Quantum Gram-Schmidt Processes and Their Application to Efficient State Read-out for Quantum Algorithms

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

(Dated: May 31, 2022)

Many quantum algorithms that claim speed-up over their classical counterparts only generate 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 read-out 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 that 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 \( \ell^2 \) norm with \( \epsilon \) error require \( \text{poly}(r, 1/\epsilon) \) copies of the output state and \( \text{poly}(r, \kappa, 1/\epsilon) \) queries to the input oracles, where \( \kappa \) is the condition number of the input matrix. With our read-out protocol, we completely characterise 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.

I. INTRODUCTION

Quantum algorithms have been popular for decades, due to the potential advantage in varying fields including physical simulations\textsuperscript{13}, combinatorial optimization\textsuperscript{4,5,18}, and linear algebra\textsuperscript{6}. Notably, the latter has induced an independent subfield known as the quantum machine learning (QML)\textsuperscript{7,8}, which involves quantum linear algebra\textsuperscript{9,11}, quantum learning protocols\textsuperscript{12,15}, and quantum neural networks\textsuperscript{16,17}. These quantum algorithms have shown to achieve speed-ups over their classical counterparts.

Despite the claimed quantum speed-up, most QML algorithms suffered from both the input and the read-out problems. Specifically, the input problem tackles the issue of efficient state preparation, namely, encoding the classical data, potentially of tantamount size, into quantum states which can be determined by local reduced density matrices\textsuperscript{1}. For example, Quantum State Tomography (QST), is one of the fundamental algorithms explored\textsuperscript{32}. State Tomography (QST), is one of the fundamental problems in quantum information science. QST has attracted significant interest from both theoretical\textsuperscript{33,38} and experimental\textsuperscript{39,45} perspectives in recent years. The best general tomography method\textsuperscript{36} could reconstruct a \( d \times d \) density matrix \( \rho \) for the unknown state with rank \( r \) by using \( n = O(rd\epsilon^{-2}) \) copies to the state, which implies \( O(de^{-2}) \) copy complexity for the pure state case \( \rho = |v\rangle \langle v| \). 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 read-out in QML is computationally expensive and would offset the gained quantum speedup. Since the required number \( n \) is proven optimal for both cases\textsuperscript{36}, any further improvement on \( n \) could be achieved only by assuming special prior knowledge on state \( \rho \). For example, QST via local measurements provides efficient estimation for states which can be determined by local reduced density matrices\textsuperscript{1}. For example, QST via local measurements provides efficient estimation for states which can be determined by local reduced density matrices.

On the other hand, the read-out problem addresses recovery of classical description from the output quantum state that contains the classical solutions. In order to preserve the quantum advantage of the underlying quantum algorithm, the output state needs to be decoded efficiently. For some quantum algorithms, such as the quantum recommendation system\textsuperscript{27}, the read-out 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 read-out problem of these quantum algorithms could be critical. However, protocols for efficiently decoding the output quantum states into classical vectors remain little explored\textsuperscript{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\textsuperscript{33,38} and experimental\textsuperscript{39,45} perspectives in recent years. The best general tomography method\textsuperscript{36} could reconstruct a \( d \times d \) density matrix \( \rho \) for the unknown state with rank \( r \) by using \( n = O(rd\epsilon^{-2}) \) copies to the state, which implies \( O(de^{-2}) \) copy complexity for the pure state case \( \rho = |v\rangle \langle v| \). 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 read-out in QML is computationally expensive and would offset the gained quantum speedup. Since the required number \( n \) is proven optimal for both cases\textsuperscript{36}, any further improvement on \( n \) could be achieved only by assuming special prior knowledge on state \( \rho \). For example, QST via local measurements provides efficient estimation for states which can be determined by local reduced density matrices.
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 $\sigma_i$ and corresponding singular vectors $|u_i\rangle$ and $|v_i\rangle$ for matrix $A = \sum_i \sigma_i u_i v_i^T$; and (2) the quantum linear system solver for linear system $Ax = b$ whose solution state $|x\rangle \propto A^{-1}b$ lies in the row space of $A$. Most machine learning problems can be reduced to these two categories [32]. Hence, finding efficient read-out 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 read-out 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 $|\psi\rangle$ lies in the row space of $A$. Instead of obtaining coefficients $\{v_i\}$ by measuring the state $|\psi\rangle = \sum_{i=1}^n v_i |i\rangle$ 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 $|\psi\rangle = \sum_{i=1}^r x_i |A_{g(i)}\rangle$, where the mapping $g(i) : [r] \rightarrow [m]$ denotes the indices of rows selected as the basis. Our state read-out protocol requires $O(\text{poly}(r))$ copies of the output states and $O(\text{poly}(r, \kappa^r))$ 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 [46-58]. 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 read-out protocol with quantum SVD or quantum linear system solver yields an end-to-end complexity that takes $O(\text{poly}(r, \kappa^r, \log(md)))$ queries to input oracles.

During the whole read-out 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 $O(\text{poly}(r, \kappa^r))$ 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. Ref. [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. Ref. [54] 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 $O(r^2 \kappa^{14r})$ 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 $\kappa$ is the rank and the condition number of $A$, respectively.

**Main Result.** The main result for our state read-out protocol is as follows.

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

Further discussion about the applications of our main result will be delayed in Section III. Instead, we will move on to formally define the framework of the state read-out protocol.

## II. STATE READ-OUT 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 $\{A_{g(i)}\}_{i=1}^r$ selected from all rows of $A$ so that the output state can be rewritten as $|\psi\rangle = \sum_{i=1}^r x_i |A_{g(i)}\rangle$. 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 rows of linearly independent rows $\{A_{g(i)}\}_{i=1}^r$ 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 1 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 non-singular, 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 $\{ \ket{\ell} \}_{\ell=1}^r$. An index set of the complete basis: $S_t = \{ g(i) \}_{i=1}^m$.

1: Initialize the index set $S_t = \emptyset$.
2: for $\ell = 1$ to $r$ do
3: Run the quantum circuit in Fig. 1. Measure the third register and post-select on result 0. Measure the first register to obtain an index $g(\ell)$. Update the index set $S_t = S_t \cup \{ g(\ell) \}$.
4: end for

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:

$$
\ket{0} \xrightarrow{V_A} \sum_{i=1}^m \frac{\ket{A_i}}{\| A \|_F} \ket{i},
$$

(1)

$$
\ket{i} \xrightarrow{U_A} \ket{i} \sum_{j=1}^d \frac{A_{ij}}{\| A_i \|} \ket{j}, \forall i \in [m],
$$

(2)

where $A_{ij}$, $A_i$, and $\| A \|_F$ denote the $(i, j)$-th element, the $i$-th 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 $\Pr(1)(i) = \| A_i \|_2^2 / \| A \|_F^2$, where $i \in [m]$. Let $\ket{t_1} := \ket{A_{g(1)}}$ be the first basis vector. The remaining basis vectors are generated inductively. Assume a set of orthogonal states $\{ \ket{t_i} \}_{i=1}^r$ has been generated in the previous $l-1$ iterations. To proceed to the $\ell$-th iteration, we perform the quantum circuit illustrated in Fig. 1 which first creates the state

$$
\ket{0} \xrightarrow{V_A} \sum_{j=1}^m \frac{\ket{A_j}}{\| A \|_F} \ket{j},
$$

(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 $R_t$ gates

$$
C(R_t) = R_t \otimes \ket{0} \bra{0} + I \otimes \ket{1} \bra{1},
$$

(4)

where the unitary $R_t = I - 2 \ket{t_i} \bra{t_i}$. Next, another Hadamard gate is applied to the third register, and the quantum state evolves into:

$$
\ket{\phi_2(\ell)} = \frac{1}{\| A \|_F} \sum_{j=1}^m \ket{A_j} \otimes \left( \left( \ket{A_j} - \sum_{i=1}^{\ell-1} \ket{t_i} \bra{t_i} \right) \ket{0} - \sum_{i=1}^{\ell-1} \ket{t_i} \bra{t_i} \right)
$$

(5)

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

$$
P_t = \frac{1}{\| A \|_F^2} \sum_{j=1}^m \| A_j \|^2 \left( \| A_j \|^2 - \sum_{i=1}^{\ell-1} \| t_i \| \| A_j \| \right),
$$

(6)

and the post-selected state (without the third register) is

$$
\ket{\phi_2(\ell)} = \frac{1}{\sqrt{\| A \|_F^2}} \sum_{j=1}^m \| A_j \| \ket{A_j} - \sum_{i=1}^{\ell-1} \| t_i \| \bra{t_i} \ket{A_j},
$$

(7)

We need roughly $1/P_t$ copies of $\ket{\phi_1(\ell)}$ to generate the state $\ket{\phi_2(\ell)}$. Finally, we measure the first register for a new basis index $g(\ell)$ and a new orthogonal state $\ket{t_\ell}$:

$$
\ket{t_\ell} = \frac{1}{Z_\ell} \left( \ket{A_{g(\ell)}} - \sum_{i=1}^{\ell-1} \ket{t_i} \bra{t_i} \ket{A_{g(\ell)}} \right),
$$

(8)

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

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

(9)

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

(10)

where $\pi_{S_t}(A_j)$ denotes the projection of the row $A_j$ on the row space of the submatrix $A(S_t, \cdot) \in \mathbb{R}^{(\ell-1) \times n}$. 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 $S_t$. After $r$ iterations, we could obtain the index set $S_t = \{ g(i) \}_{i=1}^r$ such that $\{ A_{g(i)} \}_{i=1}^r$ forms a linearly independent basis. We remark that orthonormal states $\{ \ket{t_i} \}_{i=1}^r$ are generated from $\{ \ket{A_{g(i)}} \}_{i=1}^r$ by performing Gram-Schmidt orthogonalization. Thus, an orthonormal basis could be also generated after the implementation of Algorithm 1.

The technical difficulty of constructing the circuit in Fig. 1 comes from efficient implementation of the controlled version of reflection $R_t = I - 2 \ket{t_\ell} \bra{t_\ell}$, since we do not have additional quantum memory to store $\{ \ket{t_\ell} \}$ generated during the algorithm. To overcome this problem, we note that the state $\ket{t_\ell}$ lies in span$\{ \ket{A_{g(i)}} \}_{i=1}^r$, so
that \( |t_\ell\rangle = \sum_{i=1}^\ell z_i|A_{g(i)}\rangle \) for some coefficients \( \{z_i\}_i=1^\ell \).
Instead, we could generate \( |t_\ell\rangle \) by the linear combination of unitary (LCU) method \[55\] with post-selections. Let \( C_\ell \) be the Gram matrix of \( \{|A_{g(i)}\rangle\}_i=1^\ell \), and let \( C_{\ell-1} \) be the submatrix of \( C_\ell \) by deleting the last row and column. The following lemma shows that the coefficient vector \( z_\ell = (z_{1\ell}, \ldots, z_{\ell\ell})^T \) has a compact expression that only depends on the Gram matrices. The proof is provided in Appendix A.

