

Variational quantum semi-supervised classifier based on label propagation

Yan-Yan Hou(侯艳艳), Jian Li(李剑), Xiu-Bo Chen(陈秀波), and Chong-Qiang Ye(叶崇强)
Citation: Chin. Phys. B, 2023, 32 (**7**): 070309. DOI: 10.1088/1674-1056/acb9fb
Journal homepage: http://cpb.iphy.ac.cn; http://iopscience.iop.org/cpb

What follows is a list of articles you may be interested in

Variational quantum semi-supervised classifier based on label propagation

Yan-Yan Hou(侯艳艳), Jian Li(李剑), Xiu-Bo Chen(陈秀波), and Chong-Qiang Ye(叶崇强)
Chin. Phys. B, 2023, 32 (**7**): 070309. DOI: 10.1088/1674-1056/acb9fb

Improved quantum key agreement protocol with authentication

Ji-Hong Guo(郭继红), Ming-Qiang Bai(柏明强), Xiao-Yan Lei(雷小燕), Jia-Xin Xie(谢佳欣), and Zhi-Wen Mo(莫智文)
Chin. Phys. B, 2023, 32 (**5**): 050310. DOI: 10.1088/1674-1056/ac9b31

Variational quantum simulation of thermal statistical states on a superconducting quantum processer

Xue-Yi Guo(郭学仪), Shang-Shu Li(李尚书), Xiao Xiao(效骁), Zhong-Cheng Xiang(相忠诚), Zi-Yong Ge(葛自勇), He-Kang Li(李贺康), Peng-Tao Song(宋鹏涛), Yi Peng(彭益), Zhan Wang(王战), Kai Xu(许凯), Pan Zhang(张潘), Lei Wang(王磊), Dong-Ning Zheng(郑东宁), and Heng Fan(范桁)
Chin. Phys. B, 2023, 32 (**1**): 010307. DOI: 10.1088/1674-1056/aca7f3

Quantum algorithm for neighborhood preserving embedding

Shi-Jie Pan(潘世杰), Lin-Chun Wan(万林春), Hai-Ling Liu(刘海玲), Yu-Sen Wu(吴宇森), Su-Juan Qin(秦素娟), Qiao-Yan Wen(温巧燕), and Fei Gao(高飞)
Chin. Phys. B, 2022, 31 (**6**): 060304. DOI: 10.1088/1674-1056/ac523a

Quantum annealing for semi-supervised learning

Yu-Lin Zheng(郑玉鳞), Wen Zhang(张文), Cheng Zhou(周诚), and Wei Geng(耿巍)
Chin. Phys. B, 2021, 30 (**4**): 040306. DOI: 10.1088/1674-1056/abe298

Variational quantum semi-supervised classifier based on label propagation

Yan-Yan Hou(侯艳艳)[1,2], Jian Li(李剑)[3,†], Xiu-Bo Chen(陈秀波)[4], and Chong-Qiang Ye(叶崇强)[2]

[1] *College of Information Science and Engineering, ZaoZhuang University, Zaozhuang* 277160, *China*
[2] *School of Artificial Intelligence, Beijing University of Posts and Telecommunications, Beijing* 100876, *China*
[3] *School of Cyberspace Security, Beijing University of Posts and Telecommunications, Beijing* 100876, *China*
[4] *Information Security Center, State Key Laboratory Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing* 100876, *China*

(Received 10 October 2022; revised manuscript received 14 January 2023; accepted manuscript online 8 February 2023)

Label propagation is an essential semi-supervised learning method based on graphs, which has a broad spectrum of applications in pattern recognition and data mining. This paper proposes a quantum semi-supervised classifier based on label propagation. Considering the difficulty of graph construction, we develop a variational quantum label propagation (VQLP) method. In this method, a locally parameterized quantum circuit is created to reduce the parameters required in the optimization. Furthermore, we design a quantum semi-supervised binary classifier based on hybrid Bell and Z bases measurement, which has a shallower circuit depth and is more suitable for implementation on near-term quantum devices. We demonstrate the performance of the quantum semi-supervised classifier on the Iris data set, and the simulation results show that the quantum semi-supervised classifier has higher classification accuracy than the swap test classifier. This work opens a new path to quantum machine learning based on graphs.

Keywords: semi-supervised learning, variational quantum algorithm, parameterized quantum circuit

PACS: 03.67.Ac, 03.67.Lx **DOI:** 10.1088/1674-1056/acb9fb

1. Introduction

Classification, one of the most common problems in machine learning, is devoted to predicting the labels of new input data based on labeled data. Classification algorithms have been widely applied in image processing, speech recognition, and other fields. With the development of artificial intelligence and cloud computing, data in classification tasks are expanding rapidly. Still, most of them are unlabeled and need to be labeled by experienced annotators in advance. As labeling data is expensive and time-consuming work,[1] researchers began to study how to add unlabeled data into the training set and utilize labeled and unlabeled data to build semi-supervised classifiers. Label propagation, a crucial semi-supervised learning method, predicts the labels of unlabeled data based on graphs. Semi-supervised classifiers based on label propagation pay more attention to the internal relevance of data and have good effects on the classification of multiple correlation data. However, as the scale of data grows, creating a graph requires a higher computational cost. Implementing the semi-supervised classifier based on label propagation becomes a challenging task for classical computers.

Quantum machine learning, the intersection of quantum physics and machine learning, offers potential speed-ups over classical machine learning algorithms. Variational quantum algorithms (VQAs) are the dominant strategy in the noisy intermediate-scale quantum (NISQ) era. It is devoted to building hybrid quantum-classical models, where parameterized quantum circuits construct the cost function of issues, and classical computers train the parameters of quantum circuits by minimizing cost functions. In VQAs, quantum devices focus on classically intractable problems and transfer the problems difficult to implement on quantum devices to classical computers. Therefore, VQAs have lower requirements for quantum resources and have become important methods for implementing quantum machine learning tasks. At present, VQAs have been applied in classification,[2–5] clustering,[6,7] generative models,[8–10] dimensionality reduction,[11–13] etc.

Quantum systems represent data in the Hilbert space of exponential dimension. Inspired by the advantages of quantum systems in processing high-dimensional data, researchers proposed a series of quantum classification algorithms.[14–17] Considering the high computational complexity of kernel computation, Rebentrostt[18] offered a quantum support vector machine (QSVM) algorithm. This algorithm adopted quantum matrix inversion[19] and a density matrix exponentiation method[20] to implement binary classifier tasks, and achieved exponential speed-ups over corresponding classical classifiers under certain conditions. Schuld[21] proposed a quantum distance-based classifier, which only used Hadamard gates and two single-qubit measurements to implement binary classification tasks. Blank[22] designed a quantum kernel classifier based on quantum swap test operation, which is called a swap test classifier. This classifier achieves good classification ac-

†Corresponding author. E-mail: lijian@bupt.edu.cn

© 2023 Chinese Physical Society and IOP Publishing Ltd

curacies in non-linear classification tasks.

The label propagation method predicts labels of unlabeled data by minimizing energy function, which is similar to cost function optimization of VQAs. Inspired by the similarity between VQAs and label propagation, we adopt VQAs to design a quantum label propagation method, and further implement a quantum semi-supervised classifier based on predicted labels. Our work has two main contributions. (i) A variable label propagation method based on locally parameterized quantum circuit is designed for the first time. The locally parameterized quantum circuit can be used to implement VQAs with only some unknown parameters. (ii) A classifier based on hybrid Bell and Z bases measurement is designed, and this measurement method reduces circuit depth and is more suitable implemented on NISQ devices. We organize the paper as follows. Section 2 gives a review of classical label propagation. Section 3 outlines the variational quantum label propagation method. Section 4 designs a quantum semi-supervised classifier based on hybrid Bell and Z bases measurement. Section 5 verifies the accuracies of label propagation and the semi-supervised classifier. Finally, we get a conclusion and discuss future research directions.

2. Review of label propagation

Label propagation begins with mapping a data set into an undirected weight graph, where nodes represent data and edges reflect the similarities between data. If two data have a great similarity, the edge between them has a higher weight; otherwise, the edge has a lower weight. Labeled data propagates their labels to the neighboring unlabeled data based on undirected weight graphs. Since a graph corresponds to a matrix, matrix operations can be used to implement semi-supervised learning. Let $D=\{D_l,D_u\}$ denote a training data set, including labeled data $D_l=\{(x_1,y_1),(x_2,y_2),\ldots,(x_l,y_l)\}$ and unlabeled data $D_u=\{x_{l+1},x_{l+2},\ldots,x_{l+u}\}$, where $x_i\in\mathbb{R}^m$ is the i^{th} data described by m real-valued attributes and $y_i\in\{+1,-1\}$ represents the corresponding label. k nearest neighbors method is a common method for constructing undirected weight graphs $G=(V,E)$, where $V=\{x_i\}_{i\in n}$ represents nodes, $E=\{e_{ij}\}_{i,j\in n}$ denotes the edges between the nodes x_i and x_j, and $n=l+u$. If x_j is one of k nearest neighbors of x_i, there is an edge between the nodes x_i and x_j; otherwise, there are no edges.

