

Single-qubit quantum classifier based on gradient-free optimization algorithm

Anqi Zhang(张安琪), Kelun Wang(王可伦), Yihua Wu(吴逸华), and Sheng-Mei Zhao(赵生妹)
Citation: Chin. Phys. B, 2023, 32 (**10**): 100308. DOI: 10.1088/1674-1056/acb75e
Journal homepage: http://cpb.iphy.ac.cn; http://iopscience.iop.org/cpb

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

Broadband multi-channel quantum noise suppression and phase-sensitive modulation based on entangled beam

Ke Di(邸克), Shuai Tan(谈帅), Anyu Cheng(程安宇), Yu Liu(刘宇), and Jiajia Du(杜佳佳)
Chin. Phys. B, 2023, 32 (**10**): 100302. DOI: 10.1088/1674-1056/acd8a8

Purification in entanglement distribution with deep quantum neural network

Jin Xu(徐瑾), Xiaoguang Chen(陈晓光), Rong Zhang(张蓉), and Hanwei Xiao(肖晗微)
Chin. Phys. B, 2022, 31 (**8**): 080304. DOI: 10.1088/1674-1056/ac6330

Deterministic remote state preparation of arbitrary three-qubit state through noisy cluster-GHZ channel

Zhihang Xu(许智航), Yuzhen Wei(魏玉震), Cong Jiang(江聪), and Min Jiang(姜敏)
Chin. Phys. B, 2022, 31 (**4**): 040304. DOI: 10.1088/1674-1056/ac2b17

Degenerate asymmetric quantum concatenated codes for correcting biased quantum errors

Ji-Hao Fan(樊继豪), Jun Li(李骏), Han-Wu Chen(陈汉武), and Wen-Jie Liu(刘文杰)
Chin. Phys. B, 2021, 30 (**12**): 120302. DOI: 10.1088/1674-1056/ac05a8

Steady and optimal entropy squeezing for three types of moving three-level atoms coupled with a single-mode coherent field

Wen-Jin Huang(黄文进) and Mao-Fa Fang(方卯发)
Chin. Phys. B, 2021, 30 (**1**): 010304. DOI: 10.1088/1674-1056/abc542

Single-qubit quantum classifier based on gradient-free optimization algorithm

Anqi Zhang(张安琪)[1], Kelun Wang(王可伦)[1], Yihua Wu(吴逸华)[1], and Sheng-Mei Zhao(赵生妹)[1,2,†]

[1] *Institute of Signal Processing and Transmission, Nanjing University of Posts and Telecommunications, Nanjing* 210003, *China*
[2] *Key Laboratory of Broadband Wireless Communication and Sensor Network Technology, Ministry of Education, Nanjing* 210003, *China*

(Received 3 October 2022; revised manuscript received 16 January 2023; accepted manuscript online 31 January 2023)

A single-qubit quantum classifier (SQC) based on a gradient-free optimization (GFO) algorithm, named the GFO-based SQC, is proposed to overcome the effects of barren plateaus caused by quantum devices. Here, a rotation gate $R_X(\phi)$ is applied on the single-qubit binary quantum classifier, and the training data and parameters are loaded into ϕ in the form of vector multiplication. The cost function is decreased by finding the value of each parameter that yields the minimum expectation value of measuring the quantum circuit. The algorithm is performed iteratively for all parameters one by one until the cost function satisfies the stop condition. The proposed GFO-based SQC is demonstrated for classification tasks in Iris and MNIST datasets and compared with the Adam-based SQC and the quantum support vector machine (QSVM). Furthermore, the performance of the GFO-based SQC is discussed when the rotation gate in the quantum device is under different types of noise. The simulation results show that the GFO-based SQC can reach a high accuracy in reduced time. Additionally, the proposed GFO algorithm can quickly complete the training process of the SQC. Importantly, the GFO-based SQC has a good performance in noisy environments.

Keywords: single-qubit quantum classifier, gradient-free, parameters optimizing, barren plateau, quantum noise

PACS: 03.67.Ac **DOI:** 10.1088/1674-1056/acb75e

1. Introduction

Classification has been one of the main issues in quantum computing.[1–6] There are several types of quantum classifiers. some are inspired by their classical counterparts with their kernel parts replaced by a high-depth quantum circuit,[7,8] and some are inspired by the neural networks with multi-layer quantum circuit structures.[9–13] These quantum classifiers always require expensive subroutines. Others are termed quantum circuit learning with a hybrid quantum-classical (HQC) structure[14–21] and train a shallow-depth parameterized quantum circuit with a classical optimizer.

By using a small amount of quantum resources, quantum classifiers based on an HQC structure have become a promising candidate in the noisy intermediate scale quantum (NISQ) era. For example, Ref. [19] built a model with only two qubits in the quantum circuit that could simultaneously manipulate two classes of training samples, which reduced the training processing time by half. Furthermore, Ref. [20] trained multi-class data, more than 2 classes, simultaneously for classification tasks in the qubit system. Reference [21] presented single-qubit processing units which can be seen as a single-qubit quantum classifier (SQC) and the quantum circuit was organized as a series of data re-uploads, in which a multiple data re-uploading method is used to increase circuit express ability. The advantage of SQCs is that they avoid the connectivity constraints and embedding problems of qubits in quantum devices. But the SQC still has a difficulty that needs to be solved when using the quantum–classical hybrid scheme to implement it.

