

EIGENPATH TRAVERSAL BY PHASE RANDOMIZATION

SERGIO BOIXO

*Institute for Quantum Information, California Institute of Technology
Pasadena, California, 91125, USA*

EMANUEL KNILL

*National Institute of Standards and Technology
Boulder, Colorado 80305, USA*

ROLANDO SOMMA

*Perimeter Institute for Theoretical Physics
Waterloo, Ontario N2L 2Y5, Canada*

Received March 11, 2009

Revised July 9, 2009

A computation in adiabatic quantum computing is implemented by traversing a path of nondegenerate eigenstates of a continuous family of Hamiltonians. We introduce a method that traverses a discretized form of the path: At each step we apply the instantaneous Hamiltonian for a random time. The resulting decoherence approximates a projective measurement onto the desired eigenstate, achieving a version of the quantum Zeno effect. If negative evolution times can be implemented with constant overhead, then the average absolute evolution time required by our method is $\mathcal{O}(L^2/\Delta)$ for constant error probability, where L is the length of the path of eigenstates and Δ is the minimum spectral gap of the Hamiltonian. The dependence of the cost on Δ is optimal. Making explicit the dependence on the path length is useful for cases where L is much less than the general bound. The complexity of our method has a logarithmic improvement over previous algorithms of this type. The same cost applies to the discrete-time case, where a family of unitary operators is given and each unitary and its inverse can be used. Restriction to positive evolution times incurs an error that decreases exponentially with the cost. Applications of this method to unstructured search and quantum sampling are considered. In particular, we discuss the quantum simulated annealing algorithm for solving combinatorial optimization problems. This algorithm provides a quadratic speed-up in the gap of the stochastic matrix over its classical counterpart implemented via Markov chain Monte Carlo.

Keywords: Adiabatic theorem, Zeno effect, dynamical decoupling, randomization.

Communicated by: R. Jozsa & B. Terhal

1 Introduction and summary of results

Quantum algorithms are often described by means of quantum circuits: the algorithm starts with a well-characterized pure state; a sequence of elementary (unitary) gates is applied; and a final projective measurement in a fixed basis extracts the result. The circuit model may not be best for describing all quantum information processing systems. Adiabatic quantum computing (AQC) [1], sometimes also called quantum annealing [2, 3, 4, 5, 6], has been proposed as an alternative.

In AQC the computation is performed by smoothly changing the interaction parameters of the Hamiltonian under which the system evolves. The initial state is a nondegenerate eigenstate of the Hamiltonian. The adiabatic theorem of quantum mechanics asserts that if the continuously related eigenstates remain nondegenerate and the Hamiltonians change sufficiently slowly, then the final state of the system is close to the continuously related eigenstate of the final Hamiltonian [7]. The last step is a standard projective measurement. AQC is polynomially equivalent to the quantum circuit model [8].

In this paper we give a method for traversing eigenstate paths of Hamiltonians that differs from AQC by the use of evolution randomization. The method is based on previous results [9, 10] in which the evolution of AQC is replaced by a sequence of projective measurements onto the instantaneous eigenstate of the Hamiltonian with the phase estimation algorithm [11], which exploits the quantum Zeno effect. Both AQC and the Zeno-based model work, in essence, because an effective level decoupling is introduced in the Hamiltonian eigenbasis by phase cancellation in AQC or projections in the Zeno case. Our method also implements a version of the quantum Zeno effect. We choose a discretization of the eigenstate path and apply the Hamiltonian corresponding to each point for a random time. The probability distribution over time may be discrete or continuous. Consequently, the *randomization method* can also be used in the case where we are given a path of efficiently implementable unitary operators and an eigenstate of the last operator on the path is to be prepared. This case occurs in the quantum simulated annealing (QSA) algorithm constructed in Ref. [12]. The probability distribution over evolution times must be chosen so as to cancel unwanted coherences and simulate the Zeno effect.

The algorithmic complexity of the randomization method is defined as the average sum of the absolute evolution times for the Hamiltonians or by the average number of times the unitaries are applied. The complexity can be bounded in terms of a lower bound Δ on the absolute value of the minimum spectral gap of the Hamiltonians or the minimum phase gap of the unitaries, the length L of the path of the states (defined below), and the desired maximum error ϵ of the final state compared to the target eigenstate. We show that the complexity is $\mathcal{O}(\log(L/\epsilon)^\alpha L^2/(\epsilon\Delta))$, where $\alpha = 0$ if we can evolve backward and forward in time, and 1 otherwise. Backward evolution is possible at the same cost by reversing quantum circuits for the forward evolution, if such evolution is circuit-based. To achieve this complexity without additional dependencies, we use a parametrization of the operators along the path for which the eigenstates move at a rate that is close to uniform (up to a constant factor). In many cases of interest, $L \in \mathcal{O}(1)$ so that the complexity is of order $1/(\epsilon\Delta)$ up to logarithmic factors. The scaling with the gap is optimal and is better than the $1/\Delta^3$ of rigorous proofs of the adiabatic theorem [13, 14, 15].

An advantage of our approach is that the only requirement on the Hamiltonians or the unitaries along the path is that the length of the desired eigenstate path is well-defined. A sufficient condition is that the time derivative of the operators exists. In terms of bounds on the Hamiltonians and their derivatives, the worst-case bound is of order $\|\dot{H}\|^2/(\epsilon\Delta^3)$ up to logarithmic factors. This bound comes from the inequality $L \leq \|\dot{H}\|/\Delta$ [see Eq. (13)] and does not depend on a reparametrization. It is better than the known worst-case bounds associated with the adiabatic theorem [13] in that it does not depend on existence of, or bounds on the second derivative of H . The scaling of the bound with the error is worse, in that it can be

made logarithmic for analytic Hamiltonians paths at the cost of less favorable dependencies on the other parameters of the problem [16, 14]. Logarithmic scaling can also be obtained with the randomization method provided the final eigenstate’s energy is known. To achieve this scaling, one can use high precision phase estimation to determine whether the desired eigenstate has been obtained and repeat the algorithm if not.

Our method is intuitively explained by the quantum Zeno effect. Suppose that the path of states $|\tilde{\psi}(l_1)\rangle, \dots, |\tilde{\psi}(l_q)\rangle$ satisfies the condition that for each j , $|\tilde{\psi}(l_{j-1})\rangle$ is sufficiently close to $|\tilde{\psi}(l_j)\rangle$. If we initialize the state $|\tilde{\psi}(l_1)\rangle$ and sequentially apply projections onto the $|\tilde{\psi}(l_j)\rangle$, we prepare $|\tilde{\psi}(l_q)\rangle$ with good probability. We first consider an idealized strategy, where the projections are replaced by quantum operations of the form

$$M_{l_j}(\rho) = P_{l_j}\rho P_{l_j} + \mathcal{E}((\mathbb{1} - P_{l_j})\rho(\mathbb{1} - P_{l_j})) , \tag{1}$$

where $P_{l_j} = |\tilde{\psi}(l_j)\rangle\langle\tilde{\psi}(l_j)|$ and \mathcal{E} is an arbitrary quantum operation that may vary from instance to instance. This can be thought of as a projective measurement of ρ onto $|\tilde{\psi}(l_j)\rangle$ followed by a process that does not affect $|\tilde{\psi}(l_j)\rangle$. The fundamental effect of M_{l_j} is to remove coherences between $|\tilde{\psi}(l_j)\rangle$ and orthogonal states. It is this decoherence that induces the quantum Zeno effect by suppressing transfer of population to orthogonal states. An approximation of this effect is achieved if we replace the M_{l_j} by random applications of Hamiltonians or unitaries with $|\tilde{\psi}(l_j)\rangle$ as an eigenstate. We formalize this claim in Sec. 2, Thm. 1, and give an upper bound for the error in the approximation in terms of the characteristic function of the probability distribution underlying the randomization.

We focus on the Hamiltonian-based version of the randomization method. The analysis for the unitary version is a straightforward discretization. In the Hamiltonian version, the randomization method takes as input a continuous path of Hamiltonians $\mathcal{H} = \{H(s), s \in [0, 1]\}$, and a nondegenerate eigenstate $|\psi(0)\rangle$ of $H(0)$. The method aims to output the corresponding nondegenerate eigenstate of $H(1)$, denoted by $|\psi(1)\rangle$, with high fidelity.

We require that the eigenstates $|\psi(s)\rangle$ are nondegenerate with Δ a lower bound on the energy gap. If $|\psi(s)\rangle$ is differentiable (see Appendix 1 for the more general case), we can assume without loss of generality that the phases of the $|\psi(s)\rangle$ are chosen geometrically, so that $\langle\partial_s\psi(s)|\psi(s)\rangle = 0$, which gives a path length

$$L = \int_0^1 \|\partial_s\psi(s)\| ds . \tag{2}$$

The quadratic cost dependence on L comes from a simple Zeno effect when an ideal decoherence process according to Eq. (1) is used. It is probably not fundamental: Coherent versions of the adiabatic path achieve scalings $\tilde{O}(L)$ [17]. The dependence of the cost on $1/\Delta$ is unavoidable for methods with only oracle access to the Hamiltonian or unitaries. This can be seen intuitively by noting that we must, in a sense, distinguish between the desired eigenstate and the others, which requires that we evolve the relative phases sufficiently far. More rigorously, an asymptotically better dependence would result in an unstructured search algorithm better than Grover’s, which is known to be impossible. See Sec. 4.1.

The paper is organized as follows. In Sec. 2 we explain how the quantum Zeno effect can be exploited, show how to approximate projective measurement operations by means of evolution randomization, and discuss several probability distributions that are useful for randomization.

