

# Quantum Algorithms to Matrix Multiplication

Changpeng Shao  
cpshao@amss.ac.cn

Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China

**Abstract.** In this paper, we study quantum algorithms of matrix multiplication from the viewpoint of inputting quantum/classical data to outputting quantum/classical data. The main target is trying to overcome the input and output problem, which are not easy to solve and many quantum algorithms will encounter, to study matrix operations in quantum computer with high efficiency. And solving matrix multiplication will be the first step. We propose three quantum algorithms to matrix multiplication based on swap test, SVE and HHL. From the point of making fewer assumptions, swap test method works the best than the other two. We also show that the quantum algorithm of matrix multiplication with classical input and output data by swap test achieves the best complexity  $\tilde{O}(n^2/\epsilon)$  with no assumptions. This is proved by giving an efficient quantum algorithm in polynomial time to solve the input problem, that is to prepare the quantum states of the classical data efficiently. Other contributions of this paper include: (1). Extending swap test to a more general form that is suitable to deal with quantum data in parallel, which will have further applications in other matrix operations. (2). Generalizing SVE technique such that it applies to any matrix (not just Hermitian) directly only with quantum data. (3). Proposing two new efficient quantum algorithms to prepare quantum states of classical data, which solves the input problem efficiently than other quantum algorithms.

**Key words.** quantum algorithm, quantum computation, matrix multiplication, quantum state preparation

## 1 Introduction

In the study of quantum algorithms (for example, see [3], [13], [28]), people usually encounter the “input” and “output” problem. The input problem is the transformation from classical data (such as complex vectors) into quantum data (such as quantum states); the output problem is the converse. These two problems generally are not easy to solve efficiently in quantum computer, and sometimes even cost more than other steps of the quantum algorithms. So mostly we just assume that we already get the quantum data by some methods when studying quantum algorithms. Another important idea is studying quantum algorithms with input and output data are all quantum, such as the quantum machine learning for quantum data [3], principal component analysis [18], quantum simulator [19], [29] and so on.

In this paper, we study the matrix multiplication in quantum computer from two different perspectives with three different techniques. The two perspectives share the same input data, that is it can be classical or quantum. The difference reflects in the output, one is quantum and the other one is classical. The three techniques we will use are swap test [4], SVE [14] and HHL [13]. In the two different perspectives, if the input data is quantum, then we can only apply swap test and SVE; if the input data is classical, then all the three techniques can play roles in. What we care more in this work is quantum data to quantum data. Although SVE contains more wide applications, its performance in matrix multiplication is not efficient than swap test. Also to make swap test works for matrix multiplication from quantum data to quantum data, a more generalized version of swap test (proposition 2) will be proposed in this paper. Note that the swap test proposed in [4] can be viewed as a procedure from quantum data to classical data, now the new version will achieve quantum data to quantum data.

The target of this study aims at extending classical matrix operations into quantum case (i.e., quantum input and output), hoping to obtain efficient matrix operations in quantum computer and so solving the classical problems more efficiently. Also for the comparison with classical algorithms to matrix multiplication, quantum algorithm to matrix multiplication achieve classical data to classical data are also studied comprehensively in this work, which contains two sub-works: the preparation of quantum states (classical data to quantum data) and the reading out from quantum data into classical data.

As for the matrix multiplication, the reading out problem is not difficult to solve mainly based on swap test. However, the preparation of quantum states is not so easy in quantum computer generally. Efficient quantum algorithms to certain special cases still exists (for instance, see [8], [12], [17], [26]). One special well known case is when the classical data is relatively uniform distributed [1], [8], [17]. Based on this special case, in the paper we will propose two new quantum algorithms (theorem 6 and 7) to prepare the quantum states of classical data in the general case. The corresponding complexities are satisfactory and better than any other quantum algorithm to achieve the same target to my knowledge. When we are given enough information (such as the maximum, the minimum, the norm, the positions of nonzero entries and so on) about the classical data, then the input problem can be solved efficiently in polynomial time. Obtaining such information may take some extra time. In the matrix multiplication problem, however, all these required information can be obtained before implementing the quantum algorithms. If the given matrix is  $n$ -by- $n$ , then to get the required information will take at most  $O(n^2)$ , which is acceptable, since  $\Omega(n^2)$  is the lower bound of matrix multiplication problem. This means, the input problem can be actually solved “efficiently” in matrix multiplication.

The obtained quantum algorithms to matrix multiplication is polynomially depends on the precision. Therefore, when the precision is bounded by  $O(1/\text{poly log } n)$ , then quantum computer can solve the matrix multiplication problem efficiently in time  $\tilde{O}(n^2)$  by swap test (see table 3). The quantum algorithms obtained from SVE and HHL depend on the condition number of the given matrices. If the condition number is bounded by  $O(\text{poly log } n)$ , then these two algorithms also achieve the best efficiency. However, if we only interested in quantum output data, then all the three quantum algorithms are rely on the condition number. With the same assumption above on precision and condition number, if the input is classical data, then this problem can be solved in polynomial time by HHL; if the input is quantum data, then it can be solved in time  $\tilde{O}(\sqrt{n})$  (see table 1) by swap test and SVE.

The structure of this paper is as follows: In section 2, 3, we consider the quantum algorithms to the matrix multiplication with quantum output and classical output under one assumption about the preparation of quantum states. Section 4 devotes to study efficient quantum algorithms to solve the assumption.

**Notations.** In this paper,  $i$  refers to the imaginary unit  $\sqrt{-1}$ . For any  $n \times n$  matrix  $A = (a_{ij})$ , the notation  $\|A\|_F := \sqrt{\sum_{i,j} |a_{ij}|^2}$  refers to the Frobenius norm of  $A$ ; for any vector  $\mathbf{x} = (x_0, \dots, x_{n-1})$ , the notation  $\|\mathbf{x}\|_p := (\sum_i |x_i|^p)^{1/p}$  refers to the  $p$  norm of  $\mathbf{x}$ .

## 2 Quantum algorithm to matrix multiplication with quantum information

This section is devoted to study matrix multiplication with quantum output data. The problem can be stated as: Given two matrices  $A, B$  or their quantum states by viewing them as vectors (see (2.1) below), we want to get the quantum state of  $AB$ . First, we make some statements of the preliminaries.

Let  $A = (a_{ij})_{n \times n}$  be a given  $n$ -by- $n$  matrix. We denote its  $i$ -th row as  $A_{i\bullet}$ , the  $j$ -th column as  $A_{\bullet j}$ . In order to study matrix multiplication in quantum computer, we should have some quantum information about given classical matrices. So we make the following assumption in this and the next section:

**Assumption:** Assume that we have efficient quantum algorithm to achieve the quantum state preparation of the rows and the columns of  $A$ , such as by QRAM [10] or the data structure introduced in [14]. That is we can get

$$|A\rangle = \frac{1}{\|A\|_F} \sum_{i,j=0}^{n-1} a_{ij} |i, j\rangle = \frac{1}{\|A\|_F} \sum_{i=0}^{n-1} \|A_{i\bullet}\|_2 |i\rangle |A_{i\bullet}\rangle = \frac{1}{\|A\|_F} \sum_{j=0}^{n-1} \|A_{\bullet j}\|_2 |A_{\bullet j}\rangle |j\rangle \quad (2.1)$$

efficiently, where  $\|A\|_F = \sqrt{\sum_{i,j} |a_{ij}|^2}$  refers to the Frobenius norm of  $A$ .

*Remark 1.* By applying the inverse of the quantum algorithm of preparing the quantum states of rows and columns of  $A$  on  $|A\rangle$ , we can obtain the following quantum states

$$|A_{F\bullet}\rangle = \frac{1}{\|A\|_F} \sum_{i=0}^{n-1} \|A_{i\bullet}\|_2 |i\rangle, \quad |A_{\bullet F}\rangle = \frac{1}{\|A\|_F} \sum_{j=0}^{n-1} \|A_{\bullet j}\|_2 |j\rangle. \quad (2.2)$$

Also note that, the density matrix of  $|A\rangle$  equals  $\frac{1}{\|A\|_F^2} \sum_{i=0}^{n-1} \|A_{i\bullet}\|_2^2 |i\rangle\langle i| A_{i\bullet}\langle A_{i\bullet}|$ . By taking the trace on the second register, we will get the density matrix of  $|A_{F\bullet}\rangle$ , which is another method to get the quantum state of (2.2) from (2.1). For convenience, we define  $A_{F\bullet}$  and  $A_{\bullet F}$  as the column vectors with entries  $\|A_{i\bullet}\|_2$  and  $\|A_{\bullet j}\|_2$  respectively.

*Remark 2.* The preparation of quantum states work efficient in practical in some cases [8], [12], [17], [26]. We will discuss this in section 4. Two new efficient quantum algorithms to prepare quantum state will be proposed then, which makes the assumption above more reliable.

Given two matrices  $A, B$ , in the following, we first consider the quantum algorithm to prepare the quantum state of  $AB$  from swap test by giving the quantum information of  $A$  and  $B$ . Certainly, this problem can be solved by the singular value estimation technique (SVE) proposed in [14] still with quantum information of  $A$  and  $B$ . Moreover, if we view  $A$  as a classical data, then the quantum state of  $AB$  can also be solved by performing the matrix multiplication algorithm obtained from HHL algorithm [13]. The complexity of the algorithms obtained in the last two ways depends on the condition number of  $A$  or  $B$ . Also in HHL algorithm, we need the Hamiltonian simulation of  $\tilde{A} = \begin{pmatrix} 0 & A \\ A^\dagger & 0 \end{pmatrix}$  to be efficient.

We only consider the case of real matrix multiplication, as for complex matrix multiplication, we just need to focus on the real and imaginary parts separately.

## 2.1 By swap test

### 2.1.1 Swap test and its generalization

Swap test was first proposed in [4] as a byproduct of quantum phase estimation algorithm and Grover searching, which can be used to compute the probability of some desired quantum states. It plays an important role in many quantum algorithms, such as in HHL algorithm [13] and some machine learning algorithms [17], [21], [24], [27], [28], to estimate the inner product of two quantum states. In the following, we first briefly review the underlying problem swap test considers and the basic procedures to solve it. Notice that swap test can only returns the classical information (i.e., the inner product of two quantum states we want) if we perform a measurement, which is not enough to deal with quantum data in matrix multiplication in parallel, even if we do not perform measuring. To overcome this problem, we extend swap test into a more general form (proposition 2 below) that can output quantum data in parallel, which may has other applications except matrix multiplication considered in this paper.

Let

$$|\phi\rangle = \sin \theta |0\rangle|u\rangle + \cos \theta |1\rangle|v\rangle \quad (2.3)$$

be a unknown quantum state that can be prepared in time  $O(T_{\text{in}})$ , where  $|u\rangle, |v\rangle$  are normalized quantum states. We want to consider the problem that how to estimate  $\theta$  in quantum computer to accuracy  $\epsilon$  with high success probability at least  $1 - \delta$ .