**Lemma 1.** The coefficients in \( |t_\ell\rangle = \sum_{i=1}^\ell z_i|A_{g(i)}\rangle \) could be written in the vector form
\[
z_\ell = \sqrt{\frac{C_{\ell-1}}{\ell}}e_\ell,
\]
where \( e_\ell = (0, 0, \cdots, 0, 1)^T \in \mathbb{R}^\ell \) and \( |X| \) denotes the determinant of a matrix \( X \).

We remark that each element in the matrix \( C_\ell \), 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 \( C_\ell \) would influence the accuracy of coefficients \( z_\ell \), and consequently, impacts the whole complexity of the state read-out protocol. Let \( \bar{t}_\ell = \sum_{i=1}^\ell \hat{z}_i \bar{A}_{g(i)}/\|A_{g(i)}\| \) be the perturbed vector of \( t_\ell \), where \( \{\hat{z}_i\}_i=1^\ell \) 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 \( C_\ell \). We have the following Lemma 2 to bound \( \|\bar{t}_\ell - t_\ell\| \), whose proof is given in Appendix B.

**Lemma 2.** If each element in \( \tilde{C}_\ell \) deviates from that in \( C_\ell \) by at most \( \epsilon_C \leq \frac{\sigma_{\min}(C_\ell)}{80\epsilon\sigma_{\max}(C_\ell)} \epsilon_R \), then for any \( \epsilon_R \in (0, 1) \), the \( \ell^2 \) norm of the error between \( t_\ell \) and \( \bar{t}_\ell \) is bounded as
\[
\|\bar{t}_\ell - t_\ell\| \leq \frac{\epsilon_R}{10},
\]
where \( \bar{t}_\ell = \sum_{i=1}^\ell \hat{z}_i \bar{A}_{g(i)}/\|A_{g(i)}\| \).

Lemmas 1 and 2 complete preconditions to generate the state \( |t_\ell\rangle \) through the LCU method. Then, given copies of \( |t_\ell\rangle \langle t_\ell| \), we can implement the controlled version of the gate \( R_\ell = I - 2|t_\ell\rangle\langle t_\ell| = e^{-i\pi S|t_\ell\rangle\langle t_\ell|} \) 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 \( |t_\ell\rangle \) could be prepared using \( O(\ell^2 \sigma_{\min}(C_\ell)) \) queries to the oracle \( U_A \) with the spectral norm error of \( R_\ell \) bounded by \( \epsilon_R \).

The proof is provided in Appendix C. As a natural corollary, the Gram-Schmidt orthonormal basis \( \{|t_\ell\rangle\}_\ell=1^\ell \) could be provided using \( O(\ell^2 \sigma_{\min}(C_\ell)) \) 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 \( C_\ell \), 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). Further analysis, we prove that the expectation of \( \sigma_{\min}(C_\ell) \) with the distribution formed by Eq. (9) is lower bounded as:
\[
\mathbb{E}_{\mathcal{P}_\ell} (\cdots) \geq \frac{r - \ell + 1}{\ell} \kappa^{2-2\ell}.
\]
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 2 is implemented with error bounded by \( \epsilon_R = \frac{1}{3\sigma_{\max}(C_\ell)} \), where \( r \) and \( \kappa \) is the rank and the condition number of \( A \), respectively, we have
\[
\mathbb{E}_{\mathcal{P}_\ell} [\sigma_{\min}(C_\ell)] \geq \frac{2}{3} \mathbb{E}_{\mathcal{P}_\ell} (\cdots) [\sigma_{\min}(C_\ell)],
\]
where the distribution
\[
\mathcal{P}(s_1, \cdots, s_\ell) = \mathcal{P}_\ell (s_1) \cdots \mathcal{P}_\ell (s_\ell)
\]
follows from Eq. (9) using noisy gates \( \mathcal{R}_i \).

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

As a result, we could perform Algorithm 1 for 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 \( O(\ell^{27} \kappa^{14\ell}) \) queries to input oracles \( V_A \) and \( U_A \), we could find a group of linearly independent states \( \{|A_{g(i)}\rangle\}_i=1^\ell \), such that the least singular value of the Gram matrix \( C_\ell \) formed by \( \{|A_{g(i)}\rangle\}_i=1^\ell \) is greater than \( \frac{S}{\ell \kappa^{14\ell}} \), where \( r \) and \( \kappa \) is 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 \( \{s_i\}_i=1^\ell \) for simplicity, the read-out problem reduces to obtaining coordinates \( \{x_i\}_i=1^\ell \) in the description \( |v\rangle = \sum_{i=1}^\ell x_i|s_i\rangle \). The steps are outlined in Algorithm 2.

**Algorithm 2 State Read-out**

**Input:** QRAM oracle \( U_A \). Copies of state |v\rangle. Orthonormal basis \( \{|t_i\rangle\}_i=1^\ell \). The precision parameter \( \epsilon \).

**Output:** Coordinates \( \{x_i\}_i=1^\ell \) in |v\rangle = \sum_{i=1}^\ell x_i|s_i\rangle that guarantees \( \epsilon \) accuracy under \( \ell^2 \) norm.

1. Estimate the value \( a_i^2 = \langle v|t_i\rangle^2 \), for \( i \in [\ell] \) by SWAP Test. Mark \( k := \text{argmax}_{i \in [\ell]} a_i^2 \).
2. Run the circuit in Fig. 2 to estimate \( a_k^2 = \langle v|t_i\rangle \langle t_i|v\rangle \) for \( i \in [\ell] \). Normalize the vector \( a^2 = a_i^2/|a_i^2| \).
3. Output the solution as \( x = Za \), where \( Z \) is given in Eq. (14).
The idea of Algorithm 2 is fairly natural. Since the QGSP algorithm generates orthonormal states \( \{|t_i\rangle\}_{i=1}^r \), we could first calculate the coordinate of state \( |v\rangle \) under the basis \( \{|t_i\rangle\}_{i=1}^r \): \( |v\rangle = \sum_{i=1}^r a_i |t_i\rangle \), and then transfer the orthonormal basis to the row basis \( \{|s_i\rangle\}_{i=1}^r \):

\[
(t_1, \ldots, t_r) = (s_1 \ldots s_r) Z,
\]

where \( Z = \frac{z_{ij}}{x_{ij}} \) is the transformation matrix. The coordinates \( |x_i\rangle \) is given as: \( x = Z a \).

The crucial part of Algorithm 2 is to calculate the coefficient \( a_i = \langle v | t_i \rangle, \forall i \in [r] \). However, the overlap estimation techniques based on the Hadamard Test 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 post-selection to yield the solution state easily. Another choice is the SWAP test \( \text{swap} \) that only requires copies of states. However, directly using the quantum SWAP test could only obtain the estimation to the value \( \langle v | t_i \rangle^2 \), while \( \text{sign}(a_i) \) remains unknown. To overcome this difficulty, we could assume that the state \( |v\rangle \) has the positive overlap with one of the basis, say \( |t_k\rangle \), and take the value

\[
a_i = \text{sign}(\langle t_k | v \rangle \langle v | t_i \rangle) = \frac{\langle t_k | v \rangle \langle v | t_i \rangle}{\langle t_k | v \rangle}
\]

as the state overlap. This assumption is equivalent to adding a global phase \( \pm 1 \) on \( |v\rangle \), and will not affect the extraction of the classical description.

We construct a variant of the SWAP Test, illustrated in Fig. 2 for estimating \( a'_i = \langle t_k | v \rangle \langle v | t_i \rangle \). It is easy to see that the probability of the measurement outcomes ‘00’ and ‘11’ yields the value \( a'_i \):

\[
P_{\text{same}} = P_{00} + P_{11} = \frac{1 + \langle t_k | v \rangle \langle v | t_i \rangle}{2} = 1 + a'_i.
\]

Similar to the SWAP Test, the proposed quantum circuit provides a \( \epsilon \)-error estimation to the value \( \langle t_k | v \rangle \langle v | t_i \rangle \) with \( \tilde{O}(\epsilon^{-2}) \) measurements. Notice that a larger \( |\langle t_k | v \rangle| \) 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 \( k := \arg \max_i |\langle t_k | v \rangle|^2 \) by running the circuit in Fig. 2.

A. Quantum singular value decomposition

We begin with the quantum singular value decomposition protocol. For a given \( r \)-rank input matrix \( A = \sum_{i=1}^r \sigma_i u_i v_i^T \in \mathbb{R}^{m \times d} \), there is:

\[
v_i = \frac{1}{\sigma_i} (u_i^T A)^T = \frac{1}{\sigma_i} \sum_{j=1}^m u_{ij} A_j, \forall j \in [m],
\]

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

\[
|0\rangle |0\rangle \sum_{i=1}^r \sum_{j=1}^m A_{ij} |0\rangle |j\rangle = \sum_{i=1}^r \sigma_i |u_i\rangle |v_i\rangle
\]
as the input to the quantum SVD algorithm to generate the state $\sum_{i=1}^{r} \sigma_i |u_i\rangle |v_i\rangle |\hat{\sigma}_i\rangle$. Then the measurement on the eigenvalue register could collapse the state to different eigenstates $|u_i\rangle |v_i\rangle$ with probability $\frac{\sigma_i^2}{\|A\|^2}$. Thus, any target state $|v_i\rangle$ could be prepared with complexity $O(\text{polylog}(md)\|A\|\|\sigma_i^{-2}\|\Delta_{\sigma}^{-1})$, where $\Delta_{\sigma}$ is the eigen gap 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 $|v_i\rangle$ of $A$ could be obtained by using $O(\kappa^{4\ell}\text{poly}(r, \log(md))\|A\|\|\sigma_i^{-2}\|\Delta_{\sigma}^{-1})$ queries to the input oracle of $A$, such that the $\ell^2$ norm error is bounded in $\epsilon$.

