Quantum Gram-Schmidt processes and their application to efficient state readout for quantum algorithms

Kaining Zhang,[1] Min-Hsiu Hsieh,[2] Liu Liu,[1] and Dacheng Tao[1]
[1]*School of Computer Science, Faculty of Engineering, University of Sydney, Australia*
[2]*Hon Hai Quantum Computing Research Center, Taipei, Taiwan*

(Received 15 April 2020; revised 17 August 2021; accepted 29 September 2021; published 1 November 2021)

Many quantum algorithms that claim speedup over their classical counterparts only generates quantum states as solutions instead of their final classical description. The additional step to decode quantum states into classical vectors normally will destroy the quantum advantage in most scenarios because all existing tomographic methods require runtime that is polynomial with respect to the state dimension. In this work, we present an efficient readout protocol that yields the classical vector form of the generated state, so it will achieve the end-to-end advantage for those quantum algorithms. Our protocol suits the case in which the output state lies in the row space of the input matrix, of rank r, that is stored in the quantum random access memory. The quantum resources for decoding the state in ℓ^2 norm with ϵ error require poly$(r, 1/\epsilon)$ copies of the output state and poly$(r, \kappa^r, 1/\epsilon)$ queries to the input oracles, where κ is the condition number of the input matrix. With our readout protocol, we completely characterize the end-to-end resources for quantum linear equation solvers and quantum singular value decomposition. One of our technical tools is an efficient quantum algorithm for performing the Gram-Schmidt orthonormal procedure, which we believe will be of independent interest.

DOI: 10.1103/PhysRevResearch.3.043095

I. INTRODUCTION

Quantum algorithms have been popular for decades because of their potential advantage in varying fields including physical simulations [1–3], combinatorial optimization [4,5], and linear algebra [6]. Notably, the latter has induced an independent subfield known as quantum machine learning (QML) [7,8], which involves quantum linear algebra [9–11], quantum learning protocols [12–15], and quantum neural networks [16,17]. These quantum algorithms have been shown to achieve speedups over their classical counterparts.

Despite the claimed quantum speedup, most QML algorithms suffer from both the input and the readout problems. Specifically, the input problem tackles the issue of efficient state preparation, namely, encoding the classical data, potentially of tantamount size, into quantum states. A few techniques [9,16,18,19] have been proposed to address this problem, and among them, the quantum random access memory (QRAM) oracle model [18] has become arguably the most popular method in the domain of machine learning applications. It has induced interesting outcomes in quantum algorithms for tasks such as the linear system solver [9,20,21], the singular value decomposition (SVD) [10], support-vector machines [12,22,23], supervised and unsupervised learning [13,15], neural networks [24,25], and other machine learning tasks [26–28]. Generally, for a data matrix $A \in \mathbb{R}^{m\times d}$, the corresponding QRAM oracle could be prepared by using $O(\text{polylog}(md))$ quantum operations with $O(md)$ physical resources [18] stored in a binary tree data structure [29]. Although the QRAM oracle is criticized for its requirement of large physical resources, recent works [30,31] have proven possible the practical implementation of the QRAM oracle.

On the other hand, the readout problem addresses recovery of a classical description from the output quantum state that contains the classical solutions. In order to preserve the quantum advantage of the underlining quantum algorithm, the output state needs to be decoded efficiently. For some quantum algorithms, such as the quantum recommendation system [27], the readout issue is relatively mild because the classical solution can be obtained by only a few measurements on the output state. In general, most machine learning problems demand classical solutions in vector form, for example, finding solutions to linear systems. Hence, the readout problem of these quantum algorithms could be critical. However, protocols for efficiently decoding the output quantum states into classical vectors remain largely unexplored [32].

The task of recovering the unknown quantum state from measurements, which is also known as quantum state tomography (QST), is one of the fundamental problems in quantum information science. QST has attracted significant interest from both theoretical [33–38] and experimental [39–45] perspectives in recent years. The best general tomography method [36] could reconstruct a $d \times d$ density matrix ρ for the unknown state with rank r by using $n = O(rd\epsilon^{-2})$ copies to the state, which implies $O(d\epsilon^{-2})$ copy complexity for the pure state case $\rho = |\boldsymbol{v}\rangle\langle\boldsymbol{v}|$. We remark that most of QML algorithms that output a d-dimensional state as the solution claim the time complexity polylogarithmical to d. Thus, directly using state tomography methods for state readout in QML is computationally expensive and would offset the gained quantum speedup. Since the required number n is proven optimal

for both cases [36], any further improvement on n could be achieved only by assuming special prior knowledge on state ρ. For example, QST via local measurements provides efficient estimation for states which can be determined by local reduced density matrices [38] or states with a low-rank tensor decomposition [37]. However, the output states generated by QML algorithms normally do not have these structures.

In contrast with the assumptions in the QST scenarios, the output states generated by most QML algorithms *do* have inherent relationship between the solution vector and the input data, commonly represented as a matrix. Specifically, the solution vector normally lies in the row space of the input data matrix. Notable examples that satisfy the aforementioned condition include (1) the quantum SVD algorithm where the singular value σ_i and corresponding singular vectors $|\boldsymbol{u}_i\rangle$ and $|\boldsymbol{v}_i\rangle$ for matrix $A = \sum_i \sigma_i \boldsymbol{u}_i \boldsymbol{v}_i^T$ and (2) the quantum linear system solver for linear system $A\boldsymbol{x} = \boldsymbol{b}$ whose solution state $|\boldsymbol{x}\rangle \propto A^{-1}\boldsymbol{b}$ lies in the row space of A. Most machine learning problems can be reduced to these two categories [32]. Hence, finding efficient readout protocols for them that go beyond the standard QST limit will be extremely desirable in the field of QML.

In this work, we design an efficient state readout protocol that works for QML algorithms which involve a r-rank input matrix $A \in \mathbb{R}^{m\times d}$ stored in the quantum random access memory (QRAM), and the output state $|\boldsymbol{v}\rangle$ lies in the row space of A. Instead of obtaining coefficients $\{v_i\}$ by measuring the state $|\boldsymbol{v}\rangle = \sum_{i=1}^{n} v_i|i\rangle$ in the standard orthonormal basis $\{|i\rangle\}$, our key technical contribution is an efficient method to obtain the classical description x_i in the complete basis spanned by the rows $\{A_{g(i)}\}_{i=1}^{r}$ of A, so that $|\boldsymbol{v}\rangle = \sum_{i=1}^{r} x_i|A_{g(i)}\rangle$, where the mapping $g(i) : [r] \to [m]$ denotes the indices of rows selected as the basis. Our state readout protocol requires $\tilde{O}(\mathrm{poly}(r))$ copies of the output states and $\tilde{O}(\mathrm{poly}(r, \kappa^r))$ queries to input oracles, where r is the rank of the input matrix and $\kappa = \sigma_{\max}(A)/\sigma_{\min}(A)$ is the condition number of the input matrix. We remark that the low-rank matrix assumption is common in machine learning models [46–48]. Compared to previous QST methods which require at least $O(d\epsilon^{-2})$ copies of pure states, our protocol is much more efficient given $r \ll n$ with small condition numbers, and more importantly, the complexity does not depend on the system dimension. Finally, combining our readout protocol with quantum SVD or quantum linear system solver yields an end-to-end complexity that takes $\tilde{O}(\mathrm{poly}(r, \kappa^r, \log(md)))$ queries to input oracles.

During the whole readout protocol, we develop a quantum generalization of the Gram-Schmidt orthonormalization process. Our quantum Gram-Schmidt process (QGSP) algorithm can construct a complete basis, by sampling a set of rows $\{A_{g(i)}\}_{i=1}^{r}$ of the input A, with $\tilde{O}(\mathrm{poly}(r, \kappa^r))$ queries to QRAM oracles. Since the vector orthonormalization is a crucial procedure in linear algebra as well as machine learning [49–51], an efficient quantum algorithm will be of independent interest. Notice that there are some related works for the construction of orthogonal states [52–55]. However, these results deviate from standard Gram-Schmidt process and their applications are also limited. Reference [52] is only applicable to the single-qubit system, while Refs. [53,54] only generate a state that is orthogonal to the input state and their complexity depends on the system dimension. Reference [55] constructs orthogonal states from original states by lifting the dimension of the original Hilbert space and cannot select a complete basis as standard Gram-Schmidt process does. Consequently, our proposed QGSP algorithm avoids all these restrictions and can be proven to be efficient.

Specifically, we have the following result for QGSP.

Theorem 1 (informal). By using $O(r^{27}\kappa^{14r})$ queries to QRAM oracles of the matrix A, we could find a group of linearly independent rows $\{A_{g(i)}\}_{i=1}^{r}$, where r and κ is the rank and the condition number of A, respectively.

Main Result. The main result for our state readout protocol is as follows.

Theorem 2. For the d-dimensional state $|\boldsymbol{v}\rangle$ lies in the row space of a matrix $A \in \mathbb{R}^{m\times d}$ with rank r and the condition number κ, the classical form of $|\boldsymbol{v}\rangle$ could be obtained by using $O(r^4\epsilon^{-2})$ queries to the state $|\boldsymbol{v}\rangle$ and $O(r^{27}\kappa^{14r} + r^{18}\kappa^{8r}\epsilon^{-2})$ queries to QRAM oracles of A, such that the ℓ^2 norm error is bounded in ϵ.

Further discussion about the applications of our main result will be given in Sec. III. Here, we will move on to formally define the framework of the state readout protocol.

II. STATE READOUT FRAMEWORK

In this section, we explain our protocol in detail. Since $A \in \mathbb{R}^{m\times d}$ is of rank r, we can identify a set of r linearly independent vectors $\{|A_{g(i)}\rangle\}_{i=1}^{r}$ selected from all rows of A so that the output state can be rewritten as $|\boldsymbol{v}\rangle = \sum_{i=1}^{r} x_i|A_{g(i)}\rangle$. Our goal is accomplished if we can determine $\{x_i\}_{i=1}^{r}$ efficiently. Following this, our algorithm consists of two major parts, a subroutine to sample a set of r linearly independent rows $\{|A_{g(i)}\rangle\}_{i=1}^{r}$ from all rows of A and a subroutine to calculate $\{x_i\}$, which will be introduced in following subsections, respectively.

A. Complete basis sampling

We begin with the first subroutine. The quantum Gram-Schmidt process (QGSP) in Algorithm I is developed to generate a complete row basis, by performing a quantum version of the adaptive sampling. The advantage of our adaptive sampling is that those rows, which have larger orthogonal part to the row space of previous sampled row submatrix, will be sampled with a larger probability. This ensures that the complete basis is nonsingular, and will improve the accuracy of the estimation of the coefficients in the second subroutine.

Algorithm 1. Quantum Gram-Schmidt process (QGSP).

Input: QRAM oracles V_A and U_A in Eqs. (1) and (2).
Output: A group of orthonormal states $\{|\boldsymbol{t}_i\rangle\}_{i=1}^{r}$. An index set of the complete basis: $S_I = \{g(i)\}_{i=1}^{r}$.
1: Initialize the index set $S_I = \emptyset$.
2: **for** $\ell = 1$ to r **do**
3: Run the quantum circuit in Fig 1. Measure the third register and postselect on result 0. Measure the first register to obtain an index $g(\ell)$. Update the index set $S_I = S_I \cup \{g(\ell)\}$.
4: **end for**

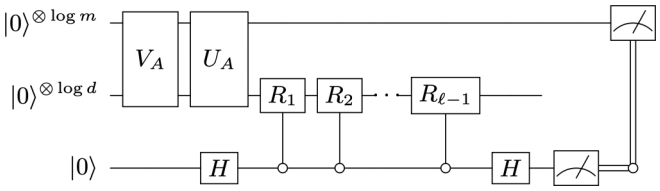

FIG. 1. Quantum circuit for the ℓth iteration in the QGSP. Oracles V_A (2) and U_A (1) are employed for encoding rows of the matrix A. Gates $C(R_i)$ (4) are used for extracting the orthogonal part of rows from existing basis rows. After unitary operations, we measure the third register and postselect on the result 0, to generate the state $|\phi_2^{(\ell)}\rangle$ (7) from the state $|\phi_1^{(\ell)}\rangle$ (5). Finally, the new index is sampled by measuring the first register of the state $|\phi_2^{(\ell)}\rangle$.

Now we analyze the QGSP in detail. We utilize QRAM oracles V_A and U_A to encode the matrix A in the amplitude of quantum states:

$$|0\rangle \xrightarrow{V_A} \sum_{i=1}^{m} \|A_i\|/\|A\|_F |i\rangle, \tag{1}$$

$$|i\rangle|0\rangle \xrightarrow{U_A} |i\rangle|A_i\rangle \equiv \sum_{j=1}^{d} A_{ij}/\|A_i\||i\rangle|j\rangle, \quad \forall i \in [m], \tag{2}$$

where A_{ij}, A_i, and $\|A\|_F$ denote the (i, j)-th element, the ith row, and the Frobenius norm of A, respectively. In the first iteration of the QGSP, an index $g(1)$ is sampled from the set $[m] := \{1, 2, \ldots, m\}$ with the probability $\Pr^{(1)}(i) = \|A_i\|^2/\|A\|_F^2$, where $i \in [m]$. Let $|\boldsymbol{t}_1\rangle := |A_{g(1)}\rangle$ be the first basis vector. The remaining basis vectors are generated inductively. Assume a set of orthogonal states $\{|\boldsymbol{t}_i\rangle\}_{i=1}^{\ell-1}$ has been generated in the previous $\ell - 1$ iterations. To proceed to the ℓth iteration, we perform the quantum circuit illustrated in Fig. 1, which first creates the state

$$\sum_{j=1}^{m} \frac{\|A_j\|}{\|A\|_F} |j\rangle|A_j\rangle|0\rangle, \tag{3}$$

with the help of input oracles U_A and V_A. Then a Hadamard gate is applied to the third register, followed by a sequence of controlled R_i gates

$$C(R_i) = R_i \otimes |0\rangle\langle 0| + I \otimes |1\rangle\langle 1|, \tag{4}$$

where the unitary $R_i = I - 2|\boldsymbol{t}_i\rangle\langle \boldsymbol{t}_i|$. Next, another Hadamard gate is applied to the third register, and the quantum state evolves into

$$|\phi_1^{(\ell)}\rangle = \frac{1}{\|A\|_F} \sum_{j=1}^{m} \|A_j\||j\rangle \otimes \left[\left(|A_j\rangle - \sum_{i=1}^{\ell-1} |\boldsymbol{t}_i\rangle\langle \boldsymbol{t}_i|A_j\rangle \right)|0\rangle - \sum_{i=1}^{\ell-1} |\boldsymbol{t}_i\rangle\langle \boldsymbol{t}_i|A_j\rangle|1\rangle \right]. \tag{5}$$

After all unitary operations, we measure the third register and postselect on result 0 with the success probability:

$$P_\ell = \frac{1}{\|A\|_F^2} \sum_{j=1}^{m} \|A_j\|^2 \left\| |A_j\rangle - \sum_{i=1}^{\ell-1} |\boldsymbol{t}_i\rangle\langle \boldsymbol{t}_i|A_j\rangle \right\|^2, \tag{6}$$

and the postselected state (without the third register) is

$$|\phi_2^{(\ell)}\rangle = \frac{1}{\sqrt{P_\ell}\|A\|_F} \sum_{j=1}^{m} \|A_j\||j\rangle \left[|A_j\rangle - \sum_{i=1}^{\ell-1} |\boldsymbol{t}_i\rangle\langle \boldsymbol{t}_m|A_j\rangle \right]. \tag{7}$$

We need roughly $1/P_\ell$ copies of $|\phi_1^{(\ell)}\rangle$ to generate the state $|\phi_2^{(\ell)}\rangle$. Finally, we measure the first register for a new basis index $g(\ell)$ and a new orthogonal state $|\boldsymbol{t}_\ell\rangle$:

$$|\boldsymbol{t}_\ell\rangle = \frac{1}{Z_\ell} \left[|A_{g(\ell)}\rangle - \sum_{i=1}^{\ell-1} |\boldsymbol{t}_i\rangle\langle \boldsymbol{t}_i|A_{g(\ell)}\rangle \right], \tag{8}$$

where Z_ℓ is the normalizing constant. Specifically, denote the probability of the outcome $g(\ell)$ being $j \in [m]$ by $\Pr^{(\ell)}(j)$, and let $S_I = \{g(i)\}_{i=1}^{\ell-1}$. We have

$$\Pr^{(\ell)}(j) = \frac{\|A_j\|^2 \left\| \left(\prod_{k=1}^{\ell-1} R_k + I \right) |A_j\rangle \right\|^2}{\sum_{i=1}^{m} \|A_j\|^2 \left\| \left(\prod_{k=1}^{\ell-1} R_k + I \right) |A_j\rangle \right\|^2}, \tag{9}$$

$$\equiv \frac{\|A_j - \pi_{S_I}(A_j)\|^2}{\sum_{i=1}^{m} \|A_i - \pi_{S_I}(A_i)\|^2}, \tag{10}$$

where $\pi_{S_I}(A_j)$ denotes the projection of the row A_j on the row space of the submatrix $A(S_I, \cdot) \in \mathbb{R}^{(\ell-1)\times n}$. In other words, the new index is sampled with the probability proportional to the norm of orthogonal part of the row $A_{g(\ell)}$ to the current basis set S_I. After r iterations, we could obtain the index set $S_I = \{g(i)\}_{i=1}^{r}$ such that $\{A_{g(i)}\}_{i=1}^{r}$ forms a linearly independent basis. We remark that orthonormal states $\{|\boldsymbol{t}_i\rangle\}_{i=1}^{r}$ are generated from $\{|A_{g(i)}\rangle\}_{i=1}^{r}$ by performing Gram-Schmidt orthogonalization. Thus, an orthonormal basis could be also generated after the implementation of Algorithm I.

The technical difficulty of constructing the circuit in Fig. 1 comes from efficient implementation of the controlled version of reflection $R_\ell = I - 2|\boldsymbol{t}_\ell\rangle\langle \boldsymbol{t}_\ell|$, since we do not have additional quantum memory to store $\{|\boldsymbol{t}_\ell\rangle\}$ generated during the algorithm. To overcome this problem, we note that the state $|\boldsymbol{t}_\ell\rangle$ lies in $\text{span}\{|A_{g(i)}\rangle\}_{i=1}^{\ell}$, so that $|\boldsymbol{t}_\ell\rangle = \sum_{i=1}^{\ell} z_{i\ell}|A_{g(i)}\rangle$ for some coefficients $\{z_{i\ell}\}_{i=1}^{\ell}$. Instead, we could generate $|\boldsymbol{t}_i\rangle$ by the linear combination of unitary (LCU) method [56] with postselections. Let C_ℓ be the Gram matrix of $\{|A_{g(i)}\rangle\}_{i=1}^{\ell}$, and let $C_{\ell-1}$ be the submatrix of C_ℓ by deleting the last row and column. The following lemma shows that the coefficient vector $\boldsymbol{z}_\ell = (z_{1\ell}, \ldots, z_{\ell\ell})^T$ has a compact expression that only depends on the Gram matrices. The proof is provided in Appendix A.

