Quantum speedup of classical mixing processes

Peter C. Richter*

Department of Computer Science, Rutgers University, Piscataway, New Jersey 08854, USA

(Received 5 April 2007; published 4 October 2007)

Most approximation algorithms for #P-complete problems (e.g., evaluating the partition function of a monomer-dimer or ferromagnetic Ising system) work by reduction to the problem of approximate sampling from a distribution π over a large set $\mathcal{S}$. This problem is solved using the *Markov chain Monte Carlo* method: a sparse, reversible Markov chain P on $\mathcal{S}$ with stationary distribution π is run to near equilibrium. The running time of this random walk algorithm, the so-called *mixing time* of P, is $O(\delta^{-1}\log 1/\pi_*)$ as shown by Aldous, where δ is the spectral gap of P and π_* is the minimum value of π. A natural question is whether a speedup of this classical method to $O(\sqrt{\delta^{-1}}\log 1/\pi_*)$ is possible using *quantum walks*. We provide evidence for this possibility using quantum walks that *decohere* under repeated randomized measurements. We show that (i) decoherent quantum walks always mix, just like their classical counterparts, (ii) the mixing time is a robust quantity, essentially invariant under any smooth form of decoherence, and (iii) the mixing time of the decoherent quantum walk on a periodic lattice $\mathbb{Z}_n^d$ is $O(nd\log d)$, which is indeed $O(\sqrt{\delta^{-1}}\log 1/\pi_*)$ and is asymptotically no worse than the diameter of $\mathbb{Z}_n^d$ (the obvious lower bound) up to at most a logarithmic factor.

DOI: 10.1103/PhysRevA.76.042306 PACS number(s): 03.67.Lx, 02.50.Ga

I. INTRODUCTION

A. Markov chain Monte Carlo and quantum walks

A rich theory has been developed for computing approximate solutions to problems in combinatorial enumeration and statistical physics which are #P-complete and therefore unlikely to have efficiently computable exact solutions. Among the highlights are randomized polynomial-time algorithms for approximating the permanent of a non-negative matrix [1], the volume of a convex polytope [2], and the partition functions of monomer-dimer and ferromagnetic Ising systems [3,4]. At the heart of these algorithms, and consuming much of the running time, is a subroutine for approximate sampling from a distribution π over a large set $\mathcal{S}$ of *states*. This problem is solved using the *Markov chain Monte Carlo* (MCMC) method: a *random walk* $(x_t)_{t=0,1,2,\ldots}$ on $\mathcal{S}$ (i.e., a random sequence of state transitions) is simulated from an arbitrary initial state x_0 to a random state x_T distributed close to π. The random walk is generated by a *Markov chain P* on $\mathcal{S}$ (i.e., an $|\mathcal{S}|\times|\mathcal{S}|$ column-stochastic matrix of state transition probabilities) whose *stationary* (fixed-point) distribution is π. The time required to guarantee closeness to π, or *mixing*, is the so-called *mixing time* τ_{mix}. Bounding τ_{mix} reduces to estimating the *spectral gap* $\delta := 1-|\lambda|$, where λ is the second-largest eigenvalue of P in magnitude, by Aldous' inequality [5],

$$\delta^{-1} \le \tau_{\text{mix}} \le \delta^{-1}(\ln 1/\pi_*), \tag{1}$$

where $\pi_* := \min_x \pi(x) > 0$.

Example 1: Monomer-dimer systems. Suppose we are given as input a graph G, consisting of vertices V and edges E, and a positive value λ. Let $\mathcal{S}_k$ be the set of all k-matchings in G, where a *k-matching* is a k-subset of edges which are vertex-disjoint. A pair of vertices connected by an edge in the matching is a *dimer*; an isolated vertex is a *monomer*. Define the *Gibbs distribution*

$$\pi(x) := \frac{1}{Z(\lambda)}\lambda^{|x|} \tag{2}$$

on the set $x \in \mathcal{S} := \cup_k \mathcal{S}_k$, where $Z(\lambda) = \Sigma_{x\in\mathcal{S}}\lambda^{|x|} = \Sigma_k |\mathcal{S}_k|\lambda^k$ is the *partition function*. Our computational task is to evaluate $Z(\lambda)$ as accurately as possible.

We cannot hope to compute $Z(\lambda)$ exactly, because this problem is #P-complete for any $\lambda > 0$. However, consider that we can write $Z(\lambda)$ as a telescoping product,

$$Z(\lambda) = \frac{Z(\lambda_r)}{Z(\lambda_{r-1})} \times \frac{Z(\lambda_{r-1})}{Z(\lambda_{r-2})} \times \cdots \times \frac{Z(\lambda_1)}{Z(\lambda_0)} \times Z(\lambda_0) \tag{3}$$

for some sequence $\lambda = \lambda_r > \lambda_{r-1} > \cdots > \lambda_1 > \lambda_0 = 0$, where $Z(\lambda_0) = 1$ is trivial to compute and the ratios $Z(\lambda_{i+1})/Z(\lambda_i)$ are easily seen to equal the expectation values $E_{x\leftarrow\pi_i}[(\lambda_{i+1}/\lambda_i)^{|x|}]$. (Here π_i is the Gibbs distribution for λ_i.) It can be shown that r need not be very large to guarantee that the expectation values $E_{x\leftarrow\pi_i}[(\lambda_{i+1}/\lambda_i)^{|x|}]$ are bounded and well-estimated using relatively few samples from the distributions π_i.

Thus we have reduced the problem of approximating $Z(\lambda)$ efficiently to the problem of sampling efficiently from the Gibbs distribution π. The sampling problem can be solved by running a random walk $(x_t)_{t=0,1,2,\ldots}$ on the Markov chain P with the following transition rule: at state $x_t \in \mathcal{S}_k$, choose an edge $e = (u,v) \in E$ uniformly at random. If e is not in x_t and $x' := x_t \cup \{e\}$ is a valid matching, set $x_{t+1} := x' \in \mathcal{S}_{k+1}$; if e is in x_t, set $x_{t+1} := x\backslash\{e\} \in \mathcal{S}_{k-1}$ with probability $1/\lambda$. Otherwise, either move to a "nearby" state $x_{t+1} \in \mathcal{S}_k$ (i.e., set x_{t+1} so that $|x_{t+1} \cap x_t| = k-1$) or simply stay put (i.e., set $x_{t+1} = x_t$). It is not too hard to see that P has π as its stationary distribution; moreover, its spectral gap is $\Omega(|V||E|)$ [3,6]. An upper bound on the mixing time follows from Aldous' inequality (1).

