Quadratic Speedup for Finding Marked Vertices by Quantum Walks

Andris Ambainis
Faculty of Computing, University of Latvia
Riga, Latvia
andris.ambainis@lu.lv

András Gilyén
QuSoft, CWI and University of Amsterdam
Amsterdam, The Netherlands
gilyen@cwi.nl

Stacey Jeffery
QuSoft and CWI
Amsterdam, The Netherlands
smjeffery@gmail.com

Martins Kokainis
Faculty of Computing, University of Latvia
Riga, Latvia
martins.kokainis@lu.lv

ABSTRACT

A quantum walk algorithm can detect the presence of a marked vertex on a graph quadratically faster than the corresponding random walk algorithm (Szegedy, FOCS 2004). However, quantum algorithms that actually find a marked element quadratically faster than a classical random walk were only known for the special case when the marked set consists of just a single vertex, or in the case of some specific graphs. We present a new quantum algorithm for finding a marked vertex in any graph, with any set of marked vertices, that is (up to a log factor) quadratically faster than the corresponding classical random walk, resolving a question that had been open for 15 years.

CCS CONCEPTS

• **Theory of computation → Quantum computation theory**; *Algorithm design techniques*; *Random walks and Markov chains.*

KEYWORDS

quantum algorithms, quantum search, quantum walks, search by random walk, Markov chains

ACM Reference Format:
Andris Ambainis, András Gilyén, Stacey Jeffery, and Martins Kokainis. 2020. Quadratic Speedup for Finding Marked Vertices by Quantum Walks. In *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing (STOC '20), June 22–26, 2020, Chicago, IL, USA.* ACM, New York, NY, USA, 13 pages. https://doi.org/10.1145/3357713.3384252

1 INTRODUCTION

As shown by Szegedy [14], quantum walks provide a quadratic speedup over classical random walks for search tasks. If a classical random walk hits a marked element in an expected number of HT steps, called the *hitting time*, then the quantum walk runs in time $O(\sqrt{\mathrm{HT}})$. However, this speedup comes with a caveat: the quantum walk does not necessarily *find* a marked element, but it can *detect* a deviation from the starting state caused by marked elements. This issue has been well known since Szegedy's work in 2004, yet so far it has eluded all attempts to solve it.

Several generalizations of Szegedy's framework have been proposed, but they only solve this issue in restricted cases. Tulsi [15] showed how to solve it for the random walk on an $N \times N$ grid with exactly one marked element. Here, the classical hitting time is $\mathrm{HT} = O(N^2 \log N)$. While Szegedy's algorithm detects the presence of a marked element in $O(\sqrt{\mathrm{HT}}) = O(N\sqrt{\log N})$ steps, measuring the final state of the algorithm only gives the marked element with probability $\Theta(1/\log N)$. Tulsi showed how to improve this to $\Theta(1)$, with the running time remaining $O(N\sqrt{\log N})$. Magniez, Nayak, Richter and Santha [12] extended this to the random walk on any vertex transitive graph with exactly one marked element.

In addition Magniez, Nayak, Roland and Santha [13] presented an alternative extension of Szegedy's work, giving a quantum algorithm for finding a marked vertex that runs in a number of steps $O(\sqrt{1/(\delta\varepsilon)})$, where δ is the eigenvalue gap of (the Markov chain corresponding to) the walk and ε is the probability that a vertex is initially marked. This bound can be as small as $O(\sqrt{\mathrm{HT}})$ in certain cases, but significantly larger in others.

Later, Krovi, Magniez, Ozols and Roland [8] proposed a new algorithm (based on a new notion of interpolated quantum walk) that achieves a quadratic advantage for finding a marked element for a random walk on any graph G with exactly one marked element. The same result was achieved by Dohotaru and Høyer [5], using a different method.

In the general case (with multiple marked elements), the algorithm of Krovi et al. finds a marked element, but takes time $O(\sqrt{\mathrm{HT}^+})$ where HT^+ is the *extended hitting time* of the walk. HT^+ is a new quantity obtained by modifying the expression for HT in terms of eigenvalues and eigenvectors of the walk. If there is only one marked element, then $\mathrm{HT}^+ = \mathrm{HT}$ and this yields the quadratic advantage for the quantum walk. However, HT^+ may be significantly larger than HT when there are multiple marked elements,[1] as we show in Section 5.

[1] The first version of the paper by Krovi et al. [8] claimed $\mathrm{HT}^+ = \mathrm{HT}$ for any number of marked elements but this turned out to be false, as corrected by the authors in later versions.

STOC '20, June 22–26, 2020, Chicago, IL, USA

ACM ISBN 978-1-4503-6979-4/20/06…$15.00
https://doi.org/10.1145/3357713.3384252

Lastly, for a two-dimensional grid, a quadratic advantage for any set of marked elements was achieved by Høyer and Komeili [7] using a divide-and-conquer approach. However, their approach is specific to the two-dimensional grid and does not seem to generalize even to grids in higher dimensions.

1.1 Our Contributions

In this paper, we finally resolve the problem of finding a marked element quadratically faster (up to a log factor) compared to the classical random walk, on any graph, for any number and any arrangement of marked elements.

Our main new algorithm combines interpolated walks with the recently invented *quantum fast-forwarding* technique of Apers and Sarlette [2]. Quantum fast-forwarding is a primitive that allows one to replace t steps of a classical random walk with $O\left(\sqrt{t}\right)$ steps of a quantum walk, in a certain sense. A caveat is that quantum fast-forwarding may only produce the final state with a very small success probability. However, in our application, it succeeds with probability $\widetilde{\Omega}(1)$. This is shown by an insightful argument that interprets the success probability of quantum fast-forwarding in terms of the classical random walk. Namely, it corresponds to the probability that the classical random walk, started in a random unmarked vertex, visits a marked vertex after t steps, but returns to an unmarked vertex after t additional steps. We show that this probability can be tuned to be $\widetilde{\Omega}(1)$ by adjusting the interpolation parameter of the walk. After describing some preliminaries in Section 2, we discuss Algorithm 1 and the main result in Section 3, and provide the details of the analysis in Section 4.

In Section 5 we show that the gap between HT^+ and HT can indeed be very large. We construct an arrangement of marked elements on an $N \times N$ grid for which $\mathrm{HT}^+ = \Omega(N^2)$ but $\mathrm{HT} = O(f(N))$ where f grows to infinity arbitrarily slowly. This shows that the algorithm of Krovi et al. can be severely suboptimal when there are multiple marked elements. The reason is that their algorithm actually solves a harder problem: it samples from the stationary distribution restricted to marked vertices (which is the uniform distribution in case of the grid). Hence, their algorithm may be slow in cases when sampling from this distribution is substantially more difficult than just finding some marked element.

In Section 6 we present a second, simpler, new algorithm, which we conjecture[2] to find a marked element in time $O\big(\sqrt{\mathrm{HT}}\big)$, for an arbitrary arrangement of marked elements (Conjecture 11). This second algorithm is also based on the idea of interpolated walks, but uses it differently from [8]. Namely, Algorithm 2 just runs the interpolated walk for $O\big(\sqrt{\mathrm{HT}}\big)$ steps (instead of using eigenvalue estimation to produce an eigenstate of the walk, as in [8]). Our numerical experiments suggest, that for any arrangement of marked vertices, there is a choice of the interpolation parameter and a choice of running time $t = O\big(\sqrt{\mathrm{HT}}\big)$ which results in the walk producing a marked vertex with probability $\Omega(1)$. We tested this conjecture for all examples with $\mathrm{HT}^+ \gg \mathrm{HT}$ that we found.

2 PRELIMINARIES

2.1 Markov Chains and Random Walks

For a random variable Z and probability distribution ρ, we will use $Z \sim \rho$ to indicate that Z is distributed according to ρ.

A sequence of random variables $Y = (Y_i)_{i=0}^{\infty}$ is a *Markov chain* if for all $i > 0$,

$$\Pr(Y_i = y_i | Y_0 = y_0, \ldots, Y_{i-1} = y_{i-1}) = \Pr(Y_i = y_i | Y_{i-1} = y_{i-1}).$$

A (time-independent) Markov chain on a discrete state space X with $|X| = n$ is specified by an $n \times n$ row-stochastic matrix $\mathcal{P}$, whose xy-entry $\mathcal{P}_{xy}$ denotes the probability that the Markov chain makes a transition from state $x \in X$ to the state $y \in X$ in one step. For a distribution ρ on X, we say that Y is a Markov chain evolving according to $\mathcal{P}$ starting from ρ if $Y_0 \sim \rho$, and for all $i > 0$ and $x, y \in X$, $\Pr(Y_i = y | Y_{i-1} = x) = \mathcal{P}_{xy}$. We will left-multiply with probability (row) vectors to follow the common conventions in the literature for Markov chains, so if $Y_0 \sim \rho$, then $Y_i \sim \rho\mathcal{P}^i$, for any $i \geq 0$.

We say that $\mathcal{P}$ is *ergodic* if for a large enough $t \in \mathbb{N}$ all elements of $\mathcal{P}^t$ are non-zero. For an ergodic $\mathcal{P}$ there exists a unique stationary distribution π such that $\pi\mathcal{P} = \pi$, and we define the *time-reversed* Markov chain as $\mathcal{P}^* := \mathrm{diag}(\pi)^{-1} \cdot \mathcal{P}^T \cdot \mathrm{diag}(\pi)$. We say that $\mathcal{P}$ is *reversible* if it is ergodic and $\mathcal{P}^* = \mathcal{P}$. Note that reversibility can be equivalently expressed by the *detailed-balance* equations:

$$\forall x, y \in X\colon \pi_x \mathcal{P}_{xy} = \pi_y \mathcal{P}_{yx}, \tag{1}$$

intuitively meaning that in the stationary distribution for each pair of states the probability of a transition between the states in both directions is that same. Moreover, it is easy to see that if $\mathcal{P}$ is reversible then so is $\mathcal{P}^t$ for every $t \in \mathbb{N}$.

For an ergodic Markov chain $\mathcal{P}$, we define the discriminant matrix D such that its xy-entry is $\sqrt{\mathcal{P}_{xy}\mathcal{P}^*_{yx}}$. It is easy to see that

$$D = \mathrm{diag}(\pi)^{\frac{1}{2}} \cdot \mathcal{P} \cdot \mathrm{diag}(\pi)^{-\frac{1}{2}}. \tag{2}$$

This form has several important consequences. First of all the spectra of $\mathcal{P}$ and D coincide, and moreover, the vector $\sqrt{\pi}$ that we get from π by taking the square root element-wise, is a left eigenvector of D with eigenvalue 1. Also from the definition $D_{xy} = \sqrt{\mathcal{P}_{xy}\mathcal{P}^*_{yx}}$ it follows that for reversible Markov chains, D is a symmetric matrix, and therefore its singular values and eigenvalues coincide up to sign.

Reversible Markov chains are equivalent to random walks on weighted graphs; for a survey on the topic see Lovász [10]. They have been used to design search algorithms in various contexts. Specifically, if $\mathcal{P}$ is a random walk on a state space X, and $M \subset X$ is a set of *marked* vertices, then a randomized algorithm that begins in any vertex $x \in X$ and repeatedly makes a step of the walk, while checking whether the current state is marked, will eventually find some $x \in M$ (assuming M is non-empty). When the algorithm starts in the stationary distribution of $\mathcal{P}$, the expected number of steps needed before a marked vertex is reached is called the *hitting time*, and is denoted $\mathrm{HT} = \mathrm{HT}(\mathcal{P}, M)$. Let Z be the smallest number such that $Y_Z \in M$, where Y is a Markov chain evolving according to $\mathcal{P}$ starting from π, then $\mathrm{HT}(\mathcal{P}, M) = \mathbb{E}[Z]$. Moreover,

[2] Very recently a slightly weaker version of our conjecture has been proven by Apers, Gilyén, and Jeffery [1]. They essentially prove Conjecture 11 up to a log factor in the success probability. Remarkably, their proof heavily builds on the correctness of our Algorithm 1, and its implementation details.

by Markov's inequality, for any positive real number c we have $\Pr(Z > c\mathrm{HT}(\mathcal{P}, M)) \leq \frac{1}{c}$.

