Bounds for mixing time of quantum walks on finite graphs

2010 J. Phys. A: Math. Theor. 43 335302

(http://iopscience.iop.org/1751-8121/43/33/335302)

IOP PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND THEORETICAL

J. Phys. A: Math. Theor. **43** (2010) 335302 (12pp) doi:10.1088/1751-8113/43/33/335302

Bounds for mixing time of quantum walks on finite graphs

Vladislav Kargin

Department of Mathematics, Stanford University, CA 94305, USA

E-mail: kargin@stanford.edu

Received 21 March 2010, in final form 15 June 2010
Published 15 July 2010
Online at stacks.iop.org/JPhysA/43/335302

Abstract

Several inequalities are proved for the mixing time of discrete-time quantum walks on finite graphs. The mixing time is defined differently than in Aharonov *et al* (2002 *Proc. 33rd STOC* (*2001*) (New York: ACM) pp 50–9 arXiv:quant-ph/0012090v2) and it is found that for particular examples of walks on a cycle, a hypercube and a complete graph, quantum walks provide no speedup in mixing over the classical counterparts. In addition, non-unitary quantum walks (i.e. walks with decoherence) are considered and a criterion for their convergence to the unique stationary distribution is derived.

PACS numbers: 03.67.a, 05.40.Fb, 02.50.Ey

1. Introduction

The origin of the concept of quantum walk lies in quantum computation theory, where a quantum version of the classical random walk was invented in an attempt to improve over classical computational algorithms. The early papers that formulated the main ideas of quantum walk are [2] and [11]. Among numerous later papers, we would like to point out [3] where the continuous-time quantum walk was defined and [1] which defined and studied the discrete-time quantum walk on finite graphs. An introductory review of quantum walks written from the prospective of quantum computation can be found in [4]. For recent developments the reader can also consult [8].

From the beginning, it became clear that quantum walks on both finite and infinite graphs have many differences from the classical walk. For example, the probability to find a particle at a particular vertex of a finite graph does not converge to a limit but in general oscillates forever. However, the average of this probability over time does converge to a limit, which can be interpreted as follows. We start the quantum walk in a certain state and measure the particle at a random time t, which is distributed uniformly over interval $[0, T]$. This measurement finds the particle at a particular vertex v with a probability $p(v, T)$, which converges to a limit as $T \to \infty$. How large should T be if we want to make sure that $p(v, T)$ is close to its limit?

Let us introduce some definitions to make this question more precise.

The quantum walk on a finite graph is a 4-tuple (G, S, ψ, U), where $G = (V, E)$ is a finite graph, S is a finite set, ψ is a function in $L^2_{\mathbb{C}}(V \times S)$ and U is a unitary operator on $L^2(V \times S)$. It is assumed that $\|\psi\| = 1$. Elements of S are called *chiralities* and the function $\psi_t = U^t\psi$ is the *wavefunction* at time $t \in \mathbb{Z}$. If a measurement is performed over the system at time t, then the walking particle is found at the vertex v in the state s with probability $|\psi_t(v, s)|^2$.

We assume that the quantum walk is *local*. That is, let x and x' denote pairs (v, s) and (v', s'), respectively. A quantum walk is local if $U_{x'x} \equiv \langle \delta_{x'}, U\delta_x \rangle \neq 0$ implies that $v \sim v'$, that is, the vertices v and v' are connected to each other. A local quantum walk is called the *general quantum walk* in [1].

A special case of the general quantum walk is the *coined quantum walk* ([1]). Here is how it is defined. Let G be a d-regular graph and let $S = \{1, \ldots, d\}$. Assume that the neighbors of each vertex v are labeled as v_i where $i = 1, \ldots, d$. In addition, assume that if $v \neq w$ and $v_i = w_j$, then $i \neq j$. (Such a labeling L always exists on Cayley graphs of finitely generated groups, where we can identify the elements of S with generators and inverses of generators of the group and write $v_g = vg$ and $v_{g^{-1}} = vg^{-1}$. In this case the choice of labeling is equivalent to the choice of ordering of generators and their inverses.)

Define U as follows. Let x and x' denote pairs (v, s) and (v', s'), respectively. If $v' \nsim v$, then $U_{x'x} = 0$. Otherwise, $v' = v_i$ and $U_{x'x} = \delta_{is}C_{s's}$, where C is a unitary matrix acting on $L^2(S)$, which is called the *coin* of the quantum walk. It is easy to check that the matrix U is unitary. Intuitively, let the particle be at vertex v in the state $|s\rangle$. Then, at the next moment the particle will be at the vertex v_s in the superposition state $C|s\rangle$. This is the coined quantum walk on G corresponding to labeling L and coin C.

A typical example of the coined quantum walk is the Hadamard quantum walk on the cycle $\mathbb{Z}_n$. In this case, the coin is the Hadamard transformation:

$$C = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}.$$

Another popular choice of the coin is Grover's transformation:

$$C_{s's} = \left(1 - \frac{2}{d}\right)\delta_{s's} + \left(-\frac{2}{d}\right)(1 - \delta_{s's}).$$

That is, the state s remains unchanged with amplitude $\left(1 - \frac{2}{d}\right)$ and moves to $s' \neq s$ with amplitude $-2/d$. We will call walks with this coin the *Grover quantum walks*.

