

## Quantum Advantage in Simulating Stochastic Processes

Korzekwa, Kamil; Lostaglio, Matteo

**DOI**

[10.1103/PhysRevX.11.021019](https://doi.org/10.1103/PhysRevX.11.021019)

**Publication date**

2021

**Document Version**

Final published version

**Published in**

Physical Review X

**Citation (APA)**

Korzekwa, K., & Lostaglio, M. (2021). Quantum Advantage in Simulating Stochastic Processes. *Physical Review X*, 11(2), Article 021019. <https://doi.org/10.1103/PhysRevX.11.021019>

**Important note**

To cite this publication, please use the final published version (if applicable). Please check the document version above.

**Copyright**

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

**Takedown policy**

Please contact us and provide details if you believe this document breaches copyrights. We will remove access to the work immediately and investigate your claim.

# Quantum Advantage in Simulating Stochastic Processes

Kamil Korzekwa<sup>1,2,‡,\*</sup> and Matteo Lostaglio<sup>3,4,‡,†</sup>

<sup>1</sup>*Faculty of Physics, Astronomy and Applied Computer Science, Jagiellonian University, 30-348 Kraków, Poland*

<sup>2</sup>*International Centre for Theory of Quantum Technologies, University of Gdańsk, 80-308 Gdańsk, Poland*

<sup>3</sup>*ICFO-Institut de Ciències Fotoniques, The Barcelona Institute of Science and Technology, Castelldefels (Barcelona), 08860, Spain*

<sup>4</sup>*QuTech, Delft University of Technology, P.O. Box 5046, 2600 GA Delft, Netherlands*

 (Received 20 May 2020; revised 5 November 2020; accepted 3 March 2021; published 22 April 2021)

We investigate the problem of simulating classical stochastic processes through quantum dynamics and present three scenarios where memory or time quantum advantages arise. First, by introducing and analyzing a quantum version of the embeddability problem for stochastic matrices, we show that quantum memoryless dynamics can simulate classical processes that necessarily require memory. Second, by extending the notion of space-time cost of a stochastic process  $P$  to the quantum domain, we prove an advantage of the quantum cost of simulating  $P$  over the classical cost. Third, we demonstrate that the set of classical states accessible via Markovian master equations with quantum controls is larger than the set of those accessible with classical controls, leading, e.g., to a potential advantage in cooling protocols.

DOI: [10.1103/PhysRevX.11.021019](https://doi.org/10.1103/PhysRevX.11.021019)

Subject Areas: Quantum Physics, Quantum Information

## I. INTRODUCTION

### A. Memory advantages

What tasks can we perform more efficiently by employing quantum properties of nature? And what are the quantum resources powering them? These are the central questions that need to be answered not only to develop novel quantum technologies but also to deepen our understanding of the foundations of physics. Over the last few decades, these questions were successfully examined in the context of cryptography [1], computing [2], simulations [3], and sensing [4], proving that the quantum features of nature can indeed be harnessed to our benefit.

More recently, an area of active theoretical and experimental interest focused on the memory advantages offered by quantum mechanics for the simulation of stochastic processes in the setting of classical causal models [5–8]. An experimentally accessible and relevant measure of such an advantage is the dimensionality of the memory required for the simulation [8,9]. These dimensional advantages have been identified experimentally (a qubit system has been

used to simulate a stochastic process that classically requires three bits [8]) and theoretically for a certain class of Poisson processes [9].

Here, we take a complementary approach starting from the following simple observation: Although all fundamental interactions are memoryless, the basic information-processing primitives (such as the bit-swap operation) cannot be performed classically in a time-continuous fashion without employing implicit microscopic states that act as a memory [10]. We show that this picture changes dramatically if instead we consider memoryless quantum dynamics. This difference is due to quantum coherence, arising from the superposition principle, which can effectively act as an internal memory of the system during the evolution.

### B. Classical vs quantum

But what do we really mean when we say that a bit-swap (or other information-processing tasks) cannot be performed classically in a memoryless way? First, when we speak of a bit, we mean a fundamentally two-level system, i.e., a system with only two microscopic degrees of freedom (e.g., a spin-1/2 particle) and not a macroscopic object with a coarse-grained description having two states (e.g., a piece of iron magnetized along or against the  $z$  axis). Otherwise, if the system merely implements a bit in a higher-dimensional state space of dimension  $d > 2$ , the internal degrees of freedom can be used as a memory and a bit swap can be performed, as illustrated in Fig. 1(a) for  $d = 3$ . Thus, when we speak of classical systems, we take them to be fundamentally  $d$  dimensional, and when we

\*korzekwa.kamil@gmail.com

†lostaglio@protonmail.com

‡These authors contributed equally to this work.

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

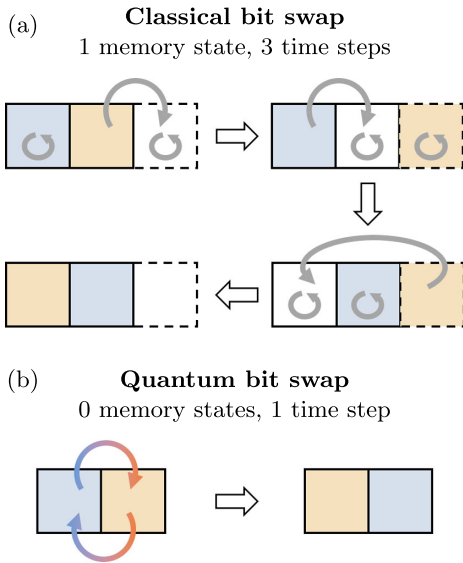


FIG. 1. Space-time cost for classical and quantum bit swap. (a) Space-time optimal realization of a bit swap, i.e., a transposition between two states (solid-line boxes), using one memory state (dashed-line boxes) and three time steps. Each time step is composed of a continuous memoryless dynamics that does not affect one of the states and maps the remaining two to one of them. (b) In the quantum regime, a bit swap can be performed without any memory, simply by a time-continuous unitary process  $\exp(i\sigma_x t)$  that continuously connects the identity operation at time  $t = 0$  with the bit swap, represented by Pauli  $x$  operator  $\sigma_x$ , at time  $t = \pi/2$ . During the process, the information about the initial state of the system is preserved in quantum coherence.

speak of memoryless dynamics, we mean probabilistic jumps between these discrete states occurring at rates independent of the system’s past. Of course, this is just the standard setting for classical Markov processes. Then, for every finite dimension  $d$ , there exist processes that cannot be performed in a memoryless fashion.

This setting should also be contrasted with classical systems with a continuous phase space. The simplest example is probably given by an isolated classical pendulum, i.e., a harmonic oscillator. The state of the system is then given by a point in (more generally, distribution over) a two-dimensional phase space  $[x, p]$  describing its position  $x$  relative to the equilibrium position, and its momentum  $p$ . If we now identify the pendulum’s state  $[-x_0, 0]$  with a bit state 0 and  $[x_0, 0]$  with a bit state 1, the time evolution of the system clearly performs a bit-swap operation within half a period. However, a pendulum is not a simple two-level system but rather a system with a continuously infinite number of states. Thus, the bit swap is performed by employing infinitely many ancillary memory states:  $[-x_0, 0]$  evolves through states with  $x > -x_0$  and positive momenta  $p$  to  $[x_0, 0]$ , while  $[x_0, 0]$  evolves through states with  $x < x_0$  and negative momenta  $p$  to  $[-x_0, 0]$ , and the momentum  $p$  effectively acts as a register that carries the

information about the past. For classical systems with a continuous phase space, it is then difficult to properly assess the number of memory states used during a given evolution. Hence, in this work, we only focus on discrete systems.

Ultimately, we are then interested in a stochastic process in which discrete outputs  $i$  are observed for given discrete inputs  $j$ . This process is characterized by a matrix of transition probabilities  $P_{ij}$ . In the classical setting, we want to know whether there exists a classical memoryless dynamics (described by a Markovian master equation) involving these states that outputs  $P_{ij}$  after some time. Quantum mechanically, we are similarly asking whether a *quantum* memoryless dynamics (described by a Markovian quantum master equation) can output  $P_{ij}$  after some time. In this work, we highlight that these two questions admit very different answers, both in terms of which  $P_{ij}$  can arise from memoryless processes and in terms of the memory required to achieve a given  $P_{ij}$ .

For concreteness, assume  $P_{ij}$  results from a thermalization process, which typically satisfies the so-called detailed balance condition. Physically, we are then asking whether the observed  $P_{ij}$  is compatible or not with a process involving no memory effects, such as information backflows from the environment [11]. Classically,  $P_{ij}$  originates from incoherent jumps induced by interacting with the environment (absorbing or emitting energy). Alternatively, we can see classical dynamics as the evolution of a quantum system that undergoes very strong decoherence at all times, so that any nonclassical effects are killed right away. In fact, standard quantum thermalization models (weak coupling with a very large thermal bath) can also be understood in this way since they are unable to generate quantum superpositions of energy states. As soon as we move away from this semiclassical limit, however, we see that more exotic thermalization processes can generate  $P_{ij}$  that, classically, would necessarily signal memory effects but that, quantum mechanically, can emerge from memoryless processes due to quantum coherence. Note that the term “memoryless” is used throughout the paper as a synonym of Markovian, i.e., that the evolution only depends on the current state of the system and not on its history. Such evolution may still require an auxiliary clock system (used, e.g., to know how much longer the system should be coupled to an external control field) and a counter system (used to record the current channel in the sequence of channels necessary to implement the given dynamics). These constitute extra resources that one may want to separately account for, e.g., using the framework of quantum clocks (see, e.g., Ref. [12]).

### C. Summary of results

In this work, we identify three aspects of potential quantum advantage in simulating stochastic processes.

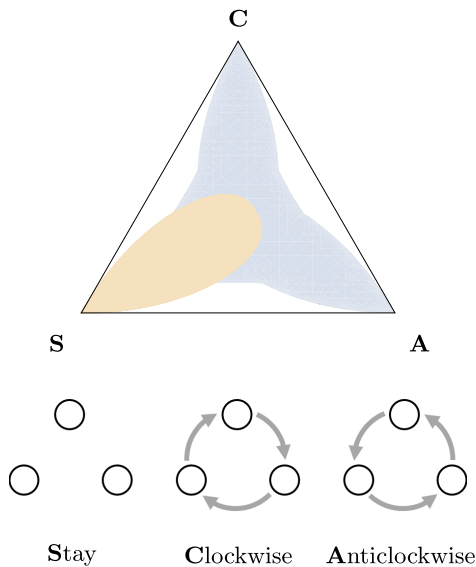


FIG. 2. Classical vs quantum memoryless processes. The vertices of the triangle correspond to deterministic processes (S, stay; C, move clockwise; A, move anticlockwise) for a random walker moving between three states. Points inside the triangle correspond to probabilistic mixtures (convex combinations) of these three deterministic processes; e.g., the center of the triangle corresponds to the maximally mixing dynamics (with S, C, and A each happening with probability  $1/3$ ). The orange petal-shaped region contains all stochastic processes that can arise from time-continuous, memoryless, classical dynamics. For time-continuous, memoryless, quantum dynamics, this set is enlarged by the remaining shaded region in blue. For details, see Sec. II C and, in particular, Fig. 4.

First, in Sec. II, we investigate the possibility to simulate classical processes requiring memory using quantum memoryless dynamics. More precisely, we compare the sets of all stochastic processes that can be generated by time-continuous memoryless dynamics in the classical and quantum domains. We prove that the latter set is strictly larger than the former one, i.e., that there exist stochastic processes that classically require memory to be implemented but can be realized by memoryless quantum dynamics. As an example, consider a random walk on a cyclic graph with three sites, where the walker can either move clockwise, move anticlockwise, or stay in place. As we present in Fig. 2, only a small orange subset of such walks can arise from a continuous classical evolution that does not employ memory (note that, differently from other investigations [13], we do not put any restriction on the classical dynamics beyond the fact that it is memoryless). However, if we allow for continuous, memoryless, quantum evolution, all stochastic processes in the much larger blue set can be achieved. Besides this particular class, in this work, we provide general constructions for whole families of stochastic processes for any finite-dimensional systems that require memory classically but can be implemented quantumly in a memoryless fashion.

Second, in Sec. III, we go beyond the simple distinction between stochastic processes that can or cannot be simulated without memory, and take a more quantitative approach, thus investigating quantum memory advantages. To this end, we employ the recent formalism of Ref. [10], which allows one to quantify the classical space-time cost of a given stochastic process, i.e., the minimal amount of memory and time steps needed to classically implement a given process. We extend this approach to the quantum domain in order to analyze the quantum space-time cost. An illustrative example is given by the bit-swap process presented in Fig. 1, which, in the classical setting, requires either one memory state and three time steps, or two memory states and two time steps. However, if one allows for quantum evolution, such a bit swap can be performed in a continuous and memoryless fashion through a simple unitary evolution  $\exp(i\sigma_x t)$ , with  $\sigma_x$  denoting the Pauli  $x$  operator. More generally, the authors of Ref. [10] have characterized the space-time cost for the family of  $\{0, 1\}$ -valued stochastic processes (i.e., all discrete functions). Their bound shows an unavoidable classical trade-off between the number of memory states  $m$  and the number of time steps  $\tau$  needed to realize a given stochastic process on  $N$  systems of dimension  $d$ . Crucially, a typical process necessarily requires extra resources, meaning that either  $m$  or  $\tau$  is exponential in  $N$ . In this paper, we prove that in the quantum regime, all such processes can be simulated with zero memory states and in, at most, two time steps, demonstrating an advantage over the best-possible classical implementation.

Third, in Sec. IV, we study memory advantages in control by comparing classical and quantum continuous memoryless dynamics in terms of the set of accessible final states. We assume a fixed point of the evolution is given, which is a realistic physical constraint in dissipative processes and typically, but not necessarily, coincides with the thermal Gibbs state. A standard example is given by a thermalization of the system to the environmental temperature. Here, we employ our recent result [14] characterizing the input-output relations of classical Markovian master equations with a given fixed point. We show how quantum memoryless dynamics with the same fixed point allows one to access a larger set of final states. This is most evident in the case of maximally mixed fixed points (corresponding to the environment in the infinite-temperature limit) since every transformation that is classically possible with arbitrary amounts of memory can be realized in a memoryless fashion in the quantum domain. For general fixed points, we prove that an analogous result holds for systems of dimension  $d = 2$ , and we argue that the set of accessible states is strictly larger in the quantum regime than in the classical one for all  $d$ . Since it is known that memory effects enhance cooling [15,16], a direct consequence of our results is that quantumly it is possible to bring the two-dimensional system below the

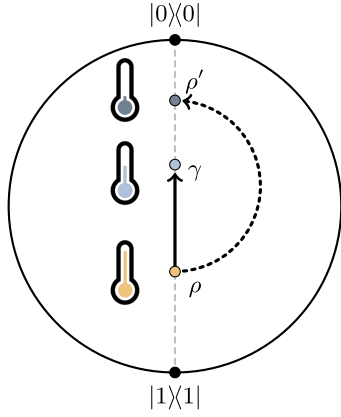


FIG. 3. Markovian cooling of a qubit. Classical memoryless processes can only cool the initial state  $\rho$  of a two-dimensional system to the thermal state  $\gamma$  at the environmental temperature (path along the solid-line arrow). Quantum memoryless dynamics with fixed point  $\gamma$  allows one to cool the system below that, all the way to the state  $\rho'$  with the lowest temperature achievable by classical processes with memory (path along the dotted-line arrow). For details, see Sec. IV B and, in particular, Fig. 8.

environmental temperature without employing memory effects, something that is impossible classically (see Fig. 3).

