universität wien

DISSERTATION / DOCTORAL THESIS

Titel der Dissertation /Title of the Doctoral Thesis

„Quantum Causality and the Indefinite Thermodynamic Arrow of Time“

verfasst von / submitted by

Giulia Rubino

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Doktorin der Naturwissenschaften (Dr. rer. nat)

Wien, 2020 / Vienna 2020

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	A 796 605 411
Dissertationsgebiet lt. Studienblatt / field of study as it appears on the student record sheet:	Physik
Betreut von / Supervisor:	Univ.-Prof. Dipl.-Ing. Dr. Philip Walther

Abstract

To date, all quantum mechanical experiments that have been performed agree with the assumption that quantum mechanical laws can be applied to any conceivable system. The working hypothesis of this thesis is that these laws can be extended to the notion of time, and, consequently, to the way in which two or more events can causally influence one another. This would have far-reaching consequences as our entire understanding of the world is based on the study of the causal relations between physical phenomena. The aim of this thesis was, hence, to investigate the consequences of the application of quantum mechanics to the causal relations among events and to the thermodynamic arrow of time.

The first part of this thesis concerns three studies falling under the umbrella of 'indefinite causality', *i.e.*, the notion according to which the causal structure between events may become genuinely indefinite. Such a scenario is expected to arise naturally in regimes wherein quantum mechanics and general relativity are both relevant and where the metric and, consequently, the causal structure may "fluctuate". In quantum optical experiments, indefinite causal structures can be realized by superposing the trajectories along which two events are occurring in alternative orders. Along this line, the first work of this thesis constitutes the first experimental demonstration of the indefinite causality of a process through the measurement of a 'causal witness', *i.e.*, a mathematical object designed to produce a certain outcome whenever a process is not consistent with a well-defined causal order. Following up on this first study, the second work of this thesis experimentally demonstrates indefinite causality outside the quantum formalism. This is achieved by showing the incompatibility of the experimental results with a class of generalised probabilistic theories complying with the assumptions of locality and definite temporal orders. The third work of this thesis looks beyond the concept of indefinite causality, to cover a variety of quantum superpositions of trajectories. Trajectories can be used as quantum control to regulate the order of different noisy communication channels, but this is not the only configuration in which the channels can be arranged. This third work, hence, experimentally compares different ways in which two trajectories can be superposed through a pair of noisy channels, demonstrating that this artifice allows for the transmission of quantum information even when standard quantum communication protocols (where a system travels along a well-defined trajectory) fail.

The second part of this thesis comprises two studies pertaining to the field of quantum thermodynamics. The linking element between the previous part of this thesis and the present one is rooted in the concept of thermodynamic arrow of time and its directionality. In fact, the second law of thermodynamics allows one to associate a positive (negative) entropy variation in a thermodynamic process with the temporal "forward" ("time-reversal") direction. The fourth work of this thesis, thus, proposes that quantum mechanics may permit quantum superpositions between thermodynamic processes yielding two opposite entropy variations. This would enable the existence of processes with a genuinely indefinite time's arrow. In more detail, this work focuses on understanding whether such superpositions entail any observable consequences, and how a well-defined temporal axis emerges upon performing

suitable measurements of entropy production. In this regard, this work shows that when very large quantities (in module) are observed in a measurement of the entropy production, this yields the effective projection of the quantum superposition of thermodynamic time's arrows onto a well-defined temporal direction. On the other hand, when small quantities of entropy production are at stake, interference effects play a prominent role in the definition of the nature of the thermodynamic process. For instance, they can lead to the observation of work probability distributions of a process that may be more or less reversible than the individual ones composing the superposition, or any classical mixture thereof. All these results revolve around the application of so-called 'thermodynamic fluctuation theorems'. Thermodynamic fluctuations relate the difference in free energy between two equilibrium states with the work performed on a system driven far from equilibrium. Since the definition of work in quantum contexts is a non-trivial concept, the last study of this thesis proposes a simple interferometric scheme that, leveraging the use of fluctuation theorems, enables a direct estimate of the work distribution and of the average work dissipated during an isothermal thermodynamic process.

The investigation of indefinite causal structures and of the arrow of time may enable novel quantum information and quantum thermodynamic tasks, and provide methodological tools for future quantum theories of gravity. To this end, proposing and implementing experimental approaches towards these goals, as undertaken in this thesis, may help to lay the groundwork for a deeper understanding of the concept of time and its role in major physical theories. In my view, this may be where future conceptual turning points will originate from.

Zusammenfassung

Alle bisher durchgeführten quantenmechanischen Experimente stimmen mit der Annahme überein, dass die Gesetze der Quantenmechanik auf jedes denkbare System anwendbar sind. Arbeitshypothese der vorliegenden Arbeit ist die Möglichkeit, diese Gesetze auf den Zeitbegriff auszudehnen, und damit auch auf die Art und Weise, in der zwei oder mehr Ereignisse sich gegenseitig kausal beeinflussen können. Dies hätte weitreichende Konsequenzen, da unser Verständnis der Welt auf der Erforschung kausaler Beziehungen zwischen physikalischen Phänomenen beruht. Ziel dieser Arbeit war es daher, die Folgen der Anwendung der Quantenmechanik auf die kausalen Beziehungen zwischen Ereignissen und auf den thermodynamischen Zeitpfeil zu untersuchen.

Der erste Teil dieser Arbeit befasst sich mit drei Studien, die unter den Begriff der "unbestimmten Kausalität" fallen, also der Vorstellung, dass die kausale Struktur zwischen Ereignissen in quantenmechanische Unbestimmtheit gelangen kann. (Es wird erwartet, dass ein solches Szenario in Systemen auftritt, für die sowohl die Quantenmechanik als auch die Allgemeine Relativitätstheorie relevant sind und in denen die Metrik und folglich die kausale Struktur "fluktuieren" kann). In quantenoptischen Experimenten können solche unbestimmten Kausalstrukturen durch Überlagerung von Trajektorien realisiert werden, entlang derer zwei Ereignisse in alternativen Ordnungen auftreten. In diesem Sinne bildet die erste Arbeit in der vorliegenden Dissertation den ersten experimentellen Nachweis der unbestimmten Kausalität eines Prozesses durch die Messung eines 'kausalen Zeugen', d.h. eines mathematischen Objekts welches ein bestimmtes Ergebnis erzeugen soll, wenn ein Prozess nicht mit einer definierten kausalen Ordnung übereinstimmt. Anknüpfend an diese erste Studie bezieht sich die zweite Arbeit auf den experimentellen Nachweis einer unbestimmten Kausalität außerhalb des Quantenformalismus. Dies wird erreicht, indem die Unvereinbarkeit der experimentellen Ergebnisse mit einer Klasse verallgemeinerter probabilistischer Theorien gezeigt wird, in denen die Lokalität und definierte zeitliche Ordnungen vorausgesetzt werden. Die als Drittes vorgestellte Arbeit erweitert nunmehr mit der Untersuchung einer Vielzahl von Quantenüberlagerungen von Trajektorien den Beobachtungsbereich über das Konzept der unbestimmten Kausalität hinaus. Es zeigt sich, dass Trajektorien für die Quantensteuerung der Ordnung verschiedener verrauschter Kommunikationskanäle anwendbar sind. Dies bildet jedoch nicht die einzige Konfiguration, in der diese letzteren angeordnet werden können. So wurden in dieser dritten Arbeit verschiedene Möglichkeiten für die Überlagerung zweier Trajektorien durch ein Paar verrauschter Kanäle experimentell verglichen. Es wurde gezeigt, dass alle diese Möglichkeiten die Übertragung von Informationen durch die Kanäle selbst dann ermöglichen, wenn standardmäßige Quantenkommunikationsprotokolle (bei denen sich ein System entlang einer wohldefinierten Trajektorie bewegt) versagen.

Der zweite Teil dieser Dissertation umfasst zwei Arbeiten, die sich auf das Gebiet der Quanten- Thermodynamik beziehen. Das Bindeglied zwischen diesem und dem vorherigen Teil der Arbeit findet sich im Konzept des thermodynamischen Zeitpfeils und seiner Richtungsabhängigkeit. Tatsächlich

erlaubt es der zweite Hauptsatz der Thermodynamik, eine positive (negative) Entropieänderung in einem thermodynamischen Prozess mit der zeitlichen “Vorwärts-” (“Zeitumkehrung”) Richtung zu assoziieren. So wird in der vierten Arbeit dieser Dissertation vorgeschlagen, dass die Quantenmechanik sogar Quantenüberlagerungen zwischen thermodynamischen Prozessen erlaubt, welche zwei entgegengesetzte Entropievariationen ergeben. Damit wird die Existenz von Prozessen mit einem unbestimmten Zeitpfeil ermöglicht. Im Detail konzentriert sich diese Arbeit auf die Frage, ob solche Überlagerungen beobachtbare Konsequenzen haben und wie eine wohldefinierte Zeitachse entsteht, wenn geeignete Messungen der Entropieproduktion durchgeführt werden. In dieser Hinsicht zeigt sich dass, wenn für die Entropieproduktion große Werte (im Absolutbetrag) gemessen werden, dies zu einer effektiven Projektion der Quantenüberlagerung der thermodynamischen Zeitpfeile auf eine wohldefinierte zeitliche Richtung führt. Geht es hingegen um kleinen Entropieproduktionswerte, so spielen Interferenzeffekte eine wichtige Rolle für die Definition der Art des thermodynamischen Prozesses. Sie können beispielsweise zur Beobachtung von Arbeitswahrscheinlichkeitsverteilungen eines Prozesses führen, der stärker oder leichter umkehrbar ist als die einzelnen Prozesse, welche die Überlagerung bilden, oder auch jede klassische Mischung daraus. Die letzte Studie dieser Dissertation schließlich knüpft an die in der vorhergehenden Arbeit verwendeten Methoden an. Bei all den genannten Arbeiten geht es um die Anwendung sogenannter ‘thermodynamischer Fluktuationssätze’. Mit Hilfe thermodynamischer Fluktuationen wird es möglich, die Differenz der freien Energie zwischen zwei Gleichgewichtszuständen mit der Arbeit an einem System in Beziehung zu setzen, welches aus einem Gleichgewicht getrieben wird. Im Kontext der Quantenmechanik bildet die Definition von Arbeit ein nicht-triviales Konzept. Daher wurde in dieser letzten Arbeit ein einfaches interferometrisches Schema aufgestellt, welches unter Ausnutzung der Fluktuationstheoreme eine direkte Schätzung der Arbeitsverteilung und der durchschnittlichen Arbeit, die während eines isothermischen thermodynamischen Prozesses dissipiert wird, erlaubt.

Die Untersuchung unbestimmter Kausalstrukturen und des Zeitpfeils führt zur Erschließung neuer Aufgabenstellungen in der Quanteninformation und der Quantenthermodynamik. Sie ist auch als methodisches Werkzeug für die Entwicklung zukünftiger Quantentheorien der Schwerkraft geeignet. Die theoretische Formulierung und experimentelle Realisierung dahingehender Ansätze, wie sie in der vorliegenden Arbeit unternommen werden, kann zur Grundlage für ein tieferes Verständnis der Rolle der Zeit in den grundlegenden physikalischen Theorien werden. In meinen Augen ist sie von zentraler Bedeutung für künftige konzeptionelle Wendepunkte in der Wissenschaft.

Acknowledgements

Being able to bring into existence the works contained in this thesis, and to learn from those who collaborated on them with me was an unforgettable experience, which enriched me as a scientist and as a person. I will always carry along with me a deep gratitude for those who have made this possible.

I am grateful to Philip for giving me what I most sought after in a supervisor: support and freedom. I think that a good supervisor should give their mentees the freedom to let their ideas thrive, but also the support to keep them from foundering. Philip has managed to do this, and to do it all with a good dose of smiles even in the grey Viennese Novembers. I feel truly thankful to him for having upheld my wish to experiment, both with experiments and with theory, with research areas, and with the wide variety of conferences which he let me partake in.

To Lee, for teaching me many of the things I learned to do in the lab, for lending an ear to my questions countless times, for fixing my bad sentences in English, and for his patience.

To Časlav, for passing on to me his deep love for science, for correcting the wrong things I said without ever belittling me, for having taught me physics both in front of the blackboard and in front of a beer, for telling me when he liked my talks but especially when he did not, for believing in me up to the point of venturing with me into research paths which were as new to him as they were to me.

To Fabio, for having been the first researcher with whom I could discuss all the questions I had on the foundations of quantum mechanics, and for introducing me to the research field which I now so deeply love.

To all my co-authors, especially Gonzalo, Mateús and Philippe, for all the discussions we had together, from which I benefited so much.

To the members of my group, especially Francesco, Marie-Christine, Maxime, Teo and Valeria, for all what I have learned from them over these years, but also for all the laughter and the time we shared.

To Andreas, Esra, Flaminia and Lorenzo, for having contributed to make my time in Vienna worthwhile.

To my parents, because they believe in me often more than I do.

Contents

Preamble

One of the most debated topics of modern science relates to one of the notions of which we first acquire an understanding in our lives: the notion of time. The questions that physicists have posed on this topic are among the most disparate, starting from the very conjecture of its existence [Rovelli, 2017]. In classical physics, time flows in a uniform way and events unfold in a well-defined background temporal order. It was later realized that time flows differently depending on how observers move and where they are, for example in relation to objects moving at relativistic speeds [Einstein, 1905], or in the vicinity of strongly gravitating matter [Einstein, 1915]. This has vast consequences, because of the way it affects the possibility of two or more events to have mutual causal relations, which is one of the features on which physical sciences rely in their understanding of Nature. Indeed, in general a pair of events can either be causally dependent (and hence be cause and effect of one another), or causally independent (no causal influence). Yet, in various approaches towards a theory of quantum gravity, it is argued that a 'fluctuating space-time' may imply indefiniteness in the causal relations [Butterfield and Isham, 2001], and thus the notions of time and causal orders among events might need to be radically revised. Understanding these phenomena is one of the greatest challenges faced by modern physics, with a major hindrance being that current technologies do not enable observations at the scales deemed necessary. It is then compelling to ask ourselves whether there is any way to conduct observations which corroborate the demand of revising our understanding of causality even at the low-energy scales.

Recently, researchers have undertaken an operational approach to address this task, applying concepts from quantum information to tackle it from a new perspective, and relate it to space-time physics [Chiribella *et al.*, 2013; Oreshkov *et al.*, 2012; Zych *et al.*, 2019]. Rather than being based on the quantities which are conventionally the core of most physical laws (*e.g.*, the position and momentum of a particle), here predictions are built according to the operations that two or more agents perform on a probe sample, and to their measurements' outcomes. If we assume that no signal can be exchanged between the two agents aside from the probe system itself (*i.e.*, the two agents act inside so-called 'local laboratories'), then the correlations which can be observed are of two types only: either they are non-signalling correlations (and hence there is no causal relation between them), or they are signalling correlations (hence one of the two operations must be in the causal past of the other). Either way, in standard classical and quantum mechanics, for every pair of events A and B, either A causes B, B causes A, or the two are independent from each other. However, recently a framework for quantum mechanics devoid of global causal order has been proposed [Oreshkov *et al.*, 2012], where the order between events can be genuinely indefinite. According to this framework, there exist bounds on the correlations among events which hold whenever these occur in a well-defined causal order. The bounds have been quantified through purposely designed 'causal inequalities' [Abbott *et al.*, 2016; Oreshkov *et al.*, 2012], and it has been shown that quantum mechanics may admit the existence of correlations which violate them.

To date, it is not clear whether these correlations are a mere mathematical artefact of the theory or if they can be observed experimentally. Nevertheless, there exist experimentally realisable schemes through which quantum superpositions of causal orders can be observed. The first explicit example of a process for which quantum mechanics predicts a genuinely indefinite causal structure[1] was proposed in 2012, and is called the 'quantum switch' [Chiribella *et al.*, 2013]. Restricting our considerations to the case of two agents exchanging one qubit of information (*i.e.*, an elementary carrier of quantum information), a quantum switch can be described as follows. The qubit is transmitted between the two agents, and the order in which the agents receive and act on it is quantum-mechanically controlled by a second system. Upon suitable measurements of the second system (*e.g.*, on a diagonal basis), and conditional on either of the two outcomes, it is possible to achieve a quantum superposition whereby operations are applied on the system in alternative orders.

A quantum switch was experimentally implemented for the first time in 2015 by superposing the order of two unitary operations [Procopio *et al.*, 2015]. That experiment was able to solve a specific computational problem more efficiently than any ordered quantum circuit, this indirectly indicated that the causal order between the performed operations was indefinite. The first objective of the present thesis was thus to devise and execute an experiment that would provide a *direct* experimental evidence of the causal non-separability of the quantum switch. In order to achieve this objective, we used a theoretical tool that had just been developed at the time, the *causal witness* [Araújo *et al.*, 2015; Oreshkov and Giarmatzi, 2016]. Loosely speaking, a causal witness is a purposely designed set of measurements whose outcome can reveal whether a given 'process' is causally ordered or not. To gain a first-hand understanding of what a process is, consider two observers A and B performing local operations on a target system. In this context, a process is what defines how the target system is exchanged between the two local laboratories (and it is hence independent of their individual operations). A 'causally-separable process' will then be a process wherein either A acts before B, B acts before A, or a classical mixture of the previous two [Oreshkov *et al.*, 2012]. Conversely, a process which does not obey these requirements is called 'causally nonseparable'. Since it has been proven that the set of causally-separable processes is convex [Araújo *et al.*, 2015; Oreshkov and Giarmatzi, 2016], one can always find a hyperplane which separates the set from any point outside it [Rockafellar, 1970]. The point outside the set would correspond to a causally-nonseparable process, and the separating hyperplane is a causal witness.

Chapter 1 of this thesis builds upon these concepts and presents the first experimental realization of a quantum superposition of orders of non-unitary channels, and the first measurement of a causal witness. In order to increase the significance of our results, it was important to adopt a witness as noise-resistant as possible. To this end, we chose to superpose the orders of a unitary and a measurement operation. Clearly, the execution of a standard measurement within the quantum switch

[1] The terms 'indefinite causality', 'indefinite causal structure', and 'indefinite causal order' all indicate a lack of well-defined causal order. They will thus be used interchangeably throughout this thesis. On the other hand, 'causal nonseparability' specifies this feature in the quantum framework.

could destroy the coherence between the elements of the superposition by revealing the time when the measurement is performed. To overcome this problem we developed a scheme wherein, although the measurement was performed inside the switch, the measurement results were encoded in some auxiliary degrees of freedom, and were read only "at the end" of the process, thereby preserving its coherence.

To provide experimental evidence of indefinite causality through the measurement of a causal witness is a result whose validity depends on multiple assumptions. Indeed, the construction of a causal witness rests on the assumptions that the system under consideration and the operations applied to it are described by quantum mechanics. Consequently, the results of the above experiment only hold as long as these assumptions are correct. So, after the completion of the first study, the question which we asked ourselves was whether it was possible to devise an experiment to certify indefinite causality whose validity would extend to a wider class of theories beyond quantum theory. Recurring to the analogy between a proof of indefinite causality and that of the entanglement of a quantum state, what we were aiming for is akin to a demonstration of the entanglement of a state through the violation of a Bell's inequality. Conversely, our first demonstration of causal indefiniteness was rather comparable to a measurement of an entanglement witness. Now, even though the quantum switch can be certified as causally-nonseparable via the measurement of a causal witness, it does not violate any causal inequality, and hence it cannot be used to draw conclusions on the lack of definite causal order in a 'device-independent' way.

A possible way around was proposed by Zych *et al.* [2019]. In their work, the authors sought a demonstration of indefinite causality which, despite not being device-independent (*i.e.*, independent of the internal functionality of experimental device), would nonetheless be theory-independent (*i.e.*, valid for a class of generalized probabilistic theories broader than quantum mechanics). To this end, the authors derived a Bell-like inequality for causal orders from the assumptions of locality of states and operations, and definite causal order. To test this inequality, they proposed to use a quantum superposition of a gravitational mass in two positions to create an entanglement between the causal orders of the operations performed by two pairs of agents (one close to each position of the gravitational mass). They then showed that a specific set of operations of the agents, together with a suitable projection of the state of the mass would lead to a violation of the Bell's inequality and, thus, to the negation of one or more of the assumptions.

Based on this result, the second aim of this thesis was to extend this novel Bell's inequality for causal orders to the case of a quantum optical experiment, and then to violate it experimentally. This extension and the corresponding results are reported in Chapter 2. To achieve a violation of the Bell's inequality for causal orders, we entangled the causal order of two quantum switches (with two pairs of agents in each, and a target state sent through each switch). The violation confirmed that the collected data could not be described by a class of generalized probabilistic theories assuming that the initial target states do not violate Bell's inequalities, that the operations on the target states are local,

and that they have a predefined order. Via this violation we could provide, as previously announced, an experimental demonstration of indefinite causality without resorting to the assumption that systems and operations should be described by the quantum formalism, that is, a theory-independent demonstration.

The two works presented above were aimed at experimentally demonstrating the existence of processes with a genuinely indefinite causal structure via "direct measurements", *i.e.*, the measurement of a witness, or the violation of an inequality). One can achieve such demonstrations also in "indirect ways", that is, by elaborating information-theoretical tasks which enable one to distinguish the class of causally indefinite processes from those outside the class, as was done, *e.g.*, in Ref. [Procopio *et al.*, 2015]. The most powerful of such examples is that, in the context of a distributed-calculation between three parties, the quantum switch provides an exponential reduction in the number of qubits communicated against causally-definite one-way quantum communication schemes [Guérin *et al.*, 2016]. This was experimentally verified by Wei *et al.* [2019]. The third objective of this thesis was to provide further evidence of indefinite causality by demonstrating another advantage that had just been proposed in the literature at that time, and which was gathering considerable interest: the enhancement of channel capacity in classical and quantum communication through noisy channels by placing the channels in a quantum switch [Chiribella *et al.*, 2018; Ebler *et al.*, 2018; Salek *et al.*, 2018].

Both a few other research teams [Goswami *et al.*, 2020; Guo *et al.*, 2020] and we could experimentally confirm this effect. Yet, we also realized that the quantum switch was not the only process allowing for such an enhancement. In particular, we noticed that causal processes which did not qualify as causally indefinite (*i.e.*, whereby channels are disposed in series and preceded by quantum-controlled operations) could provide the same, or even better advantages [Guérin *et al.*, 2019]. Independently of us, another group came to similar conclusions around that time. While they did not refer to a causal process, they showed that even quantum superposing two parallel channels leads to the proposed enhancement [Abbott *et al.*, 2020]. This finding naturally changed the kind of objective that our experimental verification was intended to achieve. Indeed, as the enhancement could not be used to give evidence of indefinite causality (since this is arguably not a distinctive feature thereof), we focused on identifying the actual key to achieving such an advantage. As presented in Chapter 3, we analysed the nature of the advantage by comparing all three schemes mentioned above: the quantum switch, the quantum-control of parallel channels and the channels in series with quantum-control operations. To this end, we compared the three schemes experimentally, applying them to various noise models. We realised that the common resource among the three schemes is the establishment of a coupling between the trajectories of the information carriers and the degree of freedom on which the noise acts. To extend our findings to a broader spectrum (*i.e.*, beyond the three noise models experimentally investigated), we carried out a numerical evaluation of the enhancements which can be achieved with the three schemes in the case of randomly-generated channels. Both the experimental tests and the simulations showed that, on average, arranging the channels in series and performing

quantum-controlled operations before them provides a better quantum communication enhancement than the other two methods.

Up until this point, we have discussed quantum superpositions between causal orders, motivated by recent developments in our understanding of the notion of causality in quantum mechanics, as well as by the prediction of the existence of a fluctuating space-time by different approaches to a theory of quantum gravity. Even though the above results allowed us to draw conclusions on the causal orders between events and their quantum superposition, they did not lead to formulating broader conjectures on the existence of a well-defined arrow of time in quantum mechanics. In fact, while the causal orders between events occurring in each laboratory were indefinite, the temporal direction of the operations carried out inside the laboratories remained well-defined. In this light, after the works presented in the first three Chapters of this thesis, we undertook a paradigm shift from the sheer notion of indefinite causality, and we focused on the study of opposing temporal directions and their quantum superpositions.

To discuss quantum superpositions between alternative temporal directions, it is necessary to refer to some physical definition of the time's arrow. The definition of an 'arrow of time' is inherently associated to the observation of time-asymmetric changes in physical processes. Such physical phenomena are numerous. In the present thesis, we have chosen to adopt a thermodynamic approach, according to which the directionality of the time's arrow can be defined using the second law, which states that the total entropy of the universe can only either increase or remain constant. At a macroscopic level, we have a clear perception of the flow of time, and the probability to observe a negative entropy production in this regime is negligible. However, this is not equally true at the microscopic level, where it has been shown that it is possible to observe so-called "fluctuations" in the entropy production. In particular, while the second law of thermodynamics states that the entropy of an isolated system tends to increase until the system reaches an equilibrium, this is only true on average —in other words, the second law is only a statistical law. So, there is a non-zero probability that the entropy of an isolated system spontaneously decreases. A positive (negative) entropy production is linked to the "forward" ("time-reversal") temporal direction. This probability may be quantifiable through the so-called "fluctuation theorems": in a system far from equilibrium, the ratio between the probability that the entropy production assumes a positive value and that of it assuming a negative value increases exponentially with the size of the entropy production. If quantum mechanics is incorporated in this picture, the definition of a temporal orientation becomes even more complex. Indeed, quantum theory allows, at least in principle, for the existence of superposition between processes with opposite entropy production. In this context, the questions which brought us to the formulation of Chapter 4 of this thesis were the following. Suppose that superpositions of alternative time's arrow exist in Nature, what could preclude us from their observation? Could their observation be hindered by an effective projection of the quantum superposition in a well-defined time direction? If so, which physical phenomenon could constitute experimental evidence for the existence of such superpositions?

To answer these questions, we formalized a method to study quantum superpositions between thermodynamic quenches, and we applied it to superpositions between quenches causing opposite entropy productions. Our approach builds upon the so-called 'two-point measurement' (TPM) scheme [Campisi *et al.*, 2011]. This is a technique allowing one to evaluate the work distribution of a given thermodynamic quench by reconstructing the difference in energy before and after the quench through ideal projective measurements of the Hamiltonian [Talkner and Hänggi, 2016; Talkner *et al.*, 2007]. From the value of the work and the variation of free energy, it is then possible to reconstruct the so-called 'dissipative work', which is equal to the product between the entropy production and the temperature of the system. In Chapter 4, we propose an extended TPM scheme where *i)* the forward quench and its time-reversal twin (each preceded and followed by projective measurements of the Hamiltonian) are quantum superposed, and *ii)* the auxiliary degree of freedom, upon which the superposition is built, is projected onto the diagonal basis. It should be noted that the projective measurements of the Hamiltonian must maintain the coherence between the forward and time-reversal thermodynamic processes. To achieve this, we adopted a technique based on the experiment performed in Chapter 1: we encoded the results of the measurements in one state of the auxiliary system and preserved the coherence between the auxiliary states until the end of the overall thermodynamic process, before finally interfering the different states with each other. With this technique, we demonstrated that quantum measurements of the dissipative work (or, equivalently, of the entropy production) yield a well-defined orientation of the time's arrow. More precisely, when the entropy production is much greater than the system's temperature (setting, for simplicity, the Boltzmann's constant to one), the superposition is effectively projected onto the forward process, whereas it is projected onto the time-reversal one when the entropy production assumes the opposite sign. This effective projection restores a definite thermodynamic arrow of time. Conversely, when the amount of entropy produced is of the order of the system's temperature, the forward and time-reversal thermodynamic processes can interfere, giving rise to work probability distributions with no classical counterpart. The observation of such work distributions would prove the existence of superpositions between opposite time's arrows. Furthermore, these considerations can be applied not only to superpositions between processes linked by time inversion, but to any pair of thermodynamic processes. With this in mind, we have also applied our results to the quantum superposition between a heat engine and a power-driven refrigerator, showing that our findings can also be harnessed to reduce undesired fluctuations in the performance of state-of-the-art quantum thermodynamic engines.

While working on this last theoretical result, we realized that a modification of our extended TPM scheme could provide a new method to estimate the work distribution. In particular, in quantum mechanics, since work is not associated with any observable [Talkner *et al.*, 2007], its definition is rather complex. Commonly, it is measured through the previously introduced TPM scheme, and this approach has been implemented in several experiments [An *et al.*, 2015; Wu *et al.*, 2019; Xiong *et al.*, 2018; Zhang *et al.*, 2018]. However, since the realisation of projective energy measurements before and after an arbitrary process may be difficult in some experimental scenarios, plus the projective

measurements may ultimately destroy the system, alternative methods have been proposed [Dorner *et al.*, 2013; Mazzola *et al.*, 2013]. Yet, the price to pay with these alternative schemes is the need to use entangling operations, which may be experimentally challenging.

Our proposed method takes advantage of the interference between two paths, one along which the system is driven out of its thermal equilibrium through the forward process, and one in which it is driven out of it through the time-reversal process. In Chapter 5, we show that injecting the initial and final energy eigenstates of the system into the two paths of the interferometer and measuring the fringe visibility at the output of the interferometer allows one to directly reconstruct the work distribution and the average dissipative work. The latter is proportional to the average entropy production, and it is a measure of thermodynamic irreversibility. In the case of a limited experimental control, when only the thermal states of the initial and final Hamiltonian of the system can be prepared, our method does not yield an exact estimation of the previous quantities but still provides useful upper limits on the average dissipative work. Our scheme does not involve any entangling operation nor energy measurements, and therefore it appears to offer a convenient playground for exploring the role of quantum coherence in thermodynamic processes.

In conclusion, this thesis is aimed at undertaking a few steps towards a deeper understanding of causal structures and the concept of time in quantum mechanics through an operational approach to these questions pursued with both experimental and theoretical analyses. Clearly, each step forward opens the way to countless further questions. On the one hand, if we consider the field of indefinite causality, the most puzzling open question is whether a physical process able to violate causal inequalities can be found. This ambitious goal, in turn, encompasses a number of additional questions, such as whether there exist physically realisable causally-nonseparable processes which are not traceable back to the quantum switch, and whether such processes can lead to further advantages in quantum communication and computation. On the other hand, as far as the question about the existence of a thermodynamic time's arrow in quantum mechanics is concerned, in this thesis we show some immediate consequences of superposing processes with opposite temporal directionality, and the effects of carrying out measurements on such superpositions. Along this line, further steps to be taken are the devising of tests, independent of the quantum theory, proving that the latter does not foresee the existence of a well-defined thermodynamic arrow of time or, on a more practical side, the application of the aforementioned results to the elaboration of further enhancements of the performance of quantum thermodynamic engines.

List of Publications

1. **Experimental Verification of an Indefinite Causal Order**

   G. Rubino, L. A. Rozema, A. Feix, M. Araújo, J. M. Zeuner, L. M. Procopio, Č. Brukner, and P. Walther

   Science Advances **3**, no. 3 e1602589 (2017)

2. **Communication through quantum-controlled noise**

   P. Allard Guérin, G. Rubino, and Č. Brukner

   Phys. Rev. A, 99, 062317 (2019)

3. **Experimental Quantum Communication Enhancement by Superposing Trajectories**

   G. Rubino, L. A. Rozema, D. Ebler, H. Kristjánsson, S. Salek, P. Allard Guérin, A. A. Abbott, C. Branciard, Č. Brukner, G. Chiribella, and P. Walther

   Accepted for publication in *Phys. Rev. Research* on 15.12.2020

Submitted:

- **Experimental Entanglement of Temporal Orders**

  G. Rubino, L. A. Rozema, F. Massa, M. Araújo, M. Zych, Č. Brukner, and P. Walther

- **Time's Arrow of a Quantum Superposition of Thermodynamic Evolutions**

  G. Rubino, G. Manzano, and Č. Brukner

In preparation:

- **Inferring Work by Quantum Superposing Forward and Time-Reversal Evolutions**

  G. Rubino, G. Manzano, L. A. Rozema, P. Walther, J. M. R. Parrondo, and Č. Brukner

- **Entropy Production and the Second Law under Coarse-Graining**

  G. Rubino, G. Manzano, and Č. Brukner

Chapter 1

Experimental Verification of an Indefinite Causal Order

G. Rubino, L. A. Rozema, A. Feix, M. Araújo, J. M. Zeuner, L. M. Procopio, Č. Brukner, and P. Walther

Abstract. Investigating the role of causal order in quantum mechanics has recently revealed that the causal distribution of events may not be a-priori well-defined in quantum theory. While this has triggered a growing interest on the theoretical side, creating processes without a causal order is an experimental task. Here we report the first decisive demonstration of a process with an indefinite causal order. To do this, we quantify how incompatible our set-up is with a definite causal order by measuring a 'causal witness'. This mathematical object incorporates a series of measurements which are designed to yield a certain outcome only if the process under examination is not consistent with *any* well-defined causal order. In our experiment we perform a measurement in a superposition of causal orders—without destroying the coherence—to acquire information both inside and outside of a 'causally non-ordered process'. Using this information, we experimentally determine a causal witness, demonstrating by almost seven standard deviations that the experimentally implemented process does not have a definite causal order.

Author contributions: G.R., L.A.R., M.A., A.F., and L.M.P. designed the experiment. G.R. built the setup, and G.R. and L.A.R. carried out data collection. G.R., L.A.R., A.F., and M.A. performed data analysis. J.M.Z. designed and built the automated components. G.R. and M.A. created the figures. P.W. and C.B. supervised the project. All authors contributed to writing the paper, based on an initial draft by G.R..

1.1 Introduction

The notion of causality is an innate concept, which defines the link between physical phenomena that temporally follow one another, one manifestly being cause of the other. Nevertheless, in quantum mechanics the concept of causality is not as straightforward. For example, when the superposition principle is applied to causal relations, situations without a definite causal-order can arise [Chiribella *et al.*, 2013; Oreshkov *et al.*, 2012]. Although this can lead to disconcerting consequences, forcing one to question concepts that are commonly viewed as the main ingredients of our physical description of the world [Brukner, 2014], these effects can be exploited to achieve improvements in computational complexity [Araújo *et al.*, 2014; Chiribella, 2012; Hardy, 2009] and quantum communications [Baumeler and Wolf, 2014; Feix *et al.*, 2015; Guérin *et al.*, 2016]. Recently, this computational advantage was experimentally demonstrated in [Procopio *et al.*, 2015]. However, the absence of a causal order was inferred from the success of an algorithm rather than being directly measured. In the present work, we explicitly demonstrate the realization of a causally non-ordered process by measuring a so-called '*causal witness*' [Araújo *et al.*, 2015].

In order to make our results stronger (*i.e.*, make the causal witness more robust to noise), we performed a superposition of the orders of a unitary gate and a measurement operation. In other words, we made a measurement *inside* a quantum process with an indefinite order of operations (the *quantum switch* [Chiribella *et al.*, 2013]). Performing a standard measurement inside the quantum switch would destroy its coherence, since it would reveal the time at which the measurement is performed, and would thus also reveal whether it is performed before or after other operations. In other words, such a measurement would reveal the causal order between the operations. In our scheme, however, the measurement outcomes are read out only "at the end" of the process, preserving its coherence. Since applications of indefinite causal orders will most likely require the superposition of orders of complex quantum operations, we believe that, in addition to the first direct demonstration of an indefinite causal order, our measurement in a quantum switch can also be considered a technological step towards such applications [Araújo *et al.*, 2014; Baumeler and Wolf, 2014; Chiribella, 2012; Feix *et al.*, 2015; Guérin *et al.*, 2016; Hardy, 2009].

In our usual understanding of causal relations, if we consider two events A and B which are connected by a time-like curve, we will have *one* of two cases: either 'A is in the past of B', or 'B is in the past of A'. However, the application of the superposition principle to such causal relations calls into question the interpretation of causal orders as a pre-existing property. In fact, the causal order can become genuinely indefinite. To see this, consider a two-qubit quantum state $|\phi\rangle$ lying in the composite Hilbert space $\mathcal{H}^C \otimes \mathcal{H}^T$ with $\mathcal{H}^C$ and $\mathcal{H}^T$ each being two-dimensional Hilbert spaces. It is possible to condition the order in which operations are applied to a target state $|\psi\rangle^T \in \mathcal{H}^T$ on the value of a control state $|\chi\rangle^C \in \mathcal{H}_C$. For example, if the state of the control qubit is $|0\rangle^C$, the two operators will be applied in the order A and then B on the state of the target qubit $|\psi\rangle^T$, and vice versa if the state

of the control qubit is $|1\rangle^C$. Therefore, if the control qubit is in a superposition state $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)^C$, a controlled quantum superposition of the situations 'A is in the past of B' and 'B is in the past of A' is established (Figure 1.1). In the above situation, the causal order is not merely in a superposition. It is, in fact, entangled with the state of the control qubit.

From this simple example we can see that the causal order between events is not always definite in quantum mechanics. One could, in the spirit of hidden-variable theories, insist that there might nonetheless be a well-defined causal order. However, this claim requires, in general, a theory to be non-local and contextual because of the Bell and Kochen-Specker theorems [Bell, 1964; Kochen and Specker, 1975]. In fact, one can even establish a no-go theorem that applies to causal orders [Zych *et al.*, 2019].

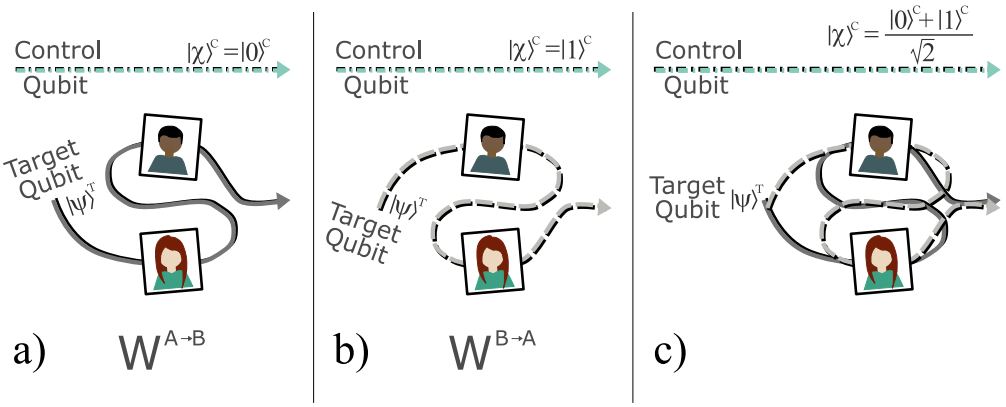


Figure 1.1 **The quantum switch.** Consider a situation wherein the order in which two parties Alice and Bob act on a target qubit $|\psi\rangle^T$ depends on the state of a control qubit in a basis $\{|0\rangle, |1\rangle\}^C$. If the control qubit is in the state $|0\rangle^C$ the target qubit is sent first to Alice and then to Bob (Panel **a**)), while if the control qubit is in the state $|1\rangle^C$, it is sent first to Bob and then to Alice (Panel **b**)). Both of these situations have a definite causal order, and are described by the process matrices $W^{A\to B}$ and $W^{B\to A}$ (Eq. 1.5). If the control qubit is prepared in a superposition state $\frac{1}{\sqrt{2}}\big(|0\rangle + |1\rangle\big)^C$, then the entire network is placed into a controlled superposition of being used in the order Alice $\to$ Bob and in the order Bob $\to$ Alice (Panel **c**)). This situation has an indefinite causal order.

The case described above, called the quantum switch, is the first explicit example wherein it was shown that quantum mechanics does not allow for a well-defined causal order [Chiribella *et al.*, 2013]. The switch was recently experimentally implemented [Procopio *et al.*, 2015] by superimposing the order in which two unitary operations acted. That experiment confirmed that a causally non-ordered quantum circuit can solve a specific computational problem more efficiently than an ordered quantum circuit. But only indirect evidence of indefinite causal order was observed through the demonstration of this computational advantage. The primary goal of our current experiment is, therefore, to provide direct experimental proof of the causal non-separability of the quantum switch. For this purpose, we used a recently developed theoretical tool: the *causal witness* [Araújo *et al.*, 2015].

1.2 Theoretical Framework

A causal witness is a carefully designed set of measurements, whose outcome will tell us if a given process is causally ordered or not. An intuitive way to introduce causal witnesses is through the well-known concept of an *entanglement witness* [Gühne and Tóth, 2009]. First, recall that a composite quantum system ρ lying in a Hilbert space $\mathcal{H}^A \otimes \mathcal{H}^B$ is *separable* or *entangled* depending on whether it can be written in the form $\rho = \sum_i p_i \rho_i^A \otimes \rho_i^B$ (with ρ_i^A and ρ_i^B states of the subsystems A and B and $0 \leq p_i \leq 1$, $\sum_i p_i = 1$) or not. Then it can be shown that for all entangled states ρ^{ent} there exists a Hermitian operator S, called an 'entanglement witness', such that $\mathrm{Tr}(S\rho^{ent}) < 0$, but $\mathrm{Tr}(S\rho^{sep}) \geq 0$ for all separable states ρ^{sep}. Hence, it follows that if one measures an entanglement witness on a state and finds a negative value the state must be entangled.

A similar quantity was recently introduced to determine if a *process matrix* W is causally separable or not[Oreshkov *et al.*, 2012]. A process matrix (the counterpart of the density matrix in the entanglement witness example) describes causal relations between local laboratories [Baumann and Brukner, 2016]. Consider two observers Alice and Bob who perform local operations M^A and M^B (M^A and M^B can be arbitrary quantum operations, from simple unitary operations to more complex measurement channels). By local operations we mean that the only connection that Alice and Bob have with the external world is given by the quantum state that they receive from it and the state that they return to it. The process matrix W then details how this quantum state moves between the two local laboratories (Figure 1.2). Hence, it is independent of the individual operations that Alice and Bob perform. In the case of the quantum switch, the process matrix first routes the input state to Alice and Bob in superposition, it then connects Alice's output to Bob's input and vice versa, and finally coherently recombines their outputs.

Since a causal witness characterizes a process rather than a state (unlike an entanglement witness), it requires a procedure akin to 'process tomography' (*i.e.*, 'causal tomography', see Section 1.5.5). Namely, we must probe the process with several different input states $\rho^{(\mathrm{in})}$. Then, for each input state, Alice and Bob implement several different *known* operations, and then we perform a final measurement $D^{(\mathrm{out})}$ (Figure 1.2). In general, M^A and M^B can include measurement operations, so each could have additional measurement outcomes associated with it. We denote the outcomes of Alice and Bob's local operations by a and b, and their choice of operation by x and y, respectively. We label the specific choice of an input state with z, and the output of a detection operation with d. With this in mind, the probability of obtaining the outcomes $M^A_{a,x}$, $M^B_{b,y}$, and $D^{(\mathrm{out})}_d$, with the input state $\rho^{(\mathrm{in})}_z$ can be written, using the *Choi-Jamiołkowski isomorphism* [Choi, 1975] (see Section 1.6.1), as

$$p(a,b,d|x,y,z) = \mathrm{Tr}\left[(\rho^{(\mathrm{in})}_z \otimes M^A_{a,x} \otimes M^B_{b,y} \otimes D^{(\mathrm{out})}_d) \cdot W\right]$$

with $\sum_{a,b,d} p(a,b,d|x,y,z) = 1$ for all the possible settings x, y, z and where W is the process matrix [Araújo *et al.*, 2015].

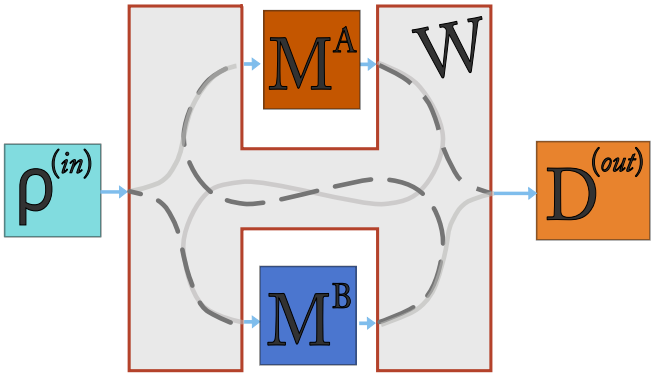


Figure 1.2 A process matrix representation of Figure 1.1. The process matrix W describes the "links" between Alice and Bob. For example, it could simply route the input state $\rho^{(\text{in})}$ to Alice M^A and then to Bob M^B (solid line), or vice versa (dashed line). In the case of the quantum switch, it creates a superposition of these two paths, conditioned on the state of a control qubit. The input state $\rho^{(\text{in})}$, the two local operations M^A and M^B, and the final measurement $D^{(\text{out})}$ must all be controllable and known a-priori. The unknown process is represented by the process matrix (shaded grey area labelled W). A causal witness quantifies the causal non-separability of W.

To calculate these probabilities for the quantum switch we must construct its process matrix, which we will call W_{switch}. To do this we will again use the Choi-Jamiołkowski isomorphism. As a first step, consider the identity channel from the output space $\mathcal{H}_{\text{out}}^{P_1}$ of a party P_1 to the input space $\mathcal{H}_{\text{in}}^{P_2}$ of a second party P_2. To describe this as a process matrix we can write it as a projector onto a process vector in the 'double-ket notation' [Braunstein *et al.*, 2000; Royer, 1991]:

$$|\mathbb{1}\rangle\rangle^{\mathcal{H}^X_{\text{in/out}}\mathcal{H}^Y_{\text{in/out}}} = \sum_j |j\rangle\rangle^{\mathcal{H}^X_{\text{in/out}}} \otimes |j\rangle\rangle^{\mathcal{H}^Y_{\text{in/out}}}, \tag{1.1}$$

where j labels a basis over the spaces. We can now use this process matrix to describe an input state passing first to Alice ($\mathcal{H}_{\text{in}}^A \to \mathcal{H}_{\text{out}}^A$), then to Bob ($\mathcal{H}_{\text{in}}^B \to \mathcal{H}_{\text{out}}^B$), and finally to the output space ($\in \mathcal{H}^{(\text{out})}$). This process is described by

$$|w^{A\to B}\rangle = |\mathbb{1}\rangle\rangle^{\mathcal{H}^{(\text{in})}\mathcal{H}^A_{\text{in}}} |\mathbb{1}\rangle\rangle^{\mathcal{H}^A_{\text{out}}\mathcal{H}^B_{\text{in}}} |\mathbb{1}\rangle\rangle^{\mathcal{H}^B_{\text{out}}\mathcal{H}^{(\text{out})}}. \tag{1.2}$$

Alice and Bob are free to perform measurements $M^A : \mathcal{H}_{\text{in}}^A \to \mathcal{H}_{\text{out}}^A$ and $M^B : \mathcal{H}_{\text{in}}^B \to \mathcal{H}_{\text{out}}^B$, respectively, but they are not part of the above process vector. Note that swapping the order of Alice and Bob is as simple as swapping the labels A and B. The vectors $|w^{A\to B}\rangle$ (describing 'Alice acts before Bob') and $|w^{B\to A}\rangle$ (describing 'Bob acts before Alice') both have a well-defined causal order (Figure 1.1, Panels **a)** and **b)**).

We are now in the position to construct the process matrix of the quantum switch. Recall that for the quantum switch the control qubit's state sets the relative amplitudes of Alice → Bob and Bob → Alice. Thus the process vector of the quantum switch (when the control qubit initially in the state

$\frac{|0\rangle^C+|1\rangle^C}{\sqrt{2}}$) is quite simply:

$$|w_{\text{switch}}\rangle = \frac{1}{\sqrt{2}}\Big(|w^{A\to B}\rangle\,|0\rangle^C + |w^{B\to A}\rangle\,|1\rangle^C\Big). \tag{1.3}$$

For the causal witness we will consider here, we will only measure the state of the control qubit after the switch. Thus, we need to construct the process matrix taking an input state, and returning the state of the control qubit. This is done by tracing over the switch output (*i.e.*, the target qubit) and fixing the state of the control qubit to be $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$. So the process matrix to compute the final state of the control qubit is represented by the process matrix:

$$W_{\text{switch}} = \text{Tr}_{\mathcal{H}^{(\text{out})}}\big(|w_{\text{switch}}\rangle\,\langle w_{\text{switch}}|\big), \tag{1.4}$$

where $\text{Tr}_{\mathcal{H}^{(\text{out})}}(\cdot)$ is the partial trace over the output system qubit.

Using the same formalism, one can also concisely describe all causally separable processes. Consider two general process matrices linking the two local laboratories A and B, $W^{A\to B}$ and $W^{B\to A}$. Here, contrary to Eq. 1.2, the link between the laboratories is in general no longer the identity channel. Then by simply taking an incoherent mixture of two, one can describe all possible causally separable processes [Araújo *et al.*, 2015]:

$$W^{\text{sep}} := pW^{A\to B} + (1-p)W^{B\to A}, \tag{1.5}$$

where $0 \leq p \leq 1$. Physically, this can be understood as each run of the process having a well-defined order, with Alice acting first with probability p and Bob acting first with probability $1-p$. From this definition it is apparent that every convex combination of causally-separable process matrices is still a causally-separable process matrix; thus, the set of causally-separable process matrices is convex.

Causal witnesses are designed to distinguish between causally separable (Eq. 1.5) and causally non-separable process matrices (such as Eq. 1.4). Indeed, for all causally non-separable process matrices $W^{\text{n-sep}}$, there exists a Hermitian operator S, called a causal witness, such that

$$\text{Tr}(SW^{\text{n-sep}}) < 0, \tag{1.6}$$

but $\text{Tr}(SW^{\text{sep}}) \geq 0$ for all causally separable process matrices W^{sep} [Araújo *et al.*, 2015], just as in the entanglement witness example. Such an operator is guaranteed to exist, because the convexity of the causally-separable process-matrices set ensures that there is always a hyperplane, which separates the set from a given causally non separable process $W^{\text{n-sep}}$ [Rockafellar, 1970] (see Figure 1.8 in the ***Materials and Methods***, Sec. V).

A causal witness can be written as a linear combination of the states at the input of the process, the quantum operations of Alice and Bob, and the measurements on the output. This is because the tensor

products of these elements span the process matrix Hilbert space (the space W_{switch} lives in). Using the notation defined in Eq. 1.1, a causal witness can be expanded as

$$S = \sum_{\substack{a,b,d \\ x,y,z}} \alpha_{a,b,d,x,y,z} \cdot \rho_z^{(\text{in})} \otimes M_{a,x}^A \otimes M_{b,y}^B \otimes D_d^{(\text{out})}, \tag{1.7}$$

where the coefficients $\alpha_{a,b,d,x,y,z}$ are real numbers which define (together with the input states and measurements) a particular witness. From the definition in Eq. 1.1, it follows that

$$\text{Tr}(SW) = \sum_{\substack{a,b,d \\ x,y,z}} \alpha_{a,b,d,x,y,z} p(a,b,d|x,y,z) \tag{1.8}$$

and therefore the evaluation of the quantity $\text{Tr}(SW)$ for a given process W translates into a determination of probabilities $p(a,b,d|x,y,z)$ for several input states and measurement choices.

In the case where there are no restrictions on which operations we are able to implement, we choose the coefficients $\alpha_{a,b,d,x,y,z}$ by maximizing the quantity $-\text{Tr}(SW)$ over the set of all possible causal witnesses, as described in Section 1.5.5. This quantity, for such a optimal witness, corresponds to the maximum 'amount of worst-case noise' that the process under examination can tolerate while remaining causally non-separable [Araújo *et al.*, 2015]. More precisely, it is the minimal $\lambda \geq 0$ for which the process matrix

$$W_\lambda = \frac{1}{1+\lambda}\left(W^{\text{n-sep}} + \lambda\Omega\right) \tag{1.9}$$

becomes causally separable, where Ω is any other process that could have been prepared instead of the desired $W^{\text{n-sep}}$. We will refer to this quantity as the 'causal non-separability' (CNS) of a process W:

$$\text{CNS}(W) := -\text{Tr}(SW). \tag{1.10}$$

When the $-\text{Tr}(SW) < 0$, we define the $\text{CNS}(W)$ to be zero.

In practice, however, we may not be able to maximize $-\text{Tr}(SW)$ over the whole set of causal witnesses, as there can be restrictions on which operations Alice and Bob have access to. To fully assess the CNS, Alice and Bob must be able to implement a complete basis of operators, which gives them access to the maximal amount of information about the process. We define, therefore, the experimentally certifiable CNS (hereafter referred to as $\text{CNS}_{\text{exp}}(W) = -\text{Tr}(S_{\text{exp}}W)$) as the maximum of $-\text{Tr}(SW)$ over the restricted set of operators. In this case, $\text{CNS}_{\text{exp}}(W)$ is no longer the amount of noise that the process can tolerate before becoming non-separable, but the maximal amount of noise for which this restricted class of witnesses can still detect its causal nonseparability.

If Alice and Bob were only able to do unitaries, for example, this would drastically diminish the attainable $\text{CNS}_{\text{exp}}(W)$—this path was chosen in Ref [Procopio *et al.*, 2015]. Indeed, since a unitary operation cannot extract any explicit information from the manipulated state (and, consequently, from

the process), neither Alice nor Bob can gain any knowledge about their received state when applying only such gates and consequently the estimated $\mathrm{CNS}_{\mathrm{exp}}(W)$ is less efficient. However, if the unitary operations are replaced with projective measurements then, roughly speaking, information about the process at different points throughout the switch can be extracted. In fact, if Alice and Bob both have access to measure and reprepare operations, one can achieve $\mathrm{CNS}_{\mathrm{exp}}(W) = \mathrm{CNS}(W)$.

Because of the experimental challenges of coherently adding measure-and-reprepare operations, in our experiment Alice performs an measure-and-reprepare operation and Bob implements a unitary channel. Remarkably, it turns out that giving one party a measure-and-reprepare operation and the other a unitary operation still increases $\mathrm{CNS}_{\mathrm{exp}}(W)$ substantially. So the causal witness we will measure depends both on Alice's outcome (performed inside the switch) and our final measurement outcome.

1.3 Experiment and Results

To experimentally implement the quantum switch, we need a control and a target qubit. In our experiment we encode a control qubit in a path degree of freedom of a photon and a target qubit in the same photon's polarization. The technique of utilizing multiple degrees of freedom has enabled many previous quantum technologies [Barreiro *et al.*, 2008; Englert *et al.*, 2001; Lanyon *et al.*, 2009]. For our present experiment this is convenient as Bob's unitary gate can be implemented easily with three waveplates, while Alice can perform a projective measurement with waveplates and a polarizing-beamsplitter. Note that there are other proposals to coherently control the causal orders of events [Araújo *et al.*, 2015; Friis *et al.*, 2014; Rambo *et al.*, 2016] (in particular, the first experimental implementation for the quantum switch was proposed by G. Chiribella, R. Ionicioiu, T. Jennewein, and D. R. Terno in a private communication). In these proposals (as in ours) the target and control system are encoded in the same particle. In principle, it is also possible to use different particles. With photons this could be done using a so-called controlled path gate [Zhou *et al.*, 2011], or potentially by using a spin qubit to control the causal order acting on a photon [Wang *et al.*, 2016].

In our experiment, the realization of the unitary channel is straightforward, but a short remark is necessary concerning Alice's measurement. It is clear that a polarizing-beamsplitter enables one to distinguish the polarization of an incoming photon. However, a polarizing-beamsplitter gives rise to additional spatial modes (*i.e.*, after the polarizing-beamsplitter there are two output paths). These two spatial modes can be considered as a new spatial qubit. Then, the action of the polarizing-beamsplitter is to couple the polarization qubit to this additional qubit. This is formally equivalent to a *von Neumann system-probe coupling*, which can model the interaction necessery for any projective measurement [Von Neumann, 1955] and has been used between path and polarization degrees of freedom in the experiment reported in Ref. [Rozema *et al.*, 2012]. In our experiment, the polarization

qubit is the system, and it is coupled (via the polarizing-beamsplitter) to an additional spatial qubit which is the probe. We can read out information about the system by measuring the probe (with a photon detector) at a later time. This solves the non-trivial problem of realizing a measurement operation *inside* a quantum switch. Most approaches to acquire information inside the switch would lead to distinguishing information about the order in which the operations were applied, destroying the quantum superposition. In our solution, however, since the probe qubit is not measured until the information about the order of application of the operations is erased, the entire process can remain coherent. This solution also works deterministically, *i.e.*, both of Alice's outcomes are retained. It also allows Alice to implement a measurement-dependent repreparation by placing different waveplates in each of the two outcome modes.

Our implementation of the quantum switch draws inspiration from a previous experiment [Procopio *et al.*, 2015], in which only orders of unitary operations were superimposed. Therefore, as in [Procopio *et al.*, 2015], our experimental skeleton consists of a Mach-Zehnder interferometer (MZI) with a loop in each arm. However, because Alice's measure-and-reprepare channel adds an additional path degree of freedom, we need an extra interferometric loop.

A scheme of our experimental apparatus is presented in Figure 1.3. The first step is to set the state of the system qubit (encoded in the polarization) with a polarizer and a half-waveplate. Then the photon impinges on a $50/50$ beamsplitter; this sets the state of the control qubit (encoded in a path degree of freedom) in $|+\rangle$. Depending on the state of the path qubit, the photon is sent to either Alice (who performs M^A) and then Bob (who performs U^B), or vice versa. As described above, M^A is a projective measurement (a sequence of two waveplates and a polarizing-beamsplitter) and a corresponding repreparation (a sequence of two waveplates in only one of the polarizing-beamsplitter outputs), and U^B is a unitary gate (a sequence of three waveplates). Since the polarizing-beamsplitter adds a second path qubit this results in four path modes, encoding both the state of the control qubit and the outcome of the measure-and-reprepare channel. Referring to Figure 1.3, the external (yellow) interferometer arises from the outcome H - also referred to as a logical 0 - and the internal (blue) one from the outcome V - a logical 1. We finalize the switch by erasing the information about the order of the gates. This can be done by applying a Hadamard gate to the control qubit. Since the control qubit is a path qubit a Hadamard gate can be implemented with a $50/50$ beamsplitter. However, in our experiment there are two path-qubits (the control qubit and Alice's ancilla measurement qubit). Thus, we must use two $50/50$ beamsplitters: one beamsplitter to interfere the control qubit when Alice's ancilla qubit is in the state $|0\rangle$, and one beamsplitter when it is in the state $|1\rangle$. Finally, each of the four outputs are coupled into single-mode fibers, which are each connected to single-photon detectors (SPD). Then detecting a photon in one of the four modes yields the result of both the measurement of the control qubit in the superposition basis and Alice's measurement (see the detector labels in Figure 1.3).

We wish to evaluate the CNS of our quantum switch by experimentally estimating the expectation

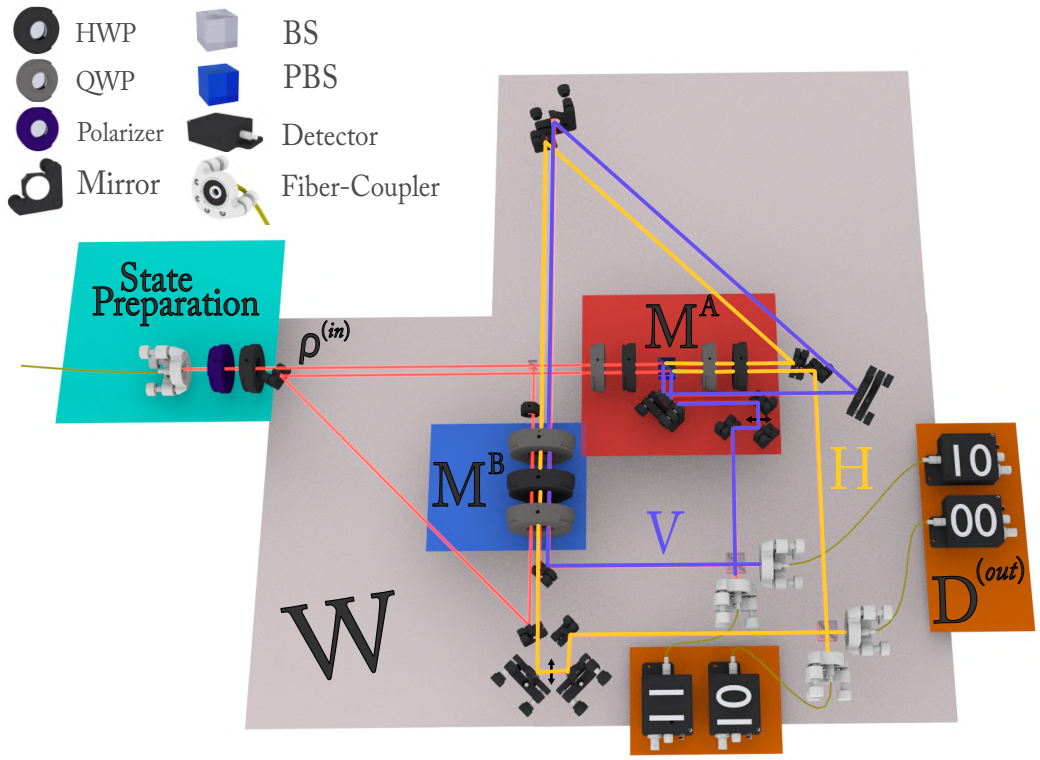


Figure 1.3 **Experimental Set-Up.** A sketch of our experiment to verify the causal non-separability of the quantum switch. We produce pairs of single photons using a Type-II spontaneous-parametric-down-conversion source (not shown here). One of the photons is used as trigger, one is sent to the experiment. The experiment body consists of two Mach-Zehnder interferometers, with loops in their arms. The qubit control, encoded in a path degree of freedom, dictates the order in which the operations M^A and M^B are applied to the target qubit (encoded in the same photon's polarization) Alice implements a measurement and repreparation (M^A), and Bob implements a unitary operation (M^B). After the interferometers the control qubit is measured, *i.e.*, we check if the photon exits port 0 or port 1. Note that there are two interferometers, each corresponds to a different outcome for Alice: the yellow path means Alice measured the photon to be horizontally polarized (a logical 0), and the blue path means she found the photon to be vertically polarized (a logical 1). The first digit written on the detector labels this outcome. The second digit refers to the final measurement outcome, which, physically, corresponds to the photon exiting either from port 0 or port 1. In this diagram port 0 (1) means the photon exits through a horizontally (vertically) drawn port. A half waveplate at 0° was used in the reflected arm of the first beamsplitter in order to compensate the acquired additional phase. Acronyms in the figure are defined as follows: QWP, quarter waveplate; HWP, half waveplate; BS, beamsplitter; PBS, polarizing-beamsplitter.

value of a causal witness S (Eq. 1.7). In other words, we want to assess $\mathrm{Tr}(S_{\mathrm{exp}} W_{\mathrm{switch}})$, where here W_{switch} refers to the process matrix describing our experiment. Since the trace is linear, this can be done by implementing one term in the sum of S (Eq. 1.7) at a time. To estimate a single term, an input state is injected into the switch, Alice and Bob each perform an operation inside, and then the outputs of the overall process are measured. Since the control qubit measurement and Alice's measurement are both single-qubit projective measurements, there are a total of four possible outcomes. For each measurement setting, different input states are sent into the switch and the probabilities of each outcome are experimentally estimated by sending multiple copies of the same input state. To compute the final value of the $\mathrm{CNS}_{\mathrm{exp}}(W_{\mathrm{switch}})$, the results of these measurements are weighted by the corresponding $\alpha_{a,b,d,x,y,z}$ and summed.

The number of terms in the sum of Eq. 1.7 is determined by the specific witness we wish to evaluate. In general, Alice and Bob must each implement a set of operators forming a basis over their channels. For Bob's unitary channel this requires 10 elements and for Alice's measure-and-reprepare channel 16 [Araújo *et al.*, 2015]. In our case, we formed Alice's basis with *four* (non-commutative) projection operators, and *three* unitary repreparation operators when the outcome was H and *one* operator (the identity operator) when the outcome was V. This corresponds to 12 measure-and-reprepare channels when the outcome of Alice's measurement is H and 4 when it is V, for a total of 16 measure-and-reprepare operators. For Bob we implement all 10 unitaries.

Varying the input state can make $\mathrm{CNS}_{\mathrm{exp}}(W_{\mathrm{switch}})$ more robust to noise. Hence, for our experiment we used *three* different input states: $|H\rangle$, $|V\rangle$ and $|+\rangle$. Finally, we implemented *two* different measurement operators $D^{(\mathrm{out})}$ on the control qubit (corresponding to the two outcomes of the projection onto basis $\left\{|\pm\rangle = \frac{|0\rangle \pm |1\rangle}{\sqrt{2}}\right\}$). Thus, for our experiment, the calculation of $\mathrm{CNS}_{\mathrm{exp}}(W_{\mathrm{switch}})$ translates into

$$\mathrm{CNS}_{\mathrm{exp}}(W_{\mathrm{switch}}) = -\sum_{z=0}^{2}\sum_{a=0}^{1}\sum_{x=0}^{11}\sum_{y=0}^{9}\sum_{d=0}^{1} \alpha_{a,d,x,y,z}\; p(a,d|x,y,z), \tag{1.11}$$

here we do not need the sum over b, since Bob's unitaries do not have an outcome. The probability in Eq. 1.11 is defined as

$$p(a,d|x,y,z) := \mathrm{Tr}\left[(\rho_z^{(\mathrm{in})} \otimes M_{a,x}^{A} \otimes U_y^{B} \otimes D_d^{(\mathrm{out})}) \cdot W_{\mathrm{switch}}\right]. \tag{1.12}$$

We must experimentally estimate all of these probabilities to evaluate $\mathrm{CNS}_{\mathrm{exp}}(W_{\mathrm{switch}})$. There are 1440 terms in this sum. However, four outcomes (two from of Alice's measurement and two from the final detection) are collected simultaneously (experimentally, this means the counts of four SPDs are collected in one setting). Therefore, we need 360 different experimental settings. However, for our witness, of the 360 pre-factors $\alpha_{a,d,x,y,z}$, 101 are equal to zero, so there are actually only 259 relevant experimental settings.

With this in place, we are able to experimentally measure the $\mathrm{CNS}_{\mathrm{exp}}(W_{\mathrm{switch}})$ (for information relating experimental visibility, stability and data taking procedure, see Sections 1.5.2-1.5.4). Figure 1.4 shows some of the probabilities $p(a,d|x,y,z)$ (Eq. 1.11) for the four outcomes, *i.e.*, for Alice $a = 0, 1$ and our final measurement $d = 0, 1$ (the remainder are shown in the Supplemental Information). In Figure 1.4, the experimentally obtained values are the blue dots and the theoretical predictions are the bars.

Our main source of error is phase fluctuations in the two interferometers. We therefore performed a separate measurement (presented in the Section 1.5.4) to characterize this error. The error bars in Figure 1.4 represent both these phase errors and Poissonian errors due to finite counts. These errors do not take into account systematic errors, such as waveplate miscalibration, since such systematic errors represent a deviation of our experimental switch from the ideal switch.

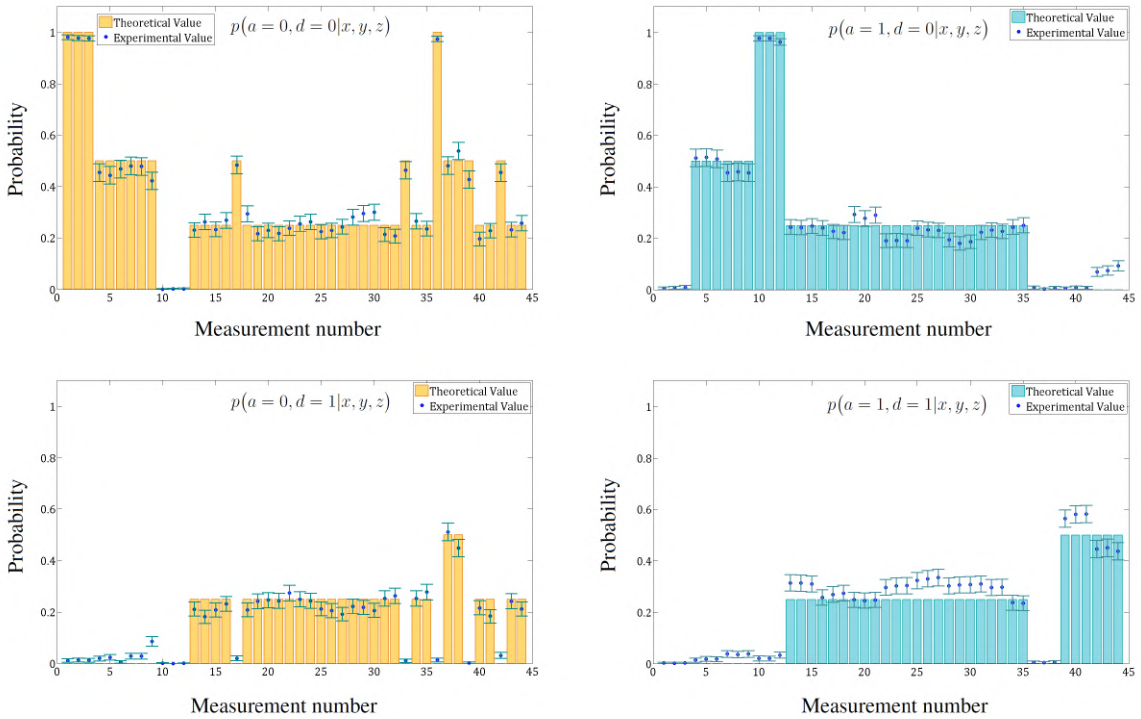


Figure 1.4 **Experimentally estimated probabilities.** Each data point represents a probability $p\big(a,d|x,y,z\big)$ in Eq. 1.11 for $a=0,1$ and $d=0,1$. The blue dots represent the experimental result and the bars the theoretical prediction. The yellow (blue) bars refer to the external (internal) interferometer. The x-axis is the measurement number, which labels a specific choice of is input state, measurement channel for Alice and Bob, and final measurement outcome. For our witness, it runs from 0 to 259, but we only show the first 44 here for brevity Alice and Bob specific choice of operator is given in Table 1 of the Supplementary Materials.

We can now obtain a value for the CNS of our process by weighting the data presented in Figure 1.4 (and Supplementary Figures 1-3) by $\alpha_{a,d,x,y,z}$ and then summing them. The result is

$$\mathrm{CNS}_{\mathrm{exp}}(W_{\mathrm{switch}}) = 0.202 \pm 0.029. \tag{1.13}$$

The error bar on $\mathrm{CNS}_{\mathrm{exp}}(W_{\mathrm{switch}})$ was calculated by Gaussian error propagation from the errors of the individual probabilities. The theoretical maximum value for $\mathrm{CNS}_{\mathrm{exp}}(W_{\mathrm{switch}})$ is 0.2842. The disagreement between this and our measured result is caused primarily by two effects. First, given the reduced visibility of the interferometers (which we will discuss in detail shortly) the maximal value for $\mathrm{CNS}_{\mathrm{exp}}(W_{\mathrm{switch}})$ is 0.2523, when the visibility is 0.9539. The remaining discrepancy comes from systematic errors, such as waveplate miscalibration, which effectively mean that the unitaries Alice and Bob implement differ slightly from their targets. For example, we estimate, using a simple Monte Carlo simulation, that a waveplate calibration error of 3° would explain this discrepancy, leading to a drop in the CNS of approximately 0.043. Still, given our measured result, we can conclude that our process is causally non-separable by a margin of approximately seven standard deviations. This large

margin demonstrates the effectiveness of performing a measurement operation inside the quantum switch.

As mentioned above, the causal non-separability (as measured using a causal witness) can be considered as a measurement of how much noise can be added to the process before it becomes causally separable. The CNS_{exp} we have discussed so far refers to a worst-case noise model [Araújo *et al.*, 2015], wherein the desired process is replaced with the process that can do most damage to its causal nonseparability with a probability

$$p_{\text{worst-case}} := \frac{\text{CNS}_{\text{exp}}(W_{\text{switch}})}{1 + \text{CNS}_{\text{exp}}(W_{\text{switch}})}. \tag{1.14}$$

Since the replacement is done with the worst-case process, this is a lower bound on the 'probability of noise' that can be tolerated (see Section 1.5.5). For our process $p_{\text{worst-case}} = 0.168 \pm 0.001$.

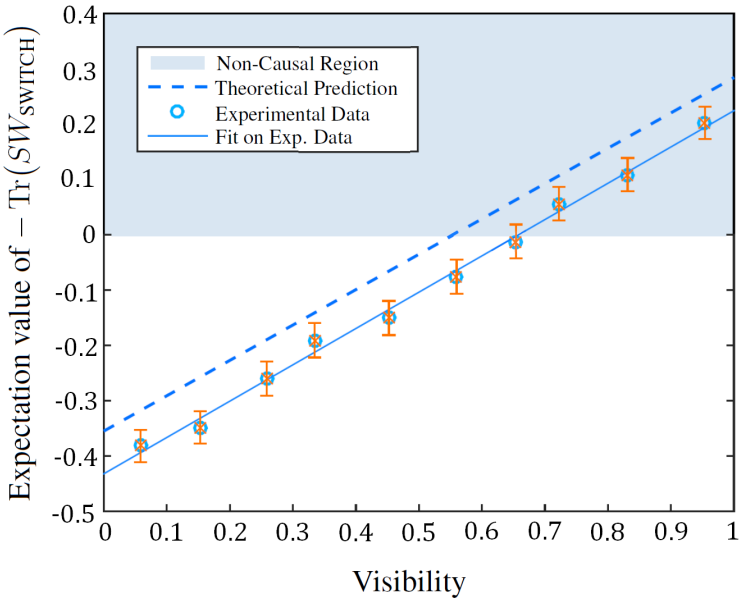


Figure 1.5 Expectation value of the causal witness ($-\text{Tr}(SW_{\text{switch}})$) in the presence of noise. As the control qubit (intitially in $|+\rangle$) is decohered, the superposition of causal orders becomes an incoherent mixture of causal orders. Hence, the causal non-separability of the switch is gradually lost. The plot shows the causal non-separability of our experimentally implemented switch as the visibility of the two interferometers is decreased (from right to left). The experimental data linearly decreases with visibility just as theory (dashed line) predicts. The gap between theory and experiment is because of systematic errors. The visibility (x axis) is a measure of the dephasing strength on the control qubit.

