Quantum anomaly detection with density estimation and multivariate Gaussian distribution

Jin-Min Liang,[1] Shu-Qian Shen,[1,*] Ming Li,[1,2] and Lei Li[1]

[1]*College of Science, China University of Petroleum, 266580 Qingdao, People's Republic of China*
[2]*Max-Planck-Institute for Mathematics in Sciences, 04103 Leipzig, Germany*

We study quantum anomaly detection with density estimation and multivariate Gaussian distribution. Both algorithms are constructed using the standard gate-based model of quantum computing. Compared with the corresponding classical algorithms, the resource complexities of our quantum algorithm are logarithmic in the dimensionality of quantum states and the number of training quantum states. We also present a quantum procedure for efficiently estimating the determinant of any Hermitian operators $\mathcal{A} \in \mathcal{R}^{N \times N}$ with time complexity $O(poly \log N)$, which forms an important subroutine in our quantum anomaly detection with multivariate Gaussian distribution. Finally, our results also include the modified quantum kernel principal component analysis (PCA) and the quantum one-class support vector machine (SVM) for detecting classical data.

I. INTRODUCTION

Anomaly detection (AD) is a pivotal topic in statistics that has been researched within diverse areas and application domains, such as fraud detection for credit cards, insurance or health care, intrusion detection for cyber-security, fault detection in safety critical systems, and military surveillance for enemy activities [1,2]. The importance of AD is due to the fact that anomalous data indicate significant information. For example, an anomalous MRI image may mean the presence of malignant tumors [3]. Traditional anomaly detection techniques focus on detecting anomalous data that differ from normal (or clean) data patterns. Over time, a variety of anomaly detection techniques have been developed in several research communities. One of the most common machine learning algorithms for AD is based on the density estimation method by estimating the probability density function [4–7]. In 2009, the authors presented a first attempt to evaluate two previously proposed methods for statistical anomaly detection in sea traffic, namely the Gaussian mixture model and the adaptive kernel density estimator [8]. Erfani *et al.* proposed a hybrid model for unsupervised anomaly detection that combines a one-class support vector machine (SVM) and a deep belief network (DBN) [9]. Carrera *et al.* presented a novel approach for detecting anomalous structures in images by learning a convolutional sparse model that describes the local structures of normal images [10]. Especially for point anomalies, Hoffmann studied kernel PCA for novelty detection in 2007 [11]. Due to the fact that abnormalities are very rare, we could not use the traditional classification method to learn an anomaly model described by the fault conditions. AD offered a solution to this situation by modeling a normal data set and using a distance measure and a threshold for determining abnormality [12].

In parallel, we are witnessing the influence of quantum information processing on its classical counterpart. Following the discovery of quantum algorithms for factoring [13], database searching [14], and quantum matrix inverse [15], decades of work have shown that quantum algorithms have the capability of outperforming existing classical methods. These algorithms have recently been employed in machine learning and form an interdisciplinary field called quantum machine learning (QML). In particular, it has proved useful in regression and classification problems that appear in machine learning [16–21].

A natural question arises of how to design a quantum machine learning algorithm for detecting anomalies in a quantum computer. Although some previous results have been reported [17,22,23], AD still requires more investigation from a quantum mechanics perspective. Quite recently, Sentís *et al.* introduced the quantum change point and devised online strategies for identifying a change point in a stream of quantum particles allegedly prepared in identical states [22,23], which is a quantum version of identifying changes in a sequence of random variables. Liu *et al.* designed two quantum algorithms for quantum anomaly detection, quantum kernel PCA and quantum one-class SVM, proved that these two quantum algorithms can be performed using resources logarithmic in the dimensionality of quantum states [17].

In the present study, we design two anomaly detection algorithms based on density estimation and multivariate Gaussian distribution. Given a training set of M quantum states which have d dimensional features, we can identify which one varies significantly from the average. We show how a training set can be modeled by using a Gaussian distribution and how the model can be used for anomaly detection. Our algorithm can be achieved using resources only logarithmic in the dimensionality of quantum states and the number of training quantum states. The organization of the paper is as follows. In Sec. II, we review the classical anomaly detection algorithm with density estimation and multivariate Gaussian distribution. In Secs. III and IV, we introduce
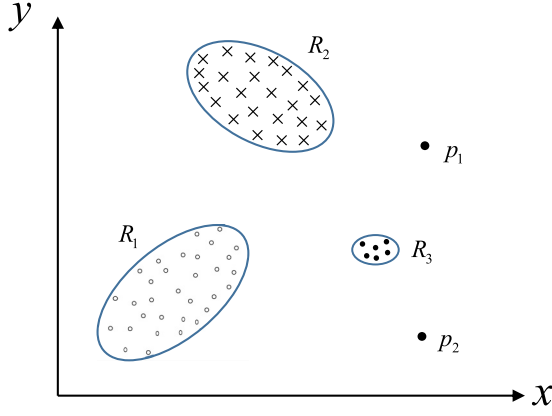
*sqshen@upc.edu.cn