Finally, in Sec. V, we discuss the potential for practical applications of our results, while Sec. VI contains the outlook for future research.

## II. EMBEDDABILITY OF STOCHASTIC PROCESSES

### A. Classical embeddability

Given a discrete state space,  $\{1, \dots, d\}$ , the state of a finite-dimensional classical system is described by a probability distribution  $\mathbf{p}$  over these states. A stochastic matrix or process  $P$  is a matrix  $P_{ij}$  of transition probabilities,

$$P_{ij} \geq 0, \quad \sum_i P_{ij} = 1, \quad (1)$$

which describes the evolution of the system from one state,  $\mathbf{p}$ , to another,  $P\mathbf{p}$ .

Classically, the  $P_{ij}$  that can be achieved without employing memory are known as embeddable. A stochastic matrix  $P$  is embeddable if it can be generated by a continuous Markov process [17]. This notion can be understood as a control problem involving a master equation. Namely, introducing a rate matrix or generator  $L$  as a matrix with finite entries satisfying

$$L_{ij} \geq 0 \quad \text{for } i \neq j, \quad \sum_i L_{ij} = 0, \quad (2)$$

a continuous one-parameter family  $L(t)$  of rate matrices generates a family of stochastic processes  $P(t)$  satisfying

$$\frac{d}{dt}P(t) = L(t)P(t), \quad P(0) = 1. \quad (3)$$

The aim of the control  $L(t)$  is to realize a target stochastic process  $P$  at some final time  $t_f$  as  $P = P(t_f)$ . If this is possible for some choice of  $L(t)$ , then  $P$  is embeddable, and if there exists a time-independent generator  $L$  such that  $P = e^{Lt_f}$ , then we say that  $P$  can be embedded by a time-homogeneous Markov process. A final technical comment is that we also consider the case  $t_f = \infty$  to be embeddable (in Ref. [10], this case was referred to as limit-embeddable). Then,  $P$  cannot be generated in any finite time, but it can be approximated arbitrarily well. This is the case, e.g., with the bit erasure process:  $0 \mapsto 0, 1 \mapsto 0$  [10].

The question of which stochastic matrices  $P$  are embeddable is a challenging open problem that has been extensively investigated for decades [17–22]. The full characterization does not go beyond  $2 \times 2$  and  $3 \times 3$  stochastic matrices; however, various necessary conditions have been found. In particular, in Ref. [21], it was proven that every embeddable stochastic matrix  $P$  satisfies the following inequalities:

$$\prod_i P_{ii} \geq \det P \geq 0. \quad (4)$$

The condition  $\det P \geq 0$  is, in fact, also known to be sufficient in dimension  $d = 2$  [19], and a time-independent rate matrix  $L$  can then be found.

**Example 1** (Thermalization). Consider a two-level system with energy gap  $E$  incoherently exchanging energy with a large environment at inverse temperature  $\beta$ . Assume  $\beta E = 1$ . Suppose one observes  $P_{ij}$  satisfying the detailed balance condition:  $P_{1|0} = P_{0|1}e^{-1}$ . Can this stochastic process originate from a memoryless dynamics? Using the condition in Eq. (4), one can verify this is the case if and only if  $P_{0|1} \leq e/(1+e) \approx 0.731$ . The dynamics that realizes  $P$  is then a standard thermalization process whereby the system's state  $\mathbf{p}$  exponentially relaxes to the equilibrium distribution  $\boldsymbol{\gamma} = (e/(1+e), 1/(1+e))$ , according to the classical master equation:

$$\frac{d}{dt}\mathbf{p}(t) = R(\boldsymbol{\gamma} - \mathbf{p}(t)), \quad (5)$$

where  $R$  denotes the thermalization rate. The stochastic process  $P$  is realized by a partial thermalization lasting for a time  $t = -\log[1 - P_{0|1}(1+e)/e]/R$ . Intuitively, the deexcitation probability  $P_{0|1}$  cannot be made larger than  $e/(1+e)$  because memoryless thermalizations need to satisfy detailed balance at every intermediate time step (so there is always some probability of absorbing an excitation from the bath). In fact,  $P_{0|1} > e/(1+e)$  can only be realized if memory effects are present. An example of such a process is a detailed balanced  $P$  with  $P_{0|1} = 1$ ,

which is the “ $\beta$ -swap” providing enhanced cooling in Ref. [15]. Indeed, the latter process can be approximated by a Jaynes-Cummings coupling of a two-level system to an “environment” given by a single harmonic oscillator initialized in a thermal state, which gives rise to a highly non-Markovian evolution of the system.

## B. Quantum embeddability

A state of a finite-dimensional quantum system is given by a density operator  $\rho$ , i.e., a positive semidefinite operator with trace one that acts on a  $d$ -dimensional Hilbert space  $\mathcal{H}_d$ . A general evolution of a density matrix is described by a quantum channel  $\mathcal{E}$ , which is a completely positive (CP) trace-preserving map from the space of density matrices to itself. Now, focusing on the computational basis  $\{|k\rangle\}_{k=1}^d$  of  $\mathcal{H}_d$ , suppose we input the quantum state  $\rho_p = \sum_k p_k |k\rangle\langle k|$ , apply the channel  $\mathcal{E}$ , and measure the resulting state  $\mathcal{E}(\rho_p)$  in the computational basis. The measurement outcomes will be distributed according to  $Pp$ , where

$$P_{ij} = \langle i | \mathcal{E}(|j\rangle\langle j|) | i \rangle. \quad (6)$$

In this way, the preparation of  $\rho_p$ , followed by a channel  $\mathcal{E}$  and the computational basis measurement, simulates the action of a stochastic process  $P$  on the classical state  $p$ .

Surprisingly, as far as we are aware, the set of processes  $P$  that can be simulated by a quantum process without employing memory has not been named or studied before. Hence, we define a stochastic matrix  $P$  as quantum embeddable if it can be simulated by a quantum process as in Eq. (6) with  $\mathcal{E}$  a Markovian quantum channel [23], i.e., a channel that can result from a Markovian master equation [the quantum analogue of Eq. (3)].

Here, we describe what it rigorously means for a quantum channel  $\mathcal{E}$  to be Markovian. Despite the difference in jargon between the two communities, Markovianity is for channels what embeddability is for stochastic matrices. It captures the fact that  $\mathcal{E}$  can be realized without employing memory effects. It can also be understood as a control problem but, this time, involving a quantum master equation. More precisely, the rate matrix  $L$  is replaced by a Lindbladian [24,25], which is a superoperator  $\mathcal{L}$  acting on density operators and satisfying

$$\mathcal{L}(\cdot) = -i[H, \cdot] + \Phi(\cdot) - \frac{1}{2} \{ \Phi^*(\mathbb{1}), \cdot \}, \quad (7)$$

with the first term describing unitary evolution and the remaining ones encoding the dissipative dynamics, e.g., due to the interaction with an external environment. Here,  $H$  is a Hermitian operator,  $[A, B] := AB - BA$  denotes a commutator,  $\Phi$  is a completely positive superoperator,  $\Phi^*$  denotes the dual of  $\Phi$  under the Hilbert-Schmidt scalar product, and  $\{A, B\} := AB + BA$  stands for the

anticommutator. In analogy with Eq. (3), a continuous one-parameter family of Lindbladians  $\mathcal{L}(t)$  generates a family of quantum channels  $\mathcal{E}(t)$  satisfying

$$\frac{d}{dt} \hat{\mathcal{E}}(t) = \hat{\mathcal{L}}(t) \hat{\mathcal{E}}(t), \quad \hat{\mathcal{E}}(0) = \hat{\mathcal{I}}, \quad (8)$$

where hats indicate superoperators (i.e., matrix representations of quantum channels that act on vectorized quantum states) and  $\mathcal{I}$  denotes the identity channel. A quantum channel  $\mathcal{E}$  is Markovian [23] if  $\mathcal{E} = \mathcal{E}(t_f)$  for some choice of the Lindbladian  $\mathcal{L}(t)$  and  $t_f$  (perhaps  $t_f = +\infty$ ). In other words,  $\mathcal{E}$  is Markovian if it is a channel that results from integrating a quantum master equation up to some time  $t_f$ . Any given Markovian channel  $\mathcal{E}$  gives a stochastic process  $P$  through Eq. (6). The aim of the control  $\mathcal{L}(t)$  is to achieve a target stochastic matrix  $P$  after some time  $t_f$ . More formally, we introduce the following definition.

**Definition 1** (Quantum-embeddable stochastic matrix). A stochastic matrix  $P$  is quantum embeddable if

$$P_{ij} = \langle i | \mathcal{E}(|j\rangle\langle j|) | i \rangle, \quad (9)$$

where  $\mathcal{E}$  is a Markovian quantum channel.

**Example 2.** Consider a two-point projective measurement scheme (TPM) [26]. First, a projective energy measurement is performed, and one finds the system in a well-defined energy state  $j$ . Then, a quantum evolution  $\mathcal{E}$  follows, and finally, a second energy measurement returns the outcome  $i$  with probability  $P_{ij}$ . We can then ask whether it is possible that the process  $\mathcal{E}$  generating  $P_{ij}$  resulted from a Markovian quantum master equation. Suppose one observes

$$P = \begin{bmatrix} 1/3 & 2/3 \\ 2/3 & 1/3 \end{bmatrix}. \quad (10)$$

Then, it is straightforward to show that the above can arise from the following unitary dynamics  $U$  (which is a Markovian channel),

$$U = \begin{bmatrix} \sqrt{1/3} & \sqrt{2/3} \\ \sqrt{2/3} & -\sqrt{1/3} \end{bmatrix}, \quad (11)$$

while according to Eq. (4), it is impossible to generate such a  $P$  using classical memoryless dynamics.

## C. Quantum advantage

One can easily see that all (classically) embeddable stochastic processes are also quantum embeddable: Given a classical generator  $L$ , one chooses the CP map  $\Phi$  defining the Lindbladian  $\mathcal{L}$  in Eq. (7) to be

$$\Phi(\cdot) = \sum_{ij} K_{ij}(\cdot) K_{ij}^\dagger, \quad K_{ij} = \sqrt{L_{ij}} |i\rangle\langle j|. \quad (12)$$

However, the converse is not true. There exist many stochastic matrices  $P$  that can be generated by a quantum, but not a classical, Markov process. The simplest example is given by a nontrivial permutation  $\Pi$ , satisfying

$$\det \Pi = \pm 1, \quad \prod_i \Pi_{ii} = 0. \quad (13)$$

Clearly, Eq. (4) is violated, and hence  $\Pi$  is not embeddable. However, noting that every unitary channel  $U(\cdot)U^\dagger$  is Markovian [by choosing the Lindbladian with no dissipative part and  $H$  such that  $U = \exp(iHt_f)$ ] and that a permutation matrix  $\Pi$  is unitary, we conclude that every permutation  $\Pi$  is quantum embeddable. This conclusion also proves that neither of the two conditions in Eq. (4) is necessary for quantum embeddability.

More generally, a larger class of stochastic matrices that are quantum embeddable is given by the set of unistochastic matrices [27,28]. These are defined as all stochastic matrices  $P$  satisfying

$$P_{ij} = |\langle j|U|i\rangle|^2 \quad (14)$$

for some unitary matrix  $U$ , and the argument for quantum embeddability is analogous to the one given for permutation matrices. The set of unistochastic matrices includes permutations but also other (classically) nonembeddable stochastic matrices. As an example, one can consider the bistochastic matrix  $P$  in Eq. (10) or, in fact, any other  $2 \times 2$  bistochastic matrix. This is because in dimension  $d = 2$ , every bistochastic matrix is unistochastic; thus, it is quantum embeddable.

Beyond these examples, we prove a simple general result that allows one to find larger families of quantum-embeddable stochastic matrices.

**Lemma 1** (Monoid property). The set of quantum-embeddable stochastic matrices contains identity and is closed under composition; i.e., if  $P$  and  $Q$  are quantum embeddable, then  $PQ$  is also.

*Proof.*—First, identity is obviously quantum embeddable as it arises from a trivial Lindbladian  $\mathcal{L} = 0$ . Now, note that the composition of Markovian quantum channels gives a Markovian quantum channel. Next, notice that a completely dephasing map

$$\mathcal{D}(\cdot) := \sum_k \langle k| \cdot |k\rangle |k\rangle\langle k| \quad (15)$$

is a Markovian quantum channel. Finally, the composition  $\mathcal{E} = \mathcal{E}_P \circ \mathcal{D} \circ \mathcal{E}_Q$ , with  $\mathcal{E}_P$  and  $\mathcal{E}_Q$  being quantum channels describing the quantum embeddings of  $P$  and  $Q$ , is a Markovian quantum channel that quantum embeds the stochastic process described by  $PQ$ . ■

Let us now discuss the consequences of Lemma 1 with increasing generality. We start with the following corollary for dimension  $d = 2$ .