*richterp@cs.rutgers.edu

Notice that Aldous' inequality (1) is tight with respect to the spectral gap. However, it is also somewhat unsatisfactory in that although the *diameter* (i.e., the maximum distance between any two vertices) of the graph underlying P—the obvious lower bound for sampling from its stationary distribution—scales like $\sqrt{\delta^{-1}}$, the dependence is δ^{-1} in Aldous' inequality due to the Gaussian-like spreading behavior of random walks. In MCMC algorithms like the monomer-dimer example above, the additional factor of $\sqrt{\delta^{-1}}$ can increase the running time significantly. Removing this factor would imply considerable improvement of such algorithms in both the *known* upper bounds (since existing estimates of δ could be used in conjunction with a sharper inequality than Aldous') and the *true* upper bounds (since Aldous' inequality is tight, and thus a sharper inequality could only come from a faster approximate sampling method). Thus a natural question is whether there is a way to modify the standard MCMC method to obtain a speedup to $O(\sqrt{\delta^{-1}}\log 1/\pi_*)$. This seems unlikely using classical randomized methods: Chen, Lovasz, and Pak [7] have shown that we can sometimes speed up mixing by *lifting* (essentially, by locally "re-routing") a Markov chain, but this requires both knowledge of the chain's global structure and its use in solving an *NP*-hard flow problem to find good routes. However, although lifting the chain seems unlikely to be practical, an idea that might work is *quantizing* the chain.

Why might a quantized Markov chain, or *quantum walk*, help us reach π more quickly? Two reasons are (i) a quantum walk is as simple to realize as its classical counterpart (i.e., it is computable locally and online at each step of the walk, unlike a classical lifting of the chain), and (ii) there is empirical (and some theoretical) evidence that quantum walks tend to propagate and "spread" probability mass across $\mathcal{S}$ in time on the order of the diameter on precisely the same low-dimensional graphs that trip up their classical counterparts. Based on these observations, the possibility of obtaining a quantum speedup for the mixing problem has been pursued by Nayak *et al.* [8,9], Aharonov *et al.* [10], Moore and Russell [11], Gerhardt and Watrous [12], and Richter [13].

We remark that a quantum speedup theorem of the sort we seek has already been proven for the *hitting* problem, in which we search (rather than sample from) the states of a Markov chain: Szegedy [14] proved a quadratic quantum speedup for the *hitting time* (i.e., the time to detect the presence of a "marked" state) of any symmetric Markov chain, generalizing considerably the celebrated search algorithm of Grover [15] and implying quantum speedups for structured search problems such as element distinctness [16], triangle finding [17], matrix product verification [18], and group commutativity testing [19]. It is this success which inspires us to investigate the possibility of a quantum speedup for the mixing problem with the goal of transferring the speedup to MCMC algorithms.

B. Our contributions

We present evidence of a possible quantum MCMC speedup to $O(\sqrt{\delta^{-1}}\log 1/\pi_*)$ using quantum walks that *decohere* under repeated randomized measurements. Decoherence (in small amounts) was first identified as a way to improve spreading and mixing properties in numerical experiments performed by Kendon and Tregenna [20] and in analytical estimates by Fedichkin *et al.* [21–23]. On the other hand, high rates of decoherence in quantum walks have been shown to degrade mixing properties substantially by the quantum Zeno effect [24]. For an excellent survey of these and other aspects of decoherent quantum walks, see Kendon [25].

Our technical contributions are as follows: First, we show that for any symmetric Markov chain P, we can generate an arbitrarily good approximation to the uniform stationary distribution π of P by subjecting the continuous-time quantum walk $U_{ct}(P)=\exp(iP)$ to reasonably "smooth" decoherence. Thus decoherent quantum walks (which are nonunitary) offer a way of circumventing an obstacle first identified by Aharonov *et al.* [10], who observed that *unitary* quantum walks often converge (in the time-averaged sense) to highly nonuniform distributions.

Second, we show that the optimal mixing time of a decoherent quantum walk is a robust quantity, in that it remains essentially invariant under any sufficiently smooth form of decoherence. In particular, decoherent quantum walks undergoing repeated Cesaro-averaged [10–12] and Bernoulli-Poisson-averaged [20,24] measurements are nearly equivalent in mixing efficiency. The proof applies more generally to a game involving time-dependent Markov chains (not necessarily describing quantum phenomena) and may be of independent interest.

Third, we prove a theorem on threshold mixing of quantum walks on (Cartesian) graph powers in order to show that the decoherent continuous-time quantum walk $U_{ct}(P(G))$ on a periodic lattice $G=\mathbb{Z}_n^d$ [where $P(G)$ denotes the standard "simple random walk" Markov chain on G] can be used to generate a good approximation to the uniform stationary distribution π of P in time $O(nd\log d)$. This upper bound is asymptotically no worse than the diameter of $\mathbb{Z}_n^d$ (the obvious lower bound) up to at most a logarithmic factor and is $O(\sqrt{\delta^{-1}}\log 1/\pi_*)$ for both high-dimensional *and* low-dimensional lattices (unlike its classical counterpart). For $d=1$, this proves a conjecture of Kendon and Tregenna [20] based on numerical experiments [26]. For $d\geq 1$, it extends the results of Fedichkin *et al.* [21–23] by confirming $O(n)$ and $O(d\log d)$ scaling (suggested by their analytical estimates and numerical experiments in regimes of both high and low decoherence) of the fastest-mixing walk, which they conjectured to be decoherent rather than unitary. We briefly discuss the prospects for extending this result to the discrete-time Grover walk $U_{dt}(P(\mathbb{Z}_n^d))$ [27–29].

Previously, mixing speedups had been proven only for the unitary quantum walks of Nayak *et al.* [8,9] and Aharonov *et al.* [10] (on the cycle) and of Moore and Russell [11] (on the hypercube). Thus our work shows that introducing a small amount of decoherence to a quantum walk can simultaneously force convergence to the uniform distribution while preserving a quantum mixing speedup, an advantageous combination for algorithmic applications [30].

II. PRELIMINARIES

A. Markov chains

Let P be a *Markov chain* (column-stochastic matrix) on the set $\mathcal{S}$ ($|\mathcal{S}|=N$) which is *irreducible* (strongly connected); then it has a unique distribution π which is *stationary* (i.e., satisfies $P\pi=\pi$). Moreover, π is strictly positive: $\pi_*:=\min_x \pi(x)$ satisfies $\pi_*>0$. In particular, if P is symmetric, then π is the uniform distribution u (the N-dimensional column vector with each component equal to $1/N$). If G is an undirected graph, we denote the standard ("simple random walk") Markov chain on G by $P(G)$.

A Markov chain P which is both irreducible and *aperiodic* (nonbipartite) is by definition *ergodic* and satisfies

$$P^t \to \pi 1^\dagger = [\pi \cdots \pi] \quad \text{as} \quad t\to\infty, \tag{4}$$

where $1^\dagger$ is the N-dimensional row vector with each component equal to 1. We can thus define the *(threshold) mixing time*

$$\tau_{\text{mix}} := \min\left\{T: \frac{1}{2}\|P^t - \pi 1^\dagger\|_1 \le \frac{1}{2e}\ \forall\, t\ge T\right\}, \tag{5}$$

where $\|\cdot\|_1$ is the matrix 1-norm. The mixing is *perfect* if $P^t=\pi 1^\dagger$. Let $\delta:=1-|\lambda|$ be the *spectral gap* of P, where λ is the second-largest eigenvalue of P in magnitude. We say that P is *reversible* if the matrix DPD^{-1} is symmetric, where D is the diagonal matrix $D(x,x)=\sqrt{\pi(x)}$. A precise statement of Aldous' inequality (1) is as follows.

