 and 1.4, I will explain the details of this source and another task it was used for.

1.2 General methodology

Here, I mention very briefly several techniques that I have used in the experiments reported in this thesis. I use single photons as the carriers of information. In Section 1.3, a short background is given on the general idea of spontaneous parametric down-conversion (SPDC) sources. The photons are transferred either in free space or via single-mode optical fibres. To encode a qubit (18) on a photon, different degrees of freedom, including polarisation and spatial modes (path), have been used. Conventionally, horizontal (H) and vertical (V) polarisation states are considered as basis
states for a polarisation qubit (92). For realising higher dimensional states (higher than two), we use time-bin and polarisation modes simultaneously to generate a hybrid state (93–105). Single-qubit operations on the polarisation degree are performed by using half-wave and quarter-wave plates (92). To do the same type of operations on a path qubit, we use a non-polarising beam splitter or interferometers (19, 37). Polarisation-path controlled unitary operations on a state are implemented by polarising beam splitters.

Two-photon, or multi-photon, operations can be challenging because photons do not directly interact. In our case, we realise the two-photon gates via quantum interference and post selection—a technique often called ‘measurement-induced nonlinearity’ (106). I implement the general form of a controlled-Z gate, as the base for the controlled-NOT gate, via a probabilistic scheme and post selection of photon coincidences (107). This is made possible using partially polarising beam splitters, which transmit H polarisation completely and V partially (107).

To detect photons, I use avalanche photodiodes in two of the experiments, where having extremely high detection efficiency does not play a significant role. However, to have higher efficiency, especially at telecom wavelength, superconducting nanowire single-photon detectors (SNSPDs) are used (108). Signals from the detectors are sent to fast electronics, counting cards, and then are analysed on computers. In order to reconstruct quantum states, a complete set of projective measurements is carried out to perform a quantum state tomography (92). Quantum process tomography is performed to characterise an unknown physical operation in a quantum circuit (109).

1.3 High-performance source

Each quantum information processing task requires a physical system that carries quantum information. In the field of optical quantum information, which borrows many well-developed techniques from classical optics, single photons are one of the main systems on which the quantum information is encoded. An ideal single-photon source should emit only a single photon at a time, on demand, with a high generation rate, where photons are well-defined in the relevant spectral, spatial, and temporal modes. The other critical requirement is that the photons should be identical—photons from a single source or multiple sources should be indistinguishable (21).

There are different approaches to generating single photons (38, 110–118); among them, spontaneous parametric down-conversion (SPDC) and quantum dots are most widely used. Quantum dots (119–122) are advancing rapidly as a strong candidate to produce the ultimate on-demand single photon sources (123). However, poor collection efficiency from the dots, and low-visibility quantum interference between single
photons coming from different dots, still remain as challenges. There are some tech-
niques, such as positioning dots in a micropillar (115) and multiplexing (124), which
are being adopted to bypass these issues, but the problems have not been fully re-
solved yet. On the other hand, SPDC sources can produce very high-quality indistin-
guishable photons. However, the basic SPDC configuration is probabilistic (not on-
demand), which makes them difficult to use for large-scale experiments. There have
been some proposals, such as either spatial or time multiplexing and then switching
to a single-mode (125-130), to overcome this challenge. Recently, some promising re-
sults were published (131). Until the ideal source has been realised, both technologies
continue to play major roles in developing optical quantum technologies.

In the experiments reported here, which need up to four photons, SPDC sources
(without multiplexing) serve our needs perfectly. They have sufficiently high count
rates with a superb quality of quantum states (38); therefore, we use this technology
in bulk optics as our photon source.

1.3.1 Spontaneous parametric down-conversion

Spontaneous parametric down-conversion is a three-wave mixing process where a
non-linear crystal with $\chi^{(2)}$ non-linearity is pumped with a pump laser with frequency
$\omega_p$ (132–134). It is a parametric process, meaning that input and output fields are
proportional. A pump photon gets absorbed and two daughter (signal and idler with
frequencies $\omega_s$ and $\omega_i$, respectively) photons are generated, as shown in Figure 1.3a.
The bi-photon state, in terms of photon numbers in daughter modes, is described
by

$$|\phi\rangle \approx |0\rangle_s |0\rangle_i + \eta |1\rangle_s |1\rangle_i + \eta^2 |2\rangle_s |2\rangle_i + ...,$$  (1.14)

where $|n\rangle_{s(i)}$ represents $n$ photons in the signal (idler) mode. For small $\eta$, this equa-
tion can be approximated by the state $|0\rangle_s |0\rangle_i + \eta |1\rangle_s |1\rangle_i$. This means that there is a
probability $\approx \eta^2$ with which a single photon is generated in each of the signal and
idler outputs. To deal with this probabilistic nature, detection of one photon is used
to herald the existence of the other one—SPDC sources are commonly known as her-
alded photon sources. We can increase $\eta$ by boosting the pump power; however, it
comes with the unwanted raised probability of having more than a single photon in
each arm (136).

In SPDC processes, energy and momentum conservation lead to frequency-matching
and phase-matching conditions, respectively (137). The former is $\omega_p = \omega_s + \omega_i$ and
the latter is $k_p = k_s + k_i$, where $|k| = \frac{\lambda}{2\pi}$ is the wave number, as shown in Figure 1.3b-
c. Chromatic dispersion—frequency dependence of the phase velocity—causes the
photons generated in different locations in the medium to be out of phase. However,

---

18 The state is approximated because it is not normalised.
the phase-matching condition guarantees a proper phase relation between the different SPDC events along the non-linear crystal, and it is crucial for achieving an effective non-linear interaction in the crystal.

Achieving the phase-matching condition can be challenging. One of the most common techniques for non-linear phase-matching is to exploit polarisation in a birefringent crystal, to cancel the phase mismatch $\Delta k = k_p - k_s - k_i$. In a birefringent material, different polarised waves are subject to different refractive indices and, therefore, different phase velocities (137). Therefore, depending on the polarisation of the input and output waves, various types of phase-matching conditions are defined. In our case, the phase-matching is type II, where signal and daughter photons have orthogonal polarisations, as depicted in Figure 1.4 (137).

From a technical point of view, polarisation phase-matching was conventionally performed using one of two techniques: i) angle tuning, a precise angular orientation of the crystal, and ii) crystal temperature tuning. However, even with these techniques, phase-matching could be very challenging and also limited in terms of the range of feasible wavelengths, angles, and temperatures, depending on the properties of the crystals. Moreover, in this approach, there is also a limitation on using long crystals to increase the SPDC conversion rate. For long crystals, spatial and temporal walk-offs negatively affect the SPDC conversion efficiency and the quality of the generated photons. Furthermore, angle-tuned phase-matching typically gives undesirable spatial mode shapes, which are not compatible with single-mode fibres.

When achieving $\Delta k = 0$ for desired wavelengths is unlikely or very difficult, there is another approach to circumvent limitations, called quasi-phase-matching (138, 139).
Chapter 1. Introduction

Figure 1.4: Schematic representation of a collinear SPDC source. This figure shows the side-view of a ppKTP crystal cut for collinear type-II phase-matching. In this setting, signal and idler beams have orthogonal polarisations. In our case, the pump is polarised horizontally, while signal and idler are polarised horizontally and vertically, respectively. The direction of the non-linearity is switched to compensate the phase mismatch from the earlier half-grating. The crystal length, $L$, and poling distance, $\Lambda$, play a significant role in defining the source characteristics, which will be discussed in Subsection 1.3.2.

The idea is to allow for some phase mismatch over a region in the crystal, and then reverse the direction of the mismatch for the next region. This is achieved by spatially modulating the non-linearity of the crystal. One example is the periodically poled crystal (with length of $L$), shown in Figure 1.4. The grating period, $\Lambda$, is given by $\frac{2m\pi}{|\Delta k|}$, and $m$ is the grating order, which is 1 in our case. In the source that we have built, we use a periodically-poled potassium titanyl phosphate (ppKTP) crystal, which is cut for a type-II collinear phase-matching (in collinear phase-matching, wave vectors of all contributing waves have the same direction).

1.3.2 Characterising SPDC sources

There are certain key measures used to evaluate the performance of an SPDC source. The heralding efficiency (140) is defined as the number of two-fold coincidences divided by the number of singles in the heralding arm:

$$\varepsilon_A = \frac{C_{AB}}{S_B}, \quad (1.15)$$

where $C_{AB}$ is the coincidence counts between detectors $A$ and $B$, and $S_B$ is the number of singles in detector $B$ (see Figure 1.3). The heralding efficiency (based on Klyshko’s definition (140)) includes loss, coupling efficiency, and detection efficiency. Having high heralding efficiency is critical in many experiments, such as unconditional quantum metrology (38), loophole-free Bell non-locality tests (141), and detection-loophole-free Einstein-Podolsky-Rosen steering (142), among others. The rate of the coincidence events (per second and per unit of pump power) defines the brightness of the source.

One of the main reasons that collinear sources are more effective than non-collinear
sources is that the collection efficiency from collinear sources is higher. In the non-collinear design, light is collected from a small portion of the generated cone, into a single-mode fibre with a Gaussian mode. This poor mode-matching leads to a limited brightness and heralding efficiency \((110)\). In collinear sources, the pump and down-conversion photon beam waists can be engineered in a way that their transverse modes best match single-mode fibre, and increase the heralding efficiency, while maintaining reasonable brightness \((143)\). However, it is known that the beam waist sizes (lens configurations) that maximise heralding efficiency are not the same as those that maximise the brightness \((144, 145)\).

The purity of the generated quantum state is another key property of photon sources, and determined by \((18)\)

\[
purity(\rho) = \text{trace}(\rho^2),
\]

where \(\rho\) is the density matrix of the quantum state. For a \(d\)-dimensional quantum state, the purity can vary from \(\frac{1}{d}\) for the maximally mixed state, to 1, for a pure state. Note that to evaluate the purity, we either reconstruct the density matrix via tomography for a state encoded in polarisation, or use the Schmidt decomposition of the joint spectral amplitude\(^{19}\) for estimating spectral purity \((145)\). When I report a figure for purity, it is the former, unless I specifically mention that it is a spectral purity.

Photons do not exert force on each other, unlike other particles such as electrons with Coulomb forces acting between them. However, two photons can interfere because they are bosons and so they tend to cluster if they are made indistinguishable. Quantum interference is essential for most of the optical quantum information protocols. Two-photon interference, which is known as Hong-Ou-Mandel (HOM) interference \((146)\), is used in most of our experiments. Two-photon interference between photons from different down-conversion events (e.g. separate sources) is called independent HOM interference, whereas the interference between photons form the same down-conversion event is called dependent HOM interference \((147)\). The visibility\(^{20}\) of the HOM interference is given by \((148)\)

\[
V = \frac{C_{\text{max}} - C_{\text{min}}}{C_{\text{max}}},
\]

where \(C\) is the number of coincidences between detections at two output modes of a 50:50 beam splitter; see Figure 1.5. The number of coincidences changes by varying the relevant time delay between the wave packets of two single photons. The maximum number of coincidences is \(C_{\text{max}}\), while the minimum is \(C_{\text{min}}\).

To have a perfect HOM interference visibility, the individual photons exiting the

\(^{19}\)The joint spectral amplitude is defined on the next page.

\(^{20}\)The definition of visibility, here, is different from the single-particle interference visibility.
Chapter 1. Introduction

beam splitter must be indistinguishable in every degree of freedom, including polarisation, spatial, spectral and temporal ones. Coupling photons to a single-mode fibre (with its Gaussian mode) makes them spatially indistinguishable. A linear polariser can be used as a polarisation filter, along with a local polarisation rotation on one of the photons in a type II configuration, to achieve indistinguishability in polarisation. For spectral indistinguishability, besides using bandpass filters, the SPDC sources can also be designed in a way that they generate spectrally indistinguishable photons.

To achieve a perfect independent HOM interference visibility, the two-photon state entering the beam splitter must be pure. Photons generated from SPDC sources may be spectrally correlated. As a result, knowing the spectral distribution of one photon reveals some information about the other photon. When two photons are strongly correlated, performing a measurement that does not resolve the spectral information on one of the photons leaves the other in a mixed state. Therefore, heralded photons from different sources do not interfere (independent HOM) with perfect visibility, unless the bi-photon states of each source are separable in frequency. Let us assume a model for the frequency distribution of the bi-photon state:

\[ |\phi(\omega_s, \omega_i)\rangle = \int_{0}^{\infty} d\omega_s d\omega_i f(\omega_s, \omega_i) g^\dagger(\omega_s) g^\dagger(\omega_i) |0_s 0_i\rangle, \]  

where \( g^\dagger(\omega_s) g^\dagger(\omega_i) \) is the creation operator in the signal (idler) mode \((110, 145, 149)\). The frequency distribution of the bi-photon state is determined by the weight function \( f(\omega_s, \omega_i) \), which is known as the joint spectral amplitude (JSA). The ideal source would be one with a factorable JSA. In other words, \( f(\omega_s, \omega_i) \) can be written as a product of two functions, \( f_1(\omega_s) \), the spectral distribution of the signal, and \( f_2(\omega_i) \), the spectral distribution of the idler\(^{21}\). Considering that the temporal distribution is

\(^{21}\)If \( f_1(\omega_s) = f_2(\omega_i) \), the photons are indistinguishable, which is known as degenerate source. However, having indistinguishable photons is not restricted to this condition.
Chapter 1. Introduction

the Fourier transform of the spectral distribution (149):

\[ g_s(t) = \int d\omega f_s(\omega_s) e^{-i\omega_s t}, \quad (1.19) \]

spectral factorability implies that photons are also temporally independent in their generation interval.

Experimentally, to achieve factorability, it is very common to use bandpass filters for filtering the down-conversion photons. In this approach, the spectral distributions of the signal and idler are determined by the spectral distribution of the filters. The drawback of this method is that it discards photons and, thus, reduces the heralding efficiency and the brightness of the sources, especially when the filters are narrow compared to the natural phase-matching bandwidth. The low brightness and heralding efficiency becomes progressively problematic in multi-photon (more than two photons) experiments, as rates and signal-to-noise ratios become poor.

The JSA can be written as a product of the pump spectral envelope and a phase-matching function. The pump part describes the energy distribution of the pump photons available for SPDC. The phase-matching function may determine the ways that the signal and idler photons are distributed with respect to pump photons. The down-conversion process has non-negligible amplitude when \( \Delta k \sim 0 \), which places constraints on the spatial and spectral distribution of the signal and idler photons. Because \( k_p(\omega) \), \( k_s(\omega) \), and \( k_i(\omega) \) are each functions of frequency, \( \Delta k L \) is also a function of frequency. Requiring \( \Delta k L \) and the distribution of \( k_p, k_s \) and \( k_i \) that satisfy this to have the desirable correlations leads to a condition that the group velocity of the pump should sit in between signal and idler group velocities—this is known as symmetric group velocity matching (GVM) (150). Meeting the phase-matching condition with GVM is a more effective approach for achieving a factorable JSA (150–156). Practically, the pump envelope and crystal properties can be engineered together in a way that, without filtering, the spectral distribution is factorable. Type II phase-matching would be ideal, because the daughter photons have different polarisations, which gives the option of engineering the refractive index individually for each of them to meet the GVM condition. The symmetric GVM condition is given by

\[ N_p = \frac{N_s + N_i}{2}, \quad (1.20) \]

where \( N = \frac{c}{v_g}, v_g \) is the group velocity, and \( c \) the speed of light in vacuum.

To sum up, for designing a high-quality SPDC source, we have a procedure to follow: i) for the specific material of the crystal (KTP in our case) and the phase-matching (type II in our case), the wavelengths that satisfy GVM are found; ii) the poling period of the crystal is determined in a way that satisfies the phase-matching condition at the desired temperature; iii) the crystal length and pump bandwidth are set in a way that the JSA is factorable; and iv) the pump and down-conversion photon beam waists are
engineered in a way to achieve the desired brightness and heralding efficiencies (157). (There is some freedom in the design—for example, one can achieve either the maximum heralding efficiency or brightness, however, one cannot have whatever combination of brightness and heralding efficiencies one might desire.)

In our group, Weston et al. (110) previously implemented the above-mentioned ideas to engineer a source that used the GVM and quasi-phase-matching techniques, realising a source with a factorable bi-photon spectrum. The results from their work—both theoretical modelling and experimental techniques—played a major role in understanding and in developing skills for building the new source that I will report. In their source, they used a ppKTP crystal pumped by a femtosecond pulsed laser at 785 nm, generating degenerate signal and idler photons at 1570 nm. They engineered the pump, signal and idler beam waists to have a very high heralding efficiency of \((65 \pm 2)\%\). (They also realised a heralding efficiency of \((82 \pm 2)\%\) for a source phase-matched to generate 1550 nm degenerate photons (38).)

My colleagues implemented the same techniques for engineering a polarisation-entangled photon source (at 1570 nm) using a Sagnac interferometer (158). The entangled source they built produced some high-quality entangled states, but the Sagnac interferometer caused significant challenges. First, it was challenging to align and optimise the interferometer due to design, component, and environmental limitations. This led to a reduction in the purity and tangle (a measure of entanglement for quantum states, ranging from 0 to 1 for the maximally entangled state (18)) of the entangled polarisation-encoded state. The other issue, perhaps the most important one, was that even for the relatively compact interferometer they built, it was not quite possible to achieve the required lens configuration for high heralding efficiency. Because the signal and idler beam waists were not quite optimal for high heralding efficiency, the value of that quantity dropped below that for the unentangled source, to \((52 \pm 2)\%\).