Let $W=\{w_{ij}\}_{i,j\in n}$ denote the weight matrix of edges,

$$w_{ij}=\begin{cases}1, & \text{if } x_j\in N_k(x_i) \text{ or } x_i\in N_k(x_j),\\ 0, & \text{otherwise}\end{cases} \tag{1}$$

represents the weight of the edge e_{ij}, where $N_k(x_i)$ means the set of k nearest neighbors of x_i. Let $f(x_i)$ represent the predicted label of x_i. Similar samples should have similar labels, the more similar the samples, the smaller the difference between the labels, then the energy function of all training data[1] is

$$\begin{aligned}E(f) &= \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}w_{ij}(f(x_i)-f(x_j))^2\\ &= \frac{1}{2}\sum_{i=1}^{n}d_i f^2(x_i)+\frac{1}{2}\sum_{j=1}^{n}d_j f^2(x_j)\\ &\quad -\sum_{i=1}^{n}\sum_{j=1}^{n}w_{ij}f(x_i)f(x_j)\\ &= \sum_{i=1}^{n}d_i f^2(x_i)-\sum_{i=1}^{n}\sum_{j=1}^{n}w_{ij}f(x_i)f(x_j),\end{aligned} \tag{2}$$

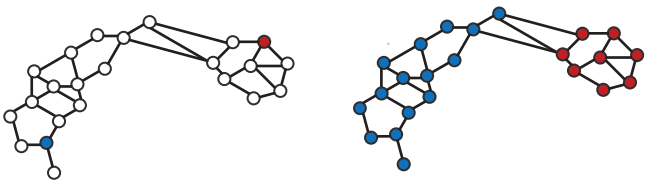
where $d_i=\sum_{j=1}^{n}w_{ij}$ represents the sum of the i^{th} row of W and $D=\text{diag}(d_1,d_2,\ldots,d_n)$ denotes the degree matrix. Equation (2) can be rewritten in the matrix form as

$$E(f)=f^{\text{T}}(D-W)f, \tag{3}$$

where $f=(f_l^{\text{T}};f_u^{\text{T}})$ represents the predicted label vector of all training data. $f_l=(f(x_1),\ldots,f(x_l))$ and $f_u=(f(x_{l+1}),\ldots,f(x_{l+u}))$ correspond to the label vectors of labeled and unlabeled data, respectively. Once $E(f)$ obtains the minimum value, $f(x_i)$ for labeled data will be equal to the correct label y_i, and $f(x_i)$ for unlabeled data will be closest to the correct label. Thus, predicting the labels of unlabeled data can be implemented by solving the optimization problem

$$\min_f(E(f))=\min_f f^{\text{T}}Lf, \tag{4}$$

where $L=D-W$ is the Laplacian matrix. Figure 1 shows a simple example of label propagation.



(a) Before label propagation (b) After label propagation

Fig. 1. Label propagation. Red nodes represent labeled data belonging to the $+1$ class. Blue nodes mean labeled data belonging to the -1 class. White nodes indicate unlabeled data. Panel (a) shows the graph before label propagation. Panel (b) gives the graph after label propagation. After label propagation, labeled data propagate their labels to their neighboring unlabeled data according to the k nearest-neighbors principal.

3. Variational quantum label propagation method

In this section, we reformulate the original label propagation and design a variational quantum label propagation (VQLP) method according to the similarity between the energy function $E(f)$ and cost functions of VQAs.

3.1. Reformulation of label propagation

Let $B=\{b_{i,(p,q)}\}\in\mathbb{R}^{n\times\frac{n(n-1)}{2}}(p<q)$ denote the incidence matrix of the graph G, where each element

$$b_{i,(p,q)}=\begin{cases}w_{pq}, & \text{if } i=p,\\ -w_{pq}, & \text{if } i=q,\\ 0, & \text{if } i\notin\{p,q\}\end{cases}\tag{5}$$

denotes the relationship between the i^{th} node and the $(p,q)^{\text{th}}$ edge.[23] Let $\widetilde{B}=\{\widetilde{b}_{i,(p,q)}\}\in\mathbb{R}^{n\times\frac{n(n-1)}{2}}(p<q)$ mean the normalized incidence matrix, where $\widetilde{b}_{i,(p,q)}=b_{i,(p,q)}/|b_i|$ is the normalized element and $|b_i|$ is the l_2-norm of the i^{th} row of B. The normalized Laplacian matrix $\widetilde{L}=D^{-\frac{1}{2}}LD^{-\frac{1}{2}}$ is equal to $\widetilde{B}\widetilde{B}^{\mathrm{T}}$, then the energy function $E(f)$ can be further written as

$$E(f)=f^{\mathrm{T}}(\widetilde{B}\widetilde{B}^{\mathrm{T}})f,\tag{6}$$

where $\widetilde{B}^{\mathrm{T}}$ denotes the transpose of $\widetilde{B}$.

Many encoding methods can be used to prepare the states for classical vectors and matrices. Amplitude encoding is an efficient encoding that encodes n-dimensional data into $\log(n)$ qubits states. Researchers proposed some efficient methods[24,25] to implement amplitude encoding. Using these amplitude methods, n-qubit states can be prepared with $O(n)$ hardware resources and $O(\log^2(n))$ operations.[26] In the quantum semi-supervised classifier, amplitude encoding is adopted to prepare the states for matrices and vectors. Based on amplitude encoding, the normalized incidence matrix $\widetilde{B}$ is encoded into $|\psi\rangle=\sum_{i=1}^{n}\sum_{j=1}^{n(n-1)/2}b_{ij}|i\rangle_1|j\rangle_{2,3}$, where the subscripts 1, 2 and 3 represent three registers. Let $\rho=|\psi\rangle\langle\psi|$ be the density matrix of $|\psi\rangle$, then $E(f)$ can be written as

$$E(f(\theta))=\langle\varphi_f(\theta)|\mathrm{tr}_{2,3}(\rho)|\varphi_f(\theta)\rangle=\langle\varphi_f(\theta)|\rho_1|\varphi_f(\theta)\rangle,\tag{7}$$

where the reduced density matrix $\rho_1=\mathrm{tr}_{2,3}(\rho)$ is got by tracing registers 2 and 3, proportional to the matrix $\widetilde{L}=\tilde{B}\tilde{B}^{\mathrm{T}}$. $|\varphi_f(\theta)\rangle$ is the state of the label vector f built by the parameterized quantum circuit, where θ is the trainable parameter set. According to the properties of energy function,[1] the optimal label vector $|\varphi_f(\theta)\rangle$ is obtained when $E(f(\theta))$ gets the minimum value. $E(f(\theta))$ satisfies the basic requirements of the cost function and can be estimated by unitary operations and measurements in quantum systems.[27] Thus, the cost function of label propagation is defined as $C(\theta)=E(f(\theta))$.

3.2. The overall structure of the VQLP algorithm

The VQLP algorithm adopts an iterative optimization method to predict the optimal label vector. In each iteration, the work includes evaluation cost function, learning parameters, and predicting label vectors. Figure 2 shows the overall structure of the VQLP algorithm. In the first stage, the state $|\psi\rangle$ for the incidence matrix $\widetilde{B}$ is prepared by conditionally accessing the state of weight matrix W. The label vector $|\varphi_f(\theta)\rangle$ represents labels of labeled data and unlabeled data. As only the labels of unlabeled data are unknown, we design a locally parameterized quantum circuit $V(\theta)$ to build the label vector $|\varphi_f(\theta)\rangle$. After preparing $|\psi\rangle$ and $|\varphi_f(\theta)\rangle$, a hybrid Bell and Z bases measurement, called as U_2, is applied on $|\psi\rangle$ and $|\varphi_f(\theta)\rangle$ to construct the cost function $C(\theta)$. The label vector $|\varphi_f(\theta)\rangle$ based on initial parameters may not correspond to the correct labels of training data. Thus, the second work is to search for the optimal label vector. The cost function value $C(\theta)$ is transmitted to a classical optimizer and minimized by tuning the parameters θ. Once $C(\theta)$ reaches the maximum iteration number τ or is less than the specified error threshold ε_τ, the optimal parameters θ^* are gotten. $V(\theta)$ consists of parameterized quantum circuit for building labels of unlabeled data and unparameterized quantum circuit for building labels of labeled data. In the third stage, parameterized quantum circuit $V'(\theta^*)$ (partial circuit of $V(\theta)$) acts on the initial state $|0\rangle$ to construct the label vector $|\varphi_u(\theta^*)\rangle$ for unlabeled data. Algorithm 1 shows the outline of the VQLP algorithm.

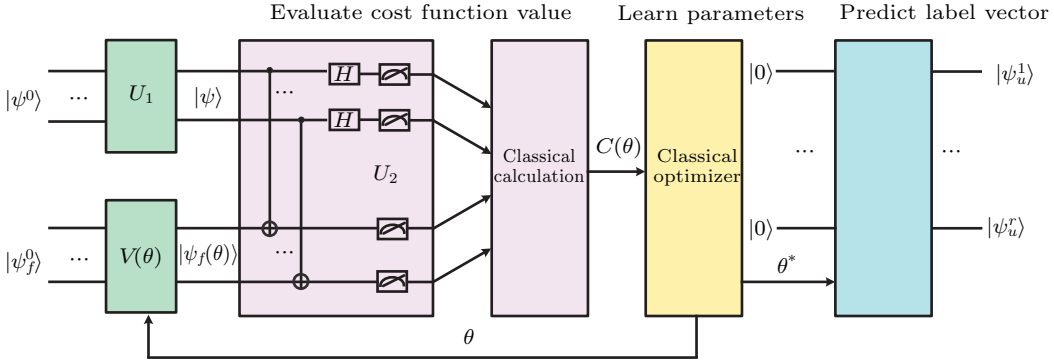


Fig. 2. The overall structure of the VQLP algorithm. U_1 acts on the initial state $|\psi^0\rangle$ to construct the state $|\psi\rangle$ of the incidence matrix $\widetilde{B}$. The parameterized quantum circuit $V(\theta)$ acts on the initial state $|\varphi_f^0\rangle$ to produce the state $|\varphi_f(\theta)\rangle$ of the label vector f. Cost function $C(\theta)$ is gotten by performing the unitary operation U_2 followed by classical post-processing, where U_2 is responsible for computing the Hibert–Schmidt inner product between $|\psi\rangle$ and $|\varphi_f(\theta)\rangle$. In each iteration, the classical computer minimizes $C(\theta)$ to get the optimized parameters θ. Once the optimal parameter θ^* is obtained, the ansatz $V'(\theta^*)$ acts on the state $|0\cdots0\rangle^r$ to build the label vector $|\varphi_u(\theta^*)\rangle=|\varphi_u^r\cdots\varphi_u^1\rangle$ for unlabeled data, where $|\varphi_u^i\rangle$ represents the i^{th} qubit of $|\varphi_u(\theta^*)\rangle$.

