SCIENCE CHINA
Information Sciences

• **Supplementary File** •

Quantum Speedup of Twin Support Vector Machines

Zekun YE[1,3], Lvzhou LI[1,4*], Haozhen SITU[2,3] & Yuyi WANG[1]

[1]*School of Data and Computer Science, Sun Yat-Sen University, Guangzhou* 510006*, China;*
[2]*College of Mathematics and Informatics, South China Agricultural University, Guangzhou* 510642*, China;*
[3]*Center for Quantum Computing, Peng Cheng Laboratory, Shenzhen* 518055*, China;*
[4] *Ministry of Education Key Laboratory of Machine Intelligence and Advanced Computing (Sun Yat-sen University), Guangzhou* 510006*, China*

Appendix A Background Knowledge of Quantum Computation

We review important notions and notations of quantum computation. Bra-Ket notations $\langle\cdot|$ and $|\cdot\rangle$ are used to denote vectors. $\langle v|$ represents the row vector $(v_1, v_2, \ldots, v_n)$, and $|v\rangle$ the conjugate transpose of $\langle v|$, i.e., $|v\rangle = \langle v|^\dagger$.

We now describe the four postulates of quantum mechanics, and a more comprehensive description on quantum computation and quantum information can be found in [1]. Firstly, a quantum system is associated with a Hibert space and its state is a unit vector in the space. The simplest quantum system is a qubit, the basic unit of quantum information, which lies in a two-dimensional state space. A qubit can be in a superposition state $|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle$, where $\alpha, \beta \in \mathbb{C}$, and $|\alpha|^2 + |\beta|^2 = 1$. Moreover, $|0\rangle$ and $|1\rangle$ are called basis states which correspond to the classical bit 0 and 1, respectively, and usually $|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. For a d-dimensional quantum system, its state space is $\mathbb{C}^d$, and its state is written as $|\psi\rangle = \sum_{i=0}^{d-1} \psi_i |i\rangle$, where $|i\rangle = (0, ..., 0, 1, 0, ..., 0)^T$ denotes a column vector with the $(i+1)$-th entry being 1 and else 0, $\psi_i \in \mathbb{C}$, and $\sum_{i=0}^{d-1} |\psi_i|^2 = 1$.

Secondly, the evolution of a closed quantum system is described by a unitary operator. A operator U is said to unitary if $UU^\dagger = U^\dagger U = I$, where $\dagger$ denotes the conjugate transpose. If the state of the system is $|\psi_1\rangle$ at time t_1, and the state of the system is $|\psi_2\rangle$ at time t_2, then there exists a unitary operator U which depends only on the time t_1 and t_2, such that $|\psi_2\rangle = U\,|\psi_1\rangle$.

Thirdly, we use the quantum measurement to obtain the information from quantum states. An important and common class of measurements is projective measurements. A projective measurement is described by a set of projective operators $\{P_i\}$, which satisfy the constraints $\sum_i P_i = I$ and $P_i P_j = P_i \delta_{ij}$. If the current state of a quantum system is $|\psi\rangle$, after the measurement, the outcome m will be observed with probability $p(m) = \langle\psi|\, P_m\,|\psi\rangle$, and the state becomes $P_m\,|\psi\rangle / \sqrt{p(m)}$ after the measurement correspondingly. For example, we measure the quantum state $|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle$ by the measurement defined by two projective operators $P_0 = |0\rangle\langle 0|, P_1 = |1\rangle\langle 1|$. Then the probability of obtaining the outcome 0 is $|\alpha|^2$, and the state becomes $|0\rangle$. Similarly, the probability of obtaining 1 is $|\beta|^2$, and then the state becomes $|1\rangle$.

Fourthly, the state space of a composite quantum system is the tensor product of the state spaces of the subsystems. If we have n quantum systems, the states of which are $|\psi_1\rangle, |\psi_2\rangle, ..., |\psi_n\rangle$, respectively, then the joint state of the total system is $|\psi_1\rangle \otimes |\psi_2\rangle \otimes ... \otimes |\psi_n\rangle$, abbreviated as $|\psi_1\rangle\,|\psi_2\rangle ...\,|\psi_n\rangle$.

Appendix B Preparation of Input Quantum States

We show the preparation of quantum state $|F^T \vec{e_2}\rangle$ in the following alogrithm.