A generalization of this concept is the *non-unitary quantum walk* [1]. A non-unitary quantum walk is specified by 4-tuple $(G, S, \rho, \mathcal{T})$, where G and S are as before a finite graph and a finite set, ρ is a *density matrix* (i.e. a positive unit-trace operator on $L^2_{\mathbb{C}}(V \times S)$) and $\mathcal{T}$ is a completely positive trace preserving operator acting on density matrices. In the literature, $\mathcal{T}$ is called a *superoperator* [14], or a *quantum channel* [14], or a trace-preserving *quantum operation* [13]. We will use these terms as synonyms. Let x denote a pair (v, s). The probability to find a particle at the vertex v in the state s at time t is given by $\langle x|\mathcal{T}^t\rho|x\rangle$. A non-unitary quantum walk is *local* if $\langle x'|\mathcal{T}(\rho)|x'\rangle > 0$ for $x' = (v', s')$ implies that there is $x = (v, s)$ with $v \sim v'$ such that $\langle x|\rho|x\rangle > 0$. (The concept of locality is more complicated in the non-unitary case and this definition is different from the definition in [1].)

An example of a non-unitary quantum walk is given by a weighted sum of unitary quantum walks. In this example, $\mathcal{T}(\rho) = \sum_{i=1}^k p_i U_i \rho U_i^*$, where U_i are the unitary operators, $p_i > 0$ and $\sum_{i=1}^k p_i = 1$. Intuitively, an operator U_i is used at each step of

the walk with probability p_i. If all U_i are local, then $\mathcal{T}$ is also local. Another example is $\mathcal{T}(\rho) = p\sum_{i=1}^{k} P_i \rho P_i + (1-p)U\rho U^*$, where P_i are the projections and $\sum_{i=1}^{k} P_i = I$. This is a walk in which with probability p the particle is measured and with probability $1-p$ it is evolved according to the unitary operator U.

First, let us consider the case of unitary quantum walks. The probability distribution $|\psi_t(x)|^2$ in general does not converge to any particular limit. Indeed, all eigenvalues of the matrix U have unit absolute value. As a consequence, every eigenvector of U corresponds to a stationary probability distribution. If the initial wavefunction ψ is a non-trivial superposition of the eigenvectors with different eigenvalues, then ψ_t continues to oscillate indefinitely. In the classical case this phenomenon occurs only when the random walk corresponds to a periodic Markov chain, and this case is not typical.

The time averages of the probabilities $|\psi_t(x)|^2$ do converge, and the limit

$$p(x) = \lim_{T\to\infty} \frac{1}{T}\sum_{t=0}^{T-1} |\psi_t(x)|^2$$

exists although may depend on the initial function ψ. We will call this limit the *time-averaged probability distribution* of the particle. In order to quantify the convergence of the initial distribution to this limit, let us define the distance of the initial distribution from its time average by the formula

$$d(T,\psi) = \sum_{x\in V\times S} \left| \frac{1}{T}\sum_{t=0}^{T-1} |\psi_t(x)|^2 - p(x) \right|.$$

This is the total variation distance between the averaged probability distribution at time T and its limit. By analogy with the classical case, the *mixing time* of a general quantum walk is defined as follows:

$$t_{\mathrm{mix}}(\varepsilon) = \sup_{\psi} \inf\{T : d(T',\psi) \leqslant \varepsilon \text{ for all } T' \geqslant T\}.$$

That is, this is the minimal time which is needed to reduce the distance between the worst initial distribution and its time-averaged limit to a quantity less than ε.

Another definition of the mixing time restricts the choice of initial wavefunctions. Namely,

$$\widetilde{t_{\mathrm{mix}}}(\varepsilon) = \sup_{\psi\in\mathcal{B}} \inf\{T : d(T') \leqslant \varepsilon \text{ for all } T' \geqslant T\},$$

where $\mathcal{B}$ is the set of basis states, that is, $\psi \in \mathcal{B}$ if $|\psi|^2$ is a delta-function concentrated at (v,s). This is the definition used in [1]. Clearly, $\widetilde{t_{\mathrm{mix}}}(\varepsilon) \leqslant t_{\mathrm{mix}}(\varepsilon)$. In the case of classical random walks, these two mixing times are always equal to each other. In the case of quantum walks, they can be different.

We intend to estimate the mixing time in terms of the distance between eigenvalues of U. Let $\lambda_k = \mathrm{e}^{\mathrm{i}\beta_k}$, $k = 1,\ldots,m$, be the distinct eigenvalues of U. We define the distance between λ_k and λ_l as the smallest distance along the unit circle:

$$d(\lambda_k,\lambda_l) = \min\{|\beta_k - \beta_l + 2\pi n|,\ n\in\mathbb{Z}\}.$$

The *relaxation time* of the operator U is defined as

$$t_{\mathrm{rel}} = \max_{k\neq l} d(\lambda_k,\lambda_l)^{-1}.$$

Finally, let us define the *overlap* of the two functions φ and ψ by the formula

$$Q(\varphi,\psi) = \sum_{x\in V\times S} |\varphi(x)\,\psi(x)|.$$

Note that if φ and ψ are two wavefunctions, then $0 \leqslant Q \leqslant 1$ by the Cauchy–Schwarz inequality.

Theorem 1.1. *Let U be the unitary transformation on $L^2(V \times S)$ associated with a discrete-time quantum walk (not necessarily coined). Let U have m distinct eigenvalues and the relaxation time t_{rel}. Then,*

$$t_{\mathrm{mix}}(\varepsilon) \leqslant 2\pi \log(2m) \frac{t_{\mathrm{rel}}}{\varepsilon}.$$

(The proofs of all theorems are in the appendix.)

It is interesting to compare this bound with the corresponding result for the classical random walk, where $t_{\mathrm{mix}}(\varepsilon) \leqslant c \log\left(\frac{1}{\varepsilon \pi_{\min}}\right) t_{\mathrm{rel}}$, where $\pi_{\min}$ is the smallest probability in the limit distribution (see for example theorems 12.3 and 12.4 on p 155 in [10]). In many cases the limit distribution is uniform and this bound can be written as $t_{\mathrm{mix}}(\varepsilon) \leqslant c \log(n/\varepsilon)\, t_{\mathrm{rel}}$, where n is the number of vertices in the graph. Note, however, that t_{rel} has a different meaning in the classical case where it denotes the inverse of the difference between 1 (the largest eigenvalue) and the second largest eigenvalue (i.e. the inverse of the 'spectral gap').