Thus, for any reversible Markov chain $\mathcal{P}$ on X, and $M \subset X$, if C is the complexity of checking whether $x \in M$ (for an arbitrary $x \in X$), U is the cost of taking one step of the walk $\mathcal{P}$, and S is the cost of sampling according to the stationary distribution, then there is a randomized algorithm that finds a marked vertex with high probability in complexity $O(\mathsf{S} + \mathrm{HT}(\mathsf{U} + \mathsf{C}))$. In the next subsection, we will consider quantum analogues of this procedure.

For simplicity in the rest of the paper we will work with reversible time-independent Markov chains, unless otherwise stated.

2.2 Interpolated Walks and Quantum Walk Search Algorithms

Interpolated walks. Some previous quantum walk algorithms are based on the notion of *interpolated walk*. Intuitively speaking such a walk works as follows: first it checks whether the current node is marked. It the node is *unmarked*, then it performs a normal step of the walk; but if it is *marked*, then it performs a normal walk step only with probability $1 - s$, and with probability s it stays at the current marked node.

Let us fix some reversible Markov chain $\mathcal{P}$ and marked set $M \subset X$. We first define the *absorbing* walk operator $\mathcal{P}'$ as the modified Markov chain that, once it hits the set of marked vertices M, stays where it is. If we arrange the states of X so that the unmarked states $U := X \setminus M$ come first, the matrices $\mathcal{P}$ and $\mathcal{P}'$ have the following block structure:

$$\mathcal{P} := \begin{pmatrix} \mathcal{P}_{UU} & \mathcal{P}_{UM} \\ \mathcal{P}_{MU} & \mathcal{P}_{MM} \end{pmatrix}, \qquad \mathcal{P}' := \begin{pmatrix} \mathcal{P}_{UU} & \mathcal{P}_{UM} \\ 0 & I \end{pmatrix}.$$

We define the *interpolated walk* operator, for $s \in [0, 1)$, as:

$$\mathcal{P}(s) := (1 - s)\mathcal{P} + s\mathcal{P}', \tag{3}$$

staying at a marked vertex with probability s. We denote the corresponding discriminant matrix by $D(s)$. Let Π_M be the projector onto marked vertices and let $\Pi_U := I - \Pi_M$ be the projector onto unmarked vertices. Then we define $\pi_U := \pi\Pi_U$ and $\pi_M := \pi\Pi_M$ as the row vectors that are obtained by restricting π to sets U and M, respectively. We denote the probability that an element is marked in the stationary distribution by $p_M := \sum_{x \in M} \pi_x = \|\pi_M\|_1$. Then $\pi' := \pi_M/p_M$ is a stationary distribution of $\mathcal{P}'$.[3] In analogy to the definition of $\mathcal{P}(s)$ in Eq. (3), let $\pi(s)$ be a convex combination of π and π', appropriately normalized:

$$\pi(s) := \frac{(1 - s)\pi + s\pi'}{(1 - s) + s p_M} = \frac{1}{1 - s(1 - p_M)}((1 - s)\pi_U + \pi_M). \tag{4}$$

Krovi et al. [8] showed that for any $s \in [0, 1)$, $\mathcal{P}(s)$ is a reversible ergodic Markov chain with unique stationary distribution $\pi(s)$.

Quantum walk operator. For a (reversible) Markov chain $\mathcal{P}$, let $V(\mathcal{P})$ be a unitary such that[4]

$$\forall x \in X\colon V(\mathcal{P})|\bar{0}\rangle|x\rangle = \sum_{y \in X} \sqrt{\mathcal{P}_{xy}}|y, x\rangle,$$

where $|\bar{0}\rangle$ is some fixed reference state. The action of $V(\mathcal{P})$ is analogous to taking one step of the random walk $\mathcal{P}$ in superposition. Let Swap be defined by the action $|x, y\rangle \mapsto |y, x\rangle$, for all $x, y \in X$, and let Ref $= (2|\bar{0}\rangle\langle\bar{0}| - I) \otimes I$. The corresponding *quantum walk operator* is

$$W(\mathcal{P}) := V^\dagger(\mathcal{P})\,\textsc{Swap}\,V(\mathcal{P})\,\textsc{Ref}.$$

Note that $\langle\bar{0}|\langle x|W(\mathcal{P})|\bar{0}\rangle|y\rangle = \sqrt{\mathcal{P}_{xy}\mathcal{P}_{yx}} = D_{xy}$.

Extended hitting time. For any $s \in [0, 1)$, suppose that $D(s)$ has eigenvalue decomposition

$$\sum_{k=1}^{n} \lambda_k(s)|v_k(s)\rangle\langle v_k(s)|,$$

with $\lambda_n(s) = 1$, so $\lambda_k(s) < 1$ for all $k < n$. Then we can define[5]

$$\mathrm{HT}(s) := \sum_{k=1}^{n-1} \frac{|\langle v_k(s)|\sqrt{\pi_U}\rangle|^2}{1 - \lambda_k(s)},$$

and

$$\mathrm{HT}^+(\mathcal{P}, M) := \lim_{s \to 1} \mathrm{HT}(s),$$

where $|\sqrt{\pi_U}\rangle = \sum_{x \in U} \sqrt{\pi_x}|x\rangle$. We call HT^+ the *extended hitting time.* To put this definition into context, note that one can show $\mathrm{HT}(\mathcal{P}, M) = \sum_{k=1}^{n-|M|} \frac{|\langle v'_k|\sqrt{\pi_U}\rangle|^2}{1-\lambda'_k}$, where λ'_k ranges over the ($\neq 1$) eigenvalues of $D(1)$ and $|v'_k\rangle$ are the corresponding eigenvectors. For a proof see, e.g., [8, Proposition 9].

Quantum walk search algorithms. We introduce the following black-box (oracle) operations:

- `Check`(M): checks if a given vertex is marked by mapping $|x\rangle|b\rangle$ to $|x\rangle|b\rangle$ if $x \notin M$ and $|x\rangle|b \oplus 1\rangle$ if $x \in M$, where $|x\rangle$ is the vertex register and $b \in \{0, 1\}$;
- `Setup`$(\mathcal{P})$: construct the superposition
$$|\sqrt{\pi}\rangle = \sum_{x \in X} \sqrt{\pi_x}|x\rangle;$$
- `Update`$(\mathcal{P})$: perform one update step. More precisely implement (separately, controlled versions of[6]) Swap, Ref, and $V(\mathcal{P})^{\pm 1}$.

Each of these operations has a corresponding associated implementation cost, which we denote by C, S, and U, respectively.

For implementing the interpolated quantum walk we define a modified version of the update operator, which is a direct quantum analogue to the interpolated classical update: if the current vertex is marked flip a coin and do noting when the result is "heads", otherwise proceed as usually. Accordingly the modified quantum update operator $V(\mathcal{P}, s)$ for all $x \in U$ acts as $I \otimes V(\mathcal{P})$ on the initial state $|0\rangle|\bar{0}\rangle|x\rangle$, and for $x \in M$ acts as

$$|0\rangle|\bar{0}\rangle|x\rangle \mapsto \sqrt{1 - s}|0\rangle V(\mathcal{P})|\bar{0}\rangle|x\rangle + \sqrt{s}|1\rangle|\bar{0}\rangle|x\rangle.$$

We define the interpolated quantum walk operator as

$$W(s) := V^\dagger(\mathcal{P}, s)\,\textsc{Swap}'\,V(\mathcal{P}, s)\,\textsc{Ref}', \tag{5}$$

[3] In fact, any distribution with support only on marked states is stationary for $\mathcal{P}'$.

[4] Note that here we swapped the role of the two registers compared to some previous works, in order to make the resemblance with block-encodings [4, 6] more apparent, see Section 2.3 for more details.

[5] Note that this definition slightly differs from the definitions of [8], namely these quantities are $(1 - p_M)$-times smaller here; this additional factor in [8] comes from conditioning on starting in an unmarked state. Our notation matches other standard definitions in the literature, and since the interesting regime is when $p_M \ll 1$, this is anyway an unimportant difference.

[6] This is mostly needed for implementing interpolated versions of the quantum walk.

where $\textsc{Swap}' := |0\rangle\langle 0| \otimes \textsc{Swap} + |1\rangle\langle 1| \otimes I$ and $\textsc{Ref}' := (2|0\rangle\langle 0| \otimes |\bar{0}\rangle\langle\bar{0}| - I) \otimes I$. It is easy to see that

$$\langle 0|\langle\bar{0}|\langle x|W(s)|0\rangle|\bar{0}\rangle|y\rangle = D_{xy}(s). \tag{6}$$

Note that $W(s)$ can be implemented[7] for any $s \in [0,1)$ in cost of order $\mathsf{C} + \mathsf{U}$, in the following way. First check whether $x \in X$ is marked, and if it is, then apply the map $|0\rangle \mapsto \sqrt{1-s}|0\rangle + \sqrt{s}|1\rangle$ to the first qubit. Controlled by the first qubit's state being $|0\rangle$, apply $V(\mathcal{P})$ to the last two registers. (From now on for simplicity we will just write $|\bar{0}\rangle$ instead of $|0\rangle|\bar{0}\rangle$ when we work with interpolated quantum walks $W(s)$.)

While a classical random walk can find a marked vertex in complexity[8] $O(\mathsf{S} + \mathrm{HT}(\mathsf{U} + \mathsf{C}))$, Krovi et al. [8] showed that using the quantum walk $W(s)$ one can find a marked vertex in complexity $O\big(\mathsf{S} + \sqrt{\mathrm{HT}^{+}}(\mathsf{U} + \mathsf{C})\big)$. However, in Section 5, we show that HT^{+} may be much larger than HT. In Section 3, we show that in fact, a quantum algorithm can find a marked vertex in complexity $\widetilde{O}\Big(\mathsf{S} + \sqrt{\mathrm{HT}}(\mathsf{U} + \mathsf{C})\Big)$, see Theorem 3 (full analysis in Section 4).

2.3 Quantum Fast-forwarding

We will use the quantum fast-forwarding technique of Apers and Sarlette [2], which allows us to, in some very "quantum" sense, apply t steps of a walk in only $\sqrt{t}$ calls to its update operation. We state their main result in a slightly adapted form.

Theorem 1 ([2]). *Let $\varepsilon \in (0,1)$, $s \in [0,1]$ and $t \in \mathbb{N}$. Let $\mathcal{P}$ be any reversible Markov chain on state space X, and let Q be the cost of implementing the (controlled) quantum walk operator $W(s)$. There is a quantum algorithm with cost $O\Big(\sqrt{t\log(1/\varepsilon)}\mathsf{Q}\Big)$ that takes input $|\bar{0}\rangle|\psi\rangle \in \mathrm{span}\{|\bar{0}\rangle|x\rangle : x \in X\}$, and outputs a state that is ε-close to a state of the form*

$$|0\rangle^{\otimes a}|\bar{0}\rangle D^{t}|\psi\rangle + |\Gamma\rangle$$

where $a = O(\log(t\log(1/\varepsilon)))$ and $|\Gamma\rangle$ is some garbage state that has no support on states containing $|0\rangle^{\otimes a}|\bar{0}\rangle$ in the first two registers.

To gain some intuition it is useful to think about the walk operator W as a block-encoding of the discriminant matrix D, i.e., a unitary matrix containing D in the top-left corner. In this terminology, fast-forwarding reads as implementing a block-encoding of D^{t} by using the block-encoding of D only roughly $\sqrt{t}$ times. By this insight one can rederive Theorem 1 via recent qubitisation [11] or quantum singular value transformation [6] result as well.

