# Fast quantum algorithms for traversing paths of eigenstates

S. Boixo,[1,*] E. Knill,[2,†] and R.D. Somma[3,‡]

[1]*California Institute of Technology, Pasadena, CA 91125, USA*

[2]*National Institute of Standards and Technology, Boulder, CO 80305, USA*

[3]*Los Alamos National Laboratory, Los Alamos, NM 87545, USA*

(Dated: May 19, 2010)



Consider a path of non-degenerate eigenstates $|\psi_s\rangle$, $0 \leq s \leq 1$, of unitary operators $U_s$ or Hamiltonians $H_s$ with minimum eigenvalue gap $\Delta$. The eigenpath traversal problem is to transform one or more copies of $|\psi_0\rangle$ into $|\psi_1\rangle$. Solutions to this problem have applications ranging from quantum physics simulation to optimization. For Hamiltonians, the conventional way of doing this is by applying the adiabatic theorem. We give "digital" methods for performing the transformation that require no assumption on path continuity or differentiability other than the absence of large jumps. Given sufficient information about eigenvalues and overlaps between states on the path, the transformation can be accomplished with complexity $\mathcal{O}((L/\Delta)\log(L/\epsilon))$, where $L$ is the angular length of the path and $\epsilon$ is a specified bound on the error of the output state. We show that the required information can be obtained in a first set of transformations, whose complexity per state transformed has an additional factor that depends logarithmically on a maximum angular velocity along the path. This velocity is averaged over constant angular distances and does not require continuity. Our methods have substantially better behavior than conventional adiabatic algorithms, with fewer conditions on the path. They also improve on the previously best digital methods and demonstrate that path length and the gap are the primary parameters that determine the complexity of state transformation along a path.

## I. INTRODUCTION

Most quantum algorithms exhibiting speedups are based on one of a small number of basic tools for quantum problem solving. These tools include phase estimation, which underlies quantum factoring, and amplitude amplification, which can be used to solve search problems. Another such tool is adiabatic state transformation (AST). Before finding applications in quantum algorithms,

*Electronic address: boixo@caltech.edu

†Electronic address: knill@boulder.nist.gov

‡Electronic address: somma@lanl.gov

AST was used in classical algorithms for quantum physics simulation [2, 12, 17]. It is now a key method for accessing low-energy states in proposed quantum simulations of physics, both in "digital" quantum algorithms based on the circuit model [5, 24] and in "analog" simulation techniques involving direct realization of Hamiltonians in systems such as optical lattices (see Ref. [19]). The potential of AST was recognized in quantum computer science when it was proposed as a powerful heuristic for solving satisfiability problems [10, 11]. This led to the idea of adiabatic quantum computing (AQC), which involves encoding the output of any quantum algorithm in an adiabatically accessible ground state [1, 11]. More recently, it has been shown that AST can be used for quantum speedups of Monte-Carlo algorithms [25, 28].

We consider AST problems that involve transforming $|\psi_0\rangle$ into $|\psi_1\rangle$, where these states are the endpoints of a path of states $|\psi_s\rangle$ for $0 \leq s \leq 1$. The $|\psi_s\rangle$ are eigenstates of operators $O_s$, which may be unitary or Hermitian. For simplicity, we focus on the unitary case $O_s = U_s$. If $O_s$ is Hermitian, we define $U_s = e^{-iO_s}$. The eigenphase of $|\psi_s\rangle$ with respect to $U_s$ is $\varphi_s$ and is assumed to be non-degenerate. The gap to the nearest other eigenphase is denoted by $\Delta_s \geq \Delta$. We make no continuity assumptions on $|\psi_s\rangle$ or $U_s$. Thus, our formulation of the AST problem can accommodate cases where the path is discrete, parameterized by integers $j = 0, \ldots, n$. It suffices to define $|\psi_s\rangle = |\psi_j\rangle$ for $j/n \leq s < (j+1)/n$.

The complexity of the AST problem depends on the available capabilities and what we know about the path. We assume that we can prepare copies of $|\psi_0\rangle$, and that we can apply quantum-controlled instances of $U_s$, both at unit cost. Additional information such as lower bounds on the gaps, eigenphase ranges and overlaps between the $|\psi_s\rangle$ may be available. An important consequence of our results is that the main parameter that determines the complexity of an AST problem is the angular length of the path, which is defined as

$$L = \sup \left\{ \sum_{j=1}^{n} \arccos(|\langle \psi_{s_j} | \psi_{s_{j-1}} \rangle|) \;\middle|\; 0 = s_0 < \ldots < s_n = 1 \right\} . \tag{1}$$

If $|\psi_s\rangle$ is differentiable in $s$, then $L = \int_0^1 \|(\mathbb{1} - |\psi_s\rangle\langle\psi_s|)|\partial_s \psi_s\rangle\| ds$. For the purpose of making complexity statements, let $\bar{L} = \max(\pi/2, L)$. Given sufficient information about the overlaps between nearby states on the path, we show that the complexity of an AST problem is $\mathcal{O}((\bar{L}/\Delta)\log(\bar{L}/\epsilon))$, where $\epsilon$ is a specified bound on the error with which $|\psi_1\rangle$ is prepared. (Our complexity statements hide constants that apply uniformly for all $L > 0$, $0 < \Delta \leq \pi$ and $0 < \epsilon < 1$.) The dependence on $L$ and $\Delta$ is within a factor of at most $\log(\bar{L})$ of optimal. That is, there are classes of AST problems for which every algorithm uses $\Omega(\bar{L}/\Delta)$ applications of the $U_s$ [8]. With less information,

we provide algorithms whose complexities have an additional factor that depends on a maximum locally averaged angular velocity

$$v_{\max} = \sup \left\{ L(s_1, s_2)/(s_1 - s_2) \;\middle|\; 0 \leq s_1 < s_2 \leq 1,\ L(s_1, s_2) \geq \theta \right\} , \tag{2}$$

where $L(s_1, s_2)$ is the angular length of the path restricted to $[s_1, s_2]$, and $\theta$ is a constant. In order for $v_{\max}$ to be finite, the path must not have jumps of angular distance $\theta$ or more. The average velocity over the whole path is $v_{\text{avg}} = L/1$. If the set in the definition of $v_{\max}$ is empty, set $v_{\max} = v_{\text{avg}}$. We find that, in general, the transformation can be accomplished with complexity $\mathcal{O}((\mathcal{L}/\Delta)\log(\mathcal{L}/\epsilon))$ per copy of $|\psi_1\rangle$ produced, where $\mathcal{L} = \bar{L}(\log(v_{\max}/v_{\text{avg}}) + 2)$. Implicit in this bound is an extra factor of $\log(\log(v_{\max}/v_{\text{avg}}) + 2)$. It is required only when it is necessary to transform multiple copies at the same time, which is the case when sufficiently small ranges for the eigenphases are not yet known. The error bound $\epsilon$ then applies to the state of all copies together, not each copy separately. All our complexities can be refined if the gap varies over the path in a known way, allowing faster traversals of parts of the path where the gap is large. We provide general tools to analyze this situation.

Many problems require the production of a large number of copies of $|\psi_1\rangle$, for example to obtain good precision on expectations of observables or values of correlation functions. In these cases, the first set of transformations can be used to get the overlap information needed to optimize future transformations. Thus, except for a first set of transformations, the complexity does not depend on $v_{\max}$.

We can compare the complexities obtained here to those of other techniques for solving AST problems. In the case where the operators $O_s$ are Hamiltonians, $O_s = H(s)$, the best known technique is based on the adiabatic approximation and involves evolving under $H(s)$, changing $s$ slowly enough to ensure the desired transformation. This technique works well in analog approaches to solving physics simulation problems. Analyses of the adiabatic approximation [1, 3, 4, 6, 9, 13, 16, 20, 23] show that the total evolution time required to obtain $|\psi_1\rangle$ with error bounded by $\epsilon$ satisfies $\tau \in \mathcal{O}\left(\sup_s \left(\|\partial_s H(s)\|^2/\Delta^3 + \|\partial_s^2 H(s)\|/\Delta^2\right)/\epsilon\right)$, where $\|\cdot\|$ is the operator norm. This assumes sufficient differentiability of $H(s)$. The dependence on $\epsilon$ is much better if $H(s)$ is highly differentiable and turned on/off slowly at the beginning and end of the path. The path length satisfies $L \leq \sup_s \|\partial_s H(s)\|/\Delta$. Examples can be constructed where $L = \Omega(\sup_s \|\partial_s H(s)\|/\Delta)$, so the bound on $\tau$ has a term that may approach $L^2/\Delta$. Besides requiring differentiability, this complexity has at least an extra factor of $L$. In the absence of differentiability, it is possible to randomize the evolution under $H(s)$ for an average complexity of $\tau = \mathcal{O}(\bar{L}^2/(\Delta\epsilon))$ [7, 25] given a

discretization of the path with sufficiently uniform angular step sizes. This technique can also be applied in an analog setting. In general, the factor of $1/\epsilon$ in the complexities can be improved to $\log(1/\epsilon)$ if we can verify $|\psi_1\rangle$ by checking its eigenvalue with respect to $H(1)$.

The randomized methods in Ref. [7] are also applicable to paths of unitaries. The paths must have a discretization with asymptotically small angular distances between adjacent states, implying a continuous underlying path. A method with better complexity is in Ref. [28]. It is based on Grover's fixed point search and if applied to our formulation of AST has a complexity of $\mathcal{O}(\bar{L}\log(\bar{L}/\epsilon)^2/\Delta)$, given an appropriate discretization of the path. This method was not analyzed for arbitrary paths, but a discretization with sufficiently large overlaps between successive states on the path works, so continuity is not required. Our work improves on this complexity, and perhaps more importantly, shows that good complexity can be obtained even when overlaps between states on the path and eigenphase ranges are unknown. It suffices to have lower bounds on the gaps, and in the least informed case, be assured that a certain eigenphase dominance condition (to be defined below) applies.

The most salient complexities are summarized in Table I. It is worth noting that many applications of AST to search and optimization problems satisfy that $L = O(1)$, in which case the complexity is dominated by $1/\Delta$ with additional logarithmic factors in $1/\epsilon$ and $v_{\max}/v_{\text{avg}}$. For example, this holds in the direct application to Grover's search problem, where $H(s)$ linearly interpolates between projectors onto the initial and final states, respectively. In this case we have sufficient knowledge of the eigenphases. If we apply our methods without using knowledge of the local behavior of the gap or the rate of change of the states, then our complexities have the expected quadratic speedup over classical search except for a logarithmic factor due to the speedup of the path near where the gap is minimal. However, in this case we are lucky: The maximum speed is attained in the exact middle of the path, which is at an angular distance of almost $\pi/4$ from either end. Because the algorithms use equal subdivision to recursively implement the state transformation (Sect. V), the transformation succeeds more quickly than expected, and the extra logarithmic factor is dropped. For applications to search and optimization there is no reason to transform more than one copy.

After we outline the conventions used in this paper, we introduce a number of basic oracles that encapsulate the elementary operations that we need to perform the state transformations. The oracles are defined to be error-free. Their actual implementations in terms of the $U_s$ are based on standard phase-estimation techniques with well understood error behavior. In Sect. IV we develop several procedures for transforming an input state $|\psi\rangle$ into $|\phi\rangle$ in one step. The procedures depend

on how much is known about the two states and their overlap. In Sect. V we compose these steps for transformations along a path. When overlaps are not known, this requires an analysis of a recursively defined tree of intervals, which we perform in sufficient detail to enable cost estimates that are sensitive to variations in the gap $\Delta_s$ along the path. Most of our algorithms are described with components whose number of steps is random and the primary complexity given is the average number of steps. To ensure that reversible versions of our algorithms can be constructed with no change in complexity, we keep track of the tail behavior of the number of steps, which always has an exponential decay. The reversible implementations have a built-in deterministic stopping criterion that is a multiple of the average. Sect. VI summarizes the complexities of our state transformation algorithms in a table and considers how our results can be generalized to the situation where the eigenphases are degenerate, and the goal is to transform a state in an eigenspace of $U_0$ into some unspecified state in a corresponding eigenspace of $U_1$.

## II. CONVENTIONS

Kets are normalized states unless explicitly stated otherwise. When writing states such as $|\varphi\rangle$ for real numbers $\varphi$, we assume that $|\varphi\rangle$ are orthonormal states for distinct $\varphi$. The numbers $\varphi$ are to be expressed in terms of labels of computational basis states for a finite system in a reasonable way. For instance, $\varphi$ could be written in binary, with the digits corresponding to basis states of qubits. The precision used should be appropriate for the context.

When writing linear expressions involving eigenphases, we always intend them to be valid modulo $2\pi$. For example, when constraining an eigenphase $\varphi$ by $\varphi \in [\varphi_0 - \delta, \varphi_0 + \delta]$, it is intended to be read with the expression "mod$(2\pi)$" appended.

For a number of parameters (such as $\Delta$, $L$, eigenphases and error bounds), we require that they have "reasonable" values. For example, eigenphase gaps should be less than $\pi$ and error bounds less than 1. We normally take such constraints for granted without specifying them explicitly. In most cases, it suffices to replace the parameter by a nearby sensible value if the parameter is out of range.

To transform states along an eigenpath of a path of unitary operators, we ultimately use controlled forms of these unitary operators and their inverses. However, we initially provide algorithms calling on idealized operators defined in terms of the unitary operators and their eigenstates of interest. We refer to these operators as oracles. When analyzing the complexity of algorithms, we do not distinguish between controlled and uncontrolled applications of unitary operators or oracles

and their inverses. In general, we do not explicitly mention the adjective "controlled" or the term "inverse", leaving them implied.

When specifying the behavior of an oracle or subroutine, we use the expression "combination of states" to refer to any superposition and/or mixture of the given states. Formally, a combination of the states $|t_i\rangle$ is a state of the form $\sum_i |t_i\rangle|e_i\rangle_E$, where $|e_i\rangle$ are unnormalized states of $E$ and $E$ is a system independent of previously introduced ones. When we say that $T$ is an operator that transforms $|\psi\rangle$ into a combination of the states $|t_{\psi,i}\rangle$, we mean that $T(|\psi\rangle) = \sum_i |t_{\psi,i}\rangle|e_{\psi,i}\rangle_E$, where $E$ is a system introduced by $T$ and $|e_{\psi,i}\rangle$ are unnormalized states of $E$. The total state associated with the combination is normalized. We always define such operators so that orthogonal states are transformed into distinguishable combinations, and require $T$ to act isometrically. The implicitly introduced system $E$ is different for each instance of $T$ in an algorithm. The particulars of the combination may also vary with instance. Thus, if an operator or oracle $T$ has been defined in terms of combinations of states for a family of input states, the symbol $T$ refers to an arbitrary operator satisfying the definition each time it is used. Formally, one can achieve this effect by transforming expressions as follows: For each occurrence of $T$ replace it with an occurrence-specific new symbol $T'$ and prefix the expression with "for some operator $T'$ satisfying the definition of $T$".

We intend $T$ to be reversible. In the absence of true decoherence processes, this can always be achieved. We may use semi-classical language to describe various computational actions such as setting a newly introduced register to a particular state, but implicitly rely on such actions having reversible forms. If the reverse of $T$ immediately follows $T$, the system $E$ is effectively eliminated by being returned to an initial state. If we performed some other action before reversing $T$, $E$ may play a decohering role. Note that since $E$ is implicit in the definition of $T$, it is not accessible to actions not involving $T$. If reversals are used, which instance of $T$ is reversed needs to be stated if it is not clear from context. When reversals are not used, the statement that $T$ maps $|\psi\rangle$ to a combination of states $|t_{\psi,i}\rangle$ is equivalent to the statement that $T$ is a quantum operation satisfying that the support of $T(|\psi\rangle\langle\psi|)$ is in the span of the $|t_{\psi,i}\rangle$. Besides allowing for reversals of instances of $T$, defining combinations in terms of implicit systems enables amplitude-based error bounds.

