

Universal Quantum Emulator

Iman Marvian¹ and Seth Lloyd^{1, 2}

¹*Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, MA 02139*

²*Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139*

(Dated: June 10, 2016)

We propose a quantum algorithm that emulates the action of an unknown unitary transformation on a given input state, using multiple copies of some unknown sample input states of the unitary and their corresponding output states. The algorithm does not assume any prior information about the unitary to be emulated, or the sample input states. To emulate the action of the unknown unitary, the new input state is coupled to the given sample input-output pairs in a coherent fashion. Remarkably, the runtime of the algorithm is logarithmic in D , the dimension of the Hilbert space, and increases polynomially with d , the dimension of the subspace spanned by the sample input states. Furthermore, the sample complexity of the algorithm, i.e. the total number of copies of the sample input-output pairs needed to run the algorithm, is independent of D , and polynomial in d . In contrast, the runtime and the sample complexity of incoherent methods, i.e. methods that use tomography, are both linear in D . The algorithm is *blind*, in the sense that at the end it does not learn anything about the given samples, or the emulated unitary. This algorithm can be used as a subroutine in other algorithms, such as quantum phase estimation.

PACS numbers:

I. INTRODUCTION

A universal quantum simulator is a machine that can be programmed to mimic the dynamics of other quantum systems [1]. The time evolution of the simulator obeys the same equations of motion as the evolution of the simulated system. A *universal quantum emulator*, on the other hand, is a machine that mimics the input-output relation of another system, by looking to the output of that system on some sample input states. Unlike a simulator, an emulator does not need to obey the same dynamical equations as of the emulated system.

In this Letter we introduce a quantum algorithm that emulates the action of an unknown unitary transformation on new given input states. The algorithm couples the new input state to multiple copies of some unknown *sample input-output pairs*, that is copies of some input states of the unitary as well as copies of the corresponding output states. We do not assume any prior information about the unitary to be emulated, or the given sample input states. The algorithm emulates the action of the unitary on any given state in the subspace spanned by the previously given input states, which could be much smaller than the system Hilbert space. Indeed, we are interested in the cases where d , the dimension of this subspace is constant or, at most, polylogarithmic in D , the dimension of the system Hilbert space.

Obviously, having multiple copies of sample input-output pairs we can perform measurements on them, and using state tomography find an approximate classical description of these states in a standard basis. This, in turn, yields the classical description of the unknown unitary transformation, which then can be used to simulate its action on the new given states. This approach, however, is highly inefficient and impractical: First of all, state tomography in a large Hilbert space is a hard task

and requires lots of copies of the sample states. Second, even if we find an approximate classical description of the unitary transformation, or if we are given its exact description, in general we cannot implement this transformation on a new given state efficiently.

More precisely, the approaches based on tomography run in time $\Omega(D)$ and need $\Omega(D)$ copies of state, where D is the dimension of the system Hilbert space. In contrast, the runtime of the algorithm proposed in this work is $\mathcal{O}(\log D)$ and polynomial in d , and its *sample complexity*, i.e. the total number of copies of the sample input-output pairs that are needed to run the algorithm, is independent of D and polynomial in d . Therefore, our algorithm is not only exponentially faster than the approaches based on tomography, its sample complexity is also dramatically lower.

It is interesting to compare this result with the scenario studied in [2], where one wants to *learn* an unknown unitary U by applying it for a finite number of times to some quantum states, so that later, when we do not have access to U , we can reproduce its effect on a new input state. It turns out that the strategy that maximizes the average fidelity, where the average is taken over all states in the Hilbert space, is an incoherent *measure-and-rotate* strategy, i.e. a method that uses tomography [2]. In contrast to this result, our work shows that under the practical assumption that the action of the unitary should be emulated in a low-dimensional subspace, and not the entire Hilbert space, the coherent methods are much more powerful than the incoherent ones.

This algorithm is *blind*, in the sense that at the end it does not learn anything about the given samples, or the emulated unitary. Although the algorithm uses a randomized technique, it always generates an output with high fidelity with the desired output state. Therefore, it can be used as a subroutine in larger algorithms, such

as quantum phase estimation. Another possible application of this algorithm is to use it to cancel the effect of an unknown unitary channel, without doing process tomography and finding the description of the unitary. Furthermore, the algorithm could be useful in the cases where we can prepare the input and the corresponding outputs of a unitary transformation efficiently, but we do not know how to implement the unitary itself.

II. PRELIMINARIES

We first present the algorithm for the special case of pure sample states, and later explain how it can be generalized to the case of mixed states as well.

Let $S_{\text{in}} = \{|\phi_k^{\text{in}}\rangle\langle\phi_k^{\text{in}}| : k = 1, \dots, K\}$ be a set of sample input states of the unitary U and $S_{\text{out}} = \{|\phi_k^{\text{out}}\rangle\langle\phi_k^{\text{out}}| = U|\phi_k^{\text{in}}\rangle\langle\phi_k^{\text{in}}|U^\dagger : k = 1, \dots, K\}$ be the corresponding outputs. Let \mathcal{H}_{in} and \mathcal{H}_{out} be the subspaces spanned by $\{|\phi_k^{\text{in}}\rangle : k = 1, \dots, K\}$ and $\{|\phi_k^{\text{out}}\rangle : k = 1, \dots, K\}$ respectively, and d be the dimension of these subspaces. We assume the set of input samples S_{in} contains sufficient number of different states to uniquely determine the action of U on the subspace \mathcal{H}_{in} (up to a global phase). It can be easily shown that having the classical description of the input and output states in S_{in} and S_{out} we can uniquely determine the action of U on any input state $|\psi\rangle \in \mathcal{H}_{\text{in}}$ (up to a global phase), if and only if the matrix algebra generated by S_{in} , that is the set of polynomials in the elements of S_{in} , is the full matrix algebra on \mathcal{H}_{in} , i.e. contains all operators with supports contained in \mathcal{H}_{in} ¹. Therefore, in the following we naturally assume this assumption is satisfied. Furthermore, we assume K the number of different sample input states in S_{in} is $\text{poly}(d)$.

To implement the algorithm, we need multiple copies of each sample state in S_{in} and S_{out} . Interestingly, at the end of the algorithm most of these states remain almost unaffected. Indeed, the main use of the given copies of sample states is to simulate *controlled-reflections* about these states.

Let $R^{\text{in}}(k) = e^{i\pi|\phi_k^{\text{in}}\rangle\langle\phi_k^{\text{in}}|}$ and $R^{\text{out}}(k) = e^{i\pi|\phi_k^{\text{out}}\rangle\langle\phi_k^{\text{out}}|}$ be the reflections about the input and output states $|\phi_k^{\text{in}}\rangle$

and $|\phi_k^{\text{out}}\rangle$, respectively. In the proposed algorithm we need to implement the controlled-reflections $R_a^{\text{in}}(k)$ and $R_a^{\text{out}}(k)$, defined as

$$R_a(k) = |0\rangle\langle 0|_a \otimes I + |1\rangle\langle 1|_a \otimes e^{i\pi|\phi_k^{\text{in}}\rangle\langle\phi_k^{\text{in}}|}, \quad (2.1)$$

where a is the label for the control qubit, and I is the identity operator on the main system. Note that we have suppressed the superscripts *in* and *out* in both sides.

Using the given copies of the sample states, we can efficiently simulate these controlled-reflections via the *density matrix exponentiation* technique of [3]. Based on this technique, in the Supplementary Material we present a new simple quantum circuit that uses n copies of an unknown state σ to simulate the unitary $e^{-it\sigma}$, or its controlled version $|0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes e^{-it\sigma}$, for any real t , with error $\epsilon = \mathcal{O}((t^2+1)/n)$, and in time $\mathcal{O}(n \log(n) \times \log(D))$, where D is the dimension of the Hilbert space (See Appendix C). This circuit only uses single qubit gates and controlled-Fredkin gates, i.e. the controlled-controlled-SWAP gates [4]. In the simplest case where the system is a qubit ($D = 2$), this circuit is basically simulating the Heisenberg interaction between the system and each given copy of state σ .

Therefore, in the following, where we present the algorithm, we assume all the controlled-reflections $\{R_a(k) : 1 \leq k \leq K\}$ can be efficiently implemented.

To simplify the presentation, we use the notation $W_a(k) \equiv R_a(k)H_aR_a(1)$, where again we have suppressed *in* and *out* superscripts in both sides. Here H_a denotes the Hadamard gate H acting on qubit a , where $H|0\rangle = |+\rangle$ and $H|1\rangle = |-\rangle$, and $|\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}$. The algorithm also uses a SWAP gate defined by $\text{SWAP}|\nu\rangle|\mu\rangle = |\mu\rangle|\nu\rangle$, for any pair of states $|\mu\rangle$ and $|\nu\rangle$.

III. THE ALGORITHM (SPECIAL CASE)

In this section we present and analyze the algorithm for the universal quantum emulator, in the special case where all the sample input-output pairs are pure states. Later, we present several generalizations of this algorithm, including to the case where the given samples contain mixed states.

Fig.(1) exhibits the quantum circuit that emulates the action of an unknown unitary transformation U on any given state $|\psi\rangle$ in the input subspace \mathcal{H}_{in} . For a general input state, which is not restricted to this subspace, this circuit first projects the state to this subspace, and if successful, then applies the unitary U to it.

In this algorithm (k_1, \dots, k_T) are T integers chosen uniformly at random from integers $1, \dots, K$, where T is a constant that determines the precision of emulation, and we choose it to be polynomial in d , and independent of D . Furthermore, state $|\phi_1^{\text{in}}\rangle$ (and $|\phi_1^{\text{out}}\rangle$) is one of the sample input states (and its corresponding output) which is chosen randomly at the beginning of the algorithm,

¹ If we know how unitary U transforms elements of S_{in} , then we also know how it transforms any operator in the matrix algebra generated by the elements of S_{in} . Therefore, if S_{in} generates the full matrix algebra on \mathcal{H}_{in} , then we know how U transforms any density operator with support contained in \mathcal{H}_{in} . On the other hand, if S_{in} does not generate the full matrix algebra on \mathcal{H}_{in} , then its commutant contains unitaries that are block-diagonal with respect to \mathcal{H}_{in} , and act non-trivially on this subspace. For any such unitary V , unitaries U and UV act exactly in the same way on all the input states in the set S_{in} . Therefore, just having the classical description of states in S_{in} and S_{out} , we cannot distinguish unitaries UV and U , for any unitary V in the commutant of S_{in} , even though they act differently on states in \mathcal{H}_{in} . This proves the claim.

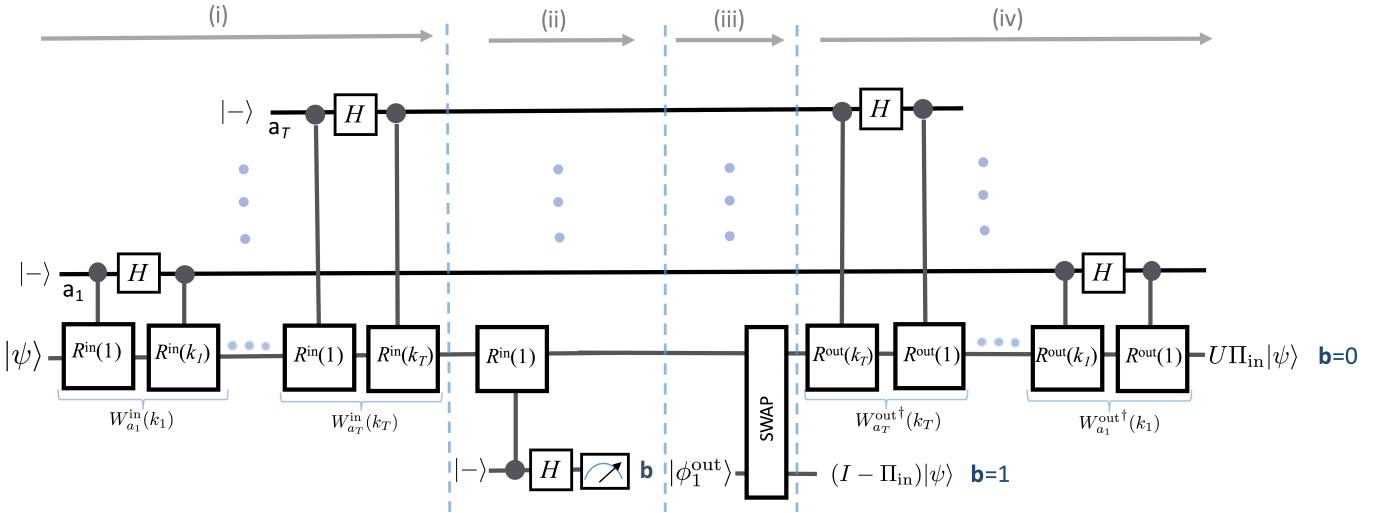


FIG. 1: The quantum circuit for emulating unitary transformation U for the special case of pure input-output sample pairs. Here k_1, \dots, k_T are $T = \text{poly}(d)$ integers chosen uniformly at random from integers $1, \dots, K$. We use the given copies of sample states in S_{in} and S_{out} to simulate the controlled-reflections $R_a^{\text{in}}(k)$ and $R_a^{\text{out}}(k)$, respectively.

and is fixed during the algorithm. In steps (i) and (iv) of the algorithm we implement, respectively, the unitaries $W_{a_i}^{\text{in}}(k_i)$ and $W_{a_i}^{\text{out}\dagger}(k_i)$ on the system and qubit a_i , for $i = 1, \dots, T$. As we explained before, all the conditional reflections $R_a^{\text{in}}(k)$ and $R_a^{\text{out}}(k)$ can be efficiently simulated using the given copies of states $|\phi_k^{\text{in}}\rangle$ and $|\phi_k^{\text{out}}\rangle$.

In step (ii) of the algorithm we perform a qubit measurement in the computational basis $\{|0\rangle, |1\rangle\}$. Then, after the measurement with probability $1 - \langle \psi | \Pi_{\text{in}} | \psi \rangle$ we get outcome $b = 1$, in which case we project the system to a state close to $(I - \Pi_{\text{in}})|\psi\rangle / \sqrt{1 - \langle \psi | \Pi_{\text{in}} | \psi \rangle}$, where Π_{in} is the projector to the subspace \mathcal{H}_{in} . On the other hand, with probability $\langle \psi | \Pi_{\text{in}} | \psi \rangle$ we get the outcome $b = 0$, in which case the final state of circuit is close to $U\Pi_{\text{in}}|\psi\rangle / \sqrt{\langle \psi | \Pi_{\text{in}} | \psi \rangle}$. In this case the algorithm consumes a copy of state $|\phi_1^{\text{out}}\rangle$, and returns a copy of state $|\phi_1^{\text{in}}\rangle$.

Note that, although the algorithm uses random integers (k_1, \dots, k_T) , for sufficiently large T it *always* transforms the input state $|\psi\rangle \in \mathcal{H}_{\text{in}}$ to a state with high fidelity with the desired output state $U|\psi\rangle$.

A. How it works

To simplify the following discussion, we first assume the initial state $|\psi\rangle$ is in the input subspace \mathcal{H}_{in} , and then we consider the general case.