Algorithm 1 Variational quantum label propagation (VQLP) algorithm

Input: input states $|\psi^0\rangle$ and $|\varphi_f^0\rangle$, maximum number of iterations τ, initial parameters θ

Output: label vector $|\varphi_u(\theta^*)\rangle$ of unlabeled data

1. Initialize iteration number $t=1$.
2. Prepare the state $|\psi\rangle$ for the incidence matrix $\widetilde{B}$.
3. **while** $t<\tau$
4. Construct the state $|\varphi_f(\theta)\rangle$ for the label vector f.
5. Build the cost function $C(\theta)$ by computing $|\langle\psi|\varphi_f(\theta)\rangle|^2$.
6. Pass the cost function value to the classical optimizer.
7. Optimize the cost function $C(\theta)$ by tuning the parameters θ.
8. Update iteration number $t=t+1$.
9. Compute the label vector $|\varphi_u(\theta^*)\rangle$ of unlabeled data.

Return The label vector $|\varphi_u(\theta^*)\rangle$

3.3. Construct normalized incidence matrix

In this subsection, our primary work is to construct the state $|\psi\rangle$ for the normalized incidence matrix $\widetilde{B}$. To compensate for scaling effects, we firstly standardize all training data to zero mean and unit variance, then normalize them into unit vectors. Let $|\phi_{x_p}\rangle=\frac{1}{|x_p|}\sum_{j=1}^{m}x_{pj}|j\rangle$ represent the amplitude encoding of the training data x_p, where $|\cdot|$ represents l_2-norm and x_{pj} means the j^{th} element of x_p. Reference [28] proposed a quantum algorithm for estimating Euclidean distances between quantum states, and this algorithm corresponds to a mapping $|p\rangle|q\rangle|0\rangle\mapsto|p\rangle|q\rangle|d^2(\phi_{x_p},\phi_{x_q})\rangle$, where $d^2(\phi_{x_p},\phi_{x_q})=|\phi_{x_p}-\phi_{x_q}|^2$ represents the square of the Euclidean distance between $|\phi_{x_p}\rangle$ and $|\phi_{x_q}\rangle$. The first work of constructing incidence matrix is to prepare the weight matrix W by the k nearest-neighbor method. To implement this work, we prepare the superposition state $\frac{1}{\sqrt{n}}\sum_{q=1}^{n}|p\rangle|q\rangle|d^2(\phi_{x_p},\phi_{x_q})\rangle$ of Euclidean distances between $|\phi_{x_p}\rangle$ and $|\phi_{x_q}\rangle$, and search k minimum of $|d^2(x_p,x_q)\rangle$ by the minimum search method. After searching k minimum values of all training data, the state

$$|\psi_W\rangle=\frac{1}{\sqrt{nk}}\sum_{p=1}^{n}\sum_{q=1}^{n}w_{pq}|p\rangle|q\rangle \tag{8}$$

for weight matrix W is constructed, where w_{pq} describes the neighbor relationship between x_p and x_q. If x_q is one of k nearest neighbors of x_q or x_p is one of k nearest neighbors of x_q, $w_{pq}=1$; otherwise, $w_{pq}=0$.

According to Eq. (15), the elements of incidence matrix $\widetilde{B}$ come from W. The second work is to build the state $|\psi\rangle$ for the incidence matrix $\widetilde{B}$ based on $|\psi_W\rangle$. We prepare the state

$$\begin{aligned}|\psi^1\rangle=&\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}|i\rangle_1\right)\\&\otimes\left(\sum_{p=1}^{n}\sum_{q=1}^{n}\frac{w_{pq}}{\sqrt{nk}}|p\rangle_2|q\rangle_3\right)\otimes|0\rangle_4|0\rangle_5|0\rangle_6\end{aligned} \tag{9}$$

as input, where register 1 stores the indexes of nodes, and registers 2 and 3 store the indexes of w_{pq}. Registers 4, 5, and 6 with initial value $|0\rangle|0\rangle|0\rangle$ will store comparison results. The specific steps of building the incidence matrix $\widetilde{B}$ are as follows.

(1) Do comparison operation (U_{C})[28] on registers 2 and 3, where the comparison result is stored in registers 4 and 6, and yield the state

$$\begin{aligned}|\psi^2\rangle=&\sum_{i=1}^{n}\widetilde{w}_{pq}|i\rangle_1\sum_{p<q}|p\rangle_2|q\rangle_3|0\rangle_4|0\rangle_5|0\rangle_6\\&+\sum_{i=1}^{n}\widetilde{w}_{pq}|i\rangle_1\sum_{p\geq q}|p\rangle_2|q\rangle_3|1\rangle_4|0\rangle_5|1\rangle_6,\end{aligned} \tag{10}$$

where $\widetilde{w}_{pq}=w_{pq}/\sqrt{n^2k}$ represents the normalized element of W. U_{C} obeys

$$U_{\text{C}}|p\rangle_2|q\rangle_3|0\rangle_4|0\rangle_6=\begin{cases}|p\rangle_2|q\rangle_3|0\rangle_4|0\rangle_6, & p<q,\\ |p\rangle_2|q\rangle_3|1\rangle_4|1\rangle_6, & p\geq q.\end{cases} \tag{11}$$

Measure register 6 with Z basis, if the measurement result is 0, then get the state

$$|\psi^3\rangle=\textstyle\sum_{i=1}^{n}\widetilde{w}_{pq}|i\rangle_1\sum_{p<q}|p\rangle_2|q\rangle_3|0\rangle_4|0\rangle_5. \tag{12}$$

(2) Extract the elements of the incidence matrix $\widetilde{B}$. Equality comparison is firstly applied on registers 1 and 2. If registers 1 and 2 have the same value, register 5 is set to $|1\rangle$. If register 5 is still $|0\rangle$ after performing the equality comparison, another equality comparison is performed on registers 1 and 3. If registers 1 and 3 have the same value, register 4 is set to $|1\rangle$. Through two equality comparisons, we get the state

$$\begin{aligned}|\psi^4\rangle=&\sum_{i=1}^{n}\widetilde{w}_{pq}|i\rangle_1\sum_{p=i,p<q}|p\rangle_2|q\rangle_3|0\rangle_4|1\rangle_5\\&+\sum_{i=1}^{n}\widetilde{w}_{pq}|i\rangle_1\sum_{q=i,p<q}|p\rangle_2|q\rangle_3|1\rangle_4|0\rangle_5\\&+\sum_{i=1}^{n}\widetilde{w}_{pq}|i\rangle_1\sum_{p,q\neq i;p<q}|p\rangle_2|q\rangle_3|0\rangle_4|0\rangle_5.\end{aligned} \tag{13}$$

(3) Perform CNOT operation on registers 5 and 4, where register 4 serves as control register, then measure register 5 with Z basis. If the measurement result is 1, the state

$$\begin{aligned}|\psi^5\rangle=&\sum_{i=1}^{n}\widetilde{w}_{pq}|i\rangle_1\sum_{p=i,p<q}|p\rangle_2|q\rangle_3|0\rangle_4\\&+\sum_{i=1}^{n}\widetilde{w}_{pq}|i\rangle_1\sum_{q=i,p<q}|p\rangle_2|q\rangle_3|1\rangle_4\end{aligned} \tag{14}$$

is gotten. All elements of $\widetilde{B}$ have been stored in the amplitudes of $|\psi^5\rangle$ by the above steps. If we want to get the state for $\widetilde{B}$ based on $|\psi^5\rangle$, the second term of $|\psi^5\rangle$ should take the opposite number according to Eq. (15).

(4) Act Hadamard operation on register 4, then measure it into $|1\rangle$. The system yields the state

$$\begin{aligned}|\psi\rangle=&\frac{1}{\sqrt{2}}\sum_{i=1}^{n}(\widetilde{w}_{pq}|i\rangle_1\sum_{p=i,p<q}|p\rangle_2|q\rangle_3\\&-\widetilde{w}_{pq}|i\rangle_1\sum_{q=i,p<q}|p\rangle_2|q\rangle_3),\end{aligned} \tag{15}$$

corresponding to the normalized incidence matrix $\widetilde{B}$. Figure 3 shows the circuit implementation, and this circuit corresponds to the module U_1 in Fig. 1.