Lemma 1. The coefficients in $|\boldsymbol{t}_\ell\rangle = \sum_{i=1}^{\ell} z_{i\ell}|A_{g(j)}\rangle$ could be written in the vector form $\boldsymbol{z}_\ell = \sqrt{\frac{|C_\ell|}{|C_{\ell-1}|}} C_\ell^{-1} \boldsymbol{e}_\ell$, where $\boldsymbol{e}_\ell = (0, 0, \ldots, 0, 1)^T \in \mathbb{R}^\ell$ and $|X|$ denotes the determinant of a matrix X.

We remark that each element in the matrix C_ℓ, i.e., the inner product between quantum states $\{|A_{g(i)}\rangle\}_{i=1}^{\ell}$, is unknown and needs to be estimated in practice. The error on elements in C_ℓ would influence the accuracy of coefficients $\boldsymbol{z}_\ell$, and consequently impacts the whole complexity of the state readout protocol. Let $\tilde{\boldsymbol{t}}_\ell = \sum_{i=1}^{\ell} \tilde{z}_{i\ell} A_{g(i)}/\|A_{g(i)}\|$ be the perturbed vector of $\boldsymbol{t}_\ell$, where $\{\tilde{z}_{i\ell}\}_{i=1}^{\ell}$ are the coefficients calculated following Lemma 1 with noisy Gram matrices $\tilde{C}_\ell$. Denote $\sigma_{\min}(C_\ell)$ as the least singular value of C_ℓ. We have the

following Lemma 2 to bound $\|\tilde{\boldsymbol{t}}_\ell - \boldsymbol{t}_\ell\|$, whose proof is given in Appendix B.

Lemma 2. If each element in $\tilde{C}_\ell$ deviates from that in C_ℓ by at most $\epsilon_C \leqslant \frac{\sigma^2_{\min}(C_\ell)}{80\ell^{5/2}}\epsilon_R$, then for any $\epsilon_R \in (0, 1)$, the ℓ^2 norm of the error between $\boldsymbol{t}_\ell$ and $\tilde{\boldsymbol{t}}_\ell$ is bounded as

$$\|\tilde{\boldsymbol{t}}_\ell - \boldsymbol{t}_\ell\| \leqslant \frac{\epsilon_R}{10}, \tag{11}$$

where $\tilde{\boldsymbol{t}}_\ell = \sum_{i=1}^{\ell} \tilde{z}_{i\ell} A_{g(i)}/\|A_{g(i)}\|$.

Lemma 1 and 2 complete preconditions to generate the state $|\boldsymbol{t}_\ell\rangle$ through the LCU method. Then, given copies of $|\boldsymbol{t}_\ell\rangle\langle\boldsymbol{t}_\ell|$, we can implement the controlled version of the gate $R_\ell = I - 2|\boldsymbol{t}_\ell\rangle\langle\boldsymbol{t}_\ell| = e^{-i\pi|\boldsymbol{t}_\ell\rangle\langle\boldsymbol{t}_\ell|}$ with the help of the Hamiltonian simulation developed in quantum PCA [57], as explained in Lemma 3.

Lemma 3. Given Eq. (11) in Lemma 2, the state $|\boldsymbol{t}_\ell\rangle$ could be prepared using $O(\ell\sigma_{\min}^{-1/2}(C_\ell))$ queries to the oracle U_A with the ℓ^2 norm error bounded by $\epsilon_R/5$. The operation $C(R_\ell)$ could be prepared using $O(\ell\sigma_{\min}^{-1/2}(C_\ell)\epsilon_R^{-1})$ queries to the oracle U_A with the spectral norm error of R_ℓ bounded by ϵ_R.

The proof is provided in Appendix C. As a natural corollary, the Gram-Schmidt orthonormal basis $\{|\boldsymbol{t}_\ell\rangle\}_{\ell=1}^r$ could be provided using $O(r^2\sigma_{\min}^{-1/2}(C_r))$ queries to the oracle U_A.

Notice that the complexity of implementing $C(R_\ell)$ depends on the least singular value of the Gram matrix C_ℓ, which is largely affected by the choice of the sampled basis $\{|A_{g(i)}\rangle\}_{i=1}^{\ell}$. A too small $\sigma_{\min}(C_\ell)$ will significantly increase the number of queries to the oracles. Notice that a group of basis with a small least singular value tends to have less probability being sampled, e.g., the probability of sampling a linearly dependent basis is 0 by Eq. (10). Through further analysis, we prove that the expectation of $\sigma_{\min}(C_\ell)$ with the distribution formed by Eq. (9) is lower bounded as

$$\mathop{\mathbb{E}}_{\Pr^{(1)}} \cdots \mathop{\mathbb{E}}_{\Pr^{(\ell)}} [\sigma_{\min}(C_\ell)] \geqslant \frac{r-\ell+1}{\ell r}\kappa^{2-2\ell}. \tag{12}$$

This statement also holds *approximately* if we take into account the error of implementing each R_i for $i \in [\ell - 1]$, as provided in Lemma 4.

Lemma 4. Given that each gate R_i in Algorithm I is implemented with error bounded by $\epsilon_R = \frac{1}{3r^5\kappa^{2r}}$, where r and κ is the rank and the condition number of A, respectively, we have

$$\mathbb{E}_{\tilde{P}}[\sigma_{\min}(C_\ell)] \geqslant \frac{2}{3} \mathop{\mathbb{E}}_{\Pr^{(1)}} \cdots \mathop{\mathbb{E}}_{\Pr^{(\ell)}} [\sigma_{\min}(C_\ell)],$$

where the distribution

$$\tilde{P}(s_1, \ldots, s_\ell) = \tilde{\Pr}^{(1)}(s_1) \ldots \tilde{\Pr}^{(\ell)}(s_\ell) \tag{13}$$

follows from Eq. (9) using noisy gates $\tilde{R}_i$.

The proof is very technical with lengthy steps. Hence, we delay their introduction to Appendix D.

As a result, we could perform Algorithm I a few times to generate a basis with bounded least singular value. The conclusion is summarized in Theorem 3 whose proof is given in Appendix E.

Theorem 3. By using $O(r^{27}\kappa^{14r})$ queries to input oracles V_A (1) and U_A (2), we could find a group of linearly independent states $\{|A_{g(i)}\rangle\}_{i=1}^r$, such that the least singular value of the Gram matrix C_r formed by $\{|A_{g(i)}\rangle\}_{i=1}^r$ is greater than $\frac{1}{2r^2\cdot\kappa^{2r-2}}$,

Algorithm 2. State readout.

Input: QRAM oracle U_A. Copies of state $|\boldsymbol{v}\rangle$. Orthonormal basis $\{|\boldsymbol{t}_i\rangle\}_{i=1}^r$. The precision parameter ϵ.
Output: Coordinates $\{x_i\}_{i=1}^r$ in $|\boldsymbol{v}\rangle = \sum_{i=1}^r x_i|\boldsymbol{s}_i\rangle$ that guarantee a ϵ accuracy under ℓ^2 norm.
1: Estimate the value $a_i^2 = |\langle\boldsymbol{v}|\boldsymbol{t}_i\rangle|^2$, for $i \in [r]$ by swap test. Mark $k := \mathrm{argmax}_{i\in[r]} a_i^2$.
2: Run the circuit in Fig. 2 to estimate $a_i' = \langle\boldsymbol{t}_k|\boldsymbol{v}\rangle\langle\boldsymbol{v}|\boldsymbol{t}_i\rangle$ for $i \in [r]$. Normalize the vector $\boldsymbol{a} = \boldsymbol{a}'/\|\boldsymbol{a}'\|$.
3: Output the solution as $\boldsymbol{x} = Z\boldsymbol{a}$, where Z is given in Eq. (14).

where r and κ are the rank and the condition number of A, respectively.

B. Coefficient calculation

Next we focus on the second subroutine. Once the row basis has been selected, which now we denote as $\{\boldsymbol{s}_i\}_{i=1}^r$ for simplicity, the readout problem reduces to obtaining coordinates $\{x_i\}_{i=1}^r$ in the description $|\boldsymbol{v}\rangle = \sum_{i=1}^r x_i|\boldsymbol{s}_i\rangle$. The steps are outlined in Algorithm II.

The idea of Algorithm II is fairly natural. Since the QGSP algorithm generates orthonormal states $\{|\boldsymbol{t}_i\rangle\}_{i=1}^r$, we could first calculate the coordinate of state $|\boldsymbol{v}\rangle$ under the basis $\{|\boldsymbol{t}_i\rangle\}_{i=1}^r$: $|\boldsymbol{v}\rangle = \sum_{i=1}^r a_i|\boldsymbol{t}_i\rangle$, and then transfer the orthonormal basis to the row basis $\{\boldsymbol{s}_i\}_{i=1}^r$:

$$(\boldsymbol{t}_1, \ldots, \boldsymbol{t}_r) = \left(\frac{\boldsymbol{s}_1}{\|\boldsymbol{s}_1\|}, \ldots, \frac{\boldsymbol{s}_r}{\|\boldsymbol{s}_r\|}\right)Z, \tag{14}$$

where $Z = [z_{ij}]_{r\times r}$ is the transformation matrix. The coordinates $\{x_i\}_{i=1}^r$ are given as $\boldsymbol{x} = Z\boldsymbol{a}$.

The crucial part of Algorithm II is to calculate the coefficient $a_i = \langle\boldsymbol{v}|\boldsymbol{t}_i\rangle$, $\forall i \in [r]$. However, the overlap estimation techniques based on the Hadamard test [58] could not be directly employed for estimating the state overlap, since the unitaries for generating the states are required. This drawback limits most quantum algorithms, e.g., the quantum linear system solver, that require postselection to yield the solution state easily. Another choice is the swap test [59] that only requires copies of states. However, directly using the quantum swap test could only obtain the estimation to the value $|\langle\boldsymbol{v}|\boldsymbol{t}_i\rangle|^2$, while $\mathrm{sign}(a_i)$ remains unknown. To overcome this difficulty, we could assume that the state $|\boldsymbol{v}\rangle$ has the positive overlap with one of the basis, say $|\boldsymbol{t}_k\rangle$, and take the value

$$a_i = \mathrm{sign}(\langle\boldsymbol{t}_k|\boldsymbol{v}\rangle\langle\boldsymbol{v}|\boldsymbol{t}_i\rangle)|\langle\boldsymbol{v}|\boldsymbol{t}_i\rangle| = \frac{\langle\boldsymbol{t}_k|\boldsymbol{v}\rangle\langle\boldsymbol{v}|\boldsymbol{t}_i\rangle}{|\langle\boldsymbol{t}_k|\boldsymbol{v}\rangle|} \tag{15}$$

as the state overlap. This assumption is equivalent to adding a global phase 0 or $e^{i\pi} = -1$ on $|\boldsymbol{v}\rangle$, and will not affect the extraction of the classical description.

We construct a variant of the swap test, illustrated in Fig. 2 for estimating $a_i' = \langle\boldsymbol{t}_k|\boldsymbol{v}\rangle\langle\boldsymbol{v}|\boldsymbol{t}_i\rangle$. It is easy to see that the probability of the measurement outcomes "00" and "11" yields the value a_i':

$$P_{\text{same}} = P_{00} + P_{11} = \frac{1 + \langle\boldsymbol{t}_k|\boldsymbol{v}\rangle\langle\boldsymbol{v}|\boldsymbol{t}_i\rangle}{2} = \frac{1 + a_i'}{2}. \tag{16}$$

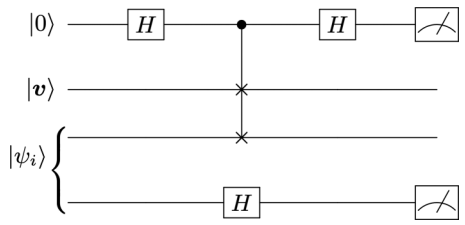

FIG. 2. Quantum circuit for estimating $\langle \boldsymbol{t}_k|\boldsymbol{v}\rangle\langle \boldsymbol{v}|\boldsymbol{t}_i\rangle$, where $|\psi_i\rangle = \frac{1}{\sqrt{2}}(|\boldsymbol{t}_k\rangle|0\rangle + |\boldsymbol{t}_i\rangle|1\rangle)$.

Similar to the swap test, the proposed quantum circuit provides a ϵ-error estimation to the value $\langle \boldsymbol{t}_k|\boldsymbol{v}\rangle\langle \boldsymbol{v}|\boldsymbol{t}_i\rangle$ with $\tilde{O}(\epsilon^{-2})$ measurements. Notice that a larger $|\langle \boldsymbol{t}_k|\boldsymbol{v}\rangle|$ is preferred to obtain more accurate estimations of a_i in Eq. (15) through the estimations of a_i' in Eq. (16). Thus, we mark $k := \text{argmax}_{i\in[r]}|\langle \boldsymbol{t}_i|\boldsymbol{v}\rangle|^2$ by using the swap test, before the estimations of $\{a_i'\}_{i=1}^r$ by running the circuit in Fig. 2.

The difficulty of implementing the quantum circuit in Fig. 2 is to efficiently prepare the state $(|\boldsymbol{t}_k\rangle|0\rangle + |\boldsymbol{t}_i\rangle|1\rangle)/\sqrt{2}$. We apply the linear combination of unitaries (LCU) method again such that $(|\boldsymbol{t}_k\rangle|0\rangle + |\boldsymbol{t}_i\rangle|1\rangle)/\sqrt{2}$ could be prepared with query complexity $O(r\sigma_{\min}^{-1/2}(C_r))$. See Appendix F for detail. By using this circuit along with the swap test, we could approximately calculate the coordinates $\{x_i\}_{i=1}^r$. The error and time complexity of Algorithm II is provided in Theorem 4, with proof given in Appendix F.

Theorem 4. Algorithm II provides a classical description $\boldsymbol{v} = \sum_{i=1}^r x_i A_{g(i)}/\|A_{g(i)}\|$ with ℓ^2 norm error bounded in ϵ, by using $O(r^4\epsilon^{-2})$ copies of state $|\boldsymbol{v}\rangle$ and $O(r^{10}\sigma_{\min}^{-4}(C_r)\epsilon^{-2})$ queries to input oracles.

Thus, our state readout protocol only requires $\tilde{O}(\text{poly}(r)\epsilon^{-2})$ copies of the unknown quantum state. The required state copy complexity is independent from the dimension of the state, which makes our algorithm more efficient than previous QST methods [36] in the low-rank case, since the latter needs at least $O(d\epsilon^{-2})$ copies. We remark that the combination of Theorems 3 and 4 yields the main result in Theorem 2.

III. APPLICATIONS

As introduced in previous text, our readout protocol suits the case that the output state of the quantum algorithm lies in the row space of the input matrix. We remark that this assumption is naturally satisfied by many proposed quantum algorithms in the field of machine learning and linear algebra. In this section, we discuss the end-to-end versions of two existing quantum algorithms: the quantum singular value decomposition (SVD) algorithm and the quantum linear system solver, when employing our state readout protocol for generating classical solutions.

A. Quantum singular value decomposition

We begin with the quantum singular value decomposition protocol. For a given r-rank input matrix $A = \sum_{i=1}^r \sigma_i \boldsymbol{u}_i \boldsymbol{v}_i^T \in \mathbb{R}^{m\times d}$, there is

$$\boldsymbol{v}_i = \frac{1}{\sigma_i}\left(\boldsymbol{u}_i^T A\right)^T = \frac{1}{\sigma_i}\sum_{j=1}^m u_i^{(j)} A_j, \quad \forall j \in [m], \tag{17}$$

so any singular vector $\boldsymbol{v}_i$ lies in the row space $\text{span}\{A_i\}_{i=1}^m$. Given QRAM oracles of the matrix A, quantum SVD allows us to perform the operation $\sum_j \beta_j|\boldsymbol{v}_j\rangle \to \sum_j \beta_j|\boldsymbol{v}_j\rangle|\hat{\sigma}_j\rangle$ with complexity $O(\text{polylog}(md)\|A\|_F\epsilon^{-1})$ such that $\hat{\sigma}_j \in \sigma_j \pm \epsilon$ with high probability. Consider the state

$$|0\rangle|0\rangle \xrightarrow{V_A U_A} \frac{\sum_{i=1}^m\sum_{j=1}^d A_{ij}|i\rangle|j\rangle}{\|A\|_F} = \frac{\sum_{i=1}^r \sigma_i|\boldsymbol{u}_i\rangle|\boldsymbol{v}_i\rangle}{\|A\|_F}$$

as the input to the quantum SVD algorithm to generate the state $\frac{1}{\|A\|_F}\sum_{i=1}^r \sigma_i|\boldsymbol{u}_i\rangle|\boldsymbol{v}_i\rangle|\hat{\sigma}_i\rangle$. Then the measurement on the eigenvalue register could collapse the state to different eigenstates $|\boldsymbol{u}_i\rangle|\boldsymbol{v}_i\rangle$ with probability $\frac{\sigma_i^2}{\|A\|_F^2}$. Thus, any target state $|\boldsymbol{v}_i\rangle$ could be prepared with complexity $O(\text{polylog}(md)\|A\|_F^3\Delta_\sigma^{-1}\sigma_i^{-2})$, where Δ_σ is the eigengap of the matrix A. Using this result along with Theorem 2, we could derive the end-to-end complexity for SVD as follows.

Corollary 1. The classical form of any eigenstate $|\boldsymbol{v}_i\rangle$ of A could be obtained by using $O(\kappa^{14r}\text{poly}(r, \log(md))\frac{\|A\|_F}{\Delta_\sigma\epsilon^2})$ queries to the input oracle of A, such that the ℓ^2 norm error is bounded in ϵ.

B. Quantum linear system solver

There has been an increasing interest in quantum machine learning [12,13,60] and linear algebra [23,28] algorithms following the quantum linear system solver proposed by Harrow *et al.* [9]. The first quantum linear system solver was proposed especially for the sparse case by Hamiltonian simulation, and several other different linear system solvers [20,61] have been proposed subsequently for the general case. Here we consider the quantum solver [20] which encodes the input matrix $A \in \mathbb{R}^{d\times d}$ into the QRAM model.

For matrix $A = \sum_{i=1}^r \sigma_i \boldsymbol{u}_i \boldsymbol{v}_i^T$, the solution could be written as

$$\boldsymbol{x} = A^+\boldsymbol{b}, \tag{18}$$

where $A^+ = \sum_{i=1}^r \frac{1}{\sigma_i}\boldsymbol{v}_i\boldsymbol{u}_i^T$ is the pseudoinverse matrix of A. Equation (18) gives $\boldsymbol{x} = \sum_{i=1}^r \frac{1}{\sigma_i}\boldsymbol{u}_i^T\boldsymbol{b}\boldsymbol{v}_i \in \text{span}\{\boldsymbol{v}_i\}_{i=1}^r$, which means $\boldsymbol{x}$ also lies in the row space $\text{span}\{A_i\}_{i=1}^n$ by using the previous conclusion about eigenvectors.

For the linear system $A\boldsymbol{x} = \boldsymbol{b}$, the solution state $|\boldsymbol{x}\rangle = |A^+\boldsymbol{b}\rangle$ could be prepared in time $O(\kappa^2\text{polylog}(d)\|A\|_F\epsilon^{-1})$ with ℓ^2 norm error bounded in ϵ, where κ is the condition number of A. Then we could derive the end-to-end complexity for the quantum linear system solver as follows.