Suppose that  $|\phi\rangle$  comes from some algorithms, that means there is a given unitary  $U$  such that  $|\phi\rangle = U|0\rangle$ . Let  $Z$  be the 2-dimensional unitary transformation that maps  $|0\rangle$  to  $-|0\rangle$  and  $|1\rangle$  to  $|1\rangle$ , which is usually called Pauli-Z matrix. Denote  $G = (2|\phi\rangle\langle\phi| - I)(Z \otimes I)$ , which is the rotation matrix used in Grover's searching algorithm. Then

$$G = \begin{pmatrix} \cos 2\theta & \sin 2\theta \\ -\sin 2\theta & \cos 2\theta \end{pmatrix}$$

in the basis  $\{|0\rangle|u\rangle, |1\rangle|v\rangle\}$ . The eigenvalues of  $G$  are  $e^{\pm i2\theta}$  and the corresponding eigenvectors are

$$|w_1\rangle = \frac{1}{\sqrt{2}}(|0\rangle|u\rangle + i|1\rangle|v\rangle), \quad |w_2\rangle = \frac{1}{\sqrt{2}}(|0\rangle|u\rangle - i|1\rangle|v\rangle)$$

respectively. Note that  $|\phi\rangle = -\frac{i}{\sqrt{2}}(e^{i\theta}|w_1\rangle - e^{-i\theta}|w_2\rangle)$ . So performing quantum phase estimation algorithm on  $G$  with initial state  $|0\rangle^n|\phi\rangle$  for some  $n = O(\log 1/\delta\epsilon)$ . We will get an approximate of the following state

$$-\frac{i}{\sqrt{2}}(e^{i\theta}|y\rangle|w_1\rangle - e^{-i\theta}|-y\rangle|w_2\rangle), \quad (2.4)$$

where  $y \in \mathbb{Z}_{2^n}$  satisfies  $|\theta - y\pi/2^n| \leq \epsilon$ . The time complexity of the above procedure is  $O(T_{\text{in}}/\epsilon\delta)$ . Sometimes,  $\delta$  will be ignored just for simplicity in the complexity analysis. Performing a measurement on (2.4), we will get an  $\epsilon$  approximate of  $\theta$ .

Furthermore, let  $f(y) = g(\theta)$  be some functions such that  $f(y) = f(-y)$  (i.e.,  $f$  is an even function), then from (2.4), we can get

$$|g(\theta)\rangle|\phi\rangle, \quad (2.5)$$

by adding a register to store  $g(\theta)$  and undoing the quantum phase estimation. This is a procedure that we want to further make use of more quantum information about  $\theta$  instead of outputting.

Now let  $|x\rangle, |y\rangle$  be two real quantum states, except a global phase, which can be prepared in time  $O(T_{\text{in}})$ . Then the above method provide us a quantum algorithm to estimate  $\langle x|y\rangle$  to accuracy  $\epsilon$  in time  $O(T_{\text{in}}/\epsilon)$ . Actually, we just need to consider the state

$$|\phi\rangle = \frac{1}{\sqrt{2}}(|+\rangle|x\rangle + |-\rangle|y\rangle) = \frac{1}{2}(|0\rangle(|x\rangle + |y\rangle) + |1\rangle(|x\rangle - |y\rangle)).$$

The probability of  $|0\rangle$  (resp.  $|1\rangle$ ) is  $(1 + \langle x|y\rangle)/2$  (resp.  $(1 - \langle x|y\rangle)/2$ ). So we can set  $\sin\theta = \sqrt{(1 + \langle x|y\rangle)/2}$  and  $\cos\theta = \sqrt{(1 - \langle x|y\rangle)/2}$ . The quantum state  $|\phi\rangle$  can be rewritten in the form (2.3), where  $|u\rangle, |v\rangle$  corresponds to the normalization of  $|x\rangle + |y\rangle, |x\rangle - |y\rangle$ . Therefore, the inner product  $\langle x|y\rangle$  can be evaluated in time  $O(T_{\text{in}}/\epsilon)$  with accuracy  $\epsilon$ . Concluding this, we get the following result

**Proposition 1.** *Let  $|x\rangle, |y\rangle$  be two quantum states, which can be prepared in time  $O(T_{\text{in}})$ , then  $\langle x|y\rangle$  can be estimated with accuracy  $\epsilon$  in time  $O(T_{\text{in}}/\epsilon)$ .*

*Remark 3.* If  $|x\rangle, |y\rangle$  are complex quantum states, then the probability of  $|0\rangle$  (resp.  $|1\rangle$ ) is  $(1 + \text{Re}\langle x|y\rangle)/2$  (resp.  $(1 - \text{Re}\langle x|y\rangle)/2$ ). So, we can only get the value of  $\text{Re}\langle x|y\rangle$ . The image part of  $\langle x|y\rangle$  can be computed by considering the inner product of  $|x\rangle$  with  $\mathbf{i}|y\rangle$ .

The above method to estimate  $\langle x|y\rangle$  is usually called swap test [4]. Note that quantum counting [2] can also used to estimate  $\langle x|y\rangle$ . They contain the same idea. Moreover, from (2.5), we actually can obtain the following quantum state

$$\frac{1}{\sqrt{2}}(|0\rangle|x\rangle + |1\rangle|y\rangle) \left| g(\langle x|y\rangle) \right\rangle,$$

for any function  $g$ , since cosine function is even.

**Proposition 2.** *Let  $|x\rangle, |y\rangle$  be two real quantum states, except a global phase, which can be prepared in time  $O(T_{\text{in}})$ . Let  $f$  be any function. Then there is a quantum algorithm within time  $O(T_{\text{in}}/\epsilon)$  to achieve*

$$\frac{1}{\sqrt{2}}(|0\rangle|x\rangle + |1\rangle|y\rangle) \mapsto \frac{1}{\sqrt{2}}(|0\rangle|x\rangle + |1\rangle|y\rangle)|f(s)\rangle, \quad (2.6)$$

where  $|\langle x|y\rangle - s| \leq \epsilon$ .

From proposition 2, it is easy to get the following result

**Corollary 1.** *For any given quantum state  $\sum_j \alpha_j |j\rangle$  prepared in time  $O(T_{\text{in}})$  and any function  $f$ , we can obtain  $\sum_j \alpha_j |j\rangle |f(\tilde{\alpha}_j)\rangle$  in time  $O(T_{\text{in}}/\epsilon)$ , where  $|\alpha_j - \tilde{\alpha}_j| \leq \epsilon$ .*

For instance, in HHL algorithm to solve the linear system  $Ax = b$ . When we get  $|x\rangle = \sum x_i |i\rangle$ , then by the above corollary, we can change it into  $\sum x_i |i\rangle |f(x_i)\rangle$ . So we can further apply the quantum information of  $x_i$  concurrently for other problems.

### 2.1.2 Matrix multiplication algorithm by swap test

With the above preliminaries about swap test, now we can consider the matrix multiplication problem. The basic idea is similar to procedure (2.6) by putting the inner product of quantum states into another register, then like

the procedure of HHL algorithm to put this value into the coefficient. At this time, the quantum parallelism will play an important role in helping us deal with the inner product in parallel.

Denote  $C = AB$ , the target of the following quantum algorithm aim at finding the quantum information of  $C$ , that is  $|C\rangle$ . Returning classical information of matrix multiplication will be studied in the next section. Note that  $C_{ij} = A_{i\bullet}^T B_{\bullet j} = \|A_{i\bullet}\|_2 \|B_{\bullet j}\|_2 \langle A_{i\bullet} | B_{\bullet j} \rangle$ . By swap test introduced above, we can estimate  $\langle A_{i\bullet} | B_{\bullet j} \rangle$  efficiently. Together with quantum parallelism, we can get the desired quantum state  $|C\rangle$  efficiently in the following five steps:

Step 1, consider the initial state, which equals the tensor product of  $|A_{F\bullet}\rangle$  and  $|B_{\bullet F}\rangle$ :

$$\frac{1}{\|A\|_F \|B\|_F} \sum_{i,j=0}^{n-1} \|A_{i\bullet}\|_2 \|B_{\bullet j}\|_2 |i, j\rangle |0, 0\rangle.$$

Step 2, by control transformation, we can prepare  $|A_{i\bullet}\rangle$  and  $|B_{\bullet j}\rangle$  in the last register:

$$\begin{aligned} & \frac{1}{\|A\|_F \|B\|_F} \sum_{i,j=0}^{n-1} \|A_{i\bullet}\|_2 \|B_{\bullet j}\|_2 |i, j\rangle \otimes \frac{1}{\sqrt{2}} (|0\rangle |A_{i\bullet}\rangle + |1\rangle |B_{\bullet j}\rangle) \\ \mapsto & \frac{1}{\|A\|_F \|B\|_F} \sum_{i,j=0}^{n-1} \|A_{i\bullet}\|_2 \|B_{\bullet j}\|_2 |i, j\rangle \otimes \frac{1}{2} (|0\rangle (\langle A_{i\bullet} | + \langle B_{\bullet j} |) + |1\rangle (\langle A_{i\bullet} | - \langle B_{\bullet j} |)). \end{aligned} \quad (2.7)$$

Denote  $|\phi_{ij}\rangle = \frac{1}{2}(|0\rangle (\langle A_{i\bullet} | + \langle B_{\bullet j} |) + |1\rangle (\langle A_{i\bullet} | - \langle B_{\bullet j} |)) = \sin \theta_{ij} |0\rangle |u_{ij}\rangle + \cos \theta_{ij} |1\rangle |v_{ij}\rangle$ , where  $\sin^2 \theta_{ij}$  (resp.  $\cos^2 \theta_{ij}$ ) is the probability of  $|0\rangle$  (resp.  $|1\rangle$ ) and  $|u_{ij}\rangle$  (resp.  $|v_{ij}\rangle$ ) is the normalization of  $|A_{i\bullet}\rangle + |B_{\bullet j}\rangle$  (resp.  $|A_{i\bullet}\rangle - |B_{\bullet j}\rangle$ ). Also denote the eigenvalues of  $G_{ij} = (2|\phi_{ij}\rangle \langle \phi_{ij}| - I)(Z \otimes I)$  as  $e^{\pm i2\theta_{ij}}$  and the corresponding eigenvectors as  $|w_{ij}^\pm\rangle$ . Then (2.7) can be written as

$$\frac{1}{\|A\|_F \|B\|_F} \sum_{i,j=0}^{n-1} \|A_{i\bullet}\|_2 \|B_{\bullet j}\|_2 |i, j\rangle (\sin \theta_{ij} |0\rangle |u_{ij}\rangle + \cos \theta_{ij} |1\rangle |v_{ij}\rangle).$$

Step 3, perform quantum phase estimation to  $G_{ij}$  with the initial state  $(\sin \theta_{ij} |0\rangle |u_{ij}\rangle + \cos \theta_{ij} |1\rangle |v_{ij}\rangle) |0\rangle$ . Together with the control operation, we can get

$$\frac{-i}{\sqrt{2}\|A\|_F \|B\|_F} \sum_{i,j=0}^{n-1} \|A_{i\bullet}\|_2 \|B_{\bullet j}\|_2 |i, j\rangle (e^{i\theta_{ij}} |w_{ij}^+\rangle |y_{ij}\rangle - e^{-i\theta_{ij}} |w_{ij}^-\rangle |y_{ij}\rangle),$$

where  $y_{ij}\pi/2^n$  is a good approximate of  $\theta_{ij}$  to accuracy  $\epsilon$ .

Step 4, apply control rotation based on the register  $|\pm y_{ij}\rangle$ , which returns the following state

$$\begin{aligned} & \frac{-i}{\sqrt{2}\|A\|_F \|B\|_F} \sum_{i,j=0}^{n-1} \|A_{i\bullet}\|_2 \|B_{\bullet j}\|_2 |i, j\rangle (e^{i\theta_{ij}} |w_{ij}^+\rangle |y_{ij}\rangle - e^{-i\theta_{ij}} |w_{ij}^-\rangle |y_{ij}\rangle) \\ & \otimes (\langle A_{i\bullet} | B_{\bullet j} \rangle |0\rangle + \sqrt{1 - \langle A_{i\bullet} | B_{\bullet j} \rangle^2} |1\rangle). \end{aligned}$$

Step 5, undo the procedure 1-3, which yields the desired state

$$\frac{1}{\|A\|_F \|B\|_F} \sum_{i,j=0}^{n-1} \|A_{i\bullet}\|_2 \|B_{\bullet j}\|_2 \langle A_{i\bullet} | B_{\bullet j} \rangle |i, j\rangle |0\rangle + |0\rangle^\perp.$$

The next thing we need to do is estimating the error and the final complexity. This procedure is quite simple just based on triangle inequality of norm, so we put all the details in appendix A. Note that the above algorithm procedure hold for all matrices, not just square. The final result can be summarized in the following

**Theorem 1.** For any two matrices  $A, B$ , the quantum state of  $AB$  can be obtained in time  $\tilde{O}(\|A\|_F^3 \|B\|_F^3 / \|AB\|_F^3 \epsilon)$  to accuracy  $\epsilon$ .