**Corollary 2.** All  $2 \times 2$  stochastic matrices are quantum embeddable.

*Proof.*—A general  $2 \times 2$  stochastic matrix  $P$  can be written as

$$P = \begin{bmatrix} a & 1-b \\ 1-a & b \end{bmatrix}. \quad (16)$$

If  $\det P \geq 0$ , then  $P$  is embeddable and hence quantum embeddable. Otherwise, if  $\det P < 0$ , we can write  $P = \Pi P'$ , with  $\Pi$  denoting the nontrivial  $2 \times 2$  permutation and  $P'$  being a stochastic matrix with  $\det P' \geq 0$ . Since  $P$  can be written as a composition of two quantum-embeddable maps, by Lemma 1, it is also quantum embeddable. ■

For  $d \geq 3$ , comparing quantum and classical embeddability becomes complicated because of the lack of a complete characterization of classical embeddability. However, one can focus on certain subclasses of stochastic processes that are better understood. For example, for the family of  $3 \times 3$  circulant stochastic matrices, defined by

$$P = \begin{bmatrix} 1-a-b & a & b \\ b & 1-a-b & a \\ a & b & 1-a-b \end{bmatrix}, \quad (17)$$

the necessary and sufficient conditions for (classical) embeddability are known. Denoting the eigenvalues of  $P$  by  $\lambda_k = r_k e^{i\theta_k}$  with  $\theta_k \in [-\pi, \pi]$ , these conditions are given by [29]

$$\forall k: r_k \leq \exp[-\theta_k \tan(\pi/3)]. \quad (18)$$

We illustrate the set of classically embeddable circulant matrices by the green region in parameter space  $[a, b]$  in Fig. 4. On the other hand, because of Lemma 1, quantum-embeddable circulant stochastic matrices also include permutations of  $P$ , i.e.,  $\Pi P$ , with  $\Pi$  denoting circulant  $3 \times 3$  permutation matrices. In fact, this set contains not only permutations of  $P$  but also compositions of  $P$  with any unistochastic matrix; however, numerical verification suggests that this does not further expand the investigated set. As a result, the set of quantum-embeddable stochastic matrices in the parameter space  $[a, b]$  contains not only the region corresponding to classically embeddable matrices but also its two copies (corresponding to two permutations), which we illustrate in blue in Fig. 4. Moreover, all unistochastic circulant matrices (which are fully characterized by the “chain-links” conditions from Ref. [28]) are also quantum embeddable. The resulting region is also plotted in Fig. 4 in orange. Thus, we clearly see that the set of quantum-embeddable stochastic circulant matrices is much larger than the classically embeddable one since it

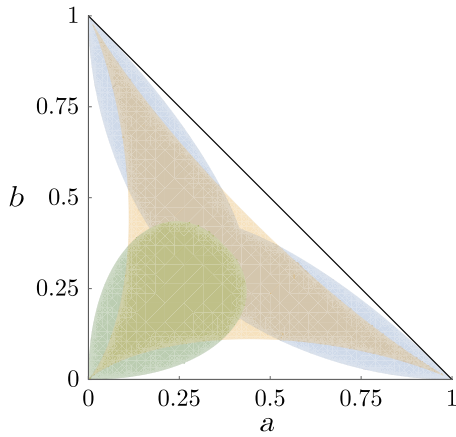


FIG. 4. Embeddability of  $3 \times 3$  circulant matrices. Every  $3 \times 3$  circulant matrix corresponds to a point within a half-square in the parameter space  $[a, b]$  according to Eq. (17). The green petal-shaped region around the origin contains all (classically) embeddable matrices. The set of quantum-embeddable matrices is larger and contains the orange triangle-like region of unistochastic matrices, as well as two, blue, petal-like regions corresponding to permutations of the classically embeddable region.

contains the union of green, blue, and orange regions. However, we do not expect that all  $3 \times 3$  circulant matrices are quantum embeddable, with the case  $a = b = 1/2$  being intuitively the least likely to arise from quantum Markovian dynamics.

For general dimension  $d$ , one can generate families of quantum-embeddable matrices using Lemma 1 in an analogous way, by composing classically embeddable matrices with unistochastic ones. Moreover, employing Corollary 2, we note that the set of quantum-embeddable stochastic matrices also includes all matrices  $P$  that can be written as products of elementary stochastic matrices  $P_{e_i}$  (also known as pinching matrices), i.e.,

$$P = P_{e_n} \dots P_{e_1}, \quad P_{e_i} = \Pi_i (P_2 \oplus I_{d-2}) \Pi_i, \quad (19)$$

where  $P_2$  is a general  $2 \times 2$  stochastic matrix,  $I_{d-2}$  denotes identity on the remaining states, and  $\Pi_i$  is an arbitrary permutation. Notably, for  $d \geq 4$ , such products of elementary stochastic matrices contain matrices that are not unistochastic [30] and hence cannot be reduced to the examples above. In conclusion, quantum embeddings allow one to achieve many stochastic processes that necessarily require memory from a classical standpoint.

#### D. Discussion

From a physical perspective, it is now natural to ask why the set of quantum-embeddable stochastic matrices is strictly larger than the set of classically embeddable ones. To address this question, let us first consider the simple example of classically nonembeddable permutation matrices,

$$\Pi_m = \sum_{n=1}^d |n \oplus m\rangle \langle n|, \quad m = 1, \dots, d-1, \quad (20)$$

where  $\oplus$  here denotes addition modulo  $d$ . A direct calculation shows that  $\Pi_m = e^{iHm}$ , with Hamiltonian

$$H = \sum_{n=1}^d \frac{2\pi(n-1)}{d} |\psi_n\rangle \langle \psi_n|, \quad (21)$$

and

$$|\psi_n\rangle = \frac{1}{\sqrt{d}} \sum_{k=1}^d e^{-i2\pi(k-1)(n-1)/d} |k\rangle. \quad (22)$$

We thus see that the continuous and memoryless Hamiltonian evolution creates a superposition of classical states  $|n\rangle$  on the way between identity and  $\Pi_m$ . The intuitive picture that emerges is that the quantum superposition between classical states created during the evolution effectively acts as a memory. For example, when we perform a rotation of the Bloch sphere around the  $y$  axis, we can implement a bit swap sending  $|0\rangle$  to  $|1\rangle$  and vice versa, but the path the state follows (going through  $|+\rangle$  if the initial state is  $|0\rangle$  and through  $|-\rangle$  if the initial state is  $|1\rangle$ ) will preserve the memory about the initial state. At the same time, a classical memoryless process moving  $(1,0)$  towards  $(0,1)$  and  $(0,1)$  towards  $(1,0)$  cannot proceed beyond the point at which the two trajectories meet.

One might wonder if we can quantify the coherent resources required for the advantage. There are various frameworks that have been put forward to quantify superposition (the resource theory of coherence in its various forms [31] or that of asymmetry [32]). However, none of them associates costs to permutations (technically, these are “free operations”) despite the fact that they carry an advantage in our setting. What we allude to in the present discussion is that, while these theories assign no cost to these operations, they can be performed in a Markovian fashion only because one can continuously connect different basis states through the creation of superpositions. Thus, current frameworks seem inadequate to capture the resources involved in the quantum memory advantage. An alternative framework would have to quantify the maximum amount of coherence that must be created at intermediate times, minimized over all Markovian realizations of a target channel. This quantity may then be given an operational meaning in terms of minimal coherent resources one must input to realize the corresponding protocol. We leave this research direction for future work.

Our results can also be naturally connected to a result by Montana [33], who proved that Markovian hidden variable models reproducing quantum mechanical predictions necessarily require a number of continuous variables that grow linearly with the Hilbert space dimension and, hence,



exponentially with the system size. Intuitively, here we are showing that this “excess baggage” [34] can be exploited to simulate memory effects. In what follows, we provide a quantification of the advantage beyond the embeddable-nonembeddable dichotomy. We will see that quantum theory allows for advantages in the simulation of stochastic processes by memoryless dynamics.

### III. SPACE-TIME COST OF A STOCHASTIC PROCESS

In this section, we first recall a recently introduced framework for the quantification of the space and time costs of simulating a stochastic process by memoryless dynamics [10]. We then extend it to the quantum domain and prove a quantum advantage in the corresponding costs.

#### A. Classical space-time cost

Let  $P$  be a nonembeddable stochastic matrix acting on  $d$  so-called visible states. We then ask, how many additional memory states  $m$  does one need to add, in order to implement  $P$  by a classical Markov process? Formally, one looks for an embeddable stochastic matrix  $Q$  acting on  $d + m$  states whose restriction to the first  $d$  rows and columns is identical to  $P$ . When this happens,  $Q$  is said to implement  $P$  with  $m$  memory states. In fact, given any  $d$ -dimensional distribution  $\mathbf{p}$ , if we take the  $d + m$ -dimensional distribution  $\mathbf{q} = (\mathbf{p}, 0, \dots, 0)$ , then  $Q\mathbf{q} = (P\mathbf{p}, 0, \dots, 0)$ . Following Ref. [10], we now have the following definition.

**Definition 2** (Space cost). The space cost of a  $d \times d$  stochastic matrix  $P$ , denoted  $C_{\text{space}}(P)$ , is the minimum  $m$  such that a  $(d + m) \times (d + m)$  embeddable matrix  $Q$  implements  $P$ .

As a technical comment, we note that the above definition can be extended to situations in which visible and memory states are not disjoint, e.g., when the visible states on which  $P$  acts are logical states defined by a coarse graining of the states on which  $Q$  acts. Since this does not change any of the results presented here, we refer to Ref. [10] for further details and adopt the simpler definition given here.

Once we find a matrix  $Q$  that implements  $P$ , the next question is as follows: What is the number of time steps required to realize  $Q$ ? The notion of a time step is meant to capture the number of independent controls that are needed to achieve  $Q$ . A natural definition would be that the number of time steps necessary to realize an embeddable stochastic matrix  $Q$  is the minimum number  $n$  such that

$$Q = e^{L^{(n)}t_n} \dots e^{L^{(1)}t_1}, \quad (23)$$

where  $L^{(1)}, \dots, L^{(n)}$  are time-independent generators; i.e., each  $L^{(k)}$  is a control applied for some time  $t_k$ . This definition captures the idea of a sequence of autonomous steps, with  $n$  the number of times an active intervention is

required to “quench”  $L^{(i)}$  to  $L^{(i+1)}$ . If  $n = 1$ , one has an almost autonomous protocol; i.e., the only requirement is the ability to switch off the controls after time  $t_1$ .

From a physical point of view, the issue with this definition is that it assigns an infinite cost to any realistic protocol in which controls are switched on and off in a continuous fashion. For example, suppose a single two-level system is kept in contact with an idealized dissipative environment while we slowly tune its energy gap. Such a protocol can be seen as the continuous limit of a sequence of steps described in Eq. (23). As such, it would be assigned an infinite time cost according to the above definition, even though it is certainly experimentally feasible. To overcome this issue, note that by Levy’s lemma [35], a crucial property of each step in the sequence of Eq. (23) is that the set of nonzero transition probabilities does not change, which naturally suggests defining the time cost as the number of times the set of “blocked” transitions changes. One could physically motivate this definition as follows. Consider a system interacting with a large thermal environment. Transitions between any pairs of the system’s energy levels are possible by absorption or emission of the corresponding energy from or to the bath. Absorptions are exponentially suppressed in the energy gap. To selectively couple only certain levels, we either need to raise and lower infinite energy barriers [10] or we need to engineer the spectrum of the environment so that only certain transitions can occur. Changing the set of the system’s energy levels involved in the interaction, by decoupling some and coupling new ones, is then a nontrivial control operation, and we hence assign a cost to it. The time cost, defined as the number of times we need to change the set of coupled energy levels, is then a good proxy for the level of required control. It solves the issue with the previous definition, and, in particular, it assigns a cost  $n = 1$  to the qubit protocol mentioned above.

To sum up, these considerations lead to the following definition of a one-step process [10].

**Definition 3** (One-step process). A stochastic matrix  $T$  is called a one-step process if

- (1) it is embeddable;
- (2) the controls  $L(t)$  that generate  $T$  at time  $t_f$  through Eq. (3) can be chosen such that the set of nonzero transition probabilities of  $P(t)$  is the same for all  $t \in (0, t_f)$ .

Putting all this together, we obtain the notion of time cost from Ref. [10].

**Definition 4** (Time cost). The time cost  $C_{\text{time}}(P, m)$  of a  $d \times d$  stochastic matrix  $P$ , while allowing for  $m$  memory states, is the minimum number  $\tau$  of one-step stochastic matrices  $T^{(i)}$  of dimension  $(d + m) \times (d + m)$  such that  $Q = T^{(\tau)} \dots T^{(1)}$  implements  $P$ .

#### B. Quantum space-time cost

The framework presented above allows one to quantify the memory and time costs of implementing a

given stochastic process by classical master equations. We now introduce a natural extension of the above to the quantum domain.

**Definition 5** (Quantum space cost). The quantum space cost of a  $d \times d$  stochastic matrix  $P$ , denoted  $Q_{\text{space}}(P)$ , is the minimum  $m$  such that the  $(d+m) \times (d+m)$  quantum-embeddable matrix  $Q$  implements  $P$ .

Concerning the definition of the quantum time cost of a stochastic matrix, the physical intuitions discussed in the classical case apply essentially unchanged. We can now add to those intuitions the standard example of a sequence of  $n$  laser pulses, each involving a fixed energy submanifold. Naturally, these will count as  $n$  time steps. In analogy with the classical counterpart, the quantum time cost admits more or less restrictive definitions. Since we eventually want to prove a quantum advantage, in the quantum regime it is convenient to adopt the simpler and more constrained definition. Thus, we use the definition involving time-independent generators only, as an advantage proven according to such a restricted scenario will persist if we allow even more freedom in the quantum protocol.

**Definition 6** (Quantum time cost) The quantum time cost  $Q_{\text{time}}(P, m)$  of a  $d \times d$  stochastic matrix  $P$ , while allowing for  $m$  memory states, is the minimum  $\tau$  such that there exist time-independent Lindbladians  $\mathcal{L}_1, \dots, \mathcal{L}_\tau$  on a  $(d+m) \times (d+m)$ -dimensional Hilbert space and

$$Q_{ij} = \langle i | e^{\mathcal{L}^{(i)}\tau} \dots e^{\mathcal{L}^{(1)}\tau} (|j\rangle\langle j|) | i \rangle \quad (24)$$

implements  $P$ .

At this point, it is worth highlighting that this notion of time step, which is a quantum version of that in Ref. [10], is disconnected from the one used in quantum computing. There, one takes the system to consist of a number of qubits, and the operations are constrained by the fact that they involve only a few qubits at once. Here, instead, we allow for transformations involving all energy levels at once. The basic disconnect is due to the fact that elementary operations in a computational setting typically do not involve dissipative dynamics, which is, however, the main subject of discussion here (for a unitary, it is not natural to act on all energy levels at once, but for a thermalization process with a collective bath, it can be). Hence, it is not obvious to what extent one can reconcile the two approaches. In this work, we adopt ours as working definitions, which carry their own physical intuitions and have the advantage of allowing a clear comparison with the classical work in Ref. [10]. Nonetheless, we believe an approach that more directly captures computational restrictions in the simulation of stochastic processing using both elementary gates and dissipative interactions would be extremely interesting.

Moving on, the central question in the classical setting is to find  $C_{\text{space}}(P)$  and then characterize  $C_{\text{time}}(P, m)$  for  $m \geq C_{\text{space}}(P)$ . The main result of Ref. [10] was to solve

this problem for stochastic matrices  $P$  that are  $\{0, 1\}$  valued or, in other words, represent a function  $f$  over the set of states  $\{1, \dots, d\}$ . How do these results compare with what can be done quantum mechanically? In the next section, we give a protocol realizing every  $\{0, 1\}$ -valued stochastic matrix that scales much better than the corresponding minimal classical cost.

