Quantum Walks On Graphs

Dorit Aharonov[*], Andris Ambainis[†], Julia Kempe[‡] and Umesh Vazirani [§]

December 17, 2000

Abstract

We initiate the study of the generalization of random walks on finite graphs to the quantum world. Such quantum walks do not converge to any stationary distribution, as they are unitary and reversible. However, by suitably relaxing the definition, we can obtain a measure of how fast the quantum walk spreads or how confined the quantum walk stays in a small neighborhood. We give definitions of mixing time, filling time, dispersion time. We show that in all these measures, the quantum walk on the cycle is almost quadratically faster then its classical correspondent. On the other hand, we give a lower bound on the possible speed up by quantum walks for general graphs, showing that quantum walks can be at most polynomially faster than their classical counterparts.

1 Introduction

Markov chains or random walks on graphs have proved to be a fundamental tool, with broad applications in various fields of mathematics, computer science and the natural sciences, such as mathematical modeling of physical systems, simulated annealing, and the Markov Chain Monte Carlo method. In the physical sciences they provide a fundamental model for the emergence of global properties from local interactions. In the algorithmic context, they provide a general paradigm for sampling and exploring an exponentially large set of combinatorial structures (such as matchings in a graph), by using a sequence of simple, local transitions.

In this paper, we initiate a study of quantum walks on graphs — the motivation, as in the case of Markov chains, is to study global properties of a certain structured set, using repeated application of local transition rules. In the quantum setting, though, the local transition rule is defined to be unitary, rather than probabilistic. A classical Markov chain is said to be a random walk on an underlying graph, if the nodes of the graph are the states in S, and a state s has non zero probability to go to t if and only if the edge (s,t) exists in the graph. To define a quantum random walk, in addition to the Hilbert space spanned by the nodes of the graph, we must explicitly introduce the Hilbert space spanned by the outcomes of the coin that control the process. Thus, the quantum walk is allowed to use an auxiliary Hilbert space, in addition to the one spanned by the nodes of the graph. Now, the quantum walk on a graph is naturally defined to be a unitary transformation on the tensor product of the Hilbert space of the graph and the auxiliary Hilbert space, and with the property that the probability amplitude (rather than the probability) is non zero only on edges of the graph.

How do the basic definitions of Markov chains carry over to quantum walks? The most fundamental property of Markov chains is the fact that they converge to a stationary distribution, independent of the initial state. However, by their very definition, quantum walks do not converge to any stationary state. This is due to the fact that unitary matrices preserve the norm of vectors, and hence the distance between the vectors describing the system at subsequent times does not converge to 0. One can ask whether the probability distribution induced

[*]E-mail: doria@cs.berkeley.edu, Computer Science Division, U.C. Berkeley, Berkeley, California, USA, supported by U.C. President's postdoctoral fellowship and NSF grant CCR-9800024

[†]E-mail: ambainis@cs.berkeley.edu, Computer Science Division, U.C. Berkeley, Berkeley, California, USA,supported by Microsoft Graduate Fellowship and NSF grant CCR-9800024

[‡]E-mail: kempe@math.berkeley.edu, Departments of Mathematics and Chemistry, U.C. Berkeley, Berkeley, California, USA

[§]E-mail: vazirani@cs.berkeley.edu, Computer Science Division, U.C. Berkeley, Berkeley, California, USA, supported by NSF grant CCR-9800024

on the nodes of the graph converges in time, but it turns out that this probability distribution does not converge either. However, we can obtain a natural notion of convergence in the quantum case, if we define the limiting distribution as the limit of the *average* of the probability distributions over time. This definition captures the amount of time the walk spends in each subset of the nodes, and moreover, it corresponds to the natural concept of sampling from the graph, since if one measures the state at a random time chosen from the interval $\{1, .., t\}$, the resulting distribution is exactly the average distribution. We show that although in general, the limiting distribution is a function of the initial state of the quantum walk, for Cayley graphs of Abelian groups the limiting distribution is independent of the initial state, and is uniform over the group elements.

The rate at which convergence takes place, called the mixing time, is of crucial importance to the algorithmic applications of classical Markov chains. Given the notion of limiting distribution in the quantum case, we can now talk about mixing times of a quantum walk. A natural definition for mixing time is the time it takes for the average probability distribution to get close to the limiting distribution. We can also talk about measures for how fast the quantum walk spreads or how long it takes the quantum walk to escape from a small neighborhood. We give definitions of quantum mixing time, sampling time, filling time, and dispersion time. How do the various mixing times of quantum walks compare with their classical counterparts? We show that the quantum walk on a cycle converges in time $O(n \log n)$, thereby giving a nearly quadratic speedup over the classical walk. For the cycle this quadratic speed up is the best possible, since the diameter of the graph is clearly a lower bound for the mixing time. How large can the quantum speed up be, for other graphs? We give a general lower bound on the various measures for the quantum mixing time, in terms of the conductance of the underlying graph. Our main result shows that quantum random walks on graphs can be at most polynomially faster than their classical counterparts, and in fact, for bounded degree graphs, the gap is at most quadratic.

It is still an open question whether quantum walks can be used to obtain a quadratic speed up for certain randomized algorithms — such as 2-SAT. Indeed, all quantum algorithms from the last decade — including Shor's celebrated factorization algorithm[1] and Grover's search algorithm[2] — use only quantum Fourier transforms and classical computation. Is it possible to use other types of unitary transformations to design new quantum algorithms? One constraint that must be met is that the unitary transformations must be poly-local — they must be a product of a polynomial number of local unitary transformations. Quantum walks on graphs might provide a good starting point to explore the effects of a sequence of local unitary transformations.

The paper is organized as follows. We first give some background regarding classical Markov chains and the quantum model. We proceed to define quantum Markov chains, and prove various general results about the limiting distribution. We then prove the speed up for the quantum walk on the cycle, which is followed by an upper bound on the mixing time for general graphs. Finally we prove the polynomial lower bound on the speed up for any graph, and conclude with a list of open questions.

Related Work: Ambainis, Bach and Watrous[3] study the various properties of the quantum walk on the infinite line, and in particular show that the variance of this walk is linear in time, as opposed to the square root behavior of the classical case. Nayak and Vishwanath[4] were able to actually calculate the asymptotic behavior of the probability distributions for the walk on the infinite line, and showed that the probability distribution at time t is within a constant in total variation distance from the uniform distribution over an interval which is of length linear in t.

2 Background

2.1 Classical Markov Chains and Random Walks

A simple random walk on an undirected graph $G(V, E)$, is described by repeated applications of a stochastic matrix P, where $P_{u,v} = \frac{1}{d_u}$ if (u, v) is an edge in G and d_u the degree of u. If G is connected and non-bipartite, then the distribution of the random walk, $D_t = P^t D_0$ converges to a stationary distribution π which is independent of the initial distribution D_0. For G which is d−regular, i.e. if all nodes have the same degree, the limiting probability distribution is uniform over the nodes of the graph. There are many definitions which capture the rate at which the convergence to the limiting distribution occurs. A survey can be found in [5].

Definition 2.1 Mixing Time: $M_\epsilon = \min\{T \,|\, \forall t \geq T, D_0 : ||D_t - \pi|| \leq \epsilon\}$, where here and throughout the paper, we use the total variation distance to measure the distance between two distributions d_1, d_2: $\|d_1 - d_2\| = \sum_i |d_1(i) - d_2(i)|$.

Definition 2.2 Filling Time: $\tau_\epsilon = \min\{T | \forall t \geq T, D_0, X \subseteq V : D_t(X) \geq (1-\epsilon)\pi(X)\}$.

Definition 2.3 Dispersion Time: $\xi_\epsilon = \min\{T | \forall t \geq T, D_0, X \subseteq V : D_t(X) \leq (1+\epsilon)\pi(X)\}$.

The mixing time is related to the gap between the (unique) largest eigenvalue $\lambda_1 = 1$ of the stochastic matrix P, and the second largest eigenvalue λ_2.

Theorem 2.4 Mixing time and spectral gap: *[6]*

$$\frac{\lambda_2}{(1-\lambda_2)\log 2\epsilon} \leq M_\epsilon \leq \frac{1}{(1-\lambda_2)}(\max_i \log \pi_i^{-1} + \log \epsilon^{-1}) \tag{1}$$

The mixing time of a random walk on a graph is strongly related to a geometric property of the graph, the **conductance**, denoted by Φ.