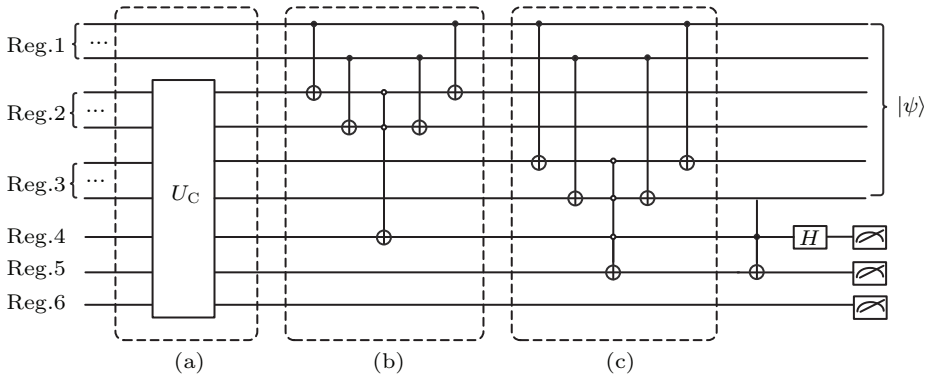


Fig. 3. Circuit of constructing the incidence matrix. Reg.1–Reg.6 represent registers 1–6. The circuit in the dotted box (a) represents the comparison operation. The circuit in the dotted box (b) represents the equality comparison between Reg.1 and Reg.2, and the circuit in the dotted box (c) denotes the equality comparison between Reg.1 and Reg.3. The output state $|\psi\rangle$ corresponds to the incidence matrix $\widetilde{B}$.

3.4. Build label vector

In this subsection, our primary work is to build the state $|\varphi_f(\theta)\rangle$ of the predicted label vector $f=(f_l,f_u)$. As the correct label vector $(y_1,\ldots,y_l)$ of labeled data is known, the predicted label vector f_l does not need to be updated in the optimization process. Let $|\varphi_l\rangle=\frac{1}{\sqrt{l}}\sum_{i=1}^{l}y_i|i\rangle$ be the amplitude encoding of f_l, where $y_i\in\{+1,-1\}$. As the predicted label vector f_u is unknown, we adopt parameterized quantum circuit (ansatz) to prepare the state

$$|\varphi_u(\theta)\rangle=V'(\theta)|0\rangle=\frac{1}{\sqrt{N_u}}\sum_{i=l+1}^{l+u}\widetilde{y}_i|i\rangle \tag{16}$$

for f_u, where $\widetilde{y}_i\in[-1,1]$ denotes the predicted label of x_i and $N_u=\sum_{i=l+1}^{l+u}(\widetilde{y}_i)^2$. For the random initial parameters θ, $|\varphi_u(\theta)\rangle$ is usually inconsistent with the correct label vector, one needs to adjust the parameters so that $|\varphi_u(\theta)\rangle$ is close to the correct label vector. The construction of the label vector $|\varphi_f(\theta)\rangle$ is as follows.

Given the state $|\varphi_f^0\rangle=\frac{1}{\sqrt{2}}|0\rangle_7(|0\rangle+|1\rangle)_8$ as the input, where register 7 is the index register, and register 8 is the flag register to distinguish $|\varphi_l\rangle$ and $|\varphi_u(\theta)\rangle$. Firstly, $|\varphi_l\rangle$ and $|\varphi_u(\theta)\rangle$ are accessed conditional on register 8, and the state

$$|\varphi_f^1(\theta)\rangle=\frac{1}{\sqrt{2l}}\sum_{i=1}^{l}y_i|i\rangle_7|1\rangle_8+\frac{1}{\sqrt{2N_u}}\sum_{i=l+1}^{l+u}\widetilde{y}_i|i\rangle_7|0\rangle_8 \tag{17}$$

is gotten. In general, the number of unlabeled data is not less than the number of labeled data in semi-supervised learning. To make two terms of $|\varphi_f^1(\theta)\rangle$ meet a particular proportion, the controlled rotation operation $R_y(2\alpha)$ is applied on added register 9 conditional on register 8 being $|1\rangle$, where $\alpha=\arccos(\sqrt{l}/\sqrt{u})$. This operation yields the state

$$\begin{aligned}|\varphi_f^2(\theta)\rangle=&\frac{1}{\sqrt{2l}}\sum_{i=1}^{l}y_i|i\rangle_7|1\rangle_8\left(\frac{\sqrt{l}}{\sqrt{u}}|0\rangle+\frac{\sqrt{u-l}}{\sqrt{u}}|1\rangle\right)_9\\&+\frac{1}{\sqrt{2N_u}}\sum_{i=l+1}^{l+u}\widetilde{y}_i|i\rangle_7|0\rangle_8|0\rangle_9.\end{aligned} \tag{18}$$

Subsequently, measure register 9 with the Z basis, if the measurement result is 0, get the state

$$\begin{aligned}|\varphi_f^3(\theta)\rangle=&\frac{1}{\sqrt{l+u}}\sum_{i=1}^{l}y_i|i\rangle_7|1\rangle_8\\&+\frac{\sqrt{u}}{\sqrt{N_u(l+u)}}\sum_{i=l+1}^{l+u}\widetilde{y}_i|i\rangle_7|0\rangle_8,\end{aligned} \tag{19}$$

where $y_i\in\{+1,-1\}$ and $N_u=\sum_{i=l+1}^{l+u}(\widetilde{y}_i)^2$. As register 8 is redundant, apply Hadamard operation on register 8 followed by measuring it with Z basis, if the measurement result is 0, finally, get the state

$$|\varphi_f(\theta)\rangle=\frac{1}{\sqrt{l+u}}\sum_{i=1}^{l}y_i|i\rangle_7+\frac{\sqrt{u}}{\sqrt{N_u(l+u)}}\sum_{i=l+1}^{l+u}\widetilde{y}_i|i\rangle_7, \tag{20}$$

where the amplitude of $|\varphi_f(\theta)\rangle$ is proportional to the predicted label vector f. Figure 4 shows the circuit of constructing $|\varphi_f(\theta)\rangle$, where U_l and $V'(\theta)$ are used to build the labels for labeled and unlabeled data, respectively. As part of the circuit is parameterized, the circuit is called as locally parameterized quantum circuit, corresponding to the module $V(\theta)$ of Fig. 1.

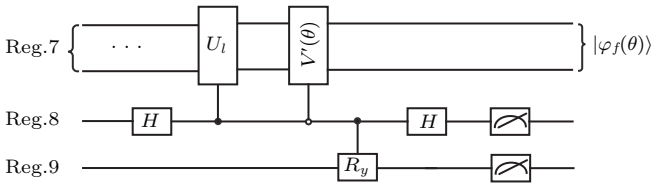


Fig. 4. Circuit of constructing the label vector $|\varphi_f(\theta)\rangle$. Reg.7, Reg.8, and Reg.9 represent registers 7, 8, and 9, respectively. U_l means quantum access to labeled vector $|\varphi_l\rangle$, and $V'(\theta)$ is the parameterized quantum circuit for building $|\varphi_u(\theta)\rangle$.

Many ansatzes can be used to implement $V'(\theta)$. Hardware-efficient ansatz uses lower circuit depth and fewer parameters to represent the solution space of problems,[29,30] and we adopt this ansatz to build $V'(\theta)$. Hardware-efficient ansatz adopts a layered layout. Each layer consists of multiple 2-qubit unitary modules $V'_{ik}(\theta_{ik})=\prod_\mu W_\mu\,\mathrm{e}^{-\mathrm{i}\theta_i^\mu H_\mu}$, where $\mathrm{e}^{-\mathrm{i}\theta_i^\mu H_\mu}$ represents 1-qubit parameterized gate, H_μ is Hermitian operator, and W_μ represents unparameterized gate. Usually, the unitary module $V'_{ik}(\theta_{ik})$ includes multiple 1-qubit parameterized gates, then $\theta_{ik}=\{\theta_i^1,\theta_i^2,\ldots,\theta_i^\mu,\ldots\}$. In the parameters optimization process, the error of $|\varphi_u(\theta^*)\rangle$ decreases

exponentially with the layers of the ansatz $V'(\theta^*)$ increasing. To get the exact number of layers, we first prepare the ansatz with fewer layers and gradually increase the layers until the ansatz satisfies the specified error tolerance.[31] With increased data scale, Hardware-efficient ansatz shows an exponentially vanishing gradient (barren plateau). The VQLP algorithm adopts the alternating layered layout to solve the barren plateau problem. In this layout, the entangled gates in each layer only act on local qubits,[32] and the cost function is the combination of local functions, so the ansatz uses shallower circuit depth to solve the vanishing gradient problem. Figure 5 shows the circuit implementation, where $\mathrm{e}^{-\mathrm{i}\theta_i^{\mu}H_{\mu}}$ and W_{μ} are implemented by by rotation R_y and CNOT, respectively.