The randomization method and its complexity are analyzed in Sec. 3. In Sec. 4.1 we show that the randomization method provides the expected quadratic quantum speed-up for the unstructured search problem. In Sec. 4.2 we describe the QSA to simulate slowly varying classical Markov chains. In Sec. 5 we show the equivalence of our randomization method with a coherent version of the quantum Zeno method implemented via the phase estimation algorithm, and briefly discuss related works. We summarize in Sec. 6.

2 Randomized evolutions

2.1 Adiabatic quantum computing using the Zeno effect

The quantum Zeno effect is based on the fact that, for a small displacement δ' , the probability of projecting $|\psi(s + \delta')\rangle$ onto $|\psi(s)\rangle$ decreases with $(\delta')^2$, while the distance between states is linear in δ' [18, 19, 20]. Therefore, for the path of states $\{|\psi(s)\rangle\}$, the final state $|\psi(1)\rangle$ can be prepared from the initial state $|\psi(0)\rangle$ with high fidelity by use of a sequence of measurement projections onto intermediate states $|\psi(s_1)\rangle, \dots, |\psi(s_q)\rangle$, $0 < s_1 < \dots < s_q = 1$. We choose s_j so that the fidelity of the final state with respect to $|\psi(1)\rangle$ is sufficiently close to unity. It is not necessary to keep track of the measurement results at intermediate steps, which gives rise to the following definition.

Definition 1 *A projective-measurement operation onto $|\tilde{\psi}(l)\rangle$ is a quantum operation of the form*

$$M_l(\rho) = P_l \rho P_l + \mathcal{E}((\mathbb{1} - P_l)\rho(\mathbb{1} - P_l)) ,$$

with $P_l = |\tilde{\psi}(l)\rangle\langle\tilde{\psi}(l)|$ and \mathcal{E} arbitrary quantum operations that may vary with l .

We assume a monotonically increasing parametrization $s(l)$, with $l \in [0, L']$, $s(0) = 0$ and $s(L') = 1$. We define $|\tilde{\psi}(l)\rangle = |\psi(s(l))\rangle$. (Objects with a tilde correspond to objects in the new parametrization). Later we consider $s(l)$ so that $L' = L$, the path length of Eq. (2). We formulate the Zeno method for quantum state preparation as follows [9, 10, 12]:

Lemma 1 (Zeno effect) *Consider a continuous path of states $\{|\tilde{\psi}(l)\rangle\}_{l \in [0, L']}$ and assume that, for fixed d and all δ ,*

$$|\langle\tilde{\psi}(l)|\tilde{\psi}(l + \delta)\rangle|^2 \geq 1 - d^2\delta^2 .$$

Then the state $|\tilde{\psi}(L')\rangle$ can be prepared from $|\tilde{\psi}(0)\rangle$ with fidelity $p > 0$ by $\lceil(L')^2 d^2 / (1 - p)\rceil$ intermediate projective-measurement operations.

Proof. Divide $[0, L']$ into $q = \lceil(L')^2 d^2 / (1 - p)\rceil$ equal segments and set $\delta = L'/q$. At every point $l_j = j\delta$, $1 \leq j \leq q$, we perform a projective-measurement operation onto $|\tilde{\psi}(l_j)\rangle$. The final state is $M_{l_q} \circ M_{l_{q-1}} \circ \dots \circ M_{l_1}(\rho)$, with $\rho = |\tilde{\psi}(0)\rangle\langle\tilde{\psi}(0)|$. The output fidelity is bounded as

$$\begin{aligned} \text{tr} [P_{l_q}(M_{l_q} \circ \dots \circ M_{l_1}(\rho))] &\geq \|P_{l_q} \dots P_{l_1} |\tilde{\psi}(0)\rangle\|^2 \\ &= \prod_{j=1}^q |\langle\tilde{\psi}(l_j)|\tilde{\psi}(l_{j-1})\rangle|^2 \\ &\geq (1 - d^2\delta^2)^q \geq 1 - d^2 L'^2 / q \geq 1 - (1 - p) = p \end{aligned} \tag{3}$$

□.

From Lemma 1 and assuming a *uniform* parametrization, defined to satisfy $L(s(l)) = \tilde{L}(l) = l$, $d = 1$, and $L' = L$ (see Appendix B), it follows that the state $|\psi(1)\rangle$ can be obtained with fidelity p starting from $|\psi(0)\rangle$ with $\mathcal{O}(L^2/(1 - p))$ projective-measurement operations.

2.2 Approximating projective-measurement operations through randomized evolutions

We assume that evolutions under $H(s)$ for time t can be implemented at a cost linear in $|t||H(s)|$, as in AQC. That is, we do not take into account the cost of simulating $H(s)$ for small time intervals. By rescaling $H(s)$ if necessary, we can assume that $\|H(s)\| \leq 1$. Thus, the cost of the randomization method is determined by the sum of the absolute evolution times. Although we consider the case where the evolution time t can be negative, one often restricts t to be nonnegative. This restriction is justified if the Hamiltonians are physical without a simple time-reversal procedure, rather than induced by quantum circuits. In the latter case, evolving for negative t is as efficient as for positive t and can be realized by reversing the quantum circuits.

We denote by $\Delta(s)$ the spectral gap for the eigenstate $|\psi(s)\rangle$ of Hamiltonian $H(s)$. The following results also apply to the unitary case where we are given operators $U(s)$ and $\Delta(s)$ is the phase gap. In the unitary case the distributions over time that are used for randomization must be concentrated at the integers, and correspond to the number of times the unitaries are applied.

According to Lemma 1, the Zeno method does not require that we keep track of intermediate measurement results. Thus, any purely dephasing mechanism in the instantaneous eigenbasis of $\tilde{H}(l)$ implements a version of M_l . A natural choice for such a decoherence mechanism is the evolution induced by $\tilde{H}(l)$ for a (unknown) random time. This is the subject of next theorem, where we bound the residual coherences in terms of the characteristic function of the random time distribution.

Theorem 1 (Randomized dephasing) *Let $|\tilde{\psi}(l)\rangle$ be a nondegenerate eigenstate of $\tilde{H}(l)$, and $\{\omega_j\}$ be the energy differences to the other eigenstates $|\tilde{\psi}_j(l)\rangle$. Let T be a random variable associated with the time of evolution under $\tilde{H}(l)$, and \mathcal{R}_l^T the corresponding quantum operation. Then there exists a quantum operation \mathcal{E} such that, for all states ρ ,*

$$\|(M_l^\mathcal{E} - \mathcal{R}_l^T)(\rho)\|_{\text{tr}} \leq \epsilon = \sup_{\omega_j} |\Phi(\omega_j)| ,$$

where $M_l^\mathcal{E}$ is the projective-measurement operation defined in Definition 1 with \mathcal{E} specified, and Φ is the characteristic function of T .

We give the proof in Appendix C. It is based on computing the coherences after the randomized evolution in terms of the characteristic function of T as

$$\mathcal{R}_l^T(|\tilde{\psi}(l)\rangle\langle\tilde{\psi}_j(l)|) = \Phi(\omega_j)|\tilde{\psi}(l)\rangle\langle\tilde{\psi}_j(l)| . \tag{4}$$

The average cost of randomization is given in terms of the random variable T as $\langle|T|\rangle$, the expected value of the absolute evolution time. If T takes only positive values the average cost is given by

$$\langle T \rangle = -i[\partial_\omega \Phi](0) ,$$

provided it is finite. Note that if $T < 0$ is allowed, then the average cost can be reduced by shifting T 's distribution so that 0 is a median of T .

We can bound the required average cost per step from below by $\Omega(1/\Delta)$, with Δ a lower bound on the smallest gap $\inf_s |\Delta(s)|$, by means of the following theorem:

Theorem 2 *Let T be a random variable with characteristic function Φ . Then, for all ω ,*

$$\text{cost}(T) = \langle |T| \rangle \geq \frac{1 - |\Phi(\omega)|}{|\omega|} .$$

Proof. From the definition of Φ we obtain

$$\begin{aligned} 1 - |\Phi(\omega)| &\leq |1 - \Phi(\omega)| \leq \int |1 - e^{i\omega t}| d\mu(t) \\ &\leq \int |\omega t| d\mu(t) = \langle |T| \rangle |\omega| , \end{aligned} \tag{5}$$

with μ the probability distribution of T \square .

We want to ensure that after the randomized evolution, the remaining coherences bounded by $|\Phi(\omega)|$ for $|\omega| \geq \Delta$ are small. Because of Thm. 2, the average absolute evolution time can be bounded by $\Omega(1/\Delta)$.

If $T < 0$ is permitted, the bound of Thm. 2 can be achieved up to a constant factor. See Example 2 and Lemma 3 in Sec. 2.3. For the case where we are given a path of unitaries and T is restricted to the integers, it suffices to consider the characteristic function on the interval $[-\pi, \pi]$. The results of this section are otherwise unchanged.

Repetition of the randomized evolution step decreases the error exponentially fast in the number of repetitions, as shown by the following argument. For independent random variables T_1 and T_2 , the characteristic function of the sum $T' = T_1 + T_2$ is

$$\Phi' = \Phi_1 \Phi_2 , \tag{6}$$

with Φ_i the characteristic function of T_i . Thus we have the following lemma (the notation is that of Thm. 1):

Lemma 2 *Let T be a random variable with characteristic function Φ , and $\sup_{\omega_j} |\Phi(\omega_j)| = \epsilon$. Let T' be the sum of n independent instances of T . Then there exists a quantum operation \mathcal{E} such that for all states ρ ,*

$$\|(M_i^{\mathcal{E}} - \mathcal{R}_i^{T'}) (\rho)\|_{\text{tr}} \leq \epsilon^n .$$

2.3 Examples of randomized evolutions

We consider some examples of randomized evolution steps involving different time distributions.

