

Simulation of quantum walks and fast mixing with classical processesSimon Apers,^{1,*} Alain Sarlette,^{1,2} and Francesco Ticozzi^{3,4}¹*Department of Electronics and Information Systems, Ghent University, 9000 Ghent, Belgium*²*QUANTIC Laboratory, INRIA Paris, 75012 Paris, France*³*Dipartimento di Ingegneria dell'Informazione, Università di Padova, Padua 35131, Italy*⁴*Department of Physics and Astronomy, Dartmouth College, Hanover, New Hampshire 03755, USA*

(Received 4 December 2017; published 20 September 2018)

We compare discrete-time quantum walks on graphs to their natural classical equivalents, which we argue are lifted Markov chains (LMCs), that is, classical Markov chains with added memory. We show that LMCs can simulate the mixing behavior of any quantum walk, under a commonly satisfied invariance condition. This allows us to answer an open question on how the graph topology ultimately bounds a quantum walk's mixing performance, and that of any stochastic local evolution. The results highlight that speedups in mixing and transport phenomena are not necessarily diagnostic of quantum effects, although superdiffusive spreading is more prominent with quantum walks. The general simulating LMC construction may lead to large memory, yet we show that for the main graphs under study (i.e., lattices) this memory can be brought down to the same size employed in the quantum walks proposed in the literature.

DOI: [10.1103/PhysRevA.98.032115](https://doi.org/10.1103/PhysRevA.98.032115)**I. INTRODUCTION**

Random walks are ubiquitous models for natural processes and a versatile algorithmic tool to explore and extract information about networks. In recent years, similar promise has been attributed to their quantum analog, quantum walks (QWs). QWs describe how the position of a “walking” quantum particle evolves on a graph, possibly entangled with other (so-called coin) quantum degrees of freedom. The joint dynamics can be discrete-time or continuous and must respect graph locality [1–4]. Following the realization that QWs on a line can beat the diffusive behavior typical of classical stochastic processes [1,5], they have been related to improved transport phenomena in biological systems [6,7], thermodynamic theories, breakdown models, and topological states of matter [8–10], and simulated in various experiments [11–15]. Furthermore, they have been studied as a paradigm for quantum computing [16,17] and to speed up algorithmic tasks [18], in particular, those related to Grover's search algorithm [19–21].

Despite impressive advances, elucidating the source and extent of quantum advantages and providing general design principles to achieve them with QWs remain open questions. A general quadratic speedup by QWs has been established for the hitting time, namely finding a marked node in any graph [21–26]. Albeit closer to the original observations on the line [5], the complementary problem of mixing, that is, converging to a particular probability distribution over the nodes, has resisted general QW speedup analysis. There is evidence for quadratic speedup, with respect to classical Markov chains, on specific graphs, including the cycle [27], hypercube [28], and torus [29]. A general characterization of QW mixing would be a fundamental step for investigating quantum vs

classical differences in statistical mechanics (thermodynamic equilibration, transport phenomena, localization defects), and its algorithmic complexity is of key relevance for applications such as sampling and Monte Carlo simulations [30].

The simulation of QWs with classical *waves* [31–34] suggests that superposition and interference effects underlie their nondiffusive spreading. In contrast, we here show that any mixing speedup can be matched using local probabilistic evolutions of *particles with memory*.

We characterize mixing performance of QWs by showing that they belong to a class of processes which can be simulated by *classical* Markov chains whose “walking” particle is locally endowed with additional memory, called *lifted Markov chains* (LMCs) [35]. For general graphs, our constructive proof is reminiscent of a classical version of the “clock Hamiltonians” by Feynman and Kitaev and used to prove universality of adiabatic computing [36–38], in combination with “stochastic bridges” generalizing [39] and [40,41] to simulate quantum channels for any fixed initial conditions. We next derive a tight bound on achievable mixing speedup, for any QW starting from any classical probability distribution over the nodes and which would keep the target distribution invariant. Our LMC approach thereby improves on the known bounds in [27,42]. Finally, for lattices, on which most QW mixing speedups have been demonstrated, we relate the QWs to fast-mixing LMCs that have not only the same mixing performance, but also the same structure [43,44].

These results provide several insights. First, an observed speedup in mixing is *not* fundamentally diagnostic of interference effects, be it classical or quantum, as it may always be explained by a classical walking particle with memory, on any graph. Second, as LMCs are just Markov chains on enlarged graphs, QW mixing is simulated within a classical probabilistic computational framework, and it shares the same fundamental bounds of algorithms relying on classical

*Corresponding author: simon.apers@ugent.be

Markov chains. Third, the search for a quantum advantage should focus on identifying efficient designs, beyond mixing on lattices, in terms of amount of memory or graph knowledge required. Whether for statistical mechanics, evolutionarily selected biological systems, or design of faster Monte Carlo algorithms, our results significantly narrow the context in which quantum effects may provide intrinsic advantages.

The remainder of the paper is organized as follows. In Sec. II we illustrate with a simple example why it is natural to compare QWs with LMCs. In Sec. III we define the general setting. In Sec. IV we state our main results: classical stochastic processes can simulate the time evolution of the distribution induced by *any* QWs over a general graph, and for typical QWs, which leave the target classical distribution invariant, this allows us to explicitly build an LMC with a mixing rate that is essentially equal to the QW mixing rate for all times (Theorem 1). As a consequence, we show that the mixing time of QWs is subject to the exact same conductance bound as for classical LMCs (Theorem 2). These results follow from the general yet more abstract analysis presented in Sec. VI, where they will be proved in full. While in proving these results we do not address how to efficiently design a fast-mixing QW or LMC under design constraints and limited resources, the issue is touched on in Sec. V by observing that known “simple” designs of fast-mixing processes, with little memory, are essentially restricted to the lattice graphs for QWs, and on those graphs a similarly simple LMC design mixes equally fast. In Sec. VI, we prove that the results of Sec. IV remain valid for a more general class of stochastic evolutions. This setting includes as special cases the previously introduced QWs and LMCs, as well as QWs whose output distribution is considered in a time-averaged sense. The latter represents a widely used way to “induce” mixing for quantum evolutions associated with unitary dynamics (so-called “Cesaro averaging”).

II. QWs AND THEIR CLASSICAL COUNTERPARTS: A PARADIGMATIC EXAMPLE

Usually, QWs are presented as the quantum analogs of, and compared to, classical random walks. We next argue that extended classical models should be considered, as QWs exhibit genuine *memory effects*. These effects are generally attributed to both quantum phases and internal degrees of freedom of the particle, such as the internal spin in [1]. Standard discrete-time QWs [23,27,45] describe the evolution of a quantum particle (“walker”) over a discrete set of graph nodes \mathcal{V} , conditioned on additional degrees of freedom \mathcal{C} (the walker’s coin or spin). The walker state is thus defined on the Hilbert space $\mathcal{H} = \mathcal{H}_{\mathcal{C}} \otimes \mathcal{H}_{\mathcal{V}} = \text{span}\{|c\rangle \otimes |v\rangle | (c, v) \in \mathcal{C} \times \mathcal{V}\}$, while the object of interest is the induced distribution p_t over \mathcal{V} . The cycle graph is a simple example where QWs provide a mixing speedup; see Fig. 1. To the nodes $\mathcal{V} = \{1, 2, \dots, N\}$ of the cycle, the QW adds a binary coin $\mathcal{C} = \{+, -\}$ [5,27]. Denoting \mathbf{P}^{\pm} the cyclic permutation of position, that is, $\mathbf{P}^{\pm}|v\rangle = |(v \pm 1) \bmod N\rangle$ for $v \in \mathcal{V}$, the unitary QW primitive reads

$$\mathbf{U} = \mathbf{S}(\mathbf{C} \otimes \mathbf{I}_N), \quad \mathbf{C} = \begin{bmatrix} e^{-i\phi} \sqrt{1-\alpha} & e^{i\theta} \sqrt{\alpha} \\ -e^{-i\theta} \sqrt{\alpha} & e^{i\phi} \sqrt{1-\alpha} \end{bmatrix},$$

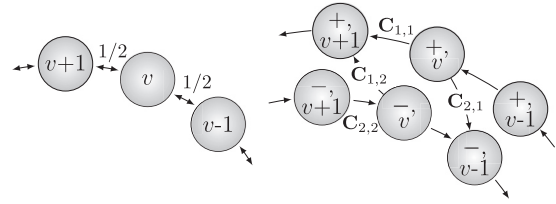


FIG. 1. Left: Random walk \mathbf{P}_0 on the cycle. Right: Unitary quantum walk with coin toss \mathbf{C} , or lifted Markov chain with a stochastic coin toss $\bar{\mathbf{C}}$, suggesting their comparison.

where $\mathbf{S} = |+\rangle \langle +| \otimes \mathbf{P}^+ + |-\rangle \langle -| \otimes \mathbf{P}^-$ expresses conditional shifts, while \mathbf{C} is a general unitary “coin toss” on $\mathcal{H}_{\mathcal{C}}$. The conditional motion can also be viewed as spin-orbit coupling. The dynamics $U^t |\psi_0\rangle$ induces a distribution p_t that spreads over \mathcal{V} in $O(N)$ steps and oscillates quasi-periodically. To perfectly *mix*, i.e., make p_t converge and remain close to the uniform distribution over \mathcal{V} , some nonunitary process must be added. One standard way of doing this in the QW literature is Cesaro averaging, where the output distribution over the nodes at time T is considered to be the uniform average of the induced probability distributions at time steps $0, 1, \dots, T$ [27,28,46]. Another, maybe operationally more direct way to induce mixing is to intersperse some measurement or slight decoherence; see [47] for a survey. For instance, one can perform with probability $q(t)$ a projective measurement in the canonical basis between consecutive applications of \mathbf{U} :

$$|\psi_{t+1}\rangle = \begin{cases} \mathbf{U} |\psi_t\rangle, & \text{probability } 1 - q(t), \\ |c, v\rangle, & \text{probability } q(t) |\langle c, v | \mathbf{U} |\psi_t\rangle|^2. \end{cases} \quad (1)$$

Taking $q(t) = 1$ for all multiples of $T \in \mathbb{N}$, $q(t) = 0$ otherwise, implements iterative application of a unitary QW with “final” measurement after T steps; $q(t) = 1 \forall t$ projects the state at each step and comes down to a classical random walk with transition matrix $\mathbf{P}_0 = (\mathbf{P}^+ + \mathbf{P}^-)/2$. With constant $q = \Theta(1/N)$, and for instance $\alpha = 1/2$, $\phi = \theta = 0$ in the above expression for \mathbf{U} , the QW converges towards a uniform p_t in $t = O(N)$ steps, from any initial distribution [29]. Applying the Cesaro averaging procedure on top of the unitary QW primitive gives the same mixing speed. In contrast, a classical random walk with transition matrix $\mathbf{P}_0 = (\mathbf{P}^+ + \mathbf{P}^-)/2$ reaches the uniform distribution only after $O(N^2)$ steps.

Compared to classical random walks, the QW above adds memory via the coin. Yet, QWs can exhibit memory effects even without coin. Consider the two-node graph $\mathcal{H} = \text{span}\{|1\rangle, |2\rangle\}$, equivalent to a qubit, and take the Hadamard gate $U_H = (\sigma_x + \sigma_z)/\sqrt{2}$ as QW primitive. Starting on a given node, after one step, the distribution is uniform over $|1\rangle, |2\rangle$, yet at the second step the initial state is perfectly recovered since $(U_H)^2$ equals identity. This behavior, impossible for any classical memoryless process, comes from the quantum state storing information in its relative phases, or coherences. Hence, to establish an intrinsic quantum advantage, QWs should be compared to classical local processes with at least some amount of additional memory.

Remarkably, a classical walker with memory that mixes fast on the cycle has already been proposed independently of

the QW literature [43,44], and it shares striking similarities. This walker moves among *classical* states in $\mathcal{C} \times \mathcal{V}$. Its probability distribution \hat{p} over $\mathcal{C} \times \mathcal{V}$ evolves as $\hat{p}_{t+1} = \mathbf{P}\hat{p}_t$, with stochastic transition matrix \mathbf{P} having the same structure as \mathbf{U} , yet with \mathbf{C} replaced by a *stochastic* coin toss:

$$\mathbf{P} = \mathbf{S}(\bar{\mathbf{C}} \otimes \mathbf{I}_N), \quad \bar{\mathbf{C}} = \begin{bmatrix} 1 - \alpha & \alpha \\ \alpha & 1 - \alpha \end{bmatrix}. \quad (2)$$

This can be seen as the mixture of two reversible evolutions: with probability $1 - \alpha$, the state follows the conditional shift \mathbf{S} , or with probability α , the coin is switched before applying \mathbf{S} . The coin allows the classical walker to use information about its previous motion direction, in physical terms its momentum. The similarity between \mathbf{U} and \mathbf{P} carries a deeper connection, as \mathbf{P} in Eq. (2) exactly describes the probabilistic evolution induced by Eq. (1) with $q(t) = 1\forall t$, and starting with $|\psi\rangle = |c, v\rangle$ for some $(c, v) \in \mathcal{C} \times \mathcal{V}$.