* Corresponding author (email: lilvzh@mail.sysu.edu.cn)

Algorithm B1 Preparation of input quantum states of QTSVM.

Input: Matrix $F \in \mathbb{R}^{m_2*(n+1)}$ with each row F_i stored in quantum RAM by the method we mention above.

Procedure:

1. Similar to the method in [2], call the training data oracles with the state $\frac{1}{\sqrt{m_2}}\sum_{i=0}^{m_2-1}|i\rangle$ to prepare the state
$$|\chi\rangle = \frac{1}{||F||}\sum_{i=0}^{m_2-1}||F_i||\,|F_i\rangle\,|i\rangle\,.$$
2. Perform the Walsh-Hadamard transformation to the second register of $|\chi\rangle$ to get state
$$\frac{1}{||F||}\sum_{i=0}^{m_2-1}||F_i||\,|F_i\rangle\sum_{j=0}^{m_2-1}(-1)^{ij}\,|j\rangle\,.$$
3. Measure the second register. If the measurement result is even (this happens with probability 50%), then the state of first register is $\frac{1}{||F||}\sum_{i=0}^{m_2-1}||F_i||\,|F_i\rangle$. Otherwise, repeat the entire procedure until the measurement result is even.

Output: Quantum state $\frac{1}{||F||}\sum_{i=0}^{m_2-1}||F_i||\,|F_i\rangle$.

Since
$$\frac{1}{||F||}\sum_{i=0}^{m_2-1}||F_i||\,|F_i\rangle = \frac{1}{||F||}\sum_{i=0}^{m_2-1}||F_i||(\frac{F_i}{||F_i||})^T = \frac{1}{||F||}\sum_{i=0}^{m_2-1}F_i^T,$$
and
$$|F^T\vec{e_2}\rangle = \frac{F^T\vec{e_2}}{||F^T\vec{e_2}||} = \frac{\sum_{i=0}^{m_2-1}F_i^T}{||\sum_{i=0}^{m_2-1}F_i^T||} = \frac{1}{||F||}\sum_{i=0}^{m_2-1}F_i^T = \frac{1}{||F||}\sum_{i=0}^{m_2-1}||F_i||\,|F_i\rangle\,,$$
we can prepare $|F^T\vec{e_2}\rangle$ by running Algorithm B1. Similarly, we can prepare $|E^T\vec{e_1}\rangle$.

Appendix C Density matrix exponentiation

We prepare a copy of $|\chi\rangle$ in Algorithm B1 and perform a partial trace operation on the first register to get the density operator
$$\begin{aligned}\mathrm{tr}_1\{|\chi\rangle\langle\chi|\} &= \frac{1}{||F||^2}\sum_{i,j=0}^{m_2-1}||F_i||\cdot||F_j||\langle F_j|F_i\rangle\,|i\rangle\langle j| \\ &= \frac{F^TF}{\mathrm{tr}(F^TF)} = \hat{K_2},\end{aligned}$$
which needs $O(\log m_2 n)$ time. Similarly, we can prepare $\hat{K_1}$ in $O(\log m_1 n)$ time. Because $m > m_1$ and $m > m_2$, the consumed time is $O(\log mn)$. By Trotter's formula [1], we have
$$\begin{aligned}e^{-i\hat{H}_1\Delta t} &= e^{-\frac{i(\frac{1}{c_1}K_1+K_2)}{\mathrm{tr}H_1}\Delta t} \\ &= e^{-\frac{1}{c_1}\frac{iK_1\Delta t}{\mathrm{tr}H_1}}e^{-\frac{iK_2\Delta t}{\mathrm{tr}H_1}} + O(\Delta t^2) \\ &= e^{-i\hat{K_1}\frac{1}{c_1}\frac{\mathrm{tr}K_1}{\mathrm{tr}H_1}\Delta t}e^{-i\hat{K_2}\frac{\mathrm{tr}K_2}{\mathrm{tr}H_1}\Delta t} + O(\Delta t^2).\end{aligned}$$
Since $\frac{\mathrm{tr}K_1}{\mathrm{tr}H_1}$, $\frac{1}{c_2}\frac{\mathrm{tr}K_2}{\mathrm{tr}H_1}$ are constant factors, and $\mathrm{tr}K_1$, $\mathrm{tr}K_2$, $\mathrm{tr}H_1$ can be efficiently estimated [2], $e^{-i\hat{H}_1\Delta t}$ can be simulated in $O(\log mn)$ time with $O(\Delta t^2)$ error. Moreover, $e^{-i\hat{H}_2\Delta t}$ can then be simulated in the same way.

