Variational Quantum Fidelity Estimation

M. Cerezo[1,2], Alexander Poremba[3], Lukasz Cincio[1], and Patrick J. Coles[1]

[1]Theoretical Division, Los Alamos National Laboratory, Los Alamos, NM, USA

[2]Center for Nonlinear Studies, Los Alamos National Laboratory, Los Alamos, NM, USA

[3]Computing and Mathematical Sciences, California Institute of Technology, Pasadena, CA, USA.

Computing quantum state fidelity will be important to verify and characterize states prepared on a quantum computer. In this work, we propose novel lower and upper bounds for the fidelity $F(\rho, \sigma)$ based on the "truncated fidelity" $F(\rho_m, \sigma)$, which is evaluated for a state ρ_m obtained by projecting ρ onto its m-largest eigenvalues. Our bounds can be refined, i.e., they tighten monotonically with m. To compute our bounds, we introduce a hybrid quantum-classical algorithm, called Variational Quantum Fidelity Estimation, that involves three steps: (1) variationally diagonalize ρ, (2) compute matrix elements of σ in the eigenbasis of ρ, and (3) combine these matrix elements to compute our bounds. Our algorithm is aimed at the case where σ is arbitrary and ρ is low rank, which we call low-rank fidelity estimation, and we prove that no classical algorithm can efficiently solve this problem under reasonable assumptions. Finally, we demonstrate that our bounds can detect quantum phase transitions and are often tighter than previously known computable bounds for realistic situations.

1 Introduction

In the near future, quantum computers will become quantum state preparation factories. They will prepare ground and excited states of molecules [1], states that simulate quantum dynamics [2], and states that encode the solutions to linear systems [3]. These states will necessarily be impure, either intentionally (e.g., when studying thermal states) or due to incoherent noise of the quantum computer (e.g., T_1 and T_2 processes). Verification and characterization of these mixed states will be important, and hence efficient algorithms will be needed for this purpose. A widely used measure for verification and characterization is the fidelity [4–8]. For example, one may be interested in the fidelity with a fixed target state (i.e., for verification) or the fidelity between subsystems of many-body states to study behavior near a phase transition (i.e., for characterization) [9]. For two states ρ and σ,

the fidelity is defined as [4–7]

$$F(\rho, \sigma) = \mathrm{Tr}\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}} = \|\sqrt{\rho}\sqrt{\sigma}\|_1, \qquad (1)$$

where $\|A\|_1 = \mathrm{Tr}\sqrt{A^\dagger A}$. See [6–8] for properties of $F(\rho, \sigma)$.

Classically computing $F(\rho, \sigma)$, or any other metric on quantum states, could scale exponentially due to the exponentially large dimension of the density matrix. This raises the question of whether a quantum computer could avoid this exponential scaling. However, $F(\rho, \sigma)$ involves non-integer powers of ρ and σ, implying that there is no exact quantum algorithm for computing it directly from the probability of a measurement outcome on a finite number of copies of ρ and σ. In addition, deciding whether the trace distance (which is closely related to fidelity [6, 7]) is large or small is QSZK-complete [10]. Here, QSZK (quantum statistical zero-knowledge) is a complexity class that contains BQP [6] (bounded-error quantum polynomial time). It is therefore reasonable to suspect that estimating $F(\rho, \sigma)$ is hard even for quantum computers.

This does not preclude the efficient estimation of fidelity for the practical case when one of the states is low rank. Low-rank states appear in condensed matter physics [11] (marginals of grounds states) and data science [12] (covariance matrices). We define Low-Rank Fidelity Estimation as the problem of estimating $F(\rho, \sigma)$ when σ is arbitrary and ρ is approximately low rank. We prove under standard assumptions that a classical algorithm cannot efficiently perform Low-Rank Fidelity Estimation.

In this work, we propose a variational hybrid quantum-classical algorithm [1, 13–21] for Low-Rank Fidelity Estimation called Variational Quantum Fidelity Estimation (VQFE). VQFE computes upper and lower bounds on $F(\rho, \sigma)$ that can be refined to arbitrary tightness. Our bounds are based on the *truncated fidelity*, which involves evaluating (1) for σ and ρ_m, a truncated version of ρ obtained by projecting ρ onto the subspace associated with its m-largest eigenvalues. Crucially, our bounds tighten monotonically in m, and eventually they equal the fidelity when $m = \mathrm{rank}(\rho)$.

This is in contrast to the state-of-the-art quantum algorithm to bound the fidelity, which employs

fixed bounds called the sub- and super-fidelity bounds (SSFB) [22–25], defined as

$$E(\rho, \sigma) = \mathrm{Tr}\rho\sigma + \sqrt{2\left[(\mathrm{Tr}\rho\sigma)^2 - \mathrm{Tr}\rho\sigma\rho\sigma\right]}, \quad (2)$$

$$G(\rho, \sigma) = \mathrm{Tr}\rho\sigma + \sqrt{(1 - \mathrm{Tr}\rho^2)(1 - \mathrm{Tr}\sigma^2)}, \quad (3)$$

respectively, such that $\sqrt{E(\rho, \sigma)} \leqslant F(\rho, \sigma) \leqslant \sqrt{G(\rho, \sigma)}$. Since the SSFB are expressed as traces of products of density matrices, they can be efficiently measured on a quantum computer [26–29]. These bounds are generally looser when both ρ and σ have high rank, and hence the SSFB likewise perform better when one of the states is low rank. Below we give a detailed comparison, showing that VQFE often outperforms the SSFB. We also note that VQFE only requires $2n + 1$ qubits while the SSFB require $4n + 1$ qubits, for n-qubit states.

To produce certified bounds from the output of VQFE, we prove several novel bounds that should be of independent interest in quantum information theory. In addition, we introduce the *fidelity spectrum*, which is the collection of all truncated fidelities. In the same sense that the entanglement spectrum provides more information than a single entanglement measure [11], we argue that the fidelity spectrum gives more information about the closeness of ρ and σ than just $F(\rho, \sigma)$.

In what follows, we first present our bounds and the VQFE algorithm. We compare its performance with the SSFB and illustrate its application to phase transitions. All proofs of our results are delegated to the Appendix.

2 Truncated fidelity bounds

Let Π_m^ρ be the projector onto the subspace spanned by the eigenvectors of ρ with the m-largest eigenvalues. Consider the sub-normalized states

$$\rho_m^\rho = \Pi_m^\rho \rho \Pi_m^\rho = \sum_{i=1}^m r_i |r_i\rangle\langle r_i|, \quad \sigma_m^\rho = \Pi_m^\rho \sigma \Pi_m^\rho, \quad (4)$$

where the eigenvalues $\{r_i\}$ of ρ are in decreasing order ($r_i \geqslant r_{i+1}$). For simplicity we denote $\rho_m^\rho = \rho_m$. These sub-normalized states can be used to define the *truncated fidelity*

$$F(\rho_m, \sigma_m^\rho) = F(\rho_m, \sigma) = \left\|\sqrt{\rho_m}\sqrt{\sigma}\right\|_1, \quad (5)$$

and the *truncated generalized fidelity*

$$F_*(\rho_m, \sigma_m^\rho) = \left\|\sqrt{\rho_m}\sqrt{\sigma}\right\|_1$$
$$+ \sqrt{(1 - \mathrm{Tr}\rho_m)(1 - \mathrm{Tr}\sigma_m^\rho)}, \quad (6)$$

where the generalized fidelity [30, 31] was defined for two sub-normalized states as $F_*(\rho, \sigma) = \|\sqrt{\rho}\sqrt{\sigma}\|_1 + \sqrt{(1 - \mathrm{Tr}\,\rho)(1 - \mathrm{Tr}\,\sigma)}$. The generalized fidelity reduces to (1) if at least one state is normalized. Equations (5) and (6) are in fact lower and upper bounds, respectively, for the fidelity.
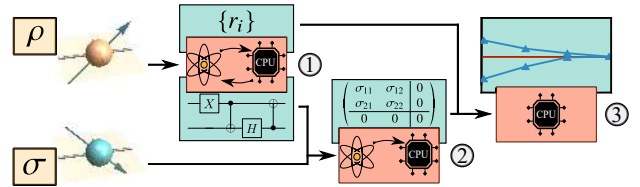

Figure 1: The VQFE algorithm. (1) First, ρ is diagonalized with a hybrid quantum-classical optimization loop, outputting the largest eigenvalues $\{r_i\}$ of ρ and a gate sequence that prepares the associated eigenvectors. (2) Second, a hybrid quantum-classical computation gives the matrix elements of σ in the eigenbasis of ρ. (3) Finally, classical processing gives upper and lower bounds on $F(\rho, \sigma)$.

Proposition 1. *The following truncated fidelity bounds (TFB) hold:*

$$F(\rho_m, \sigma_m^\rho) \leqslant F(\rho, \sigma) \leqslant F_*(\rho_m, \sigma_m^\rho). \quad (7)$$

We refer to the collection of TFB for different values of m as the *fidelity spectrum*. The TFB are satisfied with equality when $m = \mathrm{rank}(\rho)$. Moreover, they monotonically get tighter as m increases. Hence they can be refined to arbitrary tightness by increasing m.

Proposition 2. *The truncated fidelity $F(\rho_m, \sigma_m^\rho)$ is monotonically increasing in m, and the truncated generalized fidelity $F_*(\rho_m, \sigma_m^\rho)$ is monotonically decreasing in m.*

Ultimately, we will consider the case when ρ is either low rank or ϵ-*low-rank*. Here we define the ϵ-rank as a generalization of the rank to within some ϵ error:

$$\mathrm{rank}_\epsilon(\rho) = \min\{m \in \{1, ..., d\} : \|\rho - \rho_m\|_1 \leqslant \epsilon\}, \quad (8)$$

where d is the Hilbert space dimension. Note that $\mathrm{rank}_0(\rho) = \mathrm{rank}(\rho)$. We remark that the looseness of the TFB is bounded by the square root of $\|\rho - \rho_m\|_1 = 1 - \mathrm{Tr}(\rho_m)$:

$$F_*(\rho_m, \sigma_m^\rho) - F(\rho_m, \sigma_m^\rho) \leqslant \|\rho - \rho_m\|_1^{1/2}, \quad (9)$$

where we have used (5), (6), and $\mathrm{Tr}\sigma_m^\rho \leqslant 1$. Hence, the TFB looseness is bounded by $\sqrt{\epsilon}$ provided that $m = \mathrm{rank}_\epsilon(\rho)$.

3 The VQFE algorithm

Figure 1 shows the overall structure of the VQFE algorithm. VQFE involves three steps: (1) approximately diagonalize ρ with a variational hybrid quantum-classical algorithm, (2) compute matrix elements of σ in the eigenbasis of ρ, (3) classically process these matrix elements to produce certified bounds on $F(\rho, \sigma)$.

