Probab. Theory Relat. Fields 133, 215–235 (2005)
Digital Object Identifier (DOI) 10.1007/s00440-004-0423-2

Julia Kempe

Discrete Quantum Walks Hit Exponentially Faster

Received: 23 March 2004 / Revised version: 1 December 2004 /
Published online: 10 February 2005 – © Springer-Verlag 2005

Abstract. This paper addresses the question: what processes take polynomial time on a quantum computer that require exponential time classically? We show that the hitting time of the discrete time quantum walk on the n-bit hypercube from one corner to its opposite is polynomial in n. This gives the first exponential quantum-classical gap in the hitting time of discrete quantum walks. We provide the basic framework for quantum hitting time and give two alternative definitions to set the ground for its study on general graphs. We outline a possible application to sequential packet routing.

1. Introduction

Random walks form one of the cornerstones of theoretical computer science as well as the basis of a broad variety of applications in mathematics, physics and the natural sciences. In computer science they are frequently used in the design and analysis of randomized algorithms. Markov chain simulations provide a paradigm for exploring an exponentially large set of combinatorial structures (such as assignments to a Boolean formula or matchings in a graph) by a sequence of simple, local transitions. As algorithmic tools they have been applied to a variety of central problems, such as approximating the permanent [JS89, JSV01], finding satisfying assignments for Boolean formulas [Sch99, HSW02] and the estimation of the volume of a convex body [DFK91]. Other well-known examples of algorithms based on random walks include 2-SAT, Graph Connectivity and probability amplification [MR95, Pap94].

Recently the study of quantum walks has been initiated, with the hope of bringing new powerful algorithmic tools into the setting of quantum computing. To this day nearly all efficient quantum algorithms are based on the Quantum Fourier Transform (QFT), like Simon's period-finding algorithm [Sim97] or Shor's celebrated algorithms for Factoring and Discrete Log [Sho97]. However, it seems that the power of the QFT might be limited as a tool to solve similar problems on non-Abelian groups, like Graph Isomorphism [HRT00, GSVV01]. It seems crucial to develop new algorithmic tools.

Several striking differences between classical and quantum discrete walks[1] have already been observed for walks on the cycle [AAKV01], the line [ABN$^+$01] and

J. Kempe: CNRS-LRI, UMR 8623, Université de Paris-Sud, 91405 Orsay, France, and Computer Science Division and Dept. of Chemistry, University of California, Berkeley, USA.
e-mail: kempe@lri.fr

[1] For a survey of quantum walks see [Kem03a].

the hypercube [MR02]. The reason for this is quantum interference. Whereas there cannot be destructive interference in a classical random walk, in a quantum walk two separate paths leading to the same point may be out of phase and cancel out. The focus of previous work has been primarily on the mixing time of a discrete quantum walk. It has been shown that quantum walks on a large class of graphs can mix nearly quadratically faster than their classical counterparts. Since mixing times are an important quantity for many classical algorithms, this has raised the question of whether quantum walks can mix exponentially faster. However in [AAKV01] a lower bound on the mixing time of any local quantum walk has been obtained, which relates the mixing behavior of the walk to the classical conductance of the underlying graph. This result implies in essence that quantum walks can mix at most quadratically faster than classical walks (this is exactly true for bounded degree graphs; for graphs of maximal degree d this speed-up may be enhanced by a factor of $1/d$). This result showed that in all likelihood quantum walks cannot drastically enhance mixing times of classical walks.

In this paper we set the stage to exactly analyze another crucial quantity of discrete time random walks: the hitting time. The hitting time is important in many algorithmic applications of classical random walks, like k-SAT or Graph Connectivity. For instance the most efficient known solution to 3-SAT is based on the hitting time of a random walk [Sch99, HSW02]. In the algorithmic context, the question whether a quantum process can achieve an exponentially faster penetration of graphs has first been raised by Farhi and Gutmann [FG98]. For the continuous time quantum random walk, a different model from the one we analyze, Farhi et al. gave a mixture of analytical and numerical evidence of an exponential gap in hitting behavior [FG98, CFG02]. After our work has been completed they have very recently succeeded to give an oracle-based algorithmic exponential speed-up between classical and quantum query complexity based on the quantum continuous-time walk [CCD$^+$03]. In their example they are able to construct a family of random graphs with two special nodes such that on average any classical algorithm that needs to find the sink node starting from the source node requires an exponential number of queries, whereas the quantum algorithm succeeds in polynomial time. The continuous-time quantum walk at the base of their example is different from the discrete time model we analyze and it is a priori not clear how both models are related.

The hitting time h_{uv} of node v starting from node u measures the expected time it takes until the walk hits v for the first time. In the quantum case we face a dilemma: as is well known, observations of the quantum system (like "Has the walk hit node v?") influence the state of the quantum system. In particular if one were to observe the position of the quantum walk at each time it would lose its quantum coherence and reduce ("collapse") to the standard classical random walk, in which case we cannot expect any non-classical behavior or speed-ups. We give two alternatives out of this dilemma and establish two different notions of "quantum hitting time". In the first case the walk is not observed at all. Started at node u the position of the walk is measured at a (previously determined) time T. If the probability p to be at node v at time T is sufficiently large (an inverse polynomial in the logarithm of the graph size) we call T a "one-shot p hitting time". In the second case ("concurrent measurement") we do not require any previous knowledge of when to measure the position of the walk. Starting from node u at every step of the walk a partial measurement is performed (only the question "Is the position v or not v?" is asked). If the walk is found to have hit node v, it is stopped, otherwise the next step follows. This measurement perturbs the walk slightly but does not kill all the quantum coherence at once. If after a time T the probability p to halt is bounded below by an inverse polynomial in the logarithm of the size of the graph, we call T a "concurrent p hitting time".

After having made these notions rigorous we are able to show that on the hypercube both definitions of quantum hitting time lead to polynomial quantities for the walk from one corner to the opposite corner. This is in stark contrast to the classical case, where the corner-to-corner hitting time is exponential. Our result provides the first fully analytical classical-quantum exponential gap for a discrete quantum walk on a graph. It opens the possibility that quantum algorithms based on random walks may significantly improve upon classical algorithms. We will state similar results for the continuous-time quantum walk and also outline a possible application of rapid hitting on the hypercube: "quantum-random" sequential routing in a network.

It is interesting to know how much the exponential speed-up of the quantum walk depends on the choice of initial and final position. We establish two bounds: a lower bound on the size of the neighborhood of one corner from which we still achieve polynomial hitting behavior to the opposite corner and an upper bound on this neighborhood. This latter derives from a lower bound on quantum unstructured search algorithms [BBBV97].

While quantum walks are very easy to describe, they appear to be quite difficult to analyze. Standard techniques for analyzing classical random walks are apparently of little use. Whereas in the classical case most quantities depend only on the gap between the first and second largest eigenvalue of the underlying chain, in the quantum case all eigenvalues seem to play an equally important role and new methods are needed. We hope that establishing the rigorous notions and necessary techniques will help to analyze quantum walks on a variety of graphs.

Related Work:

A partial (proceedings) version of this work has appeared in [Kem03b].

Various quantum variants have previously been studied by several authors. In [Mey96, Wat01, AAKV01, ABN$^+$01] the general framework for discrete quantum walks is introduced, yet the focus and results of their work is different from ours. The mixing time of the quantum walk on the hypercube has been analysed in [MR02], both in the discrete and continuous time setting. We use the spectral decomposition they obtain for the discrete time walk. However, the results in [MR02] regard only the mixing time of the walk and do not deal with hitting times. In [ABN$^+$01] a notion of "halting" and intermediate partial measurement similar to our concurrent measurement is used, but the results regard the total halting probability of the quantum walk, and not the expected hitting time. Numerical studies of the hitting time on the hypercube have been communicated to us by Tomohiro Yamasaki [Yam] (published in [YKI02] after the work presented here has been completed) and have been reconfirmed in joint work with Neil Shenvi [SKW03].

A different model of quantum walks, so called continuous time walks, has been introduced by Farhi and Gutmann [FG98]. These are defined via a Hamiltonian that stems from the generating matrix of the classical continuous random walk. Until now it is not clear how their model is related to the discrete case we analyze. For their random walk model Farhi and Gutmann first exhibited an infinite tree and a walk that hits a set of leaves with inverse polynomial probability in polynomial time (similar to our notion of "one-shot hitting time"), where the classical analog has exponential hitting time. Later in [CFG02] another finite graph with a similar property is presented; both proofs are partly analytic and partly numeric, however. After the completion of the present work Childs et al. [CCD$^+$03] were able to construct a family of graphs based on the one in [CFG02] and to show that the continuous-time quantum walk gives rise to an exponential algorithmic speed-up between average case classical query complexity and its quantum version for the problem to find a very specific node in this graph. Even though their beautiful result proves a rigorous separation between the classical and the quantum setting, the wider applicability of their example is questionable at the moment. It is important to rigorously establish the notions and methods for hitting behaviour of quantum walks, in particular in the discrete case, and to analyze it for other graphs and structures. Our work provides a step in this direction.