### C. Quantum advantage

Let  $P_f$  be a  $\{0, 1\}$ -valued  $d \times d$  stochastic matrix defined by a function  $f: \mathbb{Z}_d \rightarrow \mathbb{Z}_d$ . Let  $\text{fix}(f)$  be the number of fixed points of  $f$ ,  $|\text{img}(f)|$  the dimension of the image of  $f$ , and  $c(f)$  the number of cycles of  $f$ , i.e., the number of distinct orbits of elements of  $\{1, \dots, d\}$  of the form  $\{i, f(i), f(f(i)), \dots, i\}$ . Recently, the following result has been shown.

**Theorem 3** (Classical cost of a function [10]). The time cost of a  $\{0, 1\}$ -valued stochastic matrix  $P_f$  described by a function  $f$  is given by

$$\begin{aligned} C_{\text{time}}(P_f, m) &= \left\lceil \frac{m+d+\max[c(f)-m, 0]-\text{fix}(f)}{m+d-|\text{img}(f)|} \right\rceil + b_f(m) \\ &\geq \left\lceil \frac{m+d-\text{fix}(f)}{m+d-|\text{img}(f)|} \right\rceil, \end{aligned} \quad (25)$$

where  $b_f(m) = 0$  or  $1$  and  $\lceil \cdot \rceil$  is the ceiling function.

Suppose that the state space is given by all bit strings of length  $s$  so that  $d = 2^s$ . Theorem 3 shows that, if  $|\text{img}(f)|$  is  $O(d)$ , then  $P_f$  is expensive to simulate by memoryless dynamics unless the number of fixed points is also  $O(d)$ . Since for a typical  $f$  we have  $|\text{img}(f)| = O(d)$  and  $\text{fix}(f) = O(1)$  (see the Appendix A), we conclude that, typically,  $C_{\text{time}}(P_f, m) = O(2^s/m)$ ; i.e., an exponential number of memory states are required to have an efficient simulation in the number of time steps. Conversely, one needs an exponential number of time steps to have an efficient simulation for a fixed number of memory states. One of the examples discussed in Ref. [10] is that of  $f_1(i) = i \oplus 1$  (addition modulo  $d$ ), which may be interpreted as keeping track of a clock in a digital computer. From Theorem 3, we see that one has  $C_{\text{time}}(P_{f_1}, m) \geq 2^s/m$ , with  $m$  the number of memory states introduced (see Fig. 5). However, as we discussed already above, any permutation is quantum embeddable by a unitary, and hence  $Q_{\text{time}}(P_{f_1}, 0) = 1$ . The existence of this advantage is generalized by the following result.

**Theorem 4** (Quantum cost of a function). For any  $m \geq 0$  and any function  $f$ , we have  $Q_{\text{time}}(P_f, m) \leq 2$ .

The explicit proof is given in Appendix B, but it is based on the simple fact that every function can be realized quantumly by a unitary process realizing a permutation followed by a classical master equation achieving an idempotent function  $f_I$  in a single time step. Hence, one can achieve every function quantumly using zero memory

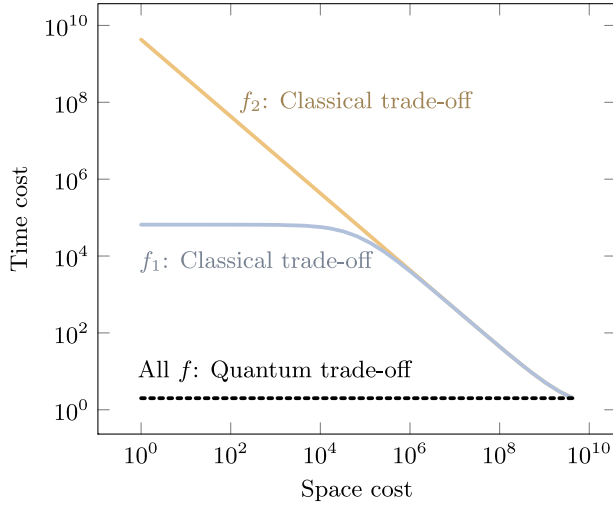


FIG. 5. Classical versus quantum space-time trade-off. The optimal trade-off between space cost and time cost of implementing stochastic matrices for a system of  $s = 32$  bits, i.e., with dimension  $d = 2^{32}$  (plotted in log-log scale). Solid colored curves correspond to optimal trade-offs for classically implementing exemplary  $\{0, 1\}$ -valued stochastic matrices described by functions  $f_1(i) = i \oplus 1$  (addition modulo  $d$ ) and  $f_2(i) = \min\{i + 2^{s/2}, 2^s - 1\}$ , as analyzed in Ref. [10]. The dashed black curve corresponds to optimal trade-offs for quantumly implementing any  $\{0, 1\}$ -valued stochastic matrix, thus illustrating a quantum advantage.

states and only two time steps, compared to the typical classical cost  $C_{\text{time}}(P_{f_1}, m) \geq d/m$ . This result, illustrated in Fig. 5, is quantitative evidence of the power of superposition to act as effective memory.

## IV. ROLE OF MEMORY IN STATE TRANSFORMATIONS

### A. Accessibility regions

In this section, we change the focus from processes to states. We investigate whether a given state transformation can be realized by either a classical or a quantum master equation. In other words, given an input distribution  $\mathbf{p}$ , is it possible to get a given final state  $\mathbf{q}$  through an embeddable (or quantum-embeddable) stochastic matrix  $P$ ? Of course, given full control, it is always possible to choose a master equation with  $\mathbf{q}$  being the unique fixed point. However, more realistically, a fixed point of the evolution is constrained rather than being arbitrary—and typically corresponds to the thermal Gibbs distribution  $\boldsymbol{\gamma}$ , with

$$\gamma_k := \frac{1}{Z} e^{-\beta E_k}, \quad Z := \sum_{k=1}^d e^{-\beta E_k}. \quad (26)$$

Here,  $E_k$  are the energy levels of the system interacting with an external environment at inverse temperature  $\beta$ . Hence, suppose some (full-rank) fixed point  $\boldsymbol{\gamma}$  is given (which may

or may not be the thermal state of the system). We then introduce the following two definitions and discuss the corresponding examples involving standard dynamical models. The former encapsulates the set of input-output relations achievable by means of general stochastic processes with a fixed point  $\boldsymbol{\gamma}$ , while the latter captures the subset achievable without exploiting memory effects, i.e., by Markovian master equations.

**Definition 7** (Classical accessibility). A distribution  $\mathbf{q}$  is accessible from  $\mathbf{p}$  by a classical stochastic process with a fixed point  $\boldsymbol{\gamma}$  if there exists a stochastic matrix  $P$ , such that

$$P\mathbf{p} = \mathbf{q} \quad \text{and} \quad P\boldsymbol{\gamma} = \boldsymbol{\gamma}. \quad (27)$$

We denote the set of all  $\mathbf{q}$  accessible from  $\mathbf{p}$  given  $\boldsymbol{\gamma}$  by  $\mathcal{C}_{\boldsymbol{\gamma}}^{\text{Mem}}(\mathbf{p})$ .

**Example 3.** A standard effective model in cavity QED and atomic physics is the Jaynes-Cummings model in the rotating wave approximation [36]. Formally, it describes the resonant interaction of a two-level system (with energy levels  $|0\rangle$  and  $|1\rangle$ ), with a single harmonic oscillator by means of the Hamiltonian

$$H_{JC} = \frac{\hbar\omega}{2} \sigma_z + \hbar\omega a^\dagger a + g(t)(\sigma^+ a + \sigma^- a^\dagger), \quad (28)$$

where  $a^\dagger$  and  $a$  are the creation and annihilation operators of the oscillator, and  $\sigma_z = |0\rangle\langle 0| - |1\rangle\langle 1|$ ,  $\sigma^+ = |1\rangle\langle 0|$ ,  $\sigma^- = |0\rangle\langle 1|$ . Suppose the oscillator is initially in a thermal state and the system is in a general state  $\rho(0)$ , and denote by  $\mathbf{p}(t)$  the populations in the energy eigenbasis of the system at time  $t$ . Then, in the reduced dynamics of the system generated by  $H_{JC}$ , the population and coherence terms decouple. Moreover, one can show that  $\mathbf{p}(t) = P_t \mathbf{p}(0)$ , where  $P_t \boldsymbol{\gamma} = \boldsymbol{\gamma}$  and  $\boldsymbol{\gamma}$  is the thermal distribution of the two-level system. Hence,  $\mathbf{p}(t) \in \mathcal{C}_{\boldsymbol{\gamma}}^{\text{Mem}}(\mathbf{p}(0))$ ; i.e., if we know  $\mathcal{C}_{\boldsymbol{\gamma}}^{\text{Mem}}(\mathbf{p}(0))$ , we can constrain the set of achievable final states. In fact, in the low-temperature regime,  $\mathcal{C}_{\boldsymbol{\gamma}}^{\text{Mem}}(\mathbf{p}(0))$  is a good approximation of the set of all states that can be achieved in the Jaynes-Cummings model after a long-enough dynamics [37].

In what follows, we denote by  $\rho_{\mathbf{p}}$  the density matrix diagonal in the computational (or energy) basis with entries given by the probability distribution  $\mathbf{p}$ :  $\rho_{\mathbf{p}} = \sum_k p_k |k\rangle\langle k|$ .

**Example 4.** Consider a system in a classical state  $\rho_{\mathbf{p}}$ . Then,  $\mathcal{C}_{\boldsymbol{\gamma}}^{\text{Mem}}(\mathbf{p})$  describes the set of classical states that can be obtained from  $\rho_{\mathbf{p}}$  by thermal operations [38,39], i.e., energy-preserving couplings of the system with arbitrary thermal baths at a temperature fixed by the choice of  $\boldsymbol{\gamma}$ . This scenario was analyzed in Refs. [38,40], and it was proven there that  $\mathcal{C}_{\boldsymbol{\gamma}}^{\text{Mem}}(\mathbf{p})$  is fully specified by the notion of thermomajorization (also known as *majorization relative to  $\boldsymbol{\gamma}$*  [41]).

The memoryless version of the above classical accessibility region is defined as follows.

**Definition 8** (Classical memoryless accessibility). A distribution  $\mathbf{q}$  is accessible from  $\mathbf{p}$  by a classical master equation with a fixed point  $\boldsymbol{\gamma}$  if there exists a continuous one-parameter family  $L(t)$  of rate matrices generating a family of stochastic matrices  $P(t)$ , such that

$$P(t_f)\mathbf{p} = \mathbf{q}, \quad L(t)\boldsymbol{\gamma} = \mathbf{0} \quad \text{for all } t \in [0, t_f]. \quad (29)$$

We denote the set of all  $\mathbf{q}$  accessible from  $\mathbf{p}$  given  $\boldsymbol{\gamma}$  by  $\mathcal{C}_\gamma(\mathbf{p})$ .

**Example 5.** Consider an  $N$ -level quantum system weakly interacting with a large thermal bath. A microscopic derivation [42] leads to a quantum Markovian master equation for the system known as a Davies process [43], which is a standard model for thermalization [44]. One can then verify that the populations  $\mathbf{p}(t)$  of the system undergoing a Davies process satisfy a classical Markovian master equation,

$$\frac{d\mathbf{p}(t)}{dt} = L\mathbf{p}(t), \quad L\boldsymbol{\gamma} = 0, \quad (30)$$

and thus  $\mathbf{p}(t) \in \mathcal{C}_\gamma(\mathbf{p}(0))$ . A characterization of  $\mathcal{C}_\gamma(\mathbf{p})$  then allows one to restrict the intermediate nonequilibrium states generated by standard thermalization processes, without the need to solve the actual dynamics. Since these processes are building blocks of more complex protocols—they form thermalization strokes of engines and refrigerators [45]—this information can be employed to identify thermodynamic protocols with optimal performance.

These definitions naturally generalize to quantum dynamics in the following way.

**Definition 9** (Quantum accessibility). A distribution  $\mathbf{q}$  is accessible from  $\mathbf{p}$  by quantum dynamics with a fixed point  $\boldsymbol{\gamma}$  if there exists a quantum channel  $\mathcal{E}$ , such that

$$\mathcal{E}(\rho_{\mathbf{p}}) = \rho_{\mathbf{q}} \quad \text{and} \quad \mathcal{E}(\rho_{\boldsymbol{\gamma}}) = \rho_{\boldsymbol{\gamma}}. \quad (31)$$

We denote the set of all  $\mathbf{q}$  accessible from  $\mathbf{p}$  given  $\boldsymbol{\gamma}$  by  $\mathcal{Q}_\gamma^{\text{Mem}}(\mathbf{p})$ .

**Example 6.** Setting  $\boldsymbol{\gamma}$  to be the thermal Gibbs state, the set of channels satisfying  $\mathcal{E}(\rho_{\boldsymbol{\gamma}}) = \rho_{\boldsymbol{\gamma}}$  (Gibbs-preserving maps) was identified as the most general set of operations that can be performed without investing work [46]. The reason is that any channel that is not Gibbs-preserving can create a nonequilibrium resource from an equilibrium state. These channels can then be taken as free operations, and the minimal work cost of a general channel  $\mathcal{F}$  can be computed from the minimal work a battery system must provide to simulate  $\mathcal{F}$  using Gibbs-preserving maps only. The set of stochastic matrices  $P$  that can be realized with Gibbs-preserving maps coincides with the set of all  $P$  with a fixed point  $\boldsymbol{\gamma}$  (see, e.g., Theorem 1 of Ref. [39]). It is not surprising that this includes  $P$  that cannot be simulated without memory effects since

Gibbs-preserving quantum channels, in general, cannot be realized by means of a Markovian master equation; i.e., they require memory effects.

It is of course natural to consider the transformations that can be realized by the subset of Gibbs-preserving maps that originate from a Markovian quantum master equation, so we introduce the following.

**Definition 10.** [Quantum memoryless accessibility] A distribution  $\mathbf{q}$  is accessible from  $\mathbf{p}$  by a quantum master equation with a fixed point  $\boldsymbol{\gamma}$  if there exists a continuous one-parameter family of Lindbladians  $\mathcal{L}(t)$  generating a family of quantum channels  $\mathcal{E}(t)$ , such that

$$\mathcal{E}(t_f)[\rho_{\mathbf{p}}] = \rho_{\mathbf{q}}, \quad \mathcal{L}(t)[\rho_{\boldsymbol{\gamma}}] = 0 \quad \text{for all } t \in [0, t_f]. \quad (32)$$

We denote the set of all  $\mathbf{q}$  accessible from  $\mathbf{p}$  given  $\boldsymbol{\gamma}$  by  $\mathcal{Q}_\gamma(\mathbf{p})$ .

