

Quantum estimation of Chernoff bound and Hoeffding bound

Honglin Chen,¹ Jicun Li,¹ Wei Xie,^{1,*} and Xiang-Yang Li^{1,2,†}

¹*School of Computer Science and Technology, University of Science and Technology of China, Hefei 230027, China*

²*Hefei National Laboratory, University of Science and Technology of China, Hefei 230088, China*



(Received 27 April 2025; accepted 27 August 2025; published 12 September 2025)

Determining the closeness of two quantum states with minimal error probability is a central problem in quantum information theory. The optimal exponents characterizing the error probability in the symmetric and asymmetric settings are given by the quantum Chernoff bound (QCB) and the quantum Hoeffding bound (QHB), respectively. In this work, we introduce the first algorithms for estimating the QCB and QHB of two unknown states. We propose a unitary encoding for diagonal density matrix based on the copy matrix, which enables unbiased estimation of nonlinear matrix functions via quantum singular value transformation (QSVT). For general density matrices, we combine QSVT with the Hadamard test to estimate both QCB and QHB. Numerical simulations demonstrate that the algorithm converges to the optimal value within a limited number of iterations. Furthermore, we establish the first integration of QSVT with convex optimization, yielding a hybrid quantum-classical framework for quantum estimation.

DOI: [10.1103/blvt-zrww](https://doi.org/10.1103/blvt-zrww)

I. INTRODUCTION

A fundamental problem in quantum information processing is to discriminate between two sources and identify the true source with minimal error probability. In both classical and quantum settings, this involves testing two hypotheses, each associated with a probability distribution or a quantum state.

Around 70 years ago, Herman Chernoff established a seminal bound that characterizes the asymptotic behavior of the minimal error probability when discriminating between two hypotheses with many observations [1]. The error probability P_e for distinguishing two probability distributions decays exponentially with the number of observations n : $P_e \sim \exp(-\zeta n)$. As n increases, the agreement improves, and the exponent ζ becomes significant in the asymptotic limit. The asymptotic efficiency corresponds to the limit of exponent ζ .

In the quantum setting, the experiment prepares n independent copies of a quantum system in an unknown state. Testing is performed via quantum measurements, and the states are represented by density matrices ρ and σ , which are positive semidefinite Hermitian matrices with trace 1. The quantum analog of the Chernoff bound was established through the results of two seminal papers [2,3]. It has been shown that the optimal error exponent for discriminating $\rho^{\otimes n}$ and $\sigma^{\otimes n}$ is given by the quantum Chernoff bound (QCB):

$$\zeta(\rho, \sigma) = -\ln \left(\inf_{0 \leq s \leq 1} \text{Tr}[\rho^s \sigma^{1-s}] \right),$$

which specifies the asymptotic rate exponent of the minimal error probability.

The Chernoff bound provides the optimal decay rate for the average error probabilities in the symmetric setting. In the asymmetric case, imposing an exponential constraint on the first error probability yields the optimal decay rate, as derived from the Hoeffding bound. The quantum Hoeffding bound (QHB) governs the rate limit for false negatives, which is given by

$$H(r) := \sup_{0 \leq s < 1} \frac{-rs - \ln(\text{Tr}[\rho^s \sigma^{1-s}])}{1-s}.$$

The estimation of QCB and QHB is crucial for property testing for quantum state and closely linked to the finite-sample regime [4]. Significant progress has been made in estimating various quantum state properties, including quantum fidelity [5–8], quantum entropy [9–13], and quantum trace distance [14,15].

We present the first algorithms for estimating both the QCB and QHB of two unknown quantum states. For efficient matrix function evaluation, our algorithm framework is based on the quantum singular value transformation (QSVT) [16], which comprises two main components: block encoding (see Definition 2) and polynomial approximation.

Direct application of the original QSVT algorithm is insufficient for computing the QCB and QHB. Since the quantum states are unknown, encoding the density matrix solely through block encoding introduces uncontrollable biases. To address this, we consider two cases. If the density matrix is diagonal (see Definition 1), we introduce an encoding scheme that performs the required computations without such errors (Sec. III B). Otherwise, we employ the Hadamard test circuit to estimate the QCB and QHB (Sec. IV B).

Second, computing the QCB and QHB requires a polynomial approximation of $\sqrt{\ln x}$. Currently, Lemma 3.3 of Ref. [9] is the only result that establishes the existence

*Contact author: xxieww@ustc.edu.cn

†Contact author: xiangyangli@ustc.edu.cn

of a polynomial of degree $O(\frac{1}{\delta} \ln(\frac{1}{\delta\epsilon}))$ that approximates $\sqrt{-\ln x}/(2\sqrt{-\ln \delta})$ over the domain $[\delta, 1 - \delta]$.

In this work, we extend standard QSVT polynomial approximation techniques: for any function $f(x)$ that admits a Taylor expansion, we construct an explicit polynomial approximation of $\sqrt{f(x)}$ (see Appendix B). Moreover, our approximation interval is $[\delta, 1]$, which is broader than the one used in Lemma 3.3 of Ref. [9].

Finally, since computing the QCB and QHB involves finding the extremum of the target functions, we use QSVT to compute their first and second derivatives, and then apply the inexact Newton method to locate the extremum [17]. With appropriately chosen forcing terms, the inexact Newton method achieves linear convergence. For further details, see Sec. IV.

By combining QSVT with the inexact Newton method, we establish a hybrid quantum-classical framework for estimating the QCB and QHB.

Theorem 1 (Informal version of Theorems 6 and 8). For diagonal density matrices ρ, σ , we provide an algorithm for estimating the QCB and QHB with query complexity $\tilde{O}(\frac{r_m^{1/\bar{s}}}{\epsilon^{1/\bar{s}+1}})$ to density matrix access. For general density matrices, the query complexity becomes $\tilde{O}(d \frac{r_m^{1/\bar{s}}}{\epsilon^{1/\bar{s}+1}})$, where $\bar{s} = \min\{s, 1 - s\}$ and $r_m = \max\{\text{rank}(\rho), \text{rank}(\sigma)\}$.

Additionally, our algorithms can also be applied to estimate quantum relative entropy and quantum α -Rényi relative entropy, partially resolving the open problem in Ref. [9]. Numerical simulations demonstrate that our method estimates QCB and QHB with high accuracy, achieving errors below 10^{-3} .

The remainder of this paper is organized as follows. Section II reviews quantum hypothesis testing and formulates QCB and QHB. Section III introduces key tools for our algorithm design, particularly quantum singular value transformation [16] for efficient matrix function evaluation. Section IV details our QCB and QHB estimation algorithms using inexact Newton method, QSVT, and amplitude amplification [5]. Section V presents numerical results, and Sec. VI concludes the paper.

II. HYPOTHESIS TESTING

Quantum state discrimination is the quantum analog of the statistical hypothesis testing problem. The simplest scenario involves discrimination between two arbitrary quantum states. For instance, consider Alice, who prepares a quantum system in an unknown state, which is either ρ with *a priori* probability p_0 or σ with *a priori* probability $p_1 := 1 - p_0$. This system is then passed to Bob, who must determine which state Alice prepared. Bob faces a hypothesis test with two hypotheses: a null hypothesis H_0 (indicating the system is in state ρ) and an alternative hypothesis H_1 (indicating the system is in state σ).

A. Quantum Chernoff bound

To distinguish between two quantum states, Bob performs a positive operator-valued measure (POVM) $\{\Pi_k\}$, with $k = 0, 1$, where $k = 0$ corresponds to the null hypothesis H_0 and $k = 1$ to the alternative hypothesis H_1 . The goal is to optimize the POVM to minimize the error probability.

This problem is asymmetric, as introduced by Neyman and Pearson, who defined type I and type II errors. Type I error (false positive), α , is the probability of rejecting H_0 when it is true:

$$\alpha_n = p(H_1|H_0) = \text{Tr}(\Pi_1 \rho^{\otimes n}).$$

Type II error (false negative), β , is the probability of failing to reject H_0 when H_1 is true:

$$\beta_n = p(H_0|H_1) = \text{Tr}(\Pi_0 \sigma^{\otimes n}).$$

In symmetric testing, although type I and type II errors may not be equally likely, Bob's POVM $\{\Pi_0, \Pi_1\}$ is chosen to minimize the overall error probability:

$$P_e = p_0 \text{Tr}(\Pi_1 \rho) + p_1 \text{Tr}(\Pi_0 \sigma).$$

The quantum Chernoff bound offers an estimate for the minimum error probability in such tests.

B. Quantum Hoeffding bound

A central problem in quantum information theory is determining the relative entropy between two quantum states, which plays a crucial role in asymmetric hypothesis testing [18].

It provides an operational interpretation as the best achievable asymptotic rate of type II errors, reflecting the asymmetry between the null and alternative hypotheses due to the different treatments of type I and type II errors. For instance, imposing a constraint $\alpha < r$ on the type I error minimizes the type II error rate. The optimal type II error rate is given by the relative entropy of the null hypothesis, independent of the constraint r . This result is formalized in Stein's lemma [19].

When the constraint r is no more than the relative entropy of ρ with respect to σ , which is defined as $D(\rho\|\sigma) := \text{Tr}(\rho(\ln \rho - \ln \sigma))$, the asymptotic exponential rate of type II errors is at most $H(r)$ [20].

In the asymptotic regime, we focus on the error-rate exponents as the number of copies increases. Bob's objective is to maximize $\lim_{n \rightarrow \infty} -(1/n) \ln \beta_n$ ensuring the fastest exponential decay of the type II error probability β_n . The QHB imposes a stronger constraint on the error probability of false positives.

III. PRELIMINARIES AND NOTATIONS

We consider a density matrix acting on a Hilbert space of dimension $d := 2^t$. Let $[d] := \{0, 1, \dots, d-1\}$ and $i := \sqrt{-1}$. To compute the QCB and QHB, we primarily require the evaluation of the functions $h(s) := \text{Tr}(\rho^s \sigma^{1-s})$ and $g(s) := (-rs - \ln \text{Tr}(\rho^s \sigma^{1-s}))/ (1-s)$, together with their first and second derivatives, which are denoted as $\dot{h}(s)$, $\ddot{h}(s)$, $\dot{g}(s)$, and $\ddot{g}(s)$. Moreover, we denote by $f_{\text{QSVT}}(x)$ the value of $f(x)$ computed via QSVT.

We use $|\cdot\rangle$ to denote normalized quantum states and $|i\rangle$ to denote the computational basis states for $i \in [d]$. The notation $\|\cdot\|$ has two meanings: for matrices, it denotes the spectral norm, and for functions, it denotes the maximum absolute value. For matrices, $\|\cdot\|_1$ denotes the trace norm, defined as $\|A\|_1 = \text{Tr}(\sqrt{A^\dagger A})$.

In this work, we assume that the copies of ρ and σ used for discrimination are obtained via quantum purified access, which is very natural when the state is the output of a quantum algorithm [21].

Definition 1 (Quantum purified access). A density matrix $\rho \in \mathbb{C}^{d \times d}$ is said to have quantum purified access if there exists a unitary oracle U_ρ (and its inverse):

$$U_\rho |0\rangle_a |0\rangle_t = |\psi_\rho\rangle = \sum_{i=0}^{d-1} \sqrt{p_i} |\phi_i\rangle |\psi_i\rangle,$$

such that $\rho = \text{Tr}_a(|\psi_\rho\rangle\langle\psi_\rho|)$. If $|\psi_i\rangle = |i\rangle$, then $\rho = \sum_{i=0}^{d-1} p_i |i\rangle\langle i|$ is called a diagonal density matrix. We refer to U_ρ as the purified matrix.

In the following subsections, we first present commonly used amplitude estimation algorithms and the unitary encoding techniques related to QSVT. Building on this foundation, we then introduce the QSVT algorithm and explain how to construct polynomial approximations of the target functions.

A. Amplitude estimation

Theorem 2 ([22], Theorem 12). Given access to a unitary U satisfying

$$U|0\rangle|0\rangle = \sqrt{p}|0\rangle|\phi\rangle + \sqrt{1-p}|1\rangle|\perp\rangle, \quad (1)$$

where $(\langle 0| \otimes I)|\perp\rangle = 0$, the amplitude estimation algorithm outputs an estimate $\tilde{p} \in [0, 1]$ such that

$$|\tilde{p} - p| \leq \frac{2\pi\sqrt{p(1-p)}}{M} + \frac{\pi^2}{M^2},$$

with a success probability of at least $8/\pi^2$, using M calls to U and $U^\dagger$.

B. Unitary encoding of density matrix

We now recall basic results from the matrix arithmetic toolbox described in Ref. [16]. A block encoding unitary [23] represents a subnormalized matrix as the top-left block of a unitary as follows:

$$U = \begin{bmatrix} A/\alpha & \\ & \end{bmatrix} \Rightarrow A = \alpha(\langle 0| \otimes I)U(|0\rangle \otimes I). \quad (2)$$

Definition 2 ([23], Definition 43). Suppose that A is a t -qubit matrix, with $\alpha, \varepsilon \in \mathbb{R}_+$ and $a \in \mathbb{N}$. We say that a $(t+a)$ -qubit unitary U is an (α, a, ε) -block encoding of A if

$$\|A - \alpha(\langle 0|_a \otimes I_t)U(|0\rangle_a \otimes I_t)\| \leq \varepsilon.$$

The matrix norm $\|\cdot\|$ is the spectral norm, and when $\alpha = 1$ and $\varepsilon = 0$, we simply refer to an (α, a, ε) -block encoding as a block encoding.