The gradient-based parameter optimization method inside a quantum classifier, such as the Adam optimizer, may lead to the barren plateau[22] (i.e., vanishing gradient) problem. There are some solutions to overcome this issue. For instance, a layer wise learning optimization method was proposed by Skolik *et al.*[23] to update only a subset of the parameters in each training iteration. A Bayesian optimization (BO) based on Gaussian process regression (GPR) and noisy expected improvement (NEI) was proposed by Iannelli *et al.*[24] to provide a more precise estimation of the ground state energy in a few iterations. An efficient method for simultaneously optimizing both the structure and the parameter values of quantum circuits was proposed by Ostaszewski *et al.*[25] It has been demonstrated that the quantum convolutional neural networks (QCNNs) proposed by Pesah *et al.*[26] don't show barren plateaus for the variance of the gradient vanishes no faster than polynomially. However, whether a single-qubit quantum classifier yields higher accuracy in solving classification tasks has not been discussed yet.

In this work, we propose a gradient-free optimization (GFO) algorithm based single-qubit quantum classifier (SQC), named the GFO-based SQC. In the GFO algorithm, a rotation gate $R_X(\phi)$ applied on one qubit initialized as $|0\rangle$ can be regarded as a binary quantum classifier. Training data and parameters of the quantum classifier are loaded into ϕ with the

†Corresponding author. E-mail: zhaosm@njupt.edu.cn

© 2023 Chinese Physical Society and IOP Publishing Ltd

form of vector-multiplication. The cost function is decreased by finding the value of each parameter that yields the minimum expectation value of measuring the quantum circuit until it meets the stop condition. After several iterations, the cost function is converged and the optimal parameters are obtained. The classification information for unclassified data can be determined by the result with the highest probability after measuring the quantum circuit of the SQC.

The advantages of the proposed GFO-based SQC are three-fold. 1) Compared with the gradient-based algorithm, the GFO algorithm requires less time for optimizing the parameters. 2) Parameters are loaded as weights of each component of the training data in the GFO-based SQC, which can save more qubits and quantum gate resources and further reduce the barren plateau problem. 3) The updated value of the parameter is calculated with the usage of the statistical average method, thus increasing the accuracy of classification.

2. The single-qubit quantum classifier based on the gradient-free optimization algorithm

The framework of the GFO-based SQC for optimizing parameters is shown in Fig. 1. Due to the nature of the GFO algorithm where only one of all the parameters in the quantum circuit is updated in each iteration while other parameters are fixed, we suppose that the former $i-1$ parameters have been updated already. In the training processes, a rotation gate $R_X(\phi)=\exp(-\mathrm{i}\frac{\phi}{2}\sigma_X)$ applied on one qubit is used as the quantum trainer, where $\phi\in(-\pi,\pi]$ is the angle of rotation and σ_X is the Pauli-X matrix. $\boldsymbol{X}=(x_1,\ldots,x_n)$ is a n-dimension training data in a batch used for updating the $(i-1)$-th parameter and $\boldsymbol{X}$ is selected randomly from the training dataset which is defined as $S=\{\boldsymbol{X},f(\boldsymbol{X})\}$, and $f(\boldsymbol{X})=0$ or 1 is the label of $\boldsymbol{X}$. The method of loading training data and parameters is to perform a multiplication operation between two n-dimensional vectors, $\boldsymbol{X}$ and $\boldsymbol{\theta}$, which makes two inputs into a single number, e.g., $\phi=\sum_{i=1}^{n}\theta_i x_i$, $\boldsymbol{\theta}=(\theta_1,\ldots,\theta_n)$ is the parameter vector that needs to be tuned for the SQC. After loading a batch of training data $\boldsymbol{X}$s and parameter vector $\boldsymbol{\theta}$ into the quantum circuit, the value of the cost function $C_{\theta_{i-1}}$ specified by the $(i-1)$-th parameter, where $C_{\theta_{i-1}}=\frac{1}{M}\sum_{m=1}^{M}(1-|\langle O\rangle_{\boldsymbol{X}_m,\theta_{i-1}}|^2)$, is obtained by using the Hermitian operator $\langle O\rangle$ to measure the quantum circuit. $\boldsymbol{X}_m$ is the m-th training data in the batch. $C_{\theta_{i-1}}$ is checked to see whether it satisfies the stop condition of the training processes. If $C_{\theta_{i-1}}$ satisfies the stop condition, the optimal parameter vector $\boldsymbol{\theta}^*$ for the testing processes is output; otherwise, the i-th parameter θ_i in $\boldsymbol{\theta}$ is continually updated. In the testing processes, unclassified data $\boldsymbol{X}_{\text{unclassified}}$ and $\boldsymbol{\theta}^*$ are loaded into the same quantum circuit, named the quantum classifier. The classification information for $\boldsymbol{X}_{\text{unclassified}}$ can be obtained after performing the measurements.