Motivated by the design implemented in References (141, 150), we use alpha-BBO beam displacers (BD), instead of a Sagnac configuration, to realise the interferometer required for producing entanglement. This design makes it possible to have better focusing and beam control, which allows for higher heralding efficiency and excellent entanglement quality. The lens configuration for the optimal heralding efficiency is mainly determined by the crystal length, which in turn is determined by the pulse length. For the BD interferometer design, the optimal lens configuration is not achievable by a femtosecond pulse, instead, it is possible by a range of picosecond pulses. Moreover, using a spectrally narrower pulse will lead to narrower down conversion, which will experience less dispersion in BDs and other optical elements.

There are other side benefits of this new design. First, we can couple the pump beam to a photonic crystal fibre before the down-conversion crystal, to decouple the laser from the SPDC and overcome the effects of laser beam pointing instability and drifts to decouple the laser from the SPDC. This alignment fibre was not practical with
Chapter 1. Introduction

a femtosecond pump because of the larger nonlinearities and dispersion associated with a shorter pulse. Second, we observed that the BD interferometer design is passively stable, generating high-quality entangled states with high heralding efficiency, for a time interval of at least one week, compared to few days for the Sagnac with no alignment fibre.

This source was used for the experiments presented in Chapters 4 and 5. The details of its application in each experiment are provided in the relevant chapters. It is worth noting that I have not required to use the full potential of this new source—I have not harnessed the high heralding efficiency, high-quality entangled states, and expected high-visibility independent HOM interference all at the same time. Here, I briefly mention the characteristics of the new source that we built.

We used a pulsed Ti-sapphire laser with a pulse repetition rate of around 72 MHz to pump the crystal at 775 nm with the spectral pulse width of \((0.61 \pm 0.03)\) nm. A short (30 cm) photonic-crystal fibre\(^{22}\) was used in-between the pump and crystal. A 15 mm long ppKTP crystal, cut for collinear type-II phase-matching, is used to generate photons. We experimentally found that the poling period of 46.20 \(\mu\)m with the crystal temperature slightly below room temperature at 21\(^{\circ}\)C, are the best configurations for producing degenerate SPDC photons at 1550 nm. We could not see any major effect on the phase-matching conditions when we slightly tilted the crystal angle.

Our source generates a singlet state \((18)\) having a quantum fidelity \((18)\) with the theoretical state of 0.9967 \(\pm\) 0.0005, a tangle \((18)\) of 0.9886 \(\pm\) 0.0002 and a purity of 0.9935 \(\pm\) 0.0002. With the addition of 8 nm FWHM filters centred at 1550 nm and by tweaking the crystal temperature, these numbers slightly improved to 0.9983 \(\pm\) 0.0005, 0.994 \(\pm\) 0.002, and 0.9966 \(\pm\) 0.0009, respectively. Moreover, the symmetric heralding efficiency \((\epsilon_A = \epsilon_B)\) of our source is 0.649 \(\pm\) 0.004, when using sub-optimal photon detectors. We estimate that this efficiency will rise to 0.82 when using our best SNSPDs. (I did not have access to those detectors at the time of characterising our source.) The heralding efficiency is barely reduced by adding 8 nm bandpass filters. This is attributed to the fact that the bandwidth of down-conversion photons, with a FWHM of about 5 nm, is narrower than the bandwidth of the filters.

We performed a dependent HOM interference, achieving visibility of 0.9847 \(\pm\) 0.0008. Performing the independent HOM interference between two identical sources will be the next step forward, which will make this new design a suitable tool for more than two-photon experiments that need HOM interference between different sources.

We also used the BD interferometer source design to realise a Werner-state \((159)\) source, essential for the experiment presented in Chapter 4. In Section 1.4, I briefly mention the background of this work.

\(^{22}\)Thorlabs ESM-12B-ESM large mode area photonic crystal fibre, with 12.2 \(\mu\)m core.
Chapter 1. Introduction

1.4 Einstein-Podolsky-Rosen steering

Entanglement is one of the most fascinating features of quantum mechanics (160). The non-local properties of entangled states were famously recognised by Einstein, Podolsky, and Rosen (EPR) (3), who wrote: “As a consequence of two different measurements performed upon the first system, the second system may be left in states with two different wave functions.” EPR posed a paradox. They observed that it is possible to find two non-commuting measurements \{A, B\}, such that measuring the first party in \(A\), the state of the second party is an eigenfunction of observable \(P\). If the first party is measured in \(B\), the second party is in an eigenfunction of \(Q\). Although \(P\) and \(Q\) are non-commuting observables, they can be certainly predicted. This is in contradiction with the Heisenberg uncertainty principle. The EPR-paradox was interpreted as suggesting that quantum mechanics was an incomplete theory, because it could not be compatible with realism (a physical quantity is an element of physical reality if it can be certainly predicted without any disturbance).

Later on, Schrödinger wrote: “It is rather discomfiting that the theory should allow a system to be steered or piloted into one or the other type of state at the experimenter’s mercy in spite of his having no access to it.” (161). He proposed that for an entangled state, if the first party, Bob, announces his measurement setting, then the other party, Alice, can steer Bob’s state to one of his eigenvalues. Therefore, Alice can certainly predict Bob’s state before his measurement (162). Schrödinger refused to accept the existence of the local hidden variable idea proposed by EPR. However, he also could not believe that non-locality existed; hence, he discomfingly agreed on the incompleteness of quantum mechanics.

Today, we are almost sure that quantum non-locality exists (141, 163, 164), after convincing experimental tests. It was not until very recently that steering was operationally defined as a two-party quantum task. In this task, there are two parties, Alice and Bob (162). One party is not trusted, meaning that whatever they announce might be fake results instead of results from quantum measurements. However, the other party and its apparatus are trusted, in a way that their results are considered to be genuine and governed by quantum mechanics. In our formalism, a potentially dishonest Alice sends a state to a trusted Bob. Then, Bob announces his choice of measurement from a specific set to Alice, and then she announces her measurement result (or null, meaning that she did not receive any state to measure) to him. Repeating this procedure many times, based on his measurement and Alice’s results, Bob calculates a parameter called the steering parameter, which characterises the correlation between the two parties. If the steering parameter violates a specific steering inequality (determined based on the state and measurement type) then Bob is convinced that the state he received is half of an entangled pair. The protocol is covered in depth in References (162, 165). This systematic approach can be quantified for different types of states and measurements, such as Werner states and projective measurements (166).
Steerable states are a subset of non-separable (entangled) states and a superset of Bell non-local states (162).

Beyond its fundamental aspects, steering has some potential technological applications. Because steering is asymmetric (only one of the two parties is trusted), it can be used for verifying that there is genuine entanglement, of a kind useful for security-based protocols, between a trusted node (e.g. at a home base) and another party that need not be in a secure location. Entanglement sharing is one of the key requirements for some of the important tasks in QIP, including quantum key distribution (QKD) (167–169), generating random numbers (170, 171) and teleportation (172). From an experimental point of view, there have been a range of tests of EPR-steering of increasing complexity. The first experiments, performed with continuous-variable states of light, demonstrated the existence of the effect that EPR described (173–176). After reformatting EPR-steering as a task that can identify non-locality, a series of new significant experimental works have been reported. I focus on those performed with photon qubits. The first experimental test was done by Saunders et al. (165), where a steering inequality was violated using Werner states and projective measurements.

In a steering task, a dishonest Alice might try to cheat by refusing to announce some of her results, and instead declaring that she did not receive the photon. Using this approach, she unfairly samples her measurement result based on Bob’s measurement choice. This is why it is crucial for Bob to be able to adjust his steering inequality to consider the fraction of trials in which Alice announces a result. This includes both genuine losses in the experiment and fake losses reported by Alice. Arbitrarily loss-tolerant EPR-steering was implemented in Reference (142). It was theoretically and experimentally shown that, with the arrangement of an untrusted Alice and a trusted Bob, the steering task is possible with high losses in Alice’s arm, as long as the state that the parties share is a high-fidelity state, and the parties are able to perform a many-setting measurement set. 23 This robustness to loss is one of the signatures of steering, not available in Bell non-locality (177). An example is the relationship of steering (178) to one-sided device independent QKD (1SDIQKD) (169). Requirements for 1SDIQKD are easier in terms of channel losses and detection efficiencies, compared to device independent QKD and its link to violation of Bell inequalities.

1.4.1 Asymmetry in quantum steering; one-way steering

One of the unique properties of quantum steering, compared to Bell non-locality tests, is its asymmetry—one party is trusted, and the other is not. This leads to an interesting and surprisingly subtle question: “Can a steering experiment be performed where Alice can steer Bob’s state but not vice versa?”. The essence of this question is whether steering remains possible if Alice and Bob exchange roles (as (un)trusted party), and

23 Indeed, the parties’ roles can be swapped—One can choose Bob as the trusted party, and Alice as the untrusted one, or vice versa.
under what conditions one- vs two-way steering occurs. One-way steering was first shown in the context of continuous variables (Gaussian states) and restricted to Gaussian measurements \((179, 180)\). However, there is theory that shows that the use of particular non-Gaussian measurements makes the states two-way steerable again \((181)\). Is there a state and protocol for which one-way steering must be the case?

For qubit states, in 2014 Bowles et al. proposed a class of two-qubit states, which are one-way steerable, where the parties’ choice of measurements is limited to arbitrary projective measurements \((182)\). However, even for this case, the range of states that are one-way steerable is very limited and, therefore, the scheme is experimentally difficult to implement. Evans et al. \((183)\) proposed an experimentally feasible class of two-qubit states that are one-way steerable—also in the limited case of only projective measurements. In a seminal work, Quintino and co-authors proved that a systematic approach exists that can be used to extend proofs for a set of states that are one-way steerable with projective measurements to proofs of one-way steerability for a new, related, set of states under the condition of general measurements (positive-operator-valued measures (POVMs)) with an arbitrary number of outcomes \((184)\). This and the former result by Evans led to knowledge of a class of one-way steerable states with general measurements, which could be practically tested in an experiment \((181)\).

The first one-way steering experiment that was not strictly limited to a very specific set of measurements was performed by Wollmann et al. \((181)\). A range of Werner states was found (theoretically \((181, 183, 184)\)) and tested (experimentally), which were one-way steerable\(^{24}\). The authors showed that with the projective measurement choice, Alice can steer Bob, with the detection loophole closed. However, if the roles are changed, such that Bob is trying to steer Alice’s state, by adding loss to Bob’s arm, he cannot steer Alice’s state, even with the potentially infinite number of POVM measurements. It is impossible to actually measure all (an infinite number) infinite-setting POVMs. Therefore, non-steerability under party exchange was established by tomographically characterising the shared state and comprehensively testing the theoretical criterion for non-steerability \(^{25}\). There were other one-way steering experiments \((186, 187)\), though they strictly limited the measurements to two-setting and multi-setting projective measurements, respectively, and these were not completely general. Indeed, all of these works also restricted their consideration to a limited class of states. The ideal one-way steering test would be one where the parties have no limitations, either on the states they can prepare or on the measurements they can carry out.

Recently, Baker et al. \((166)\) came up with a practical\(^{26}\) and necessary condition for

\(^{24}\)The family of Werner states offer tuneability from a completely mixed state to a maximally entangled state. They are also invariant under equal one-qubit local unitary operations on each party’s qubit. These properties make them popular to work with in the context of steering.

\(^{25}\)For more details, see References \((166, 185)\).

\(^{26}\)A practical condition is one that can be feasibly solved as an optimisation problem, and the answer provides a steering inequality.
any two-qubit state to be steerable, in the presence of loss, for each party—i.e. a necessary condition for steerability. In this paper, they found two simple criteria, which are the sufficient conditions for two-qubit state non-steerability for restricted projective measurements and POVMs. One of the practical considerations in their work, compared to others (188), was to include particle loss, which is inevitable for optical experimental realisations.

These new criteria were tested on the experimental results from Reference (181), in which it was assumed that the entangled state was a two-qubit Werner state subjected to loss\textsuperscript{27}. The results reported in Reference (181) did not strictly satisfy the POVM non-steerability conditions in Reference (166). The experiment I report in Chapter 4 is motivated to, for the first time, experimentally demonstrate one-way steering where there are neither assumptions about two-qubit states nor allowed measurements.

In our scheme, we improved the non-steerability criteria of Reference (166) to include all projective measurements and POVMs, with no assumption about the measurement that the parties can choose. Experimentally, we prepared a very high-quality tunable Werner state, which to our knowledge is the best of its type, thanks to the new source design. Then we showed that Alice can steer Bob’s state by violating a steering inequality, with the detection loophole closed (142). At the same time from the measured experimental results, we tomographically reconstructed the prepared state and showed that the state we prepared satisfies non-steerability (Bob to Alice) for POVMs (POVMs include projective measurements). Note that although we do not make any assumption about the prepared state, theoretically and experimentally we found that having a high-quality entangled state that is tunable is essential for implementing the rigorous experiment reported in our work. In other words, the validity of a steering inequality requires no assumption about the prepared state, but simultaneously discounting steerability in one direction, while observing steering in the other, requires a high-quality entangled state, which is very close to the ideal state. The result of our work is presented in Chapter 4 of this thesis.

\textsuperscript{27}The actually created states in Reference (181) were Werner-like states. These states are equivalent to Werner states, up to a local unitary.
Chapter 2

Witnessing the ambiguity of simplicity via quantum simulations of an Ising spin chain

2.1 Statement of contribution to co-authored published paper

This chapter includes a co-authored paper. The bibliographic details and publication status for this paper including all authors is:

My contribution to the paper involved some of the theory and most of the experiment, including:

• Working on some of the theory of an error-tolerant approach for quantum simulation of stochastic processes
• Building, characterising and optimising the experimental setup
• Acquiring the experimental data and analysing them
• Writing and editing the manuscript.

(Signed) ______________________ (Date) 11/02/2019
Name of Student: Farzad Ghafari Jouneghani

(Countersigned) ______________________ (Date) 11 Feb 2019
Supervisor and Corresponding author of paper: Geoff Pryde

(Countersigned) ______________________ (Date) 11/02/2019
Corresponding author of paper: Farzad Ghafari
Chapter 2. Witnessing the ambiguity of simplicity via quantum simulations of an Ising spin chain

(Countersigned) ____________________________  (Date) __________________________

Supervisor: Nora Tischler

11/02/2019
Chapter 2. Witnessing the ambiguity of simplicity via quantum simulations of an Ising spin chain

2.2 Quantum simulations of an Ising spin chain

Stochastic processes are employed to study a large number of different phenomena where the quantity of interest varies discretely or continuously through time or space in a probabilistic fashion. As discussed in Chapter 1, one of the crucial considerations in implementing the computer models of complex processes is the information storage requirements of the simulator. A key metric for this is the minimum amount of memory the model (and its corresponding physical simulator) must store about the past to simulate the process’ future. In complexity science, this quantity has been formalised as the statistical complexity (41), and it is regarded as a key quantifier of the structure (the repeated pattern that can be used to predict the future of a process, based on the past observations). Remarkably, quantum mechanics allows the simulation of many processes with drastically reduced memory (17), providing the potential for significantly more memory-efficient simulators.

In this work, we experimentally demonstrate the benefits of quantum-encoded storage in reducing statistical complexity for simulating an interesting physical process, namely a 1D Ising spin chain. Moreover, for the first time, we experimentally show that the relative complexity of simulating two stochastic processes is dependent on the fundamental physical properties of the simulator being used—even if they produce the same statistical outputs—rather than being an inherent property of the systems. This phenomenon was coined the ambiguity of simplicity. To observe this effect, we use an encoding in quantum states to simulate a relatively large range of the non-trivial physical processes of Ising chain. Moreover, a new error-tolerant approach is implemented that considers experimental imperfections, which is crucial for realising practical simulators with quantum-encoded memory. The Ising system parameters (temperature and magnetic field, T and B) are chosen to allow for error-tolerant simulation in a range where the turning point in quantum statistical complexity was observable, which is necessary for observing the ambiguity of simplicity.

This work is of three-fold importance. First, it is a quantum simulation of a physical process, achieving a reduction in the statistical complexity. Second, it is an experimental demonstration of the ambiguity of simplicity. Third, it uses a new method for dealing with experimental errors in stochastic simulation.

2.2.1 Witnessing the ambiguity of simplicity via quantum simulations of an Ising spin chain

The remainder of this page is intentionally left blank. The paper appears on the following pages.
Witnessing the ambiguity of simplicity via quantum simulations of an Ising spin chain

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

1Centre for Quantum Dynamics, Griffith University, Brisbane, 4111, Australia
2School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore 639793, Republic of Singapore
3Centre for Quantum Technologies, National University of Singapore,
3 Science Drive 2, Singapore, Republic of Singapore
4Centre for Quantum Computation and Communication Technology (Australian Research Council), Griffith University
(Dated: February 11, 2019)

The modelling of stochastic processes is ubiquitous throughout the natural and social sciences. An ideal model produces the correct statistical output without any unnecessary complexity. This minimal-complexity criterion is important conceptually—entailing the least number of causes of effects (Occam’s razor)—and practically—entailing the least stored information in the model. Recent works have established that quantum models can store significantly less information than their optimal classical counterparts, and can thus redefine our notion of what is complex. Here, we experimentally observe such quantum reduction in simulating the Ising spin chain. This leads us to experimentally observe a recently conjectured effect, ambiguity of simplicity, that the relative complexity of two difference processes can depend on whether we attempt to model it using classical or quantum means of information processing. Meanwhile our error-tolerant techniques account for inevitable imperfections in realizing quantum simulators, thus providing the technological milestone needed to simulate increasingly complex stochastic processes.

Introduction.—Stochastic processes arise frequently in nature and society [1, 2]. Even if a system is deterministic in principle, it may appear stochastic because of the near impossibility of tracking all the microscopic variables and transformations. Thus, we consider the system to be a black box where only its outputs are observed. A mathematical model describing such a system probabilistically generates data whose statistics match those of the future outputs of the system being simulated. To implement the model, a valid physical simulator, like a computer, is required.