Such \mathbf{P} mixes over the cycle in $O(N)$ steps [43,44], provided $\alpha = \Theta(1/N)$. This speedup, only due to classical memory, matches the one provided by the QW in Eq. (1) with $q = \Theta(1/N)$. In both cases, a $\Theta(1/N)$ nonunitarity provides a good tradeoff between fast (unitary) motion and losing correlation with the initial condition. The same $\Theta(1/N)$ nonunitary factor appears in Cesaro averaging, if one considers that mixing is obtained when $T = \Theta(N)$.

From these observations, it appears most natural to compare QWs like Eq. (1) to classical evolutions with memory like Eq. (2), which are formalized as LMCs [35]. Establishing this comparison, and deriving associated results for QWs, is the main purpose of the present paper.

III. MIXING WITH QWs AND LMCs: GENERAL DEFINITION

Consider a graph \mathcal{G} with node set \mathcal{V} and edges $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$. Nodes could represent energy levels and edges allowed transitions. The QW and LMC constructions both start by building a lifted graph, where each node of \mathcal{G} is split into “lifted nodes” or “sublevels.” This is done without loss of generality by introducing a coin set \mathcal{C} , defining the lifted nodes $\mathcal{C} \times \mathcal{V} = \{(c, v)\}$ and selecting lifted edges in $\{(c, v), (c', v') \mid (v, v') \in \mathcal{E}\}$, thus without introducing transitions that were not allowed before lifting.

The evolution mechanism of general QWs, including for instance open quantum walks [48], is then described by a quantum channel over the space generated by viewing coin and node states as quantum, i.e.,

$$\rho_{t+1} = \Gamma_t[\rho_t] = \sum_k \mathbf{M}_k(\mathbf{t})\rho_t\mathbf{M}_k(\mathbf{t})^\dagger, \quad (3)$$

where ρ_t is a density operator on $\mathcal{H} = \text{span}\{|c, v\rangle \mid (c, v) \in \mathcal{C} \times \mathcal{V}\}$ and the \mathbf{M}_k satisfy $\sum_k \mathbf{M}_k(\mathbf{t})^\dagger\mathbf{M}_k(\mathbf{t}) = \mathbf{I}_{\mathcal{C} \times \mathcal{V}}$ for all \mathbf{t} , with \mathbf{I} the identity. This, together with (3), makes each Γ_t a completely positive, trace-preserving map over the joint graph-coin space. Graph locality imposes $\langle c', v' | \mathbf{M}_k | c, v \rangle = 0$ if $(v, v') \notin \mathcal{E}$.

For mixing over a graph, the objective is to study whether and, crucially, how fast an arbitrary initial probability distribution p_0 over the primary set of nodes \mathcal{V} converges towards some “mixed” target distribution. To address mixing with a

QW-based algorithm, any distribution p_0 must first be translated to a quantum state on the lifted nodes (or sublevels). This is typically done by a map

$$\mathbf{F} : p_0 \mapsto \rho = \sum_{v \in \mathcal{V}} p_0(v)(\rho_v \otimes |v\rangle\langle v|), \quad (4)$$

where each ρ_v is a density operator on \mathcal{C} . This map is implicit in most of the QW literature, but clearly in line with the primary earlier work on QW mixing [27,29,49]. Indeed, it specifies a way to map the problem input to an initial state for the algorithm (a quantum state on \mathcal{H}). The main reason for the explicit introduction of such a map in our context is that we consider QWs as algorithms with classical input and output, namely the probability distribution over \mathcal{V} induced by the QW. A few remarks are in order: While the QW literature does include cases in which the initial quantum state is not in the form specified by (4), like coherent quantum superpositions over the nodes of \mathcal{V} , our results will show that a classical process can still *simulate* the QW behavior over any finite time interval and for any such given initial state. On the other hand, the *mixing* task requires convergence from any initial condition. This must include initial distributions with p_0 concentrated on a single node v , in which case there is no loss of generality in assuming that for such initial distributions the map \mathbf{F} is of type (4), associating with v some quantum state ρ_v over initial coin states c . Towards general initial distributions p_0 over \mathcal{V} , if the mixing algorithm gets no further information from the process that generates p_0 , then the initialization must be linear in p_0 and this leads directly to the map \mathbf{F} being of the form presented in (4). Our results on bounding QW mixing times will be valid for any setting where the possible initial states of the QW include those generated by such \mathbf{F} .

As we anticipated, the object of interest is a distribution over \mathcal{V} , the main nodes or levels [50]. As noted in the previous section, there have been two main approaches to derive such distribution in the literature. One of them just considers the output obtained with the partial trace, as

$$p_t = f(\rho_t) = \text{diag}(\text{trace}_{\mathcal{C}}(\rho_t)),$$

where we call f the marginalization, associating with ρ_t an induced probability distribution over \mathcal{V} . Another one considers as output the *Cesaro average* or *time average* $p_t = \frac{1}{T} \sum_{t=0}^{T-1} f(\rho_t)$. This involves an additional operational element on top of the dynamics (3). For simplicity, in Sec. IV we will first introduce the main results and proof ideas with the output $p_t = f(\rho_t)$, the standard one in the LMC setting. In Sec. VI, however, we show that our results also hold with Cesaro averaging.

Similarly, an LMC is described by

$$\hat{p}_{t+1} = \mathbf{P}\hat{p}_t,$$

where \hat{p}_t is a vector representing the probability distribution over $\mathcal{C} \times \mathcal{V}$, and \mathbf{P} is a stochastic matrix expressing the jump probabilities among sublevels. Namely, denoting by $p = e_v$ and $\hat{p} = e_{(c,v)}$ the distributions with probability 1 of being on v and on (c, v) , respectively, $e_{(c',v')}^\dagger \mathbf{P} e_{(c,v)} = \mathbf{P}((c', v'), (c, v))$ is the transition probability from (c, v) to (c', v') . Graph locality imposes $\mathbf{P}((c', v'), (c, v)) = 0$ if $(v, v') \notin \mathcal{E}$. Initial lifted nodes can be assigned by a map $\mathbf{F} : p_0 \mapsto \hat{p}_0 =$

$\sum_v p_0(v) q_v \otimes e_v$, where each q_v is a distribution over the coin degrees of freedom. The distribution of interest, $p_t = f(\hat{p}_t)$, is obtained by marginalizing over \mathcal{C} , which for all $v \in \mathcal{V}$ is given by

$$p_t(v) = (f \hat{p}_t)(v) = \sum_{c \in \mathcal{C}} \hat{p}_t(c, v),$$

where now f marginalizes from a distribution over $\mathcal{C} \times \mathcal{V}$ to a distribution over \mathcal{V} . An LMC is a particular QW where populations evolve without coherences, i.e., where ρ_t remains diagonal at all times and $\mathbf{M}_k = \sqrt{\mathbf{P}((c', v'), (c, v))} |c', v'\rangle \langle c, v|$, with index k running over all nonzero elements of \mathbf{P} . The key to our results will be to observe how, conversely, any QW can be simulated by some LMC (with possibly higher-dimensional coin). In other words, the non-Markovian evolution of p_t under a QW can be described as a classical Markovian evolution of sublevel populations.

We focus on comparing the *mixing* induced by QWs and LMCs. A QW or LMC is said to mix to some distribution \bar{p} over \mathcal{V} if, for all the allowed initial conditions, the associated distribution p_t asymptotically converges to \bar{p} , that is, if for any $\epsilon > 0$ there exists a finite $\tau(\epsilon)$ such that

$$\frac{1}{2} \sum_{v \in \mathcal{V}} |p_t(v) - \bar{p}(v)| \leq \epsilon, \quad \forall p_0, \quad t \geq \tau(\epsilon).$$

The quantity $\tau(\epsilon)$ is called the ϵ -*mixing time*, quantifying the time needed to get ϵ -close to \bar{p} . With this definition, if the mixing time takes a particular value for the set of initial states resulting from the initialization map \mathbf{F} in (4), then it can only be worse when considering a more general set of initial quantum states. For readers more familiar with the setting of unitary quantum walks and Cesaro averaging, we recall that in this case mixing is defined as the convergence of the distribution $\frac{1}{T} \sum_{t=0}^{T-1} f(\rho_t)$ towards \bar{p} , with $\rho_t = |\psi_t\rangle \langle \psi_t|$ following a unitary evolution. This is covered in principle by the above definition, if we allow to define $p_t = \frac{1}{T} \sum_{t=0}^{T-1} f(\rho_t)$. We will rather work with $p_t = f(\rho_t)$ and nonunitary channels first, but Sec. VI clarifies that Theorem 1 and Theorem 2 are also valid for the Cesaro mixing time.

A standard ‘‘stabilizing’’ requirement for a process that converges to \bar{p} is that \bar{p} itself is invariant: $p_0 = \bar{p}$ should imply $p_t = \bar{p}$ for all $t \geq 0$. As we work with lifted graphs, this condition, which we call \bar{p} invariance, also depends on the choice of the initialization map \mathbf{F} . The \bar{p} -invariance condition is satisfied by essentially all QWs proposed in the mixing literature: unital QWs, decohering quantum channels, and QWs on vertex-transitive graphs. We elaborate these cases in Appendix, where for completeness we also construct some examples of QWs that are not \bar{p} -invariant.

In the next section we will show that the mixing speed of any QW can be tightly bounded when it is \bar{p} -invariant, in the same way as an LMC. As an intermediate step, we show that QWs which are not \bar{p} -invariant can still be simulated by classical processes; the bound on mixing speed however then becomes trivial, i.e., the diameter of the graph, in the absence of further constraints [51].

IV. QWs AND LMCs HAVE EQUIVALENT MIXING PERFORMANCE

In this section we present the main result from this paper: the mixing time of any \bar{p} -invariant QW can be closely matched by an LMC. We do so by explicitly constructing such an LMC. Thereto we first show how, for any given initial quantum state and any (not necessarily \bar{p} -invariant) QW, we can construct a classical process that simulates the induced dynamics. Next, we show how this can be framed as a finite LMC that simulates the QW over some fixed time interval and initialized from an arbitrary initial distribution according to (4). We note that the main obstacle towards generalizing this procedure to other initializations than (4) is to formulate a corresponding LMC initialization [52]. Finally we explain how we can pursue the LMC convergence beyond this time interval, provided that the QW is \bar{p} -invariant. As a corollary of this result, existing conductance bounds on the mixing time of LMCs provide a lower bound on the mixing time of \bar{p} -invariant QWs. We here discuss the main points, and defer the mathematical details of the proof to Secs. VIA and VIB.

A. Local stochastic simulation of a QW

First we will consider a general QW described by a quantum channel Γ_t , which is local with respect to a graph \mathcal{G} . At this point we allow for general initial quantum states ρ_0 ; later on we will introduce the initialization map \mathbf{F} .

In [39] it was shown that, for unitary quantum dynamics, the probability distribution induced on the *fixed measurement basis* associated with the nodes can be simulated using just classical means. This showed that such partial evolution is not subject to the general no-go results which regard the simulation of the full quantum state, and classical local hidden-variables theories can be built to explain observations. We start by extending this result to the evolution p_t induced by a general QW of type (3). For a general initial quantum state ρ_0 , we let $p_t^{(\rho_0)}$ denote the induced QW probability distribution starting from ρ_0 , so that

$$p_t^{(\rho_0)} = f(\rho_t) = f(\Gamma_t \Gamma_{t-1} \dots \Gamma_1[\rho_0]).$$

It is then possible to follow $p_t^{(\rho_0)}$ step by step by building a sequence of local stochastic matrices $\mathbf{P}_1^{(\rho_0)}, \mathbf{P}_2^{(\rho_0)}, \dots, \mathbf{P}_t^{(\rho_0)}$ on \mathcal{V} such that

$$p_t^{(\rho_0)} = \mathbf{P}_t^{(\rho_0)} p_{t-1}^{(\rho_0)} = \mathbf{P}_t^{(\rho_0)} \dots \mathbf{P}_2^{(\rho_0)} \mathbf{P}_1^{(\rho_0)} p_0,$$

where $p_0 = f(\rho_0)$. We use the max-flow min-cut theorem from graph theory to prove that such construction always exists. This observation can be traced back to a property that holds for local stochastic processes independently of the underlying physical mechanism (classical, quantum, or other): a node cannot contain more population at time $t + 1$ than the population at time t on itself and on its neighbors [see later, Eq. (6)].

The result shows that the induced dynamics of a general QW, starting from any given initial quantum state, can be mediated by classical stochastic transitions that preserve the locality of the original QW. For a finite set of initial quantum states, and for any distribution over these states, a finite set

of stochastic transition rules then allows us to simulate the induced probability distribution for any fixed time interval. This observation presages our main result, which builds an LMC from such stochastic simulation, and with a mixing time closely related to the QW mixing time.