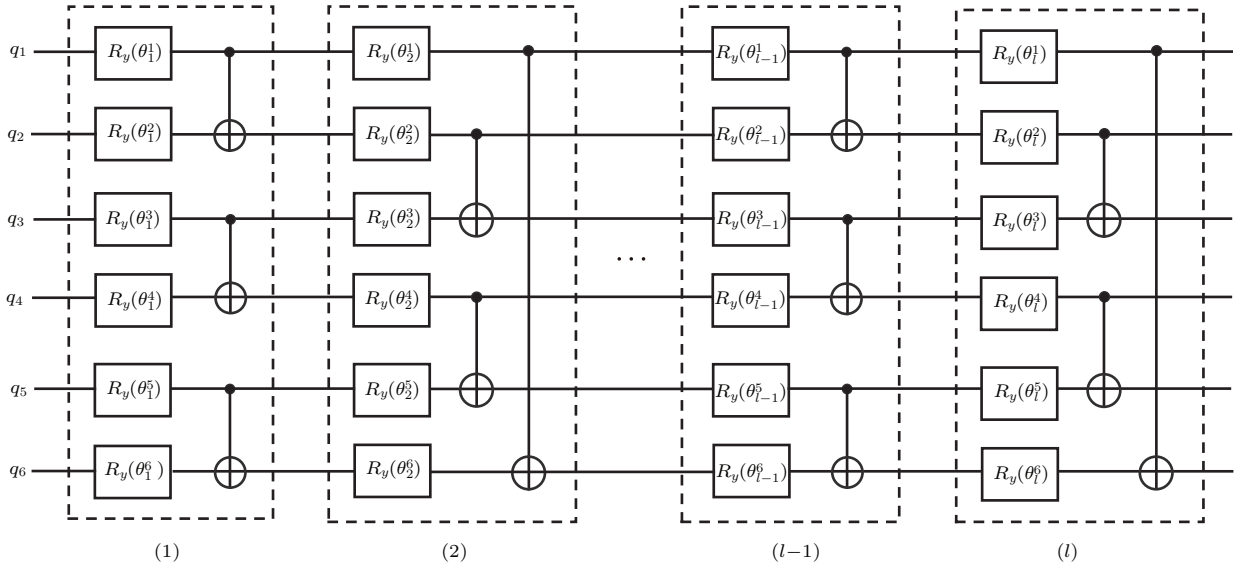


Fig. 5. Parameterized quantum circuit $V'(\theta)$ (6-qubit input). This circuit includes l layers, where $\{q_1,\ldots,q_6\}$ represents the qubit sequence of $|\varphi_u(\theta)\rangle$. The i^{th} dashed box indicates the unitary operation in the i^{th} layer. Each layer is composed of multiple unitary modules $V'_{ij}(\theta_i^j)$, consisting of single qubit rotations $R_y(\theta_i^j)$ and CNOT gates acting on neighboring qubit pairs. The circuit uses alternating layered layout, and unitary modules of adjacent layers act on alternating qubit pairs.

3.5. Compute cost function

After getting the incidence matrix $|\psi\rangle$ and the label vector $|\varphi_f(\theta)\rangle$, the sequent work is to compute the cost function $C(\theta)$. According to $\rho_1 = \mathrm{tr}_{2,3}(\rho)$, the cost function in Eq. (7) can be rewritten as

$$C(\theta) = \mathrm{tr}_{2,3}(|\varphi_f(\theta)\rangle\langle\varphi_f(\theta)||\psi\rangle\langle\psi|), \tag{21}$$

which is the Hilbert–Schmidt inner product between $|\varphi_f(\theta)\rangle$ and $|\psi\rangle$. We adopt the Bell basis measurement method[33] to compute the cost function $C(\theta)$. As $|\varphi_f(\theta)\rangle$ is stored in register 7 and $|\psi\rangle$ is stored in registers 1, 2, and 3, Eq. (21) is equal to

$$C(\theta) = \langle CZ\rangle_{1,7} = \mathrm{tr}_{2,3}(H_1|\varphi_f(\theta)\rangle\langle\varphi_f(\theta)||\psi\rangle\langle\psi|), \tag{22}$$

where $H_1 = \frac{1}{g}\bigotimes_{j=1}^{g} CZ_{1,7}^j \bigotimes I_{2,3}$ represents the operator for $|\varphi_f(\theta)\rangle$ and $|\psi\rangle$, $CZ_{1,7}^j$ denotes the controlled-Z operator acting on the j^{th} qubit pair of registers 1 and 7, and $I_{2,3}$ means the identity operator acting on registers 2 and 3.

Figure 6 shows the circuit for computing the cost function $C(\theta)$. First, CNOT operation is applied on qubit pairs of registers 1 and 7, where register 1 serves as control register. Second, Hadamard operation is performed on register 1. Third, all qubit pairs of registers 1 and 7 are measured based on Bell basis. Finally, we get the cost function $C(\theta)$ by classical post-processing. Let $|a_ib_i\rangle^{\otimes g}$ be the reordered qubit sequence of registers 1 and 7, where a_i and b_i represent the i^{th} qubits of registers 1 and 7, respectively, and g is the number of qubits in each register. We define the output probability vector of qubit sequence $|a_ib_i\rangle^{\otimes g}$ as $\overrightarrow{p} = (p_{00}^1, p_{01}^1, p_{10}^1, p_{11}^1) \otimes \cdots \otimes (p_{00}^g, p_{01}^g, p_{10}^g, p_{11}^g)$, where p_j^i represents the probability that the i^{th} qubit pair is j. Let

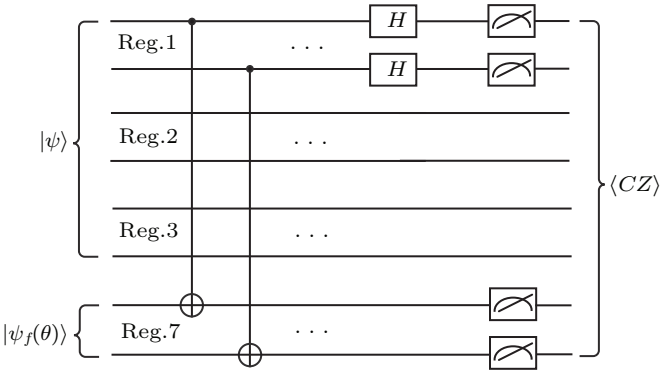


Fig. 6. The circuit of computing the cost function $C(\theta)$. Reg.1, Reg.2, and Reg.3 store the incidence matrix $|\psi\rangle$, and Reg.6 stores the label vector $|\varphi_f(\theta)\rangle$. Perform CNOT operation on registers 1 and 7, followed by Hadamard on register 7. After measuring the 2-qubit operator CZ on registers 1 and 7, the expectation value $\langle CZ\rangle_{1,7}$, corresponding to the cost function value $C(\theta)$, can be obtained by further classical computation.

$\overrightarrow{c} = (1,1,1,-1)^{\otimes g}$ denote the post-processing vector of Bell basis measurement, and the cost function value can be computed by $C(\theta) = \overrightarrow{p} \cdot \overrightarrow{c}$.

4. Semi-supervised binary classifier

Quantum label propagation gets the labels of unlabeled data in training data set. In this section, we design a quantum semi-supervised binary classifier based on all training data. Let $|\phi_{x_i}\rangle$ represent training data, where $i \in \{1,\ldots,n\}$ and $n = l+u$. If $i \in \{1,\ldots,l\}$, $|\phi_{x_i}\rangle$ represents labeled data; otherwise, $|\phi_{x_i}\rangle$ denotes unlabeled data. $f' = \{f'_1, f'_2, \ldots, f'_n\}$ denotes the predicted label vector got by label propagation, where $f'_i = 0$ means x_i belonging to the $+1$ class and $f'_i = 1$ represents x_i belonging to the -1 class. Let $|\phi_{x_*}\rangle$ be test data, and $k_{*i} = |\langle\phi_{x_*}|\phi_{x_i}\rangle|^2$ represents the overlap between $|\phi_{x_*}\rangle$ and $|\phi_{x_i}\rangle$. We adopt the weighted sum of overlaps between test and training data to predict the label

$$\begin{aligned} f_* \;=\; & c_1 \sum_{i=1}^{l} (-1)^{f'_i} |\langle\phi_{x_*}|\phi_{x_i}\rangle|^2 \\ & + c_2 \sum_{i=l+1}^{n} (-1)^{f'_i} |\langle\phi_{x_*}|\phi_{x_i}\rangle|^2 \end{aligned} \tag{23}$$

for the test data $|\phi_{x_*}\rangle$, where c_1 and c_2 are the weight coefficients determined by the importance of labeled and unlabeled data. In semi-supervised learning, labeled data is more important than unlabeled data, then $c_1 > c_2$.

Quantum swap test operation can be used to compute $|\langle\phi_{x_*}|\phi_{x_i}\rangle|^2$, so it is also applied to predict the label f_*. However, this method needs multiple Toffoli gates, and the circuit has a higher depth as input qubits increase. Thus, implementing the classifier based on quantum swap test operation is not easy for current quantum devices. The Bell basis measurement method[33] is a novel method for computing the overlap of quantum states, which has a shallower circuit depth and is easy to implement on NISQ devices. Inspired by this method, we design a hybrid Bell and Z bases measurement method to build f_*.

