

On the adiabatic condition and the quantum hitting time of Markov chains

Hari Krovi,^{1,*} Maris Ozols,^{1,2,†} and Jérémie Roland^{1,‡}

¹*NEC Laboratories America, Inc.*

²*University of Waterloo and Institute for Quantum Computing*

Abstract

We present an adiabatic quantum algorithm for the abstract problem of searching marked vertices in a graph, or spatial search. Given a random walk (or Markov chain) P on a graph with a set of unknown marked vertices, one can define a related absorbing walk P' where outgoing transitions from marked vertices are replaced by self-loops. We build a Hamiltonian $H(s)$ from the interpolated Markov chain $P(s) = (1-s)P + sP'$ and use it in an adiabatic quantum algorithm to drive an initial superposition over all vertices to a superposition over marked vertices. The adiabatic condition implies that for any reversible Markov chain and any set of marked vertices, the running time of the adiabatic algorithm is given by the square root of the classical hitting time. This algorithm therefore demonstrates a novel connection between the adiabatic condition and the classical notion of hitting time of a random walk. It also significantly extends the scope of previous quantum algorithms for this problem, which could only obtain a full quadratic speed-up for state-transitive reversible Markov chains with a unique marked vertex.

PACS numbers: 03.67.Lx 05.40.Fb

*Electronic address: krovi@engr.uconn.edu; Present affiliation: University of Connecticut

†Electronic address: marozols@yahoo.com

‡Electronic address: jroland@nec-labs.com

Introduction

Adiabatic quantum computation was introduced by Farhi *et al.* in [1]. It proceeds as follows. Suppose that the solution of a computational problem can be encoded in the ground state of a problem Hamiltonian H_P . We start in the ground state of an initial Hamiltonian H_0 which is easy to construct. Then we slowly change the Hamiltonian from H_0 to H_P , so that the instantaneous Hamiltonian at any point in the evolution is $H(s) = (1-s)H_0 + sH_P$, where $0 \leq s \leq 1$. If this is done slowly enough, then the Adiabatic Theorem of Quantum Mechanics [2] guarantees that the state at all points in the evolution stays close to the ground state of $H(s)$. Note that the validity of the *folk* version of the Adiabatic Theorem, as stated in many books of Quantum Mechanics such as [2], has recently been the subject of much debate [3–6], and there is no rigorous proof of it that holds under full generality. It is nevertheless possible to state a more rigorous version of the theorem, so that the *folk* adiabatic condition can be proved to be sufficient in many interesting cases [7], such as the adiabatic version of Grover’s algorithm [8–10].

Many classical randomized algorithms rely heavily on random walks or Markov chains. The notion of hitting time is a useful characterization of Markov chains used when searching for a marked vertex. Quantum walks are natural generalizations of classical random walks. The notion of hitting time has been carried over to the quantum case in [11–17], by generalizing the classical notion in different ways. It is intimately related to the problem of spatial search. Suppose that we are given a graph where some vertices are marked. Classically, a simple algorithm to find a marked vertex is to repeatedly apply some random walk P on the graph until one of the marked vertices is reached. The hitting time of P is precisely the expected number of repetitions necessary to reach a marked vertex, starting from the stationary distribution of P . The notions of quantum hitting time are based on different quantum analogues of this algorithm. They usually show some quadratic improvement of the quantum hitting time over the classical hitting time. However, until the present paper, they could only show such a relation under restricted conditions: either the quantum algorithm could only detect marked elements [13], or it could only be applied to state-transitive reversible Markov chains with a unique marked element [16]. Whether this quadratic speed-up for finding a marked element also holds for any reversible Markov chain and for multiple marked elements was an open question. In this paper, we answer this

question by the positive, by providing an adiabatic quantum algorithm for this problem. In addition to being more general, the algorithm is also conceptually very clean, it implements a simple rotation in a two dimensional subspace based on a quantum walk on the edges of the graph. Moreover, it reveals a close connection between the adiabatic condition and the notion of hitting time.

The paper is structured as follows. In Section I we describe related work and in Section II we state our main result. In Section III we introduce the necessary concepts such as Markov chains, the discriminant matrix, and the quantization of Markov chains. In Section IV we evaluate the spectrum of the interpolating Hamiltonian and in Section V we impose the adiabatic condition to calculate the running time of the adiabatic quantum algorithm. In Section VI we relate this to the hitting time of the Markov chain we started from, and show that the running time of the adiabatic evolution is at most the square root of the classical hitting time.

I. RELATED WORK

Inspired by Ambainis' quantum walk algorithm for Element Distinctness [18], Szegedy [13] introduced a powerful way of quantizing Markov chains which led to new quantum walk-based algorithms. He showed that for any symmetric Markov chain a quantum walk could detect the presence of marked vertices in at most the square root of the classical hitting time. However, showing that a marked vertex could also be found in the same time (as is the case for the classical algorithm) proved to be a very difficult task. Magniez *et al.* [15] extended Szegedy's approach to the larger class of ergodic Markov chains, and proposed a quantum walk-based algorithm to find a marked vertex, but its complexity may be larger than the square root of the classical hitting time. A typical example where their approach fails to provide a quadratic speed-up is the 2D grid, where their algorithm has complexity $\Theta(n)$, whereas the classical hitting time is $\Theta(n \log n)$. Ambainis *et al.* [11] and Szegedy's [13] approaches yield a complexity of $\Theta(\sqrt{n} \cdot \log n)$ in this special case, for a unique marked vertex. Childs and Goldstone [19, 20] also obtained a similar result using a continuous-time quantum walk. However, whether a full quadratic speed-up was possible remained an open question, until Tulsi [21] proposed a solution involving a new technique. Magniez *et al.* [16] extended Tulsi's technique to any reversible state-transitive Markov chain, showing that for

such chains, it is possible to find a unique marked vertex with a full quadratic speed-up over the classical hitting time. However, the state-transitivity is a strong symmetry condition, and furthermore their technique cannot deal with multiple marked vertices. It would be strange if one had to rely on involved techniques to solve the finding problem under such restricted conditions, while the classical analogue algorithm is conceptually very simple and works under very general conditions.

In this paper we show that these issues can be resolved by combining the idea of the quantization of Markov chains with adiabatic quantum computation (note that Childs and Goldstone [19, 20] showed that their algorithm for spatial search on the grid could also be translated into an adiabatic algorithm, but this failed to give a quadratic speed-up for low dimensions).

II. MAIN RESULT

Before describing our quantum algorithm, let us first recall the classical algorithm on which it will provide a quadratic speed-up. This very simple algorithm just consists in walking randomly on the graph until a marked vertex is reached. More precisely, it relies on a Markov chain P with stationary distribution π , and works as follows.

Random Walk Algorithm

1. Sample a vertex $x \in X$ according to distribution π .
2. Check if x is marked.
3. If so, output x .
4. Otherwise, update x according to P , and go back to step 2.