We may want to compare $T$ to an implementation $W$. Suppose $T$ has been defined on input states $|\psi\rangle$ as above. We say that the error amplitude of $W$ with respect to $T$ is $a$ if $W$ transforms the input states into the specified combinations up to a term of absolute amplitude at most $a$. (We omit the adjective "absolute" if it is clear from context.) That is, $W(|\psi\rangle) = \sum_i |t_{\psi,i}\rangle|f_{\psi,i}\rangle_F + |r_\psi\rangle$, where $\||r_\psi\rangle\| \leq a$. Note that the error amplitude is defined as an upper bound, not an exact error or distance. We use the fact that error amplitudes are subadditive under composition of in-principle

reversible processes. Specifically, this holds whenever the processes can be realized unitarily by addition of ancillary systems. This may require replacing explicit measurements by steps that reversibly record the measurement outcome in the ancillary systems, and performing steps that are conditioned on previous measurement outcomes by the appropriate quantum-controlled steps. After this change, subadditivity of the error amplitudes follows by writing the final state as a combination of the error free state and the errors introduced by each step propagated through the subsequent ones. Observe that subsequent steps do not change the amplitude of the propagated error.

Most of the procedures that we analyze are not explicitly formulated in reversible form, and the primary complexity measure is an average cost. Because state transformation procedures have applications as subroutines in larger quantum algorithms, it is desirable to have reversible versions that can exploit quantum parallelism without introducing unwanted decoherence. However, the average cost is determined with respect to stopping criteria associated with measurements. When the procedure is reversified, one cannot have a stopping criterion that is input dependent. In principle, this may require running the procedure much longer than suggested by the average cost to ensure that all possible computation paths terminate. Suppose the average cost is $\bar{C}$. If we allow for some error amplitude, we can set an absolute termination criterion by stopping when the total cost has exceeded $C_{\max}$. This can be done in a reversible way and introduces an error amplitude bounded by $\sqrt{\bar{C}/C_{\max}}$ (Markov's inequality for non-negative random variables converted to amplitude). This bound is undesirably large and, without additional assumptions, can be approached. So we seek procedures where the probability distribution of $C$ decays exponentially after some multiple of $\bar{C}$. If $\mathrm{Prob}(C \geq c) \leq x^{c-\lambda\bar{C}}$ for some $0 < x < 1$ and $\lambda \geq 1$, then the error amplitude for $C_{\max} > \lambda\bar{C}$ is bounded by $x^{(C_{\max}-\lambda\bar{C})/2}$. This implies that for an error amplitude of $\epsilon$, we can set $C_{\max} = \lambda\bar{C} + 2\ln(1/\epsilon)/\ln(1/x)$, so that the error dependence of the cost has an additive term that is only logarithmic in $\epsilon$.

In order to keep track of the exponential decay of costs, we use the large-deviation technique of bounding the expectation $\langle\Gamma^C\rangle$ of $\Gamma^C$. Thus, whenever it matters, we specify the tail behavior of the probability distribution of $C$ by an inequality of the form $\langle\Gamma^C\rangle \leq \Gamma^{\tilde{C}}$ for $1 \leq \Gamma < \Gamma_{\max}$. Rather than trying to optimize the inequality, we generally choose convenient, simple expressions for $\Gamma_{\max}$ and $\tilde{C}$, ensuring that $\tilde{C}$ is bounded by a constant multiple of the average cost. This suffices for stating bounds on complexities while having reasonable estimates for the hidden constants. Given such a bound, we can use Markov's inequality to show that $\mathrm{Prob}(C \geq c) = \mathrm{Prob}(\Gamma^C \geq \Gamma^c) \leq \Gamma^{\tilde{C}-c}$. This is of the desired form, with $x = 1/\Gamma$. When proving tail bounds, we liberally use the fact that

$F(\lambda) = \langle e^{\lambda C} \rangle$ is log-convex in $\lambda$. In particular, if $\langle \Gamma_1^C \rangle \le \Gamma_1^{\tilde{C}}$ for some $\Gamma_1 \ge 1$, then this inequality automatically holds for all $\Gamma$ between 1 and $\Gamma_1$.

The main reason to use the large-deviation technique of the previous paragraph is to simplify the estimation of tail bounds for total costs of compositions of procedures. For this purpose we have the following lemmas:

**Lemma II.1.** *Consider a sequence of procedures $S_j$ with costs $C_j$ satisfying $\langle \Gamma^{C_j} | C_1, \ldots, C_{j-1} \rangle \le \Gamma^{\tilde{C}_j}$ for $1 \le \Gamma \le \Gamma_{\max}$. Define $C_{\mathrm{tot},l} = \sum_{i=1}^{l} C_i$ and $\tilde{C}_{\mathrm{tot},l} = \sum_{i=1}^{l} \tilde{C}_i$. Then $\langle \Gamma^{C_{\mathrm{tot},l}} \rangle \le \Gamma^{\tilde{C}_{\mathrm{tot},l}}$ for $1 \le \Gamma \le \Gamma_{\max}$.*

For random variables $A$ and $B$, the expression $\langle A|B \rangle$ used in the lemma denotes the conditional expectation of $A$ given $B$.

*Proof.* The proof is by induction on $l$ using a standard large-deviations approach. Let $\mu$ denote the measure for the probability distribution of its arguments.

$$\begin{aligned}
\langle \Gamma^{C_{\mathrm{tot},l+1}} \rangle &= \int \langle \Gamma^{C_{l+1}} | C_1, C_2, \ldots C_l \rangle \Gamma^{C_{\mathrm{tot},l}} d\mu(C_1, \ldots, C_l) \\
&\le \Gamma^{\tilde{C}_{l+1}} \int \Gamma^{C_{\mathrm{tot},l}} d\mu(C_1, \ldots, C_l) \\
&\le \Gamma^{\tilde{C}_{l+1}} \Gamma^{\tilde{C}_{\mathrm{tot},l}} \\
&= \Gamma^{\tilde{C}_{\mathrm{tot},l+1}} \ .
\end{aligned}$$

□

Lemma II.1 implies that if each component procedure has exponentially decaying cost above a multiple of the average, so does the composition. The next lemma generalizes this result to the case where the number of $S_j$ invoked is not deterministic. In the lemma's statement, the binary random variable $W_j$ can be thought of as "$S_j$ was successful". The lemma is intended to be applied when $S_k$ depends only on which of the $S_j$ with $j < k$ where successful. For a sequence of random variables $X_i$, we write $\mathbf{X}_i = (X_1, \ldots, X_i)$.

**Lemma II.2.** *Consider a sequence of procedures $S_j$ with costs $C_j$. Let $W_j$ be a binary random variable such that $C_j$ and $W_j$ are conditionally independent of $\mathbf{C}_{j-1}$ given $\mathbf{W}_{j-1}$. Let $V_j = 1$ if $C_j > 0$ and $V_j = 0$ otherwise. Define $m = \sum_j V_j$ and suppose that $\langle \Lambda^m \rangle \le \Lambda^{\tilde{m}}$ for $1 \le \Lambda \le \Lambda_{\max}$ and $\langle \Gamma^{C_j} | \mathbf{W}_j, V_j \rangle \le \Gamma^{\tilde{C}}$ for $1 \le \Gamma \le \Gamma_{\max}$. Then $\langle \Gamma^{C_{\mathrm{tot},\infty}} \rangle \le \Gamma^{\tilde{m}\tilde{C}}$ for $1 \le \Gamma \le \min\left(\Lambda_{\max}^{1/\tilde{C}}, \Gamma_{\max}\right)$.*

The proof is given in Appendix A.

## III. ORACLES

We introduce four oracles implementing idealized quantum operations. We can implement versions of these oracles with low error amplitude in terms of subroutines with full access to the unitary operators defining the eigenpath.

We say that an oracle is $U$-controlled if it commutes with any $X$ that commutes with $U$ and acts on the same system as $U$. Such an oracle necessarily preserves eigenstates $|\psi\rangle$ of $U$ in the sense that it transforms $|\psi\rangle|a\rangle$ to a state of the form $|\psi\rangle|f(a,\psi)\rangle$ for any state $|a\rangle$ of other systems. In particular, if an instance of the oracle is reversed on a state of the form $|\psi\rangle|b\rangle$ where $|\psi\rangle$ is an eigenstate of $U$ and $|b\rangle$ is arbitrary, the result is of the form $|\psi\rangle|a'\rangle$. A strictly $U$-controlled oracle is obtained if it is implemented solely in terms of controlled-$U$ operations. The implementations described for oracles below that are required to be $U$-controlled satisfy this property.

Our oracle implementations have a specified error amplitude but are still strictly $U$-controlled when this is required. It is worth noting that for strictly $U$-controlled implementations, if a bound on the error amplitude is specified only for eigenstates of $U$, then it holds in general. To see that it holds for arbitrary superpositions of eigenstates, it is necessary to use the fact that the error amplitudes for orthogonal eigenstates are orthogonal. This holds because of being $U$-controlled: For eigenstate inputs, the decomposition of the output state into an error-free part and the error amplitude results in both being a product of the eigenstate with states of other systems. We note that in general, given error amplitudes only for a basis of the state space, errors in a superposition can add as amplitudes rather than probabilities. This may result in a $\sqrt{d}$ error enhancement, where $d$ is the dimension.

**Definition III.1.** *Given a unitary operator $U$ and resolution $\delta$, a phase estimation oracle* $\mathrm{PE}(U,\delta)$ *is an isometry that transforms eigenstates $|\psi\rangle$ of $U$ with eigenphase $\varphi$ into a combination of states of the form*

$$|\psi\rangle|\varphi_x\rangle_A \ , \tag{3}$$

*where $\varphi_x - \varphi \in [-\delta,\delta]$. We require that* PE *is $U$-controlled. We can implement* PE *with error amplitude $\epsilon$ using $\mathcal{O}(\log(1/\epsilon)/\delta)$ applications of $U$.*

To implement PE it suffices to use a high-confidence version of phase estimation. Such a phase estimation technique and its analysis are in Ref. [18]. PE is $U$-controlled because all actions involving $U$'s system are ancilla-controlled $U$ operations.

**Definition III.2.** *Given a unitary operator $U$, a phase $\varphi_0$ and a resolution $\delta$, a phase detection oracle* $\mathrm{PD}(U, \varphi_0, \delta)$ *is an isometry that acts on eigenstates $|\psi\rangle$ of $U$ with eigenphase $\varphi \in [\varphi_0 - \delta/4, \varphi_0 + \delta/4]$ or $\varphi \notin [\varphi_0 - 3\delta/4, \varphi_0 + 3\delta/4]$ as*

$$\mathrm{PD}(U, \varphi_0, \delta)|\psi\rangle = |\psi\rangle|b\rangle_A \ , \tag{4}$$

*where $b = 1$ if $\varphi \in [\varphi_0 - \delta/4, \varphi_0 + \delta/4]$ and $b = 0$ otherwise. Eigenstates $|\psi\rangle$ not satisfying the eigenphase constraint are mapped to arbitrary combinations of states of the form $|\psi\rangle|b\rangle_A$. We require that* PD *is $U$-controlled. We can implement* PD *with error amplitude $\epsilon$ using $\mathcal{O}(\log(1/\epsilon)/\delta)$ applications of $U$.*

To implement the phase detection oracle with the stated complexity, apply $\mathrm{PE}(U, \delta/4)$, labeling its output register $A'$. Reversibly set register $A$ to $|1\rangle$ if $\varphi_x \in [\varphi_0 - \delta/2, \varphi_0 + \delta/2]$, otherwise set it to $|0\rangle$. Then reverse the instance of PE used.

**Definition III.3.** *Given a unitary operator $U$, a phase $\varphi_0$ and a resolution $\delta$, a reflection oracle* $\mathrm{R}(U, \varphi_0, \delta)$ *is an isometry that acts on $|\psi\rangle$ in the subspace spanned by eigenstates of $U$ with eigenphase $\varphi \in [\varphi_0 - \delta/4, \varphi_0 + \delta/4]$ or $\varphi \notin [\varphi_0 - 3\delta/4, \varphi_0 + 3\delta/4]$ as*

$$\mathrm{R}(U, \varphi_0, \delta)|\psi\rangle = (-1)^b|\psi\rangle \ , \tag{5}$$

*where $b = 1$ if $\varphi \in [\varphi_0 - \delta/4, \varphi_0 + \delta/4]$ and $b = 0$ otherwise. Eigenstates $|\psi\rangle$ not satisfying the eigenphase constraint are mapped to a combination of $|\psi\rangle$. We require that* R *is $U$-controlled. We can implement* R *with error amplitude $\epsilon$ using $\mathcal{O}(\log(1/\epsilon)/\delta)$ applications of $U$.*

To implement the reflection oracle with the stated complexity, apply $\mathrm{PD}(U, \varphi_0, \delta)$. Conditionally on the bit in register $A$ change the phase of the input state. Then reverse the instance of PD used.

A reflection oracle preserves eigenstates but can decohere the phases for eigenstates not satisfying the constraint by correlating them with the instance-dependent system implied by our convention for combinations.

We note that the implementation of the reflection oracle is a special case of the functional calculus of $U$ implemented via phase estimation. For functions $f$ preserving the unit circle in the complex plane and slowly varying except in known eigenphase gaps, one can implement $f(U)$ by using phase estimation with sufficient resolution, changing the phase according to $f$ and reversing the phase estimation. The error can be bounded in terms of the resolution used. A version of this observation for more general $f$ (which requires postselection) but not using high-confidence phase estimation can be found in Ref. [14].

**Definition III.4.** *Given states $|\psi\rangle$, $|\phi\rangle$, an overlap threshold $\alpha$ and a resolution $\delta$, an overlap detection oracle* $\mathrm{OV}(\psi, \phi, \alpha, \delta)$ *is an isometry that transforms $|\psi\rangle$ into a combination of states of the form*

$$|\psi'\rangle|b\rangle_A \ , \tag{6}$$

*where the following holds: If* $\arccos(|\langle\psi|\phi\rangle|) < \alpha - \delta$*, then $b = 1$ and $|\psi'\rangle = |\psi\rangle$. If* $\arccos(|\langle\psi|\phi\rangle|) > \alpha + \delta$*, then $b = 0$ and $|\psi'\rangle = |\psi\rangle$. Otherwise, $|\psi'\rangle$ is in the subspace spanned by $|\psi\rangle$ and $|\phi\rangle$. We can implement* $\mathrm{OV}(\psi, \phi, \alpha, \delta)$ *with error amplitude $\epsilon$ using $\mathcal{O}(\log(1/\epsilon)/\delta)$ applications of reflections around $|\psi\rangle$ and $|\phi\rangle$.*

Let $R_\psi$ and $R_\phi$ be reflections around $|\psi\rangle$ and $|\phi\rangle$, respectively. The overlap oracle is implemented by applying the phase estimation oracle $\mathrm{PE}(U = R_\psi R_\phi, 2\delta)$. The operator $U$ preserves the subspace spanned by $|\psi\rangle$ and $|\phi\rangle$, and its two eigenvalues on this subspace are $e^{\pm i 2 \arccos(|\langle\psi|\phi\rangle|)}$. We set register $A$ to $|1\rangle$ if the phase $\varphi_x$ returned by the phase estimation oracle satisfies $\varphi_x \in [-2\alpha, 2\alpha]$, and to $|0\rangle$ otherwise. We then reverse the instance of PE used. See Ref. [18] for more details.