We studied the effect of the noise most relevant to our experiment. Namely, dephasing the control qubit but not the system qubit. This noise is the strongest in our setup because the control qubit is encoded in a path degree of freedom, which must remain interferometrically stable (see Ref. [Branciard, 2016] for the formal definition of this noise model). We realized this noise by imbalancing the path length of the interferometers by more than the photons' coherence length. The experimental signature of this imbalance is a reduced visibility of the interferometer. We measured the CNS for

several visibilities between 0.95 and 0.06. Figure 1.5 shows a decrease in the expectation value of $-\operatorname{Tr}(SW_{\text{switch}})$ as the noise increases. There is an offset between the experimental data and the the theoretical prediction due to systematic errors. However, both theory and experiment follow the same trend. By extrapolating our fit of the experimental data to $-\operatorname{Tr}(SW_{\text{switch}}) = 0$ (where the process becomes causally separable), we observe a 'noise tolerance' of 0.342 for this type of noise. As expected this is larger than our experimentally measured $p_{\text{worst-case}}$, indicating that it is indeed a lower bound.

1.4 Discussion

Our experiment demonstrates how to perform a measurement *inside* a quantum switch without destroying the superposition of causal orders. The task was only assumed to be possible in Ref. [Araújo *et al.*, 2015], but no method to accomplish it was proposed. The difficulty is that performing a standard measurement necessarily reveals the time at which it is performed, and thus whether it is performed before or after the partner's operation. Consequently, the superposition of causal orders becomes incoherent. Our way around this is to break the measurement into two steps: first the system coherently interacts with an ancilla through a unitary operation (namely, the additional path modes introduced by the local operation in our experiment). Then, after finalizing the quantum switch (interfering these modes), the ancilla is measured. This allows us to make a "coherent measurement at different times" and then erase the ordering information.

We demonstrated the causal non-separability of our experimental apparatus by measuring a causal witness. With the ability to perform a measurement inside the switch we were able to increase the robustness of the causal witness to noise. Previous experimental work only indirectly accessed the causal non-separability of the switch and, moreover, only used unitary gates in the switch[Procopio *et al.*, 2015]. Although some other experiments [Fitzsimons *et al.*, 2015; MacLean *et al.*, 2017; Ried *et al.*, 2015; Ringbauer *et al.*, 2016] have also studied the topic of causal relations in quantum mechanics, they focused on a different aspect. For example, in [Fitzsimons *et al.*, 2015; MacLean *et al.*, 2017; Ried *et al.*, 2015] they distinguish between different causal structures, rather than creating a genuinely indefinite causal order, as in our work. In fact, the incoherent mixture [Ried *et al.*, 2015] and a quantum superposition [MacLean *et al.*, 2017] of different causal relations reported previously are both compatible with one party in the past and the other in the future. Thus, in our language they correspond to causally-ordered processes.

Our work represents the first experimental realization of a quantum superposition of orders of non-unitary channels, and the first measurement of a causal witness. We believe this will be an important step towards the realization of quantum superpositions of the order of more elaborate processes. Since it has been theoretically demonstrated that causally non-ordered processes can give rise to a reduction in the query complexity of certain tasks [Araújo *et al.*, 2014; Chiribella, 2012; Hardy, 2009], and lead

to more efficient communication channels [Feix *et al.*, 2015; Guérin *et al.*, 2016], it is important to study new techniques to create more complex causally non-ordered processes. In fact, we already see an advantage in our current work. Making a measurement inside the quantum switch made our experiment more robust to noise and allowed us to demonstrate, by approximately seven standard deviations, that our set-up cannot be described by a causally ordered process.

1.5 Methods

1.5.1 Single Photon Source

We generate heralded single-photons using a Type-II spontaneous parametric down-conversion (SPDC) process in a Sagnac loop [Kim *et al.*, 2006]. The Sagnac loop is realized using a dual-wavelength polarizing-beamsplitter and two mirrors. The SPDC crystal is a 20-mm-long *periodically-poled crystal Potassium Titanyl Phosphate* (PPKTP) crystal. The crystal is pumped by a 23.7 mW diode laser centred at 395 nm. The polarization of the laser is set to be horizontal. With this, we generate degenerate pairs of single photons centred at 790 nm, in a separable polarization state $|H\rangle |V\rangle$. Polarizers in the signal and idler modes are used to ensure that the polarization is in a well-defined state. The down-converted photons are coupled into single-mode fibers. One photon is sent directly to an single-photon detector, and is used to herald the other photon's presence for the experiment, while the other is sent to our experiment. After passing though the experiment we observe a coincidence rate between the herald detector and the four final-measurement detectors of 3750 pairs per second.

1.5.2 Implementing Alice and Bob's Channels

As discussed in the main text, in order to experimentally measure a causal witness Alice and Bob need to implement a series of quantum channels on a polarization qubit inside the quantum switch. Alice must perform a measure-and-reprepare channel, while Bob implements a unitary channel. Alice measures in four different bases. We define her different bases by a unitary operator preceding a projective measurement in the basis $\{|0\rangle, |1\rangle\}$. Alice's pre-measurement operators are listed in the first column of Table 1.1. When her outcome is $|0\rangle$ (in a given basis), Alice implements one of three different repreparation operators (second column of Table 1.1). On the other hand, when her outcome is $|1\rangle$ she performs the identity channel. Thus she has 16 different measure-and-reprepare maps. Bob simply implements 10 different unitary operators (third column of Table 1.1).

Table 1.1 List of operators performed by the two parties. The table shows Alice's four measurement operators, and her three repreparation operators which she applies when her outcome is $|0\rangle$; when her outcome is $|1\rangle$ she performs the identity. Bob's ten unitary operators are shown in column three.

Alice - Meas.	**Alice - Reprep.**	**Bob - Unit. Op.**	
(1) $\begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}$	(1) $\begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}$	(1) $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	(6) $\frac{1}{\sqrt{2}}\begin{pmatrix} -i & -1 \\ -i & 1 \end{pmatrix}$
(2) $\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix}$	(2) $\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix}$	(2) $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$	(7) $\frac{1}{\sqrt{2}}\begin{pmatrix} -i & 1 \\ i & 1 \end{pmatrix}$
(3) $\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -i \\ i & -1 \end{pmatrix}$	(3) $\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -i \\ i & -1 \end{pmatrix}$	(3) $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$	(8) $\frac{1}{\sqrt{2}}\begin{pmatrix} i & i \\ -1 & 1 \end{pmatrix}$
(4) $\begin{pmatrix} 0 & -i \\ 1 & 0 \end{pmatrix}$		(4) $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$	(9) $\frac{1}{\sqrt{2}}\begin{pmatrix} -i & -i \\ -1 & 1 \end{pmatrix}$
		(5) $\frac{1}{\sqrt{2}}\begin{pmatrix} i & -1 \\ i & 1 \end{pmatrix}$	(10) $\frac{1}{\sqrt{2}}\begin{pmatrix} -i & i \\ 1 & 1 \end{pmatrix}$

We experimentally implement both Alice's measurement operators and repreparation operators through a sequence of two waveplates (quarter-, then half-waveplate), and Alice's projective measurement in a polarizing-beamsplitter measuring in $\{|H\rangle, |V\rangle\}$. Bob's operators are implemented via three waveplates (quarter-, then half-, then quarter-waveplate). In Table 1.2 we show the specific waveplates angles we use for each operator.

1.5.3 Experimentally Estimating Probabilities

Since Alice makes a two-outcome measurement and our final measurement has two outcomes, for each setting of Alice and Bob there are four different outcomes. Experimentally, each outcome corresponds to a different single-photon detector. For each setting we collect approximately 7500 counts in total after 2 seconds of data acquisition. From these counts we estimate the four corresponding output probabilities through the formula

$$p_{mn} = \frac{C_{mn}}{C_{\text{tot}} \cdot \eta_m \cdot \eta_n^m}, \tag{1.15}$$

where C_{mn} is the number of counts collected at one of the detectors and the η factors are different relative detector efficiencies, described below. Here, m labels Alice's outcome (experimentally, this labels in the internal (blue) or external (yellow) interferometer), and n the outcome of the final measurement (experimentally, port 0 or port 1 of either interferometer). The total number of (efficiency-corrected) counts, appearing in Eq. 1.15, is

$$C_{\text{tot}} = \sum_{m=0}^{1} \sum_{n=0}^{1} \frac{C_{mn}}{\eta_m \cdot \eta_n^m}. \tag{1.16}$$

Table 1.2 Set of waveplates angles. A list of all of the waveplate angles used to perform the operators listed in Table 1.1. In our experiment, all combinations of these settings were used, which, together with our three input states, results in 360 measurement settings. Acronyms in the table are defined as follows: QWP, quarter waveplate; HWP, half waveplate.

Alice - Meas.	**Alice - Reprep.**	**Bob - Unit. Op.**
		(1) $0^\circ_{\text{QWP}}, 0^\circ_{\text{HWP}}, 0^\circ_{\text{QWP}}$
(1) $0^\circ_{\text{HWP}}, 0^\circ_{\text{QWP}}$	(1) $0^\circ_{\text{HWP}}, 0^\circ_{\text{QWP}}$	(2) $0^\circ_{\text{QWP}}, 45^\circ_{\text{HWP}}, 0^\circ_{\text{QWP}}$
		(3) $90^\circ_{\text{QWP}}, 45^\circ_{\text{HWP}}, 0^\circ_{\text{QWP}}$
(2) $22.5^\circ_{\text{HWP}}, 45^\circ_{\text{QWP}}$	(2) $22.5^\circ_{\text{HWP}}, 0^\circ_{\text{QWP}}$	(4) $90^\circ_{\text{QWP}}, 0^\circ_{\text{HWP}}, 0^\circ_{\text{QWP}}$
		(5) $90^\circ_{\text{QWP}}, 0^\circ_{\text{HWP}}, 45^\circ_{\text{QWP}}$
(3) $0^\circ_{\text{HWP}}, -45^\circ_{\text{QWP}}$	(3) $0^\circ_{\text{HWP}}, -45^\circ_{\text{QWP}}$	(6) $90^\circ_{\text{QWP}}, 45^\circ_{\text{HWP}}, 45^\circ_{\text{QWP}}$
		(7) $0^\circ_{\text{QWP}}, 0^\circ_{\text{HWP}}, 45^\circ_{\text{QWP}}$
(4) $45^\circ_{\text{HWP}}, 0^\circ_{\text{QWP}}$		(8) $45^\circ_{\text{QWP}}, 0^\circ_{\text{HWP}}, 90^\circ_{\text{QWP}}$
		(9) $45^\circ_{\text{QWP}}, 45^\circ_{\text{HWP}}, 90^\circ_{\text{QWP}}$
		(10) $45^\circ_{\text{QWP}}, 0^\circ_{\text{HWP}}, 0^\circ_{\text{QWP}}$

The efficiency factors in the above equations are defined as follows. The single-subscript factor η_m refers to relative efficiencies between the internal ($m = 1$) and external ($m = 0$) interferometer (Figure 1.3). The other factors η_n^m refer to the relative efficiencies between the two ports $n = 0$ and $n = 1$, of interferometer m. Then the absolute efficiency of a given detector is $\eta_m \cdot \eta_n^m$. Roughly speaking, to estimate the relative efficiencies we must send the same number of photons between the detectors and compare the measured count rates.

To estimate η_n^m within each interferometer we send the photons between the two ports by scanning the phase (when all of the internal waveplates are set to 0) by means of a piezo-electrically driven translation stage. Plots of representative interference fringes (already efficiency corrected) for each interferometer are shown in Figure 1.6. By requiring the total counts out of each port to be constant, we can obtain a relative efficiency between the two ports in each interferometer. In practice, we obtain the efficiency by plotting the counts out of one port versus the counts out of the other port. If the two efficiencies are equal, the slope of this line will be 1. However, due to different coupling and detector efficiencies must enforce this by we requiring

$$K_0 = \frac{C_{00}}{\eta_0^0} + \frac{C_{01}}{\eta_1^0} \qquad \text{and} \qquad K_1 = \frac{C_{10}}{\eta_0^1} + \frac{C_{11}}{\eta_1^1}, \tag{1.17}$$

where K_0 and K_1 are constants. We will set one efficiency of each pair to 1, since we are interested in the relative efficiency. Setting (arbitrarily) $\eta_0^m = 1$ means that the slope of C_{m1} versus C_{m0} will be η_1^m. These plots, for both interferometers, are shown in Figure 1.7.

If we next estimate η_m, the relative efficiency between two interferometers, we are able to estimate the required probabilities (Eq. 1.15). To do this we use the state preparation waveplate (Figure 1.3) to send the photons all to one interferometer or the other. In each case we scan the phase. Then, using the previously-discussed efficiencies η_n^m, we have K_0 and K_1 (Eq. 1.17). As before, we can set one of the relative efficiencies to 1, we choose $\eta_0 = 1$. Then we can calculate the final efficiency as

$$\eta_1 = \frac{\text{mean value}(K_1)}{\text{mean value}(K_0)}. \tag{1.18}$$

This works because using the waveplates and the polarizing-beamsplitter we can send nearly all of the incident photons one way or the other.

Using this procedure we now have relative efficiencies between all of the detectors. Note that $\eta_0 \cdot \eta_0^0 = 1$; however, this does not matter as even if we had the absolute efficiency of each detector it would cancel out in the calculation of the probability (Eq. 1.15), because we must normalise by C_{tot}. After evaluating p_{00}, p_{01}, p_{10}, and p_{11} for each of Alice and Bob's settings we weight each by the corresponding $\alpha_{a,d,x,y,z}$ (Eq. 1.11), and sum them all up. This gives us our experimental value of the causal non-separability.

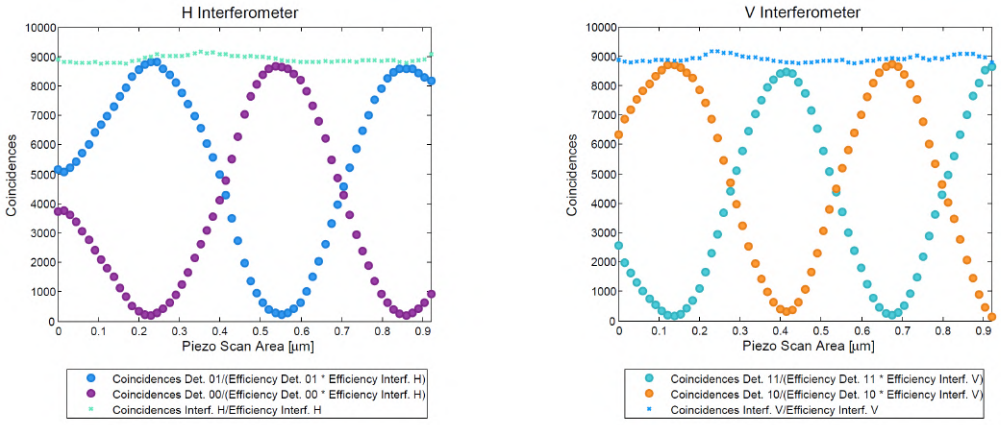


Figure 1.6 Efficiency-corrected interferometer fringes out of the two interferometers. A plot of the coincidences between the herald and the two detectors at the output of each interferometers as the interferometer phase is varied.

1.5.4 Stability and Visibility of the Interferometers

Central to our experiment are two interferometers whose overall size is approximately 80 cm $\times$ 120 cm. The visibility of the two interferometers is 95%, this is apparent in the interferograms shown in Figure 1.7. This error can be interpreted as dephasing noise on the control qubit (see the discussion in the main text).

In addition to the reduced visibility, the phase of the interferometer fluctuates. If the phase fluctuates

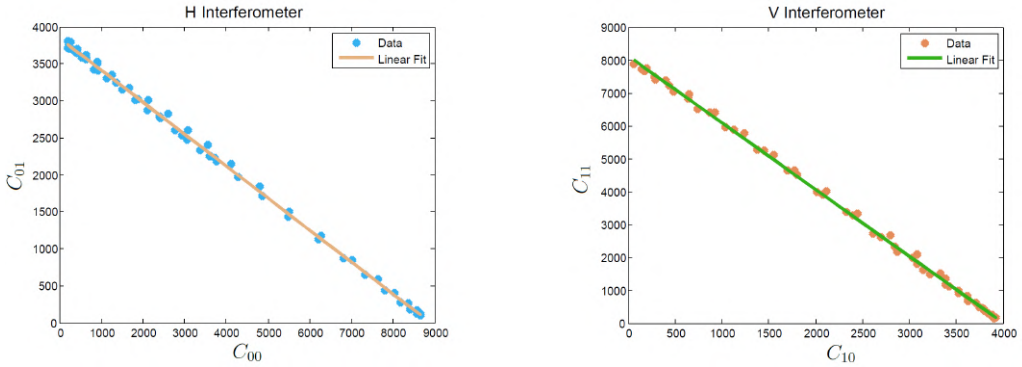


Figure 1.7 Determination of detection efficiency. Triggered coincidences detected in port 1 plotted against those detected in port 0 for both interferometers. Since total number of photons exiting the interferometer should be constant, the relative collection/detection efficiency can be determined from the slope of this line.

on the the time scale of the acquisition time this would further decrease the visibility. However, we find the phase drifts rather slowly, by approximately 0.01 rad in one minute.

To measure the causal witness we need to set 259 different waveplate settings. Moving the waveplates from one setting to the next takes approximately 30 seconds. Combined with the measurement time of two seconds this means it takes approximately 30 seconds per measurement setting. Therefore, after 30 measurements the phase drifts enough to cause a noticeable error. To combat this, we automatically reset the phase to 0 rad every 20 measurement settings by setting the waveplates to $0°$, scanning the piezo-electrically driven translation stage, and moving to the maximum of the fringe. In spite of this action there is still residual phase drift.

We performed a seperate measurement, mimicking our experimental procedure, to characterize this remaining phase drift. We set the waveplates to $0°$ so that we could directly observe the drift phase drift. As above, we counted for 2 seconds, and reset the phase to 0 rad every 20 measurements. However, the waveplates remained set to $0°$ the entire time. Therefore, in the absence of phase drift, the fringe would have remained at a maximum. By measuring the deviation from the ideal values we estimate that, over the course of our entire data run, we have a residual phase fluctuation of approximately 0.04 rad. We then propagate this error in order to estimate an error on each probability that we measure. These are the errorbars drawn in Figure 1.4 and Supplementary Figures 1-3.

1.5.5 Causal Witness Derivation for our Set-Up

Here we define what a 'causal witness' is, and sketch the algorithm that was used to compute the witness suitable for our experimental set-up. We refer the reader to Ref. [Araújo *et al.*, 2015] for an exhaustive introduction to the subject. Throughout this section we will use the Choi-Jamiołkowski isomorphism, which we introduce briefly in the ***Supplementary Materials***, Sec. I.

A process matrix $W^{\text{sep}} \in \mathcal{H}^{(\text{in})} \otimes \mathcal{H}^A \otimes \mathcal{H}^B \otimes \mathcal{H}^{(\text{out})}$ (where the Hilbert spaces refer both to the input and the output of the laboratories) is 'causally separable' if it can be written as convex combination of processes compatible with the causal order $A \to B$ and $B \to A$, that is, as $W^{\text{sep}} = pW^{A\to B} + (1-p)W^{B\to A}$.

A 'causal witness' $S \in \mathcal{H}^{(\text{in})} \otimes \mathcal{H}^A \otimes \mathcal{H}^B \otimes \mathcal{H}^{(\text{out})}$ is a Hermitian operator such that for all 'causally non-separable' process matrices $W^{\text{n-sep}}$, $\text{Tr}(S\,W^{\text{n-sep}}) < 0$, but for any process W^{sep}, $\text{Tr}(S\,W^{\text{sep}}) \geq 0$. The existence of such Hermitian operator S is justified by the separating hyperplane theorem [Rockafellar, 1970]. As a consequence of this theorem, and because the set of causally separable processes is convex, for every causally non-separable process $W^{\text{n-sep}}$ there exists a causal witness S such that $\text{Tr}(SW^{\text{n-sep}}) < 0$. This is illustrated graphically in Fig. 1.8.

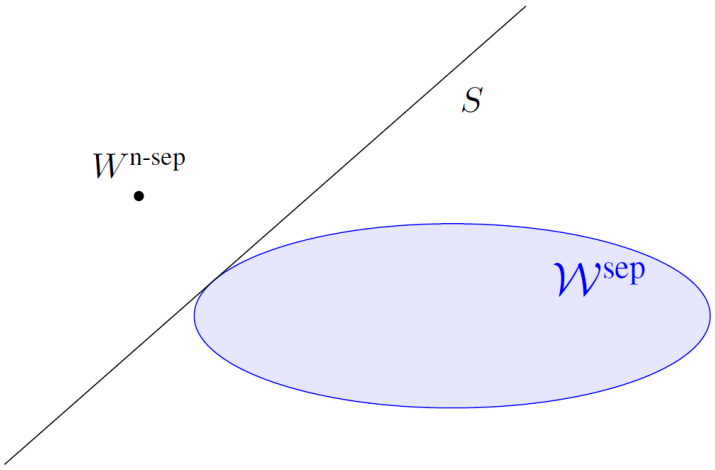


Figure 1.8 **Schematic representation of a causal witness.** In this two-dimensional representation, the causal witness is represented by the line (actually, a hyperplane) S. It separates the convex set of process matrices $\mathcal{W}^{\text{sep}}$ from a given causally non-separable process matrix $W^{\text{n-sep}}$. Because the set of causally separable processes (Eq. 1.5) is convex, the separating hyperplane theorem [Rockafellar, 1970] implies that one can always draw a hyperplane to separate it from any point outside the set (which corresponds to a causally non-separable process). This hyperplane is the causal witness.

The *optimal* causal witness S_{opt} for a given process W can be computed efficiently using a 'semidefinite program' (SDP) [Araújo *et al.*, 2015]:

$$\begin{aligned} &\min \text{Tr}(S\,W) \\ \text{s.t.} \quad S \in \mathcal{S} \qquad &\mathcal{I}^{\mathcal{H}^{(\text{in})}\otimes\mathcal{H}^A\otimes\mathcal{H}^B\otimes\mathcal{H}^{(\text{out})}}/d_{\text{out}} - S \in \mathcal{W}^* \end{aligned} \tag{1.19}$$

where $\mathcal{S}$ and $\mathcal{W}^*$ are, respectively, the set of causal witnesses and set of Hermitian operators that have non-negative trace with process matrices, as defined in [Araújo *et al.*, 2015], and $\mathcal{I}^{\mathcal{H}^{(\text{in})}\otimes\mathcal{H}^A\otimes\mathcal{H}^B\otimes\mathcal{H}^{(\text{out})}}/d_{\text{out}}$ is the identity operator on $\mathcal{H}^{(\text{in})} \otimes \mathcal{H}^A \otimes \mathcal{H}^B \otimes \mathcal{H}^{(\text{out})}$ divided by the dimension of the output spaces $d_{\text{out}} := d_{H^{(\text{in})}} d_{\mathcal{H}^A_{\text{out}}} d_{\mathcal{H}^B_{\text{out}}}$ for normalization.

The causal non-separability $\text{CNS}(W^{\text{n-sep}}) = -\text{Tr}(S_{\text{opt}}W^{\text{n-sep}})$ is the minimal $\lambda \geq 0$ such that the

process matrix

$$W_\lambda = \frac{1}{1+\lambda}\left(W^{\text{n-sep}} + \lambda\Omega\right) \tag{1.20}$$

is causally separable, optimized over all valid process matrices Ω. This means that it is the minimum amount of worst-case noise necessary to make $W^{\text{n-sep}}$ causally separable or, equivalently, the maximum (or rather the supremum) amount of worst-case noise that $W^{\text{n-sep}}$ can tolerate before becoming causally separable. Noting that $\frac{1}{1+\lambda} + \frac{\lambda}{1+\lambda} = 1$, we see that $\frac{\lambda}{1+\lambda}$ can be interpreted as the probability that the worst-case process Ω is prepared instead of the desired process W^{n-sep}, and therefore that $\frac{\text{CNS}(W)}{1+\text{CNS}(W)}$ is the maximal such probability that still allows us to see causal non-separability.

Any witness S (in particular, S_{opt}) can be decomposed with respect to a *basis* for the space $\mathcal{H}^{(\text{in})} \otimes \mathcal{H}^A \otimes \mathcal{H}^B \otimes \mathcal{H}^{(\text{out})}$. Such a basis consists of the Choi-Jamiołkowski representations of general state preparations on $\mathcal{H}^{(\text{in})}$, general measurement and repreparation operations on $\mathcal{H}^A$ and $\mathcal{H}^B$, and general measurements on $\mathcal{H}^{(\text{out})}$. Having access to such a basis of operations means being able to perform full 'causal tomography'.

However, in our experimental set-up, Alice can implement general measure-and-reprepare operations $M^A_{a,x}$, but Bob can implement only unitary operations U^B_y, and measurements are done only in the superposition basis. Thus S_{opt} will not necessarily be experimentally achievable, and in our case it is not. To compute the best witness that we can experimentally implement, we add a restriction on the decomposition of the witness as an additional constraint in the SDP, which then outputs the optimal *experimentally accessible* witness S_{exp}:

$$\begin{aligned} &\min \operatorname{Tr}(S\,W) \\ &\text{s.t.} \quad S \in \mathcal{S} \qquad \mathcal{I}^{\mathcal{H}^{(\text{in})}\otimes\mathcal{H}^A\otimes\mathcal{H}^B\otimes\mathcal{H}^{(\text{out})}}/d_{\text{out}} - S \in \mathcal{W}^* \\ &S = \sum_{z=0}^{2}\sum_{a=0}^{1}\sum_{x=0}^{11}\sum_{y=0}^{9}\sum_{d=0}^{1} \alpha_{a,d,x,y,z}\cdot \rho^{(\text{in})}_z \otimes M^A_{a,x} \otimes U^B_y \otimes D^{(\text{out})}_d \end{aligned} \tag{1.21}$$

where $\{M^A_{a,x}\}$ are the 24 Choi-Jamiołkowski representations of measurement-repreparation maps, among which 16 are linearly independent, $\{U^B_y\}$ the 10 linearly independent Choi-Jamiołkowski representations unitaries, which are listed in Sec. II, and $\{D^{(\text{out})}_d\}$ the two projectors onto the superposition basis.

The algorithm 1.21 returns the coefficients $\alpha_{a,d,x,y,z}$, which are used to weigh the experimental probabilities $p(a,d|x,y,z)$ corresponding to $\operatorname{Tr}\left[(\rho^{(\text{in})}_z \otimes M^A_{a,x} \otimes U^B_y \otimes D^{(\text{out})}_d)\cdot W_{\text{switch}}\right]$ to compute the experimental value for $\operatorname{Tr}(S_{\text{exp}}W_{\text{switch}})$.

Analogously to the ideal case, the 'experimentally accessible causal non-separability' (*i.e.*, $\text{CNS}_{\text{exp}}(W_{\text{switch}}) = -\operatorname{Tr}(S_{\text{exp}}W_{\text{switch}})$) is the maximal amount of worst-case noise that can be admixed to W_{switch} before our experimental set-up becomes incapable of certifying that W_{switch} is

causally non-separable, and $\frac{\text{CNS}_{\text{exp}}(W_{\text{switch}})}{1+\text{CNS}_{\text{exp}}(W_{\text{switch}})}$ the maximal probability of preparing the worst-case noise process instead of the ideal W_{switch}.

1.6 Supplemental Information

1.6.1 Choi-Jamiołkowski Isomorphism

In this section we define the *Choi-Jamiołkowski isomorphism* [Choi, 1975], as used in the main text. For the sake of clarity, it is convenient to first introduce the isomorphism for unitary operators and then to generalize it to linear maps.

1.6.1.1 Unitary operators.

Consider an operator $U \in \mathcal{L}(\mathcal{H}_{pur}^{\text{in}})$ (where $\mathcal{L}(\mathcal{H}_{pur})$ is the space of linear operators in the Hilbert space of pure states $\mathcal{H}_{pur}$) such that $U : \mathcal{H}_{pur}^{\text{in}} \to \mathcal{H}_{pur}^{\text{out}}$. It is always possible to rewrite the operator U acting on the basis $\{|j\rangle\} \in \mathcal{H}_{pur}^{\text{in}}$ into a new basis $\{|k\rangle\} \in \mathcal{H}_{pur}^{\text{out}}$

$$U = \sum_{j,k} U_{k,j} \, |k\rangle \langle j| \tag{1.22}$$

where $U_{k,j} = \langle k| \, U \, |j\rangle$.

According to the Choi-Jamiołkowski isomorphism, a linear operator acting on a Hilbert space $\mathcal{H}_{pur}^{\text{in}}$ is isomorphic to a vector in $\mathcal{H}_{pur}^{\text{in}} \otimes \mathcal{H}_{pur}^{\text{out}}$ [Araújo *et al.*, 2015; Leifer and Spekkens, 2013]:

$$\text{CJ} : \mathcal{L}(\mathcal{H}_{pur}^{\text{in}}) \to \mathcal{H}_{pur}^{\text{in}} \otimes \mathcal{H}_{pur}^{\text{out}} \tag{1.23}$$

We define a (non-normalized) maximally entangled state via the double-ket notation, as

$$|\mathbb{1}\rangle\rangle^{\text{in, in}} = \sum_j |j\rangle \otimes |j\rangle \in \mathcal{H}_{pur}^{\text{in}} \otimes \mathcal{H}_{pur}^{\text{in}} \tag{1.24}$$

The application of the Choi-Jamiołkowski isomorphism on the linear operator U gives

$$\text{CJ}(U) = (\mathbb{1} \otimes U^*) \, |\mathbb{1}\rangle\rangle^{\text{in, in}} = \sum_j |j\rangle \otimes U^* \, |j\rangle = \sum_{j,k} U_{k,j}^* \, |j\rangle \otimes |k\rangle := |U^*\rangle\rangle^{\text{in, out}} \tag{1.25}$$

where U^* is the complex conjugate and $|U^*\rangle\rangle$ is the Choi-Jamiołkowski representation (or Choi-Jamiołkowski vector) of U.

By comparing Eqs. 1.22 and 1.25, one can see the direct correspondence between the operator $|k\rangle\langle j|$ and the vector $|j\rangle \otimes |k\rangle$. Thus, a unitary operator is represented as a pure state in the Choi-Jamiołkowski isomorphism.

1.6.1.2 Linear maps.

Consider a linear map $\mathcal{M} : \mathcal{H}^{\text{in}} \to \mathcal{H}^{\text{out}}$ (where $\mathcal{H}$ is the Hilbert space of the density matrices, *i.e.*, $\mathcal{L}(\mathcal{H}_{pur})$). The corresponding Choi-Jamiołkowski representation is

$$M^{\text{in, out}} := \left[\mathcal{I} \otimes \mathcal{M}(|\mathbb{1}\rangle\rangle\langle\langle\mathbb{1}|)\right]^T \tag{1.26}$$

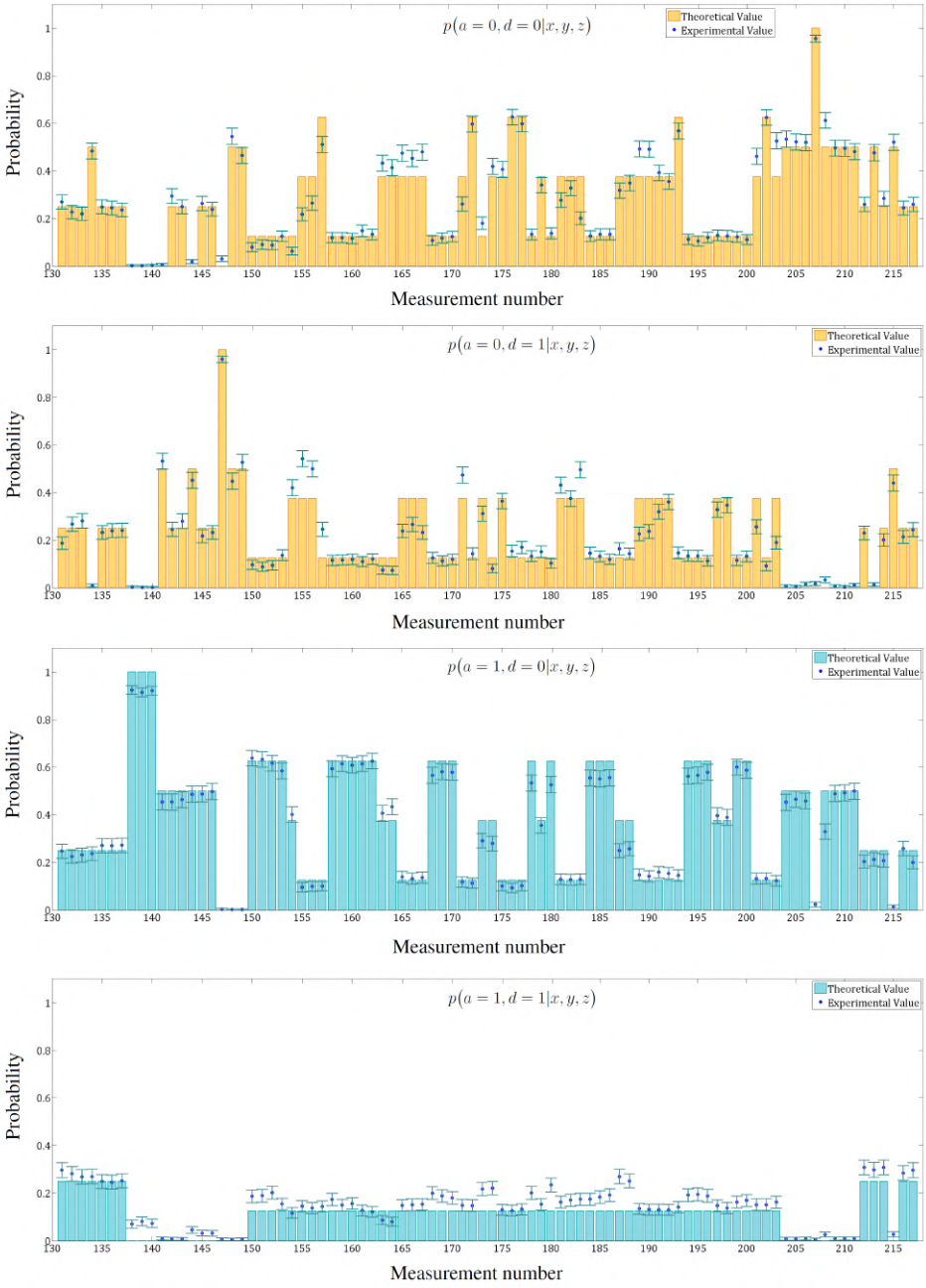
This equation can be led back to Eq. 1.25 when $\mathcal{M}(\rho) = U\rho U^\dagger$:

$$\begin{aligned} M^{\text{in, out}} :&= \left[\mathcal{I} \otimes U \left(\sum_{j,j_1} |j\rangle \otimes |j\rangle\langle j_1| \otimes \langle j_1|\right) U^\dagger\right]^T = \\ &= \left[\sum_{j,j_1,k,k_1} U^*_{k,j} U_{j_1,k_1} |j\rangle \otimes |k\rangle\langle j_1| \otimes \langle k_1|\right]^T = |U^*\rangle\rangle\langle\langle U^*| \end{aligned} \tag{1.27}$$

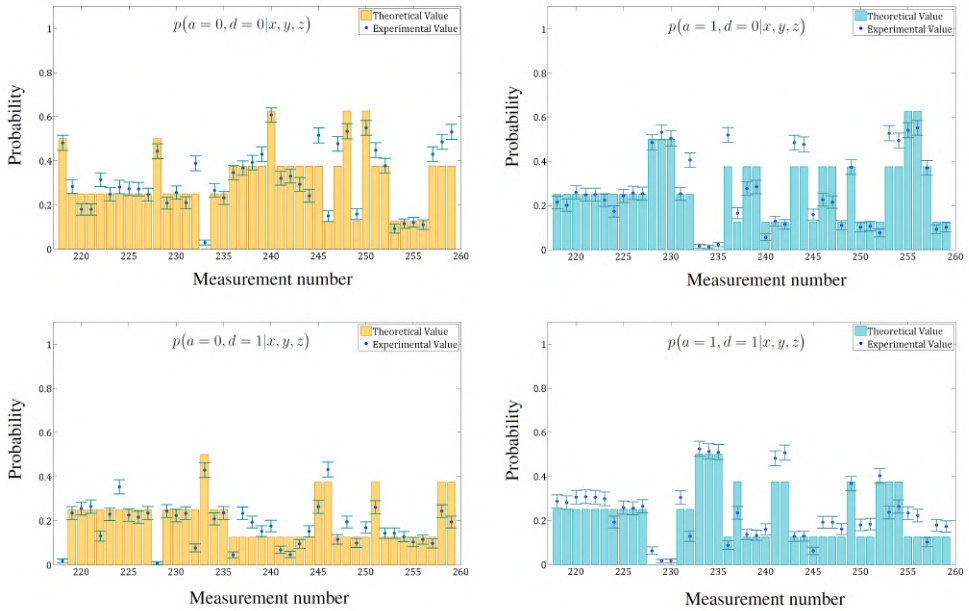
Thus, an arbitrary linear map is represented by a matrix in the Choi-Jamiołkowski isomorphism.

By looking at the action of a linear operator, such as a time evolution $|j\rangle \mapsto |k\rangle$, it turns out that in the Choi-Jamiołkowski notation it is not required to ascribe a temporal order to the application of different operators. Consequently, analyzing situations in which the causal order of events is not well-defined (such as in the quantum switch) in the Choi-Jamiołkowski isomorphism is very convenient.

S 1.1 Experimentally estimated probabilities. Continuation of Figure 4.



S 1.2 Experimentally estimated probabilities. Continuation of Figure 4.



S 1.3 Experimentally estimated probabilities. Continuation of Figure 4.

S 1.4 List of all the experimental measurement settings and the corresponding coefficients. i is the setting number (corresponding to the x-axis in Figure 4, and in Supplementary Figures 1.1-1.3), z is the input state ($1 = |H\rangle$, $2 = |V\rangle$, and $3 = |+\rangle$), y is Bob's choice of unitary (corresponding to the third column of Tables 1 and 2), x_M (x_R) is Alice's choice of measurement (repreparation) operator (corresponding to the first (second) column of Tables 1 and 2), the $\alpha_{m,n,x,y,z}$ are the weights of outcome mn required for calculating the witness.

i	z	y	x_M	x_R	$\alpha_{0,0,x,y,z}$	$\alpha_{0,1,x,y,z}$	$\alpha_{1,0,x,y,z}$	$\alpha_{1,1,x,y,z}$
1	1	1	1	1	0.005427	0.054415	0.116229	0.145418
2	1	1	1	2	0.037881	0.000000	0.000000	0.000000
3	1	1	1	3	-0.017971	0.000000	0.000000	0.000000
4	1	1	2	1	0.020993	0.000000	-0.011450	-0.013343
5	1	1	2	2	-0.009789	0.000000	0.000000	0.000000
6	1	1	2	3	0.000000	0.009420	0.000000	0.000000
7	1	1	3	1	-0.017971	0.000000	0.009578	-0.035943
8	1	1	3	2	0.027549	0.000000	0.000000	0.000000
9	1	1	3	3	0.000000	-0.035943	0.000000	0.000000
10	1	1	4	1	0.074656	0.112073	0.149339	0.180312
11	1	1	4	2	0.014995	0.000000	0.000000	0.000000
12	1	1	4	3	-0.008551	0.000000	0.000000	0.000000

S 1.5 List of all the experimental measurement settings and the corresponding coefficients. i is the setting number (corresponding to the x-axis in Figure 4, and in Supplementary Figures 1.1-1.3), z is the input state ($1 = |H\rangle$, $2 = |V\rangle$, and $3 = |+\rangle$), y is Bob's choice of unitary (corresponding to the third column of Tables 1 and 2), x_M (x_R) is Alice's choice of measurement (repreparation) operator (corresponding to the first (second) column of Tables 1 and 2), the $\alpha_{m,n,x,y,z}$ are the weights of outcome mn required for calculating the witness.

i	z	y	x_M	x_R	$\alpha_{0,0,x,y,z}$	$\alpha_{0,1,x,y,z}$	$\alpha_{1,0,x,y,z}$	$\alpha_{1,1,x,y,z}$
13	1	2	1	1	-0.078315	0.010375	0.064196	0.097329
14	1	2	1	2	0.120973	0.000000	0.000000	0.000000
15	1	2	1	3	-0.003813	0.000000	0.000000	0.000000
16	1	2	2	1	0.051055	0.000000	-0.052753	-0.024583
17	1	2	2	2	-0.067166	0.000000	0.000000	0.000000
18	1	2	2	3	0.000000	0.009420	0.000000	0.000000
19	1	2	3	1	-0.003813	0.000000	0.023736	-0.007626
20	1	2	3	2	0.027549	0.000000	0.000000	0.000000
21	1	2	3	3	0.000000	-0.007626	0.000000	0.000000
22	1	2	4	1	-0.023959	0.045125	0.088305	0.119648
23	1	2	4	2	0.098087	0.000000	0.000000	0.000000
24	1	2	4	3	0.005607	0.000000	0.000000	0.000000
25	1	3	1	1	0.000000	0.073456	0.000000	0.000000
26	1	3	1	2	0.000000	-0.023402	0.000000	0.000000
27	1	3	1	3	0.000000	-0.070162	0.000000	-0.076716
28	1	3	2	1	0.000000	-0.067408	0.000000	0.000000
29	1	3	2	2	0.000000	0.045902	0.000000	0.000000
30	1	3	2	3	0.000000	0.009420	-0.054865	-0.010342
31	1	3	3	1	0.000000	-0.070162	0.000000	0.000000
32	1	3	3	2	0.000000	-0.027549	0.000000	0.000000
33	1	3	3	3	0.000000	0.140324	-0.000000	0.042613
34	1	3	4	1	0.000000	0.015468	0.000000	0.000000
35	1	3	4	3	0.000000	-0.079582	-0.054865	-0.119495
36	1	4	1	1	0.000000	0.088690	0.000000	0.000000
37	1	4	1	2	0.000000	-0.032990	0.000000	0.000000
38	1	4	1	3	0.000000	0.003813	0.000000	-0.009079
39	1	4	2	1	0.000000	-0.029026	0.000000	0.000000
40	1	4	2	2	0.000000	0.021395	0.000000	0.000000
41	1	4	2	3	0.000000	0.009420	-0.016783	-0.010642
42	1	4	3	1	0.000000	0.003813	0.000000	0.000000
43	1	4	3	2	0.000000	-0.027549	0.000000	0.000000
44	1	4	3	3	0.000000	-0.007626	-0.000000	-0.031362

i	z	y	x_M	x_R	$\alpha_{0,0,x,y,z}$	$\alpha_{0,1,x,y,z}$	$\alpha_{1,0,x,y,z}$	$\alpha_{1,1,x,y,z}$
45	1	4	4	1	0.000000	0.047055	0.000000	0.000000
46	1	4	4	2	0.000000	-0.010104	0.000000	0.000000
47	1	4	4	3	0.000000	-0.005607	-0.036642	-0.071253
48	1	5	1	1	0.020724	0.000000	-0.011369	0.000000
49	1	5	1	2	-0.044086	0.000000	0.000000	0.000000
50	1	5	1	3	0.055770	0.000000	0.000000	0.000000
51	1	5	2	1	-0.041202	0.000000	-0.002876	-0.000522
52	1	5	3	1	-0.010092	0.000000	0.035178	-0.033673
53	1	5	4	1	0.069645	0.000000	-0.038195	0.022095
54	1	5	4	2	-0.041871	0.000000	0.000000	0.000000
55	1	5	4	3	-0.003725	0.000000	0.000000	0.000000
56	1	6	1	1	0.000000	0.017241	0.000000	-0.062504
57	1	6	1	2	0.000000	-0.012843	0.000000	0.000000
58	1	6	1	3	0.000000	-0.039821	0.000000	0.000000
59	1	6	2	1	0.000000	0.020875	0.038326	0.061007
60	1	6	2	3	0.000000	0.040654	0.000000	0.000000
61	1	6	3	1	0.000000	-0.081788	-0.153383	-0.071596
62	1	6	3	2	-0.113858	0.000000	0.000000	0.000000
63	1	6	4	1	0.000000	-0.001932	-0.018467	-0.061797
64	1	6	4	2	0.000000	-0.015057	0.000000	0.000000
65	1	6	4	3	0.000000	-0.039821	0.000000	0.000000
66	1	7	1	1	0.000000	0.083643	0.000000	0.000000
67	1	7	1	2	0.000000	-0.056929	0.000000	0.000000
68	1	7	1	3	0.000000	-0.095591	0.000000	-0.008747
69	1	7	2	1	0.000000	-0.020327	0.000000	0.000000
70	1	7	2	3	0.000000	0.040654	-0.000000	0.020327
71	1	7	3	1	0.000000	-0.071696	0.000000	0.000000
72	1	7	3	2	0.000000	0.113858	0.005744	0.018681
73	1	7	4	1	0.000000	0.053896	0.000000	0.000000
74	1	7	4	2	0.000000	-0.056929	0.000000	0.000000
75	1	7	4	3	0.000000	-0.036096	-0.000000	0.021000
76	1	8	1	1	0.104368	0.000000	-0.004718	0.000000
77	1	8	1	2	0.000000	0.043897	0.000000	0.000000
78	1	8	1	3	0.000000	0.039821	0.000000	0.000000
79	1	8	2	1	0.000000	-0.019174	-0.018467	0.002921
80	1	8	3	1	0.000000	0.081788	0.039525	0.071596
81	1	8	4	1	0.042838	0.000000	-0.027922	-0.084733

i	z	y	x_M	x_R	$\alpha_{0,0,x,y,z}$	$\alpha_{0,1,x,y,z}$	$\alpha_{1,0,x,y,z}$	$\alpha_{1,1,x,y,z}$
82	1	9	1	1	0.000000	-0.066402	0.000000	0.017154
83	1	9	1	2	0.000000	0.073277	0.000000	0.000000
84	1	9	1	3	0.000000	0.055770	0.000000	0.000000
85	1	9	2	1	0.000000	0.010574	-0.018467	-0.026826
86	1	9	2	3	0.000000	-0.059495	0.000000	0.000000
87	1	9	3	1	0.000000	-0.010092	0.013490	0.023581
88	1	9	3	2	0.058759	0.000000	0.000000	0.000000
89	1	9	4	1	0.000000	-0.025200	0.038326	0.014277
90	1	9	4	2	0.000000	0.029720	0.000000	0.000000
91	1	9	4	3	0.000000	0.055770	0.000000	0.000000
92	1	10	1	1	0.000000	-0.083643	0.000000	0.000000
93	1	10	1	2	0.000000	0.029380	0.000000	0.000000
94	1	10	1	3	0.000000	0.095591	0.000000	-0.018802
95	1	10	2	1	0.000000	0.029747	0.000000	0.000000
96	1	10	2	3	0.000000	-0.059495	-0.000000	-0.029747
97	1	10	3	1	0.000000	0.071696	0.000000	0.000000
98	1	10	3	2	0.000000	-0.058759	-0.005744	0.036418
99	1	10	4	1	0.000000	-0.063316	0.000000	0.000000
100	1	10	4	2	0.000000	0.029380	0.000000	0.000000
101	1	10	4	3	0.000000	0.054937	-0.000000	-0.039129
102	2	1	1	1	0.128541	0.113366	0.046187	0.080971
103	2	1	1	2	-0.027217	0.000000	0.000000	0.000000
104	2	1	1	3	-0.011443	0.000000	0.000000	0.000000
105	2	1	2	1	0.021849	0.000000	0.009266	0.008230
106	2	1	2	2	-0.009789	0.000000	0.000000	0.000000
107	2	1	2	3	0.000000	0.009420	0.000000	0.000000
108	2	1	3	1	-0.082301	0.000000	0.030043	-0.079808
109	2	1	3	2	0.027549	0.000000	0.000000	0.000000
110	2	1	3	3	0.000000	-0.035943	0.000000	0.000000
111	2	1	4	1	0.148433	0.122433	0.055408	0.091227
112	2	1	4	2	-0.050102	0.000000	0.000000	0.000000
113	2	1	4	3	-0.002023	0.000000	0.000000	0.000000
114	2	2	1	1	0.060564	0.093553	0.081765	0.112029
115	2	2	1	2	0.055875	0.000000	0.000000	0.000000
116	2	2	1	3	0.002715	0.000000	0.000000	0.000000
117	2	2	2	1	0.051912	0.000000	-0.032037	-0.003010
118	2	2	2	2	-0.067166	0.000000	0.000000	0.000000

i	z	y	x_M	x_R	$\alpha_{0,0,x,y,z}$	$\alpha_{0,1,x,y,z}$	$\alpha_{1,0,x,y,z}$	$\alpha_{1,1,x,y,z}$
119	2	2	2	3	0.000000	0.009420	0.000000	0.000000
120	2	2	3	1	-0.068143	0.000000	0.044201	-0.051491
121	2	2	3	2	0.027549	0.000000	0.000000	0.000000
122	2	2	3	3	0.000000	-0.007626	0.000000	0.000000
123	2	2	4	1	0.073033	0.091625	0.077693	0.100960
124	2	2	4	2	0.032990	0.000000	0.000000	0.000000
125	2	2	4	3	0.012135	0.000000	0.000000	0.000000
126	2	3	1	1	0.000000	0.013468	0.000000	0.000000
127	2	3	1	2	0.000000	0.041695	0.000000	0.000000
128	2	3	1	3	0.000000	-0.076690	0.000000	-0.075298
129	2	3	2	1	0.000000	-0.068265	0.000000	0.000000
130	2	3	2	2	0.000000	0.045902	0.000000	0.000000
131	2	3	2	3	0.000000	0.009420	-0.035006	0.010374
132	2	3	3	1	0.000000	-0.005832	0.000000	0.000000
133	2	3	3	2	0.000000	-0.027549	0.000000	0.000000
134	2	3	3	3	0.000000	0.140324	0.084795	0.063078
135	2	3	4	1	0.000000	0.012945	0.000000	0.000000
136	2	3	4	2	0.000000	0.064581	0.000000	0.000000
137	2	3	4	3	0.000000	-0.086110	0.035006	-0.085671
138	2	4	1	1	0.000000	-0.031252	0.000000	0.000000
139	2	4	1	2	0.000000	0.032108	0.000000	0.000000
140	2	4	1	3	0.000000	-0.002715	0.000000	0.052293
141	2	4	2	1	0.000000	-0.029883	0.000000	0.000000
142	2	4	2	2	0.000000	0.021395	0.000000	0.000000
143	2	4	2	3	0.000000	0.009420	0.003077	0.010074
144	2	4	3	1	0.000000	0.068143	0.000000	0.000000
145	2	4	3	2	0.000000	-0.027549	0.000000	0.000000
146	2	4	3	3	0.000000	-0.007626	0.084795	-0.010897
147	2	4	4	1	0.000000	-0.069392	0.000000	0.000000
148	2	4	4	2	0.000000	0.054993	0.000000	0.000000
149	2	4	4	3	0.000000	-0.012135	0.016783	0.040049
150	2	5	1	1	-0.055127	0.000000	0.030570	0.000000
151	2	5	1	2	0.041871	0.000000	0.000000	0.000000
152	2	5	1	3	0.003725	0.000000	0.000000	0.000000
153	2	5	2	1	-0.002354	0.000000	-0.040680	0.000522
154	2	5	3	1	0.068851	0.000000	0.023581	0.033673
155	2	5	4	1	-0.006207	0.000000	0.003744	0.022095

i	z	y	x_M	x_R	$\alpha_{0,0,x,y,z}$	$\alpha_{0,1,x,y,z}$	$\alpha_{1,0,x,y,z}$	$\alpha_{1,1,x,y,z}$
156	2	5	4	2	0.044086	0.000000	0.000000	0.000000
157	2	5	4	3	-0.055770	0.000000	0.000000	0.000000
158	2	6	1	1	0.000000	0.079644	0.000000	-0.038949
159	2	6	1	2	0.000000	-0.098800	0.000000	0.000000
160	2	6	1	3	0.000000	-0.039821	0.000000	0.000000
161	2	6	2	1	0.000000	-0.017973	-0.038326	0.023203
162	2	6	2	3	0.000000	0.040654	0.000000	0.000000
163	2	6	3	1	0.000000	-0.081788	-0.153383	-0.071596
164	2	6	3	2	-0.113858	0.000000	0.000000	0.000000
165	2	6	4	1	0.000000	0.060471	-0.018467	-0.038243
166	2	6	4	2	0.000000	-0.101015	0.000000	0.000000
167	2	6	4	3	0.000000	-0.039821	0.000000	0.000000
168	2	7	1	1	0.000000	0.097092	0.000000	0.000000
169	2	7	1	2	0.000000	-0.056929	0.000000	0.000000
170	2	7	1	3	0.000000	-0.043546	0.000000	-0.074241
171	2	7	2	1	0.000000	-0.020327	0.000000	0.000000
172	2	7	2	3	0.000000	0.040654	-0.000000	0.020327
173	2	7	3	1	0.000000	-0.150638	0.000000	0.000000
174	2	7	3	2	0.000000	0.113858	-0.084795	0.007084
175	2	7	4	1	0.000000	0.067345	0.000000	0.000000
176	2	7	4	2	0.000000	-0.056929	0.000000	0.000000
177	2	7	4	3	0.000000	0.015948	-0.000000	-0.044493
178	2	8	1	1	0.058177	0.000000	-0.041754	0.000000
179	2	8	2	1	0.000000	0.021388	0.018467	-0.000707
180	2	8	3	1	0.000000	0.032070	-0.039525	0.042262
181	2	8	4	1	-0.003353	0.000000	-0.064957	-0.084733
182	2	8	4	2	0.000000	-0.043897	0.000000	0.000000
183	2	8	4	3	0.000000	-0.039821	0.000000	0.000000
184	2	9	1	1	0.000000	-0.064564	0.000000	0.059553
185	2	9	1	2	0.000000	0.029039	0.000000	0.000000
186	2	9	1	3	0.000000	0.055770	0.000000	0.000000
187	2	9	2	1	0.000000	0.051135	0.018467	-0.030454
188	2	9	2	3	0.000000	-0.059494	0.000000	0.000000
189	2	9	3	1	0.000000	-0.010092	0.013490	0.023581
190	2	9	3	2	0.058759	0.000000	0.000000	0.000000
191	2	9	4	1	0.000000	-0.023362	0.038326	0.056677

i	z	y	x_M	x_R	$\alpha_{0,0,x,y,z}$	$\alpha_{0,1,x,y,z}$	$\alpha_{1,0,x,y,z}$	$\alpha_{1,1,x,y,z}$
192	2	9	4	2	0.000000	-0.014518	0.000000	0.000000
193	2	9	4	3	0.000000	0.055770	0.000000	0.000000
194	2	10	1	1	0.000000	-0.039290	0.000000	0.000000
195	2	10	1	2	0.000000	0.029380	0.000000	0.000000
196	2	10	1	3	0.000000	0.056602	0.000000	-0.024166
197	2	10	2	1	0.000000	0.029747	0.000000	0.000000
198	2	10	2	3	0.000000	-0.059494	-0.000000	-0.029747
199	2	10	3	1	0.000000	0.021979	0.000000	0.000000
200	2	10	3	2	0.000000	-0.058759	-0.084795	0.007084
201	2	10	4	1	0.000000	-0.018963	0.000000	0.000000
202	2	10	4	2	0.000000	0.029380	0.000000	0.000000
203	2	10	4	3	0.000000	0.015948	-0.000000	-0.044493
204	3	1	1	1	0.006362	0.000000	0.005704	0.000000
205	3	1	1	2	-0.005538	0.000000	0.000000	0.000000
206	3	1	1	3	-0.006528	0.000000	0.000000	0.000000
207	3	1	2	1	-0.000747	0.000000	0.024700	0.023953
208	3	1	3	1	0.064330	0.000000	-0.020465	0.043865
209	3	1	4	1	0.007219	0.000000	0.026420	0.021573
210	3	1	4	2	-0.005538	0.000000	0.000000	0.000000
211	3	1	4	3	-0.006528	0.000000	0.000000	0.000000
212	3	2	1	1	0.062140	0.000000	0.058734	0.000000
213	3	2	1	2	-0.114346	0.000000	0.000000	0.000000
214	3	2	1	3	-0.006528	0.000000	0.000000	0.000000
215	3	2	2	1	-0.003495	0.000000	0.049929	0.046434
216	3	2	3	1	0.064330	0.000000	-0.020465	0.043865
217	3	2	4	1	0.062997	0.000000	0.079450	0.021573
218	3	2	4	2	-0.114346	0.000000	0.000000	0.000000
219	3	2	4	3	-0.006528	0.000000	0.000000	0.000000
220	3	3	1	1	0.000000	-0.002186	0.000000	0.000000
221	3	3	1	2	0.000000	-0.059532	0.000000	0.000000
222	3	3	1	3	0.000000	0.006528	0.000000	0.055190
223	3	3	2	1	0.000000	0.057465	0.089871	0.032405
224	3	3	3	1	0.000000	-0.064330	-0.084795	-0.020465
225	3	3	4	1	0.000000	-0.003043	0.000000	0.000000
226	3	3	4	2	0.000000	-0.059532	0.000000	0.000000
227	3	3	4	3	0.000000	0.006528	0.019859	0.075906

i	z	y	x_M	x_R	$\alpha_{0,0,x,y,z}$	$\alpha_{0,1,x,y,z}$	$\alpha_{1,0,x,y,z}$	$\alpha_{1,1,x,y,z}$
228	3	4	1	1	0.000000	0.002101	0.000000	0.000000
229	3	4	1	2	0.000000	-0.015849	0.000000	0.000000
230	3	4	1	3	0.000000	0.006528	0.000000	0.007220
231	3	4	2	1	0.000000	0.005209	0.013706	0.008497
232	3	4	3	1	0.000000	-0.064330	-0.084795	-0.020465
233	3	4	4	1	0.000000	0.001244	0.000000	0.000000
234	3	4	4	2	0.000000	-0.015849	0.000000	0.000000
235	3	4	4	3	0.000000	0.006528	0.019859	0.027937
236	3	5	1	1	-0.001838	0.000000	0.001838	0.000000
237	3	5	4	1	-0.042400	0.000000	-0.001790	-0.044190
238	3	6	1	1	0.000000	-0.042515	0.000000	0.003526
239	3	6	1	3	-0.038989	0.000000	0.000000	0.000000
240	3	6	3	1	-0.049717	0.000000	0.029334	-0.020384
241	3	6	4	1	0.000000	-0.001953	0.036934	-0.000102
242	3	6	4	3	-0.038989	0.000000	0.000000	0.000000
243	3	7	1	1	0.000000	-0.044353	0.000000	0.000000
244	3	7	1	3	0.000000	0.038989	0.000000	0.005364
245	3	7	3	1	0.000000	0.049717	0.079051	0.029334
246	3	7	4	1	0.000000	-0.044353	0.000000	0.000000
247	3	7	4	3	0.000000	0.038989	0.000000	0.005364
248	3	8	1	1	-0.062403	0.000000	0.062403	0.000000
249	3	8	4	1	-0.023555	0.000000	0.024599	0.001044
250	3	9	1	1	0.000000	0.048954	0.000000	0.003091
251	3	9	1	3	0.052045	0.000000	0.000000	0.000000
252	3	9	3	1	-0.078942	0.000000	0.011597	-0.067346
253	3	9	4	1	0.000000	0.010106	-0.076652	-0.034713
254	3	9	4	3	0.052045	0.000000	0.000000	0.000000
255	3	10	1	1	0.000000	-0.013449	0.000000	0.000000
256	3	10	1	3	0.000000	-0.052045	0.000000	0.065494
257	3	10	3	1	0.000000	0.078942	0.090539	0.011597
258	3	10	4	1	0.000000	-0.013449	0.000000	0.000000
259	3	10	4	3	0.000000	-0.052045	-0.000000	0.065494

Chapter 2

Experimental Entanglement of Temporal Orders

G. Rubino, L. A. Rozema, F. Massa, M. Araújo, M. Zych, Č. Brukner, and P. Walther

Abstract. The study of causal relations has recently been applied to the quantum realm, leading to the discovery that not all quantum processes have a definite causal structure. While such processes have previously been experimentally demonstrated, these demonstrations relied on the assumption that quantum theory can be applied to causal structures and laboratory operations. Here, we present the first demonstration of entangled temporal orders beyond the quantum formalism. We do so by proving the incompatibility of our experimental outcomes with a class of generalized probabilistic theories which satisfy the assumptions of locality and definite temporal orders. To this end, we derive physical constraints (in the form of a Bell-like inequality) on experimental outcomes within such a class of theories. We then experimentally invalidate these theories by violating the inequality, thus providing an experimental proof, outside the quantum formalism, that nature is incompatible with the assumption that the temporal order between events is definite locally.

Author contributions: G.R., L.A.R., Č.B. and P.W. designed the experiment. G.R. built the set-up and carried out the data collection. G.R. and L.A.R. performed data analysis. F.M. built the single-photon source. G.R., M.Z., M.A. and Č.B. developed the theoretical idea. L.A.R., P.W. and Č.B. supervised the project. All authors contributed to writing the paper, based on an initial draft by G.R..

2.1 Introduction

Bell's theorem revolutionized the foundations of physics, leading to experiments which could demonstrate that nature cannot be described by a local-causal theory, and paving the way for modern quantum information [Bell, 1964; Brunner *et al.*, 2014]. One of the strengths of Bell's theorem is that it allows one to draw conclusions about nature without referring to the underlying physical theory. Over the past decades, tests of Bell's theorem have been performed with many different physical systems thereby entangling various observables (such as spin [Hensen *et al.*, 2015; Lamehi-Rachti and Mittig, 1976; Rowe *et al.*, 2001], polarization [Aspect *et al.*, 1982; Freedman and Clauser, 1972; Giustina *et al.*, 2015; Shalm *et al.*, 2015], position [Howell *et al.*, 2004], and energy [Kwiat *et al.*, 1993; Rarity and Tapster, 1990]) of two or more particles. However, since there is no observable associated to the measurement of the temporal order between events, this test has never been applied to the study of causal structures.

Thus far, in all the well-established physical theories, it was assumed that the order between events is pre-defined. Nevertheless, it was recently realized that quantum mechanics also allows for the existence of processes that are neither causally ordered, nor a probabilistic mixture of causally ordered processes. For example, in quantum mechanics, quantum channels and quantum states are processes with a definite causal order, meaning that they enable either one-way-signalling (*i.e.*, from a 'cause' to an 'effect') quantum channels, or no-signaling. Under 'processes', we define the set of causal relations between operations performed in different local laboratories [Brukner, 2014; Chiribella *et al.*, 2013; Oreshkov *et al.*, 2012]. More precisely, a quantum process is called *causally separable* if it can be decomposed as a convex combination of causally ordered processes, otherwise it is *causally non-separable*. (Note that the term 'temporal' order is used here to refer to operations which cannot be used to receive signals — in particular, to unitary ones — whereas 'causal' order refers to more general operations which allow for the exchange of information.) Recently, a method for certifying causal separability, based on 'causal witnesses', was developed [Araújo *et al.*, 2015; Branciard, 2016; Oreshkov and Giarmatzi, 2016], and used to experimentally demonstrate that a certain process — a quantum switch [Chiribella, 2012] — is causally non-separable [Goswami *et al.*, 2018; Rubino *et al.*, 2017a].

In the quantum switch, a qubit is transmitted between two parties, and the order in which the parties receive and act on it is entangled with a second system. This can result in a superposition of temporal orders in which operations are applied on the system in different orders. The existence of such a superposition has been experimentally demonstrated [Goswami *et al.*, 2018; Procopio *et al.*, 2015; Rubino *et al.*, 2017a]. However, the certification of this 'indefiniteness' of temporal orders was theory-dependent, requiring the assumption that the system under investigation and the applied operations were described by quantum theory. In more detail, Ref. [Goswami *et al.*, 2018; Rubino *et al.*, 2017a] reported the measurement of a value for a causal witness that could not be explained

by any model making the following three assumptions: there was a definite causal order between the parties, each party acted only once, and the quantum description of their operations was the correct one. Nevertheless, the results of these experiments could potentially have also been explained in a causal manner within a different theory (*i.e.*, outside the quantum theory). Thus, the nature of indefinite causal orders has not yet been probed without the assumption that quantum formalism provides an adequate description of the indefinite causal orders.

In addition to theory-dependent causal witnesses, there are also device-independent ways of certifying indefinite causal orders via 'causal inequalities' [Branciard *et al.*, 2016; Oreshkov *et al.*, 2012]. These inequalities only require one to measure the probabilities of outcomes for different parties in the process under study. Any probabilities that show signalling in only one direction — which can be interpreted as an influence from the past to the future —, or that is a convex mixture of those which allow signalling only in one direction (from A to B or from B to A), satisfy causal inequalities. Nevertheless, it can be shown that the quantum switch satisfies all such causal inequalities (see Refs. [Araújo *et al.*, 2015; Oreshkov and Giarmatzi, 2016] or the Suppl. Information for details), and, currently, it is not known how to realize a process which violates a causal inequality. The question then arises if it is at all possible to prove the existence of an indefinite causal order in a manner that applies to a broader class of theories, not only to quantum theory. Here, we provide an affirmative answer to this question by experimentally violating a Bell inequality applied to temporal orders, thereby demonstrating that the order of events in our experiment is incompatible with a class of so-called '*generalized probabilistic theories*' in which the states and the laboratory operations are local, and the operations are applied in a definite order. We stress that, while our inequality is valid for a class of generalized probabilistic theories, it does depend on the internal functionality of experimental devices —*i.e.*, it exploits information concerning the experimental devices composing the set-up—, and in this sense it does not have the same 'device-independent' state as the original Bell's theorem.

In our work, we generalize a Bell inequality for temporal orders [Zych *et al.*, 2019], and then experimentally violate it. The experimental violation of the Bell inequality presented here demonstrates, independent of quantum formalism, that nature is incompatible with a class of theories which assumes the order of events as locally pre-defined.

2.2 Theoretical Framework and Experiment

We now introduce a *no-go theorem* for definite temporal orders that applies to a class of generalized probabilistic theories (GPTs) in which the order of local events is assumed to be pre-defined. GPTs are a general framework that specifies a set of operations which can be applied on physical systems, assigns probabilities to experimental outcomes [Dakic and Brukner, 2010; Hardy, 2001, 2011; Masanes and Müller, 2011], and which encompasses all operational theories – including classical probability

theory and quantum theory as special cases. The no-go theorem which we present here was previously derived in the context of gravity [Zych *et al.*, 2019]. Our derivation uses an assumption about the initial state of the systems which is weaker than that in Ref. [Zych *et al.*, 2019] (we consider Bell-local states rather than separable states, which are a subset of Bell-local states), and a different notion of locality. (The relation between the assumptions and implications of the current work and those of Ref. [Zych *et al.*, 2019] are analyzed in Methods - Sec. 2.5.1-2.5.2.)

We first define what we mean by a causal order in a GPT. Consider a system in the state $\omega \in \Omega$ of a GPT state space Ω and imagine two parties, Alice and Bob, who perform some operations on this state. For example, suppose that the operation in Alice's laboratory is given by a transformation $\mathcal{A}$ and that in Bob's laboratory is given by a transformation $\mathcal{B}$. Alice's and Bob's operations are said to undergo a process that is 'causally separable' in GPTs whenever Alice's operation happens before or simultaneously to Bob's ($\mathcal{A} \preceq \mathcal{B}$), Bob's operation happens before or simultaneously to Alice's ($\mathcal{B} \preceq \mathcal{A}$), or there is a convex mixture of these two cases:

$$\mathcal{S}(\omega) = \zeta \cdot \mathcal{B}\big(\mathcal{A}(\omega)\big) + (1-\zeta) \cdot \mathcal{A}\big(\mathcal{B}(\omega)\big), \tag{2.1}$$

where $0 \leqslant \zeta \leqslant 1$ is the probability with which one or the other order is chosen and $\mathcal{Y}\big(\mathcal{X}(\cdot)\big)$ is a composition of operations $\mathcal{X}$ and $\mathcal{Y}$. (While in the current work we limit our analysis to the case of only $N = 2$ parties, an analogue relation can be established for $N > 2$ parties, giving rise to a classical mixture of all possible permutations among the N parties, or to a dynamical causal order, where the causal order between operations may depend on operations performed beforehand [Abbott *et al.*, 2017].) If a process cannot be written in the form of Eq. (2.1), it is called a 'causally non-separable process'.

Within the GPT framework, we now consider ω to be a state of the following composite system: one system (the *control* system) governing the order in which the operations $\mathcal{A}$ and $\mathcal{B}$ are applied, and another system (the *target* system) on which the operations are performed. We will further consider that there are two parties, S1 and S2, each possessing one such composite system. No restrictions are applied to the state of the control system (thus, for instance, the composite control state may violate a Bell inequality).

In the Methods - Sec. 2.5.1 we prove a *no-go theorem*, stating that any two-party system obeying the following three assumptions cannot violate a Bell inequality (below we briefly summarize our theorem, saving the detailed version for Methods - Sec. 2.5.1).

I) The initial joint state of the two target subsystems is *local* (*i.e.*, it does not violate a Bell inequality).

II) The laboratory operations are *local transformations* of the target subsystems (*i.e.*, they do not increase the amount of a violation of Bell inequalities between the two target subsystems neither

through the interaction between them, nor through the interaction between individual control and target subsystems).

III) The order of local transformations on the two target subsystems is well-defined.

We will briefly comment on assumption II here, and refer to Sec. 2.3 for an in-depth analysis of all three assumptions. Assumption II has two implications: IIa) The laboratory operations cannot increase the non-local correlations between the target subsystems of parties S1 and S2 (*i.e.*, there is no non-local interaction between the two subsystems). IIb) Within a single party Si, $i = 1, 2$, the laboratory operations do not 'couple' the control and target system. Such a 'coupling' could transfer existing non-local correlations between the pair of controls to the pair of targets, thereby enabling a violation of Bell inequalities.

In the next section, we will present a quantum mechanical process that violates this no-go theorem. Hence, at least one of the assumptions must not hold for this process. In Sec. 2.3, we will analyse our experimental data testing a Bell-like inequality to provide evidence in support of assumption I within the framework of GPTs. Consequently, either assumption II does not hold, assumption III does not hold, or both assumptions are invalid. On the basis of the data collected for the quantum switch of system S1 (or S2) individually, we will show that it is not possible to describe our results by violating only assumption II. Thus, the only viable conclusion is that the order of operations applied on each system Si is indefinite (*i.e.*, that assumption III is necessarily false).

2.2.1 Entangled quantum switch

To understand a single quantum switch, first imagine two parties, Alice and Bob, who are in two *closed laboratories*, *i.e.*, their only interaction with the external environment is through input and output systems. Suppose that each of the parties performs an operation on the same qubit (a 'target' qubit), and that this qubit may be sent first to Alice and then to Bob, or vice versa. Now, in a quantum switch, one governs the order of the operations on the target qubit according to the state of a second quantum system, a 'control' qubit. If the control qubit is placed in a superposition, this establishes a quantum-superposition of the order of the two operations. For instance, if the control qubit is in the state $|0\rangle^c$, the target qubit is sent first to Alice and then to Bob, and vice versa if the control qubit is in the state $|1\rangle^c$. When the control qubit is prepared in the state $(|0\rangle^c + |1\rangle^c)/\sqrt{2}$, the resulting process has been shown to be causally non-separable within quantum mechanics [Araújo *et al.*, 2015; Chiribella, 2012; Chiribella *et al.*, 2013; Rubino *et al.*, 2017a].

Next, consider two quantum switches (S1 and S2), each containing an Alice and a Bob. S1 and S2 are prepared in a state where their control qubits are entangled, but their target qubits are in a product

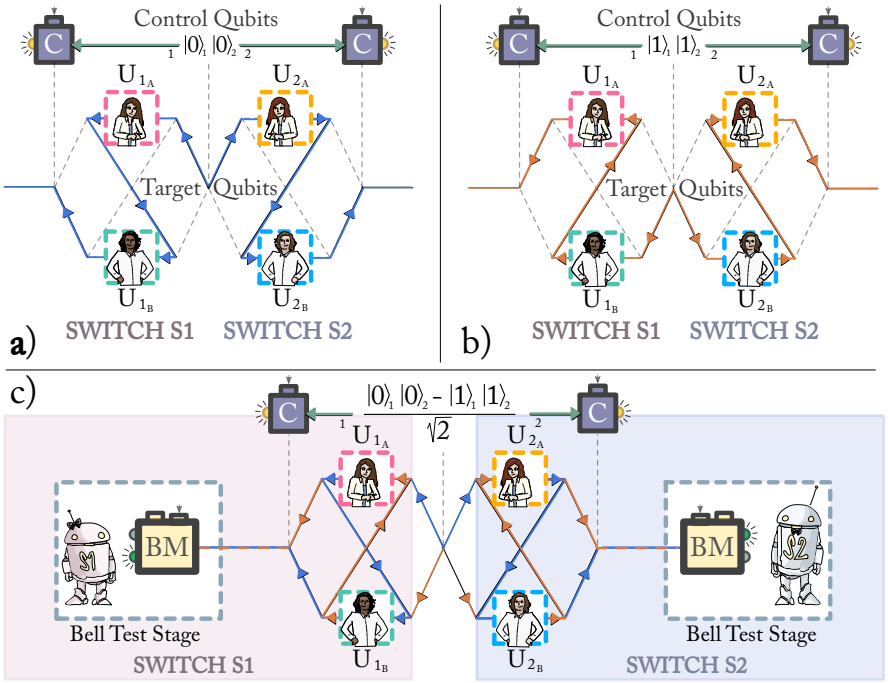


Figure 2.1 **Entangled quantum switch**. Our work is based on two quantum switches (S1 and S2). In each quantum switch, there are two parties, Alice (U_{i_A}) and Bob (U_{i_B}). A target qubit is first sent to one party, and then to the other. The order in which the qubit is sent to the two parties is governed by the state of an additional qubit: if the state of the control qubit is $|0\rangle_i^c$, the target qubit is sent first to Alice and then Bob (Panel **a**), and vice versa if the control qubit is in the state $|1\rangle_i^c$ (Panel **b**). In our work, we entangle the control qubits (Panel **c**). In this case, the order in which the target qubit in quantum switch S1 passes through U_{1_A} and U_{1_B} is entangled with the order in which the target qubit in quantum switch S2 passes through U_{2_A} and U_{2_B}. The control qubits are measured in the basis $\{|+\rangle_i^c, |-\rangle_i^c\}$. If the orders inside the two quantum switches are entangled, it will be possible to violate a Bell inequality by measuring the target qubits after the quantum switches (BM). This is possible even if the target qubits start in a separable state and only local operations are applied within each quantum switch.

state (see Fig. 2.1):

$$|0\rangle_1^t \otimes |0\rangle_2^t \otimes \left(\frac{|0\rangle_1^c \otimes |0\rangle_2^c - |1\rangle_1^c \otimes |1\rangle_2^c}{\sqrt{2}}\right) \tag{2.2}$$

The superscripts c and t refer to the control and target qubits within one quantum switch, respectively, while the subscripts 1 and 2 refer to quantum switch S1 and S2. Since we will attempt to observe a Bell violation with the target qubits, which are in a separable state, this initial condition satisfies assumption I in quantum theory.