To understand step (i) of the algorithm, we first focus on the reduced state of the system. From this point of view, during step (i) we are trying to erase the state of system and push it into $|\phi_1^{\text{in}}\rangle$, a state which is chosen randomly from the sample input set S_{in} . This erasing is done by repetitive *measuring and conditional mixing*.

Consider the first round of step (i), where we have

randomly chosen integer k_1 in $1, \dots, K$, and apply the unitary $W(k_1) = R_{a_1}(k_1)H_{a_1}R_{a_1}(1)$ to the system and qubit a_1 initially prepared in state $|-\rangle$. The effect of this transformation on the reduced state of system can be interpreted in the following way: we perform a measurement in $\{P = |\phi_1^{\text{in}}\rangle\langle\phi_1^{\text{in}}|, P^\perp = I - P\}$ basis, and if the system is found in state $|\phi_1^{\text{in}}\rangle$, we leave it unchanged; otherwise, we apply the random reflection $R^{\text{in}}(k_1)$ to it. Tracing over qubit a_1 , and averaging over the values of k_1 , we find that the overall effect of this transformation on the reduced state of system can be described by the quantum channel $\mathcal{W}(\rho) = P\rho P + \mathcal{D}(P^\perp\rho P^\perp)$, where

$$\mathcal{D}(\rho) = \frac{1}{K} \sum_{k=1}^K R^{\text{in}}(k) \rho R^{\text{in}}(k) . \quad (3.1)$$

In step (i) we repeat the above procedure for T times with T different ancillary qubits and random integers. Using the facts that (1) ancillary qubits are initially uncorrelated with each other, and (2) different random integers k_1, \dots, k_T are statistically independent of each other, we find that at the end of step (i) the average reduced state of system is described by state $\mathcal{W}^T(|\psi\rangle\langle\psi|)$.

Next, recall the assumption that the input set S_{in} generates the full matrix algebra in \mathcal{H}_{in} , an assumption which is crucial for being able to uniquely determine the action of U on \mathcal{H}_{in} . Interestingly, this assumption now translates to the fact that channels \mathcal{D} and \mathcal{W} have unique fixed point states with support contained in \mathcal{H}_{in} , namely the totally mixed state in \mathcal{H}_{in} , and $|\phi_1^{\text{in}}\rangle\langle\phi_1^{\text{in}}|$ respectively². Furthermore, since the reflections $R^{\text{in}}(k)$ are

² The fixed points of the unital channel \mathcal{D} are the commutants

block diagonal with respect to the input subspace \mathcal{H}_{in} , channels \mathcal{D} and \mathcal{W} map any initial state inside \mathcal{H}_{in} to a state with support contained in this subspace. It follows that for any initial state $|\psi\rangle \in \mathcal{H}_{\text{in}}$, at the end of step (i) with probability almost one the system should be in state $|\phi_1^{\text{in}}\rangle$. More precisely, the probability of finding the system in a state orthogonal to $|\phi_1^{\text{in}}\rangle$ is exponentially small in T .

To summarize, in step (i) we *erase* the initial state of system and push it into state $|\phi_1^{\text{in}}\rangle$. The fact that we have enough resources to uniquely determine the action of unitary U on any state in the input subspace \mathcal{H}_{in} , translates to the fact that any state in this subspace can be erased in a coherent unitary fashion.

But quantum information is conserved during a unitary evolution. This means that all the information about state $|\psi\rangle$ should now be encoded in the ancillary qubits $\mathbf{a} = a_1 \cdots a_T$. Furthermore, since we start with a global pure state, and since at the end of step (i) the reduced state of system is close to a pure state, we find that the joint reduced state of ancillary qubits \mathbf{a} should also be close to a pure state, denoted by $|\Psi(\mathbf{k})\rangle_{\mathbf{a}}$. Note that in addition to state $|\psi\rangle$, state $|\Psi(\mathbf{k})\rangle_{\mathbf{a}}$ also depends on the sample set S_{in} , and the random integers $\mathbf{k} \equiv (k_1, \dots, k_T)$. We conclude that at the end of step (i) with high probability the system and ancillary qubits \mathbf{a} are in the product state

$$[W_{a_T}^{\text{in}}(k_T) \cdots W_{a_1}^{\text{in}}(k_1)] |\psi\rangle |-\rangle^{\otimes T} \approx |\phi_1^{\text{in}}\rangle |\Psi(\mathbf{k})\rangle_{\mathbf{a}} . \quad (3.2)$$

Next, in step (ii) we basically perform a measurement on the system in $\{P = |\phi_1^{\text{in}}\rangle\langle\phi_1^{\text{in}}|, P^\perp = I - P\}$ basis. At this point we know that for any initial state $|\psi\rangle \in \mathcal{H}_{\text{in}}$ with probability almost one the system should be in state $|\phi_1^{\text{in}}\rangle$, in which case the measurement projects the ancillary qubit to state $|0\rangle$. On the other hand, if we project the qubit to state $|1\rangle$ it means the erasing has not been successful.

Finally, assuming $|\psi\rangle \in \mathcal{H}_{\text{in}}$, we find that applying steps (iii) and (iv) maps the system to state $U|\psi\rangle$. This can be seen by multiplying both sides of Eq.(3.2) in unitary U , and using the facts that $U|\phi_1^{\text{in}}\rangle = |\phi_1^{\text{out}}\rangle$, and

$(U \otimes I_a)W_a^{\text{in}}(k)(U^\dagger \otimes I_a) = W_a^{\text{out}}(k)$. This implies

$$\left[W_{a_1}^{\text{out}\dagger}(k_1) \cdots W_{a_T}^{\text{out}\dagger}(k_T) \right] |\phi_1^{\text{out}}\rangle |\Psi(\mathbf{k})\rangle_{\mathbf{a}} \approx U|\psi\rangle |-\rangle^{\otimes T} . \quad (3.3)$$

Eq.(3.3) means that, after step (ii) by preparing the system in the output sample state $|\phi_1^{\text{out}}\rangle$, which is given to us, and running all the operations in step (i) backward, with unitaries $R_a^{\text{in}}(k)$ replaced by $R_a^{\text{out}}(k)$, we get state $U|\psi\rangle$ at the end.

Using the fact that all unitaries $W_a(k) = R_a(k)H_aR_a(1)$ act trivially on the subspace orthogonal to \mathcal{H}_{in} , it can be easily seen that for a general input $|\psi\rangle$, which is not contained in \mathcal{H}_{in} , the algorithm first performs a projective measurement that projects the input state to the subspace \mathcal{H}_{in} , or the orthogonal subspace. Then, if the system is found to be in \mathcal{H}_{in} , which corresponds to outcome $b = 0$ in the measurement in step (ii), it applies the unitary U to the component of state in this subspace.

In the Supplementary Material we prove the following quantitative version of the above argument: Suppose we implement the quantum circuit presented in Fig. 1, using perfect controlled-reflections $R_a(k)$. Let \mathcal{E}_U be the quantum channel that describes the overall effect of the circuit on the system, in the case where we do not post-select based on the outcome of measurement in step (ii), which means we do not care if erasing has been successful or not. Then, for an arbitrary input state ρ , the (Uhlmann) fidelity [5, 6] of $\mathcal{E}_U(\rho)$ and the desired state $U\rho U^\dagger$ satisfies

$$F(\mathcal{E}_U(\rho), U\rho U^\dagger) \geq p_{\text{erase}}(\rho) = \langle\phi_1^{\text{in}}|\mathcal{W}^T(\rho)|\phi_1^{\text{in}}\rangle , \quad (3.4)$$

where $p_{\text{erase}}(\rho)$ is the probability that we have successfully erased the state of system (and pushed it into state $|\phi_1^{\text{in}}\rangle$), which corresponds to outcome 0 in the measurement (See Appendix A). On the other hand, if we post-select to the cases where the erasing has been successful, then for pure input state ρ , the fidelity between the output of the algorithm and the desired state $U\rho U^\dagger$ is lower bounded by $\sqrt{p_{\text{erase}}(\rho)}$.

Interestingly, as we show in the Supplementary Material, Eq.(3.4) holds in a much more general setting: suppose we run the above algorithm with any other choice of unitaries $W_a^{\text{in}}(\lambda)$ and $W_a^{\text{out}}(\lambda) = (U \otimes I_a)W_a^{\text{in}}(\lambda)(U^\dagger \otimes I_a)$ that couple the system to a qubit a , where λ is a random parameter chosen according to a probability distribution $p(\lambda)$. Then, Eq.(3.4) still holds for the channel $\mathcal{W}(\tau) = \sum_{\lambda} p(\lambda) \text{Tr}_a(W_a^{\text{in}}(\lambda)[\tau \otimes |-\rangle\langle-|_a]W_a^{\text{in}\dagger}(\lambda))$. In the Supplementary Material we use this generalization to extend the algorithm to the case where the samples contain mixed states.

Eq.(3.4) best captures the working principle behind this algorithm, which can be called *emulating via coherent erasing*. Note that using this equation we can determine for which input states, the emulation works well: if we have the required resources to coherently erase state ρ and bring the system to a pure state which we know

of its Kraus operators, i.e. the commutants of the reflections $\{R^{\text{in}}(k) = I - 2|\phi_k^{\text{in}}\rangle\langle\phi_k^{\text{in}}| : k = 1, \dots, K\}$. Since $\{|\phi_k^{\text{in}}\rangle\langle\phi_k^{\text{in}}| : k = 1, \dots, K\}$ generates the full matrix algebra on \mathcal{H}_{in} , it turns out that the only fixed points of \mathcal{D} with support contained in \mathcal{H}_{in} are multiples of the identity operator on this subspace. In the case of \mathcal{W} , definition $\mathcal{W}(\tau) = P\tau P + \mathcal{D}(P^\perp\tau P^\perp)$ implies that any fixed point τ of this channel satisfies $P^\perp\mathcal{D}(P^\perp\tau P^\perp)P^\perp = P^\perp\tau P^\perp$. Since \mathcal{D} is trace-preserving, this can hold if and only if $P^\perp\tau P^\perp = 0$, or $P^\perp\tau P^\perp$ is a fixed point of \mathcal{D} . The fact that the only fixed points of \mathcal{D} with support contained in \mathcal{H}_{in} are the multiples of the identity operator on \mathcal{H}_{in} , implies that the latter case is impossible. We conclude that the unique fixed point state of \mathcal{W} with supports contained in \mathcal{H}_{in} is $|\phi_1^{\text{in}}\rangle\langle\phi_1^{\text{in}}|$.

how transforms under unitary U , then we can emulate the action of unitary U on ρ .

B. Coordinates of the input state relative to the samples

As we saw before, at the end of step (i) all the information about the input state $|\psi\rangle$ is encoded in the ancillary qubits. Finding the explicit form of this encoded version of state clarifies an interesting interpretation of this algorithm: step (i) of the algorithm finds the *coordinates* of the input state $|\psi\rangle$ relative to the frame defined by the input samples $|\phi_{k_1}^{\text{in}}\rangle, \dots, |\phi_{k_T}^{\text{in}}\rangle$. Then, step (iv) reconstructs the state with exactly the same coordinates relative to the frame defined by the output samples $|\phi_{k_1}^{\text{out}}\rangle, \dots, |\phi_{k_T}^{\text{out}}\rangle$.

Let $|t\rangle_{\mathbf{a}} = |1\rangle^{\otimes t} \otimes |0\rangle^{\otimes(T-t)}$ be the state of qubits $a_1 \dots a_T$, in which $a_{t+1} \dots a_T$ are all in state $|0\rangle$, and the rest of qubits are in state $|1\rangle$. Then, at the end of step (i) the joint state of the system and $\mathbf{a} = a_1 \dots a_T$ is given by

$$|\phi_1^{\text{in}}\rangle \sum_{t=0}^{T-1} \langle \phi_1^{\text{in}} | \psi(t, \mathbf{k}) \rangle |t\rangle_{\mathbf{a}} + |\psi(T, \mathbf{k})\rangle |T\rangle_{\mathbf{a}}, \quad (3.5)$$

where the (unnormalized) vectors $|\psi(t, \mathbf{k})\rangle$ are defined via the recursive relation

$$|\psi(t+1, \mathbf{k})\rangle = R^{\text{in}}(k_{t+1}) P^{\perp} |\psi(t, \mathbf{k})\rangle, \quad (3.6)$$

and $|\psi(0, \mathbf{k})\rangle = |\psi\rangle$ ³. The argument in the previous section implies that for initial state $|\psi\rangle \in \mathcal{H}_{\text{in}}$, the typical norm of $|\psi(t, \mathbf{k})\rangle$ is exponentially small in t , and hence for sufficiently large T the last term in Eq.(3.5) is negligible. It follows that this expansion indeed describes a general recursive method for specifying any vector $|\psi\rangle \in \mathcal{H}_{\text{in}}$ in terms of the scalars $\{\langle \phi_1^{\text{in}} | \psi(t, \mathbf{k}) \rangle : t = 0, \dots, T-1\}$. These scalars only depend on the relation between $|\psi\rangle$ and states $|\phi_{k_1}^{\text{in}}\rangle, \dots, |\phi_{k_T}^{\text{in}}\rangle$, i.e. applying the same unitary transformation on $|\psi\rangle$ and these states leaves them invariant. We can interpret these scalars as the *coordinates* of vector $|\psi\rangle$ relative to the frame defined by states $|\phi_{k_1}^{\text{in}}\rangle, \dots, |\phi_{k_T}^{\text{in}}\rangle$. Then, in this language, the step (i) of the algorithm is a circuit for finding the coordinates of the given state $|\psi\rangle$ relative to the input frame $|\phi_{k_1}^{\text{in}}\rangle, \dots, |\phi_{k_T}^{\text{in}}\rangle$. Note that because of the no-cloning theorem [7], in order to find the coordinates of a quantum state in a reversible fashion and encode it in the ancillary qubits, we need to erase the state of system.

Therefore, the step (i) of this algorithm provides an efficient reversible method for finding the coordinates of

a given state with respect to a general frame, using multiple copies of states corresponding to that frame⁴. This method can be useful for other applications, where instead of implementing operations on the system directly, we first find the coordinates of state of system with respect to other quantum states, implement an operation on the coordinates, and then transform the state back to the physical space.

C. Runtime and error analysis

It follows from Eq.(3.4) that if we run the circuit in Fig.(1) with ideal controlled-reflections, then for any initial state $|\psi\rangle \in \mathcal{H}_{\text{in}}$, the trace distance between the output of the circuit and the desired state $U|\psi\rangle$ is less than or equal to ϵ , provided that we choose

$$T \geq \frac{d \times \log(8d\epsilon^{-2})}{1 - |\lambda_{\mathcal{D}}|}, \quad (3.7)$$

where $\lambda_{\mathcal{D}}$ is the eigenvalue of channel \mathcal{D} with the second largest magnitude (See Appendix B). Therefore, as one expects from the discussion in the previous section, the runtime is mainly determined by the the mixing time of the random unitary channel \mathcal{D} , or equivalently, its spectral gap $1 - |\lambda_{\mathcal{D}}|$.