The first subroutine employs Variational Quantum State Diagonalization (VQSD) [18], a variational hybrid algorithm that takes in two copies of ρ and

outputs approximations of the m-largest eigenvalues $\{r_i\}$ and a gate sequence U that prepares the associated eigenvectors $\{|r_i\rangle\}$. This subroutine involves a quantum-classical optimization loop that minimizes a cost function C that quantifies how far $\tilde{\rho} = U\rho U^\dagger$ is from being diagonal. Namely, $C = D_{\mathrm{HS}}(\tilde{\rho}, \mathcal{Z}(\tilde{\rho}))$, with $D_{\mathrm{HS}}(\rho, \sigma) = \mathrm{Tr}[(\rho - \sigma)^2]$ the Hilbert-Schmidt distance and $\mathcal{Z}$ the dephasing channel in the computational basis. When $C = 0$ we have exact diagonalization, whereas for $C \neq 0$, VQSD outputs the eigenvalues and eigenvectors of

$$\rho' = U^\dagger \mathcal{Z}(\tilde{\rho})U = \sum_i r_i'|r_i'\rangle\langle r_i'|, \qquad (10)$$

such that $C = D_{\mathrm{HS}}(\rho, \rho')$.

The second subroutine measures the matrix elements $\sigma_{ij} = \langle r_i|\sigma|r_j\rangle$, or more precisely $\sigma_{ij}' = \langle r_i'|\sigma|r_j'\rangle$ if $C \neq 0$. This is done by preparing superpositions of the eigenvectors of the form $(|r_i'\rangle + |r_j'\rangle)/\sqrt{2}$ (by means of VQSD's eigenstate preparation circuit [18]) and computing the inner product of this superposition with σ via a Swap Test. (For example, see [19, 28] for the precise circuit.) For a fixed m, one needs to measure $m(m+1)/2$ matrix elements. However, when incrementing m to $m+1$, one only needs to measure $2m+1$ new quantities.

The third subroutine involves only classical computation, combining the outputs of the previous subroutines to produce bounds on $F(\rho, \sigma)$. These bounds employ the TFB described before, for which one needs to compute

$$\left\| \sqrt{\rho_m} \sqrt{\sigma} \right\|_1 = \mathrm{Tr}\sqrt{\sum_{i,j} T_{ij}|r_i\rangle\langle r_j|}. \qquad (11)$$

Here we have defined the $m \times m$ matrix T with elements $T_{ij} = \sqrt{r_i r_j}\langle r_i|\sigma|r_j\rangle = \sqrt{r_i r_j}\sigma_{ij}$ such that $T \geqslant 0$. Note that the T_{ij} can be computed directly from the outputs of the first two subroutines. One can then classically diagonalize T to compute (11). (If shot noise leads to a non-positive matrix T, one can efficiently find the maximum-likelihood matrix $\tilde{T} \geqslant 0$ which gives the observed results with highest probability [32], and use $\tilde{T}$ in place of T.) Note that since $\mathrm{Tr}\rho_m = \sum_{i=1}^m r_i$ and $\mathrm{Tr}\sigma_m^\rho = \sum_{i=1}^m \sigma_{ii}$, the TFB are completely determined by $\{r_i\}$ and $\{\sigma_{ij}\}$. Explicitly, the TFB are computed via $F(\rho_m, \sigma_m^\rho) = \sum_i \sqrt{\lambda_i}$ and $F_*(\rho_m, \sigma_m^\rho) = \sum_i \sqrt{\lambda_i} + \sqrt{(1 - \sum_i r_i)(1 - \sum_i \sigma_{ii})}$ with λ_i the eigenvalues of T and $i = 1, \dots, m$.

Technically, the aforementioned procedure computes $F(\rho_m', \sigma_m^{\rho'})$ and $F_*(\rho_m', \sigma_m^{\rho'})$. If $C \approx 0$, then $\rho' \approx \rho$ and these quantities are actually bounds on $F(\rho, \sigma)$. However, if C is appreciable then to produce certified bounds we must account for $\rho' \neq \rho$. Here we present two such certified bounds.

Proposition 3. *The following certified cost function bounds (CCFB) hold:*

$$F(\rho_m', \sigma_m^{\rho'}) - \delta \leqslant F(\rho, \sigma) \leqslant F_*(\rho_m', \sigma_m^{\rho'}) + \delta, \quad (12)$$
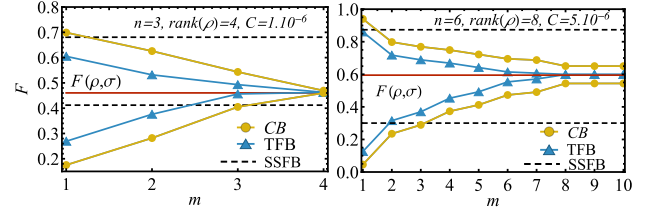

Figure 2: Implementations of VQFE on IBM's quantum computer simulator. The left (right) panel shows bounds on $F(\rho, \sigma)$ (solid straight line) versus m for a random state σ and a tensor product state ρ of $n = 3$ ($n = 6$) qubits with rank$(\rho) = 4$ (rank$(\rho) = 8$). Dashed lines depict the SSFB. In both cases the TFB converge to $F(\rho, \sigma)$. The Certified Bounds (CB) become tighter than the SSFB in both cases, but for $n = 6$ case the certified bounds remain gapped for large m since the cost C is non-negligible.

with $\delta = \min(\delta_1, \delta_2)$, and where $\delta_1 = (4 rank(\rho)C)^{1/4}$ and $\delta_2 = (2\epsilon_m' + \sqrt{2mC})^{1/2}$ with $\epsilon_m' = 1 - \mathrm{Tr}\rho_m'$.

The CCFB show the operational meaning of C, which not only bounds the VQSD eigenvalue and eigenvector error [18] but also the VQFE error. Note that δ_2 is directly computable from the VQSD experimental data, whereas δ_1 is useful if one has a promise that ρ is low rank. Alternatively one can use the following certified bounds based on the triangle inequality.

Proposition 4. *Let $D\left(F(\rho, \sigma)\right)$ be a distance measure that is monotonically decreasing in $F(\rho, \sigma)$. Let ρ, σ, and ρ' be three arbitrary quantum states. Then the following certified triangular inequality bounds (CTIB) hold:*

$$D_m^{LB} \leqslant D\left(F(\rho, \sigma)\right) \leqslant D_m^{UB}, \qquad (13)$$

with

$$D_m^{LB} = D\left(F_*(\rho_m', \sigma_m^{\rho'})\right) - D\left(F(\rho_m', \rho_m^{\rho'})\right), (14)$$

$$D_m^{UB} = D\left(F(\rho_m', \sigma_m^{\rho'})\right) + D\left(F(\rho_m', \rho_m^{\rho'})\right), (15)$$

where ρ_m', $\sigma_m^{\rho'}$, and $\rho_m^{\rho'}$ are projections onto the subspace of the m largest eigenvectors of ρ' analogous to (4).

One obtains bounds on $F(\rho, \sigma)$ from Proposition 4 by inverting D. Moreover, from Proposition 2, one can show that these bounds are refinable: D_m^{LB} monotonically increases in m, and D_m^{UB} monotonically decreases in m. The CTIB are valid for the Bures angle $D_A\left(F(\rho, \sigma)\right) = \arccos F(\rho, \sigma)$, Bures distance $D_B\left(F(\rho, \sigma)\right) = \sqrt{2 - 2F(\rho, \sigma)}$, and the Sine distance $D_S\left(F(\rho, \sigma)\right) = \sqrt{1 - F(\rho, \sigma)^2}$ [33, 34], and hence one can take the metric that gives the tightest bounds. Furthermore, combining Propositions 3 and 4, we use the term Certified Bounds (CB in Fig. 2) to refer to the minimum (maximum) of our certified upper (lower) bounds. In other words, the Certified Bounds are obtained by taking the tighter of the two bounds provided by Propositions 3 and 4.
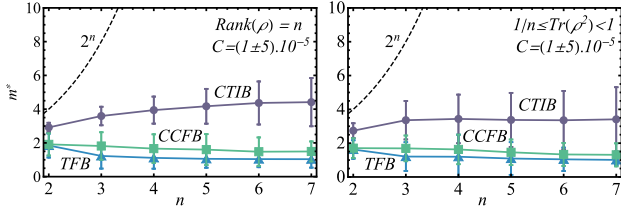
Figure 3: Minimum value of m needed for our bounds (TFB, CCFB, and CTIB) to become tighter than the SSFB (m^*) versus n. Results were averaged over 2000 random states σ (with uniformly distributed rank and purity), and random states ρ with: low rank (rank$(\rho) = n$, left), or high purity ($1/n \leqslant \text{Tr}(\rho^2) < 1$, right). Error bars depict standard deviation. A non-zero cost C was obtained by applying a random unitary close to the identity to the diagonal form of ρ. As n increases, m^* remains $O(n)$ while the dimension grows exponentially (dashed curve).
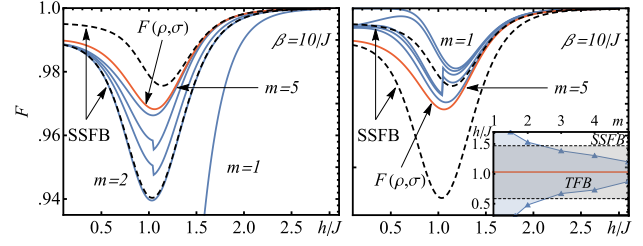

Figure 4: Fidelity spectrum ($m = 1, \ldots, 5$) between thermal states $\rho(h)$ and $\rho(h + \delta h)$ of a cyclic Ising chain with $N = 8$ spins-$1/2$ in a magnetic field h at inverse temperature $\beta = 10/J$, with J the coupling strength and $\delta h = 0.01/J$. Dashed lines indicate the SSFB. The left (right) panel depicts the lower (upper) TFB of (7). The fidelities present a dip at $h \approx 1$ indicating the zero temperature transition of the model. Inset: Bounds for the position of the minimum of the fidelity (and hence the transition) derived from the SSFB and TFB.

4 Implementations

Figure 2 shows our VQFE implementations on IBM's quantum computer simulator. The left and right panels show representative results for $n = 3$ and $n = 6$ qubits, respectively. We used the Constrained Optimization By Linear Approximation (COBYLA) algorithm [35] in the VQSD optimization loop and achieved a cost of $C \sim 10^{-6}$ (see Ref. [36] for a comparison of state-of-the-art optimizers for variational algorithms). We chose σ as a random state and $\rho = \bigotimes_{j=1}^{n} \rho_j$ as a tensor product state, where the latter can be diagonalized with a depth-one quantum circuit ansatz. As one can see, as m increases the TFB rapidly converge to $F(\rho, \sigma)$, and since C is small ($\sim 10^{-6}$), the TFB can be viewed as bounds on $F(\rho, \sigma)$. Nevertheless we also show our Certified Bounds, and in both cases the Certified Bounds are significantly tighter than the SSFB.