Structure of the paper: We begin by reviewing in Sec. 2 the necessary background on classical random walks, quantum computation and quantum discrete time walks on graphs and in particular on the hypercube. In Sec. 3 we introduce the relevant definitions of quantum hitting times, and state and prove the upper bounds on quantum hitting times on the hypercube. In Sec. 4 we provide upper and lower bounds on the size of the neighborhood of a node from which the quantum walk has polynomial hitting behavior to the opposite corner. In Sec. 5 we outline a quantum routing application. In App. A we compare continuous-time quantum walks to discrete quantum walks and establish analogous results for their hitting time.

2. Background

2.1. Random Walks

Here we will state a few specific definitions and theorems as they are relevant to the present work to compare the behavior of classical and quantum walks (for a more complete treatment see e.g. [MR95, AF]).

Simple Random Walk: 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 after t steps, $D^t := D^0 P^t$, converges to a stationary distribution π which is independent of the initial distribution D^0. If a simple random walk on a bipartite graph has some periodicity (there is a state i and an initial distribution D^0 such that $D^t_i > 0$ iff t belongs to the arithmetic progression $\{a + ms | m \geq 0\}$ for some integer a) the introduction of a resting probability will make the walk aperiodic and convergent to π. For $d-$regular graphs G (all nodes of same degree d), the limiting probability distribution is uniform over the nodes of the graph.

Hitting Time: Given an initial state i, the probability that the first transition *into* a state j occurs at time t is denoted by r_{ij}^{t}. The hitting time h_{ij} is the expected number of steps to reach state j starting from state i and is given by $h_{ij} = \sum_{t>0} t r_{ij}^{t}$. For *aperiodic* simple random walks the Fundamental Theorem of Markov Chains implies that the number of times a state i is visited in the stationary state is $1/\pi_i$ and $h_{ii} = 1/\pi_i$.

Hypercube: The stationary distribution of the simple aperiodic random walk on the n-bit hypercube is given by $\pi_i = 1/2^n$. The hitting time from one node i to the opposite corner of the cube j is exponential in n, $h_{ij} = 2^n(1 + \frac{1}{n} + \frac{1}{O(n^2)})$.

Continuous time walk: The theory of continuous time Markov chains closely parallels discrete time chains. A continuous chain is specified by non-negative transition rates q_{ij}. Given that the state of the system at time t is $X_t = i$, the probability that $X_{t+dt} = j$ is $q_{ij}dt$. One can define $q_{ii} = -\sum_{j \neq i} q_{ij}$ to obtain a matrix Q. At time t the state of the system with initial state D^0 is then given by $D^t := D^0 exp(Qt)$. All the results on convergence and hitting essentially carry over to the continuous case with only slight modifications. To transition from discrete to continuous one can "discretize" a continuous chain by setting $P = exp(Q)$ or make a discrete chain continuous by setting $q_{ij} = p_{ij}$ for $i \neq j$. Stationary distribution and mean hitting times remain unchanged.

2.2. *Quantum Computation*

The model. Consider a finite dimensional 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 $\sum_{s \in \Omega} |a_s|^2 = 1$. $|\alpha^*\rangle$ denotes the conjugate and $\langle\alpha|$ denotes the conjugate transpose of $|\alpha\rangle$. $\langle\beta|\alpha\rangle$ denotes the inner product of $|\alpha\rangle$ and $|\beta\rangle$. For more details on quantum computing see e.g. [NC00].

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 projective (von Neuman) measurements, defined by a set of orthogonal projectors $\{\Pi_i : i \in I\}$ ($\Pi_i^\dagger = \Pi_i$, $\Pi_i^2 = \Pi_i$ and $\Pi_i \Pi_j = \delta_{ij}\Pi_i$) such that $\sum_{i \in I} \Pi_i = \mathbf{1}$. The output of the measurement of the state $|\alpha\rangle$ is an element $i \in I$ with probability $||\Pi_i|\alpha\rangle||^2$, we then say that Π_i was measured. Moreover, the new state of the system after the measurement with outcome i is the (normalized) state $(||\Pi_i|\alpha\rangle||)^{-1}\Pi_i|\alpha\rangle$. We denote the projectors on one basis state $|s\rangle$ by $|s\rangle\langle s|$.

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 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.

2.3. Discrete - Time Quantum Walk

It is not possible to define the quantum walk naïvely in analogy to the classical walk as a move in all directions "in superposition". It is easy to verify [Mey96] that a translationally invariant walk which preserves unitarity is necessarily proportional to a translation in one direction. If the particle has an extra degree of freedom that assists in its motion, however, then it is possible to define more interesting homogeneous local unitary processes. Following [AAKV01] we call the extra space the "coin-space" alluding to the classical coin that decides upon the direction of the walk.

More specifically 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. We will only consider undirected d-regular graphs G here, but slightly modified definitions can be made in the general case. Let $\mathcal{H}_C$ be the "coin"-Hilbert space of dimension d spanned by the states $|1\rangle$ through $|d\rangle$. Let $\mathbf{C}$ be a unitary transformation on $\mathcal{H}_C$ (the "coin-tossing operator" which we will define later). For each vertex (of degree d) label each outgoing edge with a number between 1 and d, such each number appears exactly once. Each edge will have two labels, corresponding to its two vertices. For Cayley graphs the labeling of an outgoing edge is simply the generator associated with the edge at that vertex. In general we will only define quantum walks for graphs with a consistent labeling. A consistent labeling has the property that for each vertex v no two of its neighbors have the same outgoing label on an edge leading to v. We can always make a labeling consistent by allowing more than d labels and adding self-loops. Now we can define a shift operator $\mathbf{S}$ on $\mathcal{H}_C \otimes \mathcal{H}_V$ 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 and the labeling is consistent, S is unitary. One step of the quantum walk is given by a local transformation acting on the coin-space only, followed by a conditional shift which leaves the coin-space unchanged [AAKV01]: $\mathbf{U} = \mathbf{S} \cdot (\mathbf{C} \otimes \mathbf{I_N})$.

Quantum Walk on the Hypercube: The hypercube of dimension n is a Cayley graph with $N = 2^n$ vertices. The position states are bit-strings $|x\rangle$ of length n. We denote by $|\overline{x}\rangle$ the vertex obtained from $|x\rangle$ by conjugating all the bits. The directions can be labeled by the n basis-vectors $\{|1\rangle, \ldots, |n\rangle\}$, corresponding to the n vectors of Hamming weight 1 $\{|e_1\rangle, \ldots, |e_n\rangle\}$, where e_i has a 1 in the ith position.

To mimic the permutation symmetry of the classical simple random walk we need to define the $n \times n$ coin operator $\mathbf{C}$ such that $\mathbf{U}$ is invariant to permutations of bits. As pointed out in [MR02] the symmetry of the hypercube defines the coin operator $\mathbf{C}$ to be of the form $C_{ij} = a$ if $i = j$ and $C_{ij} = b$ if $i \neq j$ with two parameters $a, b \in \mathbb{C}$. Unitarity of $\mathbf{C}$ further imposes two quadratic constraints on a and b, so that finally up to an overall phase all symmetric coins are characterized by one real parameter $1 - 2/n \leq |a| \leq 1$. Among all these coins the one farthest away from the identity operator $\mathbf{1}_n$ is given by $a = 2/n - 1$ and $b = 2/n$ [MR02]. We will call this latter coin $\mathbf{G}$ and use it as our coin in the rest of this paper. It is not hard to see that using another coin (with constant a, b) from the set of permutation invariant coins (except $\mathbf{1}_n$ of course) only slows down the walk by a constant factor and does not change the order of magnitude of the hitting behavior. To respect symmetry we will also impose permutation invariance for the initial state of the walk.

Definition 2.1 (Discrete time walk on the hypercube). *The symmetric discrete time walk* $\mathbf{U}$ *on the n - dimensional hypercube is acting on a* $n \cdot 2^n$ *dimensional space* $\mathcal{H}_n \otimes \mathcal{H}_2^{\otimes n}$ *as* $\mathbf{U} = \mathbf{S} \cdot (\mathbf{C} \otimes \mathbf{1_N})$ *where the shift operator* $\mathbf{S}$ *is defined as* $\mathbf{S} : |i, x\rangle \Rightarrow |i, x \oplus e_i\rangle$*, i.e.* $\mathbf{S} = \sum_{i=1}^{n} |i\rangle\langle i| \otimes S_i$ *with* $S_i|x\rangle = |x \oplus e_i\rangle$ *and* $\mathbf{C}$ *respects the symmetry of the hypercube. The initial state of the walk is chosen to be symmetric with respect to bit-permutations. For a walk starting in* $|x\rangle$ *the initial state is* $\frac{1}{\sqrt{n}} \sum_{i=1}^{n} |i\rangle \otimes |x\rangle$.

Note that this discrete-time quantum walk reduces to the classical symmetric walk if we perform a measurement in the coin-space in the direction-basis after every step of the walk. The resulting classical walk with last step in direction i will uniformly change to one of the $n - 1$ directions $j \neq i$ with probability $|b|^2 = 4/n^2$ and will return back to the node it came from (direction i) with probability $|a|^2 = 1 - 4/n + 4/n^2$. This type of classical random walk has a "direction-memory" one step back in time, but can be modeled by a (memoryless) Markov chain if we add a directional space to the position space. In other words each node v is blown up into n nodes v_i where i is the direction the walk came from. This resulting walk has a preference to oscillate back and forth between two adjacent nodes and has obviously still an exponential hitting time from one corner to its opposite.