It turns out that in the actual algorithm for the universal quantum emulator, where we need to simulate the controlled-reflections using the given copies of sample states, the dominant source of error is due to the imperfections in these simulations. In the Supplementary Material we show that for initial state $|\psi\rangle \in \mathcal{H}_{\text{in}}$, the transformation $|\psi\rangle \rightarrow U|\psi\rangle$ can be implemented with error $\epsilon > 0$ in the trace distance, using

$$N_{\text{tot}} = \tilde{\mathcal{O}}\left(\frac{d^2 \times \epsilon^{-1}}{(1 - |\lambda_{\mathcal{D}}|)^2}\right) \quad (3.8)$$

total copies of sample states, and in time $t_{\text{tot}} = \tilde{\mathcal{O}}(N_{\text{tot}} \times \log D)$, where $\tilde{\mathcal{O}}$ suppresses more slowly-growing terms (See Appendix B 2). Note that the only place where D , the dimension of Hilbert space, shows up in the analysis is in the simulation of the controlled-reflections.

D. Optimality

In contrast with the approaches based on tomography, whose runtime and sample complexity both scales (at

³ Note that we only use a $(T+1)$ -dimensional subspace of the 2^T -dimensional Hilbert space of the ancillary qubits. Indeed, this algorithm can be easily modified to be implemented using only $\mathcal{O}(\log T)$ ancillary qubits.

⁴ The given copies of each sample state can be interpreted as a Quantum Reference Frame (QRF) for a *direction* in the Hilbert space. A QRF usually refers to a reference frame for physical degrees of freedom, such as position or time, which is treated quantum mechanically [8–12]. In contrast, here we are using the concept of QRF in a more abstract sense. Therefore, the relevant symmetry group, which for the standard QRF’s is usually the groups such as $\text{SO}(3)$, $\text{U}(1)$ or Z_N [8–12], in this case is $\text{SU}(D)$.

least) linearly with D , the runtime t_{tot} of the above algorithm is only logarithmic in D , whilst its sample complexity N_{tot} is independent of D .

In general, the lowest achievable runtime is lower bounded by $\Omega(\log^\gamma D)$, where γ is a constant of order one depending on the circuit architecture. This is because, in general, the state of each qubit in the desired output state $U|\psi\rangle$ depends non-trivially on the state of all other qubits in state $|\psi\rangle$, as well as the state of all qubits in the sample states. This means that the runtime of the algorithm is bounded by the minimum time required for information about one qubit in the system to propagate to all other qubits. Since the quantum circuit is formed from local unitary gates each acting on few qubits, this time grows with $\log D$, the number of qubits. For instance, in the case where qubits are all on a line and interact only with their nearest neighbors, this time is of the order $\log D$.

Furthermore, it can be easily seen that the lowest achievable runtime scales, at least, linearly with d , the dimension of the input subspace. Indeed, just to check if the given state is inside the input subspace spanned by the sample states, one needs to interact with at least d different sample states, which requires time of order d . We conclude that the runtime of the proposed algorithm cannot be improved drastically.

IV. GENERALIZATIONS

As we explained at the end of Sec.(IIIA), the proposed algorithm for the universal quantum emulator works based on a simple and general principle, namely emulating via coherent erasing. Hence, it turns out that the algorithm can be generalized in several different ways. Some of the possible generalizations are presented in the Supplementary Material (See Appendix D). In particular,

1) We show that the algorithm can be generalized to the case where the sample input-output sets contain mixed states. More precisely, as long as the sample input set, and hence the sample output set, contains (at least) one state close to a pure state, we can still coherently erase the state of system and push it into this pure state. Then, we can emulate the action of the unknown unitary, using the same approach we used in the original algorithm. The main difference with the original version of algorithm is that, instead of the controlled-reflections, in the case of mixed states we need to implement *controlled-translations* with respect to the sample states, i.e. the unitaries $|0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes e^{-it\sigma_k}$, for a random value of t and a sample state σ_k .

2) We present a more efficient algorithm which works under certain extra assumptions about the sample input states. Namely, this algorithm assumes the input samples are states in an (unknown) orthonormal basis for the input subspace, plus one or more states in the conjugate

basis. The working principle behind this algorithm is again emulating via coherent erasing. In this algorithm the state of system is erased via a coherent measurement in the orthonormal basis followed by another coherent measurement in the conjugate basis.

3) We show the algorithm can be generalized to emulate the controlled version of unitary U , i.e. to implement the unitary $|0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes U$. This generalization is crucial for some applications such as quantum phase estimation.

4) We show that if the sample input states can only approximately be transformed to the sample output states via a unitary transformation, with an error bounded by δ , then the proposed algorithm emulates this unitary with error $\text{poly}(d) \times \delta$.

A. Emulating Projective Measurements

We can use the algorithm presented in this paper to emulate projective measurements, using copies of sample states where each sample comes with a (classical) label that specifies the outcome of the measurement for this state.

Suppose in the algorithm presented in Fig.(1) we use the same set of states as both the input and the output samples. In this case, the algorithm basically performs the two-outcome projective measurement that projects the given state to the subspace spanned by the sample states, or its complement. Then, any arbitrary projective measurement can be implemented as a sequence of these two-outcome projective measurements.

This approach, however, only works if the set of sample states in each subspace generates the full matrix algebra in that subspace. But, in general, to specify a projective measurement we only need to specify the subspaces that correspond to different outcomes, and therefore this extra assumption is unjustified in this context. In the Supplementary Material we present a different efficient algorithm for emulating projective measurements, which does not require this extra assumption. This algorithm also uses random controlled-reflections about the given sample states.

V. DISCUSSION

We presented an efficient algorithm for emulating unitary transformations and projective measurements. The algorithm uses a novel randomized technique, and works based on a simple principle, which can be called emulating via coherent erasing.

The important problem of efficient emulation of general quantum channels is left open. It is interesting to see if there are some physically relevant assumptions under which one can efficiently emulate quantum channels.

This could have applications, e.g. in the context of quantum error correction.

Another open question is that whether there exists an efficient method for emulating unitary transformations in the general case where all the given sample states are mixed. In the approach taken in this work it seems crucial that (at least) one of the sample states should be close to a pure state. Of course, one can use the given copies of a mixed state to purify the state, e.g. using the

method used in [13]. But these methods seem to have large sample complexity.

VI. ACKNOWLEDGMENTS

This work was supported by grants AFOSR No 6929347 and ARO MURI No 6924605 .

-
- [1] S. Lloyd, Science **273**, 1073 (1996).
 - [2] A. Bisio, G. Chiribella, G. M. D'Ariano, S. Facchini, and P. Perinotti, Physical Review A **81**, 032324 (2010).
 - [3] S. Lloyd, M. Mohseni, and P. Rebentrost, Nature Physics **10**, 631 (2014).
 - [4] M. Nielsen and I. Chuang, *Quantum Computation and Quantum Information*, Cambridge Series on Information and the Natural Sciences (Cambridge University Press, 2000), ISBN 9780521635035.
 - [5] A. Uhlmann, Reports on Mathematical Physics **9**, 273 (1976).
 - [6] R. Jozsa, Journal of modern optics **41**, 2315 (1994).
 - [7] W. K. Wootters and W. H. Zurek, Nature **299**, 802 (1982).
 - [8] S. D. Bartlett, T. Rudolph, and R. W. Spekkens, Reviews of Modern Physics **79**, 555 (2007).
 - [9] I. Marvian and R. W. Spekkens, New Journal of Physics **15**, 033001 (2013).
 - [10] I. Marvian and R. W. Spekkens, Phys. Rev. A **90**, 062110 (2014), URL <http://link.aps.org/doi/10.1103/PhysRevA.90.062110>.
 - [11] G. Gour and R. W. Spekkens, New Journal of Physics **10**, 033023 (2008).
 - [12] I. Marvian and R. Mann, Physical Review A **78**, 022304 (2008).
 - [13] J. Cirac, A. Ekert, and C. Macchiavello, Physical review letters **82**, 4344 (1999).

Supplementary Material

Appendix A: Fidelity of emulation for the generalized algorithm

In this section we present a generalized version of the algorithm discussed in the paper. We also prove Eq.(3.4) for this generalized algorithm.

1. Preliminaries

Let $W_a^{\text{in}}(\lambda)$ be an arbitrary unitary acting on the system and qubit a , where the parameter λ is an element of a set Λ . Let

$$W_a^{\text{out}}(\lambda) = (U \otimes I_a) W_a^{\text{in}}(\lambda) (U \otimes I_a) . \quad (\text{A1})$$

Consider states $|\phi^{\text{in}}\rangle$, and $|\phi^{\text{out}}\rangle = U|\phi^{\text{in}}\rangle$. Let $R(\phi^{\text{in}}) \equiv e^{i\pi|\phi^{\text{in}}\rangle\langle\phi^{\text{in}}|}$ be reflection about state $|\phi^{\text{in}}\rangle$, and

$$R_c(\phi^{\text{in}}) = |0\rangle\langle 0|_c \otimes I + |1\rangle\langle 1|_c \otimes e^{i\pi|\phi^{\text{in}}\rangle\langle\phi^{\text{in}}|} \quad (\text{A2})$$

be the controlled-reflection about state $|\phi^{\text{in}}\rangle$, acting on the system and qubit c .

2. The generalized algorithm

Here we list the steps of the algorithm. The quantum circuit for this algorithm is presented in Fig.(2).

- (i) Let $\lambda_1, \dots, \lambda_T$ be T random elements of set Λ , chosen independently according to the distribution $p(\lambda)$. Consider T ancillary qubits labeled by a_1, \dots, a_T , all initialized in the state $|-\rangle$. Apply the unitary $W_{a_1}^{\text{in}}(\lambda_1)$ on the system and the ancillary qubit a_1 , unitary $W_{a_2}^{\text{in}}(\lambda_2)$ on the system and a_2 , and so on, until the last unitary $W_{a_T}^{\text{in}}(\lambda_T)$, which is applied to the system and a_T .
- (ii) Apply the controlled-reflection $R_c(\phi^{\text{in}})$ on the system and qubit c , initially prepared in state $|-\rangle$. Then, apply a Hadamard gate to the qubit c and measure it in the computational basis.
- (iii) Swap a copy of state $|\phi^{\text{out}}\rangle$ with the state of system, i.e. prepare the system in state $|\phi^{\text{out}}\rangle$.
- (iv) Recall $\lambda_1, \dots, \lambda_T$ chosen in step (i) and apply the sequence of unitaries $W_{a_T}^{\text{out}\dagger}(\lambda_T), \dots, W_{a_1}^{\text{out}\dagger}(\lambda_1)$ on the system and the T ancillary qubits $a_T \dots a_1$, in the following order: First apply $W_{a_T}^{\text{out}\dagger}(\lambda_T)$ to the system and qubit a_T , then apply $W_{a_{T-1}}^{\text{out}\dagger}(\lambda_{T-1})$ to the system and a_{T-1} , and so on, until the last unitary $W_{a_1}^{\text{out}\dagger}(\lambda_1)$ which is applied on the system and qubit a_1 .

3. Fidelity of emulation (Proof of Eq.(3.4))

Recall the definition of (Uhlmann) Fidelity, between two density operator σ_1 and σ_2 ,

$$F(\sigma_1, \sigma_2) = \|\sqrt{\sigma_1} \sqrt{\sigma_2}\|_1 = \text{Tr} \left(\sqrt{\sqrt{\sigma_1} \sigma_2 \sqrt{\sigma_1}} \right) , \quad (\text{A3})$$

where $\|\cdot\|_1$ is l_1 norm, defined as the sum of the singular values of the operator [5, 6]. In the following we prove

Theorem 1 *Let \mathcal{E}_U be the quantum channel that describes the overall effect of the algorithm presented in Fig.(2) on the state of system, in the case where we do not postselect based on the outcome of measurement in step (ii) (In other words, we do not care if erasing has been successful or not). Then for any input state ρ the Uhlmann fidelity of $\mathcal{E}_U(\rho)$ and the desired state $U\rho U^\dagger$ satisfies*

$$F(\mathcal{E}_U(\rho), U\rho U^\dagger) \geq p_{\text{erase}}(\rho) = \langle \phi^{\text{in}} | \mathcal{W}^T(\rho) | \phi^{\text{in}} \rangle , \quad (\text{A4})$$

where $p_{\text{erase}}(\rho)$ is the probability that we have successfully erased the state of system (Corresponding to outcome 0 in the measurement in step(ii)), and

$$\mathcal{W}(\tau) = \sum_{\lambda \in \Lambda} p(\lambda) \text{Tr}_a(W^{\text{in}}(\lambda)[\tau \otimes |-\rangle\langle-|_a]W^{\text{in}\dagger}(\lambda)) . \quad (\text{A5})$$

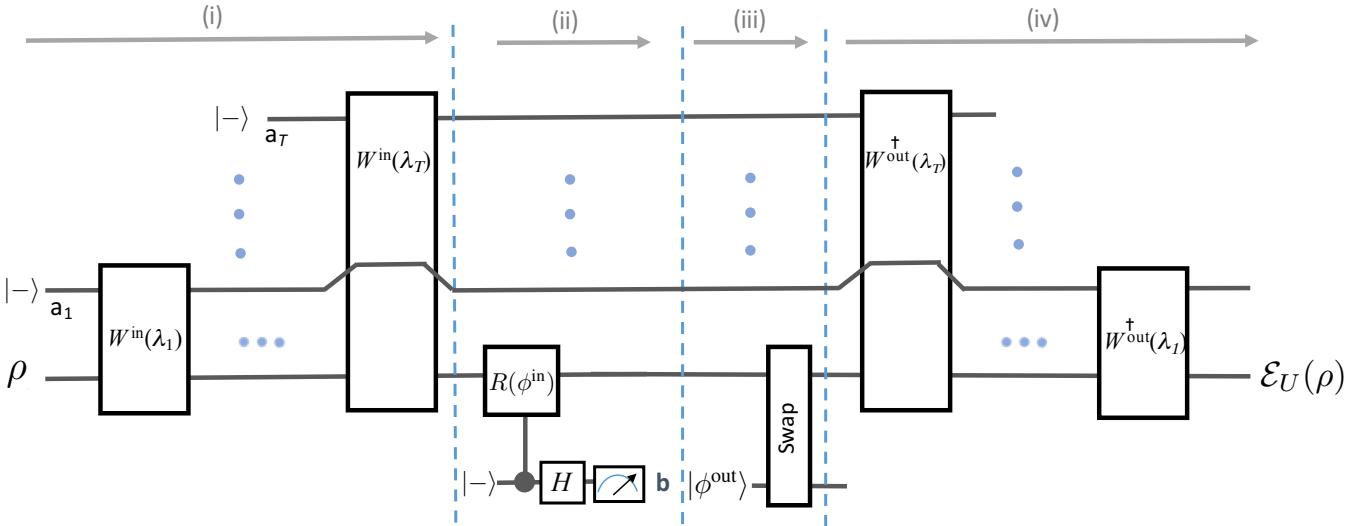


FIG. 2: The generalized quantum circuit for emulating unitary U . .

On the other hand, if we postselect to the cases where the erasing has been successful, then for pure input state ρ the fidelity of the output state $\tilde{\rho}_{post}$ with the desired state $U\rho U^\dagger$ is lower bounded by

$$F(\tilde{\rho}_{post}, U\rho U^\dagger) \geq \sqrt{p_{erase}(\rho)} = \sqrt{\langle \phi^{in} | \mathcal{W}^T(\rho) | \phi^{in} \rangle}. \quad (\text{A6})$$

Proof. We prove the theorem for the case of initial pure state $\rho = |\psi\rangle\langle\psi|$. The result for general mixed states follows from the joint concavity of Uhlmann fidelity.