5 Heuristic scaling

Let m^* denote the minimum value of m needed for our bounds to become tighter than the SSFB. Figure 3 plots m^* for the TFB, CCFB, and CTIB for systems with $n = 2, \ldots, 7$ qubits. The results were obtained by averaging m^* over 2000 random states ρ and σ. We considered two cases of interest: when ρ is a low-rank state, and when it has an exponentially decaying spectrum leading to full rank but high purity. In both scenarios m^* is $O(n)$ with $m^* \approx 2$ for the TFB and CCFB, implying that our bounds can outperform the SSFB by only considering a number of eigenvalues in the truncated states which do not scale exponentially with n.

6 Example: Quantum phase transition

It has been shown that the fidelity can capture the geometric distance between thermal states of condensed matter systems nearing phase transitions and can provide information about the zero-temperature phase diagram [37–39]. We now demonstrate the application of VQFE to the study of phase transitions. Consider a cyclic Ising chain of $N = 8$ spins-$1/2$ in a uniform field. The Hamiltonian reads $H = -\sum_j h S_j^z + J S_j^x S_{j+1}^x$, with $S_j^{x,z}$ the spin components at site j, h the magnetic field, and J the exchange coupling strength. While this model is exactly solvable, it remains of interest as its thermal states can be exactly prepared on a quantum computer [40, 41].

Figure 4 shows the fidelity spectrum for this example. We first verify that the SSFB and the TFB present a pronounced dip near $h = 1$, implying that they can detect the presence of the zero-temperature transition. Moreover, VQFE allows a better determination of the critical field: The TFB give a range for the transition which monotonically tightens as m increases and outperforms the SSFB already for $m = 3$ (see inset). The TFB also provide information regarding the closeness of eigenvectors (e.g., upper-TFB which are ≈ 1 for $m = 1, 2$), and can detect level crossings where the structure of the subspace spanned by $\{|r_i\rangle\}$ drastically changes. For instance, near $h = 1$ the $m = 3, 4$ TFB present a discontinuity from the crossing between a uniform eigenstate and a pair of exactly degenerate non-uniform symmetry-breaking states. Hence, the fidelity spectrum provides information about the structure of the states beyond the scope of the SSFB or even $F(\rho, \sigma)$.

7 Complexity analysis

Recent dequantization results [42, 43] suggest that exponential speedup over classical algorithms can disappear under low-rank assumptions. Here we show this is not the case for Low-Rank Fidelity Estimation. Low-Rank Fidelity Estimation remains classically hard despite the restriction that one or both of the states are low rank. We formally define Low-Rank Fidelity Estimation as follows.

Problem 1. (LOW-RANK FIDELITY ESTIMATION):
Input: Two $\mathrm{poly}(n)$-sized quantum circuit descriptions U and V that prepare n-qubit states ρ and σ on a subset of qubits all initially in the $|0\rangle$ state, and a parameter γ.
Promise: The state ρ is low rank (i.e., $\mathrm{rank}(\rho) = \mathrm{poly}(n)$) or ϵ-low rank (i.e., $\mathrm{rank}_\epsilon(\rho) = \mathrm{poly}(n)$ for $\epsilon = 1/\mathrm{poly}(n)$).
Output: Approximation of $F(\rho, \sigma)$ up to additive accuracy γ.

We remark that VQFE does not require knowledge of the circuit descriptions of ρ and σ. Rather, allowing for such knowledge is useful when considering classical algorithms.

Next, we analyze the hardness of Low-Rank Fidelity Estimation. Recall that the complexity class DQC1 consists of all problems that can be efficiently solved with bounded error in the one-clean-qubit model of computation [44]. The classical simulation of DQC1 is impossible unless the polynomial hierarchy collapses to the second level [45, 46], which is not believed to be the case. Hence, given the following proposition, one can infer that a classical algorithm *cannot efficiently perform Low-Rank Fidelity Estimation.*

Proposition 5. *The problem* LOW-RANK FIDELITY ESTIMATION *to within precision* $\pm\gamma = 1/\mathrm{poly}(n)$ *is* DQC1-*hard.*

On the other hand, let us now consider the complexity of the VQFE algorithm. The following proposition demonstrates that VQFE is efficient, so long as the VQSD step is efficient.

Proposition 6. *Let ρ and σ be quantum states, and suppose we have access to a subroutine that diagonalizes ρ up to error bounded by $C = D_{HS}(\rho, \rho')$. Let m and ζ be parameters. Then, VQFE runs in time $\widetilde{O}(m^6/\zeta^2)$ and outputs an additive $\pm\gamma$-approximation of $F(\rho, \sigma)$, for*

$$\gamma = (\epsilon + \sqrt{mC})^{1/2} + (2\epsilon + (2+\sqrt{2})\sqrt{mC})^{1/2} + \zeta\|T\|, \tag{16}$$

and where ϵ is determined by $m = \mathrm{rank}_\epsilon(\rho)$.

Consider the implications of Proposition 6 when ρ is ϵ-low rank, meaning that $\epsilon = 1/\mathrm{poly}(n)$ for $m =$

$\mathrm{poly}(n)$. In this case, VQFE can solve Low-Rank Fidelity Estimation with precision $\gamma = 1/\mathrm{poly}(n)$, assuming that the diagonalization error C is small enough such that $mC = 1/\mathrm{poly}(n)$ and also that $\zeta = 1/\mathrm{poly}(n)$, which suffices since $\|T\| \leqslant 1$. Under these assumptions, the run time of VQFE will be $\mathrm{poly}(n)$.

8 Conclusion

In this work, we introduced novel bounds on the fidelity based on truncating the spectrum of one of the states, and we proposed a hybrid quantum-classical algorithm to compute these bounds. We furthermore showed that our bounds typically outperform the sub- and super-fidelities, and they are also useful for detecting quantum phase transitions. Our algorithm will likely find use in verifying and characterizing quantum states prepared on a quantum computer.

Let us recall that we have proposed employing maximum likelihood methods to reconstruct the T matrix when noise makes it non-positive. However, in the context of state tomography, such procedures can can lead to biases that change the entanglement properties of the estimated state [47, 48]. Hence, we leave for future work the analysis of the effect that such reconstruction methods can have on our bounds.

A strength of VQFE is that it does not require access to purifications of ρ and σ, although future research could study whether having such access could simplify the estimation of $F(\rho, \sigma)$ (e.g., using Uhlmann's theorem [6]). Another important future direction is to use VQFE to compute distance between quantum operations, e.g., as in [49]. In addition it would be of interest to extend VQFE to the sandwiched Renyi relative entropies $D_\alpha(\rho||\sigma)$ [50, 51], defined by:

$$2^{(\alpha-1)D_\alpha(\rho||\sigma)} = \mathrm{Tr}\left(\rho^{\frac{1-\alpha}{2\alpha}}\sigma\rho^{\frac{1-\alpha}{2\alpha}}\right)^\alpha. \tag{17}$$

Note that $\alpha = 1/2$ corresponds to $F(\rho, \sigma)$. Expanding (17) in the eigenbasis of ρ, one can see that $D_\alpha(\rho||\sigma)$ depends only on $\{r_i\}$ and $\{\sigma_{ij}\}$, i.e., the same quantities used in VQFE. Hence it may be possible to use the same strategy as in VQFE to estimate the entire Renyi relative entropy family.

9 Acknowledgements

References

[1] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'Brien. A variational eigenvalue solver on a photonic quantum processor. *Nature Communications*, 5:4213, 2014. DOI: https://doi.org/10.1038/ncomms5213.

[2] Katherine L Brown, William J Munro, and Vivien M Kendon. Using quantum computers for quantum simulation. *Entropy*, 12(11):2268–2307, 2010. DOI: https://doi.org/10.3390/e12112268.

[3] Aram W Harrow, Avinatan Hassidim, and Seth Lloyd. Quantum algorithm for linear systems of equations. *Physical review letters*, 103(15):150502, 2009. DOI: https://doi.org/10.1103/PhysRevLett.103.150502.

[4] Armin Uhlmann. The "transition probability" in the state space of a⋆-algebra. *Reports on Mathematical Physics*, 9(2):273–279, 1976. DOI: https://doi.org/10.1016/0034-4877(76)90060-4.

[5] Richard Jozsa. Fidelity for mixed quantum states. *Journal of modern optics*, 41(12):2315–2323, 1994. DOI: https://doi.org/10.1080/09500349414552171.

[6] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information.* December 2010. DOI: https://doi.org/10.1017/CBO9780511976667.

[7] M. M. Wilde. *Quantum Information Theory.* Cambridge University Press, 2 edition, 2017. DOI: https://doi.org/10.1017/9781316809976.

[8] Yeong-Cherng Liang, Yu-Hao Yeh, Paulo E M F Mendonça, Run Yan Teh, Margaret D Reid, and Peter D Drummond. Quantum fidelity measures for mixed states. *Reports on Progress in Physics*, 82(7):076001, jun 2019. DOI: https://doi.org/10.1088/1361-6633/ab1ca4.

[9] Markus Hauru and Guifre Vidal. Uhlmann fidelities from tensor networks. *Physical Review A*, 98(4):042316, 2018. DOI: https://doi.org/10.1103/PhysRevA.98.042316.

[10] John Watrous. Quantum statistical zero-knowledge. *arXiv:quant-ph/0202111*, 2002. URL https://arxiv.org/abs/quant-ph/0202111.

[11] Hui Li and F. D. M. Haldane. Entanglement spectrum as a generalization of entanglement entropy: Identification of topological order in non-abelian fractional quantum hall effect states. *Phys. Rev. Lett.*, 101:010504, Jul 2008. DOI: https://doi.org/10.1103/PhysRevLett.101.010504.

[12] Christopher M Bishop. *Pattern recognition and machine learning.* Springer, 2006. URL https://www.springer.com/gp/book/9780387310732.

[13] Edward Farhi, Jeffrey Goldstone, and Sam Gutmann. A quantum approximate optimization algorithm. *arXiv:1411.4028*, 2014. URL https://arxiv.org/abs/1411.4028.

[14] J. R. McClean, J. Romero, R. Babbush, and A. Aspuru-Guzik. The theory of variational hybrid quantum-classical algorithms. *New Journal of Physics*, 18(2):023023, 2016. DOI: https://doi.org/10.1088/1367-2630/18/2/023023.

[15] J. Romero, J. P. Olson, and A. Aspuru-Guzik. Quantum autoencoders for efficient compression of quantum data. *Quantum Science and Technology*, 2:045001, December 2017. DOI: https://doi.org/10.1088/2058-9565/aa8072.

[16] Ying Li and Simon C Benjamin. Efficient variational quantum simulator incorporating active error minimization. *Physical Review X*, 7(2):021050, 2017. DOI: https://doi.org/10.1103/PhysRevX.7.021050.

[17] S. Khatri, R. LaRose, A. Poremba, L. Cincio, A. T. Sornborger, and P. J. Coles. Quantum-assisted quantum compiling. *Quantum*, 3: 140, May 2019. ISSN 2521-327X. DOI: https://doi.org/10.22331/q-2019-05-13-140.