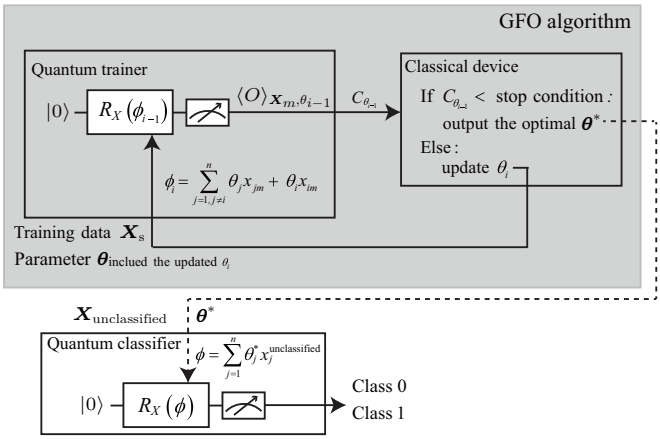


Fig. 1. The framework for optimizing parameters of the SQC based on the GFO algorithm. The arrows in the figure represent the flow of information. In the training processes (gray box), the training data $\boldsymbol{X}$ and the parameters $\boldsymbol{\theta}$ are fed into ϕ of the rotation gate in the form of vector-multiplication every iteration. The value of the cost function $C_{\theta_{i-1}}$ is obtained by measuring the quantum circuit with the $(i-1)$-th parameters and checking whether it's less than the given stop condition. If not, continue the next parameter-optimizing iteration where other parameters are fixed except θ_i, which is updated in the i-th iteration; otherwise, begin the testing processes, where the optimal parameters vector $\boldsymbol{\theta}^*$ and unclassified data $\boldsymbol{X}_{\text{unclassified}}$ are fed into the same circuit. The classification information of $\boldsymbol{X}_{\text{unclassified}}$ can be specified with the highest probability by measurement of the quantum circuit.

Here, we detail the findings of the optimal parameter vector $\boldsymbol{\theta}^*$. Inspired by the optimization strategies in Ref. [25], we consider a special case where a subset only contains one parameter to optimize in the parameterized quantum gate for the quantum circuit. Firstly, the expectation value of the quantum circuit with structure U and initial state ρ can be rewritten as a function of angle ϕ [25] which has sinusoidal form as $\langle O\rangle_\phi=\mathrm{Tr}(OU\rho U^\dagger)=A\sin(\phi+B)+C$, where A, B and C are coefficients that need to be determined. Then, we consider the task of minimizing the value of the cost function C_{θ_i} by maximizing every $|\langle O\rangle_{\boldsymbol{X}_m,\theta_i}|$ part with the i-th parameter, which can be ensured if the coefficients are estimated. Therefore, using the definition of the matrix exponential, the expectation value can be expressed as

$$\begin{aligned}\langle O\rangle &= \mathrm{Tr}[OR_X(\phi)\rho R_X^\dagger(\phi)]=\mathrm{Tr}[O\,\mathrm{e}^{-\mathrm{i}\frac{\phi}{2}\sigma_X}\rho\,\mathrm{e}^{\mathrm{i}\frac{\phi}{2}\sigma_X}]\\ &= \mathrm{Tr}\left(O(\cos\frac{\phi}{2}I-\mathrm{i}\sin\frac{\phi}{2}\sigma_X)\rho\left(\cos\frac{\phi}{2}I+\mathrm{i}\sin\frac{\phi}{2}\sigma_X\right)\right)\\ &= \cos^2\frac{\phi}{2}\mathrm{Tr}(OI\rho I)+\mathrm{i}\sin\frac{\phi}{2}\cos\frac{\phi}{2}\mathrm{Tr}(O[\rho,\sigma_X])\\ &\quad+\sin^2\frac{\phi}{2}\mathrm{Tr}(O\sigma_X\rho\sigma_X).\end{aligned}\tag{1}$$

We set $\phi=(\theta_{im}x_{im}+\boldsymbol{\theta}'_m\boldsymbol{X}'_m)$, $\boldsymbol{\theta}'_m\boldsymbol{X}'_m=\sum_{j=1}^{n}\theta_{jm}x_{jm}, j\neq i$, and other $n-1$ parameters, except θ_{im}, are fixed to their current values. We note that $\mathrm{Tr}(OI\rho I)$ is simply the expectation when the circuit is evaluated at $\phi=0$. That is, $\langle O\rangle_0=\mathrm{Tr}(OI\rho I)=\mathrm{Tr}(O\rho)$. Likewise, $\langle O\rangle_\pi=\mathrm{Tr}(O\sigma_X\rho\sigma_X)$ when $\phi=\pi$, and $\mathrm{i}\mathrm{Tr}(O[\rho,\sigma_X])=\langle O\rangle_{\frac{\pi}{2}}-\langle O\rangle_{-\frac{\pi}{2}}$ when $\phi=\frac{\pi}{2},-\frac{\pi}{2}$ respectively. Then, the expectation value characterized by the i-th parameter θ_{im} when the training data is $\boldsymbol{X}_m$ can be ex-

pressed as

$$\begin{aligned}\langle O\rangle_{\theta_{im}} &= \frac{1}{2}[(\langle O\rangle_0-\langle O\rangle_\pi)\cos(\boldsymbol{\theta}'_m\boldsymbol{X}'_m)\\&+(\langle O\rangle_{\frac{\pi}{2}}-\langle O\rangle_{-\frac{\pi}{2}})\sin(\boldsymbol{\theta}'_m\boldsymbol{X}'_m)]\cos(\theta_{im}x_{im})\\&+\frac{1}{2}[(\langle O\rangle_{\frac{\pi}{2}}-\langle O\rangle_{-\frac{\pi}{2}})\cos(\boldsymbol{\theta}'_m\boldsymbol{X}'_m)\\&+(\langle O\rangle_0-\langle O\rangle_\pi)\sin(\boldsymbol{\theta}'_m\boldsymbol{X}'_m)]\sin(\theta_{im}x_{im})\\&+\frac{1}{2}(\langle O\rangle_0+\langle O\rangle_\pi).\end{aligned}\tag{2}$$

