Quantum correlation alignment for unsupervised domain adaptation

Xi He*

Institute of Fundamental and Frontier Sciences, University of Electronic Science and Technology of China, Chengdu, Sichuan 610054, China

(Received 12 May 2020; accepted 17 August 2020; published 16 September 2020)

The correlation alignment algorithm (CORAL), a representative domain adaptation algorithm, decorrelates and aligns a labeled source domain dataset to an unlabeled target domain dataset to minimize the domain shift such that a classifier can be applied to predict the target domain labels. In this paper, we implement the CORAL on quantum devices by two different methods. One method utilizes quantum basic linear algebra subroutines to implement the CORAL with exponential speedup in the number and dimension of the given data samples. The other method is achieved through a variational hybrid quantum-classical procedure. In addition, the numerical experiments of the CORAL with three different types of data sets, namely, the synthetic data, the synthetic-Iris data, and the handwritten digit data, are presented to evaluate the performance of our paper. The simulation results prove that the variational quantum correlation alignment algorithm can achieve competitive performance compared with the classical CORAL.

DOI: 10.1103/PhysRevA.102.032410

I. INTRODUCTION

Quantum computation is demonstrated to have the potential to improve the performance of classical computation problems [1–6]. In addition, quantum computation can be applied to accomplish machine learning tasks with quantum speedup [7–10]. Originally, many quantum shallow machine learning algorithms were proposed such as quantum principal component analysis [11], quantum classification [12–16], quantum data fitting [17,18], quantum clustering [19,20], and quantum dimensionality reduction [21–23]. In recent years, quantum autoencoders [24], quantum Boltzmann machines [25,26], quantum generative adversarial networks [27,28], and quantum feedforward neural networks [29] are the representative quantum deep learning models. For transfer learning, a significant research subfield of machine learning, it can also be combined with quantum computation to implement machine learning tasks in a different but related domain with the acquired knowledge of a well-studied domain [30–32].

In the field of machine learning, labeled data sets are actually dreadfully scarce compared with the available huge amount of unlabeled data. In most cases, the collected unprocessed data are labeled by the extremely time-consuming manual labeling method. Domain adaptation (DA), a crucial research branch of transfer learning, aims to predict the labels of an unprocessed target domain dataset with a labeled source domain dataset [33]. It has various applications in computer vision [34], natural language processing [35], and reinforcement learning [36]. It can be mainly categorized into the semisupervised DA, with few labels in the target domain, and the unsupervised DA, with no labels available in the target domain. For the unsupervised DA, the data distribution adaptation [37–39] which attempts to approximate the data distributions of the source and target domain datasets is one of the most representative domain adaptation methods. In addition, subspace projection [40–42] is another common method for the DA. It first projects the original given data to a specified subspace and subsequently reduces the domain shift by aligning the subspaces. Different from the two methods above, the correlation alignment algorithm (CORAL) [43,44] is a simpler but efficient DA algorithm.

The CORAL first decorrelates the labeled source domain data to eliminate its unique data characteristics. Subsequently, it aligns the decorrelated labeled source domain data to the unlabeled target domain data to reduce the domain shift. The goal of the CORAL is to minimize the discrepancy between the source and target domain datasets by aligning their second-order statistics, namely, the covariance matrices [43]. The CORAL directly aligns the datasets without projecting the data to their corresponding subspaces, resulting in a much more concise procedure than other DA methods. After the data decorrelation and alignment, a classifier will be trained on the aligned labeled source domain dataset and applied to the unlabeled target domain dataset to predict the target domain labels. With the CORAL, the labels of an unprocessed target domain can be obtained efficiently without the need for costly manual labeling. However, the algorithmic complexity of the CORAL can be prohibited with the increase of the number and dimension of the given data.

In our paper, two different types of quantum implementations of the CORAL are presented. One implementation, namely, the quantum basic linear algebra subroutines (QBLAS)-based CORAL, can be performed on a universal quantum computer achieving exponential speedup in the number and dimension of the given data. The other implementation, the variational quantum correlation alignment algorithm (VQCORAL), can be performed on near-term quantum devices through a variational hybrid quantum-classical procedure. Concretely, the VQCORAL can be

*xihe@std.uestc.edu.cn

realized in two different ways called the end-to-end VQCORAL and the matrix-multiplication-based VQCORAL which is inspired from the variational quantum eigensolver [45,46] and the variational quantum matrix multiplication [47]. To evaluate the performance of the VQCORAL, three different numerical experiments are provided. Specifically, the no adaptation (NA) model set as the baseline model, the classical CORAL, and the VQCORAL are the models selected in the experiments. For the two synthetic data sets generated from different distributions, the VQCORAL outperforms the classical CORAL and the NA with a two-qubit eight-layer variational quantum circuit. For the synthetic-Iris data sets [48,49], the VQCORAL also shows outstanding performance with a two-qubit eight-layer parametrized quantum circuit compared to the other two models. For the handwritten digit datasets, namely, the Modified National Institute of Standards and Technology dataset (MNIST) [50] and United States Postal Service dataset (USPS) [51] data sets, the DA procedure can be implemented by an eight-qubit 16-layer parametrized quantum circuit to achieve comparable performance to the classical CORAL and better performance than the baseline model.

The arrangement of this paper is as follows. In Sec. II, the classical CORAL will be briefly overviewed. Subsequently, the quantum correlation alignment algorithm (QCORAL) is presented in Sec. III. The QBLAS-based CORAL and the VQCORAL are shown in Secs. III B and III C, respectively, in detail. Then, the numerical experiments are provided in Sec. IV. Finally, we make a conclusion and discuss some open questions in Sec. V.

II. CLASSICAL CORRELATION ALIGNMENT

Assume a labeled source domain dataset $D_s = \{x_i^{(s)}\}_{i=1}^{n_s} \in \mathbb{R}^D$ with labels $L_s = \{y_i^{(s)}\}_{i=1}^{n_s}$ and an unlabeled target domain dataset $D_t = \{x_j^{(t)}\}_{j=1}^{n_t} \in \mathbb{R}^D$ generated from different data distributions. $X_s = (x_1^{(s)}, \ldots, x_{n_s}^{(s)}) \in \mathbb{R}^{D\times n_s}$ and $X_t = (x_1^{(t)}, \ldots, x_{n_t}^{(t)}) \in \mathbb{R}^{D\times n_t}$ refer to the source and target domain dataset matrices, respectively. Assume u_s (u_t) and C_s (C_t) are the mean and covariance matrix of the source (target) domain, respectively. The data in both domains have been zero-centered, namely, $u_s = u_t = 0$, and normalized but $C_s \neq C_t$. In addition, the data in the CORAL are assumed to depend on a lower-dimensional manifold, meaning that X_s, X_t, C_s, and C_t are all low-rank matrices where r_{C_s} and r_{C_t} represent the rank of C_s and C_t, respectively.

