Search via Quantum Walk

Frédéric Magniez
LRI, Univ. Paris-Sud, CNRS
F-91405 Orsay, France
magniez@lri.fr

Ashwin Nayak
C&O and IQC, U. Waterloo
and Perimeter Institute
200 University Ave. W.
Waterloo, Ontario N2L 3G1,
Canada
anayak@uwaterloo.ca

Jérémie Roland
Computer Science Division
U.C. Berkeley
Berkeley, CA 94720, USA
jroland@berkeley.edu

Miklos Santha
LRI, Univ. Paris-Sud, CNRS
F-91405 Orsay, France
santha@lri.fr

ABSTRACT

We propose a new method for designing quantum search algorithms for finding a "marked" element in the state space of a classical Markov chain. The algorithm is based on a quantum walk *à la* Szegedy [25] that is defined in terms of the Markov chain. The main new idea is to apply quantum phase estimation to the quantum walk in order to implement an approximate reflection operator. This operator is then used in an amplitude amplification scheme. As a result we considerably expand the scope of the previous approaches of Ambainis [6] and Szegedy [25]. Our algorithm combines the benefits of these approaches in terms of being able to find marked elements, incurring the smaller cost of the two, and being applicable to a larger class of Markov chain. In addition, it is conceptually simple, avoids several technical difficulties in the previous analyses, and leads to improvements in various aspects of several algorithms based on quantum walk.

Categories and Subject Descriptors

F.2 [**Theory of Computation**]: Analysis of Algorithms and Problem Complexity

General Terms

Algorithms,Theory

Keywords

Search, Markov chain, hitting time, quantum walk, phase estimation, amplitude amplification, spectral gap, phase gap, reflection operator, recursive amplitude amplification

STOC'07, June 11–13, 2007, San Diego, California, USA.

1. INTRODUCTION

1.1 Background

At an abstract level, any search problem may be cast as the problem of finding a "marked" element from a set X with n elements. Let $M \subseteq X$ be the set of marked elements. One approach to finding from M, if it is not empty, is to repeatedly sample from X uniformly until a marked element is picked. A more cost-effective approach re-uses resources expended in generating the first sample (time, random bits, black-box queries, etc.) by simulating the steps of a Markov chain with state space X to generate the next sample. This approach often takes advantage of some structure present in the ground set X and the Markov chain, and leads to a more efficient algorithm. In this article, we study quantum analogues of this randomized scheme.

Discrete time quantum walks emerged gradually in the field of quantum algorithms (see [25] for a detailed introduction.) On the line they are related to the quantum cellular automaton model of Meyer [21]. Watrous [27] introduced quantum walks on regular graphs, and used them to show that randomized logarithmic space is included in quantum logarithmic space. Afterwards notions related to quantum walks, such as mixing time, and deviation from starting state, were studied for restricted graphs by several researchers [22, 7, 3, 23], suggesting the possibility of speed-up of classical algorithms based on random walk.

Shenvi, Kempe, and Whaley [24] pointed out the algorithmic potential of quantum walks by designing a walk based algorithm to emulate Grover Search [16]. The first algorithm using quantum walks that goes beyond the capability of Grover Search is due to Ambainis [6] for Element Distinctness. In his seminal paper he resolved the quantum query complexity of the problem, settling a difficult question that had been open for several years [12, 1]. Finally Szegedy [25] developed a theory of quantum walk based algorithms. He designed a quantum search algorithm based on any symmetric, ergodic Markov chain that detects the presence of a marked element. He defined a notion of quantum hitting time that is quadratically smaller than the classical average hitting time. Since then, in the framework of Ambainis or Szegedy, numerous new algorithms emerged with substantially better complexity in a variety of contexts [5, 20, 13,

19].

This work develops a new schema for quantum search algorithms, based on any ergodic Markov chain. We extend the Szegedy quantization of classical Markov chains to possibly non-symmetric Markov chains, but its use is more in the style of the Ambainis algorithm. Departing from the two algorithms, however, we use quantum walks only indirectly. In conjunction with the well known phase estimation algorithm [18, 15], the quantum walk helps us implement an approximate reflection operator. This operator may then be used within amplitude amplification algorithms [16, 10, 17] for search. As a result, our work generalizes previous ones by extending the class of possible Markov chains, and improving the complexity in terms of its relation with the eigenvalue or singular value gap of the related Markov chain. In addition, our approach is conceptually simple, avoids several technical difficulties in the analysis of the earlier approaches, and leads to improvements in various aspects of the algorithms.

1.2 Two subtly different search algorithms

The optimal quantum algorithm for Element Distinctness discovered by Ambainis [6] recasts the problem in terms of search for a marked state in a Johnson graph defined by the problem instance. The algorithm may be viewed as a quantum analogue of the following search process, where P is the transition matrix of a Markov chain defined on state space X.

Search Algorithm 1

1. Initialize x to a state sampled from a probability distribution s over X.
2. Repeat for t_2 steps
   (a) If the state y reached in the previous step is marked, then stop and output y.
   (b) Else, simulate t_1 steps of the Markov chain P starting with the current state y.
3. If the algorithm has not terminated, stop, and output 'no marked element exists'.

The parameters t_1 and t_2 in the algorithm are determined by the properties of the Markov chain and the marked subset M. The idea behind this algorithm is illustrated by considering an ergodic Markov chain P. If t_1 is chosen large enough, the state y in step (2a) above will be distributed (approximately) according to the stationary distribution of P. Then, the outer loop represents sampling from the stationary distribution until a marked element is found. If t_2 is chosen to be inversely proportional to the probability that a state is marked according to the stationary distribution, the algorithm will succeed with high probability.

The analysis of the Ambainis quantum algorithm depends heavily on the form of marked states, and was presented for subsets M arising out of k-Collision, a generalization of Element Distinctness, with the assumption of a unique collision. Inspired by this algorithm, Szegedy [25] designed a quantum search algorithm with uniform initial distribution, based on any symmetric, ergodic Markov chain. The Szegedy algorithm may be viewed as a quantum analogue of a subtly different, but more natural, classical process.

Search Algorithm 2

1. Initialize x to a state sampled from a probability distribution s over X.
2. Repeat for t steps
   (a) If the state y reached in the previous step is marked, then stop and output y.
   (b) Else, simulate *one* step of the Markov chain P from the current state y.
3. If the algorithm has not terminated, stop, and output 'no marked element exists'.

The parameter t is also determined by the Markov chain P, and the set M of marked states. This algorithm is a greedy version of the first algorithm: a check is performed after every step of the Markov chain to determine if a marked state has been reached, irrespective of whether the Markov chain has mixed.

Let us formally derive the complexity of the two algorithms to clarify their differences. Assume that the search algorithms maintain a data structure d that associates some data $d(x)$ with every state $x \in X$. From $d(x)$, we would like to determine if $x \in M$. When operating with d, we distinguish three types of cost.

Setup cost S: The cost to sample $x \in X$ according to the initial distribution s, and to construct the data structure $d(x)$ for the state x.
Update cost U: The cost to simulate a transition from x to y for a state $x \in X$ according to the Markov chain P, and to update $d(x)$ to $d(y)$.
Checking cost C: The cost of checking if $x \in M$ using $d(x)$.

The cost may be thought of as a vector listing all the measures of complexity of interest, such as query and time complexity. We may now state generic bounds on the efficiency of the two search algorithms in terms of our cost parameters.

PROPOSITION 1. *Let $\delta > 0$ be the eigenvalue gap of an ergodic, symmetric Markov chain P on a state space X of size n, and let $\frac{|M|}{|X|} \geq \varepsilon > 0$ whenever $M \subset X$ is non-empty. For the uniform initial distribution s,*

1. ***Search Algorithm 1** determines if a marked element exists and finds one such element with high probability if $t_1 = \mathrm{O}(\frac{1}{\delta})$ and $t_2 = \mathrm{O}(\frac{1}{\varepsilon})$. The cost incurred is of order $\mathsf{S} + \frac{1}{\varepsilon}\left(\frac{1}{\delta}\mathsf{U} + \mathsf{C}\right)$.*
2. ***Search Algorithm 2** determines if a marked element exists and finds one such element with high probability if $t = \mathrm{O}(\frac{1}{\delta\varepsilon})$. The cost incurred is of order $\mathsf{S} + \frac{1}{\delta\varepsilon}\left(\mathsf{U} + \mathsf{C}\right)$.*

PROOF. The stopping time of **Search Algorithm 2** is the average hitting time of the set M for the Markov chain P. We may therefore take t to be a constant factor more than this hitting time. As mentioned before, this time is bounded above by the stopping time for the first algorithm. Therefore part 2 of the proposition follows from part 1.

In the first algorithm, we may take t_2 to be proportional to the average hitting time of the set M for the Markov chain P^{t_1}. The quantity $\lambda^*(P)$, the magnitude of the eigenvalue of P with maximum absolute value among the eigenvalues different from 1, is bounded by $1 - \delta$ by hypothesis. The analogous quantity $\lambda^*(P^{t_1})$ is therefore bounded

by $(1-\delta)^{t_1} \leq \mathrm{e}^{-\delta t_1}$. Taking $t_1 = 1/\delta$, we get a spectral gap $\tilde{\delta}$ of at least $1/2$ for P^{t_1}. We may now bound the average hitting time of M for P^{t_1} by, for example, Equation 15 in Ref. [25] and Lemma 1 in Ref. [11] (also stated as Lemma 10 in Ref. [25]). This bound evaluates to $\frac{1}{\varepsilon\tilde{\delta}} \leq \frac{2}{\varepsilon}$. The expression for the cost of the algorithm now follows. □

For special classes of graph, for example for the 2-d toroidal grid, the hitting time may be significantly smaller than the generic bound $t = \mathrm{O}(1/\delta\varepsilon)$ given in part 2 (see [4, Page 11, Chapter 5]).