[18] Ryan LaRose, Arkin Tikku, Étude O'Neel-Judy, Lukasz Cincio, and Patrick J Coles. Variational quantum state diagonalization. *npj Quantum Information*, 5(1):8, 2019. DOI: https://doi.org/10.1038/s41534-019-0167-6.

[19] A. Arrasmith, L. Cincio, A. T. Sornborger, W. H. Zurek, and P. J. Coles. Variational consistent histories as a hybrid algorithm for quantum foundations. *Nature communications*, 10(1):3438, 2019. DOI: https://doi.org/10.1038/s41467-019-11417-0.

[20] Kunal Sharma, Sumeet Khatri, M Cerezo, and Patrick J Coles. Noise resilience of variational quantum compiling. *New Journal of Physics*, 2020. DOI: https://doi.org/10.1088/1367-2630/ab784c.

[21] Cristina Cirstoiu, Zoe Holmes, Joseph Iosue, Lukasz Cincio, Patrick J Coles, and Andrew Sornborger. Variational fast forwarding for quantum simulation beyond the coherence time. *arXiv:1910.04292*, 2019. URL https://arxiv.org/abs/1910.04292.

[22] Jaroslaw Adam Miszczak, Zbigniew Puchala, Pawel Horodecki, Armin Uhlmann, and Karol Zyczkowski. Sub- and super-fidelity as bounds for quantum fidelity. *Quantum Information*

& Computation, 9(1):103–130, 2009. DOI: https://doi.org/10.26421/QIC9.1-2.

[23] Jing-Ling Chen, Libin Fu, Abraham A Ungar, and Xian-Geng Zhao. Alternative fidelity measure between two states of an n-state quantum system. Physical Review A, 65(5):054304, 2002. DOI: https://doi.org/10.1103/PhysRevA.65.054304.

[24] Paulo EMF Mendonça, Reginaldo d J Napolitano, Marcelo A Marchiolli, Christopher J Foster, and Yeong-Cherng Liang. Alternative fidelity measure between quantum states. Physical Review A, 78(5):052330, 2008. DOI: https://doi.org/10.1103/PhysRevA.78.052330.

[25] Zbigniew Puchała and Jarosław Adam Miszczak. Bound on trace distance based on superfidelity. Phys. Rev. A, 79:024302, Feb 2009. DOI: https://doi.org/10.1103/PhysRevA.79.024302.

[26] Todd A Brun. Measuring polynomial functions of states. Quantum Information & Computation, 4(5):401–408, 2004. URL https://arxiv.org/abs/quant-ph/0401067.

[27] Karol Bartkiewicz, Karel Lemr, and Adam Miranowicz. Direct method for measuring of purity, superfidelity, and subfidelity of photonic two-qubit mixed states. Phys. Rev. A, 88:052104, Nov 2013. DOI: https://doi.org/10.1103/PhysRevA.88.052104.

[28] L. Cincio, Y. Subaşı, A. T. Sornborger, and P. J. Coles. Learning the quantum algorithm for state overlap. New J. Phys., 20:113022, Nov 2018. DOI: https://doi.org/10.1088/1367-2630/aae94a.

[29] J. C. Garcia-Escartin and P. Chamorro-Posada. Swap test and Hong-Ou-Mandel effect are equivalent. Phys. Rev. A, 87:052330, May 2013. DOI: https://doi.org/10.1103/PhysRevA.87.052330.

[30] Marco Tomamichel, Roger Colbeck, and Renato Renner. Duality between smooth min- and max-entropies. IEEE Transactions on Information Theory, 56(9):4674–4681, 2010. DOI: https://doi.org/10.1109/TIT.2010.2054130.

[31] Marco Tomamichel. Quantum Information Processing with Finite Resources: Mathematical Foundations, volume 5. Springer, 2015. DOI: https://doi.org/10.1007/978-3-319-21891-5.

[32] John A. Smolin, Jay M. Gambetta, and Graeme Smith. Efficient method for computing the maximum-likelihood quantum state from measurements with additive gaussian noise. Phys. Rev. Lett., 108:070502, Feb 2012. DOI: https://doi.org/10.1103/PhysRevLett.108.070502.

[33] Alexei Gilchrist, Nathan K. Langford, and Michael A. Nielsen. Distance measures to compare real and ideal quantum processes. Phys. Rev. A, 71:062310, Jun 2005. DOI: https://doi.org/10.1103/PhysRevA.71.062310.

[34] Alexey E Rastegin. Sine distance for quantum states. quant-ph/0602112, 2006. URL https://arxiv.org/abs/quant-ph/0602112.

[35] M J. D. Powell. The bobyqa algorithm for bound constrained optimization without derivatives. Technical Report, Department of Applied Mathematics and Theoretical Physics, 01 2009. URL http://www.damtp.cam.ac.uk/user/na/NA_papers/NA2009_06.pdf.

[36] Jonas M Kübler, Andrew Arrasmith, Lukasz Cincio, and Patrick J Coles. An adaptive optimizer for measurement-frugal variational algorithms. arXiv preprint arXiv:1909.09083, 2019. URL https://arxiv.org/abs/1909.09083.

[37] Paolo Zanardi, H. T. Quan, Xiaoguang Wang, and C. P. Sun. Mixed-state fidelity and quantum criticality at finite temperature. Phys. Rev. A, 75:032109, Mar 2007. DOI: https://doi.org/10.1103/PhysRevA.75.032109.

[38] Paolo Zanardi, Paolo Giorda, and Marco Cozzini. Information-theoretic differential geometry of quantum phase transitions. Phys. Rev. Lett., 99:100603, Sep 2007. DOI: https://doi.org/10.1103/PhysRevLett.99.100603.

[39] H. T. Quan and F. M. Cucchietti. Quantum fidelity and thermal phase transitions. Phys. Rev. E, 79:031101, Mar 2009. DOI: https://doi.org/10.1103/PhysRevE.79.031101.

[40] Frank Verstraete, J Ignacio Cirac, and José I Latorre. Quantum circuits for strongly correlated quantum systems. Physical Review A, 79(3):032316, 2009. DOI: https://doi.org/10.1103/PhysRevA.79.032316.

[41] Alba Cervera-Lierta. Exact Ising model simulation on a quantum computer. Quantum, 2: 114, December 2018. ISSN 2521-327X. DOI: https://doi.org/10.22331/q-2018-12-21-114.

[42] Ewin Tang. A quantum-inspired classical algorithm for recommendation systems. pages 217–228, 2019. DOI: https://doi.org/10.1145/3313276.3316310.

[43] András Gilyén, Seth Lloyd, and Ewin Tang. Quantum-inspired low-rank stochastic regression with logarithmic dependence on the dimension. arXiv preprint arXiv:1811.04909, 2018. URL https://arxiv.org/abs/1811.04909.

[44] E. Knill and R. Laflamme. Power of one bit of quantum information. Physical Review Letters, 81:5672–5675, 1998. DOI: https://doi.org/10.1103/PhysRevLett.81.5672.

[45] K. Fujii, H. Kobayashi, T. Morimae, H. Nishimura, S. Tamate, and S. Tani. Impossibility of Classically Simulating One-Clean-Qubit Model with Multiplicative Error. Physical Review Letters, 120:200502, May 2018. DOI: https://doi.org/10.1103/PhysRevLett.120.200502.

[46] Tomoyuki Morimae. Hardness of classically sampling one clean qubit model with constant

total variation distance error. 2017. DOI: https://doi.org/10.1103/PhysRevA.96.040302.

[47] Christian Schwemmer, Lukas Knips, Daniel Richart, Harald Weinfurter, Tobias Moroder, Matthias Kleinmann, and Otfried Gühne. Systematic errors in current quantum state tomography tools. *Physical review letters*, 114(8):080403, 2015. DOI: https://doi.org/10.1103/PhysRevLett.114.080403.

[48] Gael Sentís, Johannes N Greiner, Jiangwei Shang, Jens Siewert, and Matthias Kleinmann. Bound entangled states fit for robust experimental verification. *Quantum*, 2:113, 2018. DOI: https://doi.org/10.22331/q-2018-12-18-113.

[49] Zbigniew Puchała, Jarosław Adam Miszczak, Piotr Gawron, and Bartłomiej Gardas. Experimentally feasible measures of distance between quantum operations. *Quantum Information Processing*, 10(1):1–12, 2011. DOI: https://doi.org/10.1007/s11128-010-0166-1.

[50] Martin Müller-Lennert, Frédéric Dupuis, Oleg Szehr, Serge Fehr, and Marco Tomamichel. On quantum rényi entropies: A new generalization and some properties. *Journal of Mathematical Physics*, 54(12):122203, 2013. DOI: https://doi.org/10.1063/1.4838856.

[51] Mark M Wilde, Andreas Winter, and Dong Yang. Strong converse for the classical capacity of entanglement-breaking and hadamard channels via a sandwiched rényi relative entropy. *Communications in Mathematical Physics*, 331(2):593–622, 2014. DOI: https://doi.org/10.1007/s00220-014-2122-x.

[52] Patrick J. Coles, M. Cerezo, and Lukasz Cincio. Strong bound between trace distance and hilbert-schmidt distance for low-rank states. *Phys. Rev. A*, 100:022103, Aug 2019. DOI: https://doi.org/10.1103/PhysRevA.100.022103.

[53] Bernhard Baumgartner. An inequality for the trace of matrix products, using absolute values. *arXiv:1106.6189*, 2011. URL https://arxiv.org/abs/1106.6189.

[54] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, 1990. DOI: https://doi.org/10.1017/CBO9780511810817.

[55] H. Weyl. Das asymptotische verteilungsgesetz der eigenwerte linearer partieller differentialgleichungen (mit einer anwendung auf die theorie der hohlraumstrahlung). *Mathematische Annalen*, 71(4):441–479, 1912. DOI: https://doi.org/10.1007/BF01456804.

[56] A. S. Kholevo. On quasiequivalence of locally normal states. *Theoretical and Mathematical Physics*, 13(2):1071–1082, Nov 1972. ISSN 1573-9333. DOI: https://doi.org/10.1007/BF01035528.

[57] Koenraad M. R. Audenaert. Comparisons between quantum state distinguishability measures. *Quantum Info. Comput.*, 14(1–2):31–38, January 2014. ISSN 1533-7146. DOI: https://doi.org/10.26421/QIC14.1-2.

<div align="center">

Appendix

</div>

Here we provide detailed proofs of the Propositions presented in the main text. In order to make the results clearer we also reiterate these Propositions.