Given this input state and the action of an individual quantum switch, it is straightforward to calculate the output of the entangled quantum switch system

$$\frac{1}{\sqrt{2}}\left(U_{1_B}U_{1_A}|0\rangle_1^t\right) \otimes |0\rangle_1^c \otimes \left(U_{2_B}U_{2_A}|0\rangle_2^t\right) \otimes |0\rangle_2^c - \frac{1}{\sqrt{2}}\left(U_{1_A}U_{1_B}|0\rangle_1^t\right) \otimes |1\rangle_1^c \otimes \left(U_{2_A}U_{2_B}|0\rangle_2^t\right) \otimes |1\rangle_2^c \tag{2.3}$$

where U_{i_A} and U_{i_B} $(i = 1, 2)$ are the unitaries performed by the two parties Alice and Bob inside each

quantum switch Si.

Next, we measure the two control qubits in the basis $\{|+\rangle\langle+|, |-\rangle\langle-|\}$. If we observe both of the control qubits in the same state (either $|+\rangle_1^c|+\rangle_2^c$ or $|-\rangle_1^c|-\rangle_2^c$), the target qubits will be in the (in general) unnormalised state

$$\frac{1}{\sqrt{2}}(U_{1_B}U_{1_A}|0\rangle_1^t \otimes U_{2_B}U_{2_A}|0\rangle_2^t - U_{1_A}U_{1_B}|0\rangle_1^t \otimes U_{2_A}U_{2_B}|0\rangle_2^t), \tag{2.4}$$

while, if we find the control qubits in orthogonal states (either $|+\rangle_1^c|-\rangle_2^c$ or $|-\rangle_1^c|+\rangle_2^c$), the sign between the two terms in the superposition in the equation above is '+'. In general, depending on the choice of the unitaries in the two quantum switches, the target qubits will be left either in a separable or in an entangled state. In particular, if we choose the gates

$$U_{1_A} = U_{2_A} = \sigma_z \qquad\qquad U_{1_B} = U_{2_B} = \frac{\mathbb{1} + i\sigma_x}{\sqrt{2}}, \tag{2.5}$$

where σ_x ans σ_z are the Pauli operators, the state of the target qubits becomes

$$\frac{1}{\sqrt{2}}(|l\rangle_1^t|l\rangle_2^t - |r\rangle_1^t|r\rangle_2^t), \tag{2.6}$$

where $|r\rangle = (|0\rangle - i|1\rangle)/\sqrt{2}$ and $|l\rangle = (|0\rangle + i|1\rangle)/\sqrt{2}$. This is a *maximally entangled state* and, as a result, one can now violate a Bell inequality on the target qubits.

Within quantum theory, the entanglement between the targets and the resulting violation of the Bell inequality can be explained in terms of the indefiniteness of the temporal orders in the two quantum switches. In other words, such entanglement is not 'generated', but rather 'transferred' from the control qubits by means of the indefinite temporal order of the unitaries applied. A related interpretation of the violation in quantum mechanics is in terms of *time-delocalized quantum operations* [Oreshkov, 2019] and *causal reference frames* [Allard Guérin and Brukner, 2018], according to which a frame can be chosen such that while Bob's operation acts at a fixed time, Alice's operation is in a superposition of being implemented before and after Bob's operation, thus resulting in an indefinite causal order between them.

In the class of GPTs considered here, the presence of non-classical correlations can be determined through a violation of a Bell inequality. In our case, the violation of a Bell inequality with the target subsystems implies the violation of the no-go theorem for temporal orders, thereby proving that no underlying GPTs where assumptions I, II and III hold can explain the experimental data. We will experimentally confirm that I holds both in quantum mechanics, and in our class of GPTs (as detailed in Methods - Sec. 2.5.3). Then, we will show, both within quantum mechanics and in our class of GPTs, that one cannot describe our results if only assumption II is invalid. We will thus conclude that

either assumption III is wrong or both assumptions II and III are false, hence proving the presence of indefinite causal orders beyond the quantum framework.

2.2.2 Experimental scheme

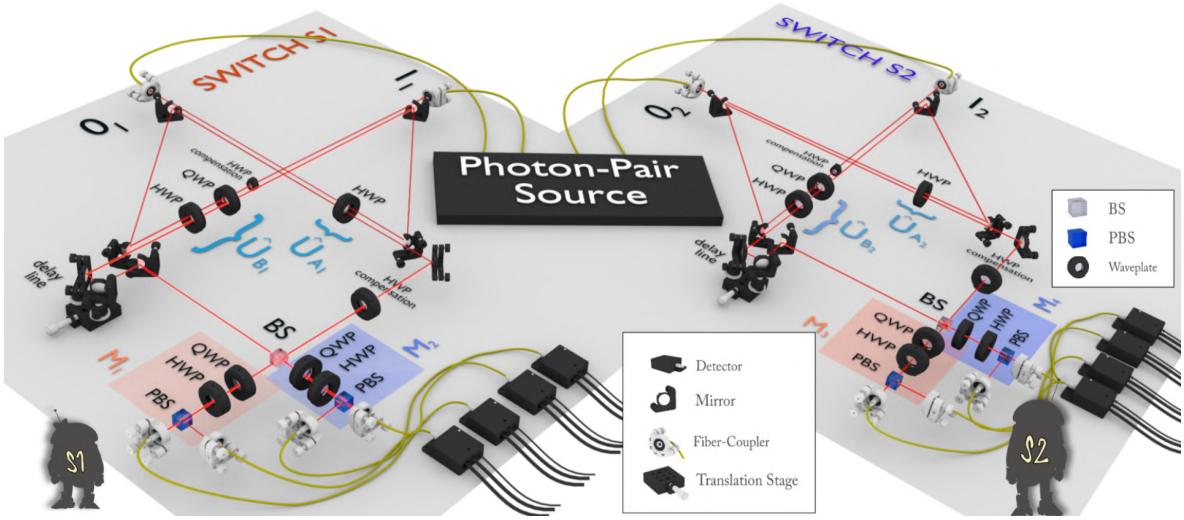


Figure 2.2 Experimental implementation of an entangled quantum switch. Each quantum switch is composed of a two-loop Mach-Zehnder interferometer. The interferometers start in the photon-pair source, wherein photon 1 and photon 2 are placed in superposition of the paths 0_1 and 1_1, and 0_2 and 1_2, respectively (see the Methods - Sec. 2.5.4). (For simplicity, we have drawn these paths as fibers, however the photons are transmitted via free-space from the source to the experiment.) These paths are routed such that path 0_i sees gate U_{i_A} and then gate U_{i_B}, and vice versa for the path 1_i. Each gate, acting on the polarization degree of freedom, is made up of waveplates (as described in the main text). The paths 0_i and 1_i are then combined on a beam splitter (BS). In quantum switch S1 (S2), the photon is detected after the polarization measurement at M_1 or M_2 (M_3 or M_4). Together with the BS (which applies a Hadamard gate to the qubit encoded in the path DOF), detecting the photon at M_1 or M_2 (M_3 or M_4) projects the path qubit on $|+\rangle$ or $|-\rangle$, respectively. Furthermore, within each measurement M_i, the polarization qubit can be measured in any basis by a combination of a quarter-waveplate (QWP), half-waveplate (HWP), and polarizing beam splitter (PBS).

We create a quantum switch with entangled control qubits using a photonic set-up. Let us first consider a single quantum switch. Each quantum switch applies gates on a target qubit, where the gates' order depends on the state of a control qubit. Experimentally, we encode the control qubit in a path degree of freedom (DOF), and the target qubit in the polarization DOF of a single photon. The photon is initially placed in a superposition of two paths (as explained in Fig. 2.2 and the Methods - Sec. 2.5.4). These paths are labeled 0_1 and 1_1 for quantum switch S1 and 0_2 and 1_2 for quantum switch S2 in Fig. 2.2. The two paths are then routed through a two-loop *Mach-Zehnder interferometer* [Procopio *et al.*, 2015; Rubino *et al.*, 2017a]. The 0_i paths lead the photons through a set of gates acting on the polarization DOF in the order $U_{i_\mathrm{A}} \preceq U_{i_\mathrm{B}}$. While the paths 1_i guide the photons through the gates in the opposite order $U_{i_\mathrm{B}} \preceq U_{i_\mathrm{A}}$. To generate the maximally entangled state between the target qubits in Eq. (2.6), we need to implement the non-commuting gates $U_{i_\mathrm{A}} = \sigma_z$ and $U_{i_\mathrm{B}} = (\mathbb{1} + i\sigma_x)/\sqrt{2}$, which we do with waveplates. In particular, a half-waveplate (HWP) at 0° for σ_z and a sequence of quarter-waveplate (QWP) and HWP both at 45° for $(\mathbb{1} + i\sigma_x)/\sqrt{2}$). After this, the two paths are recombined

on a 50/50 beamsplitter (BS) — which projects the path DOF in the basis $\{|+\rangle\langle+|, |-\rangle\langle-|\}$. The path lengths and the relative phases are set by means of a piezo-driven trombone-arm delay line. At the two outputs of each interferometer, QWPs, HWPs and polarizing beam splitters (PBSs) are used to perform arbitrary polarization measurements on the target qubits.

To entangle the two quantum switches, we first entangle the path DOFs of the two photons. As explained in the Methods - Sec. 2.5.4, we generate path-entangled photon pairs that are separable in their polarization DOF:

$$|\Phi^-\rangle_{1,2}^{\text{path}} \otimes (|H\rangle_1 |H\rangle_2)^{\text{polar.}} = \left(\frac{|0\rangle_1 |0\rangle_2 - |1\rangle_1 |1\rangle_2}{\sqrt{2}}\right)^{\text{path}} \otimes (|H\rangle_1 |H\rangle_2)^{\text{polar.}}. \tag{2.7}$$

Each photon is thus delocalized over two paths. The two photons are then sent to their respective quantum switches, and, since the control qubits began in an entangled state, the order in which the gates act on the two target qubits becomes entangled.

2.3 Results

Our goal is to demonstrate that the order of application of the gates within the two quantum switches is genuinely indefinite without assuming that the laboratory operations and the states of the systems are described by quantum theory. We can arrive at this conclusion in three steps. We will first show experimental data that violate a Bell inequality. From this we can assert that *at least one* of the three assumptions must be false. We will then prove that assumption I is satisfied in our experiment using both quantum theory and a class of GPTs. Thus, one of the remaining assumptions (*i.e.*, assumptions II and/or III) must not hold. We will analyse the case in which only assumption II does not hold both in quantum mechanics and within the set of GPTs. By acquiring additional measurements on a single quantum switch, we will show that such scenario cannot reproduce the results of our experiment. Consequently, the only two possible explanations are that either assumption III does not hold, or that both assumptions II and III are false. In either case, assumption III must be false, and hence the local operations within the two quantum switches have been applied in an indefinite temporal order.

We begin by performing a Bell test between the target states at the output of the apparatus. This allows us to experimentally probe a conjunction of all three assumptions.

We first perform polarization-state tomography on the two-qubit output target state after the quantum switches (which, of course, requires quantum mechanics), using four equivalent measurement set-ups (orange and blue boxes in Fig. 2.2). Since the 50/50 BSs apply a Hadamard gate on the path qubits, we post-select the control qubits in the same state (either $|+\rangle_1^c |+\rangle_2^c$ or $|-\rangle_1^c |-\rangle_2^c$) by grouping the results of M_1 with M_3 (orange boxes) and M_2 with M_4 (blue boxes). The resulting density matrix is presented in Fig. 2.4, and it shows a clear presence of entanglement. The reconstructed state has

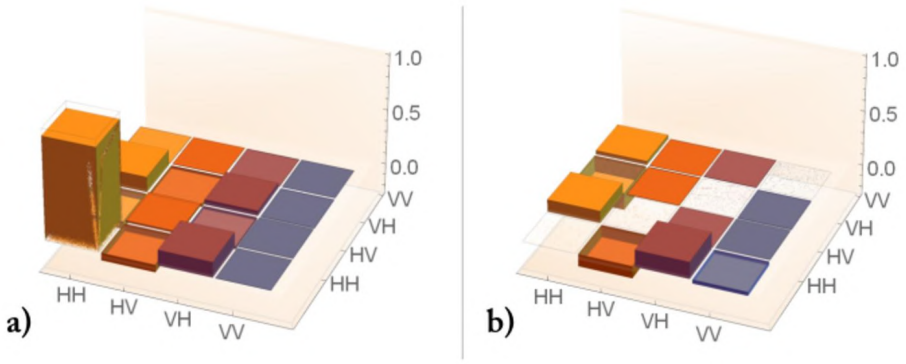


Figure 2.3 Input control state characterization: State tomography of the target qubits. The real (Panel **a**) and imaginary (Panel **b**) parts of the two-photon polarization state are measured before the two photons enter the quantum switches. This state has a fidelity 0.935 ± 0.004 with the ideal state $|HH\rangle$, and a concurrence of 0.001 ± 0.010.

a fidelity of 0.922 ± 0.005 with the ideal one [Eq. (2.6)], and a concurrence of 0.95 ± 0.01. Finally, to perform a theory-independent measurement, we perform a Bell test (more specifically, we measure a Clauser-Horne-Shimony-Holt (CHSH) inequality [Clauser *et al.*, 1969]) on the polarization DOF, obtaining $S_{\text{target}} = 2.55 \pm 0.08$. This violates the inequality, and thus also the no-go theorem, by almost 7 standard deviations. Hence, in our class of GPTs, no theory satisfying assumptions I, II and III is compatible with the experimental data.

We now proceed to test the validity of assumption I, which says that the joint target state (shared between system S1 and S2) does not initially violate a Bell inequality. Within quantum theory, one can show this by demonstrating that the state is separable; this can be done using quantum state tomography, for example. To this end, we performed tomography on the target states before the quantum switches. The resulting density matrix is shown in Fig. 2.3, Panels **a** and **b**. For our experiment, the target state was nominally prepared in $|HH\rangle$; our measured state has a fidelity of 0.935 ± 0.004 with $|HH\rangle$. Furthermore, the concurrence of the estimated state is 0.001 ± 0.010, indicating that, within experimental error, the initial target state is separable, in agreement with assumption I. The error bars are computed using a Monte Carlo simulation of our experiment; the dominant contribution comes from errors in setting the WPs, and cross-talk in the polarizing BSs.

We will now consider assumption I within a class of GPTs. We assume that the set of 'fiducial measurements' of the class of GPTs contains 'quantum fiducial measurements' as a subset. In this sense, quantum theory is embeddable in the GPT. This is similar to how classical theory can be embedded in quantum theory (*i.e.*, classical theory has one fiducial measurement in the 'computational basis'). In particular, we consider a class of GPTs wherein the state space of a single two-level system is described by a d-dimensional Bloch ball [Dakic and Brukner, 2010; Masanes *et al.*, 2014] (*i.e.*, there are three quantum fiducial measurements, and $d-3$ non-quantum fiducial measurements), with $d > 3$ in general. For this class of theories, it was shown that a single system is in a pure state if there exists

a measurement for which the system returns a given result with probability one. Similarly, a bipartite system is in a pure product state, if the above statement applies to each individual system. In more detail, in GPTs a state is pure if it cannot be written as a non-trivial mixture of other states. Moreover, a bipartite system is in a product state if, for all local measurements, the probabilities for outcome pairs on a bipartite-state are equal to the product of the two marginal probabilities of each subsystem. Such a state has perfect correlations only for a pair of fiducial measurements, it exhibits no further correlations in any other pair of fiducial measurements, and it cannot violate a Bell inequality [Dakic and Brukner, 2010; Hardy, 2001, 2011; Masanes and Müller, 2011].

For our target photon pair we demonstrated that both photons return value H with certainty. This means that, already from a pair of quantum fiducial measurements, one can conclude that, up to experimental imperfections, the state is a pure product state, and hence it cannot violate a Bell inequality. This supports the validity of assumption I in the special case of Bloch-vector theories (see Methods - Sec. 2.5.1). In Table S1, we compare the probabilities for outcome pairs on a bipartite-state to the product of the two marginal probabilities of each subsystem. The excellent agreement between the two probability distributions indicates that the joint target state is indeed a pure product state, and cannot violate a Bell inequality. This proves that assumption I of our no-go theorem holds for the class of GPTs under consideration. We quantify to what extent the two distributions agree by calculating the *root-mean-square* (RMS) difference between the two distributions, resulting in an average difference of $0.6 \cdot 10^{-2} \pm 2.7 \cdot 10^{-2}$. Although this value is consistent with zero, one could imagine that this small difference is in fact caused by correlations between the two target systems. In Methods - Sec. 2.5.3 we show, however, that such small correlations can only give rise to a vanishingly small violation of Bell's inequality. The possible level of violation from this amount of potential coupling is insufficient to explain our experimentally observed violation. Therefore, we have confirmed that the joint target system starts in an (approximately) separable state. Additionally, in Methods - Sec. 2.5.4, we experimentally show that the joint control system is initially entangled. We then send this joint state into our two quantum switches and perform measurements on the output state.

Having proven, both in quantum theory and within the class of GPTs, that our no-go theorem is violated and that assumption I is justified, we can conclude that either assumption II, or assumption III, or both must be false. We will now consider the case in which only assumption II is false.

The second assumption of our no-go theorem says that the laboratory operations are local transformations acting on the target states. As discussed earlier, this has two implications. First, it implies that the laboratory operations performed in the two quantum switches cannot transform the joint state of the target systems of S1 and S2 from a local state to a non-local one (IIa). This could be ensured by performing the operations with a space-like separation in which case the condition would be guaranteed in any theory obeying relativistic locality. However, in our experiment, we make the (well-justified) device-dependent assumption that the laboratory operations are local transformations

within S1 and S2 in GPTs, since the transformations of the systems take place at spatially separated parts of the optical table. As a consequence, the first implication of assumption II arguably holds.

Let us consider now the second, more substantial, implication. Since we are analysing the case in which only assumption II is false while assumption III holds, we will now study the scenario in which the laboratory operations occur only in a causally-ordered manner.

The second implication of assumption II states that the laboratory operations transform product states of the control and the target subsystems into product states within each quantum switch (IIb). This means that the laboratory operations of one party Si do not 'couple' the control and the target systems. Such a coupling would make it possible to transfer non-local correlations from the control systems of the two parties S1 and S2 to their target systems, and therefore a violation of Bell's inequalities would be possible.

We can experimentally prove that our experiment satisfies assumption IIb in a class of GPTs using a similar technique as for assumption I. We start by placing bounds on the degree of coupling between the target and the control qubits within a single quantum switch in the presence of *only one* of the two operations (*i.e.*, only U_A or only U_B) inside the quantum switch. We perform the full set of quantum fiducial measurements. With this, we show that the joint probabilities of the target and the control subsystems are factorizable into the products of the two marginal probabilities of each subsystem in the case where either only U_A or U_B is inserted inside a single quantum switch. From this we can conclude that the joint probabilities must also remain factorizable after U_A and U_B are applied in a well-defined order (or in a classical mixture of the two orders). In other words, in our GPTs, the optical elements do not couple the control and the target subsystem in the 'quantum subspace' of the GPT state space. Moreover, the marginal probabilities measured on the control and the target subsystem correspond to a pure (product) state. From this we also conclude that there can be no coupling in the 'non-quantum subspace'. We analyse this by performing a set of measurements on the joint control-target system, and by showing that the joint probabilities can be described by the product of the marginal probabilities (see Tables S2-S3, and Methods - Sec. 2.5.5 for more details). The RMS difference between the two distributions is, on average, 0.02 ± 0.03. This value is within one standard deviation of zero, confirming that the probability distribution is consistent with that of a product state. As we discussed for assumption 1, this small discrepancy could be caused by correlations between the control and the target systems. However, as we show in Methods - Sec. 2.5.5, these correlations are too weak to explain our experimentally observed violation of Bell's inequality.

Hence, the laboratory operations do not couple the control and target systems when they are applied in a well-defined order. This means that, for our experiment, whenever assumption III holds, assumption II must also hold. Because the statement $a \Rightarrow b$ is logically equivalent to *not* $b \Rightarrow$ *not* a, this further implies that if assumption II is invalid, then assumption III must also be invalid. It follows that the only two possible scenarios are that (1) assumption II is wrong, and therefore III is also wrong, or (2)

assumption III is false, independently of assumption II. In either case, it is not possible to explain our experimental data unless assumption III is discarded, and we can conclude that the local operations in our experiment were applied in an indefinite order.

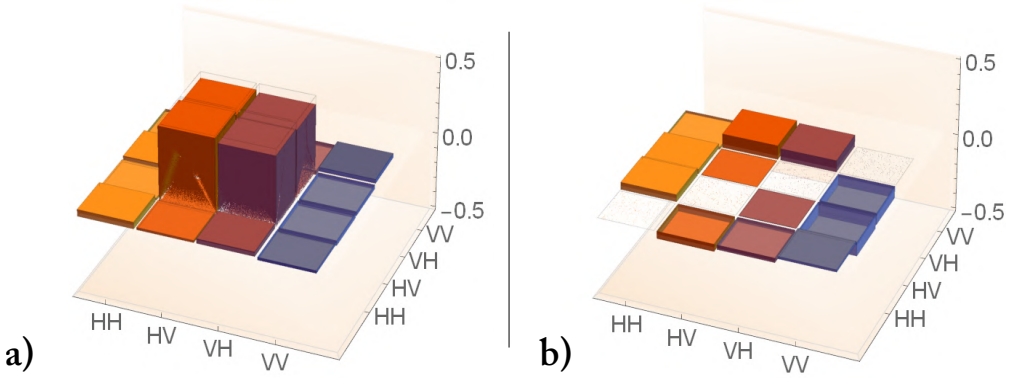


Figure 2.4 **Output state characterization.** Panels **a** and **b** show the real and imaginary parts, respectively, of the two-photon polarization state measured after the photons leave the quantum switches. For the data shown here, the two control qubits were found to be in the same state (either $|+\rangle_1^c |+\rangle_2^c$ or $|-\rangle_1^c |-\rangle_2^c$). This state has a fidelity of 0.922 ± 0.005 with the target state $(|HV\rangle + |VH\rangle)/\sqrt{2}$, and a concurrence of 0.95 ± 0.01. Performing a Bell measurement directly using this state results in a CHSH parameter of 2.55 ± 0.08.

To summarize, in our work we engineered a situation wherein the only way entanglement can be transferred from one pair of systems to another is by means of causally non-separable processes. In our experiment, this transfer takes place between different DOFs of photon pairs. Although it is often easy to transfer the entanglement from one DOF to another, this is typically done with a device that directly couples the two DOFs; *e.g.*, in the case of path-polarization transfer, a PBS could be used. In our experiment, we used an entangled quantum switch to accomplish this interchange. Our quantum switches do not contain any device which directly couples these DOFs (only waveplates, which act solely on the polarization state, and 50/50 BSs, which act solely on the path state). Rather, here the interchange occurs because the control qubit (the path) governs the order of the application of gates on the target qubit (the polarization). Then, since we begin with an entangled state of the control qubits, this state is transferred to the target qubits via an indefinite order of the application of the gates. In other words, by choosing a specific set of operations, the temporal superposition of the application of these operations is mapped onto a superposition of orthogonal states. As a result, this transfer of entanglement is the signature of an indefinite temporal order.

2.4 Discussion

In this work, we entangled the temporal orders between operations applied by two parties and experimentally showed that the resulting temporal order is indefinite, by violating a Bell inequality using the joint target system after the quantum switches. We thus verified that the data collected by entangling

temporal orders in the quantum switches cannot be described by a class of (generalized probabilistic) theories under the assumption that the initial joint target state does not violate Bell's inequalities, the operations on the target states are local, and they have a pre-defined order. This did not require the assumption that the systems and the operations are described by the quantum formalism. Clearly, for our demonstration to be loophole-free (as proposed in Ref. [Zych *et al.*, 2019]), the standard Bell loopholes (fair-sampling and locality) would need to be closed. Further loopholes can arise related to the implementation of the quantum switch. In fact, it is known that experimental data produced by the quantum switch can be simulated by a causally-separable process if at least one of the operations (either A or B) is performed two or more times. In relation to the experimental implementations of the quantum switch, this is the so-called 'single-usage loophole'. Closing this loophole would require an operational verification that each operation in the quantum switch is performed only once. For example, this could include implementation of a "counter" that would estimate the number of times an operation is performed, or a process tomography on time-delocalized quantum systems [Oreshkov, 2019]. However, our experiment is immune to the single-usage loophole as even a multiple usage of local operations on either side of the Bell test cannot result in a violation of Bell's inequality provided that the operations are performed in a definite causal order.

All previous studies involving quantum processes with indefinite temporal orders achieved their goal by superposing the order of operations, rather than entangling them. The first proposal to entangle the temporal orders was made only recently [Zych *et al.*, 2019]. Here we show that the basis of this theoretical concept is in fact experimentally accessible. Moreover, we exploit this resource as a new means to validate indefinite causal structures. Techniques to characterize these structures are becoming increasingly relevant, as it is known that these processes can lead to linear advantages in query complexity, and exponential advantages in quantum communication tasks [Araújo *et al.*, 2014; Chiribella *et al.*, 2013; Feix *et al.*, 2015; Guérin *et al.*, 2016].

2.5 Methods

2.5.1 Proof of No-Go Theorem for Temporal Orders

All previous experimental studies of causally non-separable processes [Procopio *et al.*, 2015; Rubino *et al.*, 2017a] were dependent on the validity of the quantum theory (*i.e.*, they were *theory-dependent*), and all known physically realizable processes satisfy all causal inequalities (see the Suppl. Information) [Araújo *et al.*, 2015; Oreshkov and Giarmatzi, 2016]. The latter means that experimental data taken from a given causally non-separable quantum process could still be understood as arising in causal manner, for example in an underlying generalized probabilistic theory (GPT). Hence, it is unknown whether a fully theory-independent experimental proof of indefinite causal order is possible.

In our current work, we relate a violation of a Bell inequality to the violation of a no-go theorem for temporal orders, as proposed in Ref. [Zych *et al.*, 2019]. This results in a proof of causal indefiniteness outside of the quantum framework as it is valid for a large class of generalized probabilistic theories. In this section, we provide a rigorous introduction to such no-go theorem for temporal orders.

We will begin by giving a brief introduction to the basic elements of GPTs which are necessary for our no-go theorem. A more detailed discussion of the GPT framework can be found in Ref. [Barrett, 2007; Hardy, 2001; Masanes and Müller, 2011].

In a GPT, a system is described by a state ω that specifies outcome probabilities for all measurements that can be performed on it. A complete representation of the state is given by specifying the outcome probabilities of a so-called 'fiducial set'. The smallest such set defines the number d of degrees of freedom of the system. We restrict our consideration here to binary systems that have two perfectly distinguishable states and no more. For example, the fiducial set for a two-level system in quantum theory consists of the (three) probability outcomes of spin projections along x, y and z. The state space is a compact and convex set Ω embedded in a vector space. The extremal states of Ω that cannot be decomposed as a convex mixture of other states are called 'pure states'. An effect e is defined as a linear functional on Ω that maps each state onto a probability, *i.e.*, $e : \Omega \to [0, 1]$, where $e(\omega)$ is the probability to obtain an outcome on the state ω. The linearity is required to preserve the convex structure of the state space.

A transformation U is a linear map from a state to a state, *i.e.*, $U : \Omega \to \Omega$. The transformation is linear for the same reason that probabilities have to be linear maps of states. The sequence of transformations U_1, ... , U_n, in which transformation U_1 'precedes' transformation U_2, which 'precedes' U_3, *etc.*, is represented by a composition of maps: $U_n \circ ... \circ U_1$. This defines a *definite order of transformations*, which we denote as $U_1 \preceq ... \preceq U_n$.

We will now introduce a generalization of the no-go theorem for temporal orders, which was originally proposed in Ref. [Zych *et al.*, 2019].

In the framework of a GPT, the state of a composite system shared between two parties S1 and S2 is given by $\omega_{1,2} \in \Omega_{1,2}$, where $\Omega_{1,2}$ is the state space of a composite system. The state of a composite system is given by a multiplet consisting of the *local states* $\omega_1 \in \Omega_1$ and $\omega_2 \in \Omega_2$ of individual systems, the *correlation tensor* $\hat{T}$ and a potential *global parameter* ξ [Dakic and Brukner, 2010; Hardy, 2001, 2011; Masanes and Müller, 2011]:

$$\omega_{1,2} = \omega_{1,2}(\omega_1, \omega_2, \hat{T}, \xi) \tag{2.8}$$

The fact that subsystems are themselves systems implies that each has a well-defined reduced state ω_1, ω_2 which does not depend on which transformations and measurements are performed on the other subsystem; this is often referred to as 'no-signaling'. We also assume that transformations and

measurements performed on subsystems commute with each other, so that *one* correlation tensor is enough to describe correlations between them. If this were not the case, we would need to introduce *two* correlation tensors, one when S1 applies operations before S2, and the other when S2 performs operations before S1. Finally, the states in GPT need not to satisfy the *local tomography* condition (stating that reduced states and correlation tensor completely describe the systems' state), but may include a global parameter ξ.

For the present case of binary systems, the components of the state in Eq. (2.8) are given by

$$\omega_1^{(i)} = p^{(i)}(o_1 = 1) - p^{(i)}(o_1 = -1), \tag{2.9a}$$

$$\omega_2^{(j)} = p^{(j)}(o_2 = 1) - p^{(j)}(o_2 = -1), \tag{2.9b}$$

$$T^{(i,j)} = p^{(i,j)}(o_1\, o_2 = 1) - p^{(i,j)}(o_1\, o_2 = -1), \tag{2.9c}$$

where $i, j = 1, ..., d$. Here, for example, $p^{(i)}(o_1 = 1)$ is the probability to obtain the outcome $o_1 = 1$ when the i-th measurement is performed on the first subsystem, and $p^{(i,j)}(o_1\, o_2 = 1)$ is the joint probability to obtain correlated results (*i.e.*, either $o_1 = o_2 = +1$ or $o_1 = o_2 = -1$) when the i-th measurement is performed on the first subsystem and the j-th measurement on the second one.

An effect e_{12} that maps a state onto a probability for a pair of *local* measurements is given by $e_{12} = e_{12}(r_1, r_2, r_1 r_2^T)$, where r_i is the effect on the state of *i*-th system, and r^T denotes transposition of r. (Note that the global parameter does not contribute to the probability for a pair of local measurements). The probability to obtain the effect e_{12} when the system is prepared in the state ω_{12} is given by

$$p(e_{12}|\omega_{12}) = \frac{1}{4}\left(1 + (\omega_1 \cdot r_1) + (\omega_2 \cdot r_2) + (r_2 \cdot \hat{T} r_1)\right), \tag{2.10}$$

where $(x \cdot y)$ is the Euclidean scalar product between two d-dimensional real vectors x and y.

The product state is represented by $\omega_p = \omega_p(\eta_1, \eta_2, \eta_1\eta_2^T, \xi_p)$, where the correlation tensor is of a product form. If we perform a pair of local measurements on the arbitrary product state, the outcome probability factorizes into the product of the local outcome probabilities.

We next introduce a pair of *local (reversible) transformations* $(U_1, U_2) : \Omega_{12} \to \Omega_{12}$ as a linear map from the space of states of a composite system to itself:

$$(U_1, U_2)(\omega_{12}) = (U_1\omega_1, U_2\omega_2, U_1\hat{T}U_2^T, \xi'), \tag{2.11}$$

where the global parameter ξ' is, in general, changed under the transformations (U_1, U_2). Since testing our Bell inequality involves only local transformations and measurements, it is sufficient to specify effects for those measurements.

In our experiment, ω_1 and ω_2 themselves are states of composite systems each consisting of a 'control' and a 'target' subsystem. Hence, the entire system under investigation consists of four subsystems, a

control and a target subsystems of S1 and a control and a target subsystems of S2. The overall state is

$$\omega_{1,2,3,4} = \omega_{1,2,3,4}(\omega_1^t, \omega_1^c, \omega_2^t, \omega_2^c, ..., \hat{T}^{ij}, ..., \hat{T}^{ijk}, ..., \hat{T}^{1234}, \Xi) \tag{2.12}$$

where c and t refer to the terms 'control' and 'target' subsystems, $\hat{T}^{ij}$, $\hat{T}^{ijk}$ and $\hat{T}^{1234}$ are correlation (sub)tensors describing correlations between pairs $\{i, j\}$, triple $\{i, j, k\}$ and quadruple $\{1, 2, 3, 4\}$ of subsystems, respectively, and Ξ is the set of all global parameters.

The no-go theorem concerns the reduced state of the two target systems as given by

$$\omega_{1,2}^t = \omega_{1,2}^t(\omega_1^t, \omega_2^t, \hat{T}^{tt}, \xi^t), \tag{2.13}$$

where ω_1^t and ω_2^t are states of the target subsystems of S1 and S2, $\hat{T}^{tt}$ is their correlation tensor, and ξ^t is the corresponding global parameter.

Leveraging these definitions, we now present three assumptions, which are the fulcrum of our no-go theorem for a definite local causal order.

1. The initial joint state of the target system $\omega_{1,2}^t$ does not violate a Bell inequality.

Suppose that the two observers can each perform a measurement $\mathcal{O}_1$ and $\mathcal{O}_2$, respectively. We label m_1 and m_2 as arbitrary measurement choices of S1 and S2, and o_1 and o_2 as the corresponding outcomes. Under these conditions, we suppose that our input state $\omega_{1,2}^t$ can be described through a local hidden variables theory (*i.e.*, in Bell's terms, a theory that satisfy 'local causality'), and therefore it is associated to the probability distribution

$$p(o_1, o_2|m_1, m_2, \omega_{1,2}^t) = \int \rho(\lambda)\ p(o_1|m_1, \lambda, \omega_{1,2}^t)\ p(o_2|m_2, \lambda, \omega_{1,2}^t)\ d\lambda, \tag{2.14}$$

where λ is often referred to as a 'hidden variable'. We implicitly assume the 'freedom of choice' condition — the assumption that the choices of the measurement settings are independent of λ — is fulfilled.

2. The laboratory operations are represented by local transformations U_i^t on the target subsystems. They do not increase the 'amount' of violation of Bell inequalities on such subsystems.

This is satisfied in the considered class of GPTs by definition because the "amount of violation of Bell inequalities" is obtained by maximization over all local transformations as in Eq. (2.11) (or convex mixtures therefrom), and our 'laboratory operations' are assumed to be of such type. For concreteness, let us consider the CHSH inequality for correlation functions [Clauser *et al.*, 1969] — a similar reasoning applies to different forms of Bell inequalities. Following the Peres-Horodecki criterion [Duan *et al.*, 2000; Simon, 2000], the maximal value of the CHSH inequality in quantum mechanics is given in terms of two largest absolute values of the correlation tensor singular values,

say t_1 and t_2, as $2\sqrt{t_1^2 + t_2^2}$. The singular-value elements cannot increase under local transformations (*i.e.*, they are invariant under reversible local operations).

3. The order of S1's and S2's operations on the target system is well defined.

Suppose first that the orders of application of the local operations performed inside quantum switch S1 ($U^t_{1_A} \preceq U^t_{1_B} \preceq \dots$) and those performed inside quantum switch S2 ($U^t_{2_A} \preceq U^t_{2_B} \preceq \dots$) are fixed. Since an ordered sequence of local transformations is still a local transformation, if a state undergoes such a transformation on S1's and S2's sides, the amount of violation of Bell's inequalities cannot be increased (and, in particular, the singular values of the correlation tensor cannot increase, neither can the violation of the CHSH inequality). The amount cannot be increased even if the order of operations is chosen with a given probability distribution due to convexity. The mutual order between S1's and S2's operations is irrelevant, since we have assumed the two classes of operations to commute.

Theorem. *No states, set of transformations and measurements which obey the assumptions I-III can result in violation of a Bell inequality.*

Proof. Following I, suppose that the initial target state $\omega^t_{1,2}$ does not violate a Bell inequality. This means that Eq. (2.14) is fulfilled. Because of III, operations in S1's and in S2's laboratories are applied in a definite order, say $U^t_{1_A} \preceq U^t_{1_B} \preceq \dots$ in S1's side, and $U^t_{2_A} \preceq U^t_{2_B} \preceq \dots$ in S2's side. The state evolves, therefore, under a composition of the local operations as

$$\dots\ (U^t_{1_\mathrm{B}}, U^t_{2_\mathrm{B}}) \circ (U^t_{1_\mathrm{A}}, U^t_{2_\mathrm{A}})(\omega^t_{1,2}).$$

Let us restrict ourselves to the case of only two transformations per quantum switch (U^t_A and U^t_B). After the pairs of operations are applied in order $U^t_{1_\mathrm{A}} \preceq U^t_{1_\mathrm{B}}$ and $U^t_{2_\mathrm{A}} \preceq U^t_{2_\mathrm{B}}$ on the two sides, the state becomes

$$\omega^t_{1,2}{}' = (U^t_{1_\mathrm{B}}, U^t_{2_\mathrm{B}}) \circ (U^t_{1_\mathrm{A}}, U^t_{2_\mathrm{A}})(\omega^t_{1,2}) = \big(U^t_{1_\mathrm{B}} \circ U^t_{1_\mathrm{A}}, U^t_{2_\mathrm{B}} \circ U^t_{2_\mathrm{A}}\big)(\omega^t_{1,2}) \tag{2.15}$$

which is still local due to I - III. Hence

$$p(o_1, o_2|m_1, m_2, \omega^t_{1,2}{}') = \int \rho(\lambda)\ p(o_1|m_1, \lambda, \omega^t_{1,2}{}') \cdot p(o_2|m_2, \lambda, \omega^t_{1,2}{}')\ d\lambda. \tag{2.16}$$

In general, the order of operations does not need to be fixed, but can be specified probabilistically by a further hidden variable ν, whose different values correspond to different permutations of the order of operations. We obtain

$$p(o_1, o_2|m_1, m_2) = \iint \rho(\lambda, \nu)\ p(o_1|m_1, \lambda, \omega^{t,\nu}_{1,2}) \cdot p(o_2|m_2, \lambda, \omega^{t,\nu}_{1,2})\ d\lambda\ d\nu, \tag{2.17}$$

where $\rho(\lambda, \nu)$ is the joint probability distribution over the two types of variables, and $\omega_{1,2}^{t,\nu}$ is the final state of the target systems upon application of the transformations in the order given by ν.

Hence, we conclude that a local target state subjected to the action of a set of local operations applied in a pre-defined order can by no means lead to the violation of Bell inequalities, even if the order is chosen probabilistically in each run of the experiment. This concludes the proof. □

2.5.2 Relation between the present work and Ref. [Zych *et al.*, 2019]

In Ref. [Zych *et al.*, 2019], the position of a massive object serves as a 'control' quantum system and a quantum system (*e.g.*, a photon) that is exchanged between Alice's and Bob's laboratory as a 'target' system. By putting the massive object in a macroscopic superposition of two positions, one closer to Alice's and the other closer to Bob's position, one induces a relative time dilation between Alice's and Bob's laboratory. The superposition of massive objects can effectively lead to 'entanglement' of the temporal order between local operations, enabling the violation of a Bell-type inequality. In the conceptual framework of general relativity, the resource for the violation is a 'non-classical space-time' created by macroscopic superposition of large masses. In the second-quantized picture, the superposition can be seen as entanglement in the Fock basis, and the scheme enables one to "swap" this entanglement to the final entanglement of the target systems. Unfortunately, the physical demands of the proposal make that experiment infeasible. However, quantum control of the order of events can also be achieved without the use of gravitational interaction. This can be done, for example, in an extended quantum circuit model, wherein the order of applied quantum gates is coherently controlled by an ancillary system (the quantum switch). The difference between the two scheme is that in the gravitational scheme, the spatio-temporal distance of *any* pair of events in a space-time region is influenced by a superposition state of the mass, whereas in the linear optical implementation, only the order of the gates applied on the propagating system (*e.g.*, photons) is indefinite.

A more detailed analysis of the differences and similarities between the gravitational quantum switch and its photonic counterpart here presented is given in the Suppl. Information - Sec. 4.

2.5.3 Experimental Proof of Assumption I in GPTs

Recall that assumption I says that the initial target states do not violate a Bell inequality. In the notation introduced above, the initial target state is $\omega_{1,2}^{t}$. Our demonstration of assumption I presented here is based solely on experimental data, and can be shown to be valid for a class of GPTs. Our goal is to prove that the input state is a product state, and thus it is local.

Let us denote the probabilities for measurement outcomes as measured on reduced states of the target system of S1 and S2 as $p(o_1|m_1, \omega_1^t)$ and $p(o_2|m_2, \omega_2^t)$, respectively. If the state is a local product

state then the probability for joint outcomes, as measured on the composite system of the two target subsystems in the initial state $\omega_{1,2}^t$, is factorisable, *i.e.*, it can be expressed as

$$p(o_1, o_2|m_1, m_2, \omega_{1,2}^t) = p(o_1|m_1, \omega_1^t) \cdot p(o_2|m_2, \omega_2^t). \tag{2.18}$$

We experimentally performed a large set of measurements on the input target states, and checked for this property. The measurements we made are tomographically complete in quantum theory. Nevertheless, in a GPT this might not be the case, as a GPT system may have more degrees of freedom than a quantum system. We thus restrict our considerations to a class of GPTs for which we assume that polarization measurements in three unbiased bases for each photon constitute a subset of the full set of 'fiducial measurements'. For example, in the case of GPTs whose systems are described by Bloch vectors of dimension *d*, three components of the vectors correspond to 'quantum fiducial measurements' of a single system. Similarly, in the GPTs, the correlation tensor is given by d^2 elements of which 9 elements (*i.e.*, 3 fiducial measurements for the first times 3 fiducial measurements for the second system) correspond to the 'quantum subspace' of the correlations that are accessible through quantum measurements.

Our measurements confirm that the joint probabilities for 'quantum fiducial measurements' are factorized for the two targets. In general, however, it might be possible that within the subset of quantum fiducial measurements for a bipartite system the joint probabilities are factorized into a product of marginal probabilities although the overall state is not a product one. This is because non-zero correlations could exist between non-quantum fiducial measurements. It would then be possible to transfer correlations from the 'non-quantum subspace' into the correlations within the 'quantum subspace' by applying some 'exotic' (*i.e.*, non-quantum) local transformations. Nevertheless, for our class of GPTs, where subsystems are represented by Bloch vectors of general dimension d, we know that if both subsystems individually return probability one for some measurement outcome, the state is a pure product one and it cannot violate assumption I (*Lemma 1* of Ref. [Dakic and Brukner, 2010]). More precisely, the state would have the form $\omega_{1,2}^t = \omega_{1,2}^t(\omega_1^t, \omega_2^t, \omega_1^t(\omega_2^t)^T, \xi^t)$, with ω_1^t and ω_2^t being in pure states, *i.e.*, $||\omega_1^t|| = ||\omega_2^t|| = 1$. In our experiment, we obtain outcomes with probability one for a pair of quantum fiducial measurements on the two target subsystems, and hence the two subsystems cannot exhibit any further correlations within the non-quantum subspace.

Table S1 shows the values of the probabilities $p(o_1, o_2|m_1, m_2, \omega_{1,2}^t)$ (which, for brevity, is indicated as $p_{1,2}$ in the Tables) in the first four columns, and the marginal probability products $p(o_1|m_1, \omega_1^t) \cdot p(o_2|m_2, \omega_2^t)$ (denoted as $p_1 \cdot p_2$ in the Table) in the last four columns, with almost perfect correlations in the $\{H, V\}$ basis. Moreover, the joint and the two marginal probabilities are all almost one for the HH outcome, confirming the high purity of the bipartite state. It can be seen that the two sets of probabilities agree well. More quantitatively, let us define the *root-mean-square* (RMS) distance

between the two sets of probabilities as

$$\mathrm{d} = \sqrt{\frac{1}{N}\sum_{o_1,o_2}\sum_{m_1,m_2}\Delta p^2_{o_1,o_2,m_1,m_2}}, \tag{2.19}$$

with

$$\Delta p_{o_1,o_2,m_1,m_2} = p(o_1,o_2|m_1,m_2,\omega^T_{1,2}) - p(o_1|m_1,\omega^T_1)\cdot p(o_2|m_2,\omega^T_2), \tag{2.20}$$

and where N is the number of data points. Evaluating this over our results, we obtain a RMS distance of $(0.6 \pm 2.7)\cdot 10^{-2}$, indicating that the two distributions are equal within error.

Although the two target systems are approximately in a product state, the small discrepancy between the two distributions allows for some correlations between the systems. From the full set of the fiducial measurements in the quantum subspace, we can estimate the singular values of the correlation tensor and the maximal possible amount of violation of the CHSH inequality for the two target systems to be $2\sqrt{t_1^2+t_2^2} = 2.12 \pm 0.04$. This is more than ten standard deviations lower than the observed violation of the inequality. Thus, this digression from assumption I cannot explain the violation of the inequality.

2.5.4 Entangled Photon Source

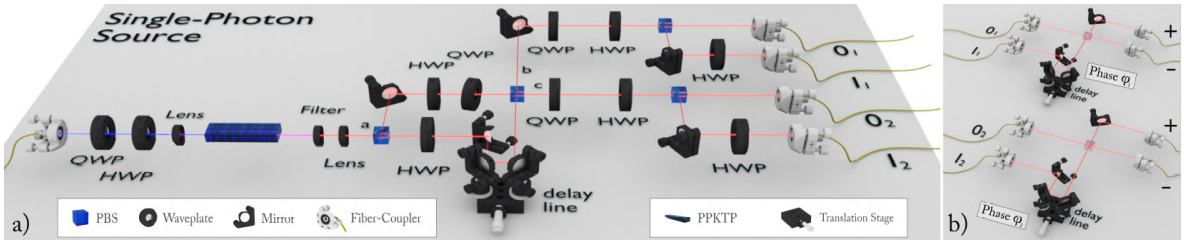


Figure 2.5 **Entangled photon-pair source. a)** *The source* — The beam from a Toptica DL Pro HP 426 laser is focused on a 30-mm-long PPKTP crystal, phase-matched for degenerate collinear type-II SPDC from 426 nm to 852 nm. The phase-matching is finely tuned by controlling the temperature of the crystal with a precision greater than $0.01K$. The emitted photons have a bandwidth of approximately 0.2 nm. After the crystal, the residual pump beam is filtered, the photons are then collimated and sent to a set-up to create entanglement by post-selection (as explained in the main text). The entanglement is first produced in polarization and then converted into path using polarizing beam splitters. The source produces ≈ 30.000 path-entangled photon pairs per second with a pump power of 8 mW. **b)** *Set-up used to measure a Bell Inequality on the path qubits* — The two paths composing each qubit are interfered on a beam splitter (BS) projecting each qubit onto a basis on the equator of the Bloch sphere (see main text for more details).

A *periodically-poled potassium titanyl phosphate* (PPKTP) crystal, phase-matched for collinear type-II *spontaneous parametric down-conversion* (SPDC), converts one photon at 426 nm into two photons at 852 nm. The photonic state after the crystal can be approximated to a Fock state of two photons in two orthogonal polarization modes $|H,a\rangle|V,a\rangle$, where a indicates the common spatial mode of the

two photons defined in Fig. 2.5. Two PBSs are used to separate and then recombine the two photons. Each photon passes through a HWP set at $\pm 45^{\circ}$. The state after the second PBS is therefore: $\left(|H,b\rangle\,|H,c\rangle - |H,b\rangle\,|V,b\rangle + |H,c\rangle\,|V,c\rangle - |V,b\rangle\,|V,c\rangle\right)/\,2$, where b and c indicate the two output spatial modes of the second PBS. By post-selecting on coincidences, only the part of the state with the photons in two different spatial modes is kept, resulting in the polarization-entangled state $\left(|H,b\rangle|H,c\rangle - |V,b\rangle|V,c\rangle\right)/\,\sqrt{2}$. We then use two PBSs and two HWPs (Fig. 2.5) to convert this state into a path-entangled state: $\left(|0\rangle_1\,|0\rangle_2 - |1\rangle_1\,|1\rangle_2\right)/\,\sqrt{2}$, where the notation is the same as specified in Fig. 2.5. A trombone delay line in between the two PBSs is used to compensate temporal delay between the two photons, and a multi-order QWP in one mode is tilted to compensate for undesired phases between the two components of the final quantum state. The delay line and the QWP can be also used to modify the final output state in a controllable way. In particular, by unbalancing the two paths by the coherence length of the down-converted photons, the entangled state can be converted into a statistical mixture of the states $|0\rangle_1|0\rangle_2$ and $|1\rangle_1|1\rangle_2$.

For our experiment, both the path and the polarization states of the photon pairs are important. To characterize the polarization state, we can perform two-qubit polarization state tomography using a QWP, a HWP and a PBS for each photon (Fig. 2.5, Panel **a**). To characterize the path entanglement, we perform a Bell measurement on the path qubits using the apparatus shown in Fig. 2.5, Panel **b**, which essentially consists of one Mach-Zehnder interferometer for each photon. The phase of the interferometers sets the measurement bases $\{\frac{1}{\sqrt{2}}(|0\rangle + e^{-i\phi_i}\,|1\rangle), \frac{1}{\sqrt{2}}(|0\rangle - e^{-i\phi_i}\,|1\rangle)\}$. Using these two interferometers we can measure all what is required for a CHSH parameter:

$$S = \left|C(o_1,o_2) + C(o_1',o_2) + C(o_1,o_2') - C(o_1',o_2')\right|, \tag{2.21}$$

where

$$C(o_1,o_2) = \frac{N_{++} - N_{+-} - N_{-+} + N_{--}}{N_{++} + N_{+-} + N_{-+} + N_{--}}. \tag{2.22}$$

Here, N_{++} is the number of coincidence events between detectors labelled $+$ for each photon in Fig. 2.5, Panel **b**, N_{+-} the number of coincidence events between detectors $+$ and $-$ for each photon, and so on.

Fig. 2.6 shows the characterization of the entanglement of the joint input control state, where we verified the initial entanglement by performing a Bell measurement on the joint control system before the quantum switch, obtaining a CHSH parameter of 2.58 ± 0.09.

2.5.5 Experimental Proof of Assumption IIb in GPTs under the assumption of III

In this section, we experimentally prove the validity of assumption IIb for the case in which assumption III holds. Assumption IIb states that the laboratory operations do not couple (*i.e.*, do not generate non-classical correlations between) the control and the target subsystems, within a given party Si,

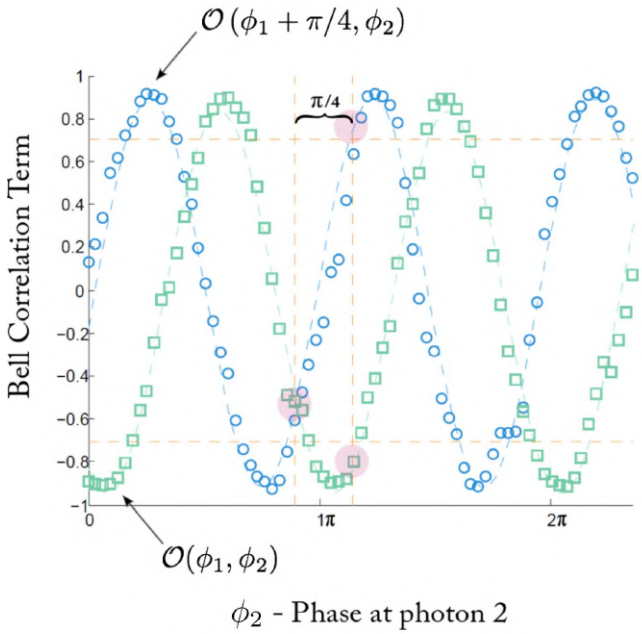


Figure 2.6 **Input control state characterization: Bell measurement on the order qubits.** Each curve is a measurement of a Bell correlation term $C\big(\mathcal{O}_1(\phi_1), \mathcal{O}_2(\phi_2)\big)$ on the control qubits, wherein the phase of ϕ_1 is fixed, and the phase ϕ_2 is scanned. As described in Eq. (2.21) of the Methods - Sec. 2.5.4, we test the Clauser-Horne-Shimony-Holt (CHSH) inequality [Clauser *et al.*, 1969] achieving a violation of 2.59 ± 0.09. For the data in the green curve, the phase ϕ_1 was nominally shifted by $\pi/4$ rad with respect to the blue curve. The red shaded areas represent the regions where values of ϕ_1 and ϕ_2 correspond with those used to construct our CHSH parameter (Eq. (2.21) of the Methods - Sec. 2.5.4). In particular, $\mathcal{O} = (\mathcal{O}_1, \mathcal{O}_2)$ where $\mathcal{O}_i(\phi_1, \phi_2) = cos(\phi_i)\,\sigma_x + cos(\phi_i)\,\sigma_z$. These data confirm that the two photons start in a path-entangled state, and the polarization state is initially separable.

$i = 1, 2$. To test it, we first prepare the control and target subsystems in a tomographically-complete set of product states within quantum theory, *i.e.*,

$$p(o^c, o^t | m^c, m^t, \omega_i = \omega^c \omega^t) = p(o^c | m^c, \omega^c) \cdot p(o^t | m^t \omega^t) \tag{2.23}$$

for all states $\omega_i = \omega^c\,\omega^t$ from a tomographically-complete set of product states. In the GPTs, Eq. (2.23) shows that, for a product state from the quantum subspace, the probability for a joint outcome factorizes into the product of the probabilities for individual outcomes. We then set a single quantum switch to have only one operation inserted, either U_{i_A} or U_{i_B}. We finally verify that, for the full set of preparations, the control and target subsystems are still in a product state after the quantum switch, when this contains only U_{i_A} or only U_{i_B}. More precisely, we verify that

$$p(o^c, o^t | m^c, m^t, U_{i_A}\omega_i) = p(o^c | m^c, U_{i_A}\omega_c)\, p(o^t | m^t, U_{i_A}\omega_t) \tag{2.24a}$$

$$p(o^c, o^t | m^c, m^t, U_{i_B}\omega_i) = p(o^c | m^c, U_{i_B}\omega_c)\, p(o^t | m^t, U_{i_B}\omega_t), \tag{2.24b}$$

for any state from a complete set of product states, and by linear extension to an arbitrary product state ω_i. We do this using the same technique we used to verify that the target qubits began in an input

state (Methods - Sec. 2.5.3). Finally, we make use of the following property: if neither operation U_{i_A} nor U_{i_B} alone couple the two subsystems, then also a sequence of the two operations cannot couple them as long as they are performed in a definite causal order. This conclusion follows directly from Eqs. (2.23)-(2.24b):

$$p(o_c, o_t|m_c, m_t, U_{i_B} \circ U_{i_A}\omega_i) = p(o_c|m_c, U_{i_B} \circ U_{i_A}\omega_c)\; p(o_t|m_t, U_{i_B} \circ U_{i_A}\omega_t) \tag{2.25}$$

Note that even under a multiple usage of U_{i_A} and U_{i_B} there can be no coupling when the operations are performed in a definite causal order. This finalizes the proof of assumption IIb.

Tables S2-S3 report the values of the probabilities $p(o_c, o_t|m_c, m_t, \omega_1)$ (which, for brevity, are indicated as $p_{c,t}$ in the Tables) compared with the marginal probability products $p(o_c|m_c, \omega_1^c)\cdot p(o_t|m_t, \omega_1^t)$ (denoted as $p_c \cdot p_t$ in the Tables).

The tomographically-complete sets of fiducial quantum measurements reported in Tables S2-S3 were performed as follows. In order to vary the input state of the control system among $|+\rangle^c$, $|-\rangle^c$, $|R\rangle^c = (|0\rangle^c - i\,|1\rangle^c)/\sqrt{2}$, and $|L\rangle^c = (|0\rangle^c + i\,|1\rangle^c)/\sqrt{2}$, we set the relative phase between the two trajectories after the first beamsplitter by means of a delay stage mounted on a calibrated piezo-actuator. Instead, by blocking either path, we prepared $|0\rangle^c$ and $|1\rangle^c$. Likewise, we measure the path qubit in the following way. To measure in $\{|+\rangle^c, |-\rangle^c\}$, or $\{|R\rangle^c, |L\rangle^c\}$, we suitably set the relative phase between the two paths before recombining them at the second beamsplitter. This can be done by adding the required phase for state preparation and subtracting the phase for state measurement. Such a phase is then converted into a path delay and sent to the piezo-actuated delay stage. (We use the same delay stage to both set the phase of the path state, and to measure it in $\{|+\rangle^c, |-\rangle^c\}$, or $\{|R\rangle^c, |L\rangle^c\}$.) To measure in the $\{|0\rangle, |1\rangle\}$ basis, we block either path before the 50/50 beamsplitter, and we then sum the counts from the two paths after the beamsplitter.

The displayed output probabilities $p(o_c, o_t|m_c, m_t, \omega_1)$ are very close to those corresponding to a product state. This is indicated by the fact that the RMS distance [Eq. (2.19)] between these two sets of probabilities (the measured joint probabilities $p(o_c, o_t|m_c, m_t, \omega_1)$, and that given by the product $p(o_c|m_c, \omega_1^c) \cdot p(o_t|m_t, \omega_1^t)$) is $(2 \pm 3) \cdot 10^{-2}$ when only operation U_{i_A} was acting on the system, and $(3 \pm 3) \cdot 10^{-2}$ when only operation U_{i_B} was present. Moreover, the displayed output probabilities $p(o_c, o_t|m_c, m_t, \omega_1)$ are very close to those of a pure (product) state, which means that there cannot be any correlations in the non-quantum subspace. More precisely, the non-vanishing discrepancy between the two probability distributions could be caused by a weak coupling between the control and the target system. Using the same technique as in Methods - Sec. 2.5.3, we can estimate that the correlations established through this coupling are too weak to violate the CHSH inequality ($2\sqrt{t_1^2 + t_2^2} = 1.76 \pm 0.04$). The coupling between each pair of the control and the target systems can "swap" the correlations from the bipartite state of the control system to that of the target system. However, assuming that the transferred amount of correlations to the target system cannot be larger

that the amount produced through the coupling between each pair of the target and control system, we conclude that the coupling cannot result in the target systems violating the Bell's inequality.

This confirms within experimental error, under the hypothesis that assumption III is valid, that assumption IIb holds in our experiment. Furthermore, this proof holds not only within quantum theory but also for our class of GPTs. In other words, our measurements imply that in both of the two quantum switches, individually, the laboratory operations do not couple the target and the control subsystems in GPTs when these operations are executed in a definite causal order. From this experimental test, we thus conclude that in our experiment assumption II cannot be false unless assumption III is also violated.

2.5.6 Data Analysis

In order to convert the coincidence counts into probabilities, we weight each measured count rate by the net detection efficiency of the corresponding detector pair. We estimate these efficiencies in two parts. First, we measure the relative coupling efficiencies between the output ports M_1 and M_2 of quantum switch S1, and M_3 and M_4 of quantum switch S2. Then, within each output port, we measure the relative efficiency of the detector in the transmitted port and the reflected port. We find relative efficiencies between ≈ 0.85 and 1. For more details, see the Methods of our previous work [Rubino *et al.*, 2017a].

The main source of error in our experiment was phase fluctuations. In the Bell measurement, this dephasing is mainly due to two contributions. (1) Undesired phase-shifts in the interferometer (which we estimated to be about 0.97°). (2) Fluctuations of the source, which produces time varying phase between the $|HH\rangle$ and $|VV\rangle$ terms. In our source, we estimate this to be approximately 1.9°, which is caused by a combination of fluctuations in the pump laser wavelength, and the phase-matching temperature. We convert these errors into an error in the Bell parameter using Gaussian error propagation. To calculate the error for the Bell measurements on the polarization qubits after the quantum switches, we consider the same error sources as above (where now the phase shifts in the measurement interferometer are replaced by phase shifts in the quantum switches). However, we also consider errors arising from setting the polarization measurements. Finally, to estimate the errors in the results extracted from tomography (*i.e.*, fidelity and concurrence), we performed a Monte Carlo simulation considering the phase fluctuations discussed above.

2.5.7 Additional consistency tests: the insertion of noise

In this section, we present two further tests of consistency of our experimental proof.

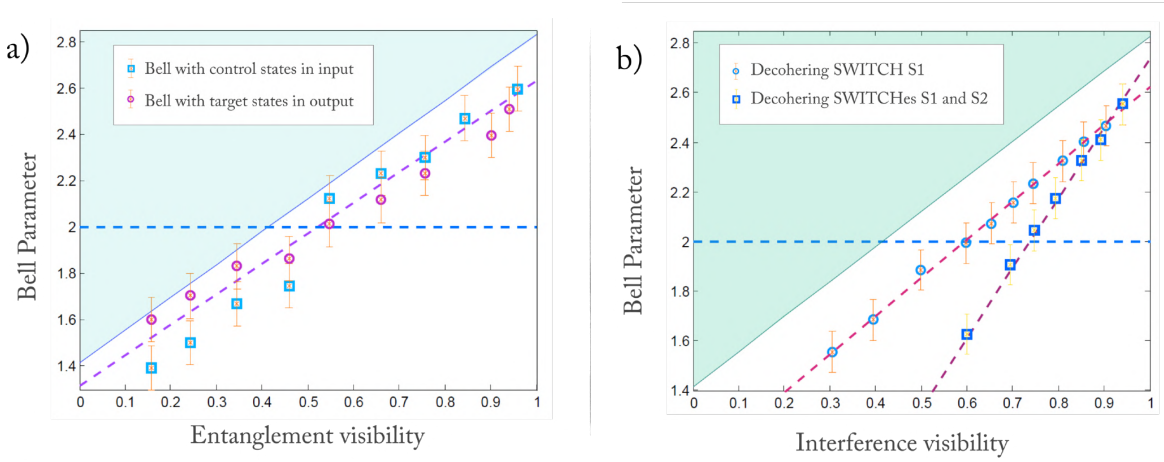


Figure 2.7 Bell parameter in the presence of various decoherence sources. a) For these data, the initial entanglement of control qubits is decreased passing from the entangled state $\frac{1}{\sqrt{2}}\left(|0,0\rangle - |1,1\rangle\right)$ to a mixture of $|0,0\rangle$ and $|1,1\rangle$. We measure the Bell parameter both on the input path qubits (squares) and output polarization qubits (circles) as the source is decohered. Here, the Bell parameter is plotted versus the visibility of the entangled state in its anti-correlated basis. The dashed line is a simulation of the experimental results. **b)** For these data, the coherence of the superposition of the orders of operations inside the quantum switches is decreased, leading to a classical mixture of orders. To control this transition, we decrease the visibility of either only one of the two interferometers (circles), or of both interferometers at the same time (squares). Each graph shows the Bell parameter plotted versus the visibility of one interferometer. The dashed lines are linear fits to the data. The horizontal dashed blue line, in both plots, is the classical limit for a Bell violation. When the state of the control qubit is too decohered, we can no longer violate a Bell inequality.

First, we decreased the entanglement of the joint control system by increasing the delay of the interferometer inside the source (see the Methods - Sec. 2.5.4). The more mixed the state of the control system becomes, the smaller is the amount of violation of a Bell inequality with the target systems which we can achieve, up until reaching the threshold of non-violation. The Bell parameter versus the 'source visibility' (*i.e.*, the two-photon visibility in its anti-correlated basis) is plotted in Fig. 2.7**a**. The dashed line is a calculation of the expected Bell parameter, including the imperfect visibility of the two interferometers. All the data points agree with the expected trend within error. The small step at an entanglement visibility of around 0.5 was caused by a lower fringe visibility which increased the systematic error in setting the phases ϕ_1 and $\phi_1 + \pi/4$ (see Fig. 2.3).

As a second test, we decreased the degree of causal non-separability of the two processes. To do this, we introduced distinguishing information between the paths corresponding to the orders $U_{i_A} \preceq U_{i_B}$ and $U_{i_B} \preceq U_{i_A}$ (in only one quantum switch, squares in Fig. 2.7**b**, and in both simultaneously, circles in Fig. 2.7**b**) by lengthening one of the paths with respect to the other, effectively reducing the visibility of the interferometers comprising the quantum switches. As this occurs, we transition from a superposition of temporal orders to a mixture of them (in other words, to a causally-separable process, which satisfies assumption III). If all three assumptions are met, one cannot violate a Bell inequality between the two systems. Indeed, we experimentally observe that as the visibility is decreased, the Bell parameter also decreases (Fig. 2.7, Panel **b**).

2.6 Supplemental Information

2.6.1 quantum switch and Causal Inequalities

The quantum switch [Procopio *et al.*, 2015; Rubino *et al.*, 2017a] has been shown not to violate causal inequalities, making it impossible to use such a violation as a theory-independent proof that the causal order of the operations in the quantum switch is indefinite. Here, we briefly re-examine such reasoning following Refs. [Araújo *et al.*, 2015; Oreshkov and Giarmatzi, 2016].

We introduce the x, y and z indices to refer, respectively, to the measurements choices of Alice, Bob and Charlie. We call a, b and c their respective measurement results. It is always possible to re-write $p(a, b, c|x, y, z)$ as

$$p^{\text{switch}}(a, b, c|x, y, z) = p(c|a, b, x, y, z)\, p(a, b|x, y, z). \tag{2.26}$$

It should be noticed that, regardless of the causal order between operations in Alice's and Bob's laboratory, the operation in Charlie's laboratory always occurs after them. In other words, his operation is in the future light cone of both Alice's and Bob's operations. Thus, a and b cannot depend on z, so

$$p(a, b|x, y, z) = p(a, b|x, y). \tag{2.27}$$

As previously observed, after tracing out Charlie's laboratory in the quantum switch, the process matrix of Alice and Bob is causally separable. Hence, one can rewrite $p(a, b|x, y)$ in the form of a convex mixture, obtaining

$$p^{\text{switch}}(a, b, c|x, y, z) = p(c|a, b, x, y, z)\big[\zeta \cdot p^{A\preceq B}(a, b|x, y) + (1 - \zeta) \cdot p^{B\preceq A}(a, b|x, y)\big], \tag{2.28}$$

with $\zeta \geq 0$. We can combine the probabilities $p^{A\preceq B}(a, b|x, y)$ ($p^{B\preceq A}(a, b|x, y)$) and $p(c|a, b, x, y, z)$ as a product of the probability respecting the order $A \preceq B$ ($B \preceq A$) with the probability respecting the order $\{A, B\} \preceq C$

$$p^{\text{switch}}(a, b, c|x, y, z) = \zeta \cdot p^{A\preceq B\preceq C}(a, b, c|x, y, z) + (1 - \zeta) \cdot p^{B\preceq A\preceq C}(a, b, c|x, y, z). \tag{2.29}$$

Therefore, the quantum switch is a process whose probabilities have a 'causal model', *i.e.*, it can always be understood as arising from events that are causally ordered, or from a convex mixture of causally ordered events. Hence, it satisfies all causal inequalities.

2.6.2 Hidden local definite causal order

In general, while experimental tests can be used to prove that the conjunction of the assumptions underlying a given no-go theorem does not describe the phenomenology observed within quantum mechanics, they do not provide information on which of the assumptions is to be discarded. In this experiment, the application of the Bell's theorem to temporal orders allowed us to test a conjunction of all our assumptions; yet, in order to verify which assumptions are valid, additional tests on a single quantum switch were necessary. This notwithstanding, it is worth noting that testing only one single quantum switch would not have provided an as stringent information. In fact, as we showed above the experimental data taken from a single quantum switch cannot violate causal inequalities, and hence can be understood as arising from an underlying causal model, in the spirit of simulation of quantum statistics by hidden variables. Such a model generates statistics compatible with operations performed on a system in a definite order, or in a convex mixture therefrom. In terms of probabilities, the statistics in the quantum switch $p^{\text{switch}}(a,b,c|x,y,z) = \int d\lambda\, \rho(\lambda)\, p^{\text{causal}}(a,b,c|x,y,z,\lambda)$, where $p^{\text{causal}}(a,b,c|x,y,z,\lambda) = p^{A\preccurlyeq B\preccurlyeq C}(a,b,c|x,y,z)$ or $p^{B\preccurlyeq A\preccurlyeq C}(a,b,c|x,y,z)$, could hence be mimicked by an underlying causal hidden variable model. The statistics obtained measuring the double quantum switch with entangled temporal orders rules out a *local* causal hidden variable model that allows for this description, *i.e.*, its statistics is incompatible with

$$\begin{aligned} p^{\text{2-switches}}&(a_1,b_1,c_1,a_2,b_2,c_2|x_1,y_1,z_1,x_2,y_2,z_2) = \\ &= \int d\lambda\, \rho(\lambda)\, p^{\text{causal}}(a_1,b_1,c_1|x_1,y_1,z_1,\lambda)\, p^{\text{causal}}(a_2,b_2c_2|x_2,y_2,z_2,\lambda). \end{aligned} \tag{2.30}$$

In other words, the causal model is called 'local' if the statistics can be understood as originating from (a convex mixture of) operations performed in each local laboratory according to a definite causal order. Our experimental data rule out the models in Eq. (2.30) for the special case where Alice and Bob both apply a single operation with a single outcome (unitary).

2.6.3 Device-Independency and Theory-Independency

Causal witnesses, violation of Bell inequalities for temporal orders, and violation of causal inequalities build a hierarchy of the notion of 'the lack of causality'. The weakest notion of the lack of causality is that of causal non-separability, which is formulated using quantum theory. A violation of a causal inequality is the strongest notion as it is formulated solely in terms of observable probabilities $p(a,b|x,y)$ without any assumption about the internal function of experimental devices — it is therefore device-independent. The violation of a Bell inequality for temporal orders should be considered, in our view, a stronger proof of lack of causality than the measurement of a causal witness, but a weaker proof than a violation of a causal inequality. The reason why it is weaker than a causal inequality violation is that, although it too is formulated in terms of the probabilities $p(a,b|x,y,\omega)$,

it also involves the notion of state ω and the assumption how laboratory operations act on it (see Methods - Sec. A of the main text) — this causes the proof to be device-dependent. However, it can be defined for a class of generalized probabilistic theories, and hence it does not rely on the quantum formalism. It is thus more general than the notion of a causal witness. Although the quantum switch violates a weaker notion of causality, shaped for quantum theory, it cannot violate the stronger (device-independent) notion of causal inequalities. The open question addressed in our work is: "Can we still use the quantum switch to perform a proof of indefinite causal orders independent of quantum formalism?" The answer is affirmative, and this study represents an experimental demonstration of this.

2.6.4 On the Physical Implementability of the quantum switch

Skepticism has been expressed about whether a tabletop experiment can demonstrate indefinite causal structures. In Ref. [MacLean *et al.*, 2017], the authors claim that it is not possible to implement the quantum switch without 'exotic physical scenarios'. In particular, they argue that one would need a closed time-like curve, and even then such an implementation would be inconsistent, being able to generate logical contradictions such as the grandfather paradox. These criticisms are based on the assumption that causal structures must be represented via directed graphs. In this representation, the quantum switch becomes a directed graph with a cycle, which could indeed be inconsistent and could generate logical contradictions.

The tension between directed acyclic graphs (DAG) and causal structures in the quantum switch is akin to the tension between classical "hidden" variable theories and quantum theory. For example, in order to describe an interferometric experiment in terms of classical variables, one is forced to say that the interfering system follows some exotic trajectories or in some non-local manner follows two classical trajectories 'at once'. However, within quantum theory one interprets interferometric tests as demonstrating that the very assumption that a system does follow a definite path is violated.

The formal sense in which the causal order of applying operations in a quantum switch is non-classical has been recently studied in Ref. [Oreshkov, 2019]. The motivation of that work was to understand where and when the operations happen in the quantum switch, which is precisely the question brought up in the context of a DAG representation. The author shows that the operations applied on systems in a quantum switch act on subsystems that are not localised in time, *i.e.*, on 'time-delocalised' subsystems. It is further shown that standard quantum theory, without exotic closed timelike-curves, is compatible with such time-delocalised operations and that they indeed realise genuine non-separable quantum processes. The work also concludes that experimental realisations of the quantum switch, including specifically its entangled version described in this work, are genuine realisations of such time-delocalised processes. In other words, there is a well-defined fashion in which temporal relations between the application of operations in a quantum switch cannot be represented with DAGs.

In fact, the Bell theorem for temporal orders [Zych *et al.*, 2019] and its version in this work can be interpreted as a limitation on achievable correlations when operations acting on a quantum system can be embedded in a causal structure compatible with an underlying DAG (or a probabilistic mixture thereof).

Therefore, the suitable conclusion to draw is that the causal structure in the quantum switch cannot be represented by a DAG since the latter can only represent what are called *definite* causal structures [Miklin *et al.*, 2017]. What our work demonstrates experimentally is that the quantum switch represents an *indefinite* causal structure incompatible with any DAG, just like experimental violations of Bell's inequalities show that there exist correlations incompatible with local hidden variables.

The authors of Ref. [MacLean *et al.*, 2017] further argue that, in a genuine quantum switch, operations must be performed in the same spatio-temporal regions in each term of the superposition, so that only their order is swapped. As mentioned above, Ref. [Oreshkov, 2019] showed that this is not necessary: in the quantum switch a single operation can be 'time-delocalised' over two (or more) spatio-temporal regions. In the originally proposed implementation of the quantum switch [Chiribella *et al.*, 2013], as well as in ours, one could in principle register the time at which the signal passes through each box, which would decohere the superposition and make the interference between the causal orders vanish. However, since we do observe coherence, such information does not exists (*i.e.*, it is not stored in any physical degree of freedom, as this would alter the results of the experiment). The above requirement of 'the same spatio-temporal regions' whose order is simply swapped is in principle realisable in a gravitational implementation of the quantum switch for a certain choice of coordinates. There, a massive object is prepared in a spatial superposition, which results in the causal order between two events being opposite in the two superposed terms. More precisely, a choice of coordinates can be found such that gate $\mathcal{A}$ is performed at a single time (the proper time of a local clock). However, an alternative choice of coordinates may as well be done such that the gate $\mathcal{A}$ is performed in a superposition of different times (according to the coordinate time or the time of a distant observer) in different superposed terms, before and after the gate $\mathcal{B}$. Therefore, even in this gravitational case, it is always possible to make a choice of coordinates where the operations appear to be performed at different times in the different superposed terms. Thus, it is in fact insubstantial to argue whether the operations are "really performed at the same times, and their order is swapped", or they are "merely performed in superposition of different times", as this depends on the choice of coordinates in which one wishes to describe the scenario (see [Allard Guérin and Brukner, 2018; Zych *et al.*, 2019] for further discussions). Furthermore, note that, contrary to the arguments of [MacLean *et al.*, 2017], this proposal does not allow for the information to travel back in time, nor does it require closed time-like curves, and it does not give rise to any logical paradoxes.