**Example 7.** Standard thermalization processes resulting from weak couplings to a large environment, such as Davies maps, are generated by a Lindbladian  $\mathcal{L}$  satisfying  $\mathcal{L}[\rho_{\boldsymbol{\gamma}}] = 0$ , as required by the definition above. However, these dynamics are unable to create quantum superpositions and hence cannot be used to show a quantum advantage. On the other hand, more exotic thermalization processes exist. Let  $|\boldsymbol{\gamma}\rangle := \sqrt{\gamma_0}|0\rangle + \sqrt{\gamma_1}|1\rangle$ ,  $\boldsymbol{\gamma} = \gamma_0|0\rangle\langle 0| + \gamma_1|1\rangle\langle 1|$  and consider the quantum master equation on a two-level system

$$\frac{d}{dt}\rho = \mathcal{L}(\rho), \quad (33)$$

where  $\mathcal{L}$  is the Lindbladian specified by Eq. (7) with a vanishing Hamiltonian  $H$  and the map  $\Phi$  given by a measure-and-prepare channel of the following form:

$$\Phi(\rho) = \frac{\rho_{00}}{\gamma_0}(\boldsymbol{\gamma} - \gamma_1|\boldsymbol{\gamma}\rangle\langle\boldsymbol{\gamma}|) + \rho_{11}|\boldsymbol{\gamma}\rangle\langle\boldsymbol{\gamma}|. \quad (34)$$

It satisfies  $\mathcal{L}(\rho_{\boldsymbol{\gamma}}) = 0$ , so if  $\mathbf{p}(t)$  is the population vector of  $\rho(t)$ , one has  $\mathbf{p}(t) \in \mathcal{Q}_\gamma(\mathbf{p}(0))$ . One can verify that the dynamics equilibrates every state to  $\boldsymbol{\gamma}$  and yet it is capable of generating coherence [for example, for  $\rho(0) = |1\rangle\langle 1|$ , one has  $d\rho_{01}/dt > 0$  around  $t = 0$ ]. This particular dynamics is unable to translate the ability to create coherence into the ability to generate a  $P$  that classically requires memory (we will later construct dynamics that do). What the example illustrates, however, is a central mechanism by which a quantum advantage can arise, i.e., generation of quantum superpositions by exotic, thermalizing, Markovian master equations. These dynamics are, in principle, allowed by quantum mechanics, but we leave the question of how they can actually be realized for future work. We note, in passing, that what is needed is a physical model where a Markovian process with a thermal fixed point naturally emerges despite the fact that the dynamics does not satisfy the secular approximation.

Note that in Definitions 8 and 10, the requirements  $L(t)\boldsymbol{\gamma} = \mathbf{0}$  and  $\mathcal{L}(t)[\rho_\gamma] = 0$  for all times ensure that  $P(t)\boldsymbol{\gamma} = \boldsymbol{\gamma}$  and  $\mathcal{E}(t)[\rho_\gamma] = \rho_\gamma$  for all intermediate times  $t \in [0, t_f]$ . By characterizing the difference between the sets  $\mathcal{C}_\gamma^{\text{Mem}}(\boldsymbol{p})$  and  $\mathcal{C}_\gamma(\boldsymbol{p})$ , one can capture the state transformations that can be achieved only through controls exploiting memory effects. In other words, all states  $\boldsymbol{q} \in \mathcal{C}_\gamma^{\text{Mem}}(\boldsymbol{p})$ , but not in  $\mathcal{C}_\gamma(\boldsymbol{p})$ , can only be achieved from  $\boldsymbol{p}$  via a transformation that employs memory. Analogous statements hold for  $\mathcal{Q}_\gamma^{\text{Mem}}(\boldsymbol{p})$  and  $\mathcal{Q}_\gamma(\boldsymbol{p})$ .

In what follows, we study relations between the accessibility regions. Our main result is that  $\mathcal{C}_\gamma(\boldsymbol{p}) \subset \mathcal{Q}_\gamma(\boldsymbol{p})$  in the case of a uniform fixed point  $\boldsymbol{\eta} := (1/d, \dots, 1/d)$  and for *general* fixed points for a qubit system. These results signal a quantum advantage; i.e., some transitions that classically require memory can be achieved through memoryless quantum dynamics.

### B. Quantum advantage at infinite temperature

A natural question that arises is whether the sets  $\mathcal{Q}_\gamma(\boldsymbol{p})$  and  $\mathcal{Q}_\gamma^{\text{Mem}}(\boldsymbol{p})$  are larger than their classical counterparts. This enlargement of the set of achievable states is another facet of the quantum advantage. More generally, one could also investigate more refined versions of memory advantages on state transformations, e.g., trying to include space and time resources in the analysis, similarly to what we did in Sec. III.

It is straightforward to prove that without the memoryless constraint, there will be no quantum advantage. In other words, we have  $\boldsymbol{q} \in \mathcal{Q}_\gamma^{\text{Mem}}(\boldsymbol{p})$  if and only if  $\boldsymbol{q} \in \mathcal{C}_\gamma^{\text{Mem}}(\boldsymbol{p})$ . The “if” part is obvious, as the set of all quantum channels with a fixed point  $\rho_\gamma$  contains as a subset the set of classical stochastic processes with the same fixed point. Conversely, take any  $\boldsymbol{q} \in \mathcal{Q}_\gamma^{\text{Mem}}(\boldsymbol{p})$ , meaning that there exists a channel  $\mathcal{E}$  such that  $\mathcal{E}(\rho_p) = \rho_q$  and  $\mathcal{E}(\rho_\gamma) = \rho_\gamma$ . Then, we can construct a stochastic process  $P$  with matrix elements  $P_{ij}$  given by  $\langle i | \mathcal{E}(|j\rangle\langle j|) | i \rangle$ . The matrix  $P$  is stochastic because  $\mathcal{E}$  is positive and trace-preserving. Furthermore, it satisfies  $P\boldsymbol{p} = \boldsymbol{q}$  and  $P\boldsymbol{\gamma} = \boldsymbol{\gamma}$ . Therefore,  $\boldsymbol{q} \in \mathcal{C}_\gamma^{\text{Mem}}(\boldsymbol{p})$ . To sum up,

$$\mathcal{Q}_\gamma^{\text{Mem}}(\boldsymbol{p}) = \mathcal{C}_\gamma^{\text{Mem}}(\boldsymbol{p}). \quad (35)$$

However, as we now prove, a quantum advantage is exhibited by  $\mathcal{Q}_\gamma(\boldsymbol{p}) \not\subset \mathcal{C}_\gamma(\boldsymbol{p})$ ; i.e., there are states classically accessible only with memory that can be achieved by quantum memoryless dynamics. In the case of a uniform fixed point, going from classical to quantum memoryless dynamics allows one to achieve the maximal quantum advantage: All transformations involving memory can be realized quantum mechanically with no memory.

**Theorem 5** (Maximal quantum advantage for uniform fixed points). For every  $\boldsymbol{p}$  and a uniform distribution  $\boldsymbol{\eta} = (1/d, \dots, 1/d)$ , one has  $\mathcal{Q}_\eta(\boldsymbol{p}) = \mathcal{C}_\eta^{\text{Mem}}(\boldsymbol{p})$ .

*Proof.*—Clearly, we have  $\mathcal{Q}_\eta(\boldsymbol{p}) \subseteq \mathcal{C}_\eta^{\text{Mem}}(\boldsymbol{p})$  because  $\mathcal{Q}_\eta(\boldsymbol{p}) \subseteq \mathcal{Q}_\eta^{\text{Mem}}(\boldsymbol{p}) = \mathcal{C}_\eta^{\text{Mem}}(\boldsymbol{p})$ . Conversely, let us show  $\mathcal{C}_\eta^{\text{Mem}}(\boldsymbol{p}) \subseteq \mathcal{Q}_\eta(\boldsymbol{p})$ . Take  $\boldsymbol{q} \in \mathcal{C}_\eta^{\text{Mem}}(\boldsymbol{p})$ . As it is well known (see Theorem B6 of Ref. [41]), this is equivalent to the majorization relation  $\boldsymbol{p} \succ \boldsymbol{q}$ , i.e.,

$$\sum_{i=1}^k p_i^\downarrow \geq \sum_{i=1}^k q_i^\downarrow, \quad k = 1, \dots, d, \quad (36)$$

where  $\boldsymbol{p}^\downarrow$  denotes the probability distribution  $\boldsymbol{p}$  sorted in a nonincreasing order. We show that every  $\boldsymbol{q}$  satisfying the above can be achieved from  $\boldsymbol{p}$  by a composition of two quantum-embeddable processes (so, according to Lemma 1, by one quantum-embeddable process), each with a uniform fixed point. First, note that every permutation is quantum embeddable, as discussed in Sec. II C. Thus, one can rearrange  $\boldsymbol{p}$  into  $\boldsymbol{p}'$  with  $\boldsymbol{p}' \succ \boldsymbol{p}$ , sorted in the same way as  $\boldsymbol{q}$ . By transitivity of majorization, we have  $\boldsymbol{p}' \succ \boldsymbol{q}$ . Now, using Theorem 11 of Ref. [47] with  $\beta \rightarrow \infty$ , it follows that  $\boldsymbol{q}$  can be achieved from  $\boldsymbol{p}'$  by applying a sequence of stochastic processes of the form

$$T^{(i,j)} := \begin{bmatrix} 1 - \lambda/2 & \lambda/2 \\ \lambda/2 & 1 - \lambda/2 \end{bmatrix} \oplus I_{\setminus(i,j)}, \quad (37)$$

where  $\lambda \in [0, 1]$  and  $I_{\setminus(i,j)}$  is the identity matrix on all states excluding  $i$  and  $j$ . The matrices  $T^{(i,j)}$  are embeddable with rate matrices

$$L^{(i,j)} = \begin{bmatrix} -1/2 & 1/2 \\ 1/2 & -1/2 \end{bmatrix} \oplus \mathbf{0}_{\setminus(i,j)}, \quad (38)$$

where  $\mathbf{0}_{\setminus(i,j)}$  is the zero matrix on all states excluding  $i$  and  $j$ . Moreover, matrices  $L^{(i,j)}$  satisfy  $L^{(i,j)}\boldsymbol{\eta} = \mathbf{0}$ . Putting everything together, we have a unitary permutation followed by a sequence of processes  $e^{L^{(i,j)}t_k}$  that map  $\boldsymbol{p}$  into  $\boldsymbol{q}$ , so  $\boldsymbol{q} \in \mathcal{Q}_\eta(\boldsymbol{p})$ . We conclude that  $\mathcal{Q}_\eta(\boldsymbol{p}) = \mathcal{C}_\eta^{\text{Mem}}(\boldsymbol{p})$ . ■

Since  $\mathcal{Q}_\eta(\boldsymbol{p}) = \mathcal{C}_\eta^{\text{Mem}}(\boldsymbol{p})$ , in order to prove a quantum advantage, we only need to show that classical memoryless dynamics is more restrictive than general classical dynamics with a fixed point, i.e., that  $\mathcal{C}_\eta(\boldsymbol{p})$  is a proper subset of  $\mathcal{C}_\eta^{\text{Mem}}(\boldsymbol{p})$ . This is indeed the case, as can be easily verified for  $d = 2$  (and it is rigorously proven in Ref. [14] for a general fixed point  $\boldsymbol{\gamma}$  in any dimension). Here, in Fig. 6, we illustrate that  $\mathcal{C}_\eta(\boldsymbol{p})$  is a proper subset of  $\mathcal{C}_\eta^{\text{Mem}}(\boldsymbol{p})$  for an exemplary case of  $d = 3$ .

### C. Quantum advantage at any finite temperature (qubit case)

Now, we generalize the considerations above to arbitrary fixed points or, if we think of  $\boldsymbol{\gamma}$  as a thermal fixed point,

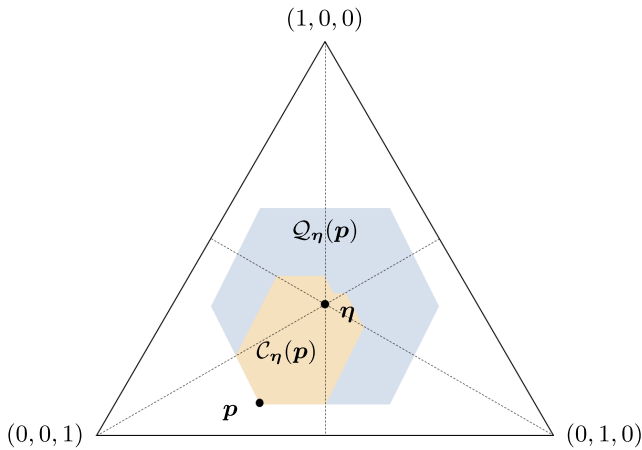


FIG. 6. Quantum advantage at infinite temperature for  $d = 3$ . We show the sets of states accessible via classical [ $\mathcal{C}_\eta(\mathbf{p})$ , smaller orange shape] and quantum [ $\mathcal{Q}_\eta(\mathbf{p})$ , larger blue hexagon] memoryless dynamics with a uniform fixed point  $\eta$  for a system of dimension  $d = 3$  and an exemplary initial state  $\mathbf{p}$  (each point inside a triangle corresponds to a probabilistic mixture of sharp distributions). The set  $\mathcal{C}_\eta(\mathbf{p})$  was constructed using the results of Ref. [14].

arbitrary temperatures. In other words, we compare states achievable by classical memoryless processes with a fixed point  $\gamma$  to states achievable by quantum memoryless processes with a fixed point  $\rho_\gamma$ . Here, we solve the case for  $d = 2$ . To ease physical intuition, we parametrize  $\gamma$  as in Eq. (26), with  $E = E_2 - E_1$  the energy gap between the two states and  $\beta$  the inverse temperature of the external environment.

Let us start by recalling the classical solution to this problem. First,  $\mathcal{C}_\eta^{\text{Mem}}(\mathbf{p})$  can be obtained using the thermo-majorization condition [38,40]. If  $\mathbf{p} = (p, 1 - p)$ , in terms of the achievable ground-state populations, we have

$$\mathcal{C}_\gamma^{\text{Mem}}(\mathbf{p}) = \begin{cases} [p, 1 - e^{-\beta E} p] & \text{if } p \leq 1/Z \\ [1 - e^{-\beta E} p, p] & \text{if } p > 1/Z. \end{cases} \quad (39)$$

This set of states can be approximately achieved by the Jaynes-Cummings interaction of Example 3, which involves memory effects. The need for memory effects to achieve the states in Eq. (39) can be understood more generally by looking at information backflows, which are a standard signature of non-Markovian effects. As one can verify from the previous expressions, any continuous trajectory  $\mathbf{r}(t)$  connecting  $\mathbf{p}$  to the other extremal point in  $\mathcal{C}_\gamma^{\text{Mem}}(\mathbf{p})$  has to pass through  $\gamma$ . Thus, the (non-equilibrium) free energy of the system,

$$F(\mathbf{p}) = \sum_i p_i E_i - \frac{1}{\beta} H(\mathbf{p}), \quad H(\mathbf{p}) = -\sum_i p_i \log p_i, \quad (40)$$

will first decrease all the way to its minimal value and then increase again, signaling an information backflow from the thermal environment. It is obvious that such a phenomenon

does not occur when dissipation is well described by a Markovian master equation since, in this case, the dynamics cannot cross the fixed state  $\gamma$ . Memory effects (e.g., in the form of non-negligible system-bath correlations) are required to access all the states in  $\mathcal{C}_\gamma^{\text{Mem}}(\mathbf{p})$ .