Another significant difference in the formulas for the mixing time is that ε enters as $\log \varepsilon$ and ε^{-1} in the classical and quantum cases, respectively. This is due to the fact that the convergence is exponentially fast in the classical case and polynomial (even linear) in the quantum case.

Finally, it is worthwhile to note that in many cases the classical bound $t_{\mathrm{mix}} \leqslant c \log(n)\, t_{\mathrm{rel}}$ is not optimal, and a large literature is devoted to improvement of this result to $t_{\mathrm{mix}} \leqslant c t_{\mathrm{rel}}$ with a sharp constant c.

For the lower bound we prove the following result.

Theorem 1.2. *Let U be the unitary transformation on $L^2(V \times S)$ associated with a discrete-time quantum walk (not necessarily coined). Suppose that U has only real eigenvectors. Let λ and λ' be two distinct eigenvalues with the corresponding eigenvectors ψ and ψ'. Assume that $d(\lambda, \lambda') \leqslant 2$ and $\varepsilon \leqslant Q(\psi, \psi')/80$. Then,*

$$t_{\mathrm{mix}}(\varepsilon) \geqslant \frac{Q(\psi, \psi')}{8\varepsilon d(\lambda, \lambda')}.$$

In particular, if λ and λ' are two eigenvalues with the smallest distance between them along the circle, then $d(\lambda, \lambda')^{-1} = t_{\mathrm{rel}}$ and we obtain the estimate

$$t_{\mathrm{mix}}(\varepsilon) \geqslant \frac{1}{8} Q\left(\psi, \psi'\right) \frac{t_{\mathrm{rel}}}{\varepsilon}$$

valid for all sufficiently small ε.

The main message of theorems 1.1 and 1.2 is that the relation of the mixing and relaxation times in the quantum case is similar to the analogous relation in the classical case. However, the relaxation time is defined differently in the quantum case. It is not the inverse of the difference between the largest and the second largest eigenvalue, but the inverse of the minimal distance between all distinct eigenvalues.

Previously, the speed of convergence of (unitary) discrete-time quantum walks was investigated in [1]. The upper bound for the quantum walks that we obtain in theorem 1.1 is similar to the bound in theorem 6.1 of [1]. The mixing time is $O(t_{\mathrm{rel}} \log(m))$ where t_{rel} is the inverse of the minimal distance between the distinct eigenvalues of the matrix U. The main difference of our result from the result in [1] is that we have $\log(m)$ instead of $\log(n)$, where m and n are the numbers of distinct and all eigenvalues, respectively.

This difference is significant for the case of the discrete walk on the hypercube, where the number of eigenvalues is 2^d and the number of distinct eigenvalues is $d+1$. In particular, we show that the mixing time on the hypercube is $O(n \log n/\varepsilon)$ and not exponential as was suggested in [12] based on previous estimates in [1].

The lower bound that we obtain is in terms of the relaxation time t_{rel}. It essentially states that the mixing time is $\Omega(t_{\text{rel}})$. This bound is different from the bound obtained in [1], which is formulated in terms of a geometrical property of the underlying graph. In addition, the mixing time is defined differently in [1]. As a result, the mixing time of the Hadamard walk on the cycle is of order $O(n \log n/\varepsilon^3)$ in theorem 4.2 of [1], and of order $O(n^2 \log n/\varepsilon)$ in our example 1. In the classical case, the mixing time is of the order $O(n^2 \log(\varepsilon^{-1}))$.

Now let us consider non-unitary quantum walks. The study of these walks helps us to understand how the decoherence affects performance of quantum algorithms. It was noted (see [7]) that decoherence in quantum walks can be useful for quantum algorithms. In particular, it appears that a small amount of decoherence can speed up the mixing of the walk. Numeric evidence in [7] was later corroborated by analytical estimates in [15]. More information about decoherence in quantum walks and additional references can be found in [6].

Let $\mathcal{M}$ denote the linear space of Hermitian linear operators acting on $L^2_{\mathbb{C}}(V \times S)$. The space $\mathcal{M}$ is a Hilbert space with respect to the norm $\|\rho\|_2 = [\text{Tr}(\rho^2)]^{1/2}$ (which we call L^2-*norm*). Other useful norms on $\mathcal{M}$ are $\|\rho\|_1 = \text{Tr}(|\rho|)$ where $|\rho| = \sqrt{\rho^2}$ and $\|\rho\|_{L^2(\mu)} = [\text{Tr}(\mu\rho^2)]^{1/2}$, where μ is a density matrix. We call these norms the *trace* and $L^2(\mu)$ norms, respectively. Superoperators are operators on $\mathcal{M}$ which possess some additional properties. Some well-known properties of superoperators are summarized in the proposition below.

Proposition 1.3. *The superoperator $\mathcal{T}$ is a contraction in the trace norm (i.e. $\|\mathcal{T}(\rho)\|_1 \leqslant \|\rho\|_1$). There exists a density matrix ρ_{st} such that $\mathcal{T}\rho_{st} = \rho_{st}$.*

This proposition is an immediate consequence of theorem 9.2 and exercise 9.9 in [13].

Note that in many cases $\mathcal{T}$ is not self-adjoint in L^2 norm. Moreover, recall that in the classical case the stochastic matrix of a random walk is always self-adjoint with respect to the norm $L^2(\mu)$, where μ is the stationary probability distribution. (This result can be traced to the fact that every random walk is a *reversible* Markov chain.) In contrast, the superoperator of a non-unitary quantum walk $\mathcal{T}$ is not necessarily self-adjoint with respect to the norm $L^2(\rho_{st})$. In fact, it appears that $\mathcal{T}$ is not even a normal operator (i.e. $\mathcal{T}^*\mathcal{T} \neq \mathcal{T}\mathcal{T}^*$) in many situations of interest.