The CORAL attempts to align the covariance matrix of the source domain to the target domain utilizing a linear transformation matrix A [43]. Thus, the objective function of the CORAL is defined as

$$\min_A \|C_{\hat{s}} - C_t\|_F^2 = \min_A \|A^T C_s A - C_t\|_F^2, \tag{1}$$

where $C_{\hat{s}} = A^T C_s A$ is the covariance matrix after the correlation alignment; $\|\cdot\|_F$ represents the Frobenius norm.

Assume $C_s = U_s \Sigma_s U_s^T$ and $C_t = U_t \Sigma_t U_t^T$ are the singular value decomposition (SVD) of C_s and C_t, respectively. The optimal solution of Eq. (1) is $C_{\hat{s}} = U_{t[1:r]} \Sigma_{t[1:r]} U_{t[1:r]}^T$ where $r = \min(r_{C_s}, r_{C_t})$; the diagonal elements of $\Sigma_{t[1:r]}$ are the r largest singular values; the columns of $U_{t[1:r]}$ are the corresponding left-singular vectors. Let

$$C_{\hat{s}} = A^T C_s A = U_{t[1:r]} \Sigma_{t[1:r]} U_{t[1:r]}^T. \tag{2}$$

Then

$$A^T U_s \Sigma_s U_s^T A = U_{t[1:r]} \Sigma_{t[1:r]} U_{t[1:r]}^T. \tag{3}$$

Hence,

$$\left(U_s^T A\right)^T \Sigma_s \left(U_s^T A\right) = E^T \Sigma_s E \tag{4}$$

where $E = \Sigma_s^{+\frac{1}{2}} U_s^T U_{t[1:r]} \Sigma_{t[1:r]}^{\frac{1}{2}} U_{t[1:r]}^T$; $\Sigma_s^{+\frac{1}{2}}$ is the Moore-Penrose pseudoinverse of $\Sigma_s^{\frac{1}{2}}$.

Finally, the optimal solution of A is

$$\begin{aligned} A_* &= U_s E \\ &= \left(U_s \Sigma_s^{+\frac{1}{2}} U_s^T\right)\left(U_{t[1:r]} \Sigma_{t[1:r]}^{\frac{1}{2}} U_{t[1:r]}^T\right). \end{aligned} \tag{5}$$

The first term $U_s \Sigma_s^{+\frac{1}{2}} U_s^T$ decorrelates the source domain dataset. The second term $U_{t[1:r]} \Sigma_{t[1:r]}^{\frac{1}{2}} U_{t[1:r]}^T$ aligns the decorrelated source domain dataset to the target domain dataset.

Therefore, the concrete steps of the CORAL are as follows.

(1) Compute the source domain covariance matrix $C_s = X_s X_s^T$ and the target domain covariance matrix $C_t = X_t X_t^T$.

(2) Decorrelate the source domain data as

$$X_{\tilde{s}} = C_s^{-\frac{1}{2}} X_s. \tag{6}$$

(3) Align the decorrelated source domain data to the target domain data as

$$X_{\hat{s}} = C_t^{\frac{1}{2}} X_{\tilde{s}}. \tag{7}$$

After the CORAL, the source domain data are transformed to the target domain data space. The classifier can be subsequently trained on the aligned source domain data $\{x_i^{(\hat{s})}, y_i^{(s)}\}_{i=1}^{n_s}$ and predict the labels $L_t = \{y_j^{(t)}\}_{j=1}^{n_t}$ of the target domain data D_t. The schematic diagram of the CORAL is presented in Fig. 1.

III. QUANTUM CORRELATION ALIGNMENT

The QCORAL can be implemented in two aspects, based on the quantum basic linear algebra subroutines and the variational hybrid quantum-classical procedure, respectively. In these two implementations, we assume that all the data have been normalized and zero-centered exactly as the classical CORAL.

A. State preparation

Assume the source domain data $X_s = \sum_{i=1}^{n_s} |x_i^{(s)}||x_i^{(s)}\rangle\langle i|$ and the target domain data $X_t = \sum_{j=1}^{n_t} |x_j^{(t)}||x_j^{(t)}\rangle\langle j|$. When the original samples are quantum data, they can be processed directly. We assume that the data stored in the quantum random access memory are accessible as described in Refs. [52,53] when the original data are given in the form of classical vectors, so that the quantum states representing the original data samples can be generated in time $O[\text{poly}\log(D)]$. The

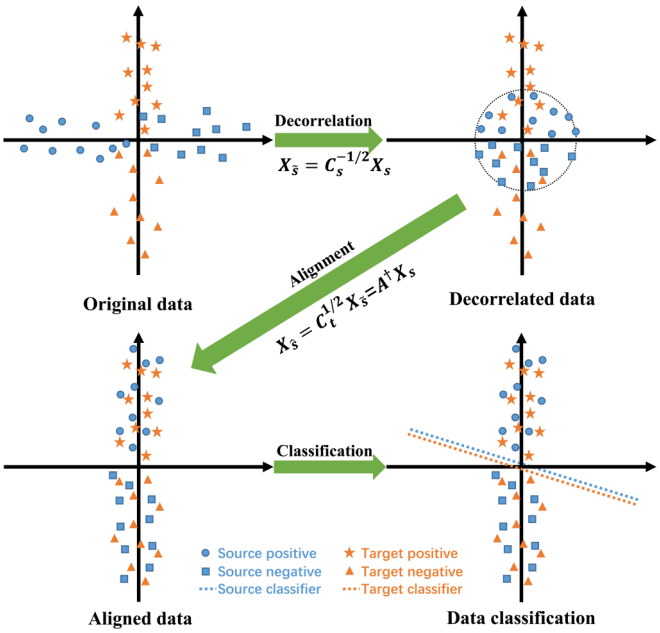

FIG. 1. The schematic diagram of the CORAL.

quantum states representing the source domain data X_s and the target domain data X_t are

$$|\psi_{X_s}\rangle = \sum_{i=1}^{n_s}\sum_{m=1}^{D} x_{mi}^{(s)}|i\rangle|m\rangle = \sum_{i=1}^{n_s}|i\rangle|x_i^{(s)}\rangle, \tag{8}$$

$$|\psi_{X_t}\rangle = \sum_{j=1}^{n_t}\sum_{m=1}^{D} x_{mj}^{(t)}|j\rangle|m\rangle = \sum_{j=1}^{n_t}|j\rangle|x_j^{(t)}\rangle, \tag{9}$$

respectively, in amplitude encoding with $\sum_{m,i}|x_{mi}^{(s)}|^2 = \sum_{m,j}|x_{mj}^{(t)}|^2 = 1$. Hence, the covariance matrices of the source and target domain data can be obtained as

$$\rho_{C_s} = \mathrm{tr}_i\{|\psi_{X_s}\rangle\langle\psi_{X_s}|\} = \sum_{m,m'=1}^{D}\sum_{i=1}^{n_s} x_{mi}^{(s)}x_{m'i}^{(s)*}|m\rangle\langle m'|, \tag{10}$$

$$\rho_{C_t} = \mathrm{tr}_j\{|\psi_{X_t}\rangle\langle\psi_{X_t}|\} = \sum_{m,m'=1}^{D}\sum_{j=1}^{n_t} x_{mj}^{(t)}x_{m'j}^{(t)*}|m\rangle\langle m'|, \tag{11}$$

respectively, by taking the partial trace over the corresponding column register.