Consider the case when we start with the subnormalised vector $|\sqrt{\pi_U}\rangle = \sum_{x\in U}\sqrt{\pi_x}|x\rangle$ and apply the "fast-forwarded" Markov chain from Theorem 1, before measuring. We show how to re-express the probability of measuring a marked element in terms of the interpolated walk $\mathcal{P}(s)$. The probability of measuring a marked state is given by the square of:[9]

$$\begin{aligned}
\left\|\Pi_M D^{t}(s)|\sqrt{\pi_U}\rangle\right\| &\geq \left\|\Pi_M D^{t}(s)|\sqrt{\pi_U}\rangle\right\|\left\|\Pi_M D^{\hat{t}}(s)|\sqrt{\pi_U}\rangle\right\| \\
&\geq \langle\sqrt{\pi_U}|D^{t}(s)\Pi_M D^{\hat{t}}(s)|\sqrt{\pi_U}\rangle \qquad \text{by Cauchy-Schwarz} \\
&= \langle\sqrt{\pi_U}|\mathrm{diag}(\pi(s))^{\frac{1}{2}}\mathcal{P}^{t}(s)\Pi_M\mathcal{P}^{\hat{t}}(s)\mathrm{diag}(\pi(s))^{-\frac{1}{2}}|\sqrt{\pi_U}\rangle \\
&\qquad\qquad \text{by Eq. (2)} \\
&= \langle\sqrt{\pi_U}|\mathrm{diag}(\pi)^{\frac{1}{2}}\mathcal{P}^{t}(s)\Pi_M\mathcal{P}^{\hat{t}}(s)\mathrm{diag}(\pi)^{-\frac{1}{2}}|\sqrt{\pi_U}\rangle \quad \text{by Eq. (4)} \\
&= \sum_{x,z\in U}\pi_x\langle x|\mathcal{P}^{t}(s)\Pi_M\mathcal{P}^{\hat{t}}(s)|z\rangle. \qquad (7)
\end{aligned}$$

In the first inequality $\hat{t}$ can be an arbitrary positive integer since $\|D(s)\| = 1$; in the penultimate equality we have used the fact from Eq. (4), that $\pi(s)$ restricted to U is proportional to π, so for some α, $\langle\sqrt{\pi_U}|\mathrm{diag}(\pi(s))^{\frac{1}{2}} = \langle\sqrt{\pi_U}|\sqrt{\alpha}\,\mathrm{diag}(\pi_U)^{\frac{1}{2}}$, and

$$\mathrm{diag}(\pi(s))^{-\frac{1}{2}}|\sqrt{\pi_U}\rangle = \frac{1}{\sqrt{\alpha}}\mathrm{diag}(\pi_U)^{-\frac{1}{2}}|\sqrt{\pi_U}\rangle.$$

The expression in (7) can be equivalently expressed as

$$\left\|\langle\pi_U|\mathcal{P}^{t}(s)\Pi_M\mathcal{P}^{\hat{t}}(s)\Pi_U\right\|_1,$$

which is the probability that upon starting from the stationary distribution of $\mathcal{P}$ and evolving according to $\mathcal{P}(s)$, the first vertex is unmarked, after t steps we are at a marked vertex, and after another $\hat{t}$ steps we are at an unmarked vertex again. We summarize this in the following lemma:

Lemma 2. *Let $s \in [0,1)$, and $\mathcal{P}$ be any reversible Markov process. Let $Y(s) = (Y_i(s))_{i=0}^{\infty}$ be the Markov chain evolving according to $\mathcal{P}(s)$ starting from $Y_0(s) \sim \pi$. Then for any $t, \hat{t} \in \mathbb{N}$, letting $t' = t + \hat{t}$:*

$$\left\|\Pi_M D^{t}(s)|\sqrt{\pi_U}\rangle\right\| \geq \Pr(Y_0(s) \in U, Y_t(s) \in M, Y_{t'}(s) \in U). \tag{8}$$

Thus, it suffices to lower bound the probability in (8) by $\widetilde{\Omega}(1)$ for some choice of s and $t = O(\mathrm{HT})$; this is established by Corollary 7 in Section 4. Note that $t' > t$ can be arbitrarily large in principle, but we will ultimately choose some $t' = O(\mathrm{HT})$. In fact, we will not even directly use Lemma 2, because we can get a slightly better bound by the direct argument presented in the proof of Corollary 8.

3 THE MAIN RESULT

Our main result is the following.

Theorem 3. *Let $\mathcal{P}$ be any reversible Markov chain on a finite state space X, and let $M \subset X$ be a marked set. Then Algorithm 1 outputs a vertex x from M with success probability at least $\frac{2}{3}$ with cost*

$$O\Big(\mathsf{S}\sqrt{\log(\mathrm{HT})} + \sqrt{\mathrm{HT}}(\mathsf{U} + \mathsf{C})\sqrt{\log(\mathrm{HT})\log\log(\mathrm{HT})}\Big),$$

where HT is a known upper bound on $\mathrm{HT}(\mathcal{P}, M)$, S is the cost of the $\mathtt{Setup}(\mathcal{P})$ operation, U is the cost of the $\mathtt{Update}(\mathcal{P})$ operation, and C is the cost of the $\mathtt{Check}(M)$ operation.

Now we sketch the proof of Theorem 3. The two key ingredients are Theorem 1 and Corollary 8, which is proven in Section 4.

[7] We note that [8, Appendix B.2] also describes a way to implement the interpolated quantum walk operator with similar complexity but additionally require (query) access to the diagonal entries of $\mathcal{P}$.

[8] We note that in the classical case, S is the cost of *classically* sampling from π, and U is the cost of classically sampling a neighbour of the current vertex. These classical sampling operations may be cheaper than Setup and Update, but in applications, they are often similar to the quantum costs.

[9] For a parametrized matrix $M(s)$ we denote $(M(s))^{t}$ simply by $M^{t}(s)$, so for example $\mathcal{P}^{t}(s) \equiv (\mathcal{P}(s))^{t}$.

Algorithm 1 Our new fast-forwarding-based search algorithm

```
Input: Oracles for the Markov chain $\mathcal{P}$ and the marked set $M$, and upper bound HT on $\mathrm{HT}(\mathcal{P}, M)$
Set $T := 72\mathrm{HT}$ and $S := \left\{1 - \frac{1}{r} : r \in \{1, 2, 4, \ldots, 2^{\lceil \log(36T) \rceil}\}\right\}$.
Use $\Theta\left(\sqrt{\log(T)}\right)$ rounds of amplitude amplification to amplify the success probability of steps 1-3:
    1.) Use Setup($\mathcal{P}$) to prepare the state
            $$\sum_{t=1}^{T} \frac{1}{\sqrt{T}} |t\rangle \sum_{s \in S} \frac{1}{\sqrt{|S|}} |s\rangle |\sqrt{\pi}\rangle.$$
    2.) Perform a binary measurement $\{\Pi_M, I - \Pi_M\}$ on the last register. If the outcome is "marked", then measure in the computational basis, and output the entry in the last register. Otherwise continue with the (subnormalised) post-measurement state
            $$\sum_{t=1}^{T} \frac{1}{\sqrt{T}} |t\rangle \sum_{s \in S} \frac{1}{\sqrt{|S|}} |s\rangle |\sqrt{\pi_U}\rangle.$$
    3.) Use quantum fast-forwarding, controlled on the first two registers, to map $|t\rangle|s\rangle|\sqrt{\pi_U}\rangle$ to $|1\rangle|t\rangle|s\rangle D^t(s)|\pi_U\rangle + |0\rangle|\Gamma\rangle$ for some arbitrary $|\Gamma\rangle$, with precision $\varepsilon = O\left(\frac{1}{\log(T)}\right)$. Finally, measure the last register and output its content if marked, otherwise output Non-marked vertex.
```

Corollary 8 shows that if $T \geq 72\mathrm{HT}(\mathcal{P}, M)$, then the success probability of the above steps 1-3 is $\Omega\left(\frac{1}{\log(T)}\right)$. Therefore, after $O\left(\sqrt{\log(T)}\right)$ steps of amplitude amplification,[10] the success probability becomes $\Omega(1)$. By Theorem 1 the complexity of step 3 is $O\left(\sqrt{T \log\log(T)}(\mathsf{U} + \mathsf{C})\right)$, since $W(s)$ can be implemented in cost $O(\mathsf{U} + \mathsf{C})$ as in (5). Thus, the complexity of steps 1-3 is

$$O\left(\mathsf{S} + \sqrt{T \log\log(T)}(\mathsf{U} + \mathsf{C})\right),$$

where S is the complexity of generating $|\sqrt{\pi}\rangle$, using Setup($\mathcal{P}$). Amplitude amplification gives a $\sqrt{\log(T)}$ multiplicative overhead.

If no upper bound is known on $\mathrm{HT}(\mathcal{P}, M)$, then one can apply the exponential search algorithm of Boyer, Brassard, Høyer and Tapp [3] (see also [8, Theorem 24] with similar analysis as in our case), where we simply run Algorithm 1 with exponentially increasing guesses of an upper bound HT. This leads to the following corollary.

Corollary 4. *Let $\mathcal{P}$ be any reversible Markov chain on a finite state space X, and let $M \subset X$ be a marked set. There is a quantum algorithm that outputs a vertex x from M with bounded error in expected cost*

$$O\left(\mathsf{S} \log^{1.5}(\mathrm{HT}) + \sqrt{\mathrm{HT}}(\mathsf{U} + \mathsf{C})\sqrt{\log(\mathrm{HT}) \log\log(\mathrm{HT})}\right),$$

where $\mathrm{HT} = \mathrm{HT}(\mathcal{P}, M)$, S *is the cost of the* Setup($\mathcal{P}$) *operation,* U *is the cost of the* Update($\mathcal{P}$) *operation, and* C *is the cost of the* Check(M) *operation.*

Finally, we briefly describe how the above corollary follows from Theorem 3. The main idea is to repeatedly run Algorithm 1 until a marked vertex is found, with 4^i as our guess of an upper bound for HT in the i^{th} round of iteration for $i = 1, 2, 3$, etc. Once $i \geq j := \lceil \log_4(\mathrm{HT}) \rceil$, Theorem 3 guarantees that a marked vertex is found with probability at least $2/3$.

Let $B_i \in \Theta\left(\sqrt{i}\mathsf{S} + 2^i \sqrt{i \log(i)}(\mathsf{C} + \mathsf{U})\right)$ denote the upper bound on the cost of the i-th round given by Theorem 3; the expected cost to find a marked element using exponential search is bounded by

$$\sum_{i=1}^{j} B_i + \sum_{k=1}^{\infty} B_{j+k} \Pr((j+k)\text{-th round reached}) \leq \sum_{i=1}^{j} B_i + \sum_{k=1}^{\infty} B_{j+k} 3^{-k}.$$

By observing that $B_{j+k} \leq O\left((k+1)2^k\right) B_j$ for $k \in \mathbb{N}$, we see that the second sum is of the order B_j, and an elementary calculation shows that the first sum can be bounded by $O\left(j\sqrt{j}\mathsf{S} + B_j\right)$, proving Corollary 4.

4 ANALYSIS OF FAST-FORWARDING – THE COMBINATORIAL LEMMA

In this section, we describe the details of the analysis of Algorithm 1 needed for proving Theorem 3.

The main goal is to understand the probabilistic expression (8) in Lemma 2 that lower bounds the success probability of the fast-forwarded walk-based search algorithm. In order to lower bound the right-hand side of (8) by $\widetilde{\Omega}(1)$, we want to prove that there is some s and some random choice of $t, t' = O(\mathrm{HT})$ with $t' > t$ (in fact, t' could also be much larger than HT) such that starting in the stationary distribution and running the chain, with constant probability, the t-th vertex is marked, and the t'-th vertex is unmarked. In this section, we reduce this problem to a simple combinatorial statement, which we prove in Lemma 5.

Let $Y = (Y_i)_{i=0}^{\infty}$ be a Markov chain evolving according to $\mathcal{P}$ starting from $Y_0 \sim \pi$.[11] In order to address interpolated walks we define $Y(s) := (Y_i(s))_{i=0}^{\infty}$ to be the same chain as Y, except that for every marked vertex in Y, $Y(s)$ stays in that vertex for a length of time that is geometrically distributed with parameter $1 - s$ (mean $\frac{1}{1-s}$), before taking a step according to Y, see Figure 1. More precisely, let $k_1 < k_2 < \ldots$ be the (random) indices such that Y_{k_j} is marked, and let $L_1, L_2, \ldots$ be geometric random variables with mean $\frac{1}{1-s}$. Let $\bar{L}_j := \sum_{j'=1}^{j}(L_{j'} - 1)$, then

$$Y_i(s) := \begin{cases} Y_i & \text{if } i \in \{0, \ldots, k_1\} \\ Y_{k_j} & \text{if } i \in \{k_j + \bar{L}_{j-1}, \ldots, k_j + \bar{L}_j\} \\ Y_{i-\bar{L}_j} & \text{if } i \in \{k_j + \bar{L}_j + 1, \ldots, k_{j+1} + \bar{L}_j\}. \end{cases}$$

It is easy to see that the marginal distribution on $Y(s)$ is a Markov chain evolving according to $\mathcal{P}(s)$ starting from π.[12]

[10] In order to avoid "over-amplification", one can use a random number of amplification steps, or alternatively use a "fixed-point" version of amplitude amplification [6, 17].