Definition 2.5 *Let the* **capacity** C_X *and the* **flow** F_X *of a subset* $X \subset G$ *of the graph* G *be defined as*

$$C_X = \sum_{u \in X} \pi_u \qquad F_X = \sum_{u \in X, v \notin X} p_{u,v} \pi_u. \tag{2}$$

where π *is the stationary distribution, and* $p_{u,v}$ *is the transition probability. Then the* **conductance** *is*

$$\Phi = \min_{\substack{0<|X|<|G| \\ C_X \leq 1/2}} \frac{F_X}{C_X} \tag{3}$$

Theorem 2.6 Conductance and spectral gap:*[Jerrum, Sinclair[7]]*

$$\frac{\Phi^2}{2} \leq (1-\lambda_2) \leq 2\Phi \tag{4}$$

Theorems 2.4 and 2.6 together imply that the mixing time of a Markov chain is bounded between $\Omega(1/\Phi)$ and $O(1/\Phi^2)$.

Example It is well known that for the simple random walk on an $n-$cycle, the mixing time is quadratic, $M_\epsilon = \theta(n^2 \cdot \log(1/\epsilon))$, and so are the filling time and the dispersion time. The conductance of this chain is $1/n$, which gives a lower bound of $\Omega(n)$ time steps for convergence, and an upper bound of $O(n^2)$.

2.2 Quantum Computation

The model. Consider a finite Hilbert space $\mathcal{H}$ with an orthonormal set of basis states $\{|s\rangle\}$ for $s \in \Omega$. The states $s \in \Omega$ may be interpreted as the possible classical states of the system described by $\mathcal{H}$. In general, the state of the system, $|\alpha\rangle$, is a unit vector in the Hilbert space $\mathcal{H}$, and can be written as $|\alpha\rangle = \sum_{s\in\Omega} a_s|s\rangle$, where $\sum_{s\in\Omega} |a_s|^2 = 1$. $\langle\alpha|$ denotes the conjugate transpose of $|\alpha\rangle$. $\langle\beta|\alpha\rangle$ denotes the inner product of $|\alpha\rangle$ and $|\beta\rangle$. A quantum system can undergo two basic operations: unitary evolution and measurement.

Unitary evolution : Quantum physics requires that the evolution of quantum states is unitary, that is the state $|\alpha\rangle$ is mapped to $U|\alpha\rangle$, where U satisfies $U \cdot U^\dagger = I$, and $U^\dagger$ denotes the transpose complex conjugate of U. Unitary transformations preserve norms, can be diagonalized with an orthonormal set of eigenvectors, and the corresponding eigenvalues are all of absolute value 1.

Measurement : We will describe here only a measurement in the orthonormal basis $|s\rangle$. The output of the measurement of the state $|\alpha\rangle$ is an element $s \in \Omega$, with probability $|\langle s|\alpha\rangle|^2$. Moreover, the new state of the system after the measurement is $|s\rangle$.

Combining two quantum systems : If $\mathcal{H}_A$ and $\mathcal{H}_B$ are the Hilbert spaces of two systems, A and B, then the joint system is described by the tensor product of the Hilbert spaces, $\mathcal{H}_A \otimes \mathcal{H}_B$. If the basis states for $\mathcal{H}_A$, $\mathcal{H}_B$ are $\{|a\rangle\}, \{|v\rangle\}$, respectively, then the basis states of $\mathcal{H}_A \otimes \mathcal{H}_B$ are $\{|a\rangle \otimes |v\rangle\}$. We will use the abbreviated notation $|a, v\rangle$ for the state $|a\rangle \otimes |v\rangle$. This coincides with the interpretation by which the set of basis states of the combined system A, B is spanned by all possible classical configurations of the two classical systems A and B.

3 Quantum Markov Chains

3.1 Definitions

Let $G(V, E)$ be a graph, and let $\mathcal{H}_V$ be the Hilbert space spanned by states $|v\rangle$ where $v \in V$. We denote by n, or $|V|$ the number of vertices in G. First assume that G is d-regular. Let $\mathcal{H}_A$ be an auxiliary Hilbert space of dimension d spanned by the states $|1\rangle$ through $|d\rangle$ (we think of this auxiliary Hilbert space as the "coin space"). Let $\mathbf{C}$ be a unitary transformation on $\mathcal{H}_A$ (which we think of as the "coin-tossing operator"). Label each directed edge with a number between 1 and d, such that for each a, the directed edges labeled a form a permutation. For Cayley graphs the labeling of a directed edge is simply the generator associated with the edge. Now we can define a shift operator $\mathbf{S}$ on $\mathcal{H}_A \otimes \mathcal{H}_V$ such that $\mathbf{S}|a, v\rangle = |a, u\rangle$ where u is the a-th neighbor of v. Note that since the edge labeling is a permutation, S is unitary. One step of the quantum walk is given by $U = S \cdot (C \otimes I)$. We call this walk a **coined quantum walk.**

Example: Coined Quantum Walk on the Cycle Consider the graph G which is a cycle with n nodes. This 2-regular graph can be viewed as the Cayley graph of the Abelian group Z_n with the generators $+1$ (denoted by R for right) and -1 (denoted by L for left). The Hilbert space of the walk would then be $\mathcal{C}^2 \otimes \mathcal{C}^n$. We choose the coin tossing operator to be the Hadamard transform,

$$\mathbf{H} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \tag{5}$$

and the shift S is defined by

$$\begin{aligned} \mathbf{S}|R, i\rangle &= |R, i+1 \ mod \ n\rangle \\ \mathbf{S}|L, i\rangle &= |L, i-1 \ mod \ n\rangle \end{aligned} \tag{6}$$

The quantum walk is then defined to be repeated applications of the Hadamard matrix operating on the first register, followed by the shift $\mathbf{S}$. Note that the coin we use corresponds to a classical " unbiased" walk, in the sense that if measured, the walk has an equal chance of moving left or right.

In our most general definition, the **general quantum walk**, we relax our restriction on the exact form of U, and require only that U *respects* the structure of the graph. In other words, we require that, for any v and a, the superposition $U|a, v\rangle$ only contains basis states $|a', v'\rangle$ with $v' \in Q(v) \cup \{v\}$, where $Q(v)$ is the set of adjacent nodes to v. This means that the quantum walk only moves to neighbors of v or stays at v.

3.2 Limiting Distribution

We now discuss the evolution of a quantum walk as a function of time. Starting with an initial state $|\alpha_0\rangle$, the state of the quantum walk at time t is $|\alpha_t\rangle = U^t|\alpha_0\rangle$. In general the limit $\lim_{t\mapsto\infty}|\alpha_t\rangle$ does not exist. The reason being that U, as a unitary transformation, preserves the norm of $|\alpha_t\rangle - U|\alpha_t\rangle$. Consider instead the probability distribution on the nodes of the graph induced by $|\alpha_t\rangle$,

Definition 3.1 $P_t(v|\alpha_0) = \sum_{a\in A} |\langle a, v|\alpha_t\rangle|^2$.

One might ask whether this probability distribution converges to a limit. However, P_t does not converge either. To see this, first observe that the eigenvalues of U are of the form $e^{i\theta}$, and therefore after a finite number of steps, t, $e^{i\theta t}$ is arbitrarily close to 1 simultaneously for all eigenvalues. Hence the evolution of the state is quasi periodic — the state of the system $U^t|\alpha_0\rangle$ is arbitrarily close to $|\alpha_0\rangle$ (and $U^{t+1}|\alpha_0\rangle$ is arbitrarily close to $|\alpha_1\rangle$) for infinitely many times t. As long as the probability distributions at time 0 and 1 are different, $P_0 \neq P_1$, this implies that P_t does not converge.

Despite the fact that the actual distribution does not converge, its average over time does. We define:

Definition 3.2 $\bar{P}_T(v|\alpha_0) = \frac{1}{T}\sum_{t=0}^{T-1} P_t(v|\alpha_0)$

It turns out, as we will see soon, that for any initial state this quantity always has a limit as T grows to infinity, which we denote by $\pi(v)$ (and sometimes write $\pi(v|\alpha_0)$ if we wish to stress its dependence on the initial vector). Intuitively, this quantity captures the proportion of time which the walk "spends" in any given node. Note that it is easy to sample according to this distribution $\bar{P}_T$ using the following process: Uniformly pick a random time t between 0 and $T-1$, let the process evolve for t time steps and then measure to see which node it is at. The node will then be distributed according to $\bar{P}_T$.