If the reflections required for the overlap oracle are implemented in terms of reflection oracles with resolution $\delta'$, the overall complexity for an error amplitude of $\epsilon$ has a factor of $\log(1/\epsilon)^2$. We generally aim for a dependence on $\epsilon$ of $\log(1/\epsilon)$, which requires bypassing direct uses of overlap oracles.

## IV. ONE-STEP STATE TRANSFORMATIONS

Suppose we are given a system in the eigenstate $|\psi\rangle$ of $U$ and we wish to transform $|\psi\rangle$ into the eigenstate $|\phi\rangle$ of $V$. We denote the eigenphases by $\varphi_U$ and $\varphi_V$, respectively. We assume that the eigenstates are unique for the eigenphases and that the gaps to the nearest other eigenphases are bounded below by $\Delta$. The overlap probability is denoted by $p = |\langle\psi|\phi\rangle|^2$. Define $q = (1 - p)$. The methods for transforming the states depend on what is known about the eigenphases, gaps and overlaps. Given reflection oracles and $p$ bounded away from 0 and 1, we can use ideas developed for fixed point quantum search [27] and QMA amplification [21, 22] for a transformation using a constant number of reflections on average. Define reflection operators by $R_\psi = \mathbb{1} - 2|\psi\rangle\langle\psi|$ and $R_\phi = \mathbb{1} - 2|\phi\rangle\langle\phi|$. Each reflection operator can be implemented with one call to the appropriate reflection oracle. The transformation from $|\psi\rangle$ into $|\phi\rangle$ is accomplished by repeatedly applying the circuit $\mathrm{RT}(\psi, \phi)$ of Fig. 1 until the measurement outcome is 1, indicating that $|\phi\rangle$ has been prepared.

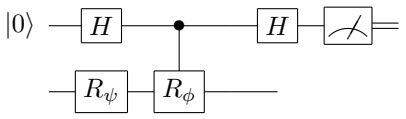


FIG. 1: Quantum circuit $\mathrm{RT}(\psi, \phi)$ for a state transformation attempt. $H$ denotes the Hadamard gate. The filled circle denotes control on the state $|1\rangle$ of the corresponding ancilla qubit.

The effect of $\mathrm{RT}(\psi, \phi)$ is to apply a reflection around $|\psi\rangle$ followed by a projection onto $|\phi\rangle$ if the measurement outcome is 1, or a projection onto the orthogonal complement otherwise. The subspace spanned by $|\psi\rangle$ and $|\phi\rangle$ is preserved by the process. Let $|\phi^\perp\rangle$ be the state orthogonal to $|\phi\rangle$ in this subspace. If the subspace is one-dimensional, choose any orthogonal state. The procedure for transforming the states can be analyzed as a Markov chain on the states $|\psi\rangle$, $|\phi\rangle$ and $|\phi^\perp\rangle$. We consider a slightly more general procedure $T(\psi, \phi)$ that can be applied to any initial state $|\psi'\rangle$ in the subspace spanned by $|\psi\rangle$ and $|\phi\rangle$. Define $p_0 = |\langle\psi'|\phi\rangle|^2$. The first step of the procedure consists of the circuit for $\mathrm{RT}(\psi, \phi)$ with the reflection around $|\psi\rangle$ omitted. Next, $\mathrm{RT}(\psi, \phi)$ is applied until the measurement outcome on the ancilla qubit is 1, indicating that $|\phi\rangle$ has been prepared. The transition probabilities for this procedure are shown in Fig. 2.

In general, when we describe a step of a procedure as a measurement of a state $|\phi'\rangle$, this is intended to be implemented by means of a controlled reflection around $|\phi'\rangle$ as in the second part of RT. The effect is a projection onto $|\phi'\rangle$ or the orthogonal complement, depending on the measurement outcome.

To simplify the notation, we omit arguments of procedures such as $T$ and RT when they are sufficiently clear from context. The arguments are typically passed on to appropriate oracle calls. We may therefore use any set of alternative arguments that are sufficient for specifying these oracles.

**Lemma IV.1.** *We can transform any state $|\psi'\rangle$ in the subspace spanned by $|\psi\rangle$ and $|\phi\rangle$ into $|\phi\rangle$ with a procedure $T$ using $\langle n\rangle = p_0 + (1 - p_0)(1 + 1/(2pq)) \leq 1 + 1/(2pq)$ reflections around the states on average. For $1 \leq \Gamma < 1/|p - q|$, define $c$ by the equation $\Gamma^2 = (1 - cpq)/(p - q)^2$. We then have $\langle \Gamma^n\rangle = \Gamma(p_0 + (1 - p_0)4\Gamma^2/c) \leq 4\Gamma^3/c$.*

*Proof.* The procedure is as described above. Consider the process of Fig. 2. Define $n'$ to be the number of reflections used after the first step, if the first step resulted in state $|\phi^\perp\rangle$. We have $\langle n\rangle = p_0 + (1 - p_0)(1 + \langle n'\rangle)$. If the first step failed, the second step can either succeed or return to $|\phi^\perp\rangle$. In the first case, we used $n' = 2$ reflections. In the second case, the expected number of

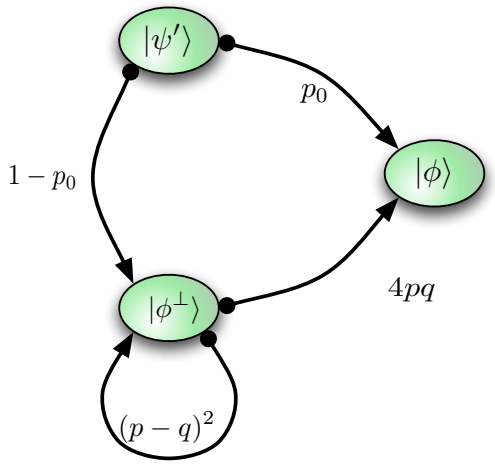


FIG. 2: State diagram for $T(\psi, \phi)$. The state $|\psi'\rangle$ is the initial state and $p_0$ is the overlap probability of $|\psi'\rangle$ with $|\phi\rangle$. For the non-trivial case of $p < 1$, the transition probabilities from $|\phi^\perp\rangle$ are obtained by explicit computation of the reflections in the two dimensional subspace of the states. For this purpose, one can write $|\psi\rangle = \sqrt{p}|\phi\rangle + \sqrt{q}|\phi^\perp\rangle$, $|\psi^\perp\rangle = \sqrt{q}|\phi\rangle - \sqrt{p}|\phi^\perp\rangle$ and $|\phi^\perp\rangle = \sqrt{q}|\psi\rangle - \sqrt{p}|\psi^\perp\rangle$.

reflections yet to be used is again $\langle n'\rangle$. This implies the equation $\langle n'\rangle = 4pq \cdot 2 + (p-q)^2(2 + \langle n'\rangle)$. Thus $\langle n'\rangle = 2/(4pq)$ and $\langle n\rangle = p_0 + (1-p_0)(1 + 1/(2pq))$.

Similarly, we can calculate $\langle \Gamma^n\rangle$ by solving the equations $\langle \Gamma^n\rangle = p_0\Gamma + (1-p_0)\Gamma\langle \Gamma^{n'}\rangle$ and $\langle \Gamma^{n'}\rangle = 4pq\Gamma^2 + (p-q)^2(\Gamma^2\langle \Gamma^{n'}\rangle)$. This gives $\langle \Gamma^{n'}\rangle = 4pq\Gamma^2/(1-(p-q)^2\Gamma^2) = 4\Gamma^2/c$. The last inequality follows because $c \leq 4$. □

According to Lemma IV.1, the number of reflections used satisfies an exponential decay with a fixed base less than 1 provided that $p$ is bounded away from both 0 and 1. In order to obtain a better behaved transformation that only requires $p$ to be bounded away from 0, we use an overlap-suppression trick to modify $T$. We first add an ancilla $A$ in the state $\sqrt{3/4}|0\rangle_A + \sqrt{1/4}|1\rangle_A$. Write $|\tilde{\psi}\rangle = |\psi\rangle(\sqrt{3/4}|0\rangle_A + \sqrt{1/4}|1\rangle_A)$ and $|\tilde{\phi}\rangle = |\phi\rangle|0\rangle_A$. Reflections around $|\tilde{\psi}\rangle$ and $|\tilde{\phi}\rangle$ in the extended statespace can be implemented with properly controlled reflections around $|\psi\rangle$ and $|\phi\rangle$. We can therefore transform $|\tilde{\psi}\rangle$ into $|\tilde{\phi}\rangle$ using $T(\tilde{\psi}, \tilde{\phi})$ After we are done, we discard the ancilla to extract the desired state $|\phi\rangle$. The overlap-suppression trick leads to the following lemma:

**Lemma IV.2.** *If $p \geq 1/3$, we can transform $|\psi\rangle$ into $|\phi\rangle$ with a procedure $T_m$ whose average number of reflections satisfies $\langle n\rangle < 4$. For $1 \leq \Gamma \leq 7/4$, we have $\langle \Gamma^n\rangle \leq \Gamma^6$.*

*Proof.* With the technique just described, the overlap probability is changed from $p$ to $3p/4$. Given the assumed lower bound, this ranges from $1/4$ to $3/4$. Here, $p_0 = 3p/4$. From Lemma IV.1, the

average number of reflections used is at most $1 + 1/(2(1/4)(3/4)) < 4$.

To prove the bound on $\langle \Gamma^n \rangle$, it suffices to show that for $\Gamma \leq 7/4$, $\Gamma^3 \geq 4/c$ and apply the last inequality of Lemma IV.1. The function $(p-q)^2/(pq)$ is symmetric in $p$ and $q$ and achieves its maximum of $1/2$ for the allowed values of $p$ at $p = 1/3$. Using the inequality $(1+x)^y \geq 1 + yx$ for $y \geq 0$ and $x > -1$, we get

$$\Gamma^3 \geq \Gamma^{\frac{2(p-q)^2}{pq}} = (1 + (\Gamma^2 - 1))^{\frac{(p-q)^2}{pq}} \geq 1 + \frac{(p-q)^2}{pq}(\Gamma^2 - 1).$$

Since $\Gamma^2 = (1 - cpq)/(p-q)^2 = 1 + (4-c)pq/(p-q)^2$, we conclude that $\Gamma^3 \geq 5 - c$. The constraints on $\Gamma$ and $p$ imply that $c \geq 1$, so $\Gamma^3 \geq 4 \geq 4/c$. □

The constants used for Lemma IV.2 have been chosen for convenience and have not been optimized. Changing the lower bound on $p$ or the parameters of the ancilla state when redefining $|\psi\rangle$ changes the bounds in the lemma but does not affect the complexities to be derived later.

To deal with the problem of low overlap probability $p$, we extend $T_m$ to a procedure $T_x$ with the property that the initial state $|\psi\rangle$ is transformed if $p$ is large enough and unchanged for $p$ too small.

**Lemma IV.3.** *We can implement a procedure $T_x$ that transforms $|\psi\rangle$ into a combination of states of the form $|\psi\rangle|0\rangle_A$ and $|\phi\rangle|1\rangle_A$, where register $A$ is $|1\rangle$ if $p > 1/2$ and $|0\rangle$ if $p < 1/3$. $T_x$ requires one overlap oracle call with a resolution of $(\arccos(\sqrt{1/3}) - \arccos(\sqrt{1/2}))/2$, a reflection around $|\psi\rangle$, and, in the case where the state is transformed into $|\phi\rangle$, one instance of $T_m$ with $p$ guaranteed to be at least $1/3$.*

*Proof.* To implement $T_x$, we use the overlap oracle $\mathrm{OV}(\psi, \phi, (\arccos(\sqrt{1/3}) + \arccos(\sqrt{1/2}))/2, (\arccos(\sqrt{1/3}) - \arccos(\sqrt{1/2}))/2)$ once to obtain a combination of $|\psi'\rangle|b\rangle_A$ with $|\psi'\rangle$ in the span of $|\psi\rangle$ and $|\phi\rangle$. We then use a reflection around $|\psi\rangle$ to measure $|\psi\rangle$. If the state is determined to be $|\psi^\perp\rangle$ or if $b = 1$, we apply $T_m$. The properties of the overlap oracle and the choice of parameters ensure that this happens only if $p \geq 1/3$. It always happens if $p > 1/2$. □

The procedure $T_x$ provides bounds on the overlap at the cost of calling an overlap oracle. As noted in Sect. III, this results in unwanted overhead when accounting for error amplitudes. This can be avoided if one does not seek guaranteed overlap bounds and accepts the possibility that even at high overlap, the transformation may not succeed.

**Lemma IV.4.** *We can implement a procedure $T'_x$ that transforms $|\psi\rangle$ into a combination of states of the form $|\psi\rangle|0\rangle_A$ and $|\phi\rangle|1\rangle_A$, where register $A$ is $|1\rangle$ with probability $1 - \frac{1-p}{1+p}(q-p)^4$, and the*

*number of reflections around* $|\psi\rangle$ *and* $|\phi\rangle$ *is bounded by* $n \leq 5 + n'$*, where* $\langle n' \rangle \leq 1/(1-p)$ *and* $\langle \Gamma^{n'} \rangle \leq (1-p)\Gamma/(1-p\Gamma)$ *for* $1 \leq \Gamma < 1/p$*.*

Register $A$'s contents indicate whether the transformation succeeded.

*Proof.* To implement $T'_x$, we apply a $|\phi\rangle$-measurement followed by at most two applications of RT, stopping if $|\phi\rangle$ is detected in the measurement. If after this, the state is $|\phi^\perp\rangle$ (as indicated by the last measurement), we alternately make $|\psi\rangle$- and $|\phi\rangle$-measurements until either $|\psi\rangle$ or $|\phi\rangle$ is detected. Register $A$ is set to $|1\rangle$ if $|\phi\rangle$ was detected in the last measurement made. The probability that register $A$ is $|1\rangle$ is at least the probability that the first three steps of the process of Fig. 2 terminate at $|\phi\rangle$, which is $1-(1-p)(q-p)^4$ (see the proof of Lemma IV.1). We can improve this by noting that conditional on the failure of these steps, the probability of success in the sequence of alternating measurements is $(1-p)p\sum_{k=}^{\infty} p^{2k} = p/(1+p)$. Multiplying by $(1-p)(q-p)^4$ and adding to $1-(1-p)(q-p)^4$ we get the overall probability of successful preparation of $|\phi\rangle$.