1. Consider the case where all the orthogonal eigenstates to $|\tilde{\psi}(l)\rangle$ are degenerate and the spectral gap of $\tilde{H}(l)$, denoted by ω_1 , is known. We then choose a random variable T_{ω_1} that takes the values $t = 0$ or $t = \pi/\omega_1$, each with probability $1/2$. The average cost is $\pi/(2\omega_1)$. The characteristic function for this distribution satisfies

$$|\Phi(\omega)| = \left| \cos \left(\frac{\pi\omega}{2\omega_1} \right) \right| . \tag{7}$$

Since $\Phi(\omega_1) = 0$, Thm. 1 implies that the projective measurement onto $|\tilde{\psi}(l)\rangle$ can be simulated exactly with this distribution. The assumptions in this example may seem

unrealistic, but it provides a basis for the randomization method in unstructured search (Sec. 4.1). It is possible to generalize the method to the case where the spectrum is known. If there are k distinct absolute eigenvalue differences ω_j , the independent sum of T_{ω_j} has the property that the characteristic function is identically zero on the eigenvalue differences. The average cost is $\sum_j \pi/(2\omega_j)$.

2. Let T 's probability density be proportional to $\text{sinc}(\lambda t)^4$, $\lambda > 0$. The function sinc is defined as $\text{sinc}(t) = \sin(t)/t$. The Fourier transform of $\lambda \text{sinc}(\lambda t)/\pi$ is the indicator function of the interval $[-\lambda, \lambda]$. The characteristic function of T is therefore proportional to the four-fold convolution of this indicator function with itself, which is continuous and has support $[-4\lambda, 4\lambda]$. There is no error in approximating the projective-measurement operation by randomized evolution if we choose $4\lambda = \Delta$, with Δ a lower bound on the minimum gap. The average cost $\langle |T| \rangle$ is proportional to $1/\lambda = \mathcal{O}(1/\Delta)$. According to Thm. 2 this is optimal. A possible problem is that the tail distribution of T is large: Moments of order greater than 2 are unbounded. Lemma 3 shows that this can be remedied. For the unitary case we modify T by restricting to the integers. That is, we set $\text{Prob}(T = n) \propto \text{sinc}(n\lambda)^4$. For $\lambda \leq \pi/4$, the restriction of the characteristic function to $[-\pi, \pi]$ is unchanged (see Lemma 4), so for the case where the eigenstate path is determined by a path of unitary operators, the same average cost of $\mathcal{O}(1/\Delta)$ is obtained.
3. When the eigenstate path is determined by unitary operators, a simple choice of T is the uniform distribution on integers between 0 and $Q - 1$, where $Q = \lceil 2\pi/\Delta \rceil$. If we repeat the randomization step n times, we can bound the error with respect to the desired projection by (Lemma 2)

$$\epsilon = \sup_{\omega_j} |\Phi(\omega_j)|^n \leq \left| \frac{1}{Q} \frac{1 - e^{i\Delta Q}}{1 - e^{i\Delta}} \right|^n \leq \frac{1}{2^n}. \tag{8}$$

The average cost is $n(Q - 1)/2 \in \mathcal{O}(n/\Delta)$. To have error at most ϵ , the cost is $\mathcal{O}(\log(1/\epsilon)/\Delta)$.

If negative T can be used, we can shift T by $-[Q/2]$. This does not affect the absolute values of the characteristic function but reduces the average cost by a constant factor near 1/2.

4. If T is unrestricted, we can consider T with Gaussian distribution $\mathcal{N}(0, \sigma)$. Note that restricting to 0-mean Gaussians minimizes $\langle |T| \rangle$ since the mean and the median coincide. The absolute value of the characteristic function is

$$|\Phi(\omega)| = \exp\left(-\frac{\sigma^2 \omega^2}{2}\right). \tag{9}$$

The error of the randomization step with respect to the desired projection is bounded by

$$|\Phi(\Delta)| = \exp\left(-\sigma^2 \Delta^2 / 2\right). \tag{10}$$

For this distribution, $\langle |T| \rangle = \sigma \sqrt{2/\pi}$. To have error at most ϵ , we need $\sigma \geq 2 \log(1/\epsilon)^{1/2}/\Delta$. This gives an average cost of $\mathcal{O}[\log(1/\epsilon)^{1/2}/\Delta]$.

If T must be positive, we can displace the Gaussian by $x > 0$ and condition on positive outcomes. The error can be estimated as the sum of the probability that the Gaussian is negative, which is bounded by $e^{-x^2/(2\sigma^2)}$, and the right-hand-side of Eq. (10). The average cost is $\mathcal{O}(x + \sigma)$. To have error at most ϵ , let $\sigma = 2 \log(2/\epsilon)^{1/2}/\Delta$ and $x = \sqrt{2\sigma} \log(2/\epsilon)^{1/2}$. The average cost is then $\mathcal{O}[\log(1/\epsilon)/\Delta]$. According to Thm. 3 (below) this is optimal for positive T .

5. In the case where T must be supported on integers, one can try to approximate the Gaussian by the shifted binomial distribution obtained from the sum of $2m$ independent $\{-1/2, 1/2\}$ mean-0 random variables. The absolute value of the characteristic function is

$$|\Phi(\omega)| = |\cos(\omega/2)|^{2m}. \quad (11)$$

This requires $m \in \Theta(\log(1/\epsilon)/\Delta^2)$ to achieve error ϵ in approximating the desired projection. The average cost is then $\mathcal{O}(\log(1/\epsilon)^{1/2}/\Delta)$. As in Example 4, we can shift the distribution by $\Theta[\log(1/\epsilon)^{1/2}/\Delta]$ and condition on positive integers to ensure that T is positive and obtain an average cost of $\mathcal{O}[\log(1/\epsilon)/\Delta]$.

Except for Example 2, the distributions above do not achieve the optimal asymptotic cost for unconstrained T . In the case of Example 2, the probability density determined by $\text{sinc}(\lambda t)^4$ has long tails and unbounded moments. This is improved by the following lemma.

Lemma 3 *There exist probability densities f for T that achieve cost $\langle |T| \rangle = \Theta(1/\Delta)$ and error $|\Phi(\omega)| = 0$ for $|\omega| \geq \Delta$, and that have bounded moments of all order, i.e. $\langle (T - \langle T \rangle)^n \rangle < \infty \forall n \geq 0$.*

We give a constructive proof in Appendix D. For these distributions and $\Delta \leq \pi$, discretization does not result in an increase in the error, see the next lemma. Note that in the discretized case we are only interested in the region of eigenphases $[-\pi, \pi]$ and the relevant gap is the eigenphase gap.

Lemma 4 *Let f be a probability density whose characteristic function has support in $(-\Delta, \Delta)$ with $\Delta \leq \pi$. Then the restriction of f to the integers is a well-defined probability distribution with $\text{prob}(k) = f(k)$ and characteristic function $\Phi(\omega) = 0$ for $|\omega| \in (\Delta, \pi]$.*

We give the proof in Appendix E.

For positive T and if only a lower bound Δ on the gap is known, it is not possible to improve asymptotically over the shifted and conditioned Gaussian distribution of Example 4:

Theorem 3 *Let T be a positive random variable with characteristic function Φ . Then $\sup_{|\omega| \geq \Delta} |\Phi(\omega)| \geq e^{-\Delta \langle T \rangle \pi/2}$.*

The proof is in Appendix F.

3 The randomization method

The goal of the randomization method is to prepare the nondegenerate eigenstate $|\psi(1)\rangle$ of $H(1)$ by traversing of the path $|\psi(s)\rangle$. This path is determined by the family $\mathcal{H} = \{H(s)\}$.

Ideally, we choose the uniform parametrization $s(l)$ discussed in Sec. 2.1 and Appendix B. Under such a parametrization the eigenstates $|\psi(s(l))\rangle$ move at a constant unit rate along the path. Finding the uniform parametrization is difficult in general. We therefore consider an arbitrary *subuniform* parametrization $l \in [0, L'] \mapsto s(l)$ so that the rate at which the states move is bounded by unity. Note that $L' \geq L$, with L the path length. A subuniform parametrization can usually be obtained from known properties of $H(s)$; see Lemma 5 and Eq. (14) below. We discretize the path using $q \in \mathcal{O}((L')^2)$ segments in order to achieve bounded error.

The randomization method uses randomized evolutions \mathcal{R}_l^T to approximate the projective-measurement operations M_l at values $s(l)$. Here, $l = k\delta$ for $k = 1, \dots, q$, with $q = L'/\delta$ and δ sufficiently small. For good asymptotic behavior, we choose T as in Lemma 3 or Example 2. If T must be positive, we use the shifted and conditioned Gaussian distribution of Example 4. If T must be restricted to the integers, as in the case of a path $U(s)$ of unitaries, we use the discretized version of T (Lemma 4). We obtain:

Theorem 4 (Randomization method) *There are choices of q and T in the randomization method such that the method outputs $|\psi(1)\rangle$ starting from $|\psi(0)\rangle$ with fidelity at least p and average cost*

$$\mathcal{O}\left(\frac{(L')^2 (\log(L'/(1-p)))^\alpha}{(1-p)\Delta}\right),$$

where $\alpha = 0$ if T can be negative and $\alpha = 1$ otherwise.