FIG. 1. A simple example of anomalies in a two-dimensional data set.

the key quantum anomaly detection algorithm with density estimation and multivariate Gaussian distribution in detail. We discuss the generalization of quantum kernel PCA and quantum one-class SVM to tackle a classical discrete data anomaly in Sec. V. A summary and discussions are included in Sec. VI.

II. ANOMALY DETECTION ALGORITHM

In this section, we provide the necessary background information to understand this paper. An anomaly detection algorithm aims to construct a data model which normal data obeys whereas anomalous data do not. Anomalies are patterns in data that do not meet a well defined notion of normal behavior [24]. For instance, Fig. 1 illustrates anomalies in a simple two-dimensional data set. The data set has two normal regions, R_1 and R_2, since most observations lie in these two regions. Points that are sufficiently far away from the regions, e.g., points p_1 and p_2, and points in region R_3, are anomalies. Typically, one can estimate a number of parameters characterized by the data model of a random variable. Parametric approaches make an assumption that data distributions are Gaussian in nature and they can be modeled statistically based on data means and covariance [12]. We will use an anomaly detection algorithm with density estimation and multivariate Gaussian distribution, described below, as a foundation to present a quantum anomaly detection algorithm.

A. Anomaly detection algorithm with density estimation

Suppose we are given a training set $\{x^i\}_{i=1}^M$ with d features and M samples which are expected to follow a Gaussian distribution. The algorithm aims to create a statistical model $p(x)$ from the training set by assuming each feature follows a Gaussian distribution and then identifying whether the new datum x is abnormal or not. Then the overall procedure is depicted as follows:

(1) Fit density parameters $\mu_1, \ldots, \mu_d, \sigma_1^2, \ldots, \sigma_d^2$, where

$$\mu_j = \frac{1}{M} \sum_{i=1}^M x_j^i, \quad \sigma_j^2 = \frac{1}{M} \sum_{i=1}^M \left(x_j^i - \mu_j\right)^2.$$

(2) Given a new example x, compute $p(x)$,

$$p(x) = \prod_{j=1}^d p\left(x_j; \mu_j, \sigma_j^2\right) = \prod_{j=1}^d \frac{1}{\sqrt{2\pi}\sigma_j} e^{-\frac{(x_j-\mu_j)^2}{2\sigma_j^2}}. \quad (1)$$

(3) If $p(x) < \epsilon$, one can flag this datum as an anomaly; otherwise, it is a normal datum.

The parameter ϵ is some threshold probability value depending on how sure we want to be. The most common method for selecting a suitable threshold is K-fold cross validation, the same as other prediction tasks [25,26]. For simplicity, we take the following computation based on Eq. (1):

$$\ln p(x) = \ln \prod_{j=1}^d \frac{1}{\sqrt{2\pi}\sigma_j} e^{-\frac{(x_j-\mu_j)^2}{2\sigma_j^2}} = \sum_{j=1}^d \ln \frac{1}{\sqrt{2\pi}\sigma_j} e^{-\frac{(x_j-\mu_j)^2}{2\sigma_j^2}}$$

$$= -\frac{d}{2} \ln 2\pi - \sum_{j=1}^d \ln \sigma_j - \sum_{j=1}^d \frac{(x_j - \mu_j)^2}{2\sigma_j^2}. \quad (2)$$

B. Anomaly detection algorithm with multivariate Gaussian distribution

The multivariate Gaussian distribution is different from the Gaussian distribution. This algorithm attempts to develop a model $p(x)$ all in one process, instead of modeling each feature separately as in Eq. (1). For a training set $\{x^i\}_{i=1}^M$ with d features and M samples, we fit model $p(x)$ by setting

$$\mu = \frac{1}{M} \sum_{i=1}^M x^i, \quad C = \frac{1}{M} \sum_{i=1}^M (x^i - \mu)(x^i - \mu)^T. \quad (3)$$

Given a new example x, compute

$$p(x) = \frac{1}{(2\pi)^{\frac{d}{2}} |C|^{\frac{1}{2}}} e^{-\frac{1}{2}(x-\mu)^T C^{-1}(x-\mu)}, \quad (4)$$

where $|C|$ represents the determinant of matrix C. If $p(x) < \epsilon$ we can flag this datum as an anomaly, otherwise, it is a normal datum.