[11] Since we start in the stationary distribution actually this distribution is also translationally invariant and is the same if we look forward or backward – due to reversibility. However, our Corollary 7 does not use these properties – by using these properties one might be able to prove a stronger $\Omega(1)$ lower bound for a well-chosen value of s.

[12] What we have actually described is a *coupling* of the random variables Y and $Y(s)$, such that $Y(s)$ uses the same randomness source for walking as Y, but it might delay transitions at marked vertices, as dictated by the additional independent geometric random variables. However, note that this is not the same kind of coupling that is commonly used between Markov chains, where the chains start and also *stay* together.

i	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Y_i	0	1	2	3	4	3	2	3	4	3	4	3	2	1	…						
$Y_i(s)$	0	1	2	3	4	4	4	4	3	2	3	4	4	4	3	4	4	4	3	2	1

Figure 1: Example of Y and $Y(s)$ when $\mathcal{P}$ is a walk on a line, $s = \frac{3}{4}$, and 4 is marked.

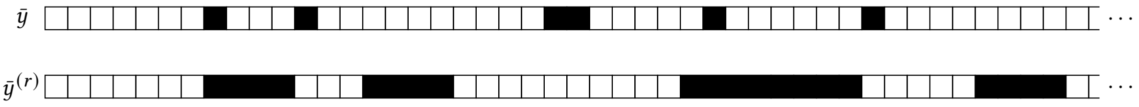

Figure 2: The first row shows a sequence $\bar{y}$ drawn from $\bar{Y}$, and the second its r-rescaling $\bar{y}^{(r)}$, for $r = 4$. The first row represents the sequence of unmarked and marked states visited by $\mathcal{P}$, and the second is an approximation of the sequence of unmarked and marked states of $\mathcal{P}(s)$ for $s = 1 - \frac{1}{r} = \frac{3}{4}$.

We are only interested in whether a state in the chain is marked or not, so we map the elements to {marked, unmarked} and denote by $\bar{Y}_i$, $\bar{Y}_i(s)$ the image of the chains after this mapping. Then we are interested in lower bounding

$$\Pr\big(\bar{Y}_0(s) = \text{unmarked}, \bar{Y}_t(s) = \text{marked}, \bar{Y}_{t'}(s) = \text{unmarked}\big). \quad (9)$$

A particular sequence $\bar{y}$ drawn from $\bar{Y}$ can then be represented visually by a sequence of boxes, each of which is either unmarked (white) or marked (black). The corresponding coupled sequence $\bar{y}(s)$ of $\bar{Y}(s)$ is essentially the same as $\bar{y}$, except that every black box is replaced with a string of black boxes, whose length is geometrically distributed with mean $r = \frac{1}{1-s}$. Thus, a good approximation of the sequence $\bar{y}(s)$ is obtained by starting with $\bar{y}$ and replacing each black box by a black box of length r, which we call an *r-rescaling* of $\bar{y}$, and denote $\bar{y}^{(r)}$, see Figure 2. Note that r need not be integral, but it is convenient and sufficient to assume that it is.

It will be sufficient to show that for some random choices $t, t' = O(\mathrm{HT})$ with $t' > t$, we have both

(m) $\bar{y}_t^{(r)}$ = marked and

(u) $\bar{y}_{t'}^{(r)}$ = unmarked,

with $\widetilde{\Omega}(1)$ probability (over $\bar{Y}$ and the random choice of t and t'), for some $r = \frac{1}{1-s}$. Let $M_{\bar{y}}^{(r)}(a, b]$ (resp. $U_{\bar{y}}^{(r)}(a, b]$) be the set of $i \in \{a+1, a+2, \dots, b\}$ such that $\bar{y}_i^{(r)}$ = marked (resp. $\bar{y}_i^{(r)}$ = unmarked). If we choose t uniformly at random from $\{a + 1, \dots, b\}$, and t' uniformly at random from $\{a' + 1, \dots, b'\}$, with $a' \geq b$, then the problem reduces to showing that for a good choice of r, with high probability over $\bar{Y}$, $|M_{\bar{y}}^{(r)}(a, b]|/(b - a)$ and $|U_{\bar{y}}^{(r)}(a', b']|/(b' - a')$ are both $\widetilde{\Omega}(1)$.

Let $T = \lceil 3\mathrm{HT} \rceil$, and suppose for the sake of this discussion that no marked vertex has a marked neighbour in $\mathcal{P}$.[13] Then for any even-length interval $\{a + 1, \dots, b\}$, the proportion of $t \in \{a + 1, \dots, b\}$ such that $\bar{y}_t$ = marked is at most $\frac{1}{2}$. As a first attempt, suppose we choose t uniformly at random from $\{1, \dots, 2T\}$ and t' uniformly at random from $\{2T + 1, \dots, 4T\}$. First note that, without any rescaling (i.e. with $r = 1$), condition (u) always holds, because $|M_{\bar{y}}^{(1)}(2T, 4T]| \leq T$. It is also easy to see that upon running the non-interpolated walk $\mathcal{P}$, with high probability there will be a marked vertex in the first subsequence of length T. Thus, if we choose $s \geq 1 - \frac{1}{T}$ so that $r \geq T$, then with high probability $|M_{\bar{y}}^{(r)}(0, 2T]| \geq T$, so condition ($m$) holds. However, after this rescaling, (u) might no longer hold. Figure 3 illustrates a $\bar{y}$, for which, before scaling, (u) holds but not (m), and after scaling by $r = T$, (m) holds but not (u).

The difficulty is that by scaling, as we create more marked boxes, we are pushing unmarked boxes out of the intervals of concern. There is a bijection between the i^{th} unmarked box in $\bar{y}$ and the i^{th} unmarked box in $\bar{y}^{(r)}$, but its index in $\bar{y}^{(r)}$ can increase. To make this precise, let $\sigma_r(i) \in \mathbb{N}$ be the position of the i^{th} unmarked box in $\bar{y}^{(r)}$. Consider $\sigma_r(i)$ as a function of r; if no marked box occurs before i, then $\sigma_r(i) \equiv i$, but otherwise it is linearly increasing in r. In particular, if $m(i)$ denotes the number of marked boxes before the i^{th} unmarked box in $\bar{y}$, then $\sigma_r(i) = i + m(i)r$. This suggests that for small enough values i, as long as $m(i) \geq 1$ — that is, there exists $j < i$ such that $\bar{y}_j$ = marked — there should be a good choice of r that pushes $\sigma_r(i)$ into the range from which we choose t'.

Our second (and final) strategy will be to choose t uniformly at random from $\{1, \dots, 3T\}$, and t' uniformly at random from $\{6T + 1, \dots, 12T\}$. We begin by scaling up by r_0, the largest scaling factor less than $3T$ such that $|M_{\bar{y}}^{(r)}(T, 3T]|/(2T) \leq \frac{3}{4}$ (for the sake of discussion, suppose it's exactly $\frac{3}{4}$). Then condition (m) holds for r_0, and this remains true even if we increase r.

It may not be the case that scaling by r_0 ensures that condition (u) holds with constant probability. However, since

$$\frac{U_{\bar{y}}^{(r)}(T, 3T]}{2T} = \frac{1}{4},$$

there are $\Theta(T)$ values i with $\sigma_{r_0}(i) = i + m(i)r_0 \in \{T + 1, \dots, 3T\}$. Increasing r will only increase the number of marked boxes (vertices) in $\{1, \dots, 3T\}$, thus increasing the probability of satisfying condition (m), but as marked boxes are being added to the window $\{1, \dots, 3T\}$, they are pushing unmarked boxes to further positions. For a high enough value of r (but not too high) we will push the i^{th} unmarked box into the window $\{6T + 1, \dots, 12T\}$. We can imagine searching for this good value r by beginning with r_0 and repeatedly doubling it, as shown in Figure 4.

We formalize this argument in the following combinatorial lemma.

[13] This could be arranged by making two copies of the graph, ensuring that each transition switches from one copy of the graph to the other, and only considering the marked vertices in one copy to be marked. However, we will ultimately not need this assumption.

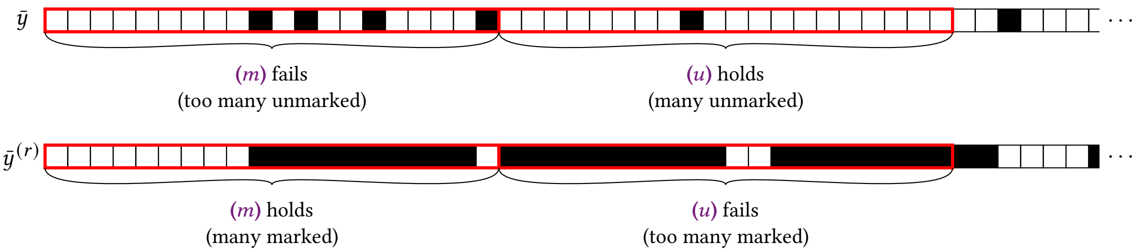

Figure 3: Illustration of the trade-off in the choice of the rescaling

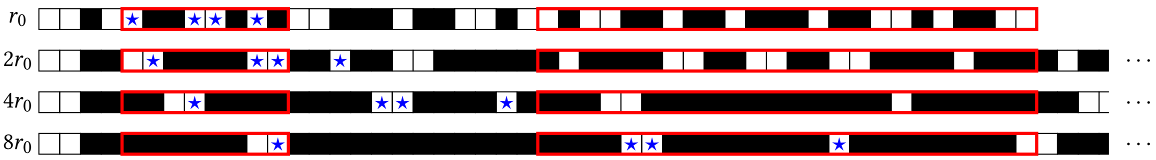

Figure 4: As we double the scaling factor, we eventually push each unmarked box (vertex) that began in the region $\{T+1,\dots,3T\}$, denoted by ★ symbols, into the region $\{6T+1,\dots,12T\}$, denoted by the right-most red rectangle. The same scaling doesn't work for every ★, but for every ★, there is some scaling that works.

Lemma 5 (Combinatorial Lemma). *Let $T \in \mathbb{N}_+$, and suppose that $y = (y_1, y_2, \dots)$ is a sequence of marked and unmarked boxes of length at least $12T$, such that*

- **(i)** *there is at least one marked among the first T boxes, and*
- **(ii)** *at most T of the boxes $(y_{T+1}, y_{T+2}, \dots, y_{3T})$ are marked.*

Let r_0 denote the largest integer such that

$$|M_y^{(r_0)}(T,3T]| < \frac{3}{2}T,$$

then $1 \le r_0 < 3T$, and for $R := \left\{1,2,\dots,2^{\lfloor \log_2(12T)\rfloor}\right\}$ we have

$$\sum_{r\in R\cap(r_0,\infty)} |U_y^{(r)}(6T,12T]| \ge \frac{1}{2}T.$$

Proof. By assumption (ii) we have $|M_y^{(1)}(T,3T]| \le T$, so $r_0 \ge 1$. By assumption (i) for any $r \ge 3T$ we have $|M_y^{(r)}(T,3T]| = 2T$, and so $r_0 < 3T$.