B. QBLAS-based QCORAL

The QBLAS-based QCORAL utilizes the quantum basic linear algebra subroutines to implement the data decorrelation and alignment procedure of the CORAL. In the spirit of Ref. [54], the source domain data X_s can be aligned to the target domain data X_t as follows.

Let $X_s = \sum_m \sigma_m^{(s)}|u_m^{(s)}\rangle\langle v_m^{(s)}|$ and $X_t = \sum_m \sigma_m^{(t)}|u_m^{(t)}\rangle\langle v_m^{(t)}|$ be the SVD of X_s and X_t, respectively. The source and target domain data X_s and X_t can be extended to

$$\tilde{X}_s = \begin{bmatrix} 0 & X_s \\ X_s^{\dagger} & 0 \end{bmatrix}, \tag{12}$$

$$\tilde{X}_t = \begin{bmatrix} 0 & X_t \\ X_t^{\dagger} & 0 \end{bmatrix}. \tag{13}$$

With the input state $|0,\psi_{X_s}\rangle|0\rangle^{\otimes\log(D+n_s)}$, the quantum state

$$\sum_{i=1}^{n_s}|i\rangle\sum_{m=1}^{D}\beta_{mi}^{(s)}|\sigma_m^{(s)}\rangle\frac{1}{\sqrt{2}}\left(|w_m^{(s)+}\rangle - |w_m^{(s)-}\rangle\right) = \sum_{i=1}^{n_s}|i\rangle\sum_{m=1}^{D}\beta_{mi}^{(s)}|\sigma_m^{(s)}\rangle|v_m^{(s)}\rangle \tag{14}$$

can be obtained by performing the quantum phase estimation (QPE)

$$\mathbf{U_{PE}}(\tilde{X}_s) = (\mathbf{QFT}^{\dagger}\otimes\mathbf{I})\left(\sum_{\tau=0}^{T-1}|\tau\rangle\langle\tau|\otimes e^{i\tilde{X}_s\tau t/T}\right)(\mathbf{H}^{\otimes n}\otimes\mathbf{I}) \tag{15}$$

as described in Refs. [3,55] where $\beta_{mi}^{(s)} = \langle u_m^{(s)}|x_i^{(s)}\rangle$; $|w_m^{(s)\pm}\rangle = \frac{1}{\sqrt{2}}(|0\rangle|u_m^{(s)}\rangle \pm |1\rangle|v_m^{(s)}\rangle)$ are the eigenvectors of $\tilde{X}_s$ corresponding to the singular value $\sigma_m^{(s)}$; $\mathbf{QFT}^{\dagger}$ represents the inverse quantum Fourier transform; and $\sum_{\tau=0}^{T-1}|\tau\rangle\langle\tau|\otimes e^{i\tilde{X}_s\tau t/T}$ is the conditional Hamiltonian evolution. Subsequently, add a new ancilla qubit and apply the controlled rotation operation $R_y[2\arcsin(\gamma_s/|\sigma_m^{(s)}|)]$ on it, resulting in

$$\sum_{i=1}^{n_s}|i\rangle\sum_{m=1}^{D}\beta_{mi}^{(s)}|\sigma_m^{(s)}\rangle|v_m^{(s)}\rangle|\psi_a^{(s)}\rangle \tag{16}$$

where the ancilla register

$$|\psi_a^{(s)}\rangle = \sqrt{1-\frac{\gamma_s^2}{|\sigma_m^{(s)}|^2}}|0\rangle + \frac{\gamma_s}{|\sigma_m^{(s)}|}|1\rangle \tag{17}$$

γ_s is a constant. By uncomputing the singular value register and measuring the ancilla register to be $|1\rangle$, the decorrelated source domain quantum state

$$|\psi_{X_{\tilde{s}}}\rangle = \sum_{i=1}^{n_s}|i\rangle\sqrt{\frac{1}{\sum_{m=1}^{D}|\gamma_s\beta_{mi}^{(s)}|^2/|\sigma_m^{(s)}|^2}}\sum_{m=1}^{D}\frac{\beta_{mi}^{(s)}\gamma_s}{|\sigma_m^{(s)}|}|v_m^{(s)}\rangle = \sum_{i=1}^{n_s}|i\rangle\frac{C_s^{-\frac{1}{2}}|x_i^{(s)}\rangle}{\sqrt{\langle x_i^{(s)}|C_s^{-\frac{1}{2}\dagger}C_s^{-\frac{1}{2}}|x_i^{(s)}\rangle}} = \sum_{i=1}^{n_s}|i\rangle|x_i^{(\tilde{s})}\rangle \tag{18}$$

representing the decorrelated source domain dataset $X_{\tilde{s}}$ is finally obtained. Hence, the source domain data can be decorrelated in time $O[\|X_s\|_{\max}^2\log^2(D+n_s)/\epsilon^3]$ where $\|X_s\|_{\max}$ is the largest absolute element of X_s and ϵ is the error parameter [54].

Similarly, we then perform the QPE $\mathbf{U_{PE}}(\tilde{X}_t)$ on $|0,\psi_{X_{\hat{s}}}\rangle|0\rangle^{\otimes \log(D+n_t)}$, resulting in

$$\sum_{i=1}^{n_s}|i\rangle\sum_{m=1}^{D}\beta_{mi}^{(t)}|\sigma_m^{(t)}\rangle\frac{1}{\sqrt{2}}\left(|w_m^{(t)+}\rangle-|w_m^{(t)-}\rangle\right)=\sum_{i=1}^{n_s}|i\rangle\sum_{m=1}^{D}\beta_{mi}^{(t)}|\sigma_m^{(t)}\rangle|v_m^{(t)}\rangle, \tag{19}$$

where $\beta_{mi}^{(t)}=\langle u_m^{(t)}|x_i^{(\tilde{s})}\rangle$; $|w_m^{(t)\pm}\rangle=\frac{1}{\sqrt{2}}(|0\rangle|u_m^{(t)}\rangle\pm|1\rangle|v_m^{(t)}\rangle)$ are the eigenvectors of $\tilde{X}_t$ corresponding to the singular value $\sigma_m^{(t)}$. By performing the controlled rotation operation $R_y[2\arcsin(\gamma_t|\sigma_m^{(t)}|)]$ on a newly added ancilla, the quantum state