1.3 Quantum analogues

As in the classical case, the quantum search algorithms look for a marked element in a finite set X, where a data structure d is maintained during the algorithm. Let X_d be the set of items along with their associated data, that is $X_d = \{(x, d(x)) : x \in X\}$. For simplicity we suppose that $\bar{0} \in X$ and that $d(\bar{0}) = \bar{0}$.

The quantum walks of Ambainis and Szegedy, and also ours, may be thought of as a walk on *edges* of the original Markov chain, rather than its vertices. Thus, that state space is a linear subspace of the vector space $\mathcal{H} = \mathbb{C}^{X\times X}$, or $\mathcal{H}_d = \mathbb{C}^{X_d \times X_d}$ when we also include the data structure. For the sake of elegance in the mathematical analyses, our data structure keeps the data for both vertices of an edge, whereas in previous works the data was kept alternately only for one of them. There is a natural isomorphism $|\psi\rangle \mapsto |\psi\rangle_d$ between $\mathcal{H}$ and $\mathcal{H}_d$, where on basis states $|x\rangle_d = |x, d(x)\rangle$. This isomorphism maps a unitary operation U on $\mathcal{H}$ into U_d on $\mathcal{H}_d$ defined by $U_d|\psi\rangle_d = (U|\psi\rangle)_d$.

Our walks will be discussed in the space $\mathcal{H}_d$ when, for implementation and cost considerations, it will be important to deal properly with the data structure. However, for convenience, we will analyze the mathematical properties of the walks without the data structure, in the space $\mathcal{H}$. This is legitimate because of the isomorphism between $\mathcal{H}_d$ and $\mathcal{H}$.

The initial state of the algorithm is explicitly related to the stationary distribution π of P. At each step, the right end-point of an edge (x, y) is "mixed" over the neighbours of x, and then the left end-point is mixed over the neighbours of the new right end-point. We will distinguish again three types of cost generalizing those of the classical search. They are of the same order as the corresponding costs in the algorithms of Ambainis and Szegedy. Some operations of the algorithms not entering into these costs will not be accounted for. This is justified by the fact that in all quantum search algorithms the overall complexity is of the order of the accounted part expressed in the specific costs.

(Quantum) Setup cost S: The cost for constructing the state $\sum_x \sqrt{\pi_x}|x\rangle_d|\bar{0}\rangle_d$
(Quantum) Update cost U: The cost to realize any of the unitary transformations and inverses

$$\begin{aligned} |x\rangle_d|\bar{0}\rangle_d &\mapsto |x\rangle_d \textstyle\sum_y \sqrt{p_{xy}}|y\rangle_d, \\ |\bar{0}\rangle_d|y\rangle_d &\mapsto \textstyle\sum_x \sqrt{p^*_{yx}}|x\rangle_d|y\rangle_d, \end{aligned}$$

where $P^* = (p^*_{xy})$ is the time-reversed Markov chain defined by the equations $\pi_x p_{xy} = \pi_y p^*_{yx}$.
(Quantum) Checking cost C: The cost to realize the unitary transformation that maps $|x\rangle_d|y\rangle_d$ to $-|x\rangle_d|y\rangle_d$ if $x \in M$, and leaves it unchanged otherwise.

The quantum search algorithms of Ambainis and Szegedy give a quadratic speed up in the times t_1, t_2 and t, with respect to the classical algorithms.

THEOREM 1 (AMBAINIS [6]). *Let P be the random walk on the Johnson graph on r-subsets of a universe of size m, with intersection size $r-1$, where $r = \mathrm{o}(m)$. Let M be either empty, or the class of all r subsets that contain a fixed subset of size a constant $k \leq r$. The eigenvalue gap of P is $\delta \in \Omega(\frac{1}{r})$, and the fraction of marked elements is either 0 or $\varepsilon = \Omega(\frac{r^k}{m^k})$. Then, there is a quantum algorithm that determines, with high probability, if M is non-empty or finds the k-subset with cost of order $\mathsf{S} + \frac{1}{\sqrt{\varepsilon}}(\frac{1}{\sqrt{\delta}}\mathsf{U} + \mathsf{C})$.*

THEOREM 2 (SZEGEDY [25]). *Let $\delta > 0$ be the eigenvalue gap of an ergodic, symmetric Markov chain P, and let $\frac{|M|}{|X|} \geq \varepsilon > 0$ whenever M is non-empty. There exists a quantum algorithm that determines, with high probability, if M is non-empty with cost of order $\mathsf{S} + \frac{1}{\sqrt{\delta\varepsilon}}(\mathsf{U} + \mathsf{C})$.*

If the checking cost C is substantially greater than that of performing one step of the walk, an algorithm with the cost structure of the Ambainis algorithm would be more efficient. Moreover, the algorithm would *find* a marked element if one exists. These advantages are illustrated by the algorithm for Triangle Finding [20]. This algorithm uses two quantum walks *à la* Ambainis recursively; the Szegedy framework seems to give a less efficient algorithm. Nonetheless, the Szegedy approach has other advantages—it applies to a wider class of Markov chain and for arbitrary sets of marked state. Moreover, the quantity $1/\sqrt{\delta\epsilon}$ in Theorem 2 may be replaced by the square-root of the classical hitting time [25]. These features make it more suitable for applications such as the near-optimal algorithm for Group Commutativity [19] which has no equivalent using the Ambainis approach.

1.4 Contribution, relation with prior work, and organization

Our algorithm is a quantum analogue of **Search Algorithm 1** and works for any ergodic Markov chain. It is most easily described for *reversible* Markov chains.

THEOREM 3 (**This paper**). *Let $\delta > 0$ be the eigenvalue gap of a reversible, ergodic Markov chain P, and let $\varepsilon > 0$ be a lower bound on the probability that an element chosen from the stationary distribution of P is marked whenever M is non-empty. Then, there is a quantum algorithm that with high probability, determines if M is empty or finds an element of M, with cost of order $\mathsf{S} + \frac{1}{\sqrt{\varepsilon}}(\frac{1}{\sqrt{\delta}}\mathsf{U} + \mathsf{C})$.*

Our algorithm considerably expands the scope of the approaches embodied in Theorems 1 and 2 above. It combines the benefits of both approaches in terms of being able to find marked elements, incurring the smaller cost of the two, and being applicable to a larger class of Markov chain. In addition, it is conceptually simple, avoids several technical difficulties in the analysis of the earlier approaches, and leads to improvements in various aspects of algorithms for Element Distinctness, Matrix Product Verification, Triangle Finding, and Group Commutativity. Namely, we give a single-shot method for any algorithm *à la* Ambainis in presence of multiple solutions, without the need for a reduction to special cases such as that of a unique solution. This applies to

Element Distinctness and Triangle Finding. For Element Distinctness, Matrix Product Verification, and Group Commutativity, where an algorithm *à la* Szegedy only detects the existence of a solution, we find one with the same time and query complexity. Finally, we improve the complexity of the best previous known algorithm for Triangle Finding from $O(n^{1.3}\text{polylog}(n))$ to $O(n^{1.3})$.

In Section 2, we describe a quantum analogue of a Markov chain based on the work of Szegedy [25] who defined such a quantum process $W(P,Q)$ for a classical bipartite walk (P,Q). By letting $Q = P$, he related the spectrum of the quantum walk $W(P)$ to that of P for symmetric Markov chains. Using an absorbing version of P as in **Search Algorithm 2**, he designed a quantum analogue of this classical scheme. Even when P is not symmetric, letting $Q = P^*$, the time-reversed Markov chain corresponding to P, leads to a natural connection between P and $W(P)$. If P is reversible, then the eigenvalues of $W(P)$ are closely related to those of P, as in the symmetric case. For an arbitrary, possibly non-reversible, ergodic Markov chain, this connection relates the eigenvalues of $W(P)$ to the singular values a "discriminant" matrix $D(P)$ associated with P.

In Section 3, we use the quantum walk $W(P)$ with the unperturbed walk P in a completely different way, more in the style of the Ambainis approach. Ambainis directly uses a power of $W(P)$ to replace the "diffusion" operator in the Grover search algorithm. The beauty of this step, and the difficulty of proving its correctness, lies in the fact that even if no power of $W(P)$ closely approximates the diffusion operator, some powers have sufficient properties to mimic its essential features. We introduce a novel way to approximate the diffusion operator. Our approach is both conceptually simpler, and more general. We observe that $W(P)$ amplifies the spectral gap of a reversible Markov chain quadratically. We translate this to an efficient approximation to the Grover diffusion operator (**Theorem 6**), using the well known phase estimation algorithm. We then begin an exposition of our algorithm by considering reversible Markov chains. To explain the basic idea of our approach, we first prove our main result with an additional logarithmic factor (**Theorem 7**).