Proof. We choose a step increment $\delta = L'/q$, with $q = \lceil 2(L')^2/(1-p) \rceil$. For this choice, Lemma 1 guarantees that, if we were to implement the projective-measurement operations exactly, the error in the preparation of $|\psi(1)\rangle$ would be bounded by $(1-p)/2$, because $d \leq 1$ for subuniform parametrizations. We need to choose T such that the additional contribution to the error due to the differences between the randomized evolutions and the projective-measurement operations is also bounded by $(1-p)/2$. Suppose that the error according to Thm. 1 is bounded by ϵ . After r steps we have

$$\begin{aligned} & \left\| (M_{l_r} \circ \dots \circ M_{l_1} - R_{l_r}^T \circ \dots \circ R_{l_1}^T)(\rho) \right\|_{\text{tr}} \\ &= \left\| (M_{l_r} \circ \dots \circ M_{l_1} - R_{l_r} \circ M_{l_{r-1}} \circ \dots \circ M_{l_1})(\rho) \right. \\ & \quad \left. + (R_{l_r}^T \circ M_{l_{r-1}} \circ \dots \circ M_{l_1} - R_{l_r}^T \circ \dots \circ R_{l_1}^T)(\rho) \right\|_{\text{tr}} \\ &\leq \left\| (M_{l_r} - R_{l_r})(\sigma) \right\|_{\text{tr}} \\ & \quad + \left\| R_{l_r}(M_{l_{r-1}} \circ \dots \circ M_{l_1} - R_{l_{r-1}}^T \circ \dots \circ R_{l_1}^T)(\rho) \right\|_{\text{tr}} \\ &\leq \epsilon + (r-1)\epsilon = r\epsilon, \end{aligned} \tag{12}$$

where we used the fact that quantum operations are trace-norm contracting, and we implicitly applied induction in the last steps. The desired bound on the error requires $\epsilon \leq (1-p)/(2q) = (1-p)^2/(4(L')^2)$. According to Lemma 3 and Example 4 this can be achieved at an average cost $\langle |T| \rangle$ of $\mathcal{O}(1/\Delta)$ if T can be negative, and $\mathcal{O}(\log(1/\epsilon)/\Delta)$ otherwise. The total cost for the procedure is $\mathcal{O}[q \log(q/(1-p))^\alpha/\Delta]$, and substitution of the value for q yields the claimed bound \square .

For differentiable $H(s)$ and eigenstate path $|\psi(s)\rangle$, we can obtain a subuniform parametrization $s(l)$ from bounds on the derivative of $H(s)$ and the gaps. For this we need the next lemma.

Lemma 5 *Suppose that $H(s)$ is differentiable and $\{|\psi(s)\rangle\}$ is a path of nondegenerate eigenstates of $\{H(s)\}$ with spectral gap $\Delta(s) \neq 0$. Then*

$$\| |\partial_s \psi(s)\rangle \| \leq \frac{\| \partial_s H(s) \|}{|\Delta(s)|}.$$

The proof is in Appendix G.

Define $\|\dot{H}\| = \sup_s \|\partial_s H(s)\|$. We obtain

$$L = \int_0^1 \| |\partial_s \psi(s)\rangle \| ds \leq L' = \frac{\|\dot{H}\|}{\Delta}, \quad (13)$$

with Δ a lower bound to the minimum absolute value of the gap. This L' is achieved for the parametrization

$$s(l) = \frac{\Delta}{\|\dot{H}\|} l, \quad (14)$$

which is subuniform in general. Using this parametrization we obtain the following corollary:

Corollary 1 *Let $H(s)$ be a differentiable path of Hamiltonians and Δ a lower bound on the minimum absolute value of the spectral gap. Then we can prepare $|\psi(1)\rangle$ from $|\psi(0)\rangle$ with bounded error probability at cost*

$$\mathcal{O} \left(\frac{\|\dot{H}\|^2}{\Delta^3} \left(\log \left(\|\dot{H}\|/\Delta \right) \right)^\alpha \right),$$

where $\alpha = 0$ if we can evolve for negative times and $\alpha = 1$ otherwise.

To conclude this section we consider the following two questions: What is the probability that the cost of the randomization method exceeds the average cost by a constant factor? How does the actual path followed by the states obtained in a given instance of the randomization compare to the adiabatic path?

The average cost of the randomization method is $\langle C \rangle = q\langle |T| \rangle$, where q is defined in the proof of Thm. 4, with T the relevant random variable. The probability $\text{prob}(C \geq a\langle C \rangle)$ is therefore at most $1/a$ (Markov's inequality). If the higher-order moments of T are bounded, better bounds can be obtained. In particular, for the distributions whose characteristic functions have smooth, compact support, $\text{prob}(C \geq a\langle C \rangle)$ decreases superpolynomially in a . For T based on Gaussians, the decrease is $e^{-\Omega(a^2)}$. Since C is determined by a sum of q independent instances of $|T|$, better bounds can be obtained for specific choices of T , particularly if q is large. In particular the variance of C is inversely proportional to q if T has finite variance and Chebyshev's inequality or, for sufficiently well-behaved T , large-deviation theory can be applied.

A distinguishing feature of the randomization method is that any given instance involves unitary evolution, which means that the sequence of states obtained is pure. What is the probability (over the randomization of the evolution times) that every state in the sequence of pure states has fidelity at least $1 - \gamma$ with respect to the corresponding eigenstate along the adiabatic path? In view of the proof of Thm. 4, the probability that the state after the

r 'th step has fidelity at least $1 - kr\epsilon$ with respect to $|\tilde{\psi}(l_r)\rangle$ is at least $1/k$ (by Markov's inequality). In particular the fidelity of the last state obtained is at least $1 - k(1 - p)$ with respect to $|\psi(1)\rangle$ with probability $1/k$. One can deduce that many of the states obtained in a typical instance of the randomization method are close to the corresponding states along the adiabatic path. Given that the deviation from the adiabatic path executes a kind of random walk, it is reasonable to conjecture that for appropriate choices of parameters, the probability that all states obtained are close to the adiabatic path is also high.

4 Examples of Quantum computations via evolution randomization

4.1 Unstructured search

In Grover's algorithm [21] we want to find a single marked element \mathcal{S} in a space of $N = 2^n$ elements. For this, we build the Hamiltonian

$$H(s) = -[s|\mathcal{S}\rangle\langle\mathcal{S}| + (1 - s)|+\rangle\langle+|], \tag{15}$$

acting on a set of n qubits. Here, $|+\rangle$ is the equal superposition state and $|\mathcal{S}\rangle$ the solution state, which is the computational basis state corresponding to the marked element. Evolving with $H(s)$ for time t can be done using $\mathcal{O}(t|^{1+\eta})$ conventional oracle calls, with $\eta > 0$ arbitrarily small [22]. For any s , $H(s)$ is nondegenerate in the subspace spanned by $\{|+\rangle, |\mathcal{S}\rangle\}$. If $|\psi(s)\rangle$ is the eigenstate with largest eigenvalue, we seek to prepare $|\psi(1)\rangle = |\mathcal{S}\rangle$ from $|\psi(0)\rangle = |+\rangle$ with sufficiently high probability. Preparation of $|\psi(1)\rangle$ using AQC was studied in Ref. [23].

The energy gap of $H(s)$ can be obtained exactly in the relevant subspace. It is

$$\Delta(s) = \sqrt{1 - 4s(1 - s)(1 - 1/N)}, \tag{16}$$

which is minimized at $s = 1/2$, giving $\Delta = \Delta(1/2) = 1/\sqrt{N}$. The path length L can also be obtained exactly and, for large N , we have $L \approx \pi/2$ (the states $|+\rangle$ and $|\mathcal{S}\rangle$ are almost orthogonal). From Thm. 4 the average cost of the randomization method for constant probability of success is $\mathcal{O}(1/\Delta) \in \mathcal{O}(1/\sqrt{N})$ if the parametrization is uniform. In the large N limit, this parametrization satisfies

$$l(s) \approx \frac{1}{2} \arctan\left(\frac{1}{\sqrt{N}} \frac{1 - s}{1/2 - s}\right), \tag{17}$$

which satisfies $0 \leq l(s) \leq \pi/2$. The randomization method then consists of a sequence of projective-measurements operations at values

$$s_j \approx \frac{1}{2} - \frac{\cot(2l_j)}{2\sqrt{N}} = \frac{1}{2} - \frac{\cot(2j\delta)}{2\sqrt{N}} \tag{18}$$

for some $\delta > 0$. Note that this is the same evolution path as the one considered in Ref. [23], and that the rate of change of s as a function of l is $\Delta(s(l))$. A possible choice for δ is $\pi/4$. At $s = 1/2$, we can implement the phase randomization by evolving under $H(1/2)$ for time 0 or π/Δ , each with probability $1/2$. This is the distribution in Example 1 of Sec. 2.3, and was also used in Ref. [9]. It outputs the desired state almost half the time.

When more than one marked element exist, the above randomization method can still be used to output a solution with bounded error probability: the main effect of adding new

projectors in $H(s)$ is an increased spectral gap $\Delta'(s) \geq \Delta(s)$. Thus, the induced decoherence still simulates an appropriate measurement in the new eigenbasis. If the uniform distribution is used for the randomization, then the algorithm is equivalent to the one discussed in Ref. [24], Sec. 8.4.

4.2 *Quantum simulated annealing*

As the previous example demonstrates, distinguishing between the cost induced by the path length and the one induced by the gap has important advantages, in particular when $L \in \mathcal{O}(1)$. Without this distinction, the actual cost of the method can be highly overestimated. In Ref. [12] we studied quantum simulations of classical annealing processes via evolution randomization. An upper bound on the path length in this case is independent of the minimum spectral gap Γ of the classical Markov chain (i.e., Γ is the difference between 1 and the second largest eigenvalue of the stochastic matrix). Furthermore, Γ can be quadratically increased using Szegedy's quantum walks [25, 26]. For bounded error probability, the randomization method using these walks has a cost $\mathcal{O}(1/\sqrt{\Gamma})$, where we are disregarding the dependency on other parameters such as error probability and path length. It provides a quantum speed-up with respect to simulated annealing using Markov Chain Monte Carlo methods, where the cost is $\mathcal{O}(1/\Gamma)$. Quantum state preparation of Gibbs' states using AQC and the Zeno method was previously studied in Ref. [10], but no quantum speed-up was obtained. Recently, a unitary version of the quantum simulated annealing algorithm (QSA), that uses Grover's fixed point method, was introduced in Ref. [17]. The unitary version improves the dependence of the cost of QSA on output fidelity compared to that in Ref. [12]. However, the scaling in the gap is the same.