We now prove a general statement about the convergence of $\bar{P}_T$. The algebra used to prove this theorem will be useful in the rest of the paper. Let $|\phi_j\rangle$, λ_j denote the eigenvectors and corresponding eigenvalues of U, respectively.

Theorem 3.3 *For an initial state* $|\alpha_0\rangle = \sum_j a_j|\phi_j\rangle$,

$$\lim_{T\mapsto\infty} \bar{P}_T(v|\alpha_0) = \sum_{i,j,a} a_i a_j^* \langle a,v|\phi_i\rangle\langle\phi_j|a,v\rangle$$

where the sum is only on pairs i,j *such that* $\lambda_i = \lambda_j$.

Proof: We start by writing down the probability to measure the basis state $|a,v\rangle$ in $|\alpha_t\rangle$, for a fixed t.

$$|\langle a,v|\alpha_t\rangle|^2 = |\sum_i a_i\lambda_i^t\langle a,v|\phi_i\rangle|^2 = \sum_{i,j} a_i a_j^*(\lambda_i\lambda_j^*)^t\langle a,v|\phi_i\rangle\langle\phi_j|a,v\rangle \tag{7}$$

We now take the average over time of (7), from $t=0$ to $T-1$. The only time dependent term in the above expression is $(\lambda_i\lambda_j^*)^t$. Hence, we are interested in

$$\frac{1}{T}\sum_{t=0}^{T-1}(\lambda_i\lambda_j^*)^t \tag{8}$$

We separate into two cases: One in which $\lambda_i\lambda_j^* = 1$, or equivalently $\lambda_i = \lambda_j$. In this case, we have that the average in Eq. (8) is equal to 1. In all other cases, we can write

$$|\frac{1}{T}\sum_{t=0}^{T-1}(\lambda_i\lambda_j^*)^t| = \frac{|1-(\lambda_i\lambda_j^*)^T|}{|1-\lambda_i\lambda_j^*|}| \leq \frac{2}{T|\lambda_i-\lambda_j|} \tag{9}$$

The latter term converges to zero, therefore the contribution to the limiting distribution comes solely from terms with $\lambda_i = \lambda_j$. Thus we get that the limiting distribution can be derived from the expression in equation 7 by summing only over pairs which correspond to equal eigenvalues. This yields the desired claim, using the fact that the probability to measure a node v is a sum over the probabilities to measure $|a,v\rangle$, and so we can let each term converge separately. □

In the case in which all eigenvalues of U are distinct, the limiting distribution takes a very simple form. Denote by $p_i(v)$ the probability to measure the node v in the eigenstate $|\phi_i\rangle$, so $p_i(v) = \sum_a |\langle a,v|\phi_i\rangle|^2$.

Corollary 3.4 *If all eigenvalues of* U *are distinct, then for an initial state* $|\alpha_0\rangle = \sum_j a_j|\phi_j\rangle$,

$$\lim_{T\mapsto\infty} \bar{P}_T(v|\alpha_0) = \sum_i |a_i|^2 p_i(v).$$

By corollary 3.4, the limiting distribution thus depends on the initial state. However, if all eigenvectors induce uniform distribution over the nodes of the graph, the limiting distribution is uniform, as is easily implied by the theorem. We show:

Theorem 3.5 *Let U be a coined quantum walk on the Cayley graph of an Abelian group, such that all eigenvalues of U are distinct. Then the limiting distribution π is uniform over the nodes of the graph, independent of the initial state $|\alpha_0\rangle$.*

Proof: We derive an explicit expression for the eigenvectors of U, which, for a coined quantum walk, is of the form $U = S \cdot (C \otimes I)$. We note that S is a matrix of dimension dn, for $n = |V|$. S is composed of d blocks, each of dimension n. The $a-$th block corresponds to applying the a-th generator g_a on the group. We note that the characters of the group, $|\chi_k\rangle = \frac{1}{\sqrt{n}} \sum_v \chi_k(v)|v\rangle$, are simultaneous eigenvectors of all the blocks. The eigenvalue associated with applying the a-th block on $|\chi_k\rangle$ is $\chi_k(g_a^{-1})$. Since the application of the coin applies an identity on $\mathcal{H}_V$, a natural guess for the form of the eigenvectors is $(\sum_{a=1}^d c_a|a\rangle) \otimes |\chi_k\rangle$. Applying C and then S on this vector, we find that this vector is an eigenvector of U if $\sum_a c_a|a\rangle$ is an eigenvector of the $d \times d$ matrix $\mathbf{H}_k = \Lambda_k \cdot C$, where Λ_k is a diagonal matrix, with $\Lambda_k(a,a) = \chi_k(g_a^{-1})$. Since $\mathbf{H}_k$, as a product of two unitary matrices, has d orthogonal eigenstates, the tensor products of these eigenstates (which depend on k) with $|\chi_k\rangle$ give d orthonormal eigenstates for U. Running over k, this gives an orthonormal set of nd eigenstates for U. It is easy to see that the probability distribution that these eigenstates induce on the group elements is uniform, since the characters $|\chi_k\rangle$ are uniformly distributed over the group, and since the eigenstates are of the form of a tensor product, the probability to measure a, v in $(\sum_{a=1}^d c_a|a\rangle) \otimes |\chi_k\rangle$ summed over a is just the probability to measure v in $|\chi_k\rangle$. This proves the theorem, using corollary 3.4. $\Box$

We claim that for any quantum walk, if the limiting distribution is independent of the initial node and state of the auxiliary space, then it must be uniform over the nodes.

Claim 3.6 *Consider a quantum walk such that for any initial basis state of the form $|a, v\rangle$, for $v \in V$, the limiting distribution over the nodes of the graph is equal to π. Then π is uniform over the nodes of the graph.*

Proof: If the initial state is chosen randomly from a uniform distribution over all basis states, then the limiting distribution is equal to the average over the limiting distributions for each initial state, but since they are all equal to π, the limiting distribution for the uniform mixture is π. However, the density matrix which represents a complete mixture, i.e. a uniformly random basis state of the space spanned by $|a, v\rangle$ is preserved under unitary transformation, since the unitary matrix maps this space into itself. Hence for any time t it induces a uniform probability distribution over the nodes in the graph, because the initial density matrix induces this distribution. This means that the limiting probability distribution starting from the complete mixture is uniform. Combining the two facts together, we get that π is uniform. $\Box$

3.3 Mixing Times

We first define the analogue of the classical notion of mixing time:

Definition 3.7 Mixing time: *The mixing time M_ϵ, of a quantum Markov chain is*

$$M_\epsilon = \min\{T | \forall t \geq T, |a, v\rangle : \|\pi(\cdot|a, v) - \bar{P}_t(\cdot|a, v)\| \leq \epsilon\}.$$

where by the notation $P(\cdot|a, v)$ we mean the probability distribution conditioned on the initial state being $|a, v\rangle$. This quantity measures the number of time steps required for the average distribution to be ϵ-close to the limiting distribution, starting from a basis state.

We next define a closely related quantity which we call sampling time:

Definition 3.8 Sampling time: *The Sampling time S_ϵ, of a quantum Markov chain is*

$$S_\epsilon = \min\{T | \forall t \geq T, |a, v\rangle, X \subseteq V : |\pi(X|a, v) - \bar{P}_t(X|a, v)| \leq \epsilon\pi(X|a, v)\}.$$

This is the time it takes for the walk to approximate the limiting distribution point-wise. Sampling at a random time between 0 and $S_\epsilon - 1$ results in a distribution which is ϵ-close point-wise to the limiting distribution, justifying the term *sampling time.* In the same sense, sampling at a random time between 0 and $M_\epsilon - 1$ results in a distribution which is ϵ-close to the limiting distribution in total variation distance.

The third quantity, namely the *filling time* of the quantum Markov chain is defined as the first time at which the walk can claim to have visited all sets with at least $(1-\epsilon)$ the correct proportion:

Definition 3.9 Filling time: *The filling time, τ_ϵ, of a quantum Markov chain is*

$$\tau_\epsilon = \min\{T | \, \forall X \subseteq V, |a,v\rangle \; \exists t \leq T : \; P_t(X|a,v) \geq (1-\epsilon)\pi(X|a,v).$$

We also define the *dispersion time*, which is in some sense the opposite definition to filling time:

Definition 3.10 Dispersion time: *The dispersion time, ξ_ϵ, of a quantum Markov chain is*

$$\xi_\epsilon = \min\{T | \; \forall X \subseteq V, |a,v\rangle \; \exists t \leq T : \; P_t(X|a,v) \leq (1+\epsilon)\pi(X|a,v).$$

This quantity measures how fast the quantum walk escapes any subset of the nodes.