For instance,

(1). If  $B = |b\rangle$ , then the quantum state  $|Ab\rangle$  can be obtained in time  $\tilde{O}(\|A\|_F^2 / \epsilon \|A|b\rangle\|_2^2) = \tilde{O}(n\kappa^2 / \epsilon)$ , where  $\kappa$  is the condition number of  $A$ . Actually, the result in theorem 1 is also bounded by  $\tilde{O}(n\kappa^2 / \epsilon)$ .

(2). If  $A = |a\rangle$  and  $B = \langle b|$ , then the quantum state of the rank 1 matrix  $|a\rangle\langle b|$  can be obtained in time  $\tilde{O}(1/\epsilon)$ . This result also holds when  $A, B$  are given in classical column and row vectors.

(3). For any two general matrices  $A, B$ , they can be decomposed by columns and rows, that is  $A = (A_{\bullet 0}, \dots, A_{\bullet (n-1)})$  and  $B = (B_{0\bullet}^T, \dots, B_{(n-1)\bullet}^T)^T$ . Then  $AB = \sum_j A_{\bullet j} B_{j\bullet}^T$ . Since the quantum state of  $|A_{\bullet j}\rangle\langle B_{j\bullet}|$  to accuracy  $\epsilon$ , denoted as  $|\tilde{C}_j\rangle$ , can be obtained in time  $\tilde{O}(1/\epsilon)$ . Then we just need to compute the quantum state  $|\tilde{C}\rangle$  proportional to the linear combination  $\sum_j \|A_{\bullet j}\|_2 \|B_{j\bullet}\|_2 |\tilde{C}_j\rangle$ . In [6, Chapter 26], there is a quantum algorithm to get  $|\tilde{C}\rangle$  in time

$$\tilde{O}\left(\frac{\sum_j \|A_{\bullet j}\|_2 \|B_{j\bullet}\|_2}{\epsilon \left\| \sum_j \|A_{\bullet j}\|_2 \|B_{j\bullet}\|_2 |\tilde{C}_j\rangle \right\|_2}\right) = \tilde{O}\left(\frac{A_{\bullet F} \cdot B_{F\bullet}}{\epsilon \|AB\|_F}\right),$$

where  $A_{\bullet F} = (\|A_{\bullet 0}\|_2, \dots, \|A_{\bullet (n-1)}\|_2)^T$  and  $B_{F\bullet} = (\|B_{0\bullet}\|_2, \dots, \|B_{(n-1)\bullet}\|_2)^T$  (see remark 1). The above result can also be changed only depending the Frobenius norm by Cauchy inequality into the form  $\tilde{O}(\|A\|_F \|B\|_F / \epsilon \|AB\|_F)$ , which is better than theorem 1.

**Theorem 2.** Let  $A$  be an  $l \times m$  matrix and  $B$  an  $m \times n$  matrix such that  $ln \neq 1$ , then the quantum state of  $AB$  can be obtained in time  $\tilde{O}(\|A\|_F \|B\|_F / \|AB\|_F \epsilon)$  to accuracy  $\epsilon$ .

## 2.2 By SVE and HHL

In [14], Kerenidis et al introduced a data structure, which is similar to QRAM [10], to store classical matrices in quantum computer efficiently. Based on this data structure, a fast quantum algorithm to the singular value estimated (SVE for brief, which is close to singular value decomposition) was obtained. More precisely, if  $A = \sum \sigma_i |u_i\rangle\langle v_i|$  is the singular value decomposition of  $A$ , then there is an efficient quantum algorithm to achieve  $\sum \alpha_i |v_i\rangle \mapsto \sum \alpha_i |v_i\rangle |\sigma_i\rangle$ . This algorithm is enough to solve certain problems relating to matrix operations, like multiplication or inversion. Moreover, their algorithm about SVE only applies the quantum information of  $A$ .

In this subsection, we first review the basic ideas of SVE, then we generalize their result into a quantum algorithm to achieve  $\sum \alpha_i |v_i\rangle \mapsto \sum \alpha_i |u_i\rangle |\sigma_i\rangle$ . So even when  $A$  is not Hermitian, we can also perform matrix multiplication or inversion directly only with the quantum information of  $A$ . Although, the data structure proposed in [14] lie in a model which is a little different from the standard quantum circuit model, their result about SVE only depends on the efficient preparation of some quantum states. So with the assumption given in the beginning of this section, their result also works in the standard quantum circuit model.

Let  $A = (a_{ij})_{n \times n}$  be a  $n \times n$  matrix, based on the assumption given in the beginning of this section, we know that quantum computer can perform the following mappings efficiently in time  $O(\text{poly log}(n))$ :

$$U_{\mathcal{M}} : |i\rangle|0\rangle \mapsto |i\rangle|A_{i\bullet}\rangle = \frac{1}{\|A_{i\bullet}\|} \sum_{j=0}^{n-1} a_{ij} |i, j\rangle, \quad U_{\mathcal{N}} : |0\rangle|j\rangle \mapsto |A_{F\bullet}\rangle|j\rangle = \frac{1}{\|A\|_F} \sum_{j=0}^{n-1} \|A_{i\bullet}\| |i, j\rangle.$$

*Remark 4.* The mapping  $U_{\mathcal{M}}$  and  $U_{\mathcal{N}}$  seem too perfect. Generally, the results will contain some other orthogonal parts or some errors in the results. However, by amplitude amplification, we can make it very close to the results given in the above formula. In the following, to make things simple, we just use these two mappings as [14] did.

Define two degenerate operators  $\mathcal{M}$  and  $\mathcal{N}$  as:  $\mathcal{M} : |i\rangle \mapsto |i\rangle|A_{i\bullet}\rangle$ , and  $\mathcal{N} : |j\rangle \mapsto |A_{F\bullet}\rangle|j\rangle$ . Then  $\mathcal{M}^\dagger \mathcal{N} = \frac{A}{\|A\|_F}$ . It is also easy to check that  $\mathcal{M}^\dagger \mathcal{M} = \mathcal{N}^\dagger \mathcal{N} = I_n$ . The reflections  $2\mathcal{M}\mathcal{M}^\dagger - I_{n^2}$  and  $2\mathcal{N}\mathcal{N}^\dagger - I_{n^2}$  can be efficiently

implemented in quantum computer. Denote  $W = (2\mathcal{M}\mathcal{M}^\dagger - I_{n^2})(2\mathcal{N}\mathcal{N}^\dagger - I_{n^2})$ . Let  $A = \sum_{i=0}^{n-1} \sigma_i |u_i\rangle\langle v_i|$  be the singular value decomposition of  $A$ , then

$$\begin{aligned} W\mathcal{N}|v_i\rangle &= \frac{2\sigma_i}{\|A\|_F} \mathcal{M}|u_i\rangle - \mathcal{N}|v_i\rangle; \\ W\mathcal{M}|u_i\rangle &= \left( \frac{4\sigma_i^2}{\|A\|_F^2} - 1 \right) \mathcal{M}|u_i\rangle - \frac{2\sigma_i}{\|A\|_F} \mathcal{N}|v_i\rangle. \end{aligned}$$

So the subspace  $\text{span}\{\mathcal{M}|u_i\rangle, \mathcal{N}|v_i\rangle\}$  is invariant under  $W$ . The matrix representation of  $W$  in this space is

$$W_i = \begin{pmatrix} \frac{4\sigma_i^2}{\|A\|_F^2} - 1 & \frac{2\sigma_i}{\|A\|_F} \\ -\frac{2\sigma_i}{\|A\|_F} & -1 \end{pmatrix}.$$

The eigenvalues of  $W_i$  are  $\exp(\pm i\theta_i)$  where  $\theta_i$  satisfies  $\cos\theta_i = 2\sigma_i^2/\|A\|_F^2 - 1$ . So  $\cos(\theta_i/2) = \sigma_i/\|A\|_F$ . The corresponding eigenvectors are  $w_i^\pm = -\mathcal{M}|u_i\rangle + e^{\mp i\theta_i/2} \mathcal{N}|v_i\rangle$ . It is easy to get the following decomposition

$$\mathcal{N}|v_i\rangle = \frac{1}{2i \sin(\theta_i/2)} (w^+ - w^-), \quad \mathcal{M}|u_i\rangle = \frac{1}{2i \sin(\theta_i/2)} (e^{i\theta_i/2} w^+ - e^{-i\theta_i/2} w^-).$$

With the above notations, we now can prove a more general result than [14].

**Proposition 3.** *Let  $A$  be a  $n \times n$  matrix with singular value decomposition  $A = \sum_{i=0}^{n-1} \sigma_i |u_i\rangle\langle v_i|$ . Then there is a quantum algorithm that runs in  $O(\text{poly log}(n)/\epsilon)$  and achieves  $\sum_{i=0}^{n-1} \alpha_i |v_i\rangle\langle 0| \mapsto \sum_{i=0}^{n-1} \alpha_i |u_i\rangle\langle \tilde{\sigma}_i|$ , where  $|\tilde{\sigma}_i - \sigma_i| \leq \epsilon \|A\|_F$  for all  $i$  with probability at least  $1 - 1/\text{poly}(n)$ .*

*Proof.* Denote the norm of  $w_i^\pm$  as  $m_i^\pm$ , the corresponding quantum states as  $|w_i^\pm\rangle$ . Since the eigenvalues and eigenvectors of  $W$  contain the information of singular value and singular vectors of  $A$ , these information can be obtained by performing quantum phase estimation on  $W$ . The desired procedure can be obtained from the following five steps:

Step 1, choose the initial state as  $\sum_{i=0}^{n-1} \alpha_i |v_i\rangle$ , then apply  $U_{\mathcal{N}}$  on it

$$\sum_{i=0}^{n-1} \alpha_i \mathcal{N}|v_i\rangle = \sum_{i=0}^{n-1} \frac{\alpha_i}{2i \sin(\theta_i/2)} (m_i^+ |w^+\rangle - m_i^- |w^-\rangle).$$

Step 2, perform the quantum phase estimation algorithm to estimate the eigenvalues and eigenvectors of  $W$ , then we get the following state

$$\sum_{i=0}^{n-1} \frac{\alpha_i}{2i \sin(\theta_i/2)} (m_i^+ |w^+\rangle |\theta_i\rangle - m_i^- |w^-\rangle |-\theta_i\rangle).$$

Step 3, change the phase and store the singular values in another register

$$\sum_{i=0}^{n-1} \frac{\alpha_i}{2i \sin(\theta_i/2)} (e^{i\theta_i/2} m_i^+ |w^+\rangle |\theta_i\rangle - e^{-i\theta_i/2} m_i^- |w^-\rangle |-\theta_i\rangle) |\tilde{\sigma}_i\rangle.$$

Step 4, undo the quantum phase estimation algorithm,

$$\sum_{i=0}^{n-1} \frac{\alpha_i}{2i \sin(\theta_i/2)} (e^{i\theta_i/2} m_i^+ |w^+\rangle - e^{-i\theta_i/2} m_i^- |w^-\rangle) |\tilde{\sigma}_i\rangle = \sum_i \alpha_i \mathcal{M}|u_i\rangle |\sigma_i\rangle.$$

Step 5, apply the inverse of  $U_{\mathcal{M}}$  and we will get the desired state  $\sum_{i=0}^{n-1} \alpha_i |u_i\rangle |\tilde{\sigma}_i\rangle$ . The complexity mainly comes from the quantum phase estimation, which is  $O(\text{poly log}(n)/\epsilon)$ .  $\square$