Let P be an ergodic Markov chain, and M be a set of marked vertices. The *hitting time* of P with respect to M , denoted by $\text{HT}(P, M)$, is the expected number of executions of step 4 during the course of the **Random Walk Algorithm** (where the expectation is calculated conditionally on the initial vertex being unmarked). Note that since the algorithm stops as soon as a marked element is reached, this is equivalent to using an *absorbing* Markov chain P' , which acts as P on all but marked vertices, where all outgoing transitions are replaced by self-loops.

Previous attempts at providing a quantum speed-up over this classical algorithm have followed one of these two approaches:

- Combining a quantum version of P with a reflection through marked vertices to mimic a Grover operation [11, 15, 18].
- Directly applying a quantum version of P' [13, 16].

The problem with these approaches is that they would only be able to find marked vertices in very restricted cases. We explain this by the different nature of random and quantum walks: while both accept a stable state, *i.e.*, the stationary distribution for the random walk and the eigenstate with eigenvalue 1 for the quantum walk, the way both walks act on other states is dramatically different. Indeed, an ergodic random walk will converge to its stationary distribution from any initial distribution. This apparent robustness may be attributed to the inherent randomness of the walk, which will smooth out any initial perturbation. After many iterations of the walk, non-stationary contributions of the initial distribution will be damped and only the stationary distribution will survive (this can be attributed to the thermodynamical irreversibility [26] of ergodic random walks). On the other hand, this is not true for quantum walks, because in the absence of measurements a unitary evolution is deterministic (and in particular thermodynamically reversible): the contributions of the other eigenstates will not be damped but just oscillate with different frequencies, so that the overall evolution is quasi-periodic. As a consequence, while iterations of P' always lead to a marked vertex, it may happen that iterations of the quantization of P' will never lead to a state with a large overlap over marked vertices, unless the walk exhibits a strong symmetry (as is the case for a state-transitive walk with only one marked element, which could be addressed by previous approaches).

The main new idea of our approach is that, instead of using a quantization of P or P' , we first modify the classical random walk, and then use a quantization of the modified walk. The original classical algorithm consists in applying P' on the stationary distribution π of P . While doing so, the system is brought far from equilibrium since π is far (in statistical distance) from any stationary distribution of P' , which only have support on marked elements. The random walk will damp any non-stationary contribution of the initial distribution, but a quantum walk based on P or P' , which is deterministic until a measurement, seems to have trouble with it. There is however a situation in Quantum Mechanics where contributions

from other eigenstates will also cancel out, similarly to what happens for a random walk: if the system starts in a state close to the ground state of its instantaneous Hamiltonian $H(s)$ (*i.e.*, close to equilibrium), and this Hamiltonian varies slowly, the Adiabatic Theorem ensures that the contributions from excited states will cancel out so that the system will remain close to its ground state. Therefore, our strategy is to first modify the classical algorithm so that the system stays close to equilibrium throughout the evolution, and then to translate it into an adiabatic quantum algorithm.

Consider the interpolated Markov chain $P(s) = (1 - s)P + sP'$ (see Section III A). Our goal is to drive the stationary distribution π of P towards a stationary distribution of P' . Instead of immediately applying P' on π , we could rather apply $P(s)$ while slowly switching s from 0 to 1, so that the system remains at all time close to the stationary distribution $\pi(s)$ of $P(s)$. It can be shown that this leads to an algorithm with only a constant overhead with respect to **Random Walk Algorithm**. Therefore, this new classical algorithm still runs in time $O(\text{HT}(P, M))$, but the difference is that at all time the system is close to equilibrium, so that we are in a better shape for designing a quantum analogue based on the Adiabatic Theorem.

Using a Hamiltonian version of Szegedy's quantization technique, proposed by Somma and Ortiz [22], we map $P(s)$ to a Hamiltonian $H(s)$ on the edge space $\mathcal{H} \otimes \mathcal{H}$, where \mathcal{H} is the Hilbert space whose basis states are labeled by the vertices of the graph (see Section III D). The eigenstate of $H(s)$ with eigenvalue 0 then corresponds to the stationary distribution of P for $s = 0$, and to a distribution over marked vertices for $s = 1$, so that this Hamiltonian may be used to solve the search problem by adiabatic evolution. The algorithm consists in preparing the first register in the state $|\pi\rangle$ corresponding to the stationary distribution π of P and applying the Hamiltonian $H(s)$ using a schedule $s(t)$ (we will specify $s(t)$ explicitly later). The resulting adiabatic evolution effectively implements a rotation on the first register at constant angular velocity from $|\pi\rangle$ to a superposition over marked vertices.

Adiabatic Search Algorithm

1. Prepare the state $|\pi\rangle|0\rangle$.
2. Apply the time-dependent Hamiltonian $H(s)$ with schedule $s(t)$ from $t = 0$ to $t = T$, where $T = \frac{\pi}{2\varepsilon}\sqrt{\text{HT}(P, M)}$.
3. Measure the first register in the vertex basis.

Under the assumption that the *folk* adiabatic condition holds in our setup, we prove the following:

Theorem 1. *For any ergodic and reversible Markov chain P with a set of marked vertices M , the Adiabatic Search Algorithm finds a marked vertex with probability at least $1 - \varepsilon^2$ in a time $T = \frac{\pi}{2\varepsilon} \sqrt{\text{HT}(P, M)}$, where $\text{HT}(P, M)$ is the hitting time of the classical Markov chain P with respect to the set of marked vertices M .*

This theorem constitutes our main result and the body of this paper will be devoted to its proof. While it relies on the *folk* adiabatic condition, a similar statement can be made for a related quantum circuit algorithm, where no such condition is necessary, as shown in [23]. Nevertheless, as explained in [23], the intuition behind the quantum circuit algorithm originates from the present adiabatic quantum algorithm.

III. PRELIMINARIES

A. Classical interpolation

Let us consider a Markov chain on a discrete state space X of size n , and let P be the $n \times n$ (row) stochastic matrix [27] describing the transition probabilities of the Markov chain. From now on, we will assume that the Markov chain is *ergodic*, meaning that it is both *irreducible* (any state can be reached from any other state by a finite number of steps) and *aperiodic* (there is no integer $k > 1$ that divides the length of every cycle of the underlying directed graph of the stochastic matrix P). Assume that a subset $M \subset X$ of elements are marked and let m be the size of M . Let P' be the Markov chain obtained from P by turning all outgoing transitions from marked elements into self-loops. We call P' the *absorbing* version of P . Note that P' differs from P only in the rows corresponding to the marked elements (where it contains all zeros on non-diagonal elements, and ones on the diagonal). If we arrange the elements of X so that the marked elements are the last ones, matrices P and P' have the following block structure:

$$P := \begin{pmatrix} P_{UU} & P_{UM} \\ P_{MU} & P_{MM} \end{pmatrix}, \quad P' := \begin{pmatrix} P_{UU} & P_{UM} \\ 0 & I \end{pmatrix}, \quad (1)$$

where P_{UU} and P_{MM} are square matrices of size $(n-m) \times (n-m)$ and $m \times m$, respectively, while P_{UM} and P_{MU} are matrices of size $(n-m) \times m$ and $m \times (n-m)$, respectively. We call

$$P(s) := (1-s)P + sP', \quad 0 \leq s \leq 1, \quad (2)$$

the *classical interpolation* of P and P' . Note that $P(0) = P$, $P(1) = P'$, and $P(s)$ has block structure

$$P(s) = \begin{pmatrix} P_{UU} & P_{UM} \\ (1-s)P_{MU} & (1-s)P_{MM} + sI \end{pmatrix}. \quad (3)$$

Moreover, note that the ergodicity of P implies that $P(s)$ is also ergodic, except for $s = 1$.