Remark: We note that one could consider all the above definitions of mixing times with an arbitrary initial state, $|\alpha_0\rangle$, and not restrict the initial state to be a basis state of the form $|a,v\rangle$. However, the mixing time could change significantly. We will see in the cycle example that the mixing time is almost linear for initial basis states of the form $|a,v\rangle$, but it is actually quadratic for general initial states.

The above definitions can be related one to another in various ways. First, it turns out that the sampling time is an upper bound on the mixing time, the filling time and the dispersion time:

Theorem 3.11 $M_\epsilon, \xi_\epsilon, \tau_\epsilon \leq S_\epsilon$.

Proof: Fix a subset of the nodes X, and an initial state $|a,v\rangle$. Suppose at all times before S_ϵ, $P_t(X|a,v) < (1-\epsilon)\pi(X|a,v)$. Then the average at time S_ϵ of the probability to measure X is less than $(1-\epsilon)\pi(X|a,v)$. But by definition of the sampling time this is a contradiction. Hence there exists some time before S_ϵ at which the probability for the measurement outcome to be a node in X is $P_t(X|a,v) \geq (1-\epsilon)\pi(X|a,v)$, and since this is true for all X, we have $\tau_\epsilon \leq S_\epsilon$. We argue in exactly the same way to prove $\xi_\epsilon \leq S_\epsilon$. The statement $M_\epsilon \leq S_\epsilon$ follows trivially from the definition of total variation distance. □

We will later define amplified versions of these quantities, and find more relations between them. Let us first proceed to give an upper bound on the mixing time M_ϵ for the quantum walk on the cycle.

4 Quantum Walk On the Cycle

In subsection 3.1, we have defined the coined quantum walk on the cycle. We restrict the discussion to cycles of an odd number of nodes n. We first show that the limiting distribution for this walk is uniform.

Theorem 4.1 *The limiting distribution π for the coined quantum walk on the n-cycle, with n odd, and with the Hadamard transform as the coin, is uniform on the nodes, independent of the initial state $|\alpha_0\rangle$.*

Proof: To prove that the limiting distribution is uniform, by theorem 3.5 it suffices to show that all eigenvalues of U are different. By the proof of theorem 3.5, the set of eigenvalues of U consists of all eigenvalues of the matrices:

$$\mathbf{H}_k = \begin{pmatrix} \omega^k & 0 \\ 0 & \omega^{-k} \end{pmatrix} \cdot \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \omega^k & \omega^k \\ \omega^{-k} & -\omega^{-k} \end{pmatrix} \tag{10}$$

where $\omega = e^{\frac{2\pi i}{n}}$. We now show that the eigenvalues of $\mathbf{H}_k$ are distinct. The eigenvalues of $\mathbf{H}_k$ are the roots of the following quadratic equation:

$$\lambda^2 - i\sqrt{2}sin(\frac{2\pi k}{n})\lambda - 1 = 0 \tag{11}$$

The two solutions to this equation are of the form $e^{i\theta_k}$, with θ_k being one of the two solutions for the following equation:

$$sin(\theta_k) = \frac{\sin(\frac{2\pi k}{n})}{\sqrt{2}}. \tag{12}$$

In particular, $|\sin(\theta_k)| \leq \frac{1}{\sqrt{2}}$ which means that the roots of the quadratic equation 11 are confined to two regions of the unit circle, $\theta_k \in [-\pi/4, \pi/4]$ and $\theta_k \in [3\pi/4, -3\pi/4]$. There are two solutions for equation 12, $\theta_{k,1}$ and $\theta_{k,2}$ where $\theta_{k,1} = \pi - \theta_{k,2}$, so they lie in different regions, and in particular, they are distinct. To get equality between eigenvalues coming from different k's, we have to have $\sin(\frac{2\pi k}{n}) = \sin(\frac{2\pi k'}{n})$, which implies that either $k = k'$ or $k + k' = n/2$. The latter equation has no solutions for odd n, which implies the theorem. $\square$

Theorem 4.2 *For the quantum walk on the n-cycle, with n odd, with the Hadamard coin, we have*

$$M_\epsilon \leq O(\frac{n \log n}{\epsilon^3}).$$

Proof: We prove an upper bound on the mixing time M_ϵ, i.e. we give an upper bound on the total variation distance between the average distribution $\bar{P}_T$ and the limiting distribution π. This is done using the following lemma which holds for any quantum walk.

Lemma 4.3 *Consider a general quantum walk specified by the unitary matrix U, and let ϕ_i, λ_i be the eigenvectors and corresponding eigenvalues of U, respectively. For any initial state $|\beta_0\rangle = \sum_i a_i|\phi_i\rangle$, the total variation distance between the average probability distribution and the limiting probability distribution satisfies*

$$||\bar{P}_T(\cdot|\beta_0) - \pi(\cdot|\beta_0)|| \leq 2 \sum_{i,j,\lambda_i \neq \lambda_j} |a_i|^2 \frac{1}{T|\lambda_i - \lambda_j|}$$

Proof: We recall that in the proof of lemma 3.3 we have already bounded the time dependent term in the average probability distribution. From equations (7) and (9) we have that

$$|\bar{P}_T(v) - \pi(v)| \leq \sum_{a,i,j,\lambda_i \neq \lambda_j} |a_i| \cdot |a_j^*| \cdot |\langle a, v|\phi_i\rangle| \cdot |\langle \phi_j|a, v\rangle| \frac{2}{T|\lambda_i - \lambda_j|} \tag{13}$$

We now use $|2ab| \leq |a|^2 + |b|^2$ twice, and summing over v we get that

$$||\bar{P}_T - \pi|| \leq \sum_{v,a,i,j,\lambda_i \neq \lambda_j} (\frac{|a_i|^2 + |a_j^*|^2}{2}) \cdot (\frac{|\langle a, v|\phi_i\rangle|^2 + |\langle \phi_j|a, v\rangle|^2}{2}) \frac{2}{T|\lambda_i - \lambda_j|} \tag{14}$$

Summing first over all v and a we get the desired bound. $\square$

We observe that in that lemma, the distances $|\lambda_i - \lambda_j|$ are of crucial importance, and they need to be large for the convergence time to be small. By the proof of theorem 4.1, the eigenvalues are distributed in two regimes (which we will call R and R') of the complex unit circle, $\theta_{k,1} \in [-\pi/4, \pi/4] = R$ and $\theta_{k,2} \in [3\pi/4, -3\pi/4] = R'$. Near the boundaries of these regimes, i.e. for those θ's coming from k's in the vicinity of $n/4, 3n/4$ modulo n, the distance between two adjacent eigenvalues can be of the order of $1/n^2$. However, we claim that the contribution of these problematic eigenvalues is small, and that for the rest of the eigenvalues, the distance is of order $1/n$. We fix $0 < \delta < 1$ (which will later be related to ϵ) and define

$$\begin{aligned} R_\delta &= [0, (1-\delta)\frac{\pi}{2}] \cup [(1+\delta)\frac{3\pi}{2}, 2\pi] \\ R'_\delta &= [(1+\delta)\frac{\pi}{2}, (1-\delta)\frac{3\pi}{2}] \end{aligned} \tag{15}$$

These two regimes together cover the entire interval $[0, 2\pi]$ except for a 2δ portion of it. We refer to k such that $\frac{2\pi k}{n}$ is in one of these regimes as "δ-good", and other k's are "δ-bad". We also refer to eigenvectors and eigenvalues associated with "δ-good" k's as "δ-good", and similarly for "δ-bad". We will later show that if the initial state for the walk is a basis state, the contribution of the bad eigenvectors is small, because the projection of basis states on bad eigenvectors is small. But first, let us restrict our attention to an initial state which is a superposition of good eigenvectors, and consider the convergence to limiting distribution in this case. We first give a lower bound on the spacing between good eigenvalues.

Definition 4.4

$$\Delta_\delta = \min_{i,j}\{|\lambda_i - \lambda_j| \ s.t. \ i \neq j\}$$

where i, j run only on δ good eigenvalues.