There will be many models that reproduce the system’s behaviour, but not all of them will be memory-efficient. A key metric for this is the minimum amount of memory the model (and its corresponding simulator) must store about the past to simulate the process’ future [3, 4]. In complexity science, this quantity has been formalized as the statistical complexity [4–6], and is regarded as a key quantifier of structure. It equals the minimal memory a model needs to generate future behaviour with correct statistics using only information available from past observations [4–6]. Remarkably, quantum mechanics allows simulation of many processes with drastically reduced memory, heralding the potential for significantly more memory-efficient stochastic simulators [7–18]. In addition, the use of quantum mechanics can drastically affect the relative order of what we consider complex—a phenomenon termed ambiguity of simplicity [9]. A process can become more complex to model classically as we increment a parameter, but easier to model quantum mechanically.

Here we demonstrate a cornerstone for developing advanced stochastic simulations, enabling us to witness such effects in realistic settings. We employ and characterize quantum memory resources in simulating a key problem in statistical physics—the one-dimensional (1D) Ising spin chain [19]. This system is complex enough for two important questions to arise. The first is how to address imperfections in a real, necessarily non-ideal simulator, allowing us to go ahead with the simulation even in the presence of small deviations from design. The second is to employ these tools to experimentally witness the ambiguity of simplicity. By simulating the same Ising system at two temperatures, TA and TB, our results confirm a significant region of temperature pairs, TA and TB, such that classical and quantum notions of relative complexity are reversed. This establishes that there is no universal ordering of simplicity between stochastic processes; which system is considered more complex fundamentally depends on what sort of information processing is allowed. In addressing these important tasks, the Ising simulation we present goes far beyond a previous demonstration of quantum-suppressed statistical complexity [12].

Framework and tools.—Consider a dynamical system that emits discrete-valued outputs x(t)—instances of random variables XI—at discrete times t ∈ Z. The output string, · · · , X−1, X0, X1, X2, · · · , is a stochastic process, described by a joint probability distribution, P = P(X, X). Here X = · · · , X−1, X0 and X = X1, X2, · · · , 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 X, the model must provide a systematic means of initializing a suitable system Ξ in some state ξ(X), such that repeated
application of a systematic action \( \mathcal{M} \) on \( \Xi \) sequentially generates \( x_1, x_2 \ldots \) governed by the conditional future \( P(\vec{X} | \Xi) = \vec{f} \). Here \( \epsilon(\cdot) \) is referred to as the encoding function, capturing precisely how the model encodes the past within its memory.

For stationary stochastic processes, the optimal, or simplest, model is defined as the one where the information entropy of \( \Xi \) in steady-state is minimized \([4,5]\). These optimal classical models are known as \( \epsilon \)-machines, and can be systematically constructed \([4,20]\). This involves assigning each past to an appropriate causal state \( S_i = \epsilon(\vec{f}) \), such that two pasts are assigned to the same state whenever their conditional futures statistics coincide. The \( \epsilon \)-machines then use a memory \( \Xi \) that associates each causal state with a different physical configuration. The subsequent dynamics of the process can be reproduced by repeated applications of a particular stochastic map that defines the probability a machine in state \( S_i \) will transition to \( S_j \) while emitting output \( x \).

The \( \epsilon \)-machine thus retains \( C_\mu = -\sum_{i=1}^{N} p_i \log_2 p_i \) bits of information about the past, where \( p_i \) is the probability that \( \vec{f} \) belongs to causal state \( S_i \). Any practical realization of such machines will thus require at least this much memory.

Despite their provable optimality, such machines still store extraneous information. For many processes, \( C_\mu \) is strictly greater than the amount of information the past contains about the future—the mutual information \( I(X; \Xi) \) \([3]\). Whenever this is true, more efficient quantum models exist \([7,8,18]\). In such quantum machines, \( \Xi \) can encode relevant past information coherently—replacing each \( S_i \) with a quantum state \( |S_i\rangle \). The resulting memory is then

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

where \( \rho = \sum p_i |S_i\rangle \langle S_i| \). In this case, repeated application of some quantum instrument on a system \( \Xi \) in state \( |S_i\rangle \) generates correct conditional future statistics, i.e. samples from \( P(\vec{X} | \Xi) = \vec{f} \), while leaving \( \Xi \) in whatever state \( |S_j\rangle \) is appropriate. Whenever \( C_\mu > E \), there exists a systematic construction realizing \( C_q < C_\mu \) (see main theorems of Gu et al. \([7]\)). Recent progress has culminated in the current state-of-the-art quantum models, known as \( q \)-machines, which are able to generate future statistics through unitary evolution \([8,18]\).

**Causal states for Ising system.**— Here we study the 1D Ising system which is an infinite spatial chain of spins with nearest-neighbor interactions. There is a one-to-one mapping such that we can replace a simulation over a series of discrete times (described above) with a simulation over 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 \([21]\). For the Ising system, the energy function is

\[
H(\vec{x}, \vec{x}) = -\sum_{n} (J x_n x_{n+1} + B x_n),
\]

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 \([19]\). We use natural units for temperature \( (k_B = 1) \) and take the coupling \( |J| \) to be the unit of energy so that \( |J| = 1 \) and \( T \) and \( B \) are dimensionless.

Here, the \( \epsilon \)-machine has two causal states, \( \{S_i\}_{i=0,1} \), with encoding function \( \epsilon(\vec{f}) \) that identifies any two pasts with coinciding \( x_0 \) \([21]\). The machine operates according to the transition probabilities \( \Gamma_{ij}(J,B,T) \), which represent the probability a simulator in state \( S_i \) will transit to \( S_j \) while emitting output \( j \) (see Supplemental materials (SM), Section 1).

The provably optimal quantum models can be established for this particular process \([14]\). The resulting quantum model has quantum causal states

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

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

where \( |0\rangle \) and \( |1\rangle \) are orthogonal qubit states. In general \( |S_0\rangle \) and \( |S_1\rangle \) are mutually non-orthogonal. Thus the quantum model requires less memory to encode past information than its optimal classical counterpart. Ideally, the quantum gate that implements the quantum \( \epsilon \)-machine and generates correct future statistics at each time-step (see Fig. 1a) is the controlled-unitary given by

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

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

To implement \( CU \) in a quantum circuit, we can decompose the unitary being controlled, \( U \), into unitary operators such that

\[
CU = (I \otimes V_0)(I \otimes V_1)(I \otimes H)CZ(I \otimes H)^{-1}(I \otimes V_1)^{-1},
\]

where \( V_0|0\rangle = |S_0\rangle \), \( V_1 \) is the rotation in the \( X-Z \) plane that rotates \( |0\rangle \) to the bisector of \( |S_0\rangle \) and \( |S_1\rangle \), \( H \) is a Hadamard operation, and \( CZ \) refers to a controlled-Z gate.

**Simulation of the Ising system.**— In the experimental set-up (Fig. 1b), we implement one complete cycle of the \( \epsilon \)-machine, comprising state preparation, a controlled-unitary operation, and read-out. 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 BiB₃O₆ (BiBO) crystal cut for type-I phase matching. We use polarization to encode classical logical states
\( |0\rangle = |H\rangle \) and \( |1\rangle = |V\rangle \), where \( H \) and \( V \) are horizontal and vertical polarization, respectively. The optical simulation circuit is based around single-qubit unitary rotations implemented with wave plates and a nondeterministic linear optics \( CZ \) gate using three partially-polarizing beam splitters (PPBSs) [22]. The polarization qubits are measured using wave plates and avalanche photodiodes. Quantum state and process tomography are implemented using the methods in ref. [23].

Small experimental imperfections mean that, instead of an ideal controlled-unitary, a more general transformation \( \mathcal{E} \) is implemented. In the previous experimental work [12] these imperfections were ignored; the simulator was assumed to be ideal. Here we go beyond that naive treatment. We use quantum process tomography to characterize errors in our real, necessarily imperfect, simulator, and based on this we determine the best mapping of the real simulator onto the Ising system parameters. This allows us to perform a simulation of a wide range of Ising parameters even in the presence of small deviations from design.

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. For the non-ideal process \( \mathcal{E} \), a different output state \( \rho^\mathcal{E}(j|i) = \mathcal{E}_j(\rho_i) \) is obtained. Here, we have broken the two-qubit map \( \mathcal{E} \) into two completely positive conditional processes \( \mathcal{E}_0 \) and \( \mathcal{E}_1 \), acting on the memory qubit, depending on whether the measurement outcome was 0 or 1. The key point is that the output state from one step of the simulator is now not, in general, equal to either of the possible input causal states. After more iterations, the states may diverge further from the ideal causal states.

To account for the device imperfections, we follow the following procedure. First, for the nominal values of the parameters \( J = 1, B = 0.3 \) and varying \( T \) (the nominal \( T \) values can be seen in SM, Section 1), we tune the experiment to realize, as best we can, the quantum \( \epsilon \)-machine for those parameters. Next, we perform quantum process tomography [23–25] of the circuit to obtain the \( \mathcal{E}_0 \) and \( \mathcal{E}_1 \) maps. From these maps, and these maps alone, we find the states and transition probabilities \( \{\rho^m_i, \Gamma^m_{ij}\} \) \((m \text{ is for "maps"})\), which describe the two-state quantum machine most closely corresponding to these maps, such that \( \mathcal{E}_j(\rho^m_i) \approx \Gamma^m_{ij}\rho^m_j \). Here, closeness is defined in terms of trace distance; see SM, Section 2. We call the \( \rho^m_i \) the fixed-point states. The stationary state of our machine is then \( \rho^m = \rho^m_0 + \rho^m_1 \), where \( \rho^m_0 = \Gamma^m_{10}/(\Gamma^m_{01} + \Gamma^m_{10}) \) and \( \rho^m_1 = 1 - \rho^m_0 \). Since the \( \Gamma_{ij} \) are functions of \( T \) and \( B \), for a fixed \( J = 1 \), 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 implements. For more details on the deviation of implemented values from nominal values, see SM, Section 3.

Using Equation (1), we can calculate the quantum statistical complexity for each of the experimentally-determined stationary state: \( C^m_q = C_q(\rho^m) \) (Fig. 3a). We observe that \( C^m_q \) lies close to the estimated theoretical range of statistical complexity values (in section 4 of SM, we explain how to estimate the uncertainty bound). The slight discrepancy between the experimental and theoretical values primarily arises from small repeatability errors in the experimental simulator settings, and from the fact that the calculated fixed-point states almost, but do not exactly, satisfy \( \mathcal{E}_j(\rho^m_i) = \Gamma^m_{ij}\rho^m_j \) for \( j = 0, 1 \).

The approach that we use, finding the fixed-point

\[ X \]

\[ 2 \]

\[ \text{Model} \]

\[ X \]

\[ \epsilon \text{-Machine} \]

\[ |S_j\rangle \]

\[ \text{System} \]

\[ X \]

\[ \text{Physical simulator} \]

\[ |0\rangle \]

\[ \text{Photon source} \]

\[ \text{Controlled-unitary} \]

\[ \text{Photon counting} \]

\[ \text{APD} \]

\[ \text{FPC} \]

\[ \text{PPBS} \]

\[ \text{BBO crystal} \]

\[ \text{Half wave plate} \]

\[ 820\text{nm bandpass filter} \]

\[ \text{GT} \]

\[ \text{GT} \]

\[ \text{APD} \]

\[ \text{FPC} \]

\[ \text{PPBS} \]

\[ \text{GT} \]

\[ \text{APD} \]

\[ \text{FPC} \]

\[ \text{PPBS} \]

\[ \text{GT} \]

\[ \text{APD} \]

\[ \text{FPC} \]

\[ \text{PPBS} \]

\[ \text{GT} \]

\[ \text{APD} \]

\[ \text{FPC} \]

\[ \text{PPBS} \]

\[ \text{GT} \]

\[ \text{APD} \]

\[ \text{FPC} \]

\[ \text{PPBS} \]

\[ \text{GT} \]

\[ \text{APD} \]

\[ \text{FPC} \]

\[ \text{PPBS} \]

\[ \text{GT} \]

\[ \text{APD} \]

\[ \text{FPC} \]

\[ \text{PPBS} \]

\[ \text{GT} \]

\[ \text{APD} \]

\[ \text{FPC} \]

\[ \text{PPBS} \]

\[ \text{GT} \]

\[ \text{APD} \]

\[ \text{FPC} \]

\[ \text{PPBS} \]

\[ \text{GT} \]

\[ \text{APD} \]

\[ \text{FPC} \]

\[ \text{PPBS} \]

\[ \text{GT} \]

\[ \text{APD} \]

\[ \text{FPC} \]

\[ \text{PPBS} \]

\[ \text{GT} \]

\[ \text{APD} \]

\[ \text{FPC} \]

\[ \text{PPBS} \]

\[ \text{GT} \]

\[ \text{APD} \]

\[ \text{FPC} \]

\[ \text{PPBS} \]

\[ \text{GT} \]

\[ \text{APD} \]

\[ \text{FPC} \]

\[ \text{PPBS} \]
states of the physical map, aims to account for the possible imperfections in the experimental implementation of the $\epsilon$-machine. However, one of the main features of the quantum $\epsilon$-machine is to encode the causal state in the input memory state and then measure binary outcome while having another causal state for the next step (Fig. 1a). To confirm that our approach is compatible with this feature of $\epsilon$-machine, we used the fixed-point states as input causal states. Because $\rho_0^m$ and $\rho_1^m$ are mixed states, we use averages over ensembles of two pure states to implement them experimentally. The binary outputs of the simulation are found by measuring the ancilla output in the logical basis, and these statistics are used to determine $\Gamma_{ij}^s$ (s is for “statistics”) which are the transition probabilities reconstructed from the collected statistics when one of the $\rho_i^m$’s is used as input causal state. This yields the corresponding stationary state probabilities $p_0^s = \Gamma_{00}^s / (\Gamma_{00}^s + \Gamma_{01}^s) = 1 - p_1^s$. Moreover, the corresponding stationary state is $\rho^s = p_0^s \rho_0^m + p_1^s \rho_1^m$.

In a similar way to $T^m$ and $B^m$, we can find $T^s$ and $B^s$. The values of $C_2^s = C_q(\rho^s)$ versus corresponding temperatures are demonstrated in Fig. 3a. We observe that $C_0^m$ and $C_0^s$ are close to each other, also in good agreement with the theoretical statistical complexity values. To compare, the $B^m$ and $B^s$ have mean values of $0.29 \pm 0.02$ and $0.28 \pm 0.05$, respectively. Moreover, the corresponding $T^m$ and $T^s$ values are shown in Fig. 3b, which also reassure that they are close to each other.

For comparison, we also implement the classical $\epsilon$-machine using the same experimental set-up. In this case, $|S_0\rangle = |0\rangle$ and $|S_1\rangle = |1\rangle$, and future statistics are generated based on introducing classical randomness [12]. Since the states are orthogonal, they do not inherently contain transition probabilities, where we implement these probabilities by preparing orthogonal states in an ensemble of experiments with numbers proportional to the probabilities $p_0$ and $p_1$, respectively. Results for the classical $\epsilon$-machine are shown in Fig. 3a, and lie close to the theoretical estimation.

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 [9]? Consider two Ising systems with different temperatures $T_A$ and $T_B$. If in the classical regime $C_1^A < C_1^B$, which means that A is simpler than B, and in the quantum regime $C_0^A < C_0^B$ as well, then there is consistency between the two classes for processes A and B. However, if the quantum model reverses their ranking compared with the classical perspective, we have the ambiguity of simplicity [9]. The basic question, “Which process is simpler?” no longer has a well-defined answer. To mathematically describe this phenomenon, we define

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

(5a)

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

(5b)

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. 3c, 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.

Conclusions.—The statistical complexity is a key measure of structure in complexity science, capturing how much information we must store about a process’s past in order to replicate its conditional future behaviour. We experimentally demonstrated that quantum mechanics allows simulation of stochastic problems of physical interest, such as the Ising spin chain using less memory than classically possible. From a technological point of view, we demonstrated that this quantum advantage persists even in the presence of experimental noise. Our results showed that the quantum resource requirements for simulation exhibit drastically different behaviour compared to their classical counterparts, peaking at finite rather than infinite temperature. Here, hence, we document the first experimental observation of ‘ambiguity of simplicity’—the relative statistical complexity between two stochastic processes (e.g. an Ising chain at two different temperatures) differs depending on whether we store information in classical or quantum memory. Our experimental advances provide a launch-pad for witnessing more memory efficient means of simulation in more sophisticated settings. These include cases where quantum advantage can be unbounded [18, 26, 27], as well as variants that apply adaptive systems, or processes that emit outputs in continuous time [13, 15, 28].

One potential extension is to consider approximate modelling. Presently, statistical complexity is defined for
the case where models are exact. In many situations, one may wish to consider an extension of statistical complexity, measuring the past information needed to replicate future statistics to a specified statistical fidelity. How robust is the ambiguity of simplicity in such scenarios? This question is non-trivial, as there presently no systematic ways to determining optimal classical or quantum models. However, one may nevertheless be able to make meaningful progress by considering upper bounds through explicit constructions.

A second important extension would be to consider single-shot scenarios. In computational mechanics, memory is most commonly measured by entropy—which is an i.i.d. quantity that reflects operational advantage in the limit of operating a large number of models simultaneously. However, it is also possible to consider the minimal dimensions (in terms of configuration space) of a memory to be able to store the past. In the case of the Ising system, both classical and quantum models required a two-level system. However, more sophisticated process have recently been found where this quantity can also diverge [29]. It would certainly be interesting to witness if ambiguity of simplicity is also present with such systems, allowing the observation of quantum advantage without the need for tomography.