In Ref. [MacLean *et al.*, 2017], another criticism follows from the observation that the quantum switch can easily be 'simulated' by using additional copies of the boxes $\mathcal{A}$ and $\mathcal{B}$, as was already noted in the original proposal of the quantum switch [Chiribella *et al.*, 2013]. In particular, one could use an

unfolded Mach-Zehnder interferometer, with gates $\mathcal{A}_1$ and $\mathcal{B}_1$ in one arm and gates $\mathcal{B}_2$ and $\mathcal{A}_2$ on the other. The straightforward response is that we do not use an unfolded Mach-Zehnder interferometer, but rather a folded one, and therefore that we use a single copy of each box instead of two. The number of applications of a box can operationally be determined by a counter (*i.e.*, a 'flag') that is raised each time the operation is applied on the system. The very fact that an unfolded Mach-Zehnder interferometer requires two copies of each box to 'simulate' the statistics of a folded one is a signature that the latter exhibits an indefinite causal order. Moreover, we also note that, in the unfolded version of the interferometer, it would be necessary to actively make $\mathcal{A}_1$ precisely equivalent to $\mathcal{A}_2$, whereas in our case this clearly follows from the implementation itself, as the gates are physically the same. It should also be emphasized that the present Bell-type proof of an indefinite causal order is valid even if the local gates are used more than once, as clarified in the main text.

Furthermore, following the reasoning in Ref. [MacLean *et al.*, 2017], one could say that it is in principle possible to make the gate $\mathcal{A}$ ($\mathcal{B}$) act differently when it comes before or after $\mathcal{B}$ ($\mathcal{A}$), as in each case the photon passes through $\mathcal{A}$ ($\mathcal{B}$) at different times. This is true, but also applies to the originally proposed implementation. To make it *locally* impossible, one could use the above mentioned superposition of a massive object to control the order of operations in two space-time regions. In such a scenario, indeed, Alice (Bob) in her (his) local laboratory cannot make the gate $\mathcal{A}$ ($\mathcal{B}$) act differently in case the operation $\mathcal{A}$ ($\mathcal{B}$) is performed before or after $\mathcal{B}$ ($\mathcal{A}$). Nevertheless, a distant observer for whom Alice's (Bob's) operation happens in a superposition of two coordinate-times could make such contingency occur with a cleverly designed set-up (*e.g.*, by sending a signal which triggers Alice's operation to change once it is received, as depicted in Fig. 2.8). As a consequence, as much as in the case of a table-top experiment the operation $\mathcal{A}$ can be made to act differently depending on whether it happens before or after $\mathcal{B}$, in its gravitational counterpart this can be achieved by a distant observer who triggers some change for certain time-coordinates. In conclusion, the requirement that operations $\mathcal{A}$ and $\mathcal{B}$ must *even in principle* be forbidden to change depending on the order has no absolute meaning (*i.e.*, it cannot be realized in all reference frames). Moreover, if the operations differed or their time was revealed (or stored in any degree of freedom), the results of the experiment would differ.

Finally, because of the differences highlighted above, one may object that the physics that describes a photonic quantum switch is not equivalent to the one which is behind the gravitational quantum switch for all possible observers. This is indeed correct. In fact, in the first case, the physics is described by Maxwell equations on Minkowski space-time, whereas in the latter case it is non-classical space-time that determines the dynamics. However, although a local as well as a global observer could tell the difference between the gravitational and the photonic quantum switch, it is not the case for a quantum particle which travels along the two superposed paths in either versions of the quantum switch. In fact, in both cases the particle experiences a genuine quantum superposition of causal orders. And this is precisely the purpose of this experimental work: we do not aim to draw conclusions concerning global/local observers, but on the system undergoing the quantum process. Therefore, neither of the

schemes (*i.e.*, the gravitational and the photonic quantum switch) is a 'simulation' of one another. They are rather two equivalent representations of the dynamics experienced by a quantum particle in presence of a quantum superposition of causal orders, *i.e.*, two representations of a quantum switch.

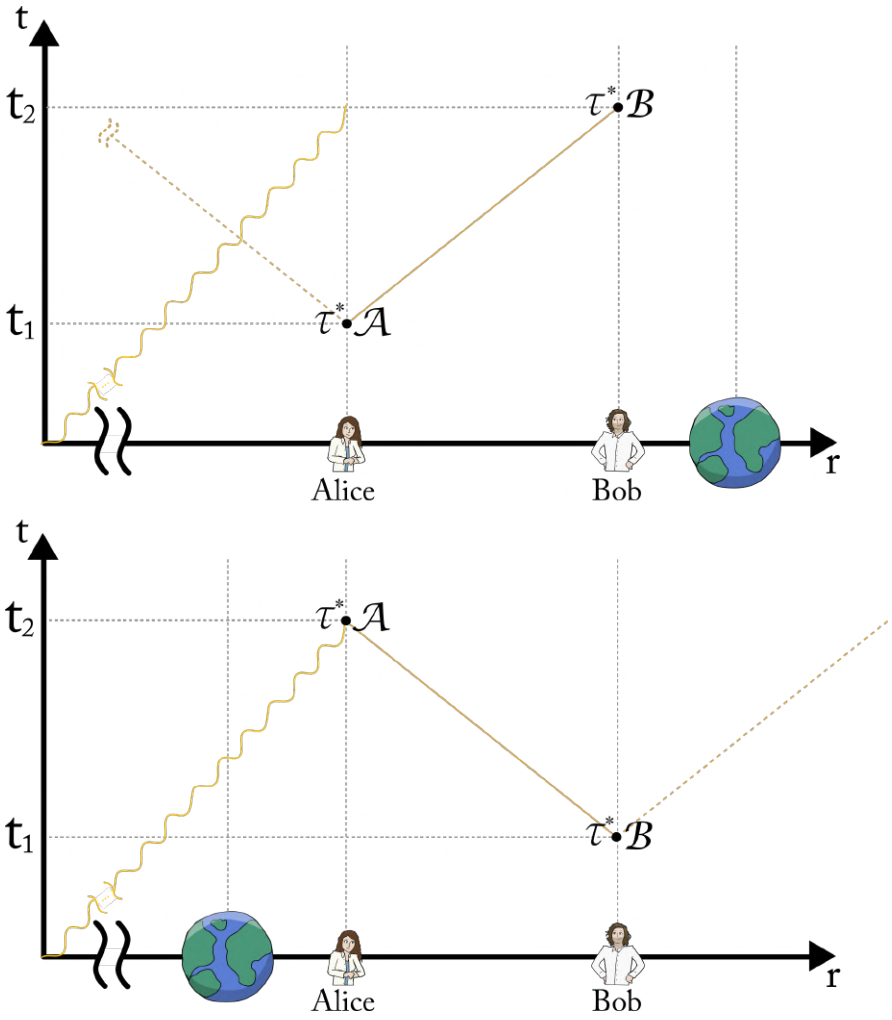


Figure 2.8 **Schematic of a gravitational quantum switch.** A quantum system is exchanged between Alice's and Bob's laboratories. The order in which such 'target' system is exchanged is governed by a second system, a 'control' system, which is encoded in the position of a massive object. By putting the massive object in a macroscopic superposition of two positions, one closer to Alice's and the other closer to Bob's position, one induces a relative time dilation between Alice's and Bob's laboratories. If an outside observer sends some system at a suitably chosen time, let us call it $t_{A\succ B}$, the observer could influence the functioning of the device that implements $\mathcal{A}$, *e.g.*, when it acts second but not when it acts first, making the operation of Alice act different depending on the order.

Table 2.1 Comparison between the two-states probabilities $p(o_1, o_2|m_1, m_2, \omega^t_{1,2})$ and the products of marginal single-state probabilities $p(o_1|m_1, \omega^t_1) \cdot p(o_2|m_2, \omega^t_2)$ for the input target states. - Part I. The compatibility between the two sets of probabilities shows the separability of the input target state $\omega^t_{1,2}$. We indicate with 'H' and 'V' the states of horizontal and vertical polarization, with 'D' and 'A' the diagonal and anti-diagonal states, with 'R' and 'L' the circular polarization states right- and left-handed. The experimental error associated to each of these probabilities is ± 0.01.

Measur. Basis	$p_{1,2}$	$p_{1,2^\perp}$	$p_{1^\perp,2}$	$p_{1^\perp,2^\perp}$	$p_1 \cdot p_2$	$p_1 \cdot p_{2^\perp}$	$p_{1^\perp} \cdot p_2$	$p_{1^\perp} \cdot p_{2^\perp}$
H, H	0.97	0.03	0.00	0.00	0.97	0.03	0.00	0.00
H, V	0.01	0.99	0.00	0.00	0.01	0.99	0.00	0.00
H, A	0.58	0.41	0.00	0.00	0.59	0.41	0.00	0.00
H, D	0.42	0.58	0.00	0.00	0.42	0.58	0.00	0.00
H, R	0.39	0.61	0.00	0.00	0.39	0.61	0.00	0.00
H, L	0.61	0.38	0.00	0.00	0.62	0.38	0.00	0.00
V, H	0.00	0.00	0.96	0.04	0.00	0.00	0.96	0.04
V, V	0.00	0.00	0.03	0.97	0.00	0.00	0.03	0.97
V, A	0.00	0.00	0.61	0.39	0.00	0.00	0.61	0.39
V, D	0.00	0.00	0.38	0.61	0.00	0.00	0.38	0.61
V, R	0.00	0.00	0.35	0.64	0.00	0.00	0.35	0.64
V, L	0.00	0.00	0.64	0.36	0.00	0.00	0.64	0.36
A, H	0.39	0.01	0.53	0.02	0.41	0.01	0.54	0.02
A, V	0.01	0.37	0.02	0.54	0.02	0.39	0.02	0.54
A, A	0.24	0.16	0.34	0.21	0.26	0.16	0.34	0.21
A, D	0.18	0.22	0.21	0.33	0.18	0.24	0.23	0.31
A, R	0.16	0.25	0.18	0.35	0.16	0.27	0.19	0.34
A, L	0.26	0.14	0.36	0.19	0.27	0.14	0.36	0.19
D, H	0.55	0.02	0.45	0.02	0.53	0.02	0.45	0.02
D, V	0.01	0.57	0.02	0.45	0.01	0.54	0.01	0.46
D, A	0.32	0.26	0.29	0.18	0.32	0.24	0.28	0.20
D, D	0.23	0.35	0.18	0.29	0.22	0.34	0.18	0.29
D, R	0.21	0.37	0.16	0.31	0.19	0.37	0.16	0.31
D, L	0.35	0.22	0.32	0.17	0.34	0.20	0.31	0.18
R, H	0.65	0.02	0.33	0.01	0.64	0.02	0.33	0.01
R, V	0.01	0.66	0.01	0.30	0.02	0.66	0.01	0.31
R, A	0.39	0.27	0.22	0.13	0.40	0.26	0.21	0.14
R, D	0.29	0.39	0.12	0.19	0.28	0.40	0.13	0.19
R, R	0.27	0.41	0.11	0.20	0.26	0.42	0.12	0.19
R, L	0.41	0.25	0.22	0.12	0.41	0.24	0.22	0.13
L, H	0.32	0.01	0.63	0.04	0.32	0.02	0.63	0.03
L, V	0.01	0.32	0.03	0.64	0.01	0.32	0.02	0.65

Table 2.2 Comparison between the two-states probabilities $p(o_1, o_2|m_1, m_2, \omega^t_{1,2})$ and the products of marginal single-state probabilities $p(o_1|m_1, \omega^t_1) \cdot p(o_2|m_2, \omega^t_2)$ for the input target states. - Part I. The compatibility between the two sets of probabilities shows the separability of the input target state $\omega^t_{1,2}$. We indicate with 'H' and 'V' the states of horizontal and vertical polarization, with 'D' and 'A' the diagonal and anti-diagonal states, with 'R' and 'L' the circular polarization states right- and left-handed. The experimental error associated to each of these probabilities is ± 0.01.

Measur. Basis	$p_{1,2}$	$p_{1,2^\perp}$	$p_{1^\perp,2}$	$p_{1^\perp,2^\perp}$	$p_1 \cdot p_2$	$p_1 \cdot p_{2^\perp}$	$p_{1^\perp} \cdot p_2$	$p_{1^\perp} \cdot p_{2^\perp}$
L, A	0.18	0.14	0.42	0.27	0.19	0.13	0.41	0.28
L, D	0.14	0.21	0.23	0.41	0.13	0.22	0.24	0.40
L, R	0.12	0.23	0.22	0.43	0.12	0.23	0.22	0.43
L, L	0.21	0.12	0.44	0.24	0.21	0.12	0.43	0.25

Table 2.3 Comparison between the two-states probabilities $p(o_c, o_t|m_c, m_t, \omega_1)$ and the products of marginal single-state probabilities $p(o_c|m_c, \omega^c_1) \cdot p(o_t|m_t, \omega^t_1)$ for the control and the target states when only operation U_{i_A} is acting on the input state. We denoted as 0, 1, $+$, $-$, l and r the analogue of the polarization states H, V, D, A, L, R in the path degree of freedom. The two sets of probabilities associated to the control and the target states in output are compatible within experimental errors. The experimental error associated to each of these probabilities is ± 0.01.

Meas. Basis (target)	Prep.-Meas. Basis (control)	$p_{c,t}$	$p_{c,t^\perp}$	$p_{c^\perp,t}$	$p_{c^\perp,t^\perp}$	$p_c \cdot p_t$	$p_c \cdot p_{t^\perp}$	$p_{c^\perp} \cdot p_t$	$p_{c^\perp} \cdot p_{t^\perp}$
H	$+$	0.95	0.00	0.04	0.00	0.95	0.00	0.04	0.00
D	$+$	0.47	0.48	0.01	0.03	0.47	0.49	0.02	0.02
R	$+$	0.48	0.47	0.01	0.03	0.47	0.48	0.02	0.02
H	$-$	0.07	0.00	0.92	0.01	0.07	0.00	0.91	0.01
D	$-$	0.04	0.04	0.48	0.44	0.04	0.04	0.48	0.44
R	$-$	0.04	0.04	0.41	0.51	0.04	0.04	0.41	0.51
H	*r*	0.55	0.00	0.44	0.01	0.55	0.01	0.44	0.00
D	*r*	0.20	0.26	0.28	0.26	0.22	0.24	0.26	0.28
R	*r*	0.28	0.24	0.18	0.30	0.24	0.28	0.22	0.26
H	*l*	0.50	0.00	0.50	0.00	0.50	0.00	0.50	0.00
D	*l*	0.30	0.28	0.21	0.21	0.30	0.28	0.22	0.21
R	*l*	0.27	0.30	0.20	0.23	0.27	0.30	0.20	0.23
H	0	0.51	0.00	0.49	0.00	0.50	0.00	0.49	0.00
D	0	0.26	0.30	0.26	0.17	0.30	0.27	0.23	0.21
R	0	0.28	0.26	0.23	0.23	0.27	0.26	0.24	0.23
H	1	0.56	0.00	0.43	0.01	0.56	0.00	0.44	0.00
D	1	0.27	0.29	0.22	0.23	0.27	0.29	0.21	0.23
R	1	0.28	0.28	0.17	0.27	0.25	0.31	0.20	0.24

Table 2.4 Comparison between the two-states probabilities $p(o_c, o_t|m_c, m_t, \omega_1)$ and the products of marginal single-state probabilities $p(o_c|m_c, \omega_1^c) \cdot p(o_t|m_t, \omega_1^t)$ for the control and the target states when only operation U_{i_B} is acting on the input state. The two sets of probabilities associated to the control and the target states in output are compatible within experimental errors. The experimental error associated to each of these probabilities is ± 0.01.

Meas. Basis (target)	Prep.-Meas. Basis (control)	$p_{c,t}$	$p_{c,t^\perp}$	$p_{c^\perp,t}$	$p_{c^\perp,t^\perp}$	$p_c \cdot p_t$	$p_c \cdot p_{t^\perp}$	$p_{c^\perp} \cdot p_t$	$p_{c^\perp} \cdot p_{t^\perp}$
H	$+$	0.47	0.33	0.11	0.09	0.47	0.33	0.12	0.08
D	$+$	0.50	0.27	0.18	0.06	0.51	0.25	0.16	0.08
R	$+$	0.75	0.02	0.23	0.01	0.75	0.02	0.23	0.00
H	$-$	0.11	0.15	0.49	0.25	0.16	0.11	0.44	0.29
D	$-$	0.12	0.12	0.60	0.16	0.17	0.07	0.55	0.21
R	$-$	0.25	0.01	0.67	0.07	0.24	0.02	0.68	0.06
H	*r*	0.43	0.44	0.10	0.03	0.46	0.41	0.07	0.06
D	*r*	0.54	0.32	0.09	0.05	0.54	0.32	0.09	0.05
R	*r*	0.86	0.01	0.11	0.02	0.84	0.03	0.13	0.00
H	*l*	0.16	0.06	0.49	0.29	0.14	0.08	0.51	0.27
D	*l*	0.13	0.09	0.62	0.15	0.17	0.05	0.59	0.19
R	*l*	0.20	0.01	0.73	0.05	0.20	0.01	0.73	0.05
H	0	0.26	0.28	0.26	0.19	0.29	0.26	0.24	0.22
D	0	0.40	0.14	0.41	0.05	0.44	0.10	0.37	0.09
R	0	0.48	0.04	0.42	0.06	0.47	0.05	0.44	0.05
H	1	0.32	0.23	0.29	0.15	0.34	0.22	0.27	0.17
D	1	0.32	0.24	0.32	0.12	0.36	0.20	0.28	0.16
R	1	0.56	0.00	0.41	0.03	0.54	0.02	0.43	0.01

Chapter 3

Experimental Quantum Communication Enhancement by Superposing Trajectories

G. Rubino, L. A. Rozema, D. Ebler, H. Kristjánsson, S. Salek, P. Allard Guérin, A. A. Abbott, C. Branciard, Č. Brukner, G. Chiribella, and P. Walther

Abstract. In quantum communication networks, wires represent well-defined trajectories along which quantum systems are transmitted. In spite of this, trajectories can be used as a quantum control to govern the order of different noisy communication channels, and such a control has been shown to enable the transmission of information even when quantum communication protocols through well-defined trajectories fail. This result has motivated further investigations on the role of the superposition of trajectories in enhancing communication, which revealed that the use of quantum-control of parallel communication channels, or of channels in series with quantum-controlled operations can also lead to communication advantages. Building upon these findings, here we experimentally and numerically compare different ways in which two trajectories through a pair of noisy channels can be superposed. We observe that, within the framework of quantum interferometry, the use of channels in series with quantum-controlled operations generally yields the largest advantages. Our results contribute to clarify the nature of these advantages in experimental quantum-optical scenarios, and showcase the benefit of an extension of the quantum communication paradigm in which both the information exchanged and the trajectory of the information carriers are quantum.

Author contributions: G.R. and L.A.R. designed the experiment. G.R. built the set-up and carried out the data collection. G.R. and L.A.R. performed data analysis. All authors contributed to the interpretation of the results and to writing the paper, based on an initial draft by G.R. and L.A.R..

3.1 Introduction

The ability to establish secure communication linkages is of paramount importance in any information technology. Quantum cryptography protocols [Bennett and Brassard, 1984; Ekert, 1991] achieve this in a stunning way, enabling a sender and receiver to communicate securely even in the presence of an eavesdropper with unlimited computational power. The crucial ingredient for this feat is the availability of reliable transmission lines for quantum particles. In this framework, any noisy process affecting the transmission is attributed to the presence of an eavesdropper, and when the noise exceeds a given threshold, the security of the communication is considered compromised. For this reason, the mitigation of any noise arising from faulty transmission lines is an integral part of the efforts to enable secure communication.

Within the quantum communication networks paradigm, quantum communication protocols encode information in quantum states, yet they treat the propagation of information carriers as classical [Chiribella *et al.*, 2009]. Nevertheless, the information carriers can propagate along non-classical trajectories, experiencing a coherent superposition of alternative quantum evolutions [Åberg, 2004; Aharonov *et al.*, 1990; Oi, 2003]. Taking advantage of this fact, Gisin *et al.* [Gisin *et al.*, 2005] realized that quantum superpositions of trajectories can be harnessed to reduce the noise induced by a pair of noisy communication channels. Therein, it was shown that when the quantum information carriers [1] are sent through two noisy channels in a quantum superposition of trajectories, interference between the two resulting noisy processes can sometimes lead to partial cancellation of the noise via post-selection.

Recently, interest in this discovery has been revived by studies emerging from quantum foundations. In particular, it was shown that the superposition of trajectories can generate setups where the order of different channels is in a quantum superposition. These setups produce the same output as a mathematical map called the 'quantum switch' [Chiribella *et al.*, 2013], a higher-order operation which takes two quantum channels as input and combines them in a quantum-controlled order. The quantum switch is an instance of a causally-indefinite process; such processes are currently the target of wide-ranging research both for fundamental reasons [Ämin Baumeler and Wolf, 2016; Hardy, 2007; Oreshkov *et al.*, 2012], and for their potential to provide advantages in quantum computation [Araújo *et al.*, 2017; Chiribella, 2012; Chiribella *et al.*, 2013; Colnaghi *et al.*, 2012; Facchini and Perdrix, 2015; Hardy, 2009; Taddei *et al.*, 2020], quantum communication complexity [Baumeler and Wolf, 2014; Feix *et al.*, 2015; Guérin *et al.*, 2016], and quantum metrology [Zhao *et al.*, 2020]. Moreover, the particular class of causally-indefinite processes based on the superposition of alternative orders can be probed via current experimental technologies, as has been recently done by encoding information in various degrees of freedom of single photons [Goswami *et al.*, 2020, 2018; Guo *et al.*, 2020; Procopio *et al.*, 2015; Rubino *et al.*, 2017a,b; Taddei *et al.*, 2020; Wei *et al.*, 2019].

[1] We use the notion of 'particles' as a synonym for quantum systems which, naturally, can be delocalised in space and time. These quantum systems are used as carriers of quantum information, and in this sense we interchangeably refer to them also as 'information carriers'.

It was further proposed [Chiribella *et al.*, 2018; Ebler *et al.*, 2018; Salek *et al.*, 2018] that the quantum switch can also reduce noise in classical and quantum communication. These findings triggered a host of subsequent proposals [Caleffi and Cacciapuoti, 2020; Procopio *et al.*, 2020, 2019; Sazim *et al.*, 2020; Wilson *et al.*, 2020], and even a few experiments [Goswami *et al.*, 2020; Guo *et al.*, 2020], highlighting the advantage of using quantum superpositions of noisy channels in alternative orders to reduce transmission noise.

However, alongside the body of work focused on superpositions of alternative orders, the use of superpositions of trajectories in quantum communication has also been investigated [Abbott *et al.*, 2020; Chiribella and Kristjánsson, 2019; Guérin *et al.*, 2019; Kristjánsson *et al.*, 2020; Loizeau and Grinbaum, 2020]. In this context, theoretical studies have pointed out that causal-indefiniteness is not necessarily required to reduce the noise in classical and quantum communication [Abbott *et al.*, 2020; Guérin *et al.*, 2019; Loizeau and Grinbaum, 2020]. In particular, similar or even better advantages can be achieved by using a quantum-control of parallel noisy channels [Abbott *et al.*, 2020], or by placing channels in series with quantum-controlled operations [Guérin *et al.*, 2019]. Indeed, in Ref. [Guérin *et al.*, 2019] it was even shown that the Shor quantum error correcting code can be used to find a channel layout in series with quantum-controlled gates which allows any arbitrary noise to be completely eliminated. This suggested the need for a thorough information-theoretic understanding of the resources in play, and a unified description of such protocols. One such approach is presented in Refs. [Chiribella and Kristjánsson, 2019; Kristjánsson *et al.*, 2020]. On the other hand, the comparison of different protocols can be also viewed as an experimental task, wherein one wishes to classify and quantify the experimental resources required for a physical implementation of the various types of superpositions of trajectories and their corresponding advantages [2].

We take the experimental approach here, focusing on three different types of superpositions of trajectories which have been identified in the literature, namely, quantum-control of parallel channels (Fig. 3.1**a**)), channels in series with quantum-controlled operations (Fig. 3.1**b**)), and quantum-control of channel order (Fig. 3.1**c**)). While previous experimental studies [Goswami *et al.*, 2020; Guo *et al.*, 2020] focused only on the reduction of noise with an indefinite causal order, no experimental work had so far implemented the other proposed schemes, nor had they compared them with indefinite causal structures to provide an exhaustive assessment of the resources in play. We find that the common resource in all the three schemes considered is the establishment of a coupling between the trajectories of the information carriers and the degree of freedom on which the noise acts. On this basis, we propose a fundamentally new understanding of the resources required for this noise reduction than that proposed in previous experimental works in this field [Goswami *et al.*, 2020; Guo *et al.*, 2020].

We experimentally apply the above three schemes to various noise models. This enables us to examine the utility and trade-offs of these different types of superpositions in the goal of communicating

[2] In this Chapter, in contrast to Ref. [Kristjánsson *et al.*, 2020], we consider resources as an experimental concept, corresponding to the physical devices and their interactions as they occur in the laboratory.

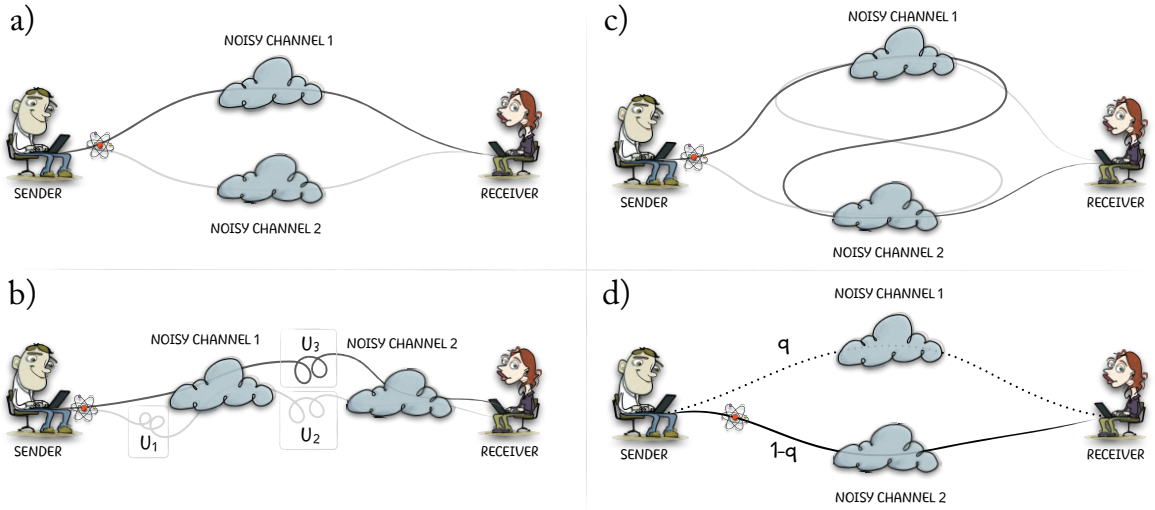


Figure 3.1 Combining two channels in a superposition of trajectories. A sender and a receiver communicate under the restriction that the information carrier must cross at least one noisy region. **a) Quantum-Control of Parallel Channels**. A quantum particle is placed in a quantum superposition of two trajectories, each branch containing a single noisy channel. **b) Channels in Series with Quantum-Controlled Operations.** Each of the branches of the superposition passes through the noisy channels in the same order, but different unitary operations are applied locally in each branch. **c) Quantum-Control of Channel Order.** The information carriers are routed through the two channels in different orders. This setup can achieve a genuinely indefinite order of the two channels. **d) Classical Trajectories.** Throughout this Chapter, we will compare the three quantum superpositions of channels above to classical trajectories. In this regard, if one has access to classical-like trajectories only, one can send the photon through one or the another noisy regions with probabilities q and $1-q$.

through a pair of noisy channels. In particular, in order to perform a comparative analysis of the performance of the three types of superpositions, we measure the coherent information (which is a lower bound for the quantum channel capacity) in the presence of *XY*, bit-flip, phase-flip and BB84-channels. We show that, within the paradigm of quantum interferometry, the use of channels in series with quantum-controlled operations generally outperforms or equals the other schemes in all the noise models which we consider. While here we study the three schemes individually in order to focus on the source of the coupling between the trajectory and the degree of freedom on which the noise acts, one could of course also combine the different types of superpositions (and, for instance, insert quantum-controlled operations also in the other two schemes), yielding different—potentially larger—advantages from those presented here.

The rest of this Chapter is structured as follows. Section 3.2.1 introduces the three different architectures for the quantum superpositions of trajectories through two noisy channels, and summarizes their performance when applied to a simple noise model. Section 3.2.2 reviews the key figures of merit that we use to quantify the performance of our experimental quantum channels, *i.e.*, the quantum capacity and the coherent information. Section 3.2.3 outlines our experiments, and Section 3.3 presents the corresponding results. Finally, Section 3.4 concludes.

3.2 Theoretical Framework and Experiment

3.2.1 Quantum Superpositions of Trajectories

For simplicity, we will focus on two communication channels and two trajectories, as this already captures the key features of the general idea.

All experiments hereafter discussed were performed using single photons, where the trajectory is naturally defined by the photon's path. Quantum information is initially encoded in one of the internal degrees of freedom of the particle (we refer to Section 3.5.1 for a discussion of the case of classical information); in our case, in its polarization. Then, using linear optical elements, it is relatively easy to place a photon in a superposition of trajectories [Goswami *et al.*, 2018; Procopio *et al.*, 2015; Rubino *et al.*, 2017a,b]. We will further consider, as does related work, that the noise acts only on the internal degree of freedom (DOF).

To introduce the basic idea, we will start by considering a particular noise model, which was studied for quantum-controlled orders in [Chiribella *et al.*, 2018]. Given some single-qubit input state ρ encoded in the internal DOF, the noisy process $\mathcal{C}$ either applies a Pauli-X or -Y operation to the internal state with equal probability:

$$\mathcal{C}_{\mathrm{EB}}(\rho) = \frac{1}{2} X\rho X + \frac{1}{2} Y\rho Y. \tag{3.1}$$

If the input to this process is a pure state $|\psi\rangle_{\mathrm{I}} = \alpha\,|0\rangle_{\mathrm{I}} + \beta\,|1\rangle_{\mathrm{I}}$ (where the subscript I denotes the internal DOF), the output is, in general, a mixed state, with all coherence in the computational basis extinguished:

$$\mathcal{C}_{\mathrm{EB}}(|\psi\rangle_{\mathrm{I}}\langle\psi|_{\mathrm{I}}) = \begin{pmatrix} |\beta|^2 & 0 \\ 0 & |\alpha|^2 \end{pmatrix}, \tag{3.2}$$

and as such, it cannot be used to transmit any quantum information. One might, of course, still employ it to transmit classical information in the computational basis. This channel is an example of a so-called 'entanglement-breaking' (EB) channel, which would destroy any preexisting entanglement between the transmitted qubit and any other system.

In a standard quantum communication scenario with a single trajectory, information, which is taken to be encoded in an internal DOF of an information carrier, must often propagate through multiple channels. Depending on the physical implementation, the channels can be linked together in different manners. With two channels and classical-like trajectories, the channels can either be put in series, or in a classical mixture of the two (depicted in Fig. 3.1**d**))—more complex combinations can also be realized, but they all perform strictly worse than a classical mixture. If two copies of the channel of Eq. (3.1) are put in series, the result is a maximally-dephasing channel $\mathcal{C}(\rho) = \frac{1}{2}\rho + \frac{1}{2} Z\rho Z$, where Z is the Pauli-Z matrix. This also destroys all coherence in the computational basis, and cannot transmit

any quantum information. Similarly, placing two of these channels in a classical mixture will not allow the transmission of any quantum information.

In a typical single-trajectory quantum communication scenario, it can be shown that, if each channel is unable to transmit quantum information (*i.e.*, its quantum capacity—to be defined later—is zero), then any combinations of the two channels should also result in a zero capacity channel. This is known as a bottleneck inequality [Yu *et al.*, 2020]. In the following subsections, we will show that this is not the case when the trajectories are superposed in a quantum fashion. Thus, the bottleneck inequality does not directly apply to communication scenarios with quantum trajectories [Guérin *et al.*, 2019; Salek *et al.*, 2018].

3.2.1.1 Quantum-Control of Parallel Channels

The first layout that we consider uses a quantum superposition of configurations where two independent channels are placed in parallel, and their use is controlled by a quantum system, as illustrated in Fig. 3.1**a)**. This was originally introduced for error filtration [Gisin *et al.*, 2005], and it was more recently reviewed in the general framework of communication through superposed channels in Refs. [Abbott *et al.*, 2020; Chiribella and Kristjánsson, 2019]. In this scheme, different independent noisy channels are placed in each branch of the superposition. In Ref. [Gisin *et al.*, 2005], it was shown that by performing a measurement on the trajectory in a suitable basis, and then post-selecting, one can non-deterministically filter out errors in the communication channel. We will now consider an initial pure state encoded in the internal DOF $|\psi\rangle_{\mathrm{I}} = \alpha\,|0\rangle_{\mathrm{I}} + \beta\,|1\rangle_{\mathrm{I}}$, independent noisy channels realised by applying a Pauli-X and -Y with equal probabilities—as described previously and resulting in Eq. (3.1)— and two trajectories in an equal superposition $|+\rangle_{\mathrm{T}} = (|0\rangle_{\mathrm{T}} + |1\rangle_{\mathrm{T}})/\sqrt{2}$ (where T refers to the trajectory DOF). It is then straightforward to calculate the output (the full calculation is presented in Section 3.5.2), and to observe that performing a measurement on the trajectory DOF in the $\{|+\rangle_{\mathrm{T}}, |-\rangle_{\mathrm{T}}\}$ basis, and finding $|-\rangle_{\mathrm{T}} = (|0\rangle_{\mathrm{T}} - |1\rangle_{\mathrm{T}})/\sqrt{2}$ (which, as shown in Section 3.5.2, occurs with probability 1/4) leaves the internal DOF in the pure state:

$$\beta\,|0\rangle_{\mathrm{I}} - i\alpha\,|1\rangle_{\mathrm{I}}\,, \tag{3.3}$$

which can be unitarily rotated back to $|\psi\rangle_{\mathrm{I}}$. On the other hand, when the trajectory state is found to be $|+\rangle_{\mathrm{T}}$ (which happens with probability 3/4), the output state is partially mixed:

$$\begin{pmatrix} |\beta|^2 & -i\alpha^*\beta/3 \\ i\alpha\beta^*/3 & |\alpha|^2 \end{pmatrix}. \tag{3.4}$$

This output state has a reduced purity, but it still maintains some coherence. Although this is not necessarily the optimal measurement strategy or the best noise model to showcase this scheme, it illustrates that a quantum-controlled superposition of noisy channels allows some coherence to reach

the receiver. Hence, the sender and the receiver can communicate some quantum information. We will quantify the amount of quantum information precisely in Section 3.2.3. Communication advantages in this case have been attributed to the ability to quantum coherently control which channel to use [Abbott *et al.*, 2020].

This type of architecture is relatively easy to imagine deploying in practice. Most modern quantum communication takes place via optical fibers. As is often the case, these fibers can be noisy, resulting in a reduced ability to transmit information. Since a photon can easily be sent through a superposition of two (or more) fibers, the use of such parallel architectures could already improve security in existing communication networks.

3.2.1.2 Channels in Series with Quantum-Controlled Operations

A different way to significantly reduce the noise produced by some channels is to let them be traversed by two trajectories in a superposition, and by allowing different operations in each branch of the superposition. In this case, we will place our two channels in series, resulting in the architecture presented in Fig. 3.1**b)**. In each branch, the channels 1 and 2 are placed in the same order, and different unitary operations may be inserted. (Such unitary operations are labeled as U_1, U_2 and U_3 in Fig. 3.1**b)**. In principle, however, more operations could be inserted along the trajectories). This scheme was originally presented in Ref. [Guérin *et al.*, 2019], where it was referred to as a 'superposition of direct pure processes'.

Let us now consider the action of the superposition of trajectories in series with the noise model of Eq. (3.1), setting, following the notation of Fig. 1**b)**, $U_1 = Y$, $U_2 = \mathcal{I}$, $U_3 = \mathcal{I}$ ($\mathcal{I}$ being the identity operator). We will again consider the initial state of the system to be $|\psi\rangle_{\mathrm{I}} |+\rangle_{\mathrm{T}}$. This time, we will imagine performing a measurement in the computational basis on the qubit stored in the internal DOF. As we show in Section 3.5.2, finding the internal qubit in $|0\rangle_{\mathrm{I}}$ projects the trajectory state into $|\psi\rangle_{\mathrm{T}}$, while finding it in $|1\rangle_{\mathrm{I}}$ projects the trajectory state into $X |\psi\rangle_{\mathrm{T}}$. Hence, this superposition of trajectories perfectly filters out the noise arising from the noisy channels. (Notice that the ability to completely restore an arbitrary initial state of the information qubit implies that, were the information carrier initially entangled with an additional qubit, due to linearity this entangled state would be completely restored in turn.)

It is easy to imagine the implementation of this scheme in a real-world scenario. The two paths (*e.g.*, optical fibers) are simply sent through a few noisy transmission channels in series. (For the scheme to work, the action of each noisy channel must be correlated along the different paths.) Since the two paths are physically distinct, the different unitary operations can easily be applied in each branch of the superposition independently. Such operations can be performed with simple linear optical elements, or even directly using calibrated optical fibers, which always implement some unitary polarization

rotation. In Ref. [Guérin *et al.*, 2019] it was also pointed out that, by superposing more than two trajectories, one can perfectly compensate for any arbitrary noise.

3.2.1.3 Quantum-Control of Channel Order

The original source of inspiration for this architecture is the quantum switch [Chiribella *et al.*, 2013], a higher-order operation which takes quantum gates and applies them in a quantum superposition of alternative orders. Within quantum-interferometry, a quantum-optical switch exploiting superposition of trajectories in flat space-time has been proposed [Araújo *et al.*, 2014; Friis *et al.*, 2014], and experimentally demonstrated [Goswami *et al.*, 2020, 2018; Guo *et al.*, 2020; Procopio *et al.*, 2015; Rubino *et al.*, 2017a,b; Taddei *et al.*, 2020; Wei *et al.*, 2019]. For two quantum operations, this is a quantum process in which a particle is placed in a superposition of two paths, each of which is routed through the two quantum operations in alternative orders (see Fig. 3.1**c**)). This scheme features all the necessary requirements for an advantage in quantum information processing over standard channels [Goswami *et al.*, 2020; Guo *et al.*, 2020], and it can be provably characterized as a causally-indefinite process [Araújo *et al.*, 2015; Branciard, 2016; Goswami *et al.*, 2018; Oreshkov, 2019; Oreshkov and Giarmatzi, 2016; Rubino *et al.*, 2017a,b].

Applying the quantum switch to two copies of the channel in Eq. (3.1), one finds that the output state is [Chiribella *et al.*, 2018]

$$\frac{1}{2}\left|\psi\right\rangle_{\mathrm{I}}\left\langle\psi\right|_{\mathrm{I}}\otimes\left|+\right\rangle_{\mathrm{T}}\left\langle+\right|_{\mathrm{T}}+\frac{1}{2}Z\left|\psi\right\rangle_{\mathrm{I}}\left\langle\psi\right|_{\mathrm{I}}Z\otimes\left|-\right\rangle_{\mathrm{T}}\left\langle-\right|_{\mathrm{T}}. \tag{3.5}$$

Analogously to the previous two examples, we will now measure the trajectory in the $\{\left|+\right\rangle_{\mathrm{T}},\left|-\right\rangle_{\mathrm{T}}\}$ basis. If the outcome is $\left|+\right\rangle_{\mathrm{T}}$, the state has been transmitted perfectly, whereas if one finds $\left|-\right\rangle_{\mathrm{T}}$, a simple phase correction is required to exactly restore the initial state.

The resources required to implement the quantum-optical switch in the laboratory are relatively minimal, it simply requires linear optical elements to route the photon through the two noisy channels in a superposition of their orders. However, in order to be effective, this layout requires the action of the two noisy channels on the photon to be suitably correlated both in space and time (as in Fig. 3 of Ref. [Kristjánsson *et al.*, 2020]). Instead, in standard communication networks [Minář *et al.*, 2008], the noisy regions are usually localized in space and fixed in time. Any such network would thus require the photon to travel back toward the sender to enter the second channel, and this scheme requires this return trip to occur without traversing any further noisy region (which could happen if the two channels introduce noise in the direction from the sender to the receiver, but not vice versa).

3.2.1.4 Comparison

Comparison of different schemes in this work — In all of the three schemes above, a particle is placed in a quantum superposition of two trajectories which are then routed through various devices and noisy communication channels. All three methods result in a coupling of the internal state to the state of the trajectory, and the two trajectories must maintain coherence in order to show a communication advantage. (Practically, this is required since the trajectory is measured in a superposition basis.)

This coupling to the trajectory DOF is a necessary requirement to achieve any advantages. In the parallel and indefinite order layouts the channels themselves or the routing through their different orders give rise to the coupling, whereas in the series scheme this coupling is created by the quantum-controlled operations. In light of this, it has been proposed that the quantum-controlled operations used in the superpositions of channels in series (Fig. 3.1**b**)) should be considered as additional resources (referred to as 'encoding', 'decoding' and 'repeaters' in Refs. [Chiribella and Kristjánsson, 2019; Kristjánsson *et al.*, 2020]), as they can couple the internal DOF to the trajectory independently of the choice of noisy channels. However, these operations do not require any additional experimental resources beyond the transmission lines themselves (for example, polarization rotations can arise from the mere twisting of optical fibers, and are effectively unavoidable), which are the same experimental resources used for the other two schemes.

It was also noted that the number of noisy channels traversed by the particle in each branch of the superposition differs between the three schemes [Abbott *et al.*, 2020; Kristjánsson *et al.*, 2020; Kristjánsson *et al.*, 2020; Loizeau and Grinbaum, 2020]: the quantum-control of parallel channels contains only one channel in each interferometer arm, whereas the other two schemes contain two channels per arm. When the information carrier crosses several noisy channels in sequence, the overall noise is always equal to (in the case, *e.g.*, of two EB channels) or greater than (*e.g.*, in the case of two depolarising channels of the form $\rho' = p\rho + (1-p)\frac{\mathcal{I}}{2}$) that introduced by one channel. However, although the quantum-control of channel order needs at least two channels to create the required coupling between the trajectory and the internal DOF, it is still able to overcome the (potentially additional) noise caused by the multiple noisy channels.

Comparison to previous work — The origin of the communication enhancement in the three schemes studied here has been a subject of recent debate in the literature [Abbott *et al.*, 2020; Chiribella and Kristjánsson, 2019; Guérin *et al.*, 2019; Kristjánsson *et al.*, 2020; Loizeau and Grinbaum, 2020]. This debate revolved around the understanding of the role of causal indefiniteness in the task of noise reduction. In fact, after it was discovered that such an enhancement could be achieved by placing the channels in an indefinite causal order, it was later found that other configurations, which did not have an indefinite causal order, could achieve the same or even a better enhancement. This called into question whether indefinite causality is necessary to achieve such effects.

In this work, we compare experimentally and numerically all the proposed setups leading to noise cancellation in quantum communication, and we quantify the achieved advantages over quantum communication schemes with classical trajectories. This represents the first experimental comparison among these different schemes, which provides an answer to the debate on experimental grounds by presenting an experimentally-relevant analysis of the resources in play. In particular, we compare the schemes with respect to their experimental requirements within an interferometric paradigm, rather than studying them as higher-order operations from a strictly theoretical viewpoint [Chiribella and Kristjánsson, 2019; Kristjánsson *et al.*, 2020]. To this end, we focus on the following four points: *i.* we illustrate that all three schemes use the same resource when considering experimental quantum interferometry, *ii.* we show that this resource is the coupling of the degree of freedom carrying the information to the trajectory degree of freedom, *iii.* we experimentally prove that, for the set of tested noisy channels, the superposition of channels in series with quantum-controlled operations features the highest performance, and *iv.* we numerically show that, in the vast majority of cases, this holds for generic randomly-generated channels.

We will now proceed to quantify the amount of quantum information which can be transmitted using these various schemes individually for different noise models. Overall, we find that, for all the types of noise considered, the use of channels in series with quantum-controlled operations exceeds or equals the performance of the quantum-control of parallel channels and quantum-control of channel order.

3.2.2 Quantifying Channel Performance

In order to rigorously compare the ability of the different schemes to transmit quantum information, an experimentally accessible figure of merit is necessary. The quantum capacity $\mathcal{Q}(\mathcal{C})$ of a channel $\mathcal{C}$ is the number of qubits that are transmitted for each use of that channel [Devetak, 2005; Lloyd, 1997]. In general, this is a rather complex function that can be difficult to even theoretically assess, making its use as a quantifier somewhat limited in practice. However, it is lower bounded by [Lloyd, 1997]:

$$\mathcal{Q}(\mathcal{C}) \geq \max_{\rho_{AB}} \mathcal{I}_{\mathrm{c}}(\mathcal{C}, \rho_{AB}), \tag{3.6}$$

where $\mathcal{I}_{\mathrm{c}}$ is the coherent information [Schumacher and Nielsen, 1996] of the channel with respect to ρ_{AB}, which is defined as

$$\mathcal{I}_{\mathrm{c}}(\mathcal{C}, \rho_{AB}) := S(\rho'_B) - S(\rho'_{AB})\,, \tag{3.7}$$

where ρ_{AB} is a bipartite state, $\rho'_{AB} := (\mathcal{I}_A \otimes \mathcal{C})(\rho_{AB})$ is the output state obtained by applying channel $\mathcal{C}$ on system B, $\rho'_B := \mathrm{Tr}_A[\rho'_{AB}]$ is its marginal state, and $S(\rho) := -\mathrm{Tr}[\rho \log \rho]$ is the von Neumann entropy. Although a comparison of the coherent information of two channels does not necessarily translate into a comparison of the quantum capacity of the channels (except, of course, when the lower bound is maximal, as in this case it coincides with the quantum capacity), we will employ

it as our quantifier for channel performance here. In addition to the practical motivation of being a more readily computable quantifier, this choice is further motivated by the fact that the coherent information has an operational meaning beyond that of the quantum capacity. Namely, it provides the one-way distillable entanglement when maximised over local operations performed by the sender, and, if maximised over LOCC operations, it gives the two-way distillable entanglement [Devetak and Winter, 2005], which is typically considered to be the amount of 'useful entanglement' which can be transmitted using the channel.

Let us briefly consider a few simple examples, assuming a 2-qubit system, with ρ_{AB} set to a maximally-entangled pure state. If the channel is unitary (*i.e.*, noiseless), then the initially pure state remains pure after the application of the channel, thus $S(\mathcal{C}(\rho'_{AB})) = 0$. However, since ρ_{AB} is maximally entangled, tracing out the subsystem A will leave the subsytem B in a maximally mixed one-qubit state with entropy $S(\mathcal{C}(\rho'_B)) = 1$. Therefore the coherent information of a unitary channel, with respect to a maximally entangled probe state is 1. If, on the other hand, the channel induces decoherence, the entangled probe state will become mixed, and the second term will increase: $S(\mathcal{C}(\rho'_{AB})) > 0$. Because the first term cannot be larger than 1, as decoherence is induced the coherent information decreases. $\mathcal{I}_{\mathrm{C}}(\mathcal{C}, \rho_{AB})$ is often maximized when ρ_{AB} is a maximally-entangled state. This was proven to be the case for the quantum switch and a specific noisy model in [Salek *et al.*, 2018]. Furthermore, numerical optimisations suggest that a maximally-entangled Bell state maximises the coherent information for the cases we study here. In any case, the evaluation of $\mathcal{I}_c$ for any arbitrary state sets a lower bound for the quantum capacity of the channel. Throughout the rest of this Chapter, when we refer to the coherent information, we do so with reference to a maximally-entangled Bell state.

One could consider estimating $\mathcal{I}_{\mathrm{C}}(\mathcal{C}, \rho_{AB})$ directly by probing the channel with an entangled state [Cuevas *et al.*, 2017]. In this case, however, the trade-off is that this state will be more prone to errors in the preparation phase (and such errors are to be considered in addition to all the others already mentioned). Thus, our experimental approach will be to first perform quantum process tomography on the superposition of communication channels. With the resulting estimate of the experimental channels, we will then be able to compute $\mathcal{I}_c$ with ideal maximally-entangled states, and will use this metric to quantify the performance of the various schemes.

In our experiment, we study single-qubit channels acting on the polarization DOF, and equal superpositions of trajectories. Since, as we have seen previously, the coherence between the two trajectories is crucial, both the internal DOF and the trajectory must be fully characterized. In general, this requires two-qubit process tomography on the path (trajectory) qubit and the polarization qubit. To perform this characterization, we use heralded single photons in order to maintain the connection to the interpretation of the quantum capacity as the information transmitted per information-carrying system.

Notice that the sender only ever encodes information in the polarization DOF, whereas the receiver

must measure both the trajectory and the polarization DOFs. Hence, this is effectively a 1-to-2 qubit channel. Because of this, performing full two-qubit process tomography provides more information than is strictly required.

3.2.3 Experiment

Implementing Noisy Channels— In our experiment, we encode and transmit information in the internal polarization DOF. We induce noise on this DOF using liquid crystal waveplates (LCWP), which can rapidly implement different polarization rotations to effectively decohere the polarization state in a precise and controllable manner [Rozema *et al.*, 2014]. The LCWP retardance can be changed between 0 rad and 2π rad in approximately 100 ms by varying the applied voltage (see Section 3.5.3 for more details). Using these fast LCWPs we can change the operations on-the-fly to actively decohere the photon's polarization, in contrast to previous experiments wherein decoherence was achieved by averaging the results during the data analysis [Goswami *et al.*, 2020; Guo *et al.*, 2020]. Nevertheless, the two methods yield the same results, so we will make use of both techniques interchangeably. Physically, the noise models we study can be understood as randomly applying one of four operations ($\mathcal{I}$, X, Y, or Z). The probability of each operation to occur defines the noisy channel.

Specifically, we implement four different noisy channels. The first is a generalization of the entanglement-breaking channel $\mathcal{C}_{\text{EB}}(\rho)$ discussed above. However, in this general case, the X and Y operations are applied with probability $1-p$ and p, respectively (one recovers the $\mathcal{C}_{\text{EB}}(\rho)$ for $p = 1/2$):

$$\mathcal{C}^{p}_{XY}(\rho) = (1-p)X\rho X + pY\rho Y. \tag{3.8}$$

We also study the well-known bit-flip (BF) $\mathcal{C}^{p}_{\text{BF}}(\rho)$ and phase-flip (PF) $\mathcal{C}^{p}_{\text{PF}}(\rho)$ (or dephasing) channels:

$$\mathcal{C}^{p}_{\text{BF}}(\rho) = (1-p)\,\rho + pX\rho X, \tag{3.9a}$$
$$\mathcal{C}^{p}_{\text{PF}}(\rho) = (1-p)\,\rho + pZ\rho Z, \tag{3.9b}$$

respectively. Finally we study a depolarizing channel $\mathcal{C}^{p}_{\text{BB84}}(\rho)$, known as the BB84-channel [Smith and Smolin, 2008]:

$$\mathcal{C}^{p}_{\text{BB84}}(\rho) = (1-p)^2\rho + (1-p)\,p\,X\rho X + (1-p)\,p\,Z\rho Z + p^2\,Y\rho Y. \tag{3.10}$$

For the BB84-channel, when the noise probability is $p = 0.5$ the channel is completely depolarizing, mapping any input to the maximally-mixed state. In Section 3.5.4, we also report a numerical estimation of the performance of the three layouts in the generic case of randomly-generated channels.

To realize a single channel we use two LCWPs. The first LCWP's optic axis is set to 0°, and can thus implement either Z or the identity operation by setting the retardance to π rad or 0 rad, respectively. The second LCWP's optic axis is set to 22.5° to execute X or the identity operation, again by setting the retardance to π rad or 0 rad, respectively. When the first LCWP performs Z and the second X, the net result is Y (up to a phase). Hence, with these two LCWPs we are able to carry out all four required unitary operations, and switch between them in about 100 ms.

In light of this, a straightforward implementation would be to generate a random number from some defined distribution before a photon enters the channel, and then to set the operations accordingly. However, the net result is the same if we allow several photons to pass through the channel for each coin flip, provided that we average over a sufficiently large number of coin flips. This is advantageous, as it allows us to increase the single-photon count rate well above the switching speed of the LCWPs. In our experiment, we employ two different methods for the data acquisition. In the first, we change the applied operation every second. Since our photon rate (detected at the output, after the experiment) is of the order of 3000 Hz, this means that approximately 3000 subsequent heralded photons experience the same unitary operation (see Section 3.5.5 for more details). Our Monte Carlo simulations show that, with these numbers, 100 seconds (and 100 different operations) per measurement setting are sufficient to achieve a process fidelity (*i.e.*, the fidelity to achieve the desired noisy channel) above 99% (for details, see Section 3.5.6). In order to ensure an optimal implementation while maintaining a reasonable duration of the data-taking procedure, we used 1000 different internal configurations for our experiment, resulting in a fidelity of 99.98% per channel. In the second technique, we simply take data for each input state and each measurement setting with the LCWPs set to implement a fixed unitary operation. We then weight the data from these different configurations according to the probability distribution of the desired noise model. (This method was also demonstrated in [Goswami *et al.*, 2020; Guo *et al.*, 2020].)

Creating Superpositions of Trajectories— As shown in Fig. 3.2, we experimentally create different superpositions of trajectories by placing single photons in an equal quantum superposition of paths using a 50/50 beamsplitter. The single photons are generated with a standard type-II down-conversion source described in Fig. 3.2**d)** and in Section 3.5.7. These two paths (trajectories) are then routed through a series of LCWPs, which implement different noisy channels, in a parallel configuration (Panel **a)**), in series (Panel **b)**), or in a quantum superposition of the two alternative orders (Panel **c)**).

All three set-ups are realised through Mach-Zehnder interferometers. In the first case (Fig. 3.2**a)**), one channel is placed in each interferometer arm. In the second case (Fig. 3.2**b)**), the channels are arranged in series in both arms of the interferometer, and additional operations are performed before each channel through waveplates. Finally, the third scheme (Fig. 3.2**c)**) is accomplished using a folded Mach-Zehnder interferometer in which the two channels appear in alternating order in each

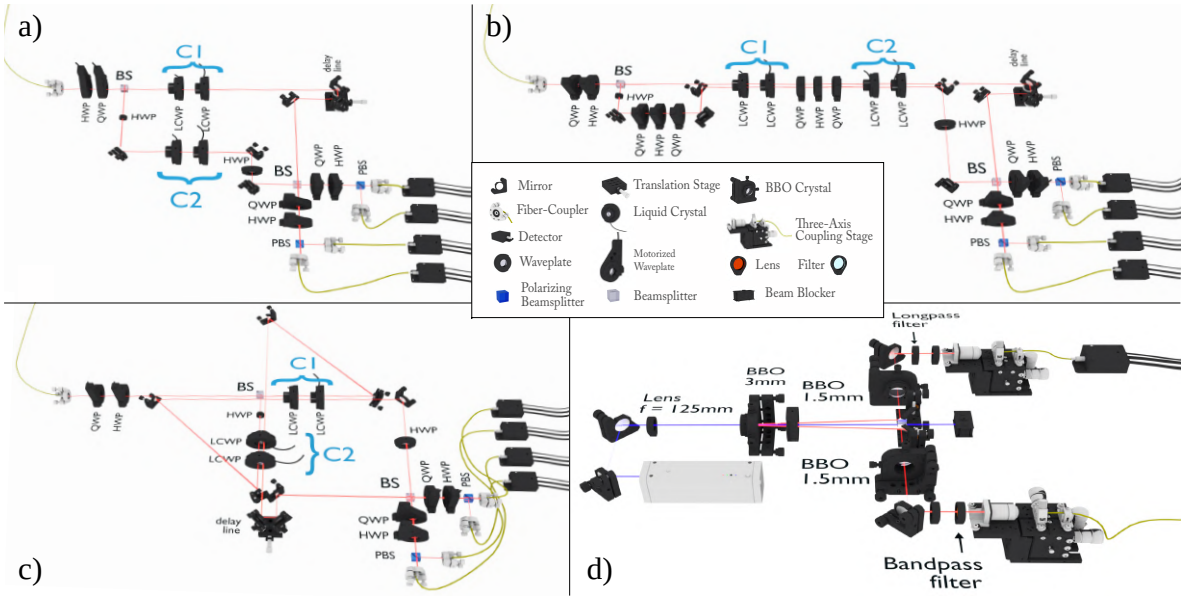


Figure 3.2 Experimental Setup. a) Quantum-Control of Parallel Channels. After their polarization is set via a half waveplate (HWP) and a quarter waveplate (QWP), single photons are injected into a Mach-Zehnder interferometer. One noisy channel is placed into each arm of the interferometer, and each channel is realized through two liquid crystal waveplates (LCWP), the first positioned at 0° (to implement $\mathcal{I}$ or Z by changing the retardance), the second at 22.5° ($\mathcal{I}$ or X). By means of a piezo-electric trombone delay line, the photon interfering on the second beamsplitter of the interferometer can be projected onto the bases $\{|+\rangle_T, |-\rangle_T\}$ or $\{|R\rangle_T, |L\rangle_T\}$ of the trajectory. Finally, the photons' polarization is measured through a sequence of QWP, HWP and a polarizing beamsplitter. **b) Channels in Series with Quantum-Controlled Operations.** As in the previous scheme, the photons are prepared in polarization via QWP and HWP and injected into a Mach-Zehnder interferometer. In this case, the two noisy channels are placed in the two superposed branches in series with the same order. Also in this case, the channels are realized through LCWPs. Furthermore, before each noisy channel, additional unitary operations are realized through sequences of QWP, HWP and QWP (before the first channel, the QWP, HWP and QWP are placed in one branch of the trajectory only, whereas between the two channels the waveplates are in both branches, since we only implement cases where $U_2 = U_3$). The rest of the setup is the same as in the previous case. **c) Quantum-Control of Channel Order.** The preparation and measurement of the photons in polarization happens as in the previous schemes, as well as the realization of the noisy channels, and the projection of the trajectory DOF. In this case, however, the Mach-Zehnder interferometer is folded into two loops so that the photon can travel through the two channels in the two alternative orders in each arm of the interferometer. **d) Heralded single-photon source.** We generate photon pairs using a type-II spontaneous-parametric-down-conversion source. One photon is directly detected with an avalanche photodiode (upper arm), whereas the other is coupled into an optical fiber and sent to one of the setups **a)**, **b)** or **c)**. The interferometers in setups **a)**, **b)**, and **c)** all contain two compensation HWP at the beginning and at the end of the reflected arm, so as to compensate for the phase shifts due to the reflection from the beamsplitter.

of the interferometer's arms. The setup presented in Fig. 3.2**c)** represents a possible realization of a quantum-optical switch wherein the system qubit is encoded in the polarization DOF, and the control qubit in the path DOF. Other encodings for this type of process have been proposed [Friis *et al.*, 2014; Rambo *et al.*, 2016] and experimentally demonstrated [Goswami *et al.*, 2018; Wei *et al.*, 2019]. Regardless of the detailed implementation, all proposals to implement a quantum-optical switch use one DOF to route a photon through channels in different orders, while the channels act on some other

DOF.

In order to perform quantum process tomography to extract the coherent information, we must prepare a tomographically-complete set of input states, and measure in a tomographically-complete number of different bases. In brief, we use waveplates before the first beamsplitter to prepare the state of the polarization qubit in either $|0\rangle_{\mathrm{I}}$, $|+\rangle_{\mathrm{I}}$, $|R\rangle_{\mathrm{I}}$, or $|L\rangle_{\mathrm{I}}$ (where $|R\rangle_{\mathrm{I}} = (|0\rangle_{\mathrm{I}} - i\,|1\rangle_{\mathrm{I}})/\sqrt{2}$, $|L\rangle_{\mathrm{I}} = (|0\rangle_{\mathrm{I}} + i\,|1\rangle_{\mathrm{I}})/\sqrt{2}$), and waveplates and polarizing beamsplitters after the second beamsplitter to measure in all bases (*i.e.*, $\{|0\rangle_{\mathrm{I}}, |1\rangle_{\mathrm{I}}\}$, $\{|+\rangle_{\mathrm{I}}, |-\rangle_{\mathrm{I}}\}$, and $\{|R\rangle_{\mathrm{I}}, |L\rangle_{\mathrm{I}}\}$). In our experiment, we set the state of the path DOF to $|+\rangle_{\mathrm{T}}$, $|-\rangle_{\mathrm{T}}$, $|R\rangle_{\mathrm{T}}$, or $|L\rangle_{\mathrm{T}}$ by varying the relative phase of the paths after the first beamsplitter using a pair of mirrors placed on a trombone-delay stage controlled by a piezo-electric actuator. To prepare the state to $|0\rangle_{\mathrm{T}}$ or $|1\rangle_{\mathrm{T}}$, we simply block one or the other path. We measure the path DOF analogously, by setting the different phases, or blocking one of the two paths. The full details of our process tomography protocol are presented in Section 3.5.8.

3.3 Results

XY-Channel— Below, we present our results for the three combinations of the noisy channels described in equations (3.8)–(3.10). We will first consider two copies of the *XY*-channel (Eq. (3.8)). In Section 3.2.1, we observed that when $p = 0.5$ both the channels in series with quantum-controlled operations and the quantum-control of channel order (with $U_1 = Y$, $U_2 = U_3 = \mathcal{I}$) are able to transmit quantum information perfectly. In Fig. 3.3, we observe that such a perfect 'activation' (in our case, the term refers to a combination of two noisy channels which enables one to communicate through such a combination with less noise than either individually) is theoretically possible for all values of p. In fact, the purple and brown lines show the coherent information for two *XY*-channels combined in series and in indefinite order, respectively. For both of these situations, the theoretical coherent information is equal to 1 for all p, meaning that one qubit per use can be transferred. In the same plot, our experimental data are presented as squares (for the quantum-control of parallel channels), circles (for the channels in series with quantum-controlled operations) and crosses (for the quantum-control of channel order) with matching colors. The dominant source of the statistical errors is the uncertainty in determining the initial states for the process tomography. (In fact, the input states were prepared and characterized at the output of the source, but they were then sent to each experiment via 3m long optical fibers, which introduced additional noise.) This uncertainty leads to the error bars on all the data sets presented in Figs. 3.3–3.5 (see Section 3.5.5 for more details on error estimation). Nevertheless, all plots display a good agreement between experiment and theory. As expected, the experimentally measured coherent information is slightly lower than that predicted theoretically. This offset is mainly due to the following systematic errors: *i.* the imperfect visibility when the two trajectories are recombined on the second beamsplitter, *ii.* phase drifts which can occur during the experimental runs, and *iii.* slight calibration errors in the LCWPs implementing the channels and the waveplates used for

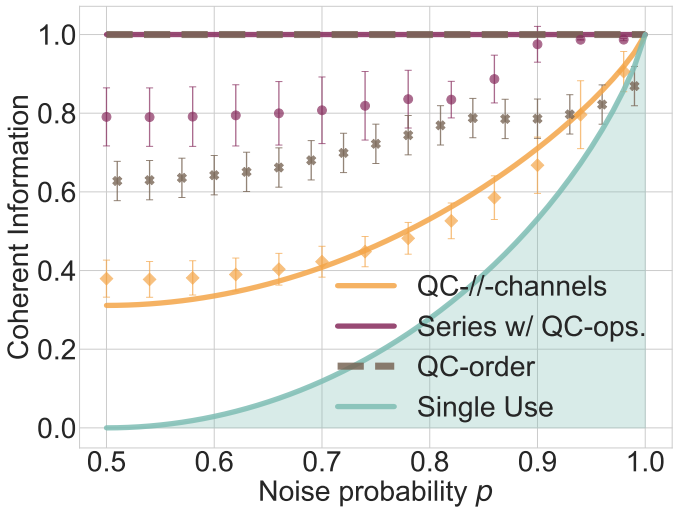


Figure 3.3 Experimental *XY*-Channel noise data. The theoretical trends associated with the channels in series with quantum-controlled operations and the quantum-control of channel order show full activation. The experimental data do not perfectly match the theoretical trends because, for $p = 0.5$, the channel produces an equal mixture of X- and Y-operations, and such case can be experimentally realised with a lower fidelity than the one in which only one of the two operations is performed (*i.e.*, when $p = 0$ or 1). It follows that, in the central region, the experimental data are further apart from the theoretical trend than they are on the upper end. The quantum-control of parallel channels does not allow full activation, and thus it is positioned below the previous two trends. In this case, the experimental data are closer to the theoretical expectation. The reason of the higher agreement is that, in the case of the disposition of noisy channels in parallel, only one channel is present in each branch of the interferometer. As a consequence, the experimental imperfections affecting each branch are smaller than in the dispositions of channels in series and in indefinite order. Finally, the coherent information associated to only one XY-channel is theoretically lower than all the other layouts. A detailed analysis of the error estimation and the systematic error is provided in Section 3.5.5. The labels 'QC-//-channels', 'Series w/ QC-ops.' and 'QC-order' stand for 'quantum-control of parallel channels', 'channels in series with quantum-controlled operations' and 'quantum-control of channel order', respectively. The same labels will be used in all plots.

state preparation and measurement. These systematic effects are not included in the calculation of our experimental errors. Full details of the measurement procedure, including photon count rates and measurement times, as well as the statistical and systematic errors affecting the data are presented in Section 3.5.5.

The orange data set reported in Fig. 3.3 corresponds to the coherent information when the two *XY*-channels are used in a quantum-controlled superposition. In Section 3.2.1, we illustrated that, when $p = 0.5$, the output still displays a partial dependence on the input state. However, calculating the coherent information reveals that this is not sufficient to transmit a single qubit per use (*i.e.*, the coherent information is less than 1). Nevertheless, the orange curve indicates that quantum information can still be transmitted, although not a the maximum rate.

The turquoise curve in Fig. 3.3 represents the coherent information of a single trajectory traversing a

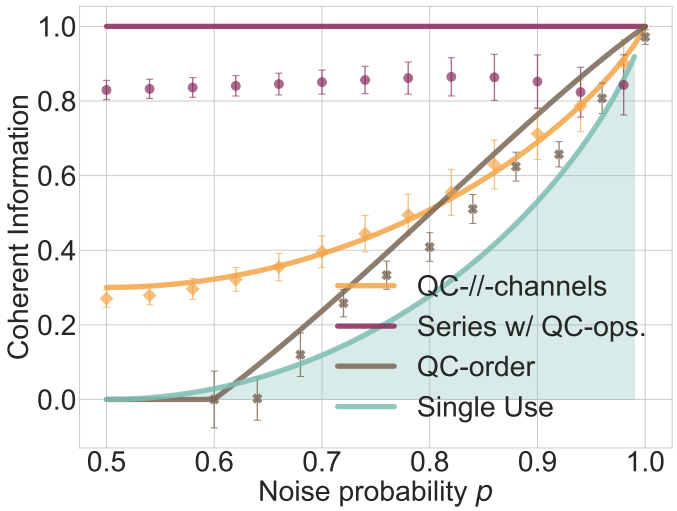


Figure 3.4 Experimental BF- and PF-noise data. The experimental data of quantum-control of parallel channels and the quantum-control of channel order are in good agreement with the theoretical trends. Conversely, the configuration of the channels in series with quantum-controlled operations shows a constant offset between the experimental data and the expected theoretical trend. This discrepancy is due to the fact that, in this case, all the liquid crystals are arranged in series, with the additional presence of waveplates realizing a Hadamard gate, and hence this configuration is the one that exhibits the greatest amount of experimental imperfections along each path. In spite of this, for most values of p the coherent information that can be achieved with the series configuration is still above all others by several standard deviations.

single copy of the channel, which is $1 - H(p)$, where $H(p) = -p\log(p) - (1-p)\log(1-p)$ is the Shannon entropy. (The shaded area underneath represents the region within which any activation by either channel layout is less effective than directly using one of the noisy channels.) Because in our experiment we assume that the noise strengths p of the two channels are always identical, using the channels in a classical mixture, as depicted in Fig 3.1**d)**, will also result in the capacity of a single use of the channel. If a single trajectory was sent through two copies of the channel in a row, the coherent information would be even lower, since the second channel would further decohere the polarization state. We see in this first case that for all values of the noise parameter p, all three superposition methods transmit more quantum information than only using a single-trajectory.

Bit-Flip and Phase-Flip Channels— Ref. [Salek *et al.*, 2018] showed that a quantum superposition of the causal order of a bit-flip and a phase-flip channel can transmit more quantum information than the amount which can travel through each channel individually. (Referring to Fig. 3.1, this corresponds to replacing channel 1 with the bit-flip channel (Eq. (3.9a)), and channel 2 with the phase-flip channel (Eq. (3.9b)). Note that, contrary to the other cases, here we consider two different types of noisy channels C1, C2, rather than two copies of the same channel). In light of this, Ref. [Guérin *et al.*, 2019] pointed out that this idea can also be applied when the noisy channels are placed in series, provided that one allows quantum-controlled operations before and between them, and that this trick

allows one to transmit quantum information perfectly (when $U_1 = Y$, $U_2 = U_3 = H$, where H is the Hadamard operation).

We experimentally confirm the predictions of Refs. [Guérin *et al.*, 2019; Salek *et al.*, 2018] in Fig. 3.4. There, we see that, regardless of the noise strength, the channels in series with quantum-controlled operations can, in theory, perfectly transmit quantum information (*i.e.*, the purple line is equal to 1). Our experimental data (purple circles) confirm this, although they do show a slight offset due to the systematic errors discussed above. In this case, the quantum-control of channel order (brown curve for theory, and crosses for experiment) does not work as well. Nonetheless, we do find that for a range of p it outperforms the single use value $1-H(p)$. For this choice of noisy channels, the quantum-control of parallel channels (orange curve) can transmit more information than their quantum-controlled order. For a large range of p, it is larger than the value achievable through the quantum-control of channel order, and the slight theoretical advantage of this latter over the quantum-control of parallel channels for large enough values of p is not observable in our experimental data.

For a fair comparison, we mention that changing the quantum-controlled operations U_i depending on the type of noise could be regarded as an additional resource. In fact, setting the optimal quantum-controlled operations requires one to characterize the noise prior to using the channels. In Section 3.5.9, we compare the performance of the channels in series with quantum-controlled operations for the same unitaries that we use for the *XY*- and BB84-channels (namely, for $U_1 = Y$, $U_2 = U_3 = \mathcal{I}$). There, we observe that setting $U_2 = U_3 = \mathcal{I}$ results in a performance that is comparable to that of the quantum-control of channel order, and which still outperforms the single-use capacity. In doing so, the quantum-controlled operations remain fixed in this configuration independently of the type of noise.

BB84-Channel— As a final example, we consider two copies of the depolarizing BB84-channel (Eq. (3.10)). These results are shown in Fig. 3.5. Also in this case, the channels in series with quantum-controlled operations (this time with $U_1 = Y$, $U_2 = U_3 = \mathcal{I}$), shown in purple, achieves the largest enhancement. While with only two trajectories it is not possible to perfectly transmit quantum information through these noisy channels, Ref. [Guérin *et al.*, 2019] showed that with additional trajectories any type of noise can be perfectly corrected with the quantum superposition of channels in series. The quantum-control of channel order in this case outperforms both the single-use coherent information ($1 - 2H(p)$) and the coherent information of the quantum-control of parallel channels.

These three examples show that, depending on the type of noise, different superpositions of channels can lead to the ability to transmit different amounts of quantum information. The physical origin of this ability is an effective coupling between the trajectory and the internal degree of freedom. In the present paper, this coupling is verified by the observed correlations between the states of the aforementioned two degrees of freedom. While these correlations were only sketched in the case of the EB channel in sections 3.2.1.1-3.2.1.3, analogous relations hold also in the case of the other noisy

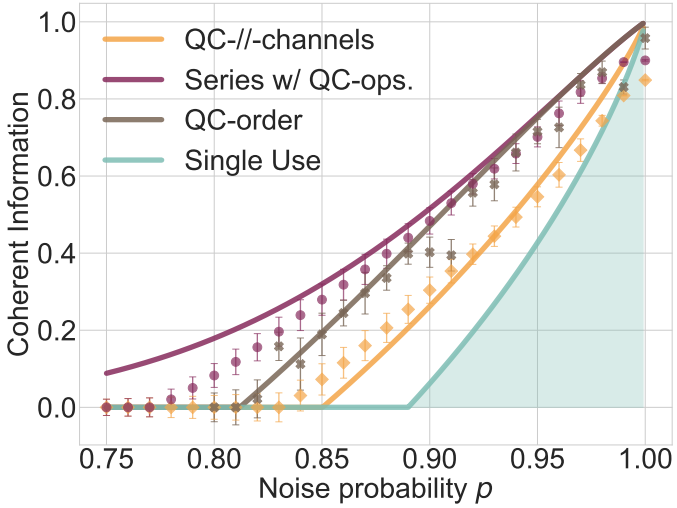


Figure 3.5 Experimental BB84-channel noise data. As in the previous plots, the continuous lines show the expected theoretical trends, while the squares, circles and crosses represent the experimental data corresponding to the quantum-control of parallel channels, the channels in series with quantum-controlled operations, and the quantum-control of channel order, respectively. All the experimental data are in high agreement with the expected theoretical trends.

channels studied in this section. In all the cases we investigated here (wherein the schemes are used individually), even in the presence of experimental imperfections, using the channels in series with quantum-controlled operations appears to be the best candidate to evade the effects of the noise.

3.4 Discussion

In this work, we experimentally and numerically explored how the degradation of quantum information due to its propagation through noisy channels can be mitigated, and in several cases fully suppressed. This was achieved by sending quantum information carriers through a pair of noisy channels in various superpositions of trajectories. In particular, we studied three types of schemes: the quantum-control of parallel noisy channels, channels in series with quantum-controlled operations, and the quantum-control of channel order.

