Predicting properties of quantum systems by regression

Andrey Kardashin[1,*], **Yerassyl Balkybek**[1], **Konstantin Antipin**[1,2]
[1] *Skolkovo Institute of Science and Technology, Moscow, Russia*
[2] *Lomonosov Moscow State University, Moscow, Russia*

Introduction

Quantum machine learning (QML) is among the most promising applications of quantum computers [1, 2]. In QML, one commonly considers data sets consisting of labeled classical data points mapped to quantum states. Depending on the task, one may want to learn how to classify these data points if the labels are discrete [3, 4], or solve a regression problem for predicting the labels when they are continuous [5, 6]. In this work, applying the notion of variational quantum circuits [7], we develop a data-agnostic method for solving regression tasks for labeled data represented by quantum states.

Problem statement

Suppose we are given the following training set:

$$\mathcal{T} = \left\{\rho_{\alpha_j}, \alpha_j\right\}_{j=1}^{T}, \tag{1}$$

where ρ_{α_j} are labeled data points and $\alpha_j \in \mathbb{R}$ are their corresponding labels. We consider the data points to be quantum states described by density operators, i.e., $\rho_\alpha \geqslant 0$ and $\operatorname{Tr}\rho_\alpha = 1$. Hereinafter, we assume that ρ_α describes a state of n qubits. Our goal is to use the given training set $\mathcal{T}$ for learning how to estimate the parameter α for an unseen datum $\rho_\alpha \notin \mathcal{T}$.

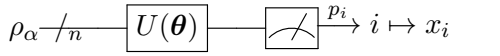
Essentially, we consider a regression problem with a peculiarity that the data points are represented by quantum states. Our aim is therefore to devise a method for predicting the labels assuming no knowledge about the connection between the states ρ_α with their labels α. Among the instances of such connection are: (i) α quantifies the entanglement of ρ_α, (ii) ρ_α is an output state of a parametrized channel $\Phi_\alpha[\rho]$ acting on some fixed input ρ, and (iii) $\rho_\alpha = |\psi_\alpha\rangle\langle\psi_\alpha|$ is the ground state of a parametrized Hamiltonian H_α.

Methods

We propose to solve the stated regression problem as follows. Given a labeled state ρ_α, we obtain the estimation $\hat{\alpha}$ of the label α as the expected value

$$\hat{\alpha}(\rho_\alpha, \boldsymbol{x}, \boldsymbol{\theta}) = \operatorname{Tr} H(\boldsymbol{x}, \boldsymbol{\theta})\, \rho_\alpha, \tag{2}$$

where the Hermitian operator H is parametrized by $\boldsymbol{x}, \boldsymbol{\theta} \subset \mathbb{R}$. We represent this observable as a spectral decomposition $H(\boldsymbol{x}, \boldsymbol{\theta}) = \sum_i x_i \Pi_i(\boldsymbol{\theta})$, where $\boldsymbol{x} = \{x_i\}_i$ are the eigenvalues, and the eigenprojectors $\Pi_i(\boldsymbol{\theta}) = U^\dagger(\boldsymbol{\theta})\,|i\rangle\langle i|\,U(\boldsymbol{\theta})$ are the projectors onto the ith state of the computational basis transformed by a variational circuit $U(\boldsymbol{\theta})$. Schematically, the estimation can be represented as follows:



That is, given an n-qubit labeled state ρ_α, we transform it by a parametrized unitary $U(\boldsymbol{\theta})$, measure the resultant state $\rho_\alpha(\boldsymbol{\theta}) \equiv U(\boldsymbol{\theta})\rho_\alpha U^\dagger(\boldsymbol{\theta})$ in the computational basis, with probability $p_i = \langle i|\rho_\alpha(\boldsymbol{\theta})|i\rangle$ get the outcome i associated with x_i, which gives the estimation (2) in the form $\hat{\alpha} = \sum_i x_i p_i$.