For any quantum state  $|b\rangle = \sum_i \alpha_i |v_i\rangle$ . To get the quantum information about  $A|b\rangle$ , in proposition 3, when we obtain  $\sum_i \alpha_i |u_i\rangle |\tilde{\sigma}_i\rangle$ , we can perform a controlled rotation on the register stores singular value and will get

$$\sum_{i=0}^{n-1} \alpha_i |u_i\rangle \left( \tilde{\sigma}_i t |0\rangle + \sqrt{1 - t^2 \tilde{\sigma}_i^2} |1\rangle \right), \quad (2.8)$$

where  $t = 1/\max_i \tilde{\sigma}_i$ . By choosing a suitable  $\epsilon$ , we will get a good approximate of  $|Ab\rangle$ . As for our problem of computing the quantum state of  $AB$ , we can choose the initial state as  $|B\rangle$ , and implement the above procedure in parallel in each column. Finally by a simple analysis about the error and complexity (details are given in appendix B), we will get the following result

**Theorem 3.** *The quantum algorithm to get the quantum state of  $AB$  to precision  $\epsilon$  costs  $\tilde{O}(\|A\|_F \|B\|_F \kappa^2 / \epsilon \|AB\|_F)$  by SVE.*

*Remark 5.* If  $B = |b\rangle$ , which only contains one column, then the complexity to obtain  $|Ab\rangle$  is  $\tilde{O}(\|A\|_F \kappa^2 \epsilon \|A|b\rangle\|_2) = \tilde{O}(\sqrt{n} \kappa^3 / \epsilon)$ . However, to get a good approximate of  $A|b\rangle$  without normalization, the procedure may not so expensive. Actually, from (2.8), we see that the error between  $\sum_{i=0}^{n-1} \alpha_i \tilde{\sigma}_i |u_i\rangle$  and  $A|b\rangle$  is bounded by  $\epsilon \|A\|_F$ . So we just need to choose  $\epsilon \|A\|_F = \epsilon_1$  small. Then the complexity to get  $A|b\rangle$  is  $\tilde{O}(\|A\|_F / \epsilon_1)$ .

The quantum algorithm to get the quantum state of  $AB$  by HHL is similar to the algorithm by SVE. We now assume that  $A$  is Hermitian, otherwise we can consider  $\tilde{A} = \begin{pmatrix} 0 & A \\ A^\dagger & 0 \end{pmatrix}$ . We also assume that the Hamiltonian simulation of  $e^{-i\tilde{A}t}$  is efficiently. For any quantum state  $|b\rangle = \sum \alpha_j |u_j\rangle$ , by HHL algorithm, we can get the state  $\alpha_j |u_j\rangle |\tilde{\sigma}\rangle$ , where  $|\tilde{\sigma}_i - \sigma_i| \leq \epsilon \max_i \sigma_i$ . Similarly, we can get (2.8). Compared to the singular value estimation to achieve the quantum state of  $|AB\rangle$ , the only change is  $\|A\|_F$ , now it becomes  $\max_i \sigma_i$ . So by HHL algorithm, we can get the quantum state of  $AB$  in time  $\tilde{O}(\kappa^3 / \epsilon)$  to accuracy  $\epsilon$ , since  $\|AB\|_F \geq \|B\|_F \min_i \sigma_i$ .

*Remark 6.* Just like HHL algorithm to solve linear system, the methods based on SVE and HHL also have the problem. More precisely, we potentially assume that each  $|B_{\bullet,j}\rangle$  lies in the nonzero components of  $A$ , i.e., the space generated by singular vectors with nonzero singular values. Otherwise, the success probability and so the final complexity will be affected. For instance, in (2.8), the success probability is  $P = \sum_{i, \tilde{\sigma}_i \neq 0} |\alpha_i \tilde{\sigma}_i t|^2$ . If we assume that  $|b\rangle$  lies in the nonzero components, then  $P \geq \min_{i, \tilde{\sigma}_i \neq 0} |\tilde{\sigma}_i t|^2 \sum_{i, \tilde{\sigma}_i \neq 0} |\alpha_i|^2 = 1/\kappa^2$ , since  $\sum_{i, \tilde{\sigma}_i \neq 0} |\alpha_i|^2 = 1$ . If there exists  $i$  such that  $\tilde{\sigma}_i \neq 0$  but  $\alpha_i = 0$ , then  $P \geq 1/\kappa^2$  may not hold anymore. So in these two algorithms, we should make this as another assumption. However, the method based on swap test do not contain such a problem.

The following table is a summary about the three quantum algorithms proposed in this section to achieve the quantum data of the multiplication of two matrices.

**Table 1.** Comparison of different quantum algorithms to achieve matrix multiplication with quantum information, where  $A, B$  are input matrices,  $\kappa$  is the condition number of  $A$ .

Methods	Complexity	Assumptions
By swap test	$\tilde{O}(\ A\ _F \ B\ _F / \epsilon \ AB\ _F) = \tilde{O}(\sqrt{n} \kappa / \epsilon)$	Efficient preparation of quantum states of $A$ and $B$
By SVE	$\tilde{O}(\ A\ _F \ B\ _F \kappa^2 / \epsilon \ AB\ _F) = \tilde{O}(\sqrt{n} \kappa^3 / \epsilon)$	
By HHL	$\tilde{O}(\kappa^3 / \epsilon)$	Efficient preparation of quantum states of $B$ and efficient Hamiltonian simulation of $\tilde{A}$

Note that swap test and SVE work for all the cases if  $A, B$  are classical or quantum data, while HHL needs  $A$  or  $B$  to be classical. All the results are related to the condition number in the worst case. Simple analysis shows



that  $\sqrt{n}/\kappa \leq \|A\|_F \|B\|_F / \|AB\|_F \leq \sqrt{n}\kappa$ , so if  $\kappa = O(\text{poly log } n)$ , then  $\|A\|_F \|B\|_F / \|AB\|_F = \tilde{O}(\sqrt{n})$ , however, the result by HHL is the best in this case. On the contrary, the complexities are not easy to determine approximately if  $\kappa = \tilde{O}(n^c)$  for some constant  $c$ .

### 3 Quantum algorithm to matrix multiplication with classical information

In this section, we consider the problem of getting classical data to the multiplication of two matrices. First, we focus on the analysis of the method based on swap test. The quantum algorithms based on SVE or HHL are similar to analyze.

Let  $\mathbf{x}, \mathbf{y}$  are two  $n$  dimensional vectors. Denote the corresponding quantum states of these two vectors as  $|x\rangle, |y\rangle$ . Then  $\mathbf{x} \cdot \mathbf{y} = \|\mathbf{x}\|_2 \|\mathbf{y}\|_2 \langle x|y\rangle$ , where  $\|\mathbf{x}\|_2, \|\mathbf{y}\|_2$  are the norms of  $\mathbf{x}, \mathbf{y}$ . By proposition 1, we can get a good approximate of  $\langle x|y\rangle$ , i.e., we can get a value  $P_{xy}$  in time  $O(T_{\text{in}}/\epsilon)$  such that  $|P_{xy} - \langle x|y\rangle| \leq \epsilon$ . However, a good approximate of  $\langle x|y\rangle$  does not imply a good approximate of  $\mathbf{x} \cdot \mathbf{y}$ . This is because  $|P_{xy} \|\mathbf{x}\|_2 \|\mathbf{y}\|_2 - \mathbf{x} \cdot \mathbf{y}| \leq \|\mathbf{x}\|_2 \|\mathbf{y}\|_2 \epsilon$ . In order to make this error small, we denote  $\|\mathbf{x}\|_2 \|\mathbf{y}\|_2 \epsilon = \tilde{\epsilon}$ , then the final complexity of estimating  $\mathbf{x} \cdot \mathbf{y}$  becomes

$$O(T_{\text{in}} \|\mathbf{x}\|_2 \|\mathbf{y}\|_2 / \tilde{\epsilon}). \quad (3.1)$$

Here we did not considered the complexity of evaluating  $\|\mathbf{x}\|_2, \|\mathbf{y}\|_2$ . From the above analysis, we conclude that proposition 1 solves the inner product problem of two classical vectors efficiently only if the norms of the vectors are small.

*Remark 7.* We should remark that the influence of norms  $\|\mathbf{x}\|_2, \|\mathbf{y}\|_2$  on the complexity (3.1) by swap test to estimate inner product of  $\mathbf{x}$  and  $\mathbf{y}$  cannot removed actually. This is all because of the optimality of Grover searching algorithm. Consider the searching problem in  $\mathbb{Z}_n$ . Assume there are  $r$  marked items, and  $f : \mathbb{Z}_n \rightarrow \mathbb{Z}_2$  is defined as  $f(i) = 1$  if and only if  $i$  is marked. Now we define  $g(x) = (-1)^{f(x)}$ . Denote  $\mathbf{x} = (g(0), \dots, g(n-1))$  and  $\mathbf{y} = (1, \dots, 1)$ . Then  $\|\mathbf{x}\|_2 = \|\mathbf{y}\|_2 = \sqrt{n}$ , and the quantum state  $|x\rangle, |y\rangle$  can be prepared efficiently. As we can see  $\mathbf{x} \cdot \mathbf{y} = \sum_{x=0}^{n-1} g(x) = n - 2r$ . Suppose the complexity of evaluating the inner product of  $\mathbf{x} \cdot \mathbf{y}$  is independent of  $\|\mathbf{x}\|_2, \|\mathbf{y}\|_2$  and can be improved into  $\tilde{O}(1/\tilde{\epsilon})$ , then we can decide whether or not there exist marked items in  $\mathbb{Z}_n$  efficiently, since we can just choose  $\tilde{\epsilon} = O(1)$ . Together with the bisection method, we can finally find one marked item if  $r > 0$  efficiently. This will contradict the optimality of Grover searching algorithm.

Let  $A, B$  be two  $n \times n$  matrices. Multiplying  $A$  and  $B$  is equivalent to evaluate  $n^2$  inner product of  $n$  dimensional vectors. Classical method to evaluate inner product of two  $n$  dimensional vectors takes time  $O(n)$ , which lead the complexity of the classical matrix multiplication to  $O(n^3)$ . However, swap test may reduce the complexity of evaluating inner product and so may reduce the complexity of matrix multiplication. The norms of  $A_{i\bullet}, B_{\bullet j}$  can be evaluated by the classical method ( $0 \leq i, j \leq n-1$ ), which costs  $O(n^2)$ . These are classical data, and so can be used as many times as we want. Since we assume that the quantum states of  $A_{i\bullet}, B_{\bullet j}$  can be prepared efficiently. Then by (3.1) and note that  $A_{F\bullet}, B_{\bullet F}$  are column vectors store the information of the 2 norms  $A_{i\bullet}, B_{\bullet j}$  (see remark 1), we have

**Theorem 4.** *There is a quantum algorithm that computes the multiplication of  $A$  and  $B$  with classical information in time  $\tilde{O}(\|A_{F\bullet}\|_1 \|B_{\bullet F}\|_1 / \epsilon + n^2)$  to accuracy  $\epsilon$ .*

*Proof.* From (3.1), we know that the complexity to multiply  $A, B$  with classical data is  $\sum_{i,j} \|A_{i\bullet}\|_2 \|B_{\bullet j}\|_2 / \epsilon = \|A_{F\bullet}\|_1 \|B_{\bullet F}\|_1 / \epsilon$ . Together with  $O(n^2)$  to compute the corresponding norms of  $A_{i\bullet}, B_{\bullet j}$ , we will get the desired result.  $\square$

The accuracy  $\epsilon$  in the theorem means that if  $C = AB = (c_{ij})$  is the exact result and  $\tilde{C} = (\tilde{c}_{ij})$  is the result obtained from the quantum algorithm, then  $|c_{ij} - \tilde{c}_{ij}| \leq \epsilon$ . This is the absolute error. It is not easy to do the analysis of relative error now, since we have no more information about the value of inner product  $\mathbf{x} \cdot \mathbf{y}$ . However, by choosing the absolute error relatively smaller than the  $\|\mathbf{x}\|_2 \|\mathbf{y}\|_2$ , the absolute error becomes closer to relative error.