A Proof of Proposition 1

Proposition 1. *The following truncated fidelity bounds (TFB) hold:*

$$F(\rho_m, \sigma_m^\rho) \leqslant F(\rho, \sigma) \leqslant F_*(\rho_m, \sigma_m^\rho). \tag{1}$$

Proof. The right-hand-side of (1) follows from the fact that the generalized fidelity is monotonous under completely positive trace non-increasing (CPTNI) maps [30, 31]. In particular, we can define the CPTNI map

$$\mathcal{E}_m(\rho) = \Pi_m^\rho \rho \Pi_m^\rho, \tag{2}$$

with Π_m^ρ the projector onto the subspace spanned by the eigenvectors with m-largest eigenvalues, which leads to

$$\begin{aligned} F(\rho, \sigma) = F_*(\rho, \sigma) &\leqslant F_*(\mathcal{E}_m(\rho), \mathcal{E}_m(\sigma)) \\ &= F_*(\rho_m, \sigma_m^\rho). \end{aligned} \tag{3}$$

The lower bound is derived from the strong concavity of the fidelity [6], as follows:

$$\begin{aligned} F(\rho, \sigma) = F(\Pi_m^\rho \rho \Pi_m^\rho + \overline{\Pi}_m^\rho \rho \overline{\Pi}_m^\rho, \sigma) &\geqslant \sqrt{p_m} F\left(\frac{\Pi_m^\rho \rho \Pi_m^\rho}{p_m}, \sigma\right) + \sqrt{1 - p_m} F\left(\frac{\overline{\Pi}_m^\rho \rho \overline{\Pi}_m^\rho}{1 - p_m}, \sigma\right) \\ &\geqslant \sqrt{p_m} F\left(\frac{\rho_m}{p_m}, \sigma\right) = F(\rho_m, \sigma) = F(\rho_m, \sigma_m^\rho), \end{aligned} \tag{4}$$

where $p_m = \mathrm{Tr}\rho_m$ and where we have expressed $\rho = \Pi_m^\rho \rho \Pi_m^\rho + \overline{\Pi}_m^\rho \rho \overline{\Pi}_m^\rho$ with $\overline{\Pi}_m^\rho$ the orthogonal complement of Π_m^ρ. $\qquad\square$

B Proof of Proposition 2

Proposition 2. *The truncated fidelity $F(\rho_m, \sigma_m^\rho)$ is monotonically increasing in m, and the truncated generalized fidelity $F_*(\rho_m, \sigma_m^\rho)$ is monotonically decreasing in m.*

Proof. The monotonicity of the truncated fidelity follows from the same procedure as the one used to derived (4). Namely, we expand $\rho_m = \Pi_{m-1}^\rho \rho \Pi_{m-1}^\rho + |r_m\rangle \rho_{mm} \langle r_m|$, with $\rho_{mm} = \langle r_m|\rho|r_m\rangle$, and we define $\hat{\rho}_m = \rho_m/p_m$ with $p_m = \mathrm{Tr}\rho_m$. Applying strong concavity gives

$$\begin{aligned} F(\rho_m, \sigma_m^\rho) = \sqrt{p_m} F(\hat{\rho}_m, \sigma) &\geqslant \sqrt{p_m} \left(\sqrt{p_{m-1}/p_m} F\left(\frac{\Pi_{m-1}^\rho \rho \Pi_{m-1}^\rho}{p_{m-1}}, \sigma\right) + \sqrt{\rho_{mm}/p_m} F\left(|r_m\rangle\langle r_m|, \sigma\right)\right) \\ &\geqslant F(\rho_{m-1}, \sigma) = F(\rho_{m-1}, \sigma_{m-1}^\rho). \end{aligned} \tag{5}$$

On the other hand, since $\rho_m = \mathcal{E}_m(\rho_{m+1})$ and $\sigma_m^\rho = \mathcal{E}_m(\sigma_{m+1}^\rho)$, then by the monotonocity under CPTNI maps of the generalized fidelity we get $F_*(\rho_{m+1}, \sigma_{m+1}^\rho) \leqslant F_*(\rho_m, \sigma_m^\rho)$. $\qquad\square$

C Proof of Proposition 3

Before proving Proposition 3, we first prove four useful lemmas. Of particular interest is Lemma 4 below since it generalizes the main result of [52] (corresponding to the special case of $\epsilon = 0$). All of these lemmas employ our notion of ϵ-rank. We remind the reader that this is defined as follows:

$$\mathrm{rank}_\epsilon(\rho) = \min\{m \in \{1, ..., d\} : \|\rho - \rho_m\|_1 \leqslant \epsilon\}, \tag{6}$$

where d is the Hilbert-space dimension.

Lemma 1. *For any positive semi-definite operators A, B, and C, where C is normalized ($\text{Tr}(C) = 1$), we have*

$$|F(A,C) - F(B,C)| \leqslant \sqrt{\text{Tr}\left[\left(\sqrt{A} - \sqrt{B}\right)^2\right]}. \tag{7}$$

Proof. We first write the left-hand-side as

$$\Delta F = F(A,C) - F(B,C) = \left\|\sqrt{A}\sqrt{C}\right\|_1 - \left\|\sqrt{B}\sqrt{C}\right\|_1. \tag{8}$$

It is well known that given a $d \times d$ matrix M with singular value decomposition $M = U_0 S U_1$, then $\|M\|_1 = |\max_{V \in \mathcal{U}_d} \text{Tr}(MV)| = \text{Tr}(MV_{\text{opt}})$ with $\mathcal{U}_d$ the set of $d \times d$ unitary matrices and $V_{\text{opt}} = U_1^\dagger U_0^\dagger$. Hence, for every pair of $d \times d$ positive semi-definite operators A and C there exists V_{opt} such that $F(A,C) = |\max_{V \in \mathcal{U}_d} \text{Tr}(\sqrt{A}\sqrt{C}V)| = |\text{Tr}(\sqrt{A}\sqrt{C}V_{\text{opt}})|$. (We remark that in the last expression the absolute value is not necessary but we include it for clarity.) Let V_{opt} and W_{opt} be the optimal unitaries for $F(A,C)$ and $F(B,C)$, respectively. Then (8) can be expressed as

$$\Delta F = \left|\text{Tr}(\sqrt{A}\sqrt{C}V_{\text{opt}})\right| - \left|\text{Tr}(\sqrt{B}\sqrt{C}W_{\text{opt}})\right|, \tag{9}$$

$$\leqslant \left|\text{Tr}(\sqrt{A}\sqrt{C}V_{\text{opt}})\right| - \left|\text{Tr}(\sqrt{B}\sqrt{C}V_{\text{opt}})\right|. \tag{10}$$

$$\leqslant \left|\text{Tr}(\sqrt{A}\sqrt{C}V_{\text{opt}}) - \text{Tr}(\sqrt{B}\sqrt{C}V_{\text{opt}})\right|, \tag{11}$$

$$= \left|\text{Tr}\left[(\sqrt{A} - \sqrt{B})\sqrt{C}V_{\text{opt}}\right]\right|. \tag{12}$$

where we have used $|\text{Tr}\sqrt{B}\sqrt{C}W_{\text{opt}}| \geqslant |\text{Tr}\sqrt{B}\sqrt{C}V_{\text{opt}}|$ and the triangular inequality for the absolute value. Then, by means of the matrix Hölder inequality [53] and the fact that C is normalized, we obtain

$$\Delta F \leqslant \sqrt{\text{Tr}\left[(\sqrt{C}V_{\text{opt}})^\dagger \sqrt{C}V_{\text{opt}}\right]}\sqrt{\text{Tr}\left[\left(\sqrt{A} - \sqrt{B}\right)^2\right]} = \sqrt{\text{Tr}\left[\left(\sqrt{A} - \sqrt{B}\right)^2\right]}. \tag{13}$$

By symmetry, one can replace ΔF with $-\Delta F$ in (13). So (13) also holds for $|\Delta F|$. $\qquad\square$

Lemma 2. *Let ρ be a quantum state and let $\rho' = U^\dagger \mathcal{Z}(U \rho U^\dagger) U$ be its VQSD approximation defined in (10) of the main text. Then*

$$rank_\epsilon(\rho) \leqslant rank_\epsilon(\rho'). \tag{14}$$

Proof. Let us first define $\tilde{\rho}_Z = \mathcal{Z}(\tilde{\rho})$, with $\tilde{\rho} = U \rho U^\dagger$ and $\mathcal{Z}$ the dephasing channel in the computational basis, such that $\rho' = U^\dagger \tilde{\rho}_Z U$. Due to Schur-Horn's theorem we have that the eigenvalues of $\tilde{\rho}$ majorize its diagonal elements, implying that $\tilde{\rho}$ majorizes $\tilde{\rho}_Z$. Moreover, since $\tilde{\rho}$ has the same eigenvalues as ρ, and ρ' has the same eigenvalues as $\tilde{\rho}_Z$, then we have that ρ majorizes ρ'. In summary,

$$\tilde{\rho} \succ \tilde{\rho}_Z, \quad \text{and} \quad \rho \succ \rho'. \tag{15}$$

Next, let us define $\hat{m}' = \text{rank}_\epsilon(\rho')$. Since $\rho \succ \rho'$, we have $\sum_{j \leqslant \hat{m}'} r_j \geqslant \sum_{j \leqslant \hat{m}'} r_j'$ and hence

$$\|\rho - \rho_{\hat{m}'}\|_1 = 1 - \sum_{j \leqslant \hat{m}'} r_j \leqslant 1 - \sum_{j \leqslant \hat{m}'} r_j' = \|\rho' - \rho_{\hat{m}'}'\|_1 \leqslant \epsilon. \tag{16}$$

The fact that $\|\rho - \rho_{\hat{m}'}\|_1 \leqslant \epsilon$ implies that $\text{rank}_\epsilon(\rho) \leqslant \hat{m}'$ and hence

$$\text{rank}_\epsilon(\rho) \leqslant \text{rank}_\epsilon(\rho'). \tag{17}$$

$\qquad\square$

Lemma 3. *Let $\Delta = \rho - \sigma = \Delta^+ - \Delta^-$. Here Δ^+ and Δ^- respectively correspond to the positive and negative part of Δ, with $\Delta^+ \geqslant 0$, $\Delta^- \geqslant 0$, and $\Delta^+ \Delta^- = 0$. Then*

$$rank_\epsilon(\Delta^+) \leqslant rank_\epsilon(\rho), \quad \text{and} \quad rank_\epsilon(\Delta^-) \leqslant rank_\epsilon(\sigma). \tag{18}$$

Proof. Let $\{r_j\}$, $\{s_j\}$, and $\{\delta_j\}$ respectively denote the eigenvalues of ρ, σ, and Δ, where the eigenvalues in each set are in decreasing order. By means of Weyl's inequality [54, 55] applied to $\rho = \Delta + \sigma$ we get

$$r_j \geqslant \delta_j + s_d, \quad \forall j. \tag{19}$$

Because $\sigma \geqslant 0$, we have $s_d \geqslant 0$, and hence

$$r_j \geqslant \delta_j, \quad \forall j. \tag{20}$$

Since Δ^+ is the positive part of Δ, then from the previous equation its eigenvalues (which we denote $\{\delta_j^+\}$ and which are in decreasing order) are such that

$$r_j \geqslant \delta_j^+ \geqslant 0, \quad \forall j. \tag{21}$$

Let us now define $\hat{m} = \mathrm{rank}_\epsilon(\rho)$. By means of the definition of ϵ-rank (6), we have that

$$\epsilon \geqslant \|\rho - \rho_{\hat{m}}\|_1 = \sum_{j > \hat{m}} r_j. \tag{22}$$

Moreover, from (21) we get $\sum_{j > \hat{m}} r_j \geqslant \sum_{j > \hat{m}} \delta_j^+ = \|\Delta^+ - \Delta_{\hat{m}}^+\|_1$, which gives

$$\epsilon \geqslant \|\rho - \rho_{\hat{m}}\|_1 \geqslant \|\Delta^+ - \Delta_{\hat{m}}^+\|_1. \tag{23}$$

This implies that $\mathrm{rank}_\epsilon(\Delta^+) \leqslant \hat{m}$ and hence that $\mathrm{rank}_\epsilon(\Delta^+) \leqslant \mathrm{rank}_\epsilon(\rho)$.