Similarly to the notation introduced before, let $y^{(r)}$ denote the r-rescaling of y and let $\sigma_r(i)$ denote the index of the i-th unmarked box in $y^{(r)}$. As discussed before, then $\sigma_r(i) = i + m(i)r$, where $m(i)$ denotes the number of marked boxes before the i-th unmarked box in y. To prove the second part of the lemma, we will show that

$$\begin{aligned}&\forall i\colon \sigma_{r_0}(i) \in \{T+1,\dots,3T\},\\ &\qquad \exists r \in R\colon r > r_0 \text{ and } \sigma_r(i) \in \{6T+1,\dots,12T\}.\end{aligned} \tag{10}$$

In other words, if the i-th marked box in $y^{(r_0)}$ is in the interval $\{T+1,\dots,3T\}$, then it gets shifted into the interval $\{6T+1,\dots,12T\}$ in $y^{(r)}$, for some $r \in R \cap (r_0,\infty)$. Note that when $\sigma_{r_0}(i) > T$, we must have $m(i) \ge 1$, by the assumption that at least one of the first T boxes is marked. We will show that the desired statement holds for $r = 2^k$, where $k = \lfloor \log_2 \frac{12T-i}{m(i)} \rfloor$, so clearly $k \le \lfloor \log_2(12T)\rfloor$. We indeed have $\sigma_r(i) = i + m(i)2^k \in \{6T+1,\dots,12T\}$, since

$$i + m(i)2^k \le i + m(i)\frac{12T-i}{m(i)} = 12T$$

and

$$i + m(i)2^k \ge i + m(i)\frac{12T-i}{2m(i)} > 6T.$$

To finish the proof of (10), note that since $3T \ge \sigma_{r_0}(i) = i + m(i)r_0$, we have $r_0 \le (3T-i)/m(i)$, therefore $r = 2^k > \frac{6T-i}{m(i)} > 2r_0$.

The second claim in the lemma follows from (10), because

$$|\{i : \sigma_{r_0}(i) \in \{T+1,\dots,3T\}\}| = |U_y^{(r_0)}(T,3T]| \ge \frac{1}{2}T$$

by the definition of r_0. By (10), each of these $\ge \frac{1}{2}T$ unmarked vertices contributes to at least one term of $\sum_{r\in R\cap(r_0,\infty)} |U_y^{(r)}(6T,12T]|$. □

Even if we replace the fixed rescalings of each marked element in $\bar{y}^{(r)}$ with independent geometric random variables, any fixed set of marked elements gets a total rescaling that is within a factor 2 of its expected length with probability at least $\frac{7}{16}$. This fact is formalised in the following lemma, proven in Appendix B:

Lemma 6. *Let $p \in (0,1]$, $t \in \mathbb{N}$ and $Z = \sum_{i=1}^{t} G_i$, where G_i are independent geometric random variables with parameter p. Then*

$$\Pr\left(\frac{t}{2p} \le Z \le \frac{2t}{p}\right) \ge \frac{7}{16}. \tag{11}$$

We can now conclude with a statement about the random walk $\mathcal{P}(s)$ that we will use to analyze our quantum algorithm. The final statement we need is proven in Corollary 8. We first prove the following corollary.

Corollary 7. *Let $\mathcal{P}$ be a (not necessarily reversible) Markov chain. Let ρ be any distribution (not necessarily stationary). Let E be the event that: the first vertex sampled according to ρ is unmarked; a marked vertex is encountered within the first T steps of $\mathcal{P}$ (equivalently $\mathcal{P}(s)$); and at most T of the next $2T$ steps of $\mathcal{P}$ (equivalently, the next $2T$ steps of $\mathcal{P}(s)$ that do not consist of staying at a marked vertex) go to a marked vertex.*

Let $r \in R := \left\{1, 2, 4, \dots, 2^{\lfloor \log_2(12T) \rfloor}\right\}$, $t \in \{1, \dots, 3T\}$ and $t' \in \{3T+1, \dots, 24T\}$ be chosen uniformly at random, then for $s = 1 - \frac{1}{r}$:

$$\mathbb{E}_{t,t',r}\left[\Pr_{Y_0(s)\sim\rho}(Y_0(s) \in U, Y_t(s) \in M, Y_{t'}(s) \in U|E)\right] = \Omega\left(\frac{1}{\log(T)}\right).$$

Proof. When sampling $Y(s)$, we distinguish between:

(i) the randomness used, when at a marked vertex, to decide whether to *skip* or take a step of the walk according to $\mathcal{P}$, and

(ii) the randomness used for choosing a neighbouring vertex to *transition* to (assuming a step is to be taken), according to $\mathcal{P}$.

The second type of randomness, (ii), is exactly the randomness of Y (recall that Y is a Markov chain that is coupled to $Y(s)$ in the sense that if $Y(s)$ does not stay at the current vertex, then it moves as Y).

We can assume without loss of generality that the Markov chain Y is terminated after $24T$ steps. This makes the treatment conceptually simpler, for example we can simply treat E as a finite set of length-$24T$ paths, and therefore we can write

$$\mathbb{E}_{t,t',r}\left[\Pr_{Y_0(s)\sim\rho}(Y_0(s) \in U, Y_t(s) \in M, Y_{t'}(s) \in U|E)\right] = \tag{12}$$
$$= \sum_{y\in E} \Pr(Y = y|E) \sum_{r\in R} \frac{1}{|R|} \sum_{y(s)} \Pr(Y(s) = y(s)|Y = y)pq,$$

where

$$p := \frac{|\{t \in \{1, \dots, 3T\} \colon y_t(s) \in M\}|}{3T},$$
$$q := \frac{|\{t' \in \{3T+1, \dots, 24T\} \colon y_{t'}(s) \in U\}|}{21T}.$$

Let us study a fixed path $y \in E$, and a fixed $r \in R$, i.e., $s = \frac{1}{1-r}$. We will examine the corresponding coupled paths $y(s) \in Y(s)$. For[14] $B \in [24T]$ let ${}_By := (y_1, y_2, \dots, y_k)$ be the shortest truncation of y such that the r-rescaling of ${}_By$ has length at least B, and let ${}_By(s)$ be the sequence where we apply the random geometric rescalings of $y(s)$ to the marked elements of ${}_By$. Let $\bar{r}(B)$ be the average rescaling applied in ${}_By(s)$. Let ℓ be the length of ${}_{3T}y(s)$; if $\bar{r}(3T) \geq r/2$, then

$$3Tp \geq |\{t \in [\min(\ell, 3T)] \colon {}_{3T}y_t(s) \in M\}| \geq \frac{1}{2}|M_y^{(r)}(0, 3T]|, \tag{13}$$

since $|\{t \in [\min(\ell, 3T)] \colon {}_{3T}y_t(s) \in U\}| \leq \frac{1}{2}|U_y^{(r)}(0, 3T]|$.

Similarly, if $\bar{r}(6T) \geq r/2$, and $\bar{r}(12T) \leq 2r$. Then the unmarked vertices of $y^{(r)}$ in $\{6T+1, 6T+2, \dots, 12T\}$ may be moved and spread out in $y(s)$, but they will all occur within the range $\{3T+1, \dots, 24T\}$:

$$|\{t' \in \{3T+1, \dots, 24T\} : y_{t'}(s) \in U\}| \geq |U_y^{(r)}(6T, 12T]|. \tag{14}$$

Let F be the event that $\bar{r}(3T) \geq r/2$, $\bar{r}(6T) \geq r/2$, and $\bar{r}(12T) \leq 2r$. Let us partition ${}_{12T}y$ to three parts according to the subsequences ${}_{3T}y$ and ${}_{6T}y$; if in all three parts the corresponding average rescalings of the marked elements of $y(s)$ are within $[\frac{r}{2}, 2r]$, then F holds. Since the geometric variables of the three parts are independent, by Lemma 6 we always get $\Pr(F) \geq (7/16)^3$. Thus, continuing from (12), we have:

$$\sum_{y\in E} \Pr(Y = y|E) \sum_{r\in R} \frac{1}{|R|} \sum_{y(s)} \Pr(Y(s) = y(s)|Y = y)pq$$
$$\geq \sum_{y\in E} \Pr(Y = y|E) \sum_{r\in R} \frac{1}{|R|} \Pr(F|Y = y) \frac{|M_y^{(r)}(0, 3T]|}{6T} \frac{|U_y^{(r)}(6T, 12T]|}{21T}$$

by Eqs. (13)-(14)

$$\geq \frac{1}{|R|}\left(\frac{7}{16}\right)^3 \sum_{y\in E} \Pr(Y = y|E) \sum_{r\in R} \frac{|M_y^{(r)}(0, 3T]|}{6T} \frac{|U_y^{(r)}(6T, 12T]|}{21T}$$

by Lemma 6

$$\geq \frac{1}{|R|}\left(\frac{7}{16}\right)^3 \sum_{y\in E} \Pr(Y = y|E) \frac{1}{4}\frac{1}{42}, \qquad \text{by Lemma 5}$$

$$= \Omega\left(\frac{1}{|R|}\right) = \Omega\left(\frac{1}{\log T}\right).$$

□

We can now conclude with the statement we needed in the analysis of our algorithm in Section 3.

Corollary 8. *Let $\mathcal{P}$ be a reversible ergodic Markov chain, and let π be its stationary distribution. If $p_M \leq 1/9$, $T \geq 3\mathrm{HT}$, and R is as defined in Corollary 7, then choosing $s \in S = \{1 - \frac{1}{r} : r \in R\}$ and $t \in [24T]$ uniformly at random, we get*

$$\mathop{\mathbb{E}}_{s,t}\left[\left\|\Pi_M D^t(s)|\sqrt{\pi_U}\rangle\right\|^2\right] = \Omega\left(\frac{1}{\log(T)}\right).$$

Proof. First we prove that the event E in Corollary 7 holds with constant probability. The probability that the initial vertex is marked is $p_M \leq 1/9$. The probability that the Markov chain does not hit a marked vertex in $T \geq 3\mathrm{HT}$ steps is at most $1/3$ by Markov's inequality. Finally, the expected number of marked sites in the first $3T$ steps is $p_M 3T \leq T/3$, therefore the probability that there are more than T marked vertices in the first $3T$ steps is at most $1/3$ by Markov's inequality. By the union bound we get the probability of the complement of E is at most $1/9 + 1/3 + 1/3 = 7/9$, therefore E holds with probability at least $2/9$.

Let us define $|v^t(s)\rangle := \Pi_M D^t(s)|\sqrt{\pi_U}\rangle$, and recall that $\hat{t} = t' - t$. Since $\Pr(E) \geq \frac{2}{9}$, by Corollary 7 we have

$$\Omega(1) = \sum_{s\in S} \sum_{t,\hat{t}\in[24T]} \sum_{x,z\in U} \frac{\pi_x \langle x|\mathcal{P}^t(s)\Pi_M\mathcal{P}^{\hat{t}}(s)|z\rangle}{(24T)^2} \qquad \text{by Corollary 7}$$
$$= \sum_{s\in S} \sum_{t,\hat{t}\in[24T]} \frac{\langle\sqrt{\pi_U}|D^t(s)\Pi_M D^{\hat{t}}(s)|\sqrt{\pi_U}\rangle}{(24T)^2} \qquad \text{by Eq. (7)}$$
$$= \sum_{s\in S} \sum_{t,\hat{t}\in[24T]} \frac{\langle v^t(s)|v^{\hat{t}}(s)\rangle}{(24T)^2} \leq \sum_{s\in S} \sum_{t,\hat{t}\in[24T]} \frac{\left\|v^t(s)\right\|\left\|v^{\hat{t}}(s)\right\|}{(24T)^2}$$

by Cauchy-Schwartz

$$= \sum_{s\in S}\left(\sum_{t\in[24T]} \frac{\left\|v^t(s)\right\|}{24T}\right)^2 \leq \sum_{s\in S} \sum_{t\in[24T]} \frac{\left\|v^t(s)\right\|^2}{24T},$$

where the last inequality follows from the fact that the arithmetic mean is always upper-bounded by the root-mean square. □

[14] For $n \in \mathbb{N}$ we use the notation $[n]$ as a shorthand for $\{1, 2, \dots, n\}$.

5 EXAMPLE WITH $\mathrm{HT}^+ \gg \mathrm{HT}$

A torus is a graph containing $n = N^2$ vertices organized in N rows and N columns; there is a vertex (x_1, x_2) for all $x_1, x_2 \in \{0, 1, \ldots, N-1\}$. A vertex (x_1, x_2) has four neighbours, (x_1+1, x_2), (x_1-1, x_2), (x_1, x_2+1) and (x_1, x_2-1), where the addition is modulo N. To prevent the graph from being bipartite, we add a self-loop at each vertex, so that at any vertex the random walker moves to any of the four neighbours with probability 0.2 and stays at the same vertex also with probability 0.2.