III. QUANTUM ANOMALY DETECTION ALGORITHM WITH DENSITY ESTIMATION

In this section, we will develop a quantum anomaly algorithm with density estimation. For classical data, our schemes require the method of encoding classical information into the amplitude of a quantum system and then applying our quantum anomaly detection algorithms directly. In other words, we encode a 2^n-dimensional vector $\vec{u} = \{u_0, u_1, \ldots, u_{2^n-1}\}$ into the 2^n amplitudes $u_0, u_1, \ldots, u_{2^n-1}$ of an n-qubit quantum system, $|u\rangle = \sum_{i=0}^{2^n-1} u_i |i\rangle$, where $\{|i\rangle\}$ is the computational basis [18]. Thus, in what follows, we only consider quantum pure-state anomaly detection. In that sense, our algorithms can detect an anomaly in both classical data and quantum states generated from quantum devices.

Suppose we are given access to the unitaries $\{U_i\}_{i=0}^M$ to obtain quantum states $\{|x^i\rangle\}_{i=0}^M$, where $U_i|0\cdots0\rangle = |x^i\rangle$. We are given a set of M *normal* quantum training states $\{|x^i\rangle\}_{i=1}^M$ and the test quantum state $\{|x^0\rangle\}$. Our task is to detect how anomalous the state $|x^0\rangle$ is compared to the normal states. We assume the runtime of our algorithm is dominated by the

number of quantum gates. First, we define the analogous mean state of the training quantum state by

$$|\mu\rangle = \frac{1}{N_\mu} \sum_{i=1}^{M} |x^i\rangle = \frac{1}{N_\mu} \sum_{i=1}^{M} \sum_{j=1}^{d} x_j^i |j\rangle, \qquad (5)$$

where x_j^i denotes the jth element of state $|x^i\rangle$. The normalization coefficient $N_\mu^2 = \sum_{k,l=1}^{M} \langle x^k | x^l \rangle$ can be found using $O(\log M)$ resources. Here, we show a method for preparing the mean state. We can apply the unitary operation $I \otimes H^{\log M}$ on the given initial state $|\Psi_1\rangle = \frac{1}{\sqrt{M}} \sum_{i=1}^{M} |x^i\rangle|i\rangle$ to obtain

$$|\psi_\mu\rangle = \frac{1}{M} \sum_{i=1}^{M} |x^i\rangle \sum_{j=1}^{M} (-1)^{i \cdot j} |j\rangle, \qquad (6)$$

where $i \cdot j$ is the bitwise inner product of i and j, modulo 2. Specifically, we perform a Hadamard gate H on each of the qubits of the second register. Then by making a projective measurement $|0 \cdots 0\rangle\langle 0 \cdots 0|$ on the second register we can recover the mean state $|\mu\rangle|0 \cdots 0\rangle$. The success probability P_μ of this measurement is

$$P_\mu = \mathrm{tr}(|0 \cdots 0\rangle\langle 0 \cdots 0| H^{\otimes \log M} |\Psi_1\rangle\langle\Psi_1| H^{\otimes \log M})$$
$$= \frac{1}{M^2} \sum_{k,l=1}^{M} \langle x^k | x^l \rangle. \qquad (7)$$

Obviously, the use of an $O(\log M)$ Hadamard gates means that our gate resource count is $O(\log M)$.

Algorithm 1: Efficiently computing $\sum_{j=1}^{d} \frac{(x_j^0 - \mu_j)^2}{2\sigma_j^2}$

Step 1. Prepare the quantum state $\sum_{j=1}^{d} |\chi_j\rangle|\chi_j^0\rangle|j\rangle$.

Step 2. Add an ancilla qubit $|0\rangle$ and perform a controlled unitary operator R_1 to obtain

$$\sum_{j=1}^{d} |\chi_j\rangle|\chi_j^0\rangle|j\rangle \left(\frac{\chi_j^0}{\chi_j}|0\rangle + \sqrt{1 - \left(\frac{\chi_j^0}{\chi_j}\right)^2}|1\rangle \right).$$

Step 3. Uncompute the second and third registers. The system state is

$$\sum_{j=1}^{d} |j\rangle \left(\frac{\chi_j^0}{\chi_j}|0\rangle + \sqrt{1 - \left(\frac{\chi_j^0}{\chi_j}\right)^2}|1\rangle \right).$$

Step 4. Measure the observable $M_1 = I \otimes |0\rangle\langle 0|$ and the expectation $\langle M_1 \rangle = \sum_{j=1}^{d} \left(\frac{\chi_j^0}{\chi_j}\right)^2$.