Appendix D Proof of Theorem 1

Theorem 1. The time complexities of Algorithm 1 and 2 are $O(\log mn)$ and $O(\log n)$, respectively, where $m = m_1 + m_2$.

Proof. We consider Algorithm 1 first. In Step 1, it needs $O(\log mn)$ time to prepare $|\chi\rangle$ and $O(\log n)$ time to perform the Walsh-Hadamard transformation. The probability that the result of measuring the second register is even is $1/2$ in Algorithm B1, so the expected number of repetitions required for the entire procedure is constant. Thus, the state preparation time is $O(\log mn)$. In Step 2, we call the quantum algorithm for systems of linear equations [3] to solve the Eq. 8 and Eq. 9, and next we analyze the time complexity and error in this procedure. The errors come from Hamitonian simulation and phase estimation. We denote the error in Hamitonian simulation by ϵ_h, the error in phase estimation ϵ_p. In Hamitonian simulation, we denote the total evolution time by t_0, and the number of evolution steps T. Then the time slice Δt of every step satisfies that $\Delta t = \frac{t_0}{T}$. We need to simulate $e^{-i\tau\hat{H}_1\Delta t}$ and $e^{-i\tau\hat{H}_2\Delta t}$ in this algorithm, where $\tau = 0, 1, 2, ..., T-1$. Since operators $e^{-i\hat{H}_1\Delta t}$ and $e^{-i\tau\hat{H}_2\Delta t}$ can be simulated with $O(\Delta t^2)$ error, operators $e^{-i\tau\hat{H}_1\Delta t}$ and $e^{-i\tau\hat{H}_2\Delta t}$ can be simulated with $O(T\Delta t^2)$ error due to the linear accumulation of error. Since $\Delta t = \frac{t_0}{T}$, we have $\epsilon_h = O(T\Delta t^2) = O(\frac{t_0^2}{T})$, thus evolution steps must satisfy that $T = O(\frac{t_0^2}{\epsilon_h})$. Because it needs $O(\log mn)$ time to simulate $e^{-i\hat{H}_1\Delta t}$, the total time of Hamiltonian simulation is $O(T \cdot \log mn) = O(\frac{t_0^2 \log mn}{\epsilon_h})$. Let $\kappa = \max\{\kappa_1, \kappa_2\}$, where κ_1 and κ_2 are the condition number of $\hat{H}_1$ and $\hat{H}_2$ respectively. In order to make the error of phase estimation no more than ϵ_p, it needs to satisfy that $t_0 = O(\kappa/\epsilon_p)$ [2]. Then we have $O(\frac{t_0^2 \log mn}{\epsilon_h}) = O(\frac{\kappa^2 \log mn}{\epsilon_h \epsilon^2})$. Finally, we need to repeat the algorithm $O(\kappa)$ times in order to get a constant success probability, so the time of solving the equations in Step 2 is $O(\frac{\kappa^3 \log mn}{\epsilon_h \epsilon_p^2})$. Therefore, the time complexity of Algorithm 1 is $O(\frac{\kappa^3 \log mn}{\epsilon_h \epsilon_p^2})$.

Next, we turn our attention to Algorithm 2. In Algorithm 2, it needs $O(\epsilon^{-2})$ copies of $|\vec{w}_i, b_i\rangle$ and $|\tilde{x}\rangle$ to estimate the values of $|\vec{w}_i\vec{x} + b_i|^2$ and $||\vec{w}_i||^2$ for $i = 1, 2$. Since it needs $O(\log n)$ time to construct state $|\tilde{x}\rangle$, the time complexity of Algorithm 2 is $O(\frac{\log n}{\epsilon^2})$.

References

1 Nielsen M A, Chuang I L. Quantum computation and quantum information. Cambridge: Cambridge University Press, 2000. 80-97

2 Rebentrost P, Mohseni M, Lloyd S. Quantum support vector machine for big data classification. Phys Rev Lett, 2014, 113(13):130503

3 Harrow A W, Hassidim A, Lloyd S. Quantum algorithm for linear systems of equations. Phys Rev Lett, 2009, 103(15):150502