B. Stationary distribution and reversibility

By definition, since P is ergodic, it has a unique and non-vanishing *stationary distribution*, *i.e.*, a probability distribution π such that $\pi P = \pi$. An ergodic Markov chain P is called *reversible* if it satisfies the so-called *detailed balance condition*

$$\forall x, y \in X : \pi_x P_{xy} = \pi_y P_{yx}. \quad (4)$$

This implies that for reversible Markov chains, the net flow of probability in the stationary distribution between every pair of states is zero.

From now on we will assume that P is reversible. Let us argue that $P(s)$ is also reversible. Let $\pi := (\pi_U \ \pi_M)$ be the stationary distribution of P , where π_U and π_M are row vectors of length $n-m$ and m , respectively. Let $p_M = \sum_{x \in M} \pi_x$ be the probability to pick a marked element from the stationary distribution and $\pi(s)$ be the following distribution:

$$\pi(s) := \frac{1}{1-s(1-p_M)} \left((1-s)\pi_U \ \pi_M \right). \quad (5)$$

One can check that $\pi(s)$ is a stationary distribution for $P(s)$ for any $s \in [0, 1]$. Moreover, for $s \in [0, 1)$ $P(s)$ is ergodic and this is therefore the unique stationary distribution, while for $s = 1$ any distribution which only has support on marked vertices is stationary. Equation (4) can be used to show that

$$\forall s \in [0, 1], \forall x, y \in X : \pi_x(s)P_{xy}(s) = \pi_y(s)P_{yx}(s), \quad (6)$$

which means that $P(s)$ is reversible for $s \in [0, 1)$. Condition (6) is also satisfied at $s = 1$, but $P(1) = P'$ is not ergodic, therefore *stricto sensu* $P(1)$ is not reversible.

C. Discriminant matrix

The *discriminant matrix* of a Markov chain $P(s)$ is

$$D(s) := \sqrt{P(s) \circ P(s)^T}, \quad (7)$$

where the Hadamard product “ \circ ” and the square root is computed entry-wise. We prefer to work with $D(s)$ rather than $P(s)$ since a Markov chain is not necessarily symmetric while its discriminant matrix is.

For a reversible Markov chain, the extended detailed balance condition (6) implies that $D_{xy}(s) = \sqrt{P_{xy}(s)P_{yx}(s)} = P_{xy}(s)\sqrt{\pi_x(s)/\pi_y(s)}$ or equivalently

$$D(s) = \text{diag}(\sqrt{\pi(s)}) P(s) \text{ diag}(\sqrt{\pi(s)})^{-1}. \quad (8)$$

For $s \in [0, 1)$ the right-hand side is well-defined so that $D(s)$ and $P(s)$ are similar and therefore have the same eigenvalues. Moreover, the entry-wise square root of the stationary distribution $\sqrt{\pi(s)^T}$ is the eigenvector of $D(s)$ with eigenvalue 1.

At $s = 1$ the right-hand side of equation (8) is not well-defined, but it can be shown that both claims still hold by expanding $P(s)$ according to equation (3) and considering the limit $s \rightarrow 1$. Then equation (8) becomes

$$D(1) = \begin{pmatrix} \text{diag}(\sqrt{\pi_U}) P_{UU} \text{ diag}(\sqrt{\pi_U})^{-1} & 0 \\ 0 & I \end{pmatrix}. \quad (9)$$

This implies that $D(1)$ is similar to $\tilde{P} := \begin{pmatrix} P_{UU} & 0 \\ 0 & I \end{pmatrix}$, and in turn to $P(1) = \begin{pmatrix} P_{UU} & P_{UM} \\ 0 & I \end{pmatrix}$ as well. To see that $\sqrt{\pi(1)^T}$ is an eigenvector of $D(1)$ with eigenvalue 1, note that $\pi(1) = (0 \ \pi_M)/p_M$, and $D(1)$ acts as the identity on marked elements (this follows from equations (5) and (9), respectively).

D. The quantum Hamiltonian

Szegedy [13] proposed a general method to map a random walk to a unitary operator that defines a quantum walk. Recently Somma and Ortiz [22] showed how Szegedy’s method may be adapted to build a Hamiltonian. We apply this method to the random walk $P(s)$.

The first step is to map the rows of $P(s)$ to quantum states. Let \mathcal{H} be a Hilbert space of dimension $n = |X|$. For every $x \in X$ we define the following state in \mathcal{H} :

$$|p_x(s)\rangle := \sum_{y \in X} \sqrt{P_{xy}(s)} |y\rangle. \quad (10)$$

Following Szegedy [13], we then define a unitary operator $V(s)$ acting on $\mathcal{H} \otimes \mathcal{H}$ as

$$V(s)|x, 0\rangle := |x\rangle |p_x(s)\rangle = \sum_{y \in X} \sqrt{P_{xy}(s)} |x, y\rangle, \quad (11)$$

when the second register is in some reference state $|0\rangle \in \mathcal{H}$, and arbitrarily otherwise.

Let S be the gate that swaps both registers. When restricted to $|0\rangle$ in the second register, the operator $V^\dagger(s)SV(s)$ acts as $D(s)$:

$$\langle x, 0 | V^\dagger(s)SV(s) | y, 0 \rangle = \langle x | p_y(s) \rangle \langle p_x(s) | y \rangle = \sqrt{P_{xy}(s)P_{yx}(s)} = D_{xy}(s). \quad (12)$$

Following [22], we now define the Hamiltonian $H(s)$ on $\mathcal{H} \otimes \mathcal{H}$ as

$$H(s) := i[V^\dagger(s)SV(s), \Pi_0], \quad (13)$$

where $\Pi_0 := I \otimes |0\rangle\langle 0|$ is the projector that keeps only the component containing the reference state $|0\rangle$ in the second register and $[A, B] := AB - BA$ is the commutator.

IV. SPECTRAL DECOMPOSITION OF $H(s)$

To understand the properties of the Hamiltonian $H(s)$, let us find its spectral decomposition. We will relate its spectrum to that of $D(s)$.