Similarly we denote the set of eigenvalues of Δ^- as $\{\delta_j^-\}$. From Weyl's inequality we now have $s_j \geqslant \delta_j^-$, $\forall j$. Let $\tilde{m} = \mathrm{rank}_\epsilon(\sigma)$ such that $\epsilon \geqslant \|\sigma - \sigma_{\tilde{m}}\|_1 = \sum_{j > \tilde{m}} s_j$. Since $\sum_{j > \tilde{m}} s_j \geqslant \sum_{j > \tilde{m}} \delta_j^-$, we then get

$$\epsilon \geqslant \|\sigma - \sigma_{\tilde{m}}\|_1 \geqslant \|\Delta^- - \Delta_{\tilde{m}}^-\|_1, \tag{24}$$

which implies $\mathrm{rank}_\epsilon(\Delta^-) \leqslant \tilde{m}$ and hence $\mathrm{rank}_\epsilon(\Delta^-) \leqslant \mathrm{rank}_\epsilon(\sigma)$. $\qquad\square$

Lemma 4. *Let $D_T(\rho, \sigma) = \frac{1}{2}\mathrm{Tr}|\rho - \sigma|$ denote the trace distance and $D_{HS}(\rho, \sigma)$ the Hilbert-Schmidt distance, then*

$$(D_T(\rho, \sigma) - \epsilon)^2 \leqslant R_\epsilon(\rho, \sigma) D_{HS}(\rho, \sigma), \tag{25}$$

where $R_\epsilon(\rho, \sigma) = rank_\epsilon(\rho) rank_\epsilon(\sigma)/(rank_\epsilon(\rho) + rank_\epsilon(\sigma))$.

Proof. Let $\Delta_{m_+}^+$ be a truncated version of Δ^+ obtained by projecting Δ^+ onto the subspace associated with its m_+-largest eigenvalues such that $m_+ = \mathrm{rank}_\epsilon(\Delta^+)$. We define $\Delta_{m_-}^-$ analogously, with $m_- = \mathrm{rank}_\epsilon(\Delta^-)$. Then, consider the following density matrices (positive semidefinite matrices with trace one):

$$\tau_{m_+}^+ = \Delta_{m_+}^+/\mathrm{Tr}(\Delta_{m_+}^+), \qquad \tau_{m_-}^- = \Delta_{m_-}^-/\mathrm{Tr}(\Delta_{m_-}^-). \tag{26}$$

Since the purity of a density matrix is lower bounded by the inverse of its rank, we have

$$\mathrm{Tr}\left[(\tau_{m_+}^+)^2\right] \geqslant 1/\mathrm{rank}(\Delta_{m_+}^+) = 1/m_+ = 1/\mathrm{rank}_\epsilon(\Delta^+) \geqslant 1/\mathrm{rank}_\epsilon(\rho), \tag{27}$$

$$\mathrm{Tr}\left[(\tau_{m_-}^-)^2\right] \geqslant 1/\mathrm{rank}(\Delta_{m_-}^-) = 1/m_- = 1/\mathrm{rank}_\epsilon(\Delta^-) \geqslant 1/\mathrm{rank}_\epsilon(\sigma), \tag{28}$$

where the right inequality follows from Lemma 3. Inserting (26) and summing the two resulting inequalities gives:

$$\mathrm{Tr}\left[(\Delta_{m_+}^+)^2\right] + \mathrm{Tr}\left[(\Delta_{m_-}^-)^2\right] \geqslant \frac{\mathrm{Tr}(\Delta_{m_+}^+)^2}{\mathrm{rank}_\epsilon(\rho)} + \frac{\mathrm{Tr}(\Delta_{m_-}^-)^2}{\mathrm{rank}_\epsilon(\sigma)}. \tag{29}$$

Let us first consider the left-hand-side of (29). Since $\mathrm{Tr}\left[(\Delta^+)^2\right] \geqslant \mathrm{Tr}\left[(\Delta_{m_+}^+)^2\right]$ and $\mathrm{Tr}\left[(\Delta^-)^2\right] \geqslant \mathrm{Tr}\left[(\Delta_{m_-}^-)^2\right]$, and using the fact that $D_{\mathrm{HS}}(\rho, \sigma) = \mathrm{Tr}\left[(\Delta^+)^2\right] + \mathrm{Tr}\left[(\Delta^-)^2\right]$, we obtain

$$D_{\mathrm{HS}}(\rho, \sigma) \geqslant \mathrm{Tr}\left[(\Delta_{m_+}^+)^2\right] + \mathrm{Tr}\left[(\Delta_{m_-}^-)^2\right]. \tag{30}$$

On the other hand, since $m_+ = \mathrm{rank}_\epsilon(\Delta^+)$ and $m_- = \mathrm{rank}_\epsilon(\Delta^-)$, then from the definition of ϵ-rank we get

$$\|\Delta^+ - \Delta_{m_+}^+\|_1 = \mathrm{Tr}(\Delta^+) - \mathrm{Tr}(\Delta_{m_+}^+) \leqslant \epsilon, \qquad \|\Delta^- - \Delta_{m_-}^-\|_1 = \mathrm{Tr}(\Delta^-) - \mathrm{Tr}(\Delta_{m_-}^-) \leqslant \epsilon. \tag{31}$$

Rewriting this and using $D(\rho, \sigma) = \text{Tr}(\Delta^+) = \text{Tr}(\Delta^-)$ gives

$$\text{Tr}(\Delta_{m_+}^+) \geqslant \text{Tr}(\Delta^+) - \epsilon = D(\rho, \sigma) - \epsilon, \qquad \text{Tr}(\Delta_{m_-}^-) \geqslant \text{Tr}(\Delta^-) - \epsilon = D(\rho, \sigma) - \epsilon. \qquad (32)$$

Hence, combining (29) with (30) and (32), we have

$$D_{\text{HS}}(\rho, \sigma) \geqslant (D_T(\rho, \sigma) - \epsilon)^2 \left(\frac{1}{\text{rank}_\epsilon(\rho)} + \frac{1}{\text{rank}_\epsilon(\sigma)} \right), \qquad (33)$$

which is equivalent to (25). $\qquad \square$

Proposition 3. *The following certified cost function bounds (CCFB) hold:*

$$F(\rho_m', \sigma_m^{\rho'}) - \delta \leqslant F(\rho, \sigma) \leqslant F_*(\rho_m', \sigma_m^{\rho'}) + \delta, \qquad (34)$$

with $\delta = \min(\delta_1, \delta_2)$, and where $\delta_1 = (4\text{rank}(\rho)C)^{1/4}$ and $\delta_2 = (2\epsilon_m' + \sqrt{2mC})^{1/2}$ with $\epsilon_m' = 1 - \text{Tr}\rho_m'$.

Proof. Let us first derive the δ_1 bound. By means of Lemma 1 we have $|F(\rho, \sigma) - F(\rho', \sigma)| \leqslant \sqrt{\text{Tr}\left[\left(\sqrt{\rho} - \sqrt{\rho'} \right)^2 \right]} = \sqrt{2 \left(1 - \text{Tr}\sqrt{\rho}\sqrt{\rho'} \right)}$. Then from the inequality $1 - \text{Tr}\sqrt{\rho}\sqrt{\sigma} \leqslant D_T(\rho, \sigma)$ [56, 57], we get

$$|F(\rho, \sigma) - F(\rho', \sigma)| \leqslant \sqrt{2D_T(\rho, \rho')}. \qquad (35)$$

Moreover, taking $\epsilon = 0$ in Lemma 4 results in the bound between the trace distance and Hilbert-Schmidt distance from [52]:

$$D_T(\rho, \rho')^2 \leqslant R(\rho, \rho') \cdot D_{\text{HS}}(\rho, \rho'), \qquad (36)$$

with $R(\rho, \sigma)$ the reduced rank defined as $R(\rho, \sigma) = \text{rank}(\rho)\text{rank}(\sigma)/(\text{rank}(\rho) + \text{rank}(\sigma))$. By noting that $R(\rho, \rho') \leqslant \text{rank}(\rho)$ and recalling that $C = D_{\text{HS}}(\rho, \rho')$ we find

$$|F(\rho, \sigma) - F(\rho', \sigma)| \leqslant (4\text{rank}(\rho)C)^{1/4}. \qquad (37)$$

Applying the TFB of Proposition 1 to $F(\rho', \sigma)$ yields

$$F(\rho_m', \sigma_m^{\rho'}) \leqslant F(\rho', \sigma) \leqslant F_*(\rho_m', \sigma_m^{\rho'}). \qquad (38)$$

Combining this result with (37) gives the δ_1 bound in (34), i.e.,

$$F(\rho_m', \sigma_m^{\rho'}) - (4\text{rank}(\rho)C)^{1/4} \leqslant F(\rho, \sigma) \leqslant F_*(\rho_m', \sigma_m^{\rho'}) + (4\text{rank}(\rho)C)^{1/4}, \qquad (39)$$

For the δ_2 bound we first apply Lemma 4 to ρ and ρ' and we specialize ϵ to be $\epsilon_m' = 1 - \text{Tr}(\rho_m')$, giving

$$(D_T(\rho, \rho') - \epsilon_m')^2 \leqslant R_{\epsilon_m'}(\rho, \rho')D_{\text{HS}}(\rho, \sigma). \qquad (40)$$

Combining this result with (35) yields

$$|F(\rho, \sigma) - F(\rho', \sigma)| \leqslant \sqrt{2\epsilon_m' + 2\sqrt{R_{\epsilon_m'}(\rho, \rho')D_{\text{HS}}(\rho, \rho')}}. \qquad (41)$$

Then, from Lemma 2 we know that $\text{rank}_{\epsilon_m'}(\rho) \leqslant \text{rank}_{\epsilon_m'}(\rho')$, which implies $R_{\epsilon_m'}(\rho, \rho') \leqslant \text{rank}_{\epsilon_m'}(\rho')/2 = m/2$, where the last equality comes from the fact that $\epsilon_m' = \|\rho' - \rho_m'\|_1$. Moreover, since $D_{\text{HS}}(\rho, \rho') = C$, we have

$$|F(\rho, \sigma) - F(\rho', \sigma)| \leqslant \sqrt{2\epsilon_m' + \sqrt{2mC}}. \qquad (42)$$

Combining this result with the TFB of Proposition 1 yields the δ_2 bound in (34). $\qquad \square$

D   Proof of Proposition 4

We remark that Proposition 4 applies to any three quantum states, although for our purposes we are interested in its application to the states ρ, σ, and ρ' discussed in the main text. Hence, for consistency, we state this proposition for these states, but we note that these states can be arbitrary.