Corollary 2. The classical form of the solution state $|A^+\boldsymbol{b}\rangle$ for the linear system $A\boldsymbol{x} = \boldsymbol{b}$ could be obtained by using $O(\kappa^{14r}\text{poly}(r, \log d)\frac{\|A\|_F}{\epsilon^3})$ queries to input oracles of A, such that the ℓ^2 norm error is bounded in ϵ.

IV. CONCLUSION AND DISCUSSION

In this work, we developed an efficient state readout framework for quantum algorithms which involve a low-rank input matrix and the output state $|\boldsymbol{v}\rangle$ lies in the row space of the input matrix. The proposed framework takes $\tilde{O}(\text{poly}(r)\epsilon^{-2})$ copies of the output state and $\tilde{O}(\text{poly}(r, \kappa^r)\epsilon^{-2})$ queries to input oracles for providing ϵ error bounded classical description. Thus, our protocol preserves the quantum speedup at the state

readout step of these quantum algorithms for the case that the rank r and the condition number κ are small, relative to the system dimension d. We analyzed the feasibility of our framework for quantum algorithms including the quantum SVD and the QRAM-based linear system solver in the low-rank case.

Recently, several quantum-inspired classical algorithms [62–65] have been developed as challenges to quantum advantage on machine learning tasks. Since QRAM oracles are employed in this work, we would like to emphasize the difference between these classical algorithms and the proposed readout protocol. Note that the state readout is a "pure quantum" task which aims to generate the classical form of the unknown quantum state. However, the quantum-inspired algorithms are developed for solving certain linear algebra problems if certain data structure and query access are allowed.

Finally, we believe that the proposed results about decoding the pure state could be extended into the mixed-state case. A quick outline of the procedure is as follows. We could first employ the quantum PCA [57] to perform the eigendecompositions, and then to decode the eigenstates using our protocol. Another future direction is to improve our readout framework such that the complexity is polynomial in both the rank and the condition number.

APPENDIX A: PROOF OF LEMMA 1

Proof. Denote $\boldsymbol{s}_i := A_{g(i)}$ for the simplicity of notation. Consider the state

$$|\boldsymbol{t}_\ell\rangle = \sum_{i=1}^{\ell} z_{i\ell}|\boldsymbol{s}_i\rangle = \frac{1}{Z_\ell}\left(|\boldsymbol{s}_\ell\rangle - \sum_{i=1}^{\ell-1}|\boldsymbol{t}_i\rangle\langle\boldsymbol{t}_i|\boldsymbol{s}_\ell\rangle\right), \tag{A1}$$

where Z_ℓ has another formulation obtained by multiplying $\langle\boldsymbol{t}_\ell|$ on both sides

$$1 = \langle\boldsymbol{t}_\ell|\boldsymbol{t}_\ell\rangle = \frac{1}{Z_\ell}\langle\boldsymbol{t}_\ell|\boldsymbol{s}_\ell\rangle = \frac{1}{Z_\ell}\sum_{i=1}^{\ell} z_{i\ell}\langle\boldsymbol{s}_i|\boldsymbol{s}_\ell\rangle. \tag{A2}$$

The restriction that $|\boldsymbol{t}_\ell\rangle$ is normalized and is orthogonal to states $|\boldsymbol{s}_1\rangle, |\boldsymbol{s}_2\rangle, \ldots |\boldsymbol{s}_{\ell-1}\rangle$ could yield

$$\langle\boldsymbol{s}_j|\boldsymbol{t}_\ell\rangle = \sum_{i=1}^{\ell} z_{i\ell}\langle\boldsymbol{s}_j|\boldsymbol{s}_i\rangle = 0, \quad \forall j \in [\ell-1], \tag{A3}$$

$$\langle\boldsymbol{t}_\ell|\boldsymbol{t}_\ell\rangle = \sum_{j=1}^{\ell}\sum_{i=1}^{\ell} z_{j\ell}z_{i\ell}\langle\boldsymbol{s}_j|\boldsymbol{s}_i\rangle = 1. \tag{A4}$$

Rewrite Eqs. (A2) and (A3) in the vector form:

$$C_\ell \boldsymbol{z}_\ell = Z_\ell \boldsymbol{e}_\ell. \tag{A5}$$

Equation (A4) could be written as

$$1 = \sum_{i,j=1}^{\ell} z_{i\ell}z_{j\ell}\langle\boldsymbol{s}_j|\boldsymbol{s}_i\rangle = \boldsymbol{z}_\ell^T C_\ell \boldsymbol{z}_\ell = Z_\ell^2 \boldsymbol{e}_\ell^T C_\ell^{-1}\boldsymbol{e}_\ell = Z_\ell^2\frac{|C_{\ell-1}|}{|C_\ell|},$$

where the third equation derives from $\boldsymbol{z}_\ell = Z_\ell C_\ell^{-1}\boldsymbol{e}_\ell$ by Eq. (A5) and the last equation is derived by noticing that the (ℓ,ℓ)-th element of C_ℓ^{-1} is $\frac{|C_{\ell-1}|}{|C_\ell|}$. Thus, we obtain

$$Z_\ell = \sqrt{\frac{|C_\ell|}{|C_{\ell-1}|}}. \tag{A6}$$

Finally, solving (A5) is trivial:

$$\boldsymbol{z}_\ell = Z_\ell C_\ell^{-1}\boldsymbol{e}_\ell = \sqrt{\frac{|C_\ell|}{|C_{\ell-1}|}}C_\ell^{-1}\boldsymbol{e}_\ell. \tag{A7}$$

■

APPENDIX B: PROOF OF LEMMA 2

Proof. We denote $\|\cdot\|$ as the ℓ^2 norm and the spectral norm for vectors and matrices.

First notice that

$$\|\boldsymbol{t}_\ell - \tilde{\boldsymbol{t}}_\ell\|^2 = \boldsymbol{t}_\ell^T\boldsymbol{t}_\ell - 2\boldsymbol{t}_\ell^T\tilde{\boldsymbol{t}}_\ell + \tilde{\boldsymbol{t}}_\ell^T\tilde{\boldsymbol{t}}_\ell \tag{B1}$$

$$= \boldsymbol{z}_\ell^T C_\ell \boldsymbol{z}_\ell - 2\boldsymbol{z}_\ell^T C_\ell\tilde{\boldsymbol{z}}_\ell + \tilde{\boldsymbol{z}}_\ell^T C_\ell\tilde{\boldsymbol{z}}_\ell \tag{B2}$$

$$= \Delta\boldsymbol{z}_\ell^T C_\ell \Delta\boldsymbol{z}_\ell \tag{B3}$$

$$\leqslant \|C_\ell\|\|\Delta\boldsymbol{z}_\ell\|^2, \tag{B4}$$

where C_ℓ in Eq. (B2) is the Gram matrix of $\{|A_{g(i)}\rangle\}_{i=1}^{\ell}$, and $\Delta\boldsymbol{z}_\ell = \tilde{\boldsymbol{z}}_\ell - \boldsymbol{z}_\ell$ in Eq. (B3). Since $\|C_\ell\| \leqslant \mathrm{Tr}[C_\ell] = \ell$, we can obtain the desired result, namely,

$$\|\boldsymbol{t}_\ell - \tilde{\boldsymbol{t}}_\ell\|^2 \leqslant \frac{\epsilon_R^2}{100}, \tag{B5}$$

if the following claim is true:

$$\|\Delta\boldsymbol{z}_\ell\| \leqslant \frac{\epsilon_R}{10\ell^{1/2}}. \tag{B6}$$

To prove Eq. (B6), let us introduce some more notation. Denote by $\tilde{C}_\ell$ and $\tilde{C}_{\ell-1}$ the perturbed Gram matrices of C_ℓ and $C_{\ell-1}$, respectively. Let $\tilde{Z}_\ell = \sqrt{\frac{|\tilde{C}_\ell|}{|\tilde{C}_{\ell-1}|}}$ and

$$\tilde{\boldsymbol{z}}_\ell = \tilde{Z}_\ell\tilde{C}_\ell^{-1}\boldsymbol{e}_\ell. \tag{B7}$$

Let $\Delta C_\ell = \tilde{C}_\ell - C_\ell$, and $\Delta Z_\ell = \tilde{Z}_\ell - Z_\ell$. Then,

$$\|\Delta\boldsymbol{z}_\ell\| = \|\tilde{C}_\ell^{-1}\tilde{Z}_\ell\boldsymbol{e}_\ell - C_\ell^{-1}Z_\ell\boldsymbol{e}_\ell\| \tag{B8}$$

$$= \|(C_\ell + \Delta C_\ell)^{-1}(Z_\ell\boldsymbol{e}_\ell + \Delta Z_\ell\boldsymbol{e}_\ell) - C_\ell^{-1}Z_\ell\boldsymbol{e}_\ell\| \tag{B9}$$

$$= \|(C_\ell + \Delta C_\ell)^{-1}[(Z_\ell\boldsymbol{e}_\ell + \Delta Z_\ell\boldsymbol{e}_\ell) - (C_\ell + \Delta C_\ell)C_\ell^{-1}Z_\ell\boldsymbol{e}_\ell]\| \tag{B10}$$

$$= \left\|(C_\ell + \Delta C_\ell)^{-1}\left(\Delta Z_\ell\boldsymbol{e}_\ell - \Delta C_\ell C_\ell^{-1}Z_\ell\boldsymbol{e}_\ell\right)\right\| \tag{B11}$$

$$\leqslant \|(C_\ell + \Delta C_\ell)^{-1}\| \cdot \left(|\Delta Z_\ell| + \|\Delta C_\ell\|\left\|C_\ell^{-1}\right\|Z_\ell\right), \tag{B12}$$

where Eq. (B12) follows from the triangular inequality.

Since each element in $\tilde{C}_\ell$ diviates from that in C_ℓ by at most $\epsilon_C \leqslant \frac{\sigma_{\min}^2(C_\ell)}{80\ell^{5/2}}\epsilon_R$, we could obtain

$$\|\Delta C_\ell\| \leqslant \|\Delta C_\ell\|_F \leqslant \sqrt{\ell^2\epsilon_C^2} = \ell\epsilon_C \tag{B13}$$

and

$$\|(C_\ell+\Delta C_\ell)^{-1}\| = \frac{1}{\sigma_{\min}(C_\ell+\Delta C_\ell)} \tag{B14}$$

$$\leqslant \frac{1}{\sigma_{\min}(C_\ell)-\|\Delta C_\ell\|} \tag{B15}$$

$$\leqslant \frac{1}{\sigma_{\min}(C_\ell)-\ell\epsilon_C} \tag{B16}$$

$$\leqslant \frac{80}{79}\sigma_{\min}^{-1}(C_\ell). \tag{B17}$$

Equation (B15) follows from the Weyl's inequality

$$|\sigma_{\min}(C_\ell+\Delta C_\ell)-\sigma_{\min}(C_\ell)| \leqslant \|\Delta C_\ell\|.$$

Equation (B16) employs Eq. (B13). Equation (B17) follows because

$$\ell\epsilon_C \leqslant \frac{\sigma_{\min}^2(C_\ell)}{80\ell^{3/2}}\epsilon_R \leqslant \frac{1}{80}\sigma_{\min}(C_\ell).$$

Together with Eqs. (B13), (B17), and $\|C_\ell^{-1}\| = \sigma_{\min}^{-1}(C_\ell)$, Eq. (B12) is upper bounded by

$$\|\Delta \boldsymbol{z}_\ell\| \leqslant \frac{80}{79}\sigma_{\min}^{-1}(C_\ell)\big[|\Delta Z_\ell| + \ell\epsilon_C\sigma_{\min}^{-1}(C_\ell)Z_\ell\big]. \tag{B18}$$

To finish the proof of Eq. (B6), we only need to bound

$$|\Delta Z_\ell| \leqslant 4\ell^2\sigma_{\min}^{-1}(C_\ell)\epsilon_C Z_\ell. \tag{B19}$$

If Eq. (B19) were true, we could further bound $\|\Delta \boldsymbol{z}_\ell\|$ from Eq. (B18) as follows:

$$\|\Delta \boldsymbol{z}_\ell\| \leqslant \frac{80}{79}\sigma_{\min}^{-2}(C_\ell)5\ell^2\epsilon_C Z_\ell \tag{B20}$$

$$\leqslant \frac{\epsilon_R}{10\ell^{1/2}}, \tag{B21}$$

where $\ell > 1$, $\epsilon_C \leqslant \frac{\sigma_{\min}^2(C_\ell)}{80\ell^{5/2}}\epsilon_R$, and

$$Z_\ell = \left\| |\boldsymbol{s}_\ell\rangle - \sum_{i=1}^{\ell-1} |\boldsymbol{t}_i\rangle\langle \boldsymbol{t}_i|\boldsymbol{s}_\ell\rangle \right\| \leqslant 1. \tag{B22}$$

The last part of this section is to prove Eq. (B19). To further analyze this term, we utilize the bound on the determinant of the perturbed matrix [66], page 113]:

$$\left|\frac{|C_\ell+\Delta C_\ell|-|C_\ell|}{|C_\ell|}\right| \leqslant \frac{\ell\|C_\ell^{-1}\|\|\Delta C_\ell\|}{1-\ell\|C_\ell^{-1}\|\|\Delta C_\ell\|}. \tag{B23}$$

We can obtain

$$\left|\frac{|C_{\ell-1}+\Delta C_{\ell-1}|-|C_{\ell-1}|}{|C_{\ell-1}|}\right| \leqslant \frac{(\ell-1)\|C_{\ell-1}^{-1}\|\|\Delta C_{\ell-1}\|}{1-(\ell-1)\|C_{\ell-1}^{-1}\|\|\Delta C_{\ell-1}\|}$$

$$\leqslant \frac{\ell\|C_\ell^{-1}\|\|\Delta C_\ell\|}{1-\ell\|C_\ell^{-1}\|\|\Delta C_\ell\|}, \tag{B24}$$

where the second inequality follows by noticing that the function $f(x) = \frac{x}{1-x}$ is monotonically increasing and the property that the range of singular values of the submatrix is contained in that of the original matrix:

$$(\ell-1)\|C_{\ell-1}^{-1}\|\|\Delta C_{\ell-1}\| \leqslant \ell\|C_{\ell-1}^{-1}\|\|\Delta C_{\ell-1}\|$$

$$= \ell\sigma_{\min}^{-1}(C_{\ell-1})\sigma_{\max}(\Delta C_{\ell-1})$$

$$\leqslant \ell\sigma_{\min}^{-1}(C_\ell)\sigma_{\max}(\Delta C_\ell)$$

$$= \ell\|C_\ell^{-1}\|\|\Delta C_\ell\|.$$

Consequently, we have the bound on the term $|\Delta Z_\ell|$:

$$|\Delta Z_\ell| = \left|\frac{\tilde{Z}_\ell}{Z_\ell}-1\right| Z_\ell \tag{B25}$$

$$= \left|\sqrt{\frac{|\tilde{C}_\ell|}{|C_\ell|}}\sqrt{\frac{|C_{\ell-1}|}{|\tilde{C}_{\ell-1}|}}-1\right| Z_\ell \tag{B26}$$

$$= \left|\sqrt{\frac{|C_\ell+\Delta C_\ell|}{|C_\ell|}}\sqrt{\frac{|C_{\ell-1}|}{|C_{\ell-1}+\Delta C_{\ell-1}|}}-1\right| Z_\ell \tag{B27}$$

$$\leqslant \max\left(\sqrt{\frac{1}{1-\ell\|C_\ell^{-1}\|\|\Delta C_\ell\|}\frac{1-\ell\|C_\ell^{-1}\|\|\Delta C_\ell\|}{1-2\ell\|C_\ell^{-1}\|\|\Delta C_\ell\|}}-1,\right.$$

$$\left.1-\sqrt{\frac{1-2\ell\|C_\ell^{-1}\|\|\Delta C_\ell\|}{1-\ell\|C_\ell^{-1}\|\|\Delta C_\ell\|}\big(1-\ell\|C_\ell^{-1}\|\|\Delta C_\ell\|\big)}\right) Z_\ell \tag{B28}$$

where Eq. (B28) is derived by employing the following equivalent form of Eqs. (B23) and (B24):

$$\frac{|C_\ell+\Delta C_\ell|}{|C_\ell|} \geqslant \frac{1-2\ell\|C_\ell^{-1}\|\|\Delta C_\ell\|}{1-\ell\|C_\ell^{-1}\|\|\Delta C_\ell\|},$$

$$\frac{|C_\ell+\Delta C_\ell|}{|C_\ell|} \leqslant \frac{1}{1-\ell\|C_\ell^{-1}\|\|\Delta C_\ell\|},$$

$$\frac{|C_{\ell-1}+\Delta C_{\ell-1}|}{|C_{\ell-1}|} \geqslant \frac{1-2\ell\|C_\ell^{-1}\|\|\Delta C_\ell\|}{1-\ell\|C_\ell^{-1}\|\|\Delta C_\ell\|},$$

$$\frac{|C_{\ell-1}+\Delta C_{\ell-1}|}{|C_{\ell-1}|} \leqslant \frac{1}{1-\ell\|C_\ell^{-1}\|\|\Delta C_\ell\|}.$$

Since $\max(A,B) \leqslant A+B$ for $A, B \geqslant 0$, Eq. (B28) yields

$$\leqslant \left(\sqrt{\frac{1}{1-2\ell\|C_\ell^{-1}\|\|\Delta C_\ell\|}}-\sqrt{1-2\ell\|C_\ell^{-1}\|\|\Delta C_\ell\|}\right) Z_\ell \tag{B29}$$

$$= \frac{2\ell\|C_\ell^{-1}\|\|\Delta C_\ell\|}{\sqrt{1-2\ell\|C_\ell^{-1}\|\|\Delta C_\ell\|}} Z_\ell \tag{B30}$$

$$\leqslant \frac{2\ell^2\sigma_{\min}^{-1}(C_\ell)\epsilon_C}{\sqrt{1-2\ell^2\sigma_{\min}^{-1}(C_\ell)\epsilon_C}} Z_\ell \tag{B31}$$

$$\leqslant 4\ell^2\sigma_{\min}^{-1}(C_\ell)\epsilon_C Z_\ell. \tag{B32}$$

Equation (B31) is derived by using Eqs. (B13), (B22), and $\|C_\ell^{-1}\| = \sigma_{\min}^{-1}(C_\ell)$. The last equation holds because

$$\sqrt{1-2\ell^2\sigma_{\min}^{-1}(C_\ell)\epsilon_C} \geqslant \sqrt{1-2\ell^2\sigma_{\min}^{-1}(C_\ell)\frac{\sigma_{\min}^2(C_\ell)}{80\ell^{5/2}}\epsilon_R}$$

$$\geqslant \sqrt{1-\frac{1}{4}} \geqslant \frac{1}{2},$$

which is obtained by using the bound of ϵ_C and $\frac{\sigma_{\min}(C_\ell)}{40\ell^{1/2}}\epsilon_R \leqslant \frac{1}{4}$. ■