$$\sum_{i=1}^{n_s}|i\rangle\sum_{m=1}^{D}\beta_{mi}^{(t)}|\sigma_m^{(t)}\rangle|v_m^{(t)}\rangle\left(\sqrt{1-\gamma_t^2|\sigma_m^{(t)}|^2}|0\rangle+\gamma_t|\sigma_m^{(t)}||1\rangle\right) \tag{20}$$

is achieved where γ_t is a constant. Ultimately, the quantum state

$$\begin{aligned}|\psi_{X_{\hat{s}}}\rangle&=\sum_{i=1}^{n_s}|i\rangle\sqrt{\frac{1}{\sum_{m=1}^{D}|\gamma_t\beta_{mi}^{(t)}\sigma_m^{(t)}|^2}}\sum_{m=1}^{D}\beta_{mi}^{(t)}\gamma_t|\sigma_m^{(t)}||v_m^{(t)}\rangle\\&=\sum_{i=1}^{n_s}|i\rangle\frac{C_t^{\frac{1}{2}}|x_i^{(\tilde{s})}\rangle}{\sqrt{\langle x_i^{(\tilde{s})}|C_t^{\frac{1}{2}\dagger}C_t^{\frac{1}{2}}|x_i^{(\tilde{s})}\rangle}}\\&=\sum_{i=1}^{n_s}|i\rangle|x_i^{(\hat{s})}\rangle\end{aligned} \tag{21}$$

can be obtained in $O[\|X_t\|_{\max}^2\log^2(D+n_t)/\epsilon^3]$ where $\|X_t\|_{\max}$ is the largest absolute element of X_t [54]. Therefore, the decorrelated source domain data are aligned to the target domain data.

After the procedure of correlation alignment, the source domain data are aligned to the target domain, resulting in the decorrelated source domain data $D_{\hat{s}}=\{|x_i^{(\hat{s})}\rangle\}_{i=1}^{n_s}$. Subsequently, a quantum classification algorithm such as the local classifier, the quantum k-nearest-neighbor algorithm, or the global classifier (the quantum support vector machine algorithm) can be applied to the decorrelated source domain quantum state $|\psi_{X_{\hat{s}}}\rangle$ with corresponding labels L_s and the target domain quantum state $|\psi_{X_t}\rangle$ directly to obtain the target domain labels L_t.

The pseudocode of the QBLAS-based QCORAL is presented in Algorithm 1. In contrast, the implementation of the classical CORAL involves SVD and matrix multiplication operations resulting in the algorithmic complexity in time $O[\text{poly}(n_s,n_t,D)]$. Thus, the QBLAS-based QCORAL presented in this subsection takes logarithmic resources in the number and dimension of the source and target domain data compared to the classical CORAL.

Algorithm 1. QBLAS-based QCORAL

Input: Source domain data X_s with labels L_s; target domain data X_t.
Output: Target domain labels L_t.
step 1: Apply the QPE $\mathbf{U_{PE}}(\tilde{X}_s)$ on the input state $|0,\psi_{X_s}\rangle|0\rangle^{\otimes\log(D+n_s)}$ resulting in Eq. (14) in $O(1/\epsilon)$ with an error ϵ.
step 2: Add a new ancilla and perform the rotation operation $R_y(2\arcsin(\gamma_s/|\sigma_m^{(s)}|))$ to obtain Eq. (16).
step 3: Uncompute the singular value register $|\sigma_m^{(s)}\rangle$ and measure the ancilla register to be $|1\rangle$ to obtain the decorrelated source domain quantum state $|\psi_{X_{\hat{s}}}\rangle$ as Eq. (18) in $O(\|X_s\|_{\max}^2\log^2(D+n_s)/\epsilon^3)$.
step 4: Perform $\mathbf{U_{PE}}(\tilde{X}_t)$ on $|0,\psi_{X_{\hat{s}}}\rangle|0\rangle^{\otimes\log(D+n_t)}$ resulting in the quantum state as Eq. (19).
step 5: Perform the rotation operation $R_y(2\arcsin(\gamma_t|\sigma_m^{(t)}|))$ on a newly added ancilla to obtain Eq. (20).
step 6: Uncompute the singular value register $|\sigma_m^{(t)}\rangle$ and measure the ancilla to be $|1\rangle$ to achieve the aligned source domain quantum state $|\psi_{X_{\hat{s}}}\rangle$ as Eq. (21) in $O(\|X_t\|_{\max}^2\log^2(D+n_t)/\epsilon^3)$.
step 7: Invoke a classifier to predict the target labels $L_t=Classifier(X_{\hat{s}},L_s,X_t)$.

C. Variational quantum correlation alignment

Although the QBLAS-based QCORAL can be performed on a universal quantum computer with exponential speedup, the implementation critically requires a high-depth quantum circuit and fully coherent evolution. Alternatively, the CORAL can be implemented on near-term noisy intermediate-scale quantum devices with a variational hybrid quantum-classical procedure. The VQCORAL combines quantum computation and classical optimization together to implement the algorithm with parametrized quantum circuits. Compared with the QBLAS-based QCORAL, the VQCORAL can achieve the procedure of correlation alignment with shallow quantum circuits. Specifically, the number of parameters utilized in the VQCORAL to achieve the procedure of correlation alignment is polynomial to the number of input quantum states. However, the cost of the classical CORAL will be exponential in the size of qubits when the original data are given in the form of quantum states. In this subsection, we will present the implementation of the VQCORAL and explore two different specific configurations in detail.

As introduced in Sec. II, the goal of the CORAL is to find a linear transformation matrix A to align the source domain data X_s to the target domain data X_t. Hence, we can approximate the linear transformation by a parametrized quantum circuit $\mathbf{U}_\theta$. The cost function of the VQCORAL can be defined as

$$L_v(\theta)=\left\|\mathbf{U}_\theta\rho_{C_s}\mathbf{U}_\theta^\dagger-\rho_{C_t}\right\|_F^2 \tag{22}$$

where

$$\mathbf{U}_\theta=\mathbf{U}_L(\theta)\cdots\mathbf{U}_l(\theta)\cdots\mathbf{U}_1(\theta) \tag{23}$$

is an L-depth parametrized quantum circuit with a set of parameters $\{\theta\}$. Then, the optimal configuration of the quantum circuit can be obtained by minimizing L_v with the optimization algorithm. Inspired by the classical neural network, this procedure can be called the end-to-end VQCORAL.

In addition to the end-to-end VQCORAL described as above, the matrix-multiplication-based VQCORAL can also

be implemented in two variational procedures successively as follows.

(1) We do not optimize the cost function L_v directly, but compute $C_s^{1/2}$ and $C_t^{1/2}$ by solving the eigenvalues and corresponding eigenvectors of C_s and C_t, respectively, through the variational quantum covariance matrix square-root solver inspired from Refs. [45,46] as presented in Algorithm 2.