All of these schemes bear much in common with error filtration [Gisin *et al.*, 2005]. More recently, this has been refined in a number of theoretical works [Abbott *et al.*, 2020; Caleffi and Cacciapuoti, 2020; Chiribella *et al.*, 2018; Chiribella and Kristjánsson, 2019; Ebler *et al.*, 2018; Guérin *et al.*, 2019; Jia and Costa, 2019; Kristjánsson *et al.*, 2020; Loizeau and Grinbaum, 2020; Procopio *et al.*, 2020; Salek *et al.*, 2018; Sazim *et al.*, 2020; Wilson *et al.*, 2020], tied into the concept of indefinite causal orders. While enhanced communication based on an indefinite causal order has been

experimentally demonstrated [Goswami *et al.*, 2020; Guo *et al.*, 2020], an experimental study comparing different superpositions of trajectories in presence of various types of noise has been lacking. Our work is aimed at bridging this gap, by suggesting common ground based on the experimental resources that each of the analysed schemes requires.

Our results suggest that, in most quantum-optical communication scenarios, creating a superposition of trajectories through channels in series with quantum-controlled operations should lead to the largest noise reduction. One can easily imagine characterizing the error introduced in various communication channels, and from there setting the unitary operations accordingly. Moreover, Ref. [Guérin *et al.*, 2019] illustrated that these types of schemes can be extended to superpositions of more than two trajectories to achieve complete error cancellation for any type of noise. We have shown experimentally that with only two trajectories it is already possible to completely cancel (after accounting for experimental errors) all the noise arising from two out of the three types of noisy channels we considered. Furthermore, the quantum-controlled operations could also be introduced in the other two schemes and, potentially, they could match the performance of the layout with channels in series.

The large experimental communication enhancements presented here highlight the practical relevance of extending the quantum communication paradigm to scenarios in which not only the information carriers, but also the trajectories along which they propagate are quantum. We expect that the relative ease of implementation of these schemes will enable them to be readily put into practice for the noise-reduction of real-world long-distance quantum communication applications.

3.5 Methods

3.5.1 Communication Advantages when Transmitting Classical Information

The present work focuses on the transmission of quantum information through channels placed in a quantum superpositions of trajectories. This is, however, not the only possible choice: the work which initiated this research direction discussed communication advantages in transmitting classical information through a quantum-control of channel order [Ebler *et al.*, 2018]. In the following, we briefly comment on our choice of figure of merit, and we explain how the scenarios in Fig. 3.1 compare for transmitting classical information.

Classical communication is determined by the amount of classical correlations which a quantum channel can maintain between its input and output. Generally, the classical capacity of a quantum channel is only zero for the class of erasure channels which replaces any input by a fixed output state. Sending quantum information, on the other hand, is a more ambitious task: here, the capability of a channel to preserve quantum correlations (*i.e.*, entanglement) during the processing quantifies the

transmissible quantum information. Indeed, quantum information was shown to be the most difficult to communicate [Wilde, 2013], and at the same time the most valuable resource for information theoretic tasks and computation. For instance, sending quantum information can ensure secure communication [Bennett and Brassard, 1984], it can be used to distill secret keys for cryptography [Wilde, 2013], and it is crucial for tasks like distributed quantum computation [Beals *et al.*, 2013; Broadbent *et al.*, 2009]. Furthermore, trivially, quantum information can also be used to communicate classical information [Holevo, 1973]. Hence, in this work we focus on advantages for the most difficult type of information transmission.

Comparing the superposition schemes from Fig. 3.1 in the case of a classical communication yields an analogous behavior to the case of quantum communication. It was already shown in Appendix G of Ref. [Gisin *et al.*, 2005], that superpositions of quantum channels [Fig. 3.1**a)**] yield at least the amount of transmissible classical information of classical trajectories [Fig. 3.1**d)**]. The same holds true for the advantages through the quantum-control of channel order [Fig. 3.1**c)**] considered in the initial work [Ebler *et al.*, 2018]. Later, Ref. [Abbott *et al.*, 2020] showed that quantum-control of parallel channels [Fig. 3.1**a)**] can outperform the quantum-control of channel order [Fig. 3.1**c)**] for certain communication tasks. Numerical simulations support generality of this claim for classical information by comparing the two scenarios for random channels [Loizeau and Grinbaum, 2020]. Finally, quantum controlled sequences of channels [Fig. 3.1**b)**] was found to allow for the highest classical communication rates, as it has a larger set of allowed encoding schemes, which allow for phase kickbacks that cause partial information exchange with the trajectory degree of freedom [Guérin *et al.*, 2019].

3.5.2 Case Study: Activation of the EB-Channels in the Three Layouts

In this section, we briefly evaluate the output state of the three superposition techniques for the noisy channel described by Eq. (3.1). We carry out our study by interpreting the action of the two channels as follows. Each channel randomly applies either X or Y with probability $1/2$. Hence, the 'internal configuration' of the superposition can be understood as either *i.* both channels 1 and 2 implement X, *ii.* channel 1 implements X, while channel 2 implements Y, *iii.* channel 1 implements Y, whereas channel 2 implements X, or *iv.* both channels 1 and 2 implement Y. The final output state will then be a mixture of the output states in these four configurations, each with probability $1/4$.

Throughout this section, we will assume that the input state is $|\psi\rangle_{\mathrm{I}} |+\rangle_{\mathrm{T}}$, where $|\psi\rangle = \alpha |0\rangle + \beta |1\rangle$, and I (T) labels the internal (trajectory) DOF.

Quantum-Control of Parallel Channels— In this scheme, one noisy channel is placed in each trajectory. The action of the two channels in parallel can be interpreted as creating a mixture of the

following four unnormalised states at the output:

$$X|\psi\rangle_{\mathrm{I}}|0\rangle_{\mathrm{T}} + X|\psi\rangle_{\mathrm{I}}|1\rangle_{\mathrm{T}} \tag{3.11a}$$
$$X|\psi\rangle_{\mathrm{I}}|0\rangle_{\mathrm{T}} + Y|\psi\rangle_{\mathrm{I}}|1\rangle_{\mathrm{T}} \tag{3.11b}$$
$$Y|\psi\rangle_{\mathrm{I}}|0\rangle_{\mathrm{T}} + X|\psi\rangle_{\mathrm{I}}|1\rangle_{\mathrm{T}} \tag{3.11c}$$
$$Y|\psi\rangle_{\mathrm{I}}|0\rangle_{\mathrm{T}} + Y|\psi\rangle_{\mathrm{I}}|1\rangle_{\mathrm{T}}\,, \tag{3.11d}$$

where the four states correspond to internal configurations of $X-X$, $X-Y$, $Y-X$, and $Y-Y$, for channel 1 and channel 2, respectively. These states can be rewritten (up to further normalisation) as:

$$X|\psi\rangle_{\mathrm{I}}|+\rangle_{\mathrm{T}} \tag{3.12a}$$
$$(X|\psi\rangle_{\mathrm{I}} + Y|\psi\rangle_{\mathrm{I}})|+\rangle_{\mathrm{T}} + (X|\psi\rangle_{\mathrm{I}} - Y|\psi\rangle_{\mathrm{I}})|-\rangle_{\mathrm{T}} \tag{3.12b}$$
$$(X|\psi\rangle_{\mathrm{I}} + Y|\psi\rangle_{\mathrm{I}})|+\rangle_{\mathrm{T}} - (X|\psi\rangle_{\mathrm{I}} - Y|\psi\rangle_{\mathrm{I}})|-\rangle_{\mathrm{T}} \tag{3.12c}$$
$$Y|\psi\rangle_{\mathrm{I}}|+\rangle_{\mathrm{T}}\,. \tag{3.12d}$$

Now, measuring the trajectory DOF in the $\{|+\rangle_{\mathrm{T}}, |-\rangle_{\mathrm{T}}\}$ basis and obtaining $|-\rangle_{\mathrm{T}}$ (which happens with probability 1/4) projects the internal DOF into $X|\psi\rangle_{\mathrm{I}} - Y|\psi\rangle_{\mathrm{I}}$, which can be rewritten (after renormalisation and up to an irrelevant global phase) as in Eq. (3.3). This is a pure state, which implies that some ability to transmit quantum information has been restored in post-selection. If, on the other hand, one obtains the result $|+\rangle_{\mathrm{T}}$ (with probability 3/4), it is straightforward to show that the internal DOF is projected in the mixed state described by Eq. (3.4).

In the recent papers on superpositions of trajectories [Abbott *et al.*, 2020; Chiribella and Kristjánsson, 2019], it was shown that the output of a quantum-controlled superposition of two channels depends on additional parameters related to the physical realisation of the channels ('transformation matrices' in [Abbott *et al.*, 2020] and 'vacuum amplitudes' in [Chiribella and Kristjánsson, 2019]). In our scheme, these additional parameters reduce to the relative phase between the vacuum and the single-photon subspace of the unitary operations (*e.g.*, the Pauli-X and -Y from above, with transformation matrix $\Gamma = (X+Y)/2$). More precisely, the vacuum extension of a qubit unitary U is $U' = e^{i\phi}|\text{vacuum}\rangle\langle\text{vacuum}| + U$, where U acts in the single-photon subspace. In the calculation above, the phase is implicitly set to zero, which is in agreement with our experiment.

Channels in Series with Quantum-Controlled Operations— Let us now consider the action of the superposition of trajectories in series with the quantum-controlled operations (Fig. 3.1**b)**), with $U_1 = Y$, $U_2 = \mathcal{I}$, $U_3 = \mathcal{I}$. In this case, the input state is transformed into $(|\psi\rangle_{\mathrm{I}}|0\rangle_{\mathrm{T}} + Y|\psi\rangle_{\mathrm{I}}|1\rangle_{\mathrm{T}})/\sqrt{2}$ before interacting with the noisy channels.

Again, we can compute the four effective unnormalised states which arise from the different internal

configurations of the noisy channels:

$$XX\left|\psi\right\rangle_{\mathrm{I}}\left|0\right\rangle_{\mathrm{T}}+XXY\left|\psi\right\rangle_{\mathrm{I}}\left|1\right\rangle_{\mathrm{T}} \tag{3.13a}$$

$$YX\left|\psi\right\rangle_{\mathrm{I}}\left|0\right\rangle_{\mathrm{T}}+YXY\left|\psi\right\rangle_{\mathrm{I}}\left|1\right\rangle_{\mathrm{T}} \tag{3.13b}$$

$$XY\left|\psi\right\rangle_{\mathrm{I}}\left|0\right\rangle_{\mathrm{T}}+XYY\left|\psi\right\rangle_{\mathrm{I}}\left|1\right\rangle_{\mathrm{T}} \tag{3.13c}$$

$$YY\left|\psi\right\rangle_{\mathrm{I}}\left|0\right\rangle_{\mathrm{T}}+YYY\left|\psi\right\rangle_{\mathrm{I}}\left|1\right\rangle_{\mathrm{T}}. \tag{3.13d}$$

The order of the above states refers to internal configurations $X-X$, $X-Y$, $Y-X$, and $Y-Y$. These states can be rewritten (up to phases) as:

$$\left|0\right\rangle_{\mathrm{I}}\left|\psi'\right\rangle_{\mathrm{T}}+i\left|1\right\rangle_{\mathrm{I}}X\left|\psi'\right\rangle_{\mathrm{T}} \tag{3.14a}$$

$$\left|0\right\rangle_{\mathrm{I}}\left|\psi'\right\rangle_{\mathrm{T}}-i\left|1\right\rangle_{\mathrm{I}}X\left|\psi'\right\rangle_{\mathrm{T}} \tag{3.14b}$$

$$\left|0\right\rangle_{\mathrm{I}}\left|\psi'\right\rangle_{\mathrm{T}}-i\left|1\right\rangle_{\mathrm{I}}X\left|\psi'\right\rangle_{\mathrm{T}} \tag{3.14c}$$

$$\left|0\right\rangle_{\mathrm{I}}\left|\psi'\right\rangle_{\mathrm{T}}+i\left|1\right\rangle_{\mathrm{I}}X\left|\psi'\right\rangle_{\mathrm{T}}, \tag{3.14d}$$

where $\left|\psi'\right\rangle=\alpha\left|0\right\rangle-i\beta\left|1\right\rangle$. As a result, we see that measuring the internal DOF in the computational basis $\{\left|0\right\rangle_{\mathrm{I}},\left|1\right\rangle_{\mathrm{I}}\}$ projects the trajectory into either $\left|\psi'\right\rangle_{\mathrm{T}}$ or $X\left|\psi'\right\rangle_{\mathrm{T}}$, upon obtaining outcomes $\left|0\right\rangle_{\mathrm{I}}$ or $\left|1\right\rangle_{\mathrm{I}}$, respectively (each with equal probabilities). Both of these states can be unitarily corrected, allowing one to achieve perfect quantum information transfer through these channels. (Even though here, for simplicity, we restricted ourselves to the case $p=1/2$, the same reasoning applies whatever the mixing probability p in the definition of the channels, Eq. (3.8)).

Quantum-Control of Channel Order— For this scheme, we make use of the fact that the output of the quantum switch for unitary operations A and B is

$$\frac{1}{2}\{A,B\}\left|\psi\right\rangle_{\mathrm{I}}\left|+\right\rangle_{\mathrm{T}}-\frac{1}{2}[A,B]\left|\psi\right\rangle_{\mathrm{I}}\left|-\right\rangle_{\mathrm{T}}, \tag{3.15}$$

where $[A,B]$ is the commutator of A and B, and $\{A,B\}$ is their anti-commutator. It is then easy to notice that the four output states (up to phases) are

$$\left|\psi\right\rangle_{\mathrm{I}}\left|+\right\rangle_{\mathrm{T}} \tag{3.16a}$$

$$Z\left|\psi\right\rangle_{\mathrm{I}}\left|-\right\rangle_{\mathrm{T}} \tag{3.16b}$$

$$Z\left|\psi\right\rangle_{\mathrm{I}}\left|-\right\rangle_{\mathrm{T}} \tag{3.16c}$$

$$\left|\psi\right\rangle_{\mathrm{I}}\left|+\right\rangle_{\mathrm{T}}. \tag{3.16d}$$

Again, the order of the above states refers to the internal configurations $X-X$, $X-Y$, $Y-X$, and $Y-Y$. This leads to the mixture described by Eq. (3.5), and it implies that measuring the trajectory in the $\{\left|+\right\rangle_{\mathrm{T}},\left|-\right\rangle_{\mathrm{T}}\}$ basis projects the internal DOF into either $\left|\psi\right\rangle_{\mathrm{I}}$, or $Z\left|\psi\right\rangle_{\mathrm{I}}$. (As above, the same

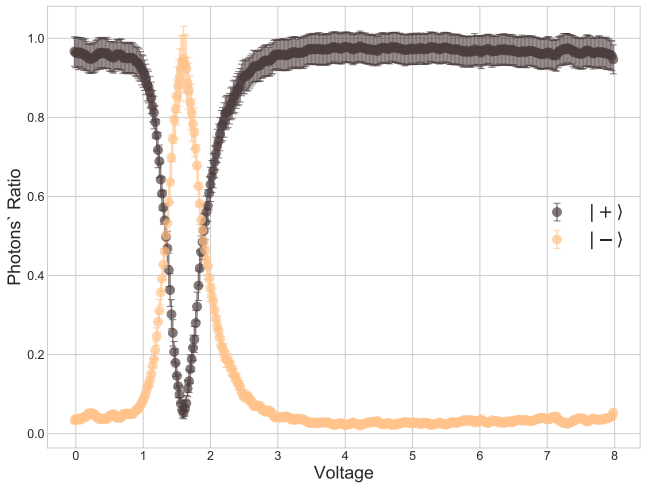


Figure 3.6 Experimental characterization of a liquid crystal waveplate (LCWP) at $0°$. Since the crystal is positioned at $0°$, it will be able to switch from an identity operation to a Pauli-Z. To characterize the voltage corresponding to a Pauli-Z, we send through it photons in the polarization basis $\{|\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}\}$, and we measure for which voltage the population inversion occurs. The estimated errors are Poissonian.

reasoning also applies whatever the mixing probability p in Eq. (3.8).)

3.5.3 Liquid Crystals Characterization

In essence, a liquid crystal waveplate (LCWP) can be understood as a standard crystalline retarder whose amount of retardance can be continuously varied by applying a voltage. Fig. 3.6 shows the characterization of one of our LCWPs. In our experiment, we used Meadowlark Liquid Crystal Variable Retarders. Their beam deviation is estimated to 2 arc min, their reflectance (per surface) is 0.5%, and their surface quality is 40-20 scratch and dig.

3.5.4 Numerical Comparison for Random Channels

To further compare the three schemes, we present a numerical evaluation of the coherent information which can be achieved with each channel layout for a large set of randomly-generated channels. The numerical procedure is carried out as follows. First, we randomly generate a quantum completely-positive and trace-preserving (CPTP) channel using the `quantinf MATLAB` package available at http://www.dr qubit.org/matlab.html. (The package uses the routine outlined in [3].) Then,

[3]We made use of the `randChan` function of the `quantinf MATLAB` package, which returns a randomly generated 2×2 quantum channel in the Kraus representation. It achieves this by generating a random isometry, and converting it onto the Kraus representation. Below, we provide further details on how these two tasks are achieved.

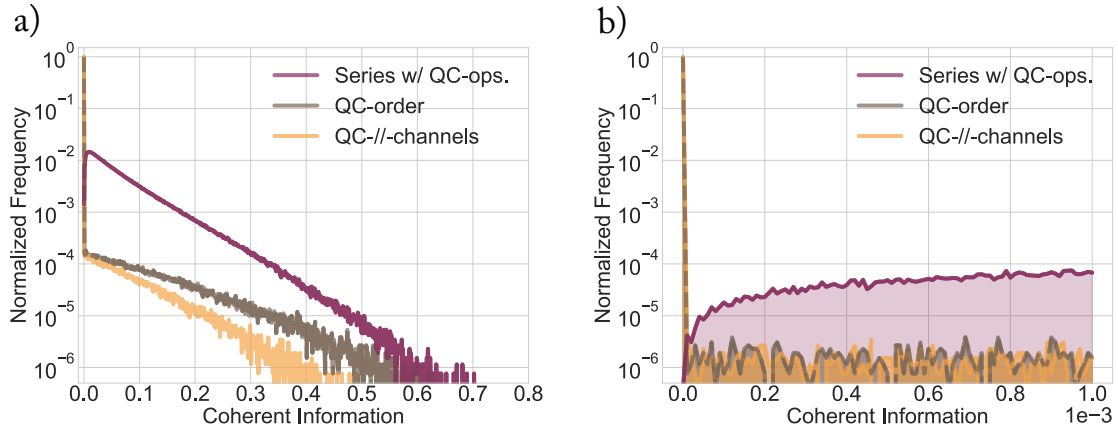


Figure 3.7 Histogram of the coherent information achieved with the three channel layouts when the two copies of the same randomly-generated channel is used in each of the three layouts. The histograms report the frequency with which a random channel (y-axis, in logarithmic scale) yields a given amount of coherent information (x-axis), normalized to the total number of channels used. **a) Histogram with 10^3 bins between a coherent information of 0 and 0.85.** As can be seen, the configuration of channels in series with quantum-controlled operations consistently achieves the highest coherent information on average. **b) Histogram of the same data with 10^5 bins displayed for values of coherent information from 0 to 0.001.** By increasing the resolution for small values of coherent information, it is possible to observe in greater detail the absence of the peak at zero for the quantum superposition of channels in series with quantum-controlled operations. In this region, the performance of the quantum-control of parallel channels and that of quantum-control of channel order is comparable.

we estimate the coherent information when *i.* two copies of the same channel (Fig. 3.7) or *ii.* two different randomly-generated channels (Fig. 3.8) are inserted in the three types of quantum superpositions of trajectories. The coherent information of the resulting superposition is then estimated using a maximally entangled state $|\Phi^+\rangle$ as input [4]. For the configuration of quantum-control of parallel channels, as shown in [Abbott *et al.*, 2020], the output state depends not only on the CPTP map, but also on the specific implementation of the channel. We therefore discuss three methods to generate different implementations for each randomly generated CPTP map, but we present results only from the one corresponding to the experimental implementation reported in the main text. For the configuration of channels in series with quantum-controlled operations, we set $U_1 = Y$, and $U_2 = U_3 = \mathcal{I}$. Further details on our numerical procedure are reported in Subsection 3.5.4.1. Histograms of our results with respect to the coherent information are presented in Figs. 3.7-3.8, for 2.6×10^6 iterations (Fig. 3.7)

Step 1) The program generates a random isometry in the following way: *Step 1.1)* it generates two 2×2 matrices, A and B, of real floating-point random entries drawn from a standard normal distribution (*i.e.*, $\mu = 0$, $\sigma = 1$, being μ the mean of the distribution, and σ its variance) using the `randn` function; *Step 1.2)* it builds $C = (A + iB)/\sqrt{2}$; *Step 1.3)* it performs an economy-size QR decomposition such that $C = Q * R$ (where Q is an orthogonal matrix, and R is an upper triangular matrix); *Step 1.4)* it diagonalises and normalises the matrix R into $R' = \mathrm{diag}\big(\mathrm{diag}(R)/|\mathrm{diag}(R)|\big)$; *Step 1.5* it creates a new matrix $V = Q * R'$, which is the desired isometry.

Step 2) The program converts the isometry V into the Kraus representation by left-multiplying it by states in the computational basis.

[4] As before, the choice of a maximally-entangled state lower bounds the quantum capacity, and numerical simulations suggest that such states maximise the amount of coherent information of a given channel.

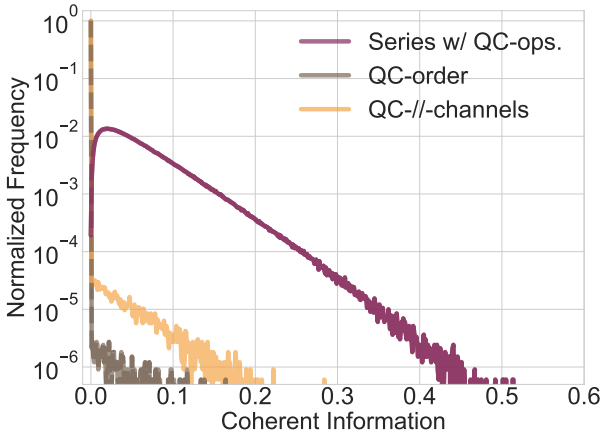


Figure 3.8 Histogram of the coherent information achieved with the three channel layouts when two different randomly-generated channels are used. The overall trend here is comparable to that of two copies of the same random channel (Fig. 3.7). However, in this case, the quantum-control of the parallel channels performs, on average, better than the quantum-control of channel order. Moreover, in general, all three layouts tend to perform worse than in the case of two copies of the same random noisy channel (*i.e.*, the maximum amount of coherent information which can be achieved through each layout is generally lower than in the case shown in Fig. 3.7).

and 3.3×10^6 iterations (Fig. 3.8). In essence, this can be interpreted as the probability to obtain a given value of coherent information with each of the three layouts. We observe that, on average, channels in series with quantum-controlled operations achieve a better performance than the other two methods, while the quantum-control of parallel channels (quantum-control of channel order) exhibits the lowest performance when operated with two copies of the same channel (two different randomly-generated channels). We also note that, in many cases, both the quantum-control of parallel channels and the quantum-control of channel order fail to obtain any activation, leading to large peaks at zero in each of their histograms. Interestingly, this peak is not present in the case of the layout in series with quantum-controlled operations. This suggests that an activation (albeit small) of the noisy channels can always be achieved using this layout with only two trajectories.

Finally, Fig. 3.9 shows a histogram, wherein the difference between the coherent information of the quantum superposition of channels in series with quantum-controlled operations and that of the quantum-control of parallel channels ($\mathrm{CI}_{\text{Series w/ QC-ops.}} - \mathrm{CI}_{\text{QC-//-channels}}$) and of quantum-control of channel order ($\mathrm{CI}_{\text{Series w/ QC-ops.}} - \mathrm{CI}_{\text{QC-order}}$) is plotted for each random pair of channels. Generally, the quantum superposition of channels in series with quantum-controlled operations can achieve coherent information values higher than two other layouts. However, the negative values in the histograms show that this is not always the case. This is in line with what illustrated in Section 3.5.9, where we highlight the fact that, if the unitaries U_1, U_2 and U_3 were not optimised for given noisy channels, a higher coherent information might be obtainable with the other layouts. While Ref. [Guérin *et al.*, 2019] proved that, by superposing a larger number of trajectories, one can always find an optimal

choice of quantum-controlled unitaries which can outperform the other two channels' layouts, we leave it as an open question whether or not this is also true for the restricted case of two trajectories only.

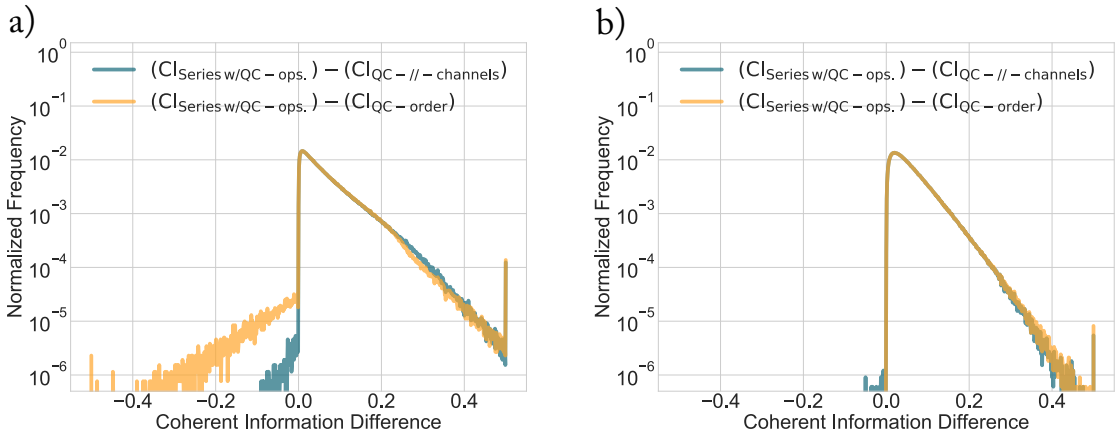


Figure 3.9 Histograms of the difference between coherent information achievable with quantum superposition of channels in series with quantum-controlled operations and the other two layouts in the case of a) two independent copies of the same random channel, and b) two different randomly-generated channels. The histograms show, for each random channel, the difference between the coherent information of the quantum superposition of channels in series with quantum-controlled operations and that of the quantum-control of parallel channels $[(\mathrm{CI}_{\text{Series w/ QC-ops.}}) - (\mathrm{CI}_{\text{QC-//-channels}})]$ and of quantum-control of channel order $[(\mathrm{CI}_{\text{Series w/ QC-ops.}}) - (\mathrm{CI}_{\text{QC-order}})]$. While, to a large extent, the layout using the channels in series with quantum-controlled operations tends to outperform the other two schemes, the negative values indicate that this is not always the case.

3.5.4.1 Summary of Numerics

In this subsection, we provide further details on how we constructed the output states in the three schemes. We start by randomly generating two single-qubit channels, using the Kraus decomposition. Since any qubit channel has a Kraus decomposition with 4 operators or less, this results in two sets of Kraus operators $\{A_0, A_1, A_2, A_3\}$ and $\{B_0, B_1, B_2, B_3\}$. We will always use the two-qubit Bell state $|\Phi^+\rangle$ to probe the channel and calculate the coherent information. Then, the full three-qubit state we consider is given by:

$$|\psi^{\text{in}}\rangle_{\text{T,I,H}} = |+\rangle_{\text{T}} \otimes |\Phi^+\rangle_{\text{I,H}}\,, \tag{3.17}$$

where T is the trajectory qubit, I is the system which will experience the noisy channel (information qubit), and H is the auxiliary (hypothetical) qubit used to evaluate the coherent information. We will use $A_i^{(\text{I})} = A_i \otimes \mathcal{I}$ as shorthand, where A_i acts on the state of the information qubit and $\mathcal{I}$ on that of the auxiliary qubit.

Quantum-Control of Parallel Channels— We construct the output state, following Ref. [Abbott *et al.*,

2020], as

$$\begin{aligned}\rho^{\text{out}}_{\text{T,I,H}} =& \frac{1}{2}\left[|0\rangle\langle 0|_{\text{T}} \otimes \mathcal{C}_A(\rho^{\text{in}}_{\text{I,H}}) + |1\rangle\langle 1|_{\text{T}} \otimes \mathcal{C}_B(\rho^{\text{in}}_{\text{I,H}})\right] \\ &+ \frac{1}{2}\left[|0\rangle\langle 1|_{\text{T}} \otimes \Gamma_A\, \rho^{\text{in}}_{\text{I,H}}\, \Gamma_B^\dagger + |1\rangle\langle 0|_{\text{T}} \otimes \Gamma_B\, \rho^{\text{in}}_{\text{I,H}}\, \Gamma_A^\dagger\right],\end{aligned} \tag{3.18}$$

where $\rho^{\text{in}}_{\text{I,H}} = |\Phi^+\rangle\langle\Phi^+|_{\text{I,H}}$. Here, $\mathcal{C}_A$ and $\mathcal{C}_B$ are the application of either channel

$$\mathcal{C}_\Xi(\rho^{\text{in}}_{\text{I,H}}) = \sum_{i=0}^{3} \Xi_i\, \rho^{\text{in}}_{\text{I,H}}\, \Xi_i^\dagger \tag{3.19}$$

with $\Xi_i = A_i, B_i$, and the transformation matrices Γ_Ξ are

$$\Gamma_\Xi = \sum_{i=0}^{3} \langle E_\Xi|i\rangle\, \Xi_i, \tag{3.20}$$

where $\{|i\rangle\}$ are orthogonal states of the environment. The transformation matrices Γ_Ξ are related to a specific purification of the channels, and they depend on the initial states of the environment $|E_A\rangle$ and $|E_B\rangle$ used in this purification. These states will be given by the actual physical implementation of the channel, and they can lead to different activations using the quantum-controlled channels. Given some Kraus representations of the channels, we numerically investigated three different states of the environment. First, as used in Ref. [Abbott *et al.*, 2020], we set $|E_A\rangle = |E_B\rangle = \frac{1}{2}\sum_{i=0}^{3}|i\rangle$. Second, we generate $|E_A\rangle$ and $|E_B\rangle$ randomly from the Haar measure for each different channel. In this case we do not optimise over $|E_A\rangle$ and $|E_B\rangle$, we simply take one random state for each. Finally, we set the weights of the environment based on the randomly-generated channel as:

$$|E_\Xi\rangle = \sum_{i=0}^{3} \sqrt{w_i^{(\Xi)}}\, |i\rangle\,, \tag{3.21}$$

where

$$w_i^{(\Xi)} = \text{Tr}\Big(\Xi_i^{(\text{I})} \rho^{\text{in}}_{\text{I,H}} \Xi_i^{(\text{I})\,\dagger}\Big), \tag{3.22}$$

which, in our case of a maximally entangled input state, reduces to $\text{Tr}(\Xi_i^{(\text{I})}\Xi_i^{(\text{I})\,\dagger})/2$.

The appropriate method to generate the states of the environment depends on the physical realisation of the channels. Within our framework, the description of quantum-control of parallel channels given in the main text, for the channels we realised experimentally, coincides with the third option, with $\Xi_i = w_i\sigma_i$, where σ_i is a Pauli unitary, and the weights w_i are given by the coefficients in Eqs. (3.8)-(3.10). Correspondingly, we present the results for this case in Figs. 3.7-3.9 [5]. Moreover, on average

[5] Although an optimisation over all environmental states may produce higher values of the coherent information, the choice of environmental states presented in Figs. 3.7-3.9 corresponds to the one which was experimentally realised with our photonic Mach-Zehnder interferometer without further control over the environment or the noisy channels. To adhere to the spirit of experimentally comparing schemes pursued in this Chapter, we thus chose not to optimise explicitly over

the final method (*i.e.*, setting the state of the environment based on the Kraus operators) performs the best among the aforementioned three. This can perhaps be explained by the fact that, compared to the other ones, this method generally leads to a larger norm of the transformation matrices Γ_{Ξ}, which is crucial for the communication advantages [Abbott *et al.*, 2020; Kristjánsson *et al.*, 2020; Kristjánsson *et al.*, 2020].

Channels in Series with Quantum-Controlled Operations— For the numerical evaluation of this layout, we must also include in the description the controlled-unitaries which are applied before the noisy channels:

$$\text{C-U} = |0\rangle\langle 0|_{\text{T}} \otimes \mathcal{I} + |1\rangle\langle 1|_{\text{T}} \otimes U. \tag{3.23}$$

For all of the numerical results presented here, we set $U = Y$, where Y is the Pauli-Y operator. Then, we construct 16 combined Kraus operators

$$K_{i,j} = B_j^{(\text{I})} A_i^{(\text{I})}, \tag{3.24}$$

and when, *e.g.*, the initial trajectory state is $|+\rangle_{\text{T}}$, we compute the output state as

$$\rho_{\text{T,I,H}}^{\text{out}} = \frac{1}{2} \sum_{k,l} U^k |k\rangle\langle l|_{\text{T}} (U^\dagger)^l \otimes \sum_{i,j} K_{i,j} \rho_{\text{I,H}}^{\text{in}} K_{i,j}^\dagger, \tag{3.25}$$

where we used the notation according to which $U^0 = \mathcal{I}$.

Quantum-Control of Channel Order— For the switch we will use a simplification. We know that for a given pair of Kraus operators, the output state when the trajectory is prepared in $|+\rangle_{\text{T}}$ and the input is one half of the maximally entangled state $|\Phi^+\rangle_{\text{I,H}}$ is:

$$\rho_{i,j} = \{A_i^{(\text{I})}, B_j^{(\text{I})}\} \rho_{\text{T,I,H}}^{+} \{A_i^{(\text{I})}, B_j^{(\text{I})}\}^\dagger + [A_i^{(\text{I})}, B_j^{(\text{I})}] \rho_{\text{T,I,H}}^{-} [A_i^{(\text{I})}, B_j^{(\text{I})}]^\dagger, \tag{3.26}$$

where $\{A_i^{(\text{I})}, B_j^{(\text{I})}\}$ and $[A_i^{(\text{I})}, B_j^{(\text{I})}]$ are the anti-commutator and commutator of $A_i^{(\text{I})}$ and $B_j^{(\text{I})}$, and $\rho_{\text{T,I,H}}^{+} = |+\rangle\langle +|_{\text{T}} \otimes |\Phi^+\rangle\langle\Phi^+|_{\text{I,H}}$ and $\rho_{\text{T,I,H}}^{-} = |-\rangle\langle -|_{\text{T}} \otimes |\Phi^+\rangle\langle\Phi^+|_{\text{I,H}}$. Then the net output state is simply

$$\rho_{\text{T,I,H}}^{\text{out}} = \sum_{i,j} \rho_{i,j}. \tag{3.27}$$

From these output states we then evaluate the coherent information as described in the main text.

3.5.5 Data Acquisition and Error Estimation

Below, we briefly outline the details of the data acquisition and the error estimation in our experiment.

all implementations of the randomly chosen channels (following the approach, *e.g.*, of Ref. [Abbott *et al.*, 2020]).

As discussed above, we followed two methods to experimentally construct of the noisy channels. In the first, we realized the noise within each channel by generating random numbers in the range $[0, 1]$. Based on this number, and on the type of noise we wanted to realize (Eqs. (3.8)–(3.10)), we assigned a unitary operation from the set $\{\mathcal{I}, X, Y, Z\}$. In order to ensure a high fidelity of the noise channel ($> 99\%$), we repeated this procedure 1000 times, measuring each configuration for 1s, and integrating the data taking procedure over these 1000 runs. In the second method, we measured all the possible combinations of unitary operations between the two noisy channels 1 and 2, and we then created the desired noise during our data analysis, following the procedure proposed in Ref. [Goswami *et al.*, 2020]. The first method was used to create the noisy channels in the indefinite order channel layout, while the second method was used for all other layouts. We did not observe any significant difference in the performance of the two methods (provided that we applied enough random unitary operations, see Section 3.5.6). However, the first method required several days of measurement, the second less than an hour. Since full QPT was not required for the indefinite order arrangement, we only used the first method for these data, and used the second method for all of the remaining channel configurations.

We collect ≈ 23000 entangled photon pairs per second directly from our source. Of these pairs, we selected only one separable polarization component (*i.e.*, $|H, V\rangle$), halving the count rate. Finally, the photons were sent through optical fibers to the different experiments. Because of experimental imperfections due to the non-zero reflectivity of the various optical elements, the non-ideal fiber coupling, and the optical fiber's losses (the distance to travel in optical fiber between the source and the various experiments is about 3m), approximately 3000 photons per second were detected at the end of the experiment.

Finally, because of the long measurement times (particularly, in the case of the physical implementation of the noise in the channels), we observed phase drifts in the two arms of the interferometer. In order to correct these drifts, so as to ensure that we always prepared and projected the desired path qubit states, we actively stabilized the interferometer by means of the delay line controlled by a piezo-actuator. We measured and reset the phase every 20 minutes (which, according to our tests, ensured phase drifts below 1%). Given these count rates, it follows that we measured about 3×10^6 photons for each internal configuration in case of physical implementation of the noisy channels, and about 3000 counts for each internal configuration in case of implementation of the noisy channels during data analysis.

The aforementioned imperfections in the path qubit had various consequences in our experiment. First, the phase drift on this qubit caused an uncertainty on the input state, since the phase of the path qubit can fluctuate over time. Moreover, if the phase drifts during the experiment, the purity of the input state can be reduced. In light of this, and of the high number of accumulated counts, the main statistical error in our experiment was related to the input state used for QPT. Therefore, to calculate all our experimental error bars, we determined the input state as follows. We performed quantum

tomography of the quantum state directly on the path and polarization qubits in absence of 'internal operations' (*i.e.*, setting to identity all the optical components meant to implement the noisy channels later on). We then observed the variation of this state as a function of time, and used this variation to analyze our data for a 'worst-case' state, a 'best-case' state, and a 'most-likely' state. This spread in the input states led to the error bars and the data-points presented in all of our experimental data.

The second main cause of experimental errors in the path qubit arises from the imperfect visibility of the Mach-Zehnder interferometers used to measure it. This reduced visibility, $\approx 0.93 - 0.96$ (depending on the amount of optical elements in each interferometer arm) essentially corresponds to a slight loss of coherence in the two trajectories and, therefore, to a decrease in the ability of the trajectories to restore quantum information. This leads to a systematic offset in our data. Likewise, a slight miscalibration of our waveplates and liquid crystal waveplates may have occurred, this would lead to further systematic errors. These systematic errors have not been included in our error bar calculation. We note, in particular, that a considerably reduced visibility of the interferometer may constitute the greatest experimental challenge in the application of our techniques to real-world quantum communication.

3.5.6 Fidelity of Channel Implementation

As described in the main text, we implemented the noisy channels in two different ways. In this section, we will discuss the first method, wherein we randomly apply either a Pauli-X, -Y, -Z or the identity operation for one second of our data acquisition time. The probability of each operation is given by the type of noisy channel we wish to implement (*i.e.*, by one of Eqs. (3.8)–(3.10)). The natural question is how many operations must we average over to ensure a faithful implementation of the noisy channels.

To answer this, we used Monte Carlo simulations to study the average 'process fidelity' F_{av} as a function of the number of applied operations. We computed the average process fidelity (defined in Ref. [Gilchrist *et al.*, 2005]) by *i.* randomly generating 10000 single qubit states from the Haar measure, *ii.* computing the ideal output state ρ_{id} using Eqs. (3.8)–(3.10), *iii.* simulating the output by applying N randomly chosen operations ρ_{sim}, and then *iv.* computing the average fidelity between ρ_{id} and ρ_{sim} for all N input states. Already for $N = 25$, the average process fidelity is larger than 0.99. As an example, a plot of the average 'process infidelity' $(1 - F_{\text{av}})$ for the BB84-channel with $p = 0.5$ is shown in Fig. 3.10. (For the infidelity, a value of 0 indicates a perfect implementation.) We chose this as a representative example since the BB84-channel takes a slightly longer time to converge than all the others (this is because the BB84-process randomly applies one of the 4 operations, while all the others only choose among 2 operations). So, the case shown in Fig. 3.10 represents the worst case among all the ones studied. Nevertheless, even such a channel converges to the ideal noisy channel

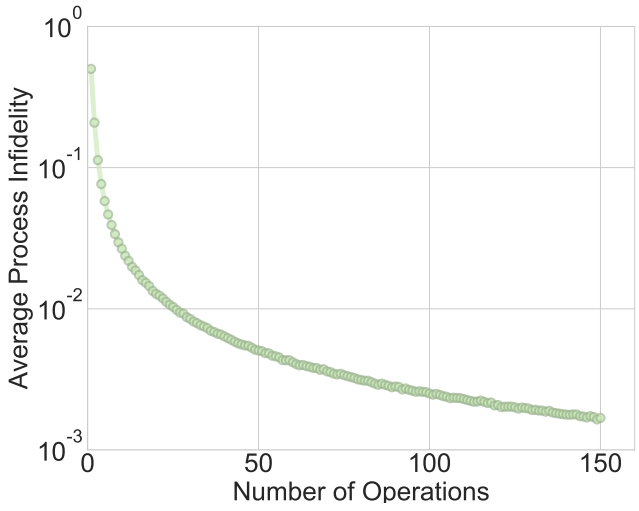


Figure 3.10 Monte Carlo simulation of the BB84-channel with $p = 0.5$. A plot of the average process infidelity between the ideal process and the simulated process versus the number of applied operations used to simulate the noisy channel. The infidelity is defined as $1 - F_{\text{av}}$, where F_{av} is the fidelity. Hence, smaller infidelities indicate a higher degree of agreement.

quite rapidly with N. Finally, since we always implement two channels simultaneously, we apply 1000 different operations, which is far beyond this limit.

3.5.7 Single-Photon Source

A CW laser centered at 392nm emits the pump beam for a source producing single photons through a process of type-II spontaneous-parametric-down-conversion. The pump beam traverses a focusing lens with $f = 12.5$cm, and then reaches, at the proper distance, a 3mm-thick beta-barium borate (BBO) crystal. Within the crystal, single photons are generated at a wavelength centred at 784nm. To compensate for the spatial and temporal walk-off of the resulting single photon pairs, they are sent each through a BBO crystal of 1.5mm thickness. They are finally filtered in polarization through a longpass filter, and a bandpass filter centered at 785nm with a full-weight-half-maximum of 10nm. The photon pairs rate is 23000/s with a pump power of 85mW.

3.5.8 Quantum Process Tomography

Our experimental measurements consist, in general, of performing two-qubit quantum process tomography (QPT) on a path and a polarization qubit. Basically, QPT requires two steps, *i.* preparing the system in a tomographically-complete set of states before the process, and *ii.* measuring the system in a complete basis set after the process. For the polarization qubit, this is relatively straightforward. In fact, in all of three superposition methods outlined in Fig. 3.2, the photons enter the experiment in

a single path. At this point a QWP and a HWP are inserted, which allow us to prepare any single-qubit polarization state. After this, the path qubit is prepared by a 50/50 beamsplitter in a quantum superposition of two paths. After the noisy channels, the paths are recombined by another 50/50 beamsplitter. On each of the output paths we place a QWP, a HWP and a polarizing beamsplitter to implement the polarization measurements. Although they are physically different elements, we ensure that the waveplates in each output arm are always set to the same angle, and hence perform the same measurement.

Even though, in all of the communication schemes presented in the main text, the trajectory is simply initialized in an equal superposition (*i.e.*, the path qubit starts out in $|+\rangle_{\mathrm{T}}$), one must prepare this qubit in a complete set of states in order to perform QPT. In order to change the input state of the path qubit between $|+\rangle_{\mathrm{T}}$, $|R\rangle_{\mathrm{T}} = (|0\rangle_{\mathrm{T}} - i\,|1\rangle_{\mathrm{T}})/\sqrt{2}$, and $|L\rangle_{\mathrm{T}} = (|0\rangle_{\mathrm{T}} + i\,|1\rangle_{\mathrm{T}})/\sqrt{2}$, we set the relative phase between the two trajectories after the first beamsplitter using a delay stage mounted on a calibrated piezo-actuator. We can also easily prepare $|0\rangle_{\mathrm{T}}$ and $|1\rangle_{\mathrm{T}}$ by blocking either path. Analogously, we measure the path qubit in two different ways. To measure in $\{|+\rangle_{\mathrm{T}}, |-\rangle_{\mathrm{T}}\}$, or $\{|R\rangle_{\mathrm{T}}, |L\rangle_{\mathrm{T}}\}$, we suitably set the relative phase between the two paths before recombining them at the second beamsplitter. We use the same delay stage to both set the phase of the path state, and to measure it in $\{|+\rangle_{\mathrm{T}}, |-\rangle_{\mathrm{T}}\}$, or $\{|R\rangle_{\mathrm{T}}, |L\rangle_{\mathrm{T}}\}$. This can be done by adding the required phase for state preparation and subtracting the phase for state measurement. Such a phase is then converted into a path delay and sent to the piezo-actuated delay stage. To measure in the $\{|0\rangle, |1\rangle\}$ basis, we block either path before the 50/50 beamsplitter, and we then sum the counts from the two paths after the beamsplitter.

To collect a complete set of data, we prepare the path qubit in $\{|0\rangle_{\mathrm{T}}, |+\rangle_{\mathrm{T}}, |R\rangle_{\mathrm{T}}, |L\rangle_{\mathrm{T}}\}$, and for each of these path states we prepare the polarization qubit in $\{|0\rangle_{\mathrm{I}}, |+\rangle_{\mathrm{I}}, |R\rangle_{\mathrm{I}}, |L\rangle_{\mathrm{I}}\}$, for a total of 16 input states. We then measure each of these 16 two-qubit states by setting 9 different two-qubit basis settings: $\{|0,0\rangle, |0,+\rangle, |0,R\rangle, |+,0\rangle, |+,+\rangle, |+,R\rangle, |R,0\rangle, |R,+\rangle, |R,R\rangle\}_{\mathrm{I,T}}$. However, for each measurement setting we measure all four outcomes. For example, when the measurement is set to $|0,0\rangle_{\mathrm{I,T}}$, we obtain the projections onto $|0,0\rangle_{\mathrm{I,T}}$, $|0,1\rangle_{\mathrm{I,T}}$, $|1,0\rangle_{\mathrm{I,T}}$, and $|1,1\rangle_{\mathrm{I,T}}$. This yields 36 different measurement results for each of the 16 input states, providing an over-complete data set, on which we perform a least-squares QPT routine.

Equipped with this mathematical description of our experimental channel, we can compute the action of our experiment on one qubit of a maximally-entangled Bell state when the path qubit is set to $|+\rangle_{\mathrm{T}}$. From this, we evaluate the coherent information (Eq. (3.7)). Fixing the state of the path qubit in this manner results in the coherent information of the effective one-to-two-qubit channel.

We carry out this method based on full QPT for the cases of quantum-control of parallel channels and channels in series with quantum-controlled operations, but for the quantum-control of channel order we can make a simplification to lower bound the coherent information which saves significant

measurement time. For these data, we only prepare the path state $|+\rangle_{\mathrm{T}}$, and then measure it in the $\{|+\rangle_{\mathrm{T}}, |-\rangle_{\mathrm{T}}\}$ basis, as described above. With these measurements, we lower bound the coherent information in our channels as follows. We first reconstruct two single-qubit χ-matrices for the target systems, $\chi^{|+\rangle}$ and $\chi^{|-\rangle}$, using single-qubit process tomography on the polarization qubit. In particular, $\chi^{|+\rangle}$ is the single-qubit effective process that the information qubit experiences when the trajectory measurement results is $|+\rangle_{\mathrm{T}}$, whereas $\chi^{|-\rangle}$ is the effective process when the trajectory measurement outcome is $|-\rangle_{\mathrm{T}}$

Next, we compute the action of the one-qubit χ-matrix on a maximally-entangled Bell state, to evaluate the two values of the coherent information $\mathcal{I}_c^{|+\rangle}$ and $\mathcal{I}_c^{|-\rangle}$ in Eq. (3.7). Afterwards, we simply calculate their average, with each term weighted by their respective post-selection probabilities $p_{|+\rangle}$ and $p_{|-\rangle}$:

$$\mathcal{I}_c^{\mathrm{LB}} = p_{|+\rangle}\mathcal{I}_c^{|+\rangle} + p_{|-\rangle}\mathcal{I}_c^{|-\rangle}. \tag{3.28}$$

In general, $\mathcal{I}_c^{\mathrm{LB}}$ sets a lower bound on $\mathcal{I}_c$ because of the data processing inequality for coherent information [Nielsen and Chuang, 2000; Schumacher and Nielsen, 1996]. Furthermore, in absence of additional errors, it can be shown that $\mathcal{I}_c^{\mathrm{LB}} = \mathcal{I}_c$ in the case of the quantum switch.

3.5.9 Fixing the Quantum-Controlled Operations Independently of the Noise

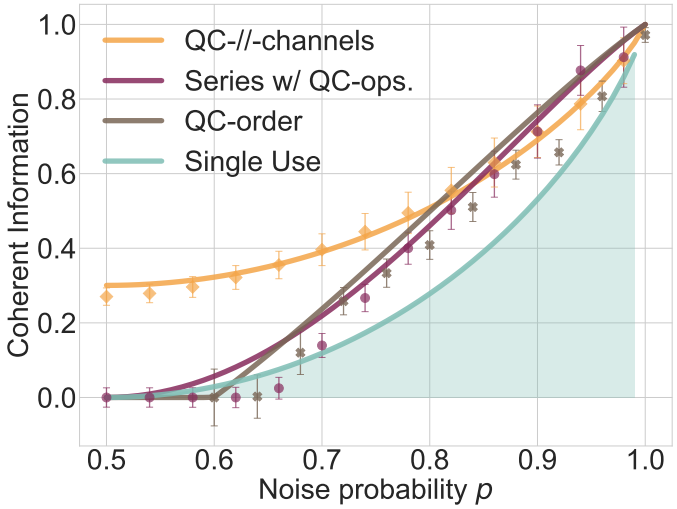


Figure 3.11 Experimental BF- and PF-noise data for sub-optimal quantum-controlled operations. The trend of the scheme featuring the channels in series with quantum-controlled operations (Series w/ QC-ops.) performs worse than the quantum-control of channel order (QC-order) for all $p \geq 0.67$, but better than the quantum-control of parallel channels (QC-//-channels) for $p \geq 0.84$. The experimental data for the sub-optimal choice of Series w/ QC-ops. is in good agreement with the expected trend.

In some cases, for instance in a rapidly-varying noise environment, it may be impossible to estimate

the type of noise and adapt the quantum-controlled operations accordingly. In these situations, one would need to fix such operations independently of the noise. For our noise varieties, the operations U_2 and U_3 were set to $\mathcal{I}$ in the cases of the *XY*- and BB84-channels, and to H for the BF-and-PF case, whereas $U_1 = Y$ in all three cases. If we were to keep the same quantum-controlled operations in the BF-and-PF case as in the *XY* and BB84 cases, the efficiency of the scheme would be reduced, and the channel activation due to the channels in series with quantum-controlled operations would result comparable to that of the two other schemes (*i.e.*, the quantum-control of parallel channels, and the quantum-control of channel order). The theoretical trend and the experimental data points corresponding to this case are shown on Fig. 3.11. Colors and data points shapes are the same as in Figs. 3.3–3.5.

Chapter 4

Time's Arrow of a Quantum Superposition of Thermodynamic Evolutions

G. Rubino, G. Manzano Paule, and Č. Brukner

Abstract. A priori, there exists no preferential temporal direction as microscopic physical laws are time-symmetric. Still, the second law of thermodynamics allows one to associate the 'forward' temporal direction to a positive variation of the total entropy produced in a thermodynamic process, and a negative variation with its 'time-reversal' counterpart. This definition of a temporal axis is normally considered to apply in both classical and quantum contexts. Yet, quantum physics admits also superpositions between forward and time-reversal processes, thereby seemingly eluding conventional definitions of time's arrow. In this work, we demonstrate that a quantum measurement of entropy production can distinguish the two temporal directions, effectively projecting such superpositions of thermodynamic processes onto the forward (time-reversal) time-direction when large positive (negative) values are measured. Remarkably, for small values (of the order of plus or minus one), the amplitudes of forward and time-reversal processes can interfere, giving rise to entropy-production distributions featuring a more or less reversible process than either of the two components individually, or any classical mixture thereof. Finally, we extend these concepts to the case of a thermal machine running in a superposition of the heat engine and the refrigerator mode, illustrating how such interference effects can be employed to reduce undesirable fluctuations.

Author contributions: *G.R., G.M. and Č.B. contributed to all aspects of the research, with the leading input of G.R..*

4.1 Introduction

In spite of it being seemingly straightforward, physics is still nowadays seeking to provide a comprehensive understanding of the apparent passage of time [Halliwell *et al.*, 1996]. The concept of time is intimately related to the observation of a change in physical systems. However, the recognition that, at their most fundamental level, physical systems generally obey time-reversal laws led to the realisation that systems' evolutions do not intrinsically differentiate between forward and backward time directions. Attempts to uphold with physical arguments the evidence of the time flow are being made on multiple fronts, mainly on the basis of empirical observations: we see that entropy in the universe increases (thermodynamic time's arrow), that the universe expands (cosmological time's arrow), that causes always precede their effects (causal time's arrow). Likewise, there have been several proposals as to the explanation of the time's arrow in a quantum-mechanical contexts [Erker *et al.*, 2017; Jennings and Rudolph, 2010a,b; Maccone, 2009; Mlodinow and Brun, 2014]. The peculiarity of the quantum framework is that it enables for processes to be placed in quantum superposition. Applied to the notion of thermodynamic time's arrow, this implies that quantum mechanics can allow the superposition of processes producing opposite variations in the entropy. This raises the question of how a well-defined thermodynamic arrow of time can be established in the quantum framework when such superpositions are in place. To address this question, in this work we show that a measurement of the entropy production has a decisive role in restoring a definite thermodynamic time's arrow. We then also investigate interference effects in such superpositions, and we apply our results to the more concrete scenario of quantum-thermodynamic engines. Our investigations bear a conceptual similarity with the field of indefinite quantum causality, wherein the order of operations is placed in a quantum superposition [Chiribella, 2012; Chiribella *et al.*, 2013; Oreshkov *et al.*, 2012]. In the present case, however, instead of exploring causal superpositions between different orders of operations, we analyse superpositions of thermodynamic processes related by time-reversal symmetry.

In thermodynamics, the time's arrow is introduced by the second law of thermodynamics, according to which the total entropy of the universe can only either increase, or remain constant. Consequently, one might think that observations of entropy changes are all we need to distinguish the past from the future: an overall increase in entropy shall be identified with the direction of time 'forward', while a overall decrease in entropy with its 'time-reversal' counterpart. Yet, for a microscopic system, fluctuations blur the direction of the time's arrow, and the time flow is only defined *on average*. This can be illustrated in terms of a 'guessing the time directionality game' which was introduced by C. Jarzynski in Ref. [Jarzynski, 2011]. There, the author supposes to record the motion of a non-equilibrium thermodynamic process, and then to toss a coin. Depending on the outcome of the coin, he either plays the movie in the order in which it took place, or in the time-reversal one. In order to determine in which order the movie is being shown, the optimal guessing strategy for a macroscopic system follows from the second law of thermodynamics: if $\langle W \rangle > \Delta F$, the movie proceeds in the correct order, while if $\langle W \rangle < \Delta F$, the movie is being run backwards. Here, $\langle W \rangle$ is

the average work performed on the system by the external driving mechanism, and ΔF the difference in free energies of the thermodynamic states at the beginning and at the end of the movie. However, for a microscopic system, the author shows that the optimal guessing strategy exploits the so-called 'fluctuation theorems' [Campisi *et al.*, 2011; Esposito *et al.*, 2009; Evans and Searles, 2002; Seifert, 2012], together with Bayesian probabilistic reasoning [Maragakis *et al.*, 2008; Shirts *et al.*, 2003]. We review this study briefly in the Supplemental Note 4.6.1.

In one of its most famous versions [Bochkov and Kuzovlev, 1977; Bochkov and Kuzovlev, 1981; Crooks, 1999; Jarzynski, 1997], the fluctuation theorem describes the fluctuations of the dissipative work $W_{\text{diss}} = W - \Delta F$ associated to the observation of a particular value of W in a single realisation of a non-equilibrium driving protocol (*i.e.*, a single shot of the movie):

$$\frac{P(+W)}{\tilde{P}(-W)} = e^{\beta W_{\text{diss}}}, \tag{4.1}$$

where $P(+W)$ represents the probability that a work W is invested along the forward thermodynamic evolution, whereas $\tilde{P}(-W)$ is the probability linked to recovering the same amount of work along the time-reversal evolution, both of which start in equilibrium with a thermal bath. Furthermore, the dissipative work W_{diss} is the work invested in a thermodynamic transformation between equilibrium states having a free energy difference ΔF, which cannot be recovered by reversing the process. The relation to the entropy production (or total entropy) ΔS_{tot} in the process is established through the relation: $\Delta S_{\text{tot}} = \beta\, W_{\text{diss}}$, where $\beta = (k_B T)^{-1}$ is the inverse temperature, with k_B being the Boltzmann constant and T the temperature of the bath [Kawai *et al.*, 2007; Parrondo *et al.*, 2009]. From this equation, it follows that both the probability of total-entropy-decreasing events ($\beta W_{\text{diss}} < 0$) in the forward evolution, and that of total-entropy-increasing ones ($\beta W_{\text{diss}} > 0$) using the time-reversal dynamics vanish exponentially with the size of the total entropy variation:

$$P(\beta W_{\text{diss}} < -\xi) \leq e^{-\xi}, \tag{4.2a}$$

$$\tilde{P}(\beta W_{\text{diss}} > +\xi) \leq e^{-\xi}, \tag{4.2b}$$

for any $\xi \geq 0$, and where the second inequality (4.2b) arises from the fact that, in the time-reversal process, the dissipative work equals $-W_{\text{diss}}$. In other words, large reductions in the total entropy are unlikely in the forward evolution, while events leading to a large entropy production are unlikely in the time-reversal one. (Notice that the sign of the entropy change is defined to match that of the dissipate work in the forward process.) Interestingly, it is evidenced that, when βW_{diss} is of the order of one, it is inherently impossible to tell in which of the two orders the process has occurred. In this region, the directionality of time flow cannot be inferred, and the time's arrow is, so to say, blurred. A clear temporal directionality is then reestablished for $\beta |W_{\text{diss}}| \gg 1$.

We remark that, here, 'forward' and 'time-reversal' are interchangeable labels since each process represents the time-inverted version of the other. Moreover, it is worth noticing that considerations on

time-inversion only take on relevance in the absence of complete time-symmetry, as this latter may lead to $\Delta S_{\rm tot}$ equal to zero in every single realisation. In order to exhibit time-asymmetry, in the present study the two conjugated processes are assumed to start from equilibrium states, a standard procedure in the derivation of fluctuations theorems [Campisi *et al.*, 2011; Esposito *et al.*, 2009]. This introduces a final (implicit) thermalization step which enables irreversibility to emerge (see, *e.g.*, Refs. [Landi and Paternostro, 2020; Manzano *et al.*, 2018; Parrondo *et al.*, 2009]).

In what follows, we will explore what confers a definite thermodynamic time's arrow to quantum superpositions between 'forward' and 'time-reversal' processes (*i.e.*, thermodynamic processes whose quenches are correlated by time-inversion symmetry). After constructing a quantum superposition between two such processes (Section 4.2.1) and the mathematical framework for their evaluation (Section 4.2.2), we will show that, in the quantum case, analogously to the classical one, quantum measurements of work (or, equivalently, entropy production) can distinguish the past from the future: when the measured dissipative work equals $\beta W_{\rm diss} \gg 1$, the superposition is effectively projected onto the forward process, whereas when $\beta W_{\rm diss} \ll -1$, it is effectively projected onto the time-reversal one, hence recovering a definite thermodynamic arrow of time (Section 4.3.1). However, when $\beta|W_{\rm diss}|$ is of the order of one, the forward and the time-reversal thermodynamic processes can quantum mechanically interfere, resulting in a work probability distribution describing work fluctuations which have no classical counterpart. More precisely, in the case of interference, the probabilities take on values which cannot be obtained by any classical (convex) mixture of the forward and the time-reversal processes (Section 4.3.2). Finally, extending our approach, we examine a quantum superposition of a work-extracting heat engine and power-driven refrigerator, and we show that such a device can achieve efficiency statistics which no thermal process acting probabilistically as one or the other engine could accomplish (Section 4.3.3).

4.2 Theoretical Framework

4.2.1 Superposition of Forward and Time-Reversal Dynamics

We start by defining the framework used to characterize thermodynamic processes and work fluctuations. First, we will introduce all the necessary elements to formally construct a state representing the quantum superposition of a thermodynamic process evolving in the forward temporal direction, and one evolving in the time-reversal direction. Then, we will discuss how to characterize work and entropy production fluctuations in such superposition states using an extended two-point-measurement (TPM) scheme, and we illustrate how the outcomes achieved through processes with well-defined time directions can be recovered inside our framework.

We consider a thermodynamic system S being, in both forward and time-reversal processes, initially in equilibrium with a thermal reservoir at inverse temperature β. The process occurring in the forward direction will be realized by a quench $U(t,0)$ induced by the time-dependent Hamiltonian $H\big(\lambda(t)\big)$ executing a controlled protocol $\Lambda \equiv \{\lambda(t); 0 \leq t \leq \tau\}$ in the time-frame $t \in [0,\tau]$. Its time-reversal twin will be described by a quench $\tilde{U}(\tau - t, 0)$ associated to the implementation of the time-reversal protocol $\tilde{\Lambda} \equiv \{\tilde{\lambda}(\tau - t); 0 \leq t \leq \tau\}$, where $\tilde{\lambda}$ is the time-reversed control parameter (associated to the parity of the Hamiltonian parameters under time-reversal). The micro-reversibility principle for non-autonomous systems establishes a strong relation between forward and time-reversal quenches lying at the core of fluctuation theorems [Andrieux and Gaspard, 2008; Campisi *et al.*, 2011]:

$$\tilde{U}(\tau - t, 0) = \Theta\, U^{\dagger}(\tau, t)\, \Theta^{\dagger}, \tag{4.3}$$

where Θ denotes the (anti-unitary) time-reversal operator acting on the system's Hilbert space, which flips the sign of observables with odd parity under time-reversal. This operator verifies the relations $\Theta\, \mathbb{1} i = -\mathbb{1} i\, \Theta$, and $\Theta\, \Theta^{\dagger} = \Theta^{\dagger}\, \Theta = \mathbb{1}$. In our configuration, the micro-reversibility principle in Eq. (4.3) holds whenever the system Hamiltonian verifies $\Theta H\big(\lambda(\tau - t)\big)\Theta^{\dagger} = H\big(\tilde{\lambda}(\tau - t)\big)$.

In order to describe superpositions of forward and time-reversal processes, the initial equilibrium states of the system S can be purified by including the environment E with a generic Hamiltonian H_E in the description (where the environment may include the thermal reservoir as well as other relevant degrees of freedom which get entangled with the system). These purifications are not unique, and they can be represented by joint states of the system and the environment of the form

$$|\psi_0\rangle_{S,E} = \sum_k \sqrt{\frac{e^{-\beta E_k^{(0)}}}{Z_0}}\, |E_k^{(0)}\rangle_S\, |\varepsilon_k^{(0)}\rangle_E, \tag{4.4a}$$

$$|\tilde{\psi}_0\rangle_{S,E} = \sum_k \sqrt{\frac{e^{-\beta E_k^{(\tau)}}}{Z_\tau}}\, \Theta |E_k^{(\tau)}\rangle_S\, |\varepsilon_k^{(\tau)}\rangle_E, \tag{4.4b}$$

where $E_k^{(0)}$ and $E_k^{(\tau)}$ are the eigenvalues of the Hamiltonian at times $t = \{0, \tau\}$, *i.e.*, $H[\lambda(0)]$ and $H[\lambda(\tau)]$, whereas $|E_k^{(0)}\rangle_S$ and $|E_k^{(\tau)}\rangle_S$ are the corresponding eigenvectors (for the sake of brevity, we will henceforth omit the subscript S in the system's energy eigenvectors). Furthermore, $|\varepsilon_k^{(0)}\rangle_E$, $|\varepsilon_k^{(\tau)}\rangle_E$ represent the set of states of the environmental degree of freedom at initial and final times, which can always be chosen as sets of orthogonal states. Notice that the environment may possess further degrees of freedom which are not entangled with the system under consideration, and which we will thus not explicitly account for.

The state $|\psi_0\rangle_{S,E}$ above corresponds to the initial state of the process evolving in the forward direction, whereas $|\tilde{\psi}_0\rangle_{S,E}$ is the initial state of the process which takes place in the time-reversal fashion. Notice that, by tracing out the environmental degrees of freedom, we recover the corresponding Gibbs thermal

states for the system $\rho_0^{\text{th}} \equiv \text{Tr}_E(|\psi_0\rangle\langle\psi_0|_{S,E}) = e^{-\beta H[\lambda(0)]}/Z_0$ and $\tilde{\rho}_0^{\text{th}} \equiv \text{Tr}_E(|\tilde{\psi}_0\rangle\langle\tilde{\psi}_0|_{S,E}) = \Theta\, e^{-\beta H[\lambda(\tau)]}\Theta^\dagger/Z_\tau = e^{-\beta H[\tilde{\lambda}(\tau)]}/Z_\tau$, being $Z_0 = \text{Tr}(e^{-\beta H[\lambda(0)]})$, and $Z_\tau = \text{Tr}(e^{-\beta H[\lambda(\tau)]})$ the partition functions.

Moreover, we introduce an auxiliary system A whose two orthogonal states $\{|0\rangle_A, |1\rangle_A\}$ govern the evolution of the process in the two temporal directions. This is a quantum analogue of the coin tossed to decide classically which process to run (forward or time-reversal). With this in place, the global Hamiltonian of the system, the environment, and the auxiliary qubit reads $\mathcal{H}(t) \equiv \big(\,|0\rangle\,\langle 0|_A \otimes H[\lambda(t)] + |1\rangle\,\langle 1|_A \otimes \tilde{H}[\lambda(\tau - t)]\big) \otimes H_E$. We then entangle each orthogonal auxiliary state to one of the initial states in Eq. (4.4). The overall initial state of system, environment and auxiliary hence reads:

$$|\Psi_0\rangle_{S,E,A} = \alpha_0\,|\psi_0\rangle_{S,E} \otimes |0\rangle_A + \alpha_1\,|\tilde{\psi}_0\rangle_{S,E} \otimes |1\rangle_A, \tag{4.5}$$

with arbitrary coefficients $\alpha_0, \alpha_1 \in \mathbb{C}$, $|\alpha_0|^2 + |\alpha_1|^2 = 1$. If, subsequently, in each branch of the superposition in Eq. (4.5) the forward and time-reversal quenches are respectively applied, the evolved state at some arbitrary instant of time $t \in [0, \tau]$ is given by $|\Psi(t)\rangle_{S,E,A} = \alpha_0\,\big[U(t,0) \otimes \mathbb{1}_{E,A}\big]|\psi_0\rangle_{S,E} \otimes |0\rangle_A + \alpha_1|\big[\tilde{U}(t,0) \otimes \mathbb{1}_{E,A}\big]|\tilde{\psi}_0\rangle_{S,E} \otimes\ |1\rangle_A$. In this expression, the first and the second amplitudes correspond to the forward and the time-reversal directions, respectively. Furthermore, we assume that the system does not interact with the environment during the timescale of the quenches (however, after the quench, the system thermalises through the interaction with the thermal reservoir). This is verified whenever the quenches are implemented in a fast timescale as compared to the characteristic relaxation time of the system in interaction with the environment. Furthermore, we will consider the quenches $U(t,0)$ and $\tilde{U}(t,0)$ in the superposition to be implemented by some external (classical) control. As we discuss in the Supplemental Note 4.6.2, this limit is adequate in our setup, and it corresponds to the case in which the control mechanism acts approximately as an ideal reservoir of energy and coherence [Åberg, 2014; Bartlett *et al.*, 2007; Korzekwa *et al.*, 2016; Malabarba *et al.*, 2015], as is the case, for instance, with lasers or radio-frequency pulses.

Taking a gas enclosed in a vessel as a pictorial example, the aforementioned state can be constructed by entangling the position of the piston with a further auxiliary quantum system, thereby establishing a quantum superposition of the following two processes: *i.* a process wherein the gas particles are initially in thermal equilibrium confined in one half of the vessel by a piston, and the piston is pulled outwards, and *ii.* the reverse process, in which the piston is pushed towards the gas, starting from an initial state where the gas occupies the entire vessel in thermal equilibrium.

4.2.2 Extended Two-Point Measurement Scheme

We will now measure the work of the system undergoing the above-mentioned superposition of forward and time-reversal dynamics. In order to implement such a measurement, we formally construct

a procedure described by a set of measurement operators forming a completely positive and trace-preserving (CPTP) map. In this regard, we will refer to a standard TPM procedure to measure work in quantum thermodynamic processes [Campisi *et al.*, 2011]. Implementations of the TPM in quantum setups [Batalhão *et al.*, 2014; De Chiara *et al.*, 2015; Dorner *et al.*, 2013; Mazzola *et al.*, 2013; Roncaglia *et al.*, 2014], as well as suitable extensions [Åberg, 2018; Debarba *et al.*, 2019; Mohammady and Romito, 2019; Perarnau-Llobet *et al.*, 2017], have recently received increasing attention. Our procedure can be seen as a generalization of the TPM scheme to situations where different thermodynamic processes are allowed to be superposed, and may consequently interfere.

In the TPM scheme, work is defined as the energy difference between the initial and final states of the system, which are measured through ideal projective measurements of the system Hamiltonian implemented before and after the thermodynamic process associated to the protocol Λ [Talkner and Hänggi, 2016; Talkner *et al.*, 2007]. This measurement scheme can be performed, individually, both for the forward and the time-reversal processes, enabling the construction of the work probability distributions $P(W)$ and $\tilde{P}(W)$, respectively.

As far as the forward process is concerned, the probability to observe a transition $|E_n^{(0)}\rangle \to |E_m^{(\tau)}\rangle$ is given by $p_{n,m} = p_{m|n}\, p_n^{(0)}$, where $p_n^{(0)} = e^{-\beta E_n^{(0)}}/Z_0$ is the probability of observing the energy $E_n^{(0)}$ at $t = 0$, and $p_{m|n} = \left|\langle E_m^{(\tau)}|\, U(\tau,0)|E_n^{(0)}\rangle\right|^2$ is the conditional probability of measuring $E_m^{(\tau)}$ at $t = \tau$ after having measured $E_n^{(0)}$ at the beginning of the process. Similarly, for the time-reversal process one has $\tilde{p}_{m,n} = \tilde{p}_{n|m}\,\tilde{p}_m^{(0)}$, where $\tilde{p}_m^{(0)} = e^{-\beta E_m^{(\tau)}}/Z_\tau$ is the probability to obtain the energy $E_m^{(\tau)}$ at the beginning of the time-reversal process, and $\tilde{p}_{n|m} = \left|\langle E_n^{(0)}|\Theta^\dagger \tilde{U}(\tau,0)\Theta|E_m^{(\tau)}\rangle\right|^2$ is the corresponding conditional probability for observing the inverse transition $\Theta\,|E_m^{(\tau)}\rangle \to \Theta\,|E_n^{(0)}\rangle$ given that one obtained $E_m^{(\tau)}$ in the first measurement. The micro-reversibility principle in Eq. (4.3) relates the conditional probabilities in the forward and time-reversal processes as $\tilde{p}_{n|m} = p_{m|n}$ [Andrieux and Gaspard, 2008; Campisi *et al.*, 2011].

The TPM scheme allows one to compute the stochastic work invested by the external driver in a single realisation of the protocol Λ, $W_{n,m} \equiv E_m^{(\tau)} - E_n^{(0)}$, associated to the outcomes of initial and final energy measurements. Its probability distribution reads:

$$P(W) = \sum_{n,m} p_{n,m} \cdot \delta(W - W_{n,m}). \tag{4.6}$$

Analogously, the probability distribution associated to the work invested in the time-reversal protocol, $\tilde{W}_{n,m} = E_n^{(0)} - E_m^{(\tau)} = -W_{n,m}$, is given by:

$$\tilde{P}(W) = \sum_{n,m} \tilde{p}_{n,m} \cdot \delta(W - \tilde{W}_{n,m}). \tag{4.7}$$

Hereafter, we consider an extension of the TPM scheme in which we include energy measurements at

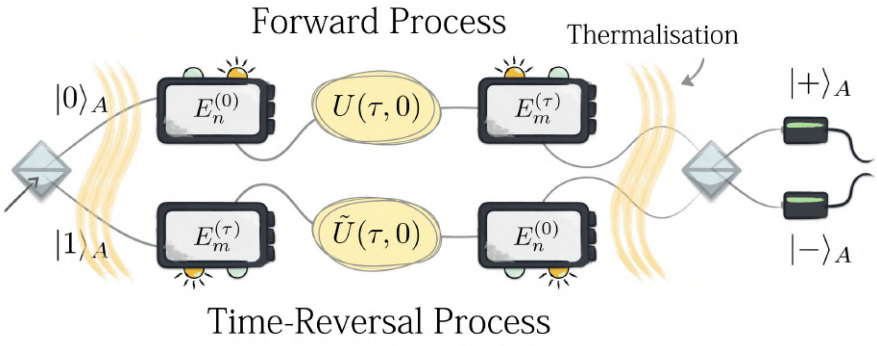


Figure 4.1 Schematic representation of a superposition of a forward thermodynamic quench with its time-reversal counterpart. A thermodynamic system S is coupled to a auxiliary system A. Depending on the state of the auxiliary system, $|0\rangle_A$) or $|1\rangle_A$, the system S is initially prepared in a thermal state of the initial or final Hamiltonians, $H(0)$ and $H(\tau)$, respectively. It is then sent through a thermodynamic quench $U(t,0)$ or its time reversal $\tilde{U}(t,0)$ in the time-frame $t \in [0,\tau]$. Before and after each quench, the system's energy is measured. The measurement outcomes $E_n^{(0)}$ and $E_m^{(\tau)}$ are found when the auxiliary system is in $|0\rangle_A$, whereas the outcomes $E_m^{(0)}$ and $E_n^{(\tau)}$ are obtained when the auxiliary system is in $|1\rangle_A$. The auxiliary system is measured in the basis $\left\{|\pm\rangle_A = (|0\rangle_A \pm |1\rangle_A)/\sqrt{2}\right\}$, while the system may eventually undergo a second thermalisation with the environment.