We start by observing that the extended hitting time HT^+ in the case of a torus can be lower bounded as follows.

Lemma 9. *Let $M \subset \{0, 1, \ldots, N-1\}^2$ be a set of marked vertices of the $N \times N$ torus. Let $m = |M|$ and $\omega = \exp(2\pi \mathrm{i}/N)$, then*

$$\mathrm{HT}^+ \geq \frac{5}{4\pi^2}\frac{N^2}{m^2}\left|\sum_{(x_1,x_2)\in M}\omega^{x_1}\right|^2. \tag{15}$$

The proof is deferred to Appendix A.

Next we describe an example of a marked set whose extended hitting time can be much larger than the hitting time.

Lemma 10. *Suppose that positive integers d_1, k_1, d, N satisfy the following requirements:*

(C1) $k_1 d_1 = o(N)$;
(C2) $N = o(k_1 d)$;
(C3) $d^2 \log d = o(N^2)$;
(C4) d_1 is a divisor of d and d is a divisor of N.

Define a marked set M on the $N \times N$ torus as $M_1 \cup M_2$, where

$$M_1 = \{(j_1 d_1, j_2 d_1) \mid 0 \leq j_1, j_2 \leq k_1 - 1\}$$

and

$$M_2 = \{(j_1 d, j_2 d) \mid 0 \leq j_1, j_2 < N/d\}.$$

Then the extended and classical hitting times for the set M satisfy

$$\mathrm{HT}^+ = \Omega(N^2) \quad \textit{and} \quad \mathrm{HT} = O\left(d^2 \log d\right) = o(\mathrm{HT}^+),$$

respectively.

In Figure 5 an illustration with $d_1 = 1$, $k_1 = 15$, $d = 6$ and $N = 36$ is depicted, with different colours for $M_1 \setminus M_2$, $M_2 \setminus M_1$ and $M_1 \cap M_2$.

An example of parameters satisfying (C1)-(C4) is $d_1 = 1$, $k_1 = a\,2^{a^2}$, $d = a^2$ and $N = a^2\,2^{a^2}$, for an integer $a > 1$. For such parameters Lemma 10 implies bounds $\mathrm{HT} = O\left(\log^2 N \log\log N\right)$ and $\mathrm{HT}^+ = \Omega(N^2)$, thus there is a $\widetilde{\Omega}\left(N^2\right)$ gap between the extended hitting time HT^+ and the classical hitting time HT.

Proof of Lemma 10. We begin by noting that the sets M_2 and M_1 overlap, since $d_1 | d$ by (C4). The set M consists of k_1^2 vertices forming a small, dense subgrid M_1, and the remaining marked vertices of M_2 forming a sparser subgrid in the rest of the torus.

Since $m = |M| \leq |M_1| + |M_2| = k_1^2 + (N/d)^2$, the constraint (C2) implies $m = O\left(k_1^2\right)$; from (C1) we conclude $m = o(N^2)$. Also by (15)

$$\mathrm{HT}^+ = \Omega\left(N^2 |\rho|^2 / m^2\right), \tag{16}$$

where ρ is defined by

$$\rho = \sum_{x\in M}\omega^{x_1} = \sum_{x\in M_1}\omega^{x_1} + \sum_{x\in M_2}\omega^{x_1} - \sum_{x\in M_1\cap M_2}\omega^{x_1}.$$

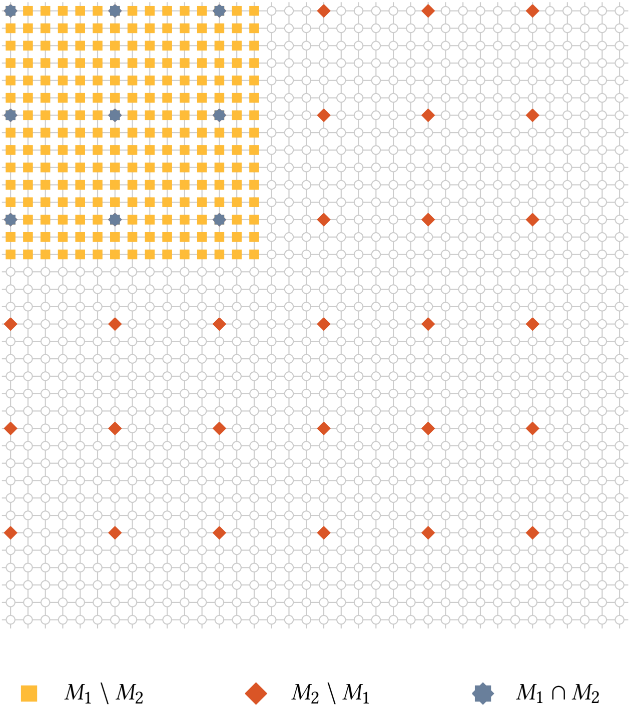

Figure 5: Illustration of the marked set with $d_1 = 1$, $k_1 = 15$, $d = 6$ and $N = 36$.

The first summand on the RHS is

$$\sum_{x\in M_1}\omega^{x_1} = k_1 \sum_{j=0}^{k_1-1}\omega^{j d_1} = k_1 \frac{\omega^{k_1 d_1} - 1}{\omega^{d_1} - 1},$$

while the second summand is a multiple of

$$\sum_{j=0}^{N/d-1}\omega^{jd} = \left(\omega^N - 1\right)/\left(\omega^d - 1\right) = 0$$

because $d | N$ by (C4). Therefore

$$\rho = k_1 \frac{\omega^{k_1 d_1} - 1}{\omega^{d_1} - 1} - \sum_{x\in M_1\cap M_2}\omega^{x_1}.$$

It is easy to see that $M_1 \cap M_2 = \{(j_1 d, j_2 d) \mid 0 \leq j_1, j_2 < k\}$, where $k = \lceil k_1 d_1 / d \rceil$, and similar arguments as previously yield

$$\rho = k_1 \frac{\omega^{k_1 d_1} - 1}{\omega^{d_1} - 1} - k \frac{\omega^{kd} - 1}{\omega^d - 1}.$$

By the reverse triangle inequality,

$$|\rho| \geq k_1 \frac{\left|\omega^{k_1 d_1} - 1\right|}{\left|\omega^{d_1} - 1\right|} - k \frac{\left|\omega^{kd} - 1\right|}{\left|\omega^d - 1\right|} = \frac{k_1 \sin\frac{\pi k_1 d_1}{N}}{\sin\frac{\pi d_1}{N}} - \frac{k \sin\frac{\pi k d}{N}}{\sin\frac{\pi d}{N}}. \tag{17}$$

From (C1) and (C3) we obtain $kd \le k_1 d_1 + d = o(N)$, therefore $\frac{k_1 d_1}{N} = o(1)$, $\frac{kd}{N} = o(1)$ and $\sin\frac{\pi k_1 d_1}{N} = \Theta\left(\frac{k_1 d_1}{N}\right)$, $\sin\frac{\pi kd}{N} = \Theta\left(\frac{kd}{N}\right)$, $\sin\frac{\pi d_1}{N} = \Theta\left(\frac{d_1}{N}\right)$, $\sin\frac{\pi d}{N} = \Theta\left(\frac{d}{N}\right)$. Consequently,

$$\frac{k_1 \sin\frac{\pi k_1 d_1}{N}}{\sin\frac{\pi d_1}{N}} = \Theta(k_1^2), \qquad \frac{k \sin\frac{\pi kd}{N}}{\sin\frac{\pi d}{N}} = \Theta(k^2) = \Theta\left(k_1^2 \frac{d_1^2}{d^2}\right) = o\left(k_1^2\right);$$

here the last bound follows from $d_1 = o(d)$, which is implied by (C1) and (C2).

Now (17) gives $|\rho| = \Omega(k_1^2)$. Combining this with (16) and the previously obtained bound $m = O\left(k_1^2\right)$, we conclude that the extended hitting time satisfies

$$\mathrm{HT}^+ = \Omega\left(\frac{N^2|\rho|^2}{m^2}\right) = \Omega\left(\frac{N^2 k_1^4}{k_1^4}\right) = \Omega(N^2). \tag{18}$$

Next we bound HT. Notice that by the linearity of expectation $\mathrm{HT} = \sum_{x\in U} \pi_x \mathrm{HT}_x(M)$ and $\mathrm{HT}_x(M)$ is the expected number of steps for the random walker to reach M for the first time, starting from vertex x. It follows that $\mathrm{HT} \le \max_{x\in U} \mathrm{HT}_x(M)$. For any fixed $x \in U$, $\mathrm{HT}_x(M)$ cannot decrease when reducing the marked set (i.e., when some marked vertices are removed from M and added to the unmarked set U), hence we have

$$\mathrm{HT} \le \max_{x\in U} \mathrm{HT}_x(M) \le \max_{x\notin M_2} \mathrm{HT}_x(M_2).$$

Therefore it suffices to show that $\mathrm{HT}_x(M_2) = O\left(d^2 \log d\right)$ when only the subgrid M_2 is marked and x is any vertex not belonging to M_2. However, the classical random walk with the marked set M_2 is equivalent to the random walk in the $d \times d$ torus with a single marked element (by identifying each vertex (x_1, x_2) with the unique vertex $(x_1^{(0)}, x_2^{(0)})$ satisfying $x_1 \equiv x_1^{(0)} \pmod{d}$, $x_2 \equiv x_2^{(0)} \pmod{d}$ and $x_1^{(0)}, x_2^{(0)} \in \{0, 1, \ldots, d-1\}$). Since, in the case of a $d \times d$ torus with a single marked element, all hitting times HT_y (with y being a non-marked vertex) are of order $O\left(d^2 \log d\right)$ [9, Eq. 10.29], the desired bound $\mathrm{HT}_x(M_2) = O\left(d^2 \log d\right)$ follows. Hence, returning to the marked set M, the classical hitting time is $\mathrm{HT} = O\left(d^2 \log d\right) = o\left(N^2\right)$ by (C3), and we conclude that $\mathrm{HT} = o\left(\mathrm{HT}^+\right)$. □

An intuitive explanation for this result is that the algorithm of Krovi et al. [8] actually solves a more difficult problem: it generates the uniform superposition over $|x\rangle, x \in M$ (with the starting state being the uniform superposition over all vertices $|x\rangle$). Almost all of the marked vertices are, however, concentrated in M_1 which is a small part of the grid. A typical component of the starting state is at a distance $\Omega(N)$ from M_1. Therefore, any algorithm that generates the uniform superposition over $|x\rangle, x \in M$ from this starting state must take $\Omega(N)$ steps, even though the classical hitting HT time is much smaller.

The running time $O(\sqrt{\mathrm{HT}^+}) = O(N\sqrt{\log N})$ achieved by the algorithm of [8] is quite close to the $\Omega(N)$ lower bound. So, in our example, this algorithm is close to being optimal for generating the uniform superposition of marked vertices but is very far from being optimal for the task of simply finding a marked vertex.

6 A SIMPLE QUANTUM WALK ALGORITHM

As discussed in Section 2.2, our quantum walk uses three registers. Register R_2 corresponds to a Hilbert space $\mathcal{H}$ containing the basis states $|x\rangle$ identified with the vertices of the graph. Register R_1 is an ancillary register initialized to the reference state $|\bar{0}\rangle$. Additionally, another ancilla register R_3 initialized to $|0\rangle \in \mathbb{C}^2$, will be attached and used for checking whether the current vertex in R_2 is marked.

Now we describe a quantum walk algorithm with a fixed interpolation parameter $s \in [0, 1)$ and a predetermined number of quantum walk steps $t \in \mathbb{N}$.