Algorithm 2. Variational quantum covariance matrix square root solver

Input: Source domain data X_s with labels L_s; target domain data X_t.
Output: The source domain covariance square root matrix $C_s^{1/2}$ and the target domain covariance square root matrix $C_t^{1/2}$.
step 1: Compute the Hamiltonians $H_s = \rho_{C_s}$, $H_t = \rho_{C_t}$ and subsequently $\tilde{H}_t = \eta I - H_t$ with a specified constant η.
step 2: Prepare the ansatz states $|\psi(\lambda_k^{(s)})\rangle$ with a set of parameters $\{\theta^{(s)}\}$. Minimize the cost function

$$F_s(\lambda_k^{(s)}) = \begin{cases} E_1^{(s)}, & k = 1, \\ E_k^{(s)} + \sum_{i=1}^{k-1} \alpha_i^{(s)} O_{ki}^{(s)}, & k = 2, \cdots, D, \end{cases}$$

to obtain the D eigenvalues of H_s and the corresponding eigenvectors where

$$\begin{cases} E_k^{(s)} = \langle\psi(\lambda_k^{(s)})|H_s|\psi(\lambda_k^{(s)})\rangle \\ O_{ki}^{(s)} = |\langle\psi(\lambda_k^{(s)})|\psi(\lambda_i^{(s)})\rangle|^2 \end{cases}$$

with the weight coefficient $\alpha_i^{(s)}$ for $i = 1, \cdots, k-1$.
step 3: Prepare the ansatz states $|\psi(\lambda_k^{(t)})\rangle$ with a set of parameters $\{\theta^{(t)}\}$. Minimize the cost function

$$F_t(\lambda_k^{(t)}) = \begin{cases} E_1^{(t)}, & k = 1, \\ E_k^{(t)} + \sum_{i=1}^{k-1} \alpha_i^{(t)} O_{ki}^{(t)}, & k = 2, \cdots, r, \end{cases}$$

to obtain the r smallest eigenvalues of $\tilde{H}_t$ and the corresponding eigenvectors where

$$\begin{cases} E_k^{(t)} = \langle\psi(\lambda_k^{(t)})|\tilde{H}_t|\psi(\lambda_k^{(t)})\rangle \\ O_{ki}^{(t)} = |\langle\psi(\lambda_k^{(t)})|\psi(\lambda_i^{(t)})\rangle|^2 \end{cases}$$

with the weight coefficient $\alpha_i^{(t)}$ for $i = 1, \cdots, k-1$.
step 4: Compute $C_s^{1/2} = U_s \Sigma_s^{1/2} U_s^T$ and $C_t^{1/2} = U_{t[1:r]} \Sigma_{t[1:r]}^{1/2} U_{t[1:r]}^T$ with the eigenvalues and eigenvectors obtained in step 2 and step 3.

In step 1, we compute the source domain covariance matrix $H_s = \rho_{C_s}$ and the target domain covariance matrix $H_t = \rho_{C_t}$. Then, the Hamiltonian $\tilde{H}_t = \eta I - H_t$ is determined with a specified constant η.

In step 2, the ansatz states $|\psi(\lambda_k^{(s)})\rangle$ are prepared by a quantum circuit with a set of parameters $\{\theta^{(s)}\}$. Subsequently, the cost function $F_s(\lambda_k^{(s)})$ is minimized to obtain the optimal ansatz states where the expectation value term $E_k^{(s)} = \langle\psi(\lambda_k^{(s)})|H_s|\psi(\lambda_k^{(s)})\rangle$ and the overlap term $O_{ki}^{(s)} = |\langle\psi(\lambda_k^{(s)})|\psi(\lambda_i^{(s)})\rangle|^2$ with the weight coefficient $\alpha_i^{(s)}$ for $i = 1, \cdots, k-1$. In the first iteration, we minimize the $F_s(\lambda_1^{(s)})$ to obtain the ground state $|\psi(\lambda_1^{(s)})\rangle$ of H_s with the corresponding eigenvalue $\lambda_1 = E_1$. In the second iteration, substitute $|\psi(\lambda_1^{(s)})\rangle$ to $F_s(\lambda_2^{(s)})$ and minimize it to obtain $|\psi(\lambda_2^{(s)})\rangle$. Then, the iteration continues until $|\psi(\lambda_D^{(s)})\rangle$ is computed by substituting $|\psi(\lambda_{D-1}^{(s)})\rangle$ to the cost function $F_s(\lambda_D^{(s)})$. Therefore, H_s's eigenstates $|\psi(\lambda_k^{(s)})\rangle$ for $k = 1, \cdots, D$ corresponding to the D eigenvalues can be obtained in $O(1/\epsilon^2)$ [46].

In step 3, the r largest eigenvalues of H_t can be obtained similarly by minimizing the cost function $F_t(\lambda_k^{(t)})$ as exactly the same procedure as in step 2 in time $O(1/\epsilon^2)$ [46] where the expectation value term $E_k^{(t)} = \langle\psi(\lambda_k^{(t)})|\tilde{H}_t|\psi(\lambda_k^{(t)})\rangle$ and the overlap term $O_{ki}^{(t)} = |\langle\psi(\lambda_k^{(t)})|\psi(\lambda_i^{(t)})\rangle|^2$ with the weight coefficient $\alpha_i^{(t)}$ for $i = 1, \cdots, k-1$.

In step 4, the matrices $C_s^{1/2} = U_s \Sigma_s^{1/2} U_s^T$ and $C_t^{1/2} = U_{t[1:r]} \Sigma_{t[1:r]}^{1/2} U_{t[1:r]}^T$ can be computed by the results of step 2 and step 3. Specifically, the D eigenvalues of H_s are the diagonal elements of Σ_s and the columns of U_s are the corresponding D eigenvectors. The nonzero diagonal elements of $\Sigma_{t[1:r]}$ are the r largest eigenvalues of H_t and the columns of $U_{t[1:r]}$ are made up of the corresponding eigenvectors.

(2) The procedure of data decorrelation and alignment can be achieved as Eqs. (6) and (7), which are actually variational processes of matrix multiplication. In the spirit of Ref. [47], we design a matrix-multiplication-based VQCORAL as in Algorithm 3.