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 research was funded, in part, by the Australian Research Council (project no. DP160101911). This work is also supported by the National Research Foundation of Singapore and in particular NRF Award No. NRF–NRF2016–02, the John Templeton Foundation Grant 54914 “Occam’s Quantum Mechanical Razor: Can Quantum theory admit the Simplest Understanding of Reality?”, and the Foundational Questions

FIG. 3. Experimental results: showing the quantum and classical statistical complexity for the ferromagnetic case with nominal values of $J = 1$, $B = 0.3$ and a range of different temperatures $T$. (a) Statistical complexity of the $\epsilon$-machine simulating a 1D Ising spin chain. Error bars are derived from Poissonian photon statistics (some of the error bars are too small to be visible). The experimentally-determined quantum statistical complexity $C^q_m$ and $C^q_o$ (see text) are plotted against the relevant temperature parameters, $T^m$ and $T^o$ respectively, for the ferromagnetic case with $J = 1$, $B = 0.3$. The black curve shows the theoretical estimation of $C_0$ for the nominal values of $J$, $B$ and $T$. The grey shaded region shows the estimated theoretical bound, using the Ising model parameter distribution corresponding to the processes implemented by our real experimental map $E$ (see SM, Section 4). The orange curve is the theoretical classical statistical complexity, $C_\mu$. (b) Implemented temperatures; $T^s$ versus $T^m$ (The corresponding $B^m$ and $B^o$ have mean values of $0.29\pm0.02$ and $0.28\pm0.05$, respectively; the nominal (design) parameter is $B = 0.3$.) (c) Consistency graph (Equation (5b)), for the experimental data, $C^m_q$ and experimental $C^o_\mu$, in part (a). For $-1 < K < 0$ the models are ambiguous and for $0 < K < 1$ they are consistent. The pale shading in the plane $K = 0$ represents a projection of the experimental-result bars onto the plane, and together with the blue curve, demonstrates the boundary between regions of ambiguity and consistency. The pale blue squares in the plane $K = 0$ indicate the areas where no experimental data exist.
Institute (FQXi). 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.

*farzad.ghafari@griffithuni.edu.au
†g.pryde@griffith.edu.au

Supplemental material: Witnessing the ambiguity of simplicity via quantum simulations of an Ising spin chain

Farzad Ghafari,1,* Mile Gu,2,3 Joseph Ho,1 Jayne Thompson,3
Whei Yeap Suen,3 Howard M. Wiseman,1,4 and Geoff J. Pryde1,†
1Centre for Quantum Dynamics, Griffith University, Brisbane, 4111, Australia
2School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore 639793, Republic of Singapore
3Centre for Quantum Technologies, National University of Singapore,
3 Science Drive 2, Singapore, Republic of Singapore
4Centre for Quantum Computation and Communication Technology (Australian Research Council), Griffith University

1. 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 $T = \{0.75, 1, 1.25, 1.5, 1.75, 2.25, 2.75, 3, 4, 5, 6, 8, 10, 12, 14\}$. 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 [S 1]:

$$
\Gamma_{00} = e^{\frac{B+J}{D}}, \quad \text{(S.1a)}
$$
$$
\Gamma_{01} = 1 - \Gamma_{00}, \quad \text{(S.1b)}
$$
$$
\Gamma_{11} = e^{\frac{B-J}{D}}, \quad \text{(S.1c)}
$$
$$
\Gamma_{10} = 1 - \Gamma_{11}, \quad \text{(S.1d)}
$$

where

$$
D = \exp\left(\frac{J}{T}\right) \cosh\left(\frac{B}{T}\right) + \sqrt{\exp\left(-\frac{2B}{T}\right) + \exp\left(\frac{2J}{T}\right)} \sinh^2\left(\frac{J}{T}\right).
$$

2. Fixed-point states

In the ideal case defined in equation (4) in the main text, if we get measurement outcome $j$ with probability $\Gamma_{ij}$, then $\bar{\mathcal{E}}_j(\rho_i)$ will be $\Gamma_{ij}\rho_j$. (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 $\bar{\mathcal{E}}_0$ and $\bar{\mathcal{E}}_1$ are the experimentally-implemented maps which are characterized through quantum process tomography [S 2–4] performed on the one-qubit process.) However, in practice, a slightly different (but very close) output state $\rho^o(j|i)$ is obtained: it turns out that $\rho^o(j|i) \neq \Gamma_{ij}\rho_j(j|i)$, motivating a theoretical question: “Given map $\mathcal{E}$, can we find $\Gamma_{ij}$ and $\rho_i$ (for $i = 0, 1$) such that $\mathcal{E}_j(\rho_i) = \Gamma_{ij}\rho_j$,

exactly, for $j = 0, 1^n$”. 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

$$
\{\rho_i, \Gamma_{ij}\} = \arg \min_{\rho_i} \left\{ \sum_{i,j=0,1} \|\mathcal{E}_j(\rho_i) - \Gamma_{ij}\rho_j\|, \{ho_i, \Gamma_{ij}\}\right\}, \quad \text{(S.2)}
$$

where $\arg \min(f)$ means the arguments that minimize the function $f$, and $\|\ldots\|$ is the trace distance [S 5].

3. Deviation of implemented Ising parameters from nominal values

Implemented mean value of $B^m$, $0.29 \pm 0.02$, is close to the nominal value (0.3). However, for $T$, the nominal values differences with implemented ones increase when we aim for high temperatures. As an example, 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 $T$ could be seen in Fig. 3.b of the main text.

The difference between fixed-point states and nominal

* farzad.ghafari@griffithuni.edu.au
† g.pryde@griffith.edu.au
causal states of the Ising chain, that one wants to simulate, mainly comes from the difference in the nominal and implemented values of $T$, given that $J$ is fixed and the implemented $B$ is close to the nominal value. One can argue that the simulation of the systems with high temperature become difficult, which is valid because of an inevitable level of noise. As it can be seen in Fig. S1, for values close to 0, the implemented $T^m$ and $T^s$ versus nominal values almost follow a linear trend, while for higher temperature it starts to saturate.

4. Theoretical prediction for the statistical complexity of the real simulator

The simulator models an Ising model with temperature $T^m$ and magnetic field $B^m$, instead of the target values of $T$ ($T \in \{0.75, 1, 1.25, 1.5, 1.75, 2.25, 2.75, 3, 4, 5, 6, 8, 10, 12, 14\}$) and constant $B = 0.3$, respectively. Although the mean value for $B^m$ is close to nominal value, 0.3, there are still some discrepancies for individual values. For a particular value of $T^m$, there exists a value and uncertainty band for $B^m(T^m)$, resulting from fitting $B^m$ values as a function of $T^m$. The theoretically predicted quantum statistical complexity for the real simulator is given by $C_q(B^m(T^m), T^m)$ for $J = 1$. The $C_q$ values corresponding to the upper and lower bound of this fit, resulted in the grey bounds in Fig. 3a in the main text.

---

Chapter 3

Interfering trajectories in experimental quantum-enhanced stochastic simulation

3.1 Statement of contribution to co-authored published paper

This chapter includes a co-authored paper. The bibliographic details and publication status for this paper including all authors are:

My contribution to the paper involved some of the theory and most of the experiment, including:

- Working on some of the theory of simulating stochastic processes with a high-dimensional state
- Designing, building, characterising and optimising the experimental setup
- Acquiring the experimental data and analysing them
- Writing and editing the manuscript.

(Signed) ____________________________________ (Date) 11/02/2019
Name of Student: Farzad Ghafari Jouneghani

(Countersigned) __________________________ (Date) 11 Feb 2019
Supervisor and Corresponding author of paper: Geoff Pryde

(Countersigned) __________________________ (Date) 11/02/2019
Corresponding author of paper: Farzad Ghafari
Chapter 3. Interfering trajectories in experimental quantum-enhanced stochastic simulation

(Countersigned) 11/02/2019

(Date) 11/02/2019

Supervisor: Nora Tischler
3.2 Quantum simulation over multiple time steps

Quantum information offers a memory advantage in simulating stochastic processes. This advantage is practically meaningful, when we can simulate a process over many steps—stochastic processes often exhibit interesting behaviour over multiple time steps. However, the previous stochastic simulation demonstrations using quantum processors looked at only a single step \((68, 91)\). Therefore, the ability to simulate multiple steps is a crucial stepping stone for solving real-world problems. In this work, we use a photonic quantum information processor to demonstrate the first quantum simulator capable of simulating a Markovian process for multiple discrete time steps. The optical system that we use allows us to experimentally implement the quantum simulation over several consecutive time steps, generating a coherent superposition over all potential future trajectories of the stochastic process. Our approach, creating a coherent quantum superposition of the different trajectories over multiple time steps, is in contrast to measuring and repreparing quantum states at each time step of the simulation \((68, 91)\), presented in Chapter 2, which would increase the entropy of the simulator.

This coherent superposition is crucial for two reasons. Firstly, it conserves the entropy throughout the simulation, which is the key to achieving the quantum advantage in memory cost over multiple time steps. Secondly, using quantum interference, we can efficiently estimate how well the statistical futures from two independent processes coincide. This highlights the advantage of the coherent superposition, by showing how it enables us to perform a second key task beyond statistical sampling. Given two (potentially different) stochastic processes, the output of our simulator lets us estimate the overlap of the statistical futures of two stochastic processes via quantum interference, without measuring the future statistics directly. Estimating the overlap between vectors is an essential task in many fields, including machine learning \((189)\). To achieve this, we produced and controlled a high-dimensional quantum state. The capability of encoding a large amount of information on a multi-dimensional photonic state makes high-dimensional states important in the context of quantum information science.

To sum up, this work contains very interesting results: i) It is the first demonstration of a multi-step quantum simulation of a stochastic process, where the statistical complexity advantage is shown. ii) We used two simulators to estimate the overlap of the future output statistics of stochastic processes, using quantum interference instead of measuring the statistics themselves. iii) High-dimensional states are produced, controlled and manipulated. The interference between two discrete high-dimensional states is achieved with a very high-quality. Our interference visibility is the highest number reported for a discrete 16-dimensional state, to the best of our knowledge.
3.2.1 Interfering trajectories in experimental quantum-enhanced stochastic simulation

The remainder of this page is intentionally left blank. The paper appears on the following pages.
Interfering trajectories in experimental quantum-enhanced stochastic simulation

Farzad Ghafari,¹,* Nora Tischler,¹ Carlo Di Franco,²,³ Jayne Thompson,⁴ Mile Gu,²,³,⁴ and Geoff J. Pryde¹,†

¹Centre for Quantum Dynamics, Griffith University, Brisbane, 4111, Australia
²School of Physical and Mathematical Sciences, Nanyang Technological University, 639793, Singapore
³Complexity Institute, Nanyang Technological University, 639773, Singapore
⁴Centre for Quantum Technologies, National University of Singapore, 117543, Singapore

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 demonstrated interference of two 16-dimensional quantum states, representing statistical futures of our process, with a visibility of 0.96 ± 0.02.

Introduction

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 [1, 2]. 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 [3], cryptography [4], sensing [5, 6] and phase estimation [7].

This advantage was first illustrated for simulating a particular stochastic process, where past information was encoded within non-orthogonal polarization states of a single photon [8]. 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 [1]. 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, $p$, can differ from the probability of remaining in tails, $q$, as presented in Fig. 1. Repetition of this procedure generates a string of 0’s and 1’s, whose statistics defines the perturbed coin
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 \( j \mid T_{ij} \), 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, \( \{|S_i\}\).}

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. In general the amount of past information stored in memory is quantified by the Shannon entropy \( C = - \sum d_s \log d_s \), where \( d_s \) is the probability that the memory is in state \( s \) and 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 [9–12]. For the perturbed coin (see 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 [13]. For the case of the perturbed coin, the quantum simulator can be implemented as shown in Fig. 2. The past of the process is encoded in a quantum system that sequentially interacts with ancillary systems. All the ancillary systems start in a fixed state, and therefore do not contain any information. The sequence of interactions produces an entangled 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.

**Experimental implementation**

Motivated by recent realisations of quantum walks in linear optical setups with time-bin encoding [14–17], 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 a 16 dimensional (8 arrival time modes × 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- (here, 16-) dimensional quantum states can be encoded and manipulated in photonic temporal and polarisation modes with high fidelity [18, 19]. This complements other related works involving hybrid optical states using spatial (path and orbital angular momentum) and polarisation modes [20–23]. This also substantiates the oft-repeated claim that combining different photonic encodings [24, 25] is a practical tool for various quantum information tasks, for example studying the remote preparation of entangled states [26], complementarity [27], Bell inequalities [21, 28, 29], QKD implementations [30] and complete optical Bell state analyser [31, 32].

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 a 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 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 [33] for each set

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 $\pm 1.5$ 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 trajectories 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 for quarter-wave plate, GT for Glan-Taylor prism, PBS for polarising beam splitter, and FPC for 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.
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 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 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 [23].

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 require the internal dynamics to be close to (ideally, completely) unitary.

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 to 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 coin flip 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 initialized 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 having higher columns in the right corners of both graphs.
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 $\Pi_1$ 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 in the centre. One beam passes through the hole unaffected, while the other passes through the wave plate regardless of plate’s azimuthal orientation. Drilling holes in the wave plate does not affect the retardance of the remaining material. This arrangement allows 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 [34], which plays an essential role in machine learning tasks such as image recognition [35].

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 [8, 36]. Further, we showed that superpositions of statistical distributions of process outcomes can be interfered. These techniques have 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 realized in the context of communication complexity [37]. 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 realized 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 [34, 37]. 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 [38]. 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 [34]. 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 such as speech recognition [35].

Methods

Theoretical background

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

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 [10, 39]) only needs to store the class $\epsilon(\vec{\epsilon})$ that $\vec{\epsilon}$ belongs to. That is, given only $\epsilon(\vec{\epsilon})$ 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 probabili-
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 $C_{\mu}$, 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 ± 0.02. In the graph, the number of measured two-fold 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, including the one studied here, there exists a resource-intensive model. The resulting $\epsilon$-machine requires [33]

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

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 [10, 40, 41], 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 $C_q < C_{\mu}$ [1]. 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

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

where $\rho = \sum_i d_i |S_i\rangle \langle S_i|$. Three-step simulation of a perturbed coin

It is easy to verify that, for the perturbed coin process, the optimal quantum causal states can be written as [1]:

$$ |S_0\rangle = \sqrt{l} |0\rangle + \sqrt{1-l} |1\rangle, \quad (3a) $$

$$ |S_1\rangle = \sqrt{1-m} |0\rangle + \sqrt{m} |1\rangle. \quad (3b) $$

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

$$ \sum_{x_n} \sqrt{p(x_1, x_2, x_3|S_1, \Pi_1)} |x_1, x_2, x_3\rangle |S_2\rangle, \quad (4) $$

where $n = \{1, 2, 3\}$ and $p(x_1, x_2, x_3|S_1, \Pi_1)$ 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 $|S_1\rangle$. The value of $p$ can be evaluated theoretically from the transition probabilities between causal states (see 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 on 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 $\rho$:

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

(5)

where

$$d_0 = \frac{\sum_{x_1, x_2} p(x_1, x_2, x_3 = 0 | S_1, \Pi_1)}{\sum_{x_1, x_2} p(x_1, x_2, x_3 = 0 | S_0, \Pi_1)}$$

$d_1 = 1 - d_0$, and $S_{\text{pol}|S}$ is the tomographically reconstructed polarisation state at each arrival time conditioned that the input memory state is 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 $\frac{1-|\langle\psi|\phi\rangle|^2}{2}$, where $\langle\psi|\phi\rangle$ is the overlap of the two states. Therefore, one can use the HOM interference $v$ to estimate overlaps, by noting that $v = |\langle\psi|\phi\rangle|^2$. 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. i.e. for two different processes $\Pi_1$ and $\Pi_2$, let $|S_i\rangle$ be a causal state of $\Pi_i$, and $|T_j\rangle$ be a causal state of $\Pi_j$. Using Eq. (4), in general the overlap between the respective outputs of the quantum simulators for $\Pi_1$ and $\Pi_2$ will be

$$\sum_{x_n} \sqrt{p(x_1, x_2, x_3 | S_1, \Pi_1) p(x_1, x_2, x_3 | T_j, \Pi_2)} |S_{x_3} T_{x_3}\rangle.$$

(6)

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. (6), actually results in an overlap

$$\sum_{x_n} \sqrt{p(x_1, x_2, x_3, x_4 | S_1, \Pi_1) p(x_1, x_2, x_3, x_4 | T_j, \Pi_2)}.$$

(7)

I.e. 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 from the source, with the state $|0\rangle$ ($|1\rangle$) which is encoded as $|H\rangle$ = horizontal ($|V\rangle$ = vertical) polarisation. The first wave plate creates the desired initial causal state of our perturbed coin, either $|S_0\rangle_{\text{pol}} = \sqrt{\frac{1}{2}} |H\rangle_{\text{pol}} + \sqrt{\frac{1}{2}} |V\rangle_{\text{pol}}$ or $|S_0\rangle_{\text{pol}} = \sqrt{\frac{1}{2}} |H\rangle_{\text{pol}} - \sqrt{\frac{1}{2}} |V\rangle_{\text{pol}}$. 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, del), with the corresponding next causal state encoded in the polarisation:

$$|S_0\rangle_{\text{pol}} |0\rangle_{\text{del}} \rightarrow \sqrt{\frac{1}{2}} |S_0\rangle_{\text{pol}} |0\rangle_{\text{del}} + \sqrt{\frac{1}{2}} |S_0\rangle_{\text{pol}} |1\rangle_{\text{del}},$$

$$|S_1\rangle_{\text{pol}} |0\rangle_{\text{del}} \rightarrow \sqrt{\frac{1}{2}} |S_1\rangle_{\text{pol}} |0\rangle_{\text{del}} + \sqrt{\frac{1}{2}} |S_1\rangle_{\text{pol}} |1\rangle_{\text{del}}.$$
two paths, a path- (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 the right output arm, at the end of the first block (see 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 [42, 43].