The walk as defined is *periodic*: nodes with even Hamming weight are visited at even times only, nodes with odd Hamming weight at odd times. The inclusion of a "resting" coin-state $|0\rangle$ and a $n + 1 \times n + 1$ coin allowing for a self-loop transition amplitude of $a = 2/(n + 1) - 1$ make this walk aperiodic. To simplify the analysis we will only show the results for the periodic case, though; they hold with very slight modification in the aperiodic case as well.

3. Hitting Times on the Hypercube

For classical random walks the *hitting time* of a node v of a walk starting at an initial node i is defined as the expected time it takes the walk to reach v for the first time starting from i. Alternatively one can let the classical walk stop upon reaching the node v and define the *stopping-time* of the walk as the expected time for this walk to stop. In the classical case both notions are clearly the same. Care has to be applied to define an analogous notion for a quantum walk. To define "reaching" v

we have to circumvent the measurement problem. Namely if we were to measure the position of the walk after each step we will kill the quantum coherences and collapse the walk onto the corresponding classical walk. There are two alternatives: either to let the walk evolve and measure the position of the walk after T iterations ("one-shot measurements"), or to perform a partial measurement, described by the two projectors $\Pi_0 = |v\rangle\langle v|$ and $\Pi_1 = \mathbf{1} - \Pi_0$ (where $|v\rangle$ is some specific position we wish to "hit") after every step of the iteration ("concurrent measurement"). A priori these two notions can be very different in the quantum case. We will show that for both definitions the hitting time from one corner to its opposite is polynomial.

Definition 3.1 (One-shot hitting time). *A quantum walk U has a (T, p) one-shot $(|\phi_0\rangle, |x\rangle)$ hitting time if the probability to measure state $|x\rangle$ at time T starting in $|\phi_0\rangle$ is larger than p, i.e. $\| \langle x|U^T|\phi_0\rangle \|^2 \geq p$.*

Definition 3.2 ($|x\rangle$-stopped walk). *A $|x\rangle$-stopped walk from U starting in state $|\phi_0\rangle$ is the process defined as the iteration of a measurement with the two projectors $\Pi_0 = \Pi_x = |x\rangle\langle x|$ and $\Pi_1 = \mathbf{1} - \Pi_0$ and, if Π_1 is measured, an application of U. If Π_0 is measured the process is stopped.*

Definition 3.3 (Concurrent hitting time). *A quantum walk U has a (T, p) concurrent $(|\phi_0\rangle, |x\rangle)$ hitting-time if the $|x\rangle$-stopped walk from U and initial state $|\phi_0\rangle$ has a probability $\geq p$ of stopping at a time $t \leq T$.*

These two notions presuppose very different behavior of an algorithm exploiting them. In the one-shot case we have to know exactly *when* to measure the walk, which usually means that we have to know the dimension of the hypercube or, in more general applications, the shape of the graph. The advantage of the concurrent case is that we do not need any knowledge of when the walk will "hit" the target state. We simply continuously query the walk at the target state until we measure a "hit". This means that we do not need to have a priori information about the graph; probably ultimately more useful for algorithmic applications.

Note also that in the concurrent case if (T, p) is a hitting-time then for $T' \geq T$ (T', p) is also a hitting-time, i.e. hitting with probability at least p is a monotone property in time. In the one-shot case this is not at all true; we will see that for the hypercube there are certain windows in time where the probability to measure a certain node is high, followed by times where this probability is very low - yet another difference to the classical case.

3.1. One-shot hitting time

We will now state and prove our first main result:

Theorem 1. *The symmetric discrete-time quantum walk on the hypercube of dimension n with coin $\mathbf{G}$ has a (T, p) one-shot $(|x\rangle, |\bar{x}\rangle)$ hitting time where T is an integer of the same parity as n with*

(1) $T = \frac{\pi}{2}n$ and $p = 1 - O(\frac{\log^3 n}{n})$ (T is either $\lfloor \frac{\pi}{2}n \rfloor$ or $\lceil \frac{\pi}{2}n \rceil$),

(2) $T = \frac{\pi}{2}n \pm O(n^\beta)$ and $p = 1 - O(\frac{\log n}{n^{1-2\beta}})$ with $0 < \beta < 1/2$,

(3) $T \in [\frac{\pi}{2}n - O(\frac{\sqrt{n}}{\log n}), \frac{\pi}{2}n + O(\frac{\sqrt{n}}{\log n})]$ and $p = 1 - O(\frac{\log \log n}{\log n})$.

The "$\sqrt{n}$"-window around the exact one-shot measurement time of $\pi n/2$ makes the algorithm more robust to slight perturbations in the exact time of the measurement.

Proof of Theorem 1. To prove the upper bound on the $(|x\rangle, |\bar{x}\rangle)$ hitting time note that by the symmetry of the hypercube and the walk U the hitting time is the same for all $(|x\rangle, |\bar{x}\rangle)$ with $x \in \{0,1\}^n$. So w.l.o.g. we set $|x\rangle = |00\ldots 0\rangle$. As already shown in [MR02], the $n \cdot 2^n$ eigenstates of U are of the form $|v_k^i\rangle \otimes |\tilde{k}\rangle$ where $|\tilde{k}\rangle = \frac{1}{\sqrt{2^n}} \sum_{x\in\{0,1\}^n} (-1)^{k\cdot x}|x\rangle$ is the $\mathbb{Z}_2^n$-Fourier transform of $|k\rangle$ for $k \in \mathbb{Z}_2^n$ and the n vectors $\{|v_k^i\rangle : i = 1\ldots n\}$ for each k are the eigenvectors of the matrix $\mathbf{S_k} \cdot \mathbf{G}$, where $\mathbf{S_k}$ is the diagonal $n \times n$ matrix with $(\mathbf{S_k})_{lm} = \delta_{lm}(-1)^{k_l}$.

The symmetric initial state is $|\Psi_{in}\rangle \otimes |00\ldots 0\rangle := \frac{1}{\sqrt{n}} \sum_{i=1}^n |i\rangle \otimes |00\ldots 0\rangle$ (see Def. 2.1). For all k, only two of the n eigenvectors $|v_k^i\rangle$ have non-zero inner product with $|\Psi_{in}\rangle$ [MR02]. These two eigenvectors are complex conjugates, call them $|w_k\rangle$ and $|w_k^*\rangle$ and their corresponding eigenvalues are λ_k and λ_k^* with $\lambda_k = 1 - \frac{2|k|}{n} + i\frac{2}{n}\sqrt{|k|\,(n-|k|)}$ where $|k|$ is the Hamming weight of k. Let $\lambda_k = e^{i\omega_{|k|}} = \cos\omega_{|k|} + i\sin\omega_{|k|}$ where $\cos\omega_m = 1 - 2m/n$. The entries of $|w_k\rangle$ are $(w_k)_l = \frac{-i}{\sqrt{2}\sqrt{n-|k|}}$ if $k_l = 0$ and $(w_k)_l = \frac{1}{\sqrt{2}\sqrt{|k|}}$ if $k_l = 1$. (If $k = 0$ and $k = n$ there is only one eigenvector, the uniform superposition over all directions, with eigenvalue $\lambda_0 = 1$ and $\lambda_n = -1$. When we write out the general eigenvectors this special case will be self-understood.) The initial state is a superposition over $2^{n+1} - 2$ eigenvectors [MR02]:

$$|\Phi_0\rangle := |\Psi_{in}\rangle \otimes |00\ldots 0\rangle = \sum_{k\in\{0,1\}^n} (a_k|w_k\rangle + a_k^*|w_k^*\rangle) \otimes |\tilde{k}\rangle \qquad (1)$$

with $a_k = \frac{1}{\sqrt{n\cdot 2^{n+1}}}(\sqrt{|k|} - i\sqrt{n-|k|})$. Let us denote by

$$|\Phi_t\rangle = U^t(|\Psi_{in}\rangle \otimes |00\ldots 0\rangle) = \sum_{x\in\{0,1\}^n} \alpha_t^x |u_t^x\rangle \otimes |x\rangle$$

the state of the system after t iterations, where $|u_t^x\rangle$ is a normalized vector in coin-space. Note that because both the walk U and its initial state preserve the bit-permutation symmetry of the hypercube, the only consistent coin-state for position $|11\ldots 1\rangle$ is the completely symmetric state over all directions: $|u_t^{11\ldots 1}\rangle = \frac{1}{\sqrt{n}}\sum_{i=1}^n |i\rangle = |\Psi_{in}\rangle$. Let us call $|f\rangle = |\Psi_{in}\rangle \otimes |11\ldots 1\rangle$ the "target" state. With these quantities in place, α_t, the amplitude at time t of the particle being in $|11\ldots 1\rangle$, the opposite corner, is

$$\begin{aligned}\alpha_t :&= \alpha_t^{11\ldots 1} = \langle f|\Phi_t\rangle = \sum_{k\in\{0,1\}^n} (a_k\lambda_k^t\langle\Psi_{in}|w_k\rangle + a_k^*\lambda_k^{*t}\langle\Psi_{in}|w_k^*\rangle)\cdot\langle 11\ldots 1|\tilde{k}\rangle \\ &= \sum_{k\in\{0,1\}^n} \frac{1}{\sqrt{n\cdot 2^{n+1}}}\,\frac{2n\cos(\omega_k t)}{\sqrt{2}\sqrt{n}}\,\frac{(-1)^{|k|}}{\sqrt{2^n}} = \frac{1}{2^n}\sum_{m=0}^n \binom{n}{m}(-1)^m\cos(\omega_m t).\end{aligned} \qquad (2)$$

Claim 1. For all $t \in [\frac{\pi}{2}n - O(n^\beta), \frac{\pi}{2}n + O(n^\beta)]$ such that $t-n$ is even, $|\alpha_t|$ is lower bounded by $1 - O(\frac{\log n}{n^{1-2\beta}})$ for $0 < \beta < 1/2$.