Let $|\Theta(\boldsymbol{\lambda})\rangle$ be the joint state of the system and the ancillary qubits $a_1 \cdots a_T$ at the end of step (i), for a particular choice of random parameters $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_T)$. This state is given by

$$|\Theta(\boldsymbol{\lambda})\rangle = [W_{a_T}^{in}(\lambda_T) \cdots W_{a_1}^{in}(\lambda_1)] |\psi\rangle \otimes |-\rangle^{\otimes T}. \quad (\text{A7})$$

Then, if we ignore the outcome of the measurement in step (ii), at the end of step (iii) the joint state of system and ancillary qubits $a_1 \cdots a_T$, for this particular choice of random parameters $\boldsymbol{\lambda}$ is given by

$$|\phi^{out}\rangle\langle\phi^{out}| \otimes \text{Tr}_S(|\Theta(\boldsymbol{\lambda})\rangle\langle\Theta(\boldsymbol{\lambda})|), \quad (\text{A8})$$

where the partial trace is over the main system Hilbert space. Therefore, at the end of the algorithm the average joint state of the system and qubits $a_1 \cdots a_T$ is given by

$$\sigma_{fin} = \sum_{\lambda_1, \dots, \lambda_T \in \Lambda} p(\lambda_1) \cdots p(\lambda_T) [W_{a_1}^{out\dagger}(\lambda_1) \cdots W_{a_T}^{out\dagger}(\lambda_T)] [\langle \phi^{out} | \phi^{out} | \otimes \text{Tr}_S(|\Theta(\boldsymbol{\lambda})\rangle\langle\Theta(\boldsymbol{\lambda})|)] [W_{a_T}^{out}(\lambda_T) \cdots W_{a_1}^{out}(\lambda_1)], \quad (\text{A9})$$

where we have averaged over iid random parameters $(\lambda_1, \dots, \lambda_T)$, where each $\lambda_i \in \Lambda$ happens with probability $p(\lambda_i)$.

Next, we look at $F_{glob}^2 = \langle - |^{\otimes T} \langle \psi | U^\dagger \sigma_{fin} U | \psi \rangle | - \rangle^{\otimes T}$, that is the squared of the fidelity of the global output state σ_{fin} with state $U|\psi\rangle|-\rangle^{\otimes T}$. Using Eq.(A9) together with the facts that $U|\phi^{in}\rangle = |\phi^{out}\rangle$, and $(U \otimes I_a)W_a^{in}(\lambda)(U^\dagger \otimes I_a) = W_a^{out}(\lambda)$, and definition of $|\Theta(\boldsymbol{\lambda})\rangle$ in Eq.(A7), we find that

$$F_{glob}^2 = \sum_{\lambda_1, \dots, \lambda_T} p(\lambda_1) \cdots p(\lambda_T) \langle \Theta(\boldsymbol{\lambda}) | [\langle \phi^{in} | \phi^{in} | \otimes \text{Tr}_S(|\Theta(\boldsymbol{\lambda})\rangle\langle\Theta(\boldsymbol{\lambda})|)] | \Theta(\boldsymbol{\lambda}) \rangle. \quad (\text{A10})$$

Next, we note that

$$|\phi^{in}\rangle\langle\phi^{in}| \otimes \text{Tr}_S(|\Theta(\boldsymbol{\lambda})\rangle\langle\Theta(\boldsymbol{\lambda})|) = |\phi^{in}\rangle\langle\phi^{in}| \otimes \text{Tr}_S(P|\Theta(\boldsymbol{\lambda})\rangle\langle\Theta(\boldsymbol{\lambda})|P) + |\phi^{in}\rangle\langle\phi^{in}| \otimes \text{Tr}_S(P^\perp|\Theta(\boldsymbol{\lambda})\rangle\langle\Theta(\boldsymbol{\lambda})|P^\perp) \quad (\text{A11a})$$

$$= [|\phi^{in}\rangle\langle\phi^{in}| \otimes I_a] |\Theta(\boldsymbol{\lambda})\rangle\langle\Theta(\boldsymbol{\lambda})| [|\phi^{in}\rangle\langle\phi^{in}| \otimes I_a] + |\phi^{in}\rangle\langle\phi^{in}| \otimes \text{Tr}_S(P^\perp|\Theta(\boldsymbol{\lambda})\rangle\langle\Theta(\boldsymbol{\lambda})|P^\perp), \quad (\text{A11b})$$

where $P = |\phi^{\text{in}}\rangle\langle\phi^{\text{in}}|$, $P^\perp = I_S - P$, and I_S and $I_{\mathbf{a}}$ are, respectively, the identity operators on the system and the ancillary qubits $a_1 \cdots a_T$. Using the fact that $\text{Tr}_S(P^\perp|\Theta(\boldsymbol{\lambda})\rangle\langle\Theta(\boldsymbol{\lambda})|P^\perp)$ is a positive operator, we find

$$F_{\text{glob}}^2 = \sum_{\lambda_1, \dots, \lambda_T} p(\lambda_1) \cdots p(\lambda_T) \langle\Theta(\boldsymbol{\lambda})| \left[|\phi^{\text{in}}\rangle\langle\phi^{\text{in}}| \otimes \text{Tr}_S(|\Theta(\boldsymbol{\lambda})\rangle\langle\Theta(\boldsymbol{\lambda})|) \right] |\Theta(\boldsymbol{\lambda})\rangle \quad (\text{A12a})$$

$$\geq \sum_{\lambda_1, \dots, \lambda_T} p(\lambda_1) \cdots p(\lambda_T) \langle\Theta(\boldsymbol{\lambda})| \left[|\phi^{\text{in}}\rangle\langle\phi^{\text{in}}| \otimes I_{\mathbf{a}} \right] |\Theta(\boldsymbol{\lambda})\rangle \langle\Theta(\boldsymbol{\lambda})| \left[|\phi^{\text{in}}\rangle\langle\phi^{\text{in}}| \otimes I_{\mathbf{a}} \right] |\Theta(\boldsymbol{\lambda})\rangle \quad (\text{A12b})$$

$$= \sum_{\lambda_1, \dots, \lambda_T} p(\lambda_1) \cdots p(\lambda_T) \left(\langle\Theta(\boldsymbol{\lambda})| \left[|\phi^{\text{in}}\rangle\langle\phi^{\text{in}}| \otimes I_{\mathbf{a}} \right] |\Theta(\boldsymbol{\lambda})\rangle \right)^2, \quad (\text{A12c})$$

where $I_{\mathbf{a}}$ is the identity operator on the qubits $a_1 \cdots a_T$. Then, using the fact that for any distribution, the variance of a random variable is non-negative, we find that

$$F_{\text{glob}}^2 \geq \sum_{\lambda_1, \dots, \lambda_T} p(\lambda_1) \cdots p(\lambda_T) \left(\langle\Theta(\boldsymbol{\lambda})| \left[|\phi^{\text{in}}\rangle\langle\phi^{\text{in}}| \otimes I_{\mathbf{a}} \right] |\Theta(\boldsymbol{\lambda})\rangle \right)^2 \quad (\text{A13a})$$

$$\geq \left(\sum_{\lambda_1, \dots, \lambda_T} p(\lambda_1) \cdots p(\lambda_T) \langle\Theta(\boldsymbol{\lambda})| \left[|\phi^{\text{in}}\rangle\langle\phi^{\text{in}}| \otimes I_{\mathbf{a}} \right] |\Theta(\boldsymbol{\lambda})\rangle \right)^2 \quad (\text{A13b})$$

$$= \text{Tr}^2 \left(|\phi^{\text{in}}\rangle\langle\phi^{\text{in}}| \sum_{\lambda_1, \dots, \lambda_T} p(\lambda_1) \cdots p(\lambda_T) \text{Tr}_{\mathbf{a}}(|\Theta(\boldsymbol{\lambda})\rangle\langle\Theta(\boldsymbol{\lambda})|) \right) \quad (\text{A13c})$$

$$= p_{\text{erase}}^2(|\psi\rangle\langle\psi|), \quad (\text{A13d})$$

where $p_{\text{erase}}(|\psi\rangle\langle\psi|)$ is the probability that at the end of step (i) the reduced state of system is in state $|\phi^{\text{in}}\rangle$.

Finally, we note that the final state of system is obtained from state σ_{fin} , by tracing over the ancillary qubits, i.e. $\mathcal{E}_U(|\psi\rangle\langle\psi|) = \text{Tr}_{\mathbf{a}}(\sigma_{\text{glob}})$. Then, using the monotonicity of Uhlmann fidelity under partial trace we find

$$F(\mathcal{E}_U(|\psi\rangle\langle\psi|), U|\psi\rangle\langle\psi|U^\dagger) \geq F_{\text{glob}} \geq p_{\text{erase}}(|\psi\rangle\langle\psi|). \quad (\text{A14})$$

Using the joint concavity of Uhlmann fidelity this proves the bound for arbitrary initial state ρ .

Next, using the definition $|\Theta(\boldsymbol{\lambda})\rangle = \left[W_{a_T}^{\text{in}}(\lambda_T) \cdots W_{a_1}^{\text{in}}(\lambda_1) \right] |\psi\rangle \otimes |-\rangle^{\otimes T}$, it can be easily shown that

$$\sum_{\lambda_1, \dots, \lambda_T} p(\lambda_1) \cdots p(\lambda_T) \text{Tr}_{\mathbf{a}}(|\Theta(\boldsymbol{\lambda})\rangle\langle\Theta(\boldsymbol{\lambda})|) = \mathcal{W}^T(|\psi\rangle\langle\psi|), \quad (\text{A15a})$$

where

$$\mathcal{W}(\tau) = \sum_{\lambda \in \Lambda} p(\lambda) \text{Tr}_a \left(W_a^{\text{in}}(\lambda) \left[\tau \otimes |-\rangle\langle-|_a \right] W_a^{\text{in}}(\lambda)^\dagger \right), \quad (\text{A16})$$

and the partial trace is over qubit a . Therefore, we conclude that for any initial state ρ it holds that $F(\mathcal{E}_U(\rho), U\rho U^\dagger) \geq p_{\text{erase}}(\rho) = \langle\phi^{\text{in}}|\mathcal{W}^T(\rho)|\phi^{\text{in}}\rangle$.

In the above argument, we assumed we ignore the outcome of measurement in step (ii), which means that we do not care if the erasing has been successful or not. Now suppose we postselect to the cases where the erasing has been successful. In this case after the measurement the joint state of the system and ancillary qubits $a_1 \cdots a_T$ is given by

$$\frac{1}{p_{\text{erase}}(|\psi\rangle\langle\psi|)} \times \sum_{\lambda_1, \dots, \lambda_T} p(\lambda_1) \cdots p(\lambda_T) \left[|\phi^{\text{in}}\rangle\langle\phi^{\text{in}}| \otimes I_{\mathbf{a}} \right] |\Theta(\boldsymbol{\lambda})\rangle \langle\Theta(\boldsymbol{\lambda})| \left[|\phi^{\text{in}}\rangle\langle\phi^{\text{in}}| \otimes I_{\mathbf{a}} \right], \quad (\text{A17})$$

where the factor $1/p_{\text{erase}}(|\psi\rangle\langle\psi|)$ is due to the postselection. Then, using an argument similar to the one we used before, we can show that the squared of the fidelity of the final joint state of the system and qubits $a_1 \cdots a_T$ at the end of the algorithm, with state $U|\psi\rangle \otimes |-\rangle^{\otimes T}$ is given by

$$F_{\text{glob}}^{\text{post}2} = \frac{1}{p_{\text{erase}}(|\psi\rangle\langle\psi|)} \times \sum_{\lambda_1, \dots, \lambda_T} p(\lambda_1) \cdots p(\lambda_T) \langle\Theta(\boldsymbol{\lambda})| \left[|\phi^{\text{in}}\rangle\langle\phi^{\text{in}}| \otimes I_{\mathbf{a}} \right] |\Theta(\boldsymbol{\lambda})\rangle \langle\Theta(\boldsymbol{\lambda})| \left[|\phi^{\text{in}}\rangle\langle\phi^{\text{in}}| \otimes I_{\mathbf{a}} \right] |\Theta(\boldsymbol{\lambda})\rangle \quad (\text{A18a})$$

$$= \frac{1}{p_{\text{erase}}(|\psi\rangle\langle\psi|)} \times \sum_{\lambda_1, \dots, \lambda_T} p(\lambda_1) \cdots p(\lambda_T) \left(\langle\Theta(\boldsymbol{\lambda})| \left[|\phi^{\text{in}}\rangle\langle\phi^{\text{in}}| \otimes I_{\mathbf{a}} \right] |\Theta(\boldsymbol{\lambda})\rangle \right)^2, \quad (\text{A18b})$$

Then, using Eq.(A13), we find

$$F_{\text{glob}}^{\text{post}} \geq \frac{p_{\text{erase}}^2(|\psi\rangle\langle\psi|)}{p_{\text{erase}}(|\psi\rangle\langle\psi|)} = p_{\text{erase}}(|\psi\rangle\langle\psi|). \quad (\text{A19})$$

This means that if we postselect to the cases where erasing has been successful, then

$$F(U|\psi\rangle\langle\psi|U^\dagger, \tilde{\rho}_{\text{post}}) \geq F_{\text{glob}}^{\text{post}} \geq \sqrt{p_{\text{erase}}(|\psi\rangle\langle\psi|)}. \quad (\text{A20})$$

■

Appendix B: Runtime and error analysis

In this section we first study the runtime and the error in the output of the algorithm, assuming we run the quantum circuit in Fig. 1 with ideal controlled-reflections $R_a(k)$. Then, we consider the runtime and error in the actual algorithm, where we use the given copies of sample states to simulate the controlled-reflections.

1. Algorithm with perfect controlled-reflections

Let \mathcal{E}_U be the quantum channel that describes the overall effect of the circuit in Fig.(1) on the system, in the case where we ignore the outcome of measurement in step (ii), and where we use ideal controlled-reflections in the circuit. Let $\|\cdot\|_1$ be the l_1 -norm, defined as the sum of the singular values of the operator. Then,

Theorem 2 Suppose in the quantum circuit presented in Fig.(1) we choose

$$T \geq \frac{d \times \log(8\epsilon_{id}^{-2}d)}{1 - |\lambda_D|}, \quad (\text{B1})$$

where λ_D is the eigenvalue of channel $\mathcal{D}(\cdot) = K^{-1} \sum_{k=1}^K e^{i\pi|\phi_k^{in}\rangle\langle\phi_k^{in}|}(\cdot) e^{i\pi|\phi_k^{in}\rangle\langle\phi_k^{in}|}$ which has the second largest magnitude. Then, for any input ρ whose support is restricted to the input subspace \mathcal{H}_{in} , the output of the circuit satisfies $\|U\rho U^\dagger - \mathcal{E}_U(\rho)\|_1 \leq \epsilon_{id}$.