Next, we consider the algorithms based on SVE and HHL. In order to compute the classical information about  $AB$ , we actually do not need to perform measurements in the quantum algorithm by SVE or HHL. As discussed in remark 5, if  $|B_j\rangle = \sum_k \alpha_{jk} |u_k\rangle$ , then (2.8) can be written as

$$\frac{1}{\max_k \sigma_k} A|B_j\rangle|0\rangle + \text{orthogonal part.}$$

The above quantum state can be obtained efficiently in time  $\tilde{O}(1/\|A\|_F \epsilon)$ . By applying swap test on the above state with  $|i, 0\rangle$ , we will get an approximate about the entries of  $AB$ . More details about the analysis of error and complexity, which are not so difficult, are given in appendix C. The final result is

**Theorem 5.** *There is a quantum algorithm that computes the multiplication of  $A$  and  $B$  with classical information in time  $\tilde{O}(\sqrt{n}\kappa\|A_{F\bullet}\|^2\|B_{\bullet F}\|_3^3/\epsilon^2 + n^2)$  by SVE and  $\tilde{O}(\kappa^2\|A\|_F^2\|B_{\bullet F}\|_3^3/\epsilon_3^2 + n^2)$  by HHL algorithm to accuracy  $\epsilon$ .*

The following table summarizes the above results about quantum matrix multiplication with classical data. Different from the algorithm to obtain quantum information, the complexity now is independent of  $\|AB\|_F$ . This is because that no measurements are needed in evaluating the classical data.

**Table 2.** Comparison of different quantum algorithms to achieve matrix multiplication with classical information, where  $A, B$  are input matrices,  $\kappa$  is the condition number of  $A$ .

Methods	Complexity	Assumptions
By swap test	$\tilde{O}(\ A_{F\bullet}\ _1\ B_{\bullet F}\ _1/\epsilon + n^2) = \tilde{O}(n\ A\ _F\ B\ _F/\epsilon + n^2)$	The same as table 1
By SVE	$\tilde{O}(\sqrt{n}\kappa\ A\ _F^2\ B_{\bullet F}\ _3^3/\epsilon^2 + n^2) = \tilde{O}(\sqrt{n}\kappa\ A\ _F^2\ B\ _F^{3/2}/\epsilon^2 + n^2)$	
By HHL	$\tilde{O}(\kappa^2\ A\ _F^2\ B_{\bullet F}\ _3^3/\epsilon^2 + n^2) = \tilde{O}(\kappa^2\ A\ _F^2\ B\ _F^{3/2}/\epsilon^2 + n^2)$	

To make the above table more easy to understand and easy to compare with classical algorithms, we assume that the singular values of  $A, B$  are smaller than 1, then  $\|A\|_F, \|B\|_F \leq \sqrt{n}$ . So the above table can be simplified into

**Table 3.** Comparison of different quantum algorithms to achieve matrix multiplication with classical information, where  $A, B$  are input matrices,  $\kappa$  is the condition number of  $A$  and the singular values of  $A, B$  are smaller than 1.

Methods	Complexity	Assumptions
By swap test	$\tilde{O}(n^2/\epsilon)$	Assumptions in table 1 and the singular values of $A, B$ are smaller than 1
By SVE	$\tilde{O}(\kappa n^{2.25}/\epsilon^2)$	
By HHL	$\tilde{O}(\kappa^2 n^{1.75}/\epsilon^2 + n^2)$	

The best classical algorithm to matrix multiplication with complexity  $O(n^{2.3728639})$  is due to Le Gall [16] at 2014. If the precision  $\epsilon$  is small in size  $O(1/\text{poly log } n)$ , then the quantum algorithm to matrix multiplication based on swap test is  $\tilde{O}(n^2)$ . Also, to make this quantum algorithm better than Le Gall's classical algorithm, the upper bound of  $\epsilon$  is  $O(n^{0.3728639}/\text{poly log } n)$ . The quantum algorithm based on SVE works better than the classical algorithm only if  $\kappa/\epsilon^2 = \tilde{O}(n^{0.1228639})$  and the quantum algorithm based on HHL works better only if  $\kappa^2/\epsilon^2 = \tilde{O}(n^{0.6228639})$ . In [5], Buhrman et al also proposed a quantum algorithm to achieve matrix multiplication, however, their complexity depends on the number of nonzero entries of  $AB$ , so we prefer not to compare with it here.

## 4 Preparation of quantum states

Let  $\mathbf{x} = (x_0, \dots, x_{n-1})$  be a complex vector, the quantum state it corresponds to equals  $|x\rangle = \frac{1}{\|\mathbf{x}\|_2} \sum_{i=0}^{n-1} x_i |i\rangle$ . The transformation from classical data  $\mathbf{x}$  into its quantum state is usually called “input problem” in quantum computer, which plays important roles in many quantum algorithms, such as [7], [8], [13], [14], [15], [17], [21], [22], [23], [27], [28], [30]. Moreover, the quantum matrix multiplication algorithms proposed above also rely on efficient preparation of quantum states. This section is devoted to study the input problem.

The most naive method is defining a unitary  $U$  such that  $U|0\rangle = |x\rangle$ . The efficiency of preparing  $|x\rangle$  is totally determined by  $U$ . In the worst case,  $U$  can be implemented in time  $O(n^2(\log n)^2 \log^c(n^2(\log n)^2/\epsilon))$  (see [20, Chapter 4]) to precision  $\epsilon$  in quantum computer, where  $c$  is some constant close to 2. So we can prepare  $|x\rangle$  within the same time in the worst case. Conclude this, we have

**Proposition 4.** *For any vector  $\mathbf{x}$ , its quantum state can be prepared in time  $O(n^2(\log n)^2 \log^c(n^2(\log n)^2/\epsilon))$  to precision  $\epsilon$ .*

Although the above method works for all cases, it is not efficient generally. We still hope the input problem can be solved efficiently in some special cases. Under certain conditions, this problem can actually solved efficiently [8], [12], [17], [26]. In the following, we focus on the one given in [17]. In this paper, Lloyd et al provided a quantum algorithm to prepare quantum state, which works very well when the given classical data are relatively uniform distributed. It has been used to solve the supervised classification problem [17] and the least square support vector machine problem [21]. In the following, we first give a detailed analysis about this technique in order to find its advantages and disadvantages. Then based on this algorithm, we will propose two new quantum algorithms with better efficiency.

Let  $f$  be a map from  $\mathbb{Z}_n$  to  $\mathbb{R}$ , denote

$$\max(f) := \max_{k \in \mathbb{Z}_n} |f(k)|; \quad \min(f) := \min_{k \in \mathbb{Z}_n, f(k) \neq 0} |f(k)|; \quad \kappa(f) := \max(f)/\min(f).$$

These notations can be similarly extended to vectors or sequences. We should remark that in [17], the authors’s main objective is solving the supervised classification problem, so their results and methods are confined to this problem. Moreover, they did not give too much analysis about the efficiency of their method. However, their method is more general then preparing quantum states. The following result is obtained by generalizing their method.

**Proposition 5.** *Let  $f$  be a map (or an oracle) from  $\mathbb{Z}_n$  to  $\mathbb{R}^* = \mathbb{R} \setminus \{0\}$ . Then for any state  $|\psi\rangle = \sum_{k=0}^{n-1} b_k |k\rangle$  with preparation complexity  $O(T_{|\psi\rangle})$ , we can construct the state  $|\psi'\rangle = \frac{1}{\sqrt{Z}} \sum_{k=0}^{n-1} f(k) b_k |k\rangle$  in time  $O(\kappa(f)^{3/2} T_{|\psi\rangle}/\epsilon)$  to accuracy  $\epsilon$ , where  $Z = \sum_{k=0}^{n-1} |f(k) b_k|^2$ .*

*Proof.* Let  $H = \sum_{k=0}^{n-1} f(k) |k\rangle\langle k|$  be a Hamiltonian, which is a diagonal matrix, so  $e^{-iHt}$  can be implemented efficiently. Consider the following procedure:

$$\begin{aligned} \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \sum_{k=0}^{n-1} b_k |k\rangle &\mapsto \frac{1}{\sqrt{2}}|0\rangle \sum_{k=0}^{n-1} b_k e^{if(k)t} |k\rangle + \frac{1}{\sqrt{2}}|1\rangle \sum_{k=0}^{n-1} b_k e^{-if(k)t} |k\rangle \\ &\mapsto \frac{1}{\sqrt{2}}|+\rangle \sum_{k=0}^{n-1} b_k e^{if(k)t} |k\rangle + \frac{1}{\sqrt{2}}|-\rangle \sum_{k=0}^{n-1} b_k e^{-if(k)t} |k\rangle \\ &= |0\rangle \sum_{k=0}^{n-1} b_k \cos(f(k)t) |k\rangle + i|1\rangle \sum_{k=0}^{n-1} b_k \sin(f(k)t) |k\rangle, \end{aligned} \tag{4.1}$$

where the first step is the result of Hamiltonian simulation and the second step applies Hadamard transformation on the first qubit.

Choosing  $t$  small enough such that there exist  $\epsilon_0, \epsilon_1$  satisfy  $\epsilon_0 \leq |f(k)t| \leq \epsilon_1 \ll 1$ , here we choose  $\epsilon_1 = O(\kappa(f)^{-1/2}\epsilon)$ . Then the state along with  $|1\rangle$  is an approximation of the state  $|\psi'\rangle$ . And the error between them is

bounded by  $O(\epsilon)$  (see more details in appendix D). Note that  $f(k) \neq 0$  for all  $k$ , so the probability of getting  $|1\rangle$  is  $P \approx Zt^2 = \sum_{k=0}^{n-1} |b_k f(k)t|^2 \geq \epsilon_0^2$ . By amplitude amplification technique, it suffices measuring  $O(1/\epsilon_0)$  times. Note that  $\epsilon_0 \leq |f(k)t| \leq \epsilon_1$ , so we have  $\epsilon_0 \approx \min(f)t$  and  $\epsilon_1 \approx \max(f)t$ . Hence  $\epsilon_0 = \kappa(f)^{-1}\epsilon_1 = O(\kappa(f)^{-3/2}\epsilon)$ . The complexity of procedure (4.1) is  $O(T_{|\psi\rangle})$ , so the complexity of getting  $|\psi'\rangle$  is  $O(T_{|\psi\rangle}/\epsilon_0) = O(\kappa(f)^{3/2}T_{|\psi\rangle}/\epsilon)$ .  $\square$

The final complexity of the algorithm given in proposition 5 is affected by  $\kappa(f)$ . A simple case is when the sequence  $\{|f(0)|, \dots, |f(n-1)|\}$  is relatively uniform. Here relatively uniform means  $\kappa(f) = \max(f)/\min(f)$  is an acceptable small constant. In this case, the complexity of proposition 5 can be further simplified into  $O(T_{|\psi\rangle}/\epsilon)$ .