Proposition 1.3 establishes the existence of the stationary density matrix. However, it does not say anything about the uniqueness or convergence properties, and we cannot expect that these properties hold in general. For example, a unitary quantum walk typically has many stationary density matrices and the convergence fails unless we average density matrices over time. The following theorem establishes the uniqueness and convergence properties provided that the quantum walk satisfies a certain condition. Let us call a density matrix ρ *strictly positive* and write $\rho > 0$, if $\langle x, \rho x\rangle = 0$ implies that $x = 0$. Next, let $\mathcal{T}$ be a linear operator acting on $\mathcal{M}$. We will call $\mathcal{T}$ *strongly positive* if for every density matrix ρ there exists an integer $n > 0$ such that $\mathcal{T}^n\rho > 0$.

(This definition is similar to a corresponding definition in the theory of Markov chains, in which it is shown that a stochastic matrix of a Markov chain is strongly positive if and only if the Markov chain is ergodic, that is, aperiodic and irreducible.)

The *multiplicity* of an eigenvalue λ is defined as $\dim\ker(\lambda I - T)$. The *rank* of λ is $\sup_{p>0}\dim\ker(\lambda I - T)^p$. The eigenvalue is called simple if its rank equals 1.

Theorem 1.4. *Let $\mathcal{T}$ be a strongly positive superoperator. Then, (i) $\mathcal{T}$ has a simple eigenvalue 1. (ii) The corresponding eigenvector ρ_{st} is a strictly positive density matrix. (iii) For every initial density matrix ρ_0, $\mathcal{T}^n\rho_0 \to \rho_{st}$, as $n \to \infty$.*

Proof is in the appendix.

After the convergence to the stationary distribution is established, it is natural to ask for an estimate on the mixing time. First, let us define the mixing time for a non-unitary quantum walk. The definition is different from the definition for the unitary walks since no time averaging is necessary. The measurement at time t finds the walking particle at the vertex v in the state s with probability $p_t(x, \rho) = \langle x|\mathcal{T}^t(\rho)|x\rangle$, where x denotes the pair (v, s) and ρ is the initial density matrix. If $\mathcal{T}$ is strongly positive, then these probabilities converge to a limit $p(x) = \langle x|\rho_{st}|x\rangle$, which does not depend on the initial density matrix. Hence, we can define the total variation distance as $d(t, \rho) = \sum_{x\in V\times S}|\langle x|\mathcal{T}^t(\rho)|x\rangle - p(x)|$. The corresponding mixing time can be defined as

$$t_{\text{mix}}(\varepsilon) = \sup_{\rho}\inf\{t : d(t', \rho) \leqslant \varepsilon \text{ for all } t' \geqslant t\}.$$

Unfortunately, while it is easy to see that the asymptotic behavior of $\mathcal{T}^t$ is governed by the spectral radius of $\mathcal{T}$, it is difficult to estimate the mixing time because of the non-normality of the operator $\mathcal{T}$. The essential difficulty is that for such operators it is hard to estimate the duration of the transient behavior. It is the same problem that makes it difficult to estimate the mixing time for non-reversible Markov chains.

(In one particular example of a non-unitary continuous-time walk on cycle this difficulty has been overcome and an estimate on the mixing time has been derived in [15].)

We consider several examples of unitary walks in this paper. The table summarizes results for unitary quantum walks on a complete graph, a cycle and a hypercube.

	Mixing time
Complete graph with n vertices	$\frac{c_1}{\varepsilon} \leqslant t_{\text{mix}}(\varepsilon) \leqslant \frac{c_2}{\varepsilon}$
Cycle with n vertices	$\frac{c_1}{\varepsilon.}n^2 \leqslant t_{\text{mix}}(\varepsilon) \leqslant \frac{c_2}{\varepsilon}n^2\log n$
Hypercube with 2^n vertices	$\frac{1}{2\varepsilon}n \leqslant t_{\text{mix}}(\varepsilon) \leqslant \frac{2\pi}{\varepsilon}n\log n.$

It appears from this table that the mixing time for quantum walks is of similar order as that for the corresponding classical random walks. In particular, the unitary quantum walks do not allow a quadratic speedup over classical walks, in contrast to the results for the mixing time in [1]. The reason for this difference is that the mixing time defined in [1] restricts the initial distributions of the particle to the class of distributions concentrated on a particular vertex of a graph, while we allow for arbitrary initial distributions. Note that this result does not rule out that the quadratic speedup can be achieved by non-unitary quantum walks. Some evidence in favor of this conjecture can be found in [7] and [15].

The rest of the paper is organized as follows. In the next section, we apply bounds on mixing times to particular examples of quantum walks on the cycle, hypercube and complete graph. The proofs of the theorems are relegated to the appendix.

2. Examples

Example 2.1. (Cycle)

Proposition 2.2. *The mixing time for the Hadamard quantum walk on the n-cycle satisfies the following inequalities:*

$$\frac{c_1}{\varepsilon}n^2 \leqslant t_{\mathrm{mix}}(\varepsilon) \leqslant \frac{c_2}{\varepsilon}n^2 \log(n),$$

where c_1 and c_2 are the positive constants.

Proof. The eigenvalues of the Hadamard walk on the cycle with n vertices were found in [1]. They are

$$t_k^{(1,2)} = \frac{1}{\sqrt{2}}\left(\cos\left(\frac{2\pi}{n}k\right) \pm \mathrm{i}\sqrt{1+\sin^2\left(\frac{2\pi}{n}k\right)}\right),$$

where $k = 0, \ldots, n-1$. In order to describe the eigenvectors, let χ_k, $0 \leqslant k \leqslant n-1$, be the functions in $L^2(\mathbb{Z}_n)$ defined by the formula $\chi_k = \sum_{r=0}^{n-1} \exp\left(2\pi \mathrm{i}\frac{kr}{n}\right)\delta_r$. Then all eigenvectors have the form $v \otimes \chi_k$ where v is a 2-vector that depends on k.