Next, we show a method for preparing the state $|\chi\rangle = \frac{1}{N_\chi} \sum_{j=1}^{d} |\chi_j\rangle|j\rangle = \frac{1}{N_\chi} \sum_{j=1}^{d} \sum_{i=1}^{M} (x_j^i - \mu_j)|i\rangle|j\rangle$ corresponding to the $x_j - \mu_j$ in classical anomaly detection with density estimation, where $N_\chi^2 = \sum_{k,l=1}^{d} \langle \chi_k | \chi_l \rangle$. We rewrite the mean state $|\mu\rangle$ as $|\mu\rangle = \frac{1}{N_\mu} \sum_{j=1}^{d} \mu_j |j\rangle$, where $\mu_j = \sum_{i=1}^{M} x_j^i$. Suppose we are given the control unitaries to create

the following superpositions of training states:

$$|\Psi_2\rangle = \frac{|0\rangle \sum_{j=1}^{d} |x^i\rangle|j\rangle + |1\rangle|\mu\rangle \sum_{i=1}^{M} |i\rangle}{\sqrt{2}}. \qquad (8)$$

Then we apply a Hadamard gate on qubit of the first register to obtain

$$\frac{1}{2} \left(|0\rangle \sum_{j=1}^{d} \sum_{i=1}^{M} (x_j^i + \mu_j)|i\rangle|j\rangle + |1\rangle \sum_{j=1}^{d} \sum_{i=1}^{M} (x_j^i - \mu_j)|i\rangle|j\rangle \right). \qquad (9)$$

As the final step measure the first register in $|1\rangle$, the remaining qubits collapse into the state

$$\sum_{j=1}^{d} \sum_{i=1}^{M} (x_j^i - \mu_j)|i\rangle|j\rangle = \sum_{j=1}^{d} |\chi_j\rangle|j\rangle. \qquad (10)$$

The success probability P_χ of this measurement is

$$P_\chi = \mathrm{tr}(|1\rangle\langle 1|H|\Psi_2\rangle\langle\Psi_2|H) = \sum_{k,l=1}^{d} \langle \chi_k | \chi_l \rangle. \qquad (11)$$

In this process, our gate resource count is $O(1)$. The same trick as above, for the test quantum state $|x^0\rangle$ and the mean state $|\mu\rangle$, we can easily prepare state $\sum_{j=1}^{d} \sum_{i=1}^{M} (x_j^0 - \mu_j)|i\rangle|j\rangle = \sum_{j=1}^{d} |\chi_j^0\rangle|j\rangle$, where $|\chi_j^0\rangle = \sum_{i=1}^{M} (x_j^0 - \mu_j)|i\rangle$.

Algorithm 2: Efficiently computing $\sum_{j=1}^{d} \ln \delta_j$

Step 1. According to Eq. (10), we obtain the quantum state $\sum_{j=1}^{d} |\chi_j\rangle|j\rangle$.

Step 2. Add an ancilla qubit $|0\rangle$ and perform a controlled unitary operator R_2 to obtain

$$\sum_{j=1}^{d} |\chi_j\rangle|j\rangle(\ln \chi_j |0\rangle + \sqrt{1 - 2\ln \chi_j}|1\rangle).$$

Step 3. Uncompute the second register. The system state is

$$\sum_{j=1}^{d} |j\rangle(\ln \chi_j |0\rangle + \sqrt{1 - 2\ln \chi_j}|1\rangle).$$

Step 4. Measure the observable $M_2 = I \otimes |0\rangle\langle 0|$ and the expectation $\langle M_2 \rangle = 2 \sum_{j=1}^{d} \ln \chi_j$.

With these preparations, we present Algorithm 1 for calculating the $\sum_{j=1}^{d} \frac{(x_j^0 - \mu_j)^2}{2\sigma_j^2}$. In step 2, the unitary operator R_1 transform $|0\rangle$ to $\frac{\chi_j^0}{\chi_j}|0\rangle + \sqrt{1 - \left(\frac{\chi_j^0}{\chi_j}\right)^2}|1\rangle$ controlled by the states $|\chi_j\rangle$ and $|\chi_j^0\rangle$. Thus, we can easily deduce $\sum_{j=1}^{d} \frac{(x_j^0 - \mu_j)^2}{2\sigma_j^2} = \frac{1}{2}\langle M_1 \rangle$.

