Quantum Higher Order Singular Value Decomposition

Lejia Gu Xiaoqiang Wang Guofeng Zhang

Abstract— **Higher order singular value decomposition (HOSVD) is an important tool for analyzing big data in multilinear algebra and machine learning. In this paper, we present a quantum algorithm for higher order singular value decomposition. Our method allows one to decompose a tensor into a core tensor containing tensor singular values and some unitary matrices by quantum computers. Compared to the classical HOSVD algorithm, our quantum algorithm provides an exponential speedup.**

Index Terms— **Quantum algorithm, Quantum machine learning, Higher order singular value decomposition (HOSVD), Tensor**

I. Introduction

Matrix computations are vital to many optimization and machine learning problems. Nowadays, due to the rise of neural networks in machine learning methods, the elements of a network are usually described by tensors which can have more than two indices. Tensors (or hypermatrices), as a higher order generalization of matrices, have found widespread applications in scientific and engineering fields. Tensor decomposition expresses a tensor as a sequence of elementary operations acting on other, often simpler tensors. Usually, key information can be extracted from the decomposed tensor, and less space is needed to store the original tensor. Tensor networks, as a countable collection of tensors connected by contractions, have been widely employed in training machine learning models. A quantum state has a tensor representation. Hence, a quantum network, namely a multipartite system, can be represented by a tensor network. Indeed, quantum circuits are a special class of tensor networks, where the arrangement of the tensors and their types are restricted [4], [7], [12].

Quantum computers are devices that perform calculations by utilizing quantum mechanical features including superposition and entanglement. Although large-scale quantum computers are not built yet, theoretical research on quantum algorithms has been conducted for several years. In 1994, Shor's algorithm [23], is proved to be able to solve integer-factorization problem with polymonial time on a quantum computer, while it is NP in classical computing. In 1996, Grover's search algorithm [9] is able to find an entry from an unstructured database quadratically faster than the corresponding classical algorithm. In 2009, Harrow, Hassidim and Lloyd put forward a quantum algorithm for solving linear systems of equations, which is famous as the HHL Algorithm [10]. Base on this algorithm, many quantum version of classical machine learning methods are designed, such as quantum least-squares linear regression [26] and support vector machines [20]. The runtimes of such algorithms are polylogarithmic in the dimensions of the matrix, so that they provide exponential speedups over their classical counterparts.

There are several types of tensor decompositions, such as canonical polyadic (CP) decomposition [5], [11], tensor-train (TT) decomposition [17], Tucker decompostion [24], and etc. However, currently there are no quantum tensor decomposition algorithms. In this paper, we propose a quantum higher order singular value decomposition (Q-HOSVD). HOSVD is a specific orthogonal Tucker decomposition, and can be considered as an extension of SVD from matrices to tensors. Our method is based upon the quantum matrix singular value decomposition algorithm [19] and several quantum computing techniques. The input can be a tensor of any order and dimension.

Classical HOSVD has been well studied, see, e.g., De Lathauwer, De Moor, and Vandewalle in 2000 [6], and it has been successfully applied to signal processing [16] and pattern recognition [25] problems. Furthermore, HOSVD has shown its strong power in quantum chemistry, especially in the second order Møller Plesset perturbation theory calculations [2]. And HOSVD is used in [27] to derive the output m photon state of a quantum linear passive system which is driven by an m photon input state; more specifically, the wave function of the output is expressed in terms of the HOSVD of the input wave function.

Since HOSVD deals with high dimensional data, it has been put into practice in some machine learning methods. In [21], HOSVD representation for neural networks is proposed. By applying the HOSVD the parameter-varying system can be expressed in a tensor product form by locally tuned neural network models. And in [13], HOSVD is applied for compressing convolutional neural networks (CNN).

By our Q-HOSVD method, it is possible to do singular value decomposition on tensors exponentially faster than the classical algorithms. It can be directly applied to quantum machine learning algorithms, and may help solve computationally challenging problems arising in quantum mechanics and chemistry.

The remainder of this paper is organized as follows. Some preliminaries are given in Section II. The quantum higher order singular value decomposition algorithm is presented in Section III. In the last section, we summarize the results and

This research is supported in part by a Hong Kong Research Grant council (RGC) grant (No. 15206915, No. 15208418).

Department of Applied Mathematics, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong {le-jia.gu,xiaoqiang.wang}@connect.polyu.hk, guofeng.zhang@polyu.edu.hk

compare the quantum HOSVD algorithm with the classical counterpart.

II. Preliminaries

First, we would like to add a comment on the notation that is used. Different symbols are used to facilitate the distinction between scalars, vectors, matrices, and tensors. Scalars are denoted by both lower-case letters $(a, b, \ldots; \alpha, \beta, \ldots)$ and capital letters $(A, B, \ldots)$. Bold-face lower-case letters $(\mathbf{a}, \mathbf{b}, \ldots)$ represent vectors. Since the algorithm we present is a quantum algorithm, vectors are represented as quantum states in Section III, ket $|\cdot\rangle$ denotes a column vector, and bra $\langle\cdot|$ denotes a row vector. Bold-face capitals $(\mathbf{A}, \mathbf{B}, \ldots)$ correspond to matrices or operators, and tensors are written as calligraphic letters $(\mathcal{A}, \mathcal{B}, \ldots)$.