For our purposes, we present concrete construction of block encoding unitary for density matrix.

Lemma 1 ([23], Lemma 45). Given a density matrix $\rho \in \mathbb{C}^{d \times d}$, we use U_ρ from Definition 1 to construct the block encoding unitary for ρ as follows:

$$\text{BE}_\rho := (U_\rho^\dagger \otimes I_t)(I_a \otimes \text{SWAP}_t)(U_\rho \otimes I_t),$$

where SWAP_t denotes the t -qubit swap matrix.

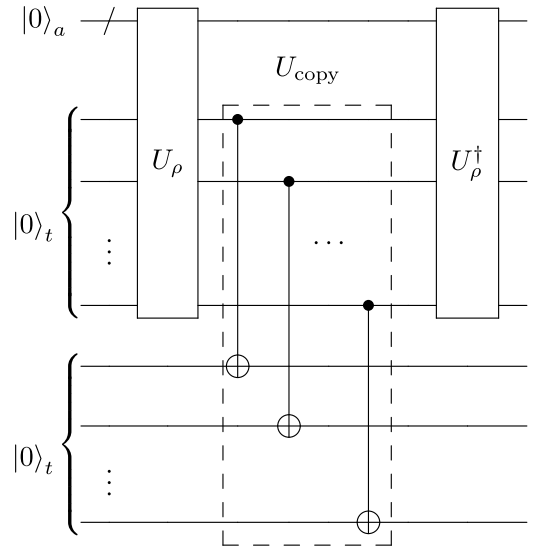


FIG. 1. The circuit of PE_ρ .

Moreover, given sampling access, one can also obtain block encoding access via density matrix exponentiation [5]. Specifically, Corollary 21 of Ref. [5] shows that it is possible to implement a quantum channel that serves as a $(4/\pi, 3, 0)$ block encoding of a density operator ρ , with error δ in the diamond norm.

However, the block encoding method proposed in Lemma 1 and Corollary 21 of Ref. [5] cannot be directly applied to the QCB and QHB estimation problem. For instance, consider the block encoding BE_ρ derived from Lemma 1. Since the eigenvectors $|\psi_i\rangle$ of ρ are unknown, applying BE_ρ to a general state $|\zeta\rangle$ leads to deviations proportional to its overlap with the unknown eigenbasis.

To address this issue, we replace the SWAP_t matrix in Lemma 1 with the copy matrix U_{copy} introduced in Ref. [24], leading to a modified unitary encoding of the density matrix. The copy matrix acts as $U_{\text{copy}}|i\rangle|0\rangle_t = |i\rangle|i\rangle$ for $i \in [d]$. We refer to the resulting unitary as the *projected encoding unitary*, since it can be applied directly to the initial state $|0\rangle$ (see Lemma 2). The following lemma presents an alternative encoding circuit for a diagonal density matrix, constructed using purified matrix together with the copy matrix.

Lemma 2 (Projected unitary encoding of density matrix). Given a diagonal density matrix $\rho \in \mathbb{C}^{d \times d}$, we use U_ρ from Definition 1 together with the copy matrix U_{copy} to define the projected unitary encoding

$$\text{PE}_\rho := (U_\rho^\dagger \otimes I_t)(I_a \otimes U_{\text{copy}})(U_\rho \otimes I_t).$$

Let $\Pi = \tilde{\Pi} = |0\rangle\langle 0|_{a+t} \otimes I_t$. Then,

$$\tilde{\Pi} \text{PE}_\rho \Pi = \sum_{i=0}^{d-1} p_i |0\rangle\langle 0|_a \otimes |0\rangle\langle 0|_t \otimes |i\rangle\langle i|, \quad (3)$$

where $\rho = \sum_{i=0}^{d-1} p_i |i\rangle\langle i|$ is the eigenvalue decomposition of ρ . The matrix U_{copy} can be implemented using CNOT gates, and the circuit for PE_ρ is shown in Fig. 1. The same construction applies to PE_σ .

Proof. This can be verified through straightforward calculations. We compute the result for PE_ρ , and the calculation for

PE_σ follows similarly:

$$\begin{aligned}
 & (|0\rangle\langle 0|_a \otimes |0\rangle\langle 0|_t \otimes I_t)(U_\rho^\dagger \otimes I_t)(I_a \otimes U_{\text{copy}})(U_\rho \otimes I_t)(|0\rangle\langle 0|_a \otimes |0\rangle\langle 0|_t \otimes I_t) \\
 &= \left(\sum_{i=0}^{d-1} \sqrt{p_i} |0\rangle\langle i| \otimes |0\rangle\langle \psi_i| \otimes I_t \right) (I_a \otimes U_{\text{copy}}) \left(\sum_{j=0}^{d-1} \sqrt{p_j} |j\rangle\langle 0| \otimes |\psi_j\rangle\langle 0| \otimes I_t \right) \\
 &= \left(\sum_{i=0}^{d-1} \sqrt{p_i} |0\rangle\langle \phi_i| \otimes |0\rangle\langle i| \otimes I_t \right) \left(\sum_{j=0}^{d-1} \sqrt{p_j} |\phi_j\rangle\langle 0| \otimes |j\rangle\langle 0| \otimes |j\rangle\langle 0| \right) \\
 &= \sum_{i=0}^{d-1} p_i |0\rangle\langle 0|_a \otimes |0\rangle\langle 0|_t \otimes |i\rangle\langle 0|.
 \end{aligned}$$

C. Quantum singular value transformation

To estimate QCB and QHB with additive error, we rely on the fundamental results of quantum singular value transformation [23]. Mathematically, the transformation of the eigenvalues of a density matrix is defined as follows:

Definition 3 (Eigenvalue transformation of density matrix). Let $f : \mathbb{R} \rightarrow \mathbb{C}$ be an even or odd function. Let $\rho \in \mathbb{C}^{d \times d}$ have the following eigenvalue value decomposition:

$$\rho = \sum_{i=0}^{d-1} p_i |\psi_i\rangle\langle \psi_i|.$$

For the function f , we define the eigenvalue transformation of ρ as

$$f^{(\text{EV})}(\rho) := \sum_{i=0}^{d-1} f(p_i) |\psi_i\rangle\langle \psi_i|.$$

Whenever no ambiguity arises, we abbreviate $f^{(\text{EV})}(\rho)$ as $f(\rho)$.

Quantum eigenvalue transformation by real polynomials can be efficiently implemented for density matrix. Here, we modify Corollary 18 of Ref. [23] to suit our specific setting.

Lemma 3 (Quantum eigenvalue transformation for density matrix). If $P = \sum_{k=0}^n a_k x^k \in \mathbb{R}[x]$ is an odd or even degree- n polynomial such that $\|P(x)\|_{[-1,1]} \leq 1$. Let $\Pi = \tilde{\Pi} = |0\rangle\langle 0|_{a+t} \otimes I_t$. Then there exists $\Phi = (\Phi_1, \dots, \Phi_n) \in \mathbb{R}^n$, such that

$$P^{(\text{EV})}(\tilde{\Pi} \text{BE}_\rho \Pi) = \begin{cases} ((+| \otimes \tilde{\Pi})(|0\rangle\langle 0| \otimes U_{(\Phi, \rho)} + |1\rangle\langle 1| \otimes U_{(-\Phi, \rho)})(|+ \rangle \otimes \Pi) & \text{if } n \text{ is odd,} \\ ((+| \otimes \Pi)(|0\rangle\langle 0| \otimes U_{(\Phi, \rho)} + |1\rangle\langle 1| \otimes U_{(-\Phi, \rho)})(|+ \rangle \otimes \Pi) & \text{if } n \text{ is even,} \end{cases}$$

where

$$\begin{aligned}
 U_{(\Phi, \rho)} &= e^{i\Phi_1(2\tilde{\Pi}-I)} \text{BE}_\rho \prod_{l=1}^{(n-1)/2} (e^{i\Phi_{2l}(2\Pi-I)} \text{BE}_\rho^\dagger e^{i\Phi_{2l+1}(2\tilde{\Pi}-I)} \text{BE}_\rho) \quad \text{if } n \text{ is odd,} \\
 U_{(\Phi, \rho)} &= \prod_{l=1}^{n/2} (e^{i\Phi_{2l-1}(2\Pi-I)} \text{BE}_\rho^\dagger e^{i\Phi_{2l}(2\tilde{\Pi}-I)} \text{BE}_\rho) \quad \text{if } n \text{ is even.}
 \end{aligned}$$

Replacing BE_ρ with BE_σ , PE_ρ , or PE_σ will yield similar conclusions.

We define $\text{BE}_{P(\rho)} := |0\rangle\langle 0| \otimes U_{(\Phi, \rho)} + |1\rangle\langle 1| \otimes U_{(-\Phi, \rho)}$, where Φ is the tuple corresponding to the polynomial $P(x)$. This definition extends similarly to other types of encoding unitary.

D. Low-degree polynomial approximation

We require polynomial approximations for positive power functions and logarithmic functions. Lemmas 4–6 and Lemma 7 provide constructions for these approximations, respectively. By combining these results with Lemma 3, we can

implement the corresponding functional transformations using QSVT.

Lemma 4 ([23], Lemma 29). Let $\delta', \varepsilon' \in (0, (1/2))$ and $t \in [-1, 1]$. There exists an even polynomial $P' \in \mathbb{R}[x]$ of degree $O(\frac{1}{\delta'} \ln(\frac{1}{\varepsilon'}))$, such that $|P'(x)| \leq 1$ for all $x \in [-1, 1]$, and

$$\begin{aligned}
 P'(x) &\in [0, \varepsilon'], & \text{for all } x \in [-1, -t - \delta'] \cup [t + \delta', 1], \\
 P'(x) &\in [1 - \varepsilon', 1], & \text{for all } x \in [-t + \delta', t - \delta'].
 \end{aligned}$$

Lemma 5 ([16], Corollary 3.4.14). Let $\delta, \varepsilon \in (0, (1/2))$ and let $f(x) := x^c/2$, then there exists even/odd polynomials $P, P' \in \mathbb{R}[x]$, such that $\|P - f\|_{[\delta, 1]} \leq \varepsilon$, $\|P\|_{[-1, 1]} \leq 1$ and

similarly $\|P' - f\|_{[\delta, 1]} \leq \varepsilon$, $\|P'\|_{[-1, 1]} \leq 1$; moreover, the degrees of the polynomials are $O(\frac{1}{\delta} \ln(\frac{1}{\varepsilon}))$.

Combing Lemmas 4 and 5, we can obtain a more detailed characterization of polynomial approximations for positive power functions on $[-1, 1]$.

Lemma 6 (*Polynomial approximation of positive power functions*). Let $\delta, \varepsilon \in (0, (1/2))$ and let $f(x) := x^c/2$, then there exists even/odd polynomials $P, P' \in \mathbb{R}[x]$, such that $\|P - f\|_{[\delta, 1]} \leq \varepsilon$, $\|P\|_{[-1, 1]} \leq 1$, and $|P(x)| \leq f(|x|) + \varepsilon$, for all $x \in [-1, 1]$. Similarly, $\|P' - f\|_{[\delta, 1]} \leq \varepsilon$, $\|P'\|_{[-1, 1]} \leq 1$, and $|P'(x)| \leq f(|x|) + \varepsilon$, for all $x \in [-1, 1]$. Moreover, the degrees of the polynomials are $O(\frac{1}{\delta} \ln(\frac{1}{\varepsilon}))$.

The proof of Lemma 6 is provided in Appendix A. In particular, we define the even polynomial function that approximates $\frac{x^{\frac{s}{2}}}{2}$ as $P_s(x)$ and $\frac{x^{\frac{1-s}{2}}}{2}$ as $P_{1-s}(x)$, as described in Lemma 6, where $s \in [0, 1]$.

Lemma 7 (*Polynomial approximations of logarithmic functions*). Let $\delta \in (0, (1/3))$, $\tilde{K} := 2\sqrt{\ln(2/\delta)}$, and $\hat{K} := 2\ln(2/\delta)$; there exists an even polynomial $\|\tilde{S}(x)\|_{[-1, 1]} \leq 1$ and $\|\hat{S}(x)\|_{[-1, 1]} \leq 1$ such that

$$\forall x \in [\delta, 1] : \left| \tilde{S}(x) - \frac{\sqrt{\ln 1/x}}{\tilde{K}} \right| \leq \varepsilon,$$

$$\forall x \in [\delta, 1] : \left| \hat{S}(x) - \frac{\ln 1/x}{\hat{K}} \right| \leq \varepsilon;$$

moreover, $\deg(\tilde{S}) = \deg(\hat{S}) = O(\frac{1}{\delta} \ln \frac{1}{\varepsilon})$.