From the identity deformation of the sinusoidal function, we can obtain a compact expression of Eq. (2) as

$$\begin{aligned}\langle O\rangle_{\theta_{im}} &= A\cos(\theta_{im}x_{im})+B\sin(\theta_{im}x_{im})+C\\&= \sqrt{A^2+B^2}\sin(\theta_{im}x_{im}+\arctan2(A,B))+C,\end{aligned}\tag{3}$$

where

$$\begin{aligned}A &= \frac{1}{2}[(\langle O\rangle_0-\langle O\rangle_\pi)\cos(\boldsymbol{\theta}'_m\boldsymbol{X}'_m)\\&+(\langle O\rangle_{\frac{\pi}{2}}-\langle O\rangle_{-\frac{\pi}{2}})\sin(\boldsymbol{\theta}'_m\boldsymbol{X}'_m)],\\B &= \frac{1}{2}[(\langle O\rangle_{\frac{\pi}{2}}-\langle O\rangle_{-\frac{\pi}{2}})\cos(\boldsymbol{\theta}'_m\boldsymbol{X}'_m)\\&+(\langle O\rangle_0-\langle O\rangle_\pi)\sin(\boldsymbol{\theta}'_m\boldsymbol{X}'_m)],\\C &= \frac{1}{2}(\langle O\rangle_0+\langle O\rangle_\pi),\end{aligned}\tag{4}$$

where $\arctan2(A,B)$ is a four quadrant arctangent of the real numbers A and B, and its value range is $(-\pi,\pi]$. It's easy to locate the maximum at $\theta_{im}x_{im}=\pi/2-\arctan2(A,B)$. Therefore, the updated value of the i-th parameter θ_{im} can be expressed as

$$\begin{aligned}\theta_{im}^{\text{upd}} &= \arg\max_{\theta_{im}}\langle O\rangle_{\theta_{im}}\\&= \frac{1}{x_{im}}\Bigg(\frac{\pi}{2}-\arctan2[(\langle O\rangle_0-\langle O\rangle_\pi)\cos(\boldsymbol{\theta}'_m\boldsymbol{X}'_m)\\&+(\langle O\rangle_{\frac{\pi}{2}}-\langle O\rangle_{-\frac{\pi}{2}})\sin(\boldsymbol{\theta}'_m\boldsymbol{X}'_m),\\&(\langle O\rangle_{\frac{\pi}{2}}-\langle O\rangle_{-\frac{\pi}{2}})\cos(\boldsymbol{\theta}'_m\boldsymbol{X}'_m)\\&+(\langle O\rangle_0-\langle O\rangle_\pi)\sin(\boldsymbol{\theta}'_m\boldsymbol{X}'_m)]\Bigg).\end{aligned}\tag{5}$$

The i-th parameter can be optimized by a batch of training data, where a strategy to obtain the average value of m i-th updated parameters in each iteration is

$$\theta_i^{\text{upd}}=\frac{1}{M}\sum_{m=1}^{M}\theta_{im}^{\text{upd}},\tag{6}$$

where M is the number of training data in a batch. The algorithm finds each updated parameter in $\boldsymbol{\theta}$ sequentially for all $i=1,\ldots,n$ to complete cycle by cycle until a stop condition is satisfied. The gradient-free method is Algorithm 1.

Algorithm 1 Gradient-free optimization algorithm for the SQC

Require: Initialize the parameters θ_i randomly, where $1\le i\le n$; set batch size M (we set M as 10 in this work); a stop condition; estimate $\langle O\rangle_0,\langle O\rangle_\pi,\langle O\rangle_{\frac{\pi}{2}},\langle O\rangle_{-\frac{\pi}{2}}$;

loop
  Calculate the value of the cost function as cost-now.
  for $i=1$ to n **do**
    For all parameters except the i-th one.
    for $m=1$ to M **do**
      Calculate $\boldsymbol{\theta}'_m\boldsymbol{X}'_m$;
      $A=\frac{1}{2}[(\langle O\rangle_0-\langle O\rangle_\pi)\cos(\boldsymbol{\theta}'_m\boldsymbol{X}'_m)+(\langle O\rangle_{\frac{\pi}{2}}-\langle O\rangle_{-\frac{\pi}{2}})\sin(\boldsymbol{\theta}'_m\boldsymbol{X}'_m)]$;
      $B=\frac{1}{2}[(\langle O\rangle_{\frac{\pi}{2}}-\langle O\rangle_{-\frac{\pi}{2}})\cos(\boldsymbol{\theta}'_m\boldsymbol{X}'_m)+(\langle O\rangle_0-\langle O\rangle_\pi)\sin(\boldsymbol{\theta}'_m\boldsymbol{X}'_m)]$;
      $\theta_{im}^{\text{opt}}\leftarrow\frac{1}{x_{im}}(\frac{\pi}{2}-\arctan2(A,B))$.
    end for
    Get the average: $\theta_i^{\text{opt}}=\frac{1}{M}\sum_{m=1}^{M}\theta_{im}^{\text{opt}}$.
    Calculate cost value: cost-new.
    if cost-new $<$ cost-now **then**
      θ_i^{opt} changed;
      cost-now $=$ cost-new.
    else
      θ_i^{opt} remains the same.
    end if
  end for if the condition is met.
end loop