For a matrix $\mathbf{M} \in \mathbb{C}^{m\times n}$, there exists a factorization called singular value decomposition (SVD):

$$\mathbf{M} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^\dagger, \tag{1}$$

where $\mathbf{U}$ is a complex $m\times m$ unitary matrix, $\mathbf{\Sigma}$ is a diagonal $m\times n$ matrix with non-negative real numbers on the diagonal, $\mathbf{V}$ is a complex $n\times n$ unitary matrix and $\mathbf{V}^\dagger$ is the conjugate transpose of $\mathbf{V}$. The decomposition can also be written as

$$\mathbf{M} = \sum_{i=1}^{r} \sigma_i \mathbf{u}_i \mathbf{v}_i^\dagger, \tag{2}$$

where r is the rank of $\mathbf{M}$, σ_i is the ith largest singular value, and $\mathbf{u}_i$ and $\mathbf{v}_i$ are the corresponding left and right singular vectors respectively.

Denote $[m] \equiv \{1, 2, \cdots, m\}$. An mth-order tensor $\mathcal{A} = (a_{i_1\cdots i_m})$ is a multi-array of $\Pi_{j=1}^m I_j$ entries, where $i_j \in [I_j]$ for $j = 1, 2, \cdots, m$, $(I_1, I_2, \ldots, I_m)$ is the dimension of $\mathcal{A}$. When $I_1 = I_2 = \cdots = I_m = n$, $\mathcal{A}$ is called an mth-order n-dimensional tensor [18].

The k-mode tensor-matrix multiplication is defined by

$$\begin{aligned}&(\mathcal{A} \times_k \mathbf{B})_{i_1 i_2 \ldots i_{k-1} j_k i_{k+1} \ldots i_m} \\ &\equiv \sum_{i_k=1}^{I_k} a_{i_1 i_2 \ldots i_{k-1} i_k i_{k+1} \ldots i_m}\, b_{j_k i_k},\end{aligned} \tag{3}$$

where matrix $\mathbf{B} \in \mathbb{C}^{J_k \times I_k}$ which produces another mth-order tensor. The inner product of two tensors $\mathcal{A}, \mathcal{B} \in \mathbb{C}^{I_1\times I_2\times\cdots\times I_m}$, denoted as $\mathcal{A}\cdot\mathcal{B}$, is defined as

$$\mathcal{A}\cdot\mathcal{B} \equiv \sum_{i_1=1}^{I_1}\sum_{i_2=1}^{I_2}\cdots\sum_{i_m=1}^{I_m} a^*_{i_1\cdots i_m} b_{i_1\cdots i_m}. \tag{4}$$

Similar to the matrix case, the induced norm $\sqrt{\mathcal{A}\cdot\mathcal{A}}$ is called the Frobenius norm of $\mathcal{A}$, denoted as $||\mathcal{A}||_F$.

For tensor $\mathcal{A} \in \mathbb{C}^{I_1\times I_2\times\cdots\times I_m}$, if there exist matrices $\mathbf{X}_k = [\mathbf{x}_1^{(k)}\mathbf{x}_2^{(k)}\cdots\mathbf{x}_{I_k}^{(k)}] \in \mathbb{C}^{I_k\times I_k}$ with $||\mathbf{x}_{i_k}^{(k)}|| = 1$ for $k \in [m]$ and $i_k \in [I_k]$ such that

$$\mathcal{A} = \mathcal{S} \times_1 \mathbf{X}_1 \times_2 \mathbf{X}_2 \cdots \times_m \mathbf{X}_m, \tag{5}$$

then (5) is said to be a Tucker decomposition of $\mathcal{A}$, and $\mathcal{S} = (s_{i_1\cdots i_m})$ is called the core tensor of $\mathcal{A}$. Higher order singular value decomposition is a specific orthogonal Tucker decomposition. For $\mathcal{A} \in \mathbb{C}^{I_1\times I_2\times\cdots\times I_m}$, the HOSVD [6] is defined as

$$\mathcal{A} = \mathcal{S} \times_1 \mathbf{U}^{(1)} \times_2 \mathbf{U}^{(2)} \cdots \times_m \mathbf{U}^{(m)}, \tag{6}$$

where the k-mode singular matrix $\mathbf{U}^{(k)} = \left[\mathbf{u}_1^{(k)}\mathbf{u}_2^{(k)}\cdots\mathbf{u}_{I_k}^{(k)}\right]$ is a complex unitary $I_k \times I_k$ matrix, the core tensor $\mathcal{S} \in \mathbb{C}^{I_1\times I_2\times\cdots\times I_m}$ and its subtensors $\mathcal{S}_{i_k=\alpha}$, of which the kth index is fixed to $\alpha \in [I_k]$, have the properties of

(i) all-orthogonality:

Two subtensors $\mathcal{S}_{i_k=\alpha}$ and $\mathcal{S}_{i_k=\beta}$ are orthogonal for $k = 1, 2, \cdots, m$:

$$\mathcal{S}_{i_k=\alpha}\cdot\mathcal{S}_{i_k=\beta} = 0 \qquad \text{when} \quad \alpha \neq \beta, \tag{7}$$

(ii) ordering:

Similar to the matrix case, the tensor singular values are defined as the Frobenius norms of the $(N-1)$th-order subtensors of the core tensor $\mathcal{S}$:

$$\sigma_\alpha^{(k)} = ||\mathcal{S}_{i_k=\alpha}||_F, \tag{8}$$

for $k = 1, \ldots, m$ and $\alpha = 1, \ldots, I_k$. Furthermore, these tensor singular values have the following ordering property

$$\sigma_1^{(k)} \geq \sigma_2^{(k)} \geq \cdots \geq \sigma_{I_k}^{(k)} \geq 0 \tag{9}$$

for $k = 1, 2, \cdots, m$. When $m = 2$, i.e. $\mathcal{A}$ is a matrix, the HOSVD is degenerated to the well-known matrix SVD.