The proof of Lemma 7 is provided in Appendix B. By combining Lemmas 3, 6, and 7, we obtain the following theorem:

Theorem 3 ($x^\alpha, \alpha \in (0, 1)$ transformation). Let $\rho \in \mathbb{C}^{d \times d}$ be arbitrary density matrices with smallest eigenvalues no less than δ . For all $\varepsilon \in (0, 1)$, given quantum purified access U_ρ and its conjugate transpose, we use $O(\frac{1}{\delta} \ln \frac{1}{\varepsilon})$ applications of U_ρ and $U_\rho^\dagger$, along with the same number of controlled reflections through $\Pi = \tilde{\Pi} = |0\rangle\langle 0|_{a+t} \otimes I_t$, to obtain

$$\left\| ((|+\rangle \otimes |0\rangle_{a+t} \otimes I_t) \text{BE}_{P_s(\rho)}(|+\rangle \otimes |0\rangle_{a+t} \otimes I_t) - \frac{\rho^\alpha}{2} \right\| \leq \varepsilon.$$

Replacing $\text{BE}_{P_s(\rho)}$ with $\text{BE}_{P_s(\sigma)}$, $\text{PE}_{P_s(\rho)}$, or $\text{PE}_{P_s(\sigma)}$ will yield similar conclusions.

Theorem 4 ($\sqrt{\ln 1/x}$ and $\ln 1/x$ transformation). Let $\rho \in \mathbb{C}^{d \times d}$ be arbitrary density matrices with smallest eigenvalues no less than δ . For all $\varepsilon \in (0, 1)$, given quantum purified access U_ρ for the density matrix ρ , we use $O(\frac{1}{\delta} \ln \frac{1}{\varepsilon})$ applications of U_ρ and $U_\rho^\dagger$, along with the same number of controlled reflections through $\Pi = \tilde{\Pi} = |0\rangle\langle 0|_{a+t} \otimes I_t$, to obtain

$$\left\| ((|+\rangle \otimes |0\rangle_{a+t} \otimes I_t) \text{BE}_{\tilde{S}(\rho)}(|+\rangle \otimes |0\rangle_{a+t} \otimes I_t) - \frac{\sqrt{\ln \frac{1}{\rho}}}{\tilde{K}} \right\| \leq \varepsilon,$$

$$\left\| ((|+\rangle \otimes |0\rangle_{a+t} \otimes I_t) \text{BE}_{\hat{S}(\rho)}(|+\rangle \otimes |0\rangle_{a+t} \otimes I_t) - \frac{\ln \frac{1}{\rho}}{\hat{K}} \right\| \leq \varepsilon.$$

Similarly, analogous conclusions hold for $\text{BE}_{\tilde{S}(\sigma)}$, $\text{BE}_{\hat{S}(\sigma)}$, $\text{PE}_{\tilde{S}(\rho)}$, $\text{PE}_{\hat{S}(\rho)}$, $\text{PE}_{\tilde{S}(\sigma)}$, and $\text{PE}_{\hat{S}(\sigma)}$.

IV. ESTIMATION OF QCB AND QHB

Let us now sketch the main idea behind our estimation algorithm, which builds on the QSVT algorithm and inexact Newton method.

To illustrate the concept of our algorithm, we use the estimation of QCB as an example. Since $h(s) = \text{Tr}(\rho^s \sigma^{1-s})$ is convex in $s \in [0, 1]$ [2], the function has a unique local minimum, which is also the global minimum. Based on this, we apply Newton's method to find the minimum value of $h(s)$ in $s \in [0, 1]$. However, since the QSVT algorithm introduces errors when computing the first and second derivatives of $h(s)$ using polynomial approximation, we use the inexact Newton method to find its minimum value:

$$x_{k+1} = x_k + s_k, \quad r_k = \ddot{f}(x_k)s_k + \dot{f}(x_k), \quad |r_k| \leq \mu_k |\dot{f}(x_k)|.$$

where s_k is the approximate step and r_k denotes the residual between the exact value $-\dot{f}(x_k)/\ddot{f}(x_k)$ and the approximate value s_k . We call μ_k a *forcing term* as the role of it is to force r_k to be small.

A. Diagonal density matrix

First, we show that we can compute the first and second derivatives of the functions $h(s)$ and $g(s)$ via QSVT. The explicit expressions for the first and second derivatives of $h(s)$ and $g(s)$ are given by

$$\begin{aligned} \dot{h}(s) &= \text{Tr}(\rho^s (\ln \rho) \sigma^{1-s} - \rho^s \sigma^{1-s} \ln \sigma), \\ \ddot{h}(s) &= \text{Tr}(\rho^s \sigma^{1-s} (\ln \rho)^2 - 2\rho^s (\ln \rho) \sigma^{1-s} \ln \sigma \\ &\quad + \rho^s \sigma^{1-s} (\ln \sigma)^2), \\ \dot{g}(s) &= \frac{-r - \ln \text{Tr}(\rho^s \sigma^{1-s})}{(1-s)^2} \\ &\quad + \frac{(\text{Tr}(\rho^s \sigma^{1-s} \ln \sigma) - \text{Tr}(\sigma^{1-s} \rho^s \ln \rho))}{(1-s) \text{Tr}(\rho^s \sigma^{1-s})}, \\ \ddot{g}(s) &= \frac{\text{Tr}[\rho^s \sigma^{1-s} (2 \ln \sigma \ln \rho - (\ln \sigma)^2 - (\ln \rho)^2)]}{(1-s) \text{Tr}(\rho^s \sigma^{1-s})} \\ &\quad + \frac{2 \text{Tr}[\rho^s \sigma^{1-s} (\ln \sigma - \ln \rho)]}{(1-s)^2 \text{Tr}(\rho^s \sigma^{1-s})} \\ &\quad + \frac{[\text{Tr}(\rho^s \sigma^{1-s} (\ln \sigma - \ln \rho))]^2}{(1-s) \text{Tr}^2(\rho^s \sigma^{1-s})} - \frac{2[r + \ln \text{Tr}(\rho^s \sigma^{1-s})]}{(1-s)^3}. \end{aligned}$$

Using the block encoding unitary and the projected encoding unitary from Lemmas 1 and 2, we can estimate $h(s)$, $g(s)$, $\dot{h}(s)$, $\dot{g}(s)$, $\ddot{h}(s)$, and $\ddot{g}(s)$ via QSVT. Since each of these quantities is a linear combination of individual terms, we can just estimate each term separately and then sum the results.

We take the estimation of the first term in $\dot{h}(s)$: $\text{Tr}(\rho^s (\ln \rho) \sigma^{1-s})$ as an example.

Theorem 5 (*Estimation of $\text{Tr}(\rho^s (\ln \rho) \sigma^{1-s})$*). Let $\rho, \sigma \in \mathbb{C}^{d \times d}$ be density matrices in computational basis. For any $\varepsilon \in (0, 1)$, we can estimate $\text{Tr}(\rho^s (\ln \rho) \sigma^{1-s})$ with accuracy ε with success probability at least $2/3$ using $O(\frac{r_m^{1/s}}{\varepsilon^{1/s+1}} (\ln \frac{1}{\varepsilon})^{1/s+2})$ calls to U_ρ , $U_\rho^\dagger$, U_σ , $U_\sigma^\dagger$, SWAP, and U_{copy} , where $\bar{s} := \min\{s, 1-s\}$, $r_m := \max\{\text{rank}(\rho), \text{rank}(\sigma)\}$.

The proof is deferred to Appendix C. Since the remaining terms in $h(s)$, $g(s)$, $\dot{h}(s)$, $\dot{g}(s)$, $\ddot{h}(s)$, and $\ddot{g}(s)$ can be derived in a

similar manner, all of these quantities can be estimated using QSVT.

Corollary 1 (Estimation of $h(s)$, $g(s)$, and their first and second derivatives). Let $\rho, \sigma \in \mathbb{C}^{d \times d}$ be density matrices in computational basis. For any $\varepsilon \in (0, 1)$, we can estimate $h(s)$, $g(s)$, $\dot{h}(s)$, $\dot{g}(s)$, $\ddot{h}(s)$, and $\ddot{g}(s)$ with accuracy ε with success probability at least $2/3$ using $O(\frac{r_m^{1/\bar{s}}}{\varepsilon^{1/\bar{s}+1}} (\ln \frac{1}{\varepsilon})^{1/\bar{s}+2})$ calls to U_ρ , $U_\rho^\dagger$, U_σ , $U_\sigma^\dagger$, SWAP, and U_{copy} , where $\bar{s} := \min\{s, 1-s\}$, $r_m := \max\{\text{rank}(\rho), \text{rank}(\sigma)\}$.

After computing the first and second derivatives, we apply the inexact Newton method to locate the extremum. Below, we show that the inexact Newton method exhibits local convergence when the forcing term is less than 1.

Lemma 8 ([17], Theorem 2.3). Assume $f: \mathbb{R} \rightarrow \mathbb{R}$ is continuously differentiable, $x^* \in \mathbb{R}$ such that $f(x^*) = 0$ and $\dot{f}(x^*) \neq 0$. Assume $\mu_k < 1$ in exact Newton method for all k , then there exists $\varepsilon > 0$, such that if $|x_0 - x^*| \leq \varepsilon$, the sequence of inexact Newton iterates $\{x_k\}$ converges to x^* and the sequence converges linearly.

For the estimation tasks of the QCB and QHB, at each iteration we can choose an appropriate forcing term η_k to satisfy the convergence conditions of the inexact Newton method.

Lemma 9 (Forcing terms in QCB and QHB estimation). We use $s_k^{\text{QCB}} = -\dot{h}_{\text{QSVT}}(x_k)/\ddot{h}_{\text{QSVT}}(x_k)$ and $s_k^{\text{QHB}} = -\dot{g}_{\text{QSVT}}(x_k)/\ddot{g}_{\text{QSVT}}(x_k)$ obtained from Corollary 1 as the iteration step for QCB and QHB estimation, respectively, such that $|\dot{h}_{\text{QSVT}}(x) - \dot{h}(x)| \leq \varepsilon_{\dot{h}}$, $|\ddot{h}_{\text{QSVT}}(x) - \ddot{h}(x)| \leq \varepsilon_{\ddot{h}}$, $|\dot{g}_{\text{QSVT}}(x) - \dot{g}(x)| \leq \varepsilon_{\dot{g}}$, and $|\ddot{g}_{\text{QSVT}}(x) - \ddot{g}(x)| \leq \varepsilon_{\ddot{g}}$. In each iteration, we set the forcing terms

$$\eta_k^{\text{QCB}} := \frac{\varepsilon_{\dot{h}}|\ddot{h}_{\text{QSVT}}(x_k)| + \varepsilon_{\ddot{h}}|\dot{h}_{\text{QSVT}}(x_k)| + 2\varepsilon_{\dot{h}}\varepsilon_{\ddot{h}}}{|\dot{h}_{\text{QSVT}}(x_k) - \varepsilon_{\dot{h}}|\ddot{h}_{\text{QSVT}}(x_k)},$$

$$\eta_k^{\text{QHB}} := \frac{\varepsilon_{\dot{g}}|\ddot{g}_{\text{QSVT}}(x_k)| + \varepsilon_{\ddot{g}}|\dot{g}_{\text{QSVT}}(x_k)| + 2\varepsilon_{\dot{g}}\varepsilon_{\ddot{g}}}{|\dot{g}_{\text{QSVT}}(x_k) - \varepsilon_{\dot{g}}|\ddot{g}_{\text{QSVT}}(x_k)}.$$

Then in each iteration, $r_k^{\text{QCB}} \leq \eta_k^{\text{QCB}}|\dot{h}(x_k)|$ and $r_k^{\text{QHB}} \leq \eta_k^{\text{QHB}}|\dot{g}(x_k)|$.

The proof of Lemma 9 is provided in Appendix D. To ensure $\eta_k < 1$ in each iteration—thereby guaranteeing linear convergence of the inexact Newton method—we adopt the following strategy in our simulations: if $\eta_k \geq 1$, we halve the tolerance and repeat the computation until the condition $\eta_k < 1$ is satisfied.

The main steps of the algorithm for estimating the QCB using the inexact Newton method and QSVT are as follows. The procedure for estimating the QHB is analogous:

- (1) Select an initial value x_0 , the maximum number of iterations N , and a tolerance ε .
- (2) Calculate the derivative of $h(s)$ by QSVT at $s = x_0$.
- (3) If the derivative is smaller than ε , conclude that the value of s maximizing the QCB has been found. If the derivative is not smaller than ε , compute the second derivative of $h(s)$ using QSVT.
- (4) Set the forcing term as in Lemma 9, and make sure it is less than 1 as discussed above. Use inexact Newton method to update the iteration step.
- (5) Repeat the process until the first derivative is smaller than ε or the maximum number of iterations N is reached.

ALGORITHM 1. Estimation of QCB and QHB with inexact Newton method.