3. Results and discussion

In this section, we demonstrate the feasibility of the proposed GFO-based SQC with the Iris and MNIST datasets. We use the PennyLane[27] module in Python to realize the proposed algorithm. In order to show the performance of the proposed GFO-based SQC, the Adam-based SQC is simulated at the same time, whose parameters are optimized in a hybrid quantum–classical way. In addition, the quantum support vector machine (QSVM), which is a quantum classifier without a classical part, is also selected for comparison.

For the Iris dataset, we use 40 training samples and 10 testing samples for each class, and they are selected randomly from the Iris dataset. Here, the optimizing parameters' time is recorded when 100% classification accuracy is achieved using the GFO-based SQC, Adam-based SQC, and QSVM. The simulations are performed 20 times, where the quantum circuit has no noise. As shown in Fig. 2, the time using the GFO-based SQC is on average smaller than that using the Adam-based SQC, except the third, the ninth, and the eighteenth times, which means that the proposed GFO algorithm not only has a higher efficiency for classification, but also has a reduced time for the parameters' optimization. Here, the time using the GFO-based SQC fluctuates because the training effect of the GFO-based SQC is highly dependent on the original parameters when the dataset is relatively simple. The time cost of using the QSVM is much higher than that of using both the GFO-based SQC and the Adam-based SQC, because the QSVM requires a lot of time for the kernel calculations. Therefore, we only compare the proposed GFO-based SQC with the Adam-based SQC in the MNIST dataset.

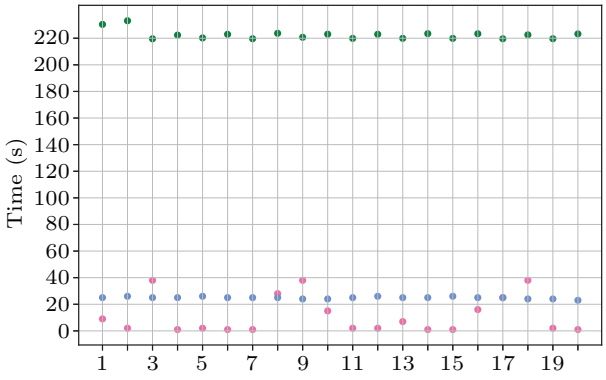


Fig. 2. The time when 100% classification accuracy is achieved using the GFO-based SQC (red), Adam-based SQC (blue), and QSVM (green) for the Iris dataset versus the simulation times.

For the MNIST dataset, we use 10000 training samples and 1000 testing samples for each class, which are selected randomly from the MNIST dataset. Each sample is processed into 32 dimensions by the rough grid feature method.[28] The simulations are performed 20 times separately based on the quantum circuit with and without noise deriving from the imperfections of quantum gates. Here, the quantum gate is affected by the five typical types of quantum noise, i.e., bit flip, depolarizing, phase flip, phase damping, and amplitude damping. Here, the error probability for the five quantum noises is assumed to be 5%. The parameters of the quantum circuit have been updated for 15 loops in each simulation.

Firstly, we compare the optimizing time and classification accuracy by using the GFO-based SQC and the Adam-based SQC for different classification tasks in the MNIST dataset. As shown in Table 1, the GFO-based SQC can complete the classification tasks for both the digits that are similarly written by hand (such as 1, 7) and those that are difficult to recognize when written (such as 2, 4, 6, 8 and 9), and takes less time than the Adam-based SQC over 15 loops.

Table 1. The optimizing time and classification accuracy using the GFO-based SQC and the Adam-based SQC for different classification tasks in the MNIST dataset.

No noise	GFO-based SQC		Adam-based SQC	
classifying	time (s)	accuracy	time (s)	accuracy
1,7	1582	97.1%	2019	99%
2,4	1785	92.6%	2213	94%
7,9	1613	86%	2029	85%
2,6	1598	80.1%	2040	82.7%
8,6	1557	90.1%	2036	93.7%

Next, we compare the time of optimizing the parameters of the two SQCs in different environments for the classification task between 0 and 1. According to Ref. [29], phase flip and phase damping have different Kraus matrices, but by simply substituting, one can check directly that the two channels yield completely the same outputs, which means that the phase flip quantum operation is exactly the same as the phase damping. Therefore, we only present the results under phase flip noise. As shown in Table 2, the optimizing time required by the GFO-based SQC is less than that of the Adam-based SQC in both no-noise and noisy environments. Furthermore, the optimizing time of the GFO-based SQC is basically stable at 1500 s, while the Adam-based SQC is heavily influenced by the environment.