In [8], Clader et al also propose a quantum algorithm to prepare quantum state, which is indirectly inspired by the work of HHL algorithm [13]. Simple analysis shows that their algorithm also contains the same problem discussed above, that is the influence of  $\kappa(f)$ . However, their algorithm is quite simple and contains no error. Note that the idea of proposition 5 can be generalized into the complex field by considering the real and image part respectively, so in the following we just need to focus on the preparation of real vectors. Moreover, if there exists  $k$  such that  $f(k) = 0$ , then the probability analysis  $P = \sum_{k=0}^{n-1} |b_k f(k)t|^2 \geq \epsilon_0^2$  may not hold anymore. However, if  $|\psi\rangle = \frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} |k\rangle$  and the sequence  $\{|f(0)|, \dots, |f(n-1)|\}$  contains  $O(z)$  nonzero elements, then the complexity of proposition 5 should be multiplied by  $O(\sqrt{n/z})$  because of amplitude amplification technique. A direct application of proposition 5 is

**Proposition 6.** *Let  $\mathbf{x} = (x_0, \dots, x_{n-1})$  be a given vector that contains  $O(z)$  nonzero elements, denote  $\kappa(\mathbf{x}) = \max_k |x_k|/\min_{k, x_k \neq 0} |x_k|$ , then  $|x\rangle = \frac{1}{\|\mathbf{x}\|_2} \sum_{k=0}^{n-1} x_k |k\rangle$  can be prepared in time  $O(\kappa(\mathbf{x})^{3/2} \sqrt{n/z} (\log n)/\epsilon)$  to accuracy  $\epsilon$ . Moreover, if  $\mathbf{x}$  is relatively uniform, then the complexity is  $O(\sqrt{n/z} (\log n)/\epsilon)$ .*

*Proof.* Just choose  $|\psi\rangle = \frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} |k\rangle$  and  $f(k) = x_k$  in proposition 5.  $\square$

However, if we know the positions of nonzero entries in  $\mathbf{x}$ , then we can just focus on these nonzero parts and apply the algorithm given in proposition 5 to prepare the quantum state of the nonzero components of  $\mathbf{x}$ , which is also equals to the quantum state of  $\mathbf{x}$  itself. So we have

**Proposition 7.** *For any given vector  $\mathbf{x}$ , there is a quantum algorithm to prepare its quantum state in time  $O(\kappa(\mathbf{x})^{3/2} (\log n)\epsilon)$ . Moreover, if  $\mathbf{x}$  is relatively uniform, then the complexity is  $O((\log n)/\epsilon)$ .*

As we can see from the above result, the quantum state preparation algorithm works efficient when  $\mathbf{x}$  is a relatively uniform distributed vector, and may performs very bad otherwise. One way to grasp the property of relatively uniform distribution is decomposing  $\mathbf{x}$  into the sum of several relatively uniform distributed vectors. In the following, we given two different such decompositions.

Let  $\mathbf{x} = (x_0, \dots, x_{n-1})$  be a real vector. Assume that  $0 < x_0 < \dots < x_{n-1}$  for simplicity. Denote  $\kappa(\mathbf{x}) = x_{n-1}/x_0 \approx 2^q$ . Then each interval  $I_j = [2^{j-1}x_0, 2^jx_0]$ , where  $1 \leq j \leq q$ , contains several values of  $x_i$ , we denote them as  $x_{j1}, \dots, x_{jt_j}$  and set the vector  $\mathbf{y}_j = (0, \dots, 0, x_{j1}, \dots, x_{jt_j}, 0, \dots, 0)$ . Then  $\mathbf{x} = \mathbf{y}_1 + \dots + \mathbf{y}_q$  and each vector  $\mathbf{y}_j$  can be prepared efficiently in time  $O((\log n)/\epsilon)$  by proposition 7. Now we see that  $|x\rangle = \lambda_1 |y_1\rangle + \dots + \lambda_q |y_q\rangle$ , where  $\lambda_j = \|\mathbf{y}_j\|_2 / \|\mathbf{x}\|_2$ . From the method given in [6, Chapter 26], the complexity to achieve such a linear combination to get  $|x\rangle$  equals  $O(C_q (\log n) \sum_{j=1}^q \|\mathbf{y}_j\|_2 / \|\mathbf{x}\|_2 \epsilon) = O(q^{5/2} (\log q)^2 \log^c(q^2 (\log q)^2 / \epsilon) (\log n)/\epsilon)$ , where  $C_q$  is the complexity to implement the unitary  $U$  such that  $U|0\rangle \propto \sum \sqrt{\lambda_j} |j\rangle$ , which is at most  $O(q^2 (\log q)^2 \log^c(q^2 (\log q)^2 / \epsilon))$  as discussed in [20, Chapter 4]. When the entries of  $\mathbf{x}$  are not all positive and increasing, we define the interval  $I_j$  based on the vector whose entries are the absolute value of the entries of  $\mathbf{x}$  and nonzero. The requirement that all entries are sorted is not so necessary, since the above analysis are also hold for the case that  $\mathbf{x}$  is not sorted, we only need to focus on positions of vectors lie in  $I_j$ . At this case, the notations will be a little complicate, but it changes nothing. Therefore, we have

**Theorem 6.** *Let  $\mathbf{x} = (x_0, \dots, x_{n-1})$  be a given vector and  $\kappa(\mathbf{x}) = \max_k |x_k|/\min_{k, x_k \neq 0} |x_k|$ . Then its quantum state can be prepared in time  $O((\log \kappa(\mathbf{x}))^{5/2} (\log \log \kappa(\mathbf{x}))^2 \log^c[(\log \kappa(\mathbf{x}))^2 (\log \log \kappa(\mathbf{x}))^2 / \epsilon] (\log n)/\epsilon)$ .*

This result is better than proposition 7. If  $\kappa(\mathbf{x})$  is too large, then we may consider giving up the components that are close to  $\min_{k, x_k \neq 0} |x_k|$  if they are not too many. Moreover, even if  $\max |x_k| = 2^{1000}$  and  $\min_{k, x_k \neq 0} |x_k| = 1$ ,

then  $\log \kappa(\mathbf{x}) = 1000$ , which is still an acceptable constant. From this point, the above result which is polynomially depending on  $\log \kappa(\mathbf{x})$  seems to be a pretty good algorithm to prepare quantum states.

Another decomposition is more direct and easy. We assume that all entries of  $\mathbf{x}$  are nonzero, otherwise, we only focus on the nonzero components. Now define  $\mathbf{y} = M(\text{sign}(x_0), \dots, \text{sign}(x_{n-1}))$ , where  $M \geq \max_i |x_i|$  and  $\text{sign}(x_i) = 1$  if  $x_i \geq 0$ ;  $\text{sign}(x_i) = -1$  if  $x_i < 0$ . Then the quantum state  $|y\rangle$  of  $\mathbf{y}$  can be obtained efficiently in time  $O(\log n)$ . Also define  $\mathbf{z} := \mathbf{x} + \mathbf{y} = (\text{sign}(x_0)M + x_0, \dots, \text{sign}(x_{n-1})M + x_{n-1})$ , which is uniformly distributed with  $n$  nonzero entries. So by proposition 7, the quantum state  $|z\rangle$  of  $\mathbf{z}$  can be obtained efficiently in time  $O((\log n)/\epsilon)$  to precision  $\epsilon$ . Since

$$|x\rangle = \frac{1}{\|\mathbf{x}\|_2}(\mathbf{z} - \mathbf{y}) = \frac{\|\mathbf{z}\|_2}{\|\mathbf{x}\|_2}|z\rangle - \frac{\|\mathbf{y}\|_2}{\|\mathbf{x}\|_2}|y\rangle.$$

What we should do next is computing the linear combination of two efficiently prepared quantum states. The linear combination of two quantum states can be obtained from a similar procedure to Hadamard test as follows: here for simplicity denote  $\|\mathbf{z}\|_2/\|\mathbf{x}\|_2$  as  $\lambda$ , denote  $\|\mathbf{y}\|_2/\|\mathbf{x}\|_2$  as  $\mu$ ,

$$\begin{aligned} \frac{1}{\sqrt{\lambda^2 + \mu^2}}(\lambda|0\rangle|z\rangle + \mu|1\rangle|y\rangle) &\mapsto \frac{1}{\sqrt{\lambda^2 + \mu^2}}(\lambda|+\rangle|z\rangle + \mu|-\rangle|y\rangle) \\ &= \frac{1}{\sqrt{2(\lambda^2 + \mu^2)}}\left[|0\rangle(\lambda|z\rangle + \mu|y\rangle) + |1\rangle(\lambda|z\rangle - \mu|y\rangle)\right]. \end{aligned}$$

The probability to get  $|x\rangle = \lambda|z\rangle - \mu|y\rangle$  is  $1/2(\lambda^2 + \mu^2) = \|\mathbf{x}\|_2^2/2(\|\mathbf{y}\|_2^2 + \|\mathbf{z}\|_2^2)$ . Then the complexity to get  $|x\rangle$  is  $O(\sqrt{(\|\mathbf{y}\|_2^2 + \|\mathbf{z}\|_2^2)/\|\mathbf{x}\|_2^2}(\log n)/\epsilon)$ . Set

$$d = \frac{\max_k (|\text{sign}(x_k)M + x_k|)}{\min_k (|\text{sign}(x_k)M + x_k|)} = \frac{M + \max_k |x_k|}{M + \min_k |x_k|}.$$

If  $d$  is a small constant, then we can just choose  $M = \max_k |x_k|$ . Note that  $\|\mathbf{x}\|_2 \geq n \min_k |x_k|$ , so

$$\frac{\|\mathbf{y}\|_2^2 + \|\mathbf{z}\|_2^2}{\|\mathbf{x}\|_2^2} = \frac{2nM^2 + 2M \sum_i x_i + \|\mathbf{x}\|_2^2}{\|\mathbf{x}\|_2^2} \leq 3\kappa(\mathbf{x})^2 + 1.$$

Hence, the complexity obtained by this decomposition is  $O(\kappa(\mathbf{x})(\log n)/\epsilon)$ , which is better by applying proposition 7 directly, however, not efficient than theorem 6. In [25], the author propose a method to achieve the linear combination, which is independent of the effect of  $\lambda$  and  $\mu$ , in time  $O(\log(n)/\epsilon^2)$ . Therefore, we have

**Theorem 7.** *For any vector  $\mathbf{x}$ , its quantum state can be prepared in time  $O(\log(n)/\epsilon^2)$  to precision  $\epsilon$ .*

In the following, we summarize all the quantum algorithms proposed above in the following table. Note that they work for all classical data, so no other assumptions are needed. However, they need a lot of information about the input data, such as the nonzero components, the maximal and minimal entries. It may take some extra time to get them (such as by quantum searching algorithm [9], [11]) and we will not consider it right here.

**Table 4.** Comparison of different quantum algorithms to prepare quantum state of  $\mathbf{x} = (x_0, \dots, x_{n-1})$ , where  $\kappa(\mathbf{x}) = \max_k |x_k| / \min_{k, x_k \neq 0} |x_k|$  and  $\epsilon$  is the precision.

Algorithms given in	Complexity
Proposition 4	$O(n \log(n) \log(1/\epsilon))$
Proposition 7	$O(\kappa(\mathbf{x})^{3/2}(\log n)/\epsilon)$
Theorem 6	$O((\log \kappa(\mathbf{x}))^{5/2}(\log \log \kappa(\mathbf{x}))^2 \log^c[(\log \kappa(\mathbf{x}))^2(\log \log \kappa(\mathbf{x}))^2/\epsilon](\log n)/\epsilon)$
Theorem 7	$O(\log(n)/\epsilon^2)$

In the quantum algorithms to achieve matrix multiplication with classical data, we can first apply searching algorithm to find the desired information to prepare quantum states, which takes at most  $O(n^2)$  steps and does not affect the final complexity of the algorithms. This means, to getting the classical data of matrix multiplication, quantum algorithms listed in table 2, except HHL, do not need the assumption listed in the beginning of section 2. Strictly speaking, actually only the quantum algorithm obtained by swap test do not need any assumptions by the analysis in remark 6. Therefore, we have

**Theorem 8.** *The multiplication of two  $n \times n$  matrices can be obtained in time  $\tilde{O}(n^2/\epsilon)$  to precision  $\epsilon$ .*

## 5 Conclusions

Quantum computer outperforms the classical computer in many problems. However, many of the quantum algorithms make one or two assumptions; the most common one is the input problem, that is we assume the given data is quantum data by some methods like QRAM. However, this problem is not easy to solve generally. As suggested in [3], we can study quantum algorithms only with quantum input and output. With this idea, we do not need to consider the input and output problem. One important task worth to study is extending classical matrix operations into quantum computer with quantum input and output data. In this paper, we only considered the problem of matrix multiplication, however, it forms the most elementary step of many other matrix operations, such as QR decomposition and LU decomposition. Until now, most matrix operations are not easy to find suitable quantum techniques to deal with them efficiently. QR decomposition is very useful, however, it seems quit difficult to make it efficient in quantum computer in polynomial time. Fortunately, we already have SVE technique and its generalized version. This will play important roles in studying quantum matrix operations. The generalized version of swap test can be viewed as another technique that we can apply.