Theorem 2 (Aldous [5]). Let P be a reversible, ergodic Markov chain with stationary distribution π and spectral gap δ. Then its mixing time satisfies

$$\delta^{-1} \le \tau_{\text{mix}} \le \delta^{-1}(\ln 1/\pi_*).$$

We will also use the *maximum pairwise column distance* $d(P):=\max_{x,x'}\frac{1}{2}\|P(\cdot,x)-P(\cdot,x')\|_1$ to estimate the mixing time. It is related to the matrix 1-norm distance by the inequality:

$$\frac{1}{2}\|P-\pi 1^\dagger\|_1 \le d(P) \le \|P-\pi 1^\dagger\|_1. \tag{6}$$

The following propositions (see [13] for proofs) can be used to estimate the mixing time of P given a common lower bound on most of the entries in each column.

Proposition 3. If $d(P)\le\alpha$, then $\tau_{\text{mix}}\le\lceil\log_{1/\alpha} 2e\rceil$.

Proposition 4. If at least βN entries in each column of P are bounded below by γ/N, where $\beta>\frac{1}{2}$ and $\gamma>0$, then $d(P)\le 1-\gamma[1-2(1-\beta)]$.

B. Quantum walks

Henceforth, a *quantum walk* is a pair $\langle U,\omega_T\rangle$ where the *transition rule* U is a unitary operator acting on a finite-dimensional Hilbert space and the *measurement rule* ω_T is a T-parametrized family of probability density functions on $[0,\infty)$ characterizing the (random) time at which a total measurement is performed on the Hilbert space [31].

The unitary transition rule determines the orbit of a pure quantum state, or *wave function* (l_2-normalized complex vector), just as a Markov chain (or stochastic transition rule) determines the orbit of a classical distribution (l_1-normalized non-negative vector). Let P be the Markov chain used in an MCMC algorithm; in particular, P is reversible. Two natural quantizations of P are (i) the $\mathcal{S}\times\mathcal{S}$ unitary *continuous-time walk* given by $U_{ct}(P)=\exp(iDPD^{-1})$ [32,33], where $H=DPD^{-1}$ is the time-independent *Hamiltonian*, and (ii) the $\mathcal{S}^2\times\mathcal{S}^2$ unitary discrete-time *Grover-Szegedy walk* given by $U_{dt}(P)=(RS)^2$ [14], where S is the involution $\Sigma_{x,y\in\mathcal{S}}|x,y\rangle\mapsto|y,x\rangle$, R is the reflection $\Sigma_{x\in\mathcal{S}}|x\rangle\langle x|\otimes(2|p_x\rangle\times\langle p_x|-I)$, and $|p_x\rangle$ is the vector $\Sigma_{y\in\mathcal{S}}\sqrt{P(y,x)}|y\rangle$. The quantization $U_{ct}(P)$ was used by Childs *et al.* [34] to solve in polynomial time a natural oracle problem for which no classical polynomial-time algorithm exists. The quantization $U_{dt}(P)$ was used by Szegedy [14] to prove a quadratic quantum speedup for the hitting time of any symmetric Markov chain.

Measurement collapses the wave function to a classical distribution according to the map $|\phi\rangle\mapsto\Sigma_x|x\rangle|\langle x|\phi\rangle|^2$. For a quantum walk $\langle U,\omega_T\rangle$, the Markov chain *generated by* the quantum walk is given by $\hat{P}_T(y,x):=E_{t\leftarrow\omega_T}[|\langle y|U^t|x\rangle|^2]$, where E denotes expected value. We say that the quantum walk *threshold mixes* if the Markov chain it generates mixes in time $O(1)$. Examples of measurement rules from the literature include the point distribution $\delta_T(t):=\delta(t-T)$ where δ is the delta function [8,9], the uniform distribution $\bar{\mu}_T:=\frac{1}{T}\chi_{[0,T]}$ and its discrete-time counterpart $\bar{\nu}_T:=\frac{1}{T}\chi_{[0,T-1]}$ where χ is the characteristic function [10], the exponential distribution $\tilde{\mu}_T(t):=\frac{1}{T}\exp(-t/T)$, and the geometric distribution $\tilde{\nu}_T(t):=\frac{1}{T}(1-\frac{1}{T})^t$. The exponential and geometric distributions are *memoryless* and describe the interarrival time between measurements in a Poisson process with measurements occurring at rate $\lambda=1/T$ and a Bernoulli process with measurements occurring with probability $p=1/T$ at each time step, respectively. These processes coincide with the decoherence models of Alagic and Russell [24] and Kendon and Tregenna [20], respectively.

Nayak *et al.* [8,9] and Aharonov *et al.* [10] showed that the so-called *Hadamard* walks $\langle U_{Had},\delta_T\rangle$ and $\langle U_{Had},\bar{\nu}_T\rangle$ on the cycle $\mathbb{Z}_n$ threshold-mix quadratically faster than the classical random walk, although their definitions of threshold mixing are slightly different than ours. Moore and Russell [11] showed that the continuous-time walk $\langle U_{ct}(P(G)),\delta_T\rangle$ and the Grover walk $\langle U_{dt}(P(G)),\delta_T\rangle$ on the hypercube $G=\mathbb{Z}_2^d$ mix perfectly and almost perfectly, respectively, in time $T=O(d)$. They also showed that the continuous-time quantum walk on the hypercube with measurement $\bar{\mu}_T$ does not mix to the uniform stationary distribution π of $P(\mathbb{Z}_2^d)$ in the limit $T\to\infty$. Gerhardt and Watrous [12] showed the same for a continuous-time quantum walk on the symmetric group with measurement $\bar{\mu}_T$. See the survey by Kempe [35] for further results on quantum walks.

We remark that there is a nice way to use the quantization $U_{dt}(P)$ to solve the mixing problem in time $O(1/\sqrt{\delta\pi_*})$ which, although prohibitively costly in π_*, exhibits the desired dependence on δ. Consider the stationary eigenstate $|\tilde{\pi}\rangle:=\Sigma_{x\in\mathcal{S}}\sqrt{\pi(x)}|x\rangle|p_x\rangle$ of $U_{dt}(P)$. It is clear that we can retrieve a good approximation to the classical distribution π by generating and then measuring a good approximation to $|\tilde{\pi}\rangle$ [36]. Magniez *et al.* [37] observe that $O(1/\sqrt{\delta})$ steps of phase estimation on $U_{dt}(P)$ enable us to reflect about $|\tilde{\pi}\rangle$. By alternating this reflection with a reflection about $|\tilde{z}\rangle := |z\rangle|p_z\rangle$ where z is the initial walk state (in particular, $\langle\tilde{z}|\tilde{\pi}\rangle \geq \sqrt{\pi_*}$), we can generate $|\tilde{\pi}\rangle$ from $|\tilde{z}\rangle$ in time $O(1/\sqrt{\delta\pi_*})$. In fact, this algorithm is described by Magniez *et al.* [37] as a *hitting* algorithm (i.e., generate the *unknown* state $|\tilde{z}\rangle$ from the fixed initial state $|\tilde{\pi}\rangle$); the idea of running a quantum hitting algorithm *in reverse* as a mixing algorithm was suggested by Childs [38].