Table 2. The average optimizing time using the GFO-based SQC and the Adam-based SQC in different environments for the MNIST dataset.

	GFO-based SQC	Adam-based SQC
No noise	1508 s	2201 s
Bit flip	1528 s	15498 s
Depolarizing	1440 s	2511 s
Phase flip	1493 s	2373 s
Amplitude damping	1521 s	13663 s

Finally, we demonstrate the details of the two-class classification task between 0 and 1 with and without quantum gate noise. The shadow is composed of 20 simulations with different initial parameters.

Figure 3 shows the value of cost (a) and the classification accuracy (b) against time for the two-class classification task. As shown in Fig. 3(a), the convergence ability of the GFO-based SQC is better than that of the Adam-based SQC with different initial parameters. As shown in Fig. 3(b), the GFO-based SQC can achieve a high classification accuracy faster than the Adam-based SQC and can be stable at a specific accuracy in the end.

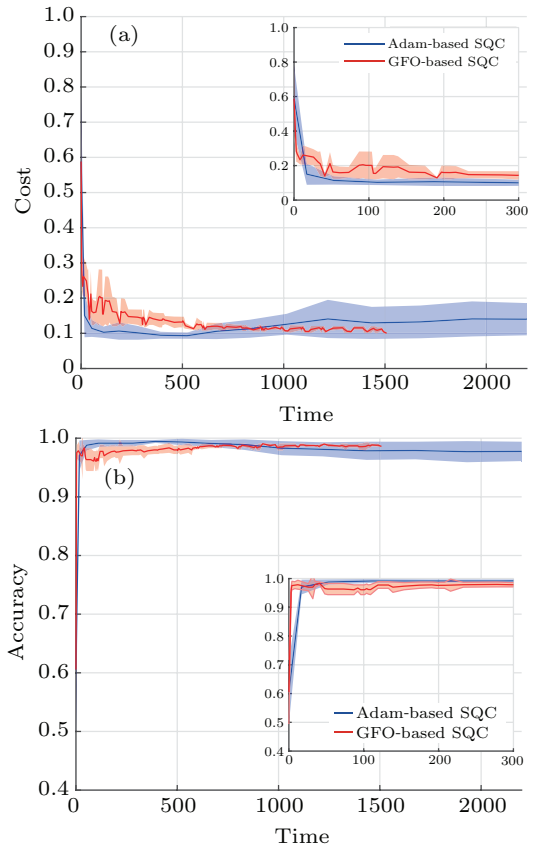


Fig. 3. Quantum gate without noise, (a) for the value of the cost function and (b) for the accuracy of classification.

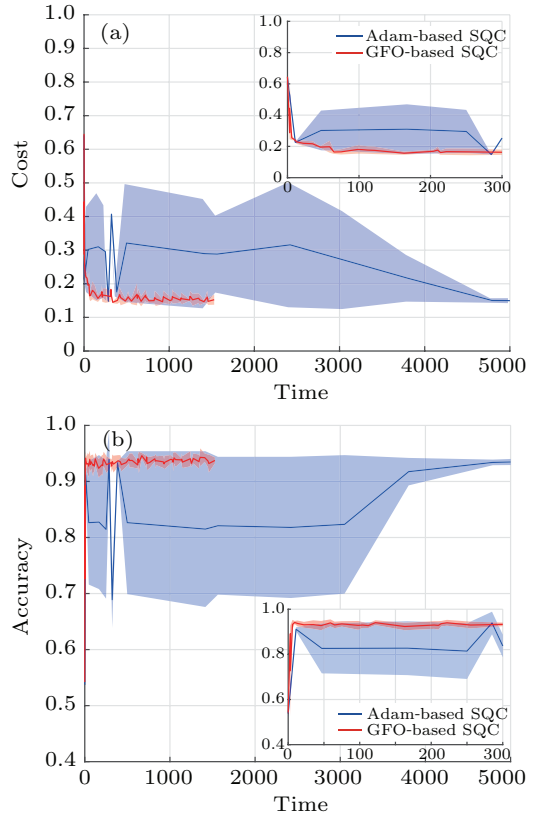


Fig. 4. Quantum gate with bit flip noise, (a) for the value of the cost function and (b) for the accuracy of classification.

Figure 4 shows the situation of bit flip noise affecting the quantum gate where (a) shows the value of cost and (b) shows the classification accuracy against time. The shaded area of the Adam-based SQC is larger than that of the GFO-based SQC, indicating that the GFO-based SQC is effective and is more adaptable in the noisy environment of bit flip and takes less time to finish 15 loops of optimizing parameters. From Fig. 4(a), the GFO-based SQC has a more stable convergence ability than the Adam-based SQC. From Fig. 4(b), the GFO-based SQC is able to reach a high accuracy faster than the Adam-based SQC.

Figure 5 shows the situation of depolarizing noise affecting the quantum gate. Figure 5(a) shows that the convergence ability of the GFO-based SQC is better than that of the Adam-based SQC, indicating that the GFO-based SQC is effective in the depolarizing noise environment and takes less time to finish 15 loops of optimizing parameters. Figure 5(b) shows the GFO-based SQC is able to reach a high accuracy faster than the Adam-based SQC and can be stable at a specific accuracy in the later stages.