Let us now look at  $\mathcal{C}_\gamma(\mathbf{p})$ , the achievable states with classical memoryless processes with a thermal fixed point. For  $d = 2$ , the set of Markovian master equations with a fixed point  $\gamma$  is limited to the thermalizations introduced in Example 1. Hence,  $\mathcal{C}_\gamma(\mathbf{p})$  is readily characterized as all states along a line connecting  $\mathbf{p}$  and the thermal state:

$$\mathcal{C}_\gamma(\mathbf{p}) = \begin{cases} [p, 1/Z] & \text{if } p \leq 1/Z \\ [1/Z, p] & \text{if } p > 1/Z. \end{cases} \quad (41)$$

The difference between Eqs. (39) and (41) is simply that memory effects allow one to go ‘‘on the other side’’ of the thermal state. Without memory, this is not possible since the thermal state is at a minimum of the free energy.

Now, we turn to the corresponding quantum mechanical problem, looking for potential advantages. Any unitary that changes the population of the state is forbidden since it does not have  $\rho_\gamma$  as a fixed point. Hence, we cannot rely on any of the previous constructions exploiting the fact that, quantum mechanically, one can generate permutations without using memory. We then need to characterize the set of diagonal quantum states achievable from a given state  $\rho_p$  via Markovian quantum master equations with a given fixed point  $\rho_\gamma$ . Even without the constraint that the channel is generated by a master equation, finding a simple characterization of the set of accessible states for  $d > 2$  has remained an open problem for decades [48]. This issue explains why we are focusing here on the simplest non-trivial case of a qubit system, where such a problem has been fully solved [48–50]. We numerically show that, in  $d = 2$ , one also achieves a maximal quantum advantage. Specifically, all states achievable classically by means of processes with thermal fixed points that involve memory, Eq. (39), can be attained by a Markovian quantum master equation with fixed point  $\rho_\gamma$ . More compactly, we have the following result:

**Result 1** (Quantum advantage at every finite temperature—numerics). For  $d = 2$ ,  $\mathcal{Q}_\gamma(\mathbf{p}) = \mathcal{C}_\gamma^{\text{Mem}}(\mathbf{p})$ .

This result showcases that the advantage of Theorem 5 is not limited to the special case of a uniform fixed point involving unitary dynamics. Superposition can substitute memory in the control of classical systems at every finite temperature.

In order to prove Result 1, we present an explicit construction and numerical evidence for an even stronger result.

**Result 2.** [Numerics] Every qubit state accessible via a qubit channel with a given fixed point can be achieved by a qubit Markovian master equation with the same fixed point.

Let us start by recalling the result of Ref. [48], where the authors provided necessary and sufficient conditions for the existence of a qubit channel  $\mathcal{E}$  satisfying

$$\mathcal{E}(\rho) = \rho', \quad \mathcal{E}(\sigma) = \sigma', \quad (42)$$

for any two pairs of qubit density matrices  $(\rho, \rho')$  and  $(\sigma, \sigma')$ . Moreover, whenever such a channel exists, the authors provided a construction of the Kraus operators of  $\mathcal{E}$ . Setting  $\sigma = \sigma' = \rho_\gamma$ , one obtains a characterization of all states accessible from  $\rho$  through arbitrary channels with a given fixed point  $\rho_\gamma$  (we choose a basis in which the fixed point is diagonal). In Ref. [50], the continuous set of conditions presented in Ref. [48] was reduced to just two inequalities:

$$R_\pm(\rho) \geq R_\pm(\rho'). \quad (43)$$

These inequalities are best understood through the standard Bloch sphere parametrization of the states involved. Recall that a general qubit state can be written as

$$\rho = \frac{\mathbb{1} + \mathbf{r}_\rho \cdot \boldsymbol{\sigma}}{2}, \quad (44)$$

where  $\boldsymbol{\sigma}$  denotes the vector of Pauli matrices  $(\sigma_x, \sigma_y, \sigma_z)$ , while  $\mathbf{r}_\rho$  is a three-dimensional real vector that uniquely represents  $\rho$  as a point inside a unit Bloch ball in  $\mathbb{R}^3$ . We parametrize the initial, final, and fixed points as follows:

$$\mathbf{r}_\rho = (x, y, z), \quad \mathbf{r}_{\rho'} = (x', y', z'), \quad \mathbf{r}_\gamma = (0, 0, \zeta). \quad (45)$$

Unitary rotations about the  $z$  axis leave  $\rho_\gamma$  unchanged. By performing such rotations before and after the channel  $\mathcal{E}$ , without loss of generality, we can set  $x \geq 0$ ,  $x' \geq 0$ , and  $y = y' = 0$ . The monotonies  $R_\pm$  from Eq. (43) are then defined as [50]

$$R_\pm(\rho) = \delta(\rho) \pm \zeta z, \quad (46)$$

where

$$\delta(\rho) := \sqrt{(z - \zeta)^2 + x^2(1 - \zeta^2)}, \quad (47)$$

with analogous (primed) definitions for  $\rho'$ . The two inequalities from Eq. (43) can then be used to find extremal states accessible from  $\rho$  via qubit channels with fixed point  $\rho_\gamma$ . As shown in Fig. 7, these are given by

- (i) states with a constant  $R_+$  lying on a circle with centre and radius  $(\mathbf{c}_0, R_0)$  if  $z' \geq z$ , where

$$R_0 = \frac{R_+ - \zeta^2}{1 - \zeta^2}, \quad \mathbf{c}_0 = [0, 0, \zeta(1 - R_0)], \quad (48)$$

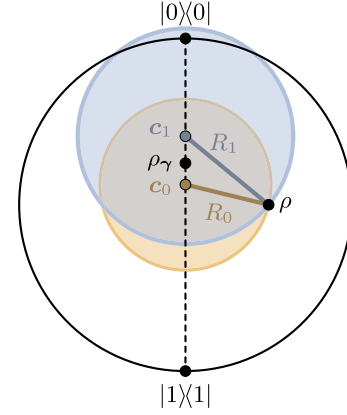


FIG. 7. Qubit accessibility region. Geometrically, states with a fixed value of  $R_+$  lie on a circle centered at  $\mathbf{c}_0$  and with radius  $R_0$  (in orange). Similarly, states with a fixed value of  $R_-$  lie on a circle centered at  $\mathbf{c}_1$  and with radius  $R_1$  (in blue). States achievable from a given initial state  $\rho$  via quantum channels with a fixed point  $\rho_\gamma$  lie inside the Bloch sphere in the intersection of two balls  $(\mathbf{c}_0, R_0)$  and  $(\mathbf{c}_1, R_1)$ . Here, the parameters for initial and fixed states are chosen to be  $x = 1/2$ ,  $z = 0$ , and  $\zeta = 1/4$ .

- (ii) states with a constant  $R_-$  lying on a circle  $(\mathbf{c}_1, R_1)$  if  $z' < z$ , where

$$R_1 = \frac{R_- + \zeta^2}{1 - \zeta^2}, \quad \mathbf{c}_1 = [0, 0, \zeta(1 + R_1)]. \quad (49)$$

The crucial observation we make here is as follows. Consider the case  $z' \geq z$ . Divide the extremal path into  $n$  parts by choosing states  $\rho_0, \dots, \rho_n$  along the  $(\mathbf{c}_0, R_0)$  circle with  $\rho_0 = \rho$ . Since Eq. (43) is satisfied, for each  $i \in \{0, \dots, n-1\}$ , there exists  $\mathcal{E}_i$  with  $\mathcal{E}_i(\rho_i) = \rho_{i+1}$  and  $\mathcal{E}_i(\rho_\gamma) = \rho_\gamma$ . Similar considerations hold for  $z' < z$  considering the  $(\mathbf{c}_1, R_1)$  circle, which suggests that there indeed exists a continuous Markov evolution that evolves the state along the extremal path.

To construct a time-dependent Lindbladian that evolves the state along the extremal path (say, the one with  $z' \geq z$ ), we fix some arbitrarily small  $\Delta > 0$  and find the state  $\rho_1$  on the extremal path with  $z' = z + \Delta$ . Using the construction of Ref. [48], we obtain an explicit form for the quantum channel  $\mathcal{E}_0$  mapping  $\rho_0$  to  $\rho_1$  while preserving  $\rho_\gamma$ . Next, we define the Lindbladian  $\mathcal{L}_0 = \mathcal{E}_0 - \mathcal{I}$  and evolve the state according to  $e^{\mathcal{L}_0}$ , obtaining  $\tilde{\rho}_1 := e^{\mathcal{L}_0} \rho_0$ . We then repeat the same procedure, but instead of  $\rho_0$ , we start with  $\tilde{\rho}_i$  for  $i > 0$ . In this way, we construct a whole set of Lindbladians  $\mathcal{L}_i$ . The procedure ends when Eq. (43) is no longer satisfied for  $z' = z + \Delta$ . Because of the extremely complicated form of the Kraus operators describing the channels  $\mathcal{E}_i$  (and hence  $\mathcal{L}_i$ ), instead of their explicit expressions, we provide their construction in Appendix C.

We have thus constructed a quantum Markovian evolution  $\prod_i e^{\mathcal{L}_i}$  passing through the points  $\tilde{\rho}_i$ . Numerical investigations show that this Markovian dynamics evolves

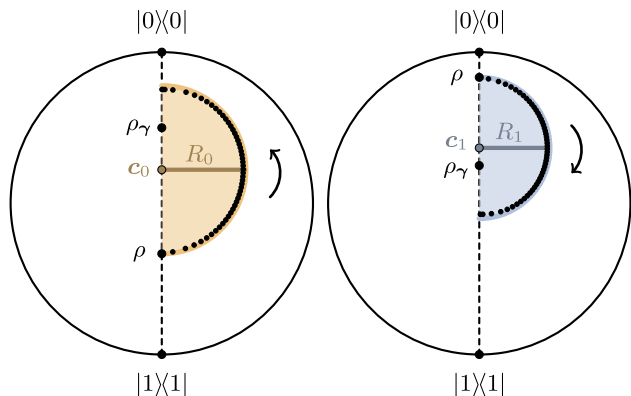


FIG. 8. Qubit memoryless accessibility region. We show the initial state  $\rho$  evolved by  $\prod_i e^{\mathcal{L}_i}$  (black dots) as described in the main text. The evolved states clearly approach the extremal paths for thermodynamic processes with memory given by circles  $(c_0, R_0)$  and  $(c_1, R_1)$ . In both panels,  $\Delta$  is chosen such that the evolution is divided into 100 equal steps (to increase readability, only the even steps are plotted). (a) Parameters:  $x = 0$ ,  $z = -1/3$  and  $\zeta = 1/2$ . (b) Parameters:  $x = 0$ ,  $z = 5/6$ ,  $\zeta = 1/4$ .

$\rho_0 = \rho$  approximately along the extremal circle  $(c_0, R_0)$  [or  $(c_1, R_1)$  for  $\Delta < 0$ ], with the approximation improving as  $\Delta \rightarrow 0$ . We illustrate these results for particular choices of initial and fixed states in Fig. 8 and note that this is strong evidence that Result 2 holds.

## V. POTENTIAL FOR PRACTICAL ADVANTAGES

While the quantum advantages described in this paper may ultimately find practical applications in information processing technologies, we do not think they can be immediately applied to improve the performance of modern-day classical information processors. Quantum advantages proven here arise only once we get to the fundamental limit of encoding classical bits in quantized levels of memory systems. As mentioned in Sec. IB, whenever a bit is encoded in macroscopic degrees of freedom (that actually consist of  $d \gg 2$  distinct quantum levels), one can employ a subset of them to act as an internal memory and, e.g., unlock the possibility of a classical Markovian bit swap. Moreover, from a practical perspective, the cost of coherent control over all degrees of freedom of a memory system is massive and, so far, seems to greatly exceed the benefits we have proven here. In some sense, the situation resembles the one with the famous Landauer’s principle [51]: While switching from irreversible to reversible computations, one can go below the fundamental  $k_B T \log 2$  limit of dissipated heat per bit of erased information by performing all logical gates reversibly [52], the technological cost related to such a change still exceeds possible gains [53]. Nevertheless, if in the future we have coherent control over quantum systems as good as incoherent control over classical systems, we will need to use fewer memory states to implement certain processes.

Despite the above outlook, we still believe that our results may find near-term applications on the side of control theory and quantum thermodynamics. As we discussed, for  $d = 2$ , any classical Markovian master equation with a fixed point  $\gamma$  evolves  $\mathbf{p}$  along the path  $\mathbf{p}(t)$  that can never go “on the other side of the fixed point” (recall Fig. 3); memory is required for that to happen. On the other hand, the corresponding quantum Markovian master equations access all states achievable under general stochastic maps with a fixed point  $\gamma$ . Creation of quantum coherence is crucial since it opens new pathways that “go around” the fixed state. What is surprising is that the creation of coherence in a qubit Markovian dissipative process can replace *all* memory effects. Even the “ $\beta$  swap,” the classical process with a thermal fixed point that requires the largest free-energy back-flow and achieves the farthest accessible state on the other side of  $\gamma$  in Eq. (41), can be approximated arbitrarily well by a quantum Markovian master equation with a thermal fixed point.

This observation potentially has significant consequences in the context of cooling, a central problem in quantum sciences. Access to pure, “cold” quantum systems is a preliminary requirement for quantum computing [54]. One possible way to achieve it is to employ heat-bath algorithmic cooling (HBAC) protocols [55–57], initially developed in the context of NMR systems. These protocols are optimized to achieve the largest possible cooling of a target system by means of a sequence of unitary operations and interactions with a heat bath. Recently, optimal HBAC were derived [15,58]. For a single qubit, these protocols involve performing a sequence of  $X$  pulses and  $\beta$  swaps, leading to an exponential convergence to the ground state in the number of cooling rounds [15]. It was also noted that if the interactions with the bath are restricted to be Markovian and destroy superpositions, then one cannot cool the target system below the environment’s temperature.

In the scheme of Ref. [15], it is the memory effects, encoded in system-bath correlations, that are central in achieving the desired cooling performance. Our Result 1 implies that this control over system-bath correlations can be replaced by the control over the coherent properties of the system. Specifically, a time-dependent Lindbladian with a thermal fixed point is able to generate the required  $\beta$  swap, which implies that qubit systems can be cooled all the way to their ground state without exploiting any memory effects. This is a purely quantum phenomenon, exploiting control over coherent resources. The practical impact of these protocols, of course, depends on the possibility of experimentally realizing these exotic dynamics, which is an important question put forward by this investigation.