III. MIXING PROPERTIES OF DECOHERENT QUANTUM WALKS

A. Two types of convergence

Let $\hat{P}_T$ be the Markov chain generated by a quantum walk $\langle U, \omega_T\rangle$. Then repeating the quantum walk T' times in succession generates the Markov chain $(\hat{P}_T)^{T'}$. The following lemma (a variant of Theorem 3.4 in Aharonov *et al.* [10]) and theorem describe the asymptotic behavior of $\hat{P}_T$ and $(\hat{P}_T)^{T'}$ in the limits $T\to\infty$ and $T'\to\infty$, respectively. For concreteness we will take $U=U_{ct}$ in this subsection and the next; it is a simple exercise to extend the results to discrete-time walk variants. Although stated explicitly for quantum walks, the results apply to any time-independent quantum dynamics on a finite-dimensional Hilbert space subjected to random destructive measurements.

Lemma 5 (The limit $T\to\infty$). Let P be a symmetric Markov chain and ω_T be a family of distributions satisfying $E_{t\leftarrow\omega_T}[e^{i\theta t}]\to 0$ as $T\to\infty$ for any $\theta\neq 0$. In the limit $T\to\infty$, the Markov chain $\hat{P}_T$ generated by the quantum walk $\langle U_{ct}(P), \omega_T\rangle$ approaches the Markov chain Π with entries

$$\Pi(y,x) := \sum_j \left|\sum_{k\in C_j} \langle y|\phi_k\rangle\langle\phi_k|x\rangle\right|^2, \tag{7}$$

where $\{\lambda_k, |\phi_k\rangle\}$ is the spectrum of P and $\{C_j\}$ is the partition of these indices k obtained by grouping together the k with identical λ_k.

Proof. Decomposing the quantum walk along spectral components gives us

$$\hat{P}_T(y,x) = E_{t\leftarrow\omega_T}\left[\left|\sum_k \langle y|\phi_k\rangle\langle\phi_k|x\rangle e^{i\lambda_k t}\right|^2\right]. \tag{8}$$

Writing $|\cdot|^2$ as a product of complex conjugates, the right hand side becomes

$$E_{\omega_T}\left[\left(\sum_k \langle y|\phi_k\rangle\langle\phi_k|x\rangle\right)\left(\sum_l \langle\phi_l|y\rangle\langle x|\phi_l\rangle\right)e^{i(\lambda_k-\lambda_l)t}\right]. \tag{9}$$

By linearity of expectation, this is equivalent to

$$\left(\sum_k \langle y|\phi_k\rangle\langle\phi_k|x\rangle\right)\left(\sum_l \langle\phi_l|y\rangle\langle x|\phi_l\rangle\right)E_{\omega_T}[e^{i(\lambda_k-\lambda_l)t}]. \tag{10}$$

Now by assumption, $E_{\omega_T}[e^{i(\lambda_k-\lambda_l)t}]$ vanishes as $T\to\infty$ for all $\lambda_k\neq\lambda_l$, so we have

$$\begin{aligned}\hat{P}_T(y,x) &\to \sum_k \langle y|\phi_k\rangle\langle\phi_k|x\rangle\left(\sum_{l:\theta_l=\theta_k}\langle\phi_k|y\rangle\langle x|\phi_k\rangle\right)\\ &= \sum_j \left|\sum_{k\in C_j}\langle y|\phi_k\rangle\langle\phi_k|x\rangle\right|^2\\ &= \Pi(y,x)\end{aligned} \tag{11}$$

in the limit $T\to\infty$.

It can be inferred from this lemma that most quantum walks converge to a distribution ρ other than the uniform stationary distribution $\pi=u$, and that ρ is not even independent of the initial walk state [39]. There are exceptions to this rule, for example quantum walks with distinct eigenvalues on Cayley graphs of Abelian groups (as observed by Aharonov *et al.* [10]), but they are not likely to arise in MCMC applications where the Markov chains have little structure. How then are we to sample from u using quantizations of these Markov chains? Here is where decoherence helps.

Theorem 6 (The limit $T'\to\infty$). Let P be a symmetric, irreducible Markov chain and ω_T be a family of distributions satisfying $E_{t\leftarrow\omega_T}[e^{i\theta t}]\to 0$ as $T\to\infty$ for any $\theta\neq 0$. For T sufficiently large (but fixed), the T'-repeated quantum walk $\langle U_{ct}(P), \omega_T\rangle$ generates a Markov chain $(\hat{P}_T)^{T'}$ approaching $u1^\dagger$ in the limit $T'\to\infty$.

Proof. We need to show that for T sufficiently large, the Markov chain $\hat{P}_T$ is ergodic with uniform stationary distribution.

That the uniform distribution is stationary is clear: each of the $P_t(y,x) := |\langle y|e^{iPt}|x\rangle|^2$ has uniform stationary distribution since the uniform classical state is invariant under unitary quantum operations and under total measurement of the system; thus any probabilistic combination $\hat{P}_T$ of them has uniform stationary distribution.

To show that $\hat{P}_T$ is ergodic for all sufficiently large T, it is sufficient (by Lemma 5) to prove that Π is ergodic. (The latter implies the former because the ergodic matrices form an open subset of the set of stochastic matrices.) Why is Π ergodic? Because the 1-eigenspace of P is precisely the space spanned by u, so it follows from Lemma 5 [by consideration of only this nondegenerate eigenspace in the expression (7)] that $\Pi(y,x)\geq 1/N^2$ for every x,y.

In fact, each of the P_t (and so $\hat{P}_T$ and Π as well) is symmetric [40]. To see this, write out the Taylor series for $\exp(iPt)$ and note that every positive integer power P^k is symmetric [since $P^2(x,y)=\Sigma_z P(x,z)\cdot P(z,y) =\Sigma_z P(y,z)\cdot P(z,x)=P^2(y,x)$]. This property will be quite useful in the next subsection: it will allow us to use Theorem 1 to relate the spectral gap and the mixing time of $\hat{P}_T$.