B. LMC construction

As a second step, we show how to build an LMC simulating such stochastic transition rules, as illustrated in Fig. 2, for the finite time interval $\bar{\tau}(\epsilon_0)$. From Sec. IV A, when the QW is initialized from a single node $p_0 = e_v$, so that the initial quantum state $\rho_0 = \rho_v = \mathbf{F}[e_v]$, we can simulate the induced dynamics $p_t^{(\rho_v)}$ by applying the transition matrices $\mathbf{P}_t^{(\rho_v)}$ (which for short we will denote by $\mathbf{P}_t^{(v)}$). If now the same QW starts from a general probability distribution p_0 and $\rho_0 = \mathbf{F}(p_0)$, then by linearity we know that the dynamics induced by the QW satisfies

$$p_t = \sum_v p_0(v) p_t^{(\rho_v)}.$$

As a consequence, the following stochastic dynamics simulate the induced probability distribution: assign an initial node v according to the distribution p_0 , and conditionally on this initial node, apply the corresponding sequence of transitions $\mathbf{P}_1^{(v)}, \mathbf{P}_2^{(v)}, \dots$ on e_v . This indeed leads to

$$\sum_v p_0(v) \mathbf{P}_1^{(v)} \dots \mathbf{P}_2^{(v)} \mathbf{P}_1^{(v)} e_v = \sum_v p_0(v) p_t^{(\rho_v)} = p_t.$$

We thus have constructed a simulator from a set of transition matrices, conditioned on a time index and on the initial node. This is not a Markovian process, but in principle it is not contrary to what a transition mechanism with memory would allow us to do. Indeed, we can encode this classical process in a (time-independent) LMC for any finite time horizon $t \leq T$, by encoding these dependencies into the coin register.

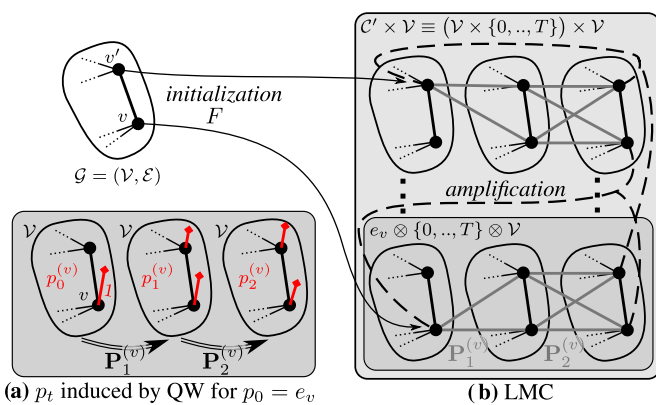


FIG. 2. Depiction of the LMC construction for $T = 2$. (a) For each node v we construct a stochastic bridge $\{\mathbf{P}_t^{(v)}\}$ that tracks the p_t induced by the QW starting on node v . (b) Each of these bridges is implemented as a time-invariant LMC on $T + 1$ copies of \mathcal{G} . These LMCs are combined into a final LMC, such that the QW starting from a general initial distribution on \mathcal{V} is simulated with the initialization \mathbf{F} , over T steps. The dashed lines represent the “amplification” edges added in the last step of the construction to pursue the LMC evolution for $t > T$.

Explicitly, we equip the nodes of the original graph with a “timer” and an extra register, containing the initial node v . This results in a lifted state space

$$\mathcal{C}' \times \mathcal{V} \equiv (\mathcal{V} \times \{0, 1, \dots, T\}) \times \mathcal{V}.$$

This lifted state space allows the LMC to conditionally apply the proper transition matrix $\mathbf{P}_t^{(v)}$. For instance, on a distribution of the form $e_v \otimes e_t \otimes e_v$, encoding the initial node v and time t , we wish the LMC to implement the transition matrix $\mathbf{P}_{t+1}^{(v)}$. This gives rise to the LMC defined by

$$\mathbf{P} = \sum_{v \in \mathcal{V}} e_v e_v^\dagger \otimes \left(\sum_{t=0}^{T-1} e_{t+1} e_t^\dagger \otimes \mathbf{P}_{t+1}^{(v)} + e_T e_T^\dagger \otimes \mathbf{I}_{\mathcal{V}} \right). \quad (5)$$

The initialization $\mathbf{F} : e_v \mapsto (e_v \otimes e_0) \otimes e_v$ simply puts the initial position in the coin register and sets the timer to zero. We illustrate how this works by writing down the dynamics of an initial state $p_0 = e_v$ on the original graph:

$$\begin{aligned} p_0 = e_v &\mapsto (e_v \otimes e_0) \otimes e_v \\ &\mapsto \mathbf{P}(e_v \otimes e_0) \otimes e_v = (e_v \otimes e_1) \otimes \mathbf{P}_1^{(v)} e_v \\ &\mapsto \mathbf{P}^2(e_v \otimes e_0) \otimes e_v = (e_v \otimes e_2) \otimes \mathbf{P}_2^{(v)} \mathbf{P}_1^{(v)} e_v \\ &\mapsto \dots \end{aligned}$$

As a consequence, if we marginalize the state $\mathbf{P}^t(e_v \otimes e_0) \otimes e_v$ over \mathcal{C}' we retrieve the correct distribution $\mathbf{P}_t^{(v)} \dots \mathbf{P}_2^{(v)} \mathbf{P}_1^{(v)} e_v = p_t^{(v)}$ for all $t \in [0, T]$. By linearity it is clear that more general initial states p_0 also get mapped to the correct state p_t . We have thus framed the simulation of a QW (3), (4) over a finite time interval into the LMC setting.

Finally, the third step of the construction deals with the limited time T over which we have built the simulation. If we choose $T = \bar{\tau}(\epsilon_0)$, then we know that after T steps the marginal distribution p_t will be ϵ_0 -close to \bar{p} for any p_0 . However, we still have to clarify what happens for $\epsilon < \epsilon_0$, as applying the LMC (5) as such would imply no convergence beyond ϵ_0 . It is here that we will use the assumption of \bar{p} invariance. Indeed, under this condition, by a standard technique in randomized algorithms, we can *amplify* the closeness to \bar{p} by simply rerunning the scheme on its output distribution. In our LMC construction, this corresponds to “resetting” the coin state of the walk when its timer reaches T . Thereto we add the operator \mathbf{P}_F which conditional on the time being T will reset the coin state, and otherwise does nothing:

$$\mathbf{P}_F(e_{v_0} \otimes e_t \otimes e_v) = \begin{cases} e_{v_0} \otimes e_t \otimes e_v, & t \neq T, \\ e_v \otimes e_0 \otimes e_v = \mathbf{F}[e_v], & t = T. \end{cases}$$

Explicitly, the final LMC construction thus becomes

$$\begin{aligned} \mathbf{P}' &= \sum_{v \in \mathcal{V}} e_v e_v^\dagger \otimes \left(\sum_{t=0}^{T-1} e_{t+1} e_t^\dagger \otimes \mathbf{P}_{t+1}^{(v)} \right) \\ &\quad + \sum_{v, v_0 \in \mathcal{V}} e_v e_{v_0}^\dagger \otimes e_0 e_T \otimes e_v e_v^\dagger. \end{aligned}$$

Since \mathbf{P}_F is a local action purely on the coin, \mathbf{P}' is still a local LMC. Every T steps the simulation is now reinitialized. When the QW is \bar{p} -invariant, the associated LMC simulator \mathbf{P} will of course be \bar{p} -invariant too, and so is \mathbf{P}' . This implies

that the resulting LMC with transition matrix \mathbf{P}' will contract towards \bar{p} at an exponential rate for all $t \geq T$, as stated in the following theorem.

Theorem 1. Given a \bar{p} -invariant QW (3), (4) with mixing time $\bar{\tau}(\epsilon_0)$ for some $\epsilon_0 \leq 1/4$, we can construct an LMC that has mixing time, for all $\epsilon > 0$,

$$\frac{\tau(\epsilon)}{\bar{\tau}(\epsilon_0)} \leq \left\lceil \frac{\ln 1/\epsilon}{\ln 1/(2\epsilon_0)} \right\rceil.$$

C. Tight conductance bound for QWs

Beyond the comparison with LMCs, the construction of Theorem 1 implies a general bound on the mixing performance of \bar{p} -invariant QWs. This tightens and generalizes the bounds of [27,42], which are restricted to uniform \bar{p} and unital quantum channels. The bound involves a quantity called the graph conductance $\Phi_{\bar{p}}$, which is a function of the graph topology and target distribution, capturing the bottlenecks that slow down mixing. Specifically, let \mathbf{P} be a Markov chain on \mathcal{V} with stationary distribution \bar{p} , and consider a partitioning of \mathcal{V} into two subsets \mathcal{X} and \mathcal{X}^c . If all the stationary population on \mathcal{X}^c is lost, the conductance $\Phi_{\mathcal{X}}(\mathbf{P})$ of \mathbf{P} with respect to \mathcal{X} counts which fraction of the remaining population $\bar{p}(\mathcal{X}) = \sum_{v \in \mathcal{X}} \bar{p}(v)$ jumps back to \mathcal{X}^c in one step:

$$\Phi_{\mathcal{X}}(\mathbf{P}) = \frac{1}{\bar{p}(\mathcal{X})} \sum_{v \in \mathcal{X}, v' \in \mathcal{X}^c} \mathbf{P}(v', v) \bar{p}(v).$$

The conductance $\Phi(\mathbf{P})$ of \mathbf{P} is defined as its conductance over the worst cut:

$$\Phi_{\bar{p}}(\mathbf{P}) = \min_{\mathcal{X}: 0 < \bar{p}(\mathcal{X}) \leq \frac{1}{2}} \Phi_{\mathcal{X}}(\mathbf{P}).$$

We define the graph conductance $\Phi_{\bar{p}}$ with respect to \mathcal{G} and \bar{p} as the maximal $\Phi_{\bar{p}}(\mathbf{P})$ over all Markov chains on \mathcal{G} that keep \bar{p} invariant.

The estimate $1/\Phi_{\bar{p}}$ is a well-known lower bound on the mixing time of any classical Markov chain; it is not tight however, as some graphs and \bar{p} do not admit a Markov chain matching the conductance bound. For instance, for converging to the uniform distribution on the cycle graph (see Sec. II), we have $1/\Phi_{\bar{p}} = N$, but the random walk can be shown to be optimal (by symmetry arguments) and needs $O(N^2)$ steps to converge. Conversely, [35] establishes a construction of LMCs that essentially saturates this bound; it does require solving a hard multicommodity flow problem over the entire graph. By exploiting the triangle inequality while computing marginal probabilities, we observe that the bound keeps holding for the marginal p_t of a \bar{p} -invariant LMC. Combining this with Theorem 1 provides a tight bound for the ultimately achievable mixing time of QWs.

Theorem 2. Any \bar{p} -invariant QW has a mixing time

$$\tau(1/4) \geq \frac{1}{8\Phi_{\bar{p}}}.$$

There exists such a \bar{p} -invariant QW with mixing time

$$\tau(\epsilon) \leq O\left(\frac{1}{\Phi_{\bar{p}}} \ln \frac{1}{\min_k \bar{p}_k} \ln \frac{1}{\epsilon}\right), \text{ for all } \epsilon > 0.$$

Proof. The lower bound is proven in detail in Theorem 5 for the generalized family of local stochastic processes. The existence result follows from the same existence result for lifted Markov chains, as formulated in [51] on the basis of [35]. Its validity for QWs follows by recognizing that lifted Markov chains are a special class of quantum channels. ■

V. EFFICIENT DESIGN OF FAST-MIXING QWs AND LMCs

Fast-mixing LMCs can often be built significantly simpler than with the general construction of Theorem 1. Accelerated mixing with QWs has been mostly demonstrated for graphs with strong symmetries, more specifically lattices [5,27–29,46,53]. Similarly to the circle example above, these QWs encode in coin values the lattice generators among which the walker can select its next move.

Remarkably, the same structure, with the same coin size as the QW, appears in a proposal of fast-mixing LMCs [44]. For a d -dimensional square lattice of size M , the coin takes $2d$ values of type \pm_k , with $k \in \{1, 2, \dots, d\}$ indicating the axis and \pm the direction of conditional motion among the nodes. At each step, the coin has a probability $\alpha = 1/(2dM)$ to switch to each of the other coin values, retaining a high probability $1 - (2d - 1)/(2dM)$ to keep the same generator. This dynamics precisely corresponds to a QW with diagonally dominant coin update \mathbf{C} that is projectively measured at each step, as in Eq. (1) with $q(t) = 1 \forall t$. For fixed dimension d , it also provides the same order of speedup as a QW with $q(t) \ll 1$ [29], and as the best possible QW according to Theorem 2, namely linear in M . Indeed, by counting the probabilities of applying to each lattice dimension, consecutively, the same ideas that yield fast mixing on the cycle, one obtains the following (possibly loose) bound [54]. The details of this counting argument can be found in the Appendix.

Theorem 3. The just described LMC on \mathbb{Z}_N^d has a mixing time $\tau(\epsilon) \leq O(Md^2 \ln(d) \ln(1/\epsilon))$.

Thus, QW and LMC have the same order of mixing time, the same structure, and they require the same graph knowledge for tuning α and/or q , namely the time $O(M)$ at which mixing will be considered accomplished. In particular, while the general construction of Theorem 1 requires a large memory in the associated LMC, we here have a construction where the LMC memory is no more than the “coin” memory of the QW.