At the same time, we present Algorithm 2 for computing $\sum_{j=1}^{d} \ln \delta_j$. In step 2, the unitary operator R_2 transform $|0\rangle$ to $\ln \chi_j |0\rangle + \sqrt{1 - 2\ln \chi_j}|1\rangle$ controlled by the states $|\chi_j\rangle$. Thus, we can easily deduce $\sum_{j=1}^{d} \ln \delta_j = \frac{1}{2}\langle M_2 \rangle$. Then, given a probability threshold ϵ, these two algorithms allow us to

detect anomalies based upon Algorithm 3. If

$$\ln p(|x^0\rangle) = -\frac{d}{2}\ln 2\pi - \frac{1}{2}(\langle M_1\rangle + \langle M_2\rangle) < \ln \epsilon,$$

one can flag $|x^0\rangle$ as a normal quantum state, otherwise, $|x^0\rangle$ is an abnormal quantum state. The overall resource complexity of Algorithm 3 is $O(\log M)$.

Algorithm 3: Quantum anomaly detection with density estimation

---

Input. The quantum state training set $\{|x^i\rangle\}_{i=1}^M$, the test quantum state $|x^0\rangle$. and a suitable threshold ϵ.

Step 1. Apply Algorithm 1 to obtain $\langle M_1\rangle$ using quantum gate resources $O(\log M)$.

Step 2. Apply Algorithm 2 to obtain $\langle M_2\rangle$ using quantum gate resources $O(1)$.

Step 3. If

$$\ln p(|x^0\rangle) = -\frac{d}{2}\ln 2\pi - \frac{1}{2}(\langle M_1\rangle + \langle M_2\rangle) < \ln \epsilon,$$

then $|x^0\rangle$ is a normal quantum state. Otherwise, $|x^0\rangle$ is an abnormal quantum state.

IV. QUANTUM ANOMALY DETECTION ALGORITHM WITH MULTIVARIATE GAUSSIAN DISTRIBUTION

Here, we show how to identify an anomaly with a multivariate Gaussian distribution on a quantum computer. For simplicity, we transform Eq. (4) to a fairly easy form,

$$p_{\text{test}} = (x - \mu)^T C^{-1}(x - \mu) > 2\ln\left((2\pi)^{\frac{n}{2}}|C|^{\frac{1}{2}}\varepsilon\right). \quad (12)$$

We also give two algorithms to efficiently compute the subitems p_{test} and the determinant of C. In subsection A, we show the computation of p_{test}. Then, we describe the computation process of estimating the determinant of a given Hermitian positive definite operator in subsection B. Since our algorithms make use of a quantum linear systems algorithm [15], we achieve an exponential speed-up over its classical counterpart.

A. Efficient computation of p_{test}

For a quantum anomaly detection algorithm with multivariate Gaussian distribution, first we define the centered state of the training quantum state by

$$|z^i\rangle = \sum_{j=1}^d z_j^i |j\rangle, \quad (13)$$

where z_j^i is the jth component of the vector z^i and $z^i = x^i - \mu$. Following the method of [19], we will assume the following oracle:

$$O_1(|i\rangle|0^{\otimes \log d}\rangle) \rightarrow |i\rangle|z^i\rangle, \quad (14)$$

to get the state $|z^i\rangle$ in time $O(\log d)$. Obviously, given a new example x^0, we can also obtain the corresponding centered state $|z^0\rangle$. The covariance matrix can be rewritten in terms of quantum states as $C = \frac{1}{M-1}\sum_{i=1}^M |z^i\rangle\langle z^i|$. It is also proportional to a density matrix $\mathcal{C} = \frac{C}{\text{tr}(C)}$, where $\text{tr}(C) =$

$\frac{1}{M-1}\sum_{i=1}^M \sum_{j=1}^d z_j^i(z_j^i)^*$. Since $\mathcal{C}$ is a Hermitian operator, we can use the method of [27] to exponentiate this operator. Assuming the operator $\mathcal{C}$ has spectral decomposition $\mathcal{C} = \sum_{k=1}^d \lambda_k |u_k\rangle\langle u_k|$, we can obtain

$$p_{\text{test}} = \langle z^0|C^{-1}|z^0\rangle = \frac{\langle z^0|\mathcal{C}^{-1}|z^0\rangle}{\text{tr}(C)} = \frac{1}{\text{tr}(C)}\sum_{j=1}^d \frac{\beta_j^2}{\lambda_j}. \quad (15)$$