$t = 0$ and $t = \tau$ in both branches of the superposition between a forward and a time-reversal processes, as illustrated in Fig. 4.1. More precisely, starting with the initial state in Eq. (4.5), and conditionally on the auxiliary state, we consider the application of the projectors $|E_n^{(0)}\rangle\langle E_n^{(0)}|$ and $\Theta\,|E_m^{(\tau)}\rangle\langle E_m^{(\tau)}|\,\Theta^\dagger$ to the initial states $|\psi_0\rangle_{S,E}$ and $|\tilde{\psi}_0\rangle_{S,E}$, respectively. Subsequently, the unitary quenches $U(\tau,0)$ and $\tilde{U}(\tau,0)$ are implemented in each branch, after which the projectors $|E_m^{(\tau)}\rangle\langle E_m^{(\tau)}|$ and $\Theta\,|E_n^{(0)}\rangle\langle E_n^{(0)}|\,\Theta^\dagger$ are respectively applied. Consequently, given the outcomes $E_n^{(0)}$ and $E_m^{(\tau)}$, a work $W_{n,m}$ is invested in the forward-dynamics branch by applying the protocol Λ, whereas the work invested in its time-reversal counterpart $\tilde{\Lambda}$ is $\tilde{W}_{n,m} = -W_{n,m}$ (that is, the same amount of work as in the forward dynamics is here extracted).

The operator representing the application of the scheme through which the work W is obtained can be written as:

$$M_W = \sum_{n,m}\Big[|E_m^{(\tau)}\rangle\langle E_m^{(\tau)}|U(\tau,0)|E_n^{(0)}\rangle\langle E_n^{(0)}| \otimes \mathbb{1}_E \otimes |0\rangle\langle 0|_A + \Theta|E_n^{(0)}\rangle\langle E_n^{(0)}|\Theta^\dagger\tilde{U}(\tau,0)\Theta|E_m^{(\tau)}\rangle\langle E_m^{(\tau)}|\Theta^\dagger \otimes \mathbb{1}_E \otimes |1\rangle\langle 1|_A\Big] \cdot \delta(W - W_{n,m}). \tag{4.8}$$

The set of operators $\{M_W\}$ forms a CPTP map, $\mathcal{E}(\rho) \equiv \int dW M_W \rho M_W^\dagger$, acting on the composite S, E, A system and fulfilling $\int dW M_W^\dagger M_W = \mathbb{1}$. The map $\mathcal{E}$ describes the average effect of the measurement scheme on an arbitrary initial state of the composite system ρ, while the operations $\mathcal{E}_W(\rho) \equiv M_W \rho M_W^\dagger$ provide the probability $\mathcal{P}(W) \equiv \mathrm{Tr}[\mathcal{E}_W(\rho)]$ to measure the work W.

It is important to stress that the operations $\mathcal{E}_W$ preserve the coherence between the forward and time-reversal thermodynamic processes. Indeed, performing a standard quantum measurement on the pro-

cess would destroy the coherence, as it would reveal the time at which the measurement has been performed, and, from this, also whether the outcome E_m was observed before (in the forward process) or after the outcome E_n (in the time-reversal process). In other words, such a measurement would reveal the time direction, and it would be equivalent to the projection of the auxiliary qubit in the basis $\{|0\rangle_A, |1\rangle_A\}$. However, there exist also measurement schemes in which the result is encoded in an auxiliary system through its entanglement with the measured system, and the result is then read only at the end of the whole evolution, thereby preserving its coherence. (Such a measurement scheme was recently used to measure the system undergoing superposition of causal orders [Rubino *et al.*, 2017a].) In such a scheme, the system on which the thermodynamic quenches act and the auxiliary system can be encoded on two different degrees of freedom of the same quantum system. If the auxiliary degree of freedom is of sufficient dimension, it is possible to encode the results of each measurement taking place within the process in a state of this system. The coherence between these states has to be maintained until the end of the overall thermodynamic process, and the different states are then interfered with each other. (For simplicity, in this study we consider only two states of the auxiliary system (Fig. 4.1). Nevertheless, all the conclusions drawn herein can be extended to the case of more than two states).

In order to evaluate the work probability distribution in the extended TMP scheme, it is also crucial to take into account the mutual phases between the conditional probabilities. We thus write, in general

$$\langle E_m^{(\tau)}|\, U(\tau,0)|E_n^{(0)}\rangle := \sqrt{p_{m|n}}\, e^{i\Phi_{n,m}}, \tag{4.9a}$$

$$\langle E_n^{(0)}|\, U^\dagger(\tau,0)|E_m^{(\tau)}\rangle := \sqrt{\tilde{p}_{n|m}}\, e^{-i\tilde{\Phi}_{m,n}}, \tag{4.9b}$$

and we notice that

$$\sqrt{\tilde{p}_{m|n}} e^{-i\tilde{\Phi}_{m,n}} = \langle E_n^{(0)}|\, U^\dagger(\tau,0)|E_m^{(\tau)}\rangle = \left(\sqrt{p_{n,m}}\, e^{i\Phi_{n,m}}\right)^* = \sqrt{p_{n|m}}\, e^{-i\Phi_{n,m}},$$

from which we get $\Phi_{n,m} = \tilde{\Phi}_{m,n}$, since $\tilde{p}_{m|n} = p_{n|m}$.

We now consider the concatenation of the operation M_W with a projection of the auxiliary qubit onto an arbitrary state $|\xi\rangle_A$. By applying this sequence of operations to the initial state in Eq. (4.5), we derive the (unnormalized) state of the composite system associated to the work outcome W and projection of the auxiliary qubit onto $|\xi\rangle_A$:

$$|\Psi_W^\xi\rangle_{S,E,A} \equiv \left(\mathbb{1}_{S,E} \otimes |\xi\rangle\langle\xi|_A\right) \circ M_W |\Psi_0\rangle_{S,E,A} = |\Xi_0^\xi\rangle + |\Xi_1^\xi\rangle\,, \tag{4.10}$$

where we identified the two branches of the superposition corresponding to the forward ($|\Xi_0^\xi\rangle$) and

the time-reversal dynamics ($|\Xi_1^\xi\rangle$). They read, respectively:

$$|\Xi_0^\xi\rangle = \alpha_0\langle\xi|0\rangle \sum_{n,m}\sqrt{p_{n,m}}\, e^{i\Phi_{n,m}}\, |E_m^{(\tau)}\rangle\, |\varepsilon_n^{(0)}\rangle_E|\xi\rangle_A \cdot \delta(W - W_{n,m}), \tag{4.11a}$$

$$|\Xi_1^\xi\rangle = \alpha_1\langle\xi|1\rangle \sum_{n,m}\sqrt{p_{n,m}}\, e^{-i\Phi_{n,m}}\, \Theta|E_n^{(0)}\rangle\, |\varepsilon_m^{(\tau)}\rangle_E|\xi\rangle_A e^{-\frac{\beta}{2}(W_{n,m}-\Delta F)} \cdot \delta(W - W_{n,m}), \tag{4.11b}$$

where, in the second equation, we made use of $\tilde{p}_{n,m} = p_{n,m}\, e^{-\beta(W_{n,m}-\Delta F)}$ (see the Supplementary Note 4.6.3), and of the relation between the forward and time-reversal phases $\tilde{\Phi}_{m,n} = \Phi_{n,m}$.

The joint probability of measuring the work W and projecting the auxiliary state onto $|\xi\rangle_A$ is hence given by $\mathcal{P}(\xi, W) = \big|\big||\Psi_W^\xi\rangle_{S,E,A}\big|\big|^2$. Furthermore, from the joint probabilities $\mathcal{P}(\xi, W)$, one can obtain the conditional ones $\mathcal{P}_\xi(W) := \mathcal{P}(W|\xi) = \mathcal{P}(\xi, W)/\mathcal{P}(\xi)$, which we will hereafter refer to as 'post-selected work probability distributions', and where $\mathcal{P}(\xi) = \int dW\, \mathcal{P}(\xi, W)$. By introducing the notation $q_0^\xi = |\alpha_0|^2\big|\langle\xi|0\rangle\big|^2/\mathcal{P}(\xi)$ and $q_1^\xi = |\alpha_1|^2\big|\langle\xi|1\rangle\big|^2/\mathcal{P}(\xi)$, we can rewrite $\mathcal{P}_\xi(W)$ as:

$$\mathcal{P}_\xi(W) = q_0^\xi\, P(W) + q_1^\xi\, \tilde{P}(-W) + 2\,\mathbb{R}\mathrm{e}\big(I_\xi(W)\big), \tag{4.12}$$

where we identified the probability distributions for the work in the forward process $P(W)$, and in the time-reversal one $\tilde{P}(-W)$ as given in Eqs. (4.6)-(4.7), respectively. From this, we obtain the interference term:

$$\begin{aligned} I_\xi(W) =& \frac{\alpha_0^*\alpha_1\langle 0|\xi\rangle\langle\xi|1\rangle}{\mathcal{P}(\xi)} \sum_{n,m}\sum_{n',m'} \sqrt{p_{n,m}\, p_{n',m'}}\, e^{-\frac{\beta}{2}(W_{n',m'}-\Delta F)}\; e^{-i(\Phi_{n,m}+\Phi_{n',m'})}\, \langle E_m^{(\tau)}|\,\Theta\,|E_{n'}^{(0)}\rangle \\ & \langle\varepsilon_n^{(0)}|\varepsilon_{m'}^{(\tau)}\rangle \cdot \delta(W - W_{n,m})\, \delta(W - W_{n',m'}). \end{aligned} \tag{4.13}$$

The functional dependence of $\mathcal{P}_\xi(W)$ on W consists of two parts: *i.* an "incoherent" part, reflecting the fact that each work value W obtained in the scheme is compatible with running the process in one or the other temporal direction with a given probability (*i.e.*, investing the work W when running the protocol Λ, and extracting the same amount of work $-W$ when executing its time-reversal counterpart $\tilde{\Lambda}$), and *ii.* a "coherent" part, which is a genuinely quantum feature arising from the superposition of the two temporal directions of the quench.

We recall that, from a quantum-mechanical perspective, there is, a priori, no preferential temporal direction. Moreover, in the case $|\alpha_0| = |\alpha_1| = 1/\sqrt{2}$, the forward state $|\Xi_0^\xi\rangle$ and the time-reversal one $|\Xi_1^\xi\rangle$ in Eq. (4.11) have the same amplitudes in the superposition. However, as in the standard scenario of well-defined temporal directions 4.6.1, one may use the properties of the work probability distribution $\mathcal{P}_\xi(W)$ together with Bayesian reasoning to infer the time's arrow of the thermodynamic process. As we will see shortly, in some cases, the thermodynamic time's arrow can be determined even in a single realisation of the process, which effectively projects the state $|\Psi_W^\xi\rangle_{S,E,A}$ onto either its forward or its time-reversal component.

4.3 Results

4.3.1 Effective Projection onto a Definite Time's Arrow

In the following, we demonstrate that measuring work values such that $W - \Delta F \gg \beta^{-1}$, or $W - \Delta F \ll -\beta^{-1}$, in single realisations of the extended TPM scheme effectively results in projecting the state $|\Psi^{\xi}_{W}\rangle_{S,E,A}$ in Eq. (4.10) onto either the forward or the time-reversal components in Eq. (4.11) (*i.e.*, $|\Xi^{\xi}_{0}\rangle$ or $|\Xi^{\xi}_{1}\rangle$, respectively). In order to show this, we consider the probabilities for the superposition state $|\Psi^{\xi}_{W}\rangle_{S,E,A}$ to be found in either $||\,|\Xi^{\xi}_{0}\rangle\,||^2$ or $||\,|\Xi^{\xi}_{1}\rangle\,||^2$, respectively. In particular, we notice that the term $||\,|\Xi^{\xi}_{1}\rangle\,||^2$ is upper bounded by

$$\begin{aligned} ||\,|\Xi^{\xi}_{1}\rangle\,||^2 &= |\alpha_0|^2 \big|\langle\xi|0\rangle\big|^2 \sum_{n,m} p_{n,m}\, e^{-\beta(W_{n,m}-\Delta F)} \\ &\cdot \delta(W - W_{n,m}) \leqslant e^{-\beta W_{\text{diss}}} \sum_{n,m} p_{n,m} = e^{-\beta W_{\text{diss}}}, \end{aligned} \tag{4.14}$$

where we used the fact that $|\alpha_0|^2\big|\langle\xi|0\rangle\big|^2 \leqslant 1$, and $\sum_{n,m} p_{n,m} = 1$. Consequently, in the limit $\beta W_{\text{diss}} \gg 1$, we have $||\,|\Xi^{\xi}_{1}\rangle\,||^2 \approx 0$, and hence $||\,|\Xi^{\xi}_{0}\rangle\,||^2 \approx 1$, that is, $|\Psi^{\xi}_{W}\rangle_{S,E,A} \simeq |\Xi^{\xi}_{0}\rangle$. Indeed, applying the detailed fluctuation theorem in Eq. (4.1) to Eq. (4.12), we obtain:

$$\begin{aligned} \mathcal{P}_{\xi}(W) &= P(W)\big(q^{\xi}_{0} + q^{\xi}_{1} e^{-\beta W_{\text{diss}}}\big) + 2\,\mathbb{Re}\big(I_{\xi}(W)\big) \\ &\approx q^{\xi}_{0}\, P(W), \end{aligned} \tag{4.15}$$

where we made use of the fact that $I_{\xi}(W) \propto e^{-\beta W_{\text{diss}}/2}$. Therefore, we obtained that, whenever one performs a measurement of the work in the extended TPM scheme and observes $W - \Delta F \gg \beta^{-1}$ (or, equivalently, $\Delta S = \beta W_{\text{diss}} \gg 1$), the state of the system is projected onto the forward component of the quantum superposition without measuring the auxiliary qubit (similarly to what one would obtain, had one projected the joint state $|\Psi(t)\rangle_{S,E,A}$ through a projective measurement $|0\rangle\langle 0|_A$ on the auxiliary system, and subsequently observed the work value W). The probability to observe this work value in the extended TPM scheme is given by Eq. (4.15).

Analogously, whenever the result of the extended TPM scheme is such that $W - \Delta F \ll -\beta^{-1}$ (or, equivalently, $\Delta S = \beta W_{\text{diss}} \ll -1$), one can neglect the term $||\,|\Xi^{\xi}_{0}\rangle\,||^2 \leq e^{\beta W_{\text{diss}}}$, and thus obtain the projection $|\Psi^{\xi}_{W}\rangle_{S,E} \simeq |\Xi^{\xi}_{1}\rangle$. In this case, we correspondingly achieve:

$$\begin{aligned} \mathcal{P}_{\xi}(W) &= \tilde{P}(-W)\big(q^{\xi}_{0} e^{\beta W_{\text{diss}}} + q^{\xi}_{1}\big) + 2\,\mathbb{Re}\big(I_{\xi}(W)\big) \\ &\approx q^{\xi}_{1}\, \tilde{P}(-W). \end{aligned} \tag{4.16}$$

Hence, here the joint state is projected onto the time-reversal component of the quantum superposition (as if a projective measurement $|1\rangle\langle 1|_A$ on the auxiliary system was performed, followed by the

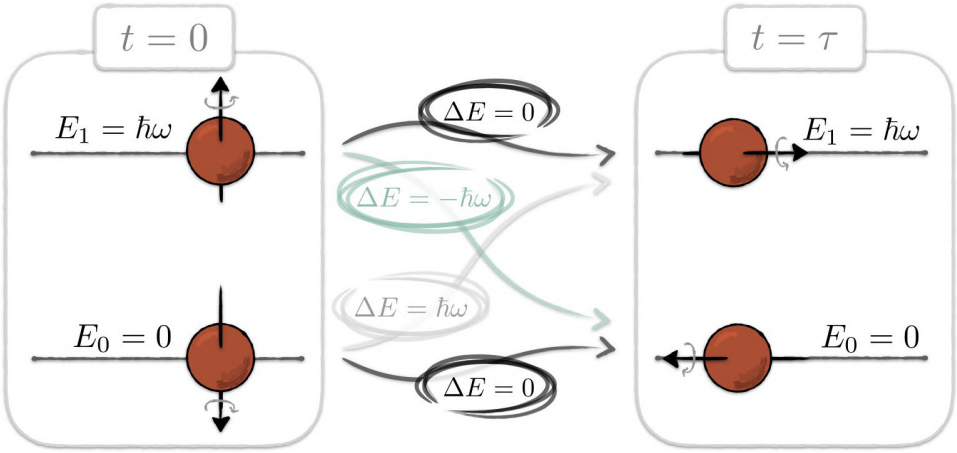


Figure 4.2 Schematic representation of the two-point measurement scheme in the forward process for our spin-$\frac{1}{2}$ system. A spin-$\frac{1}{2}$ particle in the thermal state of the initial Hamiltonian is measured in its eigenbasis $\{|z_\pm\rangle\}$ at time $t = 0$. After the action of the quench described by the time-dependent Hamiltonian in Eq. (4.19), it is measured in the eigenbasis $\{|x_\pm\rangle\}$ of the final Hamiltonian at time $t = \tau$. Depending on the measured states at the two times, the thermodynamic quench causes an energy change $\Delta E = 0, \pm\hbar\omega$, with ω being the spin's natural frequency, and $\hbar$ the reduced Planck constant.

observation of the work value W). Similarly to the previous case, Eq. (4.16) provides the probability to get such an outcome in an estimation of the work.

4.3.2 Interference Effects in the Work Distribution

In the previous section we observed that, for individual runs of the process' superposition, whenever the observed entropy production is of the order $|\Delta S| \gg 1$ (or, equivalently, $|W - \Delta F| \gg \beta^{-1}$), the system is effectively projected onto a state with a definite thermodynamic time's arrow. Conversely, if the measured entropy production is $|\Delta S| \lesssim 1$ (or equivalently $|W - \Delta F| \lesssim \beta^{-1}$), the superposition state Eq. (4.10) resulting from the application of the extended TPM scheme lacks a definite time's arrow, exhibiting interference effects.

As an illustrative example, we study the effect of interference in the work distribution in the case of a spin-$\frac{1}{2}$ system, as illustrated in Fig. 4.2. In particular, in the forward quench, the spin system is subjected to a magnetic field whose direction is rotating within the $x-z$ plane at constant angular velocity Ω around the y-axis (ω being the spin's natural frequency) $H(\Omega t) = \frac{\hbar\omega}{2}\left[\mathbb{1} + \cos(\Omega t)\,\sigma_z + \sin(\Omega t)\,\sigma_x\right]$. In the extended TPM scheme, we superpose the forward quench and its time-reversal twin, and we project the auxiliary system onto the diagonal basis $\left\{|\pm\rangle_A = (|0\rangle_A \pm |1\rangle_A)/\sqrt{2}\right\}$. This leads to the work probability distributions $\mathcal{P}_\pm(W)$, which illustrates the role played by the interference term. In the limit of a rapid quench ($\omega \ll \Omega$) (and hence of a large degree of irreversibility), the distributions are presented in Fig. 4.3 (yellow and blue bars), together with the one corresponding to a classical mixture of the forward and time-reversal processes (turquoise bars), where here $P(W) = \tilde{P}(-W)$.

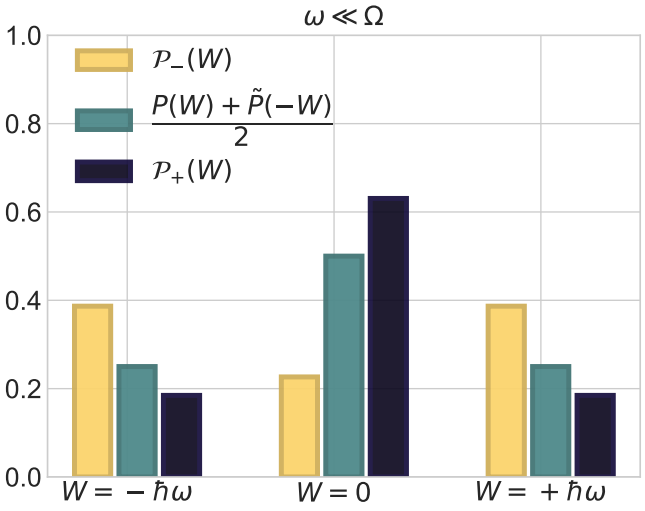


Figure 4.3 Work probability distribution for a spin-1/2 system undergoing a superposition of forward and its time-reversal thermodynamic process. The coherent work probabilities $\mathcal{P}_{\pm}(W)$ and the work probabilities of a classical mixture $\big(P(W)+\tilde{P}(-W)\big)/2$ are compared in the limit of the rapid quench $\omega \ll \Omega$ for $\varphi = \pi$. The results are temperature-independent.

While the classical mixture displays large fluctuations in the work probability distributions, the contribution of the interference term in $\mathcal{P}_{\pm}(W)$ can sharpen [$\mathcal{P}_+(W)$] or flatten [$\mathcal{P}_-(W)$] the coherent work distribution, effectively increasing or decreasing the degree of reversibility, respectively. Specifically, the probability that the process will occur in a reversible fashion (*i.e.*, that $W = 0$) is higher for $P_+(W = 0)$ [lower for $P_-(W = 0)$] than for a classical mixture (see Methods-Section 4.5.1). Since in this example, reversibility means adiabaticity, in the post-selected case, we can obtain a probability distribution $\mathcal{P}_+(W)$ corresponding to that of a slower realisation of the quench. In this sense, through our protocol, one can achieve a net "speed-up" of the realisation of an adiabatic quench.

4.3.3 Interference of cycles in a SWAP engine

Thus far, we discussed the possibility of generating a superposition of two thermodynamic processes linked together by temporal inversion, *i.e.*, a process and its time-reversal version. However, nothing prevents us from applying the methods developed above to a pair of generic processes, not necessarily temporally related to each other. In Methods-Section 4.5.2, we apply the concepts introduced above to the more practical scenario of quantum thermal devices performing thermodynamic tasks [Maslennikov *et al.*, 2019; Peterson *et al.*, 2019; Roßnagel *et al.*, 2016; Van Horne *et al.*, 2020]. In particular, we superpose two processes that correspond to two different modes of operation of a thermal machine, namely, a heat engine and a refrigerator (see Fig. 4.6). To this end, we build upon one of the simplest models of a cyclic SWAP engine operating with just two qubits and using two thermal reservoirs at

different temperatures [Campisi *et al.*, 2015].

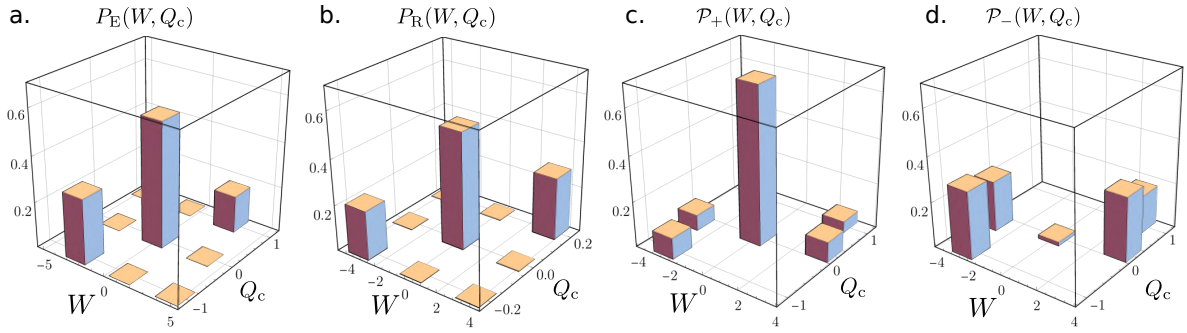


Figure 4.4 Joint probability distributions of performing work W and absorbing heat Q_{c} from the cold reservoir. The four plots correspond to **a.** the heat engine cycle, **b.** the refrigerator cycle, **c.** the superposition of both engine and refrigerator when the auxiliary system is projected onto $|+\rangle_A$, and **d.** the same superposition when the auxiliary system is projected onto $|-\rangle_A$. We used the following set of parameters: $\epsilon_1 = \epsilon_1^* = 3.5\epsilon_2$, $\epsilon_2^* = 0.2\epsilon_2$, $k_B T_1 = 8\epsilon_2$, and $k_B T_2 = \epsilon_2$ (see Methods-Section 4.5.2).

Depending on the ratio between the energies of the two qubits, our device can either function as a heat engine, extracting work $W < 0$ out of a heat current from the hot to the cold reservoirs $Q_{\mathrm{c}} < 0$, or as a power-driven refrigerator, extracting heat from the cold reservoir $Q_{\mathrm{c}} > 0$ at the price of an input work $W > 0$. Their respective joint probability distributions $P_E(W, Q_{\mathrm{c}})$ and $P_R(W, Q_{\mathrm{c}})$ for performing work W and absorbing heat Q_{c} are shown in Fig. 4.4 Panels **a.-b.**. Moreover, with the use of a auxiliary qubit, one can run the machine in a superposition of the heat-engine and the refrigerator modes. Upon measuring the auxiliary qubit in the diagonal basis $\{|+\rangle_A, |-\rangle_A\}$, one obtains processes whose joint probability distributions $\mathcal{P}_{\pm}(W, Q_{\mathrm{c}})$ display interference effects, as shown in Fig. 4.4 Panels **c.-d.**. More specifically, the interference can increase or reduce the probability of the device not performing any task (*i.e.*, $W = Q_{\mathrm{c}} = 0$), in a way that no convex mixture of the two processes is able to match, $\mathcal{P}_{+}(0,0) > |\alpha_0|^2 \mathcal{P}_H(0,0) + |\alpha_1|^2 \mathcal{P}_R(0,0) > \mathcal{P}_{-}(0,0)$. This implies that the interference effect in $\mathcal{P}_{-}(W, Q_{\mathrm{c}})$ can diminish the probability that the machine fails to perform either of the two tasks, effectively increasing the single-shot performance of the resulting process. It should be noted that this enhancement in performance cannot be found on average, since both distributions $\mathcal{P}_{\pm}(W, Q_{\mathrm{c}})$ contribute to that. However, the benefit of the post-selection is, as in its standard application in quantum information, that depending on the result obtained on the auxiliary system, one can conditionally run the machine with an improved performance.

4.4 Discussion

Viewed in isolation, a thermodynamic system coupled to a reservoir undergoes a dynamic which is generally non-unitary, even though the joint state of the system and the environment evolves in a unitary, reversible fashion. Depending on whether this dynamics favours events involving a positive or a negative change in the total entropy, it is possible to establish the temporal direction of the quench

which the system has been subjected to (*i.e.*, the time's arrow is aligned along the direction where the total entropy increases [1]). However, it can be expected that the joint state of the system and the environment may as well evolve in an arbitrary superposition of the two, whereby the direction of evolution is controlled by a further quantum system. The core questions behind this work are *i.* how a definite (thermodynamic) arrow of time can emerge in such a picture, and *ii.* whether forward-in-time and backward-in-time thermodynamic processes can interfere and what the signature of this shall be.

In this work, we showed that the coherence between the two temporal directions is effectively lost when the entropy production in the process is measured: the observation of a large increase (decrease) of dissipative work effectively projects the system in the forward (time-reversal) temporal direction. Conversely, for small values of the observed dissipative work (of the order of β^{-1}), the system and the auxiliary state display interference effects. This aspect bears important implications, insofar as, by measuring the state of the control, the system can exhibit a work (entropy production) distribution which is classically impossible. This feature can be best observed when both the forward and the time-reversal processes are, to a high degree, irreversible (*i.e.*, the probability of zero entropy production is low). In this case, indeed, the quantum superposition between the two irreversible processes can result in a dynamics which is no longer such (*i.e.*, the above probability can be significantly increased due to constructive interference). Formally, this means that when the distribution of the work $\mathcal{P}_{\pm}(W)$ is affected by interference effects, this can result in a probability distribution radically different from any classic mixture of $\mathcal{P}(W)$ and $\tilde{\mathcal{P}}(-W)$. As a consequence, $\mathcal{P}_{\pm}(W)$ does not generally satisfy the fluctuation theorem (4.1). This is not extremely surprising given that the process generating $\mathcal{P}_{\pm}(W)$ does not verify the requirements needed for the work fluctuation theorems. In particular, the initial state in Eq. (4.5) is not a thermal state neither of the system alone, nor of the system together with the control, and the work performed is defined differently in the two quenches of the superposition. Nevertheless, this violation has a crucial implication: it entails that the distribution $\mathcal{P}_{\pm}(W)$ cannot be generated by any thermodynamic process starting in equilibrium with the environment, and being subsequently driven out of it by means of any given protocol Λ. Consequently, our procedure provides a recipe to generate thermodynamic processes with a work probability distribution which cannot be reproduced within the standard framework of fluctuation theorems.

Finally, we also presented how the aforementioned results can be extended beyond the case of superpositions of mutually time-reversal processes. We exemplified this by studying the quantum superposition between two alternative thermodynamic tasks, namely, a thermal machine running in a quantum superposition of a heat engine, and a power-driven refrigeration. In this regard, we proved that our findings can also lead to practical implications, as interference effects can be used to reduce undesired fluctuations, paving the way for their adoption in more general contexts for enhancing the performance of standard thermodynamic engines.

[1] Notice that, under our sign convention, this means that the time's arrow matches a positive entropy change in the case of the forward process, and a negative entropy change for the time-reversal process.

4.5 Methods

4.5.1 Case Study: A Spin-$\frac{1}{2}$ System

In this section, we detail on the interference effects between forward and time-reversal thermodynamic evolution of a spin-$\frac{1}{2}$ system. To this end, we further develop the general expression of Eq. (4.12). Specifically, we project the auxiliary system onto the diagonal basis $|\xi\rangle_A = \Big\{|\pm\rangle_A = (|0\rangle_A \pm |1\rangle_A)/\sqrt{2}\Big\}$. This leads to the joint state of the system and the environment $|\Psi_W^{\pm}\rangle_{S,E,A} \equiv \big(\mathbb{1}_{S,E} \otimes |\pm\rangle\langle\pm|_A\big) \circ M_W|\Psi_0\rangle_{S,E,A}$.

The corresponding post-selected work probability distribution, conditioned on the projection of the auxiliary system onto $|\pm\rangle_A$, reads:

$$\mathcal{P}_{\pm}(W) = q_0^{\pm} P(W) + q_1^{\pm} \tilde{P}(-W) + 2\,\mathbb{Re}\big(I_{\pm}(W)\big), \tag{4.17}$$

where the interference term $I_{\pm}(W)$ is given by Eq. (4.13) with $\langle 0|\pm\rangle_A = 1/\sqrt{2}$ and $\langle \pm|1\rangle_A = \pm 1/\sqrt{2}$. We recall that the states $\Theta\,|E_n^{(0)}\rangle$ in the above expressions are the eigenstates of the Hamiltonian $\Theta H[\lambda(0)]\Theta^{\dagger} = H[\tilde{\lambda}(0)]$. Moreover, we notice that the distribution $\mathcal{P}_{\pm}(W)$ in Eq. (4.17) differs by the term $I_{\pm}(W) \neq 0$ from what one would have obtained by applying the extended TPM scheme to a (classical) convex mixture $|\alpha_0|^2\,|0\rangle\langle 0|_A \otimes \rho_0^{\text{th}} + |\alpha_1|^2\,|1\rangle\langle 1|_A \otimes \tilde{\rho}_0^{\text{th}}$ of the initial states.

For the outcome $W = 0$, the interference term in Eq. (4.13) can be simplified when $\Delta F = 0$, and the sets of eigenvalues of the initial and final Hamiltonians coincide, *i.e.*, $E_n^{(0)} = E_n^{(\tau)}$. In that case:

$$\begin{aligned} I_{\pm}(W=0) = \pm\,\frac{\alpha_0^*\alpha_1}{2\,\mathcal{P}(\xi)} \sum_{n,n'} \sqrt{p_{n,n}p_{n',n'}}\, e^{-i(\Phi_{n,n}+\Phi_{n',n'})} \\ \langle \varepsilon_{n'}^{(0)}|\varepsilon_n^{(\tau)}\rangle\ \langle E_{n'}^{(\tau)}|\,\Theta\,|E_n^{(0)}\rangle\,. \end{aligned} \tag{4.18}$$

As a result, it emerges that the interference effects can increase (decrease) the probability of observing the work value $W = 0$. This yields to a work probability distribution $\mathcal{P}_{\pm}(W)$ analogous to the one potentially generated by a more reversible (irreversible) process than the forward and time-reversal processes themselves, or any classical mixture therefrom. We remark that the interference term $I_{\pm}(W)$ may show non-zero values for $W \neq 0$ in general, as we will see below.

We conclude by evaluating Eq. (4.17) in the concrete example sketched in the main text. We consider a spin system with natural frequency ω in a magnetic field $\vec{\lambda}(t)$ whose direction is rotating within the $x-z$ plane at constant angular velocity around the y-axis:

$$H\big[\vec{\lambda}(t)\big] = \frac{\hbar\omega}{2}\Big[\mathbb{1} + \vec{\lambda}(t)\cdot\vec{\sigma}\Big] = \frac{\hbar\omega}{2}\Big[\mathbb{1} + \cos(\Omega t)\,\sigma_z + \sin(\Omega t)\,\sigma_x\Big], \tag{4.19}$$

where $\vec{\lambda}(t) = (\lambda_0 \sin(\Omega t), 0, \lambda_0 \cos(\Omega t))$ and $\lambda_0 = 1$ is the dimensionless magnetic field, and where the protocol reads $\Lambda = \{\vec{\lambda}(t)\ ;\ 0 \leq t \leq \pi/(2\Omega)\}$. We notice that $\Theta H[\vec{\lambda}(t)]\Theta^\dagger = H[-\vec{\lambda}(t)]$, implying that the time-reversal of the control parameter corresponds to a flip of the magnetic field. At the initial and final times of the protocol, the Hamiltonian is diagonal in the $\{|z_\pm\rangle\}$ and $\{|x_\pm\rangle\}$ bases, respectively. Therefore, $|E_n^{(0)}\rangle = \{|z_\pm\rangle_S\}$, with corresponding eigenvalues $E_n^{(0)} = \{0, \hbar\omega\}$, and $|E_m^{(\tau)}\rangle = \{|x_\pm\rangle_S = \frac{1}{\sqrt{2}}(|z_-\rangle_S \pm |z_+\rangle_S)\}$, with eigenvalues $E_m^{(\tau)} = \{0, \hbar\omega\}$ (we shifted the lower energy level by $\hbar\omega/2$ to avoid negative energy eigenvalues). As a result, $F_0 = F_\tau = -\log(1+e^{-\beta\hbar\omega})$ and $W_{n,m} = \{\hbar\omega, 0, -\hbar\omega\}$.

In the frame rotating around the y-axis at frequency Ω, the Hamiltonian becomes time-independent, and the unitary governing the evolution can be obtained straightforwardly. Turning back to the Schrödinger picture, the applied unitary $U(t,0)$ reads:

$$U(t,0) = e^{-\frac{i}{2}\Omega\sigma_y t} e^{-\frac{i}{2}[\,\omega\,(\mathbb{1}+\sigma_z)-\Omega\sigma_y\,]\,t}. \tag{4.20}$$

This is used below to compute the work distribution.

4.5.1.1 Effect of Interference on Reversibility

In this subsection, we will represent the environment as a spin-$\frac{1}{2}$ system which is left unaffected during the quench. For instance, we can assume that the purification of the thermal states in Eq. (4.4a)-(4.4b) read

$$|\psi_0\rangle_{S,E} = \sqrt{\frac{1}{Z_0}}\,|z_-\rangle_S\,|z_-\rangle_E + \sqrt{\frac{e^{-\beta\hbar\omega}}{Z_0}}\,|z_+\rangle_S\,|z_+\rangle_E, \tag{4.21a}$$

$$|\psi_\tau\rangle_{S,E} = \sqrt{\frac{1}{Z_0}}\,|x_-\rangle_S\,|z_-\rangle_E + \sqrt{\frac{e^{-\beta\hbar\omega}}{Z_0}}\,|x_+\rangle_S\,|z_+\rangle_E. \tag{4.21b}$$

Furthermore, we will assume to begin the protocol in the state in Eq. (4.5) with $\alpha_0 = 1/\sqrt{2}$, $\alpha_1 = e^{-i\varphi}/\sqrt{2}$, with φ being a controllable phase between the forward and the time-reversal processes.

Next, we compute $\mathcal{P}_\pm(W)$:

$$\begin{aligned}\mathcal{P}_\pm(W=0) &= \frac{1}{2\mathcal{P}(\pm)}(p_{0,0}+p_{1,1}) \\ &\mp \frac{1}{2\sqrt{2}\mathcal{P}(\pm)}\big[\,p_{0,0}\cos(2\Phi_{0,0}+\varphi) + p_{1,1}\cos(2\Phi_{1,1}+\varphi)\big],\end{aligned} \tag{4.22}$$

where we used the fact that $\langle E_{n'}^{(\tau)}|\Theta|E_n^{(0)}\rangle = 1/\sqrt{2}$ for $n=1, n'=0$, and $\langle E_{n'}^{(\tau)}|\Theta|E_n^{(0)}\rangle = -1/\sqrt{2}$

otherwise [2], whereas $\langle\varepsilon_{n'}^{(0)}|\varepsilon_{n}^{(\tau)}\rangle_E = \delta_{n,n'}$, and where the marginal probability of the auxiliary system reads $\mathcal{P}(\pm) = \frac{1}{2} \pm \frac{1}{2\sqrt{2}}\big[\, p_{0,0}\cos\big(2\Phi_{0,0}+\varphi\big) + p_{1,1}\cos\big(2\Phi_{1,1}+\varphi\big)\big]$, with $p_{0,0} = \frac{|\langle x_-|\,U(\tau,0)|z_-\rangle|^2}{1+e^{-\beta\hbar\omega}}$, $e^{i\Phi_{0,0}} = \frac{\langle x_-|\,U(\tau,0)|z_-\rangle}{\sqrt{|\langle x_-|\,U(\tau,0)|z_-\rangle|}}$, and $p_{1,1} = \frac{|\langle x_+|\,U(\tau,0)|z_+\rangle|^2}{1+e^{-\beta\hbar\omega}}\, e^{-\beta\hbar\omega}$, $e^{i\Phi_{1,1}} = \frac{\langle x_+|\,U(\tau,0)|z_+\rangle}{\sqrt{|\langle x_+|\,U(\tau,0)|z_+\rangle|}}$. From this result, we deduce that it is possible to observe interference between thermodynamic processes occurring in the forward and time-reversal temporal directions. Following the same procedure for the cases $W = \pm\hbar\omega$, we get

$$\mathcal{P}_\pm(W=\hbar\omega) = \frac{p_{0,1}}{4\mathcal{P}(\pm)}\big(\,1+e^{-\beta\hbar\omega}\big), \tag{4.23a}$$

$$\mathcal{P}_\pm(W=-\hbar\omega) = \frac{p_{1,0}}{4\mathcal{P}(\pm)}\big(\,1+e^{\beta\hbar\omega}\big), \tag{4.23b}$$

which do not feature interference. In the last expressions, $p_{0,1} = \frac{|\langle x_+|\,U(\tau,0)|z_-\rangle|^2}{1+e^{-\beta\hbar\omega}}$, $e^{i\Phi_{0,1}} = \frac{\langle x_+|\,U(\tau,0)|z_-\rangle}{\sqrt{|\langle x_+|\,U(\tau,0)|z_-\rangle|}}$, and $p_{1,0} = \frac{|\langle x_-|\,U(\tau,0)|z_+\rangle|^2}{1+e^{-\beta\hbar\omega}}\, e^{-\beta\hbar\omega}$, $e^{i\Phi_{1,0}} = \frac{\langle x_-|\,U(\tau,0)|z_+\rangle}{\sqrt{|\langle x_-|\,U(\tau,0)|z_+\rangle|}}$. We illustrate the probability distribution in Eq. (4.22)-(4.23) in Fig. 4.3 of the main text.

4.5.1.2 Interference Terms for Varying $\pm\hbar\omega$

In the previous case study, we represented the environment as a spin-$\frac{1}{2}$ system which is left unmodified by the thermodynamic quench. This caused the cancellation of all interference terms in $\mathcal{P}_\pm(W = \pm\hbar\omega)$. In this subsection, on the contrary, we suppose that the environment undergoes a spin-flip during the quench:

$$|\psi_0\rangle_{S,E} = \sqrt{\frac{1}{Z_0}}\,|z_-\rangle_S\,|z_-\rangle_E + \sqrt{\frac{e^{-\beta\hbar\omega}}{Z_0}}\,|z_+\rangle_S\,|z_+\rangle_E, \tag{4.24a}$$

$$|\psi_\tau\rangle_{S,E} = \sqrt{\frac{1}{Z_0}}\,|x_-\rangle_S\,|z_+\rangle_E + \sqrt{\frac{e^{-\beta\hbar\omega}}{Z_0}}\,|x_+\rangle_S\,|z_-\rangle_E. \tag{4.24b}$$

This change results in $\langle\varepsilon_{n'}^{(0)}|\varepsilon_{n}^{(\tau)}\rangle_E = 0$, for $n' = n$. For the sake of simplicity, below we will also set $\varphi = \pi$.

The three probabilities discussed in the previous section become therefore:

$$\begin{aligned}
\mathcal{P}_\pm(W=0) &= \frac{1}{2\mathcal{P}(\pm)}\big(p_{0,0}+p_{1,1}\big), \\
\mathcal{P}_\pm(W=\hbar\omega) &= \frac{p_{0,1}}{4\mathcal{P}(\pm)}\Big[\,1+e^{-\beta\hbar\omega} \pm \sqrt{2}\,e^{-\frac{\beta\hbar\omega}{2}}\cos\big(2\Phi_{0,1}\big)\Big], \\
\mathcal{P}_\pm(W=-\hbar\omega) &= \frac{p_{1,0}}{4\mathcal{P}(\pm)}\Big[\,1+e^{\beta\hbar\omega} \mp \sqrt{2}\,e^{\frac{\beta\hbar\omega}{2}}\cos\big(2\Phi_{1,0}\big)\Big],
\end{aligned} \tag{4.25}$$

[2] To evaluate $\langle E_{n'}^{(\tau)}|\Theta|E_n^{(0)}\rangle_S$, we made use of the fact that the time-reversal operator Θ for a spin-$\frac{1}{2}$ system acts as $\Theta = i\sigma_y K$, where K is the complex conjugation operator. Thus, $\Theta|z_\pm\rangle = \mp|z_\mp\rangle$.

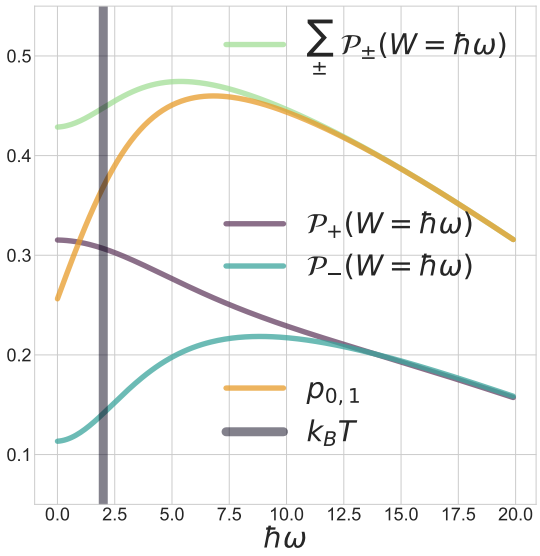


Figure 4.5 Work probabilities of a spin-1/2 system under the time-dependent Hamiltonian with a varying amount $\hbar\omega$ of work invested. For values of $\hbar\omega$ smaller or of the order of $\beta^{-1} = k_BT = 1/2$ ($k_B = \hbar = 1$), the work probabilities $\mathcal{P}_+(W = \hbar\omega)$ and $\mathcal{P}_-(W = \hbar\omega)$ (see Eq. (4.25); turquoise and purple curves) strongly depend on the interference terms. For values $\hbar\omega \gg \beta^{-1}$, $\mathcal{P}_+(W = \hbar\omega) + \mathcal{P}_-(W = \hbar\omega)$ (green curve) tends to the value $p_{0,1}$ (yellow curve), which is obtained by projecting the process to the forward direction and obtaining the work difference $\hbar\omega$. This illustrates that observing large work values $\hbar\omega \gg \beta^{-1}$ ($\hbar\omega \ll -\beta^{-1}$) effectively projects the process onto the forward (time-reversal) direction.

where the marginal probability of the auxiliary system is now $\mathcal{P}(\pm) = \frac{1}{2} \pm \frac{1}{2\sqrt{2}} \left[p_{0,1}\, e^{-\frac{\beta\hbar\omega}{2}} \cos(2\Phi_{0,1}) - p_{1,0}\, e^{\frac{\beta\hbar\omega}{2}} \cos(2\Phi_{1,0}) \right]$, and where $p_{0,0}$, $\Phi_{0,0}$, $p_{1,1}$, and $\Phi_{1,1}$ are the same as in case study 4.5.1.1.

In Fig. 4.5, we show the work probability distributions for varying $\hbar\omega$. For work values $\hbar\omega$ smaller than, or of the order of β^{-1}, we observe strong interference effect, as shown by the difference between $\mathcal{P}_+(W = \hbar\omega)$ and $\mathcal{P}_-(W = \hbar\omega)$. For work values $\hbar\omega \gg \beta^{-1}$, this difference vanishes, and the probability $\mathcal{P}(W = \hbar\omega) := \mathcal{P}_+(W = \hbar\omega) + \mathcal{P}_-(W = \hbar\omega)$ to obtain the work value $\hbar\omega$ tends to the probability $p_{0,1}$ of first projecting the auxiliary system onto the forward direction, and then obtaining the work value $\hbar\omega$. This trend shows that the observation of large work values effectively projects the system into a well-defined temporal direction.

4.5.2 Superposing cycles in a quantum heat engine

We consider one of the simplest models of a cyclic heat engine operating with two strokes [Campisi *et al.*, 2015]. This consists of two qubits only, each with energy eigenstates $\{|0\rangle_i, |1\rangle_i\}$ for $i = 1, 2$, and with different energy spacing, $H_1 = \epsilon_1 |1\rangle\langle 1|_1$ and $H_2 = \epsilon_2 |1\rangle\langle 1|_2$. The engine works as follows. In the first stroke, the qubits start in equilibrium at different inverse temperatures $\beta_1 = 1/k_BT_1$ and

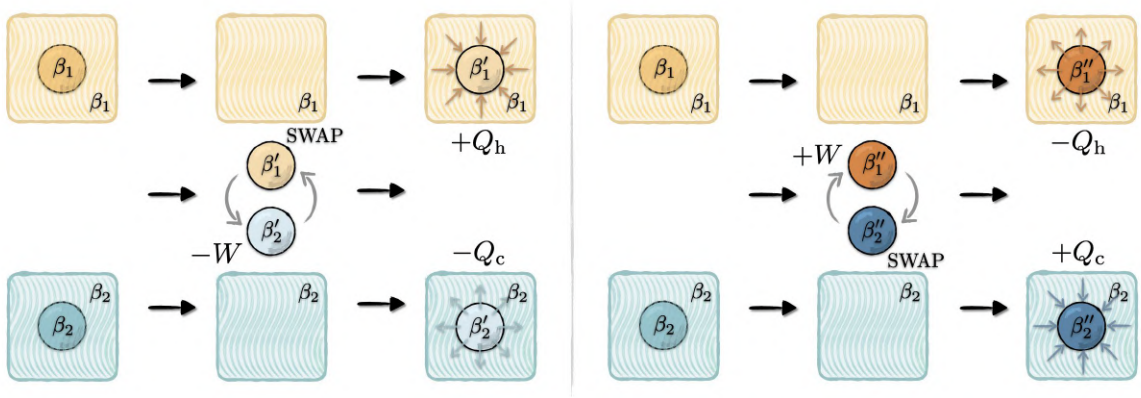


Figure 4.6 Schematic representation of the SWAP engine cycle operating as a heat engine (*Left*) and as a refrigerator (*Right*). Two qubits (circles) start the cycle in equilibrium with their corresponding reservoirs (squares) at inverse temperature β_1 and β_2. Subsequently, a SWAP operation is applied to the two systems. In the heat engine mode of operation **(*Left*)**, a work $-W$ is extracted leading to a decrease of temperature in the first qubit, $\beta_1' > \beta_1$, and an increase in the second one, $\beta_2' < \beta_2$. The qubits are then equilibrated with their respective reservoirs, resulting in a net transfer of heat from the hot reservoir $+Q_{\rm h}$ and a release of heat into the cold one $-Q_{\rm c}$. Vice versa, in the refrigeration mode **(*Right*)**, a work $+W$ is externally supplied to overheat the first qubit $\beta_1'' < \beta_1$, and reduce the temperature of the second one, $\beta_2'' > \beta_2$. This leads to a final release of heat into the hot reservoir $-Q_{\rm h}$, and the absorption of heat from the cold one $+Q_{\rm c}$, hence refrigerating it.

$\beta_2 = 1/k_B T_2$, where we assume $\beta_1 \geq \beta_2$ for concreteness, and they interact through a unitary SWAP operation of the form:

$$U_{\rm SWAP} = \begin{pmatrix} 1&0&0&0 \\ 0&0&1&0 \\ 0&1&0&0 \\ 0&0&0&1 \end{pmatrix}. \tag{4.26}$$

In the second stroke, the qubits are put in contact with respective local reservoirs at β_1 and β_2 until they thermalize. The cyclic application of these two stokes leads to a systematic operation of the device which consumes/extracts an average amount of work $\langle W \rangle$ in the first stroke, while absorbing/releasing an amount of heat $\langle Q_i \rangle$ $i = {\rm h}, {\rm c}$ from the hot and cold thermal environments in the second ones. This device functions as a heat engine, $\langle W \rangle \leq 0$, whenever $\beta_1/\beta_2 < \epsilon_2/\epsilon_1 < 1$. On the other hand, a power-driven refrigerator extracting heat from the cold reservoir, $\langle Q_c \rangle \geq 0$, can be obtained when $\epsilon_2/\epsilon_1 < \beta_1/\beta_2$. The functioning of the thermal engine is illustrated in Fig. 4.6.

The simplicity in the operation of this device makes it specially appealing for the extended TPM scheme introduced above. We consider a bipartite system S consisting on the two engine qubits, and we identify the thermodynamic process with the SWAP operation in Eq. (4.26). In order to cast the operation of the engine in the form of a standard thermodynamic protocol Λ, we introduce an interaction Hamiltonian between the two qubits of the form $H_{\rm int} \equiv g(|01\rangle\langle 10| + |01\rangle\langle 10|)$. This interaction term is switched on and off at the initial and final times of the protocol, $t = 0$ and $t = \tau$, respectively, *i.e.*, $H[\lambda(t)] = H_1 + H_2 + \lambda(t)\, H_{\rm int}$, with $\lambda(t) = 1\ \forall\, t \in (0, \tau)$, while $\lambda(0) = \lambda(\tau) = 0$. By tuning the coupling strength g and the final time τ, the quench $U(\tau, 0) = \exp\left[-i(H_1 + H_2 + H_{\rm int})\tau/\hbar\right]$ reproduces the SWAP unitary in Eq. (4.26), which belongs to the class of optimal unitaries for work extraction in the setup [Campisi *et al.*, 2015].

We now consider the superposition of two cycles, both of them as described above with the very same quench $U_{\rm SWAP}$. However, in the first one, we denote the energy spacings of the qubits ϵ_1 and ϵ_2, chosen such that $\beta_1/\beta_2 < \epsilon_2/\epsilon_1 < 1$ (work extraction). Instead, in the second cycle, the spacings of the qubits are modified to be ϵ_1^* and ϵ_2^*, where we chose the new ones to verify $\epsilon_2^*/\epsilon_1^* < \beta_1/\beta_2$ (refrigeration). In any case, the temperatures are fixed, and the two qubits start the evolution in a bipartite thermal state of the form:

$$\rho_0^{\rm th} = \frac{e^{-\beta_1 H_1}}{Z_1} \otimes \frac{e^{-\beta_2 H_2}}{Z_2}, \tag{4.27}$$

where $Z_1 = 1 + e^{-\beta_1 \epsilon_1}$ and $Z_2 = 1 + e^{-\beta_2 \epsilon_2}$ are the local partition functions for each qubit. Since the protocol is cyclic, $H(0) = H(\tau)$, the initial and final system eigenstates coincide $|E_k^{(0)}\rangle = |E_k^{(\tau)}\rangle$, so that the superscripts can be dropped for simplicity. The purified initial states of system and environment are then given by:

$$|\psi_0^{\rm E}\rangle_{S,E} = \sum_{k,l=0,1} \sqrt{\frac{e^{-\beta_1 \epsilon_1 k}}{Z_1} \frac{e^{-\beta_2 \epsilon_2 l}}{Z_2}}\, |k\rangle_1 |l\rangle_2 \, |\varepsilon_k\rangle_{R_1} |\varepsilon_l\rangle_{R_2},$$
$$|\psi_0^{\rm R}\rangle_{S,E} = \sum_{k,l=0,1} \sqrt{\frac{e^{-\beta_1 \epsilon_1^* k}}{Z_1^*} \frac{e^{-\beta_2 \epsilon_2^* l}}{Z_2^*}}\, |k\rangle_1 |l\rangle_2 \, |\varepsilon_k\rangle_{R_1} |\varepsilon_l\rangle_{R_2},$$

where we introduced the reservoir states $|\varepsilon_k\rangle_{R_1}$ and $|\varepsilon_l\rangle_{R_2}$ to purify the local mixed states in Eq. (4.27), and where we denoted the corresponding partition functions as $Z_1^* = 1 + e^{-\beta_1 \epsilon_1^*}$ and $Z_2^* = 1 + e^{-\beta_2 \epsilon_1^*}$. In order to incorporate also the auxiliary system, we consider the following initial state:

$$|\Psi_0\rangle_{S,E,A} = \alpha_0 |\psi_0^{\rm E}\rangle_{S,E} \, |0\rangle_A + \alpha_1 |\psi_0^{\rm R}\rangle_{S,E} \, |1\rangle_A \,, \tag{4.28}$$

where we associated the initial states for the heat engine and the heater to the two orthogonal states $|0\rangle_A$ and $|1\rangle_A$ of the control, respectively.

We consider an extension of the TPM scheme as in the previous sections, in which we include independent energy measurements on both engine qubits in both branches of the superposition. However, the role played before by the forward and time-reversal evolutions is now played by the heat engine and refrigeration processes. More precisely, conditional on the state $|0\rangle_A$ and $|1\rangle_A$ of the quantum control, we consider the application of the projectors $\Pi_{n_1,n_2} \equiv |n_1\rangle \langle n_1|_1 \otimes |n_2\rangle \langle n_2|_2$ and $\Pi_{n_2,n_1} \equiv |n_2\rangle \langle n_2|_1 \otimes |n_1\rangle \langle n_1|_2$ to the initial states $|\psi_0\rangle_{S,E}^{\rm E}$ and $|\tilde{\psi}_0\rangle_{S,E}^{\rm R}$, respectively. Then, we apply the SWAP unitary (4.26) in both branches. Finally, the projectors $\Pi_{m_1,m_2} = |m_1\rangle \langle m_1|_1 \otimes |m_2\rangle \langle m_2|_2$ and $\Pi_{m_2,m_1} = |m_2\rangle \langle m_2|_1 \otimes |m_1\rangle \langle m_1|_2$ are respectively applied to each branch. In this way, the changes in energy of the qubits in the work extraction cycle are $\Delta E_{n_1,m_1}^{(1)} = \epsilon_1(m_1 - n_1)$ and $\Delta E_{n_2,m_2}^{(2)} = \epsilon_2(m_2 - n_2)$, while in the refrigeration cycle they read $\Delta E_{n_2,m_2}^{(1)*} = \epsilon_1^*(m_2 - n_2)$ and $\Delta E_{n_1,m_1}^{(2)*} = \epsilon_2^*(m_1 - n_1)$.

The implementation of the scheme leads to the following operators depending on the energy changes in the two qubits:

$$M_{\Delta E_1,\Delta E_2} = \sum_{n_1,m_1}\sum_{n_2,m_2}\Big[|0\rangle\langle 0|_A \otimes \Pi_{m_1,m_2} U \Pi_{n1,n2} \otimes \mathbb{1}_E \cdot \delta(\Delta E_1 - \Delta E^{(1)}_{n_1,m_1}) \cdot \delta(\Delta E_2 - \Delta E^{(2)}_{n_2,m_2}) \\ + |1\rangle\langle 1|_A \otimes \Pi_{m_2,m_1} U \,\Pi_{n_2,n_1} \otimes \mathbb{1}_E \cdot \delta(\Delta E_1 - \Delta E^{(1)*}_{n_2,m_2}) \cdot \delta(\Delta E_2 - \Delta E^{(2)*}_{n_1,m_1})\Big]. \tag{4.29}$$

As before, these are Kraus operators fulfilling $\int d\Delta E_1\, d\Delta E_2\, M^\dagger_{\Delta E_1,\Delta E_2} M_{\Delta E_1,\Delta E_2} = \mathbb{1}$, associated to the occurrence of the outcomes ΔE_1 and ΔE_2 for the energy changes in the engine qubits during the protocol. The operator in Eq. (4.29) is then applied to the initial state in Eq. (4.28) giving rise to the (unnormalized) state $M_{\Delta E_1,\Delta E_2}\,|\Psi_0\rangle_{S,E,A}$.

Finally, by including the final projection of the auxiliary system onto the state $|\xi\rangle_A$, we can calculate the conditional probability distribution analogous to Eq. (4.12). To this end, it is useful to introduce the coefficients $q_0^\xi = |\alpha_0|^2|\langle\xi|0\rangle|^2/P(\xi)$ and $q_1^\xi = |\alpha_1|^2|\langle\xi|1\rangle|^2/P(\xi)$, where $P(\xi)$ denotes the marginal probability for post-selecting on $|\xi\rangle_{\mathrm{A}}$:

$$P(\xi) = |\alpha_0|^2\big|\langle\xi|0\rangle\big|^2 + |\alpha_1|^2\big|\langle\xi|1\rangle\big|^2 + 2\,\mathbb{Re}\big[\,\alpha_0^*\alpha_1\langle 0|\xi\rangle\langle\xi|1\rangle\,\big]\cdot f, \tag{4.30}$$

and $f := \left(\frac{1+\sqrt{e^{-\beta_1(\epsilon_1+\epsilon_1^*)-\beta_2(\epsilon_2+\epsilon_2^*)}}}{\sqrt{Z_1 Z_1^* Z_2 Z_2^*}}\right)$. The conditional probability of obtaining ΔE_1 and ΔE_2 in the interferometric scheme, given that the auxiliary system is found in the state $|\xi\rangle_A$, can then be written as:

$$P_\xi(\Delta E_1,\Delta E_2) = q_0^\xi P_{\mathrm{E}}(\Delta E_1,\Delta E_2) + q_1^\xi P_{\mathrm{R}}(\Delta E_1,\Delta E_2) + 2\,\mathbb{Re}\big[I(\Delta E_1,\Delta E_2)\big]. \tag{4.31}$$

Here, we identified the distributions of a heat engine and a refrigerator operating in independent cycles as:

$$P_{\mathrm{E}}(\Delta E_1,\Delta E_2) = \sum_{n,m}\frac{e^{-\beta_1\epsilon_1 n}}{Z_1}\frac{e^{-\beta_2\epsilon_2 m}}{Z_2}\cdot\delta(\Delta E_1 - \Delta E^{(1)}_{n,m})\cdot\delta(\Delta E_2 + \Delta E^{(2)}_{n,m}), \tag{4.32a}$$

$$P_{\mathrm{R}}(\Delta E_1,\Delta E_2) = \sum_{n,m}\frac{e^{-\beta_1\epsilon_1^* m}}{Z_1^*}\frac{e^{-\beta_2\epsilon_2^* n}}{Z_2^*}\cdot\delta(\Delta E_1 + \Delta E^{(1)*}_{n,m})\cdot\delta(\Delta E_2 - \Delta E^{(2)*}_{n,m}), \tag{4.32b}$$

whereas the interference term is:

$$I(\Delta E_1,\Delta E_2) = \frac{\alpha_0^*\alpha_1\,\langle 0|\xi\rangle\langle\xi|1\rangle}{P(\xi)}\; f\cdot\delta(\Delta E_1)\cdot\delta(\Delta E_2), \tag{4.33}$$

which only affects the case in which both energy changes are zero.

The work performed by the machine during the unitary quench U_{SWAP} equals the total energy change of the system, $W = \Delta E_1 + \Delta E_2$, which is distributed among the two qubits. This energy is dissipated

as heat into the reservoirs at the end of the cycle, so that we can identify $Q_{\rm h} = \Delta E_1$ as the heat input from the hot reservoir and $Q_{\rm c} = -\Delta E_2$ as the heat extracted from the cold one. Applying a change of variable $\Delta E_1 \to W - \Delta E_2$, and identifying $\Delta E_2 \to -Q_{\rm c}$ in $P_\xi(\Delta E_1, \Delta E_2)$ [Eq. (4.31)], we obtain the joint probability of the machine to perform work W, and extract heat $Q_{\rm c}$ from the cold reservoir during the superposition of both cycles conditioned on the projection $|\xi\rangle_A$ of the control, namely, $P_\xi(W, Q_{\rm c})$.

In Fig. 4.4 we show the probability distributions of the engine [$P_{\rm E}(W, Q_{\rm c})$], the refrigerator [$P_{\rm R}(W, Q_{\rm c})$], and the superposition of both of them [$P_\xi(W, Q_{\rm c})$, with $|\xi\rangle_A \equiv |\pm\rangle_A$], for the choice of parameters indicated in the caption. As can be appreciated in Fig. 4.4 Panels **a.**-**b.**, both the engine and refrigeration processes are characterized by a joint probability with three peaks, one corresponding to extracting work $W < 0$ and dumping heat into the cold reservoir $Q_{\rm c} < 0$ (left peak), one where the engine does not consume work nor exchange heat with the reservoirs, $W = Q_{\rm c} = 0$ (middle peak), and finally a third one where an external input of work $W > 0$ helps to refrigerate the cold reservoir $Q_{\rm c} > 0$ (right peak). The difference between the heat engine and refrigerator cycles is that, while in the former the left peak is higher than the right one, warrantying work extraction on average $\langle W \rangle < 0$, in the latter case the situation is the opposite, enforcing refrigeration on average $\langle Q_{\rm c} \rangle > 0$. However, in both cases the highest peak in the distribution is the middle one ($W = Q_{\rm c} = 0$), which implies that the machine will most probably perform no tasks at all.

Remarkably, this issue can be handled by means of the interference effects as given by Eq. (4.33). Superposing the engine and refrigeration cycles, we obtain either the constructive addition of the middle peak in both cycles, or their destructive suppression, depending on the auxiliary system post-selection. As can be seen in Figs. 4.4 Panels **c.-d.** the central peak can be either enhanced in $\mathcal{P}_+(W, Q_{\rm c})$ or almost suppressed $\mathcal{P}_-(W, Q_{\rm c})$, while the other four peaks in the probability distribution decrease or increase accordingly. We note that the other four peaks are placed as the left and right peaks in Fig. 4.4 Panels **a.**-**b.**, two of which correspond to the heat engine mode of operation, and the other two to the refrigerator mode. The case $\mathcal{P}_-(W, Q_{\rm c})$ is particularly interesting as it ensures that, in each cycle of the (superposition) process, with high probability we will perform either one useful task or the other, but certainly one of the two. The relative weights of the two tasks may be balanced at convenience by, *e.g.*, selecting different coefficients α_0 and α_1 in the initial superposition in Eq. (4.28), at the price of loosing amplitude in the interference term. By continuously varying the state $|\xi\rangle_{\rm A}$ onto which the auxiliary qubit is projected, we can interpolate between the distributions $\mathcal{P}_+(W, Q_{\rm c})$ and $\mathcal{P}_-(W, Q_{\rm c})$. Finally, it is worth noticing that this effect cannot be generated by any convex mixture of cycles, which would simply lead to a convex mixture of the probability distributions $P_{\rm E}(W, Q_{\rm c})$ and $P_{\rm R}(W, Q_{\rm c})$.

4.6 Supplemental Information

4.6.1 Guessing the Time's Direction in a Thermodynamic Process

In 1927, Sir A. Eddington introduced the notion of 'arrow of time' [Eddington, 1928] to refer to the temporal directionality that he saw as deeply rooted in the second law of thermodynamics. He explained that, according to this law, in order to determine the direction in which time is flowing for a macroscopic system subjected to an irreversible process, it is sufficient to examine the relation between the work W performed on the system, and the variation of its free energy ΔF: time must flow in the direction in which $W > \Delta F$. This apparently unequivocal description weakens in the microscopic case, where it is possible to occasionally observe 'fluctuations' from the Clausius inequality. It follows that, in the microscopic case, it is no longer possible to univocally determine the direction of time from the sign of $W - \Delta F$.

With the aim to refine these considerations, in Ref. [Jarzynski, 2011], C. Jarzynski evaluated the possibility of defining the temporal direction of a thermodynamic process from a given set of data. In the following, we go over his reasoning briefly.

Let us imagine filming a microscopic system that, subjected to a thermodynamic process $\Lambda(t)$, varies from an initial state at time $t = 0$, to a final state at time $t = \tau$. We will suppose that *i.* the camera is able to record the motion of each particle constituting the system, *ii.* we are given full knowledge of the Hamiltonian function of the system $H\big(\Lambda(t)\big)$, and of the value of $\Delta F = F_\tau - F_0$. Depending on the result obtained from the coin toss, the movie will be shown to us in either the correct or reverse order. Our goal is to determine, based on the given information, whether the movie is shown in the correct or reverse order.

This problem can be addressed using statistical inference. We call $L(F\,|\,\gamma)$ the likelihood that the process is shown in the forward direction F if the microscopic trajectory γ is shown. Likewise, $L(R\,|\,\gamma)$ is the likelihood that the process is shown in the time-reversal direction R given the microscopic trajectory γ. Obviously, the two terms sum up to one:

$$L(F\,|\,\gamma) + L(R\,|\,\gamma) = 1. \tag{4.34}$$

We call W the work performed on the system for the trajectory γ. For a macroscopic system, according to the Clausius inequality, we have that, if $W > \Delta F$, we are observing the process F, whilst we are observing R if $W < \Delta F$. In this case, then, $L(F\,|\,\gamma) = \theta\,(W - \Delta F)$, with $\theta(\,\cdot\,)$ being the unity step function. We now evaluate the likelihood corresponding to the microscopic case. From Bayesian theory, we know that

$$L(F\,|\,\gamma) = \frac{P(\gamma\,|\,F)\cdot P(F)}{P(\gamma)}, \tag{4.35}$$

where $P(F)$ is the probability that we have been shown the process in the forward direction (*i.e.*, $1/2$), while $P(\gamma) = P(F)\,P(\gamma|F) + P(R)\,P(\gamma|R)$ is a normalization constant. We write the analogous formula for $L(R\,|\,\gamma)$, and combine them together in Eq. (4.34):

$$\begin{aligned}&\frac{P(\gamma\,|\,F)\cdot P(F)}{P(\gamma)} + \frac{P(\gamma\,|\,R)\cdot P(R)}{P(\gamma)} = \\ &= \frac{P(\gamma\,|\,F)}{2P(\gamma)}\big[\,1 + e^{-\beta(W-\Delta F)}\,\big] = 1.\end{aligned} \tag{4.36}$$

In addition, we used the fact that $P(\gamma\,|\,R) = e^{-\beta(W-\Delta F)}\,P(\gamma\,|\,F)$, which is one of the main formulations of the fluctuation theorems, and which can be justified as follows:

$$\begin{aligned}\frac{P(\gamma\,|\,F)}{P(\gamma\,|\,R)} &= \frac{e^{-\beta H(\Lambda(0))}}{Z\big(\Lambda(\tau)\big)}\left(\frac{e^{-\beta H(\tilde{\Lambda}(0))}}{Z\big(\tilde{\Lambda}(\tau)\big)}\right)^{-1} \\ &= \frac{Z\big(\tilde{\Lambda}(\tau)\big)}{Z\big(\Lambda(\tau)\big)}e^{\beta[H(\Lambda(\tau))-H(\Lambda(0))]} = e^{\beta(W-\Delta F)},\end{aligned} \tag{4.37}$$

where we assumed $H(\tilde{\Lambda}(0)) = H(\Lambda(\tau))$, and where, in the second-last equality, we used the fact that, from Liouville's theorem, we know that the volume occupied by the system in the phase space does not change, and therefore $Z\big(\tilde{\Lambda}(\tau)\big) = Z\big(\Lambda(\tau)\big)$. Note that, while this argumentation applies to the case of classical physics, one can arrive at Eq. (4.37) also by using the quantum formalism [Campisi *et al.*, 2011].

From Eq. (4.36), we obtain that

$$L(F\,|\,\gamma) = \left(1 + e^{-\beta(W-\Delta F)}\right)^{-1} = \left(1 + e^{-\beta W_{\text{diss}}}\right)^{-1}, \tag{4.38}$$

where we have called $W_{\text{diss}} = W - \Delta F$. This equation has been experimentally tested recently in a driven quantum dot setup [Hofmann *et al.*, 2017].

Fig. 4.7 shows the discrepancy between the function $\theta(W_{\text{diss}})$, valid in the macroscopic case, and Eq. (4.38), true in the microscopic one. While in the macroscopic case the direction of time is always well-defined, in the microscopic scenario there is a region in which this directionality is genuinely indefinite, and the region widens with increasing the system's temperature.

4.6.2 Description of the External Control Parameter

It is commonly considered that the unitary $U(t,0)$ acting on the system is implemented by some externally-controlled parameter of the Hamiltonian (classical field). In our case, such an external control could also be included explicitly in the quantum description of the joint state [Eq. (5) of the

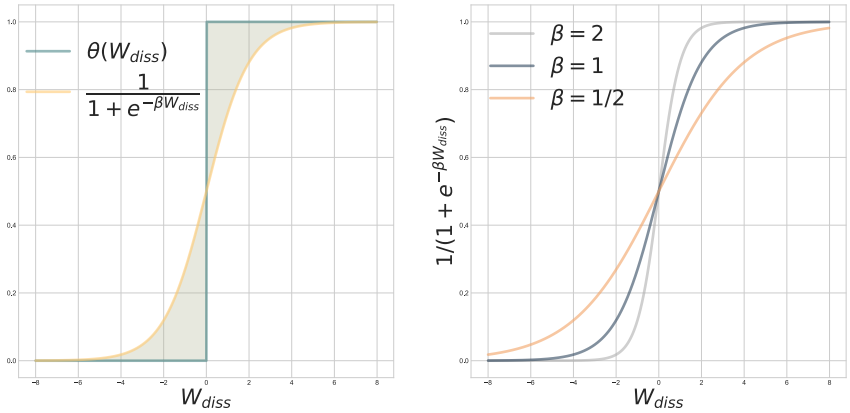


Figure 4.7 **Likelihood $L(F\,|\,\gamma)$ as a function of the dissipative work W_{diss}. (*Left*) Comparison between the trend of the unity step function, valid in the macroscopic case, and the likelihood $L(F\,|\,\gamma)$ depicting the microscopic case.** The shaded areas represent the regions where the time's arrow in the microscopic case is not univocally defined, conversely to the macroscopic one. **(*Right*) Trend of $L(F\,|\,\gamma)$ for different values of the inverse temperature β.** It is interesting to notice that this function does not depend on the features of the system, nor on the thermodynamic protocol $\Lambda(t)$. In fact, we observe that the lower the inverse temperature (and hence the higher the system's temperature), the wider the region within which the time's arrow is not well-defined.

main text] as an energy reservoir (or 'battery') interacting with the system:

$$|\Psi_0\rangle_{S,E,A,B} = \left(\alpha_0\,|\psi_0\rangle_{S,E} \otimes |0\rangle_A + \alpha_1\,|\tilde{\psi}_0\rangle_{S,E} \otimes |1\rangle_A\right) \otimes |b\rangle_B\,, \tag{4.39}$$

where $|b\rangle_B$ corresponds to an arbitrary initial state of the battery system used for the implementation of both quenches U and $\tilde{U}$. It is clear that, in order to allow coherent operations on the system, such a battery needs to be also a source of coherence [Åberg, 2014; Korzekwa *et al.*, 2016; Malabarba *et al.*, 2015]. In the limit of the battery acting as an 'unbounded' reference frame [Bartlett *et al.*, 2007] (*i.e.*, infinite source of coherence), a classical-driving is recovered. More precisely, given an arbitrary unitary U acting on the system alone, one can find an energy-preserving unitary $V(U)$ acting on the enlarged Hilbert space of the system and the battery, such that $\mathrm{Tr}_B[V(U)(\rho_S \otimes |b\rangle\langle b|_B)V(U)^\dagger] \approx U\rho_S U^\dagger$, where ρ_S is an arbitrary state of the system (see, *e.g.*, Ref. [Åberg, 2014] for a detailed proof). Furthermore, taking the battery to be initially in a strong coherent state, the states of the battery before and after the application of $V(U)$ become almost indistinguishable [Åberg, 2014; Dakić and Brukner, 2016].