Proof of Claim 1. Let us split the sum (2) into two parts, one where the index $m \in M := [(1-\delta)n/2, (1+\delta)n/2]$ and one where $m \notin M$, with $\delta < 1$ specified later. By standard Chernoff bounds on the tail probabilities of the binomial distribution we can upper-bound the absolute value of all the contributions from $m \notin M$ as

$$\left|\frac{1}{2^n}\sum_{m\notin M}\binom{n}{m}(-1)^m\cos(\omega_m t)\right| \leq \frac{1}{2^n}\sum_{m\notin M}\binom{n}{m} \leq 2e^{-\frac{\delta^2 n}{2}}. \tag{3}$$

Let us set $\delta = \sqrt{\frac{g(n)}{n}}$ with $g(n) = \Omega(\log n)$, in which case (3) is upper bounded by $2e^{-\Omega(\log n)/2}$. Let us write $t = \frac{\pi}{2}n \pm \epsilon$ (i.e. $\epsilon = O(n^\beta)$). The second term in the sum will come from contributions $m \in M$, so the terms $\cos\omega_m = 1 - 2m/n \in [-\delta, \delta]$ will be small. Call $\nu_m = \frac{\pi}{2} - \omega_m$, so $\cos\omega_m = \cos(\frac{\pi}{2} - \nu_m) = \nu_m - O(\nu_m^3)$ which means $\nu_m = 1 - 2m/n \pm O(\delta^3)$. Then

$$\begin{aligned}\cos(\omega_m t) &= \cos[\left(\frac{\pi}{2} - 1 + \frac{2m}{n} \pm O(\delta^3)\right)\left(\frac{\pi}{2}n \pm \epsilon\right)]\\ &= \cos[\left(\frac{t-n}{2} + m\right)\pi \mp \epsilon(1 - \frac{2m}{n}) \pm tO(\delta^3)]\\ &= (-1)^{\frac{t-n}{2}+m}\cos[\mp\epsilon\left(1 - \frac{2m}{n}\right) \pm O(n\delta^3)]\\ &= (-1)^{\frac{t-n}{2}+m}[1 - O(\epsilon^2\delta^2) - O(n^2\delta^6)]\end{aligned} \tag{4}$$

and the second sum $\frac{1}{2^n}\sum_{m\in M}\binom{n}{m}(-1)^m\cos(\omega_m t) = (-1)^{\frac{t-n}{2}}[1 - O(\epsilon^2\delta^2) - O(n^2\delta^6)]\frac{1}{2^n}\sum_{m\in M}\binom{n}{m}$. Since $\frac{1}{2^n}\sum_{m\in M}\binom{n}{m} \geq 1 - 2e^{-g(n)/2}$ we have

$$\begin{aligned}|\alpha_t| &\geq \left|\frac{1}{2^n}\sum_{m\in M}\binom{n}{m}(-1)^m\cos(\omega_m t)\right| - 2e^{-g(n)/2}\\ &\geq 1 - O(\frac{g(n)}{n^{1-2\beta}}) - O(\frac{g^3(n)}{n}) - 4e^{-g(n)/2}\end{aligned} \tag{5}$$

Set $g(n) = 2\log n$ to prove the claim for $0 < \beta < 1/2$.

To prove Theorem 1 note that the probability of measuring the system in $|11\ldots1\rangle$ is $p = |\alpha_t|^2$. Set $\beta = \frac{1}{2}(1 - \frac{\log\log n}{\log n})$ and use Eq. (5) with $g(n) = 2\log\log n$ to get $p \geq 1 - O(\frac{\log\log n}{\log n})$. For $\beta = 0$ set $g(n) = 2\log n$ to get a lower bound of $1 - O(\log^3 n/n)$.

Remark. Note that if we set $T = (2m+1)n\pi/2$ we obtain a similar result to the $m = 0$ case as long as T is sufficiently small so that $O(T^2\delta^6)$ terms do not matter, i.e. $m = O(1)$. We can think of the walk returning to $|11\ldots\rangle$ every πn steps, which is in stark contrast to the classical case where the expected number of times a walk returns to some node i is $1/\pi_i = 2^n$ (see Sec. 2.1).

3.2. Concurrent hitting time

Our second result relates to the concurrent version of hitting time. It implies that even without information on when to measure we retain a polynomial hitting behavior:

Theorem 2. *The symmetric discrete time quantum walk on the hypercube of dimension n has a* $(\frac{\pi}{2}n, \Omega(\frac{1}{n\log^2 n}))$ *concurrent* $(|x\rangle, |\overline{x}\rangle)$ *hitting time.*

Amplification: If the probability p in Defs. 3.1 and 3.3 is an inverse polynomial $p(n)$ in the size of the instance, we can use standard classical amplification to boost this probability to be exponentially close to 1. We just restart the quantum walk from scratch and repeat it $O(1/p(n))$ times. With amplification the coined symmetric discrete-time quantum walk on the hypercube of dimension n has a $(O(n^2\log^2 n), 1-2^{-O(n)})$ concurrent $(|x\rangle, |\overline{x}\rangle)$ hitting time.

Remark. To be fair we should compare our results to *tail-bounds* for the hitting time in the classical case. It is very easy to show, however, that for the simple random walk on the hypercube starting in a node i the probability to hit the opposite corner j in a polynomial number of steps is exponentially small since each of the probabilities r_{ij}^t to be at j at time t (see Sec. 2.1) is exponentially small.

Proof of Theorem 2. The strategy of the proof is to compare the hitting probabilities at time t of the $|11\dots1\rangle$-stopped walk to the unmeasured walk and to show that the perturbation caused by the measurement of the walk only gives a polynomial "loss" in hitting amplitude.

For the $|11\dots1\rangle$-stopped walk (see Def. 3.2) the same symmetry arguments as before apply, since the measurement projectors Π_0 and $\Pi_1 = I - \Pi_0$ are also symmetric with respect to bit permutations. So the only possible "target" state is again $|f\rangle = \frac{1}{\sqrt{n}}\sum_{i=0}^n |i\rangle \otimes |11\dots1\rangle$ and we may assume that we measure with $\{\Pi_0 = |f\rangle\langle f|, \Pi_1 = \mathbf{1} - \Pi_0\}$. As above let $|\Phi_t\rangle$ be the states of the *unmeasured* walk after t applications of U with $|\Phi_0\rangle = |\Psi_{in}\rangle \otimes |00\dots0\rangle$ and $\alpha_t = \langle f|\Phi_t\rangle$. Since the walk has non-zero transition amplitude only between nearest neighbors, the first time $\alpha_t \neq 0$ is for $t = n$ and since the walk is 2-periodic $\alpha_t = 0$ whenever t and n have different parity.

Let us define $|\tilde{\Phi}_t\rangle = (U\Pi_1)^t(|\Psi_{in}\rangle \otimes |00\dots0\rangle)$ as the *non-normalised* state we get at time t given the walk has not stopped before t and $\beta_t := \langle f|\tilde{\Phi}_t\rangle$. Note that for $t \leq n$ we have $|\Phi_t\rangle = |\tilde{\Phi}_t\rangle$ and $\alpha_t = \beta_t$. Also note that $\beta_{n+t} = 0$ for odd $(t+n)$, if the initial state of the walk is located in $|0..0\rangle$.

Claim 2. The probability to stop at some time $t \leq T$ is given by $p_T = \sum_{t=0}^T \left|\langle f|\tilde{\Phi}_t\rangle\right|^2 = \sum_{t=0}^T |\beta_t|^2$.

Proof of Claim 2. As in previous work [ABN$^+$01] it is easy to see that calculating with the unnormalized state gives the *unconditional* probability to stop. If we do not renormalize our states we get exactly the conditional probability to stop at time t given we have not stopped before.

We now want to relate the α_t from the *unmeasured* walk to the actual β_t of the *measured* walk.

Claim 3. $|\tilde{\Phi}_{n+k}\rangle = |\Phi_{n+k}\rangle - \sum_{i=0}^{k-1} \beta_{n+i} U^{k-i} |f\rangle$ and $\beta_{n+k} = \alpha_{n+k} - \sum_{i=1}^{k} \beta_{n+k-i} \cdot \gamma_i$ with $\gamma_t = \langle f|U^t|f\rangle$.