Given the superposition state

$$\begin{aligned} |\phi_X\rangle_{BCD} \;=\; & \frac{1}{\sqrt{N_X}} \Big(\sum_{i=1}^{l} c_1 |\phi_{x_i}\rangle_B |i\rangle_C |f'_i\rangle_D \\ & + \sum_{i=l+1}^{n} c_2 |\phi_{x_i}\rangle_B |i\rangle_C |f'_i\rangle_D \Big) \end{aligned} \tag{24}$$

for all training data $|x_i\rangle$, where $N_X = lc_1^2 + (n-l)c_2^2$ is the normalization coefficient, $n = l+u$, $|i\rangle$ means the index of training data, $|f'_i\rangle$ denotes the corresponding label, and subscripts B, C, and D represent registers B–D. Firstly, we prepare the state $|\omega^0\rangle = |\phi_{x_*}\rangle_A |\phi_X\rangle_{BCD}$ as input, where $|\phi_{x_*}\rangle_A$ is the test data stored in register A. Let $|\phi_{x_*}\rangle_A = (\alpha_*|0\rangle + \beta_*|1\rangle)_A$ represent test data, where $(\alpha_*)^2 + (\beta_*)^2 = 1$. Let $|\phi_{x_i}\rangle_B = (\alpha_i|0\rangle + \beta_i|1\rangle)_B$ denote training data, where $(\alpha_i)^2 + (\beta_i)^2 = 1$. In order to facilitate the derivation of the following operations, we express the coefficients c_1 and c_2 as the unified form d_i. If $i \in \{1,2,\ldots,l\}$, set $d_i = c_1$; otherwise, $d_i = c_2$, then $|\omega^0\rangle$ can be written as $|\omega^0\rangle = \frac{1}{\sqrt{N_X}} \sum_{i=1}^{n} d_i (\alpha_*|0\rangle + \beta_*|1\rangle)_A (\alpha_i|0\rangle + \beta_i|1\rangle)_B |i\rangle_C |f'_i\rangle_D$. The quantum semi-supervised classifier based on Bell and Z bases measurement can be implemented by the following operations:

(1) Perform CNOT operation on registers A and B, where register A serves as control register, then get the state

$$\begin{aligned} |\omega^1\rangle \;=\; & \frac{1}{\sqrt{N_X}} \sum_{i=1}^{n} d_i (\alpha_*\alpha_i|0\rangle|0\rangle + \beta_*\alpha_i|1\rangle|1\rangle \\ & + \alpha_*\beta_i|0\rangle|1\rangle + \beta_*\beta_i|1\rangle|0\rangle)_{AB} |i\rangle_C |f'_i\rangle_D. \end{aligned} \tag{25}$$

(2) Apply Hadamard operation on register A and get

$$\begin{aligned} |\omega^2\rangle \;=\; & \frac{1}{\sqrt{2N_X}} \sum_{i=1}^{n} d_i ((\alpha_*\alpha_i + \beta_*\beta_i)|0\rangle|0\rangle \\ & + (\beta_*\alpha_i + \alpha_*\beta_i)|0\rangle|1\rangle) + (\alpha_*\alpha_i - \beta_*\beta_i)|1\rangle|0\rangle \\ & + (\alpha_*\beta_i - \beta_*\alpha_i)|1\rangle|1\rangle)_{AB} |i\rangle_C |f'_i\rangle_D. \end{aligned} \tag{26}$$

Through the first two steps, quantum state overlaps have been stored in the amplitudes of $|\omega^2\rangle\langle\omega^2|$.

(3) Measure the expectation value of controlled-Z operator for registers A and B and σ_Z operator for register D, and this operation can be written as $\langle\omega^2|CZ^{AB}\sigma_Z^D|\omega^2\rangle$, where CZ^{AB} denotes controlled-Z operator on registers A and B, and σ_Z^D denotes σ_Z operator on register D. As $CZ^{AB} = (|00\rangle\langle 00| + |01\rangle\langle 01| + |10\rangle\langle 10| - |11\rangle\langle 11|)^{AB}$ and $\sigma_Z^D = (|0\rangle\langle 0| - |1\rangle\langle 1|)^D$, the expectation value is

$$\begin{aligned} & \langle CZ^{AB}\sigma_Z^D\rangle \\ = \; & \frac{1}{2N_X} \sum_{i=1}^{n} (d_i)^2 (\langle\omega^2|000\rangle\langle 000|\omega^2\rangle + \langle\omega^2|010\rangle\langle 010|\omega^2\rangle \\ & + \langle\omega^2|100\rangle\langle 100|\omega^2\rangle - \langle\omega^2|110\rangle\langle 110|\omega^2\rangle \\ & - \langle\omega^2|001\rangle\langle 001|\omega^2\rangle - \langle\omega^2|011\rangle\langle 011|\omega^2\rangle \\ & - \langle\omega^2|101\rangle\langle 101|\omega^2\rangle + \langle\omega^2|111\rangle\langle 111|\omega^2\rangle), \end{aligned} \tag{27}$$

where the superscripts ABD of operators are omitted for simplicity. When $f'_i = 0$, register D is $|0\rangle$; when $f'_i = 1$, register D is $|1\rangle$. According to the two values of register D, Eq. (27) can be rewritten as

$$\begin{aligned} & \langle CZ^{AB}\sigma_Z^D\rangle \\ = \; & \frac{1}{2N_X} \sum_{i=1, f'_i=0}^{n} (d_i)^2 ((\alpha_*\alpha_i + \beta_*\beta_i)^2 + (\beta_*\alpha_i + \alpha_*\beta_i)^2 \\ & + (\alpha_*\alpha_i - \beta_*\beta_i)^2 - (\alpha_*\beta_i - \beta_*\alpha_i)^2) \\ & - \frac{1}{2N_X} \sum_{i=1, f'_i=1}^{n} (d_i)^2 ((\alpha_*\alpha_i + \beta_*\beta_i)^2 + (\beta_*\alpha_i + \alpha_*\beta_i)^2 \\ & + (\alpha_*\alpha_i - \beta_*\beta_i)^2 - (\alpha_*\beta_i - \beta_*\alpha_i)^2) \\ = \; & \frac{1}{N_X} \sum_{i=1}^{n} (d_i)^2 (-1)^{f'_i} (\alpha_*\alpha_i + \beta_*\beta_i)^2. \end{aligned} \tag{28}$$

As $d_i = c_1$ when $i \in \{1,2,\ldots,l\}$, $d_i = c_2$ when $i \in \{l+1, l+2,\ldots,n\}$, and $\alpha_*\alpha_i + \beta_*\beta_i = |\langle\phi_{x_*}|\phi_{x_i}\rangle|$, Eq. (28) can be further written as $\langle CZ^{AB}\sigma_Z^D\rangle = \frac{1}{N_X}\sum_{i=1}^{l} {c_1}^2(-1)^{f_i'}|\langle\phi_{x_*}|\phi_{x_i}\rangle|^2 + \frac{1}{N_X}\sum_{i=l+1}^{n} {c_2}^2(-1)^{f_i'}|\langle\phi_{x_*}|\phi_{x_i}\rangle|^2$, proportional to the label f_*.

Quantum swap test can be used to compute the overlap between two quantum states with n qubits, where the qubits that form states can be entangled. According to the circuit equivalence of the Bell basis measurement and quantum swap test,[22] the Bell basis measurement in the quantum semi-supervised classifier can be generalized to compute the overlap between $|\phi_{x_*}\rangle_A$ and $|\phi_{x_i}\rangle_B$ with n qubits. Figure 7 shows the circuit implementation.

The semi-supervised classifier involves the expectation values of different operators, so the classical post-processing is somewhat complicated. Let $|a_ib_i\rangle_{AB}^{\otimes l}|d\rangle_D$ be the qubit sequence of registers A, B, and D, where a_i and b_i represent the i^{th} qubits of registers A and B sorted in descending order, and d is the qubit of register D. Let p_d^0 denote the probability of d being $|0\rangle$, and $\overrightarrow{g} = (g_{00}^1, g_{01}^1, g_{10}^1, g_{11}^1) \otimes \cdots \otimes (g_{00}^l, g_{01}^l, g_{10}^l, g_{11}^l)$ represents the probability of $|a_ib_i\rangle_{AB}^{\otimes l}$ when d is $|0\rangle$, where g_j^i represents the probability of the i^{th} qubit pair being j. In the same way, let p_d^1 denote the probability of d being $|1\rangle$, and $\overrightarrow{q} = (q_{00}^1, q_{01}^1, q_{10}^1, q_{11}^1) \otimes \cdots \otimes (q_{00}^l, q_{01}^l, q_{10}^l, q_{11}^l)$ represents the probability of $|a_ib_i\rangle_{AB}^{\otimes l}$ when d is $|1\rangle$, where q_j^i means the probability of the i^{th} qubit pair being j. Let $\overrightarrow{d_0} = \overrightarrow{d_1} = (1,1,1,-1)^{\otimes l}$ denote the post-processing vector of the Bell basis measurement. Then the label f_* can be computed based on $f_* = p_d^0 \cdot (\overrightarrow{g} \cdot \overrightarrow{d_0}) - p_d^1 \cdot (\overrightarrow{q} \cdot \overrightarrow{d_1})$.

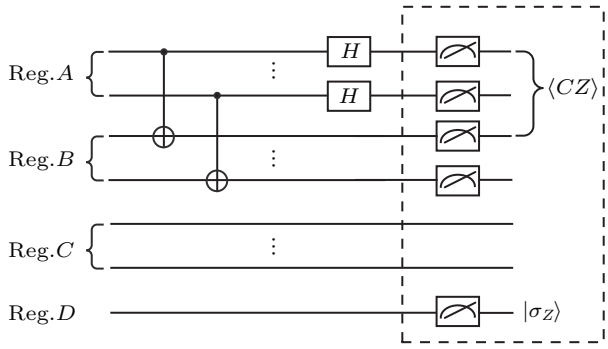


Fig. 7. The circuit of the semi-supervised binary classifier. Reg.A stores $|\phi_{x_*}\rangle$. Reg.B, Reg.C, and Reg.D store $|\phi_{x_i}\rangle$, the index $|i\rangle$, and the corresponding label $|f_i\rangle$, respectively. CNOT is performed on registers A and B, followed by Hadamard operation acts on register A. The label vector f_* is gotten by measuring the expectation value of controlled-Z operator for registers A, B, and the expectation value of σ_Z operator for register D.