In Section 4, using a technique developed by Høyer, Mosca, and de Wolf [17] we show how to eliminate the logarithmic factor in the previous theorem, thus proving **Theorem 3**.

In Section 5, we extend the algorithm to a possibly non-reversible Markov chain whose discriminant $D(P)$ has non-zero singular value gap (**Theorem 9**). The complexity of the algorithm in the general case is similar to the one for reversible Markov chains. The sole difference is that the singular value gap of $D(P)$ takes the place of the spectral gap of P. While the eigenvalues of Markov chains are well studied, we are not aware of a similar theory for singular values of this matrix. Nonetheless, we believe that such a general result may prove useful for future applications.

For simplicity, in our algorithms we assume that the number of marked elements is known in advance. This assumption may be removed using standard techniques, without increasing the asymptotic complexity of the algorithms [9].

2. QUANTUM ANALOGUE OF A CLASSICAL MARKOV CHAIN

Let $P = (p_{xy})$ be the transition matrix of any irreducible Markov chain on a finite space X with $|X| = n$. By the Perron-Frobenius theorem, the chain has a unique stationary distribution $\pi = (\pi_x)$, that is a unique left eigenvector π with eigenvalue 1 and positive coordinates summing up to 1. Since we will only consider irreducible Markov chains in this paper, from now on we may omit the term 'irreducible' in our discussions. We will also often identify a Markov chain with its transition matrix P.

Let $P^* = (p^*_{yx})$ be the time-reversed Markov chain defined by the equations $\pi_x p_{xy} = \pi_y p^*_{yx}$. We will define a quantum analogue of an arbitrary irreducible Markov chain P, based on and extending the notion of quantum Markov chain due to Szegedy [25]. The latter was inspired by an earlier notion of quantum walk due to Ambainis [6]. We also point out that a similar process on regular graphs was studied by Watrous [27].

For a state $|\psi\rangle \in \mathcal{H}$, let $\Pi_\psi = |\psi\rangle\langle\psi|$ denote the orthogonal projector onto $\mathsf{Span}(|\psi\rangle)$, and let $\text{ref}(\psi) = 2\Pi_\psi - \text{Id}$ denote the reflection through the line generated by $|\psi\rangle$, where Id is the identity operator on $\mathcal{H}$. If $\mathcal{K}$ is a subspace of $\mathcal{H}$ spanned by a set of mutually orthogonal states $\{|\psi_i\rangle : i \in I\}$, then let $\Pi_\mathcal{K} = \sum_{i\in I} \Pi_{\psi_i}$ be the orthogonal projector onto $\mathcal{K}$, and let $\text{ref}(\mathcal{K}) = 2\Pi_\mathcal{K} - \text{Id}$ be the reflection through $\mathcal{K}$.

Let $\mathcal{A} = \mathsf{Span}(|x\rangle|p_x\rangle : x \in X)$ and $\mathcal{B} = \mathsf{Span}(|p^*_y\rangle|y\rangle : y \in X)$ be vector subspaces of $\mathcal{H} = \mathbb{C}^{X\times X}$, where

$$|p_x\rangle = \textstyle\sum_{y\in X} \sqrt{p_{xy}}\,|y\rangle \quad \text{and} \quad |p^*_y\rangle = \sum_{x\in X} \sqrt{p^*_{yx}}\,|x\rangle.$$

Definition 1 (Quantum walk). *The unitary operation $W(P)$ defined on $\mathcal{H}$ by $W(P) = \text{ref}(\mathcal{B})\cdot\text{ref}(\mathcal{A})$ is called the* quantum walk *based on the classical chain P.*

This quantum walk extends to a walk $W(P)_d$ on the space $\mathcal{H}$ augmented with data structures, as explained in Section 1.3. Recall that U is the quantum update cost as defined in the same section.

Proposition 2. *The quantum walk with data $W(P)_d$ can be implemented at cost $4\mathsf{U}$.*

Proof. The reflection $\text{ref}(\mathcal{A})_d$ is implemented by mapping states $|x\rangle_d|p_x\rangle_d$ to $|x\rangle_d|\bar{0}\rangle_d$, applying $\text{ref}(\mathcal{H}_d \otimes |\bar{0}\rangle_d)$, and undoing the first transformation. In our accounting we charge unit cost for the second step since it does not depend on the database. Therefore the implementation of $\text{ref}(\mathcal{A})_d$ is of cost $2\mathsf{U}$. The reflection $\text{ref}(\mathcal{B})$ may be implemented similarly. □

The eigen-spectrum of the transition matrix P plays an important role in the analysis of a classical Markov chain. Similarly, the behaviour of the quantum process $W(P)$ may be inferred from its spectral decomposition. For this, motivated by Szegedy [25], we consider the *discriminant* matrix $D(P) = (\sqrt{p_{xy}p^*_{yx}})$. Since $\sqrt{p_{xy}p^*_{yx}} = \sqrt{\pi_x}p_{xy}/\sqrt{\pi_y}$, the discriminant matrix is equal to

$$D(P) \quad = \quad \text{diag}(\pi)^{1/2}\cdot P\cdot\text{diag}(\pi)^{-1/2},$$

where $\text{diag}(\pi)$ is the invertible diagonal matrix with the coordinates of the distribution π in its diagonal. Since the singular values of $D(P)$ all lie in the range $[0,1]$, we may express them as $\cos\theta$, for some angles $\theta \in [0, \frac{\pi}{2}]$. For later reference, we rewrite Theorem 1 of Szegedy [25] which relates the singular value decomposition of $D(P)$ to the spectral decomposition of $W(P)$.

THEOREM 4 (SZEGEDY [25]). *Let P be an irreducible Markov chain, and let $\cos\theta_1, \ldots, \cos\theta_l$ be an enumeration of those singular values (possibly repeated) of $D(P)$ that lie in the open interval $(0,1)$. Then:*

1. *On $\mathcal{A}+\mathcal{B}$ those eigenvalues of $W(P)$ that have non-zero imaginary part are exactly $e^{\pm 2i\theta_1}, \ldots, e^{\pm 2i\theta_l}$, with the same multiplicity.*
2. *On $\mathcal{A}\cap\mathcal{B}$ the operator $W(P)$ acts as the identity* Id. *$\mathcal{A}\cap\mathcal{B}$ coincides with the left (and right) singular vectors of $D(P)$ with singular value* 1.
3. *On $\mathcal{A}\cap\mathcal{B}^\perp$ and $\mathcal{A}^\perp\cap\mathcal{B}$ the operator $W(P)$ acts as* $-$Id. *$\mathcal{A}\cap\mathcal{B}^\perp$ (respectively, $\mathcal{A}^\perp\cap\mathcal{B}$) coincides with the set of left (respectively, right) singular vectors of $D(P)$ with singular value* 0.
4. *$W(P)$ has no other eigenvalues on $\mathcal{A}+\mathcal{B}$; on $\mathcal{A}^\perp\cap\mathcal{B}^\perp$ the operator $W(P)$ acts as* Id.

We define $\Delta(P)$, the *phase gap* of $W(P)$ as 2θ, where θ is the smallest angle in $(0, \frac{\pi}{2})$ such that $\cos\theta$ is a singular value of $D(P)$. This definition is motivated by the previous theorem: the angular distance of 1 from any other eigenvalue is at least $\Delta(P)$.

3. FROM QUANTUM WALK TO SEARCH

3.1 Outline of search algorithm

We now describe a search algorithm that may be viewed as a quantum analogue of **Search Algorithm 1** of Section 1.2. Consider the following quantum state in the Hilbert space $\mathcal{H}$:

$$|\pi\rangle = \sum_{x\in X} \sqrt{\pi_x}\,|x\rangle|p_x\rangle = \sum_{y\in X} \sqrt{\pi_y}\,|p_y^*\rangle|y\rangle.$$

This state will serve as the initial state for our algorithm, and corresponds to starting in the stationary distribution π in the classical search algorithms. Assume that $M \neq \emptyset$. Let $\mathcal{M} = \mathbb{C}^{M\times X}$ denote the subspace with marked items in the first register. We would like to transform the initial state $|\pi\rangle$ to the target state $|\mu\rangle$, which is the (normalized) projection of $|\pi\rangle$ onto the "marked subspace" $\mathcal{M}$:

$$|\mu\rangle \quad = \quad \frac{\Pi_{\mathcal{M}}|\pi\rangle}{\|\Pi_{\mathcal{M}}|\pi\rangle\|} \quad = \quad \frac{1}{\sqrt{\varepsilon}} \sum_{x\in M} \sqrt{\pi_x}\,|x\rangle|p_x\rangle,$$

where $\varepsilon = \|\Pi_{\mathcal{M}}|\pi\rangle\|^2 = \sum_{x\in M}\pi_x$ is the probability of a set M of marked states under the stationary distribution π. Roughly speaking, we will effect this transformation by implementing a rotation *à la* Grover [16] in the two-dimensional real subspace $\mathcal{S} = \mathsf{Span}(|\pi\rangle,\ |\mu\rangle)$ generated by the states.