A. Diagonalization of $D(s)$

Recall from equation (7) that $D(s)$ is real and symmetric. Therefore, its eigenvalues are real and its eigenvectors form an orthonormal basis of \mathcal{H} with real amplitudes. Let

$$D(s) = \sum_{i=1}^n \lambda_i(s) |v_i(s)\rangle\langle v_i(s)| \quad (14)$$

be the spectral decomposition of $D(s)$. We can make the eigenvalues of $P(s)$ and hence also of $D(s)$ to be non-negative by replacing P with $(P + I)/2$. Note that this will only modify

the hitting time of the Markov chain by a factor of 2. Hence, from now on without loss of generality we assume that all eigenvalues of $D(s)$ are non-negative. In addition, we can arrange them so that

$$0 \leq \lambda_1(s) \leq \lambda_2(s) \leq \cdots \leq \lambda_n(s). \quad (15)$$

From the Perron–Frobenius theorem we have that $\forall i : \lambda_i(s) \leq 1$ and $\lambda_n(s) = 1$. In addition, for any $s \in [0, 1)$ the Markov chain $P(s)$ is ergodic and $\forall i \neq n : \lambda_i(s) < 1$. On the other hand, at $s = 1$ the Markov chain is not ergodic and has eigenvalue 1 with multiplicity m . We may summarize this as follows:

$$\lambda_{n-1}(s) < \lambda_n(s) = 1, \quad \forall s \in [0, 1), \quad (16)$$

$$\lambda_{n-m}(1) < \lambda_{n-m+1}(1) = \cdots = \lambda_n(1) = 1. \quad (17)$$

Recall from Section III C that $\sqrt{\pi(s)^\top}$ is an eigenvector of $D(s)$ with eigenvalue 1 for any $s \in [0, 1]$, so let us choose $|v_n(s)\rangle$ in equation (14) as

$$|v_n(s)\rangle := \sqrt{\pi(s)^\top}. \quad (18)$$

B. Diagonalization of $H(s)$

Now, let us express the eigenvalues and eigenvectors of $H(s)$ in terms of those of $D(s)$. First, let us break up the Hilbert space $\mathcal{H} \otimes \mathcal{H}$ into mutually orthogonal subspaces that are invariant under $H(s)$. Let

- $\mathcal{B}_k(s) := \text{span} \{ |v_k(s), 0\rangle, V^\dagger(s)SV(s)|v_k(s), 0\rangle \}, \forall k \neq n,$
- $\mathcal{B}_n(s) := \text{span} \{ |v_n(s), 0\rangle \},$
- $\mathcal{B}^\perp(s) := (\oplus_{k=1}^n \mathcal{B}_k(s))^\perp.$

Note that $\langle v_k(s), 0 | V^\dagger(s)SV(s) | v_k(s), 0 \rangle = \langle v_k(s) | D(s) | v_k(s) \rangle = \lambda_k(s)$ by equations (12) and (14). Recall that for $s < 1$, we have $\lambda_k(s) \neq 1$ for any $k \neq n$ and thus

$$V^\dagger(s)SV(s)|v_k(s), 0\rangle = \lambda_k(s)|v_k(s), 0\rangle + \sqrt{1 - \lambda_k(s)^2}|v_k(s), 0\rangle^\perp \quad (19)$$

for some [28] unit vector $|v_k(s), 0\rangle^\perp$ orthogonal to $|v_k(s), 0\rangle$ and lying in the subspace $\mathcal{B}_k(s)$. We also define by continuity $|v_k(1), 0\rangle^\perp := \lim_{s \rightarrow 1} |v_k(s), 0\rangle^\perp$.

Following Somma and Ortiz, who were in turn relying on Szegedy’s work, we may now characterize the spectrum of $H(s)$.

Lemma 1 ([13, 22]). $H(s)$ accepts the following eigenvalues and eigenstates.

- On $\mathcal{B}_k(s)$, $\forall k \neq n$:

$$E_k^\pm(s) := \pm\sqrt{1 - \lambda_k(s)^2}, \quad |\Psi_k^\pm(s)\rangle := \frac{|v_k(s), 0\rangle \pm |v_k(s), 0\rangle^\perp}{\sqrt{2}}, \quad (20)$$

- On $\mathcal{B}_n(s)$:

$$E_n(s) := 0, \quad |\Psi_n(s)\rangle := |v_n(s), 0\rangle. \quad (21)$$

- On $\mathcal{B}^\perp(s)$:

$$F_j(s) := 0, \quad |\Phi_j(s)\rangle, \quad j \in \{1, \dots, (n-1)^2\}, \quad (22)$$

where $\{|\Phi_j(s)\rangle\}$ defines an arbitrary basis of $\mathcal{B}^\perp(s)$.

Proof. We consider the case $s \in [0, 1)$, the case $s = 1$ follows by continuity. Since $V^\dagger(s)SV(s)|v_n(s), 0\rangle = D(s)|v_n(s)\rangle|0\rangle = |v_n(s), 0\rangle$, we immediately have that $|v_n(s), 0\rangle$ is an eigenstate of $H(s)$ with eigenvalue 0. For $k \neq n$, note that

$$V^\dagger(s)SV(s)\Pi_0|v_k(s), 0\rangle = \lambda_k(s)|v_k(s), 0\rangle + \sqrt{1 - \lambda_k(s)^2}|v_k(s), 0\rangle^\perp, \quad (23)$$

$$\Pi_0V^\dagger(s)SV(s)|v_k(s), 0\rangle = \lambda_k(s)|v_k(s), 0\rangle. \quad (24)$$

By combining these expressions we get

$$H(s)|v_k(s), 0\rangle = i\sqrt{1 - \lambda_k(s)^2}|v_k(s), 0\rangle^\perp, \quad (25)$$

$$H(s)|v_k(s), 0\rangle^\perp = -i\sqrt{1 - \lambda_k(s)^2}|v_k(s), 0\rangle, \quad (26)$$

where the second line follows from the fact that $H(s)$ is Hermitian and traceless. In other words, $H(s)$ acts on subspace $\mathcal{B}_k(s)$ as $\sqrt{1 - \lambda_k(s)^2}\sigma_y$, where $\sigma_y := \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$ is the Pauli y matrix, which yields equation (20).

Since $\mathcal{B}^\perp(s)$ is the orthogonal complement of the union of invariant subspaces, it is also an invariant subspace for $H(s)$. Note that $H(s)$ restricted to this subspace is equal to zero, hence the remaining $n^2 - (2n - 1) = (n - 1)^2$ eigenvalues of $H(s)$ are zero. \square

V. THE QUANTUM ADIABATIC THEOREM

In adiabatic quantum computing it is a common practice to associate the intermediate state of the computation with the ground state (*i.e.*, the lowest energy eigenstate) of the

Hamiltonian. However, in our case the spectrum of $H(s)$ is symmetric about zero and the state that we are interested in lies in the middle of the spectrum. Hence, we will not use the ground state of $H(s)$, which has negative energy, but instead we will use the zero-eigenvector $|\Psi_n(s)\rangle$ given in equation (21). Indeed, equation (18) shows that this state is closely related to the stationary distribution $\pi(s)$ of $P(s)$. In particular, the problem would be solved if we can reach the state $|\Psi_n(1)\rangle$, as measuring the first register of this state yields a vertex distributed according to $\pi(1)$, which only has support on marked vertices.