For an mth-order tensor $\mathcal{A} \in \mathbb{C}^{I_1\times I_2\times\ldots\times I_m}$, the matrix unfolding $\mathbf{A}^{(k)} \in \mathbb{C}^{I_k\times(\Pi_{j\neq k} I_j)}$ contains the element $a_{i_1\cdots i_m}$ at the position with row number i_k and column number

$$\begin{aligned}&(i_{k+1}-1)I_{k+2}I_{k+3}\cdots I_m I_1 I_2 \cdots I_{k-1} \\ &+(i_{k+2}-1)I_{k+3}I_{k+4}\cdots I_m I_1 I_2 \cdots I_{k-1} + \cdots \\ &+(i_m-1)I_1 I_2 \cdots I_{k-1} + (i_1-1)I_2 I_3 \cdots I_{k-1} \\ &+(i_2-1)I_3 I_4 \cdots I_{k-1} + \cdots + i_{k-1}.\end{aligned}$$

By the above construction, the rank of $\mathbf{A}^{(k)}$ is at most I_k. Clearly, the elements of tensor $\mathcal{A}$ and matrix $\mathbf{A}^{(k)}$ have a one-to-one correspondence to each other.

In HOSVD, the columns of $\mathbf{U}^{(k)}$ have been sorted such that the jth column $\mathbf{u}_j^{(k)}$ corresponds to the jth largest nonzero singular value of $\mathbf{A}^{(k)}$. Then, we can similarly define the truncated HOSVD. For $k \in [m]$, we take the first r_k columns of $\mathbf{U}^{(k)}$, then $\mathbf{U}^{(k)} \in \mathbb{C}^{I_k\times r_k}$. Finally, the core tensor $\mathcal{S}$ is now of size $r_1 \times r_2 \times \cdots \times r_m$.

In quantum computing, suppose we have a bipartite system, whose state is described by a density operator ρ. The reduced density operator for the first subsystem is defined by

$$\rho_1 \equiv \mathrm{tr}_2(\rho), \tag{10}$$

where tr_2 is a map of operators known as the partial trace [15] over the second subsystem. The partial trace is defined by

$$\begin{aligned}\mathrm{tr}_2\left(|a_1\rangle\langle a_2| \otimes |b_1\rangle\langle b_2|\right) &\equiv |a_1\rangle\langle a_2|\,\mathrm{tr}(|b_1\rangle\langle b_2|) \\ &= \langle b_2|b_1\rangle\,|a_1\rangle\langle a_2|,\end{aligned} \tag{11}$$

where $|a_1\rangle$ and $|a_2\rangle$ are two states in the first subsystem, $|b_1\rangle$ and $|b_2\rangle$ are two states in the second subsystem.

III. Q-HOSVD Algorithm

In this section, we first present our Q-HOSVD algorithm.

Algorithm 1 Quantum Higher Order Singular Value Decomposition (Q-HOSVD)

Input: $\mathcal{A}, \epsilon, |b\rangle, |00\cdots0\rangle$
Output: $\mathcal{S}, \mathbf{U}^{(1)}, \mathbf{U}^{(2)}, \ldots, \mathbf{U}^{(m)}$
Step 1. Load $\mathcal{A}$ and ancilla $|b\rangle|00\cdots0\rangle$ into the quantum register
Step 2. Unfold tensor $\mathcal{A}$ to matrix $\mathbf{A}^{(k)}$
Step 3. Extend $\mathbf{A}^{(k)}$ to Hermitian matrix $\tilde{\mathbf{A}}^{(k)}$
Step 4. Apply phase estimation to obtain $|\psi\rangle$
Step 5. Perform measurement on $|\tilde{\lambda}_j/N\rangle$ and extract $\mathbf{u}_j^{(k)}$ to compose $\mathbf{U}^{(k)}$
Repeat Steps 2-5 for $k = 1, \ldots, m$
Step 6. $\mathcal{S} \leftarrow \mathcal{A} \times_1 \mathbf{U}^{(1)^\dagger} \times_2 \mathbf{U}^{(2)^\dagger} \cdots \times_m \mathbf{U}^{(m)^\dagger}$
Return $\mathcal{S}, \mathbf{U}^{(1)}, \mathbf{U}^{(2)}, \ldots, \mathbf{U}^{(m)}$

In the following we explain the implementation of Algorithm 1. For simplicity, we assume $I_1 = I_2 = \cdots = I_m = n$, then $\mathcal{A}$ is an mth-order n-dimensional tensor.

A. Step 1.

A matrix $\mathbf{A} = (a_{ij})$ can be loaded into a quantum register by an oracle named quantum random access memory (qRAM) [8]:

$$|i\,j\rangle|00\cdots0\rangle \mapsto |i\,j\rangle|a_{ij}\rangle. \tag{12}$$

Similarly tensor $\mathcal{A} = (a_{i_1 i_2 \cdots i_m})$ can be accessed by the following operation

$$|i_1 i_2 \cdots i_m\rangle|\underbrace{00\cdots0}_{\ell}\rangle \mapsto |i_1 i_2 \cdots i_m\rangle|a_{i_1 i_2 \cdots i_m}\rangle, \tag{13}$$

where $i_k \in [n]$ for $k = 1, 2, \ldots, m$. This procedure can be achieved using $m \log n + \ell$ qubits and $T_{\mathcal{A}} = O(\log^m n)$ operations.