Algorithm 2 Quantum walk algorithm

```
Input: Oracles for the Markov chain $\mathcal{P}$ and the marked set $M$,
interpolation parameter $s$, and the number of iterations $t$
  1.) Prepare the state $|\bar{0}\rangle|\sqrt{\pi}\rangle$ with Setup($\mathcal{P}$).
  2.) Apply $t$ times the operator $W(s)$ on $\mathsf{R}_1\mathsf{R}_2$.
  3.) Attach $\mathsf{R}_3$, apply Check($M$) on $\mathsf{R}_2\mathsf{R}_3$, measure $\mathsf{R}_3$.
  4.) If $\mathsf{R}_3 = 1$, then measure $\mathsf{R}_2$ in the vertex basis, output the
      outcome. Otherwise, output No marked vertex found.
```

It is obvious that the complexity of the algorithm is of the order $\mathsf{S} + t \cdot (\mathsf{U} + \mathsf{C})$. We conjecture that (under the assumption that the probability to draw a marked vertex from the stationary distribution is at most 0.5) there always exists an interpolation parameter s such that Algorithm 2 finds a marked vertex with high probability in $t = O\left(\sqrt{\mathrm{HT}}\right)$ steps:

Conjecture 11. *Let $\mathcal{P}$ be a reversible, ergodic Markov chain with stationary distribution π; suppose that M is a set of marked states which satisfies $p_M = \sum_{x\in M} \pi_x < 0.5$. Then there exists a value $s \in [0, 1)$ and a positive integer $t = O\left(\sqrt{\mathrm{HT}}\right)$ such that Algorithm 2 succeeds with probability $\Omega(1)$.*

The success probability can be lower-bounded by a quantity expressible in terms of the discriminant matrix $D(s)$. Let

$$p_{\text{success}} = \left\|(I \otimes \Pi_M) W^t(s) |\bar{0}\rangle|\sqrt{\pi}\rangle\right\|^2$$

be the probability of obtaining a marked vertex in the last step of Algorithm 2. This can be lower-bounded by

$$\left\|(I \otimes \Pi_M)\Pi_0 W^t(s) |\bar{0}\rangle|\sqrt{\pi}\rangle\right\|^2 =: q_t(s), \tag{19}$$

where $\Pi_0 := |\bar{0}\rangle\langle\bar{0}| \otimes I$. The following lemma[15] implies that $q_t(s) = \left\|\Pi_M D_t(s)|\sqrt{\pi}\rangle\right\|^2$, where $D_t(s) = T_t(D(s))$ for T_t the Chebyshev polynomial of the first kind of degree t.

Lemma 12. *The quantum walk operator $W^t(s)$, when restricted to $|\bar{0}\rangle$ in the first register, acts as the t-th Chebyshev polynomial of the first kind applied to the discriminant matrix $D(s)$, i.e.,*

$$\Pi_0 W^t(s) \Pi_0 = |\bar{0}\rangle\langle\bar{0}| \otimes D_t(s),$$

where $D_t(s) = T_t(D(s))$ and T_t is the Chebyshev polynomial of the first kind of degree t, applied (in the matrix function sense) to the matrix $D(s)$. Equivalently, $D_t(s)$ can be defined via the recurrence relations

$$D_0(s) = I, \quad D_1(s) = D(s), \tag{20}$$

$$D_{t+1}(s) = 2D_t(s) \cdot D(s) - D_{t-1}(s), \quad t \in \mathbb{N}. \tag{21}$$

[15] For a generalization of this claim see [6, Lemma 9 & Theorem 17].

Proof. Recall that $W(s) = \widetilde{W}(s) \cdot (2\Pi_0 - I \otimes I)$ where $\widetilde{W}(s) = V^{\dagger}(\mathcal{P}, s)\,\mathrm{SWAP}'\,V(\mathcal{P}, s)$. Moreover, the idempotence of Π_0 gives

$$W(s)\Pi_0 = \widetilde{W}(s) \cdot (2\Pi_0 - I \otimes I)\Pi_0 = \widetilde{W}(s)\Pi_0. \tag{22}$$

For the proof by induction on t, notice that the claim trivially holds for $t = 0$. When $t = 1$, the statement (due to (22)) is equivalent to Eq. (6). Suppose that the claim has been proven for all integers between 1 and t, and consider $\Pi_0 W^{t+1}(s)\Pi_0$. We have

$$\begin{aligned}
\Pi_0 W^{t+1}(s)\Pi_0 &= \Pi_0 W^{t-1}(s) \cdot \Big(\widetilde{W}(s) \cdot (2\Pi_0 - I \otimes I)\Big) W(s)\Pi_0 \\
&= 2\Pi_0 W^{t-1}(s)\widetilde{W}(s)\Pi_0 W(s)\Pi_0 - \Pi_0 W^{t-1}(s)\widetilde{W}(s)W(s)\Pi_0 \\
&= 2\Pi_0 W^{t-1}(s)W(s)\Pi_0 W(s)\Pi_0 - \Pi_0 W^{t-1}(s)\widetilde{W}(s)\widetilde{W}(s)\Pi_0 \\
&\qquad\qquad \text{(by Eq. (22))} \\
&= 2\Pi_0 W^{t}(s)\Pi_0 \cdot \Pi_0 W(s)\Pi_0 - \Pi_0 W^{t-1}(s)\Pi_0. \\
&\qquad\qquad \text{(since } \widetilde{W}^2(s) = I \text{ and } \Pi_0^2 = \Pi_0\text{)}
\end{aligned}$$

By the inductive hypothesis, the obtained quantity equals $|\bar{0}\rangle\langle\bar{0}| \otimes (2D_t(s) \cdot D_1(s) - D_{t-1}(s))$. We conclude that indeed

$$\Pi_0 W^{t+1}(s)\Pi_0 = |\bar{0}\rangle\langle\bar{0}| \otimes D_{t+1}(s),$$

where $D_{t+1}(s)$ is defined by the recurrence relations (20)-(21). We finish by noting that these recurrence relations define the Chebyshev polynomials of the first kind. □

In the following we describe some examples illustrating the dependence of $q_t(s)$ on the interpolation parameter s.

Example 6.1. Consider the example described in Section 5, with parameter $a = 3$ (i.e., $d_1 = 1$, $k_1 = 1536$, $d = 9$, and $N = 4608$). It can be calculated that the classical hitting time of the marked set is $\mathrm{HT} = 162.98\ldots$, whereas the extended hitting time is $\mathrm{HT}^+ = 1.01\ldots \cdot 10^7$ (the lower bound in Lemma 10 gives $\mathrm{HT}^+ \geq 1.69 \cdot 10^6$, by (15) and (17)).

In Figure 6, we plot the lower bound (19) on the success probability of Algorithm 2. As we will also see in Section 4, it is natural to replace the interpolation parameter $s \in [0, 1)$ with $r = 1/(1 - s) \in [1, \infty)$. (The parameter r is also equal to the expected number of steps until the interpolated walk makes a transition according to the original random walk at a marked vertex.)

Figure 6 shows two quantities (as functions of r):

- the maximum of the bound (19) over $t \leq \lceil 3\sqrt{\mathrm{HT}}\rceil$, denoted $q(r)$ (units on the left axis);
- the minimal value of t which achieves $q(r)$, denoted by
$$\tau(r) := \min\left\{ t \geq 0 \,\middle|\, q_t\left(1 - \frac{1}{r}\right) = q(r) \right\}$$
(with units on the right axis; represented in $\sqrt{\mathrm{HT}}$ units).

Furthermore, we indicate parameter values $r_1 = \frac{1-p_M}{p_M}$ (which corresponds to the value of s used in [8] for their $\Theta(\sqrt{\mathrm{HT}^+})$-time algorithm) and $r_2 = \mathrm{HT}$ (a plausible upper bound on the optimal r) by vertical dash-dotted and dashed lines, respectively.

From Figure 6 it can be noticed that the optimal value is $r = 96.61\ldots \approx d^2$ and it allows Algorithm 2 to find a marked vertex in $t = 21 \approx 1.65\sqrt{\mathrm{HT}}$ steps with probability exceeding 0.98. This value of r is substantially bigger than the value $r_1 \approx 7.191$ corresponding to the algorithm of [8].

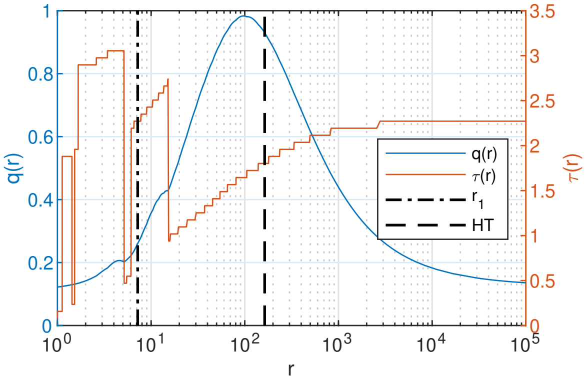

Figure 6: Bounds on Algorithm 2 in Example 6.1. The horizontal axis (r) represents the interpolation parameter $s = 1 - \frac{1}{r}$; $\tau(r)$ denotes the best choice of time t and $q(r)$ denotes the best lower bound on the success probability of Algorithm 2, as described below.

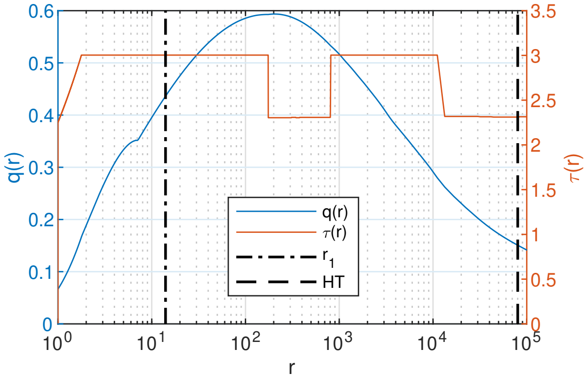

Figure 7: Some properties of the family of interpolated quantum walks in Example 6.2. For notation and explanation of the plotted quantities see Figure 6.

Example 6.2. Let G_k be the graph consisting of a single central node x_0 and k paths of length k^2; all paths have a common endpoint x_0 and the remaining vertices are distinct (i.e., G_k is a modified version of the star graph with k rays of length k^2). In each vertex the random walker stays in the same vertex with probability 0.5 and with probability 0.5 moves to a neighbour vertex (in case of several neighbours, the probability 0.5 splits evenly among them to move to a particular neighbour). Let M be one of the k paths, not including the central node.

When $k = 15$, the classical hitting time is $\mathrm{HT} = 80090.95\ldots$, whereas the extended hitting time is $\mathrm{HT}^+ = 1016848.98\ldots$. As previously, we change variables $r = 1/(1-s)$ and plot $q(r)$ and $\tau(r)$ on the left and right axis of Figure 7, respectively. Again, values $r_1 = \frac{1-p_M}{p_M}$ and $r_2 = \mathrm{HT}$ are indicated by vertical lines. As indicated by Figure 7, at $r \approx k^2$ Algorithm 2 finds a marked vertex with probability at least 0.59 in less than $2.31\sqrt{\mathrm{HT}}$ steps.