Basically, QSA is designed to traverse a coherent version of the classical-state path traversed by classical simulated annealing. The quantum state path is in a Hilbert space of dimension corresponding to the size of the classical state space. The classical annealing path we consider is determined by $\pi_x(\beta) = e^{-\beta E[x]}/\mathcal{Z}(\beta)$, where π_x is the probability of configuration x in the stationary (Gibbs) distribution. E is the associated energy or cost function, β is the inverse temperature, and $\mathcal{Z}(\beta)$ the partition function. The corresponding path in Hilbert space is given by the *quantum Gibbs states* $|\psi(\beta)\rangle = \sum_x \sqrt{\pi_x(\beta)} |x\rangle$. Note that a measurement in the computational basis samples x with probability $\pi_x(\beta)$. Since

$$|\partial_\beta \psi(\beta)\rangle = \sum_x (\langle E \rangle - E[x]) \sqrt{\pi_x}/2 |x\rangle, \quad (19)$$

we obtain the following lemma.

Lemma 6 For $\beta \in (0, \beta_f)$,

$$\| |\partial_\beta \psi(\beta)\rangle \| = \sigma(\beta)/2,$$

where $\sigma(\beta)$ is the standard deviation of E at inverse temperature β . The path length satisfies $L \leq \beta_f \sigma/2$, with $\sigma = \sup_\beta \sigma(\beta)$.

If d' is the size of the classical state space and γ is the spectral gap of E , then the state $|\psi(\beta_f)\rangle$, for $\beta_f = \mathcal{O}((\log d')/\gamma)$, has high probability amplitude in the configuration that minimizes E . With this β_f , we have

$$L \in \mathcal{O}\left(\frac{\sigma \log d'}{\gamma}\right). \quad (20)$$

That L is bounded independently of Γ is fundamental for the success of QSA. Using Szegedy's quantum walks we can boost the gap towards $\mathcal{O}(\sqrt{\Gamma})$ and achieve the desired cost. The details of this procedure are explained in Refs. [26, 12].

The QSA is basically a sequence of steps, each constructed to prepare the states $|\psi(\delta)\rangle, |\psi(2\delta)\rangle, \dots, |\psi(\beta_f)\rangle$ from the initial state $|\psi(0)\rangle$, $\delta \ll 1$. According to Ref. [12], these states can be prepared by a version of the Zeno effect in which, at each step, the corresponding Szegedy walk is applied a random number of times (see Example 3, Sec. 2.3). For this distribution the cost of the QSA is

$$\mathcal{O}\left(\frac{L^2}{\sqrt{\Gamma}} \log L\right) \in \mathcal{O}\left(\frac{\sigma^2 \log^2 d'}{\gamma^2 \sqrt{\Gamma}} \log(\sigma \log d'/\gamma)\right),$$

with Γ the minimum gap of the Markov chain along the path. The results in Sec. 2.3 show that using the inverses of the quantum walk steps, the second logarithmic factor can be dropped. Because of the way the quantum walk is constructed, circuits for the inverses can be obtained by direct reversal of the circuits for the quantum walk steps. In Ref. [17] the authors show that a coherent (non-monotonic) path traversal that uses Grover's fixed point method for this case can be implemented with an improved cost $\mathcal{O}(L \log^2 L / \sqrt{\Gamma})$.

5 Relation to other work

It has been noted previously [9, 10, 27, 12, 17] that the projective-measurement operations M_l can be simulated using Kitaev's phase estimation algorithm [11] in the discrete-time case. This requires implementing unitaries $U_l = e^{-i\tilde{H}(l)}$ controlled on r ancillary qubits initialized in the equal superposition state. Then the inverse of the quantum Fourier transform is applied to the ancillary qubits, and a projective measurement on the computational basis of the ancillae is performed [see Fig. 1(a)]. The phase estimation algorithm needs to resolve the desired eigenphase from other eigenphases to be able to project the state of the system into the desired eigenstate. This requires $2^r \in \Omega(1/\tilde{\Delta}(l))$ uses of controlled- U_l 's for constant error. The error per step has to be small. If one of the high-confidence versions of the phase estimation algorithm [28] is used, the overhead to achieve error ϵ is logarithmic in $1/\epsilon$. The overall cost is then similar to that of the randomization method when T is restricted to be positive.

Interestingly, the phase-estimation-based algorithm produces the same effect on the system as the randomization method if we sample the evolution time from the uniform distribution on an interval. This is because the phase estimation ancillary qubits can be traced out after each step. As a result, the inverse quantum Fourier transform can also be dropped. Consequently, the coherence in the state of the ancillary qubits, initialized in the equal superposition state, plays no role and these qubits can be replaced by classical bits, each being 0 or 1 with probability 1/2. This equivalence was also studied in Ref. [24]. We illustrate it in Fig. 1(b).

Repeating the phase estimation algorithm n times is equivalent to randomizing with the sum of n independent uniform distributions. This was considered in Example 3, Sec. 2.3. The unwanted coherences reduce exponentially in n .

There are previously noted relationships between the Zeno effect and coherent evolutions similar to the continuous or discrete evolutions used in the randomization method. For example, the effect of a strong interaction with another system, such as might occur in the

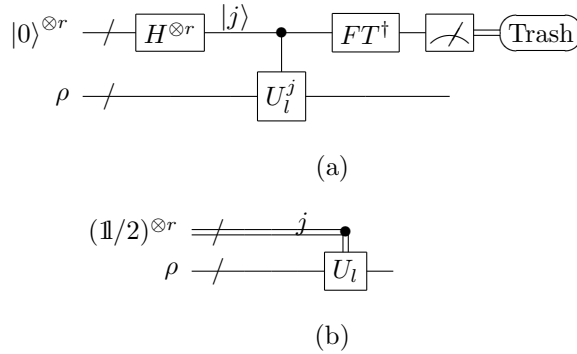


Fig. 1. (a) Phase estimation algorithm. At the end of the algorithm, the top r -qubit register encodes a r -bit approximation to an eigenphase of U_l on readout. It is initialized with Hadamard gates to an equal superposition state. A sequence of $2^r - 1$ controlled U_l^j operations is applied, and the first register is measured after an inverse quantum Fourier transform. If the measurement outcome approximates an eigenphase of U_l , the second register (system) is approximately projected onto the corresponding eigenstate. (b) Randomized evolution. If the phase estimation algorithm outcome is ignored, the overall effect is equivalent to the one induced by initializing a set of r bits (first register) in a random state j , with $j \in [0..2^r - 1]$, and by acting with U_l^j . Double lines indicate classical information.

coupling to a measurement apparatus, is to restrict the natural Hamiltonian to the eigenspaces of the interaction [29]. The suppression of coherent transitions by randomization with the interaction Hamiltonian would have a similar effect. A discrete version of this observation relevant to the analysis of dynamical decoupling was considered in [30].

There is a relationship between the way in which interactions are averaged away in dynamical decoupling, particularly randomized dynamical decoupling [31, 32], and how transitions between the adiabatic path and the other eigenstates are suppressed in the randomization method. The relationship can be made explicit by changing to an s -dependent frame in which the Hamiltonians $H(s)$ are diagonal. In this frame, the transitions show up explicitly due to the frame changes with s . Strong or randomized evolution under $H(s)$ suppresses these transitions by averaging them to zero. Dynamical decoupling typically uses operators that have stronger averaging effects.

A feature of the randomization method is the use of phase decoherence to ensure a more efficient transfer to a state of physical or computational interest. There are other ways in which decoherence can play a role in preparing states for quantum computing. Early examples proposed the synthesis of *decoherence free subspaces* from an environment-induced quantum Zeno effect [33, 34, 35]. The use of decoherence to decrease the mixing time of quantum walks was proposed in Ref. [36]. A related phenomenon has been studied in the context of energy transfer [37, 38, 39, 40] as realized in certain biological molecules. More generally, it may be that decoherence or thermal noise can enhance the success probability in adiabatic quantum computing [41, 42, 43, 44]. Note that the required thermal noise is different from the phase-decoherence associated with the randomization method in that it has the potentially desirable effect of transferring population to lower-energy eigenstates of the currently active Hamiltonian. Whether the requirements for effective exploitation of this situation can be met in realistic devices is not clear.

Finally, the engineering of a dissipative process using feedback techniques to stabilize a desired quantum state has been studied extensively in the quantum control literature [45, 46, 47, 48]. Reference [49] advances that the final state of AQC can be made the unique steady state of the dissipation, even without feedback. In essence, the process' Lindblad operators encode the gates of the quantum computation and appropriate updates to a logical clock register. Again, the necessary dissipation requires more than the phase decoherence realized by the randomization method.

6 Conclusions

We have described a method for state preparation in the spirit of AQC, but based exclusively on randomized evolutions. The idea is to perform a discrete sequence of projective measurement operations onto the desired (instantaneous) eigenstate of a given Hamiltonian or unitary path. These operations are induced via evolution randomization, which realizes the necessary decoherence in the eigenbasis. We bound the residual coherences after the randomization in terms of the characteristic function of the random time.

We obtained the following exact bounds on the dephasing achieved by randomized evolutions: First, to induce enough decoherence, the average evolution time per step scales with the inverse of the minimum absolute value of the spectral (or eigenphase) gap. Second, repetition of the randomization reduces the coherences exponentially in the amount of repetitions. Third, if negative-time evolutions are implementable with constant overhead, logarithmic factors depending on the error can be reduced to constant factors, even for discretized evolutions. Fourth, for non-negative evolutions and if only a lower bound on the absolute value of the gaps is known, the logarithmic overhead is unavoidable.