B. Step 2.

The quantum unfolding matrix $\mathbf{A}^{(k)} = (a'^{(k)}_{i_k j_k})$ can be directly processed in the following way

$$\begin{aligned}&\sum_{i_1,i_2,\ldots,i_m=0}^{n-1} |i_k i_{k+1} \cdots i_m i_1 \cdots i_{k-1}\rangle|a_{i_k i_{k+1} \cdots i_m i_1 \cdots i_{k-1}}\rangle \\ \to &\sum_{i_k=0}^{n-1}\sum_{j_k=0}^{n^{m-1}-1} |i_k j_k\rangle|a'^{(k)}_{i_k j_k}\rangle,\end{aligned} \tag{14}$$

where $|j_k\rangle = |i_{k+1} \cdots i_m i_1 \cdots i_{k-1}\rangle$. For example, for a $2 \times 2 \times 2$ tensor $\mathcal{A}$, the mode-1 unfolding $\mathbf{A}^{(1)}$ corresponds to $\mathcal{A}$ by

$$\begin{aligned}|000\rangle|a_{000}\rangle &\to |00\rangle|a'^{(1)}_{00}\rangle \\ |001\rangle|a_{001}\rangle &\to |01\rangle|a'^{(1)}_{01}\rangle \\ &\vdots \\ |111\rangle|a_{111}\rangle &\to |13\rangle|a'^{(1)}_{13}\rangle.\end{aligned}$$

After unfolding, in Step 3 below we show how to use the quantum singular value decomposition for matrices [19] to find the singular matrices of the original tensor $\mathcal{A}$.

C. Step 3.

Since $\mathbf{A}^{(k)}$ is not a Hermitian matrix, we consider the following extended matrix

$$\tilde{\mathbf{A}}^{(k)} \equiv \begin{bmatrix} 0 & \mathbf{A}^{(k)} \\ \mathbf{A}^{(k)^\dagger} & 0 \end{bmatrix}, \tag{15}$$

then $\tilde{\mathbf{A}}^{(k)}$ is an $(n + n^{m-1}) \times (n + n^{m-1})$ Hermitian matrix. For Hermitian matrices, the singular values are the absolute value of eigenvalues, so that phase estimation [15] can be used to apply the singular value decomposition. Since $\mathrm{rank}(\mathbf{A}^{(k)}) \le n$, $\mathrm{rank}(\tilde{\mathbf{A}}^{(k)}) \le 2n$. In the following we use $\mathbf{A}$ and $\tilde{\mathbf{A}}$ to represent $\mathbf{A}^{(k)}$ and $\tilde{\mathbf{A}}^{(k)}$ respectively when k is fixed and $N = n + n^{m-1}$ to represent the dimension of $\tilde{\mathbf{A}}$.

D. Step 4.

Define a SWAP operator $\mathbf{S}_{\tilde{A}} \in \mathbb{C}^{N^2 \times N^2}$:

$$\mathbf{S}_{\tilde{A}} = \sum_{j,k=1}^{N} \tilde{\mathbf{A}}_{jk} |k\rangle\langle j| \otimes |j\rangle\langle k|. \tag{16}$$

This SWAP matrix is one-sparse in a quadratically bigger space, therefore, the matrix exponentiation $e^{-\mathrm{i}\mathbf{S}_{\tilde{A}}\Delta t}$ is efficiently simulatable [3].

We use quantum principal component analysis (qPCA) [14] to implement $\tilde{\mathbf{A}}$ using $\mathbf{S}_{\tilde{A}}$. Let ρ_1 and ρ_2 be two distinct density matrices. Performing $\mathbf{S}_{\tilde{A}}$ for small Δt on $\rho_1 \otimes \rho_2$:

$$\begin{aligned}&\mathrm{tr}_1\{e^{-\mathrm{i}\mathbf{S}_{\tilde{A}}\Delta t}\rho_1 \otimes \rho_2 e^{\mathrm{i}\mathbf{S}_{\tilde{A}}\Delta t}\} \\ =&\mathrm{tr}_1\{(\mathbf{I} - \mathrm{i}\mathbf{S}_{\tilde{A}}\Delta t)\rho_1 \otimes \rho_2(\mathbf{I} + \mathrm{i}\mathbf{S}_{\tilde{A}}\Delta t)\} + O(\Delta t^2) \\ =&\rho_2 - \mathrm{i}\,\mathrm{tr}_1\{\mathbf{S}_{\tilde{A}}\rho_1 \otimes \rho_2\}\Delta t + \mathrm{i}\,\mathrm{tr}_1\{\rho_1 \otimes \rho_2 \mathbf{S}_{\tilde{A}}\}\Delta t + O(\Delta t^2).\end{aligned} \tag{17}$$