Claim 4.5 *For the quantum walk on the odd n cycle with the Hadamard coin, $\Delta_\delta \geq \frac{\pi\delta}{n}$.*

Proof: First observe that if λ_i, λ_j originate from the same k, then they lie in two different regimes R and R', which means that $|\lambda_i - \lambda_j|$ is at least $\sqrt{2}$. Hence, we can restrict our attention to eigenvalues coming from different k's. Let λ_i, λ_j originate from k, k', respectively. Then using equation 12 we have

$$|\lambda_i - \lambda_j| \geq |\sin(\theta_i) - \sin(\theta_j)| = \frac{1}{\sqrt{2}}|\sin(\frac{2\pi k}{n}) - \sin(\frac{2\pi k'}{n})| \tag{16}$$

We separate the proof to two cases. In the first case, k, k' lie in the same regime, R_δ or R'_δ. Recall the intermediate value theorem, which states that for a continuous function, for any $x \leq y$, there exists $x \leq z \leq y$ such that $|f(x) - f(y)| = |f'(z)(x - y)|$. Applying this theorem with $f(x) = \sin(x)$, we get

$$|\sin(\frac{2\pi k}{n}) - \sin(\frac{2\pi k'}{n})| = |\cos(\gamma)(\frac{2\pi(k-k')}{n})| \geq |\cos(\gamma)\frac{2\pi}{n}| \tag{17}$$

for some $\frac{2\pi k}{n} \leq \gamma \leq \frac{2\pi k'}{n}$. Since k, k' are in the same regime, then $\gamma \in R_\delta$ or $\gamma \in R'_\delta$, and by monotonicity of the cos function, we have:

$$|\cos(\gamma)| \geq |\cos(\frac{\pi(1-\delta)}{2})| = |\sin(\frac{\delta\pi}{2})| \geq \delta \tag{18}$$

where the last equality follows from the fact that $\sin(0) = 0, \sin(\pi/2) = 1$, and sin is convex in the regime $[0, \pi/2]$. If k, k' belong to different regimes, then we can no longer claim that $\cos(\gamma)$ is large. Instead, we write

$$|\sin(\frac{2\pi k}{n}) - \sin(\frac{2\pi k'}{n})| = |\sin(\frac{2\pi k}{n}) - \sin(\pi - \frac{2\pi k'}{n})| \geq |\cos(\gamma')(\frac{2\pi(k-k')}{n} - \pi)| \geq |\cos(\gamma')\frac{\pi}{n}| \tag{19}$$

for some γ', between $\frac{2\pi k}{n}$ and $\pi - \frac{2\pi k'}{n}$. Now, $\frac{2\pi k}{n}$ and $\pi - \frac{2\pi k'}{n}$ lie in the same regime, and so using the same argument as before, the lemma follows. □

We can now use claim 4.5 to give a better lower bound on the distance between two eigenvalues.

Claim 4.6 *Let us order the eigenvalues such that $0 \leq Arg(\lambda_1) \leq Arg(\lambda_2) .... \leq Arg(\lambda_{2n}) \leq 2\pi$. Consider λ_i and λ_j which lie in the same regime, R_δ or R'_δ. Then*

$$|\lambda_i - \lambda_j| \geq \frac{2\sqrt{2}}{\pi}|i - j|\Delta_\delta$$

Proof: Consider i, j as in the requirements of the claim. Let $L_{i,j}$ be the length of the shorter arc on the unit circle that connects λ_j to λ_i. We first claim that $|\lambda_i - \lambda_j| \geq \frac{2\sqrt{2}}{\pi}L_{i,j}$ in our regime. This is true since the ratio $|\lambda_i - \lambda_j|/L_{i,j}$ is monotonically decreasing in $\theta_i - \theta_j$ in the regime $\theta_i - \theta_j \in [0, \pi/2]$, and so we can bound $|\lambda_i - \lambda_j|/L_{i,j}$ from below by its value on the boundary, $|\theta_i - \theta_j| = \pi/2$, which gives $|\lambda_i - \lambda_j|/L_{i,j} = \frac{2\sqrt{2}}{\pi}$. Hence, to bound $|\lambda_i - \lambda_j|$ we give a lower bound on $L_{i,j}$. We have $L_{i,j} = |i - j|L_{i,i+1}$, so it suffices to bound $L_{i,i+1}$. This is done by noticing that $L_{i,j} \geq 2|\sin L_{i,j}/2| = |\lambda_i - \lambda_j| \geq \Delta_\delta$, where the first inequality uses $\sin(x) \leq x$, the second equality uses simple trigonometry, and the last inequality uses claim 4.5. Hence, $L_{i,j} = |i - j|L_{i,i+1} \geq |i - j|\Delta_\delta$ which combined with $|\lambda_i - \lambda_j| \geq \frac{2\sqrt{2}}{\pi}L_{i,j}$ gives the claim. □

Claim 4.7 *Let $|\beta\rangle = \sum_i a_i|\phi_i\rangle$ such that all coefficients of δ bad eigenvectors are zero. Then*

$$\|\bar{P}_T^\beta - \pi^\beta\| \leq \frac{\sqrt{2}n(ln(n)+2)}{T\delta}$$

Proof: Using lemma 4.3 we write

$$\|\bar{P}_T^\beta - \pi^\beta\| \leq 2\sum_k \sum_{i,j,|i-j|=k}^{\prime} |a_i|^2 \frac{\pi}{Tk2\sqrt{2}\Delta_\delta} + 2\sum_{i,j}^{\prime\prime} |a_i|^2 \frac{1}{T\sqrt{2}} \tag{20}$$

where in the first sum the prime indicates the fact that we sum over pairs i,j in the same regime R_δ or R'_δ, such that $\lambda_i \neq \lambda_j$, and in the second sum the double prime indicates that the sum is over pairs i,j such that λ_i, λ_j are in the different regimes. We have used claim 4.6 in the first sum, and the fact that $|\lambda_i - \lambda_j| \geq \sqrt{2}$ for eigenvalues from different regimes in the second sum.

To bound the first term, we observe that for each i there are at most two eigenvalues j such that $|i-j| = k$. We first sum over i,j. Then, summing over k, we use the fact that the sum of the first n terms of the harmonic series is less than $\ln(n)+1$. Thus, the first term is at most $\frac{\sqrt{2}\pi(\ln(n)+1)}{T\Delta_\delta}$. For the second term we get an upper bound $\frac{\sqrt{2}n}{T}$. Using claim 4.5 to bound Δ_δ in the first term we get the desired claim. □

We now prove that the contribution of the δ bad vectors is small. This follows from the following two claims.

Claim 4.8 *The projection of any basis state on the bad eigenvectors is of norm squared at most 2δ.*

Proof: The $2n$ dimensional Hilbert space of the quantum walk can be viewed as a direct sum of the two dimensional subspaces L_k, where L_k is the space spanned by the two eigenvectors originating from k. The projection of a basis state on L_k is of norm squared exactly $\frac{1}{n}$. The claim follows from the fact that there are $2\delta n$ bad $k's$. □

Claim 4.9 *Consider two initial states, $|\alpha_0\rangle, |\beta_0\rangle$. Denote by $\bar{P_T}^\alpha$,$\bar{P_T}^\beta$ the average distributions in the quantum walk starting with $|\alpha_0\rangle, |\beta_0\rangle$, respectively. Then for all T, the total variation distance between the average distribution is bounded by the distance between the initial states:*

$$\|\bar{P_T}^\alpha - \bar{P_T}^\beta\| \leq 2\||\alpha_0\rangle - |\beta_0\rangle\|$$

Proof: Denote by $|\alpha_t\rangle, |\beta_t\rangle$ the states at time t starting with $|\alpha_0\rangle, |\beta_0\rangle$ as initial states. Denote by P_t^α, P_t^β the induced distributions of $|\alpha_t\rangle, |\beta_t\rangle$ on the nodes of the graph. Clearly, $\|\bar{P_T}^\alpha - \bar{P_T}^\beta\| \leq \max_{t\leq T}\|P_t^\alpha - P_t^\beta\|$. Due to unitarity of the walk, the distance is preserved: $\||\alpha_t\rangle - |\beta_t\rangle\| = \||\alpha_0\rangle - |\beta_0\rangle\|$. By lemma 11 in [9], the total variation distance between the two probability distributions resulting from a measurement on two states which are ϵ apart, is at most 2ϵ. This proves the claim. □

Claim 4.10 *Let $|\alpha_0\rangle$ be the initial basis state, and $|\beta_0\rangle$ be the initial basis state projected on the $\delta-$good eigenvectors, and renormalized. Then*

$$\|\bar{P}_T^\alpha - \pi^\alpha\| \leq 8\sqrt{2\delta} + \|\bar{P}_T^\beta - \pi^\beta\|$$