Ideally, we would like to effect the rotation $\mathrm{ref}(\pi)_d \cdot \mathrm{ref}(\mu^\perp)_d$ in $\mathcal{S}_d$, where $|\mu^\perp\rangle$ is the state in $\mathcal{S}$ orthogonal to $|\mu\rangle$ which makes an acute angle with $|\pi\rangle$. The angle φ between $|\pi\rangle$ and $|\mu^\perp\rangle$ is given by $\sin\varphi = \langle\mu|\pi\rangle = \sqrt{\varepsilon}$. The product of the two reflections above is a rotation by an angle of 2φ within the space $\mathcal{S}$. Therefore, after $\mathrm{O}(1/\varphi) = \mathrm{O}(1/\sqrt{\varepsilon})$ iterations of this rotation starting with the state $|\pi\rangle$, we will have approximated the target state $|\mu\rangle$.

Restricted to the subspace $\mathcal{S}$, the operators $\mathrm{ref}(\mu^\perp)$ and $-\mathrm{ref}(\mathcal{M})$ are identical. Therefore, if we ensure that the state of the algorithm remains close to the subspace $\mathcal{S}$ throughout, we would be able to implement $\mathrm{ref}(\mu^\perp)_d$: it involves checking at cost C whether an item in the first register is marked.

The reflection $\mathrm{ref}(\pi)_d$ is computationally harder to perform. The straightforward strategy would be to rotate $|\pi\rangle_d$ to the state $|\bar{0}\rangle_d|\bar{0}\rangle_d$, use $\mathrm{ref}(|\bar{0}\rangle_d|\bar{0}\rangle_d)$, and then undo the first rotation. However, rotating $|\pi\rangle_d$ to $|\bar{0}\rangle_d|\bar{0}\rangle_d$ is exactly the inverse operation of the preparation of the initial state $|\pi\rangle_d$ from $|\bar{0}\rangle_d|\bar{0}\rangle_d$, and therefore requires the same cost $\mathsf{S}+\mathsf{U}$. This may be much more expensive than the update cost $4\mathsf{U}$ incurred by the walk $W(P)_d$. To use $W(P)_d$ instead, our idea is to apply phase estimation to it, and exploit this procedure to approximate the required diffusion operator on $\mathcal{A}_d + \mathcal{B}_d$ which contains the subspace $\mathcal{S}_d$.

3.2 Diffusion operator from quantum walk

To explain our approach, in the rest of this section, and in the next one, we assume that the classical Markov chain P is ergodic and reversible. Let us recall that a finite Markov chain is ergodic if it is irreducible and aperiodic. The Markov chain P is said to be *reversible* if $P = P^*$, where P^* is the time reversed chain. For a reversible chain the corresponding discriminant $D(P)$ is symmetric. Symmetry implies that the singular values of $D(P)$ equal the absolute values of its eigenvalues. Since $D(P) = \mathrm{diag}(\pi)^{1/2}\cdot P\cdot\mathrm{diag}(\pi)^{-1/2}$ is similar to the matrix P, their spectra are the same. Therefore, we only study the spectrum of P. The Perron-Frobenius theorem and the ergodicity of P imply that the eigenvalue 1 is the only eigenvalue of P with absolute value 1 and has multiplicity 1. The corresponding eigenvector of $D(P)$ is $(\sqrt{\pi_x})$, and every singular (or eigen-) vector of $D(P)$ orthogonal to this has singular value strictly less than 1. Transferring this property to the quantum walk $W(P)$ via Theorem 4, $|\pi\rangle$ is the unique eigenvector of the unitary operator $W(P)$ in $\mathcal{A}+\mathcal{B}$ with eigenvalue 1, and the remaining eigenvalues in $\mathcal{A}+\mathcal{B}$ are bounded away from 1. We use this observation to identify the component of any state $|\psi\rangle \in \mathcal{S}$ perpendicular to $|\pi\rangle$.

The main idea in our implementation of the above approach is to use phase estimation [18, 15].

THEOREM 5 (PHASE ESTIMATION, CLEVE ET AL. [15]). *For every pair of integers $m, r \geq 1$, and a unitary operator U of dimension $m\times m$, there exists a quantum circuit $C(U)$ that acts on $m+s$ qubits, where $s = r + \mathrm{O}(1)$ and satisfies the following properties:*

1. *The circuit $C(U)$ uses $2s$ Hadamard gates, $\mathrm{O}(s^2)$ controlled phase rotations, and makes 2^{s+1} calls to the controlled unitary operator $\mathrm{c}{-}U$.*
2. *For any eigenvector $|\psi\rangle$ of U with eigenvalue* 1, *i.e., if $U|\psi\rangle = |\psi\rangle$, then $C(U)|\psi\rangle|0^s\rangle = |\psi\rangle|0^s\rangle$.*
3. *If $U|\psi\rangle = \mathrm{e}^{\mathrm{i}\theta}|\psi\rangle$, where $\theta \in (0, 2\pi)$, then rounding off the number in the last s qubits of the state $C(U)|\psi\rangle|0^s\rangle$ to the r most significant bits gives the best r-bit approximation to $\theta/2\pi$ with probability at least* 2/3.

Moreover the family of circuits C parametrized by m and s is uniform.

In the following theorem we show how phase estimation can be used to design a quantum circuit $R(P)$ which implements an operation that is close to the reflection $\mathrm{ref}(\pi)$.

THEOREM 6. *Let P be an ergodic Markov chain on a state space of size n, and let $m = n^2$. Then there is a constant $c > 0$ such that for any integer k, there exists a quantum circuit $R(P)$ that acts on $m+ks$ qubits, where $s = \log_2(\frac{1}{\Delta(P)}) + \mathrm{O}(1)$, and satisfies the following properties:*

1. *The circuit $R(P)$ uses $2ks$ Hadamard gates, $\mathrm{O}(ks^2)$ controlled phase rotations, and makes at most $k\,2^{s+1}$ calls to the controlled quantum walk $\mathrm{c}{-}W(P)$, and its inverse $\mathrm{c}{-}W(P)^\dagger$.*
2. *If $|\pi\rangle$ is the unique 1-eigenvector of $W(P)$ as defined above, then $R(P)|\pi\rangle|0^{ks}\rangle = |\pi\rangle|0^{ks}\rangle$.*
3. *If $|\psi\rangle$ lies in the subspace of $\mathcal{A}+\mathcal{B}$ orthogonal to $|\pi\rangle$, then $\|(R(P)+\mathrm{Id})|\psi\rangle|0^{ks}\rangle\| \le 2^{1-ck}$.*

Moreover the family of circuits $R(P)$ parametrized by n and k is uniform.

PROOF. We describe the circuit $R(P)$. We start by applying the phase estimation theorem to the quantum walk $W(P)$, a unitary operator of dimension $m \times m$. Then using standard methods, we amplify the success probability of the resulting algorithm. We repeat it k times and take the median of the answers, increasing the success probability to $1-2^{-ck}$, for some constant $c>0$. Observe that only the number of ancillary qubits increase from s to ks in this process; we do not need additional copies of the eigenvector $|\psi\rangle$.

This approximately resolves any state $|\psi\rangle$ in $\mathcal{A}+\mathcal{B}$ along the eigenvectors of $W(P)$ by labeling them with the corresponding eigenvalue phases. We now flip the phase of all states with a non-zero estimate of the phase, that is flip the phase of all eigenvectors other than $|\pi\rangle$. Finally, we reverse the phase estimation.

The state $|\pi\rangle|0^{ks}\rangle$ stays unchanged. When $|\psi\rangle$ is orthogonal to $|\pi\rangle$ then it is the linear combination of eigenvectors of $W(P)$ whose eigenvalues are of the form $\mathrm{e}^{\pm 2i\theta}$, where $\Delta(P)/2 \le \theta < \pi/2$. Since $2\theta/2\pi > \Delta(P)/8$, the $(\log_2(1/\Delta(P))+3)$-bit phase estimation is non-zero with probability at least $1-2^{-ck}$. We can therefore decompose $|\psi\rangle|0^{ks}\rangle$ into the sum $|\psi_0\rangle+|\psi_1\rangle$, where the phase estimation is zero on $|\psi_0\rangle$, non-zero on $|\psi_1\rangle$, and $\||\psi_0\rangle\| \le 2^{-ck}$. Then $R(P)|\psi\rangle|0^{ks}\rangle = |\psi_0\rangle - |\psi_1\rangle$, and $(R(P)+\mathrm{Id})|\psi\rangle|0^{ks}\rangle = 2|\psi_0\rangle$ whose norm is at most 2^{1-ck}. □

3.3 The search algorithm for reversible Markov chains

Let us consider the following quantum procedure.

Quantum Search(P)
1. Start from the initial state $|\pi\rangle_d$.
2. Repeat $\mathrm{O}(1/\sqrt{\varepsilon})$–times:
   (a) For any basis vector $|x\rangle_d|y\rangle_d$, flip the phase if $x \in M$.
   (b) Apply circuit $R(P)_d$ of Theorem 6 with $k = \mathrm{O}(\log(1/\sqrt{\varepsilon}))$.
3. Observe the first register.
4. Output x if $x \in M$, otherwise output 'no marked element exists'.

THEOREM 7. *Let $\delta > 0$ be the eigenvalue gap of a reversible, ergodic Markov chain P, and let $\varepsilon > 0$ be a lower bound on the probability that an element chosen from the stationary distribution of P is marked whenever M is non-empty. Then, with high probability, the procedure* ***Quantum Search***(P) *determines if M is empty or finds an element of M with cost of order* $\mathsf{S} + \frac{1}{\sqrt{\varepsilon}}\left[\left(\frac{1}{\sqrt{\delta}}\log\frac{1}{\sqrt{\varepsilon}}\right)\mathsf{U}+\mathsf{C}\right]$.