Indeed, if S and S^* are the left and right shift operators on $L^2(\mathbb{Z}_n)$, respectively, then we can write U as a 2-by-2 block matrix, with blocks U_{11} and U_{12} equal to $S/\sqrt{2}$, and blocks U_{21} and U_{22} equal to $-S^*/\sqrt{2}$ and $S^*/\sqrt{2}$, respectively. It follows that $U(v \otimes \chi_k) = A_k v \otimes \chi_k$, where

$$A_k = \frac{1}{\sqrt{2}}\begin{pmatrix} \exp\left(-2\pi \mathrm{i}\frac{k}{n}\right) & \exp\left(-2\pi \mathrm{i}\frac{k}{n}\right) \\ -\exp\left(2\pi \mathrm{i}\frac{k}{n}\right) & \exp\left(2\pi \mathrm{i}\frac{k}{n}\right) \end{pmatrix}.$$

Let

$$t_k^{(1,2)} = \frac{1}{\sqrt{2}}\left(\cos\left(\frac{2\pi}{n}k\right) \pm \mathrm{i}\sqrt{1+\sin^2\left(\frac{2\pi}{n}k\right)}\right).$$

Then, eigenvectors of A_k can be written as $v_k^{(\alpha)} = (c, 1)$ with $c = 1 - \sqrt{2}t_k^{(\alpha)}\exp(-2\pi \mathrm{i}k/n)$. The corresponding eigenvectors of U are $v_k^{(\alpha)} \otimes \chi_k$ with eigenvalues $\lambda_k^{(\alpha)} = t_k^{(\alpha)}$ for $\alpha = 1, 2$.

Note that

$$\beta_k^{(1,2)} := \arg\left(t_k^{(1,2)}\right) = \pm\arccos\left[\frac{1}{\sqrt{2}}\cos\left(\frac{2\pi}{n}k\right)\right].$$

The smallest difference between β_k occurs when $k = 0$ and 1 and it can be estimated by c/n^2 for a suitable constant c. It follows that the relaxation time is $t_{\mathrm{rel}} \sim cn^2$, and by theorem 1.1 the mixing time is

$$t_{\mathrm{mix}}(\varepsilon) \leqslant \frac{c_2}{\varepsilon}n^2 \log(n),$$

with a certain constant $c_2 > 0$.

It is easy to estimate the overlap of eigenvectors that correspond to eigenvalues β_0 and β_1. It is greater than 0.97 for all n. By theorem 1.2, we have

$$\frac{c_1}{\varepsilon}n^2 \leqslant t_{\mathrm{mix}}(\varepsilon).$$

□

Example 2.3. (Hypercube)

The mixing time of the quantum walk on a hypercube was previously studied in [12], and we use their setup in the definition of the quantum walk. The quantum walk on the hypercube is also analyzed in [5] with emphasis on the hitting time of the walk.

Consider a hypercube graph $(\mathbb{Z}_2)^n$ with 2^n vertices. We think about vertices as indexed by numbers from 0 to $2^n - 1$ in the binary representation with n digits. The edges of the graph are put between numbers that are different in one bit only. The set of the states S consists of n elements. We consider the Grover quantum walk. That is, a particle at the vertex v in the state s goes to the vertex w which is different from the vertex v only in the bit s. It remains in the state s with amplitude $2/n - 1$ and goes to the state s' with amplitude $2/n$.

Proposition 2.4. *The mixing time for the Grover quantum walk on the n-dimensional hypercube satisfies the following inequalities:*

$$\frac{1}{2\varepsilon} n \leqslant t_{\mathrm{mix}}(\varepsilon) \leqslant \frac{2\pi}{\varepsilon} n \log(n).$$

Proof. The eigenvalues and eigenvectors of the Grover quantum walk on the n-dimensional hypercube were found by Moore and Russell in [12]. The eigenvalues are

$$\lambda_k^{\pm} = 1 - \frac{2k}{n} \pm 2\mathrm{i}\frac{\sqrt{k(n-k)}}{n},$$

where $k = 0, \ldots, n$. We describe eigenvectors below. For the convenience of the reader, we also give a short verification of the result.

For each sequence $t = (t_1, t_2, \ldots, t_n)$ of 0 and 1, define $\chi_t \in L^2\left(\mathbb{Z}_2^n\right)$ by the formula $\chi_t\left(x_1, \ldots, x_n\right) = 2^{-n/2}(-1)^{\sum t_i x_i}$. All eigenvectors of the matrix U have the form $v_t^{(j)} \otimes \chi_t$, where $v_t^{(j)}$ is an n-vector that depends on t, and $j = 1, \ldots, n$.

Indeed, the unitary matrix U can be written as a n-by-n block matrix, in which the ijth block is bS_j if $i \neq j$ and aS_j if $i = j$. Here $a = 2/n - 1$, $b = 2/n$ and $S_k : L^2\left(\mathbb{Z}_2^n\right) \to L^2\left(\mathbb{Z}_2^n\right)$ is the shift operator which acts as follows: $\left(S_k f\right)\left(x_1, \ldots, x_n\right) = f\left(x_1, \ldots, x_k + 1, \ldots, x_n\right)$, where addition is modulo 2. Note that $S_k \chi_t = (-1)^{t_k} \chi_t$.

A computation shows that $U(v \otimes \chi_t) = A(v) \otimes \chi_t$, where A is an n-by-n matrix (which depends on t) with entries $A_{ij} = [\delta_{ij} a + (1 - \delta_{ij}) b](-1)^{t_j}$. In other form, $A = D + bP$, where $D = -\mathrm{diag}((-1)^{t_1}, (-1)^{t_2}, \ldots, (-1)^{t_n})$, and $P = |1, 1, \ldots, 1\rangle \langle(-1)^{t_1}, (-1)^{t_2}, \ldots, (-1)^{t_n}|$.