A. The adiabatic condition

We initially prepare the system in the zero-eigenvector $|\Psi_n(0)\rangle$ of $H(0)$ and then start to change the Hamiltonian $H(s)$ by slowly increasing the parameter s from 0 to 1 according to some *schedule* $s(t)$. If the schedule $s(t)$ is chosen so that it satisfies certain conditions, the system is guaranteed to stay close to the intermediate zero-eigenstate $|\Psi_n(s)\rangle$. Then, at $s = 1$, the state will be close to $|\Psi_n(1)\rangle = |v_n(1), 0\rangle$, where the first register only has overlap over marked vertices, so that a measurement yields a marked vertex with high probability. Note that in our case the zero-eigenspace $\mathcal{B}_n(s) \cup \mathcal{B}^\perp(s)$ of the Hamiltonian $H(s)$ has a huge dimension, so we have to make sure that the non-relevant part $\mathcal{B}^\perp(s)$ is totally decoupled from $|\Psi_n(s)\rangle$ (the only zero-eigenvector that is relevant for our algorithm) before we apply the adiabatic condition. In particular, we want to show that

$$\langle \Phi_j(s) | \cdot \frac{d}{dt} |\Psi_n(s)\rangle = 0 \quad (27)$$

for any $j \in \{1, \dots, (n-1)^2\}$, since this would imply that during the evolution $|\Psi_n(s)\rangle$ is not leaked into the subspace $\mathcal{B}^\perp(s)$ spanned by states $|\Phi_j(s)\rangle$. To see that this is indeed the case, note that

$$|\Phi_j(s)\rangle \perp \text{span}\{|\Psi_1^\pm(s)\rangle, \dots, |\Psi_{n-1}^\pm(s)\rangle, |\Psi_n(s)\rangle\} \quad (28)$$

for any $j \in \{1, \dots, (n-1)^2\}$, since the eigenvectors of $H(s)$ form an orthonormal basis. In particular,

$$\begin{aligned} |\Phi_j(s)\rangle &\perp \text{span}\{|\Psi_1^+(s)\rangle + |\Psi_1^-(s)\rangle, \dots, |\Psi_{n-1}^+(s)\rangle + |\Psi_{n-1}^-(s)\rangle, |\Psi_n(s)\rangle\} \\ &= \text{span}\{|v_1(s), 0\rangle, \dots, |v_{n-1}(s), 0\rangle, |v_n(s), 0\rangle\}. \end{aligned} \quad (29)$$

Recall from equation (21) that $\frac{d}{dt}|\Psi_n(s)\rangle = \frac{d}{dt}|v_n(s)\rangle|0\rangle$. Hence, the inner product in equation (27) indeed vanishes. Thus, we can safely apply the adiabatic condition only for the relevant subspace in which the zero-eigenstate is not degenerate.

The *folk* version of the Adiabatic Theorem [2] states that during the evolution the state of the system $|\psi(t)\rangle$ is guaranteed to stay close to the intermediate zero-eigenstate $|\Psi_n(s)\rangle$, more precisely,

$$\forall t : |\langle \Psi_n(s(t)) | \psi(t) \rangle|^2 \geq 1 - \varepsilon^2, \quad (30)$$

as long as the *adiabatic condition*

$$\forall t : \sum_{\sigma=\pm 1} \sum_{k=1}^{n-1} \frac{|\langle \Psi_k^\sigma(s) | \cdot \frac{d}{dt} |\Psi_n(s)\rangle|^2}{(E_k^\sigma(s) - E_n(s))^2} \leq \varepsilon^2 \quad (31)$$

is satisfied. While this condition is known not to be sufficient in full generality (see *e.g.* the discussion in [7]), we will assume that it can be applied in our setup. We will discuss about how this assumption may be suppressed in Section VII.

If we insert all eigenvalues and eigenvectors from equations (20) and (21), the adiabatic condition (31) can be written purely in terms of the eigenvalues and eigenvectors of the discriminant matrix $D(s)$:

$$\forall t : \sum_{k: \lambda_k(s) \neq 1} \frac{|\langle v_k(s) | \cdot \frac{d}{dt} |v_n(s)\rangle|^2}{1 - \lambda_k^2(s)} \leq \varepsilon^2. \quad (32)$$

B. Rotation in a two-dimensional subspace

In this section we will provide a simple interpretation of the evolution of the eigenvector $|v_n(s)\rangle$. First, let us define the following superpositions over all elements, marked elements, and unmarked elements, respectively:

$$|\pi\rangle := \sum_{x \in X} \sqrt{\pi_x} |x\rangle, \quad |M\rangle := \frac{1}{\sqrt{p_M}} \sum_{x \in M} \sqrt{\pi_x} |x\rangle, \quad |U\rangle := \frac{1}{\sqrt{1-p_M}} \sum_{x \notin M} \sqrt{\pi_x} |x\rangle. \quad (33)$$

Now, we show that $|v_n(s)\rangle$ is subject to a rotation in the two-dimensional subspace $\mathcal{R} := \text{span}\{|U\rangle, |M\rangle\}$. From equations (5) and (18), we obtain

$$|v_n(s)\rangle = \sum_{x \in X} \sqrt{\pi_x(s)} |x\rangle = \frac{1}{\sqrt{1-s(1-p_M)}} \left(\sqrt{1-s} \sum_{x \notin M} \sqrt{\pi_x} |x\rangle + \sum_{x \in M} \sqrt{\pi_x} |x\rangle \right). \quad (34)$$

Using the definition of $|U\rangle$ and $|M\rangle$ we can rewrite this simply as

$$|v_n(s)\rangle = \cos \theta(s)|U\rangle + \sin \theta(s)|M\rangle, \quad (35)$$

where

$$\theta(s) := \arcsin \sqrt{\frac{p_M}{1 - s(1 - p_M)}}. \quad (36)$$

Let us choose the schedule $s(t)$ so that the evolution of $|v_n(s)\rangle$ as defined by equation (35) corresponds to a rotation with constant angular velocity in the subspace \mathcal{R} . If T is the length of the evolution and $s : [0, T] \rightarrow [0, 1]$ is defined as

$$s(t) := \frac{1}{1 - p_M} \left(1 - \frac{p_M}{\sin^2(\omega t + \theta_0)} \right), \quad \theta_0 := \arcsin \sqrt{p_M}, \quad \omega := \frac{1}{T} \arccos \sqrt{p_M}, \quad (37)$$

then

$$\theta(s(t)) = \omega t + \theta_0. \quad (38)$$

Let us choose a vector $|v_n^\perp(s)\rangle$ such that $\{|v(s)\rangle, |v_n^\perp(s)\rangle\}$ is an orthonormal basis of \mathcal{R} for every s :

$$|v_n^\perp(s)\rangle := -\sin \theta(s)|U\rangle + \cos \theta(s)|M\rangle. \quad (39)$$

Then from equations (35) and (38) we easily find that

$$\frac{d}{dt}|v_n(s(t))\rangle = \omega|v_n^\perp(s(t))\rangle = \frac{1}{T} \arccos \sqrt{p_M} |v_n^\perp(s(t))\rangle. \quad (40)$$

Note that $\arccos \sqrt{p_M} \leq \frac{\pi}{2}$. Therefore, we can rewrite the adiabatic condition (32) as follows:

$$\forall s : \frac{\pi^2}{4\varepsilon^2} \sum_{k: \lambda_k(s) \neq 1} \frac{|\langle v_k(s) | v_n^\perp(s) \rangle|^2}{1 - \lambda_k^2(s)} \leq T^2. \quad (41)$$

If this condition is satisfied, equation (30) implies that we obtain at time $t = T$ a state $|\psi(T)\rangle$ close to $|\Psi_n(1)\rangle = |v_n(1)\rangle|0\rangle = |M\rangle|0\rangle$, so that measuring the first register yields a marked vertex with probability at least $1 - \varepsilon^2$.