At the same time, we decompose $|z^0\rangle$ in the eigenbasis of $\mathcal{C}$. To avoid exponential small eigenvalues, we define an effective condition number κ and take into account only eigenvalues in the interval $[1/\kappa, 1]$. Here, we give a quantum Algorithm 4 for tackling this problem based on the phase estimation [28]. The expectation value for the measurement in the final state $|\Psi_3\rangle$ is $\langle M_3\rangle = \sum_j \frac{\beta_j^2}{\lambda_j}$. Thus, we can easily deduce the value of $p_{\text{test}} = \langle M_3\rangle$.

Algorithm 4: Efficiently computing p_{test}

Step 1. Prepare the quantum states $|0\rangle$ and $|z^0\rangle$ in registers A, B

Step 2. Since the operator $\mathcal{C}$ is Hermitian positive semidefinite, use the well-known technique of phase estimation [28] to obtain the eigenvalue/eigenvector, in time $O(\log d)$:

$$\sum_{j=1}^d \beta_j|\lambda_j\rangle_A|u_j\rangle_B, \beta_j = \langle z^0|u_j\rangle.$$

Step 3. Add a joint qubit $|0\rangle$ and invert the square root of eigenvalues to obtain

$$\sum_{j=1}^d \beta_j|\lambda_j\rangle_A|u_j\rangle_B\left(\frac{1}{\sqrt{\lambda_j}}|0\rangle + \sqrt{1 - \frac{1}{\lambda_j}}|1\rangle\right).$$

Step 4. Uncompute the output of the phase estimation and obtain the state

$$|\Psi_3\rangle = \sum_{j=1}^d \beta_j|u_j\rangle_B\left(\frac{1}{\sqrt{\lambda_j}}|0\rangle + \sqrt{1 - \frac{1}{\lambda_j}}|1\rangle\right).$$

Step 5. Measure the observable $M_3 = I_B \otimes |0\rangle\langle 0|$ and the expectation $\langle M_3\rangle = \sum_{j=1}^d \frac{\beta_j^2}{\lambda_j}$.

B. Efficient estimation $\ln |C|^{\frac{1}{2}}$

In this subsection, we will describe a method to estimate the determinant of a Hermitian positive definite operator, which plays key roles in machine learning and optimization [29]. However, the determinant of matrices has been less investigated from a quantum information perspective. For a Hermitian operator $\mathcal{B} \in \mathcal{R}^{n \times n}$ with spectral decomposition $\mathcal{B} = \sum_{i=1}^n \lambda_i u_i u_i^\dagger$, the determinant of $\mathcal{B}$ is given by

$$|\mathcal{B}| = \Pi_{i=1}^n \lambda_i = e^{\sum_{i=1}^d \ln \lambda_i}. \quad (16)$$

Thus, we only need to achieve the exponential of $\sum_{i=1}^d \ln \lambda_i$. Motivated by the Harrow-Hassidim-Lloyd algorithm [15], we design Algorithm 5 to estimate the determinant of Hermitian positive definite operator. Next, we would like to establish error bounds on the final result in Algorithm 5. We first define

---

Algorithm 5: Estimating the determinant of $\mathcal{B}$

Step 1. Since the operator $\mathcal{B}$ is Hermitian positive semidefinite, use the the well-known technique of phase estimation [28] to obtain the eigenvalue/eigenvector.

Step 2. Apply the map taking $|\lambda_i\rangle$ to $\ln \lambda_i |\lambda_i\rangle$ by an ancilla qubit. Once it succeeds, we go to the next step.

Step 3. Uncompute the eigenvalue register and obtain the state $\sum_{i=0}^{N} |u_i\rangle(\ln \lambda_i|0\rangle + \sqrt{1 - (\ln \lambda_i)^2}|1\rangle)$.

Step 4. Measure the observable $M = I_B \otimes |0\rangle\langle 0|$ and the expectation $\langle M \rangle = \sum_{i=0}^{N} \ln \lambda_i$. Then the determinant $|\mathcal{B}| = e^{\langle M \rangle}$.