The term $\mathrm{tr}_1\{\mathbf{S}_{\tilde{A}}\rho_1 \otimes \rho_2\}$ in (17) can be rewritten as

$$\begin{aligned}&\mathrm{tr}_1\{\mathbf{S}_{\tilde{A}}\rho_1 \otimes \rho_2\} \\ =&\mathrm{tr}_1\left\{\sum_{j,k=0}^{N-1} \tilde{\mathbf{A}}_{jk} |k\rangle\langle j|\,\rho_1 \otimes |j\rangle\langle k|\,\rho_2\right\} \\ =&\sum_{j,k=1}^{N} \tilde{\mathbf{A}}_{jk} \langle j|\,\rho_1\,|k\rangle \otimes |j\rangle\langle k|\,\rho_2.\end{aligned} \tag{18}$$

If we choose $\rho_1 = |\vec{1}\rangle\langle\vec{1}|$, with $|\vec{1}\rangle \equiv \frac{1}{\sqrt{N}}\sum_{k=0}^{N-1}|k\rangle$, i.e.

$$\rho_1 = \frac{1}{N}\sum_{j,k=0}^{N-1} |j\rangle\langle k|, \tag{19}$$

then

$$\mathrm{tr}_1\{\mathbf{S}_{\tilde{A}}\rho_1\otimes\rho_2\}=\frac{1}{N}\sum_{j,k=0}^{N-1}\tilde{\mathbf{A}}_{jk}\left|j\right\rangle\left\langle k\right|\rho_2=\frac{\tilde{\mathbf{A}}}{N}\rho_2. \quad (20)$$

Similarly, the second $O(\Delta t)$ term $\mathrm{tr}_1\{\rho_1\otimes\rho_2\mathbf{S}_{\tilde{A}}\}$ in (17) becomes

$$\mathrm{tr}_1\{\rho_1\otimes\rho_2\mathbf{S}_{\tilde{A}}\}=\rho_2\frac{\tilde{\mathbf{A}}}{N}. \quad (21)$$

Therefore,

$$\begin{aligned}\mathrm{tr}_1\{e^{-\mathrm{i}\mathbf{S}_{\tilde{A}}\Delta t}\rho_1\otimes\rho_2 e^{\mathrm{i}\mathbf{S}_{\tilde{A}}\Delta t}\}&=\rho_2-\mathrm{i}\frac{\Delta t}{N}[\tilde{\mathbf{A}},\rho_2]+O(\Delta t^2)\\&\approx e^{-\mathrm{i}\frac{\tilde{\mathbf{A}}}{N}\Delta t}\rho_2 e^{\mathrm{i}\frac{\tilde{\mathbf{A}}}{N}\Delta t}. \quad (22)\end{aligned}$$

Let ϵ_0 be the trace norm of the error term $O(\Delta t^2)$. For s steps, the resulting error is $\epsilon_1=s\epsilon_0\le 2s||\mathcal{A}||_{\max}^2\Delta t^2$, where $||\mathcal{A}||_{\max}=\max_{i_1,\ldots,i_m}|a_{i_1\cdots i_m}|$. The proof is similar to that in [19]. The simulated time is $t=s\Delta t$. Then,

$$\frac{\epsilon_1}{s}\le 2||A||_{\max}^2\left(\frac{t}{s}\right)^2. \quad (23)$$

Thus,

$$s=O\left(\frac{t^2}{\epsilon_1}||\mathcal{A}||_{\max}^2\right) \quad (24)$$

steps are required to simulate $e^{-\mathrm{i}\frac{\tilde{\mathbf{A}}}{N}\Delta t}$ if ϵ_1 and t are fixed. Assume $||\mathcal{A}||_{\max}=O(1)$, then $s=O(t^2/\epsilon_1)$. Applying the output in Eq. (22) again in the second register, we obtain

$$\begin{aligned}&\mathrm{tr}_1\left\{e^{-\mathrm{i}\mathbf{S}_{\tilde{A}}\Delta t}\rho_1\otimes\left(\rho_2-\mathrm{i}\frac{\Delta t}{N}[\tilde{\mathbf{A}},\rho_2]+O(\Delta t^2)\right)e^{\mathrm{i}\mathbf{S}_{\tilde{A}}\Delta t}\right\}\\=&\mathrm{tr}_1\{e^{-\mathrm{i}\mathbf{S}_{\tilde{A}}\Delta t}\rho_1\otimes\rho_2 e^{\mathrm{i}\mathbf{S}_{\tilde{A}}\Delta t}\}\\&-\mathrm{i}\frac{\Delta t}{N}\mathrm{tr}_1\{e^{-\mathrm{i}\mathbf{S}_{\tilde{A}}\Delta t}(\rho_1\otimes[\tilde{\mathbf{A}},\rho_2])e^{\mathrm{i}\mathbf{S}_{\tilde{A}}\Delta t}\}+O(\Delta t^2)\\=&\rho_2-\mathrm{i}\frac{\Delta t}{N}[\tilde{\mathbf{A}},\rho_2]-\mathrm{i}\frac{\Delta t}{N}\mathrm{tr}_1\{\rho_1\otimes[\tilde{\mathbf{A}},\rho_2]\}+O(\Delta t^2)\\=&\rho_2-\mathrm{i}\frac{2\Delta t}{N}[\tilde{\mathbf{A}},\rho_2]+O(\Delta t^2). \quad (25)\end{aligned}$$

Thus, by continuously using k copies of ρ_1 we can simulate $e^{-\mathrm{i}(\tilde{\mathbf{A}}/N)k\Delta t}$.