Moreover, the phenomenon identified in the previous section for qubit systems may be relevant in a much broader setup, as the following general mechanism illustrates. Suppose, for simplicity, that there are no degeneracies in



the Bohr spectrum of the system; i.e., the allowed energy differences,  $\{E_i - E_j\}_{i \neq j}$ , for the studied system are all distinct. Given the initial classical state of the system,  $\rho(0) = \rho_p$ , and its evolution described by  $\rho(t)$ , use the following decomposition:

$$\rho(t) = \rho_{r(t)} + C(t), \quad (50)$$

where  $r(t)$  is the population in the energy eigenbasis and  $C(t)$  denotes the off-diagonal terms (“coherences”) at time  $t$ . Any classical Markovian evolution with a thermal fixed point requires  $(d/dt)F(r(t)) \leq 0$  and  $C(t) = 0$  at all  $t \geq 0$ . Any quantum Markovian dynamics requires  $(d/dt)F_Q(\rho(t)) \leq 0$  for all  $t \geq 0$ , where  $F_Q$  is the quantum nonequilibrium free energy:

$$F_Q(\rho) = \text{Tr}(\rho H_S) - \beta^{-1} S(\rho), \quad (51)$$

with  $H_S$  denoting the Hamiltonian of the system and  $S(\rho) = -\text{Tr}(\rho \log \rho)$  being the von Neumann entropy. Recall that  $F_Q(\rho)$  can be additively decomposed into two non-negative components [39]:

$$F_Q(\rho(t)) = F(r(t)) + \beta^{-1} A(\rho(t)). \quad (52)$$

The first term is the classical nonequilibrium free energy, and the second one is a quantum component (called “asymmetry”), which measures the coherent contribution to  $F_Q$  [59]. At  $t = 0$ , we have  $r(0) = p$  and  $C(0) = 0$ , which implies  $A(\rho(0)) = 0$ . Hence, both classical and quantum free energies for the initial state are equal to  $F(p)$ . However, a Markovian quantum evolution can store some free energy in coherences at times  $t \geq 0$  since only the *sum* of the classical and quantum components of the free energy must monotonically decrease in time. Thus, at time  $t_*$  when  $r(t_*) = \gamma$ , one is classically stuck in a free-energy minimum  $F(\gamma)$  and cannot proceed further. But quantum mechanically, one can have  $C(t_*) \neq 0$ , and hence  $F_Q(\rho(t_*)) > F(\gamma)$ . Markovian quantum dynamics can thus access other states at  $t > t_*$  by converting back some of the quantum component of free energy into classical free energy. This process allows one to achieve the required backflow into the classical component of the free energy, as illustrated in Fig. 9. Storing free energy in coherence is of course a nontrivial task (it requires the aid of an external source of coherence, and, certainly, it cannot be done with thermal operations [59]), but our qubit construction has shown that it is possible with quantum Markovian master equations.

## VI. OUTLOOK

The central task of this paper was to introduce a unifying framework to prove that quantum dynamics offers memory improvements over the classical stochastic evolution in a variety of settings, from the more computational ones to the

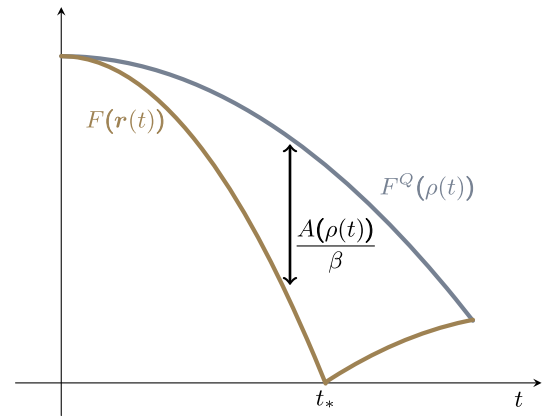


FIG. 9. Free energy stored in coherence. Under a Markovian master equation with a thermal fixed point, the quantum free energy  $F^Q$  is monotonically decreasing in time. Since  $F^Q = F + A/\beta$ , part of the classical component  $F$  can be stored in the coherent component  $A$  at times  $t \leq t_*$  and recovered later. At time  $t = t_*$ , the population is thermal, so the classical free energy is at a minimum; however,  $A \neq 0$ , so  $F^Q$  is above the minimum. For  $t \geq t_*$ , part of this coherent free energy is converted back into classical free energy. Hence, the latter undergoes a backflow, which classically would require memory effects.

more physical ones. The unifying notion is that of the underlying Markovian master equation and the advantages one gains with quantum controls as compared to the classical ones. The driving force behind these advantages is the superposition principle, which provides a wider arena for memoryless evolutions to unfold. While Holevo’s theorem [2] prevents us from retrieving more than  $n$  bits of information from  $n$  qubits, we see that, in many other respects, superpositions can take over the role played by a classical memory.

This result is most clearly captured by the notion of quantum-embeddable stochastic processes introduced here: processes that do not require any memory quantum mechanically, but classically they do. We found several classes of such processes, but the full characterization is left as a big, open problem for future research. It may be especially difficult to taking into account that the classical version of the problem is still unsolved for  $d > 3$ ; however, recent progress on accessibility of quantum channels via the Lindblad semigroup [60] is promising. Moreover, one may still hope for a partial characterization; e.g., it would be of particular interest to identify the outer limits to the quantum advantage by means of necessary conditions for quantum embeddability.

We have also proved the quantum advantage in terms of memory and time-step cost of implementing a given stochastic process. By means of computational-basis input states and measurements, a quantum Markovian master equation on a  $(d + m_Q)$ -dimensional quantum system realizes a  $d \times d$  stochastic process  $P$  in  $\tau_Q$  time steps.

We compared  $(m_Q, \tau_Q)$  with the minimal  $(m_C, \tau_C)$  required in order to simulate the same  $P$  through any classical Markovian master equation. When  $P$  is deterministic, i.e., a function over a discrete state space, we saw that, typically, one has a gap between  $(m_Q, \tau_Q)$  and  $(m_C, \tau_C)$ .

From a technological perspective, our investigations led to the question of whether these in-principle simulation advantages translate into practical ones. There is a long history, dating back to Landauer, of associating a cost to erasure [51] because of the unavoidable dissipation of entropy required for the implementation of such processes. Notably, these costs are common to classical and quantum scenarios. On the other hand, we showed here that the memory and time-step costs of computations under memoryless quantum dynamics are typically lower than the corresponding classical costs. However, practical advantages can be expected once we overcome technological barriers concerning (a) information density and (b) coherent control of quantum systems. As we do not expect these barriers to be overcome within the next few decades, we decided not to discuss this point as a potential application.

It would be of interest to extend our framework to sequential scenarios, where one samples  $P_{ij}(t_k)$  over the visible states  $i, j$  at discrete times  $t_k$ , and looks for a predictive dynamical model matching the observed data. We showed that quantum models can reproduce the same data with advantages in terms of the number of hidden states and time steps. It would be worthwhile to investigate whether such an advantage becomes larger in sequential scenarios. Whether these scenarios translate to practical advantages that can be realized in near-term quantum devices is an intriguing open question of the present investigation.

Finally, our results may also find applicability in the realm of quantum thermodynamics. Our proof that qubit systems can be quantum mechanically cooled below the environmental temperature using neither memory effects nor ancillas (something that is impossible classically) suggests that we should look for realistic Markovian master equations in which this phenomenon can be observed. One could also investigate possible memory advantages in control (and especially cooling) for higher-dimensional systems.

### ACKNOWLEDGMENTS

The authors would like to thank Stephen Bartlett for hosting them at the initial stage of the project, as well as Iacopo Carusotto and Claudio Fontanari for their kind hospitality in Trento, where part of this project was carried out. K. K. would also like to thank K. Życzkowski for helpful discussions. K. K. acknowledges financial support by the Foundation for Polish Science through the IRAP project co-financed by EU within the Smart Growth Operational Programme (Contract No. 2018/MAB/5) and through the TEAM-NET project (Contract No. POIR.04.04.00-00-17C1/18-00). M. L. acknowledges financial support

from the European Union’s Marie Skłodowska-Curie individual fellowships (No. H2020-MSCA-IF-2017 and No. GA794842), Spanish MINECO (Severo Ochoa SEV-2015-0522 and Project No. QIBEQI FIS2016-80773-P), Fundacio Cellex and Generalitat de Catalunya (CERCA Programme and SGR 875), and ERC Grant EQEC No. 682726.

### APPENDIX A: TYPICAL FUNCTIONS

Consider a function  $f: \mathbb{Z}_d \rightarrow \mathbb{Z}_d$  sampled uniformly from the set of all such functions. First, focus on the dimension of the image. The probability that a given  $a \in \mathbb{Z}_d$  is not in the image of  $f$  is  $(1 - 1/d)^d \approx 1/e$  for large  $d$ . The average dimension of the image is hence a binomial with average  $d[1 - (1/e)]$  and variance  $(d/e)[1 - (1/e)]$ . Hence, for large  $d$ , the size of the image of  $f$  is  $O(d)$ . Second, focus on the number of fixed points. The probability that a given  $a \in \mathbb{Z}_d$  is a fixed point is  $1/d$ . The number of fixed points is hence a binomial with average  $(1/d) \times d = 1$  and variance  $1 \times [1 - (1/d)]$ . As such, for large  $d$ , the number of fixed points is  $O(1)$ .

### APPENDIX B: PROOF OF THEOREM 4

*Proof.* Given a function  $f: \mathbb{Z}_d \rightarrow \mathbb{Z}_d$ , let us denote the size of the image of  $f$  by  $r = |\text{img}(f)|$ . Next, we denote the elements of  $\text{img}(f)$  by  $\{y_k\}_{k=1}^r$ , and the remaining elements belonging to  $\mathbb{Z}_d \setminus \text{img}(f)$  by  $\{y_k\}_{k=r+1}^d$ . Moreover, for each  $y_k$  with  $k \leq r$ , i.e., for each of the  $r$  elements of the image of  $f$ , let us denote the corresponding preimage as follows:

$$f^{-1}(y_k) = \{x_j^k\}_{j=1}^{d_k}. \quad (\text{B1})$$

Note that the sets  $\{x_j^k\}_{j=1}^{d_k}$  are disjoint and that their union is the full set  $\mathbb{Z}_d$ .

Now, we construct a permutation function  $f_\pi$  and an idempotent function  $f_I$ , both mapping  $\mathbb{Z}_d$  to  $\mathbb{Z}_d$ , such that

$$f = f_I \circ f_\pi. \quad (\text{B2})$$

First,  $f_\pi$  is defined by

$$f_\pi(x_j^k) = \begin{cases} y_k & \text{if } j = 1 \\ y_{r + \sum_{l=1}^{k-1} (d_l - 1) + j - 1} & \text{otherwise,} \end{cases} \quad (\text{B3})$$

where the convention is that  $\sum_{l=1}^0 \equiv 0$ . Then, introducing  $n_s := r + \sum_{l=1}^{s-1} (d_l - 1)$ , the idempotent map  $f_I$  is given by

$$f_I(y_k) = \begin{cases} y_k & \text{for } k \leq r \\ y_s & \text{for } k \in \{n_s + 1, n_s + d_s - 1\}. \end{cases} \quad (\text{B4})$$

With the above definitions, it is a straightforward calculation to show that Eq. (B2) holds.

Finally, we need to show that there exist time-independent Lindbladians  $\mathcal{L}_\pi$  and  $\mathcal{L}_I$  that generate  $P_{f_\pi}$  and  $P_{f_I}$ , i.e.,  $\{0, 1\}$ -valued stochastic matrices realizing functions  $f_\pi$  and  $f_I$ , respectively. Thus, by Definition 6, we will prove that any function  $f$  can be realized quantumly without the use of memory and in at most two time steps. First, since  $P_{f_\pi}$  is a permutation, its generator  $\mathcal{L}_\pi$  exists and is simply given by the commutator with the Hamiltonian (see discussion in Sec. II C). Now, in the case of  $f_I$ , notice that it is a function sending  $r$  disjoint sets  $Y_k$  of size  $d_k$ ,

$$Y_k = y_k \cup \{y_l\}_{l=n_k+1}^{n_k+d_k-1}, \quad (\text{B5})$$

to a single element  $y_k$  of the given set  $Y_k$ . This mapping can be easily realized for  $t_f \rightarrow \infty$  by a classical generator  $L$  (so also by the corresponding quantum Lindbladian) given by

$$L_{y_k|y_l} = \begin{cases} -1 & \text{for } k = l \text{ and } l > r \\ 1 & \text{for } k = f_I(l) \text{ and } l > r \\ 0 & \text{otherwise.} \end{cases} \quad (\text{B6})$$

### APPENDIX C: EXTREMAL PATH

Consider the initial qubit state  $\rho$  described by the Bloch vector  $(x, 0, z)$  and a fixed state  $\rho_\gamma$  with the Bloch vector  $(0, 0, \zeta)$ . Here, we show how to construct quantum channels  $\mathcal{E}_0$  and  $\mathcal{E}_1$  with a fixed point  $\rho_\gamma$ , evolving  $\rho$  along the extremal circles  $(\mathbf{c}_0, R_0)$  and  $(\mathbf{c}_1, R_1)$ , as derived in Ref. [50] and described in Sec. IV B. More precisely, for a given  $\Delta > 0$ , we look for  $\mathcal{E}_0$  that evolves  $\rho$  to  $\rho'$  with

$$z' = z + \Delta, \quad x' = \sqrt{R_0^2 - (z' - \zeta(1 - R_0))^2}. \quad (\text{C1})$$

Similarly, for a given  $\Delta > 0$ , we look for  $\mathcal{E}_1$  that evolves  $\rho$  to  $\rho'$  with

$$z' = z - \Delta, \quad x' = \sqrt{R_1^2 - (z' - \zeta(1 + R_1))^2}. \quad (\text{C2})$$

Note that, in both cases, there is a maximal value of  $\Delta$  for  $x'$  to stay real, and we assume that  $\Delta$  is below that maximal value (otherwise, the map we are looking for does not exist). Below, we explain how to construct Kraus operators  $\{A_i, B_i, C_i\}$  for  $\mathcal{E}_i$  so that

$$\mathcal{E}_i(\cdot) = A_i(\cdot)A_i^\dagger + B_i(\cdot)B_i^\dagger + C_i(\cdot)C_i^\dagger. \quad (\text{C3})$$

The construction is based on the general construction for channels mapping between pairs of qubit states as provided by Alberti and Uhlmann in Ref. [48].