Input: Max iter N , tol ε , init $x_0 = y_0$, $r = 0.5$,
#queries k

Output: ε -biased estimation of QCB & QHB

while $\dot{g}(r) < 0$ **do**
 $r \leftarrow r/2$;
 $k \leftarrow 0$;
 Compute $\dot{h}(x_k), \ddot{h}(x_k), \dot{g}(y_k), \ddot{g}(y_k)$;
 while $k < N \wedge |\dot{h}(x_k)| > \varepsilon \wedge |\dot{g}(y_k)| > \varepsilon$ **do**
 repeat // Ensure $\eta_k^{\text{QCB}}, \eta_k^{\text{QHB}} < 1$
 if $\eta_k^{\text{QCB}} \geq 1$ **or** $\eta_k^{\text{QHB}} \geq 1$ **then**
 $\varepsilon \leftarrow \varepsilon/2$;
 Recompute $\dot{h}(x_k), \ddot{h}(x_k), \dot{g}(y_k), \ddot{g}(y_k)$;
 until Compute $\eta_k^{\text{QCB}}, \eta_k^{\text{QHB}}$ using Lemma 9;
 ;
 $x_{k+1} \leftarrow x_k - \frac{\dot{h}(x_k)}{\ddot{h}(x_k)}$;
 $y_{k+1} \leftarrow y_k - \frac{\dot{g}(y_k)}{\ddot{g}(y_k)}$;
 if $|x_{k+1} - x_k| < \varepsilon$ **then**
 $\text{return } x^*$
 if $|y_{k+1} - y_k| < \varepsilon$ **then**
 $\text{return } y^*$
 $k \leftarrow k + 1$;
 Compute $\dot{h}(x_k), \ddot{h}(x_k), \dot{g}(y_k), \ddot{g}(y_k)$;
 return $x^* = x_N, y^* = y_N$ (or failure)

(6) Denote the obtained value as $\hat{x}$. Finally, the QCB is obtained by evaluating $h_{\text{QSVT}}(\hat{x})$.

We present the pseudocode for solving the QCB and QHB using the QSVT algorithm and the inexact Newton method in Algorithm 1. By combining Theorem 5 and Lemma 9 with the convergence result of the inexact Newton method from Lemma 8, we obtain the following theorem.

Theorem 6 (Estimation of QCB and QHB for diagonal density matrices). Let $\rho, \sigma \in \mathbb{C}^{d \times d}$ be diagonal density matrices in computational basis with smallest eigenvalues no less than δ . For any $\varepsilon \in (0, 1)$, we can estimate QCB and QHB with accuracy ε with success probability at least $2/3$ using $O(\frac{r_m^{1/\bar{s}}}{\varepsilon^{1/\bar{s}+1}} (\ln \frac{1}{\varepsilon})^{1/\bar{s}+3})$ calls to U_ρ , $U_\rho^\dagger$, U_σ , $U_\sigma^\dagger$, SWAP, and U_{copy} , where $\bar{s} = \min\{s, 1-s\}$, $r_m = \max\{\text{rank}(\rho), \text{rank}(\sigma)\}$.

Note that the requirement on the minimum eigenvalue of the density matrices does not arise from the polynomial approximation itself, but from the fact that the value returned by the inexact Newton method may slightly deviate from the true extremum. To tightly control the final result, we impose a Lipschitz condition, which necessitates a lower bound on the smallest eigenvalues of ρ and σ . For the detailed proof of Theorem 6, see Appendix E.

When applying the inexact Newton method to estimate the QHB, we must ensure that the parameter r does not exceed the relative entropy between ρ and σ , defined as $D(\rho||\sigma) = \text{Tr}(\rho(\ln \rho - \ln \sigma))$ (see Sec. II B). In practice, we initialize r

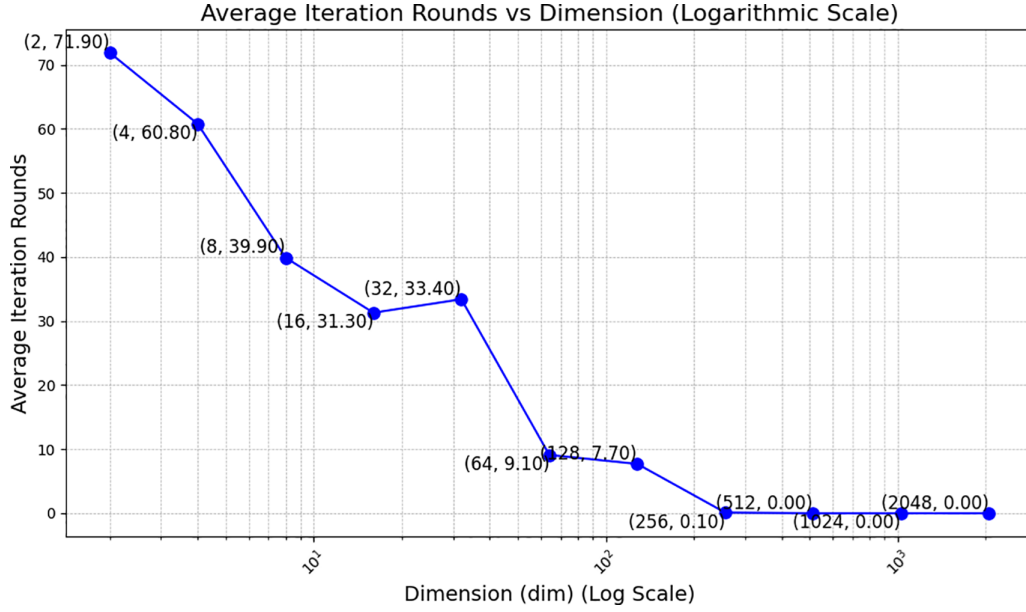


FIG. 2. QCB estimation. The horizontal axis represents the dimension of the matrix, and the vertical axis represents the average number of iterations, repeated 10 times at the same dimension.

randomly. If $\dot{g}(0)$ is found to be negative, this indicates that $r > D(\rho \parallel \sigma)$. In such cases, we iteratively halve the value of r until $\dot{g}(0) \geq 0$.

Although we do not formally prove the convexity of the function $g(s) = (-rs - \ln \text{Tr}(\rho^s \sigma^{1-s})) / (1-s)$, numerical simulations suggest that $g(s)$ is concave over $s \in [0, 1)$ under suitable conditions on r . Consequently, a similar procedure can be used for estimating the QHB. For further details, see Sec. V.

B. General density matrix

When the density matrix is not diagonal, we employ the Hadamard test circuit to estimate the QCB and QHB, thereby eliminating the redundant and unknown information introduced by the block encoding matrix. Given a real polynomial P and purified access U_ρ , the Hadamard test provides an unbiased estimator of $\text{Tr}(P(\rho))$.

By Lemma 1, we can construct a unitary

$$\text{BE}_{P(\rho)} = \sum_{i=0}^{d-1} P(p_i) |0\rangle\langle 0|_{a+t} \otimes |\psi_i\rangle\langle \psi_i|_t,$$

using the QSVT algorithm together with Lemma 3. A controlled version of $\text{BE}_{P(\rho)}$ is given by ([23], Lemma 19)

$$C\text{-BE}_{P(\rho)} := |0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes \text{BE}_{P(\rho)}.$$

The Hadamard test proceeds as follows:

(1) Prepare the initial state $|+\rangle \otimes |0\rangle_{a+t} \otimes |\Phi\rangle$, where $|\Phi\rangle = (1/\sqrt{d}) \sum_{i=0}^{d-1} |i\rangle_t |i\rangle_b$. Here, b denotes the ancillary qubits, and $|\Phi\rangle$ can be generated using Hadamard and CNOT gates.

(2) Apply $C\text{-BE}_{P(\rho)}$ to the initial state.

(3) Apply a Hadamard gate on the control qubit and then measure it in the Z basis to obtain $Y \in \{\pm 1\}$.

It follows that $\mathbb{E}[Y]$ is an unbiased estimator of $(1/d)\text{Tr}(P(\rho))$.

Theorem 7 (Unbiased estimation of $\text{Tr}(P(\rho))$). For any $\varepsilon \in (0, 1)$, $\text{Tr}(P(\rho))$ can be estimated to accuracy ε with success probability at least $2/3$ using $O(d/\varepsilon)$ calls to $C\text{-BE}_{P(\rho)}$.

The proof of Theorem 6 is given in Appendix F. Since the query complexity of polynomial approximation depends only on the polynomial degree, independent of the encoding scheme, we adopt the degree bound from Appendix C. Combining the convergence of the inexact Newton method (Lemma 8), Lemma 9, and Theorem 6, we obtain the following.

Theorem 8 (Estimation of QCB and QHB in the general case). Let $\rho, \sigma \in \mathbb{C}^{d \times d}$ be arbitrary density matrices with minimum eigenvalue at least δ . For any $\varepsilon \in (0, 1)$, the QCB and QHB can be estimated to accuracy ε with success probability at least $2/3$ using

$$O\left(d \frac{r_m^{1/\bar{s}}}{\varepsilon^{1/\bar{s}+1}} \left(\ln \frac{1}{\varepsilon}\right)^{1/\bar{s}+3}\right)$$

calls to $U_\rho, U_\rho^\dagger, U_\sigma, U_\sigma^\dagger$, and SWAP, where $\bar{s} = \min\{s, 1-s\}$ and $r_m = \max\{\text{rank}(\rho), \text{rank}(\sigma)\}$.

V. NUMERICAL RESULTS

Our method is implemented in Python. In the QSVT-based estimation of the QCB and QHB, we do not construct purification matrices as input to simulate the QSVT quantum circuit; instead, we randomly generate density matrices and directly apply the target-function transformation.

To account for potential errors in derivative estimation via QSVT, we add random noise within the range $[-\varepsilon, \varepsilon]$ to the matrix on which the function transformation is directly performed, consistent with theoretical error bounds.

A. Parameter settings

For QCB estimation, we set the initial value x_0 to 0.5, as numerical experiments indicate that $h(s)$ typically attains its minimum near $s = 0.5$ as the number of qubits increases. This choice of initial value helps reduce the number of iterations significantly.

For QHB estimation, we also set the initial value y_0 to 0.5, since the value of s that maximizes $g(s)$ tends to vary unpredictably with changes in ρ and σ . Consequently, regardless of the initial choice, the number of iterations tends to grow with the dimension of the density matrices.

Moreover, the estimation error for the QCB, as well as for computing $\dot{h}(s)$, $\ddot{h}(s)$, $\dot{g}(s)$, and $\ddot{g}(s)$ using QSVT, is set to 10^{-3} in all cases. The maximum number of iterations is set to 100.

B. Numerical results of QCB and QHB estimation

From Fig. 2, we observe that the average number of iterations required for convergence does not exceed 100. When the matrix dimension is below 64×64 , the number of iterations needed to find the optimal value varies significantly. However, as the dimension increases, the relative average number of iterations tends to decrease.

Moreover, we find that for sufficiently large density matrices, the optimal point of the objective function is consistently attained at $s = 0.5$. This phenomenon is illustrated in Figs. 3(a), 3(b), and 3(c), corresponding to dimensions 256, 512, and 1024, respectively.

From Fig. 4, we observe that the average number of iterations required for convergence does not exceed 100. When the matrix dimension is below 16×16 , the number of iterations needed to find the optimal value is unstable. However, the overall trend indicates that the number of iterations increases with the matrix dimension.

The above simulation experiments are conducted using density matrices generated from a uniform distribution.

VI. CONCLUSIONS

In this work, we develop hybrid quantum-classical algorithms for computing the quantum Chernoff Bound and quantum Hoeffding Bound by integrating quantum singular value transformation with convex optimization. Using QSVT, we introduce a unitary encoding scheme for diagonal density matrix based on the copy matrix, which eliminates the bias present in earlier approaches. For general density matrices, we combine QSVT with the Hadamard test to enable estimation. In addition, we construct an explicit polynomial approximation of $\sqrt{f(x)}$ for any function $f(x)$ that admits a Taylor expansion. This synergy between QSVT and optimization techniques might have broad applications in quantum parameter estimation and quantum machine learning.

To identify optimal points, we employ the inexact Newton method, which exploits the structure of the objective functions. Numerical simulations indicate that the algorithm typically converges to the optimum within 100 iterations. In QCB estimation, convergence is faster for larger matrix dimensions because the initial value is already close to the optimum. These results suggest that, asymptotically, the optimal value of QCB is achieved at $s = 0.5$.