For example, following Ref. [Åberg, 2014], the battery may be approximated by a doubly-infinite ladder $H_B = \sum_z \omega\,|z\rangle\,\langle z|_B$ (*e.g.*, an harmonic oscillator far from its ground state), which is assumed to be almost continuous in comparison with any energy spacing in the system, $\omega \ll E_m^{(t)} - E_n^{(t)}$ for all m, n and $t \in [0, \tau]$. Then, arbitrary unitaries on the system may be implemented by inducing rigid translations in the energy ladder. In particular, we consider the global unitary $\mathcal{V} \equiv |0\rangle\,\langle 0|_A \otimes V(U) + |1\rangle\,\langle 1|_A \otimes \tilde{V}(\tilde{U})$, which, conditionally on the state of the auxiliary system, applies to each branch in

Eq. (4.39) the corresponding system-battery unitaries:

$$V(U) \equiv \sum_{n,m} |E_m^{(\tau)}\rangle \langle E_m^{(\tau)}| U(\tau,0) |E_n^{(0)}\rangle \langle E_n^{(0)}| \otimes \Delta(W_{n,m}/\omega), \tag{4.40a}$$

$$\tilde{V}(\tilde{U}) \equiv \sum_{n,m} \Theta |E_n^{(0)}\rangle \langle E_n^{(0)}| \Theta^\dagger \tilde{U}(\tau,0) \Theta |E_m^{(\tau)}\rangle \langle E_m^{(\tau)}| \Theta^\dagger \otimes \Delta(-W_{n,m}/\omega), \tag{4.40b}$$

where we introduced the battery translation operator $\Delta(\delta) \equiv \sum_k |k+\delta\rangle \langle k|_B$, verifying $\Delta(\delta)^\dagger = \Delta(-\delta)$ and $\Delta(\delta_1)\Delta(\delta_2) = \Delta(\delta_1+\delta_2)$. Notice that, in $V(U)$, any energetic transition $|E_n^{(0)}\rangle \to |E_m^{(\tau)}\rangle$ induced by U on the system is exactly compensated by a translation on the battery of proportional magnitude, $\Delta(W_{n,m}/\omega)$, where we recall that $W_{n,m} = E_m^{(\tau)} - E_n^{(0)}$. Similarly, in $\tilde{V}(\tilde{U})$ the time-reversed transitions $\Theta|E_m^{(\tau)}\rangle \to \Theta|E_n^{(0)}\rangle$ are compensated by the opposite battery translations. As a consequence, the global Hamiltonian of the system, the auxiliary qubit and the battery at initial and final times, $\mathcal{H}_0 \equiv |0\rangle\langle 0|_A \otimes (H[\lambda(0)] + H_B) + |1\rangle\langle 1|_A \otimes (\tilde{H}[\lambda(\tau)] + H_B)$ and $\mathcal{H}_\tau \equiv |0\rangle\langle 0|_A \otimes (H[\lambda(\tau)] + H_B) + |1\rangle\langle 1|_A \otimes (\tilde{H}[\lambda(0)] + H_B)$, generate exactly the same energy distribution when applied, respectively, to the global initial and final states $|\Psi_0\rangle_{S,E,A,B}$ and $\mathcal{V}|\Psi_0\rangle_{S,E,A,B}$ (assuming the auxiliary qubit internal states $|0\rangle$ and $|1\rangle$ to have the same energy). This guarantees energy conservation.

In this situation, it is convenient to consider that the battery is initiated in a coherent state of the form $|\eta(L,l_0)\rangle_B = \sum_{l=0}^{L-1} |l+l_0\rangle_B / \sqrt{L}$, corresponding to a highly coherent state of length L. These states verify [Åberg, 2014]:

$$\langle \eta(L,l_0)|\Delta(\delta)|\eta(L,l_0)\rangle = \max(0, 1-|\delta|/L), \tag{4.41}$$

and hence the displaced state $\Delta(\delta)|\eta(L,l_0)\rangle$ becomes indistinguishable from the original state $|\eta(L,l_0)\rangle$ when $L \gg |\delta|$. Therefore, whenever we choose in Eq. (4.39) the initial state of the battery as $|b\rangle_B \equiv |\eta(L,l_0)\rangle$ for $L \gg \max_{n,m}(|W_{n,m}|/\omega)$, the back-reaction over the battery due to the implementation of the quenches U and $\tilde{U}$ may be safely neglected. This implies that, in any such protocol and for any outcomes m and n of the extended two-point measurement scheme introduced in Sec. II, the associated changes in the state of the battery would be unnoticeable.

In light of this, assuming the battery to be in a strong coherent state in the amplitude corresponding to both the forward and the time-reversal directions, we conclude that the battery ends in a nearly-indistinguishable state from its initial one, and that it can be, to a good approximation, factorized from the rest. Consequently, considering explicitly the battery in the quantum description of the joint state does not introduce any extra source of decoherence in our interferometric scheme, and hence the battery can be fully replaced by a classical external control.

Finally, we also remark that in our interferometric setup we do not require the battery to be used 'catalytically' [Åberg, 2014], hence avoiding the accumulation of (finite-size) errors leading to the battery degradation [Vaccaro *et al.*, 2018]. For instance, the battery could be reprepared in its ready-to-work state at the beginning of every realization of our scheme.

4.6.3 Relation between Entropy Production and Work

In the main body, all our results are formulated in terms of the work performed during the quench. There is, however, a link between this latter and the entropy production [Batalhão *et al.*, 2015; Kawai *et al.*, 2007; Parrondo *et al.*, 2009]. Indeed, the *stochastic entropy production* can be constructed from the stochastic work as:

$$\Delta S_{n,m} := \beta(W_{n,m} - \Delta F), \tag{4.42}$$

where $\Delta F := F_\tau - F_0 = -\log(Z_\tau/Z_0)$ is the difference in free energies between the equilibrium states at times $t = \{0, \tau\}$. Again, as a consequence of $\tilde{p}_{n|m} = p_{m|n}$, a generalized version of the fluctuation theorem for the stochastic entropy production in Eq. (4.42) can be obtained [Kawai *et al.*, 2007; Sagawa, 2012]:

$$\ln\left(\frac{p_{n,m}}{\tilde{p}_{m,n}}\right) = \ln\left(\frac{p_n^{(0)}}{\tilde{p}_m^{(0)}}\right) = \Delta S_{n,m}. \tag{4.43}$$

This equation conveys a well-defined meaning to the entropy production in terms of irreversibility by linking it to the ratio between the probability of transitions $|E_n^{(0)}\rangle \to |E_n^{(\tau)}\rangle$ in the forward dynamics, and the probability of the inverse transition $\Theta\,|E_m^{(\tau)}\rangle \to \Theta\,|E_n^{(0)}\rangle$ in the time-reversal dynamics. Moreover, following Eq. (4.43), reversible processes, for which $p_{n,m} = \tilde{p}_{m,n}$, necessarily produce zero entropy for every single realization of the protocol Λ, *i.e.*, $\Delta S_{n,m} = 0$ (or, equivalently, $W_{n,m} = \Delta F$) for all n, m.

Chapter 5

Inferring Work by Quantum Superposing Forward and Time-Reversal Evolutions

G. Rubino, G. Manzano, L. A. Rozema, P. Walther, J. M. R. Parrondo, and Č. Brukner

Abstract. The study of thermodynamic fluctuations allows to relate the free energy difference between two equilibrium states with the work done on a system through processes far from equilibrium. This finding plays a crucial role in the quantum regime, where the definition of work becomes non-trivial. Based on these relations, here we propose and experimentally realize a simple interferometric scheme allowing a direct estimation of the work distribution and the average work dissipated during an isothermal thermodynamic process. We show that our scheme provides useful upper bounds on the average work dissipated even in experimental scenarios which do not consent full control over the thermodynamic process, and we propose methodological variations depending on the possible experimental limitations encountered. Finally, we exemplify its applicability by implementing our method on a photonic system, on which the thermodynamic process acts through the polarization rotations induced by liquid crystals acting in a discrete temporal regime.

Author contributions: G.R. and L.A.R. designed the experiment. G.R. built the set-up and is carrying out the data collection. G.R., G.M., J.M.R.P. and Č.B. developed the theoretical idea. P.W., J.M.R.P. and Č.B. supervised the project. All authors contributed to writing the paper, based on an initial draft by G.R..

5.1 Introduction

While microscopic dynamical physical laws of both classical and quantum physics are time-symmetric, and hence reversible, the dynamics of macroscopic mean quantities exhibits a preferred temporal direction. The physical law formalizing this concept is the second law of thermodynamics, whereby the 'arrow of time' [Eddington, 1928] is associated with a production of positive average entropy [Callen, 1985]. According to this law, for instance, if we take a vessel divided by a wall, and put a gas in one half of the vessel, when removing the wall we will observe with a near-unity probability the gas expanding and occupying the whole vessel. Because of its unidirectional temporal evolution, this phenomenon has often been used to differentiate between past and future. There is, however, a non-zero probability that at a time all the molecules may happen to visit a half of the vessel. In this regard, the development of so-called 'fluctuation theorems', both for classical [Crooks, 1999; Evans *et al.*, 1993; Jarzynski, 1997, 2011; Seifert, 2012] and quantum [Åberg, 2018; Albash *et al.*, 2013; Alhambra *et al.*, 2016; Campisi *et al.*, 2011; Esposito *et al.*, 2009; Funo *et al.*, 2018; Manzano *et al.*, 2015, 2018; Rastegin and Życzkowski, 2014] systems, has led to the sharpening of our understanding of the second law as a statistical law, where the entropy of a system away from equilibrium can spontaneously decrease rather than increase with non-zero probability. As specified by those theorems, the ratio between the probability of entropy-decreasing events and that of entropy-increasing ones vanishes exponentially with the size of the fluctuations, and can hence be neglected in the macroscopic limit [Jarzynski, 2011].

Fundamental and empirical basis for the study of entropy production and thermodynamic irreversibility in isothermally driven systems is typically provided by the notion of dissipated work, $W_{\text{diss}} \equiv W - \Delta F$ (namely, the work invested in a thermodynamic transformation between equilibrium states having a free energy difference ΔF, which cannot be recovered by reversing the driving protocol) [Crooks, 1999; Deffner and Lutz, 2011; Jarzynski, 1997; Kawai *et al.*, 2007; Parrondo *et al.*, 2009]. The fluctuations of the dissipated work in any such processes can be characterized by constructing the work probability distribution, $P(W)$, associated to the observation of a particular value of W in a single realization of the driving protocol. Such fluctuations are constrained by a refined version of the second law: the *Crook's fluctuation theorem*, according to which

$$\frac{P(W)}{\tilde{P}(-W)} = e^{\beta W_{\text{diss}}}, \tag{5.1}$$

where $\tilde{P}(-W)$ is the probability of recovering the work W in the time-reversal dynamics, $\beta = 1/k_B T$ is the inverse temperature of the surrounding thermal environment, and k_B is the Boltzmann constant. According to Eq. (5.1), this probability ratio decreases exponentially with the amount of dissipated work, W_{diss}, in the realization. Furthermore, Eq. (5.1) implies the famous Jarzynski equality $\langle e^{-\beta W_{\text{diss}}} \rangle = 1$, where the brackets denote the statistical average with respect to $P(W)$. The Jarzynski's equality has severe implications by itself, such as the exponential decay of the probability to

observe negative values of $W_{\rm diss}$ in the forward dynamics (explicitly, $P(W_{\rm diss} < -\zeta) \leq e^{-\beta\zeta}$ for any $\zeta \geq 0$) [Jarzynski, 2011].

Work fluctuations have been measured in small classical systems leading to both testing the Crook's theorem and the Jarzynski equality, and developing applications like measurements of free-energy [Collin *et al.*, 2005; Gieseler *et al.*, 2014; Liphardt *et al.*, 2002; Tietz *et al.*, 2006; Toyabe *et al.*, 2010; Wang *et al.*, 2002]. In quantum physics, since work is not associated to any observable [Talkner *et al.*, 2007], its definition becomes more complex, and it usually demands the use of the so-called 'two-point measurement (TPM) scheme' [Campisi *et al.*, 2011]. In the TPM scheme, work is represented as the difference between the initial and final energies of the system, obtained by performing two projective measurements of the Hamiltonian at the beginning and at the end of the forward as well as of the time-reversal process (extensions to non-ideal measurements have been also considered recently [Debarba *et al.*, 2019; Ito *et al.*, 2019]). This approach has been directly implemented in several experiments [An *et al.*, 2015; Wu *et al.*, 2019; Xiong *et al.*, 2018; Zhang *et al.*, 2018]. However, since implementing projective energy measurements before and after an arbitrary process may be challenging in certain experimental scenarios, and the measurement might be destructive for the system measured, alternative methods for extracting the work distribution were proposed to circumvent this requirement. For example, in Refs. [Dorner *et al.*, 2013; Mazzola *et al.*, 2013], a scheme based on Ramsey interferometry using a single probe qubit was proposed, and subsequently implemented [Batalhão *et al.*, 2015, 2014], to extract the characteristic function of work in a NMR platform. A similar method to sample the work probability distribution from a generalized measurement scheme was introduced in Ref. [Roncaglia *et al.*, 2014], and tested experimentally on an ensemble of cold atoms [Cerisola *et al.*, 2017]. Nevertheless, the price to pay in these alternative schemes is the need of entangling operations which are standardly experimentally demanding. Developing new and simpler methods allowing for a direct estimation of the work probability distribution and irreversibility (thus refraining from the TPM scheme) is therefore of prime interest in quantum thermodynamics.

In this Chapter, we theoretically propose a simple interferometric method for quantifying the work distribution and the average dissipative work associated to a given driving protocol Λ during a isothermal thermodynamic process. The method enables to directly read out the relevant transition probabilities between eigenstates of the initial and the final Hamiltonians, which are needed to build the work probability distribution and the relative entropy (or Kullback-Leibler divergence) between the density operators in forward and time-reversal dynamics, following a driving protocol Λ. The latter was shown to be equal to the dissipative work, reinforcing the deep connection between thermodynamic and information-theoretical irreversibility put forward in Refs. [Kawai *et al.*, 2007; Parrondo *et al.*, 2009]. Remarkably, the method requires no entangling operations, and it does not require to run the scheme twice (*i.e.*, once for the forward, and once for the time-reversal process). More precisely, in the proposed method we superpose two interferometric paths: along one path the system is affected by a non-equilibrium quench implementing the first half of the driving protocol Λ (*i.e.*, from $t = 0$

to $t = \tau/2$), while along the other the system is affected by the time-reversal version of the second half of the protocol (from $t = \tau/2$ to $t = \tau$). We show that the fringe visibility in the interferometer allows the quantification of the full work probability distribution associated to any arbitrary protocol Λ, and of the relative entropy between the states in the forward and the time-reversal dynamics at any instant of time, assuming that the initial and final Hamiltonians are known. We will also demonstrate our scheme experimentally with a single photon interferometer, where the thermodynamic system is represented in terms of the polarization of a single photon. In our experiment, the thermality of the photon will be given by the degree of entanglement with an auxiliary photon, and the time-dependent Hamiltonian will be realized by traveling of the photon through a sequence of liquid crystal polarization rotators, whose rotation is such that each crystal will simulate a different time instant of the (discretised) thermodynamic process.

5.2 Theoretical Framework and Experiment

Consider a thermodynamic system S that is driven by a time-dependent Hamiltonian $H(\Lambda(t))$ depending on some control parameter $\Lambda(t)$ which varies from $t = 0$ to $t = \tau$, according to a protocol $\Lambda = \{\Lambda(t) : 0 \leq t \leq \tau\}$. The system starts the evolution in a thermal state $\rho_0^{\text{th}} = \exp[-\beta(H_0 - F_0)]$ in equilibrium with a thermal reservoir at inverse temperature β, where F_0 is the free energy corresponding to the initial Hamiltonian $H_0 \equiv H(\Lambda(0))$. The system is then isolated from the environment, and the driving protocol Λ is applied, bringing the system to an out-of-equilibrium state $\rho(t) = U(t,0)\,\rho_0^{\text{th}}\,U^\dagger(t,0)$, where $U(t,0) = \overrightarrow{T}\exp[-\frac{i}{\hbar}\int_0^t H(\Lambda(t'))dt']$, $\overrightarrow{T}$ being the so-called 'time-ordering' operator resulting from the Dyson decomposition. Once the driving protocol is ended at time τ, the system may eventually equilibrate again from $\rho_\tau = \rho(\tau)$ to the reservoir temperature, thereby reaching the thermal state $\rho_\tau^{\text{th}} = \exp\big[-\beta(H_\tau - F_\tau)\big]$, corresponding to the final Hamiltonian $H_\tau \equiv H(\Lambda(\tau))$ and the free energy F_τ.

Together with the above thermodynamic process, we consider its time-reversal twin. In the reverse process, the system starts the evolution at time $t = 0$ with Hamiltonian $\Theta H_\tau \Theta^\dagger$ in equilibrium with the thermal reservoir, that is, $\tilde{\rho}_0^{\text{th}} \equiv \Theta\rho_\tau^{\text{th}}\Theta^\dagger = \exp\big[-\beta(\Theta H_\tau\Theta^\dagger - F_\tau)\big]$. Here, Θ is the (anti-unitary) time-reversal operator, responsible for changing the sign of observables with odd parity (such as momentum, or spin under time-reversal). The time-reversal operator fulfills $\Theta\,\mathbb{1}i = -\mathbb{1}i\,\Theta$ and $\Theta\,\Theta^\dagger = \Theta^\dagger\,\Theta = \mathbb{1}$. The system is then driven according to the time-reversal protocol $\tilde{\Lambda} = \{\tilde{\Lambda}(t) \equiv \Lambda(\tau - t) : 0 \leq t \leq \tau\}$, corresponding to the inverse sequence of values of the control parameter. This brings the system out-of-equilibrium to the state $\tilde{\rho}(t) = \tilde{U}(t,0)\,\tilde{\rho}_0^{\text{th}}\,\tilde{U}^\dagger(t,0)$ at intermediate times, where now $\tilde{U}(t,0) = \overrightarrow{T}\exp\big[-\frac{i}{\hbar}\int_0^t \Theta H(\tilde{\Lambda}(t'))\Theta^\dagger dt'\big]$. After completing the protocol $\tilde{\Lambda}$, the system may return back to equilibrium at time $t = \tau$, reaching $\tilde{\rho}_\tau^{\text{th}} = \Theta\,\rho_0^{\text{th}}\,\Theta^\dagger = \exp[-\beta(\Theta H_0\Theta^\dagger - F_0)]$.

We denote by $\big|E_n^{(0)}\big\rangle$ the initial energy eigenstates of the system in the forward process, and by

$p_n^{(0)} = e^{-\beta(E_n^{(0)}-F_0)}$ the probability that the system has energy $E_n^{(0)}$. Analogously, the initial eigenstates of the system in the time-reversal process read $\Theta|E_m^{(\tau)}\rangle$, with $\tilde{p}_m^{(0)} = e^{-\beta(E_m^{(\tau)}-F_\tau)}$ being the corresponding probabilities to measure the energy $E_m^{(\tau)}$. The work probability distribution in the TPM scheme results then [Campisi *et al.*, 2011]:

$$P(W) = \sum_{m,n} p_n^{(0)}\, p_{m|n}\, \delta(W - (E_m^{(\tau)} - E_n^{(0)})), \tag{5.2}$$

where we introduced the conditional probabilities $p_{m|n} = \left|\langle E_m^{(\tau)}|\, U(\tau,0)|E_n^{(0)}\rangle\right|^2$ to find the system in the eigenstate $|E_m^{(\tau)}\rangle$ in the second projective energy measurement after the unitary evolution $U(\tau,0)$, given that it was found to be in $|E_n^{(0)}\rangle$ in the first measurement. Likewise, the work distribution in the time-reversal process reads $\tilde{P}(W) = \sum_{m,n} \tilde{p}_m^{(0)}\, \tilde{p}_{n|m}\, \delta(W - (E_n^{(0)} - E_m^{(\tau)}))$, for which $\tilde{p}_{n|m} = \left|\langle E_n^{(0)}|\Theta^\dagger \tilde{U}(\tau,0)\,\Theta|E_m^{(\tau)}\rangle\right|^2$. The *micro-reversibility relation* for non-autonomous systems [Campisi *et al.*, 2011] reads:

$$\Theta^\dagger\, \tilde{U}(\tau - t, 0)\, \Theta = U^\dagger(\tau, t). \tag{5.3}$$

Using the micro-reversibility relation (4.3), we obtain $\tilde{p}_{n|m} = p_{m|n}$. This relation is the key property to obtain Crook's theorem in Eq. (5.1) [1]. Furthermore, we assume that the Hamiltonian is invariant under time-reversal (*i.e.*, $\Theta H(t) = H(t)\Theta$). As a consequence, the relations $\Theta\,|E_n^{(0)}\rangle = |E_n^{(0)}\rangle$ and $\Theta\,|E_m^{(\tau)}\rangle = |E_m^{(\tau)}\rangle$ are also verified.

In Refs. [Kawai *et al.*, 2007; Parrondo *et al.*, 2009], the authors derived an important relation closely connected to Crook's theorem linking the dissipative work produced during the protocol Λ with the relative entropy between the density operators in forward and time-reversal dynamics at any intermediate instant of time:

$$\beta\langle W_{\rm diss}\rangle = S\big(\rho(t)\,||\,\Theta^\dagger\tilde{\rho}(\tau - t)\,\Theta\big), \tag{5.4}$$

where $S(\rho||\sigma) := \mathrm{Tr}\big[\rho\ln(\rho) - \rho\ln(\sigma)\big] \geq 0$ is the relative entropy between two generic states ρ and σ. Reversible processes for which the state in the forward dynamics is statistically indistinguishable from the one generated in the time-reversal dynamics do not dissipate work, $\langle W_{\rm diss}\rangle = 0$, and therefore all the work performed during the protocol Λ, $\langle W\rangle = \Delta F$, can be recovered back implementing the time-reversal protocol $\tilde{\Lambda}$. Importantly, the equality in Eq. (5.4) is obtained in the case of a closed system following unitary dynamics, as in the TPM scheme presented above. For open systems, the equality above is instead replaced by an inequality after tracing out environmental degrees of freedom [Kawai *et al.*, 2007; Parrondo *et al.*, 2009].

In the following, we present an interferometric scheme that allows us to directly measure the conditional probabilities $p_{m|n}$ (and therefore $\tilde{p}_{n|m}$) without implementing the TPM scheme, but resorting to the visibility of fringes in the interferometer. This enables us to construct $P(W)$ and $\tilde{P}(W)$, and

[1] The same result can be extended to Hamiltonians which are not invariant under time-reversal. In such a case, the initial state of the time-reversal process needs to incorporate the broken symmetry, that is, $\tilde{\rho}_0^{\rm th} = \Theta\,\rho_\tau^{\rm th}$

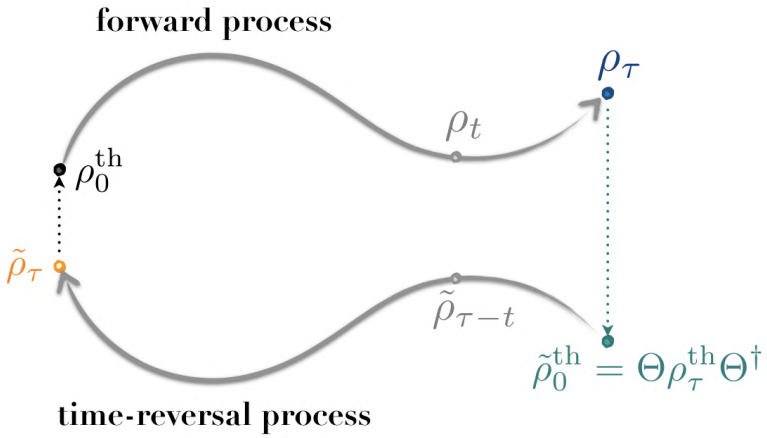


Figure 5.1 Diagrammatic representation of the forward and time-reversal evolutions of the thermodynamic system. An initial thermal state ρ_0^{th} with Hamiltonian H_0 is driven into a final, non-equilibrium state ρ_τ. It then eventually equilibrates at the reservoir temperature, reaching the thermal state ρ_τ^{th}. (If the driving process was reversible, quasi-static, the system would have ended in the state ρ_τ^{th}, immediately after the drive.) Along the driving process the Hamiltonian is changed from H_0 to H_τ. Analogously, in the process' time-reversal twin a thermal state $\tilde{\rho}_0^{\text{th}} = \rho_\tau^{\text{th}}$ with Hamiltonian H_τ evolves into a state $\tilde{\rho}_\tau$ and then it eventually equilibrates to the state $\tilde{\rho}_\tau^{\text{th}} = \rho_0^{\text{th}}$.

hence the relative entropy $S\big(\rho(t)\,||\,\Theta^\dagger\tilde{\rho}(\tau-t)\,\Theta\big)$ in Eq. (5.4).

5.3 Results

With the aim to propose a practical implementation for our alternative scheme, we introduce a second degree of freedom, which we will refer to as 'auxiliary system'. This is placed at a time $t < 0$ in a quantum superposition. For example, we can consider the system to be put in a quantum superposition of two different paths $\frac{1}{\sqrt{2}}(|0\rangle_A + |1\rangle_A)$, as shown in Fig. 5.2, where with $\{|0\rangle_A, |1\rangle_A\}$ we denote the basis in which the auxiliary system has been encoded. Suppose now that, in one of the two states of the superposition (say, $|0\rangle_A$), the system is prepared in state $|E_n^{(0)}\rangle$, while on the other state of the superposition ($|1\rangle_A$) the preparation is $|E_m^{(\tau)}\rangle$ for a certain choice of n and m.

The operation $U(\tau/2,0)$ is then applied to S in the path $|0\rangle_A$, while, on the path $|1\rangle_A$, the operation $\tilde{U}(\tau/2,0)$ is performed, followed by the time-inversion operation $\Theta^\dagger$. Therefore, at the time $\tau/2$, the state of system and path degree of freedom will read:

$$\begin{aligned}\rho_{S,A}(\tau/2) :=& \frac{1}{2}\Big\{|0\rangle_A\langle 0|_A \otimes U(\tau/2,0)\,|E_n^{(0)}\rangle\langle E_n^{(0)}|\,U^\dagger(\tau/2,0) \\ &+ |0\rangle_A\langle 1|_A \otimes U(\tau/2,0)\,|E_n^{(0)}\rangle\langle E_m^{(\tau)}|\,\tilde{U}^\dagger(\tau/2,0)\,\Theta \\ &+ |1\rangle_A\langle 0|_A \otimes \Theta^\dagger\tilde{U}(\tau/2,0)\,|E_m^{(\tau)}\rangle\langle E_n^{(0)}|\,U^\dagger(\tau/2,0) \\ &+ |1\rangle_A\langle 1|_A \otimes \Theta^\dagger\tilde{U}(\tau/2,0)\,|E_m^{(\tau)}\rangle\langle E_m^{(\tau)}|\,\tilde{U}^\dagger(\tau/2,0)\,\Theta\Big\}.\end{aligned} \tag{5.5}$$

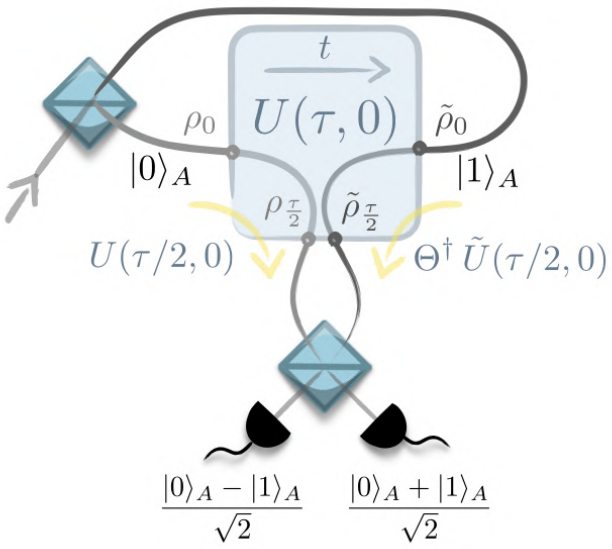


Figure 5.2 **Schematic representation of the interferometric technique to directly estimate the work dissipation.** A thermodynamic quantum system S is in its thermal state ρ_0^{th} at time $t = 0$. An additional degree of the quantum system, *e.g.*, the path degree of freedom, is put in a quantum superimposition of the two states $|0\rangle_A$ and $|1\rangle_A$. The operation $U(\tau/2,0)$ is then applied to S when traveling along the path $|0\rangle_A$, resulting in state $\rho_{\frac{\tau}{2}}$, while the operation $\tilde{U}(\tau/2,0)$ followed by the time-inversion operation $\Theta^\dagger$ is applied to S in the thermal state $\tilde{\rho}_0^{\text{th}}$ along the path $|1\rangle_A$, resulting in state $\tilde{\rho}_{\frac{\tau}{2}}$. The two quantum superimposed amplitudes are then interfered with each other, and the auxiliary system is measured in the $\left\{|\pm\rangle_A = \left(|0\rangle_A \pm |1\rangle_A\right)/\sqrt{2}\right\}$ basis.

If now we marginalize on the path degree of freedom (*i.e.*, we trace out the thermodynamic system), we obtain

$$\begin{aligned}\rho_A(\tau/2) := \mathrm{Tr}_S[\rho_{S,A}(\tau/2)] = \frac{1}{2}\Big\{&|0\rangle_A\langle 0|_A + |1\rangle_A\langle 1|_A \\ &+ |0\rangle_A\langle 1|_A \,\mathrm{Tr}_S\Big[U(\tau/2,0)\,|E_n^{(0)}\rangle\langle E_m^{(\tau)}|\,\tilde{U}^\dagger(\tau/2,0)\,\Theta\Big] + \\ &+ |1\rangle_A\langle 0|_A \,\mathrm{Tr}_S\Big[\Theta^\dagger\,\tilde{U}(\tau/2,0)\,|E_m^{(\tau)}\rangle\langle E_n^{(0)}|\,U^\dagger(\tau/2,0)\Big]\Big\}. \end{aligned} \tag{5.6}$$

Similarly, if we trace out the auxiliary system, we achieve a mixture between the state of the driven system at time $\tau/2$ in forward and time-reversal processes:

$$\rho_S(\tau/2) = \frac{1}{2}\Big[\overbrace{U(\tau/2,0)\,|E_n^{(0)}\rangle\langle E_n^{(0)}|\,U^\dagger(\tau/2,0)}^{=\rho_n(\tau/2)} + \overbrace{\Theta^\dagger\,\tilde{U}(\tau/2,0)\,|E_m^{(\tau)}\rangle\langle E_m^{(\tau)}|\,\tilde{U}^\dagger(\tau/2,0)\,\Theta}^{=\Theta^\dagger\,\tilde{\rho}_m(\tau/2)\,\Theta}\Big]. \tag{5.7}$$

Ultimately, our aim is to relate the information gained by measuring the output ports of the interferometer to the work statistics and the 'degree of reversibility' of the thermodynamic processes. To this end, we will exploit the expression of the visibility $\mathcal{V}_{m,n}$ of the interferometer fringes, which depends

on the amount of overlap between the two terms that are interfered:

$$\mathcal{V}_{m,n} := \left|\mathrm{Tr}_S\left[U(\tau/2,0)\,|E_n^{(0)}\rangle\,\langle E_m^{(\tau)}|\,\tilde{U}^\dagger(\tau/2,0)\,\Theta\right]\right|. \tag{5.8}$$

Now, we crucially apply the micro-reversibility relation in Eq. (4.3), to realize that $\Theta^\dagger\,\tilde{U}^\dagger(\tau/2,0)\,\Theta = U(\tau,\tau/2)$. Inserting this into Eq. (5.8), and using the cyclic property of the trace, we obtain the main result of our proposal:

$$\mathcal{V}_{m,n} = \left|\mathrm{Tr}_S\left[U(\tau,0)\,|E_n^{(0)}\rangle\,\langle E_m^{(\tau)}|\right]\right| = \left|\langle E_m^{(\tau)}|\,U(\tau,0)\,|E_n^{(0)}\rangle\right| = \sqrt{p_{m|n}}, \tag{5.9}$$

where, in the last equality, we identified the expression of the conditional probabilities $p_{m|n}$ of the TPM scheme.

Running this scheme for the N^2 different initial states, $n,m = 1,2,...,N$ (where N is the dimension of the system Hilbert space), and assuming that we know the eigenenergies $E_n^{(0)}$, $E_m^{(\tau)}$, and the equilibrium free energies F_0 and F_τ (or, equivalently, the initial probabilities $p_n^{(0)}$ and $\tilde{p}_m^{(0)}$), we can readily reconstruct the full probability distribution in Eq. (4.6):

$$P(W) = \sum_{m,n} p_n^{(0)}\mathcal{V}_{m,n}^2\,\delta(W-(E_m^{(\tau)}-E_n^{(0)})), \tag{5.10}$$

and its time-reversal twin $\tilde{P}(W)$. Furthermore, we can rewrite the r.h.s. of Eq. (5.4) in terms of known quantities:

$$\beta\langle W_{\mathrm{diss}}\rangle = S(\rho(t)\,||\,\Theta^\dagger\tilde{\rho}(\tau-t)\,\Theta) = \sum_n p_n^{(0)}\log p_n^{(0)} - \sum_{m,n} p_n^{(0)}\mathcal{V}_{m,n}^2\log\tilde{p}_m^{(0)}, \tag{5.11}$$

which can be alternatively obtained from the average of the work probability distribution in Eq. (5.10), $\langle W\rangle = \int_{-\infty}^{\infty} W P(W)\,dW$, and the free energy difference between the initial equilibrium states, $\Delta F = F_\tau - F_0$. As a consequence, this scheme allows, through Eqs. (5.10) and (5.11), the direct estimation of the work dissipation, and the testing of the Jarzynski equality.

5.3.1 Limited preparation and bounds on work dissipation

In the previous section, we assumed that we have the ability to prepare the energy eigenstates of the initial and final Hamiltonians of the system. Nonetheless, it could be the case that, due to technological limitations, one may not be able to prepare these pure states in the laboratory. For instance, if we do not have full control over the system in its preparation stage, and cannot isolate it from the reservoir, we may only be able to prepare the thermal states ρ_0^{th} and ρ_τ^{th}. In the following we explore what we can still learn about the work dissipation by exploiting our interferometric scheme in such a situation.

We anticipate that, although the full work probability distribution is no longer recoverable in this case, we are still able to provide useful upper bounds on the dissipative work done in the process.

As before, we prepare our auxiliary degree of freedom in a quantum superposition $\frac{1}{\sqrt{2}}(|0\rangle_A + |1\rangle_A)$ at $t < 0$. The initial states for the system in the two branches will now be, in general, the mixed thermal states ρ_0^{th} and ρ_τ^{th}. However, hereafter we will make use of their 'purifications', which can be considered as useful mathematical tools, and may correspond physically to all the environmental degrees of freedom E, such that the overall joint state of the system and these degrees of freedom is pure. (Notice that here the environment includes, but it is not limited to, the thermal reservoir). We denote the purifications of the thermal states, respectively, as $|\psi^{(0)}\rangle_{S,E}$ and $|\tilde{\psi}^{(0)}\rangle_{S,E}$, and they verify $\text{Tr}_E[|\psi^{(0)}\rangle_{S,E}\langle\psi^{(0)}|_{S,E}] = \rho_0^{\text{th}}$ and $\text{Tr}_E[|\tilde{\psi}^{(0)}\rangle_{S,E}\langle\tilde{\psi}^{(0)}|_{S,E}] = \tilde{\rho}_0^{\text{th}} = \rho_\tau^{\text{th}}$.

Again, we perform the operation $U(\tau/2,0)$ in the path $|0\rangle_A$ according to the protocol Λ, and $\tilde{U}(\tau/2,0)$ in the path $|1\rangle_A$ according to $\tilde{\Lambda}$, followed by $\Theta^\dagger$. Notice that the unitaries $U(\tau/2,0)$ and $\tilde{U}(\tau/2,0)$ only act on the system of interest, with no interactions with the environment. We can then compute the global state of the system, the environment and the auxiliary system at $\tau/2$ similarly as before, and obtain the marginal states for the auxiliary degree of freedom and the composite system consisting of the system and environment. For the latter, we obtain a mixture over the states of the system and the environment at $\tau/2$ in the forward and time-reversal dynamics:

$$\rho_{S,E}(\tau/2) = \frac{1}{2}\Big[\rho_{S,E}^{(+)} + \rho_{S,E}^{(-)}\Big], \tag{5.12}$$

where

$$\rho_{S,E}^{(+)} = \Big(U(\tau/2,0)\otimes\mathbb{1}_E\Big)\,|\psi^{(0)}\rangle_{S,E}\langle\psi^{(0)}|_{S,E}\,\Big(U^\dagger(\tau/2,0)\otimes\mathbb{1}_E\Big), \tag{5.13a}$$

$$\rho_{S,E}^{(-)} = \Big(\Theta^\dagger\,\tilde{U}(\tau/2,0)\otimes\mathbb{1}_E\Big)\,|\tilde{\psi}^{(0)}\rangle_{S,E}\langle\tilde{\psi}^{(0)}|_{S,E}\,\Big(\tilde{U}^\dagger(\tau/2,0)\,\Theta\otimes\mathbb{1}_E\Big). \tag{5.13b}$$

The corresponding state of the system only, will be then an equal probability mixture of the states $\rho_S(\tau/2) = \text{Tr}_E[\rho_{S,E}^{(+)}]$ and $\Theta^\dagger\,\tilde{\rho}_S(\tau/2)\,\Theta = \text{Tr}_E[\rho_{S,E}^{(-)}]$.

The visibility, determined by the off-diagonal elements of the auxiliary degree of freedom, reads in this case:

$$\begin{aligned}\mathcal{V} = \Big|\text{Tr}_{S,E}\Big[&(U(\tau/2,0)\otimes\mathbb{1}_E)\,|\psi^{(0)}\rangle_{S,E}\langle\tilde{\psi}^{(0)}|_{S,E}\\ &(\tilde{U}^\dagger(\tau/2,0)\,\Theta\otimes\mathbb{1}_E)\Big]\Big|\,\Big|\langle\tilde{\psi}^{(0)}|_{S,E}\,(U(\tau,0)\otimes\mathbb{1}_E)\,|\psi^{(0)}\rangle_{S,E}\Big|,\end{aligned} \tag{5.14}$$

which can no longer be related to the different outcomes of a TPM scheme. This notwithstanding, as we will shortly see, one can still make use of this information in an alternative way.

From Ref. [Englert, 1996], we know that the visibility $\mathcal{V}$ of the interferometer fringes and the distinguishability $D(\rho,\sigma)$ between two "which-path detector states" ρ and σ (*i.e.*, two states from which we

can optimally infer the which-path information, would we perform a measurement to distinguish between them) are mutually exclusive. In particular, it has been shown that these two quantities respect the *complementarity relationship*

$$\mathcal{V}^2 + D^2(\rho, \sigma) \leqslant 1, \tag{5.15}$$

and that this relation becomes an equality if the "detector states" are in pure states, as it is in our case. The distinguishability between the two states is given by the *trace-norm distance* between them, *i.e.*, $D(\rho,\sigma) := \frac{1}{2}||\rho - \sigma|| := \frac{1}{2}\mathrm{Tr}[\sqrt{(\rho-\sigma)^\dagger\,(\rho-\sigma)}]$.

In our case, $D(\rho^{(+)}_{S,E}, \rho^{(-)}_{S,E})$ gives us an estimation of how well one can distinguish between the two paths in the interferometer by measuring the system and the environment. However, we are interested in the trace-norm distance between the marginal states of the system only. We can therefore use the fact that the trace distance is non-increasing under partial trace, *i.e.*, $D(\rho^{(+)}_{S,E}, \rho^{(-)}_{S,E}) \geq D\big(\rho_S(\tau/2), \Theta^\dagger\,\tilde{\rho}_S(\tau/2)\,\Theta\big)$, to get:

$$\mathcal{V}^2 + D^2\big(\rho_S(\tau/2), \Theta^\dagger\,\tilde{\rho}_S(\tau/2)\,\Theta\big) \leqslant \mathcal{V}^2 + D^2(\rho^{(+)}_{S,E}, \rho^{(-)}_{S,E}) = 1. \tag{5.16}$$

Finally, we relate the distinguishability between the system states at $\tau/2$ in the forward and time-reversal dynamics with the relative entropy in Eq. (5.4), and hence to the average dissipated work during the protocol Λ. This can be done using the upper bounds obtained in Eqs. (17) and (19) of Ref. [Audenaert and Eisert, 2005]. Minor manipulations of these equations lead to the formulation of the following theorem:

Theorem. *Let ρ and σ be two strictly positive density operators in a finite-dimensional Hilbert space $\mathcal{H}$. Then*

$$S(\rho||\sigma) \leqslant \frac{||\rho-\sigma||_2^2}{\alpha_\sigma} \leqslant \frac{||\rho-\sigma||^2}{\alpha_\sigma} \tag{5.17}$$

where $\alpha_\sigma \in (0,1]$ is the smallest eigenvalue of σ, and $||\varrho||_2 = \sqrt{\mathrm{Tr}[\varrho^\dagger\,\varrho]}$ denotes the Frobenius (or Euclidean) norm, which verifies $||\varrho||_2 \leq ||\varrho||$.

Furthermore, setting the dimension of the Hilbert space to $\dim(\mathcal{H}) \equiv d$, we also have:

$$S(\rho||\sigma) \leqslant ||\rho-\sigma||\log(d/\sqrt{\alpha_\sigma}) + e^{-1} = D(\rho,\sigma)\log(d^2/\alpha_\sigma) + e^{-1}. \tag{5.18}$$

Combining Eqs. (5.16) and the bounds (5.17)-(5.18), we obtain the following two bounds for the dissipative work during the original thermodynamic process:

$$\langle W_{\mathrm{diss}}\rangle \leqslant k_B T\; 4(1-\mathcal{V}^2)/\alpha \equiv \mathcal{B}_2, \tag{5.19}$$

$$\langle W_{\mathrm{diss}}\rangle \leqslant k_B T\;[\sqrt{1-\mathcal{V}^2}\,\log(d^2/\alpha) + e^{-1}] \equiv \mathcal{B}_{\log}, \tag{5.20}$$

where we have used the relation between the dissipative work and the relative entropy in Eq. (5.4). Additionally, we denoted $\alpha \equiv \alpha_{\tilde{\rho}_S(\tau/2)} = \alpha_{\rho_\tau^{\text{th}}} = e^{-\beta(E^\tau_{\max}-F_\tau)}$, where $E^\tau_{\max}$ is the maximum eigenvalue of the Hamiltonian H_τ. This follows from the fact that the states $\tilde{\rho}_S(\tau/2)$ and $\tilde{\rho}_S(0) = \rho_\tau^{\text{th}}$ have the same spectrum due to their *unitary equivalence*, that is, $\tilde{\rho}_S(\tau/2) = \tilde{U}(\tau/2,0)\,\tilde{\rho}_S(0)\,\tilde{U}^\dagger(\tau/2,0)$.

5.3.2 Limited control over experimental settings

In this section, we consider the situation where the ability to control the application of the protocol Λ is heavily affected by experimental limitations such as *(i).* impossibility to split the protocol Λ in two halves and invert the second half, or *(ii).* difficulties in applying the time-reversal operation $\Theta^\dagger$ at the end of the second branch of the interferometer. If any of these circumstances applies, the requirements for the usability of the interferometric scheme proposed above may not be met. In light of this, here we propose an alternative set-up to be applied in such situations. The main price to pay is that the time needed to run the scheme for any initial state is doubled.

In this alternative scheme, we will take advantage of the unitary equivalence of the system states in the forward and time-reversal dynamics. In addition, the relation between the dissipated work and the relative entropy in Eq. (5.4) is verified for any intermediate instant of time $t \in [0,\tau]$. As a consequence, we can observe interference between the states in the forward and time-reverse dynamics also at the extremes of the interval, where one of the two states is thermal. In the following, we present the scheme in the case of interference at time τ in the forward dynamics (corresponding to $t = 0$ in the time-reversal dynamics), but an analogous scheme can be developed for interference at time $t = 0$ in the forward dynamics (corresponding to $t = \tau$ in the time-reversal one).

As in the previous case, we start by preparing the auxiliary degree of freedom in the quantum superposition $\frac{1}{\sqrt{2}}(|0\rangle_A + |1\rangle_A)$ at $t < 0$. Once again, the initial states of the system in the two branches may either be the pure states $|E_n^{(0)}\rangle$ along the path $|0\rangle_A$ and $|E_m^{(\tau)}\rangle$ along $|1\rangle_A$, or the mixed thermal states ρ_0^{th} and ρ_τ^{th}, respectively, depending on whether we have full control over the system in the preparation stage. However, in contrast to the previous case, we implement the whole protocol Λ over the system in the path $|0\rangle_A$, while the branch $|1\rangle_A$ remains unaffected.

Assuming, for concreteness, initial pure states, the global state of the system and the auxiliary system after time τ can be evaluated and, tracing the system degrees of freedom, we obtain:

$$\begin{aligned}\rho_A(\tau) := \text{Tr}_S[\rho_{S,A}(\tau)] = \frac{1}{2}\Big\{&|0\rangle_A\langle 0|_A + |1\rangle_A\langle 1|_A \\ &+ |0\rangle_A\langle 1|_A \,\text{Tr}_S\Big[U(\tau,0)\,|E_n^{(0)}\rangle\langle E_m^{(\tau)}|\Big] + |1\rangle_A\langle 0|_A\,\text{Tr}_S\Big[|E_m^{(\tau)}\rangle\langle E_n^{(0)}|\,U^\dagger(\tau,0)\Big]\Big\}. \end{aligned} \tag{5.21}$$

Consequently, in this case the visibility directly give us the conditional probabilities for the work probability distribution:

$$\mathcal{V}_{m,n} := \left|\mathrm{Tr}_S\left[U(\tau,0)\,|E_n^{(0)}\rangle\,\langle E_m^{(\tau)}|\right]\right| = \sqrt{p_{m|n}}, \tag{5.22}$$

and we recover Eqs. (5.10) and (5.11).

Likewise, when the initial states in the two interferometer paths are the mixed thermal states, we find again, for the visibility:

$$\mathcal{V} = \left|\mathrm{Tr}_{S,E}[(U(\tau,0)\otimes\mathbb{1}_E)\,|\psi^{(0)}\rangle_{S,E}\,\langle\tilde{\psi}^{(0)}|_{S,E}]\right| = \left|\langle\tilde{\psi}^{(0)}|_{S,E}\,(U(\tau,0)\otimes\mathbb{1}_E)\,|\psi^{(0)}\rangle_{S,E}\right|, \tag{5.23}$$

which is equivalent to Eq. (5.14). Consequently, the bounds developed in Eqs. (5.19) and (5.20) for the dissipated work apply also in this situation.

5.3.3 Photonic Realization

We will apply our scheme to an illustrative experimental set-up in which the thermodynamic system is represented by a single qubit realized through the polarization degree of freedom of a single photon, the auxiliary qubit is encoded in its path, and the time-dependent thermodynamic process is performed in discrete time-steps t_k by sending the photon through a sequence of liquid crystal waveplates each executing the (time-independent) Hamiltonian $H\big(\Lambda(t_k)\big)$ at time step t_k. As we are still in the data acquisition phase of this experiment, we will only present our proposal for the implementation of this scheme, omitting the presentation of the corresponding experimental results.

The Hamiltonian of the qubit system can be written as follows:

$$H\big(\Lambda(t)\big) = \frac{\hbar\omega}{2}\Big[\mathbb{1} + \cos\big(\Lambda(t)\big)\,\sigma_z + \sin\big(\Lambda(t)\big)\,\sigma_x\Big], \tag{5.24}$$

where the control parameter reads $\Lambda(t) = \Omega t$ for the range $\Lambda(0) = 0$ to $\Lambda(\tau) = \frac{\pi}{2}$. Consequently, the Hamiltonian is given by the spin operator within the $x-z$ plane, which rotates at constant angular velocity Ω around the y-axis, ω being the qubit's natural frequency. At the initial and final times of the protocol, the Hamiltonian is diagonal in the σ_z and σ_x bases, respectively. Therefore, $|E_n^{(0)}\rangle = \{|z_-\rangle, |z_+\rangle\}$ with corresponding energies $E_n^{(0)} = \{0, \hbar\omega\}$, and $|E_m^{(\tau)}\rangle = \{|x_-\rangle, |x_+\rangle\}$, where $|x_-\rangle = 1/\sqrt{2}(|z_-\rangle - |z_+\rangle)$ and $|x_+\rangle = 1/\sqrt{2}(|z_-\rangle + |z_+\rangle))$, with same eigenvalues $E_m^{(\tau)} = \{0, \hbar\omega\}$. This implies that $F_0 = F_\tau = -\log(1 + e^{-\beta\hbar\omega})$, and thus $\Delta F = F_\tau - F_0 = 0$ such that $W_{\text{diss}} = W$.

In the frame rotating around the y-axis at frequency Ω, the Hamiltonian becomes time-independent, and the unitary governing the evolution can be obtained straightforwardly (see Section 5.5.1 for de-

tails). Turning back to the Schrödinger picture, the applied unitary $U(t,0)$ reads:

$$U(t,0) = e^{-i\frac{\Omega t}{2}\sigma_y} e^{-\frac{i}{2}\left[\omega\,(\mathbb{1}+\sigma_z)-\Omega\sigma_y\right]t}. \tag{5.25}$$

In order to simulate the time-dependence of the Hamiltonian $H\big(\Lambda(t)\big)$ in Eq. (5.24), we take N discrete time-steps, where at each step we implement a different time-independent Hamiltonian. Through this artifice, we can discretise the Hamiltonian time dependence. In the k-th step, the control parameter takes a fixed value $\lambda_k \equiv \Omega\tau k/N$, where $\tau = \frac{\pi}{2\Omega}$ is the final time. Inserting $\lambda_k = \frac{k\pi}{2N}$ in Eq. (5.24), we have, for each step $k = 1, ..., N$:

$$H_k = \frac{\hbar\omega}{2}\left[\mathbb{1} + \cos\Big(\frac{k\pi}{2N}\Big)\,\sigma_z + \sin\Big(\frac{k\pi}{2N}\Big)\,\sigma_x\right]. \tag{5.26}$$

Therefore, any initial state $|\psi_R(0)\rangle$ evolves according to

$$e^{-\frac{i}{\hbar}H_N\,\Delta t}\cdots e^{-\frac{i}{\hbar}H_2\,\Delta t} e^{-\frac{i}{\hbar}H_1\,\Delta t}\,|\psi_R(0)\rangle\,, \tag{5.27}$$

where $\Delta t = \frac{\pi}{2N\Omega}$. We can view the Hamiltonian at each step as a rotation for an angle $\theta = \frac{\pi}{2N}\frac{\omega}{\Omega}$ around the axis whose direction $\vec{d}_k = (\sin(\frac{k\pi}{2N}), 0, \cos(\frac{k\pi}{2N}))$ changes from step to step.

This evolution can be implemented by means of a sequence of liquid crystal wave-plates (LCWPs). The k-th LCWP rotates the photon's polarization about an axis $\vec{d}_k$, and the angle of rotation is given by the retardance which we can charge by an externally applied voltage. Hence, to simulate the full Hamiltonian we can use a series of N LCWPs, each with an optic axis set at $\vartheta_k = \frac{k\pi}{4N} \in [0, \pi/2]$, and with the same retardance for all LCWPs (*i.e.*, θ).

We apply our scheme inserting pairs of the eigenstates of the Hamiltonians H_0 and H_τ to the interferometer. In particular, along path $|0\rangle_A$ we apply the discretised Hamiltonian $H\big(\Lambda(t_k)\big)$ for $k = 1, 2, 3$, while along path $|1\rangle_A$ we perform $H\big(\Lambda(t_k)\big)$ for $k = 6, 5, 4$. In this case, we recover the whole work probability distribution, together with the average dissipative work during the process, which we use to test the fluctuation relations.

In the second case, we test the upper bounds to the dissipative work obtained in Eqs. (5.19) and (5.20) by inserting the thermal states of the two Hamiltonians. More precisely, we insert a single photon from a pair of photons in a partially entangled state $|\psi\rangle_0^{\text{th}} = a\,|z_+\rangle\,|z_+\rangle + b\,|z_-\rangle\,|z_-\rangle$, where $a, b \in \mathbb{C}$. The state of injected photon is obtained by tracing out the second photon, $\rho_0^{\text{th}} = |a|^2\,|z_+\rangle\,\langle z_+| + |b|^2\,|z_-\rangle\,\langle z_-|$. This corresponds to a thermal state for the choice $|a|^2 = \frac{\exp(-\beta\hbar\omega)}{Z_0}$ and $|b|^2 = \frac{1}{Z_0}$, with $Z_0 = 1 + \exp(-\beta\hbar\omega)$. Specifically, if $T \to 0$, $\rho = |z_-\rangle\,\langle z_-|$, while if $T \to \infty$, $\rho = \mathbb{1}/2$. As in the previous case, we apply $H\big(\Lambda(t_k)\big)$ for $k = 1, 2, 3$ along the path $|0\rangle_A$ to realize the fist half of the forward process, and $H\big(\Lambda(t_k)\big)$ for $k = 6, 5, 4$ along the path $|1\rangle_A$ to implement the second half of the time-reversal one.

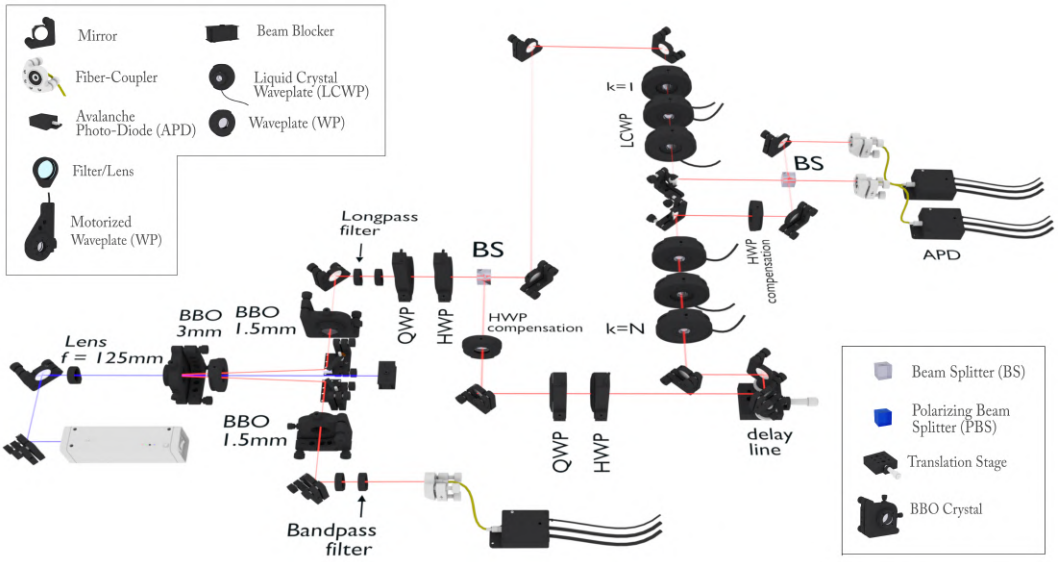


Figure 5.3 Experimental set-up of the photonic realization of the thermodynamic process. A CW laser diode emitting at 392nm pumps a BBO crystal (3 mm), giving rise to the generation of single photon pairs at 784nm via spontaneous parametric down conversion (SPDC). The pump beam is focused inside the crystal by a focusing lens. The single photons in each pair are made indistinguishable by compensating their spatio-temporal walk-off through BBO crystals (1.5 mm), and longpass and bandpass filters. The produced photons are indistinguishable and entangled in polarization. To give rise to pure (mixed) states, one of the two photons is detected with (without) polarization resolution. Both cases are tested in our experiment. The remaining single photon is sent through the set-up simulating the quantum test of thermodynamic irreversibility. The state ρ_0^{th} is prepared in polarization (via a quarter- (QWP) and a half- (HWP) waveplates), and then injected in the Mach-Zehnder interferometer. One arm of the interferometer performs the unitary $U(\tau/2,0)$ by means of a sequence of liquid crystal waveplates (LCWPs), the other prepares the photon in the state $\tilde{\rho}_0^{\text{th}} = \rho_\tau^{\text{th}}$ and then implements the unitary $\Theta^\dagger \tilde{U}(\tau/2,0)$ through LCWPs. After the two paths are recombined on a beam splitter (BS), the interference fringes are measured by varying the length of the trombone delay-line positioned along one of the two interferometric paths. The single photons are finally detected through avalanche-photodiodes (APDs).

The ratio ω/Ω determines the adiabaticity of the realized process. In the limit $\omega/\Omega \gg 1$, we obtain a fully adiabatic process, where the populations of Hamiltonian eigenstates remain constant through the entire evolution (see Section 5.5.2 for a detailed analysis). Moreover, since the Hamiltonian $H\big(\Lambda(t)\big)$ has the same eigenvalues at all times, we conclude (see Eq. (5.35) in Section 5.5.2) that, under adiabatic evolution, a system starting in a thermal state at $t = 0$ will remain in equilibrium at the same temperature at all later times. On the contrary, in the limit of a sudden quench of the Hamiltonian, $\omega/\Omega \ll 1$, the system remains in its initial state.

In the upper part of Fig. 5.4, we show the theoretical work probability distribution associated to the rotation protocol Λ for a fixed inverse temperature $\beta = 0.3(\hbar\Omega)^{-1}$ and three values of the frequency ratio $\omega/\Omega = \{0.3, 1.5, 3.5\}$ (black, blue, red). Since the eigenvalues of the Hamiltonian $H\big(\Lambda(t)\big)$ are constant, the work probability distribution can only take on three values $\{-\hbar\omega, 0, \hbar\omega\}$. When approaching the adiabatic limit ($\omega \gg \Omega$), the external values spread along the W axis and the central peak at $W = 0$ increases until it becomes a delta peak. On the contrary, in the sudden quench limit ($\omega \ll \Omega$), the central peak is quickly eliminated, and the external ones become closer to $W = 0$.

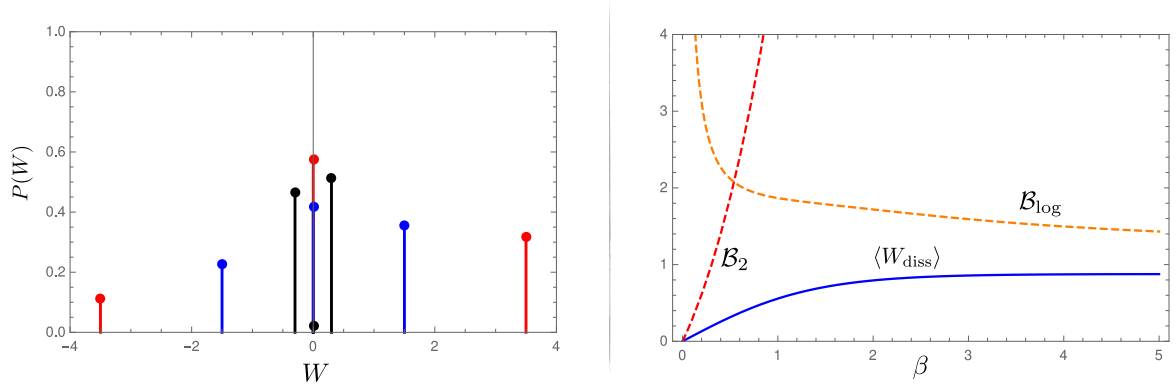


Figure 5.4 **(Left) Work probability distribution and (Bottom) bounds on the dissipative work for the qubit system under the action of the time-dependent Hamiltonian $H\big(\Lambda(t)\big)$. (Top)** Work probability distribution in Eq. (4.6) for three different protocol velocities (black, blue, red) corresponding to $\omega/\Omega = \{0.3, 1.5, 3.5\}$ for a fixed inverse temperature $\beta = 0.3(\hbar\Omega)^{-1}$. **(Right)** Dissipative work as defined in Eq. (5.4) (blue solid line) and bounds (5.19) and (5.20) (red and orange dashed lines) as a function of the inverse environmental temperature β for the case $\omega = 1.5\Omega$. Work is expressed in units of $\hbar\Omega$.

In the lower part of Fig. 5.4 we show the performance of the bounds for the dissipative work in Eqs. (5.19) and (5.20). We assume the equality in Eq. (5.16), and take a rotation frequency $\Omega = 1.5\omega$. As it can be appreciated in the plot, in the low temperatures regime (right side) the logarithmic bound $\mathcal{B}_{\log}$ becomes the best option while $\mathcal{B}_2$ diverges due to the exponential decrease of $\alpha_{\rho_\tau^{\rm th}} = e^{-\beta[E^\tau_{\rm max} - F(\tau)]}$ with temperature. On the contrary, when temperature is increased (left part), $\mathcal{B}_2$ starts to perform better as soon as k_BT becomes higher than the system energy splitting ($k_BT > \hbar\omega$).

When increasing Ω (not shown in the Figure), logarithmic and quadratic bounds become tighter in their respective temperature regimes of performance. This tendency becomes even more dramatic when approaching the sudden quench limit ($\omega \ll \Omega$). In the opposite limit of a near adiabatic processes (where the dissipative work vanishes), the quadratic bound still performs good for high temperatures, but, contrary to previous cases, the logarithmic bound becomes worst even in the limit of small temperatures. Nevertheless, the bounds do not appear to become saturated in any of the parameters' regime.

5.4 Discussion

In this work, we have developed a new method to measure the thermodynamical irreversibility of a process which a quantum system undergoes by using interferometric tools. The method utilizes the intereference between two paths, one along which the system is driven out of thermal equilibrium in the forward, and one where it is driven in the time-reversal process. We demonstrated that inserting the energy eigenstates of the initial and final Hamiltonians of the system in the two paths of the interferometer and measuring the fringe visibility enable us to reconstruct directly the work distribution

and the average dissipative work. The latter is known to be equal to a production of positive average entropy, and it is a measure of thermodynamical irreversibility. In the case of a limited experimental control, when only the thermal states of the initial and final Hamiltonians of the system can be prepared, our method provides useful upper bounds on the average dissipative work. The scheme involves no entangling operations and no energy measurements, and it might thus offer a suitable playground for studying the role of quantum coherence in thermodynamic processes.

5.5 Methods

5.5.1 Time-Independent Hamiltonian Description

In order to reach a description in terms of a time-independent Hamiltonian, we use a picture in which the states rotate at the same rate as the Hamiltonian around the $\vec{y}$-axis: $|\psi(t)\rangle = e^{-i\frac{\Omega t}{2}\sigma_y}|\psi_0(t)\rangle$.

We write the time-dependent Schrödinger equation $i\hbar\frac{d}{dt}|\psi(t)\rangle = H\big(\Lambda(t)\big)|\psi(t)\rangle$ with this substitution for $|\psi(t)\rangle$, as

$$\frac{\hbar\Omega}{2}\sigma_y|\psi_0(t)\rangle + i\hbar\frac{d}{dt}|\psi_0(t)\rangle = e^{i\frac{\Omega t}{2}\sigma_y}H\big(\Lambda(t)\big)e^{-i\frac{\Omega t}{2}\sigma_y}|\psi_0(t)\rangle\,. \tag{5.28}$$

We focus now on the r. h. s. of this equation. By substituting the expression Eq. (5.24) for the Hamiltonian, we get

$$\begin{aligned} e^{i\frac{\Omega t}{2}\sigma_y}H\big(\Lambda(t)\big)e^{-i\frac{\Omega t}{2}\sigma_y}|\psi_0(t)\rangle = \frac{\hbar\omega}{2}\Big[\mathbb{1} + \cos(\Omega t)e^{i\frac{\Omega t}{2}\sigma_y}\sigma_z e^{-i\frac{\Omega t}{2}\sigma_y} + \\ + \sin(\Omega t)e^{i\frac{\Omega t}{2}\sigma_y}\sigma_x e^{-i\frac{\Omega t}{2}\sigma_y}\Big]|\psi_0(t)\rangle \end{aligned} \tag{5.29}$$

We now write the Pauli matrices in the σ_y operator's eigenbasis, and we correspondingly evaluate the two terms in Eq. (5.29):

$$\cos(\Omega t)\sigma_z - \sin(\Omega t)\sigma_x = e^{i\frac{\Omega t}{2}\sigma_y}\sigma_z e^{-i\frac{\Omega t}{2}\sigma_y} \tag{5.30a}$$

$$\cos(\Omega t)\sigma_x + \sin(\Omega t)\sigma_z = e^{i\frac{\Omega t}{2}\sigma_y}\sigma_x e^{-i\frac{\Omega t}{2}\sigma_y} \tag{5.30b}$$

From this, Eq. (5.29) becomes $e^{i\frac{\Omega t}{2}\sigma_y}H\big(\Lambda(t)\big)e^{-i\frac{\Omega t}{2}\sigma_y} = \frac{\hbar\omega}{2}(\mathbb{1}+\sigma_z)$. By substituting this result into Eq. (5.28), we obtain

$$i\hbar\frac{d}{dt}|\psi_0(t)\rangle = \frac{\hbar}{2}\Big[\omega(\mathbb{1}+\sigma_z) - \Omega\sigma_y\Big]|\psi_0(t)\rangle\,. \tag{5.31}$$

We have thus reduced the Schrödinger equation with a time-dependent Hamiltonian into one with a

time-independent Hamiltonian. By calling

$$\sin\xi = \frac{\omega}{\sqrt{\omega^2+\Omega^2}} \qquad \cos\xi = \frac{-\Omega}{\sqrt{\omega^2+\Omega^2}}, \tag{5.32}$$

where by ξ we defined the angle between the direction $\vec{n} = (0, \cos\xi, \sin\xi)$ and $\vec{y}$-axis within z-y plane, we can therefore rewrite Eq. (5.31) as

$$i\hbar\frac{d}{dt}\left|\psi_0(t)\right\rangle = \frac{\hbar\omega}{2}\mathbb{1} + \frac{\hbar}{2}\sqrt{\omega^2+\Omega^2}\,\vec{n}\cdot\vec{\sigma}\left|\psi_0(t)\right\rangle. \tag{5.33}$$

The solution of this equation is $\left|\psi_0(t)\right\rangle = e^{-i\frac{\omega}{2}t - \frac{i}{2}\sqrt{\omega^2+\Omega^2}\,\vec{n}\cdot\vec{\sigma}\,t}\left|\psi_0(0)\right\rangle$. Neglecting the global phase ($e^{-i\frac{\omega}{2}t}$), we thus get

$$\left|\psi(t)\right\rangle = e^{-i\frac{\Omega t}{2}\sigma_y}\left|\psi_0(t)\right\rangle = \exp\Big(-i\frac{\Omega t}{2}\,\sigma_y\Big)\exp\Big(-\frac{i}{2}\sqrt{\omega^2+\Omega^2}\,\vec{n}\cdot\vec{\sigma}\,t\Big)\left|\psi(0)\right\rangle.$$

We now introduce an eigenstate basis $\{|\vec{n}_\pm\rangle\}$ of $\vec{n}\cdot\vec{\sigma}$ (*i.e.*, $\vec{n}\cdot\vec{\sigma}\,|\vec{n}_\pm\rangle = \pm\,|\vec{n}_\pm\rangle$), where $|\vec{n}_+\rangle = \cos(\xi/2)\,|y_+\rangle + \sin(\xi/2)\,|y_-\rangle$, and $|\vec{n}_-\rangle = -\sin(\xi/2)\,|y_+\rangle + \cos(\xi/2)\,|y_-\rangle$, with $\{|y_\pm\rangle\}$ being the eigenbasis of σ_y.

If we write the initial state $|\psi_0(0)\rangle$ in terms of this new basis in the general form $|\psi_0(0)\rangle = c_1\,|\vec{n}_+\rangle + c_2\,|\vec{n}_-\rangle$, where c_1 and c_2 are complex numbers, it is straightforward to obtain a solution for Eq. (5.34):

$$\begin{aligned}\left|\psi(t)\right\rangle = &\Big[c_1\,e^{-\frac{i}{2}\sqrt{\omega^2+\Omega^2}\,t}\cos(\xi/2) - c_2\,e^{\frac{i}{2}\sqrt{\omega^2+\Omega^2}\,t}\sin(\xi/2)\Big]\,e^{-i\frac{\Omega t}{2}}\,|y_+\rangle \\ &+\Big[c_1\,e^{-\frac{i}{2}\sqrt{\omega^2+\Omega^2}\,t}\sin(\xi/2) + c_2\,e^{\frac{i}{2}\sqrt{\omega^2+\Omega^2}\,t}\cos(\xi/2)\Big]\,e^{i\frac{\Omega t}{2}}\,|y_-\rangle\,.\end{aligned} \tag{5.34}$$

Finally, to set c_1 and c_2, we suppose that the initial state $|\psi_0(0)\rangle$ was an eigenstate of the initial Hamiltonian $H(t=0) = \frac{\hbar\omega}{2}(\mathbb{1}+\sigma_z)$, *i.e.*, $|z_+\rangle = (|y_+\rangle + |y_-\rangle)/\sqrt{2}$ and $|z_-\rangle = -i(|y_+\rangle - |y_-\rangle)/\sqrt{2}$. Then $c_1 = \big[\cos(\xi/2)+\sin(\xi/2)\big]/\sqrt{2}$ and $c_2 = -\big[\sin(\xi/2)-\cos(\xi/2)\big]/\sqrt{2}$ for $|z_+\rangle$ and $c_1 = -i\big[\cos(\xi/2)-\sin(\xi/2)\big]/\sqrt{2}$ and $c_2 = i\big[\sin(\xi/2)+\cos(\xi/2)\big]/\sqrt{2}$ for $|z_-\rangle$.

Let us denote by $|\psi_\pm(\tau)\rangle$ the final states evolved from the two initial states $|z_\pm\rangle$. If initially the system was in the thermal state, *i.e.*, $\rho_0^{\text{th}} = \frac{1}{Z_0}\big[|z_-\rangle\langle z_-| + e^{-\hbar\omega\beta}\,|z_+\rangle\langle z_+|\big]$, then the final state will be

$$\rho_\tau = \frac{1}{Z_0}\big[|\psi_-(\tau)\rangle\langle\psi_-(\tau)| + e^{-\hbar\omega\beta}\,|\psi_+(\tau)\rangle\langle\psi_+(\tau)|\big]. \tag{5.35}$$

5.5.2 Slow- and Fast-Varying Hamiltonians

Let us separate the two cases of a slow-varying Hamiltonian $\omega \gg \Omega$ (which will correspond to an adiabatic thermodynamic process), and that of a rapidly-varying Hamiltonian $\omega \ll \Omega$. We will

assume below that the initial state is always $|z+\rangle$.

Slow-Varying Hamiltonian ($\omega \gg \Omega$). Recalling the definitions given in Eq. (5.32), we obtain $\sin\xi \simeq 1$ and $\cos\xi \simeq 0$, and hence $\xi \simeq \pi/2$. In this case $|\psi(t)\rangle_{\omega\gg\Omega} \simeq \left(e^{-i\frac{\Omega t}{2}}|y_+\rangle + e^{i\frac{\Omega t}{2}}|y_-\rangle\right)/\sqrt{2}$, where we have neglected the global phase $e^{-i\omega t}$. We recall that, in the σ_y eigenbasis, $\sigma_x|y_\pm\rangle = \pm i|y_\mp\rangle$ and $\sigma_z|y_\pm\rangle = |y_\mp\rangle$, the action of the Hamiltonian $H\big(\Lambda(t)\big)$ onto the state $|\psi(t)\rangle$ in case of an adiabatic evolution of the system will result in $H\big(\Lambda(t)\big)|\psi(t)\rangle_{\omega\gg\Omega} \simeq \hbar\omega|\psi(t)\rangle_{\omega\gg\Omega}$. We conclude that, if a system is in an eigenstate of the time-dependent Hamiltonian (5.24) at time $t = 0$, it remains in the eigenstate of the Hamiltonian at any later time t. Since the Hamiltonian $H\big(\Lambda(t)\big)$ has the same eigenvalues at all times, we conclude from Eq. (5.35) that, under the adiabatic evolution, a system starting in a thermal state of the Hamiltonian (5.24) at $t = 0$ at temperature T will remain in the thermal state of the Hamiltonian at the temperature T at all later times. Specifically, the system will end up in the thermal state of σ_x at $t = \tau$. In this sense, the system 'thermalizes' under a slow change of the Hamiltonian.

Rapidly-Varying Hamiltonian ($\omega \ll \Omega$). In this case, $\sin\xi \simeq 0$ and $\cos\xi \simeq -1$, thus $\xi \simeq \pi$:

$$|\psi(t)\rangle_{\omega\ll\Omega} \simeq \frac{|y_+\rangle + |y_-\rangle}{\sqrt{2}}. \tag{5.36}$$

We therefore conclude that the state does not dependent on time, which indicates that, under a rapid change of the Hamiltonian, the system remains in its initial state.

Concluding Discussion

At its inception, quantum mechanics was conceived as a theory of atomic physics. Since then it has gradually been expanded to include almost all sectors of modern physics, except gravity and its related aspects [Hardy and Spekkens, 2010]. Naturally, one of the primary objectives of contemporary physical sciences is to formulate a theory meant to unify these two. Should this ambitious goal ever be achieved, it is to be expected that this will involve a radical revision[1] of both theories that it is intended to reconcile —just as the emergence of general relativity has led to a fundamental modification of Newtonian gravity and special-relativistic field theory. The present thesis aimed to investigate the quantum-mechanical definition of quantities that transcend those which are commonly "quantized", and whose reformulation could be the basis of a future revision (or extension) of quantum theory. Namely, the causal order between events and the (thermodynamic) arrow of time.

Over the last decade, it has emerged that it is possible to construct physical theories without the notions of a well-defined causal order between the events, a feature that could be an important ingredient in the construction of a theory of quantum gravity. Since space-time relations can be understood as deriving from more fundamental causal relations, this indeterminacy, in turn, might require a revision of the notions of space and time in quantum mechanics. The aim of the first part of this thesis was to experimentally investigate some of the main results of the novel research area which has stemmed from these considerations, *i.e.*, indefinite causality. In particular, in Chapter 1 we demonstrated the causal non-separability of an experimental photonic quantum switch by measuring a causal witness. In Chapter 2, we experimentally entangled the causal orders between the operations of two pairs of agents by extending the apparatus of the previous work to a so-called entangled quantum switch. This entanglement enabled the violation of a Bell's inequality for causal orders, and thus provided a proof of indefinite causality beyond the quantum formalism. In Chapter 3, we compared schemes exhibiting indefinite causality and schemes lacking such characteristics. We studied the practical task of reducing the noise introduced by noisy channels in quantum communication where the channels are embedded in either of the above schemes. We found that indefinite causality does not appear to be a necessary property to accomplish the task, as causally-definite schemes yield equal or better performances than causally-indefinite ones.

The second part of this thesis aimed to extend quantum-mechanical indefiniteness to the concept of the (thermodynamic) arrow of time. In Chapter 4, we presented a framework for the study of quantum superpositions between opposite thermodynamic time directions. In particular, we investigated the operational consequences of the quantum interference between processes with opposing time's arrows, and we found that a quantum measurement of the entropy production can reestablish a well-defined

[1] I shall remark that this is not the only possible way of looking at it. An alternative view is that the quantum formalism is a framework which can be applied to a variety of fields (*e.g.*, quantum mechanics, quantum electrodynamics, quantum field theory, *etc.*). According to this understanding, the quantum theory will not be revised, but rather extended to new domains, thereby reshaping our current understanding of these scientific sectors.

temporal direction in such a superposition. Finally, in Chapter 5, we proposed a new interferometric method to estimate the irreversibility of thermodynamic processes in the quantum regime.