PROOF. For the mathematical analysis we reason in the Hilbert space $\mathcal{H}$, without the data structures, and also suppress the ancilla qubits used by the circuit $R(P)$. We know that in the two-dimensional subspace $\mathcal{S} = \mathsf{Span}(|\pi\rangle,\ |\mu\rangle)$, the Grover algorithm, consisting in $\mathrm{O}(1/\sqrt{\varepsilon})$ iterations of $\mathrm{ref}(\pi)\cdot\mathrm{ref}(\mu^\perp)$, turns the vector $|\pi\rangle$ into a state whose inner product with $|\mu\rangle$ is $1-\mathrm{O}(\sqrt{\epsilon})$. Using a hybrid argument as in [8, 26], we prove that the algorithm **Quantum Search**(P) simulates, with an arbitrarily small error, the Grover algorithm, and therefore finds a marked element with high probability, whenever such an element exists.

For $i \ge 0$, we define $|\phi_i\rangle$ as the result of i Grover iterations applied to $|\pi\rangle$, and $|\psi_i\rangle$ as the result of i iterations of step (2) in **Quantum Search**(P) applied to $|\pi\rangle$. We show by induction on i, that $\||\psi_i\rangle - |\phi_i\rangle\| \le i2^{1-ck}$, where c is the constant of Theorem 6. Indeed, we can write $|\psi_i\rangle$ as $|\phi_i\rangle + (|\psi_i\rangle - |\phi_i\rangle)$. The action of $\mathrm{ref}(\mu^\perp)$ and $-\mathrm{ref}(\mathcal{M})$ are identical on $|\phi_i\rangle$ since it is in $\mathcal{S}$. Set $|\tau\rangle = |\phi_{i+1}\rangle + R(P)\cdot\mathrm{ref}(\mathcal{M})|\phi_i\rangle$. Since $\mathrm{ref}(\mathcal{M})|\phi_i\rangle$ is in $\mathcal{S}$, and $\mathcal{S}$ is a subspace of $\mathcal{A}+\mathcal{B}$, conclusion (3) of Theorem 6 can be applied, which implies that $\||\tau\rangle\| \le 2^{1-ck}$. Using that $\||\psi_{i+1}\rangle - |\phi_{i+1}\rangle\| \le \||\tau\rangle\| + \||\psi_i\rangle - |\phi_i\rangle\|$, the statement follows. This implies that $\||\psi_k\rangle - |\phi_k\rangle\|$ is an arbitrarily small constant when k is chosen to be $\mathrm{O}(c^{-1}\log(1/\sqrt{\varepsilon}))$.

Let us now turn to the cost of the procedure. Initialization costs $\mathsf{S}+\mathsf{U}$, and in each iteration the single phase flip costs C. In the circuit $R(P)_d$, the controlled quantum walk and its inverse can be implemented with four update operations, each of cost U. Indeed, the implementation of $W(P)$, described in the proof of Proposition 2 works also for the controlled quantum walk if we replace $\mathrm{ref}(\mathcal{H}_d\otimes|\bar{0}\rangle_d)$ by the controlled operator $\mathrm{c}{-}\mathrm{ref}(\mathcal{H}_d\otimes|\bar{0}\rangle_d)$. Since the controlled reflection is also of unit cost, this change does not alter the cost of the implementation.

In $R(P)_d$ the number of controlled quantum walks and its inverse is $\mathrm{O}((1/\Delta(P))\log(1/\sqrt{\epsilon}))$. We claim that $\Delta(P) = \Omega(\sqrt{\delta})$. Let $\lambda_0, ..., \lambda_{n-1}$ be the eigenvalues, with multiplicity, of P such that $1=\lambda_0 > |\lambda_1| \ge \ldots \ge |\lambda_{n-1}|$. Since the discriminant $D(P)$ is similar to P, their spectra are the same, and therefore the singular values of $D(P)$ are $|\lambda_0|, |\lambda_1|, \ldots$. By definition, $\Delta(P) = 2\theta_1$, where $\cos\theta_1 = |\lambda_1|$. The following (in)equalities can easily be checked: $\Delta(P) \ge \left|1-\mathrm{e}^{2i\theta_1}\right| = 2\sqrt{1-|\lambda_1|^2} \ge 2\sqrt{\delta(P)}$. This finishes the cost analysis. □

Let us observe that the origin of the quadratic speed-up due to quantum walks may be traced to the quadratic relationship between the phase gap $\Delta(P)$ of the quantum walk $W(P)$ and the eigenvalue gap δ of the classical Markov chain P, observed at the end of the previous proof.

4. SEARCH WITH APPROXIMATE REFLECTION OPERATORS

In this section, we describe how our approximate reflection operator may be incorporated into a search algorithm without incurring additional cost for reducing its error. The basic idea is to adapt the recursive amplitude amplification algorithm of Høyer, Mosca, and de Wolf [17] to our setting. To describe it, we use the notation from Section 3.1 where we discussed how the Grover algorithm works to rotate a starting state $|\pi\rangle$ into a target state $|\mu\rangle$, where $\langle\mu|\pi\rangle = \sin\varphi = \sqrt{\varepsilon}$. We define procedures A_i recursively. Let the procedure A_0 be the identity map Id, and for $i>0$, let

$$A_i \quad = \quad A_{i-1}\cdot\mathrm{ref}(\pi)\cdot A_{i-1}^\dagger\cdot\mathrm{ref}(\mu^\perp)\cdot A_{i-1}.$$

We define the states $|\pi_i\rangle$ as $A_i|\pi\rangle$. Then $|\pi_i\rangle$ forms an angle $3^i\varphi$ with $|\mu^\perp\rangle$, and therefore the state $|\pi_t\rangle$ is close to $|\mu\rangle$ when $t = \log_3 \frac{1}{\varphi} + \mathrm{O}(1)$. The final recursive algorithm is thus A_t.

We may estimate the cost $\mathsf{Cost}(t)$ of this search algorithm in terms of the cost c of implementing the two original reflections. We have $\mathsf{Cost}(0) = 0$, and for $i \geq 1$, $\mathsf{Cost}(i) = 3 \cdot \mathsf{Cost}(i-1) + c$, and therefore the cost of A_t is $\mathrm{O}(c/\sqrt{\varepsilon})$.

The recursive amplitude amplification algorithm is more suitable for situations where we have imperfect procedures that implement the basic reflections $\mathrm{ref}(\pi), \mathrm{ref}(\mu^\perp)$. Høyer *et al.* [17] demonstrated this when there is an ideal (error-free) procedure for $\mathrm{ref}(\pi)$, and a procedure for $\mathrm{ref}(\mu^\perp)$ that has ideal behaviour only with high probability. We extend this to the case where the first reflection may also be imperfect. In the context of quantum walk based search, an imperfection appears in the form given by Theorem 6. The basic idea is to create an analogue of the recursive algorithms A_i when $\mathrm{ref}(\pi)$ is replaced by increasingly fine approximations based on Theorem 6. For the sake of simplicity, we only deal with the case when the implementation of $\mathrm{ref}(\mu^\perp)$ is ideal.

THEOREM 8. *Let $|\pi\rangle$ be some state in a Hilbert space $\mathcal{H}$, let $\mathcal{M} \subseteq \mathcal{H}$ be a subspace of $\mathcal{H}$, and let $\varepsilon = \|\Pi_{\mathcal{M}}|\pi\rangle\|^2$, where $\Pi_{\mathcal{M}}$ is the projector onto $\mathcal{M}$. Assume that for any $\beta > 0$, we have a quantum circuit $R(\beta)$ acting on $\mathcal{H} \otimes \mathcal{K}$, where $\mathcal{K}$ is an extra register of s qubits ($s = s(\beta)$ may depend on β), with the following properties:*

1. *The circuit $R(\beta)$ has a cost of order $\mathrm{O}(c_1 \log \frac{1}{\beta})$.*
2. *$R(\beta)|\pi\rangle|0^s\rangle = |\pi\rangle|0^s\rangle$.*
3. *$\|(R(\beta)+\mathrm{Id})|\psi\rangle|0^s\rangle\| \leq \beta$ when $|\psi\rangle$ is orthogonal to $|\pi\rangle$.*

Further, assume that we are able to apply $-\mathrm{ref}(\mathcal{M})$ with a cost $\mathrm{O}(c_2)$. Then, for any $\gamma > 0$, there exists a quantum circuit that maps $|\pi\rangle$ to a state that has projection of length at least $\frac{1}{\sqrt{2}} - \gamma$ in $\mathcal{M}$, and incurs a cost of order $\frac{1}{\sqrt{\varepsilon}} \cdot (c_1 \log \frac{1}{\gamma} + c_2)$.