APPENDIX C: PROOF OF LEMMA 3

Proof. The main idea is to firstly derive the error analysis of $|\boldsymbol{t}_\ell\rangle$ and R_ℓ, followed by the development of the LCU protocol. Denote $\boldsymbol{s}_i := A_{g(i)}$ for the simplicity of notation. We begin from the assumption that

$$\|\tilde{\boldsymbol{t}}_\ell - \boldsymbol{t}_\ell\| \leqslant \frac{\epsilon_R}{10}, \tag{C1}$$

where $\tilde{\boldsymbol{t}}_\ell = \sum_{i=1}^{\ell} \tilde{z}_{i\ell}\boldsymbol{s}_i/\|\boldsymbol{s}_i\|$. Then the ℓ^2 norm of the error of the state $|\boldsymbol{t}_\ell\rangle$ is bounded as follows:

$$\||\tilde{\boldsymbol{t}}_\ell\rangle - |\boldsymbol{t}_\ell\rangle\| \leqslant \|\tilde{\boldsymbol{t}}_\ell - |\tilde{\boldsymbol{t}}_\ell\rangle\| + \|\tilde{\boldsymbol{t}}_\ell - |\boldsymbol{t}_\ell\rangle\| \tag{C2}$$

$$= |\|\tilde{\boldsymbol{t}}_\ell\| - 1|\||\tilde{\boldsymbol{t}}_\ell\rangle\| + \|\tilde{\boldsymbol{t}}_\ell - \boldsymbol{t}_\ell\| \tag{C3}$$

$$= |\|\tilde{\boldsymbol{t}}_\ell\| - \|\boldsymbol{t}_\ell\|| + \|\tilde{\boldsymbol{t}}_\ell - \boldsymbol{t}_\ell\| \tag{C4}$$

$$\leqslant \|\tilde{\boldsymbol{t}}_\ell - \boldsymbol{t}_\ell\| + \|\tilde{\boldsymbol{t}}_\ell - \boldsymbol{t}_\ell\| \tag{C5}$$

$$\leqslant \frac{\epsilon_R}{5}. \tag{C6}$$

Equations (C2)–(C6) are derived by using $\|\boldsymbol{t}_\ell\| = 1$, the triangular inequality, and Eq. (C1). We could further provide the spectral norm of the error of the gate R_ℓ:

$$\|\tilde{R}_\ell - R_\ell\| = \|(I - 2|\tilde{\boldsymbol{t}}_\ell\rangle\langle\tilde{\boldsymbol{t}}_\ell|) - (I - 2|\boldsymbol{t}_\ell\rangle\langle\boldsymbol{t}_\ell|)\| \tag{C7}$$

$$= 2\||\tilde{\boldsymbol{t}}_\ell\rangle\langle\tilde{\boldsymbol{t}}_\ell| - |\boldsymbol{t}_\ell\rangle\langle\boldsymbol{t}_\ell|\| \tag{C8}$$

$$\leqslant 2\||\tilde{\boldsymbol{t}}_\ell\rangle\langle\tilde{\boldsymbol{t}}_\ell| - |\tilde{\boldsymbol{t}}_\ell\rangle\langle\boldsymbol{t}_\ell|\|_F + 2\||\tilde{\boldsymbol{t}}_\ell\rangle\langle\boldsymbol{t}_\ell| - |\boldsymbol{t}_\ell\rangle\langle\boldsymbol{t}_\ell|\| \tag{C9}$$

$$= 2\||\tilde{\boldsymbol{t}}_\ell\rangle - |\boldsymbol{t}_\ell\rangle\| + 2\||\tilde{\boldsymbol{t}}_\ell\rangle - |\boldsymbol{t}_\ell\rangle\| \tag{C10}$$

$$\leqslant \frac{4}{5}\epsilon_R. \tag{C11}$$

Equation (C7) is derived due to the definition of R_ℓ. Equation (C9) is derived by using the triangular inequality. Equation (C11) is derived by using Eq. (C6).

Now we provide a framework to implement operations $C(\tilde{R}_\ell)$ using coefficients $\{\tilde{z}_{j\ell}\}_{j=1}^{\ell}$. We could first prepare the pure state $\tilde{\rho}_\ell = |\tilde{\boldsymbol{t}}_\ell\rangle\langle\tilde{\boldsymbol{t}}_\ell|$ by the linear combination of unitaries method as follows. Firstly, initialize the state $|0\rangle^{\otimes \log m}|0\rangle^{\otimes \log n}|0\rangle$. Then, we apply Hadamard operations on the last $\log \ell$ qubits in the first register to create the state

$$\frac{1}{\sqrt{\ell}}\sum_{i=1}^{\ell}|i\rangle|0\rangle|0\rangle.$$

Next, we employ the operation

$$U_{\text{index}} = \prod_{i=1}^{\ell}(I - |i\rangle\langle i| - |g(i)\rangle\langle g(i)| + |i\rangle\langle g(i)| + |g(i)\rangle\langle i|) \tag{C12}$$

to swap states $|i\rangle$ and $|g(i)\rangle$, $\forall i \in [\ell]$, to yield the state

$$\frac{1}{\sqrt{\ell}}\sum_{i=1}^{\ell}|g(i)\rangle|0\rangle|0\rangle.$$

The unitary U_{index} could be implemented by $O(\ell)$ operations. Then we employ the oracle U_A on the first and the second register, followed by the unitary $U_{\text{index}}^\dagger$, to yield

$$\frac{1}{\sqrt{\ell}}\sum_{i=1}^{\ell}|i\rangle|A_{g(i)}\rangle|0\rangle \equiv \frac{1}{\sqrt{\ell}}\sum_{i=1}^{\ell}|i\rangle|\boldsymbol{s}_i\rangle|0\rangle.$$

Denote $\tilde{z}_\ell \equiv \max_i|\tilde{z}_{i\ell}|$. Then we perform the controlled rotation

$$\sum_{i=1}^{\ell}|i\rangle\langle i| \otimes e^{-i\sigma_y \arccos(\tilde{z}_{i\ell}/\tilde{z}_\ell)} + \sum_{i=\ell+1}^{m}|i\rangle\langle i| \otimes I$$

on the third register, conditioned on the first register $|i\rangle$, to obtain

$$\frac{1}{\sqrt{\ell}}\sum_{i=1}^{\ell}|i\rangle|\boldsymbol{s}_i\rangle\left(\frac{\tilde{z}_{i\ell}}{\tilde{z}_\ell}|0\rangle + \sqrt{1 - \frac{\tilde{z}_{i\ell}^2}{\tilde{z}_\ell^2}}|1\rangle\right).$$

Finally, we employ Hadamard operations on last $\log \ell$ qubits in the first register, to obtain the state

$$\frac{1}{\ell}\sum_{i=1}^{\ell}|0\rangle\frac{\tilde{z}_{i\ell}}{\tilde{z}_\ell}|\boldsymbol{s}_i\rangle|0\rangle + \text{ orthogonal garbage state}$$

$$= \frac{\|\tilde{\boldsymbol{t}}_\ell\|}{\ell\cdot\tilde{z}_\ell}|0\rangle|\tilde{\boldsymbol{t}}_\ell\rangle|0\rangle + \text{ orthogonal garbage state.}$$

The measurement on the first and the third registers of the final state could yield state $|\tilde{\boldsymbol{t}}_\ell\rangle$ with success probability $\|\tilde{\boldsymbol{t}}_\ell\|^2/\ell^2\tilde{z}_\ell^2$, so we could prepare the state $|\boldsymbol{t}_\ell\rangle$ with $O(\ell\tilde{z}_\ell/\|\tilde{\boldsymbol{t}}_\ell\|)$ queries to U_A by using the amplitude amplification method [67].

Note that operations $\tilde{R}_\ell = I - 2|\tilde{\boldsymbol{t}}_\ell\rangle\langle\tilde{\boldsymbol{t}}_\ell|$ can be viewed as the unitary with Hamiltonian $\tilde{\rho}_\ell = |\tilde{\boldsymbol{t}}_\ell\rangle\langle\tilde{\boldsymbol{t}}_\ell|$:

$$\begin{aligned} e^{-i\pi\tilde{\rho}_\ell} &= 1 + (-i\pi\tilde{\rho}_\ell) + \frac{1}{2!}(-i\pi\tilde{\rho}_\ell)^2 + \cdots \\ &= 1 - \tilde{\rho}_\ell + \tilde{\rho}_\ell\left[1 + (-i\pi) + \frac{1}{2!}(-i\pi)^2 + \cdots\right] \\ &= 1 - \tilde{\rho}_\ell + \tilde{\rho}_\ell e^{-i\pi} \\ &= I - 2|\tilde{\boldsymbol{t}}_\ell\rangle\langle\tilde{\boldsymbol{t}}_\ell|. \end{aligned}$$

Therefore, by using the Hamiltonian simulation method developed in quantum PCA [57], the controlled version of $\tilde{R}_\ell$ could be performed with error $\epsilon_R/5$ consuming $O(5\pi^2/\epsilon_R) = O(1/\epsilon_R)$ copies of $\tilde{\rho}_\ell$. Taking the complexity of generating state $|\tilde{\boldsymbol{t}}_\ell\rangle$ into account, we could implement operation $C(\tilde{R}_\ell)$ with the error of $\tilde{R}(\ell)$ bounded as $\epsilon_R/5$, by using $O(\ell\max_i|\tilde{z}_{i\ell}|/(\|\tilde{\boldsymbol{t}}_\ell\|\epsilon_R))$ queries to U_A. We remark that the ℓ^2 norm of vector $\tilde{\boldsymbol{z}}_\ell$ is bounded as

$$1 = \langle\tilde{\boldsymbol{t}}_\ell|\tilde{\boldsymbol{t}}_\ell\rangle = \frac{\tilde{\boldsymbol{z}}_\ell^T C_\ell \tilde{\boldsymbol{z}}_\ell}{\|\tilde{\boldsymbol{t}}_\ell\|^2} \geqslant \frac{\|\tilde{\boldsymbol{z}}_\ell\|^2}{\|\tilde{\boldsymbol{t}}_\ell\|^2}\sigma_{\min}(C_\ell),$$

which yields

$$\frac{\max_i|\tilde{z}_{i\ell}|}{\|\tilde{\boldsymbol{t}}_\ell\|} \leqslant \frac{\|\tilde{\boldsymbol{z}}_\ell\|}{\|\tilde{\boldsymbol{t}}_\ell\|} \leqslant \sigma_{\min}^{-1/2}(C_\ell). \tag{C13}$$

So the query complexity for implementing $C(\tilde{R}_\ell)$ could be bounded as $O(\ell\sigma_{\min}^{-1/2}(C_\ell)\epsilon_R^{-1})$. By considering the distance between R_ℓ and $\tilde{R}_\ell$ in Eq. (C11), we could then implement the controlled version of the gate R_ℓ with error bounded by ϵ_R. Now we have proved Lemma 3. ∎

APPENDIX D: PROOF OF LEMMA 4

In this section, we prove Lemma 4. Before we detail main technical procedures, we first provide some useful theoretical bounds in Lemmas 5 and 6.

Lemma 5. The probability P_ℓ defined in Eq. (6) is bounded by

$$\frac{\sum_{i=\ell}^{r}\sigma_i^2}{\|A\|_F^2} \leqslant P_\ell \leqslant \frac{\sum_{i=1}^{r-\ell+1}\sigma_i^2}{\|A\|_F^2},$$

where $\sigma_1 \geqslant \sigma_2 \geqslant \cdots \geqslant \sigma_r$ are singular values of A.

Proof. Denote the singular value decomposition

$$A = \sum_{i=1}^{r}\sigma_i \boldsymbol{u}_i \boldsymbol{v}_i^T.$$

Since the state $|\boldsymbol{t}_i\rangle$ is the linear sum of rows $\{A_j\}_{j=1}^m$, while each row is the linear sum of singular vectors

$$A_j = \sum_{i=1}^{r}\sigma_i u_i^{(j)}\boldsymbol{v}_i, \tag{D1}$$

we can further write

$$|\boldsymbol{t}_i\rangle = \sum_{j=1}^{r} w_{ij}|\boldsymbol{v}_j\rangle. \tag{D2}$$

Rewrite Eq. (6) as

$$P_\ell = \frac{1}{\|A\|_F^2}\sum_{j=1}^{m}\left[\|A_j\|^2 - \sum_{i=1}^{\ell-1}\|A_j\|^2|\langle \boldsymbol{t}_i|A_j\rangle|^2\right] \tag{D3}$$

$$= 1 - \frac{1}{\|A\|_F^2}\sum_{j=1}^{m}\sum_{i=1}^{\ell-1}\left[\sum_{k=1}^{r} w_{ik}\sigma_k u_k^{(j)}\right]^2, \tag{D4}$$

where Eq. (D4) comes from Eqs. (D1) and (D2). Expanding the square term in Eq. (D4) yields

$$P_\ell = 1 - \frac{1}{\|A\|_F^2}\sum_{j=1}^{m}\sum_{i=1}^{\ell-1}\left[\sum_{k=1}^{r} w_{ik}^2\sigma_k^2\left(u_k^{(j)}\right)^2 + \sum_{k\neq k'}^{r} w_{ik}w_{ik'}\sigma_k\sigma_{k'}u_k^{(j)}u_{k'}^{(j)}\right] \tag{D5}$$

$$= 1 - \frac{1}{\|A\|_F^2}\sum_{i=1}^{\ell-1}\sum_{k=1}^{r} w_{ik}^2\sigma_k^2 \tag{D6}$$

$$= 1 - \frac{1}{\|A\|_F^2}\sum_{k=1}^{r} c_k\sigma_k^2, \tag{D7}$$

where Eq. (D6) follows because $\sum_{j=1}^{m} u_k^{(j)}u_{k'}^{(j)} = \boldsymbol{u}_k^T\boldsymbol{u}_{k'} = \delta_{kk'}$, and we denote $c_k = \sum_{i=1}^{\ell-1} w_{ik}^2$ in Eq. (D7).

Define the r-dimensional vector $\boldsymbol{w}_i = \sum_{k=1}^{r} w_{ik}\boldsymbol{e}_k$. Since $\langle \boldsymbol{t}_i|\boldsymbol{t}_j\rangle = \delta_{ij} = \sum_{k=1}^{r} w_{ik}w_{jk} = \boldsymbol{w}_i^T\boldsymbol{w}_j$, vectors in set $\{\boldsymbol{w}_i\}_{i=1}^{\ell-1}$ are orthogonal with each other. We can add $\boldsymbol{w}_\ell, \ldots, \boldsymbol{w}_r$ such that $\{\boldsymbol{w}_i\}_{i=1}^{r}$ forms an orthonormal basis in the r-dimensional space. Denote the matrix $W = (\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_r)$. Since $W^TW = I$, we have

$$0 \leqslant c_k = \sum_{i=1}^{\ell-1} w_{ik}^2 \leqslant \sum_{i=1}^{r} w_{ik}^2 = [WW^T]_{kk} = 1, \quad \forall k \in [r]. \tag{D8}$$

Note that

$$\sum_{k=1}^{r} c_k = \sum_{i=1}^{\ell-1}\sum_{k=1}^{r} w_{ik}^2 = \sum_{i=1}^{\ell-1}[WW^T]_{ii} = \ell - 1. \tag{D9}$$

Hence, by using Eqs. (D7)–(D9) and $\|A\|_F^2 = \sum_{i=1}^{r}\sigma_i^2$, we could obtain the lower and upper bounds for P_ℓ as follows:

$$P_\ell \geqslant 1 - \frac{1}{\|A\|_F^2}\sum_{i=1}^{\ell-1}\sigma_i^2 = \frac{\sum_{i=\ell}^{r}\sigma_i^2}{\|A\|_F^2}, \tag{D10}$$

$$P_\ell \leqslant 1 - \frac{1}{\|A\|_F^2}\sum_{i=r-\ell+2}^{r}\sigma_i^2 = \frac{\sum_{i=1}^{r-\ell+1}\sigma_i^2}{\|A\|_F^2}. \tag{D11}$$

■

Lemma 6. Denote P to be the distribution of the adaptive sampling following from the Eq. (10):

$$P(s_1, \ldots, s_\ell) = \Pr^{(1)}(s_1)\Pr^{(2)}(s_2)\ldots\Pr^{(\ell)}(s_\ell), \tag{D12}$$

where $s_\ell \in [m]$ denotes the index of the row $\boldsymbol{s}_\ell$ in the matrix $A \in \mathbb{R}^{m\times d}$. Then

$$\mathbb{E}_P[\sigma_{\min}(C_\ell)] \geqslant \frac{r-\ell+1}{\ell r}\kappa^{2-2\ell}.$$

Proof. By the Cauchy-Schwarz inequality, we have

$$\mathbb{E}_P[\sigma_{\min}(C_\ell)]\cdot\mathbb{E}_P\left[\sigma_{\min}^{-1}(C_\ell)\right] \geqslant 1. \tag{D13}$$

If the following inequality were true,

$$\mathbb{E}_P\left[\sigma_{\min}^{-1}(C_\ell)\right] \leqslant \frac{\ell r}{r-\ell+1}\kappa^{2\ell-2}, \tag{D14}$$

then we could reach the conclusion of this lemma:

$$\mathbb{E}_P[\sigma_{\min}(C_\ell)] \geqslant \frac{1}{\mathbb{E}_P\left[\sigma_{\min}^{-1}(C_\ell)\right]} \tag{D15}$$

$$\geqslant \frac{r-\ell+1}{\ell r}\kappa^{2-2\ell}. \tag{D16}$$

To prove Eq. (D14), we first rewrite it as follows:

$$\mathbb{E}_P\left[\sigma_{\min}^{-1}(C_\ell)\right] = \sum_{s_1=1}^{m}\cdots\sum_{s_\ell=1}^{m} P(s_1, \ldots, s_\ell)\sigma_{\min}^{-1}(C_\ell) \tag{D17}$$

$$= \sum_{s_1=1}^{m}\cdots\sum_{s_\ell=1}^{m}\frac{\|\boldsymbol{s}_1\|^2\ldots\|\boldsymbol{s}_\ell\|^2}{\Sigma^{(1)}\ldots\Sigma^{(\ell)}}|C_\ell|\sigma_{\min}^{-1}(C_\ell). \tag{D18}$$

In Eq. (D18), we rewrite $P(s_1, \ldots, s_\ell)$ with Eq. (D12) and

$$\Pr^{(\ell)}(s_\ell) = \frac{\|\boldsymbol{s}_\ell\|^2\||\boldsymbol{s}_\ell\rangle - \sum_{i=1}^{\ell-1}|\boldsymbol{t}_i\rangle\langle \boldsymbol{t}_i|\boldsymbol{s}_\ell\rangle\|^2}{\Sigma^{(\ell)}} \tag{D19}$$

$$= \frac{\|\boldsymbol{s}_\ell\|^2 Z_\ell^2}{\Sigma^{(\ell)}} \tag{D20}$$

$$= \frac{\|\boldsymbol{s}_\ell\|^2}{\Sigma^{(\ell)}}\frac{|C_\ell|}{|C_{\ell-1}|}, \tag{D21}$$

where, in Eq. (D19), we denote

$$\Sigma^{(\ell)} = \sum_{s_\ell=1}^{m} \|\boldsymbol{s}_\ell - \sum_{i=1}^{\ell-1} \boldsymbol{t}_i \boldsymbol{t}_i^T \boldsymbol{s}_\ell\|^2, \tag{D22}$$

Eq. (D20) is derived from $Z_\ell^2 = \||\boldsymbol{s}_\ell\rangle - \sum_{i=1}^{\ell-1} |\boldsymbol{t}_i\rangle\langle \boldsymbol{t}_i|\boldsymbol{s}_\ell\rangle\|^2$, and Eq. (D21) is due to Eq. (A6).