We show that the complexity of path traversal algorithms is best expressed in terms of the path length L . The explicit dependence of the complexity on L can be very helpful when L does not depend on the gap. This happens, for example, in the Hamiltonian version of an algorithm for unstructured search, where we showed that a simple choice of step size and random time distributions rotates into the solution state with probability $1/2$. One further advantage of the path-length formulation is that we do not require the relatively strong differentiability requirements on H as in the proofs of the adiabatic condition [13] with explicit bounds as needed for AQC.

Another case where L does not depend on the gap is in the quantum simulated annealing algorithm, which we also analyzed. This algorithm provides a quadratic quantum speed-up in terms of the gap with respect to classical simulated annealing implemented via Markov Chain Monte Carlo methods. The path is determined by an annealing schedule in which a parameter β , related to the inverse temperature of a classical system, is slowly increased in equal-size steps. The quantum simulated annealing algorithm allows us to reach the optimal configuration in time $\mathcal{O}(1/\sqrt{\Gamma})$ for constant probability of success and path length, with Γ being the minimum gap of the stochastic matrix (and the corresponding Hamiltonian) along the path. The improved randomization methods given here remove a logarithmic factor for the version of the algorithm given in Ref. [12].

The similarities of the randomization method with AQC are clear: A typical instantiation (choice of evolution times) of the randomization method is, with high probability, an approximation to an adiabatic path. We find, as is often the case, that it is easier to prove error

bounds for random instances than for the worst case. Whether the existence statement can be “derandomized” efficiently is still an interesting question.

Acknowledgements

We thank Howard Barnum for discussions. This work was supported by Perimeter Institute for Theoretical Physics, by the Government of Canada through Industry Canada and by the Province of Ontario through the Ministry of Research and Innovation. This work was also supported by the National Science Foundation under grant PHY-0803371 through the Institute for Quantum Information at the California Institute of Technology. Contributions to this work by NIST, an agency of the US government, are not subject to copyright laws. SB thanks the Laboratory Directed Research and Development Program at Los Alamos National Laboratory for support during the initial stages of this work.

References

1. E. Farhi, J. Goldstone, S. Gutmann, and M. Sipser (2000), *Quantum computation by adiabatic evolution*, arxiv:quant-ph/0001106.
2. B. Apolloni, N. Cesa-Bianchi, and D. de Falco (1988), *A numerical implementation of quantum annealing*, in Proceedings of the Ascona-Locarno Conference, World Scientific.
3. B. Apolloni, C. Carvalho, and D. de Falco (1988), *Quantum stochastic optimization*, Stochastic Processes and their Applications, vol. 33, p. 233.
4. T. Kadowaki and H. Nishimori (1998), *Quantum annealing in the transverse ising model*, Phys. Rev. E, vol. 58, p. 5355.
5. G. E. Santoro, R. Martonak, E. Tosatti, and R. Car (2002), *Theory of quantum annealing of an ising spin glass*, Science, vol. 295, pp. 2427–2430.
6. A. Das and B. K. Chakrabarti (2005), *Quantum Annealing and Related Optimization Methods*, vol. 679. Springer, Heidelberg.
7. A. Messiah (1999), *Quantum Mechanics*. Dover Publications.
8. D. Aharonov, W. van Dam, J. Kempe, Z. Landau, S. Lloyd, and O. Regev (2007), *Adiabatic quantum computation is equivalent to standard quantum computation*, SIAM Journal on Computing, vol. 37, pp. 166–194.
9. A. M. Childs, E. Deotto, E. Farhi, J. Goldstone, S. Gutmann, and A. J. Landahl (2002), *Quantum search by measurement*, Phys. Rev. A, vol. 66, p. 032314.
10. D. Aharonov and A. Ta-Shma (2003), *Adiabatic quantum state generation and statistical zero knowledge*, Proceedings of the thirty-fifth ACM symposium on Theory of computing, pp. 20–29.
11. A. Y. Kitaev (1995), *Quantum measurements and the abelian stabilizer problem*, arxiv:quant-ph/9511026.
12. R. D. Somma, S. Boixo, H. Barnum, and E. Knill (2008), *Quantum simulations of classical annealing processes*, Phys. Rev. Lett., vol. 101, pp. 130504–4.
13. S. Jansen, M.-B. Ruskai, and R. Seiler (2006), *Bounds for the adiabatic approximation with applications to quantum computation*, arxiv:quant-ph/0603175.
14. D. A. Lidar, A. T. Rezakhani, and A. Hamma (2008), *Adiabatic approximation with better than exponential accuracy for many-body systems and quantum computation*, arxiv:0808.2697.
15. J. Goldstone (2008), *Quantum computation beyond the circuit model*, in S. P. Jordan PhD thesis, arxiv:0809.2307.
16. G. A. Hagedorn and A. Joye (2002), *Elementary exponential error estimates for the adiabatic approximation*, Journal of Mathematical Analysis and Applications, vol. 267, pp. 235–246.
17. A. Abeyesinghe and P. Wocjan (2008), *Speed-up via quantum sampling*, 0804.4259.
18. B. Misra and E. C. G. Sudarshan (1977), *The zeno’s paradox in quantum theory*, Journal of Mathematical Physics, vol. 18, pp. 756–763.

19. D. Home and M. A. B. Whitaker (1997), *A conceptual analysis of quantum zeno; paradox, measurement, and experiment*, *Annals of Physics*, vol. 258, pp. 237–285.
20. H. Nakazato, M. Namiki, and S. Pascazio (1996), *Temporal behavior of quantum mechanical systems*, *Int. J. Mod. Phys B*, vol. 10, no. 3, pp. 247–295.
21. L. K. Grover (1996), *A fast quantum mechanical algorithm for database search*, *Proceedings of the twenty-eighth annual ACM symposium on Theory of computing*, pp. 212–219.
22. R. Cleve, D. Gottesman, M. Mosca, R. D. Somma, and D. L. Yonge-Mallo (, 2008), *Efficient discrete-time simulations of continuous-time quantum query algorithms*, arxiv:0811.4428.
23. J. Roland and N. J. Cerf (2002), *Quantum search by local adiabatic evolution*, *Phys. Rev. A*, vol. 65, p. 042308.
24. P. Kaye, R. Laflamme, and M. Mosca (2007), *An Introduction to Quantum Computing*. Oxford University Press.
25. A. Ambainis (2007), *Quantum walk algorithm for element distinctness*, *SIAM Journal on Computing*, vol. 37, pp. 210–239.
26. M. Szegedy (2004), *Quantum speed-up of markov chain based algorithms*, *Proceedings of the 45th Annual IEEE Symposium on Foundations of Computer Science*, pp. 32–41.
27. F. Magniez, A. Nayak, J. Roland, and M. Santha (2006), *Search via quantum walk*, arxiv:quant-ph/0608026.
28. E. Knill, G. Ortiz, and R. D. Somma (2007), *Optimal quantum measurements of expectation values of observables*, *Phys. Rev. A*, vol. 75, pp. 012328/1–13.
29. P. Facchi and S. Pascazio (2002), *Quantum zeno subspaces*, *Phys. Rev. Lett.*, vol. 89, no. 8, p. 080401.
30. P. Facchi, D. A. Lidar, and S. Pascazio (2004), *Unification of dynamical decoupling and the quantum zeno effect*, *Phys. Rev. A*, vol. 69, p. 032314.
31. L. Viola and E. Knill (2005), *Random decoupling schemes for quantum dynamical control and error suppression*, *Phys. Rev. Lett.*, vol. 94, pp. 060502–4.
32. L. F. Santos and L. Viola (2008), *Advantages of randomization in coherent quantum dynamical control*, *New J. of Phys.*, vol. 10, no. 8, p. 083009.
33. A. Beige, D. Braun, B. Tregenna, and P. L. Knight (2000), *Quantum computing using dissipation to remain in a Decoherence-Free subspace*, *Phys. Rev. Lett.*, vol. 85, no. 8, p. 1762.
34. B. Tregenna, A. Beige, and P. L. Knight (2002), *Quantum computing in a macroscopic dark period*, *Phys. Rev. A*, vol. 65, p. 032305.
35. A. Beige (2003), *Quantum computing using dissipation (proceedings)*, *Inst. Phys. Conf. Ser.*, vol. 173, no. 35.
36. P. C. Richter (2007), *Quantum speedup of classical mixing processes*, *Phys. Rev. A*, vol. 76, pp. 042306–8.
37. M. Mohseni, P. Rebentrost, S. Lloyd, and A. Aspuru-Guzik (2008), *Environment-assisted quantum walks in photosynthetic energy transfer*, *J. Chem. Phys.*, vol. 129, pp. 174106–9.
38. P. Rebentrost, M. Mohseni, and A. Aspuru-Guzik (2008), *Role of quantum coherence in chromophoric energy transport*, arxiv:0806.4725.
39. P. Rebentrost, M. Mohseni, I. Kassal, S. Lloyd, and A. Aspuru-Guzik (2008), *Environment-Assisted quantum transport*, arxiv:0807.0929.
40. F. Caruso, A. W. Chin, A. Datta, S. F. Huelga, and M. B. Plenio (2009), *Fundamental mechanisms of noise supported energy transfer in biological systems*, arxiv:0901.4454.
41. A. M. Childs, E. Farhi, and J. Preskill (2001), *Robustness of adiabatic quantum computation*, *Phys. Rev. A*, vol. 65, p. 012322.
42. M. H. S. Amin, P. J. Love, and C. J. S. Truncik (2008), *Thermally assisted adiabatic quantum computation*, *Phys. Rev. Lett.*, vol. 100, no. 6, p. 60503.
43. M. H. S. Amin, C. J. S. Truncik, and D. V. Averin (2008), *The role of single qubit decoherence time in adiabatic quantum computation*, arxiv:0803.1196.
44. S. Lloyd (2008), *Robustness of adiabatic quantum computing*, arxiv:0805.2757.
45. J. Wang and H. M. Wiseman (2001), *Feedback-stabilization of an arbitrary pure state of a two-level*

- atom, Phys. Rev. A, vol. 64, p. 063810.
46. H. M. Wiseman, S. Mancini, and J. Wang (2002), *Bayesian feedback versus markovian feedback in a two-level atom*, Phys. Rev. A, vol. 66, p. 013807.
 47. F. Ticozzi and L. Viola (2008), *Quantum markovian subsystems: Invariance, attractivity, and control*, IEEE Transactions on Automatic Control, vol. 53, no. 9, pp. 2048–2063.
 48. F. Ticozzi and L. Viola (2008), *Analysis and synthesis of attractive quantum markovian dynamics*, arxiv:0809.0613.
 49. F. Verstraete, M. M. Wolf, and J. I. Cirac (2008), *Quantum computation, quantum state engineering, and quantum phase transitions driven by dissipation*, arxiv:0803.1447.
 50. B. Porat (1996), *A Course in Digital Signal Processing*. John Wiley & Sons.
 51. H. Dym and H. P. McKean (1972), *Fourier Series and Integrals*. New York and London: Academic Press.