Next, we use the quantum phase estimation algorithm [15] to estimate the eigenvalues of $e^{-\mathrm{i}(\tilde{\mathbf{A}}/N)\Delta t}$. Given an initial quantum state

$$\left|\psi\right\rangle=\left|\underbrace{0\cdots0}_{d}\right\rangle\left|\vec{1}\right\rangle\left|b\right\rangle \quad (26)$$

with $d=O(\lceil\log(1/\epsilon_2)\rceil)$ control qubits, where $\left|b\right\rangle$ is the superposition of eigenvectors $\left|\tilde{u}_j\right\rangle$ corresponding to $\tilde{\lambda}_j$:

$$\left|b\right\rangle=\sum_{j=0}^{N-1}\beta_j\left|\tilde{u}_j\right\rangle,\qquad\sum_{j=0}^{N-1}|\beta_j|^2=1, \quad (27)$$

ϵ_2 is the accuracy for approximating the eigenvalues. Let $\rho_2=\left|b\right\rangle\left\langle b\right|$. We first apply Hadamard gates to the first register, then the state (26) becomes

$$\frac{1}{\sqrt{2^d}}\sum_{\ell=0}^{2^d-1}\left|\ell\right\rangle\left|\vec{1}\right\rangle\left|b\right\rangle, \quad (28)$$

whose density matrix has the following form

$$\frac{1}{2^d}\sum_{\ell=0}^{2^d-1}\left|\ell\right\rangle\left\langle\ell\right|\otimes\rho_1\otimes\rho_2. \quad (29)$$

Then we multiply $\sum_{\ell=0}^{2^d-1}\left|\ell\right\rangle\left\langle\ell\right|\otimes(e^{-\mathrm{i}\mathbf{S}_A\Delta t})^\ell$ and $\sum_{\ell=0}^{2^d-1}\left|\ell\right\rangle\left\langle\ell\right|\otimes(e^{\mathrm{i}\mathbf{S}_A\Delta t})^\ell$ to both sides of (29) to obtain

$$\sum_{\ell=0}^{2^d-1}\left|\ell\right\rangle\left\langle\ell\right|\otimes\left((e^{-\mathrm{i}\mathbf{S}_A\Delta t})^\ell\rho_1\otimes\rho_2(e^{\mathrm{i}\mathbf{S}_A\Delta t})^\ell\right). \quad (30)$$

Next, we perform a partial trace to the second register using (22) resulting in

$$\sum_{\ell=0}^{2^d-1}\left|\ell\right\rangle\left\langle\ell\right|\otimes\left((e^{-\mathrm{i}\frac{\tilde{\mathbf{A}}}{N}\Delta t})^\ell\rho_2(e^{\mathrm{i}\frac{\tilde{\mathbf{A}}}{N}\Delta t})^\ell\right). \quad (31)$$

After that, we apply the phase estimation to obtain the eigenvalues of $\tilde{\mathbf{A}}/N$, since

$$\begin{aligned}e^{-\mathrm{i}\frac{\tilde{\mathbf{A}}}{N}\Delta t}\left|b\right\rangle&=\sum_{j=0}^{N-1}\beta_j e^{-\mathrm{i}\frac{\tilde{\mathbf{A}}}{N}\Delta t}\left|\tilde{u}_j\right\rangle\\&=\sum_{j=0}^{N-1}\beta_j e^{-\mathrm{i}\lambda_j(\frac{\tilde{\mathbf{A}}}{N})\Delta t}\left|\tilde{u}_j\right\rangle. \quad (32)\end{aligned}$$

At last, we implement the inverse quantum Fourier transform [15] and remove the first register, the final state

$$\left|\psi\right\rangle=\sum_{j=0}^{N-1}\beta_j\left|\tilde{\lambda}_j/N\right\rangle\left|\tilde{u}_j\right\rangle \quad (33)$$

is obtained, where $\left|\tilde{\lambda}_j/N\right\rangle$ is the eigenvalue of $\tilde{\mathbf{A}}/N$ encoded in basis qubits. The corresponding eigenvector $\left|\tilde{u}_j\right\rangle$ is proportional to $(\mathbf{u}_j;\pm\mathbf{v}_j)\in\mathbb{C}^N$, where $\mathbf{u}_j$ and $\mathbf{v}_j$ are the left and right singular vectors of $\tilde{\mathbf{A}}$, the norm of each subvector $\mathbf{u}_j$ and $\mathbf{v}_j$ is $1/\sqrt{2}$, independent of their respective lengths n and n^{m-1}.

E. Step 5.

Since $\mathbf{A}$ is of size $n\times n^{m-1}$, $\mathbf{A}$ has at most n singular values $\{\sigma_j\}$. As a result, $\tilde{\mathbf{A}}$ has at most $2n$ nonzero eigenvalues $\tilde{\lambda}_j\in\{\pm\sigma_j\}$. Next, we measure the first register of state (33) in the computational basis $\{\left|0\right\rangle,\cdots,\left|2^d-1\right\rangle\}$, all eigenpairs $\left|\tilde{\lambda}_j/N\right\rangle\left|\tilde{u}_j\right\rangle$ are obtained with probability $|\beta_j|^2$. Discarding the first register, and projecting $\left|\tilde{u}_j\right\rangle$ onto the $\mathbf{u}_j$ part by using projection operators $\mathbf{P}_u=\sum_{i=0}^{n-1}\left|i\right\rangle\left\langle i\right|$ and $\mathbf{P}_v=\sum_{i=n}^{n^{m-1}+n-1}\left|i\right\rangle\left\langle i\right|$ results in $\left|\mathbf{u}_j\right\rangle$ with probability $\langle\tilde{u}_j|\mathbf{u}_j,0\rangle=\frac{1}{2}$. Then, the singular matrix $\mathbf{U}$ is calculated by

$$\mathbf{U}=\sum_{j=1}^{n}\left|\mathbf{u}_j\right\rangle\left\langle j\right|. \quad (34)$$

Repeating measurements with the initial state $\left|b\right\rangle=\left|0\right\rangle,\left|1\right\rangle,\cdots,\left|n-1\right\rangle$ and applying amplitude amplification [1], we can obtain all the singular vectors in $T_U=O(n^{3/2})$ times with probability close to 1. Thus, the singular matrix $\mathbf{U}^{(k)}$ is reconstructed.