The first part of the procedure uses at most five reflections. The probability of failure to produce an acceptable outcome in a given measurement in the second part is $p$. Thus, if the first part fails, the expected number of reflections used in the second part is $1/(1-p)$. Setting $n'$ to the number of reflections used if the first part fails (event $F$), we have $\langle \Gamma^{n'}|F\rangle = (1-p)\Gamma + p\Gamma\langle\Gamma^{n'}|F\rangle$. To obtain the last statement of the lemma, it suffices to solve this equation. □

For later use, we note the following refinement of the lemma:

**Corollary IV.5.** *When* $T'_x$ *is invoked on state* $|\psi\rangle$*, the distribution of the number of reflections used by* $T'_x$ *and whether* $T'_x$ *succeeds is independent of any previous events. We also have* $\langle \Gamma^{n'}|A\rangle \leq (1-p^2)\Gamma^2/(1-\Gamma^2p^2)$ *for* $1 \leq \Gamma < 1/p$*, where* $A$ *indicates success or failure of* $T'_x$*.*

*Proof.* Let $E_\phi$ and $E_\psi$ be the events that the first part of $T'_x$ fails and $|\phi\rangle$ or $|\psi\rangle$ are eventually obtained, respectively. The probability of $E_\phi$ is $p/(1+p)$. Conditional on $E_\phi$, $n'$ is even, $n' \geq 2$, and we get

$$\langle \Gamma^{n'}|E_\phi\rangle = \frac{1+p}{p} \sum_{n' \in \{2,4,\ldots\}} (1-p)p^{n'-1}\Gamma^{n'} = \frac{\Gamma^2(1-p^2)}{1-\Gamma^2p^2}$$

Similarly, conditional on $E_\psi$, $n'$ is odd, and $\langle \Gamma^{n'}|E_\psi\rangle = \Gamma(1-p^2)/(1-\Gamma^2p^2)$. The claim follows because $\Gamma \geq 1$. □

For concreteness, we give the following corollary:

**Corollary IV.6.** *For $p \geq 1/4$, $T'_x$ transforms into $|\phi\rangle|1\rangle_A$ with probability at least 19/20. For $p \leq 3/4$, the average number of reflections satisfies $\langle n \rangle \leq 9$, and $\langle \Gamma^n | A \rangle \leq \Gamma^{11}$ for $1 \leq \Gamma \leq 8/7$, where $A$ indicates success or failure of $T'_x$.*

*Proof.* We use the lower bound $1 - (1-p)(q-p)^4$ for the probability of success. The minimum of $1 - (1-p)(q-p)^4$ for $p \in [1/4, 1]$ is at $p = 9/10$, and one can check that the value is above 19/20. Since $\Gamma \geq 1$, $f(p) = (1-p^2)/(1-p^2\Gamma^2)$ achieves its maximum on $p \in [0, 3/4]$ at $p = 3/4$. We show the last bound for $\Gamma = 8/7$, which is sufficient by log-convexity. In this case $\Gamma^4 \geq f(3/4)$ and, since $n \leq 5 + n'$, Cor. IV.5 gives the bound. □

We let $T'_{mx}$ be the procedure obtained by combining the overlap-suppression trick with $T'_x$. The procedures $T$, $T_m$, $T_x$, $T'_x$ and $T'_{mx}$ can be implemented in terms of the reflection oracles R when the states $|\psi\rangle$ and $|\phi\rangle$ are specified to be eigenstates of $U$ and $V$ with isolated eigenphases known to satisfy $\varphi_U \in [\tilde{\varphi}_U - \Delta/4, \tilde{\varphi}_U + \Delta/4]$ and $\varphi_V \in [\tilde{\varphi}_V - \Delta/4, \tilde{\varphi}_V + \Delta/4]$, and with gaps lower bounded by $\Delta$. Here $\tilde{\varphi}_U$ and $\tilde{\varphi}_V$ are assumed to be known, but not $\varphi_U$ and $\varphi_V$. In this situation, we say that we know $\Delta/2$-ranges for the eigenvalues. The instances of the reflection oracles used by the procedures are given by $\mathrm{R}(U, \tilde{\varphi}_U, \Delta)$ and $\mathrm{R}(V, \tilde{\varphi}_V, \Delta)$. We choose the third parameter to be as large as possible, because this decreases the complexity of implementing the reflection oracles in terms of the unitary operators. For the rest of this paper, whenever we use one of $T$, $T_m$ and $T_x$, we assume that the reflections used are based on reflection oracles R with the third argument given by a known lower bound on the minimum gap, by default $\Delta$.

We are interested in the situation where the overlap probability $p$ is bounded away from 0, but we do not know a sufficiently small range for the eigenphase of $|\phi\rangle$ with respect to its defining operator $V$ to use the requisite reflection oracle. We show that a sufficiently small eigenphase range can be obtained with low error if we transform many copies of $|\psi\rangle$ in parallel into $|\phi\rangle$, provided $p$ is sufficiently large. The statement of the lemma includes an optional projection $\Pi_0$ such that $\Pi_0|\phi\rangle = 0$. This is later used in the context of the overlap-suppression trick so the 0 eigenvalue of the modified unitary does not get confused with the original phases (for instance, see Lemma IV.8).

**Lemma IV.7.** *Suppose that $p > 1/2 + \gamma$ with $\gamma > 0$, and $V = V_0 \oplus \Pi_0$, where $\Pi_0$ is a projector that we can use to control other operations and $\Pi_0|\phi\rangle = 0$. Then, using $2r$ instances of phase estimation oracles $\mathrm{PE}(V, \Delta/5)$, we can implement an isometry $\mathrm{ER}(V, \Pi_0, \Delta, \gamma)$ with the following property up to an error amplitude of $e^{-r\gamma^2}$: ER transforms $|\psi\rangle^{\otimes r}$ into a combination of states of the form*

$$|\phi\rangle^{\otimes j} |\phi^{\perp}\rangle^{\otimes (r-j)} |\varphi\rangle_A |j\rangle_B \ , \tag{7}$$

*where $j > r/2$ and $\varphi - \varphi_V \in [-\Delta/5, \Delta/5]$.*

*Proof.* Use system label $i$ for the $i$'th copy of $|\psi\rangle$. We can modify the phase estimation oracles $\mathrm{PE}(V, \Delta/5)$ so that the eigenphase register contains a non-eigenphase # if the input state is in the support of $\Pi_0$. We implement ER by first using these modified phase estimation oracles $\mathrm{PE}'$ independently on each $|\psi\rangle_i$, placing the eigenphase in register $A_i$. We can express $|\psi\rangle$ as $|\psi\rangle = \sqrt{p}|\phi\rangle + \sqrt{1-p}|\phi^\perp\rangle$. By definition, the $i$'th instance of $\mathrm{PE}'$ acts as $|\phi\rangle_i \mapsto |\phi\rangle_i |w_i\rangle_{A_i E_i}$ for some $|w_i\rangle_{A_i E_i}$ and some instance-specific additional system $E_i$.

If we make conceptual $|\phi\rangle_i$-measurements, we detect $|\phi\rangle_i$ with probability $p > 1/2 + \gamma$. Thus, with probability at least $1 - e^{-2r\gamma^2}$ (Hoeffding's inequality [15]), $j > r/2$ of the values $\varphi_i$ in registers $A_i$ are in $[\varphi_V - \Delta/5, \varphi_V + \Delta/5]$. Consider this case. Because of the gap condition, other eigenphases are outside $(\varphi_V - 4\Delta/5, \varphi_V + 4\Delta/5)$. Thus, there is a $\varphi'$ with the property that more than half of the $\varphi_i$ are in $[\varphi' - \Delta/5, \varphi' + \Delta/5]$, and we are guaranteed that $\varphi_V$ is within $2\Delta/5$ of $\varphi'$. Furthermore, any $\varphi_i$ within $\Delta/5$ of $\varphi'$ is associated with $|\phi_i\rangle$ and therefore within $\Delta/5$ of $\varphi_V$.

We do not make the measurements of the previous paragraph. Instead we reversibly (and unitarily) determine whether an interval $[\varphi' - \Delta/5, \varphi' + \Delta/5]$ containing more than half of the $\varphi_i$ exists. Except for an error amplitude of $e^{-r\gamma^2}$, the state satisfies the condition, which we now assume. We reversibly compute the number $j$ and the median $\varphi$ of the $\varphi_i$ in the interval found. As discussed in the previous paragraph, $j > r/2$ and $\varphi$ is within $\Delta/5$ of $\varphi_V$. We place $j$ in register $B$ and $\varphi$ in register $A$ and reverse all classical reversible computations that were required since invoking the phase estimation oracles. Next we reorder systems so that for $i \leq j$, the $i$'th group is in state $|\phi\rangle_i |w_i'\rangle_{A_i E_i}$ and for $i > j$, it is in the state $|\eta\rangle_{A_i E_i} = \mathrm{PE}'|\phi^\perp\rangle_i$. Here $|w_i'\rangle_{A_i E_i}$ may be different from $|w_i\rangle_{A_i E_i}$ because of possible correlations with $\varphi$. The states $|\eta\rangle_{A_i E_i}$ have not changed because the eigenphases encoded in these states have no correlation with $\varphi$.

For the last step of ER, we reverse the appropriate instance of $\mathrm{PE}'$ on each register $|\eta\rangle_{A_i E_i}$. In order to do this we need to have computed the reordering permutation into an additional register, which we retain only if we need to reverse this instance of ER. The desired combination of states has now been obtained. □

From the proof of Lemma IV.7, it can be seen that the error is such that the $r$ input systems' state remains in the tensor product of the span of $|\psi\rangle$ and $|\phi\rangle$.

To complete the transformation from $|\psi\rangle^{\otimes r}$ to $|\phi\rangle^{\otimes r}$ after applying ER without using an excessive number of reflections when $p$ is close to 1, we can use the overlap-suppression trick.

**Lemma IV.8.** *Suppose that $p > (4/3)(1/2+\gamma)$ with $\gamma > 0$ given. Then there is a procedure $T_p(U, V, \Delta, \gamma)$ that transforms $|\psi\rangle^{\otimes r}$ into a combination of states of the form $|\phi\rangle^{\otimes r}|\varphi\rangle_A$ up to an error amplitude of $e^{-r\gamma^2}$, where $\varphi - \varphi_V \in [-\Delta/5, \Delta/5]$. $T_p$ uses less than $2r$ instances of $\mathrm{PE}(V, \Delta/5)$ and an average of $\langle n\rangle < 2r$ reflections. We have $\langle \Gamma^n\rangle \leq \Gamma^{3r}$ for $1 \leq \Gamma \leq 7/4$.*

*Proof.* We begin $T_p$ by adding ancilla qubits in state $\sqrt{3/4}|0\rangle + \sqrt{1/4}|1\rangle$. Define $|\tilde{\psi}\rangle = |\psi\rangle(\sqrt{3/4}|0\rangle + \sqrt{1/4}|1\rangle)$ and $|\tilde{\phi}\rangle = |\phi\rangle|0\rangle$. For the purposes of using ER, let $\Pi_0 = \mathbb{1} \otimes |1\rangle\langle 1|$, where the second factor acts on the ancilla. We apply $\mathrm{ER}((V \otimes |0\rangle\langle 0|) \oplus \Pi_0, \Pi_0, \Delta, \gamma)$ and, provided $j > r/2$, we use $T(\tilde{\psi}, \tilde{\phi})$ (defined in IV.1) on the registers whose state is now indicated to be $|\tilde{\phi}^{\perp}\rangle$. The number of instances of $T$ applied is less than $r/2$. The bounds on the number of reflections used follow by Lemmas IV.2 and II.1. □

We need a method $T_{px}$ for transforming states as in Lemma IV.8 that is well-behaved even for small $p$. In order for the method to work we require that the probabilities of eigenphases other than $\varphi_V$ in $|\psi\rangle$ have a boundedness property. This will ensure that if $|\psi\rangle$ has a big overlap with an eigenstate of $V$, then this eigenstate is $|\phi\rangle$.

**Definition IV.9.** *We say that $\varphi_V$ is a $(\gamma, \delta)$-dominant eigenphase of $V = V_0 \oplus \Pi_0$ in $|\psi\rangle$ if for every $\varphi$ and associated projector $\Pi$ onto eigenspaces of $V_0$ with eigenphases in $I = [\varphi - \delta, \varphi + \delta]$, $|\Pi|\psi\rangle|^2 > \gamma$ implies $\varphi_V \in I$. We refer to $\delta$ as the resolution at which $\varphi_V$ is dominant and take $\Pi_0$ to be zero-dimensional if it is not specified.*

$T_{px}$ performs the transformation in the following steps. The first determines whether it is possible to confidently find a small interval for $\varphi_V$ without changing each of the $r$ copies of $|\psi\rangle$ by much. We then measure each copy so as to project it onto $|\psi\rangle$ or $|\psi^{\perp}\rangle$. We ensure that the probability of recovering a large number of copies of $|\psi\rangle$ is high. If some $|\psi^{\perp}\rangle$ are found or if we found that $\varphi_V$ can be determined sufficiently well, we use the recovered copies of $|\psi\rangle$ to learn a small interval containing $\varphi_V$ and then transform the states. The first step is encapsulated by the next lemma.

**Lemma IV.10.** *Suppose that $\varphi_V$ is a $(p_m - 3\gamma, \delta)$-dominant eigenphase of $V = V_0 \oplus \Pi_0$ in $|\psi\rangle$, where $\Pi_0$ is a projector that we can use to control other operations, $\Pi_0|\phi\rangle = 0$, $p_m - 3\gamma > 1/2$ and $p_m < 1$. Let $\delta' = \min(\delta/2, \Delta/4)$. Then, using $2r$ instances of phase estimation oracles $\mathrm{PE}(V, \delta'/2)$ and $r$ reflections around $|\psi\rangle$, we can implement an isometry $\mathrm{ER}_x(V, \Pi_0, \Delta, \delta, p_m, \gamma)$ with the following property up to an error amplitude of $5e^{-r\gamma^2}$: $\mathrm{ER}_x$ transforms $|\psi\rangle^{\otimes r}$ into a*

*combination of states of the form*

$$|\psi\rangle^{\otimes j}|\psi^{\perp}\rangle^{\otimes(r-j)}|b\rangle_A|j\rangle_B \ , \tag{8}$$

*where if* $p > p_m$*, then* $j = r$ *and* $b = 1$*; if* $p \leq p_m - 2\gamma$*, then* $j = r$ *and* $b = 0$*; and otherwise* $j \geq r/20$*.*

*Proof.* We use the notation introduced in the proof of Lemma IV.7 and apply $r$ instances of $\mathrm{PE}'(V, \delta'/2)$ to accommodate the special subspace associated with $\Pi_0$. We look (reversibly) for the first interval $I_l = [(l-1)\delta', (l+1)\delta']$, $l \in \{0, 1, \ldots, \lceil 2\pi/\delta' \rceil - 1\}$, containing at least $(p_m - \gamma)r$ of the phases in registers $A_i$. If no such interval exists, we set the state of $A$ to $|0\rangle_A$, else we set it to $|1\rangle_A$. Any temporary storage required in the reversible classical computation of the content of $A$ is erased. We then reverse the instances of $\mathrm{PE}'$ used and make $|\psi\rangle$-measurements to determine which of the $r$ input registers are in state $|\psi\rangle$. Finally, we move the $j$ registers in this state to the front and set the state of $B$ to $|j\rangle_B$.

As in the proof of Lemma IV.7, after phase estimation, for some set $S$ the state is a combination of products of $|\phi\rangle_i|w_i\rangle_{A_iE_i}$ for $i \in S$ and $|\eta\rangle_{iA_iE_i}$ for $i \notin S$. There exists $l_0$ such that $I_{l_0} \supset [\varphi_V - \delta'/2, \varphi_V + \delta'/2]$. Suppose that we conceptually measure the registers $A_i$ before the reversal of the phase estimation oracles. Let $k_l$ be the number of measured phases that are in $I_l$. Because $2\delta' < \Delta$, $k_{l_0} = |S|$. In particular, the measured phases in principle determine the members of $S$. (We can use this for the analysis but not for the procedure.) We consider the three cases $p > p_m$, $p \leq p_m - 2\gamma$ and $p_m \geq p > p_m - 2\gamma$. First, if $p > p_m$, then, from Hoeffding's inequality applied to $|S|$, the probability that $|S| > (p_m - \gamma)r$ is at least $1 - e^{-2r\gamma^2}$. Hence, with error amplitude at most $e^{-r\gamma^2}$, there is a $k_l \geq (p_m - \gamma)r$, and register $A$ contains 1 before the reversals of the phase estimation oracles. The reversals successfully restore the initial state up to the given error.