### 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 specifically 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 $A = \sum_{i=1}^{r} \sigma_i u_i v_i^T$, the solution could be written as:

$$x = A^+ b,$$

where $A^+ = \sum_{i=1}^{r} \frac{1}{\sigma_i} v_i u_i^T$ is the pseudo-inverse matrix of $A$. Equation (18) gives $x = \sum_{i=1}^{r} \frac{1}{\sigma_i} u_i^T b v_i \in \text{span}\{|v_i\rangle\}_{i=1}^{r}$, which means $x$ also lies in the row space $\text{span}(|v_i\rangle)_{i=1}^{r}$ by using the previous conclusion about eigenvectors.

For the linear system $Ax = b$, the solution state $|x\rangle = |A^+ b\rangle$ could be prepared in time $O(\kappa^{4\ell}\text{poly}(d)\|A\|\|x\|_2^{-1})$ with $\ell^2$ norm error bounded in $\epsilon$, where $\kappa$ 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 $|A^+ b\rangle$ for the linear system $Ax = b$ could be obtained by using $O(\kappa^{4\ell}\text{poly}(r, \log d)\|A\|\|\sigma_i^{-2}\|\Delta_{\sigma}^{-1})$ queries to input oracles of $A$, such that the $\ell^2$ norm error is bounded in $\epsilon$.

### IV. CONCLUSION AND DISCUSSION

In this work, we developed an efficient state read-out framework for quantum algorithms which involve a low-rank input matrix and the output state $|v\rangle$ lies in the row space of the input matrix. The proposed framework takes $O(\text{poly}(r)\epsilon^{-2})$ copies of the output state and $O(\text{poly}(r, \kappa^\epsilon)\epsilon^{-2})$ queries to input oracles for providing $\epsilon$ error bounded classical description. Thus, our protocol preserves the quantum speed-up at the state read-out step of these quantum algorithms for the case that the rank $r$ and the condition number $\kappa$ 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 read-out protocol. Note that the state read-out 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 eigen-decompositions, and then to decode the eigenstates using our protocol. Another future direction is to improve our read-out framework such that the complexity is polynomial in both the rank and the condition number.

### Appendix A: Proof of Lemma 1

**Proof.** Denote $s_i := A^i(i)$ for the simplicity of notation. Consider the state:

$$|t_\ell\rangle = \sum_{i=1}^{\ell} z_i \langle s_i | = \frac{1}{Z_\ell} (|s_\ell\rangle - \sum_{i=1}^{\ell-1} |t_i\rangle \langle t_i | s_\ell\rangle), \tag{A1}$$

where $Z_\ell$ has another formulation obtained by multiplying $|t_\ell\rangle$ on both sides

$$1 = \langle t_\ell | t_\ell \rangle = \frac{1}{Z_\ell} \langle t_\ell | s_\ell \rangle = \frac{1}{Z_\ell} \sum_{i=1}^{\ell} z_i \langle s_i | s_\ell \rangle. \tag{A2}$$

The restriction that $|t_\ell\rangle$ is normalized and is orthogonal to states $|s_1\rangle, |s_2\rangle, \ldots |s_{\ell-1}\rangle$ could yield:

$$\langle s_j | t_\ell \rangle = \sum_{i=1}^{\ell} z_i \langle s_j | s_i \rangle = 0, \forall j \in [\ell - 1], \tag{A3}$$

$$\langle t_\ell | t_\ell \rangle = \sum_{j=1}^{\ell} \sum_{i=1}^{\ell} z_j z_i \langle s_j | s_i \rangle = 1. \tag{A4}$$

Rewrite Equation (A2) and (A3) in the vector form:

$$C_\ell z_\ell = Z_\ell e_\ell. \tag{A5}$$
Equation (A4) could be written as:

\[
1 = \sum_{i,j=1}^{\ell} z_{i\ell} z_{j\ell}^* \langle s_j | s_i \rangle = z^T \ell C_{\ell} z_{\ell} = \sum_{\ell} |C_{\ell-1}| \frac{C_{\ell-1}}{|C_{\ell}|},
\]

where the third equation derives from \( z_{\ell} = Z_{\ell} C_{\ell-1}^{-1} e_{\ell} \) by Equation (A5) and the last equation is derived by noticing that the \((\ell, \ell)\)-th element of \( C_{\ell-1}^{-1} \) 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

\[
z_{\ell} = Z_{\ell} C_{\ell-1}^{-1} e_{\ell} = \sqrt{\frac{|C_{\ell}|}{|C_{\ell-1}|}} C_{\ell-1}^{-1} e_{\ell}. \tag{A7}
\]

\[\square\]

Appendix B: Proof of Lemma 2

\textbf{Proof.} We denote \( \| \cdot \| \) as the \( \ell^2 \) norm and the spectral norm for vectors and matrices.

First notice that

\[
\begin{align*}
\| t_{\ell} - \hat{t}_{\ell} \|^2 &= t_{\ell}^T t_{\ell} - 2 t_{\ell}^T \hat{t}_{\ell} + \hat{t}_{\ell}^T \hat{t}_{\ell} \tag{B1} \\
&= Z_{\ell}^T C_{\ell} z_{\ell} - 2 Z_{\ell}^T C_{\ell} \hat{z}_{\ell} + \hat{z}_{\ell}^T C_{\ell} \hat{z}_{\ell} \\
&= \Delta z_{\ell}^T C_{\ell} \Delta z_{\ell} \tag{B2} \\
&\leq \| C_{\ell} \| \| \Delta z_{\ell} \|^2, \tag{B3}
\end{align*}
\]

where \( C_{\ell} \) in Eq. (B2) is the Gram matrix of \( \{ |A_{j(i)} \rangle \}_{i=1}^{\ell} \), and \( \Delta z_{\ell} = \hat{z}_{\ell} - z_{\ell} \) in Eq. (B3). Since \( \| C_{\ell} \| \leq \text{Tr} | C_{\ell} | = \ell \), we can obtain the desired result; namely,

\[
\| t_{\ell} - \hat{t}_{\ell} \|^2 \leq \frac{\epsilon_R^2}{100}. \tag{B5}
\]

if the following claim is true:

\[
\| \Delta z_{\ell} \| \leq \frac{\epsilon_R}{10 \ell^{1/2}}. \tag{B6}
\]

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

\[
\hat{z}_{\ell} = \hat{Z}_{\ell} \hat{C}_{\ell-1}^{-1} e_{\ell}. \tag{B7}
\]

Let \( \Delta C_{\ell} = \hat{C}_{\ell} - C_{\ell} \), and \( \Delta Z_{\ell} = \hat{Z}_{\ell} - Z_{\ell} \). Then,

\[
\| \Delta z_{\ell} \| = \| \hat{C}_{\ell-1}^{-1} \hat{Z}_{\ell} e_{\ell} - C_{\ell-1}^{-1} Z_{\ell} e_{\ell} \| \tag{B8}
\]

\[
= \| (C_{\ell} + \Delta C_{\ell})^{-1} (Z_{\ell} e_{\ell} + \Delta Z_{\ell} e_{\ell}) - C_{\ell-1}^{-1} Z_{\ell} e_{\ell} \| \tag{B9}
\]

\[
= \| (C_{\ell} + \Delta C_{\ell})^{-1} (Z_{\ell} e_{\ell} + \Delta Z_{\ell} e_{\ell}) - (C_{\ell} + \Delta C_{\ell}) C_{\ell-1}^{-1} Z_{\ell} e_{\ell} \| \tag{B10}
\]

\[
= \| (C_{\ell} + \Delta C_{\ell})^{-1} (\Delta Z_{\ell} e_{\ell} - \Delta C_{\ell-1} C_{\ell-1}^{-1} Z_{\ell} e_{\ell}) \| \tag{B11}
\]

\[
\leq \| (C_{\ell} + \Delta C_{\ell})^{-1} \| \cdot (\| \Delta Z_{\ell} \| + \| C_{\ell} \| \| C_{\ell-1}^{-1} \| \| Z_{\ell} \|) \tag{B12}
\]

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

Since each element in \( \hat{C}_{\ell} \) deviates from that in \( C_{\ell} \) by at most \( \epsilon_C \leq \frac{\sigma_{\text{min}}^2(C_{\ell})}{80 \ell^{3/2}} \epsilon_R \), we can obtain

\[
\| \Delta C_{\ell} \| \leq \| \Delta C_{\ell} \|_F \leq \sqrt{\ell^2 \epsilon_C^2} = \ell \epsilon_C, \tag{B13}
\]

and

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

\[
\leq \frac{1}{\min \{ C_{\ell} \} - \| \Delta C_{\ell} \|} \tag{B15}
\]

\[
\leq \min \{ C_{\ell} \} - \ell \epsilon_C \tag{B16}
\]

\[
\leq \frac{80}{79} \frac{\sigma_{\text{min}}^{-1}(C_{\ell})}{\epsilon_R}. \tag{B17}
\]

Eq. (B15) follows from the Weyl’s inequality

\[
| \min \{ C_{\ell} + \Delta C_{\ell} \} - \min \{ C_{\ell} \} | \leq \| \Delta C_{\ell} \|. \tag{B18}
\]

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

\[
\ell \epsilon_C \leq \frac{80}{79} \frac{\sigma_{\text{min}}^{-1}(C_{\ell})}{\epsilon_R} \leq \frac{1}{80} \frac{\sigma_{\text{min}}^{-1}(C_{\ell})}{\epsilon_R}. \tag{B19}
\]

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

\[
\| \Delta Z_{\ell} \| \leq 4 \ell^2 \epsilon_C^{-1} (C_{\ell}) \epsilon_C Z_{\ell}. \tag{B20}
\]

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

\[
\| \Delta z_{\ell} \| \leq \frac{80}{79} \frac{\sigma_{\text{min}}^{-2}(C_{\ell})}{\epsilon_R} 5 \ell^2 \epsilon_C Z_{\ell} \tag{B20}
\]

\[
\leq \frac{\epsilon_R}{10 \ell^{1/2}}, \tag{B21}
\]

where \( \ell > 1 \), \( \epsilon_C \leq \frac{\sigma_{\text{min}}(C_{\ell})}{80 \ell^{3/2}} \epsilon_R \) and

\[
Z_{\ell} = \frac{\| s_\ell \|}{\sum_{i=1}^{\ell-1} |t_i\rangle \langle t_i | s_\ell \rangle} \leq 1. \tag{B22}
\]
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 [69, page 113]:

\[ \frac{|C_\ell + \Delta C_\ell|}{|C_\ell|} \leq \frac{\ell \|C_\ell^{-1}\| \|\Delta C_\ell\|}{1 - \ell \|C_\ell^{-1}\| \|\Delta C_\ell\|}. \]  