F. Step 6.

After we get all $\mathbf{U}^{(k)}$ for $k = 1, 2, \ldots, m$, in this step we calculate the core tensor $\mathcal{S}$:

$$\mathcal{S} = \mathcal{A} \times_1 \mathbf{U}^{(1)^\dagger} \times_2 \mathbf{U}^{(2)^\dagger} \cdots \times_m \mathbf{U}^{(m)^\dagger}. \tag{35}$$

Similar to the quantum matrix multiplication algorithm by swap test [22], we may calculate the tensor-matrix multiplication $\mathcal{A} \times_k \mathbf{U}^{(k)^\dagger}$ through the following state

$$\frac{1}{||\mathcal{A}||_F||\mathbf{U}^{(k)}||_F} \sum_{i_1,\ldots,i_{k-1},j_k,i_{k+1},\ldots,i_m=0}^{n-1} ||\mathbf{U}^{(k)}_{\bullet j_k}||_2 \\ ||\mathcal{A}_{i_1\ldots i_{k-1}\bullet i_{k+1}\ldots i_m}||_2 \langle \mathcal{A}_{i_1\ldots i_{k-1}\bullet i_{k+1}\ldots i_m} | \mathbf{U}^{(k)}_{\bullet j_k} \rangle \\ |i_1, \ldots, i_{k-1}, j_k, i_{k+1}, \ldots, i_m\rangle |0\rangle + |0\rangle^\perp, \tag{36}$$

where $|\mathcal{A}_{i_1\ldots i_{k-1}\bullet i_{k+1}\ldots i_m}\rangle$ is an n-level quantum state (n-entry vector) if $i_1, \ldots, i_{k-1}, i_{k+1}, \ldots, i_m$ are all fixed. Post-selecting the state $|0\rangle$, we obtain the final state.

By (36), the success probability is

$$\frac{\sum ||\mathbf{U}^{(k)}_{\bullet j_k}||_2^2 ||\mathcal{A}_{i_1\ldots i_{k-1}\bullet i_{k+1}\ldots i_m}||_2^2 \langle \mathcal{A}_{i_1\ldots i_{k-1}\bullet i_{k+1}\ldots i_m} | \mathbf{U}^{(k)}_{\bullet j_k} \rangle^2}{||\mathcal{A}||_F^2 ||\mathbf{U}^{(k)}||_F^2} \\ = \frac{||\mathcal{A} \times_k \mathbf{U}^{(k)^\dagger}||_F^2}{||\mathcal{A}||_F^2 ||\mathbf{U}^{(k)}||_F^2}. \tag{37}$$

After applying amplitude amplification [1], the final computational complexity $T_M = \tilde{O}(||\mathcal{A}||_F ||\mathbf{U}^{(k)}||_F / \epsilon_3 ||\mathcal{A} \times_k \mathbf{U}^{(k)^\dagger}||_F)$ to accuracy ϵ_3. Since unitary matrices preserve norms,

$$||\mathcal{A}||_F = ||\mathcal{A} \times_k \mathbf{U}^{(k)^\dagger}||_F.$$

Thus,

$$T_m = \tilde{O}\left(\frac{||\mathbf{U}^{(k)}||_F}{\epsilon_3}\right) = \tilde{O}\left(\frac{\sqrt{n}}{\epsilon_3}\right). \tag{38}$$

Without loss of generality, we let $\epsilon_1 = \epsilon_2 = \epsilon_3 = \epsilon$.

IV. Summary and Discussion

We have described a quantum algorithm for higher order singular value decomposition. The input can be a general tensor of any order and dimension. The output is a core tensor including tensor singular values and singular matrices stored in the quantum register.

In our method, the computational complexity mainly comes from matrix exponential simulation, data access, phase estimation, quantum measurement, and tensor-matrix multiplication. For an mth-order n-dimensional tensor, the complexity of the classical HOSVD is $O(m\, n^{m+1})$, while the complexity of our quantum HOSVD is

$$msT_U T_A/\epsilon + mT_M = O(mn^{3/2} \log^m n/\epsilon^4) + \tilde{O}(m\sqrt{n}/\epsilon), \tag{39}$$

where ϵ is the accuracy for matrix exponentiating, phase estimation and tensor-matrix multiplication. Generally, $1/\epsilon$ can be considered as $O(\mathrm{polylog}(m, n))$. In this sense, our quantum HOSVD algorithm provides an exponential speedup over the classical counterpart.

References

[1] A. Ambainis, "Variable time amplitude amplification and quantum algorithms for linear algebra problems," in *29th International Symposium on Theoretical Aspects of Computer Science (STACS 2012)*, Dagstuhl, Germany, 2012, vol. 14, pp. 636-647.