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, \(t_3 = 8\) 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 (\(l = 0.397\) and \(m = \{0.101, 0.197, 0.297, 0.391, 0.490, 0.588, 0.685, 0.784, 0.882, 0.994\}\)). 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 statement

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

References


Acknowledgements

We thank Raj B. Patel for helpful contributions. This research was funded, in part, by the Australian Research Council (project no. DP160101911), the Singapore National Research Foundation (Singapore Fellowship No. NRF-NRF2016-02), the Singapore Ministry of Education Tier 1 RG190/17 and the Foundational Questions Institute ‘The role of quantum effects in simplifying adaptive agent’. 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.

Contributions

FG, NT and CDF designed the experimental setup; FG and NT performed the experiment and analysed the data. CDF, MG and JT conducted the theory of the
project, as well as contributing to the data analysis. GJP 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.

Competing interests

The authors declare no competing interests.
Chapter 4

Conclusive experimental demonstration of one-way Einstein-Podolsky-Rosen steering

4.1 Statement of contribution to co-authored published paper

This chapter includes a co-authored paper. The bibliographic details and publication status for this paper including all authors are:

My contribution to the paper involved some of the theory and the experiment, including:

- Building the source and the experimental setup
- Acquiring the experimental data and analysing them
- Conceiving the interpretation of the experimental results
- Preparing and editing the manuscript.

(Signed) ___________________________ (Date) ___________________________ 11/02/2019
Name of Student: Farzad Ghafari Jouneghani

(Countersigned) ___________________________ (Date) ___________________________ 11 Feb 2019
Supervisor and Corresponding author of paper: Geoff Pryde

(Countersigned) ___________________________ (Date) ___________________________ 11/02/2019
Supervisor: Nora Tischler
Chapter 4. Conclusive experimental demonstration of one-way Einstein-Podolsky-Rosen steering
Chapter 4. Conclusive experimental demonstration of one-way Einstein-Podolsky-Rosen steering

4.2 One-way Einstein-Podolsky-Rosen steering

Quantum non-locality is certainly one of the most interesting features of quantum mechanics. The notion of one-way EPR steering is a unique and fundamentally interesting asymmetric phenomena in quantum non-locality. In this experiment, we realise the first conclusive experimental demonstration of one-way EPR steering. Our experiment stands out because there is no assumption either on the shared two-qubit state between parties or on the measurements they can choose, unlike previous works (181, 186, 187).

To do this experiment, we designed and implemented a photonic Werner-state source, which is the best of its type. This was possible due to our new entangled single-photon source (185). The one-way steering criterion by Baker et al. (166) placed certain requirements on the state; the condition to be able to steer in at least one direction placed requirements on the correlations and the heralding efficiency, and the source and apparatus needed to satisfy these strict conditions.

Designing and building high-quality entangled single-photon sources with high heralding efficiency is in much demand. The entangled source used in the current work is also part of a long-term project in our group. It is a stepping stone for building more similar sources for experiments which need more than two photons, and require very high-quality interference between those photons. One of the milestones would be to build two of these sources and interfere photons from different sources with high visibility, preferably without spectral filtering, while maintaining exceptional heralding efficiency.

To emphasise the importance of this work: 1. It is the first conclusive test of one-way EPR steering. 2. We realised a very high-quality Werner-state source. 3. The entangled source, pumped by a picosecond pulse, is a cornerstone for building more entangled sources with high heralding efficiency that can be used in multi-photon experiments. 4. The same design is used to realise the four-photon unentangled source employed in the experiment presented in Chapter 5.

4.2.1 Conclusive experimental demonstration of one-way Einstein-Podolsky-Rosen steering

The remainder of this page is intentionally left blank. The paper appears on the following pages.
Conclusive Experimental Demonstration of One-Way Einstein-Podolsky-Rosen Steering

Nora Tischler,1 Farzad Ghafari,1 Travis J. Baker,1 Sergei Slussarenko,1 Raj B. Patel,1,2 Morgan M. Weston,1 Sabine Wollmann,1,3 Lynden K. Shalm,4 Varun B. Verma,4 Sae Woo Nam,4 H. Chau Nguyen,4 Howard M. Wiseman,1 and Geoff J. Pryde1,*

1Centre for Quantum Computation and Communication Technology (Australian Research Council), Centre for Quantum Dynamics, Griffith University, Brisbane, Queensland 4111, Australia
2Centre for Quantum Computation and Communication Technology (Australian Research Council), Quantum Photonic Laboratory, School of Engineering, RMIT University, Melbourne, Victoria 3000, Australia
3Quantum Engineering Technology Labs, H. H. Wills Physics Laboratory and Department of Electrical and Electronic Engineering, University of Bristol, Bristol BS8 1FD, United Kingdom
4National Institute of Standards and Technology, 325 Broadway, Boulder, Colorado 80305, USA
5Naturwissenschaftlich-Technische Fakultät, Universität Siegen, Walter-Flex-Straße 3, D-57068 Siegen, Germany

(Received 31 May 2018; revised manuscript received 2 August 2018; published 7 September 2018)

Einstein-Podolsky-Rosen steering is a quantum phenomenon wherein one party influences, or steers, the state of a distant party’s particle beyond what could be achieved with a separable state, by making measurements on one-half of an entangled state. This type of quantum nonlocality stands out through its asymmetric setting and even allows for cases where one party can steer the other but where the reverse is not true. A series of experiments have demonstrated one-way steering in the past, but all were based on significant limiting assumptions. These consisted either of restrictions on the type of allowed measurements or of assumptions about the quantum state at hand, by mapping to a specific family of states and analyzing the ideal target state rather than the real experimental state. Here, we present the first experimental demonstration of one-way steering free of such assumptions. We achieve this using a new sufficient condition for nonsteerability and, although not required by our analysis, using a novel source of extremely high-quality photonic Werner states.

DOI: 10.1103/PhysRevLett.121.100401

Introduction.—One of the most noteworthy and fundamental features of quantum mechanics is the fact that it admits stronger correlations between distant objects than what would be possible in a classical world. Quantum correlations can be categorized into the following classes, which form a strict hierarchy [1–3]: Entanglement is a superset of Einstein-Podolsky-Rosen (EPR) steerability, which, in turn, is a superset of Bell nonlocality. Out of these, steering is special in that it allows for, and, in fact, intrinsically contains, asymmetry. Steering is operationally defined as a quantum information task, where one untrusted party (for instance, called Alice) tries to convince another distant, trusted party (Bob) that they share entanglement. Bob asks Alice to make certain measurements on her quantum system (e.g., particle) and to announce the measurement outcomes but is not sure whether Alice answers honestly or indeed even has a particle. He also makes corresponding measurements on his particle and checks whether the correlations of their measurement outcomes rule out a so-called local hidden state model for his particle, thereby proving shared entanglement [1].

Interestingly, the steering task allows for the case of one-way steerable states, for which steering is possible in one direction but impossible in the reverse direction [4]. One-way steering is of foundational interest, since it is a striking manifestation of asymmetry that does not exist for entanglement and Bell nonlocality. It also has applications in device-independent quantum key distribution [5]. To observe one-way steering, one needs to demonstrate steering in one direction, by violating a steering inequality. In addition, one must establish that it would be impossible to achieve steering in the opposite direction. Our scheme, which allows for arbitrary measurements and rigorously takes into account losses and the real experimental quantum state, is illustrated in Fig. 1.

FIG. 1. Scheme for demonstrating one-way steering. A two-qubit quantum state is distributed to Alice and Bob, with a lossy channel on the way to Bob, such that his probability of obtaining his qubit is $\epsilon_B$. A detection-loophole-free steering test demonstrates that Alice can steer Bob’s state. At the same time, it is established that Bob cannot steer Alice’s state for any choice of measurements, based directly on the reconstructed experimental quantum state $\rho$ and the measured efficiency $\epsilon_B$. 

0031-9007/18/121(10)/100401(6) 100401-1 © 2018 American Physical Society
Since the question of whether one-way steering is possible was first raised in the seminal paper of Ref. [1], considerable progress has been made on the topic [2,4, 6–17]. First, the original question was answered in the affirmative, and this gave rise to the quest to fully understand and demonstrate the phenomenon. An overarching effort of these works has been the elimination of assumptions.

On the theory side, the ultimate, so far unattained, goal would be to establish practical necessary and sufficient conditions for the steerability of arbitrary quantum states using arbitrary measurements, which are described by positive operator-valued measures (POVMs). Examples of one-way steerable states have been identified assuming projective measurements [4,9,11] and for POVMs [2,8,10]. While specific example states provide conclusive proof that one-way steering is possible in principle, they are challenging to work with in real experiments. Real states in the laboratory generally deviate from the ideal target states, so the ability to account for these deviations is crucial. A practical, sufficient condition for the nonsteerability of arbitrary two-qubit states under the assumption of projective measurements is known [11]. Recently, a practical, sufficient condition for the nonsteerability of generic two-qubit states with loss was also established for restricted projective measurements [17] (see explanation in Supplemental Material, Sec. I [18]).

The elimination of assumptions has also been a key development on the experimental side. Several experiments relied on assumptions about the measurements. The first demonstration of one-way steering was restricted to the case of Gaussian measurements [7]. This was followed by a demonstration that was restricted to two-setting projective measurements [12] and another one assuming multisetting projective measurements [14]. In contrast to these post-selection-based experiments, an experiment by some of us and co-workers had no detection loophole, and, therefore, its analysis could take into account vacuum state contributions to the quantum state [10]. Also, unlike the previous experiments, it made no assumptions about the measurement. It did, however, make an assumption about the type of quantum state; an analysis for Werner states was applied to the experimentally achieved state, which exhibited a high fidelity with a Werner state. However, drawing conclusions based on high fidelities can be problematic, in general [21], and caution is also warranted for the case at hand [17] (see also Sec. V in Ref. [18]).

Here, we present the first fully rigorous experimental demonstration of two-qubit one-way steering. In one direction, we demonstrate the violation of a steering inequality with the detection loophole closed. Using the recent theory result of Ref. [17] and the new theory developed in Supplemental Material, Sec. I [18], we provide a sufficient condition for nonsteerability, valid for general POVMs performed on arbitrary two-qubit states with loss, and conclusively show that our state is not steerable in the opposite direction. We further demonstrate the impact of different experimental parameters, which highlights the delicate nature of experimental one-way steering. Although the formalism does not assume it, our experimental states are very close to two-qubit Werner states. Two-qubit Werner states comprise a one-parameter family of states written as $\rho_W = \mu |\Psi^+\rangle\langle\Psi^+| + (1 - \mu)/4 I$, where $|\Psi^+\rangle = (|01\rangle - |10\rangle)/\sqrt{2}$ is the singlet state and $I$ is the $4 \times 4$ identity matrix. These states represent a well-known example of mixed states [22], with their purity determined by the Werner state parameter $\mu \in [0,1]$.

A number of sources of photonic two-qubit Werner states have been reported in the past [23–27]. Here, we use a new type of photon source, producing high-quality states that have unprecedented fidelities with Werner states.

**Werner state source.**—Our photonic source of Werner states is based on spontaneous parametric down-conversion (SPDC) with a picosecond pulsed pump laser, producing photon pairs at a telecom wavelength, with the quantum state encoded in the polarization degree of freedom ($|H\rangle \equiv |0\rangle$, $|V\rangle \equiv |1\rangle$, where $H$ and $V$ stand for horizontal and vertical polarization, respectively). It is constructed as an incoherent superposition of a singlet state source and a source of maximally mixed photon pairs. Our design provides high heralding efficiencies and full control of the Werner state parameter $\mu$.

The detailed setup is illustrated in Fig. 2. A 775 nm pulsed laser with variable power and a pulse length of 1 ps acts as the pump for the two individual sources comprising the overall source. After passing through a focusing lens, the pump beam is divided between the two sources with a controllable splitting ratio by using a half-wave plate (HWP) and polarizing beam splitter (PBS).

The singlet state source is based on the design of Ref. [28] and essentially implements a superposition of two SPDC events within a beam displacer interferometer. The pump passes through a HWP that sets its polarization to an equal superposition of $H$ and $V$ components, which are then horizontally split into two beams by the first beam displacer (BD). The next two HWPs act to make the polarizations of both beams $H$, appropriate for the subsequent down-conversion process, while matching the path lengths of the two beams. The beams then pump the 15-mm-long periodically poled potassium titanyl phosphate (ppKTP) crystal in two places, enabling degenerate type-II SPDC. The second BD separates signal and idler photons vertically, resulting in a total of four down-converted photon beams for the one photon pair. The next three HWPs modify the polarizations of the beams such that the left two beams are $H$ polarized, while the right two beams are $V$ polarized. This allows overlapping the signal photon from the two different down-conversion beams with the third BD, and likewise for the idler photon. A $D$-shaped mirror separates the propagation directions of the signal and idler photon beams, each of which are collimated, have
the pump light filtered out with a long-pass filter, and are coupled into single-mode fiber. To transform the maximally entangled state of $|HH\rangle$ and $|VV\rangle$ to one of $|HV\rangle$ and $|VH\rangle$, a 90° polarization rotation for one of the two photons is implemented with in-fiber polarization controllers, and the phase $\phi$ of the target state $(|HV\rangle - e^{i\phi}|VH\rangle)/\sqrt{2}$ can be controlled through a slight tilting of the first BD or by adjusting the crystal temperature.

The design of the mixed state source is such that a separable photon pair is created, and then each photon is fully depolarized, yielding the target state $|\bar{i}\rangle/4$. The pump beam passes through a ppKTP crystal identical to the one in the entangled state source, creating one $H$ and one $V$ polarized photon, which are collimated with a lens. The two photons are vertically separated into two beams with a BD, and subsequently their polarization is rotated by 45° with a HWP. An imbalanced BD interferometer, in which one polarization component passes straight through and the other component undergoes spatial walk-off twice (in opposite directions), decoheres the polarization of each of the signal and idler photon completely. A long-pass filter discards the pump light, before the propagation directions of the signal and idler beams are separated with a D-shaped mirror and they are fiber coupled.

The two individual sources are mixed using 50:50 fiber beam splitters, which combine the signal photon contributions coming from the two sources, and likewise for the idler photon. This mixing is incoherent, since the path lengths through the two sources are sufficiently different. Finally, a bandpass filter in Bob’s arm narrows the biphoton spectrum (see Sec. III in Ref. [18]), which enhances the polarization state quality for the singlet source. By tuning the relative power of the pump in the two individual sources, the parameter $\mu$ can be controlled. For a range of relative power values, we perform quantum state tomography of the photon pairs using a combined pump power setting of $\sim75$ mW and determine the fidelities with the closest Werner states, as detailed in Table I. These fidelities are the highest reported values to date.

A further noteworthy feature of our source is its high heralding efficiency. Despite a 50% loss due to the mixing of the two individual sources via 50:50 beam splitters and the additional components in the measurement apparatus, we still obtain typical heralding efficiencies (defined as detected coincidences divided by the detected singles of the opposite arm, also called Klyshko efficiency [29]) of $0.3100 \pm 0.0003$ and $0.2345 \pm 0.0002$, for the arm without and with the bandpass filter, respectively.

### Table I. Tunability and quality of the experimental quantum state.

<table>
<thead>
<tr>
<th>$\mu$</th>
<th>0.9978 ± 0.0003</th>
<th>0.797 ± 0.001</th>
<th>0.603 ± 0.001</th>
<th>0.398 ± 0.002</th>
<th>0.198 ± 0.002</th>
<th>0.007 ± 0.002</th>
</tr>
</thead>
<tbody>
<tr>
<td>State fidelity</td>
<td>0.9981 ± 0.0002</td>
<td>0.9964 ± 0.0004</td>
<td>0.9983 ± 0.0002</td>
<td>0.9985 ± 0.0001</td>
<td>0.9986 ± 0.0001</td>
<td>0.9983 ± 0.0001</td>
</tr>
</tbody>
</table>
high heralding efficiencies are made possible by the choices of the pump beam waist, the detection beam waist, and high-efficiency superconducting nanowire single-photon detectors [30] (see also Sec. III in Ref. [18]).

One-way steering.—To demonstrate one-way steering, we use the same setup as before, with some minor modifications. To add controllable loss, we insert a multi-setting neutral density filter before the detection apparatus in Bob’s arm, which lowers his overall heralding efficiency to $\epsilon_B$. We also increase the total pump power to $\sim 300$ mW in order to maintain a sufficiently high signal-to-noise ratio with the attenuated beam against the detector dark counts, which are $\sim 100$ per second (see Sec. III in Ref. [18]). As shown in Fig. 3 and explained below, the output of our Werner state source together with the added loss creates one-way steerable states, provided that the values of $\mu$ and $\epsilon_B$ are suitably chosen. Note that, in the steering experiment, some of the fidelities with the closest Werner states are lower than the results shown in Table I, but our subsequent analysis is robust, as it makes no assumption of the experimental states being Werner states.