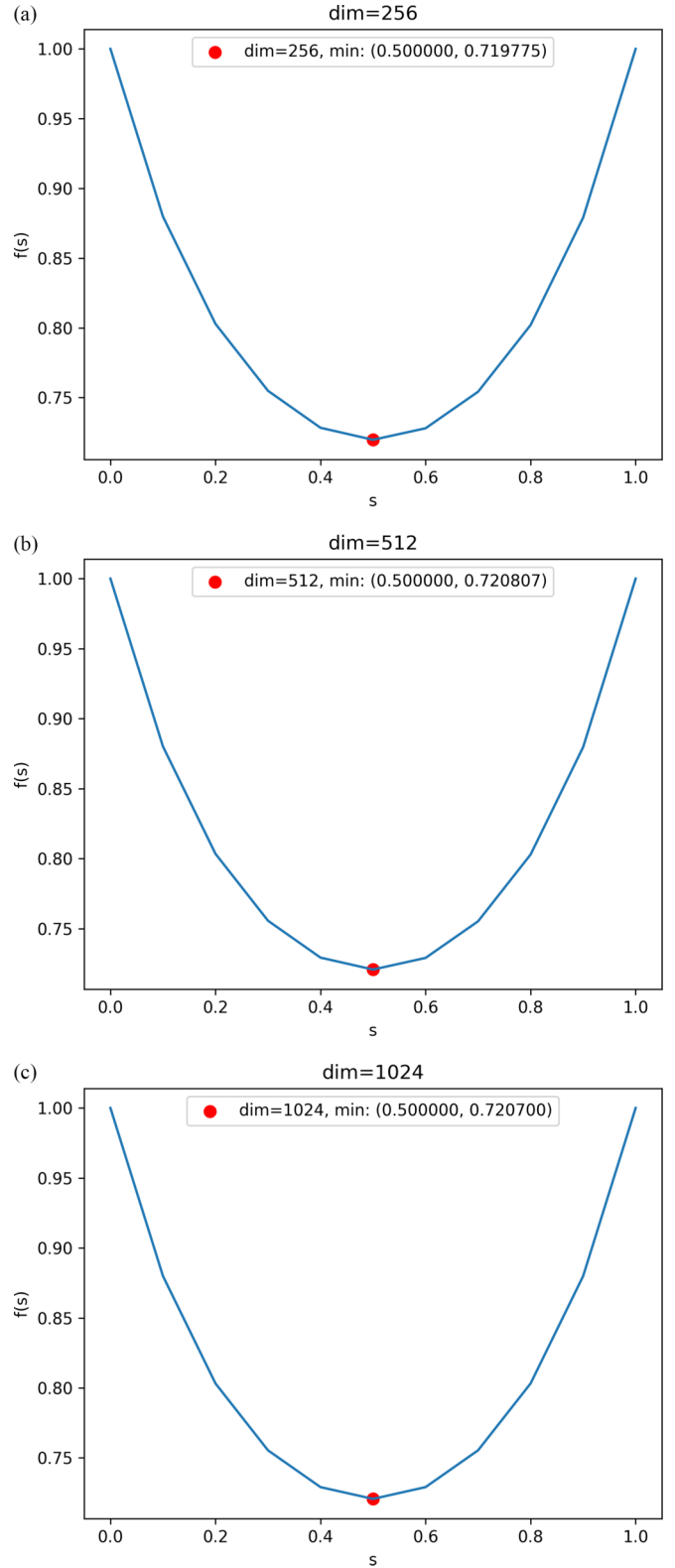


FIG. 3. QCB estimation in different dimensions. The horizontal axis represents the value of s , and the $f(s)$ in vertical axis represents $\text{Tr}(\rho^s \sigma^{1-s})$. If the dimension of density matrix is large enough, the minimum value of $f(s)$ is attained at $s = 0.5$. (a) QCB estimation of a random 256×256 density matrix. (b) QCB estimation of a random 512×512 density matrix. (c) QCB estimation of a random 1024×1024 density matrix.

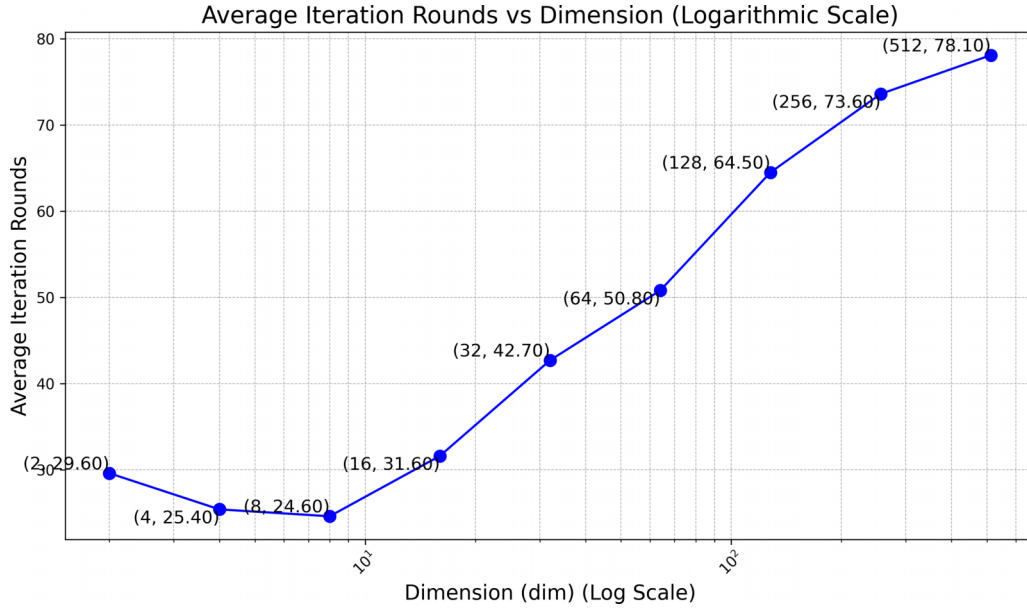


FIG. 4. QHB estimation. The horizontal axis represents the dimension of the matrix, all being powers of 2, and the vertical axis represents the average number of iterations for solving QHB, repeated 10 times at the same dimension.

ACKNOWLEDGMENTS

We acknowledge helpful comments from the anonymous referees, which improved our paper. This work was partially supported by the Innovation Program for Quantum Science and Technology (Grant No. 2021ZD0302901) and the National Natural Science Foundation of China (Grant No. 62102388).

DATA AVAILABILITY

No data were created or analyzed in this study.

APPENDIX A: THE PROOF OF LEMMA 6

Proof. We consider the case where the polynomial degree in Lemma 6 is even; the odd-degree case follows similarly. Let $t := 3\delta/2$, $\delta' := \delta/2$, and $\varepsilon' := \varepsilon/2$, and define $P''(x) := 1 - P'(x)$ as in Lemma 5. Thus, $P''(x)$ is an even polynomial such that $|P''(x)| \leq 1$ for all $x \in [-1, 1]$, and satisfies the following:

$$\begin{aligned} P''(x) &\in \left[0, \frac{\varepsilon}{2}\right], & \text{for all } x \in [-\delta, \delta], \\ P''(x) &\in \left[1 - \frac{\varepsilon}{2}, 1\right], & \text{for all } x \in [-1, -2\delta] \cup [2\delta, 1]. \end{aligned}$$

Next, choose the parameters in Lemma 5 as $\delta' := \delta$ and $\varepsilon' := \varepsilon$. Define $\tilde{P}(x) := P''(x) \cdot P(x)$. We claim that $\tilde{P}(x)$ satisfies the desired properties. It is evident that $\|\tilde{P}(x)\| \leq 1$ by applying the triangle inequality. Moreover, we have

$$\begin{aligned} \|\tilde{P}(x) - f(x)\|_{[0, \delta]} &\leq \|P''(x)(P(x) - f(x))\|_{[0, \delta]} \\ &\quad + \|(P''(x) - 1)f(x)\|_{[0, \delta]} \\ &\leq \|P''(x)\|_{[0, \delta]} \|P(x) - f(x)\|_{[0, \delta]} \\ &\quad + \|P''(x) - 1\|_{[0, \delta]} \|f(x)\|_{[0, \delta]} \end{aligned}$$

$$\leq \frac{\varepsilon}{2} \cdot 2 + 1 \cdot f(x)$$

$$= f(x) + \varepsilon,$$

$$\begin{aligned} \|\tilde{P}(x) - f(x)\|_{[\delta, 1]} &\leq \|P''(x)(P(x) - f(x))\|_{[\delta, 1]} \\ &\leq \|P''(x)\|_{[\delta, 1]} \|P(x) - f(x)\|_{[\delta, 1]} \\ &\leq \varepsilon. \end{aligned}$$

Note that $\tilde{P}(x)$ is an even polynomial, so for all $x \in [-1, 1]$, we have $|\tilde{P}(x)| \leq f(|x|) + \varepsilon$ by the inequalities above. ■

APPENDIX B: THE PROOF OF LEMMA 7

We extend Lemma 37 of Ref. [16] to fit our situation.

Lemma B1 (Low weight approximation Fourier-like series). Let $\delta, \varepsilon \in (0, 1)$ and $f : \mathbb{R} \rightarrow \mathbb{C}$ such that $|f(x) - \sum_{k=0}^K a_k x_k| \leq \varepsilon/4$ for all $x \in [-1 + \delta, 1 - \delta]$. Then, there exists $c \in \mathbb{C}^{2M+1}$ such that

$$\left| f(x) - \sum_{m=-M}^M c_m e^{i\pi m/2 x} \right| \leq \varepsilon,$$

for all $x \in [-1 + \delta, 1 - \delta]$, where $M = \max(2\lceil \ln(\frac{4\|a\|_1}{\varepsilon}) \rceil, 0)$ and $\|c\|_1 \leq \|a\|_1$. Moreover c can be efficiently calculated on a classical computer in time $\text{poly}(K, M, \ln(1/\varepsilon))$.

First, we show that $\sqrt{f(x)}$ can be expanded in the form of $\sum_{m=0}^{\infty} c_m x^m$.

Lemma B2 (Expansion of $\sqrt{f(x)}$). Let $x_0 \in [-1, 1]$, $r' \in (0, 2]$, $\delta \in (0, r']$ and let $f : [x_0 - r' - \delta, x_0 + r' + \delta] \rightarrow \mathbb{C}$ and be such that $f(x_0 + x) = \sum_{l=0}^{\infty} a_l x^l$ for all $x \in [-r' - \delta, r' + \delta]$. Then, we have $\sqrt{f(x_0 + x)} = \sum_{m=0}^{\infty} c_m x^m$, where $c_0 = \sqrt{a_0}$, $c_m = \frac{1}{m\sqrt{a_0}} \sum_{l=1}^m (l/2 - m + l) a_l c_{m-l}$, $m \geq 1$.

Proof. We have that

$$\begin{aligned}\sqrt{f(x_0 + x)} &= \sqrt{\sum_{l=0}^{\infty} a_l x^l} \\ &= \sqrt{a_0} \sqrt{\sum_{l=0}^{\infty} a'_l x^l} \\ &= \sum_{m=0}^{\infty} c_m x^m,\end{aligned}$$

where in the second equation $a'_l = a_l/a_0$. Moreover, in the third equation we use the expansion for power series raised to powers: $(\sum_{l=0}^{\infty} a'_l x^l)^{1/2} = \sum_{m=0}^{\infty} c'_m x^m$, in which $c'_0 = a'_0{}^{1/2} = 1$, $c'_m = \frac{1}{m} \sum_{l=1}^m (l/2 - m + l) a'_l c'_{m-l}$, $m \geq 1$. Thus, we have $c_0 = \sqrt{a_0}$, $c_m = \sqrt{a_0} c'_m = \frac{1}{m\sqrt{a_0}} \sum_{l=1}^m (l/2 - m + l) a_l c_{m-l}$. ■

To prove our conclusion, we modify the Corollary 66 of Ref. [23] to extend it to the case of $\sqrt{f(x)}$.

Lemma B3 (Bounded polynomial approximations based on a local Taylor series). Assume we have the function $f(x)$ from Lemma B2 such that $\sqrt{f(x_0 + x)} = \sum_{m=0}^{\infty} c_m x^m$. Suppose $B > 0$ is such that $\sum_{m=0}^{\infty} (r' + \delta)^m |c_m| \leq B$. Let $\varepsilon \in (0, (1/2B)]$, then there is an efficiently computable polynomial $P \in \mathbb{C}[x]$ of degree $O(\frac{r'}{\delta} \ln(\frac{B}{\varepsilon}))$ such that

$$\|\sqrt{f(x)} - P(x)\|_{[x_0 - r', x_0 + r']} \leq \varepsilon,$$

$$\|P(x)\|_{[-1, 1]} \leq \varepsilon + \|\sqrt{f(x)}\|_{[x_0 - r' - \delta/2, x_0 + r' + \delta/2]} \leq \varepsilon + B,$$

$$\|P(x)\|_{[-1, 1] \setminus [x_0 - r' - \delta/2, x_0 + r' + \delta/2]} \leq \varepsilon.$$

Proof. Our proof is similar to Corollary 66 of Ref. [23]. Let $L(x) := (x - x_0)/(r' + \delta)$ be the linear transformation taking $[x_0 - r' - \delta, x_0 + r' + \delta]$ to $[-1, 1]$. Let $g(y) := \sqrt{f(L^{-1}(y))}$, and $b_m := c_m(r' + \delta)^m$ such that $g(y) = \sum_{m=0}^{\infty} b_m y^m$. Let $\delta' := \delta/2(r' + \delta)$, $J := \lceil (1/\delta') \ln(24B/\varepsilon) \rceil$, then for all $y \in [-1 + \delta', 1 - \delta']$ we have that

$$\begin{aligned}\left| g(y) - \sum_{j=0}^{J-1} b_j y^j \right| &= \left| \sum_{j=J}^{\infty} b_j y^j \right| \\ &\leq \left| \sum_{j=J}^{\infty} b_j (1 - \delta')^j \right|\end{aligned}$$

$$\begin{aligned}&\leq (1 - \delta')^J \sum_{j=J}^{\infty} |b_j| \\ &\leq e^{-\delta' J} \cdot B \\ &\leq \frac{\varepsilon}{24}.\end{aligned}$$

Now, we construct a Fourier approximation of g of all $y \in [-1 + \delta', 1 - \delta']$, with precision $\varepsilon' = \varepsilon/3$. Let $b' := (b_0, b_1, \dots, b_{J-1})$ and $\|b'\|_1 \leq \|b\|_1 \leq B$.