It is easy to verify that the following vectors are the eigenvectors of A. Let k be the number of non-zero entries in the vector t. First, assume that $n > k \geqslant 1$ and define $x = \pm\,\mathrm{i}\sqrt{\frac{k}{n-k}}$ and $v_r = x^{1-t_r}$. Then $v = (v_1, \ldots, v_n)$ is an eigenvector of A with eigenvalue

$$\lambda_k^{\pm} = 1 - \frac{2k}{n} \pm 2\mathrm{i}\frac{\sqrt{k(n-k)}}{n}.$$

In addition, note that every non-zero vector v such that $v_r = 0$ if $t_r = 0$ and $\sum_{r=1}^n v_r = 0$ is an eigenvector of A with eigenvalue 1. The set of such vectors forms an eigenspace of dimension $k - 1$. Similarly, every non-zero v such that $v_r = 0$ if $t_r = 1$ and $\sum_{r=1}^n v_r = 0$ is an eigenvector of A with eigenvalue -1. The set of such vectors forms an eigenspace of dimension $n - k - 1$.

For the case when $t = 0$, the vector $(1, \ldots, 1)$ is an eigenvector with eigenvalue 1 and its orthogonal complement is eigenspace of -1. For $t = (1, 1, \ldots, 1)$, the situation is reverse.

By counting dimensions of eigenspaces, it is clear that these are all eigenvalues of the matrix A. Since there are 2^n different choices of the vector t, we also found all eigenvalues of the matrix U. It follows that these eigenvalues are ± 1, and $\lambda_k^{\pm}$ for $k = 1, \ldots, n-1$.

From the formula for eigenvalues, the distance between distinct eigenvalues can be estimated from below as $\Delta > \frac{2}{n}$. Hence, $t_{\text{rel}} < \frac{n}{2}$.

By applying theorem 1.1, we find

$$t_{\text{mix}}(\varepsilon) \leqslant \frac{2\pi}{\varepsilon} n \log(n).$$

For the lower bound, consider for simplicity the case of even $n = 2m$. (The case of odd n is similar.) Let $v_{x,k}$ denote the value of the function $v \in L^2((\mathbb{Z}_2)^n \times \mathbb{Z}_n)$ on the vertex x and the state k, and consider the eigenvectors that correspond to eigenvalues λ_m^+ and λ_{m+1}^+ respectively. Then, it is easy to compute the overlap of these eigenvectors as $\sqrt{1-(m+1)/(2m^3)} \sim 1$ for large n. The distance between arguments of the eigenvalues λ_m^+ and λ_{m+1}^- is approximately $2/n$. Hence, by theorem 1.2 we have the inequality

$$t_{\text{mix}}(\varepsilon) \gtrsim \frac{n}{2\varepsilon}.$$

□

Example 2.5. (Complete graph)

There are several ways to define a discrete-time walk on the complete graph with n vertices. We will consider the following variant. Let $|S| = n$. Define the entries of the unitary matrix as follows:

$$U(ws', vs) = \delta_{ws} \left\{ \left(1 - \frac{2}{n}\right) \delta_{vs'} + \left(-\frac{2}{n}\right) (1 - \delta_{vs'}) \right\}.$$

In other words, let the particle start at the vertex v in the state s. Then at the next moment of time it will be at the vertex $w = s$. The particle moves to the state s' with amplitude $(-2/n)$ for $s' \neq v$. If $s' = v$, then the amplitude of the transition is $(1 - 2/n)$.

Proposition 2.6. *The mixing time for the quantum walk on the complete graph satisfies inequalities*

$$\frac{c_1}{\varepsilon} \leqslant t_{\text{mix}}(\varepsilon) \leqslant \frac{c_2}{\varepsilon},$$

where c is a positive constant.

Proof. We will show that the eigenvalues of U are 1, -1, i and $-i$ with multiplicities $n(n-1)/2$, $1 + (n-1)(n-2)/2$, $n-1$ and $n-1$, respectively. Let $X_{vs} = \psi(v, s)$. Then, the action of U can be written as follows:

$$U : X \to X^T \left(I - \frac{2}{n} 1_n 1_n^T \right),$$

where X^T is the transposed matrix X, 1_n is the column n-by-1 vector that consists of all ones and 1_n^T is the corresponding row vector.

If $X = X^T$ and all columns of X sum to 0, then $U(X) = X$. This gives us an eigenspace of the operator U with eigenvalue 1 and dimension $n(n-1)/2$.

Similarly if $X^T = -X$ and all columns of X sum to 0, then $U(X) = -X$. This gives us an eigenspace of U with eigenvalue -1. In addition, $U(I) = -I$. Hence, the dimension of the eigenspace with eigenvalue -1 is $1 + (n-1)(n-2)/2$. In order to find the eigenspaces with eigenvalues $\pm i$, consider U^2. It acts as follows:

$$U^2 \colon X \to \left(I - \frac{2}{n} 1_n 1_n^T \right) X \left(I - \frac{2}{n} 1_n 1_n^T \right).$$

Let $c_1, \ldots, c_n$ and $r_1, \ldots, r_n$ be arbitrary numbers satisfying the conditions $\sum c_i = \sum r_j = 0$, and define $X_{ij} = r_i + c_j$. Then $U^2(X) = -X$. Hence, these matrices belong to the eigenspace of U^2 with eigenvalue -1. The dimension of this space is $2n-2$. It follows that U has two eigenspaces of dimension $n-1$ which correspond to eigenvalues i and $-i$, respectively. By counting dimensions we confirm that we have found all eigenvalues and eigenspaces of the matrix U.