To demonstrate one-way steering, we perform two sets of measurements. The purpose of the first set is to show steering from Alice to Bob. This is done via a steering test with $n = 6$ measurement settings, using a platonic-solid measurement scheme [31]. Detection-loophole-free steering is demonstrated if the correlations of the measurement outcomes are sufficiently large, resulting in a steering parameter that exceeds the $n = 6$ steering bound (the definition of the steering parameter is provided in Supplemental Material, Sec. IV [18]). The bound is a function of Alice’s heralding efficiency $\epsilon_A$, because, in this task, she is the person who is attempting to steer her opponent’s state. Our experiment thus necessarily closes the detection efficiency loophole, though we make no claim to close the spacelike-separation loophole.

The purpose of the second set of measurements is to establish nonsteerability from Bob to Alice, for general POVMs. This is achieved via a quantum state tomography, through which our experimental density matrix is reconstructed. Based on the density matrix and Bob’s experimentally measured heralding efficiency, we test the criterion for nonsteerability derived in Supplemental Material, Sec. I [18]:

$$N_{\text{POVM}} \leq 1.$$  

Here $N_{\text{POVM}}$ is defined as

$$N_{\text{POVM}} = \max_{\mathbf{x} \in \mathbb{R}^3} \left[ (1 - 3\epsilon_B) |\mathbf{b} \cdot \mathbf{x}| + \frac{3\epsilon_B}{2} \left[ 1 + (\mathbf{b} \cdot \mathbf{x})^2 \right] + \|T\mathbf{x}\| \right],$$  

where $\mathbf{b}$ is Bob’s local Bloch vector, $T$ is the correlation matrix of the quantum state in its canonical form, and $\|\ldots\|$ denotes the 2-norm. The maximization is carried out over all unit vectors $\mathbf{x}$ in three dimensions. This criterion is stronger than that in Ref. [17], and its derivation (see Supplemental Material, Sec. I [18]) is more rigorous: It ensures nonsteerability from Bob to Alice without restricting Bob’s measurements to POVMs on the photonic qubit subspace.

Obtaining a steering parameter in one direction above the steering bound and showing, based on the density matrix and heralding efficiency, that the corresponding quantum
state is unsteerable in the opposite direction successfully demonstrates one-way steering. We perform the measurements for two sets of quantum states. In the first set, we keep the loss added by the neutral density filter fixed such that Bob’s heralding efficiency is $\varepsilon_B = (2.52 \pm 0.03) \times 10^{-3}$ while varying $\mu$. The results from the steering test are shown in Fig. 3(a), and the results from the corresponding test of the sufficient condition for nonsteerability in the opposite direction are depicted in Fig. 3(b). Of all the $\mu$ values shown, only one, marked by the red triangle, is conclusively one-way steerable [steering bound violation from Alice to Bob by 3.8 standard deviations (s.d.) and fulfilment of the sufficient condition for nonsteerability from Bob to Alice with a margin of 5.3 s.d.].

For the second set of quantum states, we keep $\mu$ fixed at $0.951 \pm 0.004$ while varying the loss added by the neutral density filter. The results of the steering test are given in Fig. 3(c), and the results from the corresponding test of the sufficient condition for nonsteerability in the opposite direction are shown in Fig. 3(d). Here, the two states corresponding to the lowest $\varepsilon_B$ values are conclusively one-way steerable (steering bound violation by 3.3 and 5.2 s.d. and nonsteerability with margins of 6.0 and 5.6 s.d., respectively). The states with higher $\varepsilon_B$ are no longer conclusively nonsteerable from Bob to Alice. The ability to further reduce $\varepsilon_B$ is limited for technical reasons only, namely, the decreasing signal-to-noise ratio due to dark counts, which reduces the ideally constant measured heralding efficiency $\varepsilon_A$ when the attenuation is very high.

The results highlight that, in practice, demonstrating one-way steering based on two-qubit states with loss requires a balance between (i) having sufficient correlations to observe steering in one direction while (ii) keeping the loss needed to conclude nonsteerability in the opposite direction at a technically feasible level.

**Discussion.**—Our experiment is based on a two-qubit state with loss. Implementing loss in a quantum information protocol is relatively straightforward. In fact, some amount of loss is generally unavoidable in practice, so, even if the loss were not actively leveraged, an experimental analysis would need to account for it in any case. Therefore, a two-qubit state with loss is well motivated from a practical perspective.

It is worth emphasizing that we establish nonsteerability by checking against our sufficient condition for nonsteerability [Eq. (1)]. This condition offers the best currently available method for demonstrating the nonsteerability of general two-qubit states with loss, allowing for general POVMs. However, the condition is not proven to be tight, so it is possible that tighter conditions will be found in the future. For example, it might be possible to show that the necessary and sufficient conditions for steerability coincide for projective measurements and POVMs, which would then make it easier to demonstrate one-way steering. However, the fact that we work with a sufficient condition for nonsteerability means that our results are conclusive now and will remain so, even in the event that tighter conditions are found in the future.

**Conclusion.**—In this work, we present a new, high-heralding-efficiency photon-pair source that produces quantum states with very large fidelities with two-qubit Werner states and provides full control of the Werner state parameter. We use the source and a new sufficient condition for nonsteerability to achieve a rigorous demonstration of two-qubit one-way steering free of previous limiting assumptions about the experimental quantum state or measurement.

This work was supported by the Australian Research Council (ARC) Centre of Excellence CE110001027 and CE170100012. F. G., T. J. B., and S. W. acknowledge financial support through Australian Government Research Training Program Scholarships. H. C. N. acknowledges support by the Deutsche Forschungsgemeinschaft (DFG) and the European Research Council (ERC) (Consolidator Grant No. 683107/TemIQ).

*g.prade@griffith.edu.au

See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.121.100401, which includes Refs. [19,20], for the derivation of the non-steerability criterion and more details about the experiment and its analysis.


I. THE SUFFICIENT CONDITION FOR NONSTEERABILITY

In this section, we provide a proof for the nonsteerability criterion of Eqs. (1) and (2) in the main text. For theoretical convenience, we consider quantum steering from Alice to Bob as in Ref. [17]; a simple permutation at the end allows us to obtain the criterion for nonsteerability from Bob to Alice as described by Eqs. (1) and (2) in the main text. In the following, we use PVMs to denote projective measurements, and n-POVMs to denote positive operator-valued measures of n outcomes. In fact, we work very often with the notion of 2-POVMs with a rank-1 projection component. Those 2-POVMs are of the form \((Q, I - Q)\) with \(I\) being the identity operator and \(Q\) being a rank-1 projection.

We start with repeating the proof of Lemma 1 of Quintino et al. in Ref. [2] in its more general form (see Ref. [19]).

**Lemma 1** (Quintino et al.). If a state \(\rho\) of \(\mathbb{C}^d \times \mathbb{C}^{d'}\) is nonsteerable for 2-POVMs with a rank-1 projection component, then the state
\[
\tilde{\rho} = \frac{1}{d} \rho + \frac{d - 1}{d} \sigma_A \otimes \rho_B
\]
with arbitrary state \(\sigma_A\) and \(\rho_B = \text{Tr}_A[\rho]\) is nonsteerable with arbitrary POVMs.

**Proof.** Since an extremal POVM \(E = (E_1, E_2, \ldots, E_n)\) has at most \(n = d^2\) nonzero components, we can fix \(n = d^2\). Moreover, we can assume that the components of the POVM \(E\) are rank-1, namely \(E_i = \alpha_i Q_i\) for some rank-1 projections \(Q_i\) and \(0 \leq \alpha_i \leq 1\) (since all other POVMs can be post-processed from these; see, e.g., Ref. [20]). In proving this lemma, it is convenient to rewrite a measurement \(E\) in the direct sum form \(E = \oplus_{i=1}^n E_i\). The steering ensemble—the set of Bob’s reduced states conditioned on \(E\)—is thus written as \(\oplus_{i=1}^n \text{Tr}_A[\rho E_i \otimes I_B]\).

Now we claim the following identity:
\[
\bigoplus_{k=1}^n \text{Tr}_A[\tilde{\rho} E_k \otimes I_B] = \sum_{i=1}^n \frac{\alpha_i}{d} \sum_{j=1}^n \frac{\beta_j}{d} \bigoplus_{k=1}^n \text{Tr}_A[\rho(\delta_{ik} Q_i + \delta_{jk}(I_A - Q_i)) \otimes I_B],
\]
where \(\beta_j = \text{Tr}(\sigma_A E_j)\). Note that \(\sum_{i=1}^n \frac{\alpha_i}{d} = \sum_{j=1}^n \beta_j = 1\). This identity can be proved straightforwardly by passing the sums over the direct sum and performing them explicitly. This identity tells us that the steering ensemble \(\oplus_{k=1}^n \text{Tr}_A[\tilde{\rho} E_k \otimes I_B]\) can be written as a convex combination of \(n^2\) steering ensembles \(\oplus_{k=1}^n \text{Tr}_A[\rho(\delta_{ik} Q_i + \delta_{jk}(I_A - Q_i)) \otimes I_B]\), each with probability \(\frac{\alpha_i \beta_j}{d^2}\). Note then that the latter ensembles \(\oplus_{k=1}^n \text{Tr}[\rho(\delta_{ik} Q_i + \delta_{jk}(I_A - Q_i)) \otimes I_B]\) correspond to steering \(\rho\) with 2-POVMs with a rank-1 projection component, \((Q_i, I_A - Q_i)\), when empty components are discarded. These ensembles can all be locally simulated from a local hidden state (LHS) ensemble by assumption. It follows that the former steering ensemble \(\oplus_{k=1}^n \text{Tr}_A[\tilde{\rho} E_k \otimes I_B]\) can also be locally simulated from the LHS ensemble.

In other words, \(\tilde{\rho}\) is nonsteerable with n-POVMs. \(\square\)

**Remark 1.** We can translate this mathematical proof to a physical one. Alice’s aim is to simulate steering of \(\tilde{\rho}\) with POVMs \(E = (\alpha_1 Q_1, \alpha_2 Q_2, \ldots, \alpha_n Q_n)\) (chosen by Bob). Alice provides Bob with the LHS ensemble that she can use to simulate steering of \(\rho\) with 2-POVMs with one rank-1 projection component. With probability \(\alpha_i \beta_j / d\), she chooses a pair \((i, j)\). She then simulates the outcomes \(i\) and \(j\) as if they are outcomes of \(Q_i\) and \(I_A - Q_i\) in the measurement \((Q_i, I_A - Q_i)\), respectively. This is slightly different from the protocol in Ref. [19], but the result is the same.

Suppose we have a two-qubit state \(\rho\) acting on \(\mathbb{C}^2 \otimes \mathbb{C}^2\). A state with loss can be described by a state on \(\mathbb{C}^3 \otimes \mathbb{C}^2\) as
\[
\rho_{\varepsilon_A} = \varepsilon_A \rho + (1 - \varepsilon_A) |v\rangle \langle v| \otimes \rho_B,
\]
where \(\varepsilon_A\) and \(\varepsilon_B\) are the loss parameters.
where $|\nu\rangle$ is the vacuum state, which is orthogonal to the standard qubit states $|0\rangle$ and $|1\rangle$ (see the main text and Ref. [17]). Here $\varepsilon_A$ is Alice’s heralding efficiency, $0 \leq \varepsilon_A \leq 1$.

In Ref. [17], some of the present authors considered in particular the steerability of $\rho_{x_A}$ with PVMs which are restricted to the form $(|\varphi_1\rangle \langle \varphi_1|, |\varphi_2\rangle \langle \varphi_2|, |\nu\rangle \langle \nu|)$, where $|\varphi_1\rangle$ and $|\varphi_2\rangle$ are two orthogonal qubit states (which are both orthogonal to the vacuum $|\nu\rangle$). These PVMs are not the most general PVMs, and will be referred to as restricted PVMs in the following. Let $\rho$ be written as

$$
\rho = \frac{1}{4} \left[ I \otimes I + \sum_{i=1}^{3} a_i \sigma_i \otimes I + \sum_{i,j=1}^{3} T_{ij} \sigma_i \otimes \sigma_j \right],
$$

(S.4)

where $a = (a_1, a_2, a_3)^T$ is Alice’s Bloch vector, and $T$ is the correlation matrix, which can be assumed to be diagonal. Theorem 1 in Ref. [17] states that $\rho_{x_B}$ is nonsteerable with restricted PVMs if

$$
\max_x \left[ (1 - \varepsilon_A) |a \cdot \hat{x}| + \frac{\varepsilon_A}{2} (1 + |a \cdot \hat{x}|^2) + \|T \hat{x}\| \right] \leq 1,
$$

(S.5)

where the maximization is taken over all unit vectors $\hat{x}$ of $\mathbb{R}^3$.

We now show that nonsteerability with restricted PVMs implies nonsteerability with all 2-POVMs with a rank-1 projection component. This then allows us to apply Lemma 1 to construct a state which is nonsteerable with arbitrary POVMs.

**Lemma 2.** Consider the two-qubit state with loss $\rho_{x_A}$ in (S.3). If $\rho_{x_A}$ is nonsteerable with restricted PVMs, then it is nonsteerable with all 2-POVMs with a rank-1 projection component.

**Proof.** Consider Alice making a measurement of the form $(|\psi\rangle \langle \psi|, I_A - |\psi\rangle \langle \psi|)$, where $I_A$ is the identity operator on Alice’s space $\mathbb{C}^3$, $|\psi\rangle$ is a state of $\mathbb{C}^3$, which may be nonorthogonal to the vacuum $|\nu\rangle$. It is sufficient to show that the steering outcome corresponding to $|\psi\rangle \langle \psi|$ can be locally simulated. We start with finding the steering outcome $\text{Tr}_A[\rho_{x_A}(|\psi\rangle \langle \psi| \otimes I_B)]$ for $|\psi\rangle \langle \psi|$, which is

$$
\varepsilon_A \text{Tr}_A(\rho |\psi\rangle \langle \psi| \otimes I_B) + (1 - \varepsilon_A) \langle \nu|\psi\rangle^2 \rho_B,
$$

(S.6)

where $I_B$ is the identity operator acting on Bob’s space. Since $\text{Tr}_B[\rho]$ has no support on $|\nu\rangle \langle \nu|$, we can insert the projection $\Pi = I_A - |\nu\rangle \langle \nu|$, which projects $\mathbb{C}^3$ to the two-qubit space $\mathbb{C}^2$, such that $\rho = (\Pi \otimes I_B)\rho (\Pi \otimes I_B)$. Therefore,

$$
\text{Tr}_A(\rho |\psi\rangle \langle \psi| \otimes I_B) = \text{Tr}_A(\rho \Pi |\psi\rangle \langle \psi| \Pi \otimes I_B).
$$

(S.7)

Now let $\Pi |\psi\rangle = r |\varphi\rangle$ with $|\varphi\rangle$ being a two-qubit state. Note that $1 = \langle \psi| (\Pi + |\nu\rangle \langle \nu|) |\psi\rangle$, so $|\langle \nu|\psi\rangle|^2 = 1 - r^2$. The steering outcome $\text{Tr}_A[\rho_{x_A}(|\psi\rangle \langle \psi| \otimes I_B)]$ therefore can be written as

$$
r^2 \varepsilon_A \text{Tr}_A(\rho |\varphi\rangle \langle \varphi| \otimes I_B) + (1 - r^2)(1 - \varepsilon_A) \rho_B.
$$

(S.8)

This is explicitly a convex combination of two steering outcomes $\varepsilon_A \text{Tr}_A(\rho |\varphi\rangle \langle \varphi| \otimes I_B)$ and $(1 - \varepsilon_A) \rho_B$. Trivially, the latter steering outcome $(1 - \varepsilon_A) \rho_B$ can be locally simulated. We therefore need only to show that the steering outcome $\varepsilon_A \text{Tr}_A(\rho |\varphi\rangle \langle \varphi| \otimes I_B)$ can be locally simulated. However, this is exactly one steering outcome of the steering ensemble made by the restricted PVM $(|\varphi\rangle \langle \varphi|, |\varphi\rangle \langle \varphi|, |\nu\rangle \langle \nu|)$, where $|\varphi\rangle$ is the two-qubit state that is orthogonal to the two-qubit state $|\varphi\rangle$ (and the vacuum $|\nu\rangle$). The latter can be locally simulated with a LHS ensemble by assumption. Therefore, $\varepsilon_A \text{Tr}_A(\rho |\varphi\rangle \langle \varphi| \otimes I_B)$ can indeed be locally simulated with the LHS ensemble and $\rho_{x_A}$ is nonsteerable with 2-POVMs with a rank-1 projection component.

Now if the two-qubit state with loss $\rho_{x_A}$ in Eq. (S.3) is nonsteerable with restricted PVMs, this lemma guarantees that it is also nonsteerable with 2-POVMs with one rank-1 projection component. Applying Lemma 1 with $\sigma_A = |\nu\rangle \langle \nu|$, we find that

$$
\frac{\varepsilon_A}{3} \rho + \left( 1 - \frac{\varepsilon_A}{3} \right) |\nu\rangle \langle \nu| \otimes \rho_B
$$

(S.9)

is nonsteerable with arbitrary POVMs. Now retrieving the criterion for nonsteerability with restricted PVMs (S.5), we find the sufficient condition for a two-qubit state with loss $\rho_{x_A}$ to be nonsteerable with arbitrary POVMs to be

$$
\max_x \left[ (1 - 3\varepsilon_A) |a \cdot \hat{x}| + \frac{3\varepsilon_A}{2} (1 + |a \cdot \hat{x}|^2) + \|T \hat{x}\| \right] \leq 1.
$$

(S.10)

This is exactly the criterion of Eqs. (1) and (2) in the main text, except that there, we considered steering from Bob to Alice, and as a consequence Alice’s Bloch vector $\mathbf{a}$ and heralding efficiency $\varepsilon_A$ were replaced by Bob’s Bloch vector $\mathbf{b}$ and heralding efficiency $\varepsilon_B$. 

67
II. DENSITY MATRICES

The density matrices that make up the data for Table I are displayed in Fig. S1.