Therefore, we can apply Lemma B1 to the function g , utilizing the polynomial approximation of g' corresponding to the truncation of the first $2J - 1$ terms with the coefficients e' . This results in a Fourier $\varepsilon/3$ approximation $\tilde{g}(y) := \sum_{m=-M}^M \tilde{c}_m e^{(i\pi m/2)y}$ for g , with

$$M = O\left(\frac{1}{\delta'} \ln\left(\frac{\|b'\|_1}{\varepsilon'}\right)\right) = O\left(\frac{r'}{\delta} \ln\left(\frac{B}{\varepsilon}\right)\right).$$

The remaining proof process is consistent with that in Corollary 66 of Ref. [23]. Therefore, the degree of the final constructed polynomial is

$$O\left(\frac{M}{r' + \delta} + \ln\left(\frac{B}{\varepsilon}\right) + \frac{1}{\delta} \ln\left(\frac{B}{\varepsilon}\right)\right) = O\left(\frac{r'}{\delta} \ln\left(\frac{B}{\varepsilon}\right)\right),$$

and by construction it satisfies the required properties. ■

We can use the above result to construct the polynomial approximation of $\sqrt{\ln x}$ up to a constant factor.

The Proof of Lemma 7. The $\hat{S}$ part has already been proven in Lemma 11 of Ref. [25]. We mainly focus on proving $\tilde{S}$ part and is similar to Lemma 11 of Ref. [25]. Denote $f(x) = \sqrt{\ln(1/x)/2} \sqrt{\ln(2/\delta)}$. We take $\varepsilon' = \varepsilon/2$, $x_0 = 2/3$, $r' = (2/3) - \delta$, $\delta' = \delta/2$, and $B = 1/2$ in Lemma B3; we have a polynomial $S \in \mathbb{C}[x]$ of degree $O(\frac{r'}{\delta} \ln(\frac{B}{\varepsilon})) = O(\frac{1}{\delta} \ln(\frac{1}{\varepsilon}))$ such that

$$\|f(x) - S(x)\|_{[\delta, \frac{4}{3} - \delta]} \leq \varepsilon', \quad (\text{B1})$$

$$\|S(x)\|_{[-1, 1]} \leq \frac{1}{2} + \varepsilon', \quad (\text{B2})$$

$$\|S(x)\|_{[-1, \frac{3\delta}{4}]} \leq \varepsilon'. \quad (\text{B3})$$

Note that $\sqrt{-\ln(1+x)} = \sum_{m=0}^{\infty} c_m x^m$, where $c_0 = \ln 2/3$, and for $m \geq 1$, $c_m = \frac{1}{m} \sum_{k=1}^m \left(\frac{3k}{2} - m\right) \frac{(-1)^k}{k+1} \left(\frac{3}{2}\right)^{k+1} c_{m-k}$. Then, we can give the bound of the expansion coefficient for $f(x)$:

$$\begin{aligned}\frac{1}{2\sqrt{\ln(2/\delta)}} \sum_{m=0}^{\infty} \left(\frac{2}{3} - \frac{\delta}{2}\right)^m |c_m| &\leq \frac{1}{2\sqrt{\ln(2/\delta)}} \left[\ln \frac{2}{3} + \sum_{m=1}^{\infty} \left| \frac{1}{m} \sum_{k=1}^m \left(\frac{3k}{2} - m\right) \frac{(-1)^k}{k+1} \left(\frac{3}{2}\right)^{k+1} c_{m-k} \right| \left(\frac{2}{3} - \frac{\delta}{2}\right)^m \right] \\ &= \frac{1}{2\sqrt{\ln(2/\delta)}} \left\{ 1 + \sum_{m=1}^{\infty} \frac{(-1)^m}{m} \left(\sum_{k=1}^m \left(\frac{3k}{2} - m\right) \frac{(-1)^k}{k+1} \left(\frac{3}{2}\right)^{k+1} c_{m-k} \right) \left[(-1)^m \left(\frac{\delta}{2} - \frac{2}{3}\right)^m \right] \right\} \\ &= \frac{1}{2\sqrt{\ln(2/\delta)}} \left[1 + \sum_{m=1}^{\infty} \frac{1}{m} \sum_{k=1}^m \left(\frac{3k}{2} - m\right) \frac{(-1)^k}{k+1} \left(\frac{3}{2}\right)^{k+1} c_{m-k} \left(\frac{\delta}{2} - \frac{2}{3}\right)^m \right] \\ &= \frac{1}{2\sqrt{\ln(2/\delta)}} \sqrt{-\ln\left(\frac{\delta}{2}\right)} \\ &= \frac{1}{2}.\end{aligned}$$

In the fifth equation, we observe that c_m is negative in the odd terms and positive in the even terms, so we remove the absolute value and add the term $(-1)^m$. Therefore, we can set B to $1/2$. However, in general $S(x)$ is not an even polynomial, so we take $\tilde{S}(x) = S(x) + S(-x)$ for all $x \in [-1, 1]$. Then by Inequalities (B1) and (B3), we have

$$\begin{aligned} \|f(x) - \tilde{S}(x)\|_{[\delta, 1]} &\leq \|f(x) - S(x)\|_{[\delta, 1]} + \|S(-x)\|_{[\delta, 1]} \\ &\leq \varepsilon' + \varepsilon' \\ &= \varepsilon. \end{aligned}$$

We can now safely claim that $\tilde{S}(x)$ is an even polynomial with degree $\deg(\tilde{S}) = O(\frac{1}{\delta} \ln(\frac{1}{\varepsilon}))$. Furthermore, Inequalities (B2) and (B3) imply that

$$\begin{aligned} \|\tilde{S}(x)\|_{[-1, 1]} &= \|\tilde{S}(x)\|_{[0, 1]} \\ &\leq \|S(x)\|_{[0, 1]} + \|S(-x)\|_{[0, 1]} \\ &\leq \frac{1}{2} + \varepsilon' + \|S(x)\|_{[-1, 0]} \\ &\leq \frac{1}{2} + \varepsilon' + \varepsilon' \\ &\leq 1, \end{aligned}$$

where $\varepsilon \leq 1/2$. Finally, we take the real part of $\tilde{S}(x)$, since $\tilde{S}(x) \in \mathbb{C}[x]$. ■

APPENDIX C: THE PROOF OF THEOREM 5

Proof. The main idea is to first construct one block encoding matrix $\text{BE}_{\tilde{S}(\rho)}$ and two projected unitary encoding $\text{PE}_{P_s(\rho)}$, $\text{PE}_{P_{1-s}(\sigma)}$ through QSVT. We multiply the three matrices and apply them to initial state $|0\rangle_{a+2t}$.

$$\begin{aligned} |\psi_{\text{final}}\rangle &= \text{PE}_{P_{1-s}(\sigma)} \text{BE}_{\tilde{S}(\rho)} \text{PE}_{P_s(\rho)} |0\rangle \\ &= (\text{PE}_{P_{1-s}(\sigma)} \text{BE}_{\tilde{S}(\rho)}) \left(\sum_i P_s(p_i) |0\rangle |0\rangle |\psi_i\rangle + |\perp\rangle \right) \\ &= \text{PE}_{P_{1-s}(\sigma)} \left(\sum_i \tilde{S}(p_i) P_s(p_i) |0\rangle |0\rangle |\psi_i\rangle + |\perp\rangle \right) \\ &= \sum_{i,j} P_{1-s}(q_j) \tilde{S}(p_i) P_s(p_i) \langle \eta_j | \psi_i \rangle |0\rangle |0\rangle |0\rangle + |\perp\rangle, \end{aligned}$$

where $\rho := \sum_i p_i |\psi_i\rangle \langle \psi_i|$, $\sigma := \sum_i q_i |\eta_i\rangle \langle \eta_i|$ are the eigenvalue decompositions for ρ and σ , respectively, and here $|\psi_i\rangle = |\eta_i\rangle = |i\rangle$. We measure the final state $|\psi_{\text{final}}\rangle$ in the Z basis, and show that the probability $p(0)$ of obtaining result 0 is an ε approximation of $(1/4K) \text{Tr}(\rho^s (\ln(1/\rho)) \sigma^{1-s})$. With the lemmas in low-degree polynomial approximation, we have

$$\begin{aligned} &\left| p(0) - \frac{1}{16\tilde{K}^2} \text{Tr} \left(\rho^s \left(\ln \frac{1}{\rho} \right) \sigma^{1-s} \right) \right| \\ &= \left| \sum_{i,j} P_{1-s}^2(q_j) \tilde{S}^2(p_i) P_s^2(p_i) |\langle \eta_j | \psi_i \rangle|^2 - \sum_{i,j} \frac{p_i^s (\ln \frac{1}{p_i}) q_j^{1-s}}{16\tilde{K}^2} |\langle \eta_j | \psi_i \rangle|^2 \right| \\ &\leq \sum_{i,j} \left| P_{1-s}^2(q_j) \tilde{S}^2(p_i) P_s^2(p_i) - \frac{p_i^s (\ln \frac{1}{p_i}) q_j^{1-s}}{16\tilde{K}^2} \right| \\ &\leq \sum_{i,j} \left| P_{1-s}^2(q_j) \tilde{S}^2(p_i) P_s^2(p_i) - \frac{q_j^{1-s} \tilde{S}^2(p_i) P_s^2(p_i)}{4} \right| + \sum_{i,j} \left| \frac{q_j^{1-s}}{4} \tilde{S}^2(p_i) P_s^2(p_i) - \frac{p_i^s (\ln \frac{1}{p_i}) q_j^{1-s}}{16\tilde{K}^2} \right| \\ &\leq \sum_{i,j} \left| P_{1-s}^2(q_j) - \frac{q_j^{1-s}}{4} \right| |\tilde{S}^2(p_i) P_s^2(p_i)| + \sum_{i,j} \left| \frac{q_j^{1-s}}{4} \right| \left| \tilde{S}^2(p_i) P_s^2(p_i) - \frac{p_i^s (\ln \frac{1}{p_i})}{4\tilde{K}^2} \right| \\ &\leq \sum_j \left| P_{1-s}(q_j) - \frac{q_j^{\frac{1-s}{2}}}{2} \right| \left(|P_{1-s}(q_j)| + \left| \frac{q_j^{\frac{1-s}{2}}}{2} \right| \right) + \frac{1}{4} \sum_i \left(\left| \tilde{S}^2(p_i) P_s^2(p_i) - \tilde{S}^2(p_i) \frac{p_i^s}{4} \right| + \left| \tilde{S}^2(p_i) \frac{p_i^s}{4} - \frac{p_i^s (\ln \frac{1}{p_i})}{4\tilde{K}^2} \right| \right) \\ &\leq \sum_{q_j < \delta} \left(|P_{1-s}(q_j)| + \frac{q_j^{\frac{1-s}{2}}}{2} \right)^2 + \sum_{q_j \geq \delta} 2\varepsilon' + \sum_i |\tilde{S}^2(p_i)| \left| P_s(p_i) - \frac{p_i^{\frac{s}{2}}}{2} \right| \left(|P_s(p_i)| + \left| \frac{p_i^{\frac{s}{2}}}{2} \right| \right) \\ &\quad + \sum_i \frac{p_i^s}{4} \left| \tilde{S}(p_i) - \frac{\sqrt{\ln \frac{1}{p_i}}}{\tilde{K}} \right| \left| \tilde{S}(p_i) + \frac{\sqrt{\ln \frac{1}{p_i}}}{\tilde{K}} \right| \\ &\leq \sum_{q_j < \delta} (q_j^{\frac{1-s}{2}} + \varepsilon')^2 + 2r_m \varepsilon' + \sum_{p_i < \delta} (p_i^{\frac{s}{2}} + \varepsilon')^2 + 2r_m \varepsilon' + \sum_{p_i < \delta} \frac{p_i^s}{4} \left(|\tilde{S}(p_i)| + \left| \frac{\sqrt{\ln \frac{1}{p_i}}}{\tilde{K}} \right| \right)^2 + \sum_{p_i \geq \delta} \frac{p_i^s}{4} \cdot \varepsilon' \cdot \left(|\tilde{S}(p_i)| + \left| \frac{\sqrt{\ln \frac{1}{p_i}}}{\tilde{K}} \right| \right). \end{aligned}$$

The first inequality follows from the absolute value inequality and the fact that $|\langle \eta_j | \psi_i \rangle| \leq 1$. The second uses the triangle inequality. The fourth and fifth inequalities apply the difference-of-squares identity and the absolute value inequality, noting that the maximum value of the approximation polynomial is at most 1. In the last inequality, we apply the same reasoning to the third term in the fifth inequality as we did for the first two terms, yielding the third and fourth terms in the final bound.