VI. GENERAL THEORY: LOCAL STOCHASTIC PROCESSES

In the following we introduce the more abstractly defined family of “local stochastic processes,” containing both QWs and LMCs but also more general processes like the Cesaro average of a QW or a time-inhomogeneous Markov chain. We will prove Theorems 1 and 2 in this generalized setting. The main ideas of the proofs remain the same as explained above.

A. Local stochastic processes

A stochastic map Ψ over \mathcal{V} is a function that maps a probability distribution p_0 to another probability distribution p_1 ; it is linear and preserves both the positivity and the sum of the components of p_0 . A general stochastic process is a

family of stochastic linear maps Ψ_t , indexed by time $t \in \mathbb{N}$, which maps an initial probability distribution p_0 over \mathcal{V} to a probability distribution p_t over \mathcal{V} at each time t , defined by

$$p_t = \Psi_t[p_0].$$

We say that the family Ψ_t is *local* [27] with respect to a graph \mathcal{G} with nodes \mathcal{V} if and only if $\forall \mathcal{X} \subseteq \mathcal{V}$, $p_0, t \geq 0$ it holds that

$$p_{t+1}(\mathcal{X}) \leq p_t(\mathcal{X}) + p_t(\mathcal{B}(\mathcal{X})). \quad (6)$$

Here $\mathcal{B}(\mathcal{X})$ denotes the in-neighborhood of set \mathcal{X} :

$$\mathcal{B}(\mathcal{X}) = \{v \in \mathcal{V} \setminus \mathcal{X} : (v, v') \in \mathcal{E} \text{ for some } v' \in \mathcal{X}\}.$$

Formula (6) expresses the intuitive statement mentioned above in the proof sketch of Theorem 1: a set of nodes \mathcal{X} cannot contain more population at time $t + 1$ than the population at time t on itself and on its neighbors $\mathcal{B}(\mathcal{X})$. The family Ψ_t leaves a distribution \bar{p} invariant, or in short is \bar{p} -invariant, if $\Psi_t[\bar{p}] = \bar{p}$ for all $t \geq 0$. This ensures that the process does not perturb \bar{p} , at all times.

We can describe a local, \bar{p} -invariant quantum walk as a local and \bar{p} -invariant stochastic process. Thereto consider a QW $\Gamma[\rho] = \sum_k \mathbf{M}_k \rho \mathbf{M}_k^\dagger$, as introduced in (3), together with the linear initialization map $\mathbf{F}[p_0] = \rho_{p_0} = \sum_{v \in \mathcal{V}} p_0(v) |c_v, v\rangle \langle c_v, v|$ and marginalization map $f(\rho_t) = \text{diag}(\text{trace}_C(\rho_t))$. We can describe this QW as a stochastic process by setting

$$\Psi_t[e_v] = f(\Gamma^t[\rho_{e_v}]).$$

From linearity, it follows that $\Psi_t[p_0] = f(\Gamma^t[\rho_{p_0}]) = p_t$. If the QW is \bar{p} -invariant, so $f(\Gamma^t[\rho_{\bar{p}}]) = \bar{p}$ for all $t \geq 0$, then this stochastic process will also be \bar{p} -invariant: $\Psi_t[\bar{p}] = \bar{p}$ for all $t \geq 0$. In the Appendix, Lemma 4, we rigorously show that the two definitions of locality are equivalent: if Γ is local in the sense that the \mathbf{M}_k have zero entries where nodes are not connected in \mathcal{G} , then so is Ψ_t in the sense of Eq. (6).

By just developing the corresponding expressions, one can see immediately that also the Cesaro average output associated with a local and \bar{p} -invariant QW or LMC, namely $p_t = \frac{1}{t} \sum_{s=0}^{t-1} f(\Gamma^s[\rho_{p_0}])$, is described by a local and \bar{p} -invariant stochastic process. Of independent interest are time-inhomogeneous Markov chains \mathbf{P}_t that leave a common limit distribution invariant. We can cast these as a local \bar{p} -invariant process by setting $\Psi_t[p_0] = \mathbf{P}_t \dots \mathbf{P}_2 \mathbf{P}_1 p_0$.

B. Linear stochastic processes and LMCs have equivalent mixing performance

Theorem 1 essentially states that QW mixing can be simulated by an LMC, and the main steps of its proof are described for this setting in Sec. IV. Here we provide a formal proof for the more general statement: the mixing performance of any local and \bar{p} -invariant stochastic process can be simulated using a suitably constructed local LMC.

Simulability of stochastic linear maps. In the following, we will show that the p_t generated by a local stochastic map, starting from any given initial distribution p_0 , can always be simulated by a sequence of local stochastic transition matrices $\mathbf{P}_1^{(p_0)}, \mathbf{P}_2^{(p_0)}, \dots, \mathbf{P}_t^{(p_0)}$ on \mathcal{V} :

$$p_t = \mathbf{P}_t^{(p_0)} p_{t-1} = \mathbf{P}_t^{(p_0)} \dots \mathbf{P}_2^{(p_0)} \mathbf{P}_1^{(p_0)} p_0.$$

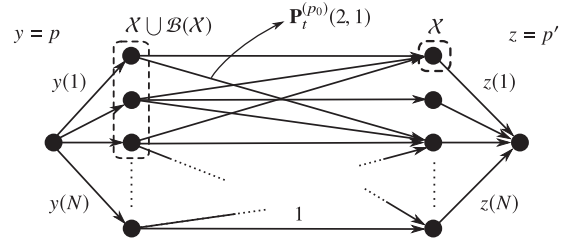


FIG. 3. Capacitated network construction used in Lemma 1.

This sequence will be in general dependent on the initial distribution p_0 . This observation and its proof are a generalization of the result by [39] from unitary evolution to abstract stochastic linear maps. The result builds on the following lemma.

Lemma 1. Consider two probability distribution p, p' over the nodes \mathcal{V} of a graph \mathcal{G} . If $\forall \mathcal{X} \subseteq \mathcal{V}$ it holds that

$$p'(\mathcal{X}) \leq p(\mathcal{X}) + p(\mathcal{B}(\mathcal{X}));$$

then there exists a local stochastic matrix \mathbf{P} such that $p' = \mathbf{P}p$.

Proof. Call $y = p$ and $z = p'$. In order to prove the above statement, it is convenient to resort to results concerning flows over capacitated networks [55], and, in particular, consider the graph shown in Fig. 3, where each edge is assigned a corresponding weight, or capacity. The network consists of a source node s , a sink node r , and two copies \mathcal{W} and \mathcal{W}' of the set of node states \mathcal{V} . Node s is connected with capacity $y(v)$ to any node $v \in \mathcal{W}$; any node $v \in \mathcal{W}$ is connected with capacity 1 to any node $v' \in \mathcal{W}'$ iff $(v, v') \in \mathcal{E}$, else the nodes are not connected; and any node $v' \in \mathcal{W}'$ is connected with capacity $z(v')$ to node r . The capacities $y(v)$ and $z(v')$, respectively from s and to r , thus reflect the probability distributions to be mapped. The key observation is the following: if this network can route a steady flow of value 1 from node s to node r , then the fraction from $v \in \mathcal{W}$ that is routed towards $v' \in \mathcal{W}'$ directly defines the entry $e_v^\dagger \mathbf{P}_t^{(p_0)} e_{v'}$ that we need, and also denoted $\mathbf{P}_t^{(p_0)}(v', v)$. Indeed, to route a flow of value 1 the edges from r to \mathcal{W} will have to be used to their full capacities $y(v)$, such that the flow through the edges from \mathcal{W}' to s becomes $z(v') = \sum_{v \in \mathcal{V}} \mathbf{P}_t^{(p_0)}(v', v) y(v)$, so we would have $\mathbf{P}_t^{(p_0)} y = z$ as claimed.

The max-flow min-cut theorem [55] states that the maximum steady flow which can be routed from node s to node r is equal to the minimum cut value of the graph, where a cut value is the sum of the capacities of a set of edges that disconnects s from r . It is clear that cutting all edges arriving at r disconnects the graph, with a cut value of 1, whereas cutting any middle edge between \mathcal{W} and \mathcal{W}' gives a cut value ≥ 1 . So the minimum cut should not include any of these “middle” edges, and it must be some combination of edges starting on s or arriving at r . Assume that we know the optimal cut, and let $\mathcal{X} \subseteq \mathcal{W}'$ such that the cut involves the edges from the complement of $\mathcal{X} \subseteq \mathcal{W}'$ to r . To block any flow from s to r while keeping all middle edges, we must then cut the edges from s to all the $l \in \mathcal{W}$ which have an edge to \mathcal{X} . This corresponds to all $l \in \mathcal{X} \cup \mathcal{B}(\mathcal{X})$. The value of this cut is thus

$$1 - z(\mathcal{X}) + y(\mathcal{X}) + y(\mathcal{B}(\mathcal{X})).$$

Recalling that $y = p_t$ and $z = p_{t+1}$, locality (6) imposes

$$z(\mathcal{X}) \leq y(\mathcal{X}) + y(\mathcal{B}(\mathcal{X})),$$

from which it follows that the minimum value of the cut is ≥ 1 . This minimum is attained (among others) with cutting all edges arriving at r , i.e., with \mathcal{X} the empty set. Hence, the minimum cut value is 1 and a solution \mathbf{P} to our problem exists. ■

By setting $p' = p_t$ and $p = p_{t-1}$, this directly leads to the following corollary.

Corollary 1 (local simulability). If Ψ_t is local, then for every pair (p_0, t) with $t > 0$ there exists a local stochastic matrix $\mathbf{P}_t^{(p_0)}$ such that $p_t = \mathbf{P}_t^{(p_0)} p_{t-1}$, where $p_t = \Psi_t[p_0]$.

Amplification lemma. Corollary 1 is instrumental in proving Theorem 1 for a finite time frame, by simulating the QW up to some given time. The following is instrumental to prove the theorem for arbitrary time, showing that a *finite-memory* process is sufficient to extend this mixing performance to arbitrarily small $\epsilon > 0$. In particular, we now show that, given a \bar{p} -invariant evolution map that mixes up to a certain total variation distance, we can iterate this map in order to mix to arbitrarily small distance, a process informally known as *amplification*. If the original process is \bar{p} -invariant, then necessarily the amplified process will also be \bar{p} -invariant.

Lemma 2 (amplification lemma). Assume that Ψ_t is a family of stochastic linear maps that mixes to an invariant \bar{p} , and admits a mixing time $\bar{\tau}(\epsilon)$ for all $\epsilon > 0$. Then for any $\epsilon < 1/2$, its amplified version defined as

$$\tilde{\Psi}_t = \Psi_{t \bmod T}(\Psi_T)^{\lfloor t/T \rfloor},$$

with $T = \bar{\tau}(\epsilon_0)$, is \bar{p} -invariant and has a mixing time $\tau(\epsilon) \leq \bar{\tau}(\epsilon_0) \lceil \ln(1/\epsilon) / \ln(1/(2\epsilon_0)) \rceil$ for all $\epsilon > 0$.

Proof. We will check that at any time $t \geq T \lceil \ln(1/\epsilon) / \ln(1/(2\epsilon_0)) \rceil = \kappa T$, $\kappa \in \mathbb{N}$, the total variation distance to \bar{p} is lower than ϵ . The proof uses invariance of \bar{p} under Ψ_t to transform $\Psi_t[p] - \bar{p}$ into $\Psi_t[p - \bar{p}]$. The \bar{p} invariance of the amplification is a direct consequence of the \bar{p} invariance of the original process.

For $t = \kappa T$, we get

$$\begin{aligned} & \max_p \|\tilde{\Psi}_{\kappa T}[p] - \bar{p}\|_{TV} \\ &= \max_p \|(\Psi_T)^\kappa[p] - \bar{p}\|_{TV} \\ &\leq \max_{p, p'} \|(\Psi_T)^\kappa[p] - (\Psi_T)^\kappa[p']\|_{TV} \\ &\leq (\max_{p, p'} \|\Psi_T[p] - \Psi_T[p']\|_{TV})^\kappa \end{aligned}$$

using the submultiplicativity of the total variation norm under any linear map in the form stated in [56]. Since $\max_{p, p'} \|\Psi_T[p] - \Psi_T[p']\|_{TV} \leq 2 \max_p \|\Psi_T[p] - \bar{p}\|_{TV} \leq 2\epsilon_0$, by our choice of $T = \bar{\tau}(\epsilon_0)$, we see that

$$\max_p \|\tilde{\Psi}_{\kappa T}[p] - \bar{p}\|_{TV} \leq (2\epsilon_0)^\kappa \leq \epsilon,$$

if $\kappa \geq \ln(1/\epsilon) / \ln(1/(2\epsilon_0))$.