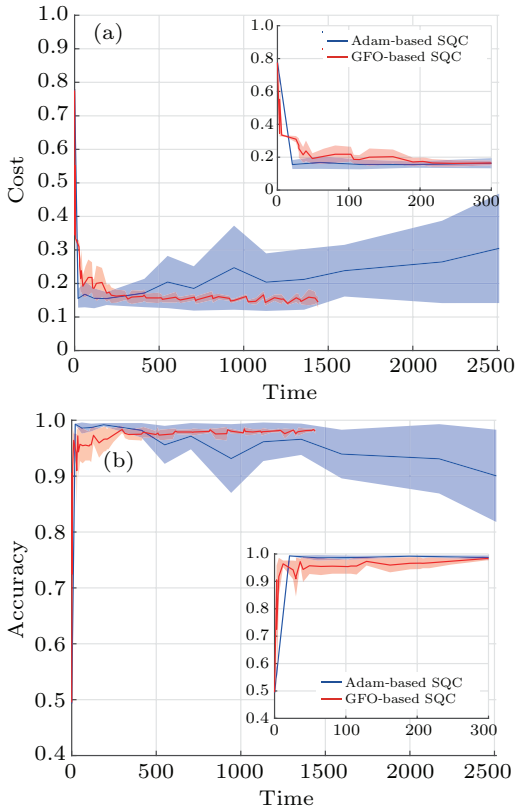


Fig. 5. Quantum gate with depolarizing noise, (a) for the value of the cost function and (b) for the accuracy of classification.

Figure 6 shows the situation of phase flip noise affecting the quantum gate. Figure 6(a) shows that the GFO-based SQC has the ability to resist the phase flip noise and keep stable continuous convergence, while the Adam-based SQC fluctuates badly. Figure 6(b) shows that the GFO-based SQC can achieve a high classification accuracy faster and can be stable at a specific accuracy in the later stages. The results in Fig. 6

show that the GFO-based SQC is effective in the noisy environment of phase flip.

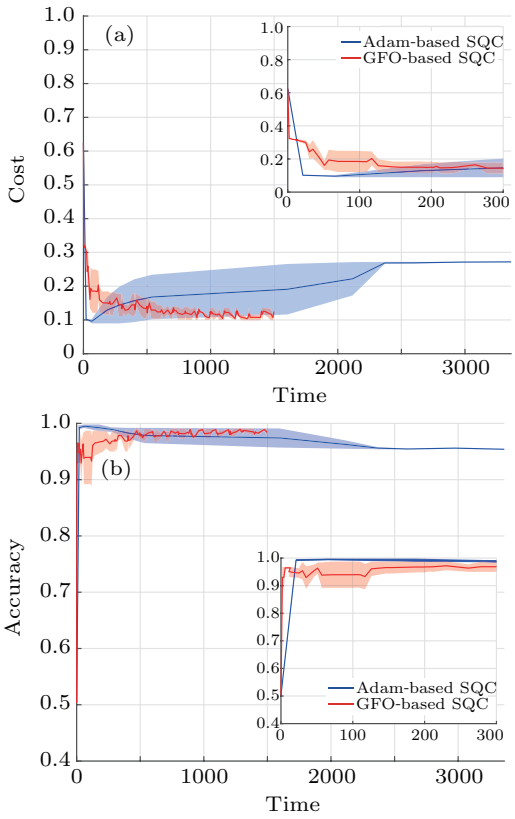


Fig. 6. Quantum gate with phase flip noise, (a) for the value of the cost function and (b) for the accuracy of classification.

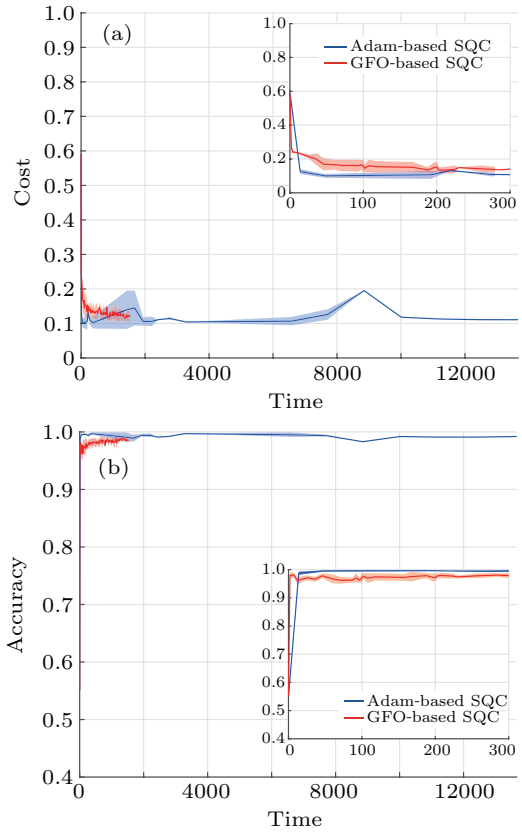


Fig. 7. Quantum gate with amplitude damping noise, (a) for the value of the cost function and (b) for the accuracy of classification.

Figure 7 shows the situation of amplitude damping noise affecting the quantum gate. As is shown in Fig. 7(a), the GFO-based SQC takes less time to finish 15 loops of optimizing parameters but the Adam-based SQC has a better convergence ability. Figure 7(b) shows that the GFO-based SQC can achieve a high classification accuracy faster, but the Adam-based SQC can reach a higher accuracy eventually.