Proof: We write

$$\|\bar{P}_T^\alpha - \pi^\alpha\| \leq \|\bar{P}_T^\alpha - \bar{P}_T^\beta\| + \|\bar{P}_T^\beta - \pi^\beta\| + \|\pi^\beta - \pi^\alpha\|. \tag{21}$$

The first term, by claim 4.9 is smaller than $2\||\alpha_0\rangle - |\beta_0\rangle\|$. The last term is also smaller than $2\||\alpha_0\rangle - |\beta_0\rangle\|$, since it is the limit of distances which are smaller than this term. We claim that $\||\alpha_0\rangle - |\beta_0\rangle\| \leq 2\sqrt{2\delta}$. This is true since by claim 4.8 we can write $|\alpha_0\rangle = a|\beta_0\rangle + |v\rangle$, where $|v\rangle$ is a vector of norm at most $\sqrt{2\delta}$ and a is larger than $\sqrt{1-2\delta}$. $\||\alpha_0\rangle - |\beta_0\rangle\| \leq |1-a| + \sqrt{2\delta} \leq 2\sqrt{2\delta}$. □

We can now combine claim 4.10 and claim 4.7 to finish the proof of the Theorem. If we now pick $\delta = \frac{1}{2}(\epsilon/16)^2$, and $T \geq \sqrt{2}n(ln(n)+2)/\epsilon\delta$, we get that

$$\|\bar{P}_T^\alpha - \pi\| \leq 8\sqrt{2\delta} + \frac{n(ln(n)+2)}{\sqrt{2}T\delta} \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon, \tag{22}$$

which is the desired result. $\Box$

Remark In the theory of classical Markov chains, the distance between the first and second eigenvalues plays a crucial role in mixing time analysis. In the quantum case, we see that the distances between eigenvalues play a similarly important role; However, unlike in the classical case, since all eigenvalues of a unitary matrix are of absolute value 1, there is no special eigenvector which plays the role of the fixed state, and all eigenvalues play an equally important role.

5 Amplification

In classical Markov chains, after approaching a certain closeness to the limiting distribution, the distance to the limiting distribution starts to drop exponentially. Theorem 4.2 gives polynomial dependence on $1/\epsilon$. This means that simply running the walk for longer time, does not achieve this exponential improvement. However, one can amplify the closeness in a very simple way. Suppose the limiting distribution $\pi(\cdot|a, v)$ is independent of the initial node v and the state a, and is equal to π. (Recall that by claim 3.6 π is uniform.) In this case, the closeness to π can be amplified in a standard way to get logarithmic dependence on $1/\epsilon$. This is done by running the walk for M_ϵ steps (i.e. for a random time between 0 and $M_\epsilon - 1$) and then measuring the node. If the measured node is v, we then initialize the state to be $|a, v\rangle$ with a random auxiliary state a, and start the walk again for one more stage of M_ϵ steps, and so on for k times. We claim:

Lemma 5.1 Amplification lemma *Running the quantum walk for k amplification steps, each lasting M_ϵ time steps, results in a distribution which is ϵ^k close to the limiting distribution.*

Proof: Define $P_{v,u}$ to be the probability to measure the node u starting from a random initial basis state $|a, v\rangle$, in one amplification step, where a is randomly chosen from all basis states of the auxiliary space. The matrix P defined by these transition probabilities is a stochastic matrix. We claim that applying one amplification step starting from the uniform distribution π one gets the distribution π again. The reason is that a uniformly random state $|a, v\rangle$ (which induces a uniform distribution over the nodes v) is a complete mixture of the Hilbert space in which the walk evolves. The unitary transformation associated with the walk is a map from this space to itself, therefore, starting from a complete mixture of this space, the state of the system remains a complete mixture, i.e. π is preserved.

We now claim that the L_1 norm of any vector orthogonal to π is shrunk by a factor of ϵ by the matrix P. To prove that $\|\tau P\| \leq \epsilon \|\tau\|$ for $\tau \perp \pi$, observe that by definition of M_ϵ, for any distribution $\sigma = \pi + \tau$, $\|\sigma P - \pi\| \leq \epsilon$. This means that for any vector τ for which the sum of elements is zero, and each coordinate is at least $-1/n$, we have $\|\tau P\| \leq \epsilon$. We can define a basis for the subspace orthogonal to π, which is composed of such vectors: These will be the vectors v_i, where v_i has $(n-1)/n$ on its i^{th} coordinate and the rest are all equal to $-1/n$. Any $\tau \perp \pi$ can be written as a sum of v_i and π: $\tau = \sum_i \tau_i (v_i + \pi) = \sum_i \tau_i v_i$. $\|\tau P\| \leq \sum_i |\tau_i| \|v_i P\| \leq \sum_i |\tau_i| \epsilon = \|\tau\| \epsilon$.

We can now prove the claim by induction. Starting from a distribution σ which is within δ total variation distance from π, we can write $\sigma = \pi + \tau$ where $\tau \perp \pi$, and $\|\tau\| \leq \delta$. Then, $\|\sigma P - \pi\| = \|\pi P + \tau P - \pi\| = \|\tau P\| \leq \|\tau\| \epsilon \leq \delta \epsilon$. $\Box$

For the cases in which the limiting distribution is independent of the initial state, we can now generalize our measures of convergence to allow the possibility of amplification. This means that in all definitions we allow a warm start, i.e. we first amplify for several amplification steps, which all together last T_A time steps, to get an initial "warm" node (the exact times at which one measures are chosen so that T_A is minimized, and the times at which the node is measured during T_A are referred to as the "amplification scheme"). Then we apply the various definitions of mixing times with the "warm start" node as the initial node. However, to account for the initial amplification stage, we add T_A to the mixing times. We denote those amplified versions of convergence with primes. M'_ϵ, S'_ϵ, and so on. We have:

Theorem 5.2 $\tau'_\epsilon, \xi'_\epsilon \leq S'_\epsilon \leq \frac{\log(1/\min_v\{\pi(v)\})}{log(1/\epsilon)} M_\epsilon$

Proof: We first prove the right inequality. Closeness to within ϵ point wise will be guaranteed if the total variation distance is at most $\epsilon \min_v\{\pi(v)\}$. For that, by the amplification lemma, it suffices to apply

$\log(1/\min \pi(v))/log(1/\epsilon)$ amplification steps, each of length M_ϵ. The proof of the left inequality is as follows. Let S'_ϵ be achieved with a certain amplification scheme. We then use the same amplification scheme for the dispersion time and the filling time, so that we start with the same distribution over initial nodes. Now, the remaining of the proof is exactly as the proof of theorem 3.11, referring only to the time interval starting at the end of the amplification stage. $\square$

Theorem 5.3 *For the quantum walk on the n-cycle, with n odd, with the Hadamard coin, we have*

$$M'_\epsilon \leq O\big(n \log n log(\frac{1}{\epsilon})\big) \ , \ S'_\epsilon, \xi'_\epsilon, \tau'_\epsilon \leq O\big(n \log^2 n \log\big(\frac{1}{\epsilon}\big)\big)$$

proof: The upper bound on $S'_\epsilon, \tau'_\epsilon$ and ξ'_ϵ follows from theorems 5.2 and 4.2, and the observation that $\min\{\pi(v)\} = 1/n$. $\square$

6 Quantum Walks on General Graphs

We now prove a general upper bound on M_ϵ for any quantum walk. This will imply upper bounds on the other mixing times by theorem 5.2.

Let $|\phi_i\rangle$ be the eigenvectors of U with eigenvalues λ_i. The upper bound will be given in terms of Δ, which is defined to be the minimal spacing between the eigenvalues.

Theorem 6.1 *Consider a general quantum walk on a graph G with n nodes, with an auxiliary space of dimension d. Then, for an initial state $|\alpha\rangle$, the total variation distance from its limiting distribution is*

$$\|\bar{P}_T(\cdot|\alpha) - \pi(\cdot|\alpha)\| \leq \frac{\pi(\ln(nd/2)+1)}{T\Delta}$$

Proof: The proof follows approximately the same lines as the proof for the upper bound for the cycle, except that the complications due to throwing away part of the system disappear. More precisely, the proof goes along the lines of the proof of claim 4.7. The main difference is that we do not have a partition of the eigenvalues into two regimes. For this reason, in the counterpart of claim 4.6, we can have $|\theta_i - \theta_j| \in [0, \pi]$ (instead of $[0, \pi/2]$). Then, $|\lambda_i - \lambda_j|/L_{i,j}$ is minimized by $|\theta_i - \theta_j| = \pi$ (instead of $\pi/2$) and we get $|\lambda_i - \lambda_j| \geq \frac{2}{\pi}|i-j|\Delta$. Also, the counterpart of equation 20 has just one summation (over all eigenvalues) instead of two (over eigenvalues in the same regime and eigenvalues in different regimes). After that, we just notice that in the general case there are nd eigenvalues, which implies that $k = |i-j|$ runs up to $nd/2$. $\square$.