Continuing from Eq. (D18), it holds

$$\mathbb{E}_P\big[\sigma_{\min}^{-1}(C_\ell)\big] \leqslant \sum_{s_1=1}^{m} \cdots \sum_{s_\ell=1}^{m} \frac{\|\boldsymbol{s}_1\|^2 \ldots \|\boldsymbol{s}_\ell\|^2}{\Sigma^{(1)} \ldots \Sigma^{(\ell)}} \sum_{i=1}^{\ell} |C_\ell^{(i)}| \tag{D23}$$

$$= \sum_{i=1}^{\ell} \sum_{s_i=1}^{m} \|\boldsymbol{s}_i\|^2 \sum_{s_j=1, j\neq i}^{m} \frac{\prod_{j=1, j\neq i}^{\ell} \|\boldsymbol{s}_j\|^2}{\Sigma^{(1)} \ldots \Sigma^{(\ell)}} |C_\ell^{(i)}|, \tag{D24}$$

where Eq. (D23) uses

$$\sigma_{\min}^{-1}(C_\ell) = \sigma_{\max}\big(C_\ell^{-1}\big) \leqslant \mathrm{Tr}\big(C_\ell^{-1}\big) = \frac{\sum_{i=1}^{\ell} |C_\ell^{(i)}|}{|C_\ell|},$$

with $C_\ell^{(i)} \in \mathbb{R}^{(\ell-1)\times(\ell-1)}$ being the principal submatrix of C_ℓ by removing the ith row and column, and Eq. (D24) follows by rearranging the sum order.

Next, we will provide a lower bound on the denominator term $\Sigma^{(1)} \cdots \Sigma^{(\ell)}$ in Eq. (D24). Note that for any $1 \leqslant j \leqslant \ell$, $\Sigma^{(j)}$ only depends on the matrix A and indices $(s_1, s_2, \ldots, s_{j-1})$, so it can be viewed as the function of $(s_1, s_2, \ldots, s_{j-1})$ when treating A as the constant matrix, namely,

$$\Sigma^{(j)} := \Sigma^{(j)}(s_1, s_2, \ldots, s_{j-1}) \tag{D25}$$

$$= \|A\|_F^2 P_j, \tag{D26}$$

where Eq. (D26) comes from the definition of P_j in Eq. (6). By employing the lower and upper bounds of P_j in Eqs. (D10) and (D11), we could bound the function $\Sigma^{(j)}$ as

$$\sum_{i=j}^{r} \sigma_i^2(A) \leqslant \Sigma^{(j)} \leqslant \sum_{i=1}^{r-j+1} \sigma_i^2(A), \tag{D27}$$

where $\sigma_1(A) \geqslant \sigma_2(A) \geqslant \cdots \geqslant \sigma_r(A)$ denotes singular values of A, and Eq. (D27) holds for any choice of linearly independent row vectors for $\Sigma^{(j)}$. Then Eq. (D27) yields

$$\Sigma^{(j)}(s_1, \ldots, s_{j-1}) \geqslant \sum_{i=j}^{r} \sigma_i^2(A) \tag{D28}$$

$$\geqslant \sum_{i=j}^{r} \sigma_i^2(A) \frac{\Sigma^{(j)}(s_1, \ldots, s_{i-1}, s_{i+1}, \ldots, s_j)}{\sum_{i=1}^{r-j+1} \sigma_i^2(A)} \tag{D29}$$

$$\geqslant \frac{\sigma_{\min}^2(A)}{\sigma_{\max}^2(A)} \Sigma^{(j)}(s_1, \ldots, s_{i-1}, s_{i+1}, \ldots, s_j) \tag{D30}$$

where Eq. (D29) holds true because of the second inequality in Eq. (D27):

$$\frac{\Sigma^{(j)}(s_1, \ldots, s_{i-1}, s_{i+1}, \ldots, s_j)}{\sum_{i=1}^{r-j+1} \sigma_i^2(A)} \leqslant 1. \tag{D31}$$

Continuing from Eq. (D24) and with the inequality $\Sigma^{(\ell)} \geqslant \sum_{k=\ell}^{r} \sigma_k^2(A)$ in Eq. (D27), we obtain the first inequality below:

$$\mathbb{E}_P\big[\sigma_{\min}^{-1}(C_\ell)\big] \leqslant \sum_{i=1}^{\ell} \sum_{s_i=1}^{m} \|\boldsymbol{s}_i\|^2 \sum_{s_j=1, j\neq i}^{m} \frac{\prod_{j=1, j\neq i}^{\ell} \|\boldsymbol{s}_j\|^2}{\Sigma^{(1)} \ldots \Sigma^{(\ell-1)} \sum_{k=\ell}^{r} \sigma_k^2(A)} |C_\ell^{(i)}| \tag{D32}$$

$$\leqslant \left(\frac{\sigma_{\max}^2(A)}{\sigma_{\min}^2(A)}\right)^{\ell-2} \sum_{i=1}^{\ell} \sum_{s_i=1}^{m} \frac{\|\boldsymbol{s}_i\|^2}{\sum_{k=\ell}^{r} \sigma_k^2(A)} \cdot \sum_{s_j=1, j\neq i}^{m} \frac{\prod_{j=1, j\neq i}^{\ell} \|\boldsymbol{s}_j\|^2}{\Sigma'^{(1)} \ldots \Sigma'^{(\ell-1)}} |C_\ell^{(i)}|, \tag{D33}$$

where in Eq. (D33) we denote

$$\Sigma'^{(j)} = \begin{cases} \Sigma^{(j)}(s_1, s_2, \ldots, s_{j-1}), & \forall j < i+1, \\ \Sigma^{(j)}(s_1, s_2, \ldots, s_{i-1}, s_{i+1}, \ldots, s_j), & \forall j \geqslant i+1. \end{cases} \tag{D34}$$

and employ Eq. (D30). Notice that in Eq. (D33),

$$\sum_{s_j=1, j\neq i}^{m} \frac{\prod_{j=1, j\neq i}^{\ell} \|\boldsymbol{s}_j\|^2}{\Sigma'^{(1)} \ldots \Sigma'^{(\ell-1)}} |C_\ell^{(i)}| = 1, \tag{D35}$$

which can be interpreted as the probability for sampling $(s_1, s_2, \ldots, s_{i-1}, s_{i+1}, \ldots, s_\ell)$ over all choice of indices. Finally, Eq. (D33) further leads to

$$\mathbb{E}_P\big[\sigma_{\min}^{-1}(C_\ell)\big] \leqslant \left(\frac{\sigma_{\max}^2(A)}{\sigma_{\min}^2(A)}\right)^{\ell-2} \sum_{i=1}^{\ell} \sum_{s_i=1}^{m} \frac{\|\boldsymbol{s}_i\|^2}{\sum_{k=\ell}^{r} \sigma_k^2(A)} \tag{D36}$$

$$\leqslant \frac{\ell \|A\|_F^2}{(r-\ell+1)\sigma_{\min}^2(A)} \left(\frac{\sigma_{\max}^2(A)}{\sigma_{\min}^2(A)}\right)^{\ell-2} \tag{D37}$$

$$\leqslant \frac{\ell r}{r-\ell+1} \left(\frac{\sigma_{\max}^2(A)}{\sigma_{\min}^2(A)}\right)^{\ell-1} \tag{D38}$$

$$= \frac{\ell r}{r-\ell+1} \kappa^{2\ell-2}, \tag{D39}$$

where $\sum_{s_i=1}^{m} \|\boldsymbol{s}_i\|^2 = \|A\|_F^2$ and $\sum_{k=\ell}^{r} \sigma_k^2(A) \geqslant (r-\ell+1)\sigma_{\min}^2(A)$ are used to derive Eq. (D37), and $\|A\|_F^2 = \sum_{i=1}^{r} \sigma_i^2(A) \leqslant r\sigma_{\max}^2(A)$ is used to derive Eq. (D38). ■

Instead of the distribution P defined in Eq. (D12), the perturbed distribution $\tilde{P}$ is employed due to noisy gates $\tilde{R}_i, \forall i \in [r]$ in Algorithm I. For simplicity, we denote $\Pi_\ell = \prod_{i=1}^{\ell} R_i$ and $\tilde{\Pi}_\ell = \prod_{i=1}^{\ell} \tilde{R}_i, \forall \ell \in [r]$, and then the sampling distributions could be rewritten as

$$\mathrm{Pr}^{(\ell)}(s_\ell) := P(s_\ell|s_1, \ldots, s_{\ell-1}) = \frac{1}{\Sigma^{(\ell)}} \|\boldsymbol{s}_\ell\|^2 \left\| \frac{\Pi_{\ell-1} + I}{2} |\boldsymbol{s}_\ell\rangle \right\|^2, \tag{D40}$$

$$\tilde{\mathrm{Pr}}^{(\ell)}(s_\ell) := \tilde{P}(s_\ell|s_1, \ldots, s_{\ell-1}) = \frac{1}{\tilde{\Sigma}^{(\ell)}} \|\boldsymbol{s}_\ell\|^2 \left\| \frac{\tilde{\Pi}_{\ell-1} + I}{2} |\boldsymbol{s}_\ell\rangle \right\|^2, \tag{D41}$$

where

$$\Sigma^{(\ell)} = \sum_{s_\ell=1}^{m} \|\boldsymbol{s}_\ell\|^2 \left\| \frac{\Pi_{\ell-1} + I}{2} |\boldsymbol{s}_\ell\rangle \right\|^2 \tag{D42}$$

$$\tilde{\Sigma}^{(\ell)} = \sum_{s_\ell=1}^{m} \|\boldsymbol{s}_\ell\|^2 \left\| \frac{\tilde{\Pi}_{\ell-1} + I}{2} |\boldsymbol{s}_\ell\rangle \right\|^2 \tag{D43}$$

are corresponding normalization factors. Now we prove Lemma 4.

Proof. The main idea is that, if the following statement holds true for any $0 \leqslant j \leqslant \ell - 1$:

$$\begin{aligned} &\mathop{\mathbb{E}}_{\tilde{\Pr}^{(j+1)}} \mathop{\mathbb{E}}_{\Pr^{(j+2)}} \cdots \mathop{\mathbb{E}}_{\Pr^{(\ell)}} [\sigma_{\min}(C_\ell)] \\ &\geqslant \left(1 - \frac{1}{6\ell}\right) \mathop{\mathbb{E}}_{\Pr^{(j+1)}} \cdots \mathop{\mathbb{E}}_{\Pr^{(\ell)}} [\sigma_{\min}(C_\ell)] - \frac{1}{6\ell} \mathbb{E}_P[\sigma_{\min}(C_\ell)], \end{aligned} \tag{D44}$$

then we could provide a lower bound on the expectation of $\sigma_{\min}(C_\ell)$ with the distribution $\tilde{P}$ inductively. Specifically, we could obtain

$$\begin{aligned} \mathbb{E}_{\tilde{P}}[\sigma_{\min}(C_\ell)] &= \mathop{\mathbb{E}}_{\tilde{\Pr}^{(1)}} \mathop{\mathbb{E}}_{\tilde{\Pr}^{(2)}} \cdots \mathop{\mathbb{E}}_{\tilde{\Pr}^{(\ell)}} [\sigma_{\min}(C_\ell)] \\ &\geqslant \mathop{\mathbb{E}}_{\tilde{\Pr}^{(1)}} \cdots \mathop{\mathbb{E}}_{\tilde{\Pr}^{(\ell-1)}} \left(1 - \frac{1}{6\ell}\right) \mathop{\mathbb{E}}_{\Pr^{(\ell)}} [\sigma_{\min}(C_\ell)] \\ &\quad - \frac{1}{6\ell} \mathbb{E}_P[\sigma_{\min}(C_\ell)] \qquad \text{(D45)} \\ &\geqslant \mathop{\mathbb{E}}_{\tilde{\Pr}^{(1)}} \cdots \mathop{\mathbb{E}}_{\tilde{\Pr}^{(\ell-2)}} \left(1 - \frac{1}{6\ell}\right)^2 \mathop{\mathbb{E}}_{\Pr^{(\ell-1)}} \mathop{\mathbb{E}}_{\Pr^{(\ell)}} [\sigma_{\min}(C_\ell)] \\ &\quad - \frac{1}{6\ell} \mathbb{E}_P[\sigma_{\min}(C_\ell)] - \frac{1}{6\ell} \mathbb{E}_P[\sigma_{\min}(C_\ell)] \qquad \text{(D46)} \\ &\quad \vdots \\ &\geqslant \left(1 - \frac{1}{6\ell}\right)^\ell \mathop{\mathbb{E}}_{\Pr^{(1)}} \cdots \mathop{\mathbb{E}}_{\Pr^{(\ell)}} [\sigma_{\min}(C_\ell)] - \frac{\ell}{6\ell} \mathbb{E}_P[\sigma_{\min}(C_\ell)] \\ &\geqslant \frac{2}{3} \mathbb{E}_P[\sigma_{\min}(C_\ell)], \qquad \text{(D47)} \end{aligned}$$

where $\tilde{\Pr}^{(j)}$ is defined in Eq. (D41), Eqs. (D45)–(D47) follow from Eq. (D44), and we employ

$$\left(1 - \frac{1}{6\ell}\right)^\ell \geqslant 1 - \frac{1}{6}$$

to obtain the last inequality.

To prove Eq. (D44), we need a lower bound on the distribution $\tilde{\Pr}^{(j+1)}$, which could be derived as follows:

$$\begin{aligned} \tilde{P}^{(j+1)}(s_{j+1}) &= \frac{\|\boldsymbol{s}_{j+1}\|^2 \langle \boldsymbol{s}_{j+1}| \frac{2I + \tilde{\Pi}_j + \tilde{\Pi}_j^\dagger}{4} |\boldsymbol{s}_{j+1}\rangle}{\sum_{s_{j+1}=1}^m \|\boldsymbol{s}_{j+1}\|^2 \langle \boldsymbol{s}_{j+1}| \frac{2I + \tilde{\Pi}_j + \tilde{\Pi}_j^\dagger}{4} |\boldsymbol{s}_{j+1}\rangle} \qquad \text{(D48)} \\ &\geqslant \frac{\|\boldsymbol{s}_{j+1}\|^2 \left(\langle \boldsymbol{s}_{j+1}| \frac{\Pi_j + I}{2} |\boldsymbol{s}_{j+1}\rangle - \| \frac{\tilde{\Pi}_j - \Pi_j}{2} \| \right)}{\sum_{s_{j+1}=1}^m \|\boldsymbol{s}_{j+1}\|^2 \left(\langle \boldsymbol{s}_{j+1}| \frac{\Pi_j + I}{2} |\boldsymbol{s}_{j+1}\rangle + \| \frac{\tilde{\Pi}_j - \Pi_j}{2} \| \right)} \qquad \text{(D49)} \\ &= \frac{\Sigma^{(j+1)} \Pr^{(j+1)}(s_{j+1}) - \|\boldsymbol{s}_{j+1}\|^2 \| \frac{\tilde{\Pi}_j - \Pi_j}{2} \|}{\Sigma^{(j+1)} + \|A\|_F^2 \| \frac{\tilde{\Pi}_j - \Pi_j}{2} \|} \qquad \text{(D50)} \\ &\geqslant \frac{\Sigma^{(j+1)} \Pr^{(j+1)}(s_{j+1}) - \|\boldsymbol{s}_{j+1}\|^2 \frac{j\epsilon_R}{2}}{\Sigma^{(j+1)} + \|A\|_F^2 \frac{j\epsilon_R}{2}}, \qquad \text{(D51)} \end{aligned}$$

where Eq. (D48) is derived by using Eqs. (D41) and (D43). Equation (D49) is obtained by noticing

$$-\left\| \frac{\tilde{\Pi}_j - \Pi_j}{2} \right\| \leqslant \langle \boldsymbol{s}_{j+1}| \frac{\tilde{\Pi}_j - \Pi_j}{2} |\boldsymbol{s}_{j+1}\rangle \leqslant \left\| \frac{\tilde{\Pi}_j - \Pi_j}{2} \right\|.$$

Equation (D50) is derived by using Eqs. (D40) and (D42). Equation (D51) is derived by noticing

$$\|\tilde{\Pi}_j - \Pi_j\| \leqslant \sum_{i=1}^{j} \|\tilde{R}_i - R_i\| \leqslant j\epsilon_R,$$

where we denote by $\epsilon_R = \frac{1}{3r^5\kappa^{2r}}$ the error bound on each R_i, as provided in the assumption of this lemma. Notice that

$$0 \leqslant \sigma_{\min}(C_\ell) \leqslant \frac{\mathrm{Tr}[C_\ell]}{\ell} \leqslant 1 \tag{D52}$$

holds for any choice of row vectors. For simplicity, in Eq. (D44), we denote

$$X := \mathop{\mathbb{E}}_{\Pr^{(j+2)}} \cdots \mathop{\mathbb{E}}_{\Pr^{(\ell)}} [\sigma_{\min}(C_\ell)] \in [0, 1],$$

and proceed as follows

$$\begin{aligned} \mathop{\mathbb{E}}_{\tilde{\Pr}^{(j+1)}} \mathop{\mathbb{E}}_{\Pr^{(j+2)}} \cdots \mathop{\mathbb{E}}_{\Pr^{(\ell)}} [\sigma_{\min}(C_\ell)] &= \mathop{\mathbb{E}}_{\tilde{\Pr}^{(j+1)}} [X] \\ &= \sum_{s_{j+1}=1}^m \tilde{P}^{(j+1)}(s_{j+1}) \cdot X \qquad \text{(D53)} \\ &\geqslant \sum_{s_{j+1}=1}^m \left(\frac{\Sigma^{(j+1)} \Pr^{(j+1)}(s_{j+1}) - \|\boldsymbol{s}_{j+1}\|^2 \frac{j\epsilon_R}{2}}{\Sigma^{(j+1)} + \|A\|_F^2 \frac{j\epsilon_R}{2}} \right) X, \qquad \text{(D54)} \end{aligned}$$

where the inequality employs Eq. (D51). Using the identity $\sum_{s_{j+1}=1}^m \|\boldsymbol{s}_{j+1}\|^2 = \|A\|_F^2$ and $X \leqslant 1$, Eq. (D54) further yields

$$\begin{aligned} &\geqslant \frac{\Sigma^{(j+1)}}{\Sigma^{(j+1)} + \|A\|_F^2 \frac{j\epsilon_R}{2}} \mathop{\mathbb{E}}_{\Pr^{(j+1)}} [X] - \frac{\|A\|_F^2 \frac{j\epsilon_R}{2}}{\Sigma^{(j+1)} + \|A\|_F^2 \frac{j\epsilon_R}{2}} \qquad \text{(D55)} \\ &= \frac{P_{j+1}}{P_{j+1} + \frac{j\epsilon_R}{2}} \mathop{\mathbb{E}}_{\Pr^{(j+1)}} [X] - \frac{\frac{j\epsilon_R}{2}}{P_{j+1} + \frac{j\epsilon_R}{2}} \qquad \text{(D56)} \\ &\geqslant \frac{P_{j+1}}{P_{j+1} + \frac{j\epsilon_R}{2}} \mathop{\mathbb{E}}_{\Pr^{(j+1)}} [X] - \frac{1}{6\ell} \mathbb{E}_P[\sigma_{\min}(C_\ell)] \qquad \text{(D57)} \\ &\geqslant \left(1 - \frac{1}{6\ell}\right) \mathop{\mathbb{E}}_{\Pr^{(j+1)}} [X] - \frac{1}{6\ell} \mathbb{E}_P[\sigma_{\min}(C_\ell)]. \qquad \text{(D58)} \end{aligned}$$

Equation (D56) is obtained by using $\Sigma^{(j+1)} = \|A\|_F^2 P_{j+1}$ in Eq. (D26). Equation (D57) is derived by noticing that

$$\frac{\frac{j\epsilon_R}{2}}{P_{j+1} + \frac{j\epsilon_R}{2}} \leqslant \frac{\frac{j\kappa^{-2r}}{6r^5}}{\frac{r-j}{r}\kappa^{-2}} \tag{D59}$$

$$= \frac{j}{r(r-j)} \frac{\kappa^{2-2r}}{6r \cdot r^2} \tag{D60}$$

$$\leqslant \frac{1}{6\ell} \mathbb{E}_P[\sigma_{\min}(C_\ell)]. \tag{D61}$$

The first inequality follows from $\epsilon_R = \frac{1}{3r^5\kappa^{2r}}$ and

$$P_{j+1} \geqslant \frac{\sum_{i=j+1}^{r} \sigma_i^2(A)}{\|A\|_F^2} \geqslant \frac{(r-j)\sigma_{\min}^2(A)}{r\sigma_{\max}^2(A)} = \frac{r-j}{r}\kappa^{-2}, \tag{D62}$$

where the first inequality uses Lemma 5. Equation (D61) holds due to the lower bound in Lemma 6.