Proof of Claim 3. By induction on k. By previous arguments we have $|\Phi_t\rangle = |\tilde{\Phi}_t\rangle$ and $\alpha_t = \beta_t$ for $t \leq n$. Further $|\tilde{\Phi}_{n+1}\rangle = U|\Phi_n\rangle - U\alpha_n|f\rangle = |\Phi_{n+1}\rangle - \beta_n U|f\rangle$ so $\beta_{n+1} = \langle f|\Phi_{n+1}\rangle - \alpha_n\langle f|U|f\rangle = \alpha_{n+1} - \beta_n\langle f|U|f\rangle$. Write $|\tilde{\Phi}_{n+k+1}\rangle = U|\tilde{\Phi}_{n+k}\rangle - \beta_{n+k}U|f\rangle$ and apply the induction hypothesis to $|\tilde{\Phi}_{n+k}\rangle$. The claim on β_{n+k} follows immediately.

Claim 4. Let $T = \lceil \frac{\pi}{2} n \rceil$ or $\lfloor \frac{\pi}{2} n \rfloor$ s.t. $T - n$ is even, let $0 \leq 2t \leq T - n$ and define $\tilde{\gamma}_{2t} = (-1)^t \gamma_{2t}$.

1. $\gamma_t = \frac{1}{2^n} \sum_{m=0}^{n} \binom{n}{m} \cos(\omega_m t)$ and $\gamma_{2t+1} = 0$,

2. $\left|\tilde{\gamma}_{2t} - \tilde{\gamma}_{2(t+1)}\right| = O(\frac{\log n}{\sqrt{n}})$,

3. $\exists c$ s.t. for $t_c = \lfloor c\sqrt{n} \rfloor$ we have $\left|\alpha_{T-2t_c}\right| \leq \frac{1}{2}$.

Proof of Claim 4. First note that by the symmetry of the states and U we have that $\gamma_t = \langle f|U^t|f\rangle = \langle \Psi_{in}| \otimes \langle 11 \ldots 1|U^t|\Psi_{in}\rangle \otimes |11 \ldots 1\rangle = \langle \Psi_{in}| \otimes \langle 00 \ldots 0|U^t|\Psi_{in}\rangle \otimes |00 \ldots 0\rangle$. To obtain 4.1 adapt Eq. (2) with $\gamma_t = \langle \Psi_{in}| \otimes \langle 00 \ldots 0|\Phi_t\rangle$. Since γ_t is the amplitude of the initial state at time t starting with the initial state and since the walk is 2-periodic, γ_{2t+1} must be 0 at odd times, which proves Claim 4.1. For part 4.2 let us write as before $\nu_m = \pi/2 - \omega_m$ and $\gamma_{2t} = \frac{1}{2^n} \sum_{m=0}^{n} \binom{n}{m} \cos(t\pi - 2t\nu_m) = (-1)^t \frac{1}{2^n} \sum_{m=0}^{n} \binom{n}{m} \cos(2t\nu_m)$. Then $\tilde{\gamma}_{2t} = \frac{1}{2^n} \sum_{m=0}^{n} \binom{n}{m} \cos(2t\nu_m)$ and

$$\begin{aligned} \tilde{\gamma}_{2t} - \tilde{\gamma}_{2t+2} &= \frac{1}{2^n} \sum_{m=0}^{n} \binom{n}{m} [\cos 2t\nu_m - \cos(2t+2)\nu_m] \\ &= \frac{-2}{2^n} \sum_{m=0}^{n} \binom{n}{m} \sin(2t+1)\nu_m \sin \nu_m \qquad (6) \end{aligned}$$

As before we split the sum in (6) into two parts ($m \in M$ and $m \notin M$), and set $\delta = \frac{\sqrt{2\log n}}{\sqrt{n}}$ such that the Chernoff-tails are $O(\frac{1}{n})$. For the part with $m \in M$ we use again that $\nu_m = 1 - 2m/n + O(\nu_m^3)$ and define $i = n/2 - m$ (i.e. $\nu_m = \frac{2i}{n} + O(\frac{i^3}{n^3})$). Then up to terms of $O(\frac{1}{n})$ and with $|\sin(2t+1)\nu_m| \leq 1$ we get

$$\begin{aligned} |\tilde{\gamma}_{2t} - \tilde{\gamma}_{2t+2}| &\leq \frac{2}{2^n} \sum_{i=-\delta n/2}^{\delta n/2} \binom{n}{n/2 - i} \left| \frac{2i}{n} + O(\frac{i^3}{n^3}) \right| \\ &= \frac{8}{n2^n} \sum_{i=0}^{\delta n/2} \binom{n}{n/2 - i} i + O(\delta^3) \end{aligned}$$

Note that $\frac{1}{2^n}\binom{n}{n/2-i} = O(\frac{1}{\sqrt{n}})$ and $\sum_{i=0}^{\delta n/2} i = O(\delta^2 n^2)$, so $|\tilde{\gamma}_{2t} - \tilde{\gamma}_{2t+2}| = O(\frac{\log n}{\sqrt{n}})$.

For part 4.3 we use expression (2) with $\nu_m = \frac{\pi}{2} - \omega_m$ and follow the reasoning and notation of Eq. (4)

$$\alpha_{T-2t} = \frac{1}{2^n}\sum_{m=0}^{n}\binom{n}{m}(-1)^m \cos\Big((T-2t)(\frac{\pi}{2} - \nu_m)\Big)$$
$$= (-1)^{\frac{T-n}{2}-t}\frac{1}{2^n}\sum_{m\in M}\binom{n}{m}\cos(2t\nu_m) + O(T\delta^3)$$

where we set $\delta = \sqrt{\log n}/\sqrt{n}$, so that the Chernoff-tails are $O(\frac{1}{\sqrt{n}})$ and $O(T\delta^3) = O(\frac{\log^{3/2}(n)}{\sqrt{n}})$. Set $i = n/2 - m$ (i.e. $\nu_m = \frac{2i}{n} + O(\frac{i^3}{n^3})$) so that up to terms of $O(T\delta^3)$ and with $\theta = 4t_c/n$,

$$\left|\alpha_{T-2t_c}\right| = \frac{2}{2^n}\left|\sum_{i=0}^{\delta n/2}\binom{n}{n/2-i}\cos(\frac{4t_c i}{n})\right| = \frac{2}{2^n}\left|\sum_{i=0}^{\delta n/2}\binom{n}{n/2-i}\cos(i\theta)\right|. \quad (7)$$

Let i_1 be the largest integer such that $i_1\theta = \frac{4t_c i_1}{n} \le \frac{\pi}{2}$, i_2 the largest integer such that $i_2\theta \le \frac{2\pi}{2}$ and so on, so $i_k = \lfloor\frac{k\pi}{2\theta}\rfloor = \lfloor\frac{k\pi n}{8\lfloor c\sqrt{n}\rfloor}\rfloor$ implying that $\frac{k\pi}{8c}\sqrt{n} + \frac{k\pi}{8c^2} \ge i_k \ge \frac{k\pi}{8c}\sqrt{n} - 1$. The index k runs from $1\ldots K$ and $i_K = \lceil\delta n/2\rceil$, which gives $\frac{4c}{\pi}\sqrt{\log n} + \frac{16c}{\pi\sqrt{n}} \ge K \ge \frac{4c}{\pi}\sqrt{\log n} - \frac{4\sqrt{\log n}}{\pi\sqrt{n}}$. This means $i_k = \frac{k\pi}{8c}\sqrt{n} \pm O(\sqrt{\log n})$. The cosine function in Eq. (7) is non-negative for $0 \le i \le i_1$ and $i_{4k-1} < i \le i_{4k+1}$ and non-positive otherwise. Then the expression in Eq. (7) becomes

$$\frac{2}{2^n}\left|\sum_{i=0}^{i_1}\binom{n}{\frac{n}{2}-i}\cos(i\theta) + \sum_{k=1}^{K/4}\left(\sum_{i=i_{4k-1}+1}^{i_{4k+1}}\binom{n}{\frac{n}{2}-i}\cos(i\theta) - \sum_{i=i_{4k-3}+1}^{i_{4k-1}}\binom{n}{\frac{n}{2}-i}|\cos(i\theta)|\right)\right|$$
$$\le \frac{2}{2^n}\left[\binom{n}{\frac{n}{2}}\sum_{i=0}^{i_1}\cos(i\theta) + \sum_{k=1}^{K/4}\binom{n}{\frac{n}{2}-i_{4k-1}}\left|\sum_{i=i_{4k-1}+1}^{i_{4k+1}}\cos(i\theta) - \sum_{i=i_{4k-3}+1}^{i_{4k-1}}|\cos(i\theta)|\right|\right] \quad (8)$$

We will make use of the following fact: $\sum_{i=0}^{J}\cos i\theta = \frac{\cos\frac{J\theta}{2}\sin\frac{(J+1)\theta}{2}}{\sin\theta/2}$. So $\sum_{i=0}^{i_1}\cos(i\theta) \le \frac{1}{\sin\theta/2} = \frac{2}{\theta} + O(\theta)$. Also $\frac{1}{2^n}\binom{n}{n/2-i} \le \frac{1}{2^n}\binom{n}{n/2} = \frac{\sqrt{2}}{\sqrt{\pi n}} + O(\frac{1}{n^{3/2}})$. Putting this together we get for the first sum in Eq. (8)

$$\frac{2}{2^n}\sum_{i=0}^{i_1}\binom{n}{\frac{n}{2}-i}\cos(i\frac{4t_c}{n}) \le \frac{2\sqrt{2}}{\sqrt{\pi n}}\frac{n}{2t_c} + O\left(\frac{1}{n}\right) = \frac{\sqrt{2}}{\sqrt{\pi c}} + O\left(\frac{1}{n}\right).$$