For $t = t' + \kappa T$ with any $t' > 0$, we know that for all p

$$\begin{aligned} \|\Psi_t[p] - \bar{p}\|_{TV} &= \|\Psi_t[p - \bar{p}]\|_{TV} \\ &= \|\Psi_{t'}[\Psi_{\kappa T}[p - \bar{p}]]\|_{TV} \leq \|\Psi_{\kappa T}[p - \bar{p}]\|_{TV}, \end{aligned}$$

thanks to contractivity of the 1-norm under stochastic maps. So finally we find that, for arbitrary $t \geq 0$,

$$\max_p \|\tilde{\Psi}_t[p] - \bar{p}\|_{TV} \leq \max_p \|\tilde{\Psi}_{\lfloor t/T \rfloor T}[p] - \bar{p}\|_{TV} \leq \epsilon,$$

if $t \geq \bar{\tau}(\epsilon_0) \lceil \ln(1/\epsilon) / \ln(1/(2\epsilon_0)) \rceil$. ■

Proof of main theorem. We can now finalize the proof of a generalized form of Theorem 1, proving that for any local and \bar{p} -invariant stochastic process Ψ_t we can construct an LMC with a closely related mixing time.

Theorem 4 (Theorem 1, generalized version). Let Ψ_t be a stochastic linear map that mixes to some distribution \bar{p} with mixing time $\bar{\tau}(\epsilon)$, satisfying some locality constraint and leaving \bar{p} invariant. Then for any $\epsilon_0 < 1/2$ we can construct a \bar{p} -invariant LMC that satisfies the same locality constraint and that mixes to \bar{p} with a mixing time $\tau(\epsilon) = \bar{\tau}(\epsilon)$ for all $\epsilon \geq \epsilon_0$, and a mixing time

$$\tau(\epsilon) \leq \bar{\tau}(\epsilon_0) \left\lceil \frac{\ln 1/\epsilon}{\ln 1/(2\epsilon_0)} \right\rceil, \text{ for all } \epsilon > 0.$$

Proof. From Corollary 1 we know that for any $v \in \mathcal{V}$ we can construct local stochastic matrices $\mathbf{P}_t^{(v)}$ so that

$$p_t^{(v)} = \Psi_t[e_v] = \mathbf{P}_t^{(v)} \dots \mathbf{P}_2^{(v)} \mathbf{P}_1^{(v)} e_v.$$

Using these matrices, we can take over verbatim the construction discussed in Sec. IV. First we lift the node space to

$$\mathcal{C}' \times \mathcal{V} \equiv (\mathcal{V} \times \{0, 1, \dots, T\}) \times \mathcal{V}.$$

Then we construct the LMC $\mathbf{P}' = \mathbf{P}_F \mathbf{P}$ with

$$\mathbf{P} = \sum_{v \in \mathcal{V}} e_v e_v^\dagger \otimes \left(\sum_{t=0}^{T-1} e_{t+1} e_t^\dagger \otimes \mathbf{P}_t^{(v)} + e_T e_T^\dagger \otimes \mathbf{I}_\mathcal{V} \right) \quad (7)$$

and

$$\mathbf{P}_F = \sum_{t=0}^{T-1} \mathbf{I}_\mathcal{V} \otimes e_t e_t^\dagger \otimes \mathbf{I}_\mathcal{V} + \sum_{v \in \mathcal{V}} e_v e_{v_0}^\dagger \otimes e_0 e_T \otimes e_v e_v^\dagger, \quad (8)$$

reinitializing the coin state when the timer equals T . Note that the combination $\mathbf{P}_F \mathbf{P}$ is indeed a valid, local transition matrix on the lifted nodes, and for all $0 \leq t < T$ we see that indeed $\mathbf{P}' \mathbf{F}[e_v] = (e_v \otimes e_t) \otimes \Psi_t[e_v]$. For $t = T$ this gives

$$(\mathbf{P}')^T \mathbf{F}[e_v] = \mathbf{P}_F((e_v \otimes e_T) \otimes \Psi_T[e_v]) = \mathbf{F}[\Psi_T[e_v]],$$

so that the output $\Psi_T[e_v]$ gets reinitialized. For general t and p , we see that

$$f((\mathbf{P}')^t \mathbf{F}[p]) = \tilde{\Psi}_t,$$

where $\tilde{\Psi}_t$ is the amplified process defined in Lemma 2. Therefore, \mathbf{P}' exactly simulates the amplified process $\tilde{\Psi}_t$ for all $t \geq 0$. The \bar{p} invariance of the constructed LMC is a direct consequence of the \bar{p} invariance of the amplified process that it simulates. Since it is easy to check that both statements about the mixing time hold for the amplified process, they also hold for the LMC. ■

C. Tight conductance bound for linear stochastic processes

Thanks to Theorem 4 we can prove a conductance bound for linear stochastic processes, and quantum walks in particular, by proving such a bound for the simulating LMC.

It proves a tightening of the bound for LMCs formulated in for instance [35] for LMCs by relaxing their assumptions on two points: (i) the LMC does not start from an arbitrary distribution over subnodes, but it can be initialized on the lifted space with some designed local map \mathbf{F} , reflecting typical algorithm initialization, and (ii) we require convergence to a limit distribution only on the *marginal* over \mathcal{V} , while standard literature considers convergence on the whole lifted space.

We will need the concept of an *induced Markov chain*, which can be seen as the projection of an LMC back to the original graph; see [57]. Let \mathbf{P} be an irreducible lifted Markov chain on the nodes of a lifted graph $\mathcal{C} \times \mathcal{V}$, having stationary distribution \hat{p} . The induced chain $\mathbf{P}_\mathcal{V}$ over \mathcal{V} is defined by

$$\mathbf{P}_\mathcal{V}(v', v) = \sum_{c, c' \in \mathcal{C}} \frac{\hat{p}(c, v)}{\bar{p}(v)} \mathbf{P}((c', v'), (c, v)),$$

where \bar{p} is defined by $\bar{p}(v) = \hat{p}(\mathcal{C} \times v)$. It is easy to see that $\mathbf{P}_\mathcal{V} \bar{p} = \bar{p}$. This definition is motivated by obtaining matching ergodic flows. We define the ergodic flow from \mathcal{X} to its complement by $\mathbf{Q}_\mathbf{P}(\mathcal{X}^c, \mathcal{X}) = \sum_{v \in \mathcal{X}, v' \in \mathcal{X}^c} \mathbf{P}(v', v) \bar{p}(v)$. By the definition of the induced chain, it then holds that $\mathbf{Q}_{\mathbf{P}_\mathcal{V}}(v', v) = \mathbf{Q}_\mathbf{P}(\mathcal{C} \times v', \mathcal{C} \times v)$. Since we can rewrite $\Phi_\mathcal{X}(\mathbf{P}) = \mathbf{Q}_\mathbf{P}(\mathcal{X}^c, \mathcal{X}) / \bar{p}(\mathcal{X})$, and $\bar{p}(\mathcal{X}) = \hat{p}(\mathcal{C} \times \mathcal{X})$, we see that

$$\Phi_\mathcal{X}(\mathbf{P}_\mathcal{V}) = \Phi_{\mathcal{C} \times \mathcal{X}}(\mathbf{P}). \quad (9)$$

This readily implies that $\Phi(\mathbf{P}_\mathcal{V}) \geq \Phi(\mathbf{P})$. We also have that $\Phi_{\bar{p}} \geq \Phi(\mathbf{P}_\mathcal{V})$, with $\Phi_{\bar{p}}$ the graph conductance associated with \bar{p} on the nonlifted graph. Indeed by definition $\mathbf{P}_\mathcal{V}$ obeys the graph locality and we saw that $\mathbf{P}_\mathcal{V} \bar{p} = \bar{p}$, so it is an element of the set over which the graph conductance $\Phi_{\bar{p}}$ is maximized. We next borrow standard techniques, as presented in for instance [56,58], to prove two instrumental lemmas.

Lemma 3. Consider an irreducible Markov chain \mathbf{P} over a set \mathcal{V} , with unique stationary distribution \bar{p} . Then

$$(\mathbf{P}^t \bar{p}_\mathcal{X})(\mathcal{X}^c) \leq t \Phi_\mathcal{X}(\mathbf{P}) \quad \text{for all } \mathcal{X} \subseteq \mathcal{V}, t \geq 0,$$

where $\bar{p}_\mathcal{X}(v) = \bar{p}(v) / \bar{p}(\mathcal{X})$, for $v \in \mathcal{X}$, and zero elsewhere.

Proof. This follows by noting that

$$(\mathbf{P}^t \bar{p}_\mathcal{X})(\mathcal{X}^c) = \sum_{v \in \mathcal{X}, v' \in \mathcal{X}^c} \frac{\mathbf{P}^t(v', v) \bar{p}(v)}{\bar{p}(\mathcal{X})} = \Phi_\mathcal{X}(\mathbf{P}^t),$$

and using Eq. (5.1) in [58] stating that $\Phi_\mathcal{X}(\mathbf{P}^t) \leq t \Phi_\mathcal{X}(\mathbf{P})$. Their proof assumes irreducibility and aperiodicity, yet the aperiodicity condition is irrelevant. ■

Theorem 5 (Theorem 2, generalized version). Any local and \bar{p} -invariant stochastic linear process has a mixing time $\tau(1/4) \geq 1/(8\Phi_{\bar{p}})$. As a consequence, any \bar{p} -invariant QW has a mixing time $\tau(1/4) \geq 1/(8\Phi_{\bar{p}})$.

Proof. Choosing $\epsilon_0 = 1/4$ in Theorem 4 shows that we can construct a local LMC whose 1/4-mixing time also equals $\tau(1/4)$. We call this LMC \mathbf{P}' , as defined in Eqs. (7) and (8), with $T = \tau(\epsilon_0)$. We will bound the mixing time of this LMC.

We will call \hat{p} the stationary distribution of \mathbf{P}' , keeping in mind that $f(\hat{p}) = \bar{p}$. First we consider a general initial distribution p over the lifted node set $\mathcal{C} \times \mathcal{V}$. With f the marginalization map, we can apply the reverse triangle in-

equality to see that

$$\begin{aligned} \|\mathbf{P}'^t p - \hat{p}\|_{TV} &= \frac{1}{2} \sum_{(c,v) \in \mathcal{C} \times \mathcal{V}} |(\mathbf{P}'^t p)(c, v) - \hat{p}(c, v)| \\ &\geq \frac{1}{2} \sum_{v \in \mathcal{V}} \left| \sum_{c \in \mathcal{C}} (\mathbf{P}'^t p)(c, v) - \hat{p}(c, v) \right| \\ &= \|f(\mathbf{P}'^t p) - f(\hat{p})\|_{TV}. \end{aligned}$$

Now take a subset $\mathcal{X} \subseteq \mathcal{V}$ such that $\hat{p}(\mathcal{C} \times \mathcal{X}) \leq 1/2$. Using this subset we define a second marginalization $g_\mathcal{X}$ mapping distributions over the nodes of \mathcal{V} to the binary property $\{v \in \mathcal{X}\}$ or $\{v \notin \mathcal{X}\}$; i.e., $g_\mathcal{X}(p)$ can be represented as a vector $[\sum_{v \in \mathcal{X}} p(v); \sum_{v \notin \mathcal{X}} p(v)]$. By a similar reasoning we get

$$\|f(\mathbf{P}'^t p) - f(\hat{p})\|_{TV} \geq \|g_\mathcal{X}(f(\mathbf{P}'^t p)) - g_\mathcal{X}(f(\hat{p}))\|_{TV}.$$

If we take $p = \hat{p}_{\mathcal{C} \times \mathcal{X}}$ as defined in Lemma 3 and we use the triangle inequality again, it follows that

$$\begin{aligned} &\|g_\mathcal{X}(f(\mathbf{P}'^t \hat{p}_{\mathcal{C} \times \mathcal{X}})) - g_\mathcal{X}(f(\hat{p}))\|_{TV} \\ &\geq \|g_\mathcal{X}(f(\hat{p}_{\mathcal{C} \times \mathcal{X}})) - g_\mathcal{X}(f(\hat{p}))\|_{TV} \\ &\quad - \|g_\mathcal{X}(f(\hat{p}_{\mathcal{C} \times \mathcal{X}})) - g_\mathcal{X}(f(\hat{\mathbf{P}}^t \hat{p}_{\mathcal{C} \times \mathcal{X}}))\|_{TV} \\ &\geq \frac{1}{2} - \mathbb{P}_{g_\mathcal{X} \hat{p}_{\mathcal{C} \times \mathcal{X}}}(\mathcal{C} \times \mathcal{X}^c) \geq \frac{1}{2} - t \Phi_{\mathcal{C} \times \mathcal{X}}(\mathbf{P}') \\ &= \frac{1}{2} - t \Phi_\mathcal{X}(\mathbf{P}_\mathcal{V}'). \end{aligned}$$

For the second inequality we have used that $\mathbb{P}_{g_\mathcal{X}(f(\hat{p}_{\mathcal{C} \times \mathcal{X}}))}(x) = 1$, while $\mathbb{P}_{g_\mathcal{X}(f(\hat{p}))}(x) \leq 1/2$; this ensures that the first term is $\geq 1/2$, while the second term is exactly the probability to be in \mathcal{X}^c . The last inequalities follow from Lemma 3 and Eq. (9). We thus find altogether that

$$\|f(\mathbf{P}'^t p) - \bar{p}\|_{TV} \geq \frac{1}{2} - t \Phi_\mathcal{X}(\mathbf{P}_\mathcal{V}'). \quad (10)$$