VI. RUNNING TIME OF THE QUANTUM ALGORITHM

A. Choice of running time T

We have to change the parameter s slowly for the evolution to be adiabatic, thus we want to choose T big enough so that condition (41) holds. Recall from Section IV A that we can

assume that $\lambda_k(s) \geq 0$. Thus, $1 - \lambda_k^2(s) = (1 + \lambda_k(s))(1 - \lambda_k(s)) \geq 1 - \lambda_k(s)$. Let us impose a slightly stronger condition on T in equation (41) by replacing $1 - \lambda_k^2(s)$ with $1 - \lambda_k(s)$. In addition, let us choose the smallest T that still satisfies the inequality and use it as the running time of our adiabatic algorithm:

$$T := \frac{\pi}{2\varepsilon} \max_{0 \leq s \leq 1} \sqrt{\sum_{k: \lambda_k(s) \neq 1} \frac{|\langle v_k(s) | v_n^\perp(s) \rangle|^2}{1 - \lambda_k(s)}}. \quad (42)$$

It turns out that there is an interesting relationship between this quantity and the hitting time of the Markov chain P .

B. Hitting time of a Markov chain

Let us first give an explicit expression for the hitting time $\text{HT}(P, M)$ as defined in Section II. Let 0_M and 1_M (resp., 0_U and 1_U) be the all-zero and all-one row vectors of dimension m (resp., $n - m$). Furthermore, let $\tilde{\pi}_M := \pi_M/p_M$ and $\tilde{\pi}_U := \pi_U/(1 - p_M)$ be the row vectors describing distributions over marked and unmarked vertices. Then, the distribution of vertices at the the first execution of step 4 of **Random Walk Algorithm** is $(\tilde{\pi}_U \ 0_M)$, and from the definition of the discriminant $D(1)$, we have

$$\text{HT}(P, M) := \sum_{t=0}^{\infty} \Pr[\text{No marked vertex found after } t \text{ steps from } (\tilde{\pi}_U \ 0_M)] \quad (43)$$

$$= \sum_{t=0}^{\infty} (\tilde{\pi}_U \ 0_M) P^t (1) (1_U \ 0_M)^\top \quad (44)$$

$$= \sum_{t=0}^{\infty} \langle U | D^t(1) | U \rangle, \quad (45)$$

where the last equality follows from equation (9). We will show that the running time of our adiabatic quantum algorithm is directly related to the square root of the hitting time $\text{HT}(P, M)$. In order to do this, we define the following extension of the hitting time. Let

$$\text{HT}(s) := \sum_{t=0}^{\infty} \langle U | (D^t(s) - |v_n(s)\rangle \langle v_n(s)|) | U \rangle. \quad (46)$$

Note that $\text{HT}(1) = \text{HT}(P, M)$ since $\langle U | v_n(1) \rangle = 0$. This justifies to consider $\text{HT}(s)$ as an extension of the hitting time. Intuitively, $\text{HT}(s)$ may be understood as the time it takes for $P(s)$ to converge to its stationary distribution, starting from $(\tilde{\pi}_U \ 0_M)$. For $s = 1$, the walk

$P(1) = P'$ converges to the (non-unique) stationary distribution $(0_U \tilde{\pi}_M)$, which only has support over marked elements.

Using the expansion $(1 - x)^{-1} = \sum_{t=0}^{\infty} x^t$ and the spectral decomposition (14) of the discriminant $D(s)$, we have

$$\text{HT}(s) = \sum_{k \neq n} \sum_{t=0}^{\infty} \lambda_k^t(s) \langle U | v_k(s) \rangle \langle v_k(s) | U \rangle \quad (47)$$

$$= \sum_{k: \lambda_k(s) \neq 1} \frac{|\langle v_k(s) | U \rangle|^2}{1 - \lambda_k(s)}. \quad (48)$$

C. Relation between the running time and the extended hitting times

Let us express the running time T from equation (42) in terms of $\text{HT}(s)$. Define

$$A(s) := \sum_{k: \lambda_k(s) \neq 1} \frac{|v_k(s)\rangle \langle v_k(s)|}{1 - \lambda_k(s)}. \quad (49)$$

Note that both T and $\text{HT}(s)$ can be expressed in terms of $A(s)$ as follows:

$$T = \frac{\pi}{2\varepsilon} \max_{0 \leq s \leq 1} \sqrt{\langle v_n^\perp(s) | A(s) | v_n^\perp(s) \rangle}, \quad \text{HT}(s) = \langle U | A(s) | U \rangle. \quad (50)$$

By definition, we have $A(s) |v_n(s)\rangle = 0$, which together with equation (35) implies that

$$A(s) |M\rangle = -\frac{\cos \theta(s)}{\sin \theta(s)} A(s) |U\rangle. \quad (51)$$

Using this and the definition of $|v_n^\perp(s)\rangle$ in equation (39), we see that $\langle v_n^\perp(s) | A(s) | v_n^\perp(s) \rangle = \langle U | A(s) | U \rangle / \sin^2 \theta(s)$. Thus we get the following relationship between T and $\text{HT}(s)$:

$$T = \frac{\pi}{2\varepsilon} \max_{0 \leq s \leq 1} \frac{\sqrt{\text{HT}(s)}}{\sin \theta(s)}. \quad (52)$$

To relate T and the usual hitting time $\text{HT}(P, M)$ of P , we first provide an explicit expression for $\text{HT}(s)$ in terms of $\text{HT}(P, M)$ (the proof is given in the appendix).

Lemma 2. $\text{HT}(s) = \text{HT}(P, M) \sin^4 \theta(s)$.

Now, recalling the definition of $\theta(s)$ in equation (36), it is easy to see that the maximum in equation (52) is attained at $s = 1$. This immediately implies that the running time of the adiabatic quantum algorithm is given by

$$T = \frac{\pi}{2\varepsilon} \sqrt{\text{HT}(P, M)}, \quad (53)$$

therefore providing a quadratic improvement over the classical hitting time. This also concludes the proof of Theorem 1.