The quantum semi-supervised classifier based on the hybrid Bell and Z bases measurement, shown in Fig. 7, contains two layers. CNOT gates implement the first layer, and Hadamard gates implement the second layer. As CNOT gates act on different qubits, all CNOT gates can be executed in parallel. Similarly, Hadamard gates act on different qubits and can also be performed in parallel. No matter how many qubits of training or test data contains, the quantum semi-supervised classifier based on hybrid Bell and Z bases measurement needs only two layers, independent of the size of the classification problem. According to the method proposed in Ref. [24], the quantum semi-supervised classifier can also be implemented by swap test operation, shown in Fig. 8. This circuit requires multiple CNOT and Toffoli gates, which cannot be implemented in parallel. One Toffoli gate requires multiple one-qubit and two-qubit gates to implement. Figure 9 shows Toffoli gate decomposition, where the Toffoli gate is implemented with CNOT, Hadamard, T, and $T^\dagger$ gates. The quantum semi-supervised classifier based on the swap test operation needs 14 layers when training data or test data contains one qubit and $14m$ layers when training data or test data has m qubits. The circuit depth is linear with the size of the classification problem. Compared with the quantum semi-supervised classifier based on swap test operation, our proposed quantum semi-supervised classifier is more suitable for implementation on near-term quantum devices.

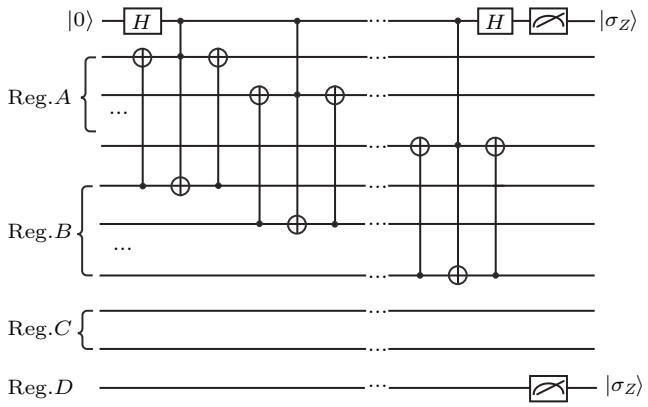


Fig. 8. Semi-supervised classifier based on swap test operation.

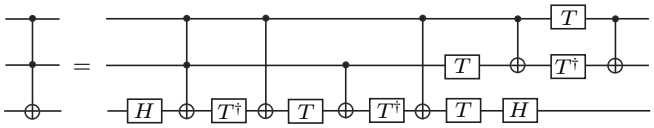


Fig. 9. Toffoli gate decomposition into Hadamard, T, $T^\dagger$, and CNOT gates.

The quantum semi-supervised classifier based on hybrid Bell and Z bases measurement requires more complex classical post-processing operations, which scale linearly with the system size l. From the realization difficulty, the classical post-processing with higher complexity is more easily implemented than the quantum circuit whose depth scales linearly with l. The speed-up of the quantum semi-supervised classifier does not come from transmitting the exponential complexity work to the classical computer but from parallel quantum operations. To reduce the complexity of classical post-processing, we can convert the complex classical post-processing into quantum operations. Figure 10 shows the circuit implementation. This circuit adds an ancilla register and Toffoli gates,

and the expectation value of σ_Z operator for the ancilla register replaces the expectation value of the controlled-Z operator for registers A and B. Compared with the circuit in Fig. 7. This circuit has a higher depth but simpler classical post-processing.

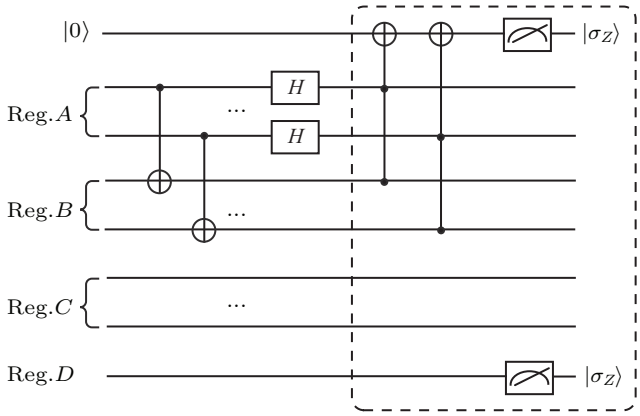


Fig. 10. The quantum semi-supervised binary classifier with simplified classical post processing. Compared with circuit Fig. 7, this circuit converts the complex classical post-processing of the semi-supervised binary classifier into the quantum operation, where f_* is obtained by measuring the expectation value of the σ_Z operator for the ancilla register ($|0\rangle$) and register D. The circuit in the dashed box has the same function as that in Fig. 7.

5. Numerical simulations and performance analysis

In this section, we adopt the Iris dataset to test the performances of the quantum semi-supervised binary classifier based on hybrid Bell and Z bases. The Iris dataset contains 150 samples, where samples 0–49 belong to class 1, samples 50–99 belong to class 2, and samples 100–149 belong to class 3. Classifying samples of classes 2 and 3 is the most difficult task for the Iris dataset, and we mainly analyze this task. We first choose 8 samples $\{x^0,x^1,x^2,x^3,x^4,x^5,x^6,x^7\}$ to demonstrate label propagation, where samples $\{x^0,x^2,x^4,x^6\}$ belong to class 2 and samples $\{x^1,x^3,x^5,x^7\}$ belong to class 3. Let samples $\{x^0,x^1,x^2,x^3\}$ represent labeled data. Assume the labels of samples $\{x^4,x^5,x^6,x^7\}$ are unlabeled data, label propagation is to predict the labels of samples $\{x^4,x^5,x^6,x^7\}$. If the sample x_i belongs to class 2, the label y_i is 1, and if the sample x_i belongs to class 3, the label y_i is -1. Figure 11 shows the predicted labels of samples $\{x^4,x^5,x^6,x^7\}$. Simulation results show that the predicted label $f(x^i)$ is close to the correct label y^i. Let $L=[\frac{1}{2},-\frac{1}{2},\frac{1}{2},-\frac{1}{2}]$ represent the correct normalized label vector for samples $\{x^4,x^5,x^6,x^7\}$, $\widetilde{L}=\frac{1}{2}[f(x^4),f(x^5),f(x^6),f(x^7)]$ represents the predicted label vector, and $S=L\cdot\widetilde{L}^{\mathrm{T}}$ denotes the accuracy of the predicted labels. Figure 12 exhibits the accuracy S under different initial parameters. We can find that the accuracy S reaches 99.5% after about 90 iterations regardless of the initial parameters.

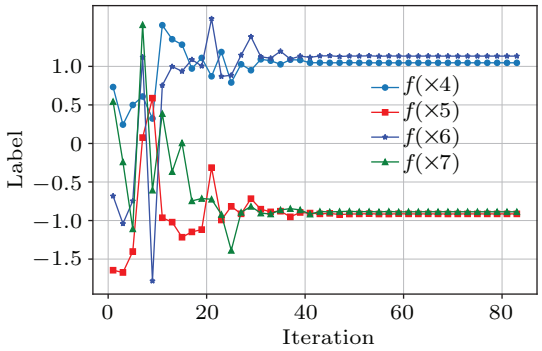


Fig. 11. Predicted labels versus the number of iterations. The classical optimizer is the COBYLA method. Curves represent the estimated labels of unlabeled data.

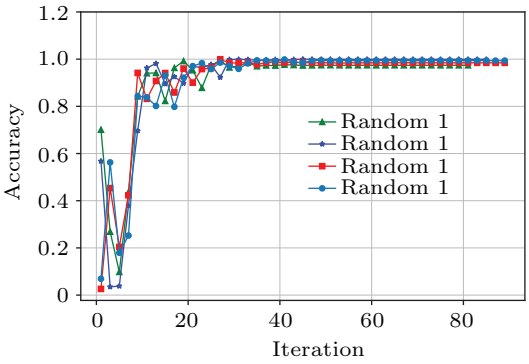


Fig. 12. Accuracy versus the number of iterations. Curves represent the accuracies under four different random initialized parameters.

Our second work is to demonstrate the accuracy of the semi-supervised binary classifier. Figure 13 presents the result for classifying samples of classes 1 and 2, where the predicted label greater than 0 represents the sample belonging to class 1; otherwise, the sample belongs to class 2. The result shows all samples of classes 1 and 2 can be correctly classified. Figure 14 shows the result for classifying samples of classes 1 and 3, where all samples can be correctly classified. Figure 15 shows the classification result of classifying samples of classes 2 and 3. As samples of classes 2 and 3 are not linearly separable, a few classification errors occurred in this task.

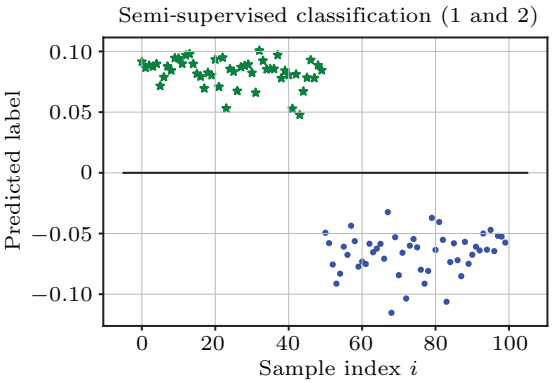


Fig. 13. Classification results of semi-supervised binary classifier (classes 1 and 2). The horizontal axis represents the index i of samples, and the vertical axis shows the predicted label $f(x_i)$. Stars represent samples from class 1, and dots represent samples from class 2.