B. Invariance of the mixing time

Consider the quantum walks $\langle U_{ct}(P), \bar{\mu}_T\rangle$ and $\langle U_{ct}(P), \tilde{\mu}_T\rangle$ where P is a symmetric Markov chain. We show that these two quantum walks mix with essentially the same efficiency. The result extends beyond $\bar{\mu}, \tilde{\mu}$ to any pair of measurements ω, ω' which are sufficiently smooth or have nontrivial overlap as distributions. Let $\bar{P}_T$ and $\tilde{P}_T$ be the Markov chains generated by the walks with measurement $\bar{\mu}$ and $\tilde{\mu}$, respectively, and let $\bar{\delta}_T$ and $\tilde{\delta}_T$ be their respective spectral gaps.

Lemma 7 (Spectral gap inequality). Let $\bar{\delta}_T$ and $\tilde{\delta}_T$ be defined as above. Then for any $k \geq 1$ we have the inequality:

$$e^{-1}\bar{\delta}_T \leq \tilde{\delta}_T \leq k(1-e^{-k})\bar{\delta}_{kT} + 2e^{-k}. \tag{12}$$

Proof. Suppose we want to simulate $\bar{P}_T$ by $\tilde{P}_T$. Scaling the distribution $\bar{\mu}_T$ by $\alpha := 1/e$ allows us to "fit it inside" the distribution $\tilde{\mu}_T$ (i.e., $e^{-1}\bar{\mu}_T \leq \tilde{\mu}_T$ pointwise), so we can express $\tilde{\mu}_T$ as the probabilistic combination $\alpha\bar{\mu}_T + (1-\alpha)\nu$ for some distribution ν, so that

$$\tilde{P}_T = E_{t\leftarrow\tilde{\mu}_T}[P_t] = \alpha E_{\bar{\mu}_T}[P_t] + (1-\alpha)E_\nu[P_t] = \alpha\bar{P}_T + (1-\alpha)Q, \tag{13}$$

where Q is stochastic with uniform stationary distribution. It follows that

$$\|\tilde{P}_T|_{u^\perp}\|_2 \leq 1/e\|\bar{P}_T|_{u^\perp}\|_2 + (1-1/e)\|Q_{u^\perp}\|_2 \tag{14}$$

which implies that $\tilde{\delta}_T \geq (1/e)\bar{\delta}_T$ since $\|Q|_{u^\perp}\|_2 \leq 1$.

Suppose we want to simulate $\tilde{P}_T$ by $\bar{P}_{kT}$. Then the basic approach is the same, but since the support of $\tilde{\mu}_T$ is not compact we have to be careful. Scaling the distribution $\tilde{\mu}_T$ by $\beta := 1/k$ allows us to fit it inside the distribution $\bar{\mu}_{kT}$ up to the point $t = kT$, and the probability mass in $\tilde{\mu}_T$ past $t = kT$ is only $\Pr_{t\leftarrow\tilde{\mu}_T}[t > kT] = e^{-k}$. So we can write

$$\tilde{\mu}_T = (1-e^{-k})\cdot\tilde{\mu}_T^{\text{head}} + e^{-k}\cdot\tilde{\mu}_T^{\text{tail}}, \tag{15}$$

where $\tilde{\mu}_T^{\text{head}}$ and $\tilde{\mu}_T^{\text{tail}}$ are the conditional distributions of $\tilde{\mu}_T$ such that $t \leq kT$ and $t > kT$, respectively; thus

$$\tilde{P}_T = (1-e^{-k})\cdot\tilde{P}_T^{\text{head}} + e^{-k}\cdot\tilde{P}_T^{\text{tail}}, \tag{16}$$

where $\tilde{P}_T^{\text{head}}$ and $\tilde{P}_T^{\text{tail}}$ are the expectations of P_t with respect to $\tilde{\mu}_T^{\text{head}}$ and $\tilde{\mu}_T^{\text{tail}}$, respectively. Since we can fit $\tilde{\mu}_T^{\text{head}}$ inside $\bar{\mu}_{kT}$ if we scale it by $1/k$, we can write

$$\bar{P}_{kT} = \frac{1}{k}\tilde{P}_T^{\text{head}} + \left(1-\frac{1}{k}\right)Q, \tag{17}$$

where Q is stochastic with uniform stationary distribution. The above equations yield

$$\bar{P}_{kT} = \frac{1}{k(1-e^{-k})}(\tilde{P}_T - e^{-k}\tilde{P}_T^{\text{tail}}) + \left(1-\frac{1}{k}\right)Q. \tag{18}$$

From the triangle inequality, $\|\bar{P}_{kT}|_{u^\perp}\|_2$ is at most

$$\frac{\|\tilde{P}_T|_{u^\perp}\|_2 + e^{-k}\|\tilde{P}_T^{\text{tail}}|_{u^\perp}\|_2}{k(1-e^{-k})} + \left(1-\frac{1}{k}\right)\|Q|_{u^\perp}\|_2. \tag{19}$$

Rearranging terms and simplifying, we have

$$\frac{(1-\|\tilde{P}_T|_{u^\perp}\|_2) - 2e^{-k}}{k(1-e^{-k})} \leq 1 - \|\bar{P}_{kT}|_{u^\perp}\|_2. \tag{20}$$

Theorem 8 (Equivalence of measurements). Let P be a symmetric Markov chain. Then (i) if the T'-repeated quantum walk $\langle U_{ct}(P), \bar{\mu}_T\rangle$ threshold-mixes, then the $T'\cdot O(\log N)$-repeated quantum walk $\langle U_{ct}(P), \tilde{\mu}_T\rangle$ threshold-mixes; (ii) if the T'-repeated quantum walk $\langle U_{ct}(P), \tilde{\mu}_T\rangle$ threshold-mixes, then the $T'\cdot O(\log T' \log N)$-repeated quantum walk $\langle U_{ct}(P), \bar{\mu}_{T\cdot O(\log T')}\rangle$ threshold-mixes.

Proof. To see (i), note that our assumption implies that $\bar{P}_T$ mixes in time T'. Therefore $\bar{\delta}_T = \Omega(1/T')$ by Theorem 2, and from Lemma 7 it follows that $\tilde{\delta}_T = \Omega(1/T')$. Applying Theorem 2 again, we obtain for $\tilde{P}_T$ a mixing time of $O(T' \log N)$.

The proof of (ii) is almost as straightforward. Our assumption implies that $\tilde{P}_T$ mixes in time T', so $\tilde{\delta}_T = \Omega(1/T')$ by Theorem 2. Set k to be the smallest integer for which $\tilde{\delta}_T \geq 3e^{-k}$; in particular, $k = \Theta(\log \tilde{\delta}_T^{-1}) = O(\log T')$. By Lemma 7,

$$\bar{\delta}_{kT} \geq \frac{1}{k(1-e^{-k})}(\tilde{\delta}_T - 2e^{-k}) \geq \frac{1}{k(1-e^{-k})}(e^{-k}). \tag{21}$$

Asymptotically, the right hand side is

$$\Theta\left(\frac{\tilde{\delta}_T}{\log \tilde{\delta}_T^{-1}}\right) = \Theta\left(\frac{1}{T' \log T'}\right). \tag{22}$$

Applying Theorem 2 again, we obtain for $\bar{P}_{kT}$ a mixing time of $O(T' \log T' \log N)$.