Just like we did in the cycle case, one can separate the eigenvectors to "good" and "bad" vectors, where the proportion of the "bad" vectors is δ, to get a better estimation of the mixing time.

7 A Polynomial Lower bound on Quantum Speed up

Here we are going to prove a lower bound on the various mixing measures of a general quantum random walk. In analogy to the classical case this bound will be stated in terms of the *conductance* Φ of the underlying graph G (cf. Chapter 2).
We will define a slightly different quantity Φ' first: Let $(X, \bar{X})$ be a cut in the graph G (i.e., a partition of vertices into two sets). Define B_X, the boundary of X, as the set of vertices in $\bar{X}$ that have an edge going to X. Let $\Phi' = \min_{|X| \leq \frac{1}{2}|V|} \frac{|B_X|}{|X|}$.

Theorem 7.1 *The filling, dispersion, mixing and sampling times of a general quantum walk with a uniform limiting distribution are $\Omega(1/\Phi')$.*

Proof: Let $(X, \bar{X})$ be the cut that achieves the minimum $(|B_X|/|X|)$. To simplify the notation, let $B = B_X$. Let $\mathcal{H}_X$ be the Hilbert space supported by nodes in X, i.e. the subspace spanned by $\{|a, v\rangle\}$ for all $v \in X$ and all basis states a of the auxiliary space. Let $\mathcal{H}_{\bar{X}}$ and $\mathcal{H}_B$ be Hilbert spaces supported by nodes in $\bar{X}$ and B

(defined similarly). We show a lower bound on filling and dispersion times by taking a random state $|\alpha\rangle$ of form $|a,v\rangle$, $v \in \bar{X}$ and showing that the projection of $U^t|\alpha\rangle$ onto $\mathcal{H}_X$ is small for all $t \leq \Theta(1/\Phi')$.
For any state $|\phi\rangle$, let $P_X|\phi\rangle$ and $P_B|\phi\rangle$ be the projections of $|\phi\rangle$ onto $\mathcal{H}_X$ and $\mathcal{H}_B$, respectively.
Let $|\alpha\rangle$ be a uniformly random basis state $|a,v\rangle$, $v \in \bar{X}$. We bound the expected projection of $|\alpha\rangle$ onto the boundary $\mathcal{H}_B$ and then use that to bound the expected projection of $U^t|\alpha\rangle$ onto $\mathcal{H}_X$. (This works because the only way to go from $\bar{X}$ to X is through the boundary B.)

Claim 7.2 *For any t, the expected value of $\|P_B U^t|\alpha\rangle\|^2$ is at most $|B|/|\bar{X}|$.*

Proof: $|\alpha\rangle$ is a uniformly random state in $|\bar{X}|d$ dimensions. Since U is unitary, $U^t|\alpha\rangle$ is a uniformly random state in some $|\bar{X}|d$-dimensional subspace of $\mathcal{H}$. The projection of this state to the $|B|d$-dimensional subspace $\mathcal{H}_B$ is at most $\frac{|B|d}{|\bar{X}|d} = \frac{|B|}{|\bar{X}|}$ (with equality if and only if $\mathcal{H}_B \subseteq U^t(\mathcal{H}_{\bar{X}})$). □

Claim 7.3 *For any state $|\alpha\rangle$ and any $k \in \mathbf{N}$,*

$$\|P_X U^k|\alpha\rangle\|^2 \leq \|P_X U^{k-1}|\alpha\rangle\|^2 + \|P_B U^{k-1}|\alpha\rangle\|^2.$$

Proof: Let $|\alpha'\rangle = U^{k-1}|\alpha\rangle$. Then, we need to prove

$$\|P_X U|\alpha'\rangle\|^2 \leq \|P_X|\alpha'\rangle\|^2 + \|P_B|\alpha'\rangle\|^2.$$

Let $|\alpha'\rangle = |\alpha'_1\rangle + |\alpha'_2\rangle$, with $|\alpha'_1\rangle$ being a superposition over vertices in $X \cup B$ and $|\alpha'_2\rangle$ being a superposition over $\bar{X} - B$.
Then, $P_X U|\alpha'_2\rangle = 0$ because $|a,v\rangle$ components of $|\alpha'_2\rangle$ get mapped to components corresponding to neighbors of v and no vertex in $\bar{X} - B$ is connected by an edge to a vertex in X. Therefore, $P_X U|\alpha'\rangle = P_X U|\alpha'_1\rangle$. Since P_X is a projection (and can only decrease norm) and U is unitary, $\|P_X U|\alpha'_1\rangle\|^2 \leq \|U|\alpha'_1\rangle\|^2 = \|\alpha'_1\|^2$. Since $|\alpha'_1\rangle$ is a superposition over vertices in $X \cup B$, $\|\alpha'_1\|^2 = \|P_X|\alpha'_1\rangle\|^2 + \|P_B|\alpha'_1\rangle\|^2$. □

Applying Claim 7.3 several times, we get

$$\|P_X U^k|\alpha\rangle\|^2 \leq \|P_B U^{k-1}|\alpha\rangle\|^2 + \|P_B U^{k-2}|\alpha\rangle\|^2 + \ldots + \|P_B U|\alpha\rangle\|^2 + \|P_B|\alpha\rangle\|^2.$$

By claim 7.2, the expected value of each term on the right-hand side (for a random $|\alpha\rangle = |a,v\rangle$, $v \in \bar{X}$) is at most $|B|/|\bar{X}|$. Therefore, the expected value of the sum on the right hand side is at most $k(|B|/|\bar{X}|)$. For some $|\alpha\rangle = |a,v\rangle$, the sum $\sum_{i=0}^{k-1} \|P_B U^i|\alpha\rangle\|^2$ is at most its expectation. For this $|\alpha\rangle$, we have

$$\|P_X U^j|\alpha\rangle\|^2 \leq \sum_{i=0}^{j-1} \|P_B U^i|\alpha\rangle\|^2 \leq \sum_{i=0}^{k-1} \|P_B U^i|\alpha\rangle\|^2 \leq k\frac{|B|}{|\bar{X}|}$$

for all $j < k$. For k to be the filling time, one of $\|P_X U^j|\alpha\rangle\|^2$ should be at least $(1-\epsilon)\frac{|X|}{|V|}$. Then, $k\frac{|B|}{|\bar{X}|} \geq (1-\epsilon)\frac{|X|}{|V|}$ and $k \geq (1-\epsilon)\frac{|X||\bar{X}|}{|B||V|}$. Since $|X| \leq \frac{1}{2}|V|$, $|\bar{X}| \geq \frac{1}{2}|V|$ and $k = \Omega(\frac{|X|}{|B|}) = \Omega(1/\Phi')$. A similar argument applies to dispersion time and sampling time. The bound on mixing time is implied by theorem 3.11 □

The quantity Φ' (which we call *boundary*) is similar but not identical to the conductance.

Lemma 7.4 *For a graph with maximal degree d,*

$$\Phi' \leq d\Phi,$$

where Φ is the conductance of a simple random walk on the graph.

Proof: For a simple random walk on G, the limiting distribution is $\pi_v = d_v / \sum_v d_v = d_v/2|E|$. Fix a cut $X, \bar{X}$, and denote the conductance of this cut by Φ_X. The capacity of X is $C_X = \frac{\sum_{v\in X} d_v}{2|E|} \leq \frac{d|X|}{2|E|}$. Let $E(X:\bar{X})$ be the set of edges going between $\bar{X}$ and X. The flow F_X satisfies $F_X = \frac{|E(X:\bar{X})|}{2|E|} \geq \frac{|B_X|}{2|E|}$. Therefore, $\Phi_X = \frac{F_X}{C_X} \geq \frac{2|B_X||E|}{2d|E||X|} \geq \frac{|B_X|}{d|X|} = \frac{\Phi'_X}{d}$. This is true for any cut X, which implies the lemma. □

Therefore, $\Omega(\frac{1}{\Phi'}) = \Omega(\frac{1}{d\Phi})$ and theorem 7.1 implies an $\Omega(\frac{1}{d\Phi})$ lower bound on For constant degree d graphs, this lower bound is $\Omega(\frac{1}{\Phi})$, the same as the classical lower bound on filling, dispersion and sampling times. Since a classical random walk converges in $O(\frac{1}{\Phi^2})$ steps, this means that a quantum walk can be at most quadratically faster.