Finally, Eq. (D58) is derived by using

$$\frac{P_{j+1}}{P_{j+1} + \frac{j\epsilon_R}{2}} \geqslant 1 - \frac{1}{6\ell}\mathbb{E}_P[\sigma_{\min}(C_\ell)] \geqslant 1 - \frac{1}{6\ell}, \tag{D63}$$

which holds due to Eqs. (D61) and (D52). ■

APPENDIX E: PROOF OF THEOREM 3

Proof. We sketch the main idea of the proof first. We could implement Algorithm I for N times to guarantee sampling out one basis which satisfies the conditions

$$\text{cond}^{(\ell)} : \left\{ \sigma_{\min}(C_\ell) \geqslant \frac{1}{2r^2\kappa^{2r-2}} \right\}, \quad \forall \ell \in [r]. \tag{E1}$$

Let T_{QGSP} be the query complexity of oracles U_A and V_A to implement Algorithm I once. Thus, the overall query complexity is

$$T_{\text{basis}} = NT_{\text{QGSP}}. \tag{E2}$$

To begin with, consider the first iteration of Algorithm I. The Gram matrix of the sampled basis has the dimension 1×1 with one element 1. Thus, the condition $\text{cond}^{(1)}$ always holds. We proceed to the general cases inductively. Suppose that a basis with $(\ell - 1)$ rows, which satisfies the condition $\text{cond}^{(\ell-1)}$ in Eq. (E1), has been obtained. Next, we move on to the ℓth iteration of Algorithm I. We accept the newly sampled row as part of the basis, if the condition $\text{cond}^{(\ell)}$ holds, and proceed to the $(\ell + 1)$-th iteration. If the condition is violated, we stop the procedure and repeat Algorithm I from the first iteration. Thus, the conditions in Eq. (E1) would hold during the procedure, with the cost that Algorithm I needs to be run N number of times in order to guarantee one basis obtained with high probability.

Now we analyze the complexity of the procedure in detail. Notice that T_{QGSP} consists of three parts: the cost of oracles U_A and V_A for encoding all rows of the input matrix A, the cost of Hadamard test for calculating coefficients $\{z_\ell\}_{\ell=1}^{r-1}$, and the cost of implementing gates $\{C(R_\ell)\}_{\ell=1}^{r-1}$. Based on Lemmas 2 and 3, the latter two complexities depend on the error in the implementation of R_ℓ. In the following proof, we provide explicit upper bounds of N and T_{QGSP}, by setting

$$\epsilon_C = \frac{1}{960r^{\frac{23}{2}}\kappa^{6r}} \tag{E3}$$

to be the error bound of each element in C_r.

Firstly we demonstrate that the sampling in each iteration of Algorithm I obeys the distribution in Eq. (13), i.e., the error of each gate $C(R_j)$ is bounded as

$$\|\tilde{R}_j - R_j\| \leqslant \epsilon_R = \frac{1}{3r^5\kappa^{2r}}, \quad \forall j \in [r-1]. \tag{E4}$$

Based on Lemma 2, the error of $\boldsymbol{t}_j$ induced by noisy coefficients is bounded by

$$\|\tilde{\boldsymbol{t}}_j - \boldsymbol{t}_j\| \leqslant \frac{8j^{\frac{5}{2}}}{\sigma_{\min}^2(C_j)}\epsilon_C \tag{E5}$$

$$\leqslant 32r^{\frac{13}{2}}\kappa^{4r}\epsilon_C \tag{E6}$$

$$= \frac{1}{30r^5\kappa^{2r}} \tag{E7}$$

$$= \frac{\epsilon_R}{10}, \tag{E8}$$

for $j \in [r-1]$. Equation (E5) follows from Eq. (11). Since the condition $\text{cond}^{(j)}$ (E1) holds, we obtain Eq. (E6). Equation (E7) is derived by using Eq. (E3). Equation (E8) is derived by using Eq. (E4).

Then, based on Lemma 3, we could implement the gate $C(R_j)$ with an error ϵ_R by using

$$T_{R_j} = O\left(j\sigma_{\min}^{-\frac{1}{2}}(C_j)\epsilon_R^{-1}\right) \tag{E9}$$

$$\leqslant O\left(r^2\kappa^r\epsilon_R^{-1}\right) \tag{E10}$$

$$\leqslant O(r^7\kappa^{3r}) \tag{E11}$$

queries to the oracle U_A. Since the condition $\text{cond}^{(j)}$ (E1) holds, we obtain Eq. (E10). Equation (E11) follows from the definition of ϵ_R in Eq. (E4).

Next we calculate N. The number of times to perform Algorithm 1 is bounded as

$$N = O\left(\frac{1}{\Pr\{\text{cond}^{(r)}\}}\right), \tag{E12}$$

where the probability follows from the distribution $\tilde{P}$ defined in Eq. (13). Now we proceed to bound $\Pr\{\text{cond}^{(r)}\}$. In fact, we have

$$(1 - \Pr\{\text{cond}^{(r)}\})\frac{1}{2r^2\kappa^{2r-2}} + \Pr\{\text{cond}^{(r)}\} \cdot 1 \tag{E13}$$

$$\geqslant (1 - \Pr\{\text{cond}^{(r)}\})\frac{1}{2r^2\kappa^{2r-2}} + \Pr\{\text{cond}^{(r)}\}\mathbb{E}_{\tilde{P}}[\sigma_{\min}(C_r)|\text{cond}^{(r)}] \tag{E14}$$

$$\geqslant (1 - \Pr\{\text{cond}^{(r)}\})\mathbb{E}_{\tilde{P}}[\sigma_{\min}(C_r)|\text{not cond}^{(r)}] + \Pr\{\text{cond}^{(r)}\}\mathbb{E}_{\tilde{P}}[\sigma_{\min}(C_r)|\text{cond}^{(r)}] \tag{E15}$$

$$= \mathbb{E}_{\tilde{P}}[\sigma_{\min}(C_r)], \tag{E16}$$

where Eq. (E14) is obtained by noticing

$$\sigma_{\min}(C_r) \leqslant \frac{\text{Tr}(C_r)}{r} = 1$$

holds for all choices of basis. Equation (E15) is derived since $\frac{1}{2r^2\kappa^{2r-2}} \geqslant \sigma_{\min}(C_r)$ when the condition $\text{cond}^{(r)}$ does not hold.

Combining Eq. (E4) with Lemma 4, we have the following statement:

$$\mathbb{E}_{\bar{P}}[\sigma_{\min}(C_r)] \geqslant \frac{2}{3} \mathop{\mathbb{E}}_{\Pr^{(1)}} \cdots \mathop{\mathbb{E}}_{\Pr^{(r)}} [\sigma_{\min}(C_r)] \tag{E17}$$

$$\geqslant \frac{2}{3r^2\kappa^{2r-2}}. \tag{E18}$$

Thus, Eq. (E18) together with Eq. (E13) yields

$$\frac{1-\Pr\{\text{cond}^{(r)}\}}{2r^2\kappa^{2r-2}} + \Pr\{\text{cond}^{(r)}\} \geqslant \frac{2}{3r^2\kappa^{2r-2}}. \tag{E19}$$

We could solve

$$\Pr\{\text{cond}^{(r)}\} \geqslant \frac{\frac{1}{6r^2\kappa^{2r-2}}}{1-\frac{1}{2r^2\kappa^{2r-2}}} \geqslant \frac{1}{6r^2}\kappa^{2-2r}, \tag{E20}$$

which induces the bound $N \leqslant O(r^2\kappa^{2r-2})$ by using Eq. (E12).

Finally we move on to analyze the query complexity T_{QGSP}. Based on Lemma 1, coefficients $\{z_\ell\}_{\ell=1}^{r-1}$ are calculated using the estimation of C_r. Denote by T_C the required query complexity of the oracle U_A to estimate each element in C_r via the Hadamard test. We have

$$T_C = O\big(r^2\epsilon_C^{-2}\big) \tag{E21}$$

$$= O(r^{25}\kappa^{12r}), \tag{E22}$$

where Eq. (E21) is derived by using Eq. (E3). Recall that in each iteration of $\ell = 1, \ldots, r$ in Algorithm I, we perform operations $U_A, V_A, R_1, R_2, \ldots, R_{\ell-1}$ for $1/P_\ell$ times. Taking the complexity of estimating C_r into account, we have

$$T_{\text{QGSP}} = T_C + \sum_{\ell=1}^{r} \frac{1}{P_\ell}\left(2 + \sum_{m=1}^{\ell-1} T_{R_m}\right) \tag{E23}$$

$$= O(r^{25}\kappa^{12r}) + \sum_{\ell=1}^{r} \frac{1}{P_\ell}\left(2 + \sum_{m=1}^{\ell-1} O(r^7\kappa^{3r})\right) \tag{E24}$$

$$\leqslant O(r^{25}\kappa^{12r}) + O(r^8\kappa^{3r}) \sum_{\ell=1}^{r} \frac{1}{P_\ell} \tag{E25}$$

$$\leqslant O(r^{25}\kappa^{12r}) + O(r^8\kappa^{3r}) \sum_{\ell=1}^{r} r\kappa^2 \tag{E26}$$

$$\leqslant O(r^{25}\kappa^{12r}). \tag{E27}$$

Equation (E24) is obtained by using Eqs. (E22) and (E11). Equation (E26) is derived by using Eq. (D62).

By considering $N \leqslant O(r^2\kappa^{2r-2})$ being the required number of times to run Algorithm I, we prove the Theorem 3.

APPENDIX F: PROOF OF THEOREM 4

We will first demonstrate that the proposed quantum circuit in Fig. 2 is similar to the swap test and provides a ϵ-error estimation to the value $a_i' = \langle \boldsymbol{t}_k|\boldsymbol{v}\rangle\langle\boldsymbol{v}|\boldsymbol{t}_i\rangle, \forall i \in [r]$, with $O(1/\epsilon^2)$ measurements.

First, after all unitary operations, the state in Fig. 2 before the measurements is

$$\begin{aligned}&\tfrac{1}{4}|0\rangle[|\boldsymbol{v}\rangle|\boldsymbol{t}_k\rangle + |\boldsymbol{v}\rangle|\boldsymbol{t}_i\rangle + |\boldsymbol{t}_k\rangle|\boldsymbol{v}\rangle + |\boldsymbol{t}_i\rangle|\boldsymbol{v}\rangle]|0\rangle\\ &+\tfrac{1}{4}|0\rangle[|\boldsymbol{v}\rangle|\boldsymbol{t}_k\rangle - |\boldsymbol{v}\rangle|\boldsymbol{t}_i\rangle + |\boldsymbol{t}_k\rangle|\boldsymbol{v}\rangle - |\boldsymbol{t}_i\rangle|\boldsymbol{v}\rangle]|1\rangle\\ &+\tfrac{1}{4}|1\rangle[|\boldsymbol{v}\rangle|\boldsymbol{t}_k\rangle + |\boldsymbol{v}\rangle|\boldsymbol{t}_i\rangle - |\boldsymbol{t}_k\rangle|\boldsymbol{v}\rangle - |\boldsymbol{t}_i\rangle|\boldsymbol{v}\rangle]|0\rangle\\ &+\tfrac{1}{4}|1\rangle[|\boldsymbol{v}\rangle|\boldsymbol{t}_k\rangle - |\boldsymbol{v}\rangle|\boldsymbol{t}_i\rangle - |\boldsymbol{t}_k\rangle|\boldsymbol{v}\rangle + |\boldsymbol{t}_i\rangle|\boldsymbol{v}\rangle]|1\rangle.\end{aligned}$$

Measuring the first and the last register could result in outcomes 00 and 11 with probability

$$P_{00} = \frac{2 + |\langle\boldsymbol{v}|\boldsymbol{t}_k\rangle|^2 + |\langle\boldsymbol{v}|\boldsymbol{t}_i\rangle|^2 + 2\langle\boldsymbol{t}_i|\boldsymbol{v}\rangle\langle\boldsymbol{v}|\boldsymbol{t}_k\rangle}{8},$$

$$P_{11} = \frac{2 - |\langle\boldsymbol{v}|\boldsymbol{t}_k\rangle|^2 - |\langle\boldsymbol{v}|\boldsymbol{t}_i\rangle|^2 + 2\langle\boldsymbol{t}_i|\boldsymbol{v}\rangle\langle\boldsymbol{v}|\boldsymbol{t}_k\rangle}{8}.$$

We remark that the statistics of outcomes 00 and 11 implies the value $a_i a_k$:

$$P_{\text{same}} = P_{00} + P_{11} = \frac{1 + \langle\boldsymbol{t}_i|\boldsymbol{v}\rangle\langle\boldsymbol{v}|\boldsymbol{t}_k\rangle}{2} = \frac{1 + a_i a_k}{2}.$$

The efficiency of the quantum circuit in Fig. 2 depends on the efficiency of preparing the input state $(|\boldsymbol{t}_k\rangle|0\rangle + |\boldsymbol{t}_i\rangle|1\rangle)/\sqrt{2}$. Lemma 7 below proves that it can be prepared with query complexity $O(r\sigma_{\min}^{-1/2}(C_r))$.

Lemma 7. Given perturbed coefficients provided in Lemma 2 for both indices k and ℓ, the state $\frac{1}{\sqrt{2}}(|0\rangle|\boldsymbol{t}_k\rangle + |1\rangle|\boldsymbol{t}_\ell\rangle)$ could be prepared with query complexity $O(r\sigma_{\min}^{-1/2}(C_r))$ with ℓ^2 norm error bounded by ϵ.

Proof. We sketch the main idea of the proof. First, we generate the superposition state of $|\tilde{\boldsymbol{t}}_\ell\rangle$ and $|\tilde{\boldsymbol{t}}_k\rangle$ using perturbed coefficients, where perturbed vectors are expressed as

$$\tilde{\boldsymbol{t}}_\ell = \sum_{i=1}^{\ell} \tilde{z}_{i\ell}\boldsymbol{s}_i/\|\boldsymbol{s}_i\|, \quad \tilde{\boldsymbol{t}}_k = \sum_{i=1}^{k} \tilde{z}_{ik}\boldsymbol{s}_i/\|\boldsymbol{s}_i\|. \tag{F1}$$

Then, we provide the error analysis. Specifically, given the coefficients $\tilde{z}_k$ and $\tilde{z}_\ell$, we prepare the state

$$\frac{1}{\sqrt{\|\tilde{\boldsymbol{t}}_\ell\|^2 + \|\tilde{\boldsymbol{t}}_k\|^2}}(\|\tilde{\boldsymbol{t}}_k\||0\rangle|\tilde{\boldsymbol{t}}_k\rangle + \|\tilde{\boldsymbol{t}}_\ell\||1\rangle|\tilde{\boldsymbol{t}}_\ell\rangle) \tag{F2}$$

by the LCU method as follows. Since the notations k and ℓ are symmetrical here, we could assume that $\ell \geqslant k$ for convenience. First, we initialize the state $\frac{|0\rangle+|1\rangle}{\sqrt{2}}|0\rangle^{\otimes \log m}|0\rangle^{\otimes \log n}|0\rangle$. Then, we apply Hadamard operations on the last $\log \ell$ qubits in the second register to create the state:

$$\frac{|0\rangle + |1\rangle}{\sqrt{2\ell}} \sum_{j=1}^{\ell} |j\rangle|0\rangle|0\rangle.$$

Next, we employ the operation U_{index} defined in (C12) to create the state

$$\frac{|0\rangle + |1\rangle}{\sqrt{2\ell}} \sum_{j=1}^{\ell} |g(j)\rangle|0\rangle|0\rangle.$$

Then we employ the oracle U_A on the first and the second registers, followed by the unitary $U_{\text{index}}^\dagger$ to yield

$$\frac{|0\rangle + |1\rangle}{\sqrt{2l}} \sum_{j=1}^{\ell} |j\rangle|A_{g(j)}\rangle|0\rangle \equiv \frac{|0\rangle + |1\rangle}{\sqrt{2l}} \sum_{j=1}^{\ell} |j\rangle|\boldsymbol{s}_j\rangle|0\rangle.$$

Denote $\tilde{z} \equiv \max(\max_j |\tilde{z}_{j\ell}|, \max_j |\tilde{z}_{jk}|)$. Next, we perform the controlled rotation

$$\begin{aligned}
&|0\rangle\langle 0| \otimes \sum_{j=1}^{k} |j\rangle\langle j| \otimes I \otimes e^{-i\sigma_y \arccos(\tilde{z}_{jk}/\tilde{z})} \\
&\quad + |0\rangle\langle 0| \otimes \sum_{j=k+1}^{\ell} |j\rangle\langle j| \otimes I \otimes \sigma_x \\
&\quad + |0\rangle\langle 0| \otimes \sum_{j=\ell+1}^{m} |j\rangle\langle j| \otimes I \otimes I \\
&\quad + |1\rangle\langle 1| \otimes \sum_{j=1}^{\ell} |j\rangle\langle j| \otimes I \otimes e^{-i\sigma_y \arccos(\tilde{z}_{j\ell}/\tilde{z})} \\
&\quad + |1\rangle\langle 1| \otimes \sum_{j=\ell+1}^{m} |j\rangle\langle j| \otimes I \otimes I, \qquad \text{(F3)}
\end{aligned}$$

to obtain the state

$$\begin{aligned}
&\frac{1}{\sqrt{2\ell}}|0\rangle \sum_{j=1}^{k} |j\rangle|\boldsymbol{s}_j\rangle \left(\frac{\tilde{z}_{jk}}{\tilde{z}}|0\rangle + \sqrt{1-\frac{\tilde{z}_{jk}^2}{\tilde{z}^2}}|1\rangle\right) \\
&\quad + \frac{1}{\sqrt{2\ell}}|0\rangle \sum_{j=k+1}^{\ell} |j\rangle|\boldsymbol{s}_j\rangle|1\rangle \\
&\quad + \frac{1}{\sqrt{2\ell}}|1\rangle \sum_{j=1}^{\ell} |j\rangle|\boldsymbol{s}_j\rangle \left(\frac{\tilde{z}_{j\ell}}{\tilde{z}}|0\rangle + \sqrt{1-\frac{\tilde{z}_{j\ell}^2}{\tilde{z}^2}}|1\rangle\right).
\end{aligned}$$

The unitary (F3) could be performed by using $O(\ell)$ quantum operations due to the $O(\ell)$ sparsity. Finally, we employ Hadamard operations on the last $\log \ell$ qubits in the second register, to obtain the state

$$\begin{aligned}
&\frac{1}{\sqrt{2}\ell\tilde{z}}|0\rangle \sum_{j=1}^{k} |0\rangle\tilde{z}_{jk}|\boldsymbol{s}_j\rangle|0\rangle + \frac{1}{\sqrt{2}\ell\tilde{z}}|1\rangle \sum_{j=1}^{\ell} \tilde{z}_{j\ell}|0\rangle|\boldsymbol{s}_j\rangle|0\rangle \\
&\quad + \text{orthogonal garbage state} \\
&= \frac{1}{\sqrt{2}\ell\tilde{z}}(\|\tilde{\boldsymbol{t}}_k\||0\rangle|0\rangle|\tilde{\boldsymbol{t}}_k\rangle|0\rangle + \|\tilde{\boldsymbol{t}}_\ell\||1\rangle|0\rangle|\tilde{\boldsymbol{t}}_\ell\rangle|0\rangle) \\
&\quad + \text{orthogonal garbage state},
\end{aligned}$$

The measurement on the second and fourth registers of the final state could yield state in (F2) with probability

$$\frac{\|\tilde{\boldsymbol{t}}_\ell\|^2 + \|\tilde{\boldsymbol{t}}_k\|^2}{2\ell^2\tilde{z}^2},$$

so we could prepare this state with

$$O\left(\frac{\ell\tilde{z}}{\sqrt{\|\tilde{\boldsymbol{t}}_\ell\|^2 + \|\tilde{\boldsymbol{t}}_k\|^2}}\right)$$

queries to U_A by using the amplitude amplification method [67]. By using Eq. (C13), the complexity is further upper bounded as $O(r\sigma_{\min}^{-1/2}(C_r))$.