**Acknowledgements.** This work is supported by the NSFC Project 11671388 and the CAS Frontier Key Project QYZDJ-SSW-SYS022.

## References

1. Aaronson, S., Quantum Machine Learning Algorithms: Read the Fine Print, *Nature Physics* 11(4): 291-293, 2015.
2. Brassard, G., Høyer, P., Tapp, A., Quantum Counting, *25th Intl. Colloquium on Automata, Languages, and Programming (ICALP)*, LNCS 1443, pp. 820-831, 1998.
3. Biamonte, J., Wittek, P., Pancotti, N., Rebentrost, P., Wiebe, N. and Lloyd, S., Quantum machine learning, *Nature* 549, 195-202, 2017.
4. Buhrman, H., Cleve, R., Watrous, J. and Wolf, R. de, Quantum Fingerprinting, *Phys. Rev. Lett.* 87, pp. 167902, 2001.
5. Buhrman, H. and Špalek, R., Quantum verification of matrix products. In *Proceedings of the 17th ACM-SIAM Symposium on Discrete Algorithms*, pp. 880-889, 2006.
6. Childs, A.M., *Lecture Notes on Quantum Algorithms*, <http://www.cs.umd.edu/~amchilds/qa/>, 2017.
7. Childs, A.M., Kothari, R., and Somma, R.D., Quantum algorithm for systems of linear equations with exponentially improved dependence on precision, *SIAM Journal on Computing* 46, 1920-1950, 2017.
8. Clader, B.D., Jacobs, B.C. and Sprouse, C.R., Preconditioned quantum linear system algorithm, *Phys. Rev. Lett.* 110, pp. 250504, 2013.
9. Dürr, C. and Høyer, P., *A quantum algorithm for finding the minimum*, arXiv:quant-qh9607014, 1996.
10. Giovannetti, V., Lloyd, S. and Maccone, L., Quantum random access memory. *Phys. Rev. Lett.*, 100, pp. 160501, 2008.
11. Grover, L.K., A fast quantum mechanical algorithm for database search, *Proceedings, 28th Annual ACM Symposium on the Theory of Computing (STOC)*, 212-219, 1996.
12. Grover, L. and Rudolph, T., *Creating superpositions that correspond to efficiently integrable probability distributions*. arXiv:quant-ph/0208112, 2002.
13. Harrow, A.W., Hassidim, A. and Lloyd, S., Quantum algorithm for solving linear systems of equations, *Phys. Rev. Lett.* 15, pp. 150502, 2009.
14. Kerenidis, I. and Prakash, A., Quantum Recommendation System, *8th Innovations in Theoretical Computer Science Conference*, 49:1-49:21, 2017.



15. Kerenidis, I. and Prakash, A., *Quantum gradient descent for linear systems and least squares*, arXiv:1704.04992v3, 2017.
16. Le Gall, F., Powers of tensors and fast matrix multiplication, in *Proc. ISSAC 2014*, 296-303, 2014.
17. Lloyd, S., Mohseni, M. and Rebentrost, P., *Quantum algorithms for supervised and unsupervised machine learning*, arXiv:1307.0411v2, 2013.
18. Lloyd, S., Rebentrost, P. and Mohseni, M. Quantum principal component analysis, *Nature Physics* 10, 631-633, 2014.
19. Marvian, I. and Lloyd, S. *Universal quantum emulator*, arXiv:1606.02734, 2016.
20. Nielsen, M.A. and Chuang, I.L., *Quantum Computation and Quantum Information*, 10th Anniversary Edition, Cambridge University Press, 2010.
21. Rebentrost, P., Mohseni, M. and Lloyd, S., Quantum support vector machine for big data classification. *Phys. Rev. Lett.* 113(13), pp. 130503, 2014.
22. Rebentrost, P., Schuld, M., Wossnig, L., Petruccione, F. and Lloyd, S., *Quantum gradient descent and Newton's method for constrained polynomial optimization*, arXiv:1612.01789v2, 2016.
23. Rebentrost, P., Steffens, A. and Lloyd, S., Quantum singular value decomposition of non-sparse low-rank matrices, *Phys. Rev. A* 97, pp. 012327, 2018.
24. Schuld, M., Sinayskiy, I. and Petruccione, F., Prediction by linear regression on a quantum computer, *Phys. Rev. A* 94, pp. 022342, 2016.
25. Shao, C.P., *From linear combination of quantum states to Grover's searching algorithm*, arXiv:1807.09693, 2018.
26. Soklakov, A.N. and Schack, R., Efficient state preparation for a register of quantum bits, *Physical review A* 73, pp. 012307, 2006.
27. Wang, G.M., Quantum algorithm for linear regression, *Phys. Rev. A* 96, pp. 012335, 2017.
28. Wiebe, N., Braun, D. and Lloyd, S., Quantum Algorithm for Data Fitting, *Phys. Rev. Lett.* 109, pp. 050505, 2012.
29. Wiebe, N., Granade, C., Ferrie, C. and Cory, D.G., Hamiltonian learning and certification using quantum resources. *Phys. Rev. Lett.* 112, pp. 190501, 2014.
30. Wossnig, L., Zhao, Z.K. and Prakash, A., A quantum linear system algorithm for dense matrices, *Phys. Rev. Lett.* 120, pp. 050502, 2018.

## A Error and complexity analysis of theorem 1

For simplicity, we denote the approximate of  $s_{ij} := \langle A_{i\bullet} | B_{\bullet j} \rangle$  obtained from the algorithm as  $\tilde{s}_{ij}$ . Then  $|s_{ij} - \tilde{s}_{ij}| \leq \epsilon$ . Denote

$$|C\rangle = \frac{1}{\|C\|_F} \sum_{i,j=0}^{n-1} C_{ij} |i, j\rangle = \frac{1}{\|C\|_F} \sum_{i,j=0}^{n-1} \|A_{i\bullet}\|_2 \|B_{\bullet j}\|_2 s_{ij} |i, j\rangle,$$

and

$$|\tilde{C}\rangle = \frac{1}{\sqrt{Z}} \sum_{i,j=0}^{n-1} \|A_{i\bullet}\|_2 \|B_{\bullet j}\|_2 \tilde{s}_{ij} |i, j\rangle$$

as the quantum state obtain from the algorithm, where

$$Z = \sum_{i,j=0}^{n-1} \|A_{i\bullet}\|_2^2 \|B_{\bullet j}\|_2^2 \tilde{s}_{ij}^2.$$

Then

$$\begin{aligned} \left| \|C\|_F^2 - Z \right| &= \sum_{i,j=0}^{n-1} \|A_{i\bullet}\|_2^2 \|B_{\bullet j}\|_2^2 |s_{ij}^2 - \tilde{s}_{ij}^2| \leq 2\epsilon \sum_{i,j=0}^{n-1} \|A_{i\bullet}\|_2^2 \|B_{\bullet j}\|_2^2 = 2\epsilon \|A\|_F^2 \|B\|_F^2; \\ \left| \|C\|_F - \sqrt{Z} \right| &= \frac{\left| \|C\|_F^2 - Z \right|}{\|C\|_F + \sqrt{Z}} \leq \frac{2\epsilon \|A\|_F^2 \|B\|_F^2}{\|C\|_F + \sqrt{Z}}. \end{aligned}$$



Finally,

$$\begin{aligned}
\| |C\rangle - |\tilde{C}\rangle \|_2^2 &\leq \frac{1}{\|C\|_F^2 Z} \sum_{i,j=0}^{n-1} \|A_{i\bullet}\|_2^2 \|B_{\bullet j}\|_2^2 \left( s_{ij} \sqrt{Z} - \tilde{s}_{ij} \|C\|_F \right)^2 \\
&\leq \frac{2}{\|C\|_F^2 Z} \sum_{i,j=0}^{n-1} \|A_{i\bullet}\|_2^2 \|B_{\bullet j}\|_2^2 \left( \tilde{s}_{ij}^2 (\sqrt{Z} - \|C\|_F)^2 + Z (s_{ij} - \tilde{s}_{ij})^2 \right) \\
&\leq \frac{2\|A\|_F^2 \|B\|_F^2}{\|C\|_F^2} \epsilon^2 + \frac{8\|A\|_F^4 \|B\|_F^4}{\|C\|_F^2 (\|C\|_F + \sqrt{Z})^2} \epsilon^2 \\
&\leq \frac{2\|A\|_F^2 \|B\|_F^2}{\|C\|_F^2} \epsilon^2 + \frac{2\|A\|_F^4 \|B\|_F^4}{\|C\|_F^4} \epsilon^2.
\end{aligned}$$

To make the error is small in size  $\epsilon_0$ , we need to choose  $\epsilon \|A\|_F^2 \|B\|_F^2 / \|C\|_F^2 = \epsilon_0$ . Since the success probability is  $Z / \|A\|_F^2 \|B\|_F^2 \approx \|C\|_F^2 / \|A\|_F^2 \|B\|_F^2$ , finally the complexity will be  $\tilde{O}(\|A\|_F \|B\|_F / \|C\|_F \epsilon) = \tilde{O}(\|A\|_F^3 \|B\|_F^3 / \|C\|_F^3 \epsilon_0)$ .