In a broader context, works in this thesis look at scenarios where the causal relations between events and the thermodynamic time's arrow may be quantum-mechanically undefined. The formulation of all of our current physical theories is based on the notion of a "linear time", according to which events unfold as a well-defined sequence in a background temporal structure. The question then arises of how to formulate future theories which incorporate indeterminacy in causal relations and time directionality, and which can nevertheless be used to make predictions about observable phenomena. One possible answer is that predictions should be based on the notion of "operational events"[2], detached of the concept of space-time. In this respect, one of the contributions of this thesis has been to provide further insights into the notion of events through dedicated experimental investigations. More specifically, while historically events were tied to space-time points, the process matrix formalism and the experimental works presented herein show that these may rather be tied to operations, even if this action entails the extension of operations to space-time regions [Oreshkov, 2019]. Such a view can then encompass scenarios where events can even take place in quantum superpositions of opposite causal orders, although these superposition are still realised within laboratory conditions with a fixed background time. One can go further and conjecture the possibility of observing signatures of quantum superpositions of alternative arrows of time. A first step towards this possibility has been pursued in this thesis through the exploration of quantum superpositions of thermodynamic arrows of time. Extensions to quantum superpositions of other notions of time's arrows, as well as a comprehensive understanding of the notion of events in space-time are likely to be an important ingredient for building the conceptual foundations for future theories of gravity.

Open questions in the fields of the foundations of quantum mechanics, of indefinite causality, and of the quantum-mechanical definition of the arrow of time are still numerous, and every time we find new answers they open the way to even more new questions. In a way, this can be viewed as one of the main objectives of this thesis and, as I see it, of fundamental scientific research in general: to help provide some answers, but also and above all to pose new and more mature questions.

[2]This view is usually referred to as the "operational approach", according to which a theory is expected to specify the probabilities of results that will be observed given certain preparation and measurement procedures. The alternative approach is the so-called "realist" one, according to which there is some deeper reality underlying the equations of quantum theory which ultimately explains why we see the relative frequencies we do [Hardy and Spekkens, 2010].

List of Figures

List of Tables

Contributions to Scientific Events

Invited Talks (in inverse chronological order):

Date & Venue	Conference	Title of contribution
04-09.12.2020 Online (host: IISER Mohali)	Conference on Quantum Foundations, Technology, and Applications 2020	Time's Arrow of a Quantum Superposition of Thermodynamic Evolutions
27.11.2020 Online (host: Perimeter Institute)	Quantum Foundations Seminar	Time's Arrow of a Quantum Superposition of Thermodynamic Evolutions
13-17.01.2020 Hong-Kong, China	Quantum Information of Space Structure (QISS)	Experimental Entanglement of Temporal Orders
17-20.09.2019 Capri, Italy	Causality in the Quantum World	Experimental Entanglement of Temporal Orders
23-24.05.2019 Gdánsk, Poland	ICTQT/KCIK Symposium 2019	Experimental Entanglement of Temporal Orders
01-04.04.2019 Bristol, UK	Bristol Quantum Information Technologies 2019	Experimental Quantum Communication Enhancement by Superposing Trajectories
26-29.11.2018 Shenzhen, China	The 2nd Hong-Kong - Shenzhen Quantum Information Workshop	Experimental Quantum Communication Enhancement by Superposing Trajectories
30.07-17.08.2018 Natal, Brasil	Modern Topics in Quantum Information	Experimental Entanglement of Temporal Orders
08-12.01.2018 Hong-Kong, China	'The Role of the Observer', FXQi conference	Experimental Entanglement of Temporal Orders
10-11.04.2017 Vienna, Austria	Quantum Causal Structures, JTF-Consortium Kick-Off meeting	Experimental Entanglement of Temporal Orders

Contributed Talks (in inverse chronological order):

Date & Venue	**Conference**	**Title of contribution**
15-18.04.2019 Stockholm, Sweden	New Directions in Quantum Information Workshop	Experimental Entanglement of Temporal Orders
01-09.04.2019 Shanghai, China	International Student Workshop on Quantum Technologies	Experimental Entanglement of Temporal Orders
11-14.06.2018 Växjö, Sweden	Towards Ultimate Quantum Theory (UQT)	Experimental Entanglement of Temporal Orders
18-22.09.2018 Vienna, Austria	CoQuS Summer School 2017	Experimental Entanglement of Temporal Orders
20.11.2017 Vienna, Austria	CoQuS Colloquium	Experimental Entanglement of Temporal Orders
21-25.11.2016 Natal, Brasil	Quantum Networks Conference	Experimental Verification of an Indefinite Causal Order
27-29.09.2016 Vienna, Austria	16th ÖPG Meeting	Experimental Verification of an Indefinite Causal Order
12-16.09.2016 Vienna, Austria	CoQuS Summer School 2016	Experimental Verification of an Indefinite Causal Order
14-15.07.2016 Innsbruck, Austria	33rd SFB Meeting	Experimental Verification of an Indefinite Causal Order

References

[Åberg 2014] ÅBERG, Johan: Catalytic Coherence. In: *Phys. Rev. Lett.* 113 (2014), Oct, S. 150402. – URL https://link.aps.org/doi/10.1103/PhysRevLett.113.150402

[Åberg 2018] ÅBERG, Johan: Fully Quantum Fluctuation Theorems. In: *Phys. Rev. X* 8 (2018), Feb, S. 011019. – URL https://link.aps.org/doi/10.1103/PhysRevX.8.011019

[Abbott *et al.* 2016] ABBOTT, Alastair A. ; GIARMATZI, Christina ; COSTA, Fabio ; BRANCIARD, Cyril: Multipartite causal correlations: Polytopes and inequalities. In: *Phys. Rev. A* 94 (2016), Sep, S. 032131. – URL https://link.aps.org/doi/10.1103/PhysRevA.94.032131

[Abbott *et al.* 2017] ABBOTT, Alastair A. ; WECHS, Julian ; COSTA, Fabio ; BRANCIARD, Cyril: Genuinely multipartite noncausality. In: *Quantum* 1 (2017), Dezember, S. 39. – URL https://doi.org/10.22331/q-2017-12-14-39. – ISSN 2521-327X

[Abbott *et al.* 2020] ABBOTT, Alastair A. ; WECHS, Julian ; HORSMAN, Dominic ; MHALLA, Mehdi ; BRANCIARD, Cyril: Communication through coherent control of quantum channels. In: *Quantum* 4 (2020), September, S. 333. – URL https://doi.org/10.22331/q-2020-09-24-333. – ISSN 2521-327X

[Åberg 2004] ÅBERG, Johan: Subspace preservation, subspace locality, and gluing of completely positive maps. In: *Annals of Physics* 313 (2004), Nr. 2, S. 326–367. – URL http://www.sciencedirect.com/science/article/pii/S0003491604000879. – ISSN 0003-4916

[Aharonov *et al.* 1990] AHARONOV, Yakir ; ANANDAN, Jeeva ; POPESCU, Sandu ; VAIDMAN, Lev: Superpositions of time evolutions of a quantum system and a quantum time-translation machine. In: *Phys. Rev. Lett.* 64 (1990), Jun, S. 2965–2968. – URL https://link.aps.org/doi/10.1103/PhysRevLett.64.2965

[Albash *et al.* 2013] ALBASH, Tameem ; LIDAR, Daniel A. ; MARVIAN, Milad ; ZANARDI, Paolo: Fluctuation theorems for quantum processes. In: *Phys. Rev. E* 88 (2013), Sep, S. 032146. – URL https://link.aps.org/doi/10.1103/PhysRevE.88.032146

[Alhambra *et al.* 2016] ALHAMBRA, Álvaro M. ; MASANES, Lluis ; OPPENHEIM, Jonathan ; PERRY, Christopher: Fluctuating Work: From Quantum Thermodynamical Identities to a Second Law Equality. In: *Phys. Rev. X* 6 (2016), Oct, S. 041017. – URL https://link.aps.org/doi/10.1103/PhysRevX.6.041017

[Allard Guérin and Brukner 2018] ALLARD GUÉRIN, Philippe ; BRUKNER, Časlav: Observer-dependent locality of quantum events. In: *New Journal of Physics* 20 (2018), oct, Nr. 10, S. 103031. – URL https://doi.org/10.1088%2F1367-2630%2Faae742

[An *et al.* 2015] AN, Shuoming ; ZHANG, Jing-Ning ; UM, Mark ; LV, Dingshun ; LU, Yao ; ZHANG, Junhua ; YIN, Zhang-Qi ; QUAN, H. T. ; KIM, Kihwan: Experimental test of the quantum

Jarzynski equality with a trapped-ion system. In: *Nature Physics* 11 (2015), Nr. 11, S. 193–199. – URL https://www.nature.com/articles/nphys3197

[Andrieux and Gaspard 2008] ANDRIEUX, David ; GASPARD, Pierre: Quantum Work Relations and Response Theory. In: *Phys. Rev. Lett.* 100 (2008), Jun, S. 230404. – URL https://link.aps.org/doi/10.1103/PhysRevLett.100.230404

[Araújo *et al.* 2015] ARAÚJO, Mateus ; BRANCIARD, Cyril ; COSTA, Fabio ; FEIX, Adrien ; GIARMATZI, Christina ; BRUKNER, Časlav: Witnessing causal nonseparability. In: *New Journal of Physics* 17 (2015), oct, Nr. 10, S. 102001. – URL https://doi.org/10.1088%2F1367-2630%2F17%2F10%2F102001

[Araújo *et al.* 2014] ARAÚJO, Mateus ; COSTA, Fabio ; BRUKNER, Časlav: Computational Advantage from Quantum-Controlled Ordering of Gates. In: *Phys. Rev. Lett.* 113 (2014), Dec, S. 250402. – URL https://link.aps.org/doi/10.1103/PhysRevLett.113.250402

[Araújo *et al.* 2017] ARAÚJO, Mateus ; GUÉRIN, Philippe A. ; BAUMELER, Ämin: Quantum computation with indefinite causal structures. In: *Phys. Rev. A* 96 (2017), Nov, S. 052315. – URL https://link.aps.org/doi/10.1103/PhysRevA.96.052315

[Aspect *et al.* 1982] ASPECT, Alain ; DALIBARD, Jean ; ROGER, Gérard: Experimental Test of Bell's Inequalities Using Time-Varying Analyzers. In: *Phys. Rev. Lett.* 49 (1982), Dec, S. 1804–1807. – URL https://link.aps.org/doi/10.1103/PhysRevLett.49.1804

[Audenaert and Eisert 2005] AUDENAERT, Koenraad M. R. ; EISERT, Jens: Continuity bounds on the quantum relative entropy. In: *Journal of Mathematical Physics* 46 (2005), Nr. 10, S. 102104. – URL https://doi.org/10.1063/1.2044667

[Barreiro *et al.* 2008] BARREIRO, Julio T. ; WEI, Tzu-Chieh ; KWIAT, Paul G.: Beating the channel capacity limit for linear photonic superdense coding. In: *Nature Physics* 4 (2008), Apr, S. 282–286. – URL https://doi.org/10.1038/nphys919

[Barrett 2007] BARRETT, Jonathan: Information processing in generalized probabilistic theories. In: *Phys. Rev. A* 75 (2007), Mar, S. 032304. – URL https://link.aps.org/doi/10.1103/PhysRevA.75.032304

[Bartlett *et al.* 2007] BARTLETT, Stephen D. ; RUDOLPH, Terry ; SPEKKENS, Robert W.: Reference frames, superselection rules, and quantum information. In: *Rev. Mod. Phys.* 79 (2007), Apr, S. 555–609. – URL https://link.aps.org/doi/10.1103/RevModPhys.79.555

[Batalhão *et al.* 2015] BATALHÃO, T. B. ; SOUZA, A. M. ; SARTHOUR, R. S. ; OLIVEIRA, I. S. ; PATERNOSTRO, M. ; LUTZ, E. ; SERRA, R. M.: Irreversibility and the Arrow of Time in a Quenched Quantum System. In: *Phys. Rev. Lett.* 115 (2015), Nov, S. 190601. – URL https://link.aps.org/doi/10.1103/PhysRevLett.115.190601

[Batalhão *et al.* 2014] BATALHÃO, Tiago B. ; SOUZA, Alexandre M. ; MAZZOLA, Laura ; AUCCAISE, Ruben ; SARTHOUR, Roberto S. ; OLIVEIRA, Ivan S. ; GOOLD, John ; DE CHIARA, Gabriele ; PATERNOSTRO, Mauro ; SERRA, Roberto M.: Experimental Reconstruction of Work Distribution and Study of Fluctuation Relations in a Closed Quantum System. In: *Phys. Rev. Lett.* 113 (2014), Oct, S. 140601. – URL https://link.aps.org/doi/10.1103/PhysRevLett.113.140601

[Baumann and Brukner 2016] BAUMANN, Veronika ; BRUKNER, Časlav: Appearance of causality in process matrices when performing fixed-basis measurements for two parties. In: *Phys. Rev. A* 93 (2016), Jun, S. 062324. – URL https://link.aps.org/doi/10.1103/PhysRevA.93.062324

[Baumeler and Wolf 2014] BAUMELER, Ä. ; WOLF, S.: Perfect signaling among three parties violating predefined causal order. In: *2014 IEEE International Symposium on Information Theory*, IEEE, June 2014, S. 526–530. – URL https://ieeexplore.ieee.org/document/6874888. – ISSN 2157-8095

[Ämin Baumeler and Wolf 2016] BAUMELER Ämin ; WOLF, Stefan: The space of logically consistent classical processes without causal order. In: *New Journal of Physics* 18 (2016), jan, Nr. 1, S. 013036. – URL https://doi.org/10.1088%2F1367-2630%2F18%2F1%2F013036

[Beals *et al.* 2013] BEALS, Robert ; BRIERLEY, Stephen ; GRAY, Oliver ; HARROW, Aram W. ; KUTIN, Samuel ; LINDEN, Noah ; SHEPHERD, Dan ; STATHER, Mark: Efficient distributed quantum computing. In: *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 469 (2013), Nr. 2153, S. 20120686. – URL https://doi.org/10.1098/rspa.2012.0686

[Bell 1964] BELL, J. S.: On the Einstein Podolsky Rosen paradox. In: *Physics Physique Fizika* 1 (1964), Nov, S. 195–200. – URL https://link.aps.org/doi/10.1103/PhysicsPhysiqueFizika.1.195

[Bennett and Brassard 1984] BENNETT, C. H. ; BRASSARD, G.: Quantum cryptography: Public key distribution and coin tossing. In: *Proceedings of IEEE International Conference on Computers, Systems, and Signal Processing*, Elsevier, 1984, S. 175. – URL https://core.ac.uk/download/pdf/82447194.pdf

[Bochkov and Kuzovlev 1977] BOCHKOV, G. N. ; KUZOVLEV, Iu. E.: General theory of thermal fluctuations in nonlinear systems. In: *Zhurnal Eksperimentalnoi i Teoreticheskoi Fiziki* 72 (1977), Januar, S. 238–247. – URL https://ui.adsabs.harvard.edu/abs/1977ZhETF..72..238B

[Bochkov and Kuzovlev 1981] BOCHKOV, G.N. ; KUZOVLEV, Yu.E.: Nonlinear fluctuation-dissipation relations and stochastic models in nonequilibrium thermodynamics: I. Generalized fluctuation-dissipation theorem. In: *Physica A: Statistical Mechanics and its Applications* 106

(1981), Nr. 3, S. 443 – 479. – URL http://www.sciencedirect.com/science/article/pii/0378437181901229. – ISSN 0378-4371

[Branciard 2016] BRANCIARD, Cyril: Witnesses of causal nonseparability: an introduction and a few case studies. In: *Scientific Reports* 6 (2016), May, S. 26018. – URL https://doi.org/10.1038/srep26018

[Branciard *et al.* 2016] BRANCIARD, Cyril ; ARAÚJO, Mateus ; FEIX, Adrien ; COSTA, Fabio ; BRUKNER, Časlav: The simplest causal inequalities and their violation. In: *New Journal of Physics* 18 (2016), Nr. 1, S. 013008. – URL http://stacks.iop.org/1367-2630/18/i=1/a=013008

[Braunstein *et al.* 2000] BRAUNSTEIN, Samuel L. ; D'ARIANO, Giacomo M. ; MILBURN, G. J. ; SACCHI, Massimiliano F.: Universal Teleportation with a Twist. In: *Phys. Rev. Lett.* 84 (2000), Apr, S. 3486–3489. – URL https://link.aps.org/doi/10.1103/PhysRevLett.84.3486

[Broadbent *et al.* 2009] BROADBENT, A. ; FITZSIMONS, J. ; KASHEFI, E.: Universal Blind Quantum Computation. In: *2009 50th Annual IEEE Symposium on Foundations of Computer Science*, IEEE, 2009, S. 517–526. – URL https://ieeexplore.ieee.org/document/5438603

[Brukner 2014] BRUKNER, Časlav: Quantum causality. In: *Nat. Phys.* 10 (2014), Apr, S. 259–263. – URL https://doi.org/10.1038/nphys2930

[Brunner *et al.* 2014] BRUNNER, Nicolas ; CAVALCANTI, Daniel ; PIRONIO, Stefano ; SCARANI, Valerio ; WEHNER, Stephanie: Bell nonlocality. In: *Rev. Mod. Phys.* 86 (2014), Apr, S. 419–478. – URL https://link.aps.org/doi/10.1103/RevModPhys.86.419

[Butterfield and Isham 2001] BUTTERFIELD, Jeremy ; ISHAM, Christopher: *Spacetime and the philosophical challenge of quantum gravity*. S. 33–89. In: CALLENDER, Craig (Hrsg.) ; HUGGETT, NickEditors (Hrsg.): *Physics Meets Philosophy at the Planck Scale: Contemporary Theories in Quantum Gravity*, Cambridge University Press, 2001. – URL https://doi.org/10.1017/CBO9780511612909.003

[Caleffi and Cacciapuoti 2020] CALEFFI, Marcello ; CACCIAPUOTI, Angela S.: Quantum switch for the quantum internet: Noiseless communications through noisy channels. In: *IEEE Journal on Selected Areas in Communications* 38 (2020), Nr. 3, S. 575–588. – URL https://ieeexplore.ieee.org/document/8966996

[Callen 1985] CALLEN, H. B.: *Thermodynamics and an Introduction to Thermostatistics*. New York : Wiley, 1985. – URL https://www.academia.edu/35013052/Herbert_B_Callen_Thermodynamics_and_an_Introduction_to_Thermostatistics

[Campisi *et al.* 2011] CAMPISI, Michele ; HÄNGGI, Peter ; TALKNER, Peter: Colloquium: Quantum fluctuation relations: Foundations and applications. In: *Rev. Mod. Phys.* 83 (2011), Jul, S. 771–791. – URL https://link.aps.org/doi/10.1103/RevModPhys.83.771

[Campisi *et al.* 2015] CAMPISI, Michele ; PEKOLA, Jukka ; FAZIO, Rosario: Nonequilibrium fluctuations in quantum heat engines: theory, example, and possible solid state experiments. In: *New Journal of Physics* 17 (2015), mar, Nr. 3, S. 035012. – URL https://doi.org/10.1088%2F1367-2630%2F17%2F3%2F035012

[Cerisola *et al.* 2017] CERISOLA, Federico ; MARGALIT, Yair ; MACHLUF, Shimon ; RONCAGLIA, Augusto J. ; PAZ, Juan P. ; FOLMAN, Ron: Using a quantum work meter to test non-equilibrium fluctuation theorems. In: *Nature Communications* 8 (2017), Nr. 8, S. 1241. – URL https://www.nature.com/articles/s41467-017-01308-7

[Chiribella 2012] CHIRIBELLA, Giulio: Perfect discrimination of no-signalling channels via quantum superposition of causal structures. In: *Phys. Rev. A* 86 (2012), Oct, S. 040301. – URL https://link.aps.org/doi/10.1103/PhysRevA.86.040301

[Chiribella *et al.* 2018] CHIRIBELLA, Giulio ; BANIK, Manik ; BHATTACHARYA, Some S. ; GUHA, Tamal ; ALIMUDDIN, Mir ; ROY, Arup ; SAHA, Sutapa ; AGRAWAL, Sristy ; KAR, Guruprasad: Indefinite causal order enables perfect quantum communication with zero capacity channel. In: *arXiv:1810.10457 [quant-ph]* (2018). – URL https://arxiv.org/abs/1810.10457

[Chiribella *et al.* 2009] CHIRIBELLA, Giulio ; D'ARIANO, Giacomo M. ; PERINOTTI, Paolo: Theoretical framework for quantum networks. In: *Phys. Rev. A* 80 (2009), Aug, S. 022339. – URL https://link.aps.org/doi/10.1103/PhysRevA.80.022339

[Chiribella *et al.* 2013] CHIRIBELLA, Giulio ; D'ARIANO, Giacomo M. ; PERINOTTI, Paolo ; VALIRON, Benoit: Quantum computations without definite causal structure. In: *Phys. Rev. A* 88 (2013), Aug, S. 022318. – URL https://link.aps.org/doi/10.1103/PhysRevA.88.022318

[Chiribella and Kristjánsson 2019] CHIRIBELLA, Giulio ; KRISTJÁNSSON, Hlér: Quantum Shannon theory with superpositions of trajectories. In: *Proceedings of the Royal Society A* 475 (2019), Nr. 2225, S. 20180903. – URL https://doi.org/10.1098/rspa.2018.0903

[Choi 1975] CHOI, Man-Duen: Completely positive linear maps on complex matrices. In: *Linear Algebra and its Applications* 10 (1975), Nr. 3, S. 285 – 290. – URL https://doi.org/10.1016/0024-3795(75)90075-0. – ISSN 0024-3795

[Clauser *et al.* 1969] CLAUSER, John F. ; HORNE, Michael A. ; SHIMONY, Abner ; HOLT, Richard A.: Proposed Experiment to Test Local Hidden-Variable Theories. In: *Phys. Rev. Lett.* 23 (1969), Oct, S. 880–884. – URL https://link.aps.org/doi/10.1103/PhysRevLett.23.880

[Collin *et al.* 2005] COLLIN, D. ; RITORT, F. ; JARZYNSKI, C. ; SMITH, S. B. ; TINOCO, I. ; BUSTAMANTE, Carlos: Verification of the Crooks fluctuation theorem and recovery of RNA folding free energies. In: *Nature* 437 (2005), Nr. 437, S. 231–234. – URL https://www.nature.com/articles/nature04061

[Colnaghi *et al.* 2012] COLNAGHI, T. ; D'ARIANO, G. M. ; FACCHINI, S. ; PERINOTTI, P.: Quantum computation with programmable connections between gates. In: *Phys. Lett. A* 376 (2012), Oct, S. 2940–2943

[Crooks 1999] CROOKS, Gavin E.: Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences. In: *Phys. Rev. E* 60 (1999), Sep, S. 2721–2726. – URL https://link.aps.org/doi/10.1103/PhysRevE.60.2721

[Cuevas *et al.* 2017] CUEVAS, Álvaro ; PROIETTI, Massimiliano ; CIAMPINI, Mario A. ; DURANTI, Stefano ; MATALONI, Paolo ; SACCHI, Massimiliano F. ; MACCHIAVELLO, Chiara: Experimental Detection of Quantum Channel Capacities. In: *Phys. Rev. Lett.* 119 (2017), Sep, S. 100502. – URL https://link.aps.org/doi/10.1103/PhysRevLett.119.100502

[Dakic and Brukner 2010] DAKIC, Borivoje ; BRUKNER, Časlav: *Quantum Theory and Beyond: Is Entanglement Special?* Contribution to "Deep beauty", Editor Hans Halvorson (Cambridge University Press, 2010. – URL https://www.semanticscholar.org/paper/Deep-beauty-%3A-understanding-the-quantum-world-Halvorson/6a169046ee4a2c4cfffd428e0ac45533d04a40ce

[Dakić and Brukner 2016] DAKIĆ, Borivoje ; BRUKNER, Časlav: *The Classical Limit of a Physical Theory and the Dimensionality of Space*. S. 249–282. In: CHIRIBELLA, Giulio (Hrsg.) ; SPEKKENS, Robert W. (Hrsg.): *Quantum Theory: Informational Foundations and Foils*. Dordrecht : Springer Netherlands, 2016. – URL https://link.springer.com/chapter/10.1007/978-94-017-7303-4_8. – ISBN 978-94-017-7303-4

[De Chiara *et al.* 2015] DE CHIARA, Gabriele ; RONCAGLIA, Augusto J. ; PAZ, Juan P.: Measuring work and heat in ultracold quantum gases. In: *New Journal of Physics* 17 (2015), mar, Nr. 3, S. 035004. – URL https://doi.org/10.1088%2F1367-2630%2F17%2F3%2F035004

[Debarba *et al.* 2019] DEBARBA, Tiago ; MANZANO, Gonzalo ; GURYANOVA, Yelena ; HUBER, Marcus ; FRIIS, Nicolai: Work estimation and work fluctuations in the presence of non-ideal measurements. In: *New Journal of Physics* 21 (2019), nov, Nr. 11, S. 113002. – URL https://doi.org/10.1088%2F1367-2630%2Fab4d9d

[Deffner and Lutz 2011] DEFFNER, Sebastian ; LUTZ, Eric: Nonequilibrium Entropy Production for Open Quantum Systems. In: *Phys. Rev. Lett.* 107 (2011), Sep, S. 140404. – URL https://link.aps.org/doi/10.1103/PhysRevLett.107.140404

[Devetak 2005] DEVETAK, I.: The private classical capacity and quantum capacity of a quantum channel. In: *IEEE Transactions on Information Theory* 51 (2005), Nr. 1, S. 44–55. – URL https://ieeexplore.ieee.org/document/1377491

[Devetak and Winter 2005] DEVETAK, Igor ; WINTER, Andreas: Distillation of secret key and entanglement from quantum states. In: *Proceedings of the Royal Society A* 461 (2005), Jan. – URL https://royalsocietypublishing.org/doi/10.1098/rspa.2004.1372

[Dorner *et al.* 2013] DORNER, R. ; CLARK, S. R. ; HEANEY, L. ; FAZIO, R. ; GOOLD, J. ; VEDRAL, V.: Extracting Quantum Work Statistics and Fluctuation Theorems by Single-Qubit Interferometry. In: *Phys. Rev. Lett.* 110 (2013), Jun, S. 230601. – URL https://link.aps.org/doi/10.1103/PhysRevLett.110.230601

[Duan *et al.* 2000] DUAN, Lu-Ming ; GIEDKE, G. ; CIRAC, J. I. ; ZOLLER, P.: Inseparability Criterion for Continuous Variable Systems. In: *Phys. Rev. Lett.* 84 (2000), Mar, S. 2722–2725. – URL https://link.aps.org/doi/10.1103/PhysRevLett.84.2722

[Ebler *et al.* 2018] EBLER, Daniel ; SALEK, Sina ; CHIRIBELLA, Giulio: Enhanced Communication with the Assistance of Indefinite Causal Order. In: *Phys. Rev. Lett.* 120 (2018), Mar, S. 120502. – URL https://link.aps.org/doi/10.1103/PhysRevLett.120.120502

[Eddington 1928] EDDINGTON, Arthur S.: *The nature of the physical world*. Cambridge, England : The University Press, 1928. – URL https://henry.pha.jhu.edu/Eddington.2008.pdf

[Einstein 1905] EINSTEIN, A.: Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt. In: *Annalen der Physik* 322 (1905), Januar, Nr. 6, S. 132–148. – URL https://ui.adsabs.harvard.edu/abs/1905AnP...322..132E

[Einstein 1915] EINSTEIN, Albert: Die Feldgleichungen der Gravitation. In: *Sitzungsberichte der Preussischen Akademie der Wissenschaften zu Berlin* (1915), S. 844–847. – URL https://doi.org/10.1007/978-3-662-48039-7_11

[Ekert 1991] EKERT, Artur K.: Quantum cryptography based on Bell's theorem. In: *Phys. Rev. Lett.* 67 (1991), Aug, S. 661–663. – URL https://link.aps.org/doi/10.1103/PhysRevLett.67.661

[Englert 1996] ENGLERT, Berthold-Georg: Fringe Visibility and Which-Way Information: An Inequality. In: *Phys. Rev. Lett.* 77 (1996), Sep, S. 2154–2157. – URL https://link.aps.org/doi/10.1103/PhysRevLett.77.2154

[Englert *et al.* 2001] ENGLERT, Berthold-Georg ; KURTSIEFER, Christian ; WEINFURTER, Harald: Universal unitary gate for single-photon two-qubit states. In: *Phys. Rev. A* 63 (2001), Feb, S. 032303. – URL https://link.aps.org/doi/10.1103/PhysRevA.63.032303

[Erker *et al.* 2017] ERKER, Paul ; MITCHISON, Mark T. ; SILVA, Ralph ; WOODS, Mischa P. ; BRUNNER, Nicolas ; HUBER, Marcus: Autonomous Quantum Clocks: Does Thermodynamics Limit Our Ability to Measure Time? In: *Phys. Rev. X* 7 (2017), Aug, S. 031022. – URL https://link.aps.org/doi/10.1103/PhysRevX.7.031022

[Esposito *et al.* 2009] ESPOSITO, Massimiliano ; HARBOLA, Upendra ; MUKAMEL, Shaul: Nonequilibrium fluctuations, fluctuation theorems, and counting statistics in quantum systems. In: *Rev. Mod. Phys.* 81 (2009), Dec, S. 1665–1702. – URL https://link.aps.org/doi/10.1103/RevModPhys.81.1665

[Evans *et al.* 1993] EVANS, Denis J. ; COHEN, E. G. D. ; MORRISS, G. P.: Probability of second law violations in shearing steady states. In: *Phys. Rev. Lett.* 71 (1993), Oct, S. 2401–2404. – URL https://link.aps.org/doi/10.1103/PhysRevLett.71.2401

[Evans and Searles 2002] EVANS, Denis J. ; SEARLES, Debra J.: The Fluctuation Theorem. In: *Advances in Physics* 51 (2002), Nr. 7, S. 1529–1585. – URL https://doi.org/10.1080/00018730210155133

[Facchini and Perdrix 2015] FACCHINI, Stefano ; PERDRIX, Simon: Quantum Circuits for the Unitary Permutation Problem. In: JAIN, Rahul (Hrsg.) ; JAIN, Sanjay (Hrsg.) ; STEPHAN, Frank (Hrsg.): *Theory and Applications of Models of Computation*. Cham : Springer International Publishing, 2015, S. 324–331. – URL https://doi.org/10.1007/978-3-319-17142-5_28

[Feix *et al.* 2015] FEIX, Adrien ; ARAÚJO, Mateus ; BRUKNER, Časlav: Quantum superposition of the order of parties as a communication resource. In: *Phys. Rev. A* 92 (2015), Nov, S. 052326. – URL https://link.aps.org/doi/10.1103/PhysRevA.92.052326

[Fitzsimons *et al.* 2015] FITZSIMONS, Joseph F. ; JONES, Jonathan A. ; VEDRAL, Vlatko: Quantum correlations which imply causation. In: *Scientific Reports* 5 (2015), Dec, Nr. 18281. – URL https://www.nature.com/articles/srep18281

[Freedman and Clauser 1972] FREEDMAN, Stuart J. ; CLAUSER, John F.: Experimental Test of Local Hidden-Variable Theories. In: *Phys. Rev. Lett.* 28 (1972), Apr, S. 938–941. – URL https://link.aps.org/doi/10.1103/PhysRevLett.28.938

[Friis *et al.* 2014] FRIIS, Nicolai ; DUNJKO, Vedran ; DÜR, Wolfgang ; BRIEGEL, Hans J.: Implementing quantum control for unknown subroutines. In: *Phys. Rev. A* 89 (2014), Mar, S. 030303. – URL https://link.aps.org/doi/10.1103/PhysRevA.89.030303

[Funo *et al.* 2018] FUNO, Ken ; UEDA, Masahito ; SAGAWA, Takahiro: *Quantum Fluctuation Theorems*. S. 249–273. In: BINDER, Felix (Hrsg.) ; CORREA, Luis A. (Hrsg.) ; GOGOLIN, Christian (Hrsg.) ; ANDERS, Janet (Hrsg.) ; ADESSO, Gerardo (Hrsg.): *Thermodynamics in the Quantum Regime: Fundamental Aspects and New Directions*. Cham : Springer International Publishing, 2018. – URL https://doi.org/10.1007/978-3-319-99046-0_10. – ISBN 978-3-319-99046-0

[Gieseler *et al.* 2014] GIESELER, Jan ; QUIDANT, Romain ; DELLAGO, Christoph ; NOVOTNY, Lukas: Dynamic relaxation of a levitated nanoparticle from a non-equilibrium steady state. In: *Nature Nanotechnology* 9 (2014), Mar, S. 358. – URL https://doi.org/10.1038/nnano.2014.40

[Gilchrist *et al.* 2005] GILCHRIST, Alexei ; LANGFORD, Nathan K. ; NIELSEN, Michael A.: Distance measures to compare real and ideal quantum processes. In: *Phys. Rev. A* 71 (2005), Jun, S. 062310. – URL https://link.aps.org/doi/10.1103/PhysRevA.71.062310

[Gisin *et al.* 2005] GISIN, N. ; LINDEN, N. ; MASSAR, S. ; POPESCU, S.: Error filtration and entanglement purification for quantum communication. In: *Phys. Rev. A* 72 (2005), Jul, S. 012338. – URL https://link.aps.org/doi/10.1103/PhysRevA.72.012338

[Giustina *et al.* 2015] GIUSTINA, Marissa ; VERSTEEGH, Marijn A. M. ; WENGEROWSKY, Sören ; HANDSTEINER, Johannes ; HOCHRAINER, Armin ; PHELAN, Kevin ; STEINLECHNER, Fabian ; KOFLER, Johannes ; LARSSON, Jan-Åke ; ABELLÁN, Carlos ; AMAYA, Waldimar ; PRUNERI, Valerio ; MITCHELL, Morgan W. ; BEYER, Jörn ; GERRITS, Thomas ; LITA, Adriana E. ; SHALM, Lynden K. ; NAM, Sae W. ; SCHEIDL, Thomas ; URSIN, Rupert ; WITTMANN, Bernhard ; ZEILINGER, Anton: Significant-Loophole-Free Test of Bell's Theorem with Entangled Photons. In: *Phys. Rev. Lett.* 115 (2015), Dec, S. 250401. – URL https://link.aps.org/doi/10.1103/PhysRevLett.115.250401

[Goswami *et al.* 2020] GOSWAMI, K. ; CAO, Y. ; PAZ-SILVA, G. A. ; ROMERO, J. ; WHITE, A. G.: Increasing communication capacity via superposition of order. In: *Phys. Rev. Research* 2 (2020), Aug, S. 033292. – URL https://link.aps.org/doi/10.1103/PhysRevResearch.2.033292

[Goswami *et al.* 2018] GOSWAMI, K. ; GIARMATZI, C. ; KEWMING, M. ; COSTA, F. ; BRANCIARD, C. ; ROMERO, J. ; WHITE, A. G.: Indefinite Causal Order in a Quantum Switch. In: *Phys. Rev. Lett.* 121 (2018), Aug, S. 090503. – URL https://link.aps.org/doi/10.1103/PhysRevLett.121.090503

[Guérin *et al.* 2016] GUÉRIN, Philippe A. ; FEIX, Adrien ; ARAÚJO, Mateus ; BRUKNER, ⋮ Exponential Communication Complexity Advantage from Quantum Superposition of the Direction of Communication. In: *Phys. Rev. Lett.* 117 (2016), Sep, S. 100502. – URL https://link.aps.org/doi/10.1103/PhysRevLett.117.100502

[Guérin *et al.* 2019] GUÉRIN, Philippe A. ; RUBINO, Giulia ; BRUKNER, ⋮ Communication through quantum-controlled noise. In: *Phys. Rev. A* 99 (2019), Jun, S. 062317. – URL https://link.aps.org/doi/10.1103/PhysRevA.99.062317

[Guo *et al.* 2020] GUO, Yu ; HU, Xiao-Min ; HOU, Zhi-Bo ; CAO, Huan ; CUI, Jin-Ming ; LIU, Bi-Heng ; HUANG, Yun-Feng ; LI, Chuan-Feng ; GUO, Guang-Can ; CHIRIBELLA, Giulio: Experimental Transmission of Quantum Information Using a Superposition of Causal Orders. In: *Phys. Rev. Lett.* 124 (2020), Jan, S. 030502. – URL https://link.aps.org/doi/10.1103/PhysRevLett.124.030502

[Gühne and Tóth 2009] GÜHNE, Otfried ; TÓTH, Géza: Entanglement detection. In: *Physics Reports* 474 (2009), Nr. 1, S. 1 – 75. – URL http://www.sciencedirect.com/science/article/pii/S0370157309000623. – ISSN 0370-1573

[Halliwell *et al.* 1996] HALLIWELL, J. J. ; PÉREZ-MERCADER, J. ; ZUREK, W. H.: *Physical Origins of Time Asymmetry Paperback*. Cambridge, England : The University Press, 1996. – URL https://www.cambridge.org/at/academic/subjects/physics/history-philosophy-and-foundations-physics/physical-origins-time-asymmetry?format=PB&isbn=9780521568371

[Hardy 2001] HARDY, Lucien: Quantum Theory From Five Reasonable Axioms. In: *Preprint at arXiv:0101012 [quant-ph]* (2001), Nr. , S. . – URL https://arxiv.org/abs/quant-ph/0101012

[Hardy 2007] HARDY, Lucien: Towards quantum gravity: a framework for probabilistic theories with non-fixed causal structure. In: *Journal of Physics A: Mathematical and Theoretical* 40 (2007), mar, Nr. 12, S. 3081–3099. – URL https://doi.org/10.1088%2F1751-8113%2F40%2F12%2Fs12

[Hardy 2009] HARDY, Lucien: *Quantum Gravity Computers: On the Theory of Computation with Indefinite Causal Structure*. S. 379–401. In: *Quantum Reality, Relativistic Causality, and Closing the Epistemic Circle: Essays in Honour of Abner Shimony*. Dordrecht : Springer Netherlands, 2009. – URL https://link.springer.com/chapter/10.1007/978-1-4020-9107-0_21. – ISBN 978-1-4020-9107-0

[Hardy 2011] HARDY, Lucien: *Foliable Operational Structures for General Probabilistic Theories*. S. 409–442. In: HALVORSON, HansEditor (Hrsg.): *Deep Beauty: Understanding the Quantum World through Mathematical Innovation*, Cambridge University Press, 2011. – URL https://doi.org/10.1017/CBO9780511976971.013

[Hardy and Spekkens 2010] HARDY, Lucien ; SPEKKENS, Robert: Why Physics Needs Quantum Foundations. In: *Physics in Canada* 66 (2010), Nr. 2, S. 73–76. – URL https://arxiv.org/abs/1003.5008

[Hensen *et al.* 2015] HENSEN, B. ; BERNIEN, H. ; DREAU, A. E. ; REISERER, A. ; KALB, N. ; BLOK, M. S. ; RUITENBERG, J. ; VERMEULEN, R. F. L. ; SCHOUTEN, R. N. ; ABELLAN, C. ; AMAYA, W. ; PRUNERI, V. ; MITCHELL, M. W. ; MARKHAM, M. ; TWITCHEN, D. J. ; ELKOUSS, D. ; WEHNER, S. ; TAMINIAU, T. H. ; HANSON, R.: Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres. In: *Nature* 526 (2015), Oct, S. 682–686. – URL http://dx.doi.org/10.1038/nature15759

[Hofmann *et al.* 2017] HOFMANN, Andrea ; MAISI, Ville F. ; BASSET, Julien ; REICHL, Christian ; WEGSCHEIDER, Werner ; IHN, Thomas ; ENSSLIN, Klaus ; JARZYNSKI, Christopher: Heat dissipation and fluctuations in a driven quantum dot. In: *physica status solidi (b)* 254 (2017), Nr. 3, S. 1600546. – URL https://doi.org/10.1002/pssb.201600546

[Holevo 1973] HOLEVO, A. S.: Bounds for the Quantity of Information Transmitted by a Quantum Communication Channel. In: *Problems Inform. Transmission* 9 (1973), S. 177–18. – URL https://www.semanticscholar.org/paper/

Bounds-for-the-quantity-of-information-transmitted-Holevo/0733393a545b62e49e9fc4a76d53d843cfb72010

[Howell *et al.* 2004] HOWELL, John C. ; BENNINK, Ryan S. ; BENTLEY, Sean J. ; BOYD, R. W.: Realization of the Einstein-Podolsky-Rosen Paradox Using Momentum- and Position-Entangled Photons from Spontaneous Parametric Down Conversion. In: *Phys. Rev. Lett.* 92 (2004), May, S. 210403. – URL https://link.aps.org/doi/10.1103/PhysRevLett.92.210403

[Ito *et al.* 2019] ITO, Kosuke ; TALKNER, Peter ; VENKATESH, B. P. ; WATANABE, Gentaro: Generalized energy measurements and quantum work compatible with fluctuation theorems. In: *Phys. Rev. A* 99 (2019), Mar, S. 032117. – URL https://link.aps.org/doi/10.1103/PhysRevA.99.032117

[Jarzynski 1997] JARZYNSKI, C.: Nonequilibrium Equality for Free Energy Differences. In: *Phys. Rev. Lett.* 78 (1997), Apr, S. 2690–2693. – URL https://link.aps.org/doi/10.1103/PhysRevLett.78.2690

[Jarzynski 2011] JARZYNSKI, Christopher: Equalities and Inequalities: Irreversibility and the Second Law of Thermodynamics at the Nanoscale. In: *Annual Review of Condensed Matter Physics* 2 (2011), Nr. 1, S. 329–351. – URL https://doi.org/10.1146/annurev-conmatphys-062910-140506

[Jennings and Rudolph 2010a] JENNINGS, David ; RUDOLPH, Terry: Comment on "Quantum Solution to the Arrow-of-Time Dilemma". In: *Phys. Rev. Lett.* 104 (2010), Apr, S. 148901. – URL https://link.aps.org/doi/10.1103/PhysRevLett.104.148901

[Jennings and Rudolph 2010b] JENNINGS, David ; RUDOLPH, Terry: Entanglement and the thermodynamic arrow of time. In: *Phys. Rev. E* 81 (2010), Jun, S. 061130. – URL https://link.aps.org/doi/10.1103/PhysRevE.81.061130

[Jia and Costa 2019] JIA, Ding ; COSTA, Fabio: Causal order as a resource for quantum communication. In: *Phys. Rev. A* 100 (2019), Nov, S. 052319. – URL https://link.aps.org/doi/10.1103/PhysRevA.100.052319

[Kawai *et al.* 2007] KAWAI, R. ; PARRONDO, J. M. R. ; BROECK, C. V. den: Dissipation: The Phase-Space Perspective. In: *Phys. Rev. Lett.* 98 (2007), Feb, S. 080602. – URL https://link.aps.org/doi/10.1103/PhysRevLett.98.080602

[Kim *et al.* 2006] KIM, Taehyun ; FIORENTINO, Marco ; WONG, Franco N. C.: Phase-stable source of polarization-entangled photons using a polarization Sagnac interferometer. In: *Phys. Rev. A* 73 (2006), Jan, S. 012316. – URL https://link.aps.org/doi/10.1103/PhysRevA.73.012316

[Kochen and Specker 1975] KOCHEN, Simon ; SPECKER, E. P.: *The Problem of Hidden Variables in Quantum Mechanics*. S. 293–328. In: HOOKER, C. A. (Hrsg.): *The Logico-Algebraic Approach*

to Quantum Mechanics: Volume I: Historical Evolution. Dordrecht : Springer Netherlands, 1975. – URL https://link.springer.com/chapter/10.1007/978-94-010-1795-4_17. – ISBN 978-94-010-1795-4

[Korzekwa *et al.* 2016] KORZEKWA, Kamil ; LOSTAGLIO, Matteo ; OPPENHEIM, Jonathan ; JENNINGS, David: The extraction of work from quantum coherence. In: *New Journal of Physics* 18 (2016), feb, Nr. 2, S. 023045. – URL https://doi.org/10.1088%2F1367-2630%2F18%2F2%2F023045

[Kristjánsson *et al.* 2020] KRISTJÁNSSON, Hlér ; CHIRIBELLA, Giulio ; SALEK, Sina ; EBLER, Daniel ; WILSON, Matthew: Resource theories of communication. In: *New Journal of Physics* 22 (2020), jul, Nr. 7, S. 073014. – URL https://doi.org/10.1088%2F1367-2630%2Fab8ef7

[Kristjánsson *et al.* 2020] KRISTJÁNSSON, Hlér ; MAO, Wen-Xu ; CHIRIBELLA, Giulio: Single-particle communication through correlated noise. In: *arXiv:2004.06090 [quant-ph]* (2020). – URL https://arxiv.org/abs/2004.06090

[Kwiat *et al.* 1993] KWIAT, P. G. ; STEINBERG, A. M. ; CHIAO, R. Y.: High-visibility interference in a Bell-inequality experiment for energy and time. In: *Phys. Rev. A* 47 (1993), Apr, S. R2472–R2475. – URL https://link.aps.org/doi/10.1103/PhysRevA.47.R2472

[Lamehi-Rachti and Mittig 1976] LAMEHI-RACHTI, M. ; MITTIG, W.: Quantum mechanics and hidden variables: A test of Bell's inequality by the measurement of the spin correlation in low-energy proton-proton scattering. In: *Phys. Rev. D* 14 (1976), Nov, S. 2543–2555. – URL https://link.aps.org/doi/10.1103/PhysRevD.14.2543

[Landi and Paternostro 2020] LANDI, Gabriel T. ; PATERNOSTRO, Mauro: Irreversible entropy production, from quantum to classical. In: *arXiv:2010.03752 [quant-ph]* (2020). – URL https://arxiv.org/abs/2010.03752

[Lanyon *et al.* 2009] LANYON, Benjamin P. ; BARBIERI, Marco ; ALMEIDA, Marcelo P. ; JENNEWEIN, Thomas ; RALPH, Timothy C. ; RESCH, Kevin J. ; PRYDE, Geoff J. ; O'BRIEN, Jeremy L. ; GILCHRIST, Alexei ; WHITE, Andrew G.: Simplifying quantum logic using higher-dimensional Hilbert spaces. In: *Nature Physics* 5 (2009), Feb, S. 134–140. – URL https://www.nature.com/articles/nphys1150

[Leifer and Spekkens 2013] LEIFER, M. S. ; SPEKKENS, Robert W.: Towards a formulation of quantum theory as a causally neutral theory of Bayesian inference. In: *Phys. Rev. A* 88 (2013), Nov, S. 052130. – URL https://link.aps.org/doi/10.1103/PhysRevA.88.052130

[Liphardt *et al.* 2002] LIPHARDT, Jan ; DUMONT, Sophie ; SMITH, Steven B. ; TINOCO, Ignacio ; BUSTAMANTE, Carlos: Equilibrium Information from Nonequilibrium Measurements in an Experimental Test of Jarzynski's Equality. In: *Science* 296 (2002), Nr. 5574, S. 1832–1835. – URL https://science.sciencemag.org/content/296/5574/1832. – ISSN 0036-8075

[Lloyd 1997] LLOYD, Seth: Capacity of the noisy quantum channel. In: *Phys. Rev. A* 55 (1997), Mar, S. 1613–1622. – URL https://link.aps.org/doi/10.1103/PhysRevA.55.1613

[Loizeau and Grinbaum 2020] LOIZEAU, Nicolas ; GRINBAUM, Alexei: Channel capacity enhancement with indefinite causal order. In: *Phys. Rev. A* 101 (2020), Jan, S. 012340. – URL https://link.aps.org/doi/10.1103/PhysRevA.101.012340

[Maccone 2009] MACCONE, Lorenzo: Quantum Solution to the Arrow-of-Time Dilemma. In: *Phys. Rev. Lett.* 103 (2009), Aug, S. 080401. – URL https://link.aps.org/doi/10.1103/PhysRevLett.103.080401

[MacLean *et al.* 2017] MACLEAN, Jean-Philippe W. ; RIED, Katja ; SPEKKENS, Robert W. ; RESCH, Kevin J.: Quantum-coherent mixtures of causal relations. In: *Nature Communications* 8 (2017), May, S. 15149. – URL https://doi.org/10.1038/ncomms15149

[Malabarba *et al.* 2015] MALABARBA, Artur S L. ; SHORT, Anthony J. ; KAMMERLANDER, Philipp: Clock-driven quantum thermal engines. In: *New Journal of Physics* 17 (2015), apr, Nr. 4, S. 045027. – URL https://doi.org/10.1088/1367-2630/17/4/045027

[Manzano *et al.* 2015] MANZANO, Gonzalo ; HOROWITZ, Jordan M. ; PARRONDO, Juan M. R.: Nonequilibrium potential and fluctuation theorems for quantum maps. In: *Phys. Rev. E* 92 (2015), Sep, S. 032129. – URL https://link.aps.org/doi/10.1103/PhysRevE.92.032129

[Manzano *et al.* 2018] MANZANO, Gonzalo ; HOROWITZ, Jordan M. ; PARRONDO, Juan M. R.: Quantum Fluctuation Theorems for Arbitrary Environments: Adiabatic and Nonadiabatic Entropy Production. In: *Phys. Rev. X* 8 (2018), Aug, S. 031037. – URL https://link.aps.org/doi/10.1103/PhysRevX.8.031037

[Maragakis *et al.* 2008] MARAGAKIS, P. ; RITORT, F. ; BUSTAMANTE, C. ; KARPLUS, M. ; CROOKS, G.E.: Bayesian estimates of free energies from nonequilibrium work data in the presence of instrument noise. In: *J. Chem. Phys.* 129 (2008), S. 024102. – URL https://doi.org/10.1063/1.2937892

[Masanes *et al.* 2014] MASANES, Ll. ; MÜLLER, M. P. ; PÉREZ-GARCÍA, D. ; AUGUSIAK, R.: Entanglement and the three-dimensionality of the Bloch ball. In: *Journal of Mathematical Physics* 55 (2014), Nr. 12, S. 122203. – URL https://doi.org/10.1063/1.4903510

[Masanes and Müller 2011] MASANES, Lluí?s ; MÜLLER, Markus P.: A derivation of quantum theory from physical requirements. In: *New Journal of Physics* 13 (2011), Nr. 6, S. 063001. – URL http://stacks.iop.org/1367-2630/13/i=6/a=063001

[Maslennikov *et al.* 2019] MASLENNIKOV, G. ; DING, S. ; HABLÜTZEL, R ; GAN, Jaren ; ROULET, Alexandre ; NIMMRICHTER, Stefan ; DAI, Jibo ; SCARANI, Valerio ; MATSUKEVICH, Dzmitry: Quantum absorption refrigerator with trapped ions. In: *Nature Communications* 10 (2019), Nr. 202, S. 2041–1723. – URL https://doi.org/10.1038/s41467-018-08090-0

[Mazzola *et al.* 2013] MAZZOLA, L. ; DE CHIARA, G. ; PATERNOSTRO, M.: Measuring the Characteristic Function of the Work Distribution. In: *Phys. Rev. Lett.* 110 (2013), Jun, S. 230602. – URL https://link.aps.org/doi/10.1103/PhysRevLett.110.230602

[Miklin *et al.* 2017] MIKLIN, Nikolai ; ABBOTT, Alastair A. ; BRANCIARD, Cyril ; CHAVES, Rafael ; BUDRONI, Costantino: The entropic approach to causal correlations. In: *New Journal of Physics* 19 (2017), Nr. 11, S. 113041. – URL http://stacks.iop.org/1367-2630/19/i=11/a=113041

[Minář *et al.* 2008] MINÁŘ, J. ; RIEDMATTEN, Hugues de ; SIMON, Christoph ; ZBINDEN, Hugo ; GISIN, Nicolas: Phase-noise measurements in long-fiber interferometers for quantum-repeater applications. In: *Phys. Rev. A* 77 (2008), May, S. 052325. – URL https://link.aps.org/doi/10.1103/PhysRevA.77.052325

[Mlodinow and Brun 2014] MLODINOW, Leonard ; BRUN, Todd A.: Relation between the psychological and thermodynamic arrows of time. In: *Phys. Rev. E* 89 (2014), May, S. 052102. – URL https://link.aps.org/doi/10.1103/PhysRevE.89.052102

[Mohammady and Romito 2019] MOHAMMADY, M. H. ; ROMITO, Alessandro: Conditional work statistics of quantum measurements. In: *Quantum* 3 (2019), August, S. 175. – URL https://doi.org/10.22331/q-2019-08-19-175. – ISSN 2521-327X

[Nielsen and Chuang 2000] NIELSEN, M.A. ; CHUANG, I.L.: *Quantum Computation and Quantum Information*. Cambridge University Press, 2000 (Cambridge Series on Information and the Natural Sciences). – URL https://doi.org/10.1017/CBO9780511976667. – ISBN 9780521635035

[Oi 2003] OI, Daniel K. L.: Interference of Quantum Channels. In: *Phys. Rev. Lett.* 91 (2003), Aug, S. 067902. – URL https://link.aps.org/doi/10.1103/PhysRevLett.91.067902

[Oreshkov 2019] ORESHKOV, Ognyan: Time-delocalized quantum subsystems and operations: on the existence of processes with indefinite causal structure in quantum mechanics. In: *Quantum* 3 (2019), Dezember, S. 206. – URL https://doi.org/10.22331/q-2019-12-02-206. – ISSN 2521-327X

[Oreshkov *et al.* 2012] ORESHKOV, Ognyan ; COSTA, Fabio ; BRUKNER, Časlav: Quantum correlations with no causal order. In: *Nat. Commun.* 3 (2012), Oct, S. 1092. – URL https://doi.org/10.1038/ncomms2076

[Oreshkov and Giarmatzi 2016] ORESHKOV, Ognyan ; GIARMATZI, Christina: Causal and causally separable processes. In: *New Journal of Physics* 18 (2016), Nr. 9, S. 093020. – URL https://doi.org/10.1088/1367-2630/18/9/093020

[Parrondo *et al.* 2009] PARRONDO, J. M. R. ; BROECK, C. V. den ; KAWAI, R.: Entropy production and the arrow of time. In: *New Journal of Physics* 11 (2009), jul, Nr. 7, S. 073008. – URL https://doi.org/10.1088/1367-2630/11/7/073008

[Perarnau-Llobet *et al.* 2017] PERARNAU-LLOBET, Martí ; BÄUMER, Elisa ; HOVHANNISYAN, Karen V. ; HUBER, Marcus ; ACIN, Antonio: No-Go Theorem for the Characterization of Work Fluctuations in Coherent Quantum Systems. In: *Phys. Rev. Lett.* 118 (2017), Feb, S. 070601. – URL https://link.aps.org/doi/10.1103/PhysRevLett.118.070601

[Peterson *et al.* 2019] PETERSON, John P. S. ; BATALHÃO, Tiago B. ; HERRERA, Marcela ; SOUZA, Alexandre M. ; SARTHOUR, Roberto S. ; OLIVEIRA, Ivan S. ; SERRA, Roberto M.: Experimental Characterization of a Spin Quantum Heat Engine. In: *Phys. Rev. Lett.* 123 (2019), Dec, S. 240601. – URL https://link.aps.org/doi/10.1103/PhysRevLett.123.240601

[Procopio *et al.* 2020] PROCOPIO, Lorenzo M. ; DELGADO, Francisco ; ENRÍQUEZ, Marco ; BELABAS, Nadia ; LEVENSON, Juan A.: Sending classical information via three noisy channels in superposition of causal orders. In: *Phys. Rev. A* 101 (2020), Jan, S. 012346. – URL https://link.aps.org/doi/10.1103/PhysRevA.101.012346

[Procopio *et al.* 2019] PROCOPIO, Lorenzo M. ; DELGADO, Francisco ; ENRÍQUEZ, Marco ; BELABAS, Nadia ; LEVENSON, Juan A.: Communication Enhancement through Quantum Coherent Control of N Channels in an Indefinite Causal-Order Scenario. In: *Entropy* 21 (2019), S. 1012. – URL https://doi.org/10.3390/e21101012

[Procopio *et al.* 2015] PROCOPIO, Lorenzo M. ; MOQANAKI, Amir ; ARAÚJO, Mateus ; COSTA, Fabio ; ALONSO CALAFELL, Irati ; DOWD, Emma G. ; HAMEL, Deny R. ; ROZEMA, Lee A. ; BRUKNER, Časlav ; WALTHER, Philip: Experimental superposition of orders of quantum gates. In: *Nat. Commun.* 6 (2015), Aug, Nr. 7913. – URL https://doi.org/10.1038/ncomms8913

[http://www.dr qubit.org/matlab.html]

[Rambo *et al.* 2016] RAMBO, Timothy M. ; ALTEPETER, Joseph B. ; KUMAR, Prem ; D'ARIANO, G. M.: Functional quantum computing: An optical approach. In: *Phys. Rev. A* 93 (2016), May, S. 052321. – URL https://link.aps.org/doi/10.1103/PhysRevA.93.052321

[Rarity and Tapster 1990] RARITY, J. G. ; TAPSTER, P. R.: Experimental violation of Bell's inequality based on phase and momentum. In: *Phys. Rev. Lett.* 64 (1990), May, S. 2495–2498. – URL https://link.aps.org/doi/10.1103/PhysRevLett.64.2495

[Rastegin and Życzkowski 2014] RASTEGIN, Alexey E. ; ŻYCZKOWSKI, Karol: Jarzynski equality for quantum stochastic maps. In: *Phys. Rev. E* 89 (2014), Jan, S. 012127. – URL https://link.aps.org/doi/10.1103/PhysRevE.89.012127

[Ried *et al.* 2015] RIED, Katja ; AGNEW, Megan ; VERMEYDEN ; JANZING, Dominik ; SPEKKENS, Robert W. ; RESCH, Kevin J.: A quantum advantage for inferring causal structure. In: *Nat. Phys.* 11 (2015), May, S. 414–420. – URL https://doi.org/10.1038/nphys3266

[Ringbauer *et al.* 2016] RINGBAUER, Martin ; GIARMATZI, Christina ; CHAVES, Rafael ; COSTA, Fabio ; WHITE, Andrew G. ; FEDRIZZI, Alessandro: Experimental test of nonlocal causality. In: *Science Advances* 2 (2016), Nr. 8. – URL https://advances.sciencemag.org/content/2/8/e1600162

[Rockafellar 1970] ROCKAFELLAR, R. T.: *Convex Analysis*. In: GRIFFITHS, Phillip A. (Hrsg.) ; MATHER, John N. (Hrsg.) ; STEIN, Elias M. (Hrsg.): *Convex Analysis*. Princeton, NJ : Princeton Landmarks, 1970. – URL https://www.jstor.org/stable/j.ctt14bs1ff. – ISBN 9780691015866

[Roncaglia *et al.* 2014] RONCAGLIA, Augusto J. ; CERISOLA, Federico ; PAZ, Juan P.: Work Measurement as a Generalized Quantum Measurement. In: *Phys. Rev. Lett.* 113 (2014), Dec, S. 250601. – URL https://link.aps.org/doi/10.1103/PhysRevLett.113.250601

[Roßnagel *et al.* 2016] ROSSNAGEL, Johannes ; DAWKINS, Samuel T. ; TOLAZZI, Karl N. ; ABAH, Obinna ; LUTZ, Eric ; SCHMIDT-KALER, Ferdinand ; SINGER, Kilian: A single-atom heat engine. In: *Science* 352 (2016), Nr. 6283, S. 325–329. – URL https://science.sciencemag.org/content/352/6283/325. – ISSN 0036-8075

[Rovelli 2017] ROVELLI, Carlo: *Is Time's Arrow Perspectival?* S. 285–296. In: CHAMCHAM, Khalil (Hrsg.) ; SILK, Joseph (Hrsg.) ; BARROW, John D. (Hrsg.) ; SAUNDERS, SimonEditors (Hrsg.): *The Philosophy of Cosmology*, Cambridge University Press, 2017. – URL https://doi.org/10.1017/9781316535783.015

[Rowe *et al.* 2001] ROWE, M. A. ; KIELPINSKI, D. ; MEYER, V. ; SACKETT, C. A. ; ITANO, W. M. ; MONROE, C. ; WINELAND, D. J.: Experimental violation of a Bell's inequality with efficient detection. In: *Nature* 409 (2001), Feb, S. 791. – URL http://dx.doi.org/10.1038/35057215

[Royer 1991] ROYER, Antoine: Wigner function in Liouville space: A canonical formalism. In: *Phys. Rev. A* 43 (1991), Jan, S. 44–56. – URL https://link.aps.org/doi/10.1103/PhysRevA.43.44

[Rozema *et al.* 2012] ROZEMA, Lee A. ; DARABI, Ardavan ; MAHLER, Dylan H. ; HAYAT, Alex ; SOUDAGAR, Yasaman ; STEINBERG, Aephraim M.: Violation of Heisenberg's Measurement-Disturbance Relationship by Weak Measurements. In: *Phys. Rev. Lett.* 109 (2012), Sep, S. 100404. – URL https://link.aps.org/doi/10.1103/PhysRevLett.109.100404

[Rozema *et al.* 2014] ROZEMA, Lee A. ; MAHLER, Dylan H. ; BLUME-KOHOUT, Robin ; STEINBERG, Aephraim M.: Optimizing the Choice of Spin-Squeezed States for Detecting and Characterizing Quantum Processes. In: *Phys. Rev. X* 4 (2014), Nov, S. 041025. – URL https://link.aps.org/doi/10.1103/PhysRevX.4.041025

[Rubino *et al.* 2017a] RUBINO, Giulia ; ROZEMA, Lee A. ; FEIX, Adrien ; ARAÚJO, Mateus ; ZEUNER, Jonas M. ; PROCOPIO, Lorenzo M. ; BRUKNER, Časlav ; WALTHER, Philip: Experimental verification of an indefinite causal order. In: *Science Advances* 3 (2017), Nr. 3

[Rubino *et al.* 2017b] RUBINO, Giulia ; ROZEMA, Lee A. ; MASSA, Francesco ; ARAÚJO, Mateus ; ZYCH, Magdalena ; BRUKNER, Časlav ; WALTHER, Philip: Experimental Entanglement of Temporal Orders. In: *arXiv:1712.06884 [quant-ph]* (2017). – URL https://arxiv.org/abs/1712.06884

[Sagawa 2012] SAGAWA, Takahiro: *Second Law-Like Inequalities with Quantum Relative Entropy: An Introduction*. S. 125–190. In: *Lectures on Quantum Computing, Thermodynamics and Statistical Physics*, World Scientific, 2012. – URL https://www.worldscientific.com/doi/abs/10.1142/9789814425193_0003

[Salek *et al.* 2018] SALEK, Sina ; EBLER, Daniel ; CHIRIBELLA, Giulio: Quantum communication in a superposition of causal orders. In: *arXiv:1809.06655 [quant-ph]* (2018). – URL https://arxiv.org/abs/1809.06655

[Sazim *et al.* 2020] SAZIM, Sk ; SINGH, Kratveer ; PATI, Arun K.: Classical Communications with Indefinite Causal Order for N completely depolarizing channels. In: *arXiv:2004.14339 [quant-ph]* (2020). – URL https://arxiv.org/abs/2004.14339

[Schumacher and Nielsen 1996] SCHUMACHER, Benjamin ; NIELSEN, M. A.: Quantum data processing and error correction. In: *Phys. Rev. A* 54 (1996), Oct, S. 2629–2635. – URL https://link.aps.org/doi/10.1103/PhysRevA.54.2629

[Seifert 2012] SEIFERT, Udo: Stochastic thermodynamics, fluctuation theorems and molecular machines. In: *Reports on Progress in Physics* 75 (2012), nov, Nr. 12, S. 126001. – URL https://doi.org/10.1088%2F0034-4885%2F75%2F12%2F126001

[Shalm *et al.* 2015] SHALM, Lynden K. ; MEYER-SCOTT, Evan ; CHRISTENSEN, Bradley G. ; BIERHORST, Peter ; WAYNE, Michael A. ; STEVENS, Martin J. ; GERRITS, Thomas ; GLANCY, Scott ; HAMEL, Deny R. ; ALLMAN, Michael S. ; COAKLEY, Kevin J. ; DYER, Shellee D. ; HODGE, Carson ; LITA, Adriana E. ; VERMA, Varun B. ; LAMBROCCO, Camilla ; TORTORICI, Edward ; MIGDALL, Alan L. ; ZHANG, Yanbao ; KUMOR, Daniel R. ; FARR, William H. ; MARSILI, Francesco ; SHAW, Matthew D. ; STERN, Jeffrey A. ; ABELLÁN, Carlos ; AMAYA, Waldimar ; PRUNERI, Valerio ; JENNEWEIN, Thomas ; MITCHELL, Morgan W. ; KWIAT, Paul G. ; BIENFANG, Joshua C. ; MIRIN, Richard P. ; KNILL, Emanuel ; NAM, Sae W.: Strong Loophole-Free Test of Local Realism. In: *Phys. Rev. Lett.* 115 (2015), Dec, S. 250402. – URL https://link.aps.org/doi/10.1103/PhysRevLett.115.250402

[Shirts *et al.* 2003] SHIRTS, Michael R. ; BAIR, Eric ; HOOKER, Giles ; PANDE, Vijay S.: Equilibrium Free Energies from Nonequilibrium Measurements Using Maximum-Likelihood Methods. In: *Phys. Rev. Lett.* 91 (2003), Oct, S. 140601. – URL https://link.aps.org/doi/10.1103/PhysRevLett.91.140601

[Simon 2000] SIMON, R.: Peres-Horodecki Separability Criterion for Continuous Variable Systems. In: *Phys. Rev. Lett.* 84 (2000), Mar, S. 2726–2729. – URL https://link.aps.org/doi/10.1103/PhysRevLett.84.2726

[Smith and Smolin 2008] SMITH, Graeme ; SMOLIN, John A.: Additive extensions of a quantum channel. In: *2008 IEEE Information Theory Workshop, ITW 2008, Porto, Portugal, May 5-9, 2008*, IEEE, 2008, S. 368–372. – URL https://doi.org/10.1109/ITW.2008.4578688

[Taddei *et al.* 2020] TADDEI, Márcio M. ; CARIÑE, Jaime ; MARTÍNEZ, Daniel ; GARCÍA, Tania ; GUERRERO, Nayda ; ABBOTT, Alastair A. ; ARAÚJO, Mateus ; BRANCIARD, Cyril ; GÓMEZ, Esteban S. ; WALBORN, Stephen P. ; AOLITA, Leandro ; LIMA, Gustavo: Experimental computational advantage from superposition of multiple temporal orders of quantum gates. In: *arXiv:2002.07817v1 [quant-ph]* (2020). – URL https://arxiv.org/abs/2002.07817

[Talkner and Hänggi 2016] TALKNER, Peter ; HÄNGGI, Peter: Aspects of quantum work. In: *Phys. Rev. E* 93 (2016), Feb, S. 022131. – URL https://link.aps.org/doi/10.1103/PhysRevE.93.022131

[Talkner *et al.* 2007] TALKNER, Peter ; LUTZ, Eric ; HÄNGGI, Peter: Fluctuation theorems: Work is not an observable. In: *Phys. Rev. E* 75 (2007), May, S. 050102. – URL https://link.aps.org/doi/10.1103/PhysRevE.75.050102

[Tietz *et al.* 2006] TIETZ, C. ; SCHULER, S. ; SPECK, T. ; SEIFERT, U. ; WRACHTRUP, J.: Measurement of Stochastic Entropy Production. In: *Phys. Rev. Lett.* 97 (2006), Aug, S. 050602. – URL https://link.aps.org/doi/10.1103/PhysRevLett.97.050602

[Toyabe *et al.* 2010] TOYABE, Shoichi ; SAGAWA, Takahiro ; UEDA, Masahito ; MUNEYUKI, Eiro ; SANO, Masaki: Experimental demonstration of information-to-energy conversion and validation of the generalized Jarzynski equality. In: *Nature Physics* 6 (2010), Nr. 6, S. 988–992. – URL https://www.nature.com/articles/nphys1821

[Vaccaro *et al.* 2018] VACCARO, Joan A. ; CROKE, Sarah ; BARNETT, Stephen M.: Is coherence catalytic? In: *Journal of Physics A: Mathematical and Theoretical* 51 (2018), sep, Nr. 41, S. 414008. – URL https://doi.org/10.1088%2F1751-8121%2Faac112

[Van Horne *et al.* 2020] VAN HORNE, N. ; YUM, D. ; DUTTA, T. ; HÄNGGI, P. ; GONG, J. ; POLETTI, D. ; MUKHERJEE, M.: Single-atom energy-conversion device with a quantum load. In: *npj Quantum Information* 6 (2020), S. 37. – URL https://doi.org/10.1038/s41534-020-0264-6

[Von Neumann 1955] VON NEUMANN, John: *Mathematical Foundations of Quantum Mechanics*. In: *originally published in 1932 as "Mathematische Grundlagen der Quantenmechanik."*. Princeton, NJ : Princeton Univ. Press, 1955. – URL https://press.princeton.edu/books/hardcover/9780691178561/mathematical-foundations-of-quantum-mechanics

[Wang *et al.* 2016] WANG, Da-Wei ; CAI, Han ; LIU, Ren-Bao ; SCULLY, Marlan O.: Mesoscopic Superposition States Generated by Synthetic Spin-Orbit Interaction in Fock-State Lattices. In: *Phys. Rev. Lett.* 116 (2016), Jun, S. 220502. – URL https://link.aps.org/doi/10.1103/PhysRevLett.116.220502

[Wang *et al.* 2002] WANG, G. M. ; SEVICK, E. M. ; MITTAG, Emil ; SEARLES, Debra J. ; EVANS, Denis J.: Experimental Demonstration of Violations of the Second Law of Thermodynamics for Small Systems and Short Time Scales. In: *Phys. Rev. Lett.* 89 (2002), Jul, S. 050601. – URL https://link.aps.org/doi/10.1103/PhysRevLett.89.050601

[Wei *et al.* 2019] WEI, Kejin ; TISCHLER, Nora ; ZHAO, Si-Ran ; LI, Yu-Huai ; ARRAZOLA, Juan M. ; LIU, Yang ; ZHANG, Weijun ; LI, Hao ; YOU, Lixing ; WANG, Zhen ; CHEN, Yu-Ao ; SANDERS, Barry C. ; ZHANG, Qiang ; PRYDE, Geoff J. ; XU, Feihu ; PAN, Jian-Wei: Experimental Quantum Switching for Exponentially Superior Quantum Communication Complexity. In: *Phys. Rev. Lett.* 122 (2019), Mar, S. 120504. – URL https://link.aps.org/doi/10.1103/PhysRevLett.122.120504

[Wilde 2013] WILDE, Mark M.: *Quantum Information Theory*. Cambridge University Press, 2013. – URL https://doi.org/10.1017/CBO9781139525343

[Wilson *et al.* 2020] WILSON, Matthew ; CHAU, Hoi-Fung ; CHIRIBELLA, Giulio: Quantum Communication Through Completely Depolarising Channels in a Superposition of Causal Orders. In: *arXiv:2005.00618* (2020). – URL https://arxiv.org/abs/2005.00618

[Wu *et al.* 2019] WU, Kang-Da ; YUAN, Yuan ; XIANG, Guo-Yong ; LI, Chuan-Feng ; GUO, Guang-Can ; PERARNAU-LLOBET, Martí: Experimentally reducing the quantum measurement back action in work distributions by a collective measurement. In: *Science Advances* 5 (2019). – URL https://advances.sciencemag.org/content/5/3/eaav4944

[Xiong *et al.* 2018] XIONG, T. P. ; YAN, L. L. ; ZHOU, F. ; REHAN, K. ; LIANG, D. F. ; CHEN, L. ; YANG, W. L. ; MA, Z. H. ; FENG, M. ; VEDRAL, V.: Experimental Verification of a Jarzynski-Related Information-Theoretic Equality by a Single Trapped Ion. In: *Phys. Rev. Lett.* 120 (2018), Jan, S. 010601. – URL https://link.aps.org/doi/10.1103/PhysRevLett.120.010601

[Yu *et al.* 2020] YU, Shang ; MENG, Yu ; PATEL, Raj B. ; WANG, Yi-Tao ; KE, Zhi-Jin ; LIU, Wei ; LI, Zhi-Peng ; YANG, Yuan-Ze ; ZHANG, Wen-Hao ; TANG, Jian-Shun ; LI, Chuan-Feng ; GUO, Guang-Can: Experimental Observation of Coherent-Information Superadditivity in a Dephrasure Channel. In: *Phys. Rev. Lett.* 125 (2020), Aug, S. 060502. – URL https://link.aps.org/doi/10.1103/PhysRevLett.125.060502

[Zhang *et al.* 2018] ZHANG, Zhenxing ; WANG, Tenghui ; XIANG, Liang ; JIA, Zhilong ; DUAN, Peng ; CAI, Weizhou ; ZHAN, Ze ; ZONG, Zhiwen ; WU, Jianlan ; SUN, Luyan ; YIN, Yi ; GUO, Guoping: Experimental demonstration of work fluctuations along a shortcut to adiabaticity with a

superconducting Xmon qubit. In: *New Journal of Physics* 20 (2018), aug, Nr. 8, S. 085001. – URL https://doi.org/10.1088/1367-2630/aad4e7

[Zhao *et al.* 2020] ZHAO, Xiaobin ; YANG, Yuxiang ; CHIRIBELLA, Giulio: Quantum Metrology with Indefinite Causal Order. In: *Phys. Rev. Lett.* 124 (2020), May, S. 190503. – URL https://link.aps.org/doi/10.1103/PhysRevLett.124.190503

[Zhou *et al.* 2011] ZHOU, Xiao-Qi ; RALPH, Timothy C. ; KALASUWAN, Pruet ; ZHANG, Mian ; PERUZZO, Alberto ; LANYON, Benjamin P. ; O'BRIEN, Jeremy L.: Adding control to arbitrary unknown quantum operations. In: *Nature Communications* 2 (2011), Aug, Nr. 413. – URL https://doi.org/10.1038/ncomms1392

[Zych *et al.* 2019] ZYCH, Magdalena ; COSTA, Fabio ; PIKOVSKI, Igor ; BRUKNER, Časlav: Bell's Theorem for Temporal Order. In: *Nature Communications* 10 (2019), 08, S. 3772. – URL https://doi.org/10.1038/s41467-019-11579-x