---

the map

$$|g(\lambda)\rangle := \ln \lambda |0\rangle + \sqrt{1 - 2\ln \lambda}|1\rangle. \tag{17}$$

This allows us to obtain the logarithm of the eigenvalue. We will make use of the following lemma to bound the error.

Lemma 1. The map $\lambda \mapsto |g(\lambda)\rangle$ is $O(\kappa^2)$-Lipschitz, meaning that for any $\lambda_i \neq \lambda_j$,

$$\||g(\lambda_i)\rangle - |g(\lambda_j)\rangle\| \leqslant c\kappa^2 |\lambda_i - \lambda_j| \tag{18}$$

for some $c = O(1)$.

Proof. Since $\lambda \mapsto |g(\lambda)\rangle$ is continuous everywhere and differentiable, it suffices to bound the norm of the derivative of $|g(\lambda)\rangle$. We consider

$$\frac{d}{d\lambda}|g(\lambda)\rangle = \frac{1}{\lambda}|0\rangle - \frac{1}{\lambda\sqrt{1-2\ln\lambda}}|1\rangle. \tag{19}$$

Next, the norm of $\frac{d}{d\lambda}|g(\lambda)\rangle$ is

$$\frac{1}{\lambda^2} + \frac{1}{\lambda^2(1-2\ln\lambda)} \leqslant \frac{2}{\lambda^2} = 2\kappa^2. \tag{20}$$

This completes the proof, with $c = 2$. ∎

If we require result error to be of $O(\epsilon)$, we need to take the phase estimation accuracy to be $\delta = O(\frac{\epsilon}{\kappa^2})$.

Using classical computation theory, the determinant can be expressed as a sum of products of entries of the operator where each product has N items, and this expression grows rapidly with the size of the operator, $O(N!)$. However, the runtime of the proposed algorithm is $O(poly(\log N))$, which is an exponential improvement over its classical counterpart.

Based on these two algorithms and Eq. (12), we can flag example quantum state $|x^0\rangle$ as an anomaly if $\langle M_3 \rangle > 2\ln[(2\pi)^{\frac{d}{2}}|C|^{\frac{1}{2}}\varepsilon]$, otherwise it is a normal state.

V. QUANTUM KERNEL PCA AND ONE-CLASS SVM FOR CLASSICAL DATA

In [17], the authors presented quantum kernel PCA and one-class SVM for quantum state anomaly detection. In this section, we apply the quantum kernel PCA and one-class SVM for detecting a classical data anomaly. After encoding the discrete data set $\{x^i\}_{i=1}^{M}$ to quantum state $\{|x^i\rangle\}_{i=1}^{M}$, we can directly use the quantum one-class SVM to identify the anomaly. Although quantum kernel PCA can also directly detect an anomaly for classical data, the resources required are much more than for our modified quantum kernel PCA discussed below.

The modified quantum kernel PCA

We can rewrite the centered quantum state as

$$|z^i\rangle = \sum_{j=1}^{d} z_j^i |j\rangle, \tag{21}$$

the same as Eq. (13) in Sec. 4 A. We can prepare the test centered state by $|z^0\rangle = \sum_{j=1}^{d} z_j^0 |j\rangle$. The centered sampled covariance matrix is $C = \frac{1}{M-1}\sum_{i=1}^{M} |z^i\rangle\langle z^i|$. The proximity measure $f(|z^0\rangle)$ obeys $0 \leqslant f(|z^0\rangle) \leqslant 1$, where

$$f(|z^0\rangle) = \langle z^0|I - C|z^0\rangle. \tag{22}$$

In [17], the inner product can be calculated by M times the modified swap test, and each of that requires $O(poly(\log d))$. Thus the proximity measure $f(|z^0\rangle)$ can be estimated using resources scaling as $O(poly(M \log d))$ [17]. Here, we present a modified algorithm to efficiently compute the proximity based on Algorithm 4 using resources scaling as $O(poly(\log d))$. More specifically, we use the method of [27] to obtain the eigenvalue of C and rotate one qubit $|0\rangle$ to $(1 - \lambda)|0\rangle + \sqrt{1 - (1 - \lambda)^2}|1\rangle$ controlled by λ. The time complexity of this process is $O(poly(\log d))$ independent of the number of training quantum states.