Proof. We start with Eq.(3.4) proven in Appendix A,

$$F(\mathcal{E}_U(\rho), U\rho U^\dagger) \geq \langle \phi_1^{in} | \mathcal{W}^T(\rho) | \phi_1^{in} \rangle. \quad (\text{B2})$$

Let $P_{in}^\perp = \Pi_{in} - P = \Pi_{in} - |\phi_1^{in}\rangle\langle\phi_1^{in}|$ be the projector to the $(d-1)$ -dimensional subspace of \mathcal{H}_{in} orthogonal to $|\phi_1^{in}\rangle$, and $\mathcal{P}_{in}^\perp(\cdot) = P_{in}^\perp(\cdot)P_{in}^\perp$. Using the definition $\mathcal{W}(\sigma) = P\sigma P + \mathcal{D}(P^\perp\sigma P^\perp)$, it can be easily seen that for any state σ whose support is restricted to \mathcal{H}_{in} , it holds that

$$\mathcal{P}_{in}^\perp \circ \mathcal{W}(\sigma) = \mathcal{P}_{in}^\perp \circ \mathcal{W} \circ \mathcal{P}_{in}^\perp(\sigma) = \mathcal{P}_{in}^\perp \circ \mathcal{D} \circ \mathcal{P}_{in}^\perp(\sigma). \quad (\text{B3})$$

Furthermore, using the fact that under channels \mathcal{W} and \mathcal{D} any state ρ whose support is restricted to \mathcal{H}_{in} is mapped to a state with support restricted to this subspace, we find that

$$\mathcal{P}_{in}^\perp \circ \mathcal{W}^T(\rho) = [\mathcal{P}_{in}^\perp \circ \mathcal{D} \circ \mathcal{P}_{in}^\perp]^T(\rho). \quad (\text{B4})$$

This implies

$$F(\mathcal{E}_U(\rho), U\rho U^\dagger) \geq \langle \phi_1^{in} | \mathcal{W}^T(\rho) | \phi_1^{in} \rangle \quad (\text{B5})$$

$$= 1 - \text{Tr}(P^\perp \mathcal{W}^T(\rho)) \quad (\text{B6})$$

$$= 1 - \text{Tr}(\mathcal{P}_{in}^\perp \mathcal{W}^T(\rho)) \quad (\text{B7})$$

$$= 1 - \text{Tr}(\mathcal{P}_{in}^\perp \circ \mathcal{W}^T(\rho)) \quad (\text{B8})$$

$$= 1 - \text{Tr}([\mathcal{P}_{in}^\perp \circ \mathcal{D} \circ \mathcal{P}_{in}^\perp]^T(\rho)), \quad (\text{B9})$$

where to get the third line we have used the fact that the support of $\mathcal{W}^T(\rho)$ is contained in \mathcal{H}_{in} , and to get the last line we have used Eq.(B3).

It turns out that quantum operations \mathcal{D} and $\mathcal{P}_{in}^\perp \circ \mathcal{D} \circ \mathcal{P}_{in}^\perp$ have several nice properties which simplify the following analysis. In particular, they both have Hermitian matrix representations: Let $\{F_\mu : \mu = 1 \cdots d^2\}$ be an orthonormal basis for the operator space acting on \mathcal{H}_{in} , such that $\text{Tr}(F_\mu^\dagger F_\nu) = \delta_{\mu,\nu}$. Consider the matrix representation of \mathcal{D} and $\mathcal{P}_{in}^\perp \circ \mathcal{D} \circ \mathcal{P}_{in}^\perp$ on \mathcal{H}_{in} , i.e. the matrices $D_{\mu,\nu} = \text{Tr}(F_\mu^\dagger \mathcal{D}(F_\nu))$ and $D_{\mu,\nu}^\perp = \text{Tr}(F_\mu^\dagger \mathcal{P}_{in}^\perp \circ \mathcal{D} \circ \mathcal{P}_{in}^\perp(F_\nu))$, respectively. Then, the fact that any reflection $e^{i\pi|\phi\rangle\langle\phi|}$ is a Hermitian operator, implies that these matrices are both Hermitian (Note that the matrix representation of channel \mathcal{W} is not Hermitian). Let λ_\perp be the largest eigenvalue of $\mathcal{P}_{in}^\perp \circ \mathcal{D} \circ \mathcal{P}_{in}^\perp$. It follows that for any T , the map $[\mathcal{P}_{in}^\perp \circ \mathcal{D} \circ \mathcal{P}_{in}^\perp]^T$ has also a Hermitian matrix representation, and its largest eigenvalue

is λ_{\perp}^T . This implies

$$\mathrm{Tr}\left(\left[\mathcal{P}_{\text{in}}^{\perp} \circ \mathcal{D} \circ \mathcal{P}_{\text{in}}^{\perp}\right]^T(\rho)\right) = \mathrm{Tr}\left(P_{\text{in}}^{\perp} \left[\mathcal{P}_{\text{in}}^{\perp} \circ \mathcal{D} \circ \mathcal{P}_{\text{in}}^{\perp}\right]^T(\rho)\right) \quad (\text{B10a})$$

$$\leq \mathrm{Tr}\left(P_{\text{in}}^{\perp} \left[\mathcal{P}_{\text{in}}^{\perp} \circ \mathcal{D} \circ \mathcal{P}_{\text{in}}^{\perp}\right]^T(P_{\text{in}}^{\perp})\right) \quad (\text{B10b})$$

$$\leq |\lambda_{\perp}|^T \times \mathrm{Tr}(P_{\text{in}}^{\perp}) \quad (\text{B10c})$$

$$= |\lambda_{\perp}|^T \times (d - 1), \quad (\text{B10d})$$

where to get the first line we use the fact that for any X , $P_{\text{in}}\mathcal{P}_{\text{in}}(X) = \mathcal{P}_{\text{in}}(X)$, and to get the third line we use the fact that λ_{\perp}^T is the largest eigenvalue of $\mathcal{P}_{\text{in}}^{\perp} \circ \mathcal{D} \circ \mathcal{P}_{\text{in}}^{\perp}$, which implies for any X , $\mathrm{Tr}\left(X^{\dagger} \left[\mathcal{P}_{\text{in}}^{\perp} \circ \mathcal{D} \circ \mathcal{P}_{\text{in}}^{\perp}\right]^T(X)\right) \leq \mathrm{Tr}(X^{\dagger}X)|\lambda_{\perp}|^T$.

Putting this into Eq.(B5) we find that

$$F(\mathcal{E}_U(\rho), U\rho U^{\dagger}) \geq 1 - \mathrm{Tr}\left(\left[\mathcal{P}^{\perp} \circ \mathcal{D} \circ \mathcal{P}^{\perp}\right]^T(\rho)\right) \geq 1 - d \times |\lambda_{\perp}|^T. \quad (\text{B11})$$

Next, we use this bound together with the Fuchs-van de Graaf inequality, to bound the trace distance of $\mathcal{E}_U(\rho)$ and $U\rho U^{\dagger}$. We find

$$\|\mathcal{E}_U(\rho) - U\rho U^{\dagger}\|_1 \leq 2\sqrt{1 - F^2(\mathcal{E}_U(\rho), U\rho U^{\dagger})} \leq 2\sqrt{2}\sqrt{1 - F(\mathcal{E}_U(\rho), U\rho U^{\dagger})} \quad (\text{B12})$$

$$\leq 2\sqrt{2d \times |\lambda_{\perp}|^T}. \quad (\text{B13})$$

This means that to achieve error ϵ_{id} in trace distance, it is sufficient to have

$$T \geq \frac{\log(8\epsilon_{\text{id}}^{-2}d)}{\log|\lambda_{\perp}|^{-1}}. \quad (\text{B14})$$

In the following we prove

$$|\lambda_{\perp}| \leq 1 - \frac{1 - |\lambda_{\mathcal{D}}|}{d}. \quad (\text{B15})$$

This together with the fact that $\log(1 + x) \leq x$ for $x > -1$, implies

$$\frac{d}{1 - |\lambda_{\mathcal{D}}|} \geq \frac{1}{\log|\lambda_{\perp}|^{-1}}. \quad (\text{B16})$$

Therefore, using Eq.(B14), we conclude that if we choose T such that

$$T \geq \frac{d \times \log(8\epsilon_{\text{id}}^{-2}d)}{1 - |\lambda_{\mathcal{D}}|}, \quad (\text{B17})$$

then $\|\mathcal{E}_U(\rho) - U\rho U^{\dagger}\|_1 \leq \epsilon_{\text{id}}$.

Bound on the maximum eigenvalue of $\mathcal{P}_{\text{in}}^{\perp} \circ \mathcal{D} \circ \mathcal{P}_{\text{in}}^{\perp}$

To complete the proof, in the following we prove Eq.(B15), which is a bound on λ_{\perp} , the largest eigenvalue of $\mathcal{P}_{\text{in}}^{\perp} \circ \mathcal{D} \circ \mathcal{P}_{\text{in}}^{\perp}$. Recall that $\mathcal{P}_{\text{in}}^{\perp} \circ \mathcal{D} \circ \mathcal{P}_{\text{in}}^{\perp}$ and \mathcal{D} both have Hermitian matrix representations. This means that \mathcal{D} has a spectral decomposition as $\mathcal{D}(X) = \sum_{\mu} \lambda_{\mu} M_{\mu}^{\dagger} \mathrm{Tr}(M_{\mu} X)$, where each λ_{μ} and M_{μ}^{\dagger} are, respectively, the eigenvalue of \mathcal{D} and its corresponding eigenvector, and $\mathrm{Tr}(M_{\mu}^{\dagger} M_{\nu}) = \delta_{\mu,\nu}$. Then, (the absolute value of) the maximum eigenvalue of $\mathcal{P}_{\text{in}}^{\perp} \circ \mathcal{D} \circ \mathcal{P}_{\text{in}}^{\perp}$ is given by

$$|\lambda_{\perp}| = \max_X \frac{|\mathrm{Tr}(X^{\dagger} \mathcal{P}_{\text{in}}^{\perp} \circ \mathcal{D} \circ \mathcal{P}_{\text{in}}^{\perp}(X))|}{\mathrm{Tr}(X^{\dagger} X)} = \max_X \left| \sum_{\mu} \lambda_{\mu} \frac{|\mathrm{Tr}(M_{\mu} P_{\text{in}}^{\perp} X P_{\text{in}}^{\perp})|^2}{\mathrm{Tr}(X^{\dagger} X)} \right|. \quad (\text{B18})$$

Note that the Hermiticity of \mathcal{D} implies that its eigenvectors form an orthonormal basis, and so

$$\sum_{\mu} \left| \text{Tr}(M_{\mu} P_{\text{in}}^{\perp} X P_{\text{in}}^{\perp}) \right|^2 = \text{Tr}^2(P_{\text{in}}^{\perp} X^{\dagger} P_{\text{in}}^{\perp} X P_{\text{in}}^{\perp}). \quad (\text{B19})$$

Then, using the fact that the operator Π_{in}/\sqrt{d} is the only (normalized) eigenvector of \mathcal{D} with eigenvalue 1, we find

$$|\lambda_{\perp}| = \max_X \left| \sum_{\mu} \lambda_{\mu} \frac{\left| \text{Tr}(M_{\mu} P_{\text{in}}^{\perp} X P_{\text{in}}^{\perp}) \right|^2}{\text{Tr}(X^{\dagger} X)} \right| \leq \max_X \left[\frac{\left| \text{Tr}(P_{\text{in}}^{\perp} X P_{\text{in}}^{\perp}) \right|^2}{d \times \text{Tr}(X^{\dagger} X)} + |\lambda_{\mathcal{D}}| \frac{\text{Tr}(P_{\text{in}}^{\perp} X^{\dagger} P_{\text{in}}^{\perp} X P_{\text{in}}^{\perp}) - \left| \text{Tr}(P_{\text{in}}^{\perp} X P_{\text{in}}^{\perp}) \right|^2 \times d^{-1}}{\text{Tr}(X^{\dagger} X)} \right], \quad (\text{B20})$$

where $\lambda_{\mathcal{D}}$ is the second largest eigenvalue of \mathcal{D} (in magnitude). This implies

$$|\lambda_{\perp}| \leq \max_X \left[\frac{\left| \text{Tr}(X P_{\text{in}}^{\perp}) \right|^2}{d \times \text{Tr}(X^{\dagger} X)} \times (1 - |\lambda_{\mathcal{D}}|) + |\lambda_{\mathcal{D}}| \right]. \quad (\text{B21})$$

Then, using the Cauchy-Schwarz inequality we find

$$|\lambda_{\perp}| \leq \max_X \left[\frac{\text{Tr}(X X^{\dagger}) \text{Tr}(P_{\text{in}}^{\perp})}{d \times \text{Tr}(X^{\dagger} X)} \times (1 - |\lambda_{\mathcal{D}}|) + |\lambda_{\mathcal{D}}| \right] = 1 - \frac{1 - |\lambda_{\mathcal{D}}|}{d}, \quad (\text{B22})$$

where we have used the fact that $\text{Tr}(P_{\text{in}}^{\perp}) = d - 1$. This completes the proof of Eq.(B15) and theorem 2. ■

2. Total error in the actual algorithm with simulated controlled-reflections

As we showed in the previous section, in the idealized case where we are given perfect controlled-reflections by choosing

$$T \geq T(d, \epsilon_{\text{id}}) \equiv \frac{d \log(8\epsilon_{\text{id}}^{-2}d)}{1 - |\lambda_{\mathcal{D}}|}, \quad (\text{B23})$$

we can implement the transformation $|\psi\rangle \rightarrow U|\psi\rangle$ on any initial state $|\psi\rangle \in \mathcal{H}_{\text{in}}$ with error less than or equal to ϵ_{id} in trace distance. But, in the main algorithm we need to simulate the controlled-reflections using the given copies of the sample states.

In the Appendix C we show that using n copies of state $|\phi\rangle$ we can implement the unitary $e^{i\pi|\phi\rangle\langle\phi|}$, or its controlled version, $|0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes e^{i\pi|\phi\rangle\langle\phi|}$, with error $\mathcal{O}(1/n)$ in trace distance, and in time $\mathcal{O}(\log D \times n \log n) = \tilde{\mathcal{O}}(\log D \times n)$, where $\tilde{\mathcal{O}}$ suppresses more slowly-growing terms. In other words, to achieve the error of order ϵ_{ref} in simulating each controlled-reflection we need time of order $\tilde{\mathcal{O}}(\epsilon_{\text{ref}}^{-1} \times \log D)$, and $\tilde{\mathcal{O}}(\epsilon_{\text{ref}}^{-1})$ copies of the corresponding state.