(B23)

We can obtain

\[ \frac{|C_{\ell-1} + \Delta C_{\ell-1} - |C_{\ell-1}|}{|C_{\ell-1}|} \leq \frac{(\ell - 1) \|C_{\ell-1}^{-1}\| \|\Delta C_{\ell-1}\|}{1 - (\ell - 1) \|C_{\ell-1}^{-1}\| \|\Delta C_{\ell-1}\|} \leq \frac{\ell \|C_{\ell-1}^{-1}\| \|\Delta C_{\ell-1}\|}{1 - \ell \|C_{\ell-1}^{-1}\| \|\Delta C_{\ell-1}\|}, \]  

(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:

\[ (\ell - 1) \|C_{\ell-1}^{-1}\| \|\Delta C_{\ell-1}\| \leq \ell \|C_{\ell-1}^{-1}\| \|\Delta C_{\ell-1}\| = \ell \sigma_{\min}(C_{\ell-1}) \sigma_{\max}(C_{\ell-1}) \leq \ell \sigma_{\min}(C) \sigma_{\max}(\Delta C_\ell) = \ell \|C_\ell^{-1}\| \|\Delta C_\ell\|. \]

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

\[ |\Delta Z_\ell| = \left| \frac{Z_\ell}{Z_\ell} - 1 \right| Z_\ell \]  

(B25)

\[ = \sqrt{\frac{|C_\ell + \Delta C_\ell|}{|C_\ell|} \sqrt{\frac{|C_{\ell-1}|}{|C_{\ell-1}|} - 1} Z_\ell} \]  

(B26)

\[ = \sqrt{\frac{|C_\ell + \Delta C_\ell|}{|C_{\ell-1} + \Delta C_{\ell-1}|} \sqrt{\frac{|C_{\ell-1}|}{|C_{\ell-1} + \Delta C_{\ell-1}|} - 1} Z_\ell} \]  

(B27)

\[ \leq \max \left( \sqrt{1 - \frac{1}{1 - \ell \|C_\ell^{-1}\| \|\Delta C_\ell\|}} \left| \frac{1}{1 - \ell \|C_\ell^{-1}\| \|\Delta C_\ell\|} \right| Z_\ell \right) \]  

(B28)

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

\[ \frac{|C_\ell + \Delta C_\ell|}{|C_\ell|} \geq 1 - \ell \|C_\ell^{-1}\| \|\Delta C_\ell\|, \]  

\[ \frac{|C_\ell + \Delta C_\ell|}{|C_\ell|} \leq \frac{1}{1 - \ell \|C_\ell^{-1}\| \|\Delta C_\ell\|}, \]  

\[ \frac{|C_{\ell-1} + \Delta C_{\ell-1}|}{|C_{\ell-1}|} \leq 1 - \ell \|C_{\ell-1}^{-1}\| \|\Delta C_{\ell-1}\|, \]  

\[ \frac{|C_{\ell-1} + \Delta C_{\ell-1}|}{|C_{\ell-1}|} \geq \frac{|C_{\ell-1}|}{|C_{\ell-1} + \Delta C_{\ell-1}|} \left| \frac{1}{1 - \ell \|C_{\ell-1}^{-1}\| \|\Delta C_{\ell-1}\|} \right| \]  

Since \( \operatorname{max}(A, B) \leq A + B \) for \( A, B \geq 0 \), Eq. (B28) yields

\[ \leq \left( \sqrt{1 - \frac{1}{1 - \ell \|C_\ell^{-1}\| \|\Delta C_\ell\|}} \right) Z_\ell \]  

(B29)

\[ = \frac{2 \ell \|C_\ell^{-1}\| \|\Delta C_\ell\|}{\sqrt{1 - 2 \ell \|C_\ell^{-1}\| \|\Delta C_\ell\|}} Z_\ell \]  

(B30)

\[ \leq \frac{2 \ell^2 \sigma_{\min}(C_\ell) \|\Delta C_\ell\|}{\sqrt{1 - 2 \ell^2 \sigma_{\min}(C_\ell)^2 \|\Delta C_\ell\|}} Z_\ell \]  

(B31)

\[ \leq 4 \ell^2 \sigma_{\min}(C_\ell) \|\Delta C_\ell\| Z_\ell \]  

(B32)

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

\[ \sqrt{1 - 2 \ell^2 \sigma_{\min}^{-1}(C_\ell)^2 \|\Delta C_\ell\|} \geq \sqrt{1 - 4 \ell^2 \sigma_{\min}^{-1}(C_\ell) \frac{\sigma_{\min}^2(C_\ell)}{8005/2}} + \epsilon_R \]  

substitution of which is obtained by using the bound of \( \epsilon_C \) and \( \frac{\sigma_{\min}(C)}{4000/2} \leq \frac{1}{4} \).

\[ \square \]

Appendix C: Proof of Lemma 3

**Proof.** The main idea is to firstly derive the error analysis of \( |t_\ell| \) and \( R_\ell \), followed by the development of the LCU protocol. Denote \( s_i := A_{ij} \) for the simplicity of notation. We begin from the assumption that

\[ \|\tilde{t}_\ell - t_\ell\| \leq \frac{\epsilon_R}{10}, \]  

(C1)

where \( \tilde{t}_\ell = \sum_{i=1}^\ell \tilde{z}_{ii} s_i / \|s_i\| \). Then the \( \ell^2 \) norm of the error of the state \( |t_\ell| \) is bounded as follows.

\[ \|\tilde{t}_\ell - |t_\ell|\| \leq \|\tilde{t}_\ell - |\tilde{t}_\ell|\| + \|\tilde{t}_\ell - t_\ell\| \]  

(C2)

\[ = \|\tilde{t}_\ell - |\tilde{t}_\ell|\| + \|\tilde{t}_\ell - t_\ell\| \]  

(C3)

\[ = \|\tilde{t}_\ell - |t_\ell|\| + \|\tilde{t}_\ell - t_\ell\| \]  

(C4)

\[ \leq \|\tilde{t}_\ell - t_\ell\| + \|\tilde{t}_\ell - t_\ell\| \]  

(C5)

\[ \leq \frac{\epsilon_R}{5}. \]  

(C6)

Eqs. (C2)-(C6) are derived by using \( \|t_\ell\| = 1 \), the triangular inequality, and Eq (C1). We could further provide the spectral norm of the error of the gate \( R_\ell \):

\[ \|\tilde{R}_\ell - R_\ell\| \]  

(C7)

\[ = \|I - 2|\tilde{t}_\ell| |t_\ell| - (I - 2|\tilde{t}_\ell| |t_\ell|)\| \]  

(C8)

\[ \leq 2 \|\tilde{t}_\ell| |t_\ell| - |t_\ell| |t_\ell|\| \]  

(C9)

\[ = 2 \|\tilde{t}_\ell| - |t_\ell|\| + 2 \|\tilde{t}_\ell| - |t_\ell|\| \]  

(C10)
\[ \leq \frac{4}{5} \epsilon_R. \]  

Eq. (C11) is derived due to the definition of \( R_\ell \). Eq. (C9) is derived by using the triangular inequality. Eq. (C11) is derived by using Eq. (C6).

Now we provide a framework to implement operations \( C(R_\ell) \) using coefficients \( \{\hat{z}_\ell\}_j \). We could first prepare the pure state \( \hat{\rho}_\ell = |\ell\rangle \langle \ell| \) 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

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

(C12)

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

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

The unitary \( U_{\text{index}} \) 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 \( U_A^\dagger \), to yield:

\[
\frac{1}{\sqrt{\ell}} \sum_{i=1}^{\ell} |i\rangle |A_{g(i)}\rangle |0\rangle = \frac{1}{\sqrt{\ell}} \sum_{i=1}^{\ell} |i\rangle |s_i\rangle |0\rangle.
\]

Denote \( \hat{z}_{\ell} \equiv \max_j |\hat{z}_{\ell,j}| \). Then we perform the controlled rotation

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

on the third register, conditioned on the first register \( |i\rangle \), to obtain:

\[
\frac{1}{\sqrt{\ell}} \sum_{i=1}^{\ell} |i\rangle |s_i\rangle \left( \frac{\hat{z}_{\ell}}{\hat{z}_{\ell}} |0\rangle + \sqrt{1 - \frac{\hat{z}_{\ell}^2}{\hat{z}_{\ell}^2}} |1\rangle \right).
\]

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

\[
\frac{1}{\ell} \sum_{i=1}^{\ell} |0\rangle \frac{\hat{z}_{\ell}}{\hat{z}_{\ell}} |s_i\rangle |0\rangle + \text{orthogonal garbage state}
\]

\[= \frac{\|\hat{t}_{\ell}\|}{\ell \cdot \hat{z}_{\ell}} |\ell\rangle |0\rangle + \text{orthogonal garbage state}.
\]

The measurement on the first and the third registers of the final state could yield state \( |\ell\rangle \) with success probability \( \|\hat{t}_{\ell}\|^2/\ell^2 \hat{z}_{\ell}^2 \), so we could prepare the state \( |\ell\rangle \) with \( O(\ell \|\hat{t}_{\ell}\|/\|\hat{t}_{\ell}\|) \) queries to \( U_A \) by using the amplitude amplification method \cite{Brassard1997}.