PROOF. Let t be the smallest non-negative integer such that $3^t \sin^{-1}\sqrt{\epsilon} \in [\pi/4, 3\pi/4]$. We have $t = \log_3 \frac{1}{\sqrt{\varepsilon}} + \mathrm{O}(1)$. The quantum circuit consists in t recursive amplitude amplification steps and acts on $\mathcal{H} \otimes \left[\bigotimes_{i=1}^t \mathcal{K}_i\right]$, where $\mathcal{K}_i$ is an extra register used at step i. Let $s_i = s(\beta_i)$ be the size of register $\mathcal{K}_i$, where β_i will be specified later. Let $S = \sum_{i=1}^t s_i$.

The quantum circuit follows exactly the recursive amplitude amplification algorithm explained above. We use $|\phi_0\rangle = |\phi\rangle = |\pi\rangle|0^S\rangle$ as the initial state and essentially replace $\mathrm{ref}(\phi)$ at step i by an approximation $R(\beta_i)$, acting on $\mathcal{H} \otimes \mathcal{K}_i$, and Id on the rest. The precise definition of the algorithm A_i in terms of the reflection operator R_i used in this recursive step is as follows. Let $A_0 = \mathrm{Id}$. For $i \geq 1$:

Approximate reflection R_i
The algorithm acts on $\mathcal{H} \otimes \left[\bigotimes_{j=1}^t \mathcal{K}_j\right]$.

1. Apply $A_{i-1}^\dagger$.
2. If the registers corresponding to $\bigotimes_{j=1}^t \mathcal{K}_j$ are in state $|0^s\rangle$, apply $R(\beta_i)$ on $\mathcal{H} \otimes \mathcal{K}_i$.
3. Otherwise apply $-\mathrm{Id}$.
4. Apply A_{i-1}.

Further, define

$$\begin{aligned} A_i &= R_i \cdot \mathrm{ref}(\mathcal{M}) \cdot A_{i-1} \\ |\phi_i\rangle &= A_i\,|\phi_0\rangle \end{aligned}$$

Since $R(\beta_i)$ is an approximation to $\mathrm{ref}(\phi_0)$, we see that R_i is an approximation to $\mathrm{ref}(\phi_{i-1})$. Let $E_i = R_i - \mathrm{ref}(\phi_{i-1})$ be the error made in our implementation of $\mathrm{ref}(\phi_{i-1})$.

FACT 1. *E_i satisfies the following properties:*
1. *$E_i|\phi_{i-1}\rangle = 0$, and*
2. *$\|E_i|\psi\rangle\| \leq \beta_i$, for all $|\psi\rangle \perp |\phi_{i-1}\rangle$ within the subspace $(\mathcal{A}+\mathcal{B}) \otimes \left[\bigotimes_{j=1}^t \mathcal{K}_j\right]$.*

To analyze this algorithm, we keep track at all steps of the projection of $|\phi_i\rangle$ on the marked subspace. The marked subspace corresponds to $\mathcal{M} \otimes \left[\bigotimes_j \mathcal{K}_j\right]$; it consists of states in which the $\mathcal{H}$-part is marked. For ease of notation, we will denote this space by $\mathcal{M}$. Define the normalized projections of $|\phi_i\rangle$ on the marked subspace $\mathcal{M}$ and on its orthogonal complement as:

$$\begin{aligned} |\mu_i\rangle &= \frac{\Pi_{\mathcal{M}}|\phi_i\rangle}{\|\Pi_{\mathcal{M}}|\phi_i\rangle\|} \\ |\mu_i^\perp\rangle &= \frac{(\mathrm{Id}-\Pi_{\mathcal{M}})|\phi_i\rangle}{\|(\mathrm{Id}-\Pi_{\mathcal{M}})|\phi_i\rangle\|}. \end{aligned}$$

We thus have

$$|\phi_i\rangle = \sin\varphi_i\,|\mu_i\rangle + \cos\varphi_i\,|\mu_i^\perp\rangle, \tag{1}$$

where $\sin^2\varphi_i = \|\Pi_{\mathcal{M}}|\phi_i\rangle\|^2$ is the success probability at step i, that is the probability of finding a marked item by measuring the first register according to $\{\Pi_{\mathcal{M}}, \mathrm{Id}-\Pi_{\mathcal{M}}\}$. For later use, let us also define $|\phi_i^\perp\rangle$ as the state in the 2-dimensional subspace spanned by $|\mu_i\rangle$ and $|\mu_i^\perp\rangle$ that is orthogonal to $|\phi_i\rangle$:

$$|\phi_i^\perp\rangle = \cos\varphi_i\,|\mu_i\rangle - \sin\varphi_i\,|\mu_i^\perp\rangle.$$

For the initial state $|\phi_0\rangle$, we have $\sin^2\varphi_0 = \varepsilon$, by the hypothesis of the theorem. If all the errors β_i were zero, our recursive algorithm would implement an amplitude amplification in the subspace spanned by $|\mu_i\rangle = |\mu_0\rangle$ and $|\mu_i^\perp\rangle = |\mu_0^\perp\rangle$, with the angles $\varphi_{i+1} = 3\varphi_i$, that is $\varphi_i = 3^i\varphi_0$. Therefore by recursively iterating our procedure for a total number of t steps, we would end up with a state whose inner product with $|\mu_0\rangle$ is at least $\frac{1}{\sqrt{2}}$.

Analysis of the errors — We show that the algorithm still works when the errors β_i are sufficiently small. In that case, the 2-dimensional subspace $\mathsf{Span}(|\mu_i\rangle, |\mu_i^\perp\rangle)$ may drift away from the initial subspace $\mathsf{Span}(|\mu_0\rangle, |\mu_0^\perp\rangle)$, and the angles φ_i may be different from the ideal value $\bar{\varphi}_i = 3^i\varphi_0$. We derive bounds on the error e_i:

$$e_i = |\sin\varphi_i - \sin\bar{\varphi}_i|\,, \tag{2}$$

the difference between the amplitude $\sin\varphi_i$ of the marked part of the state $|\phi_i\rangle$ and the ideal amplitude, $\sin\bar{\varphi}_i$.

We assume without loss of generality that $0 < \gamma \leq \frac{1}{\sqrt{2}}$ since the case $\gamma > \frac{1}{\sqrt{2}}$ is vacuous. We prove that after t steps $e_t \leq \gamma$. This will conclude the error analysis since $\frac{1}{\sqrt{2}} \leq \sin\bar{\varphi}_t \leq 1$.

We have

$$\begin{aligned} |\phi_{i+1}\rangle &= R_{i+1} \cdot \mathrm{ref}(\mathcal{M})\,|\phi_i\rangle \\ &= \mathrm{ref}(\phi_i) \cdot \mathrm{ref}(\mathcal{M})\,|\phi_i\rangle + E_{i+1} \cdot \mathrm{ref}(\mathcal{M})\,|\phi_i\rangle \\ &= \sin 3\varphi_i\,|\mu_i\rangle + \cos 3\varphi_i\,|\mu_i^\perp\rangle + |\omega_{i+1}\rangle, \end{aligned} \tag{3}$$

where we used the fact that $\mathrm{ref}(\phi_i)\cdot\mathrm{ref}(\mathcal{M})$ implements a perfect amplitude amplification step, and we introduced an error state $|\omega_i\rangle$, defined as

$$\begin{aligned}|\omega_{i+1}\rangle &= E_{i+1}\cdot\mathrm{ref}(\mathcal{M})\,|\phi_i\rangle\\ &= E_{i+1}\cdot\mathrm{ref}(\mathcal{M})\left(\sin\varphi_i\,|\mu_i\rangle+\cos\varphi_i\,|\mu_i^\perp\rangle\right)\\ &= E_{i+1}\left(-\sin\varphi_i\,|\mu_i\rangle+\cos\varphi_i\,|\mu_i^\perp\rangle\right)\\ &= E_{i+1}\left(\cos 2\varphi_i\,|\phi_i\rangle-\sin 2\varphi_i\,|\phi_i^\perp\rangle\right)\\ &= -\sin 2\varphi_i\;E_{i+1}\,|\phi_i^\perp\rangle,\end{aligned}$$

where we used Fact 1, property 1. Moreover, $|\phi_i^\perp\rangle\perp|\phi_i\rangle$, so $\||\omega_{i+1}\rangle\|\le\beta_{i+1}\sin 2\varphi_i$. Finally, comparing Eq. (1) and Eq. (3), we get

$$|\sin\varphi_{i+1}-\sin 3\varphi_i| \;\le\; \beta_{i+1}\sin 2\varphi_i.$$

We may now bound the error defined in Eq. (2) as:

$$\begin{aligned}e_{i+1} &\le |\sin\varphi_{i+1}-\sin 3\varphi_i|+|\sin 3\varphi_i-\sin\bar\varphi_{i+1}|\\ &\le \beta_{i+1}\sin 2\varphi_i+|\sin 3\varphi_i-\sin 3\bar\varphi_i|\\ &\le \beta_{i+1}(\sin 2\bar\varphi_i+|\sin 2\varphi_i-\sin 2\bar\varphi_i|)\\ &\qquad +\;|\sin 3\varphi_i-\sin 3\bar\varphi_i|\\ &\le \beta_{i+1}(\sin 2\bar\varphi_i+2e_i)+3e_i\\ &\le 2\beta_{i+1}(\bar\varphi_i+e_i)+3e_i,\end{aligned}$$

where we have used the triangle inequality and the following trigonometric inequalities

$$\begin{aligned}|\sin 2A-\sin 2B| &\le 2|\sin A-\sin B|\\ |\sin 3A-\sin 3B| &\le 3|\sin A-\sin B|\\ \sin A &\le A\end{aligned}$$

that hold for any angles $A,B\in[0,\pi/4]$.