Figure S1. Experimental quantum states from the Werner-state source that are used for Table I. The plots depict the real parts of the density matrices, obtained through quantum state tomography, in order of decreasing effective $\mu$ value. The absolute values of the imaginary parts are all below 0.02.

III. EXPERIMENTAL DETAILS

Details of optical elements in the setup of Fig. 2.— The ppKTP crystal has a poling period of 46.20 $\mu$m. The focal lengths of the lenses are $f_1 = 75$ cm and $f_2 = 25$ cm. The BDs are made of alpha-BBO, with a displacement of 1.4 mm for BD1, and 3.0 mm for BD2. The bandpass filter is centered at 1550 nm and has a full width at half maximum (FWHM) of 8.8 nm.

Heralding efficiencies.— The pump and detection beam waists are approximately 300 $\mu$m and 115 $\mu$m, respectively. If the aim was to achieve a fixed value of $\mu \neq 0.5$, it would generally be possible to further increase the heralding efficiency in the following way: Using beam splitters with a splitting ratio other than 50:50, the output of one of the individual sources (singlet state source or maximally mixed state source) would suffer more loss, while the output of the other one would undergo less loss. This asymmetry could be exploited to lower the loss for the singlet state or the maximally mixed state, whichever has the heavier weight in the Werner state. For the extreme cases where $\mu = 0$ or $\mu = 1$, this approach would lead to the use of fully reflective or transmissive “beam splitters”, which is intuitive since only one of the individual sources would be required. However, using splitting ratios other than 50:50, the overall heralding efficiency would depend on the value of $\mu$ because the heralding efficiencies from the two individual sources would not be equal, and $\mu$ determines the contribution from each. This means that the improved overall heralding efficiency at a given $\mu$ value would be achieved at the expense of a reduced tunability of $\mu$. We choose a 50:50 splitting ratio to maintain tunability over the full range of parameter values with fixed heralding efficiencies. Indeed, Figs. 3(a) and (b) show that $\varepsilon_A$ is nearly constant over the whole range of $\mu$ values.

Gating.— To reduce the effect of dark counts, we gate the detection of photons (singles and coincidences) based on emission events of the pump laser, by detecting in a 3 ns window around a synchronization signal from the laser. This results in average dark count values (per second) of 99 and 156 for the detectors in Alice’s arm, and 55 and 113 for the detectors in Bob’s arm.
IV. STEERING PARAMETER

The steering parameter for the number of measurement settings \( n = 6 \) is defined as \( S = \frac{1}{5} \sum_{k=1}^{6} \langle a_k \hat{\sigma}^B_k \rangle \). Here, \( a_k \in \{-1,1\} \) is the measurement outcome announced by Alice for those measurements where Bob has chosen the measurement setting corresponding to the Pauli observable \( \hat{\sigma}^B_k \) in the direction \( u_k \). The measurement setting is chosen out of a predetermined set of six options [31]. The steering bound is a nontrivial function of \( \varepsilon_A \) and is based on Alice’s optimal cheating strategy [31].

V. COMPARISON WITH RESULT FROM REF. [10]

In Ref. [10], the conclusion of one-way steering was reached based on an analysis for Werner states. To this end, the Werner states closest to the experimentally obtained density matrices were used. Enabling an analysis of general states, the work of Ref. [17] has since indicated that even small deviations from a Werner state can have a significant bearing on proving the nonsteerability of quantum states. Here, we use our new condition for nonsteerability (Eq. (1)) to compare one of our states with the quantum state from Ref. [10] that was thought to be one-way steerable for POVMs. We show that in contrast to our quantum state, the state from Ref. [10] does not conclusively meet our condition for nonsteerability, and reveal the key experimental improvements in the present work.

Figure S2. Comparison of the quantum state of Ref. [10] (in blue) with one of our one-way steerable states (in green), with regard to the sufficient condition for nonsteerability from Bob to Alice of Eq. (1). The lines represent the sufficient condition for nonsteerability, given the experimental \( \varepsilon_B \) values of 0.005 \( \pm \) 0.003 (blue) and 0.00269 \( \pm \) 0.00001 (green). The shaded regions, visible for the blue line but too small to be discernible for the green line, show the corresponding uncertainties, based on the uncertainties of the \( \varepsilon_B \) values. Each ensemble of points is associated with one quantum state, and the sufficient condition for nonsteerability (Eq. (1)) is equivalent to the whole ensemble being to the left of the corresponding line. An ensemble captures relevant properties of the quantum state, with individual points corresponding to specific choices of the unit vector \( \hat{x} \), of which a representative sample of 625 choices is shown. The choice of \( \hat{x} \) that constitutes the solution to the maximization in Eq. (1) is marked by the larger point symbol. For the quantum state from Ref. [10], the sufficient condition for nonsteerability is not satisfied with statistical significance (note the shaded region), whereas it is for our quantum state. The quantum state used as the one-way steerable state example from the present experiment is the one corresponding to the left-most data points in Fig. 3(c) and (d). For the case of an exact Werner state, the ensemble would collapse to a single point in the plot. An assumption made in Ref. [10] but not in this work, namely the mapping to a Werner state, is illustrated by the star. The uncertainty in its position, based on the uncertainty in the quantum state, is indicated by the horizontal error bar.
A useful way to visualize the relevant properties of a quantum state is through its correlation matrix $T$ and Bob’s local Bloch vector $b$, as detailed in Sec. I of this Supplemental Material. Specifically important are $||T\hat{x}||$, the norm of the correlation matrix multiplied by arbitrary unit vectors $\hat{x}$, and $b \cdot \hat{x}$, the projection of the reduced Bloch vector onto the same unit vectors, plotted in Fig. S2. If the points for all choices of $\hat{x}$ lie to the left of a bound, the state is nonsteerable from Bob to Alice. The bound depends on Bob’s heralding efficiency $\varepsilon_B$.

Fig. S2 shows our quantum state and that of Ref. [10], each with its corresponding bound calculated from the measured values of $\varepsilon_B$. Our bound is slightly to the right of the bound from Ref. [10] because we work with a lower value of $\varepsilon_B$, and this makes witnessing nonsteerability a little easier. However, the main differences between the two cases can be ascribed to the two ensembles of points. The ensemble corresponding to our state (i) lies at lower values of $||T\hat{x}||$, and (ii) has a smaller spread in $b \cdot \hat{x}$.

The lower values of $||T\hat{x}||$ are due to lower correlations in the state, also evidenced by a lower effective $\mu$ value. This generally helps to show nonsteerability from Bob to Alice, while at the same time making it more challenging to demonstrate steering from Alice to Bob. We have the option to operate at these lower correlation values and still violate a detection-loophole-free steering inequality from Alice to Bob, thanks to Alice’s high heralding efficiency. It should also be noted that it is the tunability of our source which gives us the freedom to move to this advantageous correlation condition.

The smaller spread in $b \cdot \hat{x}$ is an indication of a smaller reduced Bloch vector, which would be 0 for the ideal case of an exact Werner state. The reason that the spread in $b \cdot \hat{x}$ matters is because the bounds do not correspond to vertical lines in the plot.

In summary, the experimental improvement can be attributed to several factors: our ability to tune the correlations of our experimental quantum state, Alice’s higher heralding efficiency that lets us demonstrate steering from Alice to Bob with lower correlations, a better state quality in the sense of a smaller reduced Bloch vector, and Bob’s lower heralding efficiency.
Chapter 5

Dimensional advantage in the simulation of stochastic processes

5.1 Statement of contribution to co-authored published paper

This chapter includes a co-authored paper. The bibliographic details and publication status for this paper including all authors are:

My contribution to the paper involved some of the theory and most of the experiment, including:

• Working on of the theory of simulating a stochastic process with three causal states
• Designing, simulating, building, characterising and optimising the experimental setup
• Acquiring the experimental data and analysing them
• Writing and editing the manuscript.

(Signed) __________________________ (Date) 11/02/2019
Name of Student: Farzad Ghafari Jouneghani

(Countersigned) __________________________ (Date) 11 Feb 2019
Supervisor and Corresponding author of paper: Geoff Pryde

(Countersigned) __________________________ (Date) 11/02/2019
Supervisor and Corresponding author of paper: Nora Tischler
5.2 Quantum memory advantage using a single simulator

In this work (190), we present the first experimental stochastic simulation where a dimensional quantum memory advantage is attained, using a single simulator. This kind of memory advantage can find practical use even if the simulation is to be performed without many simulators running in parallel. The optimal classical simulator of the process we consider needs a three-dimensional memory register (requiring more than one bit to encode), whereas the quantum simulator requires only a two-dimensional memory (a qubit). In the previous works, the practical memory advantage as quantified by the statistical complexity, would be realisable in the scale of a very large number of simulators, whereas in the current work, the memory advantage is scale-independent—it can be realised with any number of simulators. In addition to this clear dimensional advantage, we also demonstrate an advantage for the statistical complexity.

The common measure of memory, statistical complexity, is based on information entropy. To determine the von Neumann entropy in a quantum simulator, quantum state tomography is needed, which is an extremely hard task, especially for high-dimensional states. However, to verify the dimensional memory measure (topological complexity), we just need to count the dimensions of the memory of the quantum simulator. Although the two measures are connected (the topological complexity is an upper bound for statistical complexity), verifying one is straightforward and the other is hard. Therefore, the dimensional memory advantage is of practical importance because it can be easily determined for stochastic simulators with high-dimensional memory registers.

To implement the simulator in the lab, we required sequential quantum gates with very high quality, for which a photonics implementation is well suited. To realise the four-photon experiment, we have also benefited from the source design which I explained in Chapter 4 and the introduction. Moreover, we devised a design which circumvents the challenge of concatenating probabilistic post-selected gates in this experiment by adding a non-demolition-measurement.

Accordingly, this work is important for two reasons: 1. It is the first demonstration of the dimensional memory advantage, which, remarkably, is a scale-independent advantage. 2. This new type of advantage is straightforward to verify, in comparison to the statistical complexity of the previous works, which needs tomographic reconstruction.

5.2.1 Dimensional advantage in the simulation of stochastic processes

The remainder of this page is intentionally left blank. The paper appears on the following pages.
Dimensional quantum memory advantage in the simulation of stochastic processes

Farzad Ghafari,1 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,1,6 and Geoff J. Pryde1,†

1Centre for Quantum Dynamics, Griffith University, Brisbane, 4111, Australia
2Centre for Quantum Technologies, National University of Singapore,
3Science Drive 2, Singapore, Republic of Singapore
4National Institute of Standards and Technology, 325 Broadway, Boulder, Colorado 80305, USA.
5Clarendon Laboratory, Department of Physics, Oxford University, Parks Road OXI 3PU Oxford, United Kingdom
6Centre for Quantum Computation and Communication Technology (Australian Research Council)

(Dated: January 22, 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 behaviour, 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 memory advantage using an individual simulator. This 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 memory registers of individual quantum simulators had the same dimensionality as their classical counterparts. Our photonic experiment thus establishes the potential of a new, practical resource saving in the simulation of complex systems.

INTRODUCTION

Stochastic processes are ubiquitous in science and technology [1, 2]. Quantum information reduces the required memory storage for simulating these processes [3–15]—a newly identified advantage [3] that complements other quantum information technological enhancements. Recent first experiments confirmed the existence of this advantage [9, 10]. The memory enhancement was with respect to an entropic information measure—however, a large number of parallel simulators would be required to exploit this benefit in a practical application. Also, verifying the entropic advantage requires quantum state tomography, with difficulty scaling exponentially with the problem size.

Here we realise the first experimental demonstration of a dimensional memory advantage for simulating stochastic processes. By “dimensional advantage” we mean that any individual quantum simulator uses a memory register with fewer dimensions than any classical counterpart. This is a scale-independent memory advantage, achievable with any number of simulators, rather than requiring an asymptotically large array of simulators [4, 16].

This realises the quantum advantage for stochastic simulation in its fullest sense. Moreover, characterising the dimensional advantage is relatively straightforward, compared to other measures of memory, which need to determine 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 it is theoretically known that the advantage holds for a range of other simulation tasks [14]. The process we simulate here can be understood as the output of a biased perturbed coin after post-processing [15](see Fig. 1A): at each discrete time step, the state of the coin is a probabilistic binary outcome, which depends on the parameters $p$ and $q$ that are defined by the process. Over multiple time steps, this produces a string of ‘zero’s and ‘one’s. Then, in post-processing, every ‘0’ that precedes a ‘1’ is replaced by a ‘2’. For classical simulation, this post-processing markedly increases the amount of past information that needs to be stored in order to generate future predictions. This is not so for quantum processors.

It is known that for the provably optimal simulators [15] in each class (classical or quantum) of this stochastic process (Fig. 1B), it suffices to classify any possible past into three different states called causal states [3, 17]. To this end, the classical processor must have three distinguishable states, $\{S_i\}_{i=0,1,2}$, as its memory. By contrast, as we experimentally demonstrate, the quantum processor works with the three required quantum states, $\{|S_i\rangle\}_{i=0,1,2}$, compressed into a two-dimensional quantum system.

RESULTS

Generally, as illustrated in Fig. 2A, 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 a unitary transforma-
The stochastic process and its simulation. (A) The perturbed coin process involves a coin in a box. At each step, the box is perturbed, which may or may not flip the coin. The probability of flipping from zero to one, $p$, can differ from the probability of flipping from one to zero, $q$, and similarly for the complementary probabilities of remaining in zero, $1-p$, and remaining in one, $1-q$. The process we study here is the post-processed data of the perturbed coin, which has three possible outputs at each time step, represented by the squares. The transition probabilities $T_{ij}$, $(i, j \in \{0, 1, 2\})$, between outputs $i$ and $j$ are the functions of $p$ and $q$ provided next to the arrows. These probabilities form the transition matrix. (B) The optimal classical simulator of the process uses causal states, as shown in the circles. The arrows represent transitions between different causal states, with the associated expressions $\begin{bmatrix} j \end{bmatrix} T_{ij}$ providing the classical output of the transition, $j$, and its probability $T_{ij}$. In this case a simple mapping exists from the past of the process to the appropriate causal state: the last output from the string of past outputs determines the causal state. The transition probability $T_{ij}$ is the probability of transitioning from causal state $i$ to $j$ while emitting $j$. The eigenvalues of the transition matrix form the probability distribution of the causal states, called the stationary distribution $\{p_i\}_{i=0,1,2}$. In the quantum case, the causal states become quantum states, $\{|S_i\rangle\}$. 

For the stochastic process of Fig. 1, the quantum memory required is a single qubit, in which the three causal states are encoded as three, non-mutually-orthogonal, pure quantum states, as described in Methods. We implement our simulator in a photonic quantum information processor. The memory qubit is encoded in the polarisation degree of freedom of a single photon. The non-trivial unitary transformations in our experiment include a mapping from the memory qubit to a qutrit space of three spatial modes (paths), followed by a controlled-\textsc{not} (C-\textsc{not}) [18, 19] and a controlled-rotation (C-rotation) gate, as detailed in Fig. 2C. 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 in the polarisation state of another photon, to the correct conditional state, which can be characterised by quantum state tomography.

We overcome constraints in the nondeterministic photonic implementation of consecutive quantum gates by introducing a non-destructive measurement realised by an additional C-\textsc{not} gate [20, 21] and a corresponding ancilla photon. The photons are generated via spontaneous parametric downconversion (SPDC) and four-fold 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 [22]) and coincidence logic modules. The detailed experimental setup is shown in Fig 3, and additional details are in Methods.

The first goal of the experiment is to verify that the quantum simulator is performing the intended simulation. For this, two criteria must be fulfilled: i) After initialisation 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). 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 the Methods section. 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 [23] 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 Methods) 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. A stochastic simulator can be used in different ways, with correspondingly different ways of analysing the memory use. The most straightforward use is as a single simulator. In this scenario, the memory size, in bits, is measured by the max-entropy, which is simply $\log_2 D$, where $D$ is the dimensionality of the memory system [4, 16]. Since the information about the past is encoded in the polarisation 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 quantum dimensional advantage in
FIG. 2. Conceptual diagram of a simulation step. (A) The quantum simulator accepts a memory qubit and an ancilla. (We use wavy lines to denote quantum objects, with the number of lines in parallel indicating the dimensionality.) The ancilla contains no information and its preparation, $P$, is fixed. The memory qubit undergoes a fan-out operation, $F$, after which the information is contained in a qutrit space. 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 realisation of the circuit in subfigure 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 non-destructive measurement (NDM). Then the unitary operation $U$ is performed by an additional two gates, C-NOT 2 and C-rotation. The preparation of the memory system (photon 3) $P(q)$, 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.

If multiple simulations are run in parallel, 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, 16]), the minimum required memory to replicate the process faithfully is given by $NC$, where $C$ is called the statistical complexity [24]. The classical statistical complexity [24], $C_\mu$, 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 Methods for mathematical definitions).

Fig. 4A illustrates the theoretically-expected statistical complexities $C_\mu$ and $C_Q$ for all possible values of $p$ and $q$, showing the potential for a significant 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 Fig. 4B-E, are determined from the density matrices of the output memory system and the transition probabilities (see Methods). The slight deviations of the experimental data compared to the theoretical curves arises from experimental imperfections such as reduced qubit purity from imperfect nonclassical interference, small imperfections and setting errors in polarisation-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.