It should be readily apparent that this equivalence holds for any two measurement rules with finite expectation and significant overlap for most T. We also remark that although the above lemma and theorem are stated in terms of quantum walks, the proofs indicate that they are merely statements about an abstract game involving a collection of symmetric Markov chains $\{P_t\}_{t\geq 0}$ and a T-parametrized family of probability measures $\{\omega_T\}$, where we seek to minimize the "cost function" $T\cdot T'$.

IV. QUANTUM SPEEDUP FOR PERIODIC LATTICES

A. A near-diameter upper bound

The classical random walk on the periodic lattice $\mathbb{Z}_n^d$ (with $N = n^d$ vertices and diameter $\lfloor n/2 \rfloor d$) has uniform stationary distribution $\pi = u$ and spectral gap $\delta = \Theta(\frac{1}{dn^2})$. It threshold-mixes in time $\Theta(n^2 d \log d)$, which is $O(\sqrt{\delta^{-1}} \log 1/\pi_*)$ only when $\mathbb{Z}_n^d$ is quite high-dimensional: in particular, when d is roughly of order n^2 or larger. We show that a few repetitions of the continuous-time quantum walk can bring this down to $O(nd \log d)$, which is $O(\sqrt{\delta^{-1}} \log 1/\pi_*)$ for any $d \geq 1$ and $n \geq 2$ and asymptotically no worse than the diameter of $\mathbb{Z}_n^d$ up to at most a logarithmic factor. First we prove a lemma governing mixing of various decoherent quantum walks on the cycle $\mathbb{Z}_n$.

Lemma 9 (Mixing on cycles). Let $\mathbb{Z}_n$ be the cycle on $n \geq 2$ vertices. The continuous-time walks $\langle U_{ct}(P(\mathbb{Z}_n)), \omega_T\rangle$ with measurement $\omega \in \{\delta, \bar{\mu}, \tilde{\mu}\}$ threshold-mix for any $T \in \mathcal{I} := \left[\frac{2}{3}\frac{n}{2}, \frac{n}{2}\right]$, and the Hadamard walks $\langle U_{Had}(\mathbb{Z}_n), \omega_T\rangle$ with measurement $\omega \in \{\bar{\nu}, \tilde{\nu}\}$ threshold-mix for any $T \in \mathcal{J} := \left[\frac{2}{3}\frac{n}{\sqrt{2}}, \frac{n}{\sqrt{2}}\right]$.

Proof. Consider first the continuous-time walk. To prove that it threshold-mixes with any of the measurements $\omega\in\{\delta,\bar{\mu},\tilde{\mu}\}$ for any $T\in\mathcal{I}$, it suffices by Proposition 3 to show that for every $t\in\mathcal{I}':=[\frac{3}{5}\frac{n}{2},\frac{4}{5}\frac{n}{2}]$, $d(P_t)$ is bounded below 1 by a positive constant, where P_t is the Markov chain generated with measurement $\omega=\delta$. (Indeed, this easily implies that $d(\bar{P}_T)$ and $d(\tilde{P}_T)$ are bounded below 1 by a smaller positive constant, where $\bar{P}_T$ and $\tilde{P}_T$ are the Markov chains generated with measurement $\omega=\bar{\mu}$ and $\omega=\tilde{\mu}$, respectively.)

Let $|\phi_t\rangle$ and $|\psi_t\rangle$ be the wave functions at time t for the continuous-time walks on $\mathbb{Z}$ and $\mathbb{Z}_n$, respectively, starting from the origin (without loss of generality, since $\mathbb{Z}$ and $\mathbb{Z}_n$ are vertex-transitive). Then for each $\bar{y}\in\mathbb{Z}_n$ we have

$$\langle\bar{y}|\psi_t\rangle=\sum_{y\equiv\bar{y}\bmod n}\langle y|\phi_t\rangle. \tag{23}$$

Childs [38] shows that $\langle y|\phi_t\rangle=(-i)^y J_y(t)$, where J_y is a Bessel function of the first kind. In particular, for $|y|\gg 1$ the quantity $|J_y(t)|$ is (i) exponentially small in $|y|$ for $t<(1-\epsilon)\cdot|y|$ and (ii) of order $|y|^{-1/2}$ for $t>(1+\epsilon)\cdot|y|$. For every $t<\frac{4}{5}\frac{n}{2}$, property (i) implies that the only term in the above summand that is non-negligible is the $\langle y|\phi_t\rangle$ with $|y|<\frac{n}{2}$ (call it $\hat{y}$ and note that $\bar{y}\leftrightarrow\hat{y}$ is a 1–1 correspondence), so we can use property (ii) to conclude that up to a negligible correction

$$|\langle\bar{y}|\psi_t\rangle|\approx|\langle\hat{y}|\phi_t\rangle|=\Theta(1/\sqrt{n}) \tag{24}$$

for every $|\hat{y}|\gg 1$ and $t>(1+\epsilon)\cdot|\hat{y}|$. In particular, the nearly $\frac{3}{5}n$ different $\bar{y}$ with $1\ll|\hat{y}|\le\frac{3}{5}\frac{n}{2}$ satisfy $|\langle\bar{y}|\psi_t\rangle|=\Omega(1/\sqrt{n})$, and therefore $P_t(\bar{y},\bar{0})=\Omega(1/n)$, for every $t\in\mathcal{I}'$. So by Proposition 4, $d(P_t)$ is bounded below 1 by a positive constant.

For the Hadamard walk, the wave function is no longer characterized by Bessel functions, but it retains the same essential asymptotic spreading behavior as its continuous-time counterpart (see Nayak *et al.* [8,9]), and the argument above works with little modification. A caveat is the emergence of a parity problem: if n is even, then $\mathbb{Z}_n$ is bipartite and the wave function is supported only on vertices of the same parity at each integer time step. Hence the Hadamard walk with $\omega=\delta$ threshold-mixes only on vertices of the same parity, but with time-averaged measurement $\omega=\bar{\nu}$ or $\omega=\tilde{\nu}$ parity is broken and threshold-mixing occurs on the entire vertex set.

Although the argument above relies on the asymptotic behavior of the wave function as $n\to\infty$, this is clearly the difficult case: if n is bounded, then it suffices to show only that there exists a time (or a pair of consecutive time steps, in the case of the Hadamard walk) in which the wave function is supported on at least 2/3 of the vertices.

For the Hadamard walk with measurement $\omega=\tilde{\nu}$, Lemma 9 resolves a conjecture of Kendon and Tregenna [20] based on numerical experiments. Let G^d denote the dth (Cartesian) power of a graph G. Examples are the d-dimensional standard lattice (the dth power of a line) and the d-dimensional periodic lattice (the dth power of a cycle). The following theorem shows how to extend a threshold-mixing result from G to G^d.