Without loss of generality, we define $\varepsilon' := (\varepsilon/85r_m) < 1/2$, $\delta := (\frac{\varepsilon}{12r_m \ln \frac{r_m}{\varepsilon}})^{1/\bar{s}}$. Then $\deg(\tilde{P}_s) = \deg(\tilde{P}_{1-s}) = \deg(\tilde{S}) = O((\frac{r_m}{\varepsilon})^{1/\bar{s}}(\ln \frac{r_m}{\varepsilon})^{1/\bar{s}+1})$. Provided that $r_m/\varepsilon \geq 24$, we have

$$\begin{aligned}\tilde{K} &= 2\sqrt{\ln \frac{2}{\delta}} \\ &= 2\sqrt{\ln 2 + \ln \left[\left(\frac{12r_m}{\varepsilon} \ln \frac{r_m}{\varepsilon} \right)^{1/\bar{s}} \right]} \\ &\geq 2\sqrt{\ln 2 + \frac{1}{\bar{s}} \ln 2} \\ &= \sqrt{\frac{4 \ln 2(\bar{s} + 1)}{\bar{s}}} \\ &\geq \sqrt{\frac{2(\bar{s} + 1)}{\bar{s}}}.\end{aligned}$$

Thus, we have

$$\begin{aligned}&\left| p(0) - \frac{1}{16\tilde{K}^2} \text{Tr} \left(\rho^s \left(\ln \frac{1}{\rho} \right) \sigma^{1-s} \right) \right| \\ &\leq r_m(\delta^{\bar{s}/2} + \varepsilon')^2 + 4r_m\varepsilon' + r_m \cdot (\delta^{\bar{s}/2} + \varepsilon')^2 \\ &\quad + \frac{1}{4\tilde{K}^2} \sum_{p_i < \delta} \left(p_i^{\bar{s}/2} \left(\tilde{K} + \sqrt{\ln \frac{1}{p_i}} \right) \right)^2 \\ &\quad + \frac{\varepsilon'}{4\tilde{K}} \sum_{p_i \geq \delta} p_i^{\bar{s}} \left(\tilde{K} + \sqrt{\ln \frac{1}{p_i}} \right) \\ &\leq 2r_m \cdot \left(\left(\frac{\varepsilon}{12r_m \ln \frac{r_m}{\varepsilon}} \right)^{1/2} + \varepsilon' \right)^2 + 4r_m\varepsilon' \\ &\quad + \frac{1}{4\tilde{K}^2} \sum_{p_i < \delta} \left(p_i^{\bar{s}/2} \left(\tilde{K} + \sqrt{\ln \frac{1}{p_i}} \right) \right)^2 \\ &\quad + \frac{\varepsilon'}{4\tilde{K}} \sum_{p_i \geq \delta} p_i^{\bar{s}} \left(\tilde{K} + \sqrt{\ln \frac{1}{p_i}} \right) \\ &\leq \frac{3\varepsilon}{20} + \frac{4\varepsilon}{85} + \frac{1}{4\tilde{K}^2} \sum_{p_i < \delta} \left(p_i^{\bar{s}/2} \left(\tilde{K} + \sqrt{\ln \frac{1}{p_i}} \right) \right)^2 \\ &\quad + \frac{\varepsilon'}{4\tilde{K}} \sum_{p_i \geq \delta} p_i^{\bar{s}} \left(\tilde{K} + \sqrt{\ln \frac{1}{p_i}} \right).\end{aligned}$$

The first term in the third inequality comes from expanding the first term in the second inequality:

$$\begin{aligned}2r_m \cdot \left(\left(\frac{\varepsilon}{12r_m \ln \frac{r_m}{\varepsilon}} \right)^{1/2} + \varepsilon' \right)^2 \\ &= \frac{\varepsilon}{6 \ln \frac{r_m}{\varepsilon}} + 4r_m\varepsilon' \left(\frac{\varepsilon}{12r_m \ln \frac{r_m}{\varepsilon}} \right)^{1/2} + 2r_m(\varepsilon')^2 \\ &\leq \frac{\varepsilon}{20} + \frac{4\varepsilon}{85} + \frac{2\varepsilon}{85} \\ &\leq \frac{3\varepsilon}{20}.\end{aligned}$$

The inequality comes from the fact that $r_m/\varepsilon \geq 24$ and $(\frac{\varepsilon}{12r_m \ln \frac{r_m}{\varepsilon}})^{1/2} \leq 1$, and $\varepsilon' < (1/2) < 1$.

Let $\tilde{f}(x) := x^{\bar{s}/2}(\tilde{K} + \sqrt{\ln(1/x)})$, $\tilde{g}(x) := x^{\bar{s}}(\tilde{K} + \sqrt{\ln(1/x)})$. Since $\tilde{f}(x) = \frac{1}{2}x^{\bar{s}/2-1}(\bar{s}\tilde{K} + \bar{s}\sqrt{\ln \frac{1}{x}} - \frac{1}{\sqrt{\ln \frac{1}{x}}})$, we know that $\tilde{f}(x)$ attains its maximum value at $x_0 = e^{-(\tilde{K}/2 + \sqrt{\tilde{K}^2/4 + 1/\bar{s}})^2}$. We show that $x_0 > \delta$ which implies that $\tilde{f}(x)$ attains its maximum value at $x = \delta$ with $x \in (0, 1]$.

Let $y = -\frac{\tilde{K}}{2} + \sqrt{\frac{\tilde{K}^2}{4} + \frac{1}{\bar{s}}}$, we have

$$\begin{aligned}y &= -\sqrt{\ln \frac{2}{\delta}} + \sqrt{\ln \frac{2}{\delta} + \frac{1}{\bar{s}}} \\ &= -\sqrt{\ln \frac{2}{\delta}} + \sqrt{\ln \frac{2}{\delta}} \sqrt{1 + \frac{1}{\ln \frac{2}{\delta} \bar{s}}} \\ &\leq -\sqrt{\ln \frac{2}{\delta}} + \sqrt{\ln \frac{2}{\delta}} \left(1 + \frac{3}{2 \ln \frac{2}{\delta} \bar{s}} \right) \\ &\leq \frac{3}{2\bar{s}} \cdot \frac{1}{\sqrt{\ln \frac{2}{\delta}}}.\end{aligned}$$

The first inequality comes from the Taylor expansions of $\sqrt{1+x} = \sum_{k=0}^{\infty} \binom{1/2}{k} x^k = 1 + \frac{x}{2} - \frac{x^2}{8} + \dots$, where we add $\frac{1}{\ln \frac{2}{\delta} \bar{s}}$ to make the inequality solid. Thus, $y^2 \leq \frac{9}{4\bar{s}^2} \frac{1}{\ln \frac{2}{\delta}}$. Next, we show $\frac{9}{4\bar{s}^2} < (\ln \frac{1}{\delta})^2$, which implies $\frac{9}{4\bar{s}^2 \ln \frac{2}{\delta}} < \ln \frac{1}{\delta}$. Then, we have $y^2 < \ln \frac{1}{\delta}$, which means $e^{-y^2} > e^{-\ln \frac{1}{\delta}}$. And we can safely claim that $\tilde{f}(x)$ attains its maximum value at $x = \delta$ with $x \in (0, 1]$.

Given that $r_m/\varepsilon > 24$, we have

$$\begin{aligned}\left(\ln \frac{1}{\delta} \right)^2 &= \left(\frac{1}{\bar{s}} \ln \left(12 \frac{r_m}{\varepsilon} \ln \frac{r_m}{\varepsilon} \right) \right)^2 \\ &> \frac{1}{\bar{s}^2} \ln^2(12 \cdot 24 \ln 24) \\ &> \frac{1}{\bar{s}^2} 6^2 \\ &> \frac{9}{4\bar{s}^2}.\end{aligned}$$

Through a similar derivation as above, we can also conclude that $\tilde{g}(x)$ only attains its maximum value at $x = \delta$ or $x = 1$.

So we know

$$\begin{aligned}
\delta^{\frac{1}{2}} \left(\tilde{K} + \sqrt{\ln \frac{1}{\delta}} \right) &= \left(\frac{\varepsilon}{12r_m} \right)^{1/2} \left(\frac{1}{\ln \frac{r_m}{\varepsilon}} \right)^{1/2} \left(\tilde{K} + \sqrt{\frac{1}{\delta} \ln \left(\frac{12r_m}{\varepsilon} \ln \frac{r_m}{\varepsilon} \right)} \right) \\
&= \left(\frac{\varepsilon}{12r_m} \right)^{1/2} \left(\frac{1}{\ln \frac{r_m}{\varepsilon}} \right)^{1/2} \left(2\sqrt{\ln 2 + \frac{1}{\delta} \ln \left(\frac{12r_m}{\varepsilon} \ln \frac{r_m}{\varepsilon} \right)} + \sqrt{\frac{1}{\delta} \ln \left(\frac{12r_m}{\varepsilon} \ln \frac{r_m}{\varepsilon} \right)} \right) \\
&\leq \left(\frac{\varepsilon}{12r_m} \right)^{1/2} \left(\frac{1}{\ln \frac{r_m}{\varepsilon}} \right)^{1/2} \left(2\sqrt{\frac{1}{\delta} \ln \left(\frac{24r_m}{\varepsilon} \ln \frac{r_m}{\varepsilon} \right)} + \sqrt{\frac{1}{\delta} \ln \left(\frac{12r_m}{\varepsilon} \ln \frac{r_m}{\varepsilon} \right)} \right) \\
&\leq \left(\frac{\varepsilon}{12r_m} \right)^{1/2} \left(\frac{1}{\ln \frac{r_m}{\varepsilon}} \right)^{1/2} \left(3\sqrt{\frac{1}{\delta} \ln \left(\frac{r_m}{\varepsilon} \right)^3} \right) \\
&\leq \sqrt{\frac{27}{\delta}} \left(\frac{\varepsilon}{12r_m} \right)^{1/2}.
\end{aligned}$$

The second inequality comes from that $(r_m/\varepsilon) > 24$ and $(r_m/\varepsilon) > \ln(r_m/\varepsilon)$. Note that $\tilde{g}(\delta) = \delta^{\frac{1}{2}}(\tilde{K} + \sqrt{\ln \frac{1}{\delta}}) \leq \sqrt{\frac{27}{\delta}} \frac{\varepsilon}{12r_m} \left(\frac{1}{\ln \frac{r_m}{\varepsilon}} \right)^{\frac{1}{2}}$, $\tilde{g}(1) = \tilde{K}$. Through simple calculations, it can be found that $\tilde{g}(\delta) < \tilde{g}(1)$.

Now we have

$$\begin{aligned}
\left| p(0) - \frac{1}{16\tilde{K}^2} \text{Tr} \left(\rho^s \left(\ln \frac{1}{\rho} \right) \sigma^{1-s} \right) \right| &\leq \frac{3\varepsilon}{20} + \frac{4\varepsilon}{85} + r_m \cdot \frac{\bar{s}}{8(\bar{s}+1)} \frac{27}{\bar{s}} \frac{\varepsilon}{12r_m} + \frac{r_m \varepsilon'}{4} \\
&\leq \frac{3}{4} \varepsilon.
\end{aligned}$$

By Theorem 1, we can estimate $p(0)$ with additive error $\varepsilon/(108 \ln(2/\delta))$ using $O(\frac{\ln \frac{1}{\delta}}{\varepsilon})$ applications of the matrices and their conjugate transposes. In total, we can estimate $\text{Tr}(\rho^s (\ln \rho) \sigma^{1-s})$ within additive error

$$\frac{\varepsilon}{108 \ln(2/\delta)} \cdot 16\tilde{K}^2 + \frac{3}{4} \varepsilon = \varepsilon,$$

with a success probability of at least $8/\pi^2$. The total complexity of the algorithm is

$$\begin{aligned}
&O\left(\frac{\ln \frac{1}{\delta}}{\varepsilon}\right) O\left(\left(\frac{r_m}{\varepsilon}\right)^{1/\bar{s}} \left(\ln \frac{r_m}{\varepsilon}\right)^{1/\bar{s}+1}\right) \\
&= O\left(\frac{r_m^{1/\bar{s}}}{\varepsilon^{1/\bar{s}+1}} \left(\ln \frac{1}{\varepsilon}\right)^{1/\bar{s}+2}\right).
\end{aligned}$$

APPENDIX D: THE PROOF OF LEMMA 9

Proof. We only give the proof for η_k^{QCB} , the conclusion for η_k^{QHB} can be obtained through a similar process.

$$\begin{aligned}
|r_k^{\text{QCB}}| &= \left| \frac{-\dot{h}_{\text{QSVT}}(x_k) \ddot{h}(x_k) + \ddot{h}_{\text{QSVT}}(x_k) \dot{h}(x_k)}{\ddot{h}_{\text{QSVT}}(x_k)} \right| \\
&\leq \left| \frac{-\ddot{h}(x_k)(\dot{h}(x_k) - \epsilon_{\tilde{h}}) + (\ddot{h}(x_k) + \epsilon_{\tilde{h}})\dot{h}(x_k)}{\ddot{h}_{\text{QSVT}}(x_k)} \right| \\
&\leq \frac{|\dot{h}(x_k)\epsilon_{\tilde{h}} + \ddot{h}(x_k)\epsilon_{\tilde{h}}|}{|\ddot{h}_{\text{QSVT}}(x_k)|}
\end{aligned}$$

$$\begin{aligned}
&\leq \frac{|\dot{h}_{\text{QSVT}}(x_k) + \epsilon_{\tilde{h}}\epsilon_{\tilde{h}} + (\ddot{h}_{\text{QSVT}}(x_k) + \epsilon_{\tilde{h}})\epsilon_{\tilde{h}}|}{|\ddot{h}_{\text{QSVT}}(x_k)|} \\
&\leq \frac{\epsilon_{\tilde{h}}|\ddot{h}_{\text{QSVT}}(x_k)| + \epsilon_{\tilde{h}}|\dot{h}_{\text{QSVT}}(x_k)| + 2\epsilon_{\tilde{h}}\epsilon_{\tilde{h}}}{|\ddot{h}_{\text{QSVT}}(x_k)|} \\
&\leq \eta_k^{\text{QCB}} |\dot{h}(x_k)|.
\end{aligned}$$