Now by definition of the (1/4)-mixing time $\tau(1/4)$ of \mathbf{P}' (recall that this is defined with respect to the subset of correctly initialized distribution $\mathbf{F}[p_0]$), we know that for any p_0 over the original node set \mathcal{V} it holds that

$$\|f(\mathbf{P}'^{\tau(1/4)} \mathbf{F}[p_0]) - \bar{p}\|_{TV} \leq 1/4.$$

We can use this to prove that $\|f(\mathbf{P}'^{2\tau(1/4)} p) - \bar{p}\|_{TV} \leq 1/4$ for any p over $\mathcal{C} \times \mathcal{V}$. To see this, note that if $p = e_v \otimes e_t \otimes e_{v'}$ for some arbitrary v, t, v' , then by construction of \mathbf{P}' it holds that $\mathbf{P}'^{T-t} p = \mathbf{F}[p']$ for some p' . Therefore, for all $t' \geq \tau(1/4)$ we find that

$$\|f(\mathbf{P}'^{t'+\tau(1/4)-t} p) - \bar{p}\|_{TV} = \|f(\mathbf{P}'^{t'} \mathbf{F}[p']) - \bar{p}\|_{TV} \leq 1/4.$$

Since $\tau(1/4) - t \leq \tau(1/4)$ this shows that $\|f(\mathbf{P}'^{2\tau(1/4)} p) - \bar{p}\|_{TV} \leq 1/4$ for any $p = e_v \otimes e_t \otimes e_{v'}$. For a general p over $\mathcal{C} \times \mathcal{V}$ we can simply apply the triangle inequality to prove that indeed $\|f(\mathbf{P}'^{2\tau(1/4)} p) - \bar{p}\|_{TV} \leq 1/4$. Combined with Eq. (10) shows that

$$\frac{1}{2} - 2\tau(1/4)\Phi_\mathcal{X}(\mathbf{P}_\mathcal{V}') \leq 1/4,$$

and so $\tau(1/4) \geq 1/[8\Phi_\mathcal{X}(\mathbf{P}_\mathcal{V}')]$. Minimizing $\Phi_\mathcal{X}$ over \mathcal{X} shows that $\tau(1/4) \geq 1/[8\Phi(\mathbf{P}_\mathcal{V})]$. The fact that $\Phi_{\bar{p}} \geq \Phi(\mathbf{P}_\mathcal{V})$ was already discussed after Eq. (9). ■

VII. CONCLUSION

In summary, we clarify that the discrete-time QWs on graphs proposed in the literature induce non-Markovian local processes whose mixing behavior can be simulated by LMCs (Theorem 1). The construction can be extended to abstract local stochastic dynamics beyond the QW model (Theorem 4). Consequently, the hierarchy LMCs \subseteq QWs \subseteq {general local processes} collapses in terms of mixing, not only regarding ultimate speedup achievable on general graphs (Theorem 2), but also regarding paradigmatic cases for which efficient designs are known (lattices, Theorem 3). In this light, at least for those QW proposals, a mixing speedup is not diagnostic of underlying quantum dynamics, but potentially just a memory effect. This prompts the question of where there remains room for a “quantum advantage” at all. Our analysis does suggest that this is not the end of the story. While \bar{p} invariance holds and stabilizes the system in typical QW proposals, it does not hold in some mixing-related applications, like simulated annealing. This distinction may be important as, without \bar{p} invariance, the conductance bound of Theorem 2 could be broken significantly [51]. We have found no direct way to exploit such design ideas in a “realistic” way for QW mixing, so the question—including whether dropping \bar{p} invariance in this context is acceptable—remains currently open. Another point is the design of simple yet fast-mixing QWs on graphs for which, unlike lattices, known simple LMCs do not meet the conductance bound. Our results suggest the importance of checking whether such QWs would have an intrinsic advantage, or be a proxy towards efficient LMC design. Furthermore, the QW of Eq. (1), taking $\alpha = 1/2$ and $q = 1/N$, mixes approximately over the t nodes closest to its starting node, for any number of iterations $t < N$ [5]. Such “multiscale” mixing cannot be achieved with the simple LMC of Eq. (2), where tuning $\alpha = 1/N$ to have good mixing at $t = N$ implies almost-deterministic motion for $t \ll N$. This could point to efficient quantum speedups for tasks related to mixing, yet not directly reducible to it.

ACKNOWLEDGMENTS

The authors want to thank Giuseppe Vallone and Lorenza Viola for valuable suggestions and comments on earlier versions of the manuscript.

APPENDIX

1. EXAMPLES OF \bar{p} INVARIANCE

One important class of QWs are the unital quantum channels. Such channels take the form $\Gamma_t[\rho] = \sum q_l(t) \mathbf{U}_l \rho \mathbf{U}_l^\dagger$ with $\sum q_l(t) = 1$ and $\{\mathbf{U}_l\}$ unitary transition operators, and converge to the fully mixed state $\bar{\rho} = \frac{1}{|\mathcal{C}||\mathcal{V}|} \mathbf{I}_N \otimes \mathbf{I}_C$. Unital channels arise for instance when taking into account a measurement mechanism [29,49], as in Eq. (1), or in the study of quantum trajectories and open quantum walks [48]. These channels cover the largest part of the current literature on QW mixing; see [47,59] for surveys. For unital channels, to ensure \bar{p} invariance and thus to be able to apply our results, one

simple option is to consider the initialization

$$\mathbf{F}[p_0] = \sum_{v \in \mathcal{V}} p_0(v) \frac{1}{|\mathcal{C}|} \mathbf{I}_C \otimes |v\rangle \langle v|,$$

i.e., locally mapping each node v to the fully mixed state $\frac{1}{|\mathcal{C}|} \mathbf{I}_C$ over the coin space.

More general “mixing quantum processes” are obtained by (possibly time-varying) sequences of quantum channels [60,61] which converge to a fixed point $\bar{\rho}$ over the full Hilbert space:

$$\lim_{t \rightarrow \infty} \Gamma_t \Gamma_{t-1} \dots \Gamma_1[\rho] = \bar{\rho}, \quad \forall \rho, \quad \Gamma_t[\bar{\rho}] = \bar{\rho}, \quad \forall t \geq 0.$$

For such maps \bar{p} invariance always holds if the QW initialization maps \bar{p} to $\bar{\rho}$. It is easy to see that this can always be achieved with an initialization map \mathbf{F} of the form (4), provided the channel is also so-called “decohering” [62]; i.e., $\bar{\rho}$ is block-diagonal in the node basis:

$$\bar{\rho} = \sum_{v \in \mathcal{V}} \bar{p}(v) \bar{\rho}_v \otimes |v\rangle \langle v|.$$

Indeed it then suffices to specify

$$\mathbf{F}[p_0] = \sum_{v \in \mathcal{V}} p_0(v) \bar{\rho}_v \otimes |v\rangle \langle v|.$$

Finally, in the mixing context a large class of QW proposals, if not all, have considered vertex-transitive graphs (e.g., Cayley graphs such as the cycle and the lattice). In this case, any initialization and quantum walk that maintains graph symmetry will automatically be \bar{p} -invariant. On the cycle example of Sec. II for instance, this amounts to assuming that we admit any initialization that attributes the same initial coin state to each node $v \in \mathcal{V}$.

For completeness, let us briefly give some (academic) examples of QWs that would not satisfy \bar{p} invariance. As a trivial example, one could consider a vertex-transitive QW with an initialization map that breaks this symmetry. Consider for instance the QW example on the cycle graph from Sec. II where initially, coin state $|+\rangle$ would be attributed to node 1 and $|-\rangle$ to node 3; this would not preserve the stationary population on node 2. There can however be more fundamental obstacles. Indeed, we can for instance have a problem to map \bar{p} to $\bar{\rho}$ with an initialization of type (4), when the fixed point $\bar{\rho}$ of a mixing quantum channel is nondiagonal in the node basis, as occurs in many thermal equilibrating systems [63]. An (artificial) example of such a quantum channel on the two-node graph, with no additional coin space, is the following:

$$\Gamma[\rho] = |\varphi_+\rangle \langle \varphi_+ | \rho | \varphi_+\rangle \langle \varphi_+ | + p |\varphi_+\rangle \langle \varphi_- | \rho | \varphi_-\rangle \langle \varphi_+ | + (1-p) |1\rangle \langle \varphi_- | \rho | \varphi_-\rangle \langle 1|, \quad (\text{A1})$$

where $|\varphi_\pm\rangle = \frac{|0\rangle \pm |1\rangle}{\sqrt{2}}$, with $|0\rangle$ and $|1\rangle$ corresponding to the two graph nodes. It is easy to check that $\Gamma^t[\rho] \xrightarrow{t \rightarrow \infty} \bar{\rho} = |\varphi_+\rangle \langle \varphi_+ |$, $\forall \rho$, and $\Gamma[\bar{\rho}] = \bar{\rho}$. The associated limit distribution $\bar{p} = f(\bar{\rho})$ is the uniform distribution over the two nodes. Since there is no coin, any initialization of the form (4) is defined by $\mathbf{F}[p_0] = p_0(0) |0\rangle \langle 0| + p_0(1) |1\rangle \langle 1|$. It is clear that this gives $f(\Gamma[\mathbf{F}[\bar{p}]]) \neq \bar{p}$, so that this channel is not \bar{p} -invariant. Our results do allow us to simulate the behavior

of a QW like (A1) with a classical stochastic process, for any given initial state and finite time interval. However, in the absence of \bar{p} invariance with the algorithm initialization (4), our conductance bound on mixing time does not apply. One could conjecture that attaining \bar{p} via a coherent superposition \bar{p} should be harder, and hence slower, than attaining a (block-)diagonal mixed \bar{p} ; but rigorously, we currently do not know whether channels like (A1) can give a speedup for mixing purposes. Our results at least identify that such a relaxation of \bar{p} invariance would be *necessary* in order to break the bounds that we present next.

2. Equivalence of “locality” definitions for quantum channels

For any subset $\mathcal{X} \subseteq \mathcal{V}$ and density matrix ρ , we introduce the notation $\mathbb{P}_\rho(\mathcal{X}) = \text{trace}(\cdot_{\mathcal{X}}\rho)$, where $\cdot_{\mathcal{X}} = \mathbf{I} \otimes \sum_{v \in \mathcal{X}} |v\rangle\langle v|$ is the projector onto the subspace associated with the subset of nodes \mathcal{X} of the original graph \mathcal{G} . Upon measuring in the node basis, $\mathbb{P}_\rho(\mathcal{X})$ is the probability of retrieving a node in \mathcal{X} .

Lemma 4. Let Γ be a quantum channel. The following statements are equivalent:

(a) $\forall c, c' \in \mathcal{C}, v, v' \in \mathcal{V}$, it holds: if $(v, v') \notin \mathcal{E}$ then $\langle c', v' | \mathbf{M}_l | c, v \rangle = 0 \forall l$.

(b) For all $\mathcal{X} \subseteq \mathcal{V}$ and $\rho \in \mathcal{D}(\mathcal{H}_{\mathcal{C} \times \mathcal{V}})$, it holds that $\mathbb{P}_{\Gamma[\rho]}(\mathcal{X}) \leq \mathbb{P}_\rho(\mathcal{X}) + \mathbb{P}_\rho(\mathcal{B}(\mathcal{X}))$.