Theorem 10 (Mixing on graph powers). Suppose the continuous-time quantum walk $\langle U_{ct}(P(G)),\delta_T\rangle$ threshold-mixes. Then the $O(\log d)$-repeated walk $\langle U_{ct}(P(G^d)),\delta_{Td}\rangle$ threshold-mixes.

Proof. The Hamiltonian $H'=P(G^d)$ is related to the Hamiltonian $H=P(G)$ by the identity

$$H'=\frac{1}{d}\sum_{j=1}^{d}I^{\otimes(j-1)}\otimes H\otimes I^{\otimes(d-j)}. \tag{25}$$

Since H' commutes with the identity I, which can introduce at most a global phase factor to the system, the Markov chain P_t' generated by the walk $\langle U_{ct}(P(G^d)),\delta_t\rangle$ is the dth tensor power of the Markov chain $P_{t/d}$ generated by the walk $\langle U_{ct}(P(G)),\delta_{t/d}\rangle$. By assumption, $d(P_T)\le\alpha$ for some constant $\alpha<1$. By Proposition 3, we can choose $T'=O(\log d)$ to ensure that

$$\frac{1}{2}\|(P_T)^{T'}-u1^{\dagger}\|_1\le d((P_T)^{T'})\le\frac{1}{6d^2}. \tag{26}$$

Then at least $n\sqrt[d]{2/3}$ entries in each column of $(P_T)^{T'}$ are bounded below by $\frac{1-1/2d}{n}$, otherwise we would have the contradiction

$$\frac{1}{2}\|(P_T)^{T'}-u1^{\dagger}\|_1=1-\sum_y\min\left\{(P_T)^{T'}(y,x),\frac{1}{n}\right\}$$
$$\times>(1-\sqrt[d]{2/3})\frac{1}{2d}\ge\frac{1}{6d^2}, \tag{27}$$

where the first equation uses the identity $\frac{1}{2}\|p-q\|_1=1-\sum_k\min\{p_k,q_k\}$ for distributions p,q and the last inequality uses simple algebra along with the fact that for any $d\ge 1$

$$\left(1-\frac{1/3}{d}\right)^d\ge\frac{2}{3}. \tag{28}$$

Since $(P_{Td}')^{T'}=((P_T)^{T'})^{\otimes d}$, at least $(n\sqrt[d]{2/3})^d=\frac{2}{3}n^d$ of the entries in each column of $(P_{Td}')^{T'}$ are bounded below by $(\frac{1-1/2d}{n})^d\ge\frac{1}{2n^d}$. It follows from Proposition 4 that $(P_{Td}')^{T'}$ threshold-mixes in time $O(1)$. We have the following corollary for the dth power $\mathbb{Z}_n^d$ of the cycle $\mathbb{Z}_n$.

Corollary 11 (Mixing on periodic lattices). Let $\mathbb{Z}_n^d$ be the d-dimensional periodic lattice with $n\ge 2$ vertices per side. The $O(\log d)$-repeated continuous-time quantum walk $\langle U_{ct}(P(\mathbb{Z}_n^d)),\omega_{nd/2}\rangle$ with measurement $\omega\in\{\delta,\bar{\mu},\tilde{\mu}\}$ threshold-mixes.

Proof. Combining Lemma 9 with Theorem 10, we conclude that $\langle U_{ct}(P(\mathbb{Z}_n^d)),\delta_T\rangle$ threshold-mixes for any $T\in[\frac{2}{3}\frac{nd}{2},\frac{nd}{2}]$. It is easy to see (cf. Lemma 9) that this is sufficient to imply the stated corollary not only for the measurement $\omega=\delta$, but also for the time-averaged measurements $\bar{\mu}$ and $\tilde{\mu}$.

For $d\ge 1$, this extends the results of Fedichkin *et al.* [21–23] by confirming $O(n)$ and $O(d\log d)$ scaling (suggested by their analytical estimates and numerical experiments in regimes of both high and low decoherence) of the fastest-mixing walk, which they conjectured to be decoherent rather than unitary.

B. The Grover walk

An important question is whether there is a T'-repeated Grover walk $\langle U_{dt}(P(\mathbb{Z}_n^d)), \omega_T\rangle$ that threshold-mixes for $T = O(nd)$ and $T' = O(\log d)$. Szegedy [14] showed that the *phase gap* (minimum nonzero eigenvalue phase from $[-\pi, \pi]$ in absolute value) of $U_{dt}(P)$ is $\Omega(\sqrt{\delta})$ and exploited this property to prove a quadratic quantum speedup for the *hitting time* of any symmetric Markov chain P. A natural adaptation of his argument to the *mixing time* setting would be something like the following: since the phase gap of $U_{dt}(P)$ is $\theta = \Omega(\sqrt{\delta})$, we expect to see by decomposing the action of $U_{dt}(P)$ along spectral components that roughly $O(1/\theta) = O(\sqrt{\delta^{-1}})$ time steps suffice for the orbit of any initial basis state to "cover" the entire state space with sufficient amplitude.

Unfortunately, this argument is incorrect: in fact, the orbit may remain quite localized around the initial basis state. This happens with dramatic effect to the Grover walk on the complete graph $G = K_N$, which mixes in time $T' = \Theta(N)$ (cf [13]) even though the classical random walk on K_N mixes in a single time step. It also happens to the Grover walk on $\mathbb{Z}_n^d$ for $d = 2$, albeit less dramatically [27–29]. [41] The probability distribution p_t induced on the vertices of $\mathbb{Z}_n^2$ at time $t \leq n/2$ is primarily localized at the initial basis state [28] but has substantial secondary spikes which propagate across the lattice in orthogonal directions [27]. In particular, the standard deviation of p_t appears to grow linearly with t [27] and the mixing time of the Grover walk on $\mathbb{Z}_n^2$ with measurement $\omega = \bar{\nu}$ appears to be $T = O(n)$. In the high-dimensional regime, Moore and Russell [11] proved that the Grover walk on $\mathbb{Z}_2^d$ with measurement $\omega = \delta$ mixes almost perfectly in time $T = O(d)$. It seems plausible that the decoherent Grover walk on $\mathbb{Z}_n^d$ with measurement $\omega = \bar{\nu}$ mixes as fast asymptotically as the continuous-time walk.

V. CONCLUSION AND OPEN PROBLEMS

We have shown that decoherent quantum walks have the potential to speed up a large class of classical MCMC mixing processes. Exactly how large this class is, and whether a generic quantum mixing speedup to $O(\sqrt{\delta^{-1}} \log 1/\pi_*)$ is possible, remain important open questions. Since it seems that quantum walks can outperform classical random walks in low-dimensional examples and underperform in very high-dimensional examples (such as the complete graph), a hybrid method may work best for generic Markov chains consisting of both low- and high-dimensional substructures. Also worth investigating is whether by randomizing the "coin" used in discrete-time quantum walks we can improve mixing on the complete graph and other adversarial examples.