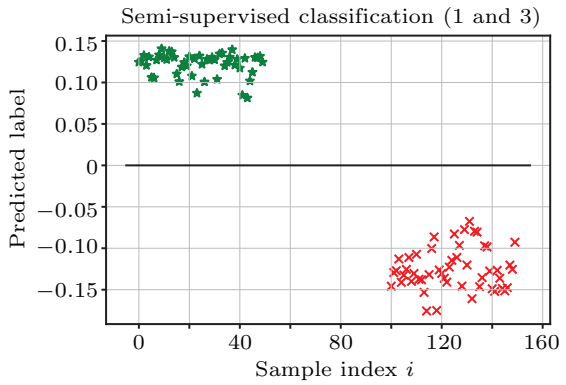


Fig. 14. Classification results of semi-supervised binary classifier (classes 1 and 3). Stars represent samples from class 1, and crosses represent samples from class 3.

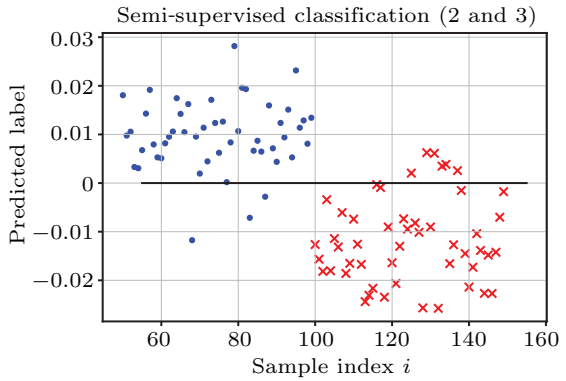


Fig. 15. Classification results of semi-supervised binary classifier (classes 2 and 3). Dots represent samples from class 2, and crosses represent samples from class 3.

Table 1 shows mean accuracies and standard deviations for the quantum semi-supervised binary classifier based on Bell and Z bases (semi-supervised classifier) and the quantum classifier based on swap test operation (swap test classifier)[22] under 10 random initial parameters, where each sample only contains two features. The cells of the format $\pm$ mean accuracy (standard deviation). For the semi-supervised classifier, the mean accuracy of classifying classes 1 and 2 is 100%, and the mean accuracy is also 100% for classifying classes 1 and 3. The mean accuracy of classifying classes 2 and 3 is 90.90%. Still, this accuracy of the semi-supervised classifier is higher than that of the swap test classifier.

Table 1. Classification accuracy and standard deviation (two features).

Classifier	1 and 2	2 and 3	1 and 3
Semi-supervised classifier	100.00 ± 0.00	100.00 ± 0.00	90.90 ± 0.02
Swap test classifier	100.00 ± 0.00	100.00 ± 0.00	87.20 ± 0.03

Table 2. Classification accuracy and standard deviation (four features) .

Classifier	1 and 2	2 and 3	1 and 3
Semi-supervised classifier	100.00 ± 0.00	100.00 ± 0.00	93.30 ± 0.02
Swap test classifier	100.00 ± 0.00	100.00 ± 0.00	88.10 ± 0.03

To improve the separability of classes 2 and 3, we extract four features from the Iris dataset. Table 2 shows mean accuracies and standard deviations for samples containing four features. Compared with Tables 1 and 2, we can find that classifying the Iris dataset containing four features has higher accuracy than classifying the Iris dataset containing two features.

6. Conclusions and future work

In this paper, we adopt a quantum method to implement a quantum semi-supervised binary classifier. By converting the incidence matrix and label vector into quantum states, we design a variational quantum label propagation (VQLP) method. This method utilizes locally parameterized quantum circuits to reduce parameters required in the optimization and is more suitable for implementation on quantum devices. Based on the predicted labels, we further design a quantum semi-supervised classifier based on hybrid Bell and Z bases measurement, which has a shallower circuit depth compared with the swap test classifier. Simulation results show that the VQLP method can predict the labels of unlabeled data with 99.5% accuracy, and the quantum semi-supervised classifier has higher classification accuracy than the swap test classifier. This algorithm assumes quantum operations under noiseless environments. However, hardware noise exists when the semi-supervised learning algorithms are implemented on near-term quantum devices. The quantum semi-supervised classifier under noise environments needs to be researched in future work. Besides, we can further investigate how to adopt multiple data copies to build quantum semi-supervised classifiers based on kernel functions. The design method of the k nearest neighbor graph in VQLP provides a novel idea for creating quantum machine learning models based on graphs. Simultaneously, this research promotes the development of VQAs in quantum semi-supervised learning fields.

Acknowledgements

Project supported by the Open Fund of Advanced Cryptography and System Security Key Laboratory of Sichuan Province (Grant No. SKLACSS-202108), the National Natural Science Foundation of China (Grant No. U162271070), Scientific Research Fund of Zaozhuang University (Grant No. 102061901).

References

[1] Grant E, Benedetti M, Cao S, Hallam A, Lockhart J, Stojevic V, Green A G and Severini S 2018 *NPJ Quantum Inform.* **4** 65
[2] Mitarai K, Negoro M, Kitagawa M and Fujii K 2018 *Phys. Rev. A* **98** 032309
[3] Grant E, Benedetti M, Cao S, Hallam A, Lockhart J, Stojevic V, Green A G and Severini S 2018 *NPJ Quantum Inform.* 4
[4] Havlek V, Crcoles A D, Temme K, Harrow A W, Kandala A, Chow J M and Gambetta J M 2019 *Nature* **567** 209
[5] Hou Y Y, Li J, Chen X B and Tian Y 2022 *Chin. Phys. B* **31** 030304
[6] Benedetti M, Garcia-Pintos D, Perdomo O, Leyton-Ortega V, Nam Y and Perdomo-Ortiz A 2019 *NPJ Quantum Inform.* **5** 1
[7] Mao W W, Lv L H, Ji Y Y, *et al.* 2020 *Chin. Phys. B* **29** 018702
[8] Hamilton K E, Dumitrescu E F and Pooser R C 2019 *Phys. Rev. A* **99** 062323

[9] Zhu D, Linke N M, Benedetti M, Landsman K A, Nguyen N H, Alderete C H, Perdomo-Ortiz A, Korda N, Garfoot A and Brecque C 2019 *Sci. Adv.* **5** eaaw9918

[10] Coyle B, Mills D, Dano V and Kashefi E 2020 *NPJ Quantum Inform.* **6** 1

[11] He X, Sun L, Lyu C and Wang X 2020 *Quantum Inf. Process.* **19** 309

[12] Liang J M, Shen S Q, Li M and Li L 2020 *Phys. Rev. A* **101** 032323

[13] Song X T, Li H W, Yin Z Q, Liang W Y, Zhang C M, Han Y G, Chen W and Han Z F 2015 *Chin. Phys. Lett.* **32** 80302

[14] Adhikary S, Dangwal S and Bhowmik D 2020 *Quantum Inf. Process.* **19** 1

[15] Park D K, Blank C and Petruccione F 2020 *Phys. Lett. A* **384** 126422

[16] Li B M, Hu M L and Fan H 2019 *Acta Phys. Sin.* **68** 030304 (in Chinese)

[17] Zhang S, Duan Q, Li T, *et al.* 2020 *Chin. Phys. B* **29** 010308

[18] Rebentrost P, Mohseni M and Lloyd S 2014 *Phys. Rev. Lett.* **113** 130503

[19] Harrow A W, Hassidim A and Lloyd S 2009 *Phys. Rev. Lett.* **103** 150502

[20] Lloyd S, Mohseni M and Rebentrost P 2014 *Nat. Phys.* **10** 631

[21] Schuld M, Fingerhuth M and Petruccione F 2017 *EPL* **119** 60002

[22] Blank C, Park D K, Rhee J K K and Petruccione F 2020 *NPJ Quantum Inform.* **6** 1

[23] Kerenidis I and Landman J 2021 *Phys. Rev. A* **103** 042415

[24] Sieberer L M and Lechner W 2018 *Phys. Rev. A* **97** 052329

[25] Long G L and Sun Y 2001 *Phys. Rev. A* **64** 014303

[26] Park D K, Petruccione F and Rhee J K K 2019 *Sci. Rep.* **9** 1

[27] Cerezo M, Sone A, Volkoff T, Cincio L and Coles P J 2021 *Nat. Commun.* **12** 1

[28] Wiebe N, Kapoor A and Svore K 2014 Arxiv 1401.2142

[29] Havlicek V, Corcoles A D, Temme K, Harrow A W, Kandala A, Chow J M and Gambetta J M 2019 *Nature* **567** 209

[30] Kandal A, Mezzacapo A, Temme K, Takita M, Brink M, Chow J M and Gambetta J M 2017 *Nature* **549** 242

[31] Zoufal C, Lucchi A and Woerner S 2019 *NPJ Quantum Inform.* **5** 1

[32] Cerezo M, Arrasmith A, Babbush R, Benjamin S C, Endo S, Fujii K, McClean J R, Mitarai K, Yuan X and Cincio L 2021 *Nat. Rev. Phys.* **3** 625

[33] Cincio L, Subasi Y, Sornborger A T and Coles P J 2018 *New J. Phys.* **20** 113022

[34] Garcia-Escartin J C and Chamorro-Posada P 2013 *Phys. Rev. A* **87** 052330