Proof. “(a) \Rightarrow (b)”: We will show that the inequality in (b) holds for a one-dimensional projection $\rho = |\psi\rangle\langle\psi|$, $|\psi\rangle \in \mathcal{H}_{\mathcal{C} \times \mathcal{V}}$; due to linearity of the involved operators in ρ , the inequality must then necessarily hold also for all density operators $\rho \in \mathcal{D}(\mathcal{H}_{\mathcal{C} \times \mathcal{V}})$, being convex combinations of projections. We can write

$$\begin{aligned} \mathbb{P}_{\Gamma[|\psi\rangle\langle\psi|]}(\mathcal{X}) &= \sum_{c' \in \mathcal{C}, v' \in \mathcal{X}} \langle c', v' | \Gamma[|\psi\rangle\langle\psi|] | c', v' \rangle \\ &= \sum_{c' \in \mathcal{C}, v' \in \mathcal{X}} \sum_l \langle c', v' | \mathbf{M}_l | \psi \rangle \langle \psi | \mathbf{M}_l^\dagger | c', v' \rangle \\ &= \sum_{c' \in \mathcal{C}, v' \in \mathcal{X}} \sum_l |\langle c', v' | \mathbf{M}_l | \psi \rangle|^2. \end{aligned}$$

Since we assume that (a) holds, we have that $\forall l, \langle c', v' | \mathbf{M}_l | c, v \rangle = 0$ for $v' \in \mathcal{X}$ and $v \in \mathcal{X}^c \setminus \mathcal{B}(\mathcal{X})$, where $\mathcal{X}^c = \mathcal{V} \setminus \mathcal{X}$. If we now write $|\psi\rangle = |\psi_{\mathcal{X}}\rangle + |\psi_{\mathcal{B}(\mathcal{X})}\rangle + |\psi_{\mathcal{X}^c \setminus \mathcal{B}(\mathcal{X})}\rangle$, where $|\psi_{\mathcal{Y}}\rangle \equiv \cdot_{\mathcal{Y}} |\psi\rangle$ for any $\mathcal{Y} \subseteq \mathcal{V}$, then (a), in particular, implies that $\forall c' \in \mathcal{C}, v' \in \mathcal{X}$ and all l , we have $\langle c', v' | \mathbf{M}_l | \psi_{\mathcal{X}^c \setminus \mathcal{B}(\mathcal{X})}\rangle = 0$. Intuitively, this expresses that $|\psi_{\mathcal{X}^c \setminus \mathcal{B}(\mathcal{X})}\rangle$ does not contribute to the probability of observing $\Pi_{\mathcal{X}}$ after the action of Γ . Inserting $\langle c', v' | \mathbf{M}_l | \psi_{\mathcal{X}^c \setminus \mathcal{B}(\mathcal{X})}\rangle = 0$ into the above sum, we thus get

$$\begin{aligned} \mathbb{P}_{\Gamma[|\psi\rangle\langle\psi|]}(\mathcal{X}) &= \mathbb{P}_{\Gamma[(|\psi_{\mathcal{X}}\rangle + |\psi_{\mathcal{B}(\mathcal{X})}\rangle)(\langle\psi_{\mathcal{X}}| + \langle\psi_{\mathcal{B}(\mathcal{X})}|)]}(\mathcal{X}) \\ &\leq \text{trace}(\Gamma[(|\psi_{\mathcal{X}}\rangle + |\psi_{\mathcal{B}(\mathcal{X})}\rangle)(\langle\psi_{\mathcal{X}}| + \langle\psi_{\mathcal{B}(\mathcal{X})}|)]) \\ &= \text{trace}((|\psi_{\mathcal{X}}\rangle + |\psi_{\mathcal{B}(\mathcal{X})}\rangle)(\langle\psi_{\mathcal{X}}| + \langle\psi_{\mathcal{B}(\mathcal{X})}|)) \\ &= \text{trace}(|\psi_{\mathcal{X}}\rangle\langle\psi_{\mathcal{X}}|) + \text{trace}(|\psi_{\mathcal{B}(\mathcal{X})}\rangle\langle\psi_{\mathcal{B}(\mathcal{X})}|) \\ &= \mathbb{P}_{|\psi_{\mathcal{X}}\rangle\langle\psi_{\mathcal{X}}|}(\mathcal{X}) + \mathbb{P}_{|\psi_{\mathcal{B}(\mathcal{X})}\rangle\langle\psi_{\mathcal{B}(\mathcal{X})}|}(\mathcal{B}(\mathcal{X})), \end{aligned}$$

where the second and third inequalities follow from the fact that Γ is trace-preserving and $|\psi_{\mathcal{X}}\rangle$ is orthogonal to $|\psi_{\mathcal{B}(\mathcal{X})}\rangle$.

“(b) \Rightarrow (a)”: Assume that (a) does not hold. Thus, there exists some l , some $c, c' \in \mathcal{C}$, some $v' \in \mathcal{X}$ and $v \in \mathcal{X}^c \setminus \mathcal{B}(\mathcal{X})$ such that $\langle c', v' | \mathbf{M}_l | c, v \rangle \neq 0$. If we now consider $|\psi\rangle = |c, v\rangle$, then $\mathbb{P}_{|c, v\rangle\langle c, v|}(\mathcal{X}) + \mathbb{P}_{|c, v\rangle\langle c, v|}(\mathcal{B}(\mathcal{X})) = 0$, whereas $\mathbb{P}_{\Gamma[|c, v\rangle\langle c, v|]}(\mathcal{X}) \geq \mathbb{P}_{\Gamma[|c, v\rangle\langle c, v|]}(\langle c', v' |) = |\langle c', v' | \mathbf{M}_l | c, v \rangle|^2 > 0$. So (b) does not hold when (a) does not; thus conversely, if (b) holds then (a) must hold too. ■

Based on this lemma, we will call a quantum channel *local* with respect to a reference lifted graph if (a), or equivalently (b), holds; and from (b) thus, the associated Ψ_t will be local in the sense of (6) too.

3. Proof of Theorem 3 for quantum walks on lattices

Our purpose is just to provide a simple, possibly intuitive argument showing that simple LMCs on lattices have the same order of mixing time as the corresponding QW. In this sense, we must note that the following estimate of mixing time could be tightened at places (as the algorithm might be somewhat improved with little modifications); we do not further insist on this issue since optimizing LMCs is not the purpose of this paper.

Consider a d -dimensional periodic lattice \mathbb{Z}_M^d of side M , encoded in a graph with node set $\mathcal{V} = \{(i_1, i_2, \dots, i_d) | 1 \leq i_k \leq M, \forall k\}$. Similarly to the QW/LMC construction for the cycle, we lift this graph by adding a set of coin states $\mathcal{C} = \{+k, -k | 1 \leq k \leq d\}$. An LMC on this graph thus takes place on the vector space $\mathcal{H}_{\mathcal{C}} \otimes \mathcal{H}_{\mathcal{V}} = \text{span}\{e_{c,v} | (c, v) \in \mathcal{C} \times \mathcal{V}\}$. With operator \mathbf{P}_k^\pm defined on $\mathcal{H}_{\mathcal{V}}$ as the cyclic permutation of the k th dimension, that is, $\mathbf{P}_k^\pm e_{\dots i_{k-1}, i_k, i_{k+1} \dots} = e_{\dots i_{k-1}, (i_k \pm 1) \bmod M, i_{k+1} \dots}$ for all $1 \leq k \leq d$, the LMC defined in [44] is written

$$\mathbf{P} = \left(\sum_k e_{+k} e_{+k}^\dagger \otimes \mathbf{P}_k^+ + e_{-k} e_{-k}^\dagger \otimes \mathbf{P}_k^- \right) (\mathbf{S} \otimes \mathbf{I}_{\mathcal{V}}), \quad (\text{A2})$$

recalling that $\mathbf{I}_{\mathcal{V}}$ is the identity matrix on $\mathcal{H}_{\mathcal{V}}$. We now specifically select $\mathbf{S} = (1 - 2d\alpha)\mathbf{I}_{\mathcal{C}} + \alpha\mathbf{1}$, with $\alpha = 1/(2dM)$ and $\mathbf{1}$ the all-ones matrix.

We will prove the mixing time for M odd. For M even, the LMC shows a parity problem: starting from a single state, at any given time the walk will be supported only on the even or only on the odd nodes. This is easily remedied by for instance modifying \mathbf{P} to $(\mathbf{P} + \mathbf{I}_{\mathcal{C} \times \mathcal{V}})/2$, which changes the mixing time only by a constant factor, or by randomizing the parity of the initial state. To facilitate its reading, we again structure the proof using two technical lemmas.

Lemma 5. Assume that the LMC in Eq. (A2) starts with any $p_0 \in \{e_{(c,v)}\}$, i.e., with all its weight concentrated on a single vertex and single coin choice $c \in \{+k, -k\}$ for some k . Then the resulting distribution after $2M$ time steps has uniformly mixed i_k with a probability $\geq 1/(16d)$, in the sense that $\mathbb{P}_{p_{2M}}(i_k = n) \geq 1/(16dM)$ for all $n \in \{1, 2, \dots, M\}$.

Proof. By symmetry, we can consider without loss of generality that the initial distribution is

$$p_0 = e_{+,1,\dots,1}.$$

From the structure of \mathbf{S} we see that at each time step with probability $2d\alpha = 1/M$ a coin toss takes place. Then the probability of a single coin toss happening over $2M$ steps is

given by

$$\begin{aligned}\mathbb{P}_{2M \text{ steps}}(1 \text{ coin toss}) &= \binom{2M}{1} \frac{1}{M} \left(1 - \frac{1}{M}\right)^{2M-1} \\ &= 2 \left(1 - \frac{1}{M}\right)^{2M-1} \geq \frac{1}{8} \quad \text{for } M \geq 2,\end{aligned}$$

where the inequality follows from the fact that $(1 - 1/M)^M$ is an increasing function of M , going from $1/4$ for $M = 2$ to $1/e$ for M large. From this, the event E_1 that a single coin toss takes place and switches the coin state from $+1$ to -1 , occurs with probability

$$\mathbb{P}_{2M \text{ steps}}(E_1) = \frac{1}{2d} \mathbb{P}_{2M \text{ steps}}(1 \text{ coin toss}) \geq \frac{1}{16d}.$$

When E_1 holds true with the single coin toss at time $T \in [1, 2M]$, the distribution at time $2M$ equals

$$\begin{aligned}p_{2M} &= e_{-1, 1+(T-1)-(2M-T+1)\bmod M, 1, \dots, 1} \\ &= e_{-1, (2T-1)\bmod M, 1, \dots, 1}.\end{aligned}$$

Yet, conditional on the fact that E_1 holds true, the timing T of the single coin toss is uniformly distributed between 1 and $2M$. As a consequence, $2T - 1$ is uniformly distributed over $1, 3, 5, \dots, M, 2, 4, 6, \dots, M - 1, 1, 3, 5, \dots, M, 2, 4, 6, \dots, M - 1$, i.e., effectively over the integers from 1 to M . Thus

$$\begin{aligned}\mathbb{P}_{p_{2M}}(i_k = n) &\geq \mathbb{P}_{E_1}(i_k = n) \mathbb{P}_{2M \text{ steps}}(E_1) \\ &= \frac{1}{16d} \frac{1}{M} \quad \forall n \in \{1, 2, \dots, M\},\end{aligned}$$

which proves the statement. \blacksquare

Our application of Lemma 5 to prove the following result is loose. Its sequential use, coordinate by coordinate, leaves further room for improvement, and this is why we think that it should be possible to win a factor d on the mixing time. However, the resulting estimate is sufficient for establishing the main message of Sec. V.

Lemma 6. Consider the LMC defined in Eq. (A2) on \mathbb{Z}_M^d , with M odd. For any initial distribution p_0 , the distribution p_T after $T = 3Md(d \ln(d) + d)$ steps satisfies

$$p_T = \mathbf{P}^T p_0 = q\bar{p} + (1 - q) \left(\frac{p_T - q\bar{p}}{1 - q} \right), \quad (\text{A3})$$

with $\tilde{p}_T = (p_T - q\bar{p})/(1 - q)$ a positive distribution; \bar{p} the stationary distribution of \mathbf{P} , which is the uniform distribution over $\mathcal{C} \times \mathcal{V}$; and $q = (1 - 1/e)/2$ where $e = \exp(1)$.

Proof. We consider time intervals of $3M$ steps, which we analyze as follows:

(1) In the first M steps: As in Lemma 5, we say that at each time step with probability $2d\alpha = 1/M$ a coin toss takes place. Now, we use that this completely randomizes the coin state. The probability that no such coin toss has happened after M steps is $(1 - 1/M)^M \leq 1/e$.

(2) In the next $2M$ steps: By Lemma 5 the coordinate corresponding to the randomized coin state is mixed uniformly with a probability $\geq 1/16d$. If that coordinate was already in a more mixed state than in the hypothesis of Lemma 5, then the resulting mixing can only be better.

Each interval of $3M$ steps will thus uniformly mix a random coordinate with probability $r \geq 1/(16ed)$. Using Cantelli's inequality for a binomial process with success probability r , we find the following bound for the number of successful mixing episodes k :

$$\mathbb{P}_{2l/r \text{ iterations}}(k \geq l) \geq 1/2.$$

So if we go through $[d \ln(d) + d]32ed$ such intervals of $3M$ steps, then with a probability $1/2$ we will have mixed $d \ln(d) + d$ coordinates; the latter are chosen randomly according to independent uniform processes with repetition. According to the coupon collector's problem, $d \ln(d) + d$ random choices selects all coordinates with a probability $(1 - 1/e)$. This implies that we can bound the state after $T = 3M[d \ln(d) + d]32ed$ steps as $p_T \geq q\pi$, with $q = (1 - 1/e)/2$. \blacksquare

We now have all the pieces to prove the actual result.