Consider next $p \leq p_m - 2\gamma$. The probability that $|S| \geq (p_m - \gamma)r$ is bounded by $e^{-2r\gamma^2}$. We show that the probability of finding a $k_l \geq (p_m - \gamma)r$ is small. For this purpose, consider the set $H$ of $l$ such that $\varphi_V \notin I'_l = [(l - 3/2)\delta', (l + 3/2)\delta']$. For $l \in H$, any measured phase in $I_l$ is associated with an eigenphase in $I'_l$ and therefore different from $\varphi_V$. Because $(3/2)\delta' < \delta$, the dominance condition ensures that such eigenphases occur with probability at most $p_m - 3\gamma$ in $|\psi\rangle$. To obtain a good bound on the mentioned probability, we consider the $k_l$'s according to $l$'s location in a small partition of $H$. For any $F \subseteq H$, let $\varphi(F)$ be the set of eigenvalues of $V$ in $\bigcup\{I'_l | l \in F\}$ and $P(F)$ the total probability of eigenphases in $\varphi(F)$ in $|\psi\rangle$. We claim that we can in principle partition $H = F_1 \cup \ldots \cup F_{16}$ such that $P(F_i) \leq p_m - 3\gamma$. First note that $\sum_l P(\{l\}) \leq 4$ because each eigenphase occurs in at most 4 of the $I'_l$. We can construct the $F_i$ greedily. Initialize $i = 1$ and set

$F_i = \emptyset$. Then step through $l \in H$. If $P(F_i \cup \{l\}) \leq p_m - 3\gamma$, add $l$ to $F_i$ and proceed to the next $l$. If not, because $p_m - 3\gamma > 1/2$, either $P(F_i) > 1/4$ or $P(\{l\}) > 1/4$. If $P(F_i) > 1/4$, initialize $F_{i+1} = \{l\}$, update $i$ to $i+1$ and proceed to the next $l$. If $P(\{l\}) > 1/4$, set $F_{i+1} = F_i$, reset $F_i$ to $\{l\}$, update $i$ to $i+1$ and proceed to the next $l$. At the end of the procedure, $P(F_{i'}) > 1/4$ for all $i' < i$. As we observed above, for all $l \in H$, $P(\{l\}) \leq p_m - 3\gamma$, and the claim follows. Let $k(F_j)$ be the number of measured phases that are in $\bigcup_{l \in F_j} I_l$. Because these phases are due to eigenphases in $\varphi(F_i)$, it follows from Hoeffding's inequality that $k(F_i) \geq (p_m - \gamma)r$ with probability at most $e^{-8r\gamma^2}$. If $l$ satisfies that $\varphi_V \in I'_l$, then the measured phases in $I_l$ are associated with $\varphi_V$. This is because $(7/2)\delta' < \Delta$. Probabilistic reasoning can be applied, so we conclude that the probability of $k_{\max} \geq (p_m - \gamma)r$ is at most $e^{-2r\gamma^2} + 16e^{-8r\gamma^2} < 17e^{-2r\gamma^2}$. Thus, with an error amplitude of at most $5e^{-r\gamma^2}$, register $A$ contains 0 before the phase estimation reversals, and the reversals successfully restore the initial state.

Finally, consider $p_m - \gamma \geq p > p_m - 2\gamma$. Because $p_m - 3\gamma > 1/2$, the probability that $|S| > r/2$ is at least $1 - e^{-2r\gamma^2}$. Therefore, except for an error amplitude of at most $e^{-r\gamma^2}$, $k_{\max} > p - \gamma > p_m - 3\gamma > 1/2$ and $k_{\max} = |S|$. We now assume this condition. After reversing the phase estimation oracles, the $i$'th system is in state $|\phi\rangle$ if $i \in S$ and $|\phi^\perp\rangle$ otherwise. Let $j$ be the number of $|\psi\rangle$ observed in the $|\psi\rangle$-measurements. The probability of detecting $|\psi\rangle$ when the state is $|\phi\rangle$ or $|\phi^\perp\rangle$ is $p$ and $1-p$, respectively. The average value of $j$ is $p|S| + (1-p)(r - |S|) \geq r/2$, since $p > 1/2$ and $|S| > r/2$. By Hoeffding's inequality, for fixed $S$, the probability of the event $E$ that $j \leq (1-x)r/2$ is bounded by $e^{-rx^2}$. To determine a bound on the overall probability $P$ of $E$, we must use amplitude addition over different $S$. Thus $\sqrt{P} \leq \sum_S \sqrt{P_S} e^{-rx^2/2}$. There are $2^r$ possible $S$, so the worst case sum of the $\sqrt{P_S}$ is $2^{r/2}$. Therefore, $P \leq e^{r(\ln(2) - x^2)}$. We set $x = 9/10$ and use $\ln(2) - (9/10)^2 < -1/10$ to see that the amplitude for having $j < r/20$ is bounded by $e^{-r/20}$. By amplitude addition, the overall error amplitude is bounded by $e^{-r\gamma^2} + e^{-r/20} < 2e^{-r\gamma^2}$, since $\gamma < 1/6$.

To complete the proof, it suffices to determine the maximum error amplitude. The maximum error bound comes from the second case and is given by $5e^{-r\gamma^2}$. □

As we noted for the error amplitude in Lemma IV.7, the error in Lemma IV.10 is such that the $r$ input systems' state remains in the tensor product of the span of $|\psi\rangle$ and $|\phi\rangle$.

The last lemma of this section gives the properties of the parallel state transformation procedure $T_{px}$ that we outlined above.

**Lemma IV.11.** *Assume that $\varphi_V$ is a $(1-4\gamma, \delta)$-dominant eigenphase of $V$ in $|\psi\rangle$ with $1-4\gamma > 2/3$.*

*Then there is a procedure $T_{px}$ that transforms $|\psi\rangle^{\otimes r}$ into a combination of $|\phi\rangle^{\otimes r}|\varphi\rangle_A$ and $|\psi\rangle^{\otimes r}|\#\rangle_A$, where $\varphi - \varphi_V \in [-\Delta/5, \Delta/5]$, A's state is $|\#\rangle$ if $p \leq 1-3\gamma$ and $|\varphi\rangle$ if $p > 1-\gamma$, and the error amplitude is bounded by $6e^{-r\gamma^2/36}$. The procedure uses less than $4r$ instances of $PE(V, \delta'/2)$, where $\delta' = \min(\delta/2, \Delta/4)$, and an average number of reflections bounded by $\langle n\rangle \leq 5r$. We have $\langle \Gamma^n\rangle \leq \Gamma^{7r}$ for $1 \leq \Gamma \leq 7/4$.*

Note that the procedure implicitly provides overlap information. That is, if the transformation succeeds, the overlap satisfies $p > 1-3\gamma$.

*Proof.* We use the overlap suppression trick and change each copy of $|\psi\rangle$ to $|\tilde{\psi}\rangle = |\psi\rangle(\sqrt{3/4}|0\rangle + \sqrt{1/4}|1\rangle)$ and define $|\tilde{\phi}\rangle = |\phi\rangle|0\rangle$. Let $\Pi_0 = \mathbb{1} \otimes |1\rangle\langle 1|$. We apply the procedure $\mathrm{ER}_x(V \otimes |0\rangle\langle 0| \oplus \Pi_0, \Pi_0, \Delta, \delta, 3(1-\gamma)/4, 3\gamma/4)$ of the previous Lemma. If the first output register (system A in Eq. (8)) contains $b = 1$ or if the second register (system B in Eq. (8)) has $j < r$, we continue. Otherwise we set the return register to $|\#\rangle$ and stop.

To continue the procedure, we apply $T_p(U, V, \Delta, (3/4)\gamma)$ (defined in IV.8) to the first $j$ registers (that now contain $|\tilde{\psi}\rangle$), omitting the initial overlap-suppressing steps as they have already been done. The specification of $ER_x$ and the assumption $1 - 4\gamma > 2/3$ ensures that the overlap is big enough to apply $T_p$, which returns $\varphi$. We then apply the appropriate instances of $T$ (defined in IV.1) to the remaining $r-j$ registers to transform $|\tilde{\psi}^{\perp}\rangle$ into $|\tilde{\phi}\rangle$. The reflections around $|\tilde{\phi}\rangle$ implicitly require $\varphi$. To finish we return the $r$ registers and $|\varphi\rangle_A$.

The error amplitudes associated with the different steps must be added. From the application of $\mathrm{ER}_x$ we get $5e^{-r(3/4)^2\gamma^2}$ to which $T_p$ adds at most $e^{-(r/20)(3/4)^2\gamma^2}$.

The number of instances of phase estimation oracles used comes from the application of $\mathrm{ER}_x$ and $T_p$. The average number of reflections is bounded by the sum of $r$ (from applying $\mathrm{ER}_x$), $2j$ (from applying $T_p$), and at most $4(r-j)$ (from using instances of $T$ as in Lemma IV.2). The reflections in $T_p$ are from applications of at most $j/2$ instances of $T$. The total number of instances of $T$ is bounded by $r$. To get the tail bounds, apply the bounds from Lemmas IV.2 and II.1, with an additional offset of $r$ for the first set of reflections. □

## V. STATE TRANSFORMATIONS ALONG A PATH

We consider paths of eigenstates $|\psi_s\rangle$ of unitary operators $U_s$ with eigenphases $\varphi_s$ and gaps $\Delta_s$ as defined in the introduction. We assume the ability to apply any $U_s$ and to prepare $|\psi_0\rangle$. If the exact gaps are difficult to obtain, we take the $\Delta_s$ to be known lower bounds on the gaps.

We define $p_{s,t} = |\langle\psi_s|\psi_t\rangle|^2$. The goal is to transform copies of the initial state $|\psi_0\rangle$ into the final state $|\psi_1\rangle$. The transformations' complexities depend on what is known about the overlaps and the eigenphases along the path. They are designed to provide such information if it is not already known, so that future transformations can be performed more efficiently.

**Theorem V.1.** *Suppose that we know a subsequence* $0 = s_0 < \ldots < s_n = 1$ *of* $[0,1]$ *such that* $p_{s_k,s_{k+1}} \geq 1/3$*, and phases* $\tilde{\varphi}_i$ *satisfying* $\tilde{\varphi}_i - \varphi_{s_i} \in [-\Delta/4, \Delta/4]$*. We can then transform* $|\psi_0\rangle$ *to* $|\psi_1\rangle$ *with* $m$ *reflections* $\mathrm{R}(U_{s_i}, \tilde{\varphi}_i, \Delta_{s_i})$ *where* $\langle m\rangle < 4n$ *and* $\langle\Gamma^m\rangle \leq \Gamma^{6n}$ *for* $1 \leq \Gamma \leq 7/4$*.*

*Proof.* It suffices to apply $T_m$ with the reflections instantiated by reflection oracles to advance from each state to the next. The complexities follow from Lemma IV.2 and II.1. □

Given sufficiently large overlaps, the phases can be inferred to sufficient precision during a parallel state transformation. Note that the eigenphase $\varphi_0$ of $|\psi_0\rangle$ for $U_0$ can be determined to within $\Delta_0/4$ by one call to a phase estimation oracle with resolution $\Delta/4$ and input state $|\psi_0\rangle$. We therefore assume that a phase sufficiently close to $\varphi_0$ is known and reflections around $|\psi_0\rangle$ can be applied.

**Theorem V.2.** *Suppose that we know a subsequence* $0 = s_0 < \ldots < s_n = 1$ *of* $[0,1]$ *such that* $p_{s_k,s_{k+1}} > (4/3)(1/2 + \gamma)$ *with* $\gamma > 0$*. We can then transform* $|\psi_0\rangle^{\otimes r}$ *into* $|\psi_1\rangle^{\otimes r}$ *with an error amplitude of* $ne^{-r\gamma^2}$*. The transformation requires* $2nr$ *instances of phase estimation* $\mathrm{PE}(U_{s_i}, \Delta_{s_i}/5)$ *and an average of* $\langle m\rangle < 2nr$ *reflections* $\mathrm{R}(U_{s_i}, \tilde{\varphi}_i, \Delta_{s_i})$*. Furthermore* $\langle\Gamma^m\rangle \leq \Gamma^{3rn}$ *for* $1 \leq \Gamma \leq 7/4$*. The transformation provides phases* $\tilde{\varphi}_i$ *satisfying* $\tilde{\varphi}_i - \varphi_{s_i} \in [-\Delta/5, \Delta/5]$ *for* $i > 0$*.*

*Proof.* It suffices to apply $T_p(U_{s_{i-1}}, U_{s_i}, \Delta_{s_i}, \gamma)$ $n$ times to transform the states. The complexities follow from Lemma IV.8 and II.1. □

The number of underlying calls to phase estimation oracles per copy of $|\psi_1\rangle$ produced by the procedure of Thm. V.2 is within a constant factor of that of Thm. V.1 (where phase estimation is used for the implementation of the reflections). The implementations of the oracles in the former case have an additional overhead to achieve the error goal, see Sect. VI.

Theorems V.1 and V.2 suggest that we can obtain state transformations along paths with complexities bounded by the path length. In particular, if the overlap probabilities $p_{s_i,s_{i+1}}$ are bounded above by $\cos(\theta)^2$, the path length is at least $n\theta$. Although it is possible for $n$ to be much smaller than the path length due to shortcuts, generically we do not expect this. If the angular rate $\chi(s) = \|(\mathbb{1} - |\psi_s\rangle\langle\psi_s|)|\partial_s\psi_s\rangle\|$ along the path is defined and constant, $\chi(s) = \chi$, then regularly

spaced $s_i = \pi/(8\chi)$ ensure that the overlap conditions for the theorems above are satisfied and $n = \lceil 8L/\pi \rceil$. On the other hand, if many overlaps are close to 1, $n$ could be large compared to $L$. We can eliminate this possibility if we have sufficient information about the overlaps, or after the first transformation by checking $n$ overlaps during the transformation, as the next lemma shows.

**Lemma V.3.** *Let $\theta, \underline{\theta}, \overline{\theta} > 0$ with $\underline{\theta} < \overline{\theta} < \underline{\theta} + \theta < \pi/2$ be given. Consider a procedure for transforming copies of $|\psi_0\rangle$ into copies of $|\psi_1\rangle$ in $n$ steps, where the steps transform from $|\psi_{t_{i-1}}\rangle$ to $|\psi_{t_i}\rangle$ with $\arccos(|\langle\psi_{t_{i-1}}|\psi_{t_i}\rangle|) < \theta$. Neither $n$ nor the $t_i$ need to be deterministic, but we assume that after $|\psi_{t_i}\rangle$ has been reached, information required to perform reflections and call overlap oracles for states $|\psi_{t_j}\rangle$ with $j \leq i$ is available. We can then modify the procedure so that it outputs a sequence $S = \{0 = s_0 < \ldots < s_k = 1\}$ satisfying $\arccos(|\langle\psi_{s_{i-1}}|\psi_{s_i}\rangle|) \leq \overline{\theta} + \theta$ for all $i$ and $\underline{\theta} \leq \arccos(|\langle\psi_{s_{i-1}}|\psi_{s_i}\rangle|)$ for $i < k$ so that $L(S) \geq \underline{\theta}(k-1)$. To do so requires $n$ calls to overlap oracles $\mathrm{OV}(|\psi_{t_l}\rangle, |\psi_{t_j}\rangle, (\overline{\theta} + \underline{\theta})/2, (\overline{\theta} - \underline{\theta})/2)$, $n$ explicit reflections on $|\psi_{t_l}\rangle$ and $k$ invocations of $T(|\psi_{t_l}\rangle, |\psi_{t_j}\rangle)$ with $\cos(\overline{\theta})^2 \leq p \leq \cos(\underline{\theta})^2$.*

The procedures we describe satisfy that the information required to call reflection and overlap oracles is available when needed by the modification in the lemma.