Given that the circuit requires $4T + 1$ controlled-reflection, it follows that we can implement the transformation $|\psi\rangle \rightarrow U|\psi\rangle$, with the total error less than or equal to

$$\epsilon_{\text{tot}} = [4T(d, \epsilon_{\text{id}}) + 1] \times \epsilon_{\text{ref}} + \epsilon_{\text{id}}, \quad (\text{B24})$$

in the total time

$$t_{\text{tot}} = [4T(d, \epsilon_{\text{id}}) + 1] \times \tilde{\mathcal{O}}(\epsilon_{\text{ref}}^{-1} \times \log D), \quad (\text{B25})$$

and using the total number of input-output sample pairs

$$N_{\text{tot}} = [4T(d, \epsilon_{\text{id}}) + 1] \times \epsilon_{\text{ref}}^{-1}. \quad (\text{B26})$$

Choosing ϵ_{id} and ϵ_{ref} such that $T(d, \epsilon_{\text{id}}) \times \epsilon_{\text{ref}} = \epsilon_{\text{id}} = \epsilon$, we find that the transformation $|\psi\rangle \rightarrow U|\psi\rangle$ can be implemented with the total error less than or equal to ϵ , in the total time

$$t_{\text{tot}} = \tilde{\mathcal{O}} \left(\frac{d^2 \epsilon^{-1} \log D}{(1 - |\lambda_{\mathcal{D}}|)^2} \right), \quad (\text{B27})$$

and using the total number of input-output sample pairs

$$N_{\text{tot}} = \tilde{\mathcal{O}} \left(\frac{d^2 \epsilon^{-1}}{(1 - |\lambda_{\mathcal{D}}|)^2} \right). \quad (\text{B28})$$

Appendix C: An efficient quantum circuit for exponentiating a density operator

In this section, we introduce an explicit simple quantum circuit that uses multiple copies of state σ , to simulate the unitary $e^{-it\sigma}$, or its controlled version $|0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes e^{-it\sigma}$, for any real t . The circuit works based on the density matrix exponentiation technique of [3].

Let S be the SWAP operator acting on the system and an ancillary system with equal Hilbert space dimensions, such that $S|\mu\rangle|\nu\rangle = |\nu\rangle|\mu\rangle$, for any pair of states $|\mu\rangle$, and $|\nu\rangle$. Note that S is a Hermitian unitary operator. Then, as it is observed in Ref.([3]),

$$\text{Tr}_{\text{RF}}(e^{-i\Delta t S}[\rho \otimes \sigma]e^{i\Delta t S}) = \rho + i\Delta t[\rho, \sigma] + \mathcal{O}(\Delta t^2) = e^{-i\sigma\Delta t}\rho e^{i\sigma\Delta t} + \mathcal{O}(\Delta t^2), \quad (\text{C1})$$

where the partial trace is over the system with state σ (which can be interpreted as a quantum reference frame [8, 12]). Then, choosing sufficiently small Δt , and repeating the above procedure for $t/\Delta t$ times with $t/\Delta t$ copies of state σ , we can simulate the unitary $e^{-i\sigma t}$ with arbitrary accuracy.

In the following section we present a new simple and efficient algorithm for simulating the unitary $e^{-i\theta S}$ for arbitrary θ , with error ϵ and in time $\mathcal{O}(\log(D) \times \log(1/\epsilon))$. Therefore, if we use this algorithm together with the above method, using one copy of state σ we can simulate the unitary $e^{-i\sigma\Delta t}$ with the total error of order $\Delta t^2 + \epsilon$, and in time $\mathcal{O}(\log(D) \times \log(1/\epsilon))$. Then, having n copies of state σ , to simulate the unitary $e^{-i\sigma t}$ we can choose $\Delta t = t/n$ and use each copy to simulate $e^{-i\sigma\Delta t}$. In this case the total error is

$$\epsilon_{\text{tot}} = \mathcal{O}(n \times (\epsilon + \Delta t^2)) = \mathcal{O}(n\epsilon + \frac{t^2}{n}), \quad (\text{C2})$$

and the total runtime is

$$t_{\text{tot}} = \mathcal{O}(\log(D) \times \log(1/\epsilon) \times n), \quad (\text{C3})$$

where ϵ is the error in implementing the unitary $e^{-i\Delta t S}$. Choosing $\epsilon = \frac{1}{n^2}$, we find that using n copies of state σ we can simulate unitary $e^{-i\sigma t}$, with the total error of order $\epsilon_{\text{tot}} = \mathcal{O}(\frac{t^2+1}{n})$ in trace distance, and in time of order $\mathcal{O}(\log(D) \times n \log(n))$.

Furthermore, using Eq.(C1), it can be easily shown that if instead of the unitary $e^{-i\Delta t S}$ we implement its controlled version, i.e. the unitary $|0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes e^{-iS\Delta t}$, then we can simulate the controlled version of $e^{-i\sigma\Delta t}$, i.e. the unitary $|0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes e^{-i\sigma\Delta t}$. To implement the unitary $|0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes e^{-iS\Delta t}$, we note that

$$e^{-iS_c\Delta t} = |0\rangle\langle 0| \otimes e^{-i\Delta t}I + |1\rangle\langle 1| \otimes e^{-iS\Delta t}, \quad (\text{C4})$$

where $S_c = |0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes S$, is the controlled-SWAP unitary. Using the algorithm presented in the next section, we can efficiently simulate this unitary with error ϵ in time $\mathcal{O}(\log(D) \times \log(1/\epsilon))$. Then, if after applying $e^{-iS_c\Delta t}$ we apply $e^{i\sigma_z\Delta t/2}$ on the control qubit, we implement the desired unitary $|0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes e^{-iS\Delta t}$, up to a global phase.

Therefore, having n copies of state σ to simulate the unitary $|0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes e^{-i\sigma t}$, we first apply the unitary $e^{-iS_c t/n}$ to the system, the control qubit and each one of these n copies, and then apply the the unitary $e^{i\sigma_z t/2}$ on the control qubit. It follows that we can implement the unitary $|0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes e^{-i\sigma t}$, with the total error less than or equal to $\mathcal{O}(\frac{t^2+1}{n})$ in time $\mathcal{O}(\log(D) \times n \log(n))$.

We conclude that

Theorem 3 *Using n copies of an unknown state σ we can simulate the unitary $e^{-i\sigma t}$, or its controlled version $|0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes e^{-i\sigma\Delta t}$ in time $\mathcal{O}(\log(D) \times n \log(n))$, with error $\mathcal{O}(\frac{t^2+1}{n})$ in trace distance.*

1. An efficient quantum circuit for simulating the time evolution generated by the (controlled-)SWAP Hamiltonian

In this section we present a simple quantum circuit to simulate the time evolution generated by the SWAP Hamiltonian on a pair of systems whose Hilbert spaces have equal dimension. This algorithm simulates any such unitary with error less than or equal to ϵ (in trace distance) in time $\mathcal{O}(\log(D) \times \log(1/\epsilon))$, where D is the dimension of the Hilbert space of each system. We also explain how the same idea can be used to simulate the controlled-SWAP Hamiltonian as well.

Let S be the SWAP operator acting on a pair of systems A and B . The fact that S^2 is the identity operator implies that any unitary generated by S has the following form

$$e^{iS\theta} = \cos \theta I + i \sin \theta S , \quad (\text{C5})$$

where I is the identity operator on the joint system. To simulate a general unitary in this form we first implement S_c , the controlled-SWAP unitary defined by

$$S_c = |0\rangle\langle 0|_c \otimes I + |1\rangle\langle 1|_c \otimes S , \quad (\text{C6})$$

where c is the label for an ancillary qubit. This unitary can be efficiently implemented, using controlled-SWAP gates for qubits, also known as *Fredkin* gate [4]. To implement S_c in this way, we need to apply $\mathcal{O}(\log(D))$ Fredkin gates, where each Fredkin gate is controlled by the ancillary qubit c , and is acting on a qubit in systems A , and its corresponding qubit in system B .

Then, using the controlled-SWAP operator S_c , we can easily implement the unitary $e^{iS\theta}$, for any θ : First, we prepare the ancillary qubit in state $\cos \theta |0\rangle + i \sin \theta |1\rangle$ and apply the controlled-SWAP S_c to the ancillary qubit and the systems A and B in an arbitrary initial state $|\gamma\rangle$. This yields

$$S_c(\cos \theta |0\rangle + i \sin \theta |1\rangle) \otimes |\gamma\rangle = \cos \theta |0\rangle |\gamma\rangle + i \sin \theta |1\rangle S |\gamma\rangle . \quad (\text{C7})$$

Next, we measure the ancillary qubit in $\{|0\rangle \pm |1\rangle\}/\sqrt{2}$ basis. Then, with probability 1/2 we project this qubit to state $(|0\rangle + |1\rangle)/\sqrt{2}$. In this case we have implemented the desired transformation on the systems, i.e. we get state

$$\frac{|0\rangle + |1\rangle}{\sqrt{2}} (\cos \theta |\gamma\rangle + i \sin \theta S |\gamma\rangle) = \frac{|0\rangle + |1\rangle}{\sqrt{2}} e^{i\theta S} |\gamma\rangle . \quad (\text{C8})$$

On the other hand, with probability 1/2 the state of control qubit is projected to $(|0\rangle - |1\rangle)/\sqrt{2}$, in which case we get state

$$\frac{|0\rangle - |1\rangle}{\sqrt{2}} (\cos \theta |\gamma\rangle - i \sin \theta S |\gamma\rangle) = \frac{|0\rangle - |1\rangle}{\sqrt{2}} \otimes e^{-i\theta S} |\gamma\rangle . \quad (\text{C9})$$

Therefore, in the latter case instead of the desired unitary $e^{i\theta S}$ we have implemented the unitary $e^{-i\theta S}$. But because these unitaries all commute with each other, we can correct this error: we repeat the above process and this time we try to implement the unitary $e^{i2\theta S}$, by preparing the ancilla in state $\cos 2\theta |0\rangle + i \sin 2\theta |1\rangle$, instead of $\cos \theta |0\rangle + i \sin \theta |1\rangle$. Then, again with probability 1/2 we will be successful and implement the overall unitary $e^{i2\theta S} e^{-i\theta S} = e^{i\theta S}$, as we desired. On the other hand, with probability 1/2 we are unsuccessful in which case we have implemented the overall unitary $e^{-iS\theta} e^{-iS2\theta} = e^{-iS3\theta}$. In this case we repeat the above process and try to implement the unitary $e^{iS4\theta}$, by preparing the ancillary qubit in state $\cos 4\theta |0\rangle + i \sin 4\theta |1\rangle$.

Repeating this process for r times we achieve the desired unitary with probability $1 - 1/2^r$. Since each time we need to apply S_c , which takes time of order $\log D$, we find that the total time it takes to implement $e^{iS\theta}$ with error ϵ is $\mathcal{O}(\log(D) \times \log(1/\epsilon))$.

Finally, note that we can use a similar method to implement the unitary

$$e^{i\theta S_c} = |0\rangle\langle 0| \otimes e^{i\theta} I + |1\rangle\langle 1| \otimes e^{iS\theta} = \cos \theta I + i \sin \theta S_c . \quad (\text{C10})$$

To implement this unitary, it suffices to replace the Fredkin gates in the above algorithm, with controlled-Fredkin gates, i.e. controlled-controlled-SWAP gates each of which acts on a pair of qubits in systems A and B , and is controlled by the original control qubit and the ancillary qubit that we need to implement the above procedure.

Appendix D: Generalizations

In this section we discuss several generalizations of the results presented in the paper. These generalizations are summarized in Sec.(IV) of the paper.

1. Input-output pairs of mixed states

The algorithm presented in the paper assumes the given samples of the input-output pairs are all pure states. However, as we explained at the end of Sec.(IIIA) this algorithm can be generalized to the case where the sample states contain mixed states. It can be easily shown that, as long as the sample input states $S_{\text{in}} = \{\rho_k^{\text{in}}, k = 1, \dots, K\}$ contain (at least) one state close to a pure state, we can still coherently erase the state of system and push it into this pure state. Then, we can emulate the action of the unknown unitary U , using the same approach we used in the original algorithm.

Let \mathcal{H}_{in} be the subspace spanned by the union of the support of density operators in $S_{\text{in}} = \{\rho_k^{\text{in}}, k = 1, \dots, K\}$. Then, having copies of sample input-output states in S_{in} and S_{out} , we can determine the action of U on any state in \mathcal{H}_{in} , if and only if the set S_{in} generates the full matrix algebra on \mathcal{H}_{in} . Therefore, in the following we naturally assume this condition is satisfied.

To erase the state of system coherently we use controlled-translations with respect to the given sample states, i.e. the unitaries

$$T_a(k, t) = |0\rangle\langle 0|_a \otimes I + |1\rangle\langle 1|_a \otimes e^{-i\rho_k t}, \quad (\text{D1})$$

where we have suppressed *in* and *out* superscript in both sides. As we discussed in Appendix C, using the given copies of sample states we can efficiently simulate these unitaries. Recall that the input set S_{in} (and hence S_{out}) contains at least one pure state. Without loss of generality, let ρ_1^{in} and $\rho_1^{\text{out}} = U\rho_1^{\text{in}}U^\dagger$ be the pure state in the sample set S_{in} and its corresponding output state in the set S_{out} , respectively. Then instead of unitaries $W_a(k) = R_a(k)H_aR_a(1)$ used in the original algorithm, we use unitaries $W_a(k, t) = T_a(k, t)H_aT_a(1, \pi)$, where we have suppressed *in* and *out* superscripts, and choose k and t uniformly at random from the sets $1, \dots, K$ and $[0, 1]$, respectively. Then, it can be easily shown that state $\rho_{\text{in}}^1 = |\phi_1^{\text{in}}\rangle\langle\phi_1^{\text{in}}|$ is the unique fixed point state of channel $\mathcal{W} = 1/K \sum_{k=1}^K \int_0^1 dt \text{Tr}_a(W_a^{\text{in}}(k, t)[\tau \otimes |-\rangle\langle -|_a] W_a^{\text{in}\dagger}(k, t))$ inside \mathcal{H}_{in} . It follows that we can coherently erase the state of system, and therefore, using the same technique we used in the main algorithm, we can emulate the action of unitary U .

2. Emulating controlled unitaries

In many quantum algorithms, such as quantum phase estimation, one needs to implement the controlled version of a unitary U , i.e. the unitary $U^c \equiv |0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes U$. Can we modify the proposed algorithm to implement the controlled version of U as well?

To answer this question, first note that if the only given resources are multiple copies of states in S_{in} and S_{out} , it is impossible to distinguish between unitaries U and $e^{i\theta}U$, for any phase $e^{i\theta}$. On the other hand, in general the controlled version of U and $e^{i\theta}U$ are distinct unitaries, which are not equivalent up to a global phase. This means that even to specify what is the controlled version of U we need extra resources that define and fix this global phase. For instance, we can use multiple copies of the input state $|\Phi^{\text{in}}\rangle = (\alpha|0\rangle + \beta|1\rangle) \otimes |\phi^{\text{in}}\rangle$, with $\alpha \neq 0$ and $\beta \neq 0$ and $|\phi^{\text{in}}\rangle \in \mathcal{H}_{\text{in}}$, together with copies of its corresponding output $|\Phi^{\text{out}}\rangle = U^c|\Phi^{\text{in}}\rangle$.