VII. CONCLUSION AND DISCUSSION

Our quantum algorithm defines a new notion of quantum hitting time, which is quadratically smaller than the classical hitting time for any reversible Markov chain and any set of marked elements. While previous approaches were subject to various restrictions, *e.g.*, the quantum algorithm could only detect the presence of marked elements [13], did not always provide full quadratic speed-up [15], or could only be applied for state-transitive Markov chains with a unique marked element [16], our adiabatic approach only requires minimal assumptions. Indeed, it can be shown that the only remaining condition, reversibility, is necessary. Let us consider the Markov chain on a cycle $P = \frac{I+C}{2}$, where C implements a clockwise shift, *i.e.*, $C|x\rangle = |(x+1) \bmod n\rangle$. This Markov chain is ergodic but not reversible. While its classical hitting time is of order $\Theta(n)$, a simple locality argument implies that any quantum operator acting locally on the cycle requires a time $\Omega(n)$ to find a marked vertex, so that a quadratic speed-up cannot be achieved. Magniez *et al.* [16] have also shown that under reasonable conditions the quadratic speed-up is optimal. This provides evidence that our result is both as strong and as general as possible.

While our result relies on the assumption that the *folk* adiabatic condition is sufficient, this assumption could be suppressed in different ways. One option would be to actually prove that the *folk* Adiabatic Theorem holds in our setup, as was previously done for the adiabatic version of Grover's algorithm [7]. Another option would be to circumvent adiabatic evolution altogether, by using the phase randomization technique of Boixo *et al.* [24]. Their technique provides a quantum circuit realizing the same evolution as the adiabatic approach with a similar running time, but without relying on the adiabatic condition. This leads to the quantum circuit algorithm described in [23].

Finally, note that in order to design the schedule $s(t)$, our algorithm requires to know p_M and the order of magnitude of $\text{HT}(P, M)$. These assumptions are standard in other quantum algorithms for this problem. In particular, a similar issue arises in Grover's algorithm when the number of marked elements is unknown. In Grover's case, there are techniques to deal with this issue [25], and similar techniques could be applied in our case. While we do not provide a full answer to these questions in the present paper, they do not present any new technical difficulty and we refer the reader to [23] where a full study of these implementation issues is provided for a related quantum circuit algorithm.

Acknowledgments

J. Roland would like to thank F. Magniez, A. Nayak, M. Santha and R. Somma for useful discussions. H. Krovi would like to thank F. Magniez for useful discussions. This research has been supported in part by ARO/NSA under grant W911NF-09-1-0569. M. Ozols acknowledges support from QuantumWorks.

- [1] E. Farhi, J. Goldstone, S. Gutmann, and M. Sipser (2000), arXiv:quant-ph/0001106.
- [2] A. Messiah, *Mécanique Quantique* (Dunod, Paris, 1959).
- [3] K.-P. Marzlin and B. C. Sanders, Phys. Rev. Lett. **93**, 160408 (2004).
- [4] M. S. Sarandy, L.-A. Wu, and D. A. Lidar, Quantum Information Processing **3**, 331 (2004).
- [5] Z. Wu, L. Zheng, and H. Yang (2004), arXiv:quant-ph/0411212.
- [6] D. M. Tong, K. Singh, L. C. Kwek, and C. H. Oh, Phys. Rev. Lett. **95**, 110407 (2005).
- [7] S. Jansen, M.-B. Ruskai, and R. Seiler, J. Math. Phys. **48**, 102111 (2007).
- [8] L. K. Grover, in *Proceedings of the 28th ACM Symposium on Theory of Computing* (ACM Press, 1996), pp. 212–219.
- [9] W. van Dam, M. Mosca, and U. Vazirani, in *Proceedings of the 42nd IEEE Symposium on Foundations of Computer Science* (IEEE Computer Society Press, 2001), pp. 279–287.
- [10] J. Roland and N. J. Cerf, Phys. Rev. A **65**, 042308 (2002).
- [11] A. Ambainis, J. Kempe, and A. Rivosh, in *Proceedings of the 16th ACM-SIAM Symposium on Discrete Algorithms* (SIAM, 2005), pp. 1099–1108.
- [12] J. Kempe, Prob. Th. Rel. Fields **133**, 215 (2005).
- [13] M. Szegedy, in *Proceedings of the 45th IEEE Symposium on Foundations of Computer Science* (IEEE Computer Society Press, 2004), pp. 32–41.
- [14] H. Krovi and T. A. Brun, Phys. Rev. A **73**, 032341 (2006).
- [15] F. Magniez, A. Nayak, J. Roland, and M. Santha, in *Proceedings of the 39th ACM Symposium on Theory of Computing* (ACM Press, 2007), pp. 575–584, ISBN 978-1-59593-631-8.
- [16] F. Magniez, A. Nayak, P. C. Richter, and M. Santha, in *Proceedings of the 19th ACM-SIAM Symposium on Discrete Algorithms* (SIAM, 2009), pp. 86–95.
- [17] M. Varbanov, H. Krovi, and T. A. Brun, Phys. Rev. A **78**, 022324 (2008).

- [18] A. Ambainis, in *Proceedings of the 45th IEEE Symposium on Foundations of Computer Science* (IEEE Computer Society Press, 2004), pp. 22–31.
- [19] A. M. Childs and J. Goldstone, Phys. Rev. A **70**, 022314 (2004).
- [20] A. M. Childs and J. Goldstone, Phys. Rev. A **70**, 042312 (2004).
- [21] A. Tulsi, Phys. Rev. A **78**, 012310 (2008).
- [22] R. D. Somma and G. Ortiz, in *Quantum Quenching, Annealing and Computation* (Springer, Heidelberg, 2010), Lecture Notes in Physics, in Press.
- [23] H. Krovi, F. Magniez, M. Ozols, and J. Roland (2010), arXiv:1002.2419 [quant-ph].
- [24] S. Boixo, E. Knill, and R. D. Somma, Quant. Inf. Comp. **9**, 0833 (2009).
- [25] M. Boyer, G. Brassard, P. Høyer, and A. Tapp, Fortschritte Der Physik **46**, 493 (1998).
- [26] Note that when we consider *reversible Markov chains* as defined in Section III B, this corresponds to a different notion of reversibility than in the usual thermodynamical sense. Actually, even a “reversible” Markov chain is thermodynamically irreversible.
- [27] Throughout the paper we use the convention that each row of a stochastic matrix P sums to one ($\forall x \in X : \sum_{y \in X} P_{xy} = 1$) and probability distributions are row vectors and hence are multiplied to the transition matrix from the left hand side (*e.g.*, πP).
- [28] Note that $|v_k(s), 0\rangle^\perp$ depends on how the operator $V(s)$ defined in equation in (11) is extended to the whole Hilbert space.

Appendix: Proof of Lemma 2

In this section, we will often use $\dot{f}(s)$ as a shorthand form of $\frac{d}{ds}f(s)$. We will show that the derivative of $\text{HT}(s)$ satisfies the following lemma.

Lemma 3. *The derivative of $\text{HT}(s)$ is related to $\text{HT}(s)$ as:*

$$\frac{d}{ds} \text{HT}(s) = 4\dot{\theta}(s) \frac{\cos \theta(s)}{\sin \theta(s)} \text{HT}(s). \quad (\text{A.1})$$

Note that Lemma 2 follows directly from there, since $\text{HT}(s) = \text{HT}(P, M) \sin^4 \theta(s)$ satisfies the differential equation for $\text{HT}(s)$ with the boundary condition $\text{HT}(1) = \text{HT}(P, M)$. Therefore, it remains to prove Lemma 3.