A PROOF OF LEMMA 9

Lemma 9. *Let $M \subset \{0, 1, \ldots, N-1\}^2$ be a set of marked vertices of the $N \times N$ torus. Let $m = |M|$ and $\omega = \exp(2\pi \mathrm{i}/N)$, then*

$$\mathrm{HT}^+ \geq \frac{5}{4\pi^2}\frac{N^2}{m^2}\left|\sum_{(x_1,x_2)\in M}\omega^{x_1}\right|^2. \tag{15}$$

Proof. While the vertices (x_1, x_2) of the torus graph can be ordered arbitrarily, we use the lexicographic ordering (i.e., $(x_1, x_2) \prec (x_1', x_2')$ iff $x_1 < x_1'$ or $x_1 = x_1'$ and $x_2 < x_2'$), Then $\mathcal{P}$ is formed accordingly to this ordering, i.e., the first row (column) of $\mathcal{P}$ corresponds to the vertex $(0, 0)$, the second row (column) corresponds to the vertex $(0, 1)$, and so on. Now $\mathcal{P}$ is an $(N^2) \times (N^2)$ BCCB (block circulant with circulant blocks) matrix [16, Definition 5.27] and can be diagonalized using the discrete Fourier transform as [16, Proposition 5.31]

$$\mathcal{P} = (F_N \otimes F_N)\mathrm{diag}(\Lambda)(F_N \otimes F_N)^\dagger,$$

where Λ is the vector of the eigenvalues of $\mathcal{P}$, $\otimes$ stands for the Kronecker product and

$$F_N = \frac{1}{\sqrt{N}}\begin{pmatrix} 1 & 1 & 1 & \ldots & 1\\ 1 & \omega & \omega^2 & \ldots & \omega^{N-1}\\ 1 & \omega^2 & \omega^4 & \ldots & \omega^{2(N-1)}\\ & & \ddots & & \\ 1 & \omega^{N-1} & \omega^{2(N-1)} & \ldots & \omega^{(N-1)(N-1)}\end{pmatrix},$$

$\omega := \exp\left(\frac{2\pi i}{N}\right)$. It can be verified by direct calculation (or by applying the two-dimensional discrete Fourier transform as described in [16, Proposition 5.31]) that the eigenvalues of the matrix $\mathcal{P}$ are

$$\lambda_{j,k} = \frac{1}{5}\left(1 + 2\cos\frac{2\pi j}{N} + 2\cos\frac{2\pi k}{N}\right), \quad j, k \in \{0, 1, \ldots, N-1\},$$

and the corresponding eigenvectors are $|v_{j,k}\rangle = w^{(j)} \otimes w^{(k)}$,

$$w^{(j)} := \frac{1}{\sqrt{N}}\begin{pmatrix} 1 & \omega^j & \omega^{2j} & \ldots & \omega^{(N-1)j}\end{pmatrix}^T.$$

By [8, Theorem 17], the extended hitting time is related to the interpolated hitting time $\mathrm{HT}(0)$ via $\mathrm{HT}^+ = p_M^{-2}\,\mathrm{HT}(0)$, where $\mathrm{HT}(0)$ is defined as

$$\mathrm{HT}(0) = \sum_{\substack{j=0..N-1\\ k=0..N-1\\ (j,k)\neq(0,0)}} \frac{\left|\langle v_{j,k}|\sqrt{\pi_U}\rangle\right|^2}{1-\lambda_{j,k}}.$$

Since the stationary distribution π is uniform, we have $|\pi_U\rangle = \frac{1}{N}\sum_{x\in U}|x\rangle$; moreover,

$$1 - \lambda_{j,k} = \frac{2 - 2\cos\frac{2\pi j}{N} + 2 - 2\cos\frac{2\pi k}{N}}{5} = \frac{4}{5}\left(\sin^2\frac{\pi j}{N} + \sin^2\frac{\pi k}{N}\right)$$

and

$$\sum_{x\in U}\langle x|v_{j,k}\rangle = \frac{1}{N}\sum_{(x_1,x_2)\in U}\omega^{jx_1+kx_2},$$

thus we arrive at

$$\mathrm{HT}^+ = \frac{5}{4}\frac{1}{m^2}\sum_{\substack{j=0..N-1\\ k=0..N-1\\ (j,k)\neq(0,0)}}\frac{\left|\sum_{(x_1,x_2)\in U}\omega^{jx_1+kx_2}\right|^2}{\sin^2\frac{\pi j}{N} + \sin^2\frac{\pi k}{N}}. \tag{23}$$

For all pairs $(j, k) \neq (0, 0)$, $0 \leq j, k \leq N-1$, we have

$$\sum_{x_1=0}^{N-1}\sum_{x_2=0}^{N-1}\omega^{jx_1+kx_2} = \sum_{(x_1,x_2)\in M}\omega^{jx_1+kx_2} + \sum_{(x_1,x_2)\in U}\omega^{jx_1+kx_2} = 0,$$

hence we can rewrite (23) as

$$\mathrm{HT}^+ = \frac{5}{4m^2}\sum_{\substack{j=0..N-1\\ k=0..N-1\\ (j,k)\neq(0,0)}}\frac{\left|\sum_{(x_1,x_2)\in M}\omega^{jx_1+kx_2}\right|^2}{\sin^2\frac{\pi j}{N} + \sin^2\frac{\pi k}{N}}. \tag{24}$$

Finally, we lower bound the RHS of (24) by a single term ($j = 1, k = 0$) of the sum, and use that $\sin\frac{\pi}{N} \leq \frac{\pi}{N}$, yielding (15). □

B CONCENTRATION OF SUMS OF GEOMETRIC RANDOM VARIABLES

Lemma 6. *Let $p \in (0, 1]$, $t \in \mathbb{N}$ and $Z = \sum_{i=1}^{t} G_i$, where G_i are independent geometric random variables with parameter p. Then*

$$\Pr\left(\frac{t}{2p} \leq Z \leq \frac{2t}{p}\right) \geq \frac{7}{16}. \tag{11}$$

Proof. Let G be a geometric random variable of parameter p, then it has expectation value $1/p$ and variance $(1-p)/p^2 \leq 1/p^2$. Moreover $\Pr(G \leq k) = 1 - (1-p)^k$ for all $k \in \mathbb{N}$. In particular,

$$\Pr(\lfloor 1/(2p)\rfloor < G \leq \lfloor 2/p\rfloor) = (1-p)^{\lfloor 1/(2p)\rfloor} - (1-p)^{\lfloor 2/p\rfloor} \geq \frac{7}{16}.$$

More generally Z has negative binomial distribution.

One can check that for every $t \in [7]$ and all $p \in (0, 1]$ we have that $\Pr(\lfloor t/(2p)\rfloor < Z \leq \lfloor 2t/p\rfloor) \geq 7/16$, see Figure 8.

On the other hand the variance of Z is at most $\frac{t}{p^2}$, so

$$\Pr\left(|Z - t/p| \geq \frac{t}{2p}\right) \leq \frac{4}{t}$$

by Chebyshev's inequality, which implies the claim for $t \geq 8$. □

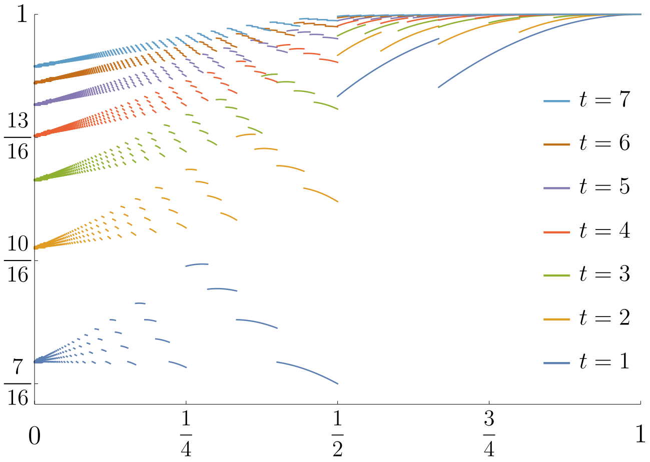

Figure 8: A plot of $\Pr\left(\frac{t}{2p} \leq Z \leq \frac{2t}{p}\right)$ as a function of p, illustrating Equation (11) for $t = 1, 2, \ldots, 7$.

ACKNOWLEDGMENTS

We thank Simon Apers, Jérémie Roland, Shantanav Chakraborty, and Ashwin Nayak for useful comments and discussions, and Ronald de Wolf for thorough feedback on the manuscript.

A.A., A.G., and M.K. were supported by QuantERA project QuantAlgo 680-91-034. A.A. also acknowledges support by Latvian Council of Science grant lzp-2018/1-0173. A.G. was also supported by ERC Consolidator Grant QPROGRESS. S.J. was supported by an NWO Veni Innovational Research Grant under project number 639.021.752 and an NWO WISE Grant.

REFERENCES

[1] Simon Apers, András Gilyén, and Stacey Jeffery. 2019. A Unified Framework of Quantum Walk Search. (2019). arXiv: 1912.04233

[2] Simon Apers and Alain Sarlette. 2019. Quantum Fast-Forwarding: Markov Chains and Graph Property Testing. *Quantum Information and Computation* 19, 3&4 (2019), 181–213. https://doi.org/10.26421/QIC19.3-4 arXiv: 1804.02321

[3] Michel Boyer, Gilles Brassard, Peter Høyer, and Alain Tapp. 1998. Tight Bounds on Quantum Searching. *Fortschritte der Physik* 46, 4–5 (1998), 493–505. https://doi.org/10.1002/(SICI)1521-3978(199806)46:4/5<493::AID-PROP493>3.0.CO;2-P arXiv: quant-ph/9605034

[4] Shantanav Chakraborty, András Gilyén, and Stacey Jeffery. 2019. The power of block-encoded matrix powers: improved regression techniques via faster Hamiltonian simulation. In *Proceedings of the 46th International Colloquium on Automata, Languages, and Programming (ICALP)*. 33:1–33:14. https://doi.org/10.4230/LIPIcs.ICALP.2019.33 arXiv: 1804.01973

[5] Cătălin Dohotaru and Peter Høyer. 2017. Controlled Quantum Amplification. In *Proceedings of the 44th International Colloquium on Automata, Languages, and Programming (ICALP)*. 18:1–18:13. https://doi.org/10.4230/LIPIcs.ICALP.2017.18

[6] András Gilyén, Yuan Su, Guang Hao Low, and Nathan Wiebe. 2019. Quantum singular value transformation and beyond: exponential improvements for quantum matrix arithmetics. In *Proceedings of the 51st ACM Symposium on the Theory of Computing (STOC)*. 193–204. https://doi.org/10.1145/3313276.3316366 arXiv: 1806.01838

[7] Peter Høyer and Mojtaba Komeili. 2017. Efficient Quantum Walk on the Grid with Multiple Marked Elements. In *Proceedings of the 34th Symposium on Theoretical Aspects of Computer Science (STACS)*. 42:1–42:14. https://doi.org/10.4230/LIPIcs.STACS.2017.42 arXiv: 1612.08958

[8] Hari Krovi, Frédéric Magniez, Maris Ozols, and Jérémie Roland. 2016. Quantum Walks Can Find a Marked Element on Any Graph. *Algorithmica* 74, 2 (2016), 851–907. https://doi.org/10.1007/s00453-015-9979-8 arXiv: 1002.2419

[9] David A. Levin and Yuval Peres. 2017. *Markov chains and mixing times* (2nd ed.). AMS, Providence, RI, USA. https://doi.org/10.1090/mbk/107

[10] László Lovász. 1996. Random Walks on Graphs: A Survey. In *Combinatorics, Paul Erdős is Eighty (Vol. 2) (Bolyai Society Mathematical Studies)*, Dezső Miklós, Tamás Szőnyi, and Vera T. Sós (Eds.). János Bolyai Mathematical Society, 1–46. http://web.cs.elte.hu/~lovasz/erdos.pdf

[11] Guang Hao Low and Isaac L. Chuang. 2017. Hamiltonian Simulation by Uniform Spectral Amplification. (2017). arXiv: 1707.05391

[12] Frédéric Magniez, Ashwin Nayak, Peter C. Richter, and Miklos Santha. 2012. On the Hitting Times of Quantum Versus Random Walks. *Algorithmica* 63, 1 (2012), 91–116. https://doi.org/10.1007/s00453-011-9521-6 Earlier version in SODA'09. arXiv: 0808.0084

[13] Frédéric Magniez, Ashwin Nayak, Jérémie Roland, and Miklos Santha. 2011. Search via Quantum Walk. *SIAM Journal on Computing* 40, 1 (2011), 142–164. https://doi.org/10.1137/090745854 Earlier version in STOC'07. arXiv: quant-ph/0608026

[14] Márió Szegedy. 2004. Quantum speed-up of Markov chain based algorithms. In *Proceedings of the 45th IEEE Symposium on Foundations of Computer Science (FOCS)*. 32–41. https://doi.org/10.1109/FOCS.2004.53 arXiv: quant-ph/0401053

[15] Avatar Tulsi. 2008. Faster quantum-walk algorithm for the two-dimensional spatial search. *Physical Review A* 78, 1 (2008), 012310. https://doi.org/10.1103/PhysRevA.78.012310 arXiv: 0801.0497

[16] Curtis R. Vogel. 2002. *Computational Methods for Inverse Problems*. SIAM, Philadelphia, PA, USA. https://doi.org/10.1137/1.9780898717570

[17] Theodore J. Yoder, Guang Hao Low, and Isaac L. Chuang. 2014. Fixed-Point Quantum Search with an Optimal Number of Queries. *Physical Review Letters* 113, 21 (2014), 210501. https://doi.org/10.1103/PhysRevLett.113.210501 arXiv: 1409.3305