Proposition 4. *Let $D\left(F(\rho,\sigma)\right)$ be a distance measure that is monotonically decreasing in $F(\rho,\sigma)$. Let ρ, σ, and ρ' be three arbitrary quantum states. Then the following certifiable triangular inequality bounds (CTIB) hold:*

$$D_m^{LB} \leqslant D\left(F(\rho,\sigma)\right) \leqslant D_m^{UB}, \tag{43}$$

with

$$D_m^{LB} = D\left(F_*(\rho'_m, \sigma_m^{\rho'})\right) - D\left(F(\rho'_m, \rho_m^{\rho'})\right), \tag{44}$$

$$D_m^{UB} = D\left(F(\rho'_m, \sigma_m^{\rho'})\right) + D\left(F(\rho'_m, \rho_m^{\rho'})\right), \tag{45}$$

where ρ'_m, $\sigma_m^{\rho'}$, and $\rho_m^{\rho'}$ are projections of ρ', σ, and ρ, respectively, onto the subspace of the m-largest eigenvectors of ρ'.

Proof. Since $D\left(F(\rho,\sigma)\right)$ is a distance measure, it satisfies the triangular inequality. Applying this to the states ρ, σ, and ρ' gives

$$D\left(F(\rho,\sigma)\right) \geqslant D\left(F(\rho',\sigma)\right) - D\left(F(\rho',\rho)\right),$$
$$D\left(F(\rho,\sigma)\right) \leqslant D\left(F(\rho',\sigma)\right) + D\left(F(\rho',\rho)\right). \tag{46}$$

Combining Proposition 1 with (46) and using the monotonicity of D yields (43). □

E   Proof of Proposition 5

Proposition 5. *The problem* LOW-RANK FIDELITY ESTIMATION *to within precision $\pm\gamma = 1/\operatorname{poly}(n)$ is* `DQC1`-*hard.*

Proof. We reduce from the problem of approximating the Hilbert-Schmidt inner-product magnitude Δ_{HS} between two quantum circuits $\tilde{U}$ and $\tilde{V}$ acting on n-qubits [17], where we define

$$\Delta_{\mathrm{HS}}(\tilde{U}, \tilde{V}) := \frac{1}{d^2}|\mathrm{Tr}(\tilde{V}^\dagger \tilde{U})|^2, \tag{47}$$

where $d = 2^n$.

Consider a specific instance of approximating Δ_{HS}. We are given as input classical instructions to prepare $\operatorname{poly}(n)$-sized quantum circuits $\tilde{U}$ and $\tilde{V}$ on n-qubits each, and the task is to approximate $\Delta_{\mathrm{HS}}(\tilde{U}, \tilde{V})$ to precision $1/\operatorname{poly}(n)$. Our reduction will identify this problem as a specific instance of LOW-RANK FIDELITY ESTIMATION (Low-Rank Fidelity Estimation) via the Choi-Jamiołkowski isomorphism over the unitary channels,

$$\mathcal{E}_{\tilde{U}} : \mathcal{D}(\mathcal{H}_d) \to \mathcal{D}(\mathcal{H}_d), \quad X \mapsto \tilde{U} X \tilde{U}^\dagger \tag{48}$$

$$\mathcal{E}_{\tilde{V}} : \mathcal{D}(\mathcal{H}_d) \to \mathcal{D}(\mathcal{H}_d), \quad X \mapsto \tilde{V} X \tilde{V}^\dagger \tag{49}$$

where $\mathcal{D}(\mathcal{H}_d)$ is the space of $d \times d$ dimensional hermitian matrices. Consider now the $2n$-qubit maximally entangled state,

$$|\Phi^+\rangle = \frac{1}{\sqrt{d}} \sum_{\boldsymbol{j}} |\boldsymbol{j}\rangle \otimes |\boldsymbol{j}\rangle = E|\boldsymbol{0}\rangle, \tag{50}$$

where $\boldsymbol{j} = (j_1, j_2, ..., j_n)$ is a binary vector taking values j_k in $\{0, 1\}$, and where E is an efficent unitary entangling gate (e.g., a depth-two circuit composed of Hadamard and CNOT gates), where $|\boldsymbol{0}\rangle = |0\rangle^{\otimes 2n}$.

A special case of Low-Rank Fidelity Estimation is when ρ and σ correspond to the Choi states of $\tilde{U}$ and $\tilde{V}$. In this case, as the input to Low-Rank Fidelity Estimation, we would be given the gate sequences $U = (\tilde{U} \otimes \mathbb{1})E$ and $V = (\tilde{V} \otimes \mathbb{1})E$, which respectively prepare the pure states ρ and σ as

$$\rho = U|\boldsymbol{0}\rangle\langle\boldsymbol{0}|U^\dagger = (\tilde{U} \otimes \mathbb{1})E|\boldsymbol{0}\rangle\langle\boldsymbol{0}|E^\dagger(\tilde{U}^\dagger \otimes \mathbb{1}), \tag{51}$$

$$\sigma = V|\boldsymbol{0}\rangle\langle\boldsymbol{0}|V^\dagger = (\tilde{V} \otimes \mathbb{1})E|\boldsymbol{0}\rangle\langle\boldsymbol{0}|E^\dagger(\tilde{V}^\dagger \otimes \mathbb{1}). \tag{52}$$

Then, the fidelity between ρ and σ is given by

$$F(\rho,\sigma) = |\langle\boldsymbol{0}|E^\dagger(\tilde{V}^\dagger \tilde{U} \otimes \mathbb{1})E|\boldsymbol{0}\rangle| = \frac{1}{d}|\mathrm{Tr}(\tilde{V}^\dagger \tilde{U})|. \tag{53}$$

We can run LOW-RANK FIDELITY ESTIMATION to estimate the above expression to within $\gamma = 1/\operatorname{poly}(n)$ precision, and thus also approximate $\Delta_{\mathrm{HS}}(U, V)$ to within $1/\operatorname{poly}(n)$. Finally, it is known that approximating $\Delta_{\mathrm{HS}}(U, V)$ to within inverse polynomial precision is `DQC1`-hard [17], and hence the result follows. □

F Proof of Proposition 6

Before proving Proposition 6, we first prove the following useful lemmas.

Lemma 5. *Let $\epsilon_m = \|\rho - \rho_m\|_1$ and $\epsilon'_m = \|\rho' - \rho'_m\|_1$. Then the following bound holds*

$$\epsilon'_m \leqslant \epsilon_m + \sqrt{mC} \,, \tag{54}$$

where $C = D_{HS}(\rho, \rho')$ and ρ' is given by Eq. (10) of the main text.

Proof. Consider the difference

$$\epsilon'_m - \epsilon_m = \|\rho' - \rho'_m\|_1 - \|\rho - \rho_m\|_1 = (1 - \mathrm{Tr}(\rho'_m)) - (1 - \mathrm{Tr}(\rho_m)) = \sum_{i=1}^m (r_i - r'_i) \leqslant \sum_{i=1}^m |r_i - r'_i| = \|\boldsymbol{\delta}\|_1 \,, \tag{55}$$

where we have defined the vector $\boldsymbol{\delta} = \boldsymbol{r} - \boldsymbol{r}'$, with $\boldsymbol{r} = \{r_1, \ldots, r_m\}$ and $\boldsymbol{r}' = \{r'_1, \ldots, r'_m\}$. From the vector norm equivalence, we have that $\|\boldsymbol{\delta}\|_1 \leqslant \sqrt{m}\|\boldsymbol{\delta}\|_2$, which then implies

$$\epsilon'_m - \epsilon_m \leqslant \sqrt{m}\|\boldsymbol{\delta}\|_2 \,. \tag{56}$$

Moreover, as shown in [18], the VQSD cost C bounds the eigenvalue error as $\|\boldsymbol{\delta}\|_2^2 = \sum_{i=1}^m (r_i - r'_i)^2 \leqslant C$. Hence, combining this with Eq. (56), we get $\epsilon'_m \leqslant \epsilon_m + \sqrt{mC}$. $\qquad\square$

The following lemma is an alternative version of Proposition 6 that may be of interest in itself, particularly when we do not have the promise that ρ is low rank. Note that this lemma does not refer to the rank properties of ρ but rather refers to the rank properties of ρ'. The ϵ-rank of ρ' is experimentally measurable and hence one can use the following lemma to guarantee a particular precision even when there is no prior knowledge about the rank properties of ρ.

Lemma 6. *Let ρ and σ be quantum states, and suppose we have access to a subroutine that diagonalizes ρ up to error bounded by $C = D_{HS}(\rho, \rho')$. Let m and ζ be parameters. Then, VQFE runs in time $\tilde{O}(m^6/\zeta^2)$ and outputs an additive $\pm\gamma'$-approximation of $F(\rho, \sigma)$, for*

$$\gamma' = \sqrt{2\epsilon'_m + \sqrt{2mC}} + \sqrt{\epsilon'_m} + \zeta\|T\| \,, \tag{57}$$

where $\epsilon'_m = \|\rho' - \rho'_m\|_1$.

Proof. Let us first define the following quantities which will be useful to bound the error in each step of the VQFE algorithm:

$$\Delta_0 = |F(\rho, \sigma) - \hat{F}(\rho'_m, \sigma)| \,, \tag{58}$$

$$\Delta_1 = |F(\rho, \sigma) - F(\rho', \sigma)| \,, \qquad\qquad B_1 = \sqrt{2\epsilon'_m + \sqrt{2mC}} \,, \tag{59}$$

$$\Delta_2 = |F(\rho', \sigma) - F(\rho'_m, \sigma)| \,, \qquad\qquad B_2 = \sqrt{\epsilon'_m} \,, \tag{60}$$

$$\Delta_3 = |F(\rho'_m, \sigma) - \hat{F}(\rho'_m, \sigma)| \,, \qquad\qquad B_3 = m \cdot \|T - \hat{T}\| \,. \tag{61}$$

where the notation $\hat{a}$ (e.g., in $\hat{T}$) indicates a fine-sampling estimate of the random variable a. Note that the following bounds hold:

$$\Delta_1 \leqslant B_1, \qquad \Delta_2 \leqslant B_2, \qquad \Delta_3 \leqslant B_3 \,, \tag{62}$$

where the first inequality comes from Eq. (42) and the second inequality can be derived from Eq. (9) of the main text as follows

$$F(\rho', \sigma) - F(\rho'_m, \sigma) \leqslant F_*(\rho'_m, \sigma^{\rho'}_m) - F(\rho'_m, \sigma^{\rho'}_m) \leqslant \|\rho' - \rho'_m\|_1^{1/2} = \sqrt{\epsilon'_m} \,. \tag{63}$$

Finally, we prove the third inequality in Eq. (73) below.