Before proving Lemma 3, let us consider the derivative of the discriminant $D(s)$. Let $\Pi_M = \sum_{x \in M} |x\rangle\langle x|$ be the projector onto the m -dimensional subspace spanned by marked vertices, and let $\{X, Y\} := XY + YX$ be the anticommutator of X and Y .

Lemma 4. $\dot{D}(s) = \frac{1}{2(1-s)}\{\Pi_M, I - D(s)\}$.

Proof. Note from equation (3) that $D(s)$ has the following block structure:

$$D(s) = \begin{pmatrix} \sqrt{P_{UU} \circ P_{UU}^T} & \sqrt{(1-s)(P_{UM} \circ P_{MU}^T)} \\ \sqrt{(1-s)(P_{MU} \circ P_{UM}^T)} & (1-s)\sqrt{P_{MM} \circ P_{MM}^T} + sI \end{pmatrix}. \quad (\text{A.2})$$

Hence,

$$\dot{D}(s) = \begin{pmatrix} 0 & -\frac{1}{2\sqrt{1-s}}\sqrt{P_{UM} \circ P_{MU}^T} \\ -\frac{1}{2\sqrt{1-s}}\sqrt{P_{MU} \circ P_{UM}^T} & I - \sqrt{P_{MM} \circ P_{MM}^T} \end{pmatrix}. \quad (\text{A.3})$$

Observe that $\dot{D}(s) + \frac{1}{2(1-s)}\{\Pi_M, D(s)\} = \frac{1}{1-s}\Pi_M$, which implies the lemma. \square

We are now ready to prove Lemma 3.

Proof of Lemma 3. In this proof we will often omit to write the dependence on s explicitly.

From equation (50) we have

$$\frac{d}{ds} \text{HT} = \langle U | \dot{A} | U \rangle. \quad (\text{A.4})$$

Note that $A = B^{-1} - \Pi_n$, where

$$B := I - D + \Pi_n, \quad \Pi_n := |v_n\rangle\langle v_n|. \quad (\text{A.5})$$

For any invertible matrix M we have $\frac{d}{ds}(M^{-1}) = -M^{-1}\dot{M}M^{-1}$. Therefore,

$$\dot{A} = -B^{-1}(-\dot{D} + \dot{\Pi}_n)B^{-1} - \dot{\Pi}_n. \quad (\text{A.6})$$

Hence, we have $\frac{d}{ds} \text{HT} = h_1 + h_2 + h_3$, where

$$h_1 := \langle U | B^{-1} \dot{D} B^{-1} | U \rangle, \quad (\text{A.7})$$

$$h_2 := -\langle U | B^{-1} \dot{\Pi}_n B^{-1} | U \rangle, \quad (\text{A.8})$$

$$h_3 := -\langle U | \dot{\Pi}_n | U \rangle. \quad (\text{A.9})$$

Let us evaluate each of these terms separately. We can use Lemma 4 and the definition of B to express the first term h_1 as follows:

$$2(1-s)h_1 = \langle U | B^{-1} \{\Pi_M, I - D\} B^{-1} | U \rangle \quad (\text{A.10})$$

$$= \langle U | B^{-1} (\{\Pi_M, B\} - \{\Pi_M, \Pi_n\}) B^{-1} | U \rangle \quad (\text{A.11})$$

$$= \langle U | \{B^{-1}, \Pi_M\} | U \rangle - \langle U | B^{-1} \{\Pi_M, \Pi_n\} B^{-1} | U \rangle. \quad (\text{A.12})$$

Note that $\Pi_M|U\rangle = 0$, thus the first term vanishes. Also note that $B^{-1}|v_n\rangle = |v_n\rangle$ and B^{-1} is Hermitian, thus

$$2(1-s)h_1 = -2\langle U|B^{-1}\Pi_M\Pi_n|U\rangle. \quad (\text{A.13})$$

Note from equation (35) that $\Pi_M|v_n\rangle = \sin\theta|M\rangle$. Moreover, using equation (51) we can simplify equation (A.13) even further:

$$2(1-s)h_1 = -2\cos\theta\sin\theta\langle U|B^{-1}|M\rangle \quad (\text{A.14})$$

$$= -2\cos\theta\sin\theta\langle U|(A + \Pi_n)|M\rangle \quad (\text{A.15})$$

$$= -2\cos\theta\sin\theta\langle U|A|M\rangle - 2\cos^2\theta\sin^2\theta \quad (\text{A.16})$$

$$= 2\cos^2\theta(\langle U|A|U\rangle - \sin^2\theta). \quad (\text{A.17})$$

Let us now consider the second term h_2 :

$$h_2 = -\langle U|B^{-1}\dot{\Pi}_nB^{-1}|U\rangle \quad (\text{A.18})$$

$$= -\cos\theta(\langle \dot{v}_n|B^{-1}|U\rangle + \langle U|B^{-1}|\dot{v}_n\rangle) \quad (\text{A.19})$$

$$= -2\cos\theta\langle U|(A + \Pi_n)|\dot{v}_n\rangle \quad (\text{A.20})$$

$$= -2\dot{\theta}\cos\theta(-\sin\theta\langle U|A|U\rangle + \cos\theta\langle U|A|M\rangle) \quad (\text{A.21})$$

$$= -2\dot{\theta}\cos\theta\left(-\sin\theta\langle U|A|U\rangle - \frac{\cos^2\theta}{\sin\theta}\langle U|A|U\rangle\right) \quad (\text{A.22})$$

$$= 2\dot{\theta}\frac{\cos\theta}{\sin\theta}\langle U|A|U\rangle, \quad (\text{A.23})$$

where we have used the fact that

$$|\dot{v}_n\rangle = \dot{\theta}(-\sin\theta|U\rangle + \cos\theta|M\rangle), \quad (\text{A.24})$$

and $\Pi_n|\dot{v}_n\rangle = 0$. Similarly, for the last term h_3 we have

$$h_3 = -\langle U|\dot{\Pi}_n|U\rangle = 2\dot{\theta}\cos\theta\sin\theta. \quad (\text{A.25})$$

Putting all the terms back together, we have

$$\begin{aligned} \langle U|\dot{A}|U\rangle &= \frac{\cos^2\theta}{1-s}(\langle U|A|U\rangle - \sin^2\theta) + 2\dot{\theta}\frac{\cos\theta}{\sin\theta}\langle U|A|U\rangle + 2\dot{\theta}\cos\theta\sin\theta \\ &= \frac{\cos\theta}{\sin\theta}\left(2\dot{\theta} + \frac{\cos\theta\sin\theta}{1-s}\right)\langle U|A|U\rangle + \cos\theta\sin\theta\left(2\dot{\theta} - \frac{\cos\theta\sin\theta}{1-s}\right). \end{aligned} \quad (\text{A.26})$$

From the definition of θ in equation (36) it is straightforward to check that

$$2\dot{\theta} = \frac{\cos\theta\sin\theta}{1-s} \quad (\text{A.27})$$

which together with equations (A.4) and (A.26) implies the lemma. \square