Thus, our quantum simulator has an advantage over its classical counterpart both for the individual and i.i.d. cases. Remarkably, we even simulate processes, marked
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 polarisation 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 polarisation using HWPs. To implement the fanning out from the memory qubit to a qutrit, a HWP and polarising 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 realised via HWPs and partially polarising 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 maximise the visibility of the quantum interference. P stands for state preparation, SMF for single mode fibre, QWP for quarter-wave plate, GT for Glan-Taylor prism, and FPC for fibre polarisation controller. For more details, see Methods.

by the shaded regions in Fig. 4B-E, where the classical statistical complexity $C_\mu$ exceeds one bit. In these cases, we have a gap between both quantum measures and both classical measures: $C_Q < \log_2 2 < C_\mu < \log_2 3$. (Note that $\log_2 D$ always forms an upper bound on the Shannon or von Neumann entropy.)

**DISCUSSION**

The present experiment allows us to study both the dimensionality and the statistical complexity of the memory system. However, for more complex processes that entail high-dimensional memory systems, the quantum state tomography required for the estimation of the statistical complexity would require increased resources (such as photons, modes, detectors), and could become prohibitively time-consuming. In contrast, verifying a dimensionality advantage remains straightforward, because it is based on counting dimensions of a Hilbert space rather than characterising quantum states. We perform a single step of the simulation in our experiment, which is already sufficient for demonstrating a quantum advantage. In the future, it would be interesting to perform multiple simulation steps with a dimensional quantum advantage.

A natural question is to ask: what is the prevalence of such dimensionality advantage? While this remains an open question, its existence is certainly not isolated to the stochastic process in this experiment. Indeed, such advantage arises naturally in the context of processes that exhibit causal asymmetry—a memory overhead (in both dimensional and entropic memory costs) between predicting the future versus retrodicting the past [25]. All such processes lead to dimensionality advantage, and there exist families of processes where this advantage can grow without bound [15].

In conclusion, we have shown that quantum information processing enables the simulation of a stochastic process with a memory that is smaller both in terms of its dimensionality (the number of orthogonal states it can support) and its von Neumann entropy, compared to the optimal classical simulator, measured by the number of states it uses and the Shannon entropy, respectively. The demonstrated decrease in the dimensionality of the memory system establishes a new type of memory saving—namely a dimensional memory advantage. This advantage becomes possible when the system being simulated has at least three causal states, in contrast to previous works with only two causal states [9, 10]. This advantage is available at any scale, in principle at the scale of a single simulator, and multiple parallel simulators are not required. Finally, as counting the dimensions of the memory register is very simple, this advantage is easy to verify.

**MATERIALS AND METHODS**

**Stochastic processes.**

A stochastic process [1, 2] evolving in discrete time is a collection of random variables \{\ldots, X_{t-1}, X_t, X_{t+1}, X_{t+2}, \ldots\}, where the previously observed variables \{\ldots, 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 behaviour 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 [17].
A classical causal states are perfectly distinguishable states, and there are three of them for this process. The classical output $X_t$, is sufficient for determining the memory state for step $t + 1$ [24]. The possible memory states are called causal states [17, 24], 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, $\{\left| S'_i \right\rangle\}_{i=0,1,2}$, can be similarly defined as

$$
\left| S'_0 \right\rangle = \sqrt{1-\overline{p}} \left| 0 \right\rangle + \sqrt{\overline{p}} \left| 2 \right\rangle \\
\left| S'_1 \right\rangle = \sqrt{\overline{q}(1-\overline{p})} \left| 0 \right\rangle + \sqrt{\overline{1-q}(1-\overline{p})} \left| 1 \right\rangle + \sqrt{\overline{pq}} \left| 2 \right\rangle \\
\left| S'_2 \right\rangle = \left| 1 \right\rangle,
$$

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

$$
\left| S_0 \right\rangle = \left| 0 \right\rangle \\
\left| S_1 \right\rangle = \sqrt{\overline{q}} \left| 0 \right\rangle + \sqrt{1-\overline{q}} \left| 1 \right\rangle \\
\left| S_2 \right\rangle = \left| 1 \right\rangle,
$$

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

**Experimental details.**

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 $t + 1$ [24]. The possible memory states are called causal states [17, 24], 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, $\{\left| S'_i \right\rangle\}_{i=0,1,2}$, can be similarly defined as

$$
\left| S'_0 \right\rangle = \sqrt{1-\overline{p}} \left| 0 \right\rangle + \sqrt{\overline{p}} \left| 2 \right\rangle \\
\left| S'_1 \right\rangle = \sqrt{\overline{q}(1-\overline{p})} \left| 0 \right\rangle + \sqrt{\overline{1-q}(1-\overline{p})} \left| 1 \right\rangle + \sqrt{\overline{pq}} \left| 2 \right\rangle \\
\left| S'_2 \right\rangle = \left| 1 \right\rangle,
$$

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

$$
\left| S_0 \right\rangle = \left| 0 \right\rangle \\
\left| S_1 \right\rangle = \sqrt{\overline{q}} \left| 0 \right\rangle + \sqrt{1-\overline{q}} \left| 1 \right\rangle \\
\left| S_2 \right\rangle = \left| 1 \right\rangle,
$$

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

**Experimental details.**

Four photons are generated via SPDC, as shown in Fig. 3. For this, two SPDC sources are realised using a 775 nm Ti:sapphire picosecond-pulse-length pump laser and ppKTP (46.20 μm poling period) crystals cut for type-II collinear degenerate phase matching [26, 27]. The photons are not entangled in polarisation. The crystal temperature is controlled at 25°C by a temperature controller. The bandpass filter is centred at 1550 nm and has a FWHM of 8.8 nm.

To run the simulator, the causal states in equation (2) are encoded in the polarisation degree of freedom of a single photon acting as the memory system. We use polarisation modes such that $\left| 0 \right\rangle = \left| H \right\rangle$ and $\left| 1 \right\rangle = \left| V \right\rangle$, where $H$ and $V$ are horizontal and vertical polarisations, respectively.

The fan-out transformation implements the basis change from equation (2) to (1), so that the three paths correspond to orthogonal states $\left| 0 \right\rangle, \left| 1 \right\rangle, \text{and} \, \left| 2 \right\rangle$. The experimental setup contains non-deterministic two-qubit gates. The C-NOT gates 1 and 2 are realised with a HWP, a PBS, and post-selective detection. This simplified version (compared to a universal photonic C-NOT gate [18]) is adequate, since the photons in the two input spatial modes always have a fixed polarisation. The controlled-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. [19], which uses three partially polarising beam splitters (PPPBSs). However, we only require two because of the fixed polarisation in one of the input spatial modes. Four-fold 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 four-fold detection combinations are formed by coincidence detections between detectors from each of the following four sets (see Fig. 2C): $\{1\}$, $\{2,3\}$, $\{4,5,6\}$ and $\{7,8\}$. This implies that 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.

**Statistical complexity.**

The statistical complexity [17, 24, 25] is the minimal memory a model needs to generate future statistics correctly using only information from past observations. The classical statistical complexity is

$$C_p = -\sum \pi_i \log_2 \pi_i,$$

where $\pi_i$ 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]:

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

where $\rho = \sum \pi_i |S_i\rangle \langle S_i|$ is the quantum stationary state.

Our simulator implements the provably optimal model, the so-called quantum epsilon machine [3, 15, 24]. Therefore, we can measure $C_Q$ by inputting the causal states described in equation (2) for a given set of $p$ and $q$ values. The stationary state, $\rho$, is calculated as:

$$\rho = d_0 \sum_{i=0}^{2} T_{0i} S_{\text{pol}|S_0} + d_1 \sum_{i=0}^{2} T_{1i} S_{\text{pol}|S_1}$$

$$+ d_2 \sum_{i=0}^{2} T_{2i} S_{\text{pol}|S_2},$$

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

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

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

**Data availability**

The datasets generated during and analysed in the current study are available from the corresponding author on reasonable request.

* n.tischler@griffith.edu.au
g.pryde@griffith.edu.au


Acknowledgements

The authors would like to thank Thomas Elliott and Chengran Yang for useful discussions. This research was funded, in part, by the Australian Research Council (project no. DP160101911 and project no. CE170100012) F.G. acknowledges support by the Australian Government Research Training Program (RTP) scholarship. This work was partly supported by the National Research Foundation of Singapore and, in particular, NRF Awards no. NRF-NRFF2016-02, no. NRF-CRP14-2014-02, and no. RF2017-NRF-ANR004 VanQuTe, the Singapore Ministry of Education Tier 1 RG190/17 and the John Templeton Foundation Grant no. 54914. We acknowledge the traditional owners of the land on which this work was undertaken at Griffith University, the Yuggera people.

Author contributions

F.G. and N.T., with contributions from R.B.P., performed the experiment, acquired the data and analysed them. L.K.S. assisted with the source design. G.J.P. supervised the experiment. N.T., F.G., G.J.P., and H.M.W. conceived the experiment, based on theoretical work by M.G. and J.T. SNSPDs were fabricated by V.B.V. and S.W.N. The manuscript was written by F.G., N.T., J.T., M.G., R.B.P., H.M.W., and G.J.P. with contributions from L.K.S., V.B.V., and S.W.N.

Additional information

Competing financial interests

The authors declare no competing financial interests.
Chapter 6

Conclusion

In this thesis, I presented four experimental works on quantum complexity and steering. The common motivation for three of them was to build experimental setups that simulate classical stochastic processes in a way that classical computers cannot—quantum simulators can run with less memory storage. In my PhD, I also worked on designing and implementing a new single-photon source that can be used for different purposes in the field of optical quantum information processing. One example of these applications is the demonstration of one-way EPR steering. Both themes of my PhD have great potential to be extended, as I will explain below.

To summarise the present work in the context of simulating stochastic processes, in Chapter 2, the quantum simulation of a 1D Ising system was presented. In this work, a novel method, based on quantum process tomography, was implemented to deal with the inevitable experimental errors in quantum simulations. Moreover, we experimentally observed the ambiguity of simplicity, a phenomenon that is about the relative complexity of simulating two processes with classical or quantum encodings.

In Chapter 3, the task of simulating a classical stochastic process with a quantum information processor was taken to a multi-step level. We demonstrated the first quantum simulator that samples the statistics of a process over three time steps, using a temporal encoding. Practical examples are often complicated and need simulation beyond a single step, in such a way that the quantum coherence is preserved during the simulation. That is why the work presented in Chapter 3 is of practical application. We performed a second task beyond statistical sampling, namely the comparison of two statistical futures. This can be mapped to the estimation of the distance between two vectors. We estimated the overlap of statistical futures from two processes, without directly measuring the processes’ futures. In this experiment, two high-dimensional states were interfered with a very good visibility, which by itself is an interesting and useful technical development in optical quantum information science.

In Chapter 5, a dimensional quantum memory advantage is reported. In this work, we demonstrated that the exact quantum simulation is feasible using a memory with information encoded in fewer dimensions than is required for classical simulation. This is practically important because the dimensional memory advantage is achievable with only a single simulator, compared to a statistical complexity advantage,
which needs an infinitely large number of parallel simulators. Moreover, comparing the dimensionality of memory systems in different simulators is straightforward. In contrast, evaluating statistical complexity (with tomographic reconstruction) is not an easy task, especially for high-dimensional memory registers.

What is the way forward in the quantum simulation of stochastic processes? One direction would be to explore less abstract stochastic processes. To simulate a stochastic process, we need two things: first, to determine the unitary operation that performs the exact quantum simulation of the process, and, second, to realise this unitary in the lab with high fidelity. Regarding the former, there exists a systematic theory approach to find a unitary operator that implements the desired quantum simulator for a discrete-time stochastic process—including physical processes (67). Considering the latter, it has been recently shown that it is possible to implement an arbitrary two-qubit quantum unitary in a photonic setup (191). Therefore, for two-qubit problems, determining the required unitary of the simulator and implementing it are feasible, even with current technology. This paves the way for investigating other examples that have a large range of application in physics, or even other branches of science. This means that we have enough tools to simulate small-scale stochastic processes using the current primitive quantum computers. So far, all experiments in this context have been performed using photonics implementations. Realising quantum simulators with other physical platforms, such as ions and superconducting qubits, to advance this field, could be another interesting area of research.

In the multi-step simulation experiment, we used three consecutive unbalanced Mach-Zehnder-like interferometers to simulate three steps of the process. We have simulated a stochastic process for three steps. There is no in-principle reason that it can’t be extended further, given the excellent coherence. However, the addition of more simulation steps using our current design would lead to two issues. First, the implementation of additional Mach-Zehnder-like interferometers requires a very long optical path (e.g. almost 5 metres for the fourth interferometer), which is challenging, considering optical alignment and bulk-optics space limitations. This issue can be avoided by implementing the same idea using optical fibres. Second, in our post-selected approach, half of the photons are lost in each simulation step due to the probabilistic implementation. Considering the exponential decrease in the number of photons, it is not feasible to expand the current design to more simulation steps.

An extension of this line of research would be to realise the multi-step simulation in a deterministic design. Like our implementation, that would also be possible in a setup that uses single photons and time-bin encoding. However, instead of having $m$ consecutive interferometers for $m$ steps of the simulation, now a single step, with $m$ different possible paths that photons can take, such as delay lines in various fibre loops with different lengths, would be employed. A fast switch should be used to guide
photons to the relevant delay line.\footnote{A general scheme for implementing an arbitrary unitary operation, using time-bin encoding, is described in Reference (192).} If one could implement sufficiently fast switches with low loss, there would be no limitations on the number of simulation steps, \( m \). This brings up an opportunity for quantum simulators to simulate the dynamics of a stochastic process for a large number of steps, which might be hard by classical means, in terms of the requirement for the memory size.

The demonstration of the dimensional memory advantage, reported in Chapter 5, is a very important experimental work in the context of stochastic quantum simulation, as previously stated. One interesting step forward from our work would be to experimentally demonstrate the unbounded dimensional memory advantage. This would be possible in an experiment where the optimal classical simulator needs an infinite number of bits, while a quantum simulator can perform the exact simulation with a limited number of memory qubits. Through private communication with our colleagues, Mile Gu and Thomas Elliott, I am aware that there exists a family of processes that offers this advantage. A quantum simulation of these processes can be performed using only a qubit, while a classical simulator needs a memory of infinite dimension. A parameter like \( \delta t \), which tunes the non-Markovianity of the process such that decreasing \( \delta t \) increases the prediction accuracy, is probed. This parameter determines the time steps in which a process is observed, where \( \delta t \to 0 \) implies that the process is observed continuously. For quantum simulators, there is no increase in memory cost for increasingly accurate simulation—it always needs a single memory qubit and an ancilla qubit. However, for the classical simulator, by increasing the accuracy, the memory dimensionality grows infinitely. Given that the required unitary for simulation is experimentally realisable, the only concern about the quantum simulator is the experimental imperfections. In principle, this problem offers a quantum supremacy that no classical computer can achieve. However, in reality with experimental noise, we would reach a point where noise will affect the accuracy of prediction, and decreasing \( \delta t \) further would make no difference. Despite this issue, the problem is still very exciting to investigate because it provides a practical scenario to see how far current quantum simulators can go, in terms of the memory advantage that they can provide in the existence of experimental imperfections.

The above example indicates a case where the model requires an infinite number of causal states. The classical simulator uses a memory system with infinite dimensionality, while the quantum simulator only needs a qubit for its memory. In this example, an infinite number of causal states are mapped to a single qubit space. From an experimental point of view, in principle, it should be possible to find a design to show the memory advantage with two qubits (one as the memory, and one as the ancilla). In general, the memory advantage and the experimental implementation depend on the process and also on the unitary required for realising the quantum simulator.
Chapter 6. Conclusion

The field of quantum computational mechanics offers a new kind of quantum advantage, a memory advantage, which is different from other advantages studied before, such as exponentially reducing the computation time for a classical problem (6). Although quantum computational mechanics has great potential, it is still a young field and there are many unanswered questions. One of them is that exact replication (prediction) of a stochastic process becomes less relevant in the presence of experimental errors. It would certainly make sense to consider an extension of statistical complexity, or other measures of memory, where the future need only be replicated up to some statistical fidelity, and observe how this changes the quantum-classical divergence in memory requirements. So far, this direction has been unexplored, and there exists no systematic construction of optimal models for processes specified up to some given fidelity. Non-Markovian processes (66), phase-enhanced models (86), and the connection between stochastic modelling and many-body systems (90, 193) are also challenging research topics that have not yet been deeply investigated. Accordingly, quantum computational mechanics is a promising field, both theoretically and experimentally, which is in urgent need of future research.

In Chapter 4, our work on an efficient entangled source and a Werner-state source was presented. Our new source was designed in a way that can produce a very high-quality entangled state with high heralding efficiency. The generated Werner states were the best reported to date, in terms of fidelity to theoretical Werner states. We performed a demonstration of one-way EPR steering. This is a task which tests a type of quantum non-locality that does not need the extreme conditions required for the Bell-type correlations—the strongest form of quantum correlations. In this work (185), we have for the first time conclusively realised one-way EPR steering (181), where the theoretical condition makes no assumption about the measurements or the form of the underlying two-qubit state. We have also improved the non-steerability criteria of Reference (166) to include all projective measurements and POVMs.

One of the plans we have for the new source is to use it in experiments with more than two photons. Often, experiments of this type involve independent HOM interferences. Our calculations show that with our current design, we would reach a high visibility, approximately 84%, with no filters. The next step for our source would be to build another copy and perform the interference to check this visibility. The reason for this limited visibility is the phase matching function which is a sinc function with side lobes around the central wavelength. These side lobes limit the spectral separability, and therefore the spectral purity of the heralded photons. That means two heralded photons cannot achieve perfect visibility in an independent HOM interference. It would be possible to increase this visibility to 1 by using narrowband filters. However, they are not the ideal solution because they limit the brightness and the heralding efficiency of the source.

Our goal is to design and implement a source that can be used to achieve the
Appendix A

List of Co-Authored Papers published during PhD not included elsewhere

Bibliography

BIBLIOGRAPHY