Note that operations \( R_\ell = I - 2 |\ell\rangle \langle \ell| \) can be viewed as the unitary with Hamiltonian \( \hat{\rho}_\ell = |\ell\rangle \langle \ell| \):

\[
e^{-i\pi \hat{\rho}_\ell} = 1 + (-i\pi \hat{\rho}_\ell) + \frac{1}{2!}(-i\pi \hat{\rho}_\ell)^2 + \cdots
\]

\[= 1 - \hat{\rho}_\ell + \hat{\rho}_\ell \left[ 1 + (-i\pi) + \frac{1}{2!}(-i\pi)^2 + \cdots \right]
\]

\[= 1 - \hat{\rho}_\ell + \hat{\rho}_\ell e^{-i\pi}
\]

\[= I - 2 |\ell\rangle \langle \ell|\]

Therefore, by using the Hamiltonian simulation method developed in Quantum PCA \cite{Brassard1997}, the controlled version of \( R_\ell \) could be performed with error \( \epsilon_R/5 \) consuming \( O(5\pi^2/\epsilon_R) = O(1/\epsilon_R) \) copies of \( \hat{\rho}_\ell \). Taking the complexity of generating state \( |\ell\rangle \) into account, we could implement operation \( C(R_\ell) \) with the error of \( \hat{R}(\ell) \) bounded as \( \epsilon_R/5 \), by using \( O(\max_i |z_{\ell,i}|/\|\ell\langle \ell| R\|) \) queries to \( U_A \). We remark that the \( \ell^2 \) norm of vector \( \hat{z}_{\ell} \) is bounded as

\[
1 = \langle \ell | \langle \ell | = \frac{\hat{z}_{\ell}^T C_{\ell} \hat{z}_{\ell}}{\|\hat{z}_{\ell}\|^2} \geq \frac{\|\hat{z}_{\ell}\|^2}{\|\hat{z}_{\ell}\|^2} \sigma_{\text{min}}(C_{\ell}),
\]

which yields:

\[
\frac{\max_i |z_{\ell,i}|}{\|\hat{z}_{\ell}\|} \leq \|\hat{z}_{\ell}\| \leq \sigma_{\text{min}}^{-1/2}(C_{\ell}).
\]

So the query complexity for implementing \( C(R_\ell) \) could be bounded as \( O(\ell \sigma_{\text{min}}^{-1/2}(C_{\ell}) \epsilon_R^{-1}) \). By considering the distance between \( R_\ell \) and \( R_\ell \) in Eq. (C11), we could then implement the controlled version of the gate \( R_\ell \) with error bounded by \( \epsilon_R \). Now we have proved Lemma 4.

\[\square\]

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 Lemma 5 and Lemma 6.

Lemma 5. The probability \( P_\ell \) defined in Eq. (D) is bounded by

\[
\sum_{i=1}^{\ell} \sigma_i^2 \leq P_\ell \leq \frac{\sum_{i=1}^{\ell+1} \sigma_i^2}{\|A\|^2},
\]

where \( \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r \) are singular values of \( A \).

Proof. Denote the singular value decomposition

\[
A = \sum_{i=1}^{r} \sigma_i u_i v_i^T.
\]

Since the state \( |\ell\rangle \) is the linear sum of rows \( \{A_j\}_{j=1}^{m} \), while each row is the linear sum of singular vectors:

\[
A_j = \sum_{i=1}^{r} \sigma_i u_{j,i} v_i,
\]

(D1)
we can further write:
\[ |t_i| = \sum_{j=1}^r w_{ij} |v_j|. \] (D2)

Rewrite Eq. (6) as:
\[ P_\ell = \frac{1}{\|A\|_F^2} \sum_{j=1}^m \left[ \|A_j\|_F^2 - \sum_{i=1}^{\ell-1} \|A_j A_j^T (t_i A_j)\|_F^2 \right] \]
\[ = 1 - \frac{1}{\|A\|_F^2} \sum_{j=1}^m \left[ \sum_{k=1}^r \sum_{\ell \neq k} w_{ik} \sigma_k \sigma_k^\ell (u_k^\ell) \right]^2, \] (D4)

where Eq. (D4) comes from Eq. (D1) and Eq. (D2). Expand the square term in Eq. (D4) yields:
\[ P_\ell = 1 - \frac{1}{\|A\|_F^2} \sum_{j=1}^m \left[ \sum_{k=1}^r \sum_{\ell \neq k} w_{ik} w_{ik} \sigma_k \sigma_k^\ell (u_k^\ell) \right] \]
\[ = 1 - \frac{1}{\|A\|_F^2} \sum_{j=1}^m \sum_{k=1}^r w_{ik} w_{ik} \sigma_k \sigma_k^\ell (u_k^\ell) \]
\[ = 1 - \frac{1}{\|A\|_F^2} \sum_{k=1}^r c_k \sigma_k \] (D7)

where Eq. (D6) follows because \( \sum_{j=1}^m w_{ik} w_{ik} = w_{ik} w_{ik} = \delta_{kk'}, \) and we denote \( c_k = \sum_{i=1}^r w_{ik} \) in Eq. (D7).

Define the \( r \)-dimensional vector \( u_i = \sum_{k=1}^r w_{ik} e_k. \) Since \( (t_i, t_j) = \delta_{ij} = \sum_{k=1}^r w_{ik} w_{jk} = w_{ij} w_{ij}, \) vectors in set \( \{w_i\}_{i=1}^r \) are orthogonal with each other. We can add \( w_1, \ldots, w_r \) such that \( \{w_i\}_{i=1}^r \) forms an orthonormal basis in the \( r \)-dimensional space. Denote the matrix \( W = (w_1, w_2, \ldots, w_r). \) Since \( W^T W = I, \) we have:
\[ 0 \leq c_k = \sum_{i=1}^{\ell-1} w_{ik} \leq \sum_{i=1}^r w_{ik} = [WW^T]_{kk} = 1, \forall k \in [r]. \] (D8)

Note that
\[ \sum_{k=1}^r c_k = \sum_{k=1}^{\ell-1} \sum_{i=1}^r w_{ik}^2 = \sum_{i=1}^{\ell-1} [WW^T]_{ii} = \ell - 1. \] (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 \( P_\ell \) as follows.
\[ P_\ell \geq 1 - \frac{1}{\|A\|_F^2} \sum_{i=1}^{\ell-1} \sigma_i^2 = \frac{\sum_{i=1}^{\ell-1} \sigma_i^2}{\|A\|_F^2}, \] (D10)
\[ P_\ell \leq 1 - \frac{1}{\|A\|_F^2} \sum_{i=r-\ell+2}^r \sigma_i^2 = \frac{\sum_{i=r-\ell+2}^r \sigma_i^2}{\|A\|_F^2}. \] (D11)

**Lemma 6.** Denote \( P \) to be the distribution of the adaptive sampling following from the Eq. (10):
\[ P(s_1, \ldots, s_\ell) = \Pr^{(1)}(s_1) \Pr^{(2)}(s_2) \cdots \Pr^{(\ell)}(s_\ell), \] (D12)
where \( s_\ell \in [m] \) denotes the index of the row \( s_\ell \) in the matrix \( A \in \mathbb{R}^{m \times d}. \) Then
\[ \mathbb{E}_P[\sigma_{\min}(C_\ell)] \geq \frac{r - \ell + 1}{\ell r} \kappa^{-2\ell}. \]

**Proof.** By the Cauchy-Schwarz Inequality, we have:
\[ \mathbb{E}_P[\sigma_{\min}(C_\ell)] \cdot \mathbb{E}_P[\sigma_{\min}(C_{\ell})] \geq 1. \] (D13)

If the following inequality were true,
\[ \mathbb{E}_P[\sigma_{\min}(C_\ell)] \leq \frac{\ell r}{r - \ell + 1} \kappa^{-2\ell}, \] (D14)
then we could reach the conclusion of this lemma:
\[ \mathbb{E}_P[\sigma_{\min}(C_\ell)] \geq \frac{1}{\mathbb{E}_P[\sigma_{\min}^2(C_\ell)]} \] (D15)
\[ \geq \frac{r - \ell + 1}{\ell r} \kappa^{-2\ell}. \] (D16)

To prove Eq. (D14), we first rewrite it as follows:
\[ \mathbb{E}_P[\sigma_{\min}^2(C_\ell)] \]
\[ = \sum_{s_{i1}=1}^m \cdots \sum_{s_{\ell \ell}=1}^m \frac{m}{s_{i1}=1} \cdots \frac{m}{s_{\ell \ell}=1} P(s_1, \ldots, s_\ell) \sigma_{\min}^{-2}(C_\ell) \]
\[ = \sum_{s_{i1}=1}^m \cdots \sum_{s_{\ell \ell}=1}^m \frac{m}{s_{i1}=1} \cdots \frac{m}{s_{\ell \ell}=1} \frac{\|s_1\|^2 \cdots \|s_\ell\|^2}{\Sigma(t) \cdots \Sigma(t)} |C_\ell| \sigma_{\min}^{-2}(C_\ell). \] (D18)

In Eq. (D18), we rewrite \( P(s_1, \ldots, s_\ell) \) with Eq. (D12) and
\[ \Pr^{(\ell)}(s_\ell) = \frac{\|s_\ell\|^2 \cdot \|s_\ell\| - \sum_{i=1}^{\ell-1} |t_i| \langle t_i, s_\ell \rangle |^2}{\Sigma(t)^\ell} \]
\[ = \frac{\|s_\ell\|^2 Z_{\ell}^2}{\Sigma(t)} \]
\[ = \frac{\|s_\ell\|^2 \cdot |C_\ell|}{\Sigma(t)^\ell |C_{\ell-1}|}, \] (D21)

where, in Eq. (D20), we denote
\[ \Sigma(t)^\ell = \sum_{s_{i1}=1}^m \cdots \sum_{s_{i\ell}=1}^m \frac{m}{s_{i1}=1} \cdots \frac{m}{s_{i\ell}=1} |t_i|^2 \langle t_i, s_\ell \rangle |^2, \] (D22)

Eq. (D20) is derived from \( Z_{\ell}^2 = \|s_\ell\| - \sum_{i=1}^{\ell-1} |t_i| \langle t_i, s_\ell \rangle |^2, \) and Eq. (D21) is due to Eq. (A6).

Continuing from Eq. (D18), it holds
\[ \mathbb{E}_P[\sigma_{\min}^2(C_\ell)] \]
\[ \leq \sum_{s_{i1}=1}^m \cdots \sum_{s_{i\ell}=1}^m \frac{m}{s_{i1}=1} \cdots \frac{m}{s_{i\ell}=1} \frac{\ell}{\Sigma(t)^\ell |C_{\ell-1}|} \] (D23)
where $\Sigma^{(j)}$ is a principal submatrix of $C_\ell$ obtained by removing the $j$-th row and column. The following inequalities hold:

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

$\|A\|_F^2 - P_j$  \quad \text{(D26)}

The inequality follows by rearranging the sum order.

The next step is to provide a lower bound on the denominator term $\Sigma^{(1)} \cdots \Sigma^{(\ell)}$ in $\Sigma_j^{(j)}$. Note that for any $1 \leq j \leq \ell$, $\Sigma^{(j)}$ 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,

$\Sigma^{(j)} := \Sigma^{(j)}(s_1, s_2, \ldots, s_{j-1})$  \quad \text{(D27)}

where $\sigma_1(A) \geq \sigma_2(A) \geq \cdots \geq \sigma_r(A)$ denote singular values of $A$.