Now we analyze the distance between the state $\frac{1}{\sqrt{2}}(|0\rangle|\boldsymbol{t}_k\rangle + |1\rangle|\boldsymbol{t}_\ell\rangle)$ and the state in (F2) as follows:

$$\begin{aligned}
&\left\| \frac{\|\tilde{\boldsymbol{t}}_k\||0\rangle|\tilde{\boldsymbol{t}}_k\rangle + \|\tilde{\boldsymbol{t}}_\ell\||1\rangle|\tilde{\boldsymbol{t}}_\ell\rangle}{\sqrt{\|\tilde{\boldsymbol{t}}_\ell\|^2 + \|\tilde{\boldsymbol{t}}_k\|^2}} - \frac{|0\rangle|\boldsymbol{t}_k\rangle + |1\rangle|\boldsymbol{t}_\ell\rangle}{\sqrt{2}} \right\| \\
&\leqslant \left\| \frac{\tilde{\boldsymbol{t}}_k}{\sqrt{\|\tilde{\boldsymbol{t}}_\ell\|^2 + \|\tilde{\boldsymbol{t}}_k\|^2}} - \frac{\boldsymbol{t}_k}{\sqrt{2}} \right\| + \left\| \frac{\tilde{\boldsymbol{t}}_\ell}{\sqrt{\|\tilde{\boldsymbol{t}}_\ell\|^2 + \|\tilde{\boldsymbol{t}}_k\|^2}} - \frac{\boldsymbol{t}_\ell}{\sqrt{2}} \right\| \qquad \text{(F4)} \\
&\leqslant \left\| \frac{\tilde{\boldsymbol{t}}_k}{\sqrt{\|\tilde{\boldsymbol{t}}_\ell\|^2 + \|\tilde{\boldsymbol{t}}_k\|^2}} - \frac{\tilde{\boldsymbol{t}}_k}{\sqrt{2}} \right\| + \left\| \frac{\tilde{\boldsymbol{t}}_k - \boldsymbol{t}_k}{\sqrt{2}} \right\| \\
&\quad + \left\| \frac{\tilde{\boldsymbol{t}}_\ell}{\sqrt{\|\tilde{\boldsymbol{t}}_\ell\|^2 + \|\tilde{\boldsymbol{t}}_k\|^2}} - \frac{\tilde{\boldsymbol{t}}_\ell}{\sqrt{2}} \right\| + \left\| \frac{\tilde{\boldsymbol{t}}_\ell - \boldsymbol{t}_\ell}{\sqrt{2}} \right\| \qquad \text{(F5)} \\
&\leqslant \left[\left(1+\frac{\epsilon}{4}\right) \left| \frac{1}{\sqrt{2\left(1-\frac{\epsilon}{4}\right)^2}} - \frac{1}{\sqrt{2}} \right| + \frac{\epsilon}{4\sqrt{2}} \right] \times 2 \qquad \text{(F6)} \\
&\leqslant \epsilon,
\end{aligned}$$

where Eq. (F6) is derived by using

$$\begin{aligned}
1 - \frac{\epsilon}{4} &\leqslant \|\boldsymbol{t}_i\| - \|\boldsymbol{t}_i - \tilde{\boldsymbol{t}}_i\| \leqslant \|\tilde{\boldsymbol{t}}_i\| \leqslant \|\boldsymbol{t}_i\| + \|\boldsymbol{t}_i - \tilde{\boldsymbol{t}}_i\| \\
&\leqslant 1 + \frac{\epsilon}{4},
\end{aligned}$$

for $i = k/\ell$. ■

Now we begin the proof of Theorem 4 that provides the error analysis of Algorithm II for reading out the state $|\boldsymbol{v}\rangle$.

Proof. We first study the error in the readout procedure and then provide the time analysis. Specifically, notice that the state $\frac{1}{\sqrt{2}}(|0\rangle|\boldsymbol{t}_k\rangle + |1\rangle|\boldsymbol{t}_i\rangle)$ generated by Lemma 7 is perturbed due to the noisy coefficients z_k and z_i. Thus, the readout error consists of two parts: the error on generating $\frac{1}{\sqrt{2}}(|0\rangle|\boldsymbol{t}_k\rangle + |1\rangle|\boldsymbol{t}_i\rangle)$ and the error induced by the statistical noise during the measurement in the Fig 2.

First, we analyze the measurement distribution of Fig. 2 which uses the perturbed input state $\frac{1}{\sqrt{2}}(|0\rangle|\boldsymbol{t}_k\rangle + |1\rangle|\boldsymbol{t}_i\rangle)$. Denote $\tilde{\boldsymbol{z}}_j$ and $\tilde{\boldsymbol{t}}_j$ as the perturbed form of $\boldsymbol{z}_j$ and $\boldsymbol{t}_j$, respectively, $\forall j \in [r]$. In this proof, we assume the ℓ^2 norm on the error of each $\boldsymbol{t}_j$ is bounded by $\epsilon_3 = \frac{1}{14r^{3/2}}\epsilon$. The final state in Fig. 2 is

$$\begin{aligned}
&\tfrac{1}{4}|0\rangle[f_k|\boldsymbol{v}\rangle|\tilde{\boldsymbol{t}}_k\rangle + f_i|\boldsymbol{v}\rangle|\tilde{\boldsymbol{t}}_i\rangle + f_k|\tilde{\boldsymbol{t}}_k\rangle|\boldsymbol{v}\rangle + f_i|\tilde{\boldsymbol{t}}_i\rangle|\boldsymbol{v}\rangle]|0\rangle \\
&+\tfrac{1}{4}|0\rangle[f_k|\boldsymbol{v}\rangle|\tilde{\boldsymbol{t}}_k\rangle - f_i|\boldsymbol{v}\rangle|\tilde{\boldsymbol{t}}_i\rangle + f_k|\tilde{\boldsymbol{t}}_k\rangle|\boldsymbol{v}\rangle - f_i|\tilde{\boldsymbol{t}}_i\rangle|\boldsymbol{v}\rangle]|1\rangle \\
&+\tfrac{1}{4}|1\rangle[f_k|\boldsymbol{v}\rangle|\tilde{\boldsymbol{t}}_k\rangle + f_i|\boldsymbol{v}\rangle|\tilde{\boldsymbol{t}}_i\rangle - f_k|\tilde{\boldsymbol{t}}_k\rangle|\boldsymbol{v}\rangle - f_i|\tilde{\boldsymbol{t}}_i\rangle|\boldsymbol{v}\rangle]|0\rangle \\
&+\tfrac{1}{4}|1\rangle[f_k|\boldsymbol{v}\rangle|\tilde{\boldsymbol{t}}_k\rangle - f_i|\boldsymbol{v}\rangle|\tilde{\boldsymbol{t}}_i\rangle - f_k|\tilde{\boldsymbol{t}}_k\rangle|\boldsymbol{v}\rangle + f_i|\tilde{\boldsymbol{t}}_i\rangle|\boldsymbol{v}\rangle]|1\rangle,
\end{aligned}$$

where we denote

$$f_k = \sqrt{\frac{2\|\tilde{\boldsymbol{t}}_k\|^2}{\|\tilde{\boldsymbol{t}}_k\|^2 + \|\tilde{\boldsymbol{t}}_i\|^2}}, \quad f_i = \sqrt{\frac{2\|\tilde{\boldsymbol{t}}_i\|^2}{\|\tilde{\boldsymbol{t}}_k\|^2 + \|\tilde{\boldsymbol{t}}_i\|^2}}.$$

Measuring the first and the last registers could result in outcomes 00 and 11 with probability:

$$\begin{aligned}
\tilde{P}_{00} &= \frac{1 + f_i f_k \langle\tilde{\boldsymbol{t}}_i|\boldsymbol{v}\rangle\langle\boldsymbol{v}|\tilde{\boldsymbol{t}}_k\rangle}{4} \\
&\quad + \frac{f_k^2|\langle\boldsymbol{v}|\tilde{\boldsymbol{t}}_k\rangle|^2 + f_i^2|\langle\boldsymbol{v}|\tilde{\boldsymbol{t}}_i\rangle|^2 + 2f_i f_k\langle\tilde{\boldsymbol{t}}_k|\tilde{\boldsymbol{t}}_i\rangle}{8},
\end{aligned}$$

$$\tilde{P}_{11} = \frac{1 + f_i f_k \langle \tilde{\boldsymbol{t}}_i|\boldsymbol{v}\rangle\langle \boldsymbol{v}|\tilde{\boldsymbol{t}}_k\rangle}{4} - \frac{f_k^2|\langle \boldsymbol{v}|\tilde{\boldsymbol{t}}_k\rangle|^2 + f_i^2|\langle \boldsymbol{v}|\tilde{\boldsymbol{t}}_i\rangle|^2 + 2 f_i f_k \langle \tilde{\boldsymbol{t}}_k|\tilde{\boldsymbol{t}}_i\rangle}{8}.$$

Thus, the perturbed statistics of outcomes 00 and 11 is

$$\tilde{P}_{\text{same}} = \tilde{P}_{00} + \tilde{P}_{11} = \frac{1 + f_i f_k \langle \tilde{\boldsymbol{t}}_i|\boldsymbol{v}\rangle\langle \boldsymbol{v}|\tilde{\boldsymbol{t}}_k\rangle}{2} = \frac{1 + \tilde{a}_i \tilde{a}_k}{2}, \quad \text{(F7)}$$

where we denote $\tilde{a}_i = f_i \langle \tilde{\boldsymbol{t}}_i|\boldsymbol{v}\rangle$.

Next, we analyze the error induced by the statistical noise. Notice that each $\tilde{a}_i\tilde{a}_k$ in Eq. (F7) is estimated via the swap test. We assume the statistical error of each value in $\{\tilde{a}_i\tilde{a}_k\}_{i=1}^r$ is bounded by $\epsilon_2 = \frac{1}{14r^{3/2}}\epsilon$, and denote $\widetilde{(\tilde{a}_i\tilde{a}_k)}$ as the approximated value of $\tilde{a}_i\tilde{a}_k$. Then, in parallel to the exact form

$$\boldsymbol{v} = \sum_{i=1}^r a_i \boldsymbol{t}_i, \quad \text{(F8)}$$

we use the expression

$$\tilde{\boldsymbol{v}} = \sum_{i=1}^r \tilde{\tilde{a}}_i \tilde{\boldsymbol{t}}_i \quad \text{(F9)}$$

as the perturbed description of the vector $\boldsymbol{v}$, where

$$\tilde{\tilde{a}}_i = \frac{\widetilde{(\tilde{a}_i\tilde{a}_k)}}{\sqrt{\sum_{i=1}^r \widetilde{(\tilde{a}_i\tilde{a}_k)}^2}}, \quad \forall i \in [r]. \quad \text{(F10)}$$

Thus, the ℓ^2 norm of the error on the vector description of the readout state could be bounded as follows:

$$\|\tilde{\boldsymbol{v}} - \boldsymbol{v}\| = \left\| \sum_{i=1}^r (\tilde{\tilde{a}}_i \tilde{\boldsymbol{t}}_i - a_i \boldsymbol{t}_i) \right\| \quad \text{(F11)}$$

$$\leqslant \left\| \sum_{i=1}^r \tilde{\tilde{a}}_i (\tilde{\boldsymbol{t}}_i - \boldsymbol{t}_i) \right\| + \left\| \sum_{i=1}^r (\tilde{\tilde{a}}_i - \tilde{a}_i)\boldsymbol{t}_i \right\| + \left\| \sum_{i=1}^r (\tilde{a}_i - a_i)\boldsymbol{t}_i \right\| \quad \text{(F12)}$$

$$\leqslant \sum_{i=1}^r |\tilde{\tilde{a}}_i|\epsilon_3 + \sqrt{\sum_{i=1}^r (\tilde{\tilde{a}}_i - \tilde{a}_i)^2} + \sqrt{\sum_{i=1}^r (\tilde{a}_i - a_i)^2} \quad \text{(F13)}$$

$$\leqslant \sqrt{r}\epsilon_3 + \sqrt{\sum_{i=1}^r (\tilde{\tilde{a}}_i - \tilde{a}_i)^2} + \sqrt{\sum_{i=1}^r (\tilde{a}_i - a_i)^2} \quad \text{(F14)}$$

$$\leqslant \frac{\epsilon}{14} + \frac{4\epsilon}{7} + \frac{2\epsilon}{7} \leqslant \epsilon, \quad \text{(F15)}$$

where Eq. (F11) is obtained by using Eqs. (F8) and (F9). Equation (F12) follows from the triangular inequality. Equation (F13) holds due to $\|\tilde{\boldsymbol{t}}_i - \boldsymbol{t}_i\| \leqslant \epsilon_3$ and $\boldsymbol{t}_i^T \boldsymbol{t}_j = \delta_{ij}$. Equation (F14) is derived by using the definition in Eq. (F10) and

$$\frac{\sum_{i=1}^r |\tilde{\tilde{a}}_i|}{r} \leqslant \sqrt{\frac{\sum_{i=1}^r |\tilde{\tilde{a}}_i|^2}{r}} = \frac{1}{\sqrt{r}}.$$

Since the term $\epsilon_3 = \frac{1}{14r^{3/2}}\epsilon$ is provided, we notice that Eq. (F15) holds if the following statements is true for any $i \in [r]$:

$$|\tilde{a}_i - a_i| \leqslant \frac{2\epsilon}{7r^{1/2}}, \quad \text{(F16)}$$

$$|\tilde{\tilde{a}}_i - \tilde{a}_i| \leqslant \frac{4\epsilon}{7r^{1/2}}. \quad \text{(F17)}$$

So we just need to bound terms $|\tilde{a}_i - a_i|$ and $|\tilde{\tilde{a}}_i - \tilde{a}_i|$ for deriving the upper bound on $\|\tilde{\boldsymbol{v}} - \boldsymbol{v}\|$, which can be obtained in Eqs. (F18)–(F23) and Eqs. (F24)–(F30), respectively, as follows:

$$|\tilde{a}_i - a_i| = |f_i \langle \tilde{\boldsymbol{t}}_i|\boldsymbol{v}\rangle - \langle \boldsymbol{t}_i|\boldsymbol{v}\rangle| \quad \text{(F18)}$$

$$\leqslant \left\| \sqrt{\frac{2}{\|\tilde{\boldsymbol{t}}_k\|^2 + \|\tilde{\boldsymbol{t}}_i\|^2}}\tilde{\boldsymbol{t}}_i - \boldsymbol{t}_i \right\| \cdot \|\boldsymbol{v}\| \quad \text{(F19)}$$

$$\leqslant \left\| \sqrt{\frac{2}{\|\tilde{\boldsymbol{t}}_k\|^2 + \|\tilde{\boldsymbol{t}}_i\|^2}}(\tilde{\boldsymbol{t}}_i - \boldsymbol{t}_i) \right\| + \left\| \sqrt{\frac{2}{\|\tilde{\boldsymbol{t}}_k\|^2 + \|\tilde{\boldsymbol{t}}_i\|^2}}\boldsymbol{t}_i - \boldsymbol{t}_i \right\| \quad \text{(F20)}$$

$$\leqslant \sqrt{\frac{2}{\|\tilde{\boldsymbol{t}}_k\|^2 + \|\tilde{\boldsymbol{t}}_i\|^2}}\epsilon_3 + \left| \sqrt{\frac{2}{\|\tilde{\boldsymbol{t}}_k\|^2 + \|\tilde{\boldsymbol{t}}_i\|^2}} - 1 \right| \quad \text{(F21)}$$

$$\leqslant \sqrt{\frac{2}{2(1-\epsilon_3)^2}}\epsilon_3 + \left| \sqrt{\frac{2}{2(1-\epsilon_3)^2}} - 1 \right| \quad \text{(F22)}$$

$$= \frac{2\epsilon_3}{1-\epsilon_3} \leqslant \frac{2\epsilon}{13r^{3/2}} \leqslant \frac{2\epsilon}{7r^{1/2}}. \quad \text{(F23)}$$

Equation (F18) follows from the definitions

$$\tilde{a}_i = f_i \langle \tilde{\boldsymbol{t}}_i|\boldsymbol{v}\rangle, \quad a_i = \langle \boldsymbol{t}_i|\boldsymbol{v}\rangle.$$

Equation (F19) is obtained by using the definition $f_i = \sqrt{\frac{2\|\tilde{\boldsymbol{t}}_i\|^2}{\|\tilde{\boldsymbol{t}}_k\|^2 + \|\tilde{\boldsymbol{t}}_i\|^2}}$. Equation (F20) follows from the triangular inequality and $\|\boldsymbol{v}\| = 1$. Equations (F21) and (F22) are derived by using $\|\boldsymbol{t}_i\| = 1$ and $\|\tilde{\boldsymbol{t}}_{k/i} - \boldsymbol{t}_{k/i}\| \leqslant \epsilon_3$. Equation (F23) is obtained by the assumption $\epsilon_3 = \frac{\epsilon}{14r^{3/2}}$.