The relaxation time is $t_{\text{rel}} = 1/\sqrt{2}$. By applying theorems 1.2 and 1.1, the mixing time satisfies inequalities

$$\frac{c_1}{\varepsilon} \leqslant t_{\text{mix}}(\varepsilon) \leqslant \frac{c_2}{\varepsilon}$$

for some positive c_1 and c_2. □

Appendix. Proofs of theorems

Proof of theorem 1.1. Let V_k be the eigenspace corresponding to the eigenvalue β_k of the operator U. Then, we can write

$$\psi = \sum_{k=1}^{m} c_k \varphi_k,$$

where $\varphi_k \in V_k$, $\|\varphi_k\| = 1$, $\sum_k |c_k|^2 = 1$. Then at time t,

$$\psi(x,t) = \sum_{k=1}^{m} c_k \varphi_k(x) (\lambda_k)^t,$$

and

$$|\psi(x,t)|^2 = \sum_{k=1}^{m} |c_k|^2 |\varphi_k(x)|^2 + \sum_{k \neq l} c_k \overline{c_l} \varphi_k(x) \overline{\varphi}_l(x) (\lambda_k \overline{\lambda}_l)^t.$$

Hence,

$$\frac{1}{T} \sum_{t=0}^{T-1} |\psi(x,t)|^2 \, dt = \sum_{k=1}^{m} |c_k|^2 |\varphi_k(x)|^2 + \frac{1}{T} \sum_{k \neq l} c_k \overline{c_l} \varphi_k(x) \overline{\varphi}_l(x) \frac{(\lambda_k \overline{\lambda}_l)^T - 1}{\lambda_k \overline{\lambda}_l - 1}.$$

It follows that

$$p(x) := \lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T-1} |\psi(x,t)|^2 = \sum_{k=1}^{m} |c_k|^2 |\varphi_k(x)|^2,$$

and

$$d(T) = \frac{1}{T} \sum_{x \in G} \left| \sum_{k \neq l} c_k \overline{c_l} \varphi_k(x) \, \overline{\varphi}_l(x) \frac{(\lambda_k \overline{\lambda}_l)^T - 1}{\lambda_k \overline{\lambda}_l - 1} \right|.$$

In order to bound this quantity, note that

$$\sum_{x \in G} |c_k \overline{c_l} \varphi_k(x) \overline{\varphi}_l(x)| \leqslant Q(\varphi_k, \varphi_l) |c_k| |c_l| \leqslant |c_k| |c_l|,$$

and, therefore,

$$d(T) \leqslant \frac{2}{T} \sum_{k \neq l} \frac{|c_k| |c_l|}{|e^{i(\beta_k - \beta_l)} - 1|}.$$

Note that $|e^{i(\beta_k-\beta_l)}-1| \geqslant \frac{2}{\pi} d(\beta_k,\beta_l)$ where $d(\beta_k,\beta_l)$ is the distance between β_k and β_l modulo 2π, that is, $d(\beta_k,\beta_l) := \min_s\{|\beta_l-\beta_k+2\pi s|\}$. Let Δ denote $\min_{k,l} d(\beta_k,\beta_l)$ and assume that $0 \leqslant \beta_1 < \beta_2 < \cdots < \beta_m < 2\pi$. Then we can write $d(\beta_k,\beta_l) \geqslant \Delta d(k,l)$, where $d(k,l) = \min\{|k-l|, |k-l+m|, |k-l-m|\}$. This inequality holds because the shortest arc of the circle between β_k and β_l contains $d(k,l)$ non-overlapping intervals whose endpoints are β_s and the length of each of these intervals is at least Δ. It follows that $|e^{i(\beta_k-\beta_l)}-1| \geqslant \frac{2\Delta}{\pi} d(k,l)$.

Next,

$$d(T) \leqslant \frac{\pi}{T\Delta}\sum_{k\neq l}\frac{|c_k||c_l|}{d(k,l)} \leqslant \frac{\pi}{T\Delta}\sum_{k\neq l}\frac{\frac{1}{2}(|c_k|^2+|c_l|^2)}{d(k,l)} = \frac{\pi}{T\Delta}\sum_{k\neq l}\frac{|c_k|^2}{d(k,l)}.$$

This sum can be estimated as follows:

$$\sum_{k\neq l}\frac{|c_k|^2}{d(k,l)} \leqslant \sum_{k=1}^{m}|c_k|^2\left(\sum_{l\neq k}\frac{1}{d(k,l)}\right) = \sum_{k=1}^{m}|c_k|^2\left(\sum_{l=2}^{m}\frac{1}{d(1,l)}\right).$$

We estimate

$$\sum_{l=2}^{m}\frac{1}{d(1,l)} \leqslant 2\sum_{k=1}^{[m/2]}\frac{1}{k} \leqslant 2\left(1+\log\left[\frac{m}{2}\right]\right) < 2\log(2m)$$

and since $\sum_{k=1}^{m}|c_k|^2 = 1$, we have

$$d(T) \leqslant \frac{2\pi}{T\Delta}\log(2m).$$

Hence,

$$t_{\mathrm{mix}}(\varepsilon) \leqslant 2\pi\log(2m)\frac{t_{\mathrm{rel}}}{\varepsilon}.$$

□

Proof of theorem 1.2. Let the initial function be $\varphi = (\psi+\psi')/\sqrt{2}$ and let $\Delta := d(\lambda,\lambda')$. Then we compute

$$d(T) = \frac{1}{2T}\left|1-\cos\Delta T + \frac{\sin\Delta T\sin\Delta}{1-\cos\Delta}\right| Q(\psi,\psi').$$

Consider an interval $I_r = \Delta^{-1}\pi(1/6+2r, 5/6+2r)$, where r is a non-negative integer. For every $T \in I_r$, $\sin(\Delta T) \geqslant 1/2$. Moreover, there is an integer $T_r \in I_r$ because $|I_r| = \Delta^{-1}(2/3)\pi > 1$. Note that the distance between T_r and T_{r+1} is less than $3\Delta^{-1}\pi$.