Equation (D27) holds for any choice of linearly independent row vectors for $\Sigma^{(j)}$. Then Eq. (D27) yields

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

$\geq \sigma_{\min}^2(A) \Sigma^{(j)}(s_1, \ldots, s_{j-1}, s_{i+1}, \ldots, s_j)$  \quad \text{(D29)}

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

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

$\Sigma^{(j)}(s_1, \ldots, s_{j-1}, s_{i+1}, \ldots, s_j) \leq \frac{\sigma_{\min}^2(A)}{\sigma_{\max}^2(A)}$  \quad \text{(D31)}

Continuing from Eq. (D24) and with the inequality $\Sigma_j^{(j)} \geq \sum_{i=1}^{\ell} r \sigma_i^2(A)$ in Eq. (D27), we obtain the first inequality below:

$\Sigma_j^{(j)} \geq \sum_{i=1}^{\ell} \sum_{s_j=1}^{\ell} \sigma_i^2(A) \leq \frac{1}{\Sigma_j^{(j)}} \Sigma^{(j)}(s_1, \ldots, s_{j-1}, s_{i+1}, \ldots, s_j)$  \quad \text{(D32)}

$\text{Pr}(s_\ell | s_1, \ldots, s_{\ell-1}) := P(s_\ell | s_1, \ldots, s_{\ell-1})$  \quad \text{(D33)}

$\text{Pr}(s_\ell | s_1, \ldots, s_{\ell-1}) := P(s_\ell | s_1, \ldots, s_{\ell-1})$  \quad \text{(D34)}

where $\text{Pr}(s_\ell | s_1, \ldots, s_{\ell-1})$ is the probability of sampling $(s_1, s_2, \ldots, s_{\ell-1}, s_{\ell+1}, \ldots, s_j)$ over all choices of indices $s_j \in [r]$.
\[ \tilde{\Sigma}^{(\ell)} = \sum_{s_j=1}^{m} \|s_{\ell}\|^2 \left\| \frac{\bar{\Pi}_{\ell-1} + I \cdot s_{\ell}}{2} \right\|^2 \]  
(D43)

are corresponding normalization factors. Now we prove Lemma [4].

**Proof.** The main idea is that, if the following statement holds true for any \(0 \leq j \leq \ell - 1:\)

\[ E_{\tilde{P}^{(j+1)}} E_{\tilde{P}^{(j+2)}} \cdots E_{\tilde{P}^{(\ell)}} \sigma_{\min}(C_{\ell}) \geq 1 - \frac{1}{6\ell} \left( 1 - \frac{1}{6\ell} \right)^2 \cdots \left( 1 - \frac{1}{6\ell} \right)^{\ell} \cdots = 1 - \frac{1}{6\ell} \sum_{s_j=1}^{m} \sigma_{\min}(C_{\ell}) \]  
(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

\[ E_{\tilde{P}}[\sigma_{\min}(C_{\ell})] = E_{\tilde{P}^{(1)}} \cdots E_{\tilde{P}^{(\ell-1)}} \sigma_{\min}(C_{\ell}) \geq \frac{1}{6\ell} E_{\tilde{P}}[\sigma_{\min}(C_{\ell})] \]  
(D45)

\[ \frac{1}{6\ell} E_{\tilde{P}}[\sigma_{\min}(C_{\ell})] \geq \frac{1}{6\ell} E_{\tilde{P}}[\sigma_{\min}(C_{\ell})] \]  
(D46)

\[ \frac{2}{3} E_{\tilde{P}}[\sigma_{\min}(C_{\ell})], \]  
(D47)

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

\[ \left( 1 - \frac{1}{6\ell} \right) \geq 1 - \frac{1}{6} \]  

(D48)

(D49)

to obtain the last inequality.

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

\[ \Sigma^{(j+1)} = \frac{\Sigma^{(j)} + \Sigma^{(j+1)} - \|s_{j+1}\|^{2} \|s_{j+1}\|}{2 \|s_{j+1}\|} \|s_{j+1}\| \]  
(D50)

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

\[ -\|\tilde{\Pi}_{j} - \Pi_{j}\| \leq \langle s_{j+1}, \tilde{\Pi}_{j} - \Pi_{j} \rangle \leq \|\tilde{\Pi}_{j} - \Pi_{j}\|. \]  

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

\[ \|\tilde{\Pi}_{j} - \Pi_{j}\| \leq \sum_{i=1}^{j} \|\tilde{R}_{i} - R_{i}\| \leq j \epsilon_{R}, \]  

where we denote by \(\epsilon_{R} = \frac{1}{3 \sigma_{\min} \epsilon_{F}}\) the error bound on each \(R_{i}\), as provided in the assumption of this Lemma. Notice that

\[ 0 \leq \sigma_{\min}(C_{\ell}) \leq \frac{\text{Tr}[C_{\ell}]}{\ell} \leq 1 \]  

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

\[ X = \sum_{s_{j+1}=1}^{m} \tilde{P}^{(j+1)}(s_{j+1}) \cdot X \]  

(D53)

(D54)

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

\[ \geq \sum_{s_{j+1}=1}^{m} \left( \frac{\Sigma^{(j+1)} + \|\tilde{\Pi}_{j+1}\| \|s_{j+1}\|^{2} \|s_{j+1}\|}{2 \|s_{j+1}\|} \right) \]  

(D55)

(D56)

(D57)

(D58)

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

\[ \frac{\Sigma^{(j+1)} + \|\tilde{\Pi}_{j+1}\| \|s_{j+1}\|^{2} \|s_{j+1}\|}{2 \|s_{j+1}\|} \leq \frac{\Sigma^{(j+1)} + \|\tilde{\Pi}_{j+1}\| \|s_{j+1}\|^{2} \|s_{j+1}\|}{2 \|s_{j+1}\|} \]  

(D59)
The first inequality follows from \( \epsilon_R = \frac{1}{3r^5\kappa^2} \) and

\[
P_{j+1} \geq \sum_{i=1}^{\ell+1} \sigma_i^2(A) \geq \frac{(r-j)^2}{r^2\sigma_{\text{max}}^2(A)} = \frac{r-j}{r^2},
\]

where the first inequality uses Lemma 5, Eq. (D61) holds due to the lower bound in Lemma 6. Finally, Eq. (D58) is derived by using

\[
P_{j+1} \geq 1 - \frac{1}{6l} \mathbb{E}_T[\sigma_{\text{min}}(C)] \geq 1 - \frac{1}{6l},
\]

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 1 for \( N \) times to guarantee sampling out one basis which satisfies the conditions

\[
\text{cond}^{(0)}: \left\{ \sigma_{\text{min}}(C_r) \geq \frac{1}{2\kappa^2 \kappa^2 r^2}, \forall \ell \in [r] \right\} \quad (E1)
\]

Let \( T_{\text{QGSP}} \) be the query complexity of oracles \( U_A \) and \( V_A \) to implement Algorithm 1 once. Thus, the overall query complexity is

\[
T_{\text{basis}} = NT_{\text{QGSP}}. \quad (E2)
\]

To begin with, consider the first iteration of Algorithm 1. The Gram matrix of the sampled basis has the dimension \( 1 \times 1 \) with one element 1. Thus, the condition \( \text{cond}^{(0)} \) 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. (11), has been obtained. Next, we move on to the \( \ell \)-th iteration of Algorithm 1. 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 1 from the first iteration. Thus, the conditions in Eq. (E1) would hold during the procedure, with the cost that Algorithm 1 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 \( T_{\text{QGSP}} \) 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 \( \{z_r\}_{r=1}^{r-1} \), and the cost of implementing gates \( \{C(R)\}_{r=1}^{r-1} \). Based on Lemma 2 and Lemma 3, the latter two complexities depend on the error in the implementation of \( R \). In the following proof, we provide explicit upper bounds of \( N \) and \( T_{\text{QGSP}} \), by setting

\[
\epsilon_C = \frac{1}{960r^{3} \kappa^6}
\]

to be the error bound of each element in \( C_r \).

Firstly we demonstrate that the sampling in each iteration of Algorithm 1 obeys the distribution in Eq. (13), i.e., the error of each gate \( C(R_j) \) is bounded as

\[
\|R_j - R\| \leq \epsilon_R = \frac{1}{3r^5\kappa^2}, \forall j \in [r - 1].
\]

Based on Lemma 2, the error of \( t_j \) induced by noisy coefficients is bounded by

\[
\|t_j - t\| \leq \frac{8j^2}{\sigma_{\text{min}}^2(C)} \epsilon_C \leq 32r^5 \kappa^2 \epsilon_C \leq \frac{1}{30r^5 \kappa^2} \epsilon_C \leq \frac{\epsilon_R}{10},
\]

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

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

\[
T_{R_j} = O(j^2 \sigma_{\text{min}}^2(C) \epsilon_R^{-1}) \leq O(r^2 \kappa^r \epsilon_R^{-1}) \leq O(r^7 \kappa^{3r})
\]

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

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

\[
N = O\left(\frac{1}{\Pr\{\text{cond}^{(r)}\}}\right),
\]

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

\[
\left(1 - \Pr\{\text{cond}^{(r)}\}\right) \frac{1}{2r^2 \kappa^2 r^2} + \Pr\{\text{cond}^{(r)}\} \cdot 1 \geq \left(1 - \Pr\{\text{cond}^{(r)}\}\right) \frac{1}{2r^2 \kappa^2 r^2} + \Pr\{\text{cond}^{(r)}\} \mathbb{E}_P[\sigma_{\text{min}}(C_r) | \text{cond}^{(r)}] \geq \left(1 - \Pr\{\text{cond}^{(r)}\}\right) \mathbb{E}_P[\sigma_{\text{min}}(C_r) | \text{not cond}^{(r)}]
\]
Taking the complexity of estimating form operations $U^\ell$ in each iteration of $T_C$. Thus, Eq. (E18) together with Eq. (E13) yields the following statement:

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

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

Combining Eq. (E4) with Lemma 1 we have the following statement:

$$\mathbb{E}_\rho[\sigma_{\text{min}}(C_r)] \geq \frac{2}{3} \mathbb{E}_\rho \left[ \prod_{r=1}^\ell \mathbb{E}_r[\sigma_{\text{min}}(C_r)] \right] (E17)$$

$$\geq \frac{2}{3r^2k^{2r-2}}. (E18)$$

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

$$1 - \Pr\{\text{cond}^{(r)}\} \geq \frac{1}{2r^2k^{2r-2}}.$$  

We could solve

$$\Pr\{\text{cond}^{(r)}\} \geq \frac{1}{1 - \frac{1}{2r^2k^{2r-2}}} = \frac{1}{6r^2k^{2-2r}},$$

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

Finally we move on to analyze the query complexity $T_{QGSP}$. Based on Lemma 1 coefficients $\{z_t\}_{t=1}^\ell$ 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

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

$$= O(r^2k^{2r-2}) + O(r^{25}k^{12r}), (E23)$$

$$\leq O(r^{25}k^{12r}) + O(r^8k^{3r}) \sum_{\ell=1}^r \frac{1}{^\ell P_t}$$

$$\leq O(r^{25}k^{12r}) + O(r^8k^{3r}) \sum_{\ell=1}^r r^\ell (E24)$$

$$\leq O(r^{25}k^{12r}) + O(r^8k^{3r}) \sum_{\ell=1}^r \frac{1}{^\ell P_t}$$

$$\leq O(r^{25}k^{12r}) + O(r^8k^{3r}) \sum_{\ell=1}^r \frac{1}{^\ell P_t}$$

$$\leq O(r^{25}k^{12r}) + O(r^8k^{3r}) \sum_{\ell=1}^r \frac{1}{^\ell P_t}$$

$$\leq O(r^{25}k^{12r}). (E27)$$

Eq. (E24) is obtained by using Eq. (E22) and Eq. (E11). Eq. (E26) is derived by using Eq. (E12).