We define a quantity $\tilde e_i$, intended to be an upper bound on e_i (it would be if $\tilde e_i\le\bar\varphi_i$). Let

$$\begin{aligned}\tilde e_0 &= 0\\ \tilde e_{i+1} &= 4\beta_{i+1}\bar\varphi_i+3\tilde e_i.\end{aligned}$$

We show that $\tilde e_i\le\gamma$ for every $i\le t$, if we take $\{\beta_i\}$ so that they define a convergent series. Indeed, let us define u_i as

$$\tilde e_i \;=\; \gamma\,\bar\varphi_i\,u_i.$$

We therefore have the following recursion for u_i

$$\begin{aligned}u_0 &= 0\\ u_{i+1} &= u_i+\frac{4}{3\gamma}\beta_{i+1},\qquad(\forall i\ge 0)\end{aligned}$$

so that

$$u_i \;=\; \frac{4}{3\gamma}\sum_{j=1}^{i}\beta_j.$$

Since we assumed that $\{\beta_i\}$ define a convergent series, the non-decreasing sequence (u_i) tends to some real u_∞ when $i\to\infty$. We therefore have $\tilde e_i\le\gamma u_\infty\varphi_i$.

We now set $\beta_i=\frac{18}{4\pi^3}\gamma/i^2$. Then $u_\infty=1/\pi$ and we have $\tilde e_i\le\gamma\bar\varphi_i/\pi\le\gamma$ since $0\le\bar\varphi_t\le\pi$ for $i\le t$.

Since $0<\gamma\le 1$, we have $\tilde e_i\le\bar\varphi_i$, and we can show by induction that $e_i\le\tilde e_i$ for all $i\le t$. This finishes the error analysis.

Complexity — We now evaluate the complexity of our algorithm. We know from the hypotheses of the theorem that applying $R(\beta_i)$ has a cost $c_1\log\frac{1}{\beta_i}$, while applying $\mathrm{ref}(\mathcal{M})$ has a cost c_2. Moreover, we see from Eq. (1) that applying A_i requires 3 calls to A_{i-1}, one call to $R(\beta_i)$ and one call to $\mathrm{ref}(\mathcal{M})$. Hence, if we denote $\mathsf{Cost}(i)$ the cost of applying A_i, we have

$$\begin{aligned}\mathsf{Cost}(0) &= 0\\ \mathsf{Cost}(i) &= 3\,\mathsf{Cost}(i-1)+c_1\log\frac{1}{\beta_i}+c_2.\end{aligned}$$

Since we have fixed $\beta_i=\frac{18}{4\pi^2}\gamma/i^2$, we find that $\mathsf{Cost}(i)$ equals

$$\begin{aligned}&\left(c_1\left(2\log\frac{1}{\gamma}+\mathrm{O}(1)\right)\right)\sum_{j=1}^{i}3^{i-j}(\log j)+c_2\sum_{j=1}^{i}3^{i-j}\\ &= 3^i\left(\left(c_1\left(2\log\frac{1}{\gamma}+\mathrm{O}(1)\right)\right)\sum_{j=1}^{i}\frac{\log j}{3^j}+c_2\sum_{j=1}^{i}\frac{1}{3^j}\right)\end{aligned}$$

where both sums converge as $i\to\infty$. After t steps we have $\mathsf{Cost}(t)=\mathrm{O}\left(\frac{1}{\sqrt{\varepsilon}}\cdot(c_1\log\frac{1}{\gamma}+c_2)\right)$.

If the cost refers to time complexity, then there is an additional term pertaining to the reflection R_i. This arises from the check to see if the ancilla used are in state $|0^S\rangle$. This does not change the asymptotic complexity of the algorithm. □

We now have all the elements to prove Theorem 3.

Proof of Theorem 3. The algorithm consists in the recursive quantum search from Theorem 8, using for the approximate reflections $R(\beta)$ the quantum phase estimation circuit $R(P)$ from Theorem 6. Setting $k=\left\lceil((\log\frac{1}{\beta})+1)/c\right\rceil$ in Theorem 6, $R(P)$ simulates a reflection with an error upper bounded by $2^{1-ck}\le\beta$. Implementing $R(P)_d$ then requires $k\,2^{s+1}$ calls to the controlled quantum walk $\mathrm{c}{-}W(P)_d$ or its inverse, where $s=\log(\frac{1}{\sqrt{\delta}})+\mathrm{O}(1)$. Since implementing $\mathrm{c}{-}W(P)_d$ or its inverse has a cost $4\mathsf{U}$, the cost of implementing the circuit $R(P)_d$ for a given error β is $c_1\log\frac{1}{\beta}$, with c_1 of order $\frac{1}{\sqrt{\delta}}\mathsf{U}$. Furthermore, preparing the initial state $|\pi\rangle_d$ has a cost $\mathsf{S}+\mathsf{U}$, and implementing $-\mathrm{ref}(\mathcal{M})_d$ has a cost $c_2=\mathsf{C}$. Therefore total cost of the recursive search algorithm from Theorem 8 is of order $\mathsf{S}+\frac{1}{\sqrt{\varepsilon}}(\frac{1}{\sqrt{\delta}}\mathsf{U}+\mathsf{C})$. □

5. NON-REVERSIBLE MARKOV CHAINS

In this section, we discuss the performance of the search algorithm presented earlier for any ergodic, but possibly non-reversible Markov chain P. For the analysis of the quantum walk $W(P)$ we directly examine the singular value decomposition of the discriminant matrix $D(P)=\mathrm{diag}(\pi)^{1/2}\cdot P\cdot\mathrm{diag}(\pi)^{-1/2}$. This matrix has the same eigenvalues as P, but the singular values of $D(P)$ may be different from the eigenvalues of P. The singular values of $D(P)$ lie in the interval $[0,1]$. The vector $v=(\sqrt{\pi_x})$ is both a left and a right eigenvector of $D(P)$ with eigenvalue 1. Therefore, $\mathsf{Span}(v)$ and $\mathsf{Span}(v)^\perp$ are invariant subspaces of $D(P)$, and we may choose v to be a left and right singular vector. If every singular vector orthogonal to v has a singular value strictly smaller than 1, that is $D(P)$ has a non-zero singular value gap, then Theorem 3 and its proof stay valid when the eigenvalue gap of P is replaced by the singular value gap of $D(P)$.

The discriminant of an irreducible walk does not necessarily have non-zero singular value gap, even if it is ergodic. Ergodicity implies a non-zero eigenvalue gap for P, but there are examples of ergodic Markov chain whose discriminant has 0 singular value gap. In the next proposition we show that if for every state, the Markov has a transition to itself with non-zero probability, then its discriminant has non-zero singular value gap. There is a standard and simple modification to any Markov chain P such that the resulting chain has this property: with some probability $\alpha \in (0,1)$, stay at the current state, and with probability $1-\alpha$, make a transition according to P.

PROPOSITION 3. *Let $P = (p_{xy})$ be an irreducible Markov chain on a finite state space X, such that $p_{xx} > 0$, for every $x \in X$. Then, the discriminant matrix $D(P)$ has exactly one singular value equal to 1.*

We first state and prove that all the singular values of $D(P)$ lie in $[0,1]$.

LEMMA 1. *Let $P = (p_{xy})_{x,y\in X}$ be an irreducible Markov chain with stationary distribution $\pi = (\pi_x)_{x\in X}$. Then the singular values of the matrix $D(P)$ given by*

$$D(P) \quad = \quad \mathrm{diag}(\pi)^{1/2} \cdot P \cdot \mathrm{diag}(\pi)^{-1/2}$$

all lie in the interval $[0,1]$.

PROOF. Singular values are by convention taken to be non-negative real. To verify that $\|D(P)\|$, the largest singular value of $D(P)$ is at most 1, consider the inner product $u^\dagger D(P) v$, for some unit vectors u, v. The maximum absolute value that this inner product achieves is the norm of $D(P)$. By the Cauchy-Schwartz inequality, the inner product may be bounded as

$$\begin{aligned} & \left| u^\dagger D(P) v \right| \\ &= \left| \sum_{xy} \bar{u}_x v_y \sqrt{\frac{\pi_x}{\pi_y}}\, p_{xy} \right| \\ &\leq \left(\sum_{xy} |u_x|^2 p_{xy} \right)^{1/2} \left(\sum_{xy} |v_y|^2 \frac{\pi_x}{\pi_y} p_{xy} \right)^{1/2} \qquad (4) \\ &\leq 1, \end{aligned}$$

since $\sum_x \pi_x p_{xy} = \pi_y$. □

PROOF OF PROPOSITION 3. From Lemma 1, we know that the singular values of $D(P)$ all lie in $[0,1]$. Further $v = (\sqrt{\pi_x})$ is a left (and right) singular vector with singular value 1. We show below that for any left and right singular vectors $u, w \in \mathbb{C}^X$, if $u^\dagger D(P) w = 1$, then $u = w = v$ (modulo an overall phase). This establishes the uniqueness of the singular value 1 and a non-zero singular value gap in $D(P)$.