VI. CONCLUSION AND DISCUSSION

In the present study, we presented quantum anomaly detection with density estimation and multivariate Gaussian distribution. For an exponentially large data set, the resources are logarithmic in the dimensionality of quantum states and the number of training quantum states. Moreover, We also developed a quantum subroutine for efficiently estimating the determinant of any Hermitian operator $\mathcal{A} \in \mathcal{R}^{N \times N}$ with time complexity $O(\log(N))$. This algorithm might be a subroutine in machine learning and optimization. Finally, our work generalized the result given in [17] to tackle a classical data anomaly while the authors had only considered quantum anomaly detection for quantum states generated from quantum devices.

Although we have seen the feasibility of detecting an anomaly on a quantum computer, some open questions still require further study in the future. For example, it would be interesting to study how to select a suitable threshold ϵ efficiently on a quantum computer. Finally, both algorithms are described in the standard gate-based model of quantum computing. In 2007, Aharonov *et al.* stated that adiabatic quantum computation [30] is as powerful as conventional quantum computation [31]. Thus, it would be interesting to consider other alternative quantum anomaly detection algorithms with adiabatic quantum computing.

[1] E. Aleskerov, B. Freisleben, and B. Rao, in *Proceedings of the IEEE/IAFE 1997 Computational Intelligence for Financial Engineering (CIFEr)* (IEEE, Piscataway, NJ, 1997).

[2] J. Laurikkala, M. Juhola, and E. Kentala, Fifth International Workshop on Intelligent Data Analysis in Medicine and Pharmacology, Berlin, Germany, August 2000.

[3] C. Spence, L. Parra, and P. Sajda, in *Proceedings of the IEEE Workshop on Mathematical Methods in Biomedical Image Analysis* (IEEE Computer Society, Washington, DC, 2003), p. 3.

[4] M. M. Breunig, H.-P. Kriegel, R. T. Ng, and J. Sander, in *Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data* (ACM Press, New York, 2000).

[5] S. Ramaswamy, R. Rastogi, and K. Shim, in *Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data* (ACM Press, New York, 2000).

[6] A. Lazarevic and V. Kumar, in *Proceedings of the Eleventh ACM SIGKDD International Conference on Knowledge Discovery in Data Mining* (ACM Press, New York, 2005).

[7] M.-L. Shyu, S.-C. Chen, K. Sarinnapakorn, and L. Chang, in *Proceedings of the Third IEEE International Conference on Data Mining, Melbourne, FL, 2003* (IEEE, Piscataway, NJ, 2003), pp. 172–179.

[8] R. Laxhammar, G. Falkman, and E. Sviestins, 12th International Conference on Information Fusion, Seattle, WA, USA, 6-9 July 2009.

[9] S. M. Erfani, S. Rajasegarar, S. Karunasekera, and C. Leckie, Pattern Recogn. **58**, 121 (2016).

[10] D. Carrera, G. Boracchi, A. Foi, and B. Wohlberg, the International Joint Conference on Neural Networks, Killarney, Ireland, 12-17 July 2015.

[11] H. Hoffman, Pattern Recogn. **40**, 863 (2007).

[12] M. Markou and S. Singh, Signal Process. **83**, 2481 (2003).

[13] P. Shor, in *Symposium on Foundations of Computer Science* (IEEE, Piscataway, NJ, 1994), pp. 124–134

[14] L. K. Grover, Phys. Rev. Lett. **79**, 325 (1997).

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

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

[17] N. Liu and P. Rebentrost, Phys. Rev. A **97**, 042315 (2018).

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

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

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

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

[22] G. Sentís, J. Calsamiglia, and R. Muñoz-Tapia, Phys. Rev. Lett. **119**, 140506 (2017).

[23] G. Sentís, E. Martinez-Vargas, and R. Muñoz-Tapia, Phys. Rev. A **98**, 052305 (2018).

[24] V. Chandola, A. Banerjee, and V. Kumar, ACM Comput. Surv. **41**, 15 (2009).

[25] R. Kohavi, in *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence* (Morgan Kaufmann, San Francisco, 1995), Vol. 2, pp. 1137–1145.

[26] J. D. Rodriguez, A. Perez, and J. A. Lozano, IEEE Trans. Pattern Anal. Mach. Intell. **32**, 569 (2010).

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

[28] A. Luis and J. Perina, Phys. Rev. A **54**, 4564 (1996).

[29] C. M. Bishop, *Pattern Recognition and Machine Learning*, (Springer, New York, 2006).

[30] E. Farhi, J. Goldstone, S. Gutmann, and M. Sipser, arXiv:quant-ph/0001106.

[31] D. Aharonov, W. van Dam, J. Kempe, Z. Landau, S. Lloyd, and O. Regev, SIAM J. Comput. **37**, 166 (2007).