Algorithm 3. Matrix-multiplication-based VQCORAL

Input: Source domain data X_s with labels L_s; target domain data X_t; $C_s^{1/2}$ and $C_t^{1/2}$.
Output: Target domain labels L_t.
step 1: Prepare the ansatz states $|x_i^{(\tilde{s})}(\theta^{(d)})\rangle$ by parametrized quantum circuits and a set of parameters $\{\theta^{(d)}\}$ to represent the data points of the decorrelated source domain data $X_{\tilde{s}}$.
step 2: Minimize the cost function

$$L_{m1} = 1 - \frac{1}{n_s} \sum_{i=1}^{n_s} \left| \frac{\langle x_i^{(s)}|C_s^{\frac{1}{2}}|x_i^{(\tilde{s})}(\theta^{(d)})\rangle}{\sqrt{\langle x_i^{(\tilde{s})}(\theta^{(d)})|C_s^{\frac{1}{2}\dagger} C_s^{\frac{1}{2}}|x_i^{(\tilde{s})}(\theta^{(d)})\rangle}} \right|^2$$

to obtain the optimal decorrelated source domain data states $|x_{i*}^{(\tilde{s})}(\theta^{(d)})\rangle$.
step 3: Prepare the ansatz states $|x_i^{(\hat{s})}(\theta^{(a)})\rangle$ by parametrized quantum circuits and a set of parameters $\{\theta^{(a)}\}$ to represent the data points of the aligned source domain data $X_{\hat{s}}$.
step 4: Minimize the cost function

$$L_{m2} = 1 - \frac{1}{n_s} \sum_{i=1}^{n_s} \left| \frac{\langle x_i^{(\hat{s})}(\theta^{(a)})|C_t^{\frac{1}{2}}|x_{i*}^{(\tilde{s})}(\theta^{(d)})\rangle}{\sqrt{\langle x_i^{(\hat{s})}(\theta^{(a)})|C_t^{\frac{1}{2}} C_t^{\frac{1}{2}\dagger}|x_i^{(\hat{s})}(\theta^{(a)})\rangle}} \right|^2$$

to obtain the optimal aligned source domain data states $|x_{i*}^{(\hat{s})}(\theta^{(a)})\rangle$.
step 5: Invoke a classifier to predict the target labels $L_t = Classifier(X_{\hat{s}}, L_s, X_t)$.

In step 1, the quantum ansatz states $|x_i^{(\tilde{s})}(\theta^{(d)})\rangle$ representing the decorrelated source domain data points $x_i^{(\tilde{s})}$ are designed by parametrized quantum circuits with a set of parameters $\{\theta^{(d)}\}$.

In step 2, the states

$$|\psi_1\rangle = \frac{C_s^{\frac{1}{2}} \left| x_i^{(\tilde{s})}(\theta^{(d)}) \right\rangle}{\sqrt{\left\langle x_i^{(\tilde{s})}(\theta^{(d)}) \right| C_s^{\frac{1}{2}\dagger} C_s^{\frac{1}{2}} \left| x_i^{(\tilde{s})}(\theta^{(d)}) \right\rangle}} \tag{24}$$

are defined to be proportional to $|x_i^{(s)}\rangle$ with a set of parameters $\{\theta^{(d)}\}$. Thus, the optimal quantum ansatz states $|x_{i*}^{(\hat{s})}(\theta^{(d)})\rangle$ representing the decorrelated source domain data can be obtained by minimizing the cost function

$$L_{m1} = 1 - \frac{1}{n_s}\sum_{i=1}^{n_s}\left|\langle x_i^{(s)}|\psi_1\rangle\right|^2 \tag{25}$$

in time $O(\kappa_s/\epsilon)$ [47] where κ_s is the conditional number of $C_s^{\frac{1}{2}}$.

In step 3, the ansatz states $|x_i^{(\hat{s})}(\theta^{(a)})\rangle$ are prepared by parametrized quantum circuits with a set of parameters $\{\theta^{(a)}\}$.

In step 4, define the states

$$|\psi_2\rangle = \frac{C_t^{\frac{1}{2}\dagger}|x_i^{(\hat{s})}(\theta^{(a)})\rangle}{\sqrt{\langle x_i^{(\hat{s})}(\theta^{(a)})|C_t^{\frac{1}{2}}C_t^{\frac{1}{2}\dagger}|x_i^{(\hat{s})}(\theta^{(a)})\rangle}} \tag{26}$$

with a set of parameters $\{\theta^{(a)}\}$. Then, we align the decorrelated source domain data $X_{\hat{s}}$ by minimizing

$$L_{m2} = 1 - \frac{1}{n_s}\sum_{i=1}^{n_s}\left|\langle\psi_2|x_{i*}^{(\hat{s})}(\theta^{(d)})\rangle\right|^2 \tag{27}$$

in time $O(\kappa_t/\epsilon)$ [47] where κ_t is the conditional number of $C_t^{-\frac{1}{2}}$. The data alignment procedure actually aims to generate the state $|\psi_2\rangle$ to be proportional to $|x_i^{(\hat{s})}(\theta^{(a)})\rangle$.

In step 5, a classifier such as the local classifier, the nearest-neighbor algorithm, or the global classifier (the support vector machine) will be applied to $X_{\hat{s}}$ with L_s and X_t to predict the target labels L_t.

IV. NUMERICAL EXPERIMENTS

In this section, three numerical experiments are presented to demonstrate the feasibility and efficiency of the VQCORAL. The NA model, the classical CORAL, and the VQCORAL are applied to the synthetic data sets, the synthetic-Iris data sets, and the handwritten digit data sets, respectively, to evaluate their performance. According to the simulation results, the VQCORAL can achieve comparable or even better performance than the classical CORAL. The VQCORAL is simulated on a classical computer using the PYTHON programming language and the Scikit-learn machine learning library [56]. The code and the selected parameters can be found in Ref. [57].

A. Basic settings

The NA model is set as the baseline model. In addition, the classical CORAL is also selected as a performance comparison to the VQCORAL. As for the VQCORAL, we design parametrized quantum circuits with hierarchical structures. Specifically, we apply the Hadamard operation on each register, respectively, as the first layer. Then, we alternately apply the rotation layer constructed by the R_y gate on each qubit and the entanglement layer constructed by the controlled-NOT gate on each pair of qubits to introduce the parameters and entanglement as shown in Fig. 2. The classical optimization algorithm, the AdaGrad [58], is selected to optimize the cost function.

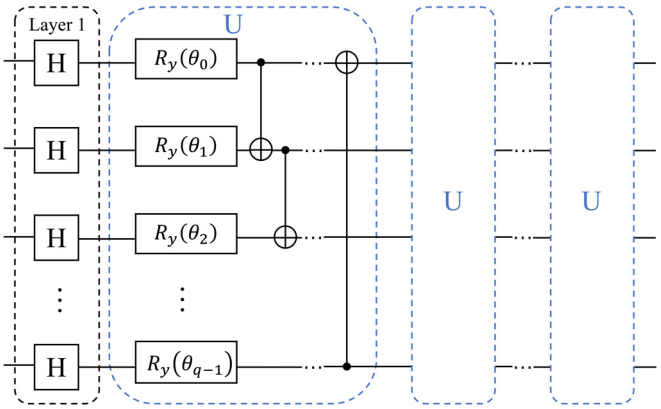

FIG. 2. The variational quantum circuit for preparing $\mathbf{U}_\theta$ where $q = \log D$.