*Proof.* The $s_j$ are elements of $\{t_i\}_i$. We begin by setting $s_0 = 0$. We ensure that at the end of the $l$'th step ($l \geq 1$) of the modified procedure, the last $s_j$ that has been determined satisfies the invariant $\arccos(|\langle\psi_{s_j}|\psi_{t_l}\rangle|) \leq \overline{\theta}$. To do so, let $s_j$ be the last member of $S$ that has been determined before the $l$'th step. If $l = n$, set $s_{j+1} = 1$. Else, after the transformation into copies of $|\psi_{t_l}\rangle$ has been accomplished, call the overlap oracle $\mathrm{OV}(|\psi_{t_l}\rangle, |\psi_{s_j}\rangle, (\overline{\theta} + \underline{\theta})/2, (\overline{\theta} - \underline{\theta})/2)$ on the first copy of $|\psi_{t_l}\rangle$. Then use a reflection around $|\psi_{t_l}\rangle$ to determine whether the state was preserved. If not, or if the overlap oracle returned 1, set $s_{j+1} = t_l$. If the state was not preserved, then call $T(|\psi_{s_j}\rangle, |\psi_{t_l}\rangle)$ to restore it. This completes the modification of the $l$'th step and ensures the desired properties for the $s_j$ determined so far and the invariant. The lower bound on the length follows by adding up the lower bounds on the angular distances between successive $|\psi_{s_i}\rangle$ for $i < k$. □

When the overlaps $p_{s,t}$ or the angular rates $\chi(s)$ are unknown, the state transformation requires a recursive procedure to find a sequence of successive states with sufficiently high overlap. We assume that such states can be found, more specifically, we require that there are no jumps of angular distance equal to some given constant or greater. Our recursive state transformations involve binary subdivision of intervals. To transform the state from $|\psi_a\rangle$ to $|\psi_b\rangle$, we check whether we can do it directly at a cost of $C(a, b)$. If not, we recursively transform from $|\psi_a\rangle$ to $|\psi_{(a+b)/2}\rangle$ and then

from $|\psi_{(a+b)/2}\rangle$ to $|\psi_b\rangle$. We are interested in the total cost of the transformation. For our purposes, the cost is the number of times the unitaries $U_s$ are used. This is determined by the number of times phase-estimation is used either directly or indirectly when applying reflections. The resolution required is typically the gap, and the cost is related to the inverse gap (Sect. III), which can depend on the position along the path. To enable taking this into account we provide general tools for analyzing the complexity of recursive path transformations based on binary subdivision in Appendix B.

We define the symmetric binary interval tree on $[a, b]$, $\mathrm{BIT}(a, b)$, as the set of intervals constructed by starting with $T = T_0 = \{[a, b]\}$ and recursively adjoining $[c, (c+d)/2]$ and $[(c+d)/2, d]$ to $T$ for every $[c, d]$ in $T$. We also define the cost of $\mathrm{BIT}(a, b)$ as

$$C(\mathrm{BIT}(a, b)) = \sum_{[c,d]\in \mathrm{BIT}(a,b)} C(c, d) \; . \tag{9}$$

With the appropriate choice of the cost function $C$, this is the cost of a recursive state transformation procedure, and we show that it depends linearly on length and at worst logarithmically on the ratio of maximum to average angular rates. For simplicity, we use this estimate to state complexity bounds for state transformations where relevant, with the understanding that local-cost-sensitive estimates can be obtained if needed.

Let $C_{\max} = \sup\{C(s_1, s_2) \mid a \le s_1 < s_2 \le b\}$, $v_{\max} = \sup\{(L(s_2) - L(s_1))/(s_2 - s_1) \mid a \le s_1 < s_2 \le b, L(s_2) - L(s_1) > \theta\}$ and $v_{\mathrm{avg}} = (L(b) - L(a))/(b - a)$. If $L \le \theta$, define $v_{\max} = v_{\mathrm{avg}}$.

**Lemma V.4.** *If $C(c, d) = 0$ for $L(d) - L(c) \le \theta$, then*

$$C(\mathrm{BIT}(a, b)) \le \frac{2(L(b) - L(a))}{\theta} \left( \log_2 \frac{v_{\max}}{v_{\mathrm{avg}}} + 3 \right) C_{\max} \; . \tag{10}$$

The proof is given in Appendix B.

Let $v_{\max}$ and $v_{\mathrm{avg}}$ be as defined in Lemma V.4, where by default $a = 0$, $b = 1$ and $\theta$ is clear from context if not specified. Note that finiteness of $v_{\max}$ requires that the path has no jumps of angular distance $\theta$ or more.

**Theorem V.5.** *Suppose that $\varphi_s$ is a $(1 - 4\gamma, \delta)$-dominant eigenphase of $U_s$ in $|\psi_r\rangle$ for all $r < s$, where $1 - 4\gamma \ge 2/3$. Let $\delta' = \min(\delta/2, \Delta/4)$, $\theta < \arccos(\sqrt{1-\gamma})$ and $C = 2L(\log_2(v_{\max}/v_{\mathrm{avg}}) + 3)/\theta + 1$. Then we can transform $|\psi_0\rangle^{\otimes r}$ into $|\psi_1\rangle^{\otimes r}$ with an error amplitude bounded by $6Ce^{-r\gamma^2/36}$ with at most $4rC$ instances of phase estimation oracles with precision at least $\delta'$, an average number of reflections bounded by $\langle n\rangle \le 5rC$ and $\langle \Gamma^n\rangle \le \Gamma^{7rC}$ for $1 \le \Gamma \le 7/4$.*

*Proof.* To implement the transformation, we apply $T_{px}$ of Lemma IV.11 recursively to the intervals of the BIT, terminating at intervals where the transformation succeeds. The lemma guarantees that the transformation succeeds if the angular length of the interval being tried is less than $\arccos(\sqrt{1-\gamma})$.

To determine the complexity of the transformation, we need to consider a modified tree cost. Let $C_\theta(a,b) = C(a,(a+b)/2) + C((a+b)/2,b)$ if $L(b) - L(a) > \theta$ and $C_\theta(a,b) = 0$ otherwise. Define

$$\tilde{C}_\theta(\mathrm{BIT})(a,b) = C(a,b) + C_\theta(\mathrm{BIT})(a,b)\ . \tag{11}$$

This accounts for the fact that the shortest intervals in the tree that require action can be associated with arbitrarily short angular lengths. It is their parents whose angular length must be too long for terminating the recursion. With $C(a,b) = 1$, $\tilde{C}_\theta(\mathrm{BIT})(0,1)$ is an upper bound on the number of intervals for which transformation is attempted. According to Lemma V.4, this is bounded by $C = 2L(\log_2(v_{\max}/v_{\mathrm{avg}})+3)/\theta+1$. The rest follows by multiplying the complexities in Lemma IV.11 by $C$ and applying Lemma II.1. For the error amplitude we used amplitude addition. □

The next theorem can be applied when little information on overlaps is available, but we know sufficient eigenphase ranges for performing the necessary reflections. For this we need to consider the case where the transformation has probability of success $p_s < 1$ when $L(b) - L(a) \le \theta$. In this case, the process of subdividing $[a,b]$ may continue indefinitely. For $p_s > 1/2$, the expected number of intervals considered is finite with an exponentially decreasing tail probability. This follows from the theory of Galton-Watson processes, but in Appendix C we give a statement and proof sufficient for our purposes.

**Theorem V.6.** *Let* $\theta = \arccos(\sqrt{1/3}) \approx 0.96$. *Suppose that we know phases* $\tilde{\varphi}_s$ *satisfying* $\tilde{\varphi}_s - \varphi_s \in [-\Delta/4, \Delta/4]$. *We can transform* $|\psi_0\rangle$ *into* $|\psi_1\rangle$ *using* $\langle n\rangle \le \bar{n} = 40\,L(\log_2(v_{\max}/v_{\mathrm{avg}}) + 3)/\theta + 10$ *reflections, with* $\langle \Gamma^n\rangle \le \Gamma^{36\bar{n}}$ *for* $1 \le \Gamma \le 14/13$.

*Proof.* To implement the transformation, we apply $T'_{mx}$ to the intervals of the BIT recursively. ($T'_{mx}$ is defined after Cor. IV.6.) The recursion terminates when a transformation succeeds. For the intervals of angular length greater than $\theta$, Cor. IV.6 characterizes the distribution of the number of reflections used whether or not the transformation succeeds. For intervals $I$ of angular length at most $\theta$, the number of subintervals that need to be tried is characterized by Lemma C.1, where the success probability satisfies $p_s \ge 19/20$. Whether an interval needs to be tried depends only on whether any of the intervals containing it (that is, above it in the BIT) succeeded. Because

of Cor. IV.5, Lemma II.2 can be applied. Thus, according to Cor. IV.6, the total number $n_I$ of reflections used in our transforming across $I$ satisfies $\langle n_I \rangle < 9p_s/(2p_s - 1) \leq 10$ and $\langle \Gamma^{n_I} \rangle \leq \Gamma^{11(1+2/(2p_s-1))} \leq \Gamma^{36}$ for $1 \leq \Gamma \leq 14/13 < \min\{(1/(2\sqrt{p_s(1-p_s)}))^{1/11}, 8/7\}$. To finish the proof, as was noted in the proof of Thm. V.5, the number of intervals of the BIT whose parents have angular length greater than $\theta$ is bounded by $C = 4L(\log_2(v_{\max}/v_{\text{avg}}) + 3)/\theta + 1$. These intervals form a subtree $\text{BIT}_\theta$. Every reflection can be associated with the smallest interval in $\text{BIT}_\theta$ for whose traversal it was used. The statistics of the number of reflections associated with each such node are bounded by the ones we obtained for intervals of angular length at most $\theta$. Thus, we can apply Lemma II.1 to complete the proof. □

## VI. SUMMARY OF COMPLEXITIES

The most salient complexities are summarized in Table I. The bounds apply uniformly for $\bar{L} = \max(\pi/2, L)$, $0 < \Delta \leq \pi$, and $0 < \epsilon < 1/2$. We also define $\mathcal{L} = \bar{L}\,(\log(v_{\max}/v_{\text{avg}}) + 2)$. In order to obtain the bounds, it is necessary to take into account the error amplitudes contributed by two sources and make sure they do not exceed the error goal of the algorithm. The first is in the implementation of phase estimation and reflection oracles, and the second in our multi-copy transformations. Calls to either oracle in our algorithms require $\mathcal{O}(\log(1/\delta)/\Delta)$ uses of the underlying unitary operator for error amplitude $\delta$. If the state transformation requires $M$ phase estimations or reflections, then we can set $\delta = \epsilon/(2M)$ to ensure that the total error is bounded by $\epsilon/2$, since error amplitudes are sub-additive. Thus the complexity in terms of uses of the relevant unitary operators is $\mathcal{O}(M \log(M/\epsilon)/\Delta)$. The additional error in our multi-copy state transfer algorithm is bounded by $e^{-\Omega(r)}$ per state transformation attempt and is given explicitly in Theorems V.2 and V.5. In our algorithms, the number of phase estimation and reflection oracle calls per copy is linearly related to the number $n$ of state transformation attempts. The latter determines the total error from this contribution, which is $ne^{-\Omega(r)}$. Thus, for an error goal of $\epsilon/2$ we can set $r = \Theta(\log(n/\epsilon))$. This requires $\mathcal{O}(n \log(n/\epsilon))$ total phase estimation and reflection oracle calls. After accounting for the error in the implementation of these oracle calls, we obtain a complexity per copy of $\mathcal{O}(n \log(n \log(n/\epsilon)/\epsilon)/\Delta) = \mathcal{O}(n(\log(n/\epsilon) + \log\log(n/\epsilon))/\Delta) = \mathcal{O}(n \log(n/\epsilon)\Delta)$.

The formal meaning of the columns in Table I for assumed knowledge can be determined from the statements of the referenced lemmas and theorems. Having knowledge of overlap approximations means knowing enough about the $p_{s,t}$ to be able to pick $0 = s_0 < \ldots < s_n = 1$ such that $p_{s_{j-1},s_j}$ is large but bounded away from 1. This ensures that transforming along the $s_i$ is possible and efficient

| Knowledge assumed | | | | | | |
|---|---|---|---|---|---|---|
| Overlap approx-imations? | Overlap lower bounds? | Eigenphase ranges? | Eigenphase dominance? | Cost per copy | Number of copies required | Reference |
| √ | | √ | | $\mathcal{O}\left(\frac{\bar{L}}{\Delta}\log\left(\frac{\bar{L}}{\epsilon}\right)\right)$ | 1 | Thm. V.1, Lemma V.3 |
| | √ | | | $\mathcal{O}\left(\frac{n}{\Delta}\log\left(\frac{n}{\epsilon}\right)\right)$ | $\Theta\left(\log\left(\frac{n}{\epsilon}\right)\right)$ | Thm. V.2 |
| | | √ | | $\mathcal{O}\left(\frac{\mathcal{L}}{\Delta}\left(\log\left(\frac{\mathcal{L}}{\epsilon}\right)\right)\right)$ | 1 | Thm. V.6 |
| | | | √ | $\mathcal{O}\left(\frac{\mathcal{L}}{\Delta}\left(\log\left(\frac{\mathcal{L}}{\epsilon}\right)\right)\right)$ | $\Theta\left(\log\left(\frac{\mathcal{L}}{\epsilon}\right)\right)$ | Thm. V.5 |

TABLE I: Path transformation complexities. The entry in the column "number of copies required" gives the minimum needed by the referred-to algorithm to achieve the desired error amplitude. In this case, the error amplitude applies to all copies simultaneously, so unless there are strong error correlations, individual copies may have substantially less error. We have not determined the extent to which the transformations of the copies can be parallelized. If eigenphase dominance applies, we assume $\Delta$ is also a lower bound on the resolution for dominance. The maximum angular velocity $v_{\max}$ can be bounded with $\theta = \Omega(1)$ in Lemma V.4. The results in Appendix D may also be helpful. $n$ is determined by $0 = s_0 < \ldots < s_n = 1$ where $p_{s_l, s_{l+1}} > 1/2 + \Omega(1)$. We have $L = \mathcal{O}(n)$, but $n$ could be substantially larger than $L$ if we do not use preprocessing as in Lemma V.3.

in terms of path length. Knowing overlap lower bounds ensures the former only. When we say that eigenphase ranges are known, we mean that for all $s$, we know an interval (or more generally, a set) containing $\varphi_s$ such that the distance from this interval to every other eigenphase of $U_s$ is at least $\Delta$. This is sufficient for implementing the reflections with low error. The eigenphase dominance condition ensures that we can statistically distinguish the wanted eigenphase when using multiple copies of the states to infer adequate eigenphase ranges. The formal definition for a path is in Thm. V.5 based on Def. IV.9.