Appendix A

Path length. For states $|\phi_1\rangle$ and $|\phi_2\rangle$, let $\Theta(|\phi_1\rangle, |\phi_2\rangle) = \arccos(|\langle\phi_2|\phi_1\rangle|)$ be the angular distance between the states. We assume that the $|\psi(s)\rangle$ form a projectively continuous path, $s \in [0, 1]$. The length is given by

$$L = \sup_{(s_k)} \sum_k \Theta(|\psi(s_{k+1})\rangle, |\psi(s_k)\rangle), \quad (\text{A.1})$$

where the ordered sequences (s_k) subdivide $[0, 1]$. Note that the expression in the limit depends monotonically on the (s_k) , increasing in the refinement order. If $|\psi(s)\rangle$ is differentiable, the expression in Eq. (A.1) reduces to the one in Eq. (2).

Appendix B

Uniform parametrization. Let $\tilde{L}(l)$ be the length of the path $|\tilde{\psi}(l')\rangle$ for $0 \leq l' \leq l$, defined as in Eq. (A.1). Suppose that $\tilde{L}(l)$ is Lipschitz continuous so that $\omega(l_1, l_2) = \sup_{l_1 \leq l' < l'' \leq l_2} (\tilde{L}(l'') - \tilde{L}(l')) / (l'' - l')$ is finite. Note that if \tilde{L} is differentiable, one can take $\omega(l_1, l_2) = \sup_{l_1 \leq l \leq l_2} \frac{d\tilde{L}(l)}{dl}$. In particular, if $|\tilde{\psi}(l)\rangle$ is differentiable, $\omega(l_1, l_2) = \sup_{l_1 \leq l \leq l_2} \|\partial_l |\tilde{\psi}(l)\rangle\|$ works [see Eq. (2)]. We obtain:

Lemma 7 *The squared overlap $|\langle\tilde{\psi}(l+\delta)|\tilde{\psi}(l)\rangle|^2$ can be bounded by*

$$|\langle\tilde{\psi}(l+\delta)|\tilde{\psi}(l)\rangle|^2 \geq 1 - \omega(l, l+\delta)^2 \delta^2.$$

Proof. We have

$$\begin{aligned} |\langle\tilde{\psi}(l+\delta)|\tilde{\psi}(l)\rangle|^2 &= \cos^2(\Theta(|\tilde{\psi}(l+\delta)\rangle, |\tilde{\psi}(l)\rangle)) \\ &\geq 1 - \Theta(|\tilde{\psi}(l+\delta)\rangle, |\tilde{\psi}(l)\rangle)^2 \\ &\geq 1 - (\tilde{L}(l+\delta) - \tilde{L}(l))^2 \\ &\geq 1 - \omega(l+\delta, l)^2 \delta^2 \end{aligned} \quad (\text{B.1})$$

□.

To take advantage of Lemma 1, it helps to parametrize the path with an $s(l)$ for which $\omega(l_1, l_2)$ is as uniform as possible. For this purpose, define $s(l) = \inf\{s : L(s) \geq l\}$ for $0 \leq l \leq L$, where the length $L(s)$ is the length of the path $|\psi(s')\rangle$, $0 \leq s' \leq s$. The function $s(l)$ is not necessarily continuous.

Continuity of states and finiteness of L implies continuity of $L(s)$. This can be shown as follows: Suppose that $L(s)$ is not continuous at s . Then either $\sup_{\delta>0} L(s - \delta) < L(s)$ or $\inf_{\delta>0} L(s + \delta) > L(s)$. Consider the first case. We have $L(s) = \limsup_{\delta>0} (\Theta(|\psi(s)\rangle, |\psi(s - \delta)\rangle) + L(s - \delta))$. The inequality implies that $\limsup_{\delta>0} \Theta(|\psi(s)\rangle, |\psi(s - \delta)\rangle) > 0$, contradicting continuity of $|\psi(s)\rangle$. For the second case, $s < 1$. Define $L(s_1, s_2)$ as the length of the path from $|\psi(s_1)\rangle$ to $|\psi(s_2)\rangle$. It can be seen from the definition, monotonicity in the refinement order of the term in the limit of the definition, and from projective continuity of $|\psi(s)\rangle$ that $L(s, 1) = \sup_{\delta>0} L(s + \delta, 1) + \Theta(|\psi(s)\rangle, |\psi(s + \delta)\rangle) = \sup_{\delta>0} L(s + \delta, 1)$ and $L(s, 1) = L(s + \delta, 1) + L(s, s + \delta)$. It follows that $\inf_{\delta>0} L(s, s + \delta) = 0$. The observation now follows from $L(s + \delta) = L(s) + L(s, s + \delta)$.

We define $\tilde{L}(l)$ as the length of the path $|\tilde{\psi}(l')\rangle = |\psi(s(l'))\rangle$ for $0 \leq l' \leq l$. We show that $\tilde{L}(l) = L(s(l)) = l$. The second inequality follows from continuity of L and the definitions. From the definition of path length and since any subdivision (l_k) of $[0, l]$ corresponds to a subdivision $(s(l_k))$ of $[0, s(l)]$, $\tilde{L}(l) \leq L(s(l))$. To show the reverse inequality, let $\bar{s} = s(L(s))$. Then $\bar{s} \leq s$ and $\Theta(|\psi(\bar{s})\rangle, |\psi(s)\rangle) = 0$. Hence for all $s' \in [\bar{s}, s]$, $|\psi(s')\rangle \propto |\psi(\bar{s})\rangle$ (that is, the two states are projectively identical). Consequently, the right-hand side of Eq. (A.1) is unchanged if we replace the s_k by \bar{s}_k . Since the \bar{s}_k are in the range of $l \mapsto s(l)$, we can choose $l_k = L(\bar{s}_k)$ to show that the defining supreme for $\tilde{L}(l)$ and for $L(s(l))$ are the same.

By the previous paragraph, $\omega(l_1, l_2) = 1$ for the parametrization $s(l)$. We therefore refer to $s(l)$ as the *uniform* parametrization.

Appendix C

Proof of Theorem 1 Let μ be the probability distribution of T . For any \mathcal{E} ,

$$\begin{aligned}
 & (\mathcal{R}_l^T - M_l^{\mathcal{E}})(|\tilde{\psi}(l)\rangle\langle\tilde{\psi}_j(l)|) \\
 &= \mathcal{R}_l^T(|\tilde{\psi}(l)\rangle\langle\tilde{\psi}_j(l)|) \\
 &= \int e^{-i\tilde{H}(l)t}(|\tilde{\psi}(l)\rangle\langle\tilde{\psi}_j(l)|)e^{i\tilde{H}(l)t}d\mu(t) \\
 &= \int e^{i\omega_j t}d\mu(t)|\tilde{\psi}(l)\rangle\langle\tilde{\psi}_j(l)| \\
 &= \Phi(\omega_j)|\tilde{\psi}(l)\rangle\langle\tilde{\psi}_j(l)|.
 \end{aligned} \tag{C.1}$$

We assume without loss of generality that ρ is pure, $\rho = |\phi\rangle\langle\phi|$. Write

$$|\phi\rangle = c_1|\tilde{\psi}(l)\rangle + \sum_{j>1} c_j|\tilde{\psi}_j(l)\rangle. \tag{C.2}$$

Let \mathcal{S} be the subspace orthogonal to $|\tilde{\psi}(l)\rangle$. The operation \mathcal{R}_l^T leaves \mathcal{S} invariant, and we can

choose $\mathcal{E} = \mathcal{R}_l^T$ in that subspace. Then

$$\begin{aligned} & \| (M_l^{\mathcal{E}} - \mathcal{R}_l^T)(|\phi\rangle\langle\phi|) \|_{\text{tr}} \\ &= \left\| \mathcal{R}_l^T \left(\sum_{j>1} c_1 c_j^* |\tilde{\psi}(l)\rangle\langle\tilde{\psi}_j(l)| + \text{h.c.} \right) \right\|_{\text{tr}} \\ &= \left\| \sum_{j>1} \left(\Phi(\omega_j) c_1 c_j^* |\tilde{\psi}(l)\rangle\langle\tilde{\psi}_j(l)| + \text{h.c.} \right) \right\|_{\text{tr}} . \end{aligned} \tag{C.3}$$