The second term in Eq. (8) can be bounded above by

$$\frac{2}{2^n}\binom{n}{n/2}\sum_{k=1}^{K/4}\sum_{i=i_{4k-3}+1}^{i_{4k+1}}\cos(i\theta)$$
$$= \left[\frac{2\sqrt{2}}{\sqrt{\pi n}} + O(\frac{1}{n^{3/2}})\right]\sum_{k=1}^{K/4}\frac{\cos\frac{i_{4k+1}\theta}{2}\sin\frac{(i_{4k+1}+1)\theta}{2} - \cos\frac{i_{4k-3}\theta}{2}\sin\frac{(i_{4k-3}+1)\theta}{2}}{\sin\theta/2}$$

Note that $\cos\frac{i_{4k+1}\theta}{2} = \cos(\frac{(4k+1)\pi}{8c}\sqrt{n}\frac{2t}{n} \pm O(\frac{\sqrt{\log n}}{\sqrt{n}})) = \cos(k\pi + \frac{\pi}{4} \pm O(\frac{\sqrt{\log n}}{\sqrt{n}})) = \frac{(-1)^k}{\sqrt{2}} \pm O(\frac{\sqrt{\log n}}{\sqrt{n}})$ and similarly $\sin\frac{(i_{4k+1}+1)\theta}{2} = \frac{(-1)^k}{\sqrt{2}} \pm O(\frac{\sqrt{\log n}}{\sqrt{n}})$, so

$$\sum_{k=1}^{K/4} \sum_{i=i_{4k-3}+1}^{i_{4k+1}} \cos(i\theta) = \sum_{k=1}^{K/4} \frac{\frac{1}{2} - \frac{1}{2} + O\left(\sqrt{\log n}/\sqrt{n}\right)}{\sin\theta/2} = O\left(\frac{\log n}{n}\right).$$

Putting all the above together we get $\left|\alpha_{T-2t_c}\right| \leq \frac{\sqrt{2}}{\sqrt{\pi c}} + O(\frac{\log^{3/2} n}{\sqrt{n}})$ which can be made smaller than $\frac{1}{2}$ with the appropriate choice of c.

We now can give a lower bound on $|\beta_t|$ in terms of the quantities of the unmeasured walk:

Claim 5. Let t_c be as in Claim 4.3. If $\sum_{i=0}^{\frac{T-n}{2}-t_c} |\beta_{n+2i}| = o(\frac{1}{\log n})$ then $|\beta_{n+2t}| \geq |\alpha_{n+2t}| - |\alpha_{n+2t-2}| - o(\frac{1}{\sqrt{n}})$ for $T - n - 2t_c \leq 2t \leq (T - n)$.

Proof of Claim 5. Call $\tilde{\beta}_{n+2t} = (-1)^t \beta_{n+2t}$ and $\tilde{\alpha}_{n+2t} = (-1)^t \alpha_{n+2t}$. Adapt Claim 3 with $\gamma_{2i+1} = 0$ (Claim 4.1.) to get

$$\begin{aligned}\tilde{\beta}_{n+2t} &= \tilde{\alpha}_{n+2t} - \sum_{i=1}^{t} \tilde{\beta}_{n+2t-2i} \cdot \tilde{\gamma}_{2i}\\ &= \tilde{\alpha}_{n+2t} - \sum_{i=0}^{t-1} \tilde{\beta}_{n+2(t-1-i)} \cdot \tilde{\gamma}_{2i} + \sum_{i=0}^{t-1} \tilde{\beta}_{n+2(t-1-i)}(\tilde{\gamma}_{2i} - \tilde{\gamma}_{2i+2})\end{aligned}$$

Note that $\tilde{\gamma}_0 = 1$ and so $\tilde{\alpha}_{n+2t-2} = \sum_{i=0}^{t-1} \tilde{\beta}_{n+2(t-1-i)} \cdot \tilde{\gamma}_{2i}$, which gives the lower bound

$$\begin{aligned}|\beta_{n+2t}| &\geq |\alpha_{n+2t}| - |\alpha_{n+2t-2}| - \left|\sum_{i=0}^{t-1} \tilde{\beta}_{n+2(t-1-i)}(\tilde{\gamma}_{2i} - \tilde{\gamma}_{2i+2})\right|\\ &\geq |\alpha_{n+2t}| - |\alpha_{n+2t-2}| - O\left(\frac{\log n}{\sqrt{n}}\right) \sum_{i=0}^{\frac{T-n}{2}} |\beta_{n+2i}|\end{aligned}$$

where we used Claim 4.2. Let us split $\sum_{i=0}^{\frac{T-n}{2}} |\beta_{n+2i}| = \sum_{i=0}^{\frac{T-n}{2}-t_c} |\beta_{n+2i}| + \sum_{i=0}^{t_c-1} |\beta_{T-2i}|$ The first sum is $o(\frac{1}{\log n})$ by assumption and to upper bound the second sum we use

$$\sum_{i=0}^{t_c-1} |\beta_{T-2i}| \leq \sqrt{t_c} \cdot \sqrt{\sum_{i=0}^{t_c-1} |\beta_{T-2i}|^2} \leq \sqrt{\lfloor c\sqrt{n} \rfloor}\sqrt{p_T}.$$

Now either $p_T = \Omega(\frac{1}{\log^2 n\sqrt{n}})$, which would prove our theorem, or $p_T = o(\frac{1}{\log^2 n\sqrt{n}})$, which establishes that the second sum is also $o(\frac{1}{\log n})$.

If the assumption of Claim 5 is not true, then

$$\Omega\left(\frac{1}{\log n}\right) = \sum_{i=0}^{\frac{T-n}{2}-t_c} |\beta_{n+2i}| \leq \sqrt{\frac{T-n}{2}\sum_{i=0}^{\frac{T-n}{2}} |\beta_{n+2i}|^2} \leq \sqrt{np_T}$$

which means $p_T = \Omega(\frac{1}{n\log^2 n})$.

The rest of Theorem 2 follows from Claim 2 and Claim 5

$$\begin{aligned} p_T &= \sum_{t=n}^{T} |\beta_t|^2 \geq \sum_{t=T-\lfloor c\sqrt{n}\rfloor}^{T} |\beta_t|^2 \geq \frac{1}{c\sqrt{n}}\left(\sum_{t=T-\lfloor c\sqrt{n}\rfloor}^{T} |\beta_t|\right)^2 \\ &\geq \frac{1}{c\sqrt{n}}\left(\sum_{t=T-\lfloor c\sqrt{n}\rfloor}^{T} |\alpha_t| - |\alpha_{t-1}| - o(\frac{1}{\sqrt{n}})\right)^2 \\ &= \frac{\left(|\alpha_T| - \left|\alpha_{T-\lfloor c\sqrt{n}\rfloor-1}\right| - o(1)\right)^2}{c\sqrt{n}} \geq \frac{(|\alpha_T| - 1/2 - o(1))^2}{c\sqrt{n}} \end{aligned} \tag{9}$$

From Theorem 1 we know $|\alpha_T| = 1 - O(\frac{\log^3 n}{n})$ which establishes $p_T \geq \frac{1/4}{c\sqrt{n}} - o(\frac{1}{\sqrt{n}}) = \Omega(\frac{1}{\sqrt{n}})$ if the assumption of Claim 5 is true or $p_T = \Omega(\frac{1}{n\log^2 n})$ if it is not, in both cases proving the theorem.

4. Dependence on the initial state

One might wonder how much this polynomial hitting time depends on the fact that the walk is from one vertex to exactly the opposite corner of the hypercube. What if the two states were not exactly in opposite corners? It is not hard to see (using the methods introduced in [AAKV01], in particular Claim 3.2 and Claim 7.2), at least in the case of one-shot hitting time, that if we start the walk in a vertex a constant Hamming distance away from $|x\rangle$ we still obtain a polynomial hitting time.

But how large can the "polynomially $|\bar{x}\rangle$ hitting" region around $|x\rangle$ be? It turns out that a polynomial hitting time can not be true in general. We give a limit that comes from the lower bound on quantum unstructured search ([BBBV97]).

Theorem 3. *The number of states $|y\rangle$ in a neighborhood of $|x\rangle$ on an n-bit hypercube (defined e.g. by a cut-off Hamming distance from $|x\rangle$) such that the quantum walk has a $(O(poly(n)), \Omega(1/poly(n))$ concurrent $(|y\rangle, |\bar{x}\rangle)$ hitting time is $O(poly(n)\cdot\sqrt{2^n})$.*

Proof of Theorem 3. Let d_c be a cut-off distance and define the neighborhood of a node $|x\rangle$ as $N_x = \{|y\rangle : d_H(x, y) \leq d_c\}$ where d_H is the Hamming distance. We can think of N_x as a ball around $|x\rangle$, but the neighborhood of a node can be defined in any arbitrary way, the arguments go through for all of them. Assume that for a ball of size M around $|x\rangle$ all $|y\rangle \in N_x$ have $(O(p(n)), \Omega(1/q(n))$ concurrent $(|y\rangle, |\bar{x}\rangle)$ hitting time, where p and q are polynomials. Let us cover the hypercube with K balls of size M, where each of the balls is centered around a node $x_1, x_2, \dots, x_K$. A simple probabilistic argument shows that we can achieve this with $K = O(n \cdot 2^n/M)$ balls. Define a quantum search algorithm as follows: starting in $|x_1\rangle$ launch an $|x\rangle$-stopped quantum walk as in Def. 3.2, where $|x\rangle$ is the marked state we are searching for. That means at every step we query the oracle with the current state of the walk and the question "Is this the marked state or not?". (We can adapt the standard oracle in Grover's algorithm [Gro96] to behave this way by measuring the auxiliary output qubit of the oracle.) We iterate this quantum walk for $p(n)$ steps and use classical amplification (repeat $q(n)$ many times). We repeat the amplified walk for each initial state $|x_i\rangle : i = 1 \dots K$. With probability close to 1 one of the walks will find the marked state. The whole algorithm takes $O(p(n) \cdot q(n) \cdot K)$ queries. From the query lower bound of $\Omega(\sqrt{2^n})$ for any unstructured quantum search algorithm [BBBV97] it follows that $K = \Omega(\sqrt{2^n}/poly(n))$ which yields the upper bound on M.