Lemma V.3 can be used to preprocess the transformation steps so that the complexity of the first row of Table I applies to subsequent transformations. Use of Lemma V.3 requires $\mathcal{O}(n)$ calls to overlap oracles, where $n$ is the number of actual transformation steps used. Thus, the complexity has an additive term of $\mathcal{O}(n \log(n/\epsilon)^2)$, according to the note after Def. III.4. However, if the recursive subdivision technique is used as in the last two rows of the table, the use of overlap oracles can be avoided.

We have given the key complexities in terms of global quantities that are simple to state. The

complexities actually depend on local aspects of the path. In particular, if the gaps $\Delta_s$ are typically large compared to the minimum, sections of the path can be traversed much more quickly. This can be taken into account by a finer complexity analysis, for example by taking advantage of Lemma B.2.

Our analyses apply to paths of non-degenerate eigenstates, but much of it can be extended to paths of eigenspaces as follows. Suppose that the path is characterized by a family of spaces $Z_s$ consisting of a union of eigenspaces of $U_s$ whose set of eigenphases have a minimum distance $\Delta_s$ to all eigenphases of states orthogonal to $Z_s$. The multi-copy transformation algorithms used when we have insufficient information about the eigenphases require that $Z_s$ is an eigenspace. Reflections around $Z_s$ can be implemented with the functional calculus of $Z_s$ as noted after Def. III.3. The goal is to transform an arbitrary initial state $|\psi_0\rangle \in Z_0$ into some $|\psi_1\rangle \in Z_1$. To generalize our analysis, it is necessary to redefine the path length. Let $\Pi_s$ be the projector onto $Z_s$. We define $L([a,b]) = \sup L(a = s_0 < \ldots < s_n = b)$, where $L(s_0 < \ldots < s_n) = \sum_{j=0}^{n-1} \theta(s_j, s_{j+1})$ and $\theta(s,t) = \max\{\arccos(|\Pi_t \psi_s|) \,|\, \psi_s \in Z_s\}$. Note that having no large jumps in the path implies that the dimension of $Z_s$ is non-decreasing. The basic transformation steps are the same, but their analysis requires the observation that the reflections around the subspaces $Z_s$ and $Z_t$ are a direct sum of reflections on two-dimensional subspaces of the space spanned by $Z_s$ and $Z_t$, see Ref. [26]. Within each such subspace, the transformation behaves as expected. The relevant overlaps now depend on the relationship between the reflection axes in the mentioned two-dimensional subspaces.

## Acknowledgments

We thank A. Harrow for discussions regarding the algorithm in the unknown-eigenphase case. This research was supported by the 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. Contributions to this work by NIST, an agency of the US government, are not subject to copyright laws. This work was supported by the National Science Foundation under grant PHY-0803371 through the Institute for Quantum Information at the California Institute of Technology. We also thank the Laboratory Directed Research and Development Program at Los Alamos National Laboratory for support.

### Appendix A: Proof of Lemma II.2

Let $\mu$ be the measure for the probability distribution of its arguments. The conditional independence assumption is equivalent to having the $C_j$ and $W_j$ generated via the sequence of probabilistic transitions $\ldots \to (C_j, \mathbf{W}_j) \to (\mathbf{W}_j) \to (C_{j+1}, \mathbf{W}_{j+1}) \to \ldots$. It follows that $C_j$ is conditionally independent of the $C_k$ for $k < j$ given $\mathbf{W}_j$. Formally we have

$$\begin{aligned}
d\mu(\mathbf{C}_k, \mathbf{W}_k) &= d\mu(C_k, W_k, \mathbf{C}_{k-1}|\mathbf{W}_{k-1})d\mu(\mathbf{W}_{k-1}) \\
&= d\mu(C_k, W_k|\mathbf{W}_{k-1})d\mu(\mathbf{C}_{k-1}|\mathbf{W}_{k-1})d\mu(\mathbf{W}_{k-1}) \\
&= d\mu(C_k, W_k|\mathbf{W}_{k-1})d\mu(\mathbf{C}_{k-1}, \mathbf{W}_{k-1}) \\
&= d\mu(C_k|\mathbf{W}_k)d\mu(W_k|\mathbf{W}_{k-1})d\mu(\mathbf{C}_{k-1}, \mathbf{W}_{k-1}) \\
&\vdots \\
&= \left(\prod_{j=1}^{k} d\mu(C_j|\mathbf{W}_j)d\mu(W_j|\mathbf{W}_{j-1})\right) \\
&= \left(\prod_{j=1}^{k} d\mu(C_j|\mathbf{W}_j)\right) d\mu(\mathbf{W}_k) \ ,
\end{aligned}$$

where the omitted identities involve applying the first steps recursively to the last term.

Given the constraints on $\Gamma$, we obtain

$$\begin{aligned}
\langle \Gamma^{C_{\mathrm{tot},k}} \rangle &= \int \Gamma^{\sum_{j=1}^{k} C_j} d\mu(\mathbf{C_k}, \mathbf{W_k}) \\
&= \int \left(\prod_{j=1}^{k} \int \Gamma^{C_j} d\mu(C_j|\mathbf{W_j})\right) d\mu(\mathbf{W_k}) \\
&= \int \left(\prod_{j=1}^{k} \int \left(1 - V_j + V_j \int \Gamma^{C_j} d\mu(C_j|\mathbf{W_j}, V_j)\right) d\mu(V_j|\mathbf{W_j})\right) d\mu(\mathbf{W_k}) \\
&\leq \int \left(\prod_{j=1}^{k} \int (1 - V_j + V_j \Gamma^{\tilde{C}}) d\mu(V_j|\mathbf{W_j})\right) d\mu(\mathbf{W_k}) \\
&= \int \left(\prod_{j=1}^{k} \int \Gamma^{\tilde{C}V_j} d\mu(V_j|\mathbf{W_j})\right) d\mu(\mathbf{W_k}) \\
&= \int \left(\prod_{j=1}^{k} \int \Gamma^{\tilde{C}V_j} d\mu(C_j|\mathbf{W_j})\right) d\mu(\mathbf{W_k}) \\
&= \int \Gamma^{\tilde{C}\sum_{j=1}^{k} V_j} d\mu(\mathbf{C}_k, \mathbf{W}_k) \\
&= \int \Gamma^{\tilde{C}\sum_{j=1}^{k} V_j} d\mu(\mathbf{C}_k) \ ,
\end{aligned}$$

because $V_j$ is a function of $C_j$. To get the desired bound, we let $k \to \infty$, use the monotone convergence theorem, and apply the bound on $\langle \Lambda^m \rangle$ with $\Lambda = \Gamma^{\tilde{C}}$:

$$\begin{aligned} \lim_{k\to\infty} \int \Gamma^{\tilde{C}\sum_{j=1}^{k} V_j} d\mu(\mathbf{C}_k) &= \int \lim_{k\to\infty} \Gamma^{\tilde{C}\sum_{j=1}^{k} V_j} d\mu(\mathbf{C}_k) \\ &= \langle \Gamma^{\tilde{C}m} \rangle \leq \Lambda^{\tilde{m}} \ . \end{aligned}$$

### Appendix B: Recursive transformations complexity

Let $L(s)$ be the length of the path $|\psi_t\rangle$ from $t = 0$ to $t = s$. Define $\underline{s}(l) = \inf\{s : L(s) \geq l, s \in [0,1]\}$ and $\overline{s}(l) = \sup\{s : L(s) \leq l, s \in [0,1]\}$. We have $\underline{s}(l) \leq \overline{s}(l)$ and the state is constant on the open interval $(\underline{s}(l), \overline{s}(l))$. The functions $\underline{s}$ and $\overline{s}$ are monotone. Given $l \in [L(a), L(b)]$ and a distance scale $\theta$, we define a local maximum speed variation at $l$ by

$$\begin{aligned} \sigma_\theta(l,[a,b]) = \ \sup\Big\{ \tfrac{s_4-s_1}{s_3-s_2} \mid\ & a \leq s_1 \leq s_2 < s_3 \leq s_4 \leq b, \\ & L(s_2) < l < L(s_3), \\ & L(s_3) - L(s_2) > \theta, \\ & L(s_4) - L(s_1) \leq 2\theta \Big\} \ . \end{aligned} \tag{B1}$$

If the set under the supremum is empty, let $\sigma_\theta(l,[a,b]) = 1$. To justify the description of $\sigma_\theta$ as a speed variation, we define average speeds between $l_1$ and $l_2 > l_1$ by $\overline{v}(l_1,l_2) = (l_2 - l_1)/(\underline{s}(l_2) - \overline{s}(l_1))$ and $\underline{v}(l_1,l_2) = (l_2 - l_1)/(\overline{s}(l_2) - \underline{s}(l_1))$, allowing for constant sections on the path. A more direct local speed variation is

$$\begin{aligned} \rho_\theta(l,[a,b]) = \ \sup\Big\{ \tfrac{\overline{v}(l_2,l_3)}{\underline{v}(l_1,l_4)} \mid\ & L(a) \leq l_1 \leq l_2 < l < l_3 \leq l_4, \\ & l_3 = l_2 + \theta, \\ & l_4 = \min(l_1 + 2\theta, L(b)) \Big\} \ . \end{aligned} \tag{B2}$$

If the set under the supremum is empty, let $\rho_\theta(l,[a,b]) = 1$. The description of $\sigma_\theta(l,[a,b])$ as a speed variation at $l$ at scale $\theta$ comes from the observation that

**Lemma B.1.** *For all intervals $I$*

$$\sigma_\theta(l,I) \leq 2\rho_\theta(l,I) \tag{B3}$$

*with equality if $l \in [L(a) + 2\theta, L(b) - 2\theta]$, the path is continuous and has no constant intervals.*

*Proof.* The set $X$ in the definition of $\sigma_\theta(l,I)$ is empty iff $L(b) - L(a) \leq \theta$ or $l \notin [L(a), L(b)]$. If the latter holds, then the set $Y$ in the definition of $\rho_\theta(l,I)$ is empty. If not, then either $Y$ is empty or

$L(b) = L(a) + \theta$. In this case $Y = \{y\}$ with $y \geq 1$. Consider $r = (s_4 - s_1)/(s_3 - s_2) \in X$, with $s_i$ as in Eq. B1. Let $l_i = L(s_i)$ and define $l'_1 = l_1$, $l'_4 = \min(l_1 + 2\theta, L(b)) \geq l_4$ and choose $l'_2$, $l'_3$ such that $l'_3 = l'_2 + \theta$, $l_2 \leq l'_2 < l < l'_3 \leq l_3$. Then

$$\begin{aligned} r &\leq \frac{\overline{s}(l'_4) - \underline{s}(l'_1)}{\underline{s}(l'_3) - \overline{s}(l'_2)} \\ &\leq 2\left(\frac{l'_3 - l'_2}{l'_4 - l'_1}\right)\frac{\overline{s}(l'_4) - \underline{s}(l'_1)}{\underline{s}(l'_3) - \overline{s}(l'_2)} \\ &\leq 2\rho_\theta(l, I). \end{aligned}$$

For the reverse inequality, given $\epsilon > 0$ arbitrarily small, choose $l_i$ such that $r' = \overline{v}(l_2, l_3)/\underline{v}(l_1, l_4) \geq \rho_\theta(l, I) - \epsilon$. The constraint on $l$ implies that $l_4 = l_1 + 2\theta$, so $r' = 2(\overline{s}(l_4) - \underline{s}(l_1))/(\underline{s}(l_3) - \overline{s}(l_2))$. Let $\delta > 0$ be arbitrarily small. Let $s_i = \overline{s}(l_i)$ for $i = 2, 4$ and $s_j = \underline{s}(l_j)$ for $j = 1, 3$. Continuity implies that $L(s_i) = l_i$. If $l_2 > l_1$, then $s_2 > s_1$ and we let $s'_2 = s_2 - \delta$. Else $l_3 < l_4$ and we let $s'_3 = s_3 + \delta$. Unassigned $s'_j$ are set to $s_j$. The assumptions imply that for $\delta$ small enough, the $s'_i$ satisfy the constraints in Eq. B1, showing that $\sigma_\theta(l, I) \geq 2\rho_\theta(l, I) - \epsilon$. Letting $\epsilon \downarrow 0$ gives the desired result. □

The next lemma gives a bound on the cost of a symmetric binary interval tree that is sensitive to local variations.

**Lemma B.2.** *If $C(c, d) = 0$ for $L(d) - L(c) \leq \theta$, then*

$$C(\mathrm{BIT}(a, b)) \leq \frac{1}{\theta}\sum_{k=0}^{\infty}\frac{1}{2^k}\int_{L(a)}^{L(b)}\Big(\lfloor \log_2\left(\sigma_{2^k\theta}(l, [a, b])\right)\rfloor + 1\Big)\overline{C}_{2^k\theta}(l)dl\,, \tag{B4}$$

*where $\overline{C}_{\theta'}(l) = \sup\{C(s_1, s_2) \mid L(s_1) < l < L(s_2), \theta' < L(s_2) - L(s_1) \leq 2\theta'\}$.*

*Proof.* The bound is obtained in three steps. First we separately sum over intervals $[s_1, s_2]$ in the tree in each length class $2^k\theta < L(s_2) - L(s_1) \leq 2^{k+1}\theta$. Second, we uniformly, randomly assign the cost $C(s_1, s_2)$ to the open interval between $L(s_1)$ and $L(s_2)$ and integrate over the length variable. Third, we use bounds on costs and numbers of intervals in the tree spanning the value of the length

variable. The steps are implemented in the following sequence of identities and inequalities:

$$\begin{aligned}
C(\mathrm{BIT}(a,b)) &= \sum_{[s_1,s_2]\in \mathrm{BIT}(a,b)} C(s_1,s_2) \\
&= \sum_{k=0}^{\infty} \sum_{\substack{[s_1,s_2]\in \mathrm{BIT}(a,b),\\ 2^k\theta < L(s_2)-L(s_1) \le 2^{k+1}\theta}} C(s_1,s_2) \\
&= \sum_{k=0}^{\infty} \sum_{\substack{[s_1,s_2]\in \mathrm{BIT}(a,b),\\ 2^k\theta < L(s_2)-L(s_1) \le 2^{k+1}\theta}} C(s_1,s_2)\frac{1}{L(s_2)-L(s_1)}\int_{L(s_1)}^{L(s_2)} dl \\
&\le \sum_{k=0}^{\infty} \sum_{\substack{[s_1,s_2]\in \mathrm{BIT}(a,b),\\ 2^k\theta < L(s_2)-L(s_1) \le 2^{k+1}\theta}} C(s_1,s_2)\frac{1}{2^k\theta}\int_{L(a)}^{L(b)} \mathbf{1}_{[L(s_1),L(s_2)]}(l)dl \\
&\le \sum_{k=0}^{\infty} \frac{1}{2^k\theta}\int_{L(a)}^{L(b)} \left( \sum_{\substack{[s_1,s_2]\in \mathrm{BIT}(a,b),\\ 2^k\theta < L(s_2)-L(s_1) \le 2^{k+1}\theta}} \mathbf{1}_{[L(s_1),L(s_2)]}(l) \right) \overline{C}_{2^k\theta}(l)dl
\end{aligned}$$

To finish the proof, note that the number in parenthesis in the last line,

$$\left| \left\{ [s_1,s_2] \in \mathrm{BIT}(a,b) \mid L(s_1) < l < L(s_2), 2^k\theta < L(s_2)-L(s_1) \le 2^{k+1}\theta \right\} \right| ,$$

is the size of a set of nested intervals in the tree. Let $[s_1,s_4]$ be the biggest and $[s_2,s_3]$ the smallest of these intervals. Because of the way the tree is constructed, the number of these intervals is given by

$$\log_2((s_4-s_1)/(s_3-s_2)) + 1 \le \sigma_{2^k\theta} + 1 ,$$

where the inequality follows from the definition of $\sigma_{2^k\theta}$ . □

Finally, we give the proof of Lemma V.4 stated in the text.