By means of multiple applications of the triangle inequality, we obtain the following result:

$$\Delta_0 \leqslant \Delta_1 + |F(\rho', \sigma) - \hat{F}(\rho'_m, \sigma)| \tag{64}$$

$$\leqslant \Delta_1 + \Delta_2 + \Delta_3 \tag{65}$$

$$\leqslant B_1 + B_2 + B_3 \,. \tag{66}$$

Note that Δ_0 is ultimately the quantity we are interested in bounding since $F(\rho, \sigma)$ is the desired fidelity while $\hat{F}(\rho'_m, \sigma)$ is the quantity that VQFE actually outputs. While (66) provides an upper bound on Δ_0, we will need to re-formulate this as a probabilistic bound since B_3 depends on the estimate $\hat{T}$ of a random variable.

To do this re-formulation, let us consider the sources of statistical noise on $\hat{T}$. In the first step of VQFE, we call the VQSD subroutine to approximately diagonalize ρ up to some error bounded by $C = D_{\mathrm{HS}}(\rho, \rho')$. Here $\rho' = \sum_i r'_i |r'_i\rangle\langle r'_i|$, where $\{r'_i\}$ and $\{|r'_i\rangle\}$ are approximations of the eigenvalues and eigenvectors of ρ, respectively. While ρ' can only be estimated up to some finite sampling precision, such error can be tuned appropriately. We refer to [18] for further details.

The second step of VQFE employs the Swap Test to measure the matrix elements $\sigma_{ij} = \langle r_i | \sigma | r_j \rangle$. If $C \neq 0$ we measure instead $\sigma'_{ij} = \langle r'_i | \sigma | r'_j \rangle$, which can be used to construct the approximate matrix T' with matrix elements $T'_{ij} = (r'_i r'_j)^{1/2} \sigma'_{ij}$. For the sake of brevity, we will simply write T instead of T'. For each entry of the symmetric matrix T that needs to be determined, we call the Swap Test $O(m^4/\zeta^2 \cdot \log(2m^2/\delta))$ times. We denote $\hat{T}$ as the finite-sampling estimate of T. Then, according to the two-sided Chernoff bound, we have that any entry (i,j) in $\hat{T}$ has relative error at most ζ/m^2 with probability greater than $1 - \delta/m^2$,

$$\Pr\left[|\hat{T}_{ij} - T_{ij}| \leqslant \frac{\zeta}{m^2} \cdot T_{ij} \right] \geqslant 1 - \delta/m^2. \tag{67}$$

By the union bound, the probability that this is the case simultaneously for all entries is then given by

$$\Pr\left[\forall (i,j) \in [m] \times [m] : \ |\hat{T}_{ij} - T_{ij}| \leqslant \frac{\zeta}{m^2} \cdot T_{ij} \right] \geqslant 1 - \delta. \tag{68}$$

Then, we conclude that

$$\Pr\left[\|T - \hat{T}\| \leqslant \frac{\zeta \|T\|}{m} \right]$$

$$= \Pr\left[\sqrt{\sum_{i,j=1}^m |T_{ij} - \hat{T}_{ij}|^2} \leqslant \frac{\zeta \|T\|}{m} \right] \tag{69}$$

$$= \Pr\left[\sqrt{\sum_{i,j=1}^m |T_{ij} - \hat{T}_{ij}|^2} \leqslant \frac{\zeta \|T\|}{m} \ \Big| \ \forall (i,j) : |\hat{T}_{ij} - T_{ij}| \leqslant \frac{\zeta \cdot T_{ij}}{m^2} \right] \cdot \Pr\left[\forall (i,j) : |\hat{T}_{ij} - T_{ij}| \leqslant \frac{\zeta \cdot T_{ij}}{m^2} \right] \tag{70}$$

$$\geqslant \underbrace{\Pr\left[\frac{\zeta}{m^2} \sqrt{\sum_{i,j=1}^m T_{ij}^2} \leqslant \frac{\zeta \|T\|}{m} \ \Big| \ \forall (i,j) : |\hat{T}_{ij} - T_{ij}| \leqslant \frac{\zeta \cdot T_{ij}}{m^2} \right]}_{=1} \cdot \Pr\left[\forall (i,j) : |\hat{T}_{ij} - T_{ij}| \leqslant \frac{\zeta \cdot T_{ij}}{m^2} \right] \tag{71}$$

$$= \Pr\left[\forall (i,j) : |\hat{T}_{ij} - T_{ij}| \leqslant \frac{\zeta}{m^2} \cdot T_{ij} \right] \quad \geqslant \quad 1 - \delta. \qquad \text{(by Eq. (68))} \tag{72}$$

The third step of VQFE requires the diagonalization of the $m \times m$ symmetric matrix $\hat{T}$, which can be accomplished in time $O(m^3)$. Let us now prove the bound $\Delta_3 \leqslant B_3$, i.e., the bound on the finite sampling error in our estimate $\hat{F}(\rho'_m, \sigma)$ of the fidelity $F(\rho'_m, \sigma)$. By denoting λ_j and $\hat{\lambda}_j$ the eigenvalues of T and $\hat{T}$, respectively, we get

$$\Delta_3 = |F(\rho'_m, \sigma) - \hat{F}(\rho'_m, \sigma)|$$

$$= |\mathrm{Tr}\sqrt{T} - \mathrm{Tr}\sqrt{\hat{T}}|$$

$$\leqslant \sum_{j=1}^m |\sqrt{\lambda_j} - \sqrt{\hat{\lambda}_j}|$$

$$\leqslant m \cdot \max_{1 \leqslant j \leqslant m} |\sqrt{\lambda_j} - \sqrt{\hat{\lambda}_j}|$$

$$\leqslant m \cdot \|T - \hat{T}\| = B_3 \,. \tag{73}$$

Then, by means of Eq. (72) and Eq. (73), we get

$$\Pr\left[\Delta_3 \leqslant \zeta \|T\| \right] = \Pr\left[|F(\rho'_m, \sigma) - \hat{F}(\rho'_m, \sigma)| \leqslant \zeta \|T\| \right] = \Pr\left[|\mathrm{Tr}\sqrt{T} - \mathrm{Tr}\sqrt{\hat{T}}| \leqslant \zeta \|T\| \right] \tag{74}$$

$$\geqslant \Pr\left[m \cdot \|T - \hat{T}\| \leqslant \zeta \|T\| \right] \tag{75}$$

$$\geqslant 1 - \delta \,. \tag{76}$$

Let us define

$$\gamma' = B_1 + B_2 + \zeta\|T\| = \sqrt{2\epsilon'_m + \sqrt{2mC}} + \sqrt{\epsilon'_m} + \zeta\|T\|\,, \tag{77}$$

and combine our previous results to bound the VQFE run time:

$$\Pr\left[|F(\rho,\sigma) - \hat{F}(\rho'_m,\sigma)| \leqslant \gamma'\right] = \Pr\left[\Delta_0 \leqslant B_1 + B_2 + \zeta\|T\|\right] \tag{78}$$

$$\geqslant \Pr\left[\Delta_1 + \Delta_2 + \Delta_3 \leqslant B_1 + B_2 + \zeta\|T\|\right] \quad \text{(by Eq. (65))} \tag{79}$$

$$\geqslant \Pr\left[B_1 + B_2 + \Delta_3 \leqslant B_1 + B_2 + \zeta\|T\|\right] \quad \text{(by Eq. (66))} \tag{80}$$

$$= \Pr\left[\Delta_3 \leqslant \zeta\|T\|\right] \tag{81}$$

$$\geqslant \quad 1 - \delta. \quad \text{(by Eq. (76))} \tag{82}$$

As previously mentioned, VQFE performs the Swap Test for each of the $O(m^2)$ matrix elements. Therefore, VQFE runs in time $O(m^6/\zeta^2 \cdot \log(2m^2/\delta))$, or just $\widetilde{O}(m^6/\zeta^2)$ when neglecting slower growing logarithmic contributions. Furthermore, from (82), we have that VQFE outputs an additive $\pm\gamma'$-approximation of the fidelity $F(\rho,\sigma)$ with precision γ' defined by (77) and with probability at least $1 - \delta$. $\qquad\square$

Proposition 6. *Let ρ and σ be quantum states, and suppose we have access to a subroutine that diagonalizes ρ up to error bounded by $C = D_{HS}(\rho,\rho')$. Let m and ζ be parameters. Then, VQFE runs in time $\widetilde{O}(m^6/\zeta^2)$ and outputs an additive $\pm\gamma$-approximation of $F(\rho,\sigma)$, for*

$$\gamma = (\epsilon + \sqrt{mC})^{1/2} + (2\epsilon + (2+\sqrt{2})\sqrt{mC})^{1/2} + \zeta\|T\|\,, \tag{83}$$

where ϵ is determined by $m = \mathrm{rank}_\epsilon(\rho)$.

Proof. This proposition is very similar to Lemma 6 except that γ' is replaced by γ. To make this replacement we show that $\gamma \geqslant \gamma'$ as follows:

$$\gamma = \sqrt{2\epsilon + (2+\sqrt{2})\sqrt{mC}} + \sqrt{\epsilon + \sqrt{mC}} + \zeta\|T\|$$

$$\geqslant \sqrt{2\|\rho - \rho_m\|_1 + (2+\sqrt{2})\sqrt{mC}} + \sqrt{\|\rho - \rho_m\|_1 + \sqrt{mC}} + \zeta\|T\|$$

$$\geqslant \sqrt{2\epsilon'_m + \sqrt{2mC}} + \sqrt{\epsilon'_m} + \zeta\|T\|$$

$$= \gamma'\,. \tag{84}$$

The first inequality follows from the fact that $m = \mathrm{rank}_\epsilon(\rho)$ and the second inequality follows from Lemma 5.

Recall from the proof of Lemma 6 (Eq. (82)) that the output of VQFE satisfies the following inequality

$$\Pr\left[|F(\rho,\sigma) - \hat{F}(\rho'_m,\sigma)| \leqslant \gamma'\right] \geqslant 1 - \delta. \tag{85}$$

Then, because $\gamma \geqslant \gamma'$ we also have that

$$\Pr\left[|F(\rho,\sigma) - \hat{F}(\rho'_m,\sigma)| \leqslant \gamma\right] \geqslant \Pr\left[|F(\rho,\sigma) - \hat{F}(\rho'_m,\sigma)| \leqslant \gamma'\right] \geqslant 1 - \delta\,. \tag{86}$$

As noted in the proof of Lemma 6, VQFE achieves this output with a run time $\widetilde{O}(m^6/\zeta^2)$. Hence, with this run time, VQFE outputs a $\pm\gamma$-approximation of the fidelity $F(\rho,\sigma)$ with precision γ defined by (83) and with probability at least $1 - \delta$.

$\qquad\square$