B. Synthetic data

In the first numerical experiment, we select two synthetic data sets $D_1 \sim \mathcal{N}(\mu_1^{(1)} = \mu_2^{(1)} = 0, \sigma_1^{(1)} = \sigma_2^{(1)} = 1)$ and $D_2 \sim \mathcal{N}(\mu_1^{(2)} = \mu_2^{(2)} = 0, \sigma_1^{(2)} = \sigma_2^{(2)} = 2)$ as the source and target domain data sets alternately. Both X_s and X_t contain 100 four-dimensional data points distributed in two different classes.

The design of the VQCORAL in this experiment is a two-qubit eight-layer quantum circuit. The simulation results of the NA, the classical CORAL, and the VQCORAL applied to the $D_1 \to D_2$ task and the $D_2 \to D_1$ task are presented in Table I.

As shown in Table I, for both the $D_1 \to D_2$ and the $D_2 \to D_1$ tasks, it is obvious that the NA (baseline model) cannot achieve a relatively high accuracy. However, the performance of the classical CORAL is comparable to the NA, meaning that the classical CORAL may not play the role of domain adaptation as we expected in some cases. Compared with the classical CORAL and the NA, the VQCORAL model achieves significantly better performance.

C. Synthetic and Iris data

In the second experiment, the synthetic data set $D_3 \sim \mathcal{N}(\mu_1^{(3)} = \mu_2^{(3)} = \mu_3^{(3)} = 0, \sigma_1^{(3)} = \sigma_2^{(3)} = \sigma_3^{(3)} = 1)$ and the Iris data set [48,49] are selected as the source and target domain data sets alternately. Both the D_3 and the Iris data set contain 150 samples evenly distributed in three different classes.

The model adopted by the VQCORAL in this experiment is a two-qubit eight-layer parametrized quantum circuit. The NA, the classical CORAL, and the VQCORAL are applied to

TABLE I. Accuracies of the NA, the classical CORAL, and the VQCORAL applied on the synthetic data sets D_1 and D_2.

	$D_1 \to D_2$	$D_2 \to D_1$
NA	50%	50%
Classical CORAL	50%	50%
VQCORAL	90%	97%

TABLE II. Accuracies of the NA, the classical CORAL, the VQCORAL applied on the synthetic data set D_3, and the Iris data set.

	$D_3 \rightarrow$ Iris	Iris $\rightarrow D_3$
NA	33.3%	4%
Classical CORAL	33.3%	14%
VQCORAL	66.6%	72.7%

the $D_3 \rightarrow$ Iris task and the Iris $\rightarrow D_3$ task, resulting in the results in Table II.

As in Table II, the accuracies of both the NA and the classical CORAL are 33.3% for the $D_3 \rightarrow$ Iris task, which are worse than the 66.6% accuracy of the VQCORAL. For the Iris $\rightarrow D_3$ task, the accuracy of the NA is only 4%. The classical CORAL shows improvement with 14% accuracy. The VQCORAL achieves significant performance improvement with the accuracy of 72.7%, indicating that the VQCORAL can exhibit more powerful expressivity in some specific tasks.

D. Handwritten digit data

The MNIST [50] and USPS [51] are the two representative handwritten digit data sets widely used for evaluating the performance of machine learning and pattern recognition. For the transfer learning task, 2000 28×28 images of the MNIST and 1800 16×16 images of the USPS are selected as the source and target domain data sets. In the data preprocessing, all the images are linearly rescaled to 16×16, meaning that the gray values of each image are represented by a 256-dimensional vector. The MNIST and USPS share the same feature space but are generated from different distributions.

Concretely, the quantum circuit adopted by the VQCORAL has an eight-qubit 16-layer structure. The simulation results of the NA, the classical CORAL, and the VQCORAL applied to the MNIST $\rightarrow$ USPS task and the USPS $\rightarrow$ MNIST task are presented in Table III.

According to Table III, both the classical CORAL and the VQCORAL show better performance than the NA, meaning that the CORAL is helpful in accomplishing transfer learning tasks. In addition, the VQCORAL can achieve a comparable accuracy, namely, 65.6%, as the classical CORAL in the MNIST $\rightarrow$ USPS task. Although in the USPS $\rightarrow$ MNIST task the accuracy of the VQCORAL is 44.5%, which is not as good as the classical CORAL, the VQCORAL still exhibits better performance than the NA. We believe that the VQCORAL can achieve at least comparable accuracy to the classical CORAL by further optimizing the design of the quantum circuit.

TABLE III. Accuracies of the NA, the classical CORAL, and the VQCORAL applied on the MNIST and USPS handwritten digit data sets.

	MNIST $\rightarrow$ USPS	USPS $\rightarrow$ MNIST
NA	64.4%	35.9%
Classical CORAL	65.6%	46.9%
VQCORAL	65.6%	44.5%

V. DISCUSSIONS

In this paper, we propose two quantum versions of the CORAL, one of the most representative domain adaptation algorithms. On the one hand, the QCORAL implemented by the QBLAS can be performed on a universal quantum computer with exponential speedup in the dimension and number of the given data. On the other hand, the VQCORAL can be performed on near-term quantum devices with low circuit depth. Specifically, the VQCORAL can be implemented in two different perspectives. From an intuitive perspective, the VQCORAL can be realized directly by an end-to-end hierarchical structure. In addition, the source domain data can be decorrelated and aligned to the target domain data by successively applying the variational quantum covariance matrix square-root solver and the variational matrix multiplication operations. To evaluate the feasibility and efficiency of our work, we design three different types of numerical experiments, namely, the synthetic data, the synthetic-Iris data, and the handwritten digit data. According to the simulation results, the VQCORAL presented in this paper can achieve at least comparable or even better performance than the classical CORAL.

However, some open questions need further study. First of all, the QBLAS-based QCORAL requires a high-depth quantum circuit and fully coherent evolution, which are actually prohibited in experiment at present. In addition, although the VQCORAL algorithm can be realized with limited quantum resources, the performance of the variational algorithm actually depends largely on the specific design of the parametrized circuits. Hence, it is well worth exploring how to design quantum circuits specifically to achieve optimal performance. Although some further exploration is required, it is demonstrated that quantum techniques can make a contribution to the field of domain adaptation.

ACKNOWLEDGMENTS

The author would like to thank Jonathan Allcock for fruitful discussions. The author also would like to thank anonymous referees for helpful feedback on this paper. This work is supported by National Key Research and Development Program of China Grant No. 2018YFA0306703.

[1] P. W. Shor, in *Proceedings of the 35th Annual Symposium on Foundations of Computer Science* (IEEE, New York, 1994), pp. 124–134.

[2] L. K. Grover, in *Proceedings of the Twenty-Eighth Annual ACM Symposium on Theory of Computing* (ACM, New York, 1996), pp. 212–219.

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

[4] S. Aaronson and A. Arkhipov, in *Proceedings of the Forty-Third Annual ACM Symposium on Theory of Computing* (ACM, New York, 2011), pp. 333–342.

[5] E. Farhi and H. Neven, arXiv:1802.06002 (2018).