This is the trace norm of a matrix having

$$\pm \sqrt{\sum_{j>1} |\Phi(\omega_j) c_1 c_j^*|^2} \tag{C.4}$$

as the only non-zero eigenvalues. Because of the normalization, $|c_1|^2 \sum_{j>1} |c_j|^2 \leq 1/4$. Thus

$$\begin{aligned} & \| (M_l^{\mathcal{E}} - \mathcal{R}_l^T)(|\phi\rangle\langle\phi|) \|_{\text{tr}} \\ &= 2 \sqrt{\sum_{j>1} |\Phi(\omega_j) c_1 c_j^*|^2} \\ &\leq \sup_{\omega_j} |\Phi(\omega_j)| 2 \sqrt{\sum_{j>1} |c_1 c_j^*|^2} \\ &\leq \sup_{\omega_j} |\Phi(\omega_j)| . \end{aligned} \tag{C.5}$$

Appendix D

Proof of Lemma 3. We start with any smooth even function \hat{h} of compact support in $(-1/2, 1/2)$. This implies that its inverse Fourier transform h is real and all its moments are bounded since

$$|\langle X^n \rangle| = \left| \int_{-\infty}^{+\infty} h(x) x^n dx \right| = \left| \frac{\partial^n \hat{h}(0)}{\partial \omega^n} \right| < \infty . \tag{D.1}$$

We define the characteristic function Φ_1 to be proportional to the convolution of \hat{h} with itself,

$$\Phi_1(\omega) \propto (\hat{h} * \hat{h})(\omega) . \tag{D.2}$$

We normalize such that $\Phi_1(0) = 1$. By construction, the inverse Fourier transform of Φ_1 , denoted by f_1 , is positive, normalized to 1, and rapidly decaying, as desired. To accommodate arbitrary spectral gaps $\Delta > 0$, we rescale the characteristic function as $\Phi_\Delta(\omega) = \Phi(\omega/\Delta)$, which has support in $(-\Delta, \Delta)$. Its inverse Fourier transform is a probability density function

$f_\Delta(t) = \Delta f_1(\Delta t)$. The cost of randomization with f_Δ is

$$\begin{aligned} \langle |T| \rangle_\Delta &= \int_{-\infty}^{+\infty} |t| f_\Delta(t) dt \\ &= \Delta \int_{-\infty}^{+\infty} |t| f_1(\Delta t) dt \\ &= \frac{\langle |T| \rangle_1}{\Delta}. \end{aligned} \tag{D.3}$$

where $\langle |T| \rangle_1$ is the cost of randomization with f_1 , and is independent of Δ . It follows that $\langle |T| \rangle_\Delta \in \Theta(1/\Delta)$, which is optimal.

Appendix E

Proof of Lemma 4. Consider a probability density f with characteristic function Φ of support in $(-\Delta, \Delta)$, where $\Delta \leq \pi$. Consider $\sum_k \Phi(\omega + 2\pi k) = (\Phi * \hat{C})(\omega)$ where $\hat{C}(\omega) = \sum_k \delta(\omega - 2\pi k)$ is a comb. As a distribution, $\hat{C}(\omega)$ is the Fourier transform of the comb $C(t) = \sum_k \delta(t - k)/(2\pi)$. See, for example, Sec. 2.4 of [50]. Using the rules for convolution under the inverse Fourier transform, we find that the distribution $(f \cdot C)(t) = \sum_k f(k)\delta(t - k)$ has Fourier transform $\Phi * \hat{C}$. Because $(\Phi * \hat{C})(0) = \Phi(0) = 1$, it follows that $f(k)$ is a probability distribution with the stated properties.

Appendix F

Proof of Theorem 3. For $\langle T \rangle$ infinite, there is nothing to prove. So assume $\langle T \rangle$ is finite, which implies that the characteristic function is differentiable. Suppose first that T has a square-integrable probability density $f(t)$. The characteristic function is then a ‘‘Hardy function’’ of class H^{2+} as defined in Ref. [51], pg. 162. By noting that for $\alpha > 0$, $\omega \mapsto \Phi(\alpha\omega)$ is also Hardy, the proof of Thm. 2 on pg. 166 of Ref. [51] shows that

$$\int_{-\infty}^{+\infty} \frac{\log |\Phi(\alpha\gamma)|}{1 + \gamma^2} d\gamma \geq \pi \log |\Phi(\alpha i)|, \tag{F.1}$$

where Φ has been analytically extended to the upper half plane. The analytical extension of Φ is obtained by using complex ω in the Fourier transform. Consequently, $d\Phi(z)/dz$ is the Fourier transform of $t \mapsto itf(t)$, where defined. In particular, $|d\Phi(z)/dz|$ is bounded by $\langle T \rangle$ for $z = i\beta$ with $\beta \geq 0$. Since $\Phi(0) = 1$, we have $\log |\Phi(\alpha i)| \geq \log(1 - \alpha\langle T \rangle)$. The integral of the inequality in Eq. (F.1) can be related to the desired supremum as follows:

$$\begin{aligned} &\int_{-\infty}^{+\infty} \frac{\log |\Phi(\alpha\gamma)|}{1 + \gamma^2} d\gamma \\ &= \alpha \int_{-\infty}^{+\infty} \frac{\log |\Phi(\gamma)|}{\alpha^2 + \gamma^2} d\gamma \\ &\leq \alpha \int_{-\Delta}^{+\Delta} \frac{\log |\Phi(\gamma)|}{\alpha^2 + \gamma^2} d\gamma \\ &\quad + \alpha \log(\sup_{|\gamma| \geq \Delta} |\Phi(\gamma)|) \int_{|\gamma| \geq \Delta} \frac{1}{\alpha^2 + \gamma^2} d\gamma \\ &\leq \log(\sup_{|\gamma| \geq \Delta} |\Phi(\gamma)|) (\pi - 2 \arctan(\Delta/\alpha)). \end{aligned} \tag{F.2}$$

To drop the first summand in the last step we used the fact that $|\Phi(\gamma)| \leq 1$ because Φ is the characteristic function of a probability distribution. We now let $\alpha \rightarrow 0^+$ and combine with the earlier inequality to get, to first order in α ,

$$-\pi\alpha\langle T \rangle \leq 2 \log\left(\sup_{|\gamma| \geq \Delta} |\Phi(\gamma)|\right)\alpha/\Delta, \quad (\text{F.3})$$

which gives $e^{-\Delta\langle T \rangle \frac{\pi}{2}} \leq \sup_{|\gamma| \geq \Delta} |\Phi(\gamma)|$.

Now consider arbitrary positive T with $\langle T \rangle < \infty$, and with probability distribution μ . Let S_δ be uniformly distributed between 0 and δ . The probability distribution of $T + S$ has cumulative distribution $F(x) = \int_0^x \min(1, (x-y)/\delta) d\mu(y)$, which is differentiable. The corresponding probability density is given by $\mu([x-\delta, x])/\delta = \int_{x-\delta}^x d\mu(y)/\delta$ and is square integrable because

$$\begin{aligned} \int \mu([y-\delta, y])^2 dy &\leq \int \mu([y-\delta, y]) dy \\ &= \int \int_{y-\delta}^y d\mu(z) dy \\ &= \int \int_z^{z+\delta} dy d\mu(z) \end{aligned} \quad (\text{F.4})$$

$$= \int \delta d\mu(z) = \delta. \quad (\text{F.5})$$

Thus $T + S$ is subject to the bound of the Theorem. The characteristic function of $T + S$ is given by $\Phi(\omega)s_\delta(\omega)$, where $s_\delta(\omega)$ is the characteristic function of S_δ . The function $s_\delta(\omega)$ converges uniformly to 1 on bounded intervals as $\delta \rightarrow 0^+$. It follows that the desired bound applies to arbitrary positive T .

Appendix G

Proof of Lemma 5. Without loss of generality, the phases of $|\psi(s)\rangle$ are geometric. Because $\Delta(s) > 0$ and $H(s)$ is differentiable, it follows that $|\psi(s)\rangle$ is differentiable. From the eigenvalue equation

$$H(s)|\psi(s)\rangle = E(s)|\psi(s)\rangle, \quad (\text{G.1})$$

we get

$$\begin{aligned} \partial_s H(s)|\psi(s)\rangle + H(s)|\partial_s \psi(s)\rangle = \\ \partial_s E(s)|\psi(s)\rangle + E(s)|\partial_s \psi(s)\rangle. \end{aligned} \quad (\text{G.2})$$

Denote by $|\psi_j(s)\rangle$, $j \in \{2, \dots, d\}$, the j -th eigenstate of $H(s)$, orthogonal to $|\psi(s)\rangle$, and with eigenvalue $E_j(s)$. We obtain

$$\langle \psi_j(s) | \partial_s \psi(s) \rangle = \frac{\langle \psi_j(s) | \partial_s H(s) | \psi(s) \rangle}{E(s) - E_j(s)}. \quad (\text{G.3})$$

Because the path $|\psi(s)\rangle$ is geometric, $\langle\psi(s)|\partial_s\psi(s)\rangle = 0$ for all s . This gives

$$\begin{aligned} \|\partial_s\psi(s)\|^2 &= \sum_{j\geq 2} \frac{|\langle\psi_j(s)|\partial_s H(s)|\psi(s)\rangle|^2}{|E(s) - E_j(s)|^2} \\ &\leq \frac{1}{\Delta(s)^2} \sum_{j\geq 2} \langle\psi(s)|\partial_s H|\psi_j(s)\rangle \langle\psi_j(s)|\partial_s H(s)|\psi(s)\rangle \\ &\leq \frac{1}{\Delta(s)^2} \langle\psi(s)|(\partial_s H(s))^2|\psi(s)\rangle \leq \frac{\|\partial_s H(s)\|^2}{\Delta(s)^2}. \end{aligned}$$