5. Quantum Routing

We have not yet succeeded to find an algorithm or exhibit an oracle[2] that provides a quantum speed-up using the polynomial hitting time on the hypercube. However we can give an application of the rapid hitting of the quantum walk to sequential routing of a packet in a noisy network with a possible adversary trying to prevent the arrival of the packet. The nodes of the network are bit-strings of length n and each node is connected to all nodes that differ by exactly one bit, so that the network has the topology of the hypercube. Consider the scenario in which a packet needs to be routed from node x to node y. The quantum routing algorithm is as follows:

(1) Let $d = d_H(x, y)$ and consider the sub-cube of dimension d spanned by the support of $x \oplus y$ (i.e. all strings z s.t. $z_i = x_i$ whenever $x_i = y_i$). The packet will be routed only on this sub-cube. The coin-space of the quantum walk is d-dimensional; call the corresponding coin operator C_d.

(2) The quantum walk is applied $T = d\frac{\pi}{2}$ times (rounded appropriately). At each time step the coin C_d acts on the appropriate directions followed by the conditional shift.

(3) After T steps the state of the system is measured. With probability $1 - O(\frac{\log^3 d}{d})$ the packet is at y.

(3') At each time step node y performs the partial measurement to see if it has received the packet or not. After T steps the probability that the packet is at y is $\Omega(\frac{1}{n \log^2 n})$. In case of failure the packet can be resent ($O(n \log^2 n)$ times) to boost the success probability close to 1.

We assume that each node v is capable to locally apply $C_d \otimes |v\rangle\langle v|$. C_d can be either given to them (as a black box operation to apply locally) or the

[2] See however [CCD$^+$03] for an exponential speed-up in query complexity using a similar rapid hitting time in the continuous walk model. Their idea can be easily adapted to give a similar polynomial query complexity using the discrete time walk on the hypercube (in fact they originally started their work with the hitting time results of the hypercube in mind). Unfortunately for the hypercube there is also a classical algorithm with polynomial query complexity.

bitpositions $x \oplus y$ of the sub-cube can be broadcast. Further nodes can locally implement the conditional shift (which requires only interactions between nearest neighbors). Both operations are local in the topology of the hypercube and can be implemented in a quantum network. The version using (3') is advantageous if T (and d) is not exactly known, like in the black box model.

Let us state the quantum advantages of this algorithm when x and y differ in $\Omega(n)$ bits (which happens almost surely when x and y are chosen at random). We are concerned here about both robustness of the algorithm against random noise (edge deletion, faulty nodes) as well as malicious attacks (adversary choses the most vulnerable edges/nodes to delete).

Classically we could route the packet deterministically (by fixing the path in advance). This strategy is fast ($T = O(d)$) but neither secure against failure of one of the routing nodes/traversed edges nor against adversarial attacks. It suffices to affect one node/edge on the fixed path and the routing will fail. A fast randomized algorithm can flip the necessary bits in some random order. This strategy is robust against deletion of a subexponential number of random edges or nodes. However it requires common knowledge of y. This in turn makes it vulnerable to adversarial attacks (it suffices to delete all the edges incident to y). A fully randomized classical routing algorithm, corresponding to a simple random walk on the cube, is robust against adversarial attacks but takes exponential time. It is here that quantum routing has an advantage. The nodes do not have to know the origin x and destination y of the packet, only $x \oplus y$. In the one-shot case even the node at y does not have to know that it is the target - only at the measurement stage will it receive the packet[3]. In the concurrent case y needs to measure at every step and hence to know it is the target, but no other node (and the adversary) will have this information. Knowledge of $x \oplus y$ alone is not sufficent to identify the most vulnerable edges (those incident or close to x and y) which reduces the adversary to random noise.

Theorem 4. *If a subexponential number of edges is deleted at random or a subexponential number of random nodes does not cooperate in the process, the success probability of the quantum routing algorithm is changed only by an exponentially small amount.*

To account for edge deletion in our model we can assume that the deleted edge is replaced by a self-loop at each of its incident nodes. A faulty node v could apply any local operation $O_v \otimes |v\rangle\langle v|$ (including measurements) instead of $C_d \otimes |v\rangle\langle v|$. Almost surely the deleted edges or faulty nodes will be in a region of the hypercube of Hamming weight $\frac{d}{2} \pm O(\sqrt{d})$. In this region there is an exponential number of nodes for each Hamming weight. Since the walk spreads symmetrically over all states of same Hamming weight, the amplitude of each single state is exponentially small and perturbing a subexponential number of them in each step can induce only an exponentially small perturbation to the state of the walk. The walk is only $O(d)$ steps long so these exponential perturbations cannot add up to anything significant. ∎

[3] This will presumably enforce a more cooperative behavior of each of the routing nodes since they all could be the target of a packet.

Note that the fact that all the adversary can do is essentially random allows us to use this type of argument. If even an exponentially small change at each step happens outside the region around Hamming weight $d/2$ the resulting perturbation can be large - this is precisely the difficulty in proving Theorem 2 from Theorem 1.

It is important to see the quantum routing algorithm not only in terms of its advantages over classical routing. It is very conceivable that quantum nets will be available in the near future and new routing strategies might have to be applied for instance to distribute qubits to establish secret keys between certain nodes in the network. Our algorithm is a first step in this spirit.

Acknowledgements. I wish to thank Ronald de Wolf for many helpful discussions and for suggesting improvements to Theorem 3, Tomohiro Yamasaki for mentioning to me his numerical results [Yam], Cris Moore for extended conversations and for sharing an updated version of his manuscript on the mixing time of the hypercube [MR02] and Wim van Dam for sharing his ideas on continuous time quantum walks. Thanks to Andris Ambainis, Daniel Gottesman, Neil Shenvi and Birgitta Whaley for very inspirational discussions. Support by ACI Sécurité Informatique, 2003-n24, projet "Réseaux Quantiques", ACI-CR 2002-40 and EU 5th framework program RESQ IST-2001-37559, as well as by DARPA and Air Force Laboratory, Air Force Materiel Command, USAF, under agreement number F30602-01-2-0524, and by DARPA and the Office of Naval Research under grant number FDN-00014-01-1-0826 is gratefully acknowledged.

References

[AAKV01] Aharonov, D., Ambainis, A., Kempe, J., Vazirani, U.: Quantum walks on graphs. In: Proceedings of 33rd ACM STOC, ACM, New York, NY, pp. 50–59, 2001

[ABN$^+$01] Ambainis, A., Bach, E., Nayak, A., Vishwanath, A., Watrous, J.: One-dimensional quantum walks. In: Proceedings of 33rd ACM STOC, ACM, New York, NY, pp. 60–69, 2001

[AF] Aldous, D., Fill, J.: Reversible markov chains and random walks on graphs. Unpublished, preprint available at http://stat-www.berkeley.edu/users/aldous/book.html.