*Proof of Lemma V.4.* We begin by bounding $\int_{L(a)}^{L(b)} \log_2(\sigma_{\theta'}(l,[a,b]))dl$ for $\theta' \ge \theta$ in terms of the maximum and average angular rates. If $L(b)-L(a) \le \theta'$, then $\sigma_{\theta'}(l,[a,b]) = 1$ so the integral is 0. Assume $L(b)-L(a) > \theta'$ and let $\epsilon > 0$ be arbitrarily small. For every $l$, choose $a \le s_1(l) \le s_2(l) < s_3(l) \le s_4(l) \le b$ such that the following hold: The $s_i = s_i(l)$ satisfy the constraints given in the definition of $\sigma_{\theta'}(l,[a,b])$ (see Eq. B1), $\sigma_{\theta'}(l,[a,b]) \le 2^{\epsilon}(s_4(l)-s_1(l))/(s_3(l)-s_1(l))$, and the $s_i(l)$ are measurable functions. We can now bound

$$\int_{L(a)}^{L(b)} \log_2(\sigma_{\theta'}(l,[a,b]))dl \;\le\; (L(b)-L(a))\epsilon + \int_{L(a)}^{L(b)} \log_2 \frac{s_4(l)-s_1(l)}{s_3(l)-s_2(l)} dl \; .$$

Jensen's inequality, applied to the concave $\log_2$ function, gives

$$\int_{L(a)}^{L(b)} \log_2 \frac{s_4(l)-s_1(l)}{s_3(l)-s_2(l)} dl \;\le\; (L(b)-L(a)) \log_2 \frac{1}{L(b)-L(a)} \int_{L(a)}^{L(b)} \frac{s_4(l)-s_1(l)}{s_3(l)-s_2(l)} dl \;.$$

We bound the denominator inside the integral with the inequality $(s_3(l)-s_2(l)) > \theta'/v_{\max}$, which does not depend on the variable of integration, so we can continue by bounding the integral of the numerator. We first change the order of integration

$$\int_{L(a)}^{L(b)} (s_4(l)-s_1(l))\, dl \;=\; \int_{L(a)}^{L(b)} dl \int_{s_1(l)}^{s_4(l)} dt = \int_a^b dt \int_{L(a)}^{L(b)} \mathbf{1}_{\{l \mid t \in [s_1(l), s_4(l)]\}}(l) dl \;.$$

If $t \in [s_1(l), s_4(l)]$, then $L(l) \le L(s_4(l)) \le L(s_1(l)) + 2\theta' \le L(t) + 2\theta'$. Similarly, $L(l) \ge L(s_1(l)) \ge L(s_4(l)) - 2\theta' \ge L(t) - 2\theta'$. It follows that the inner integral is bounded by $4\theta'$. This gives the bound

$$\frac{1}{L(b)-L(a)} \int_{L(a)}^{L(b)} (s_4(l)-s_1(l)) dl \;\le\; \frac{1}{L(b)-L(a)} \int_a^b 4\theta' dt = \frac{4\theta'(b-a)}{L(b)-L(a)} = \frac{4\theta'}{v_{\text{avg}}} \;.$$

Combining the bounds and letting $\epsilon$ go to 0, we obtain

$$\int_{L(a)}^{L(b)} \log_2(\sigma_{\theta'}(l, [a,b])) \le (L(b)-L(a)) \left( \log_2 \frac{v_{\max}}{v_{\text{avg}}} + 2 \right) \;.$$

Substituting into the bound of Lemma B.2, we get

$$\begin{aligned} C(\text{BIT}(a,b)) \;&\le\; \frac{1}{\theta} \sum_{k=0}^{\infty} \frac{1}{2^k} (L(b)-L(a)) \left( \log_2 \frac{v_{\max}}{v_{\text{avg}}} + 3 \right) C_{\max} \\ &\le\; \frac{2(L(b)-L(a))}{\theta} \left( \log_2 \frac{v_{\max}}{v_{\text{avg}}} + 3 \right) C_{\max} \;, \end{aligned}$$

as desired. □

## Appendix C: A Galton-Watson lemma

**Lemma C.1.** *Consider the random process starting from* $S = \{([a,b], \text{active})\}$ *which is defined recursively as follows: For all* $N = ([c,d], \text{active})$ *in* $S$*, replace* $N$ *with with* $([c,d], \text{inactive})$ *and with probability* $1-p_s$ *add* $([c,(c+d)/2], \text{active})$ *and* $([(c+d)/2, d], \text{active})$ *to* $S$*. Let* $S_\infty$ *be the possibly infinite set obtained by running this process countably many times. For* $p_s > 1/2$*, the expected size of* $S_\infty$ *is* $\langle |S_\infty| \rangle = p_s/(2p_s-1)$ *and* $\langle \Gamma^{|S_\infty|} \rangle \le \Gamma^{1+2/(2p_s-1)}$ *for* $1 \le \Gamma \le 1/(2\sqrt{p_s(1-p_s)})$*.*

*Proof.* We outline the proof, omitting some necessary existence arguments. Either $S_\infty = \{([a,b], \text{inactive})\}$, which happens with probability $p_s$, or $S_\infty = \{([a,b], \text{inactive})\} \cup S_l \cup S_r$, where $S_l$ and $S_r$ are independent with the same statistics as $S_\infty$. The latter happens with probability $1-p_s$.

Thus $\langle |S_\infty| \rangle = p_s + 2(1-p_s)\langle |S_\infty| \rangle$, or $\langle |S_\infty| \rangle = p_s/(2p_s - 1)$. Similarly, using the independence of $S_l$ and $S_r$ again, we find $\langle \Gamma^{|S_\infty|} \rangle = \Gamma \left( p_s + (1-p_s)\bar{\langle} \Gamma^{|S_\infty|} \rangle^2 \right)$. This is solved by

$$\langle \Gamma^{|S_\infty|} \rangle = \frac{1}{2(1-p_s)\Gamma} \left( 1 \pm \sqrt{1 - 4\Gamma^2 p_s(1-p_s)} \right) .$$

The relevant solution is the negative branch, as can be seen by checking that $\langle 1^{|S_\infty|} \rangle = 1$. The maximum value of $\Gamma$ for which it is defined is $\Gamma = 1/(2\sqrt{p_s(1-p_s)})$. For this $\Gamma$, $\langle \Gamma^{|S_\infty|} \rangle = 2p_s\Gamma$. The following sequence of inequalities together with log-convexity completes the proof:

$$\begin{aligned} 2p_s &= 1 + (2p_s - 1) \\ &\leq 1 + \frac{2p_s - 1}{4p_s(1-p_s)} \\ &= 1 + \left( \frac{1}{2p_s - 1} \right) \left( \frac{(2p_s-1)^2}{4p_s(1-p_s)} \right) \\ &\leq \left( 1 + \frac{(2p_s-1)^2}{4p_s(1-p_s)} \right)^{\frac{1}{2p_s-1}} \\ &= \left( \frac{1}{4p_s(1-p_s)} \right)^{\frac{1}{2p_s-1}} = \Gamma^{\frac{2}{2p_s-1}} \end{aligned}$$

at the maximum $\Gamma$. □

### Appendix D: Angular rate estimation

Here are some tools for estimating average angular rates for paths of eigenstates of normal operators.

**Lemma D.1.** *Let $H$ be a normal operator with eigenstate $|\psi\rangle$ having eigenvalue $\lambda$. Suppose $H+S$ is normal and has an eigenspace $V$ with eigenvalue $\lambda + \delta$ gapped by $\Delta$. Then the maximum angle from $|\psi\rangle$ to $V$ is bounded by* $\arcsin(\|(S-\delta)|\psi\rangle\|/\Delta)$.

*Proof.* Let $\Pi$ be the projector onto the orthogonal complement of $V$. It suffices to show that $|\langle \psi' | \Pi | \psi \rangle| \leq \|(S-\delta)|\psi\rangle\|/\Delta$ for all $|\psi'\rangle$. We can use the "resolvent trick" to express

$$\begin{aligned} \Pi &= \frac{1}{2\pi i} \int \frac{1}{z} - (z - ((H+S) - (\lambda + \delta)))^{-1} dz \\ &= \frac{1}{2\pi i} \int (z - ((H+S) - (\lambda+\delta)))^{-1} ((H+S) - (\lambda+\delta)) dz/z \, , \end{aligned}$$

where the integral is over a circle of radius $d$ less than $\Delta$ around 0. Thus

$$\begin{aligned} |\langle \psi' | \Pi | \psi \rangle| &= \left| \langle \psi' | \frac{1}{2\pi} \int (z - ((H+S) - (\lambda+\delta)))^{-1} (S-\delta) dz/z | \psi \rangle \right| \\ &\leq \|(S-\delta)|\psi\rangle\| (\Delta - d)^{-1} \, . \end{aligned}$$

The result follows by letting $d$ go to 0. □

**Corollary D.2.** *Let $H(s)$ be a family of normal operators, $H(s)|\psi(s)\rangle = \lambda(s)|\psi(s)\rangle$ with all objects differentiable at $s = t$. Let $\Pi_t^\perp = \mathbb{1} - |\psi(t)\rangle\langle\psi(t)|$. If the eigenvalues $\lambda(s)$ are gapped with gap $\Delta$ in a neighborhood of $s = t$, then*

$$\left\| \Pi_t^\perp \frac{d|\psi(s)\rangle}{ds}\bigg|_{s=t} \right\| \leq \left\| \frac{dH(s)}{ds}\bigg|_{s=t} \right\| / \Delta \, . \tag{D1}$$

*Proof.* It suffices to apply first-order perturbation theory to Lemma D.1, noting that $H(s) = H(t) + (t-s)(dH(s)/ds|_{s=t}) + o(|t-s|)$ and $\lambda(s) = \lambda(t) + (t-s)\langle\psi(t)|(dH(s)/ds|_{s=t})|\psi(t)\rangle + o(|t-s|)$. □

---


[1] D. Aharonov and A. Ta-Shma. Adiabatic quantum state generation and statistical zero knowledge. *Proceedings of the 35th Annual ACM STOC*, pages 20–29, 2003.

[2] P. Amara, D. Hsu, and J. E. Straub. Global energy minimum searches using an approximate solution of the imaginary time schroedinger equation. *The J. of Phys. Chem.*, 97:6715–6721, 1993.

[3] A. Ambainis and O. Regev. An elementary proof of the quantum adiabatic theorem. *arXiv:quant-ph/0411152*, 2004.

[4] M. H. S. Amin. Consistency of the adiabatic theorem. *Phys. Rev. Lett.*, 102(22):220401–4, 2009.

[5] A. Aspuru-Guzik, A. D. Dutoi, P. J. Love, and M. Head-Gordon. Simulated quantum computation of molecular energies. *Science*, 309:1704–1707, 2005.

[6] J. E. Avron, R. Seiler, and L. G. Yaffe. Adiabatic theorems and applications to the quantum hall effect. *Commun. Math. Phys.*, 110(1):33–49, 1987.

[7] S. Boixo, E. Knill, and R. D. Somma. Eigenpath traversal by phase randomization. *Quantum Inf. and Comp.*, 9:833–855, 2009.

[8] S. Boixo and R. D. Somma. Necessary condition for the quantum adiabatic approximation. *Phys. Rev. A*, 81(3):032308, 2010.

[9] A. M. Childs, E. Deotto, E. Farhi, J. Goldstone, S. Gutmann, and A. J. Landahl. Quantum search by measurement. *Phys. Rev. A*, 66:032314, 2002.

[10] E. Farhi, J. Goldstone, S. Gutmann, J. Lapan, A. Lundgren, and D. Preda. A quantum adiabatic evolution algorithm applied to random instances of an NP-complete problem. *Science*, 292:472–476, 2001.

[11] E. Farhi, J. Goldstone, S. Gutmann, and M. Sipser. Quantum computation by adiabatic evolution. *arXiv:quant-ph/0001106*, 2000.

[12] A. B. Finnila, M. A. Gomez, C. Sebenik, C. Stenson, and J. D. Doll. Quantum annealing: A new method for minimizing multidimensional functions. *Chem. Phys. Lett.*, 219:343–348, 1994.

[13] G. A. Hagedorn and A. Joye. Elementary exponential error estimates for the adiabatic approximation. *J. Math. Anal. Appl.*, 267(1):235–246, 2002.

[14] A. W. Harrow, A. Hassidim, and S. Lloyd. Quantum algorithm for solving linear systems of equations. arXiv:0811.3171, 2008.

[15] W. Hoeffding. Probability inequalities for sums of bounded random variables. *J. Am. Stat. Assoc.*, 58:13–30, 1963.

[16] S. Jansen, M-B Ruskai, and R. Seiler. Bounds for the adiabatic approximation with applications to quantum computation. *J. Math. Phys.*, 48(10):102111–15, 2007.

[17] T. Kadowaki and H. Nishimori. Quantum annealing in the transverse ising model. *Phys. Rev. E*, 58:5355–5363, 1998.

[18] E. Knill, G. Ortiz, and R. Somma. Optimal quantum measurements of expectation values of observables. *Phys. Rev. A*, 75:012328, 2007.

[19] M. Lewenstein, A. Sanpera, V. Ahufinger, B. Damski, A. Sende, and U. Sen. Ultracold atomic gases in optical lattices: Mimicking condensed matter physics and beyond. *Adv. in Phys.*, 56:243–379, 2007.

[20] D. A Lidar, A. T Rezakhani, and A. Hamma. Adiabatic approximation with better than exponential accuracy for many-body systems and quantum computation. *J. Math. Phys.*, page 102106, 2009.

[21] C. Marriott and J. Watrous. Quantum Arthur-Merlin games. *Computational Complexity*, 14(2):122–152, 2005.

[22] D. Nagaj, P. Wocjan, and Y. Zhang. Fast amplification of QMA. *Quant. Inf. Comp.*, 9:1053, 2009.

[23] O. Oreshkov and J. Calsamiglia. Adiabatic markovian dynamics. *arXiv:1002.2219*, 2010.

[24] G. Ortiz, J. E. Gubernatis, E. Knill, and R. Laflamme. Quantum algorithms for fermionic simulations. *Phys. Rev. A*, 64:022319/1–14, 2001.

[25] R. D. Somma, S. Boixo, H. Barnum, and E. Knill. Quantum simulations of classical annealing processes. *Phys. Rev. Lett.*, 101:130504–4, 2008.

[26] M. Szegedy. Quantum speed-up of markov chain based algorithms. *Foundations of Computer Science, 2004. Proceedings. 45th Annual IEEE Symposium on*, pages 32–41, 2004.

[27] T. Tulsi, L. Grover, and A. Patel. A new algorith for fixed point quantum search. *Quant. Inf. Comp.*, 6:483, 2006.

[28] P. Wocjan and A. Abeyesinghe. Speedup via quantum sampling. *Phys. Rev. A*, 78(4):042336, 2008.