In general, the GFO-based SQC has a certain resistance to noises and can keep the ability of continuous convergence. Compared with the Adam-based SQC, the GFO-based SQC can finish 15 loops of optimizing parameters more quickly. Although the GFO-based SQC is affected by the randomly initialized parameters, which makes the optimization performance of the GFO-based SQC fluctuate in the optimizing processes, the GFO-based SQC can always achieve a high classification accuracy faster than the Adam-based SQC and maintain stability in the end. As gradients do not need to be calculated, the GFO-based SQC saves lots of calculation time and can quickly complete the optimizing processes, while the Adam-based SQC is affected badly by the different noises and takes different times to finish the optimizing processes in different noise environments.

4. Conclusion

In this work, we have proposed a GFO-based SQC. By loading training data and parameters in a summation form, only one rotation gate has been used in the quantum circuit. The proposed GFO-based SQC has been used to optimize only one parameter in each iteration, resulting in a shorter time for completing the optimizing processes of the SQC. Furthermore, we have simulated the classification tasks with the GFO-based SQC, and have compared the results with those using the Adam-based SQC and QSVM. The simulation results have shown that the GFO-based SQC has obtained a higher accuracy and a shorter optimizing time than the Adam-based SQC and QSVM. Additionally, the GFO-based SQC has a good performance in noisy environments.

For further work, we will focus on how to increase the expression power of quantum circuits where parameters can be optimized by this GFO-based SQC for other complex classification tasks.

Acknowledgements

Project supported by the National Natural Science Foundation of China (Grant No. 62375140) and Postgraduate Research & Practice Innovation Program of Jiangsu Province (Grant No. KYCX19 0900).

References

[1] Biamonte J, Wittek P, Pancotti N, Rebentrost P, Wiebe N and Lloyd S 2017 *Nature* **549** 195

[2] Benedetti M, Lloyd E, Sack S and Fiorentini M 2019 *Quantum Sci. Technol.* **4** 043001
[3] Teresa S L, Juan R R and David Z 2022 *Phys. Rev. A* **105** 042432
[4] Naoko K M and Kei M 2021 *Phys. Rev. A* **104** 062411
[5] Wang X M, Zhang A Q, Xu P and Zhao S M 2021 *Chin. Phys. B* **30** 030307
[6] Hou Y Y, Li J, Chen X B and Tian Y 2022 *Chin. Phys. B* **31** 030304
[7] Wan K H, Dahlsten O, Kristj´ansson H, Gardner R and Kim M S 2017 *npj Quantum Inform.* **3** 36
[8] Torrontegui E and Garcia-Ripoll J J 2019 *Europhys. Lett.* **125** 30004
[9] Killoran N, Bromley T R, Arrazola J M, Schuld M, Quesada N and Lloyd S 2019 *Phys. Rev. Research* **1** 033063
[10] Mari A, Bromley T R, Izaac J, Schuld M and Killoran N 2020 *Quantum* **4** 340
[11] Schuld M and Killoran N 2019 *Phys. Rev. Lett.* **122** 040504
[12] Gilyén A, Arunachalam S and Wiebe N 2019 *Proceedings of the 2019 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)* 1425
[13] Campos E, Nasrallah A and Biamonte J 2021 *Phys. Rev. A* **103** 032607
[14] Schuld M, Bocharov A, Svore K M and Wiebe N 2020 *Phys. Rev. A* **101** 032308
[15] Adhikary S, Dangwal S and Bhowmik D 2020 *Quantum Inf. Process.* **19** 89
[16] Chalumuri A, Kune R and Manoj B S 2021 *Quantum Inf. Process.* **20** 119
[17] Chen S Y C, Huang C M, Hsing C W and Kao Y J 2021 *Mach. Learn-Sci. Techn.* **2** 045021
[18] Bhatia A S, Saggi M K, Kumar A and Jain S 2019 *arXiv:1905.01426*
[19] Adhikary S 2020 *arXiv:2006.13302*
[20] Zhang A Q, He X Y and Zhao S M 2022 *arXiv:2203.04097*
[21] Adri´an P S, Alba C L, Elies G F and José I L 2020 *Quantum* **4** 226
[22] Holmes Z, Sharma K, Cerezo M and Coles P J 2022 *PRX Quantum* **3** 010313
[23] Skolik A, McClean J R, Mohseni M, Smagt P V D and Leib M 2021 *Quant. Mach. Intell.* **3** 5
[24] Iannelli G, Jansen K 2021 *arXiv:2112.00426*
[25] Ostaszewski M, Grant E and Benedetti M 2021 *Quantum* **5** 391
[26] Pesah A, Cerezo M, Wang S, Volkoff T, Sornborger A T and Coles P J 2021 *Phys. Rev. X* **11** 041011
[27] Bergholm V, Izaac J, Schuld M, *et al.* 2018 *arXiv:1811.04968*
[28] Comelli P, Ferragina P, Granieri M N and Stabile F 1995 *IEEE Transactions on Vehicular Technology* **44** 790
[29] Nielsen M A and Chuang I 2002 *American Journal of Physics* **70** 558