On the other hand, the term $\tilde{\tilde{a}}_i$ is bounded around $\tilde{a}_i$ as follows:

$$|\tilde{\tilde{a}}_i - \tilde{a}_i| = \left| \frac{\widetilde{(\tilde{a}_i\tilde{a}_k)}}{\sqrt{\sum_{i=1}^r \widetilde{(\tilde{a}_i\tilde{a}_k)}^2}} - \tilde{a}_i \right| \quad \text{(F24)}$$

$$\leqslant \left| \frac{|\tilde{a}_i\tilde{a}_k| + \epsilon_2}{|\tilde{a}_k|\|\tilde{\boldsymbol{a}}\| - \sqrt{r}\epsilon_2} - |\tilde{a}_i| \right| \quad \text{(F25)}$$

$$= \frac{|\tilde{a}_i||\tilde{a}_k|(1 - \|\tilde{\boldsymbol{a}}\|) + (\sqrt{r} + |\tilde{a}_i|)\epsilon_2}{|\tilde{a}_k|\|\tilde{\boldsymbol{a}}\| - \sqrt{r}\epsilon_2} \quad \text{(F26)}$$

$$\leqslant \frac{\sqrt{r}\frac{2\epsilon_3}{1-\epsilon_3} + \left(\sqrt{r} + \frac{1+\epsilon_3}{1-\epsilon_3}\right)\epsilon_2}{\left(\frac{1}{\sqrt{r}} - \frac{2\epsilon_3}{1-\epsilon_3}\right)\left(1 - \sqrt{r}\frac{2\epsilon_3}{1-\epsilon_3}\right) - \sqrt{r}\epsilon_2} \quad \text{(F27)}$$

$$\leqslant \frac{(2r\epsilon_3 + 2r\epsilon_2)(1-\epsilon_3)}{(1 - \epsilon_3 - 2\sqrt{r}\epsilon_3)^2 - r\epsilon_2(1-\epsilon_3)^2} \quad \text{(F28)}$$

$$\leqslant \frac{2r\epsilon_3 + 2r\epsilon_2}{1 - 6r\epsilon_3 - r\epsilon_2} \tag{F29}$$

$$= \frac{4\epsilon}{7r^{1/2}}. \tag{F30}$$

Equation (F24) follows from the definition in Eq. (F10). Equation (F25) is derived by using $|\widetilde{(\tilde{a}_i\tilde{a}_k)} - \tilde{a}_i\tilde{a}_k| \leqslant \epsilon_2$. Equation (F27) is derived by using $|\tilde{a}_i\tilde{a}_k| \leqslant 1$ [Eq. (F7)] and

$$|\tilde{a}_i| \leqslant |a_i| + \frac{2\epsilon_3}{1-\epsilon_3} \leqslant \frac{1+\epsilon_3}{1-\epsilon_3},$$

$$|\tilde{a}_k| \geqslant |a_k| - \frac{2\epsilon_3}{1-\epsilon_3} \geqslant \frac{1}{\sqrt{r}} - \frac{2\epsilon_3}{1-\epsilon_3},$$

$$\|\tilde{\boldsymbol{a}}\| \geqslant \|\boldsymbol{a}\| - \|\tilde{\boldsymbol{a}} - \boldsymbol{a}\| \geqslant 1 - \sqrt{r}\frac{2\epsilon_3}{1-\epsilon_3}.$$

Equation (F28) is derived by multiplying $\sqrt{r}(1-\epsilon_3)^2$ on both the numerator and the denominator, and noticing that

$$r\epsilon_2(1-\epsilon_3)^2 + \sqrt{r}\epsilon_2\left(1-\epsilon_3^2\right) \leqslant 2r\epsilon_2(1-\epsilon_3).$$

Equation (F29) is obtained since

$$(1-\epsilon_3) \leqslant 1, \quad (1-\epsilon_3 - 2\sqrt{r}\epsilon_3)^2 \geqslant 1 - 6r\epsilon_3.$$

We further derive Eq. (F30) by inserting

$$\epsilon_2 = \epsilon_3 = \frac{1}{14r^{3/2}}\epsilon.$$

Finally, we analyze the time complexity of the protocol. Notice that the error $\|\tilde{\boldsymbol{t}}_i - \boldsymbol{t}_i\| \leqslant \epsilon_3$ could be achieved for all $i \in [r]$ by using

$$r^2 \cdot O\left(r^5\sigma_{\min}^{-4}(C)\epsilon_3^{-2}\right) = O\left(r^{10}\sigma_{\min}^{-4}(C)\epsilon^{-2}\right)$$

queries to input oracles due to Lemma 2. Besides, the error ϵ_2 induced as the statistical noise during the measurement in Fig. 2 could be achieved by using $\epsilon_2^{-2} = O(r^3\epsilon^{-2})$ copies of states $|\boldsymbol{v}\rangle$ and $\frac{1}{\sqrt{2}}(|0\rangle|\boldsymbol{t}_k\rangle + |1\rangle|\boldsymbol{t}_i\rangle)$ for $i \in [r]$, where the latter state could be prepared by using $O(r\sigma_{\min}^{-1/2}(C_r))$ queries to input oracles. By summing for each $i \in [r]$, we need $O(r^4\epsilon^{-2})$ copies of the state $|\boldsymbol{v}\rangle$ and $O(r^5\sigma_{\min}^{-1/2}(C)\epsilon^{-2})$ queries to input oracles in the measurement stage. By counting the required resources in two stages, we have proved Theorem 4. ■

[1] C. Hempel, C. Maier, J. Romero, J. McClean, T. Monz, H. Shen, P. Jurcevic, B. P. Lanyon, P. Love, R. Babbush, A. Aspuru-Guzik, R. Blatt, and C. F. Roos, Quantum Chemistry Calculations on a Trapped-Ion Quantum Simulator, Phys. Rev. X **8**, 031022 (2018).

[2] S. McArdle, S. Endo, A. Aspuru-Guzik, S. C. Benjamin, and X. Yuan, Quantum computational chemistry, Rev. Mod. Phys. **92**, 015003 (2020).

[3] B. Nachman, D. Provasoli, W. A. de Jong, and C. W. Bauer, Quantum Algorithm for High Energy Physics Simulations, Phys. Rev. Lett. **126**, 062001 (2021).

[4] G. G. Guerreschi and A. Y. Matsuura, QAOA for max-cut requires hundreds of qubits for quantum speed-up, Sci. Rep. **9**, 6903 (2019).

[5] Y. R. Sanders, D. W. Berry, P. C. S. Costa, L. W. Tessler, N. Wiebe, C. Gidney, H. Neven, and R. Babbush, Compilation of fault-tolerant quantum heuristics for combinatorial optimization, PRX Quantum **1**, 020312 (2020).

[6] A. Prakash, Quantum algorithms for linear algebra and machine learning, Ph.D. thesis, University of California, Berkeley, 2014.

[7] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd, Quantum machine learning, Nature (London) **549**, 195 (2017).

[8] V. Havlicek, A. D. Corcoles, K. Temme, A. W. Harrow, A. Kandala, J. M. Chow, and J. M. Gambetta, Supervised learning with quantum-enhanced feature spaces, Nature (London) **567**, 209 (2019).

[9] A. W. Harrow, A. Hassidim, and S. Lloyd, Quantum Algorithm for Linear Systems of Equations, Phys. Rev. Lett. **103**, 150502 (2009).

[10] P. Rebentrost, A. Steffens, I. Marvian, and S. Lloyd, Quantum singular-value decomposition of nonsparse low-rank matrices, Phys. Rev. A **97**, 012327 (2018).

[11] R. D. Somma and Y. Subasi, Complexity of quantum state verification in the quantum linear systems problem, PRX Quantum **2**, 010315 (2021).

[12] P. Rebentrost, M. Mohseni, and S. Lloyd, Quantum Support Vector Machine for Big Data Classification, Phys. Rev. Lett. **113**, 130503 (2014).

[13] S. Lloyd, M. Mohseni, and P. Rebentrost, Quantum algorithms for supervised and unsupervised machine learning, arXiv:1307.0411.

[14] S. Arunachalam and R. de Wolf, Guest column: A survey of quantum learning theory, SIGACT News **48**, 41 (2017).

[15] I. Kerenidis, J. Landman, A. Luongo, and A. Prakash, q-means: A quantum algorithm for unsupervised machine learning, in *Proceedings of the 33rd International Conference on Neural Information Processing Systems*, Vol. 32 (Curran Associates, Inc., New York, 2019).

[16] A. Kapoor, N. Wiebe, and K. Svore, Quantum perceptron models, in *Proceedings of the 30th International Conference on Neural Information Processing Systems*, Vol. 29 (Curran Associates, Inc., New York, 2016).

[17] J. Bausch, Recurrent quantum neural networks, in *Proceedings of the 34th International Conference on Neural Information Processing Systems*, Vol. 33 (Curran Associates, Inc., New York, 2020).

[18] V. Giovannetti, S. Lloyd, and L. Maccone, Quantum Random Access Memory, Phys. Rev. Lett. **100**, 160501 (2008).

[19] A. Gilyén and T. Li, Distributional property testing in a quantum world, in *11th Innovations in Theoretical Computer Science Conference* (Schloss Dagstuhl–Leibniz Center for Informatics, Dagstuhl, Germany, 2020).

[20] L. Wossnig, Z. Zhao, and A. Prakash, Quantum Linear System Algorithm for Dense Matrices, Phys. Rev. Lett. **120**, 050502 (2018).

[21] I. Kerenidis and A. Prakash, Quantum gradient descent for linear systems and least squares, Phys. Rev. A **101**, 022316 (2020).

[22] Z. Li, X. Liu, N. Xu, and J. Du, Experimental Realization of a Quantum Support Vector Machine, Phys. Rev. Lett. **114**, 140504 (2015).

[23] I. Kerenidis, A. Prakash, and D. Szilágyi, Quantum algorithms for second-order cone programming and support vector machines, Quantum **5**, 427 (2021).

[24] J. Allcock, C.-Y. Hsieh, I. Kerenidis, and S. Zhang, Quantum algorithms for feedforward neural networks, ACM Trans. Quantum Comput. **1**, 1 (2020).

[25] I. Kerenidis, J. Landman, and A. Prakash, Quantum algorithms for deep convolutional neural networks, in *International Conference on Learning Representations (ICLR)* (2020), https://openreview.net/forum?id=Hygab1rKDS.

[26] I. Kerenidis, A. Luongo, and A. Prakash, Quantum expectation-maximization for Gaussian mixture models, in *Proceedings of the 37th International Conference on Machine Learning (ICML)*, Vol. 119 (PMLR, 2020), pp. 5187–5197, https://proceedings.mlr.press/v119/kerenidis20a.html.

[27] I. Kerenidis and A. Prakash, Quantum recommendation systems, in *8th Innovations in Theoretical Computer Science Conference* (Schloss Dagstuhl–Leibniz Center for Informatics, Dagstuhl, Germany, 2017).

[28] I. Kerenidis and A. Prakash, A quantum interior point method for LPS and SDPS, ACM Trans. Quantum Comput. **1**, 1 (2020).

[29] Another implementation of the QRAM oracle is proposed in Ref. [27], which requires $O(k\text{polylog}(md))$ quantum operations and physical resources for $m \times d$ matrix with k nonzero elements.

[30] C. T. Hann, C.-L. Zou, Y. Zhang, Y. Chu, R. J. Schoelkopf, S. M. Girvin, and L. Jiang, Hardware-Efficient Quantum Random Access Memory with Hybrid Quantum Acoustic Systems, Phys. Rev. Lett. **123**, 250501 (2019).

[31] C. T. Hann, G. Lee, S. M. Girvin, and L. Jiang, Resilience of quantum random access memory to generic noise, PRX Quantum **2**, 020311 (2021).

[32] S. Aaronson, Read the fine print, Nat. Phys. **11**, 291 (2015).

[33] D. Gross, Y.-K. Liu, S. T. Flammia, S. Becker, and J. Eisert, Quantum State Tomography via Compressed Sensing, Phys. Rev. Lett. **105**, 150401 (2010).

[34] A. Kyrillidis, A. Kalev, D. Park, S. Bhojanapalli, C. Caramanis, and S. Sanghavi, Provable compressed sensing quantum state tomography via non-convex methods, npj Quantum Inf. **4**, 36 (2018).

[35] J. Haah, A. W. Harrow, Z. Ji, X. Wu, and N. Yu, Sample-optimal tomography of quantum states, IEEE Trans. Inf. Theory **63**, 5628 (2017).

[36] R. O'Donnell and J. Wright, Efficient quantum tomography, in *Proceedings of the 48th Annual ACM Symposium on Theory of Computing* (ACM, New York, NY, 2016), pp. 899–912.

[37] M. Cramer, M. B. Plenio, S. T. Flammia, R. Somma, D. Gross, S. D. Bartlett, O. Landon-Cardinal, D. Poulin, and Y.-K. Liu, Efficient quantum state tomography, Nat. Commun. **1**, 149 (2010).

[38] T. Xin, D. Lu, J. Klassen, N. Yu, Z. Ji, J. Chen, X. Ma, G. Long, B. Zeng, and R. Laflamme, Quantum State Tomography via Reduced Density Matrices, Phys. Rev. Lett. **118**, 020401 (2017).

[39] H. Häffner, W. Hänsel, C. F. Roos, J. Benhelm, D. Chek-al-kar, M. Chwalla, T. Körber, U. D. Rapol, M. Riebe, P. O. Schmidt, C. Becher, O. Gühne, W. Dür, and R. Blatt, Scalable multiparticle entanglement of trapped ions, Nature (London) **438**, 643 (2005).

[40] M. Riebe, K. Kim, P. Schindler, T. Monz, P. O. Schmidt, T. K. Korber, W. Hansel, H. Haffner, C. F. Roos, and R. Blatt, Process Tomography of Ion Trap Quantum Gates, Phys. Rev. Lett. **97**, 220407 (2006).

[41] A. I. Lvovsky and M. G. Raymer, Continuous-variable optical quantum-state tomography, Rev. Mod. Phys. **81**, 299 (2009).

[42] R. Gupta, R. Xia, R. D. Levine, and S. Kais, Maximal entropy approach for quantum state tomography, PRX Quantum **2**, 010318 (2021).

[43] X. Bonet-Monroig, R. Babbush, and T. E. O'Brien, Nearly Optimal Measurement Scheduling for Partial Tomography of Quantum States, Phys. Rev. X **10**, 031064 (2020).

[44] N. Bent, H. Qassim, A. A. Tahir, D. Sych, G. Leuchs, L. L. Sánchez-Soto, E. Karimi, and R. W. Boyd, Experimental Realization of Quantum Tomography of Photonic Qudits via Symmetric Informationally Complete Positive Operator-Valued Measures, Phys. Rev. X **5**, 041006 (2015).

[45] G. I. Struchalin, Y. A. Zagorovskii, E. V. Kovlakov, S. S. Straupe, and S. P. Kulik, Experimental estimation of quantum state properties from classical shadows, PRX Quantum **2**, 010307 (2021).

[46] B. Kulis, M. Sustik, and I. Dhillon, Learning low-rank kernel matrices, in *Proceedings of the 23rd International Conference on Machine Learning* (ACM, New York, NY, 2006), pp. 505–512.

[47] Q. Yao, J. T. Kwok, T. Wang, and T.-Y. Liu, Large-scale low-rank matrix learning with nonconvex regularizers, IEEE Trans. Pattern Anal. Machine Intell. **41**, 2628 (2018).

[48] M. Udell and A. Townsend, Why are big data matrices approximately low rank?, SIAM J. Math. Data Sci. **1**, 144 (2019).

[49] D. Wang, H. Zhang, R. Liu, X. Liu, and J. Wang, Unsupervised feature selection through Gram-Schmidt orthogonalization—a word co-occurrence perspective, Neurocomputing **173**, 845 (2016).

[50] Y.-P. Zhao, Z.-Q. Li, P.-P. Xi, D. Liang, L. Sun, and T.-H. Chen, Gram-Schmidt process based incremental extreme learning machine, Neurocomputing **241**, 1 (2017).

[51] N. Bansal, X. Chen, and Z. Wang, Can we gain more from orthogonality regularizations in training deep networks?, in *Advances in Neural Information Processing Systems* (Curran Associates Inc., New York, 2018), pp. 4261–4271.

[52] M. R. Vanner, M. Aspelmeyer, and M. S. Kim, Quantum State Orthogonalization and a Toolset for Quantum Optomechanical Phonon Control, Phys. Rev. Lett. **110**, 010504 (2013).

[53] M. Ježek, M. Mičuda, I. Straka, M. Mikova, M. Dušek, and J. Fiurášek, Orthogonalization of partly unknown quantum states, Phys. Rev. A **89**, 042316 (2014).

[54] A. S. Coelho, L. S. Costanzo, A. Zavatta, C. Hughes, M. S. Kim, and M. Bellini, Universal Continuous-Variable State Orthogonalizer and Qubit Generator, Phys. Rev. Lett. **116**, 110501 (2016).

[55] H. Havlicek and K. Svozil, Dimensional lifting through the generalized Gram-Schmidt process, Entropy **20**, 284 (2018).

[56] A. M. Childs and N. Wiebe, Hamiltonian simulation using linear combinations of unitary operations, Quantum Inf. Comput. **12**, 901 (2012).

[57] S. Lloyd, M. Mohseni, and P. Rebentrost, Quantum principal component analysis, Nat. Phys. **10**, 631 (2014).

[58] D. Aharonov, V. Jones, and Z. Landau, A polynomial quantum algorithm for approximating the Jones polynomial, Algorithmica **55**, 395 (2009).

[59] H. Buhrman, R. Cleve, J. Watrous, and R. de Wolf, Quantum Fingerprinting, Phys. Rev. Lett. **87**, 167902 (2001).

[60] N. Wiebe, D. Braun, and S. Lloyd, Quantum Algorithm for Data Fitting, Phys. Rev. Lett. **109**, 050505 (2012).

[61] A. M. Childs, R. Kothari, and R. D. Somma, Quantum algorithm for systems of linear equations with exponentially improved dependence on precision, SIAM J. Comput. **46**, 1920 (2017).

[62] E. Tang, A quantum-inspired classical algorithm for recommendation systems, in *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing* (ACM, New York, 2019), p. 217–228.

[63] N.-H. Chia, H.-H. Lin, and C. Wang, Quantum-inspired sublinear classical algorithms for solving low-rank linear systems, arXiv:1811.04852.

[64] D. Jethwani, F. Le Gall, and S. K. Singh, Quantum-inspired classical algorithms for singular value transformation, in *45th International Symposium on Mathematical Foundations of Computer Science* (Schloss Dagstuhl–Leibniz Center for Informatics, Dagstuhl, Germany, 2020).

[65] Y. Du, M.-H. Hsieh, T. Liu, and D. Tao, Quantum-inspired algorithm for general minimum conical hull problems, Phys. Rev. Research **2**, 033199 (2020).

[66] S. Godunov, *Guaranteed Accuracy in Numerical Linear Algebra* (Springer Science & Business Media, Berlin, Germany, 1993).

[67] G. Brassard, P. Hoyer, M. Mosca, and A. Tapp, Quantum amplitude amplification and estimation, Contemp. Math. **305**, 53 (2002).