Proof of Theorem 3. We have shown in Lemma 6 that for any initial distribution p_0 over \mathcal{V} , with a fixed probability q the state will be uniformly mixed after $T = 3Md(d \ln(d) + d)$ steps; i.e., p_T will be of the form Eq. (A3). Then after $2T$ steps, we get

$$\begin{aligned}p_{2T} &= \mathbf{P}^T p_T = \mathbf{P}^T (q\bar{p} + (1 - q)\tilde{p}_T) \\ &= q\bar{p} + q(1 - q)\bar{p} + (1 - q)^2 \tilde{p}_{2T}.\end{aligned}$$

And after another $(k - 2)T$ steps we find by an iterative argument that

$$p_{kT} = [1 - (1 - q)^k] \bar{p} + (1 - q)^k \tilde{p}_{kT}.$$

This shows that

$$\|\mathbf{P}^t p_0 - \bar{p}\|_{TV} \leq (1 - q)^{\lfloor t/T \rfloor} \quad \forall t \geq 0, p_0,$$

and thus $\|\mathbf{P}^t p_0 - \bar{p}\|_{TV} \leq \epsilon$ provided $t \geq T(1 + \frac{\ln \epsilon^{-1}}{\ln(1 - q)})$. As q is a fixed constant below 1, and $T \in O(Md^2 \ln(d))$, this proves the claimed mixing time. \blacksquare

- [1] Y. Aharonov, L. Davidovich, and N. Zagury, Quantum random walks, *Phys. Rev. A* **48**, 1687 (1993).
- [2] E. Farhi and S. Gutmann, Quantum computation and decision trees, *Phys. Rev. A* **58**, 915 (1998).
- [3] J. Watrous, Quantum simulations of classical random walks and undirected graph connectivity, *J. Comput. Syst. Sci.* **62**, 376 (2001).
- [4] J. Kempe, Quantum random walks: An introductory overview, *Contemp. Phys.* **44**, 307 (2003).

- [5] A. Ambainis, E. Bach, A. Nayak, A. Vishwanath, and J. Watrous, One-dimensional quantum walks, *Proceedings of the Thirty-Third Annual ACM Symposium on Theory of Computing (STOC '01)* (ACM, New York, NY, 2001), pp. 37–49.
- [6] G. S. Engel, T. R. Calhoun, E. L. Read, T.-K. Ahn, T. Mančal, Y.-C. Cheng, R. E. Blankenship, and G. R. Fleming, Evidence for wavelike energy transfer through quantum coherence in photosynthetic systems, *Nature (London)* **446**, 782 (2007).

- [7] M. Mohseni, P. Rebentrost, S. Lloyd, and A. Aspuru-Guzik, Environment-assisted quantum walks in photosynthetic energy transfer, *J. Chem. Phys.* **129**, 174106 (2008).
- [8] A. Romanelli, R. Donangelo, R. Portugal, and F. L. Marquezino, Thermodynamics of N -dimensional quantum walks, *Phys. Rev. A* **90**, 022329 (2014).
- [9] T. Oka, N. Konno, R. Arita, and H. Aoki, Breakdown of an Electric-Field Driven System: A Mapping to a Quantum Walk, *Phys. Rev. Lett.* **94**, 100602 (2005).
- [10] T. Kitagawa, M. A. Broome, A. Fedrizzi, M. S. Rudner, E. Berg, I. Kassal, A. Aspuru-Guzik, E. Demler, and A. G. White, Observation of topologically protected bound states in photonic quantum walks, *Nat. Commun.* **3**, 1872 (2012).
- [11] M. Karski, L. Förster, J.-M. Choi, A. Steffen, W. Alt, D. Meschede, and A. Widera, Quantum walk in position space with single optically trapped atoms, *Science* **325**, 174 (2009).
- [12] A. Peruzzo, M. Lobino, J. C. F. Matthews, N. Matsuda, A. Politi, K. Poulios, X.-Q. Zhou, Y. Lahini, N. Ismail, K. Wörhoff, *et al.*, Quantum walks of correlated photons, *Science* **329**, 1500 (2010).
- [13] M. Genske, W. Alt, A. Steffen, A. H. Werner, R. F. Werner, D. Meschede, and A. Alberti, Electric Quantum Walks with Individual Atoms, *Phys. Rev. Lett.* **110**, 190601 (2013).
- [14] P. M. Preiss, R. Ma, M. E. Tai, A. Lukin, M. Rispoli, P. Zupancic, Y. Lahini, R. Islam, and M. Greiner, Strongly correlated quantum walks in optical lattices, *Science* **347**, 1229 (2015).
- [15] E. Flurin, V. V. Ramasesh, S. Hacoheh-Gourgy, L. S. Martin, N. Y. Yao, and I. Siddiqi, Observing Topological Invariants Using Quantum Walks in Superconducting Circuits, *Phys. Rev. X* **7**, 031023 (2017).
- [16] A. M. Childs, Universal Computation by Quantum Walk, *Phys. Rev. Lett.* **102**, 180501 (2009).
- [17] N. B. Lovett, S. Cooper, M. Everitt, M. Trevers, and V. Kendon, Universal quantum computation using the discrete-time quantum walk, *Phys. Rev. A* **81**, 042330 (2010).
- [18] A. Ambainis, Quantum walks and their algorithmic applications, *Int. J. Quantum Inf.* **1**, 507 (2003).
- [19] N. Shenvi, J. Kempe, and K. Birgitta Whaley, Quantum random-walk search algorithm, *Phys. Rev. A* **67**, 052307 (2003).
- [20] A. M. Childs and J. Goldstone, Spatial search by quantum walk, *Phys. Rev. A* **70**, 022314 (2004).
- [21] F. Magniez, A. Nayak, J. Roland, and M. Santha, Search via quantum walk, *SIAM J. Comput.* **40**, 142 (2011).
- [22] A. M. Childs, R. Cleve, E. Deotto, E. Farhi, S. Gutmann, and D. A. Spielman, Exponential algorithmic speedup by a quantum walk, *Proceedings of the Thirty-Fifth Annual ACM Symposium on Theory of Computing (STOC '03)* (ACM, New York, NY, 2003), pp. 59–68.
- [23] M. Szegedy, Quantum speed-up of Markov chain based algorithms, *45th Annual IEEE Symposium on Foundations of Computer Science, Rome Italy, 2004* (IEEE, Piscataway, NJ, 2004), pp. 32–41.
- [24] J. Kempe, Discrete quantum walks hit exponentially faster, *Probab. Theory Relat. Fields* **133**, 215 (2005).
- [25] H. Krovi, F. Magniez, M. Ozols, and J. Roland, Quantum walks can find a marked element on any graph, *Algorithmica* **74**, 851 (2016).
- [26] P. Hoyer and M. Komeili, Efficient quantum walk on the grid with multiple marked elements, *34th Symposium on Theoretical Aspects of Computer Science (STACS)*, Vol. 66 (Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2017), p. 42.
- [27] D. Aharonov, A. Ambainis, J. Kempe, and U. Vazirani, Quantum walks on graphs, *Proceedings of the Thirty-Third Annual ACM Symposium on Theory of Computing (STOC '01)* (ACM, New York, NY, 2001), pp. 50–59.
- [28] C. Moore and A. Russell, Quantum walks on the hypercube, in *Randomization and Approximation Techniques in Computer Science RANDOM 2002*, Lecture Notes in Computer Science, edited by J. D. P. Rolim and S. Vadhan, Vol. 2483 (Springer, Berlin, Heidelberg, 2002), pp. 164–178.
- [29] P. C. Richter, Quantum speedup of classical mixing processes, *Phys. Rev. A* **76**, 042306 (2007).
- [30] A. Sinclair, *Algorithms for Random Generation and Counting: A Markov Chain Approach* (Birkhauser Verlag, Basel, Switzerland, 1993).
- [31] P. L. Knight, E. Roldán, and J. E. Sipe, Quantum walk on the line as an interference phenomenon, *Phys. Rev. A* **68**, 020301 (2003).
- [32] H. Jeong, M. Paternostro, and M. S. Kim, Simulation of quantum random walks using the interference of a classical field, *Phys. Rev. A* **69**, 012310 (2004).
- [33] E. Roldán and J. C. Soriano, Optical implementability of the two-dimensional quantum walk, *J. Mod. Opt.* **52**, 2649 (2005).
- [34] S. K. Goyal, F. S. Roux, A. Forbes, and T. Konrad, Implementation of multidimensional quantum walks using linear optics and classical light, *Phys. Rev. A* **92**, 040302(R) (2015).
- [35] F. Chen, L. Lovász, and I. Pak, Lifting Markov chains to speed up mixing, in *Proceedings of the Thirty-First Annual ACM Symposium on Theory of Computing (STOC '99)* (ACM, New York, NY, 1999), pp. 275–281.
- [36] R. P. Feynman, Simulating physics with computers, *Int. J. Theor. Phys.* **21**, 467 (1982).
- [37] A. Y. Kitaev, A. Shen, and M. N. Vyalyi, *Classical and Quantum Computation* (American Mathematical Society, Providence, 2002).
- [38] D. Aharonov, W. Van Dam, J. Kempe, Z. Landau, S. Lloyd, and O. Regev, Adiabatic quantum computation is equivalent to standard quantum computation, *SIAM Rev.* **50**, 755 (2008).
- [39] S. Aaronson, Quantum computing and hidden variables, *Phys. Rev. A* **71**, 032325 (2005).
- [40] M. Pavon and F. Ticozzi, Discrete-time classical and quantum Markovian evolutions: Maximum entropy problems on path space, *J. Math. Phys.* **51**, 042104 (2010).
- [41] T. T. Georgiou and M. Pavon, Positive contraction mappings for classical and quantum Schrödinger systems, *J. Math. Phys.* **56**, 033301 (2015).
- [42] K. Temme, M. J. Kastoryano, M. B. Ruskai, M. M. Wolf, and F. Verstraete, The χ^2 -divergence and mixing times of quantum Markov processes, *J. Math. Phys.* **51**, 122201 (2010).
- [43] P. Diaconis, S. Holmes, and R. M. Neal, Analysis of a non-reversible Markov chain sampler, *Ann. Appl. Probab.* **10**, 726 (2000).
- [44] P. Diaconis and L. Miclo, On the spectral analysis of second-order Markov chains, *Annales de la Faculté des Sciences de Toulouse. Mathématiques. Série 6* **22**, 573 (2013).
- [45] D. A. Meyer, From quantum cellular automata to quantum lattice gases, *J. Stat. Phys.* **85**, 551 (1996).

- [46] F. L. Marquezino, R. Portugal, and G. Abal, Mixing times in quantum walks on two-dimensional grids, *Phys. Rev. A* **82**, 042341 (2010).
- [47] V. Kendon, Decoherence in quantum walks: A review, *Math. Struct. Comput. Sci.* **17**, 1169 (2007).
- [48] S. Attal, F. Petruccione, C. Sabot, and I. Sinayskiy, Open quantum random walks, *J. Stat. Phys.* **147**, 832 (2012).
- [49] P. C. Richter, Almost uniform sampling via quantum walks, *New J. Phys.* **9**, 72 (2007).
- [50] As we will compare this to LMCs, we must note that standard LMC literature such as [35] considers the *joint* distribution over $\mathcal{C} \times \mathcal{V}$ as the object of interest, without specifying the output map or the initialization map \mathbf{F} . This does not affect our QW results; in fact it implies no significant difference in general. See [51].
- [51] S. Apers, F. Ticozzi, and A. Sarlette, Lifting Markov chains to mix faster: Limits and opportunities, [arXiv:1705.08253](https://arxiv.org/abs/1705.08253).
- [52] For instance, if a QW distinguishes the initial states $\rho_0 = (|1\rangle + |2\rangle)(\langle 1| + \langle 2|)/2$ and $\rho_0 = (|1\rangle - |2\rangle)(\langle 1| - \langle 2|)/2$, which reasonable initialization element can we attribute to an LMC in order to make a comparable distinction?
- [53] F. L. Marquezino, R. Portugal, G. Abal, and R. Donangelo, Mixing times in quantum walks on the hypercube, *Phys. Rev. A* **77**, 042312 (2008).
- [54] This bound was conjectured for all Abelian Cayley graphs in [44], where the authors provide a concrete proof only for the cycle.
- [55] L. R. Ford and D. R. Fulkerson, Maximal flow through a network, *Can. J. Math.* **8**, 399 (1956).
- [56] D. A. Levin and Y. Peres, *Markov Chains and Mixing Times* (American Mathematical Society, Providence, Rhode Island, 2017).
- [57] D. Aldous and J. Fill, Reversible Markov Chains and Random Walks on Graphs, 2002, <https://www.stat.berkeley.edu/users/aldous/RWG/book.html>.
- [58] R. Montenegro and P. Tetali, Mathematical aspects of mixing times in Markov chains, *Foundations and Trends® in Theoretical Computer Science*, Vol. 1 (Now Publishers Inc., Boston, Delft, 2006), p. 237.
- [59] S. Attal, A. Joye, and C.-A. Pillet, *Open Quantum Systems I: The Hamiltonian Approach* (Springer-Verlag, Berlin, Heidelberg, 2006).
- [60] F. Ticozzi, L. Zuccato, P. D. Johnson, and L. Viola, Alternating projections methods for discrete-time stabilization of quantum states, *IEEE Transactions on Automatic Control* **63**, 819 (2018).
- [61] D. Burgarth, G. Chiribella, V. Giovannetti, P. Perinotti, and K. Yuasa, Ergodic and mixing quantum channels in finite dimensions, *New J. Phys.* **15**, 073045 (2013).
- [62] M. A. Schlosshauer, *Decoherence and the Quantum-to-Classical Transition*, The Frontiers Collection (Springer Science & Business Media, Berlin, Heidelberg, 2007).
- [63] N. Linden, S. Popescu, A. J. Short, and A. Winter, Quantum mechanical evolution towards thermal equilibrium, *Phys. Rev. E* **79**, 061103 (2009).