Suppose $u^\dagger D(P) w = 1$. This implies that the Cauchy-Schwartz inequality in Equation (4) in the proof of Lemma 1 is tight. Then necessarily, the two unit vectors $u', w' \in \mathbb{C}^{X\times X}$ given by $u' = (u_x \sqrt{p_{xy}})_{x,y\in X}$ and $w' = (w_y \sqrt{\pi_x p_{xy}/\pi_y})_{x,y\in X}$ are equal. This means that for every pair $x, y \in X$ such that $p_{xy} > 0$, $u_x = w_y \sqrt{\pi_x/\pi_y}$. In particular, since $p_{xx} > 0$, $u_x = w_x$ for every x, and so $u_x = u_y \sqrt{\pi_x/\pi_y}$ for every neighbour y of x in the graph underlying the Markov chain P.

Furthermore, for any path $x_1, x_2, \ldots, x_k$ in the graph, chaining together the equations

$$u_{x_{i+1}} \quad = \quad u_{x_i} \sqrt{\frac{\pi_{x_{i+1}}}{\pi_{x_i}}},$$

for $i = 1, \ldots, k-1$, we get that

$$u_{x_i} \quad = \quad u_{x_1} \sqrt{\frac{\pi_{x_i}}{\pi_{x_1}}},$$

for every i. Since the chain P is irreducible, i.e., the underlying graph is strongly connected, there is path from x_1 to y for every $y \in X$. Thus,

$$u_y \quad = \quad u_{x_1} \sqrt{\frac{\pi_y}{\pi_{x_1}}},$$

for every y. Since the vector u is a unit vector, this implies that $u = w = (\sqrt{\pi_x})_{x\in X} = v$ (up to an unimportant global phase). □

Let us finally state the theorem on our quantum search algorithm when the underlying Markov chain is not necessary reversible.

THEOREM 9. *Let $P = (p_{xy})$ be an irreducible Markov chain on a finite state space X, such that $D(P)$ has exactly one singular value equal to 1. Let $\delta > 0$ be the singular value gap of $D(P)$, and let $\varepsilon > 0$ be a lower bound on the probability that an element chosen from the stationary distribution of P is marked whenever M is non-empty. Then, there is a quantum algorithm that determines, with high probability, if M is empty or finds an element of M with cost of order $\mathsf{S} + \frac{1}{\sqrt{\varepsilon}}(\frac{1}{\sqrt{\delta}}\mathsf{U} + \mathsf{C})$.*

6. CONCLUSION

We proposed a new method for designing quantum search algorithms for finding a "marked" element in the state space of a classical Markov chain. Our algorithm expands the scope of earlier work due to Ambainis [6] and Szegedy [25] and combines their advantages into a unified approach. An upshot is the improvement in various aspects of many known algorithms based on quantum walk. The method is also conceptually and technically simpler, and thus seems to demystify the role of quantum walks in search algorithms.

Theorem 2 of Szegedy may be strengthened to give a quadratic relationship between a notion of quantum hitting time and the classical hitting time [25] (cf. the remarks at the end of Section 1.3). Our schema trades this relationship with classical hitting time for a form of run time that is more conducive to composition of algorithms. As a result, a direct application of our schema to search in special classes of graph such as the d-dimensional toroidal grid (for any constant d) leads to algorithms that are asymptotically slower than known ones [2, 5, 14, 25]. We believe that these results could be recast in our framework, but leave it open for the moment.

7. ACKNOWLEDGMENTS

A part of this work was done while the authors were visiting Institut Henri Poincaré, Paris, France, during the Programme on Quantum Information, Computation, and Complexity, January-April 2006.

F. M., J. R., and M. S. were in part supported by the European Commission IST project QAP 015848 and by the ANR Blanc AlgoQP grant of the French Research Ministry. A. N. was in part supported by NSERC Canada, CIAR, MITACS, CFI, and OIT (Canada) and ARO/NSA (USA). Research at Perimeter Institute for Theoretical Physics is supported in part by the Government of Canada through NSERC and by the Province of Ontario through MRI. J. R. acknowledges support from NSF Grant CCF-0524837 and ARO Grant DAAD 19-03-1-0082, and during this work he was supported by the Belgian FNRS as a Postdoctoral Researcher at LRI, Université Paris XI, and QuIC, Université Libre de Bruxelles.

8. REFERENCES

[1] S. Aaronson and Y. Shi. Quantum lower bounds for the collision and the element distinctness problems. *Journal of the ACM*, pages 595–605, 2004.

[2] S. Aaronson and A. Ambainis. Quantum search of spatial regions. *Theory of Computing*, 1:47–79, 2005.

[3] D. Aharonov, A. Ambainis, J. Kempe, and U. Vazirani. Quantum walks on graphs. In *Proceedings of the 33rd ACM Symposium on Theory of computing*, pages 50–59, 2001.

[4] D. Aldous and J. Fill. *Reversible Markov Chains and Random Walks on Graphs*. http://www.stat.berkeley.edu/~aldous/RWG/book.html. Monograph in preparation, August 2006 version.

[5] A. Ambainis, J. Kempe, and A. Rivosh. Coins make quantum walks faster. In *Proceedings of the 16th ACM-SIAM symposium on Discrete algorithms*, pages 1099–1108, 2005.

[6] A. Ambainis. Quantum Walk Algorithm for Element Distinctness. In *Proceedings of the 45th Symposium on Foundations of Computer Science*, pages 22–31, 2004.

[7] A. Ambanis, E. Bach, A. Nayak, A. Vishwanath, and J. Watrous. One-dimensional quantum walks. In *Proceedings of the 33rd ACM Symposium on Theory of computing*, pages 37–49, 2001.

[8] C. Bennett, E. Bernstein, G. Brassard, and U. Vazirani. Strengths and weaknesses of quantum computing. *SIAM Journal on Computing*, 26(5):1510–1523, 1997.

[9] M. Boyer, G. Brassard, P. Høyer, and A. Tapp. Tight bounds on quantum searching. *Fortschritte Der Physik*, 46(4-5):493–505, 1998.

[10] G. Brassard, P. Høyer, M. Mosca, and A. Tapp. Quantum amplitude amplification and estimation. In Jr. Samuel J. Lomonaco and Howard E. Brandt, editors, *Quantum Computation and Quantum Information: A Millennium Volume*, volume 305 of *Contemporary Mathematics Series*. American Mathematical Society, 2002.

[11] A. Broder and A. Karlin. Bounds on the cover time. *Journal of Theoretical Probability*, 2(1):101–120, 1989.

[12] H. Buhrman, C. Dürr, M. Heiligman, P. Høyer, M. Santha F. Magniez, and R. de Wolf. Quantum algorithms for element distinctness. *SIAM Journal of Computing*, 34(6):1324–1330, 2005.

[13] H. Buhrman and R. Spalek. Quantum verification of matrix products. In *Proceedings of the 17th ACM-SIAM Symposium on Discrete Algorithms*, pages 880–889, 2006.

[14] A. Childs and J. Goldstone. Spatial search and the Dirac equation. *Physical Review A*, 70, 2004. Article No. 42312.

[15] R. Cleve, A. Ekert, C. Macchiavello, and M. Mosca. Quantum algorithms revisited. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 454(1969):339–354, 1998.

[16] L. Grover. A fast quantum mechanical algorithm for database search. In *Proceedings of the 28th ACM Symposium on the Theory of Computing*, pages 212–219, 1996.

[17] P. Høyer, M. Mosca, and R. de Wolf. Quantum search on bounded-error inputs. In Verlag, editor, *Proceedings of the 30th International Colloquium on Automata, Languages and Programming*, volume 2719 of *Lecture Notes in Computer Science*, pages 291–299. Verlag, 2003.

[18] A. Kitaev. Quantum measurements and the abelian stabilizer problem. Technical Report quant-ph/9511026, arXiv, 1995.

[19] F. Magniez and A. Nayak. Quantum complexity of testing group commutativity. In *Proceedings of the 32nd International Colloquium on Automata, Languages and Programming*, volume 1770 of *Lecture Notes in Computer Science*, pages 1312–1324, 2005.

[20] F. Magniez, M. Santha, and M. Szegedy. Quantum Algorithms for the Triangle Problem. In *Proceedings of the 16th ACM-SIAM Symposium on Discrete Algorithms*, 2005.

[21] D. Meyer. From quantum cellular automata to quantum lattice gases. *Journal of Statistical Physics*, 85(5-6):551–574, 1996.

[22] A. Nayak and A. Vishwanath. Quantum walk on the line. Technical Report quant-ph/0010117, arXiv, 2000.

[23] P. Richter. Almost uniform sampling via quantum walks. *New Journal of Physics*, 2007. To appear.

[24] N. Shenvi, J. Kempe, and K.B. Whaley. Quantum random-walk search algorithm. *Physical Review A*, 67(052307), 2003.

[25] M. Szegedy. Quantum Speed-Up of Markov Chain Based Algorithms. In *Proceedings of the 45th IEEE Symposium on Foundations of Computer Science*, pages 32–41, 2004.

[26] U. Vazirani. On the power of quantum computation. *Philosophical Transactions of the Royal Society of London, Series A*, 356:1759–1768, 1998.

[27] J. Watrous. Quantum simulations of classical random walks and undirected graph connectivity. *Journal of Computer and System Sciences*, 62(2):376–391, 2001.