Corollary 7.5 *For a general quantum walk on a bounded degree graph, the filling, dispersion, sampling and mixing times are at most quadratically faster than the mixing time of the simple classical random walk on that graph.*

For unbounded d, the factor-d gap between the two lower bounds (quantum and classical) is important. This gap can be quite large: we did not rule out the case in which the quantum filling time is $O(\log^c n)$ but d is $\Theta(n)$. We suspect that the bound can be improved and one can show that quantum walks are at most quadratically faster on any graph. We can prove that for the special case of coined quantum walks.

Theorem 7.6 *For a coined quantum walk, the filling, dispersion and sampling times are $\Omega(1/\Phi)$.*

Proof: To simplify the proof, we assume that the unitary transformation U is of form $C \circ S$, not $S \circ C$ (i.e. we first do the shift S and then the coin flip C). This assumption can be removed by replacing the starting state $|\alpha\rangle$ by $C^{-1}|\alpha\rangle$ and adding an extra C at the end.
The proof is similar to Theorem 7.1. We take the set X which achieves the conductance, $\Phi = \frac{F_X}{C_X}$. Let $\mathcal{H}_X$ and $\mathcal{H}_{\bar{X}}$ be similar to the proof of Theorem 7.1 and $\mathcal{H}_C$ be the Hilbert space spanned by *edges* in the cut (i.e., the space spanned by $|b, v\rangle$, $v \in \bar{X}$, $b \circ v \in X$). Let P_X and P_C be the projections onto $\mathcal{H}_X$ and $\mathcal{H}_C$, respectively.

Claim 7.7 *For any t, the expected value of $\|P_C U^t|\alpha\rangle\|^2$ (for a uniformly random $|\alpha\rangle = |a, v\rangle$, $v \in \bar{X}$) is at most $\frac{|E(X:\bar{X})|}{|\bar{X}|d}$.*

Proof: $|\alpha\rangle$ is a uniformly random state in $|\bar{X}|d$ dimensions. Since U is unitary, $U^t|\alpha\rangle$ is a uniformly random state in some $|\bar{X}|d$-dimensional subspace of $\mathcal{H}$. The projection of this state to the $|E(X : \bar{X})|$-dimensional subspace $\mathcal{H}_C$ is at most $\frac{|E(X:\bar{X})|}{|\bar{X}|d}$. □

Claim 7.8 *For any state $|\alpha\rangle$ and any $k \in \mathbf{N}$,*

$$\|P_X U^k|\alpha\rangle\|^2 \leq \|P_X U^{k-1}|\alpha\rangle\|^2 + \|P_C U^{k-1}|\alpha\rangle\|^2.$$

Proof: Define $|\alpha'\rangle = U^{k-1}|\alpha\rangle$. Let $|\alpha'\rangle = |\alpha'_1\rangle + |\alpha'_2\rangle$, with $|\alpha'_1\rangle$ being a superposition over $|b, v\rangle$ with $v \in X$ or $b \circ v \in X$ and $|\alpha'_2\rangle$ being a superposition over all other $|b, v\rangle$.
Then, $S|\alpha'_2\rangle \in \mathcal{H}_{\bar{X}}$ because $b \circ v \notin X$ for all $|b, v\rangle$ that appear in $|\alpha'_2\rangle$. Since the coin flip C does not change v, this also means that $U|\alpha'_2\rangle = C \cdot S|\alpha'_2\rangle \in \mathcal{H}_{\bar{X}}$. Therefore, $P_X U|\alpha'\rangle = P_X U|\alpha'_1\rangle$.
Similarly to the proof of claim 7.3,

$$\|P_X U|\alpha'_1\rangle\|^2 \leq \||\alpha'_1\rangle\|^2 = \|P_X|\alpha'_1\rangle\|^2 + \|P_C|\alpha'_1\rangle\|^2 = \|P_X|\alpha'\rangle\|^2 + \|P_C|\alpha'\rangle\|^2.$$

This completes the proof. □

The rest of proof is identical to Theorem 7.1. □

8 Conclusions and Open Questions

In this paper we have set up the basic definitions for quantum walks on graphs. However, the foundations of the theory of quantum walks on graphs still await discovery. We list here a few selected open problems.

The first open question is for which graphs is quantum speed up achievable. More generally, can the $1/\Phi$ lower bound always be achieved quantumly? In [8] it was shown that for any Markov chain, there exists a lifted version of it which achieves this bound, but no lifting can give better than $1/\Phi$ convergence. To achieve the lower bound of $1/\Phi$ by lifting, one has to be able to solve the multi-commodity flow on the graph, a task which is in general extremely hard. Therefore the lifting speed-up is an existence proof, rather than an algorithmic one. It would therefore be very interesting to know whether convergence in time $1/\Phi$ can be achieved by quantum walks for graphs other than the cycle in an efficiently constructible way.

An open question is to make our two bounds tight. We have shown how to improve the factor of $1/d$ in the lower bound for coined quantum walks, and this needs to be generalized to general quantum walks, or else find a counter example. One possible candidate is the graph consisting of two complete graphs connected by one edge. It is not clear that quantumly one cannot achieve convergence in time $O(n)$ which matches the $1/d\Phi = 1/n\Phi$ lower bound.

The characterization of the limiting distribution for general quantum walks still needs to be understood. For Abelian groups, we have shown that coined quantum walks converge to the uniform distribution. On the other hand, we know one example in which a quantum walk does not converge to the same limiting distribution as the classical simple random walk. This is a quantum walk on the Cayley graph of the symmetric group S_3. Is there a simple description, perhaps via representation theory, of the limiting distribution in the case of quantum walks on Cayley graphs of non-Abelian groups?

A very interesting question is how to use quantum walks in order to speed up algorithms. One way to do that is via speeding up the convergence time, however it is still an open question to give an example in which fast sampling cannot be done in an easy way classically. Another direction to pursue is to find other ways of using the various curious features of quantum walks, rather than speeding up the convergence time. For example, one might try to use quantum walks which converge to limiting distributions which are different than those of the corresponding classical walks. Another way might be to investigate which quantum states can be generated using quantum walks. Generating interesting quantum states is an important primitive for quantum algorithms. A well known example is the graph isomorphism problem which can be reduced to the problem of generating a certain quantum state efficiently.

9 Acknowledgements

We wish to thank John Watrous for introducing us to his model of a quantum walk on a line. We are grateful to Barbara Terhal for useful discussions. We are most grateful to Alesha Kitaev for pointing out to us an error in a previous version of this paper.

References

[1] D. Aharonov, A. Kitaev, and N. Nisan. Quantum circuits with mixed states. In *Proceedings of the 30th Annual ACM Symposium on Theory of Computing (STOC-98)*, pages 20–30, New York, May 23–26 1998. ACM Press.

[2] A. Ambainis, E. Bach and J. Watrous, Quantum Random Walks and the Analysis of Discrete Quantum Processes, Private Communication, October 2000.

[3] F. Chen, L. Lovasz and I. Pak, Lifting Markov Chains to Speed up Mixing, Proceedings of STOC'99, 275-281.

[4] Grover L K, A fast quantum mechanical algorithm for database search *Proc. of the 28th Annual ACM Symposium on Theory of Computing (STOC)* 212–219, 1996, ACM press.

[5] L. Lovasz and P. Winkler, Mixing Times, in *Microsurveys in Discrete Probability, Dimacs Series in Discrete Mathematics and Theoretical Computer Science*, **41**, eds. D. Aldous and J. Propp.

[6] A. Nayak and A. Vishwanath, Quantum walk on a line, Private Communication, October 2000.

[7] Shor P W, Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer, *SIAM J. Comp.*, **26**, No. 5, pp 1484–1509, October 1997

[8] A. Sinclair, Algorithms for Random Generation and counting, a Markov Chain approach, Birkauser press, 1993.

[9] A. Sinclair and M. Jerrum, Approximate counting, generation, and rapidly mixing Markov chains, in *Information and Computation*, **82**, 1989, pages 93-133.