By considering $N \leq O(r^2k^{2r-2})$ being the required number of times to run Algorithm 1 we prove the Theorem 3.

$$\square$$

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 $c$-error estimation to the value $a_i = \langle t_k | v | t_i \rangle$, for all $i \in [r]$, with $O(1/c^2)$ measurements.

Firstly, after all unitary operations, the state in Fig. 2 can be written as:

$$\frac{1}{4} (| v \rangle | t_k \rangle + | v \rangle | t_i \rangle + | t_k \rangle | v \rangle + | t_i \rangle | v \rangle)$$

$$+ \frac{1}{4} (| v \rangle | t_k \rangle - | v \rangle | t_i \rangle + | t_k \rangle | v \rangle - | t_i \rangle | v \rangle)$$

$$+ \frac{1}{4} (| v \rangle | t_k \rangle + | v \rangle | t_i \rangle - | t_k \rangle | v \rangle - | t_i \rangle | v \rangle)$$

$$+ \frac{1}{4} (| v \rangle | t_k \rangle - | v \rangle | t_i \rangle - | t_k \rangle | v \rangle + | t_i \rangle | v \rangle).$$

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

$$P_{00} = \frac{2 + |\langle v | t_k \rangle|^2 - |\langle v | t_i \rangle|^2 + 2 |\langle t_k | v \rangle |\langle t_i | v \rangle|}{8}$$

$$P_{11} = \frac{2 - |\langle v | t_k \rangle|^2 - |\langle v | t_i \rangle|^2 + 2 |\langle t_k | v \rangle |\langle t_i | v \rangle|}{8}.$$

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

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

The efficiency of the quantum circuit in Fig. 2 depends on the efficiency of preparing the input state $| t_k \rangle | 0 \rangle + | t_i \rangle | 1 \rangle$ with $\ell^2$ norm error bounded by $\epsilon$. Lemma 7 below proves that it can be prepared with query complexity $O(r \sigma_{\text{min}}^{-1/2}(C_r))$.

**Lemma 7.** Given perturbed coefficients provided in Lemma 4 for both indices $k$ and $\ell$, the state $\frac{1}{\sqrt{2}} (| t_k \rangle + | t_i \rangle)$ could be prepared with query complexity $O(r \sigma_{\text{min}}^{-1/2}(C_r))$ with $\ell^2$ norm error bounded by $\epsilon$.

**Proof.** We sketch the main idea of the proof. Firstly, we generate the superposition state of $| t_i \rangle$ and $| t_k \rangle$ using perturbed coefficients, where perturbed vectors are expressed as

$$t_\ell = \sum_{i=1}^r \tilde{t}_\ell s_i / \| s_i \|, \quad t_k = \sum_{i=1}^r \tilde{t}_k s_i / \| s_i \|. (F1)$$
Then, we provide the error analysis. Specifically, Given the coefficients $\tilde{z}_k$ and $\tilde{z}_\ell$, we prepare the state

$$\frac{1}{\sqrt{\|\hat{k}\|^2 + \|\hat{\ell}\|^2}}(\|\hat{k}\|\langle 0|\hat{k}\rangle + \|\hat{\ell}\|\langle 1|\hat{\ell}\rangle)$$

by the LCU method as follows. Since the notation $k$ and $\ell$ are symmetrical here, we could assume that $\ell \geq k$ for convenience. Firstly, we initialize the state $\frac{1}{\sqrt{2}}|0\rangle \otimes \log m|0\rangle \otimes \log n|0\rangle$. 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 $U_{\text{index}}$ defined in Eq. (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 register, followed by the unitary $U_{\text{index}}^\dagger$ to yield:

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

Denote $\tilde{z} \equiv \max_j \{\tilde{z}_j\}$, Next, we perform the controlled rotation

$$|0\rangle \otimes \sum_{j=1}^k |j\rangle |j\rangle \otimes I \otimes e^{-i\sigma_y \arccos(\tilde{z}_j/\tilde{z})}$$

$$+ |0\rangle |0\rangle \otimes \sum_{j=k+1}^\ell |j\rangle |j\rangle \otimes I \otimes \sigma_x$$

$$+ |0\rangle |0\rangle \otimes \sum_{j=\ell+1}^m |j\rangle |j\rangle \otimes I \otimes I$$

$$+ |1\rangle |1\rangle \otimes \sum_{j=1}^k |j\rangle |j\rangle \otimes I \otimes e^{-i\sigma_y \arccos(\tilde{z}_j/\tilde{z})}$$

$$+ |1\rangle |1\rangle \otimes \sum_{j=k+1}^\ell |j\rangle |j\rangle \otimes I \otimes I,$$

to obtain the state

$$\frac{1}{\sqrt{2\ell}} \sum_{j=1}^k |j\rangle |s_j\rangle \left(\frac{\tilde{z}_j}{\tilde{z}} |0\rangle + \sqrt{1 - \frac{\tilde{z}_j^2}{\tilde{z}^2}} |1\rangle\right)$$

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

$$+ \frac{1}{\sqrt{2\ell}} \sum_{j=\ell+1}^m |j\rangle |s_j\rangle \left(\frac{\tilde{z}_j}{\tilde{z}} |0\rangle + \sqrt{1 - \frac{\tilde{z}_j^2}{\tilde{z}^2}} |1\rangle\right).$$

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

$$\frac{1}{\sqrt{2\ell^2}} \sum_{j=1}^k |0\rangle \sum_{j=1}^\ell \tilde{z}_j |s_j\rangle |0\rangle + \frac{1}{\sqrt{2\ell^2}} \sum_{j=1}^\ell \tilde{z}_j |s_j\rangle |0\rangle$$

$$+ \text{orthogonal garbage state}$$

$$= \frac{1}{\sqrt{2\ell^2}} \left(\|\hat{k}\|\langle 0|\hat{k}\rangle |0\rangle + \|\hat{\ell}\|\langle 1|\hat{\ell}\rangle |0\rangle\right)$$

$$+ \text{orthogonal garbage state},$$

The measurement on the 2-nd and the 4-th registers of the final state could yield state in [F2] with probability

$$\frac{\|\hat{\ell}\|^2 + \|\hat{k}\|^2}{2\ell^2},$$

so we could prepare this state with

$$O \left(\frac{\ell^2}{\sqrt{\|\hat{\ell}\|^2 + \|\hat{k}\|^2}}\right)$$

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

Now we analyze the distance between the state $\frac{1}{\sqrt{2}}(|0\rangle |t_k\rangle + |1\rangle |t_\ell\rangle)$ and the state in [F2] as follows.

$$\|\|\hat{k}\|\| |\hat{k}\rangle + \|\|\hat{\ell}\|\| |\hat{\ell}\rangle - \frac{|0\rangle |t_k\rangle + |1\rangle |t_\ell\rangle}{\sqrt{2}}\|$$

$$\leq \|\|\hat{k}\|\| - \|\hat{k}\| - \|\hat{\ell}\| \|\hat{\ell}\| - \frac{|t_k\rangle - |t_\ell\rangle}{\sqrt{2}}\|$$

$$\leq \|\|\hat{k}\|\| - \|\hat{k}\| - \|\hat{\ell}\| \|\hat{\ell}\| - \frac{|t_k\rangle - |t_\ell\rangle}{\sqrt{2}}\|$$

$$\leq \left(\frac{1 + \epsilon}{4}\right) \frac{1}{\sqrt{2} \left(1 - \frac{\epsilon}{4}\right)^2} - \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}^2} \leq \epsilon,$$

where Eq. (F6) is derived by using

$$1 - \frac{\epsilon}{4} \leq \|t_i\| - \|t_i - \bar{t}_i\| \leq \|t_i\| \leq \|t_i\| + \|t_i - \bar{t}_i\| \leq 1 + \frac{\epsilon}{4},$$

for $i = k/\ell$. 

$\Box$
Now we begin the proof of Theorem 4 that provides the error analysis of Algorithm 2 for reading out the state $|v\rangle$.

Proof. We firstly study the error in the read-out procedure and then provide the time analysis. Specifically, notice that the state $\frac{1}{\sqrt{2}} \left( |0\rangle |t_k\rangle + |1\rangle |t_i\rangle \right)$ generated by Lemma 4 is perturbed due to the noisy coefficients $z_k$ and $z_i$. Thus, the read-out error consists of two parts: the error on generating $\frac{1}{\sqrt{2}} \left( |0\rangle |t_k\rangle + |1\rangle |t_i\rangle \right)$, and the error induced by the statistical noise during the measurement in the Fig 2.

Firstly, we analyze the measurement distribution of Fig. 2 which uses the perturbed input state $\frac{1}{\sqrt{2}} \left( |0\rangle |t_k\rangle + |1\rangle |t_i\rangle \right)$. Denote $\tilde{z}_j$ and $\tilde{t}_j$ as the perturbed form of $z_j$ and $t_j$, respectively, $\forall j \in [r]$. In this proof, we assume the $\ell_2$ norm on the error of each $t_j$ is bounded by $\epsilon_3 = \frac{1}{14\sqrt{r}} \epsilon$.