For every integer T_r we have

$$d(T_r) \geqslant \frac{1}{4T_r}\left(\frac{\sin\Delta}{1-\cos\Delta}\right)Q(\psi,\psi') \geqslant \frac{1}{4T_r\Delta}Q(\psi,\psi').$$

(The second inequality holds because $\Delta \leqslant 2$, and therefore $\sin\Delta/(1-\cos\Delta) \geqslant \Delta^{-1}$.)

It follows that if $\varepsilon > 0$ is sufficiently small then there exists an integer T_r between $Q(\psi,\psi')/4\varepsilon\Delta - 3\Delta^{-1}\pi$ and $Q(\psi,\psi')/4\varepsilon\Delta$, such that $d(T_r) \geqslant \varepsilon$. In particular, if

$\varepsilon \leqslant Q(\psi, \psi')/80$, then $T_r \geqslant Q(\psi, \psi')/4\varepsilon\Delta - 3\Delta^{-1}\pi \geqslant Q(\psi, \psi')/8\varepsilon\Delta$, and we conclude

$$t_{\text{mix}}(\varepsilon) \geqslant \frac{Q(\psi, \psi')}{8\varepsilon\Delta}.$$

□

Proof of theorem 1.4. The proof of (i) and (ii) is an application of results by Krein and Rutman from [9]. In this paper a cone K in a Banach space X is fixed and the operator $A \in L(X)$ is called strongly positive if for every non-zero $x \in K$, there is an integer $n > 0$, such that $A^n x$ is in the interior of K. Theorem 6.3 of this paper (on page 70 of the English translation) shows that if A is compact and strongly positive, then there exists one and only one eigenvector of A in the interior of K and the corresponding eigenvalue exceeds all others in absolute value. Moreover, the proof of the theorem shows that this eigenvalue is simple. The claim of our theorem follows if we apply the Krein–Rutman theorem to the cone of positive-definite matrices. Indeed, by 1.3 there exists ρ_{st} such that $\mathcal{T}\rho_{st} = \rho_{st}$. Since $\mathcal{T}$ is strongly positive, hence ρ_{st} is in the interior of K (i.e. strictly positive); by the Krein–Rutman theorem it is the only eigenvector in the interior of K, and its eigenvalue 1 is simple.

For (iii), let Z be the space of Hermitian matrices with zero trace, $Z = \{\rho : \text{tr}(\rho) = 0\}$. Then, $\mathcal{T}Z \subset Z$ and all the eigenvalues of $\mathcal{T}|_Z$ are less than 1 in absolute value, because 1 is a simple eigenvalue and $\rho_{st} \notin Z$. It follows that the spectral radius η of $\mathcal{T}|_Z$ is smaller than 1. Hence, $\lim_{n\to\infty} \|(\mathcal{T}|_Z)^n\|^{1/n} = \eta < 1$, which implies that $\mathcal{T}^n z \to 0$ for every $z \in Z$. Since $\rho_0 - \rho_{st}$ belongs to Z for every density matrix ρ_0, we conclude that $\mathcal{T}^n\rho_0 \to \rho_{st}$ for every ρ_0. □

References

[1] Aharonov D, Ambainis A, Kempe J and Vazirani U 2002 Quantum walks on graphs *Proc. 33rd STOC (2001)* (New York: ACM) pp 50–9 (arXiv:quant-ph/0012090v2)

[2] Aharonov Y, Davidovich L and Zagury N 1993 Quantum random walks *Phys. Rev.* A **48** 1687–90

[3] Farhi E and Gutmann S 1998 Quantum computation and decision trees *Phys. Rev.* A **58** 915–928

[4] Kempe J 2003 Quantum random walks—an introductory overview *Contemp. Phys.* **44** 302–27 (arXiv:quant-ph/0303081v1)

[5] Kempe J 2005 Discrete quantum walks hit exponentially faster *Probab. Theory Relat. Fields* **133** 215–35 (arXiv:quant-ph/0205083v1)

[6] Kendon V 2007 Decoherence in quantum walks—a review *Math. Struct. Comput. Sci.* **17** 1169–220

[7] Kendon V and Tregenna B 2003 Decoherence can be useful in quantum walks *Phys. Rev.* A **67** 042315

[8] Konno N 2008 Quantum walks *Quantum Potential Theory* (*Lecture Notes in Mathematics* vol 1954) (Berlin: Springer) pp 309–452

[9] Krein M G and Rutman M A 1948 Linear operators leaving invariant a cone in a Banach space *Uspekhi Matematicheskih Nauk* **3** 3–95
Krein M G and Rutman M A 1950 *Uspekhi Matematicheskih Nauk* **26** 1–128 (Engl. Transl.)

[10] Levin D A, Peres Y and Wilmer E L 2009 *Markov Chains and Mixing Times* (Providence, RI: American Mathematical Society)

[11] Meyer D 1996 From quantum cellular automata to quantum lattice gases *J. Stat. Phys.* **85** 551–74

[12] Moore C and Russell A 2002 Quantum walks on the hypercube *Proc. 6th Int. Workshop on Randomization and Approximation Techniques in Computer Science (RANDOM) (Lecture Notes in Computer Science)* (New York: Springer) (arXiv:quant-ph/0104137)

[13] Nielsen M A and Chuang I L 2000 *Quantum Computation and Quantum Information* (Cambridge: Cambridge University Press)

[14] Preskill J 1999 Lecture notes on quantum computation Available at http://www.theory.caltech.edu/people/preskill/ph229/

[15] Richter P C 2007 Quantum speedup of classical mixing processes *Phys. Rev.* A **76** 042306