[2] F. Bell, D. S. Lambrecht, and M. Head-Gordon, "Higher order singular value decomposition in quantum chemistry," *Molecular Physics*, vol. 108, no. 19-20, pp. 2759-2773, Oct. 2010.

[3] D. W. Berry, G. Ahokas, R. Cleve, and B. C. Sanders, "Efficient Quantum Algorithms for Simulating Sparse Hamiltonians," *Commun. Math. Phys.*, vol. 270, no. 2, pp. 359-371, Mar. 2007.

[4] J. Biamonte and V. Bergholm, "Tensor Networks in a Nutshell," *arXiv:1708.00006 [quant-ph]*, Jul. 2017.

[5] J. D. Carroll and J.-J. Chang, "Analysis of individual differences in multidimensional scaling via an n-way generalization of 'Eckart-Young' decomposition," *Psychometrika*, vol. 35, no. 3, pp. 283-319, Sep. 1970.

[6] L. De Lathauwer, B. De Moor, and J. Vandewalle, "A Multilinear Singular Value Decomposition," *SIAM J. Matrix Anal. Appl.*, vol. 21, no. 4, pp. 1253-1278, Jan. 2000.

[7] V. Dunjko and H. J. Briegel, "Machine learning & artificial intelligence in the quantum domain: a review of recent progress," *Rep. Prog. Phys.*, vol. 81, no. 7, p. 074001, Jun. 2018.

[8] V. Giovannetti, S. Lloyd, and L. Maccone, "Quantum Random Access Memory," *Phys. Rev. Lett.*, vol. 100, no. 16, p. 160501, Apr. 2008.

[9] L. K. Grover, "A Fast Quantum Mechanical Algorithm for Database Search," in *Proceedings of the Twenty-eighth Annual ACM Symposium on Theory of Computing*, New York, NY, USA, 1996, pp. 212-219.

[10] A. W. Harrow, A. Hassidim, and S. Lloyd, "Quantum Algorithm for Linear Systems of Equations," *Phys. Rev. Lett.*, vol. 103, no. 15, p. 150502, Oct. 2009.

[11] R. A. Harshman, "Foundations of the PARAFAC procedure: Models and conditions for an 'explanatory' multi-modal factor analysis," *UCLA Working Papers in Phonetics*, vol. 16, 1970.

[12] W. Huggins, P. Patil, B. Mitchell, K. B. Whaley, and E. M. Stoudenmire, "Towards quantum machine learning with tensor networks," *Quantum Sci. Technol.*, vol. 4, no. 2, p. 024001, Jan. 2019.

[13] M. Kholiavchenko, "Iterative Low-Rank Approximation for CNN Compression," *arXiv:1803.08995 [cs]*, Mar. 2018.

[14] S. Lloyd, M. Mohseni, and P. Rebentrost, "Quantum principal component analysis," *Nature Physics*, vol. 10, no. 9, pp. 631-633, Sep. 2014.

[15] M. A. Nielsen and I. Chuang. *Quantum computation and quantum information*. Cambridge University Press, 2000.

[16] L. Omberg, G. H. Golub, and O. Alter, "A tensor higher-order singular value decomposition for integrative analysis of DNA microarray data from different studies," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 104, no. 47, pp. 18371-18376, Nov. 2007.

[17] I. Oseledets, "Tensor-Train Decomposition," *SIAM J. Sci. Comput.*, vol. 33, no. 5, pp. 2295-2317, Jan. 2011.

[18] L. Qi and Z. Luo, *Tensor analysis: spectral theory and special tensors*. SIAM, 2017.

[19] P. Rebentrost, A. Steffens, I. Marvian, and S. Lloyd, "Quantum singular-value decomposition of nonsparse low-rank matrices," *Phys. Rev. A*, vol. 97, no. 1, p. 012327, Jan. 2018.

[20] P. Rebentrost, M. Mohseni, and S. Lloyd, "Quantum Support Vector Machine for Big Data Classification," *Phys. Rev. Lett.*, vol. 113, no. 13, p. 130503, Sep. 2014.

[21] A. Rövid, L. Szeidl, and P. Várlaki, "On tensor-product model based representation of neural networks," in *2011 15th IEEE International Conference on Intelligent Engineering Systems*, Poprad, Slovakia, 2011, pp. 69-72.

[22] C. Shao, "Quantum Algorithms to Matrix Multiplication," *arXiv:1803.01601 [quant-ph]*, Mar. 2018.

[23] P. W. Shor, "Algorithms for quantum computation: discrete logarithms and factoring," in *Proceedings 35th Annual Symposium on Foundations of Computer Science*, Santa Fe, NM, USA, 1994, pp. 124-134.

[24] L. R. Tucker, "Some mathematical notes on three-mode factor analysis," *Psychometrika*, vol. 31, no. 3, pp. 279-311, Sep. 1966.

[25] M. A. O. Vasilescu and D. Terzopoulos, "Multilinear Analysis of Image Ensembles: TensorFaces," in *Proceedings of the 7th European Conference on Computer Vision (ECCV '02)*, Copenhagen, Denmark, 2002, pp. 447-460.

[26] N. Wiebe, D. Braun, and S. Lloyd, "Quantum Algorithm for Data Fitting," *Phys. Rev. Lett.*, vol. 109, no. 5, p. 050505, Aug. 2012.
[27] G. Zhang, "Dynamical analysis of quantum linear systems driven by multi-channel multi-photon states," *Automatica*, vol. 83, pp. 186-198, Sep. 2017.