The first step is to define the following projectors:

$$|\psi_0\rangle\langle\psi_0| = \frac{1}{R_0}(\rho - (1 - R_0)\rho_\gamma), \quad (\text{C4a})$$

$$|\psi'_0\rangle\langle\psi'_0| = \frac{1}{R_0}(\rho' - (1 - R_0)\rho_\gamma), \quad (\text{C4b})$$

$$|\psi_1\rangle\langle\psi_1| = -\frac{1}{R_1}(\rho - (1 + R_1)\rho_\gamma), \quad (\text{C4c})$$

$$|\psi'_1\rangle\langle\psi'_1| = -\frac{1}{R_1}(\rho' - (1 + R_1)\rho_\gamma). \quad (\text{C4d})$$

The above four projectors are used to define four unitary matrices,

$$U_i = |0\rangle\langle\psi_i| + |1\rangle\langle\psi_i^\perp|, \quad (\text{C5})$$

with  $i \in \{0, 1\}$  and an analogous primed definition for  $U'_i$ . These matrices are then employed to define four rotated fixed states  $\Gamma_i := U_i\rho_\gamma U_i^\dagger$  and analogously for  $\Gamma'_i$ . Let us parametrize these states as follows:

$$\Gamma_i = \begin{pmatrix} a_i & \epsilon_i\sqrt{a_i(1-a_i)} \\ \epsilon_i\sqrt{a_i(1-a_i)} & 1-a_i \end{pmatrix}. \quad (\text{C6})$$

These eight parameters are then used to calculate the following eight new parameters:

$$\alpha_i = \sqrt{\frac{a_i(1-a'_i)}{a'_i(1-a_i)} \frac{\epsilon_i\epsilon'_i}{1-\frac{a_i}{a'_i}(1-\epsilon_i^2)}}, \quad (\text{C7a})$$

$$\beta_i = \sqrt{\frac{(a'_i - a_i)(1-a'_i)}{(1-a_i)a'_i} \frac{\epsilon'_i}{1-\frac{a_i}{a'_i}(1-\epsilon_i^2)}}, \quad (\text{C7b})$$

$$\gamma_i = \sqrt{\frac{(1-\epsilon_i^2) - \frac{a_i}{a'_i}(1-\epsilon_i^2)}{1-\frac{a_i}{a'_i}(1-\epsilon_i^2)}} \sqrt{\frac{1-a'_i}{1-a_i}}, \quad (\text{C7c})$$

$$\omega_i = \sqrt{\frac{a'_i - a_i}{1-a_i}}. \quad (\text{C7d})$$

These parameters, in turn, allow us to introduce the following operators:

$$A_i^* = \begin{pmatrix} 1 & 0 \\ 0 & \alpha_i \end{pmatrix}, \quad B_i^* = \begin{pmatrix} 0 & \omega_i \\ 0 & \beta_i \end{pmatrix}, \quad C_i^* = \begin{pmatrix} 0 & 0 \\ 0 & \gamma_i \end{pmatrix}, \quad (\text{C8})$$

which, after unitary rotations, yield the final Kraus operators we seek:

$$A_i = U_i^{\dagger} A_i^* U_i, \quad B_i = U_i^{\dagger} B_i^* U_i, \quad C_i = U_i^{\dagger} C_i^* U_i. \quad (\text{C9})$$

[1] N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden, *Quantum Cryptography*, *Rev. Mod. Phys.* **74**, 145 (2002).

- [2] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, England, 2010).
- [3] I. M. Georgescu, S. Ashhab, and F. Nori, *Quantum Simulation*, *Rev. Mod. Phys.* **86**, 153 (2014).
- [4] C. L. Degen, F. Reinhard, and P. Cappellaro, *Quantum Sensing*, *Rev. Mod. Phys.* **89**, 035002 (2017).
- [5] W. Y. Suen, J. Thompson, A. J. Garner, V. Vedral, and M. Gu, *The Classical-Quantum Divergence of Complexity in Modelling Spin Chains*, *Quantum* **1**, 25 (2017).
- [6] J. Thompson, A. J. Garner, J. R. Mahoney, J. P. Crutchfield, V. Vedral, and M. Gu, *Causal Asymmetry in a Quantum World*, *Phys. Rev. X* **8**, 031013 (2018).
- [7] F. C. Binder, J. Thompson, and M. Gu, *Practical Unitary Simulator for Non-Markovian Complex Processes*, *Phys. Rev. Lett.* **120**, 240502 (2018).
- [8] F. Ghafari, N. Tischler, J. Thompson, M. Gu, L. K. Shalm, V. B. Verma, S. W. Nam, R. B. Patel, H. M. Wiseman, and G. J. Pryde, *Dimensional Quantum Memory Advantage in the Simulation of Stochastic Processes*, *Phys. Rev. X* **9**, 041013 (2019).
- [9] T. J. Elliott, C. Yang, F. C. Binder, A. J. Garner, J. Thompson, and M. Gu, *Extreme Dimensional Compression with Quantum Modelling*, *Phys. Rev. Lett.* **125**, 260501 (2020).
- [10] D. H. Wolpert, A. Kolchinsky, and J. A. Owen, *A Space-Time Tradeoff for Implementing a Function with Master Equation Dynamics*, *Nat. Commun.* **10**, 1727 (2019).
- [11] H.-P. Breuer, E.-M. Laine, and J. Piilo, *Measure for the Degree of Non-Markovian Behavior of Quantum Processes in Open Systems*, *Phys. Rev. Lett.* **103**, 210401 (2009).
- [12] P. Erker, M. T. Mitchison, R. Silva, M. P. Woods, N. Brunner, and M. Huber, *Autonomous Quantum Clocks: Does Thermodynamics Limit Our Ability to Measure Time?*, *Phys. Rev. X* **7**, 031022 (2017).
- [13] V. Gultieri, C. Benedetti, and M. G. Paris, *Quantum-classical Dynamical Distance and Quantumness of Quantum Walks*, *Phys. Rev. A* **102**, 012201 (2020).
- [14] M. Lostaglio and K. Korzekwa, *Fundamental Constraints of Quantum Thermodynamics in the Markovian Regime* (to be published).
- [15] Á. M. Alhambra, M. Lostaglio, and C. Perry, *Heat-Bath Algorithmic Cooling with Optimal Thermalization Strategies*, *Quantum* **3**, 188 (2019).
- [16] P. Taranto, F. Bakhshinezhad, P. Schüttelkopf, and M. Huber, *Exponential Improvement for Quantum Cooling through Finite Memory Effects*, *Phys. Rev. Applied* **14**, 054005 (2020).
- [17] E. B. Davies, *Embeddable Markov Matrices*, *Electron. J. Pro* **15**, 1474 (2010).
- [18] G. Elfving, *Zur Theorie der Markoffschen Ketten*, *Acta Soc. Sci. Fennicae*, n. Ser. A2 **8**, 1 (1937).
- [19] J. F. C. Kingman, *The Imbedding Problem for Finite Markov Chains*, *Probab. Theory Relat. Fields* **1**, 14 (1962).
- [20] J. T. Runnenburg, *On Elfving's Problem of Imbedding a Time-Discrete Markov Chain in a Time-Continuous One for Finitely Many States I*, *P. K. Ned. Akad. A Math* **65**, 536 (1962).
- [21] G. S. Goodman, *An Intrinsic Time for Non-Stationary Finite Markov Chains*, *Probab. Theory Relat. Fields* **16**, 165 (1970).
- [22] P. Carette, *Characterizations of Embeddable  $3 \times 3$  Stochastic Matrices with a Negative Eigenvalue*, *New York J. Math.* **1**, 129 (1995), <http://nyjm.albany.edu/j/1995/1-8p.pdf>.
- [23] M. M. Wolf, J. Eisert, T. S. Cubitt, and J. I. Cirac, *Assessing Non-Markovian Quantum Dynamics*, *Phys. Rev. Lett.* **101**, 150402 (2008).
- [24] V. Gorini, A. Kossakowski, and E. C. G. Sudarshan, *Completely Positive Dynamical Semigroups of  $N$ -Level Systems*, *J. Math. Phys. (N.Y.)* **17**, 821 (1976).
- [25] G. Lindblad, *On the Generators of Quantum Dynamical Semigroups*, *Commun. Math. Phys.* **48**, 119 (1976).
- [26] M. Campisi, P. Hänggi, and P. Talkner, *Colloquium: Quantum Fluctuation Relations: Foundations and Applications*, *Rev. Mod. Phys.* **83**, 771 (2011).
- [27] I. Bengtsson, *The Importance of Being Unistochastic*, [arXiv:quant-ph/0403088](https://arxiv.org/abs/quant-ph/0403088).
- [28] I. Bengtsson, Å. Ericsson, M. Kuś, W. Tadej, and K. Życzkowski, *Birkhoff's Polytope and Unistochastic Matrices,  $n = 3$  and  $n = 4$* , *Commun. Math. Phys.* **259**, 307 (2005).
- [29] P. Lencastre, F. Rauschel, T. Rogers, and P. G. Lind, *From Empirical Data to Time-Inhomogeneous Continuous Markov Processes*, *Phys. Rev. E* **93**, 032135 (2016).
- [30] Y.-T. Poon and N.-K. Tsing, *Inclusion Relations between Orthostochastic Matrices and Products of Pinching Matrices*, *Linear Multilinear Algebra* **21**, 253 (1987).
- [31] A. Streltsov, G. Adesso, and M. B. Plenio, *Colloquium: Quantum Coherence as a Resource*, *Rev. Mod. Phys.* **89**, 041003 (2017).
- [32] I. Marvian and R. W. Spekkens, *How to Quantify Coherence: Distinguishing Speakable and Unspeakable Notions*, *Phys. Rev. A* **94**, 052324 (2016).
- [33] A. Montina, *Exponential Complexity and Ontological Theories of Quantum Mechanics*, *Phys. Rev. A* **77**, 022104 (2008).
- [34] L. Hardy, *Quantum Ontological Excess Baggage*, *Stud. Hist. Philos. Sci. B* **35**, 267 (2004).
- [35] D. Freedman, *Approximating Countable Markov Chains* (Springer-Verlag, Berlin, 1983).
- [36] H.-P. Breuer and F. Petruccione, *The Theory of Open Quantum Systems* (Oxford University Press, New York, 2002).
- [37] M. Lostaglio, Á. M. Alhambra, and C. Perry, *Elementary Thermal Operations*, *Quantum* **2**, 52 (2018).
- [38] M. Horodecki and J. Oppenheim, *Fundamental Limitations for Quantum and Nanoscale Thermodynamics*, *Nat. Commun.* **4**, 2059 (2013).
- [39] M. Lostaglio, *An Introductory Review of the Resource Theory Approach to Thermodynamics*, *Rep. Prog. Phys.* **82**, 114001 (2019).
- [40] E. Ruch, R. Schraner, and T. H. Seligman, *The Mixing Distance*, *J. Chem. Phys.* **69**, 386 (1978).
- [41] A. W. Marshall, I. Olkin, and B. C. Arnold, *Inequalities: Theory of Majorization and Its Applications* (Springer, New York, 2010).
- [42] E. B. Davies, *Markovian Master Equations*, *Commun. Math. Phys.* **39**, 91 (1974).
- [43] W. Roga, M. Fannes, and K. Życzkowski, *Davies Maps for Qubits and Qutrits*, *Rep. Math. Phys.* **66**, 311 (2010).

- [44] R. Alicki and R. Kosloff, *Introduction to Quantum Thermodynamics: History and Prospects*, in *Thermodynamics in the Quantum Regime: Fundamental Aspects and New Directions*, edited by F. Binder, L. A. Correa, C. Gogolin, J. Anders, and G. Adesso (Springer Nature, Switzerland, 2018), pp. 1–33.
- [45] R. Uzdin, A. Levy, and R. Kosloff, *Equivalence of Quantum Heat Machines, and Quantum-Thermodynamic Signatures*, *Phys. Rev. X* **5**, 031044 (2015).
- [46] P. Faist and R. Renner, *Fundamental Work Cost of Quantum Processes*, *Phys. Rev. X* **8**, 021011 (2018).
- [47] C. Perry, P. Ćwikliński, J. Anders, M. Horodecki, and J. Oppenheim, *A Sufficient Set of Experimentally Implementable Thermal Operations for Small Systems*, *Phys. Rev. X* **8**, 041049 (2018).
- [48] P. Alberti and A. Uhlmann, *A Problem Relating to Positive Linear Maps on Matrix Algebras*, *Rep. Math. Phys.* **18**, 163 (1980).
- [49] G. Gour, *Quantum Resource Theories in the Single-Shot Regime*, *Phys. Rev. A* **95**, 062314 (2017).
- [50] K. Korzekwa, *Structure of the Thermodynamic Arrow of Time in Classical and Quantum Theories*, *Phys. Rev. A* **95**, 052318 (2017).
- [51] R. Landauer, *Irreversibility and Heat Generation in the Computing Process*, *IBM J. Res. Dev.* **5**, 183 (1961).
- [52] C. H. Bennett, *Logical Reversibility of Computation*, *IBM J. Res. Dev.* **17**, 525 (1973).
- [53] J. Koomey, S. Berard, M. Sanchez, and H. Wong, *Implications of Historical Trends in the Electrical Efficiency of Computing*, *IEEE annals of the history of computing* **33**, 46 (2010).
- [54] D. P. DiVincenzo, *The Physical Implementation of Quantum Computation*, *Fortschr. Phys.* **48**, 771 (2000).
- [55] L. J. Schulman and U. V. Vazirani, *Molecular Scale Heat Engines and Scalable Quantum Computation*, in *Proceedings of the 31st Annual ACM Symposium on Theory of Computing* (ACM, 1999), pp. 322–329, <https://doi.org/10.1145/301250.301332>.
- [56] P. O. Boykin, T. Mor, V. Roychowdhury, F. Vatan, and R. Vrijen, *Algorithmic Cooling and Scalable NMR Quantum Computers*, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 3388 (2002).
- [57] N. A. Rodríguez-Briones, *Novel Heat-Bath Algorithmic Cooling Methods*, Ph.D. thesis, University of Waterloo (2020).
- [58] N. A. Rodríguez-Briones, E. Martín-Martínez, A. Kempf, and R. Laflamme, *Correlation-Enhanced Algorithmic Cooling*, *Phys. Rev. Lett.* **119**, 050502 (2017).
- [59] M. Lostaglio, D. Jennings, and T. Rudolph, *Description of Quantum Coherence in Thermodynamic Processes Requires Constraints Beyond Free Energy*, *Nat. Commun.* **6**, 6383 (2015).
- [60] F. Shahbeigi, D. Amaro-Alcalá, Z. Puchała, and K. Życzkowski, *Log-Convex Set of Lindblad Semigroups Acting on  $N$ -Level System*, [arXiv:2003.12184](https://arxiv.org/abs/2003.12184).