Therefore, in the following we assume in addition to the multiple copies of states in S_{in} and S_{out} , we are also given multiple copies of states $|\Phi^{\text{in}}\rangle$ and $|\Phi^{\text{out}}\rangle$. Again, we naturally assume the set S_{in} generates the full matrix algebra on \mathcal{H}_{in} . This together with the fact that $|\phi^{\text{in}}\rangle \in \mathcal{H}_{\text{in}}$ implies the set

$$(\{|0\rangle\langle 0|, |1\rangle\langle 1| \otimes S_{\text{in}}\} \cup \{|\Phi^{\text{in}}\rangle\langle\Phi^{\text{in}}|\}) \quad (\text{D2})$$

generates the full matrix algebra on $\mathbb{C}^2 \otimes \mathcal{H}_{\text{in}}$, where \mathbb{C}^2 denotes the Hilbert space of the controlled qubit. It follows that, given these resources we can now implement the algorithm proposed in the paper to emulate the controlled-unitary U^c on $\mathbb{C}^2 \otimes \mathcal{H}_{\text{in}}$. In other words, if instead of states in the sets S_{in} and S_{out} we choose states from the sets $(\{|0\rangle\langle 0|, |1\rangle\langle 1| \otimes S_{\text{in}}\} \cup \{|\Phi^{\text{in}}\rangle\langle\Phi^{\text{in}}|\})$ and $(\{|0\rangle\langle 0|, |1\rangle\langle 1| \otimes S_{\text{out}}\} \cup \{|\Phi^{\text{out}}\rangle\langle\Phi^{\text{out}}|\})$ respectively, we implement unitary U^c .

3. More efficient algorithm with prior information about samples

The algorithm presented in the paper does not assume any prior information about the sample input states, or the relation between them. On the other hand, as we show in the following, making some assumptions about the sample input states, we can emulate the unknown unitary more efficiently. Note that we do not assume any prior information about any single sample states; the assumption is only about the relation between them. More precisely, the assumption is about the pairwise inner product between the states in the input set.

The main idea behind this version of algorithm is again emulating via coherent erasing. We use the fact that by measuring the system in two conjugate bases, we can completely erase the state of system. Therefore, we assume we are given multiple copies of states in an orthonormal basis for the input subspace, and their corresponding output states. We also assume we are given multiple copies of one (or more) state in the conjugate basis, and their corresponding outputs. Then using these sample states we can simulate coherent measurements in both basis, and coherently erase the state of system.

Let $\{|\theta_k^{\text{in}}\rangle : k = 0 \dots d-1\}$ be d unknown orthonormal states in the d -dimensional *input subspace* \mathcal{H}_{in} , and

$$|\alpha_j^{\text{in}}\rangle = \frac{1}{\sqrt{d}} \sum_{k=0}^{d-1} e^{i2\pi kj/d} |\theta_k^{\text{in}}\rangle \quad : j = 0 \dots d-1 , \quad (\text{D3})$$

be the orthonormal basis for \mathcal{H}_{in} , which is conjugate to $\{|\theta_k^{\text{in}}\rangle : k = 0 \dots d-1\}$. We assume we are given multiple copies of the input states $\{|\theta_k^{\text{in}}\rangle : k = 0 \dots d-1\}$ and $\{|\alpha_j^{\text{in}}\rangle : j = 0 \dots d-1\}$, and their corresponding output states $\{|\theta_k^{\text{out}}\rangle = U|\theta_k^{\text{in}}\rangle : k = 0 \dots d-1\}$ and $\{|\alpha_j^{\text{out}}\rangle = U|\alpha_j^{\text{in}}\rangle : j = 0 \dots d-1\}$. Note that, even if we are only given multiple copies of states $\{|\theta_k^{\text{in}}\rangle : k = 0 \dots d-1\}$ and multiple copies of *one* of the states in $\{|\alpha_k^{\text{in}}\rangle : k = 0 \dots d-1\}$, we can efficiently generate all states in this set (and similarly, for the set $\{|\alpha_k^{\text{out}}\rangle : k = 0 \dots d-1\}$).⁵

In the first step of this algorithm we perform a coherent measurement in $\{|\theta_k^{\text{in}}\rangle : k = 0 \dots d-1\}$ basis. To do this we couple the system to an ancillary system with a d -dimensional Hilbert space. The ancillary system is initially prepared in the state $|\Gamma\rangle = \sqrt{d^{-1}} \sum_{t=0}^{d-1} |t\rangle$, where $\{|t\rangle : t = 0 \dots d-1\}$ is a standard orthonormal basis. Then, we use the given copies of sample states $\{|\theta_k^{\text{in}}\rangle : k = 0 \dots d-1\}$ to simulate the unitary

$$V_P^{\text{in}} = \sum_{k,t=0}^{d-1} e^{itk2\pi/d} |\theta_k^{\text{in}}\rangle\langle\theta_k^{\text{in}}| \otimes |t\rangle\langle t| = \sum_{t=0}^{d-1} e^{itP^{\text{in}}2\pi/d} \otimes |t\rangle\langle t| , \quad (\text{D4})$$

where $P^{\text{in}} = \sum_{k=0}^{d-1} k |\theta_k^{\text{in}}\rangle\langle\theta_k^{\text{in}}|$. To efficiently simulate this unitary we first note that it has a decomposition as

$$V_P^{\text{in}} = \prod_{k=0}^{d-1} \sum_{t=0}^{d-1} e^{it|\theta_k^{\text{in}}\rangle\langle\theta_k^{\text{in}}|2\pi k/d} \otimes |t\rangle\langle t| . \quad (\text{D5})$$

Therefore, to simulate V_P^{in} we can simulate the (commuting) unitaries $\sum_{t=0}^{d-1} e^{it|\theta_k^{\text{in}}\rangle\langle\theta_k^{\text{in}}|2\pi k/d} \otimes |t\rangle\langle t|$, for $k = 0, \dots, d-1$. Each unitary $\sum_{t=0}^{d-1} e^{it|\theta_k^{\text{in}}\rangle\langle\theta_k^{\text{in}}|2\pi k/d} \otimes |t\rangle\langle t|$ can be efficiently simulated using the given copies of state $|\theta_k^{\text{in}}\rangle$. Using the results presented in Appendix C, this simulation can be done in time $\mathcal{O}(\log D)$, where D is the dimension of the Hilbert space. Then, it follows from the decomposition of V_P^{in} given by Eq.(D5) that we can simulate V_P^{in} in time $\mathcal{O}(d \log D)$, using the given copies of sample states $\{|\theta_k^{\text{in}}\rangle, k = 0 \dots d-1\}$.

Applying the unitary V_P^{in} to the system in state $|\psi\rangle \in \mathcal{H}_{\text{in}}$ and the ancillary system prepared in state $|\Gamma\rangle$, we get state

$$|\psi\rangle \otimes |\Gamma\rangle = \sum_{k=0}^{d-1} \psi_k |\theta_k^{\text{in}}\rangle \otimes |\Gamma\rangle \longrightarrow V_P^{\text{in}} |\psi\rangle \otimes |\Gamma\rangle = \sum_{k=0}^{d-1} \psi_k |\theta_k^{\text{in}}\rangle \otimes \frac{1}{\sqrt{d}} \sum_{t=0}^{d-1} e^{itk2\pi/d} |t\rangle , \quad (\text{D6})$$

⁵ Let $P^{\text{in}} = \sum_{k=0}^{d-1} k |\theta_k^{\text{in}}\rangle\langle\theta_k^{\text{in}}|$. Having multiple copies of states $\{|\theta_k^{\text{in}}\rangle : k = 0 \dots d-1\}$ we can simulate the unitary $e^{-iP^{\text{in}}s}$ for any real s , and by applying this unitary, for different values of s , we can transform one element of the set $\{|\alpha_k^{\text{in}}\rangle : k = 0 \dots d-1\}$ to the other elements. To simulate the unitary $e^{-iP^{\text{in}}s}$, we note that $e^{-iP^{\text{in}}s} = \prod_{k=0}^{d-1} e^{-isk|\theta_k^{\text{in}}\rangle\langle\theta_k^{\text{in}}|}$. As we have seen in Appendix C each unitary $e^{-isk|\theta_k^{\text{in}}\rangle\langle\theta_k^{\text{in}}|}$ can be efficiently simulated using multiple copies of state $|\theta_k^{\text{in}}\rangle$, in time $\mathcal{O}(\log(D))$. Therefore, we can simulate the unitary $e^{-iP^{\text{in}}s}$ in time $\mathcal{O}(d \log(D))$.

where we have used the decomposition $|\psi\rangle = \sum_{k=0}^{d-1} \psi_k |\theta_k^{\text{in}}\rangle$.

Then, performing quantum Fourier transform on the ancillary system we transform the joint state to

$$V_P^{\text{in}} |\psi\rangle \otimes |\Gamma\rangle = \sum_{k=0}^{d-1} \psi_k |\theta_k^{\text{in}}\rangle \otimes \sum_{t=0}^{d-1} e^{itk2\pi/d} |k\rangle \xrightarrow{\text{QFT}} \sum_{k=0}^{d-1} \psi_k |\theta_k^{\text{in}}\rangle \otimes |k\rangle . \quad (\text{D7})$$

Then, to erase the information in the system we implement the unitary

$$V_Q^{\text{in}} = \sum_{k,t=0}^{d-1} e^{itk2\pi/d} |\alpha_k^{\text{in}}\rangle \langle \alpha_k^{\text{in}}| \otimes |t\rangle \langle t| = \sum_{t=0}^{d-1} e^{itQ^{\text{in}}2\pi/d} \otimes |t\rangle \langle t| , \quad (\text{D8})$$

where $Q^{\text{in}} = \sum_{l=0}^{d-1} l |\alpha_l^{\text{in}}\rangle \langle \alpha_l^{\text{in}}|$. Note that we can efficiently simulate V_Q^{in} , using a similar approach we used to simulate V_P^{in} .

Applying V_Q^{in} to state in Eq.(D7) we find

$$\sum_{k=0}^{d-1} \psi_k |\theta_k^{\text{in}}\rangle \otimes |k\rangle \longrightarrow V_Q^{\text{in}} \sum_{k=0}^{d-1} \psi_k |\theta_k^{\text{in}}\rangle \otimes |k\rangle = \sum_{k=0}^{d-1} \psi_k e^{ikQ^{\text{in}}2\pi/d} |\theta_k^{\text{in}}\rangle \otimes |k\rangle = |\theta_0^{\text{in}}\rangle \otimes \sum_{k=0}^{d-1} \psi_k |k\rangle , \quad (\text{D9})$$

where we have used the fact that

$$e^{ikQ^{\text{in}}2\pi/d} |\theta_k^{\text{in}}\rangle = \sum_{l=0}^{d-1} e^{i2\pi lk/d} |\alpha_l^{\text{in}}\rangle \langle \alpha_l^{\text{in}}| \theta_k^{\text{in}}\rangle = \frac{1}{\sqrt{d}} \sum_{l=0}^{d-1} e^{i2\pi lk/d} e^{-i2\pi kl/d} |\alpha_l^{\text{in}}\rangle = |\theta_0^{\text{in}}\rangle . \quad (\text{D10})$$

Therefore, after these three steps for any input state $|\psi\rangle \in \mathcal{H}_{\text{in}}$ we have

$$[V_Q^{\text{in}}(I \otimes F_a) V_P^{\text{in}}] |\psi\rangle |\Gamma\rangle = |\theta_0^{\text{in}}\rangle \otimes \sum_{k=0}^{d-1} \psi_k |k\rangle , \quad (\text{D11})$$

where F_a denotes the quantum Fourier transform on the ancillary system.

At this point we have completely erased the state of system and transferred all its information to the ancillary system. Now using a method similar to the one used in the main algorithm, we can transform this state to state $U|\psi\rangle |\Gamma\rangle$: We replace $|\theta_0^{\text{in}}\rangle$ with $|\theta_0^{\text{out}}\rangle$ and apply $V_P^{\text{out}}(I \otimes F_a) V_Q^{\text{out}} \otimes I$ to state $|\theta_0^{\text{out}}\rangle \otimes \sum_{k=0}^{d-1} \psi_k |k\rangle$, where

$$V_Q^{\text{out}} = \sum_{k,t=0}^{d-1} e^{itk2\pi/d} |\alpha_k^{\text{out}}\rangle \langle \alpha_k^{\text{out}}| \otimes |t\rangle \langle t| = \sum_{t=0}^{d-1} e^{itQ^{\text{out}}2\pi/d} \otimes |t\rangle \langle t| , \quad (\text{D12})$$

$$V_P^{\text{out}} = \sum_{k,t=0}^{d-1} e^{itk2\pi/d} |\theta_k^{\text{out}}\rangle \langle \theta_k^{\text{out}}| \otimes |t\rangle \langle t| = \sum_{t=0}^{d-1} e^{itP^{\text{out}}2\pi/d} \otimes |t\rangle \langle t| . \quad (\text{D13})$$

Note that these unitaries can be efficiently simulated using the same method we used to simulate the unitary V_P^{in} .

It follows that we can emulate the action of unitary U on the input subspace in time $\mathcal{O}(d \log(D))$.

4. Approximate transformations

In the above discussions we always assumed there exists a unitary U for which $U|\phi_k^{\text{in}}\rangle = |\phi_k^{\text{out}}\rangle$, for all $k = 1, \dots, K$. Now suppose we only know that these transformations are possible approximately. I.e. there exists a constant $\delta > 0$ and a unitary U such that $\|U|\phi_k^{\text{in}}\rangle - |\phi_k^{\text{out}}\rangle\| \leq \delta$, for $k = 1, \dots, K$. Then, it can be easily shown that if we run the proposed algorithm on an input state $|\psi\rangle \in \mathcal{H}_{\text{in}}$, with the given sample input-output pairs, in the output we generate a quantum state whose trace distance from state $U|\psi\rangle$ is bounded by $\mathcal{O}(T \times \delta) = \delta \times \text{poly}(d)$. This follows from the fact that, under the above assumption, we can repeat the argument that is used to derive Eq.(3.3) from Eq.(3.2), and use $\|U|\phi_1^{\text{in}}\rangle - |\phi_1^{\text{out}}\rangle\| \leq \delta$ and $\|(U \otimes I_a) W_a^{\text{in}}(k)(U^\dagger \otimes I_a) - W_a^{\text{out}}(k)\| \leq 4\delta$, which follows from $\|U|\phi_k^{\text{in}}\rangle - |\phi_k^{\text{out}}\rangle\| \leq \delta$, for $k = 1, \dots, K$. This implies Eq.(3.3) is satisfied with an additional error of order $\mathcal{O}(T\delta)$, which proves the claim. Note that, for a general input-output sets S_{in} and S_{out} and a general input state $|\psi\rangle \in \mathcal{H}_{\text{in}}$, the lowest achievable error in this transformation is $\mathcal{O}(d \times \delta)$.