[6] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. Brandao, D. A. Buell *et al.*, Nature (London) **574**, 505 (2019).

[7] S. Lloyd, M. Mohseni, and P. Rebentrost, arXiv:1307.0411 (2013).

[8] V. Havlíček, A. D. Córcoles, K. Temme, A. W. Harrow, A. Kandala, J. M. Chow, and J. M. Gambetta, Nature (London) **567**, 209 (2019).

[9] M. Schuld and N. Killoran, Phys. Rev. Lett. **122**, 040504 (2019).

[10] M. Schuld and F. Petruccione, *Supervised Learning with Quantum Computers*, Quantum Science and Technology (QST) Vol. 17 (Springer, New York, 2018).

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

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

[13] M. Schuld, A. Bocharov, K. M. Svore, and N. Wiebe, Phys. Rev. A **101**, 032308 (2020).

[14] M. Schuld, M. Fingerhuth, and F. Petruccione, Europhys. Lett. **119**, 60002 (2017).

[15] M. Schuld, I. Sinayskiy, and F. Petruccione, in *Pacific Rim International Conference on Artificial Intelligence* (Springer, New York, 2014), pp. 208–220.

[16] M. Schuld and F. Petruccione, Sci. Rep. **8**, 1 (2018).

[17] N. Wiebe, D. Braun, and S. Lloyd, Phys. Rev. Lett. **109**, 050505 (2012).

[18] M. Schuld, I. Sinayskiy, and F. Petruccione, Phys. Rev. A **94**, 022342 (2016).

[19] E. Aïmeur, G. Brassard, and S. Gambs, Mach. Learn. **90**, 261 (2013).

[20] N. Wiebe, A. Kapoor, and K. M. Svore, Quantum Inf. Comput. **15**, 3 (2018).

[21] I. Cong and L. Duan, New J. Phys. **18**, 073011 (2016).

[22] B. Duan, J. Yuan, J. Xu, and D. Li, Phys. Rev. A **99**, 032311 (2019).

[23] X. He, L. Sun, C. Lyu, and X. Wang, Quantum Inf. Process. **19**, 1 (2020).

[24] J. Romero, J. P. Olson, and A. Aspuru-Guzik, Quantum Sci. Techn. **2**, 045001 (2017).

[25] N. Wiebe, A. Kapoor, and K. M. Svore, Quantum Inf. Comput. **16**, 541 (2016).

[26] M. H. Amin, E. Andriyash, J. Rolfe, B. Kulchytskyy, and R. Melko, Phys. Rev. X **8**, 021050 (2018).

[27] S. Lloyd and C. Weedbrook, Phys. Rev. Lett. **121**, 040502 (2018).

[28] P.-L. Dallaire-Demers and N. Killoran, Phys. Rev. A **98**, 012324 (2018).

[29] K. H. Wan, O. Dahlsten, H. Kristjánsson, R. Gardner, and M. S. Kim, npj Quant. Info. **3**, 36 (2017).

[30] A. Mari, T. R. Bromley, J. Izaac, M. Schuld, and N. Killoran, arXiv:1912.08278 (2019).

[31] X. He, C. Lyu, M.-H. Hsieh, and X. Wang, arXiv:1912.09113 (2019).

[32] X. He, arXiv:2001.02472 (2020).

[33] S. J. Pan and Q. Yang, IEEE Transactions on Knowledge and Data Engineering **22**, 1345 (2009).

[34] G. Csurka, *Domain Adaptation in Computer Vision Applications* (Springer, New York, 2017), Vol. 2.

[35] X. Glorot, A. Bordes, and Y. Bengio, in *ICML'11: Proceedings of the 28th International Conference on International Conference on Machine Learning* (Omni Press, Madison, WI, 2011), pp. 513–520.

[36] T. Carr, M. Chli, and G. Vogiatzis, in *Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems* (International Foundation for Autonomous Agents and Multiagent Systems, 2019), pp. 1859–1861.

[37] S. J. Pan, I. W. Tsang, J. T. Kwok, and Q. Yang, IEEE Transactions on Neural Networks **22**, 199 (2011).

[38] M. Long, J. Wang, G. Ding, J. Sun, and P. S. Yu, in *Proceedings of the IEEE International Conference on Computer Vision* (IEEE, New York, 2013), pp. 2200–2207.

[39] J. Wang, Y. Chen, S. Hao, W. Feng, and Z. Shen, in *2017 IEEE International Conference on Data Mining (ICDM)* (IEEE, New York, 2017), pp. 1129–1134.

[40] B. Fernando, A. Habrard, M. Sebban, and T. Tuytelaars, in *Proceedings of the IEEE International Conference on Computer Vision* (IEEE, New York, 2013), pp. 2960–2967.

[41] R. Gopalan, R. Li, and R. Chellappa, in *2011 International Conference on Computer Vision* (IEEE, New York, 2011), pp. 999–1006.

[42] B. Gong, Y. Shi, F. Sha, and K. Grauman, in *2012 IEEE Conference on Computer Vision and Pattern Recognition* (IEEE, New York, 2012), pp. 2066–2073.

[43] B. Sun, J. Feng, and K. Saenko, in *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence* (AAAI Press, Phoenix, AR, 2016), pp. 2058–2065.

[44] B. Sun, J. Feng, and K. Saenko, arXiv:1612.01939 (2016).

[45] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'Brien, Nat. Commun. **5**, 4213 (2014).

[46] O. Higgott, D. Wang, and S. Brierley, Quantum **3**, 156 (2019).

[47] C. Bravo-Prieto, R. LaRose, M. Cerezo, Y. Subasi, L. Cincio, and P. J. Coles, arXiv:1909.05820 (2019).

[48] R. Fisher, Ann. Eugenics **7**, 179 (1936).

[49] E. Anderson, Ann. Mo. Bot. Gard. **23**, 457 (1936).

[50] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, Proc. IEEE **86**, 2278 (1998).

[51] Y. LeCun, B. E. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. E. Hubbard, and L. D. Jackel, in *Advances in Neural Information Processing Systems*, edited by D. S. Touretzky (Morgan-Kaufmann, Denver, CO, 1990), pp. 396–404.

[52] V. Giovannetti, S. Lloyd, and L. Maccone, Phys. Rev. Lett. **100**, 160501 (2008).

[53] A. Prakash, Quantum algorithms for linear algebra and machine learning, Ph.D. thesis, University of California Berkeley, 2014.

[54] P. Rebentrost, A. Steffens, I. Marvian, and S. Lloyd, Phys. Rev. A **97**, 012327 (2018).

[55] B. Duan, J. Yuan, Y. Liu, and D. Li, Phys. Rev. A **96**, 032301 (2017).

[56] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, J. Mach. Learn. Res. **12**, 2825 (2011).

[57] https://github.com/xihechn/QCORAL.

[58] J. Duchi, E. Hazan, and Y. Singer, J. Mach. Learn. Res. **12**, 2121 (2011).