The first inequality follows from the convexity of $h(s)$ and the conclusion of Corollary 1. The third inequality also follows from the result of Corollary 1. ■

APPENDIX E: THE PROOF OF THEOREM 6

Proof. Each term in $h(s)$, $\dot{h}(s)$, $\ddot{h}(s)$, $g(s)$, $\dot{g}(s)$, $\ddot{g}(s)$ can be derived through a proof similar to that in Theorem 5, and the query complexity for $U_\rho, U_\rho^\dagger, U_\sigma, U_\sigma^\dagger, \text{SWAP}, U_{\text{copy}}$ will not exceed that of $O(\frac{r_m^{1/\bar{s}}}{\varepsilon^{1/\bar{s}+1}} (\ln \frac{r_m}{\varepsilon})^{1/\bar{s}+1})$. Since $h(s)$, $\dot{h}(s)$, $\ddot{h}(s)$, $g(s)$, $\dot{g}(s)$, and $\ddot{g}(s)$ are obtained through a linear combination of estimated terms, the query complexity to compute $h(s)$ and $g(s)$ is $O(\frac{r_m^{1/\bar{s}}}{\varepsilon^{1/\bar{s}+1}} (\ln \frac{r_m}{\varepsilon})^{1/\bar{s}+1})$. Using the forcing terms from Lemma 9, the inexact Newton method converges in $O(\ln(1/\varepsilon))$ iterations to an estimate $\hat{x}$ satisfying $|x^* - \hat{x}| \leq (\varepsilon/4) |\ln \delta|$, as shown in Lemma 8. So the total query complexity of QCB and QHB estimation is $O(\frac{r_m^{1/\bar{s}}}{\varepsilon^{1/\bar{s}+1}} (\ln \frac{r_m}{\varepsilon})^{1/\bar{s}+1} \ln \frac{1}{\varepsilon})$.

In the following discussion, we show that $h(s)$ is $2|\ln \delta|$ -Lipschitz continuous, an analogous result holds for $g(s)$. For all $x, y \in [0, 1]$, we have $|h(x) - h(y)| \leq \sup_{u \in [0, 1]} |h(u)| |x - y|$.

$y]$. Since $\dot{h}(s) = \text{Tr}(\rho^s \sigma^{1-s} (\ln \rho - \ln \sigma))$, we have

$$\begin{aligned} |\dot{h}(s)| &\leq \|\rho^s \sigma^{1-s}\|_1 \|\ln \rho - \ln \sigma\| \\ &\leq \|\rho^s\|_1 \|\sigma^{1-s}\| \|\ln \rho - \ln \sigma\| \\ &\leq 1 \cdot \|\ln \rho - \ln \sigma\| \\ &\leq \|\ln \rho\| + \|\ln \sigma\| \\ &\leq 2|\ln \delta|. \end{aligned}$$

The first and the second inequality comes from the fact that $|\text{Tr}(\rho \sigma)| \leq \|\rho\|_1 \|\sigma\|$ [26]. Thus, we have $|h(x) - h(y)| \leq 2|\ln \delta||x - y|$, for all $x, y \in [0, 1]$. Finally, given the estimation $h_{\text{QSVT}}(\hat{x})$ from Corollary 1 such that $|h_{\text{QSVT}}(\hat{x}) - h(\hat{x})| \leq$

$(\varepsilon/2)$, we can safely claim that we have a good estimation of $h(x^*)$:

$$\begin{aligned} |h_{\text{QSVT}}(\hat{x}) - h(x^*)| &\leq |h_{\text{QSVT}}(\hat{x}) - h(\hat{x})| + |h(\hat{x}) - h(x^*)| \\ &\leq \frac{\varepsilon}{2} + 2|\ln \delta| \cdot \frac{\varepsilon}{4|\ln \delta|} \\ &\leq \varepsilon. \end{aligned}$$

■

APPENDIX F: THE PROOF OF THEOREM 7

Proof. Use Hadamard test with initial state set to be $|0\rangle \otimes |0\rangle_{a+t} \otimes |\Phi\rangle$, where $|\Phi\rangle := \frac{1}{\sqrt{d}} \sum_{i=0}^{d-1} |i\rangle_t |i\rangle_b$. The final state is

$$\begin{aligned} &[(H \otimes I_{a+2t+b})(|0\rangle\langle 0| \otimes I_{a+2t+b} + |1\rangle\langle 1| \otimes \text{BE}_{P(\rho)} \otimes I_b)(H \otimes I_{a+2t+b})]|0\rangle \otimes |0\rangle_{a+t} \otimes |\Phi\rangle \\ &= |0\rangle \frac{I + \text{BE}_{P(\rho)} \otimes I_b}{2} |0\rangle_{a+t} |\Phi\rangle + |1\rangle \frac{I - \text{BE}_{P(\rho)} \otimes I_b}{2} |0\rangle_{a+t} |\Phi\rangle. \end{aligned} \quad (\text{F1})$$

With Z measurement on the ancilla qubit, the probability of getting result 0 is

$$\begin{aligned} \text{Pr}(0) &= \frac{1}{4} \langle 0|_{a+t} \langle \Phi| (I + \text{BE}_{P(\rho)}^\dagger \otimes I_b) (I + \text{BE}_{P(\rho)} \otimes I_b) |0\rangle_{a+t} |\Phi\rangle \\ &= \frac{1}{2} + \frac{1}{4} \langle 0|_{a+t} \langle \Phi| \left[\left(\sum_{i=0}^{d-1} P(p_i) |0\rangle\langle 0|_{a+t} \otimes |\psi_i\rangle\langle \psi_i|_t \right) \otimes I_b \right] |0\rangle_{a+t} |\Phi\rangle \\ &\quad + \frac{1}{4} \langle 0|_{a+t} \langle \Phi| \left[\left(\sum_{i=0}^{d-1} P(p_i) |0\rangle\langle 0|_{a+t} \otimes |\psi_i\rangle\langle \psi_i|_t \right) \otimes I_b \right] |0\rangle_{a+t} |\Phi\rangle \\ &= \frac{1}{2} + \frac{1}{2d} \left(\sum_{k=0}^{d-1} \langle k|_t \langle k|_b \right) \left[\left(\sum_{i=0}^{d-1} P(p_i) |\psi_i\rangle\langle \psi_i|_t \right) \otimes I_b \right] \left(\sum_{j=0}^{d-1} |j\rangle_t |j\rangle_b \right) \\ &= \frac{1}{2} + \frac{1}{2d} \sum_{i,k=0}^{d-1} P(p_i) \langle k|\psi_i\rangle \langle \psi_i|k\rangle. \end{aligned} \quad (\text{F2})$$

Similarly, we know

$$\text{Pr}(1) = \frac{1}{2} - \frac{1}{2d} \sum_{i,k=0}^{d-1} P(p_i) \langle k|\psi_i\rangle \langle \psi_i|k\rangle.$$

So, the expectation of the measurement result $Y \in \{+1, -1\}$ is

$$\begin{aligned} \mathbb{E}(Y) &= (+1) \cdot \left(\frac{1}{2} + \frac{1}{2d} \sum_{i,k=0}^{d-1} P(p_i) \langle k|\psi_i\rangle \langle \psi_i|k\rangle \right) + (-1) \cdot \left(\frac{1}{2} - \frac{1}{2d} \sum_{i,k=0}^{d-1} P(p_i) \langle k|\psi_i\rangle \langle \psi_i|k\rangle \right) \\ &= \frac{1}{d} \text{Tr} \left(P(\rho) \sum_{k=0}^{d-1} |k\rangle\langle k| \right) \\ &= \frac{1}{d} \text{Tr}(P(\rho)). \end{aligned} \quad (\text{F3})$$

By running the Hadamard test and applying amplitude estimation with additive error ε/d , we obtain an estimation of $\text{Tr}(P(\rho))$ within ε , with success probability at least $2/3$, using $O(d/\varepsilon)$ calls to the controlled block encoding unitaries. ■

[1] H. Chernoff, A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations, *Ann. Math. Stat.* **23**, 493 (1952).

[2] K. M. Audenaert, J. Calsamiglia, R. Muñoz-Tapia, E. Bagan, L. Masanes, A. Acín, and F. Verstraete, Discriminating states: The quantum Chernoff bound, *Phys. Rev. Lett.* **98**, 160501 (2007).

- [3] M. Nussbaum and A. Szkoła, The Chernoff lower bound for symmetric quantum hypothesis testing, *Ann. Stat.* **37**, 1040 (2009).
- [4] J. J. Meyer, S. Khatri, D. S. França, J. Eisert, and P. Faist, Quantum metrology in the finite-sample regime, *PRX Quantum* **6**, 030336 (2025).
- [5] A. Gilyén and A. Poremba, Improved quantum algorithms for fidelity estimation, [arXiv:2203.15993](https://arxiv.org/abs/2203.15993).
- [6] M. Cerezo, A. Poremba, L. Cincio, and P. J. Coles, Variational quantum fidelity estimation, *Quantum* **4**, 248 (2020).
- [7] Q. Wang, Z. Zhang, K. Chen, J. Guan, W. Fang, J. Liu, and M. Ying, Quantum algorithm for fidelity estimation, *IEEE Trans. Inf. Theory* **69**, 273 (2022).
- [8] X. Zhang, M. Luo, Z. Wen, Q. Feng, S. Pang, W. Luo, and X. Zhou, Direct fidelity estimation of quantum states using machine learning, *Phys. Rev. Lett.* **127**, 130503 (2021).
- [9] Q. Wang, J. Guan, J. Liu, Z. Zhang, and M. Ying, New quantum algorithms for computing quantum entropies and distances, *IEEE Trans. Inf. Theory* **70**, 5653 (2024).
- [10] Y. Wang, L. Zhang, Z. Yu, and X. Wang, Quantum phase processing and its applications in estimating phase and entropies, *Phys. Rev. A* **108**, 062413 (2023).
- [11] S. Subramanian and M.-H. Hsieh, Quantum algorithm for estimating α -Renyi entropies of quantum states, *Phys. Rev. A* **104**, 022428 (2021).
- [12] T. Gur, M.-H. Hsieh, and S. Subramanian, Sublinear quantum algorithms for estimating von Neumann entropy, [arXiv:2111.11139](https://arxiv.org/abs/2111.11139).
- [13] Y. Lu and K. Fang, Estimating quantum relative entropies on quantum computers, [arXiv:2501.07292](https://arxiv.org/abs/2501.07292).
- [14] R. Chen, Z. Song, X. Zhao, and X. Wang, Variational quantum algorithms for trace distance and fidelity estimation, *Quantum Sci. Technol.* **7**, 015019 (2021).
- [15] Q. Wang and Z. Zhang, Fast quantum algorithms for trace distance estimation, *IEEE Trans. Inf. Theory* **70**, 2720 (2023).
- [16] A. Gilyén, Quantum singular value transformation & its algorithmic applications, Ph.D. thesis, University of Amsterdam, 2019.
- [17] R. S. Dembo, S. C. Eisenstat, and T. Steihaug, Inexact Newton methods, *SIAM J. Numer. Anal.* **19**, 400 (1982).
- [18] V. Vedral, The role of relative entropy in quantum information theory, *Rev. Mod. Phys.* **74**, 197 (2002).
- [19] T. M. Cover, *Elements of Information Theory* (John Wiley & Sons, Hoboken, NJ, USA, 1999).
- [20] H. Nagaoka, The converse part of the theorem for quantum Hoeffding bound, [arXiv:quant-ph/0611289](https://arxiv.org/abs/quant-ph/0611289).
- [21] J. van Apeldoorn, A. Cornelissen, A. Gilyén, and G. Nannicini, Quantum tomography using state-preparation unitaries, in *Proceedings of the 2023 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)* (SIAM, Philadelphia, PA, USA, 2023), pp. 1265–1318.
- [22] G. Brassard, P. Hoyer, M. Mosca, and A. Tapp, Quantum amplitude amplification and estimation, *Contemp. Math.* **305**, 53 (2002).
- [23] A. Gilyén, Y. Su, G. H. Low, and N. Wiebe, Quantum singular value transformation and beyond: Exponential improvements for quantum matrix arithmetics, in *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing (STOC)* (Association for Computing Machinery, New York, NY, USA, 2019), pp. 193–204.
- [24] J. Luo, Q. Wang, and L. Li, Succinct quantum testers for closeness and k -wise uniformity of probability distributions, *IEEE Trans. Inf. Theory* **70**, 5092 (2024).
- [25] A. Gilyén and T. Li, Distributional property testing in a quantum world, in *11th Innovations in Theoretical Computer Science Conference (ITCS 2020)* (Schloss Dagstuhl–Leibniz-Zentrum für Informatik, Wadern, Germany, 2020), pp. 25:1–25:20.
- [26] R. Bhatia, Matrix Analysis in *Graduate Texts in Mathematics* (Springer Science & Business Media, New York, NY, USA, 2013), Vol. 169.