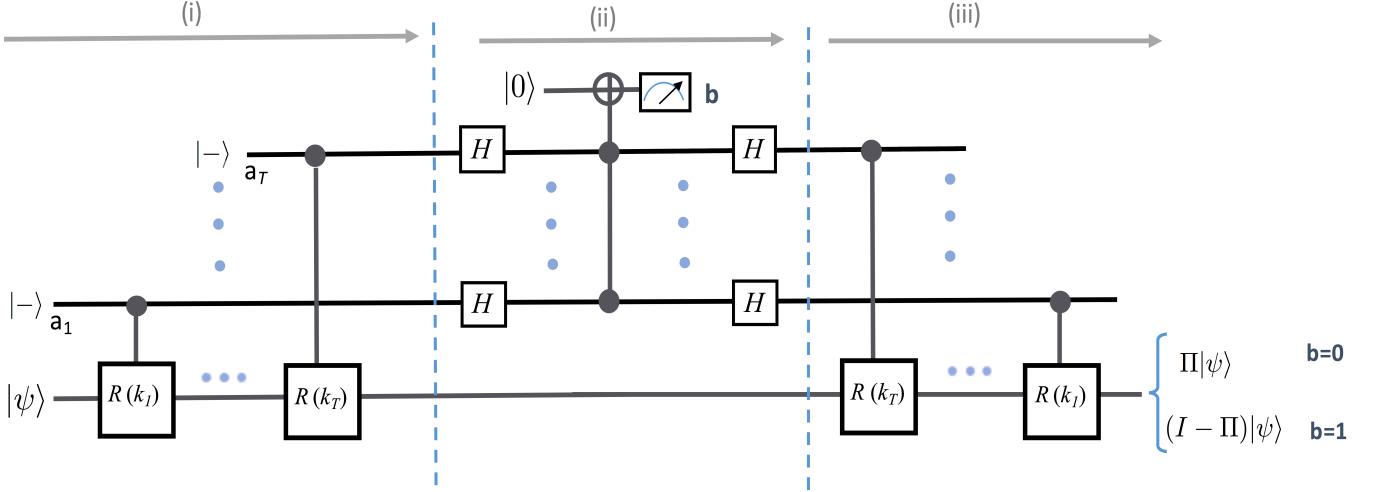


FIG. 3: The circuit for emulating the two-outcome projective measurement with projectors $\{\Pi, I - \Pi\}$, where Π is the projector to the subspace spanned by the sample states $\{|\phi_k\rangle : 1, \dots, K\}$.

Appendix E: Emulating projective measurements

In this section we consider the problem of emulating projective measurements. We assume we are given multiple copies of states that belong to different subspaces of the Hilbert space, with the labels which specify these subspaces. Then, we are interested to simulate the projective measurement that projects any given input state to one of these subspaces. Note that any projective measurement can be realized as a sequence of two-outcome measurements. Therefore, in the following we focus on implementing two-outcome measurements.

Consider the set of sample states $\{|\phi_k\rangle : 1, \dots, K\}$. We do not make any assumption about the sample states, or the relation between them. Let \mathcal{H}_Π be the subspace spanned by these states, and Π be the projector to this subspace. Assuming we are given multiple copies of each state in this set we are interested to implement the projective measurement described by the projectors $\{\Pi, I - \Pi\}$.

In the algorithm we use the controlled-reflections

$$R_a(k) = |0\rangle\langle 0|_a \otimes I + |1\rangle\langle 1|_a \otimes e^{i\pi|\phi_k\rangle\langle\phi_k|}. \quad (\text{E1})$$

As we explained in Appendix C, using n copies of state $|\phi_k\rangle$ we can implement this unitary in time $\mathcal{O}(\log(D) \times n \log(n))$, with error $\mathcal{O}(\frac{1}{n})$ in trace distance, where D is the dimension of the Hilbert space.

a. The Algorithm

- The quantum circuit for this algorithm is presented in Fig.(3). The algorithm has the following three steps:
- (i) Consider T qubits $a_1 \dots a_T$, which are all initially prepared in state $|-\rangle = (|0\rangle - |1\rangle)/\sqrt{2}$. Let k_1, \dots, k_T be T independent random integers, each chosen uniformly at random from $1, \dots, K$. Apply the controlled-reflection $R_{a_1}(k_1)$ to the system and qubit a_1 , $R_{a_2}(k_2)$ to the system and qubit a_2 , until the last controlled-reflection $R_{a_T}(k_T)$, which is applied to the system and qubit a_T .
 - (ii) Perform the two-outcome projective measurement $\{|-\rangle\langle -|^{\otimes T}, I - |-\rangle\langle -|^{\otimes T}\}$ on qubits $a_1 \dots a_T$. (This can be implemented efficiently, e.g., by first applying the Hadamard gates on qubits $a_1 \dots a_T$, then applying a T -bit Toffoli gate controlled by $a_1 \dots a_T$, that is acting on an ancillary qubit initialized in state $|0\rangle$, and finally applying the Hadamard gates on qubits $a_1 \dots a_T$ again. Then, measuring this ancillary qubit in the computational basis, realizes the desired measurement.)
 - (iii) Recall the random integers k_1, \dots, k_T chosen in step (i). Apply the controlled-reflection $R_{a_T}(k_T)$ to the system and qubit a_T , then apply the controlled-reflection $R_{a_{T-1}}(k_{T-1})$ to the system and qubit a_{T-1} , until the last controlled-reflection $R_{a_1}(k_1)$ which is applied to the system and qubit a_1 .

Then, at the end of the algorithm, the initial state $|\psi\rangle$ is projected to state $\Pi|\psi\rangle/\sqrt{\langle\psi|\Pi|\psi\rangle}$ with probability $\langle\psi|\Pi|\psi\rangle$ (corresponding to outcome $b = 0$ in the measurement in step (ii)), and is projected to state $(I - \Pi)|\psi\rangle/\sqrt{1 - \langle\psi|\Pi|\psi\rangle}$ with probability $1 - \langle\psi|\Pi|\psi\rangle$ (corresponding to outcome $b = 1$ in the measurement).

b. How it works

In the following we assume the circuit is run with perfect controlled-reflections. Then, the additional error due to the imperfections in the simulations of the controlled-reflections can be taken into account, using the same approach we used for the main algorithm.

Let \mathcal{E} be the quantum channel that describes the overall effect of circuit in Fig.(3) in the case where we ignore the output of measurement, i.e. we do not postselect. Let \mathcal{H}_Π be the subspace spanned by the sample states $\{|\phi_k\rangle : k = 1, \dots, K\}$.

Theorem 4 *Let λ_{\min} be the minimum non-zero eigenvalue of the density operator $K^{-1} \sum_{k=1}^K |\phi_k\rangle\langle\phi_k|$. If the support of ρ , the state of system, is contained in either the subspace \mathcal{H}_Π or the orthogonal subspace, then the probability of error in emulating the measurement $\{I - \Pi, \Pi\}$ is less than or equal to $(1 - \lambda_{\min})^T$. Furthermore, the fidelity of $\mathcal{E}(\rho)$, the output of the circuit, with ρ the desired output state, satisfies $F(\mathcal{E}(\rho), \rho) \geq 1 - (1 - \lambda_{\min})^T$.*

Proof. First, consider the case where ρ , the initial state of system, does not have any support on \mathcal{H}_Π , i.e. $\Pi\rho\Pi = 0$. Then, all the reflections $e^{i\pi|\phi_k\rangle\langle\phi_k|}$ act as the identity operator on this state. Therefore, after step (i) the state of qubits $a_1 \cdots a_T$ remain unchanged. In this case, in step (ii) with probability one we project these qubits to state $|-\rangle^{\otimes T}$. It follows that in this case the algorithm emulates the action of measurement perfectly.

Next, we consider the case where $\Pi\rho\Pi = \rho$. Using the fact that

$$\langle - | R_a(k) | - \rangle_a = \frac{I + e^{i\pi|\phi_k\rangle\langle\phi_k|}}{2} = I - |\phi_k\rangle\langle\phi_k|, \quad (\text{E2})$$

we find that for a general state ρ (not necessarily restricted to \mathcal{H}_Π) the probability that in step (ii) we find the qubits $a_1 \cdots a_T$ in state $|-\rangle^{\otimes T}$ is given by

$$p(|-\rangle^{\otimes T}) = \frac{1}{K^T} \sum_{k_1, \dots, k_T}^K \text{Tr}\left([|-\rangle\langle -|^{\otimes T} \otimes I] R_{a_T}(k_T) \cdots R_{a_1}(k_1) [\rho \otimes |-\rangle\langle -|^{\otimes T}] R_{a_1}(k_1) \cdots R_{a_T}(k_T)\right) \quad (\text{E3a})$$

$$= \frac{1}{K^T} \sum_{k_1, \dots, k_T}^K \text{Tr}\left((I - |\phi_{k_T}\rangle\langle\phi_{k_T}|) \cdots (I - |\phi_{k_1}\rangle\langle\phi_{k_1}|) \rho (I - |\phi_{k_1}\rangle\langle\phi_{k_1}|) \cdots (I - |\phi_{k_T}\rangle\langle\phi_{k_T}|)\right) \quad (\text{E3b})$$

$$= \text{Tr}(\mathcal{M}^T(\rho)), \quad (\text{E3c})$$

where $\mathcal{M}(\cdot) = \frac{1}{K} \sum_{k=1}^K P_k^\perp(\cdot) P_k^\perp$, and $P_k^\perp = I - |\phi_k\rangle\langle\phi_k|$.

Next, note that for any operator X we have

$$\text{Tr}(\mathcal{M}(X)) = \frac{1}{K} \sum_{k=1}^K \text{Tr}(XP_k^\perp) = \text{Tr}(X) - \text{Tr}\left(X \sum_{k=1}^K \frac{1}{K} |\phi_k\rangle\langle\phi_k|\right) = \text{Tr}(X) - \text{Tr}(X\sigma_{\text{avg}}), \quad (\text{E4})$$

where

$$\sigma_{\text{avg}} \equiv \frac{1}{K} \sum_K |\phi_k\rangle\langle\phi_k|. \quad (\text{E5})$$

The subspace \mathcal{H}_Π is defined as the subspace spanned by $\{|\phi_k\rangle : k = 1, \dots, K\}$. Therefore, the density matrix σ_{avg} is automatically full-rank in this subspace. Let λ_{\min} be the minimum eigenvalue of σ_{avg} in this subspace, i.e. the minimum nonzero eigenvalue of σ_{avg} . Note that for any positive operator X whose support is restricted to \mathcal{H}_Π we have

$$\text{Tr}(X\sigma_{\text{avg}}) \geq \lambda_{\min} \text{Tr}(X). \quad (\text{E6})$$

Then, using Eq.(E4) this means that for any positive operator X , whose support is restricted to \mathcal{H}_Π we have

$$\text{Tr}(\mathcal{M}(X)) = \text{Tr}(X) - \text{Tr}(X\sigma_{\text{avg}}) \leq \text{Tr}(X)(1 - \lambda_{\min}). \quad (\text{E7})$$

Next, note that if X is a positive operator with support restricted to \mathcal{H}_Π , then $\mathcal{M}(X)$ is also a positive operator with support restricted to \mathcal{H}_Π . It follows that if X is a positive operator with support restricted to \mathcal{H}_Π , then

$$\text{Tr}(\mathcal{M}^T(X)) \leq \text{Tr}(X) \times (1 - \lambda_{\min})^T. \quad (\text{E8})$$

This together with Eq.(E3) imply that if the support of the initial state ρ is restricted to \mathcal{H}_{Π} , then the probability that at the end of algorithm we find qubits $a_1 \cdots a_T$ in state $|-\rangle^{\otimes T}$ is given by

$$p(|-\rangle^{\otimes T}) = \text{Tr}(\mathcal{M}^T(\rho)) \leq (1 - \lambda_{\min})^T. \quad (\text{E9})$$

This completes the proof of the first part of the theorem.

Finally, the second part of the theorem, i.e. the bound on the fidelity, follows from the following lemma, which is proven in the same way we proved Theorem 1.

Lemma 5 Consider the following quantum operation: (i) We apply a unitary $V(\boldsymbol{\lambda})$ to the system, where $\boldsymbol{\lambda}$ is a random parameter chosen according to the probability distribution $p(\boldsymbol{\lambda})$. (ii) We perform a projective measurement on the system with projectors $\{P_\alpha : \alpha \in A\}$. (iii) We apply the unitary $V^\dagger(\boldsymbol{\lambda})$.

Let \mathcal{G} be the quantum channel that describes the overall effect of this operation on the system, in the case where we ignore the outcome of the projective measurement. Then, for any input state ρ , the fidelity of ρ and $\mathcal{G}(\rho)$ satisfies

$$F(\rho, \mathcal{G}(\rho)) \geq \max_{\alpha \in A} p_\alpha(\rho) \quad (\text{E10})$$

where $p_\alpha(\rho)$ is the (average) probability of outcome α in the measurement.

Proof. Quantum operation \mathcal{G} is given by

$$\mathcal{G}(\cdot) = \sum_{\boldsymbol{\lambda}} p(\boldsymbol{\lambda}) \sum_{\alpha \in A} V^\dagger(\boldsymbol{\lambda}) P_\alpha V(\boldsymbol{\lambda})(\cdot) V^\dagger(\boldsymbol{\lambda}) P_\alpha V(\boldsymbol{\lambda}). \quad (\text{E11})$$

Consider a pure input state $|\gamma\rangle$. Then, the squared fidelity of $\mathcal{G}(|\gamma\rangle\langle\gamma|)$ and $|\gamma\rangle\langle\gamma|$ is given by

$$F^2(\mathcal{G}(|\gamma\rangle\langle\gamma|), |\gamma\rangle\langle\gamma|) = \langle\gamma|\mathcal{G}(|\gamma\rangle\langle\gamma|)|\gamma\rangle \quad (\text{E12})$$

$$= \sum_{\alpha \in A} \sum_{\boldsymbol{\lambda}} p(\boldsymbol{\lambda}) \langle\gamma|V^\dagger(\boldsymbol{\lambda}) P_\alpha V(\boldsymbol{\lambda})|\gamma\rangle^2 \quad (\text{E13})$$

$$\geq \max_{\alpha \in A} \sum_{\boldsymbol{\lambda}} p(\boldsymbol{\lambda}) \langle\gamma|V^\dagger(\boldsymbol{\lambda}) P_\alpha V(\boldsymbol{\lambda})|\gamma\rangle^2 \quad (\text{E14})$$

$$\geq \max_{\alpha \in A} \left(\sum_{\boldsymbol{\lambda}} p(\boldsymbol{\lambda}) \langle\gamma|V^\dagger(\boldsymbol{\lambda}) P_\alpha V(\boldsymbol{\lambda})|\gamma\rangle \right)^2 \quad (\text{E15})$$

$$= \max_{\alpha \in A} p_\alpha^2(|\gamma\rangle\langle\gamma|), \quad (\text{E16})$$

where to get the fourth line we have used the fact that the variance of any random variable is non-negative, and to get the last line we have used the fact that the average density operator of the system before the measurement is $\sum_{\boldsymbol{\lambda}} p(\boldsymbol{\lambda}) V(\boldsymbol{\lambda}) |\gamma\rangle\langle\gamma| V^\dagger(\boldsymbol{\lambda})$, and so the probability of outcome α is $\sum_{\boldsymbol{\lambda}} p(\boldsymbol{\lambda}) \langle\gamma|V^\dagger(\boldsymbol{\lambda}) P_\alpha V(\boldsymbol{\lambda})|\gamma\rangle$. This proves the lemma for the special case of pure states. The result for general mixed states follows from the joint concavity of fidelity. ■

This lemma together with the fact that for input states contained in \mathcal{H}_{Π} , or the complement subspace the outcome of the ideal measurement is deterministic, and the circuit generates this outcome with probability of error less than $(1 - \lambda_{\min})^T$, proves the bound

$$F(\mathcal{E}(\rho), \rho) \geq 1 - (1 - \lambda_{\min})^T. \quad (\text{E17})$$

This completes the proof of theorem.

■