[BBBV97] Bennett, C.H., Bernstein, E., Brassard, G., Vazirani, U.: Strengths and weaknesses of quantum computing. Siam J. Comput. **26**, 1510–1523, (1997)

[CCD$^+$03] Childs, A. M., Cleve, R., Deotto, E., Farhi, E., Gutmann, S., Spielman, D. A.: Exponential algorithmic speedup by a quantum walk. In: Proceedings of 35th ACM STOC, pp. 59–68, 2003

[CFG02] Childs, A., Farhi, E., Gutmann, S.: An example of the difference between quantum and classical random walks. Quantum Information Processing **1**, 35 (2002)

[DFK91] Dyer, M., Frieze, A., Kannan, R.: A random polynomial-time algorithm for approximating the volume of convex bodies. J. ACM **38** (1), 1–17 January (1991)

[FG98] Farhi, E., Gutmann, S.: Quantum computation and decision trees. Phys. Rev. A **58**, 915–928 (1998)

[Gro96] Grover, L.: A fast quantum mechanical algorithm for database search. In: Proceedings of the 28th ACM STOC, pages 212–219, Philadelphia, Pennsylvania, ACM Press, 1996

[GSVV01] Grigni, M., Schulman, L., Vazirani, M., Vazirani, U.: Quantum mechanical algorithms for the nonabelian hidden subgroup problem. In: Proceedings of the 33rd ACM STOC, pp. 68–74, 2001

[HRT00] Hallgren, S., Russell, A., Ta-Shma, A.: Normal subgroup reconstruction and quantum computation using group representations. In: Proceedings of the 32nd ACM STOC, pp. 627–635, 2000

[HSW02] Hofmeister, T., Schöning, U., Watanabe, O.: A probabilistic 3-SAT algorithm further improved. In: Helmut Alt and Afonso Ferreira, (eds.), STACS 2002, 19th Annual Symposium on Theoretical Aspects of Computer Science, Antibes - Juan les Pins, France, March 14-16, 2002, Proceedings, volume 2285 of Lecture Notes in Computer Science pp. 192–202. Springer, 2002

[JS89] Jerrum, M., Sinclair, A.: Approximate counting, uniform generation and rapidly mixing Markov chains. Information and Computation **82** (1) 93–133 (1989)

[JSV01] Jerrum, M., Sinclair, A., Vigoda, E.: A polynomial-time approximation algorithm for the permanent of a matrix with non-negative entries. In: Proceedings of the 33rd ACM STOC, ACM, New York, NY, pp. 712–721, 2001

[Kem03a] Kempe, J.: Quantum random walks - an introductory overview. Contemporary Physics **44** (4), 302–327 (2003)

[Kem03b] Kempe, J.: Quantum walks hit exponentially faster. In: RANDOM-APPROX 2003, Lecture Notes in Computer Science, Heidelberg, Springer, pp. 354–369, 2003

[Mey96] Meyer, D.: From quantum cellular automata to quantum lattice gases. J. Stat. Phys. **85**, 551–574 (1996)

[MR95] Motwani, R., Raghavan, P.: Randomized Algorithms. Cambridge University Press, 1995

[MR02] Moore, C., Russell, A.: Quantum walks on the hypercube. In: Proc. RANDOM 2002, Lecture Notes in Computer Science, Cambridge, MA, Springer, pp. 164–178, 2002

[NC00] Nielsen, M.A., Chuang, I.L.: Quantum Computation and Quantum Information. Cambridge University Press, Cambridge, UK, 2000

[Pap94] Papadimitriou, C.: Computational Complexity. Addison Wesley, Reading, Massachusetts, 1994

[Sch99] Schöning, U.: A probabilistic algorithm for k-SAT and constraint satisfaction problems. In: 40th Annual Symposium on Foundations of Computer Science, IEEE, pp. 410–414, 1999

[Sho97] Shor, P.W.: Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. SIAM J. Comp. **26** (5), 1484–1509 (1997)

[Sim97] Simon, D.: On the power of quantum computation. SIAM J. Comp. **26** (5), 1474–1483 (1997)

[SKW03] Shenvi, N., Kempe, J., Whaley, K.B.: A quantum random walk search algorithm. Phys. Rev. A **67** (5), 052307 (2003)

[Wat01] Watrous, J.: Quantum simulations of classical random walks and undirected graph connectivity. J. Comp. Sys. Sci. **62** (2), 376–391 (2001)

[Yam] Yamasaki, T.: private communication.

[YKI02] Yamasaki, T., Kobayashi, H., Imai, H.: An analysis of absorbing times of quantum walks. In: C. Calude, M.J. Dinneen, and F. Peper, editors, Unconventional Models of Computation, Third International Conference, UMC 2002, Kobe, Japan, October 15-19, 2002, Proceedings, volume 2509 of Lecture Notes in Computer Science, Springer, pp. 315–330, 2002

A. Continuous - Time Quantum Random Walk

The continuous-time walk has been defined by Farhi and Gutmann [FG98] as a quantum version of the classical continuous-time walk (see Sec. 2.1). To make the classical continuous walk with generator Q quantum one simply sets $U(t) = exp(iQt)$, which is unitary as long as $Q = Q^\dagger$ (which is the case for simple random walks on undirected graphs). This walk works directly with the space formed by the nodes of the graph and does not require auxiliary coin spaces. In general, however, it is hard to see how to carry out such a walk in a generically programmable way using only local information about the graph. Instead the continuous time walk might correspond to special purpose analog computers, where we build in interactions corresponding to the desired Hamiltonian Q.

For the hypercube the continuous time quantum walk is described by the following transformation on the space spanned by n-bit strings [MR02]:

$$U_{walk}(t) = e^{i\frac{t}{n}(X_1+X_2+\cdots+X_n)} = e^{i\frac{t}{n}X_1} \cdot e^{i\frac{t}{n}X_2} \cdot \ldots \cdot e^{i\frac{t}{n}X_n} \qquad (10)$$

where X_i acts only on the ith bit as $X|0\rangle = |1\rangle$ and $X|1\rangle = |0\rangle$. The expression in the exponential corresponds to the adjacency matrix of the hypercube. The unitary transformation $U_{walk}(t)$ can be simulated uniformly by a quantum circuit with $O(n)$ local gates.

One - shot hitting time:

Theorem 5. *The continuous time quantum random walk has a* $(T = \frac{\pi n}{2}, 1)$ *and a* $(T = \frac{\pi n}{2} \pm n^\beta, 1 - O(1/n^{1-2\beta}))$ *one shot hitting time for* $\beta = const < 1/2$.

Proof. From $e^{i\frac{t}{n}X} = \cos\frac{t}{n}\mathbf{1} + i\sin\frac{t}{n}X$ it is easy to calculate the amplitude α_t of the state $|11\ldots1\rangle$ in the state $|\Phi_t\rangle := U(t)|00\ldots0\rangle$. It gives $|\alpha_t| = (\sin\frac{t}{n})^n$. Write $T = \pi n/2 \pm \epsilon$ with $\epsilon = O(n^\beta)$. Then $\sin\frac{t}{n} = \sin(\pi/2 \pm \epsilon/n) = 1 - O(\epsilon^2/n^2)$. This gives $|\alpha_t| = (1 - O(n^{2\beta}/n^2))^n$ which is $1 - O(1/n^{1-2\beta})$ for $\beta = const < 1/2$. □

This corresponds exactly to what we have shown in the discrete case Theorem 1.

Concurrent hitting time:

There is some arbitrariness in defining an $|x\rangle$-stopped continuous time walk. If the walk is to be continuous one could argue that the measurement ("Is the state $|x\rangle$ or not?") should also be continuous. In this case the measurements should not be projective, but rather "weak" measurements. We do not wish at this stage to introduce a new apparatus of notations and tools, in particular since it is not obvious how to model weak measurements on a quantum computer generically. To compare the two models we chose to measure the continuous time walk at discrete time intervals $(t = 1, 2, \ldots)$.

Definition A.1 **($|x\rangle$-stopped walk and concurrent hitting time:).** *The* $|x\rangle$*-stopped walk is the iterative process where first a measurement with* $\{\Pi_0 = |x\rangle\langle x|, \Pi_1 = \mathbf{1} - \Pi_0\}$ *is performed. If* $|x\rangle$ *is measured the walk is stopped, otherwise* $U_{walk}(1)$ *is applied and the procedure is repeated. The walk has a* (T, p) *concurrent hitting time if the probability to stop before time* T *is* $> p$.

Theorem 6. *The continuous time walk on the hypercube has a* $(T = \frac{\pi n}{2}, \Omega(\frac{1}{\sqrt{n}}))$ *concurrent hitting time.*

Proof. We adapt the notations and claims of the proof of theorem 2. Let α_t and β_t be defined as the amplitudes of the target state $|f\rangle = |11\ldots 1\rangle$ in the unmeasured resp. measured walk at integer times and let the unnormalized state of the unmeasured walk at time t be $|\tilde{\Phi}_t\rangle$. Then Claim 2 and Claim 3 hold without change with $\gamma_k = \langle f|U(k)|f\rangle$.

The quantities here are easy to calculate: $\alpha_t = i^n(\sin\frac{t}{n})^n$ and $\gamma_t = (\cos\frac{t}{n})^n$. This means that $-i^n\alpha_t$ are monotonically increasing and γ_t are monotonically decreasing for $t < T = \frac{\pi n}{2}$. This in turn suffices to prove Claim 5 with the modification that now $-i^n\beta_k \geq 0$ and $|\beta_{t+1}| \geq |\alpha_{t+1}| - |\alpha_t|$. As in Claim 4.3 we can set $t_c = c\sqrt{n}$ and note that $\left|\alpha_{T-t_c}\right| = (\sin\frac{\pi}{2} - \frac{c}{\sqrt{n}})^n = (1 - \frac{c^2}{n} + O(\frac{1}{n^2}))^n = e^{-c^2}$ up to exponentially small terms. Pick c such that $\left|\alpha_{T-t_c}\right| = 1/2$. Then we can adopt Eq. (9) to

$$p_T = \sum_{t=n}^{T}|\beta_t|^2 \geq \sum_{t=T-c\sqrt{n}+1}^{T}|\beta_t|^2 \geq \frac{(\sum_{t=T-c\sqrt{n}}^{T}|\beta_t|)^2}{c\sqrt{n}}$$

$$\geq \frac{(\sum_{t=T-c\sqrt{n}}^{T}|\alpha_t| - |\alpha_{t-1}|)^2}{c\sqrt{n}} = \frac{(1-1/2)^2}{c\sqrt{n}} = \frac{1}{4c\sqrt{n}}$$

which proves the theorem. □