The final state in Fig. 2 is:

$$\frac{1}{4} \left[ f_{x_k} |v\rangle |\tilde{t}_k\rangle + f_{x_i} |v\rangle |\tilde{t}_i\rangle + f_{x_k} |\tilde{t}_k\rangle |v\rangle + f_{x_i} |\tilde{t}_i\rangle |v\rangle \right] |0\rangle$$

$$+ \frac{1}{4} \left[ f_{x_k} |v\rangle |\tilde{t}_k\rangle - f_{x_i} |v\rangle |\tilde{t}_i\rangle + f_{x_k} |\tilde{t}_k\rangle |v\rangle - f_{x_i} |\tilde{t}_i\rangle |v\rangle \right] |1\rangle$$

$$+ \frac{1}{4} \left[ f_{x_k} |v\rangle |\tilde{t}_k\rangle + f_{x_i} |v\rangle |\tilde{t}_i\rangle - f_{x_k} |\tilde{t}_k\rangle |v\rangle - f_{x_i} |\tilde{t}_i\rangle |v\rangle \right] |0\rangle$$

$$+ \frac{1}{4} \left[ f_{x_k} |v\rangle |\tilde{t}_k\rangle - f_{x_i} |v\rangle |\tilde{t}_i\rangle - f_{x_k} |\tilde{t}_k\rangle |v\rangle + f_{x_i} |\tilde{t}_i\rangle |v\rangle \right] |1\rangle,$$

where we denote

$$f_{x_k} = \sqrt{\frac{2||\tilde{t}_k||^2}{||\tilde{t}_k||^2 + ||\tilde{t}_i||^2}}, f_{x_i} = \sqrt{\frac{2||\tilde{t}_i||^2}{||\tilde{t}_k||^2 + ||\tilde{t}_i||^2}}.$$

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

$$\tilde{P}_{00} = \frac{1}{4} f_{x_k} f_{x_i} (\tilde{t}_k |v\rangle \langle \tilde{t}_k| + f_{x_k} f_{x_i} (\tilde{t}_i |v\rangle \langle \tilde{t}_i| + f_{x_k} f_{x_i} (\tilde{t}_k |v\rangle \langle \tilde{t}_i| + f_{x_k} f_{x_i} (\tilde{t}_i |v\rangle \langle \tilde{t}_k|)$

$$\tilde{P}_{11} = \frac{1}{4} f_{x_k} f_{x_i} (\tilde{t}_k |v\rangle \langle \tilde{t}_k| + f_{x_k} f_{x_i} (\tilde{t}_i |v\rangle \langle \tilde{t}_i| + f_{x_k} f_{x_i} (\tilde{t}_k |v\rangle \langle \tilde{t}_i| + f_{x_k} f_{x_i} (\tilde{t}_i |v\rangle \langle \tilde{t}_k|)$

Thus, the perturbed statistics of outcomes 00 and 11 is:

$$\tilde{P}_{\text{same}} = \tilde{P}_{00} + \tilde{P}_{11} = \frac{1}{2} f_{x_k} f_{x_i} (\tilde{t}_k |v\rangle \langle \tilde{t}_k| + f_{x_k} f_{x_i} (\tilde{t}_i |v\rangle \langle \tilde{t}_i| = \frac{1}{2} \tilde{a}_i \tilde{a}_k,$$

where we denote $\tilde{a}_i = f_{x_i} (\tilde{t}_i |v\rangle$.

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 $\epsilon_2 = \frac{1}{14\sqrt{r}} \epsilon$, and denote $(\tilde{a}_i \tilde{a}_k)$ as the approximated value of $\tilde{a}_i \tilde{a}_k$. Then, in parallel to the exact form

$$v = \sum_{i=1}^{r} a_i t_i,$$

we use the expression

$$\tilde{v} = \sum_{i=1}^{r} \tilde{a}_i t_i,$$

as the perturbed description of the vector $v$, where

$$\tilde{a}_i = \frac{\langle \tilde{a}_i \tilde{a}_k \rangle}{\sum_{i=1}^{r} \langle \tilde{a}_i \tilde{a}_k \rangle}, \forall i \in [r].$$

Thus, the $\ell^2$ norm of the error on the vector description of the read-out state could be bounded as follows.

$$\|\tilde{v} - v\| = \left\| \sum_{i=1}^{r} (\tilde{a}_i t_i - a_i t_i) \right\|$$

$$\leq \sum_{i=1}^{r} \langle \tilde{a}_i | a_i \rangle + \sum_{i=1}^{r} \langle \tilde{a}_i | \tilde{a}_i \rangle$$

$$\leq \sum_{i=1}^{r} \langle \tilde{a}_i | a_i \rangle + \sum_{i=1}^{r} \langle \tilde{a}_i | a_i \rangle$$

$$\leq \epsilon_2 + \epsilon_3 \leq \epsilon,$$

where Eq. (F11) is obtained by using Eqs. (F9), Eq. (F2), follows from the triangular inequality. Eq. (F13) holds due to $\| \tilde{t}_i - t_i \| \leq \epsilon_3$ and $\tilde{t}_i \tilde{t}_j = \delta_{ij}$.

Eq. (F14) is derived by using the definition in Eq. (F10) and

$$\frac{\sum_{i=1}^{r} |\tilde{a}_i|^2}{r} \leq \frac{1}{\sqrt{r}}.$$
Eq. (F18) follows from the definitions by using \( \epsilon \) obtained by the assumption \( |\epsilon| \leq 1 \) in Eqs. (F21) and (F22) are derived by using \( \|\| \leq 1 \) and \( \|\| \leq 1 \) in Eq. (F23) is obtained by the assumption \( \epsilon_3 = 0 \).

On the other hand, the term \( \tilde{a} \) is bounded around \( \tilde{a} \) as follows,

\[
\tilde{a}_k = f_i(\tilde{t}_i|v), \quad a_i = (\tilde{t}_i|v).
\]

Eq. (F19) is obtained by using the definition \( f_i = \sqrt{2|\tilde{t}_i|/|\tilde{t}_i|} \). Eq. (F20) follows from the triangular inequality and \( \|v\| = 1 \). Eqs. (F21) and (F22) are derived by using \( \|\| = 1 \) and \( \|\| = 1 \) in Eq. (F23) is obtained by the assumption \( \epsilon_3 = 0 \).

Finally, we analyze the time complexity of the protocol. Notice that the error \( \|\| \leq \epsilon_3 \) could be achieved for all \( i \in [r] \) by using

\[
(1 - \epsilon_3) \leq 1,
\]

\[
(1 - \epsilon_3) \leq 1,
\]

queries to input oracles due to Lemma 2. Besides, the error \( \epsilon_2 \) induced as the statistical noise during the measurement in Fig. 2 could be achieved by using \( \epsilon_2 = O(r^3 \epsilon^{-2}) \) copies of states \( |v\rangle \) and \( \frac{1}{\sqrt{2}}(|0\rangle|\tilde{t}_k\rangle + |1\rangle|\tilde{t}_i\rangle) \) for \( i \in [r] \), where the latter state could be prepared by using \( O(r^2 \sigma_{\min}^{-1/2}(C_i)) \) queries to input oracles. By summing for each \( i \in [r] \), we need \( O(r^2 \epsilon^{-2}) \) copies of the state \( |v\rangle \) 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 2.

\[ \]

[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. Provansol, W. A. de Jong, and C. W. Bauer, Quantum algorithm for high energy 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, 1 (2019).

[5] Y. R. Sanders, D. W. Berry, P. C. 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, UC Berkeley (2014).

[7] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd, Quantum machine learning, Na-
V. Giovannetti, A. Harrow, and S. Lloyd, Quantum algorithm for linear systems of equations, Phys. Rev. Lett. 103, 150502 (2009).

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

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

P. Rebentrost, M. Mohseni, and S. Lloyd, Quantum support vector machine for big data classification, Phys. Rev. Lett. 113, 130503 (2014).

S. Lloyd, M. Mohseni, and P. Rebentrost, Quantum algorithms for supervised and unsupervised machine learning, arXiv:1307.0411 (2013).

I. Kerenidis and A. Prakash, Quantum gradient descent, Phys. Rev. Lett. 105, 200502 (2010).

I. Kerenidis and A. Prakash, A quantum interior point method for lps and sdps, ACM Transactions on Quantum Computing 1, 1 (2020).

Another implementation of the QRAM oracle is proposed in Ref. [24], which requires $O(k\text{polylog}(md))$ quantum operations and physical resources for $m \times d$ matrix with $k$ non-zero elements.

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).

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

S. Aaronson, Read the fine print, Nature Physics 11, 291 (2015).

D. Gross, Y.-K. Liu, S. T. Flammia, S. Becker, and E. Jens, Quantum state tomography via compressed sensing, Phys. Rev. Lett. 105, 150401 (2010).

A. Kyrillidis, A. Kalev, D. Park, S. Bhojanapalli, C. Caramanis, and S. Sanghavi, Provably compressed sensing quantum state tomography via non-convex methods, npj Quantum Information 4, 36 (2018).

J. Haah, A. W. Harrow, Z. Ji, X. Wu, and N. Yu, Sample-optimal tomography of quantum states, IEEE Transactions on Information Theory 63, 5628 (2017).

R. O’Donnell and J. Wright, Efficient quantum tomography, in Proceedings of the 48th annual ACM symposium on Theory of Computing (ACM, 2016) pp. 899–912.

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, Nature communications 1, 149 (2010).

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).

H. Haffner, W. Hansel, C. Roos, et al., Scalable multiparticle entanglement of trapped ions, Nature 438, 643 (2005).

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).

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

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

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).

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).

G. Struchalin, Y. A. Zagorovskii, E. Kovlakov, S. Straupe, and S. 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 (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 transactions on pattern analysis and machine intelligence 41, 2628 (2018).
[48] M. Udell and A. Townsend, Why are big data matrices approximately low rank?, SIAM Journal on Mathematics of Data Science 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 (2018) pp. 4261–4271.
[52] M. Vanner, M. Aspelmeyer, and M. 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. Svobid, 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 Information & Computation 12, 901 (2012).
[57] S. Lloyd, M. Mohseni, and P. Rebentrost, Quantum principal component analysis, Nature Physics 10, 631–633 (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 Journal on Computing 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 (Association for Computing Machinery, New York, NY, USA, 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 (2018), 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 (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.
[67] G. Brassard, P. Hoyer, M. Mosca, and A. Tapp, Quantum amplitude amplification and estimation, Contemporary Mathematics 305, 53 (2002).