ACKNOWLEDGMENTS

I would like to thank Mario Szegedy, Viv Kendon, Frédéric Magniez, Miklos Santha, Iordanis Kerenidis, and Julia Kempe for useful discussions on quantum walks and Todd Brun for helpful comments on the presentation. This material is based upon work supported by the National Science Foundation under Grant No. 0523866. Part of this research was done while the author was visiting LRI, Université Paris-Sud, Orsay, France.

[1] M. Jerrum, A. Sinclair, and E. Vigoda, J. ACM **51**, 671 (2004).
[2] M. Dyer, A. Frieze, and R. Kannan, J. ACM **38**, 1 (1991).
[3] M. Jerrum and A. Sinclair, SIAM J. Comput. **18**, 1149 (1989).
[4] M. Jerrum and A. Sinclair, SIAM J. Comput. **22**, 1087 (1993).
[5] D. Aldous, J. London Math. Soc. **25**(3), 564 (1982).
[6] We say that $f(n) = \Omega(g(n))$ if $g(n) = O(f(n))$, and $f(n) = \Theta(g(n))$ if $f(n) = O(g(n))$ and $f(n) = \Omega(g(n))$.
[7] F. Chen, L. Lovasz, and I. Pak, *Proceedings of the ACM Symposium on Theory of Computing* (1999).
[8] A. Nayak and A. Vishwanath, e-print arXiv:quant-ph/0010117.
[9] A. Ambainis, E. Bach, A. Nayak, A. Vishwanath, and J. Watrous, *Proceedings of the ACM Symposium on Theory of Computing* (2001).
[10] D. Aharonov, A. Ambainis, J. Kempe, and U. Vazirani, *Proceedings of the ACM Symposium on Theory of Computation (STOC'01)* (2001).
[11] C. Moore and A. Russell, *Proceedings of the International Workshop on Randomization and Computation* (2002), e-print arXiv:quant-ph/0104137.
[12] H. Gerhardt and J. Watrous, *Proceedings of the International Workshop on Randomization and Computation* (2003), e-print arXiv:quant-ph/0305182.
[13] P. Richter, New J. Phys. **9**, 72 (2007).
[14] M. Szegedy, *Proceedings of the IEEE Symposium on Foundations of Computer Science* (2004), e-print arXiv:quant-ph/0401053.
[15] L. Grover, *Proceedings of the ACM Symposium on Theory of Computing* (1996), e-print arXiv:quant-ph/9605043.
[16] A. Ambainis, *Proceedings of the IEEE Symposium on Foundations of Computer Science* (2004), e-print arXiv:quant-ph/0311001.
[17] F. Magniez, M. Santha, and M. Szegedy, *Proceedings of the ACM/SIAM Symposium on Discrete Algorithms* (2005), e-print arXiv:quant-ph/0310134.
[18] H. Buhrman and R. Spalek, *Proceedings of the ACM/SIAM Symposium on Discrete Algorithms* (2006), e-print arXiv:quant-ph/0409035.
[19] F. Magniez and A. Nayak, *Proceedings of the International Colloquium on Automata, Languages, and Programming* (2005), e-print arXiv:quant-ph/0506265.
[20] V. Kendon and B. Tregenna, Phys. Rev. A **67**, 042315 (2003).
[21] L. Fedichkin, D. Solenov, and C. Tamon, Quantum Inf. Com-

put. **6**, 263 (2006).
[22] D. Solenov and L. Fedichkin, Phys. Rev. A **73**, 012313 (2006).
[23] D. Solenov and L. Fedichkin, Phys. Rev. A **73**, 012308 (2006).
[24] G. Alagic and A. Russell, Phys. Rev. A **72**, 062304 (2005).
[25] V. Kendon, e-print arXiv:quant-ph/0606016.
[26] In fact this is not quite true: Kendon and Tregenna [20] conjectured this for the discrete-time Hadamard walk. We prove it for both walks.
[27] T. Mackay, S. Bartlett, L. Stephenson, and B. Sanders, J. Phys. A **35**, 2745 (2002).
[28] N. Inui, Y. Konishi, and N. Konno, Phys. Rev. A **69**, 052323 (2004).
[29] B. Tregenna, W. Flanagan, R. Maile, and V. Kendon, New J. Phys. **5**, 83 (2003).
[30] Alagic and Russell [24] exhibit a decoherent quantum walk on the hypercube which converges to the uniform distribution and preserves the quantum mixing speedup proven by Moore and Russell [11], but in contrast to our work, it is their single (carefully chosen) final measurement that *forces* uniform convergence, not the decoherence itself (which only adds "noise" that if small enough does not destroy the speedup).
[31] Though repeatedly measured quantum walks generate mixed states, our analysis is simplest without the introduction of density matrices. We refer the reader lacking sufficient background in quantum computing to the excellent text of Kitaev, Shen, and Vyalyi [42].
[32] E. Farhi and S. Gutmann, Phys. Rev. A **58**, 915 (1998).
[33] D. Aharonov and A. Ta-Shma, *Proceedings of the ACM Symposium on Theory of Computing* (2003), e-print arXiv:quant-ph/0301023.
[34] A. Childs, R. Cleve, E. Deotto, E. Farhi, S. Gutmann, and D. Spielman, *Proceedings of the ACM Symposium on Theory of Computing* (2003) e-print, arXiv:quant-ph/0209131.
[35] J. Kempe, Contemp. Phys. **44**, 307 (2003).
[36] Similarly, we can retrieve a good approximation to π by generating and then measuring a good approximation to the ground state $|\pi\rangle=\Sigma_{x\in\mathcal{S}}\sqrt{\pi(x)}|x\rangle$ of the Hamiltonian $H=DPD^{-1}$ [33].
[37] F. Magniez, A. Nayak, J. Roland, and M. Santha, *Proceedings of the ACM Symposium on Theory of Computing* (2007) e-print, arXiv:quant-ph/0608026.
[38] A. Childs, Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, MA2004.
[39] For a walk on the symmetric group S_n, Gerhardt and Watrous [12] showed that $\frac{1}{2}\|\Pi-u1^{\dagger}\|_1\geq\frac{1}{n}-\frac{1}{n\cdot n!}\binom{2n-2}{n-1}$; for a walk on the hypercube $\mathbb{Z}_2^n$, Moore and Russell showed that there exists an $\epsilon>0$ such that $\frac{1}{2}\|\Pi-u1^{\dagger}\|_1\geq\epsilon$.
[40] Π is also positive semidefinite: it is the Gram matrix of $\{f_s\}$ with $f_s(kl):=\langle s|\phi_k\rangle\langle\phi_l|s\rangle$ if $\lambda_k=\lambda_l$, 0 otherwise.
[41] For $d=1$ it does not; the Grover walk on $\mathbb{Z}_n$ with measurement $\omega=\bar{\nu}$ mixes perfectly in $T=n$ time steps.
[42] A. Kitaev, A. Shen, and M. Vyalyi, *Classical and Quantum Computation* (AMS, Providence, RI2002).