## B Error and complexity analysis of theorem 3

Denote  $|B_{\bullet j}\rangle = \sum_{k=0}^{n-1} \alpha_{jk} |u_k\rangle$ , then similar to the procedure (2.8), by choosing the initial state as

$$|B\rangle = \frac{1}{\|B\|_F} \sum_{j=0}^{n-1} \|B_{\bullet j}\|_2 |B_{\bullet j}\rangle |j\rangle = \frac{1}{\|B\|_F} \sum_{j,k=0}^{n-1} \|B_{\bullet j}\|_2 \alpha_{jk} |u_k\rangle |j\rangle,$$

we can get the following state

$$\frac{1}{\|B\|_F \max_k \tilde{\sigma}_k} \sum_{j,k=0}^{n-1} \|B_{\bullet j}\|_2 \alpha_{jk} \tilde{\sigma}_k |u_k\rangle |j\rangle |0\rangle + \text{orthogonal parts.} \quad (2.1)$$

Now denote

$$|\phi\rangle = \frac{1}{\sqrt{Z}} \sum_{j,k=0}^{n-1} \|B_{\bullet j}\|_2 \alpha_{jk} \tilde{\sigma}_k |u_k\rangle |j\rangle, \quad |\psi\rangle = \frac{1}{\sqrt{W}} \sum_{j,k=0}^{n-1} \|B_{\bullet j}\|_2 \alpha_{jk} \sigma_k |u_k\rangle |j\rangle,$$

where  $Z = \sum_{j,k=0}^{n-1} \|B_{\bullet j}\|_2^2 |\alpha_{jk}|^2 \tilde{\sigma}_k^2$  and  $W = \sum_{j,k=0}^{n-1} \|B_{\bullet j}\|_2^2 |\alpha_{jk}|^2 \sigma_k^2$ . Since  $|\tilde{\sigma}_k - \sigma_k| \leq \epsilon \|A\|_F$  and  $\|B\|_F^2 = \sum_{j,k=0}^{n-1} \|B_{\bullet j}\|_2^2 |\alpha_{jk}|^2$ , we have

$$|Z - W| \leq \sum_{j,k=0}^{n-1} \|B_{\bullet j}\|_2^2 |\alpha_{jk}|^2 |\tilde{\sigma}_k^2 - \sigma_k^2| \leq \epsilon \|A\|_F \|B\|_F^2 \max_{0 \leq k \leq n-1} |\tilde{\sigma}_k + \sigma_k|.$$

Therefore,

$$\begin{aligned}
\| |\phi\rangle - |\psi\rangle \|_2^2 &= \frac{1}{ZW} \left\| \sum_{j,k=0}^{n-1} \|B_{\bullet j}\|_2 \alpha_{jk} (\sqrt{W} \tilde{\sigma}_k - \sqrt{Z} \sigma_k) |u_k\rangle |j\rangle \right\|_2^2 \\
&\leq \frac{2}{ZW} \sum_{j,k=0}^{n-1} \|B_{\bullet j}\|_2^2 |\alpha_{jk}|^2 \left( W (\tilde{\sigma}_k - \sigma_k)^2 + (\sqrt{Z} - \sqrt{W})^2 \sigma_k^2 \right) \\
&\leq \frac{2\epsilon^2 \|A\|_F^2 \|B\|_F^2}{Z} + \frac{2\epsilon^2 \|A\|_F^2 \|B\|_F^4 \max_k |\tilde{\sigma}_k + \sigma_k|^2}{Z (\sqrt{Z} + \sqrt{W})^2}.
\end{aligned} \quad (2.2)$$

By choosing  $\epsilon_1 = \epsilon \|A\|_F$  small, then  $\tilde{\sigma}_k \approx \sigma_k$ , so  $|Z - W| \leq 2\epsilon_1 \|B\|_F^2 \max_k \sigma_k$ . Now choosing  $\epsilon_2 = \epsilon_1 \|B\|_F^2 \max_k \sigma_k$ , we have  $Z \approx W \leq \|B\|_F^2 \max_k \sigma_k^2$ , so the upper bound given in (2.2) is close to

$$\begin{aligned} \frac{2\epsilon^2 \|A\|_F^2 \|B\|_F^2}{W} + \frac{\epsilon^2 \|A\|_F^2 \|B\|_F^4 \max_k \sigma_k^2}{W^2} &= \frac{2\epsilon_1^2 \|B\|_F^2}{\|AB\|_F^2} + \frac{\epsilon_2^2}{\|AB\|_F^4} \\ &= \frac{\epsilon_2^2}{\|AB\|_F^2} \left( \frac{2}{\|B\|_F^2 \max_k \sigma_k^2} + \frac{1}{\|AB\|_F^2} \right) \\ &\leq \frac{3\epsilon_2^2}{\|AB\|_F^4}. \end{aligned}$$

Finally, we set  $\epsilon_2 = \epsilon_3 \|AB\|_F^2$  which makes the error between  $|\phi\rangle$  and  $|\psi\rangle$  smaller than  $\epsilon_3$ . The probability of getting  $|\phi\rangle$  in (2.1) equals  $Z/\|B\|_F^2 \max_k \sigma_k^2 \approx \|AB\|_F^2/\|B\|_F^2 \max_k \sigma_k^2$ . So the complexity to get  $|\phi\rangle$  is

$$\frac{\|B\|_F \max_k \sigma_k}{\|AB\|_F \epsilon} = \frac{\|A\|_F \|B\|_F^3 \max_k \sigma_k^2}{\|AB\|_F^3 \epsilon_3} \leq \frac{\|A\|_F \|B\|_F \kappa^2}{\|AB\|_F \epsilon_3},$$

since  $\|AB\|_F^2 \geq \|B\|_F^2 \min_k \sigma_k^2$ .

## C Error and complexity analysis of theorem 5

Set  $|B_{\bullet,j}\rangle = \sum_{k=0}^{n-1} \alpha_{jk} |u_k\rangle$  and  $\epsilon \|A\|_F = \epsilon_1$ , then in (2.8), we obtain

$$\frac{1}{\max_k \sigma_k} \sum_k \alpha_{jk} \tilde{\sigma}_k |u_k\rangle |0\rangle + \text{orthogonal part}$$

in time  $\tilde{O}(\|A\|_F/\epsilon_1)$ . Apply swap test on the state  $|i, 0\rangle$  and the above state, we will get a value  $L$  in time  $\tilde{O}(\|A\|_F/\epsilon_1 \epsilon_2)$ , such that

$$\left| L - \frac{1}{\|B_{\bullet,j}\|_2 \max_k \sigma_k} \sum_k \|B_{\bullet,j}\|_2 \alpha_{jk} \tilde{\sigma}_k \langle i | u_k \rangle \right| \leq \epsilon_2.$$

Note that  $c_{ij} = \sum_k \|B_{\bullet,j}\|_2 \alpha_{jk} \sigma_k \langle i | u_k \rangle$ , so

$$\begin{aligned} \left| L \|B_{\bullet,j}\|_2 \max_k \sigma_k - c_{ij} \right| &\leq \left| L \|B_{\bullet,j}\|_2 \max_k \sigma_k - \sum_k \|B_{\bullet,j}\|_2 \alpha_{jk} \tilde{\sigma}_k \langle i | u_k \rangle \right| \\ &\quad + \left| \sum_k \|B_{\bullet,j}\|_2 \alpha_{jk} \sigma_k \langle i | u_k \rangle - \sum_k \|B_{\bullet,j}\|_2 \alpha_{jk} \tilde{\sigma}_k \langle i | u_k \rangle \right| \\ &\leq \epsilon_2 \|B_{\bullet,j}\|_2 \max_k \sigma_k + \epsilon_1 \|B_{\bullet,j}\|_2 \sum_k |\alpha_{jk} \langle i | u_k \rangle|. \end{aligned}$$

We choose  $\epsilon_1$  and  $\epsilon_2$  such that  $\epsilon_2 \|B_{\bullet,j}\|_2 \max_k \sigma_k = \epsilon_3$  and  $\epsilon_1 \|B_{\bullet,j}\|_2 \sum_k |\alpha_{jk} \langle i | u_k \rangle| = \epsilon_3$ . Finally, the complexity is

$$O\left(\frac{\|A\|_F}{\epsilon_1 \epsilon_2}\right) = O\left(\frac{\|A\|_F \|B_{\bullet,j}\|_2^2 \max_k \sigma_k \sum_k |\alpha_{jk} \langle i | u_k \rangle|}{\epsilon_3^2}\right).$$

This is the complexity to compute  $c_{ij}$  to accuracy  $\epsilon_3$ . Therefore, the total complexity to compute all entries of  $C$  equals

$$O\left(\frac{\|A\|_F \max_k \sigma_k}{\epsilon_3^2} \sum_{i,j,k} \|B_{\bullet,j}\|^2 |\alpha_{jk} \langle i | u_k \rangle| + n^2\right) = O\left(\frac{n \|A\|_F \|B_{\bullet,F}\|_3^3 \max_k \sigma_k}{\epsilon_3^2} + n^2\right), \quad (3.1)$$

which is due to

$$\sum_{i,j,k} \|B_{\bullet j}\|_2^2 |\alpha_{jk}| |\langle i|u_k\rangle| = \sum_{j,k} \|B_{\bullet j}\|_2^2 |\alpha_{jk}| \|u_k\|_1 \leq \sqrt{n} \sum_j \|B_{\bullet j}\|_2^2 \sum_k |\alpha_{jk}| = \sqrt{n} \sum_j \|B_{\bullet j}\|_2^2 \|B_{\bullet j}\|_1 \leq n \sum_j \|B_{\bullet j}\|_3^3.$$

Since  $\|A\|_F \geq \sqrt{n} \min_k \sigma_k$ , the above result can be changed into

$$O\left(\frac{\sqrt{n}\kappa\|A\|_F^2\|B_{\bullet F}\|_3^3}{\epsilon_3^2} + n^2\right).$$

If we apply HHL algorithm, then  $\|A_{F\bullet}\|$  should be changed into  $\max_k \sigma_k$ , so the complexity in (3.1) becomes

$$O\left(\frac{n\|B_{\bullet F}\|_3^3 \max_k \sigma_k^2}{\epsilon_3^2} + n^2\right) = O\left(\frac{n\kappa^2\|B_{\bullet F}\|_3^3 \min_k \sigma_k^2}{\epsilon_3^2} + n^2\right) = O\left(\frac{\kappa^2\|B_{\bullet F}\|_3^3\|A\|_F^2}{\epsilon_3^2} + n^2\right).$$

## D Estimating the error in proposition 5

We only consider the case when all given data are real. By measuring (4.1), if we get  $|1\rangle$ , then the post measurement state is

$$|\psi''\rangle = \frac{1}{\sqrt{Y}} \sum_{k=0}^{n-1} b_k \sin(f(k)t) |k\rangle,$$

where  $Y = \sum_{k=0}^{n-1} |b_k \sin(f(k)t)|^2$ . The desired state is

$$|\psi'\rangle = \frac{1}{\sqrt{Z}} \sum_{k=0}^{n-1} b_k f(k) |k\rangle,$$

where  $Z = \sum_{k=0}^{n-1} |b_k f(k)|^2$ . Then

$$\begin{aligned} \langle \psi'' | \psi' \rangle &= \frac{1}{\sqrt{Y}Z} \sum_{k=0}^{n-1} b_k^2 f(k) \sin(f(k)t) \geq \frac{1}{\sqrt{Y}Z} \sum_{k=0}^{n-1} b_k^2 f(k) \left(f(k)t - \frac{\epsilon_1^3}{6}\right) \\ &= \frac{1}{\sqrt{Y}Z} \left(Zt - \frac{\epsilon_1^3}{6} \sum_{k=0}^{n-1} b_k^2 f(k)\right) \geq \frac{1}{\sqrt{Y}Z} \left(Zt - \frac{\epsilon_1^3}{6} \sqrt{Z}\right) = \frac{1}{\sqrt{Y}} \left(\sqrt{Z}t - \frac{\epsilon_1^3}{6}\right) \\ &\geq \frac{1}{\sqrt{Z}t} \left(\sqrt{Z}t - \frac{\epsilon_1^3}{6}\right) = 1 - \frac{\epsilon_1^3}{6\sqrt{Z}t} \geq 1 - \frac{\epsilon_1^2 \kappa(f)}{6}, \end{aligned}$$

where in the second step, we use the fact  $\sin x \geq x - x^3/6$  and  $|f(k)t| \leq \epsilon_1$ . In the fourth step, we applies Cauchy-Schwarz inequality to

$$\sum_{k=0}^{n-1} b_k^2 f(k) \leq \sqrt{\sum_{k=0}^{n-1} b_k^2} \sqrt{\sum_{k=0}^{n-1} b_k^2 f(k)^2} = \sqrt{Z}.$$

In the sixth step, we use the fact that  $Y \leq Zt^2$ . And in the final step, we applies  $|f(k)t| \geq \epsilon_0$  for all  $k$  and  $\epsilon_0 = \kappa(f)^{-1}\epsilon_1$ . Therefore, we have

$$\| |\psi''\rangle - |\psi'\rangle \|_2 = \sqrt{2(1 - \langle \psi'' | \psi' \rangle)} \leq \sqrt{\frac{\kappa(f)}{3}} \epsilon_1 = O(\epsilon),$$

if we choose  $\epsilon_1 = O(\epsilon/\sqrt{\kappa(f)})$ .