To find optimal parameters $\boldsymbol{x}^*$ and $\boldsymbol{\theta}^*$, we propose to solve the following minimization problem:

$$\boldsymbol{x}^*, \boldsymbol{\theta}^* = \arg\min_{\boldsymbol{x},\boldsymbol{\theta}} \Big(w_{\text{ls}} f_{\text{ls}}(\boldsymbol{x}, \boldsymbol{\theta}) + w_{\text{var}} f_{\text{var}}(\boldsymbol{x}, \boldsymbol{\theta}) \Big), \tag{3}$$

*andrey.kardashin@skoltech.ru

where

$$f_{\text{ls}}(\boldsymbol{x},\boldsymbol{\theta}) = \sum_{j=1}^{T} \Big(\alpha_j - \hat{\alpha}_j\big(\rho_\alpha, \boldsymbol{x}, \boldsymbol{\theta}\big)\Big)^2, \quad f_{\text{var}}(\boldsymbol{x},\boldsymbol{\theta}) = \sum_{j=1}^{T} \Delta^2_{\rho_{\alpha_j}} H(\boldsymbol{x},\boldsymbol{\theta}),$$

with $w_{\text{ls}}, w_{\text{var}} > 0$ being some weights. Here, f_{ls} is essentially the sum of the squared differences between the given labels α and our estimations $\hat{\alpha} = \langle H\rangle_{\rho_\alpha} \equiv \operatorname{Tr} H\rho_\alpha$, while f_{var} is the sum of variances $\Delta^2_{\rho_\alpha} H \equiv \langle H^2\rangle_{\rho_\alpha} - \langle H\rangle^2_{\rho_\alpha}$.

So, we seek to train an observable H which gives accurate estimations $\hat{\alpha}$ with presumably low variances. The estimation accuracy can be characterized by the mean squared error $\Delta^2\hat{\alpha} \equiv \big\langle(\alpha-\hat{\alpha})^2\big\rangle_{\hat{\alpha}}$, for which one can write [8]

$$\Delta^2\hat{\alpha} = \frac{\Delta^2_{\rho_\alpha} H}{\mu\big|\partial_\alpha\langle H\rangle_{\rho_\alpha}\big|^2} \geqslant \frac{1}{\mu I_c(\Pi,\rho_\alpha)} \geqslant \frac{1}{\mu I_q(\rho_\alpha)}. \tag{4}$$

where μ is the number of measurements. In this expression, the equality is known as the error propagation formula. Additionally, the first and the second inequalities are, respectively, the classical and the quantum Cramer-Rao bounds (CRB), where $I_c(\Pi,\rho_\alpha)$ is the classical Fisher information (FI), which is a function of both the state ρ_α and measurements $\Pi = \{\Pi_i\}_i$, and $I_q(\rho_\alpha)$ is the quantum FI, which depends solely on the state ρ_α.

Results

In this section, we report the results of the numerical application of the proposed method. To represent the parametrized unitary $U(\boldsymbol{\theta})$, we used a two-layered hardware-efficient ansatz [9]. As the weights in (3), if not stated otherwise, we set $w_{\text{ls}} = 1$ and $w_{\text{var}} = 10^{-4}$.

First, we demonstrate the performance of our method in predicting the label α of a labeled state $\rho_\alpha = |\psi_\alpha\rangle\langle\psi_\alpha|$ being the ground state of the transverse field Ising Hamiltonian

$$H_\alpha = -\sum_{i=1}^{n}\big(\sigma_z^i\sigma_z^{i+1} + \alpha\sigma_x^i\big) \tag{5}$$

of $n = 8$ qubits and with the periodic boundary conditions $\sigma_z^{n+1} \equiv \sigma_z^1$. In other words, given a collection of the form (1) of the ground states $|\psi_\alpha\rangle$ of (5), we want to solve (3) to find an observable H the expectation $\langle H\rangle_{\psi_\alpha}$ of which gives an estimation $\hat{\alpha}$ of the field α. In Figure 1, we show the performance for the observables trained with different weights w_{var} in (3) and setting $w_{\text{ls}} = 1$. As might be expected, the greater is the weight w_{var}, the less accurate predictions we get. However, with a greater weight w_{var}, we can almost saturate the quantum CRB in (4).

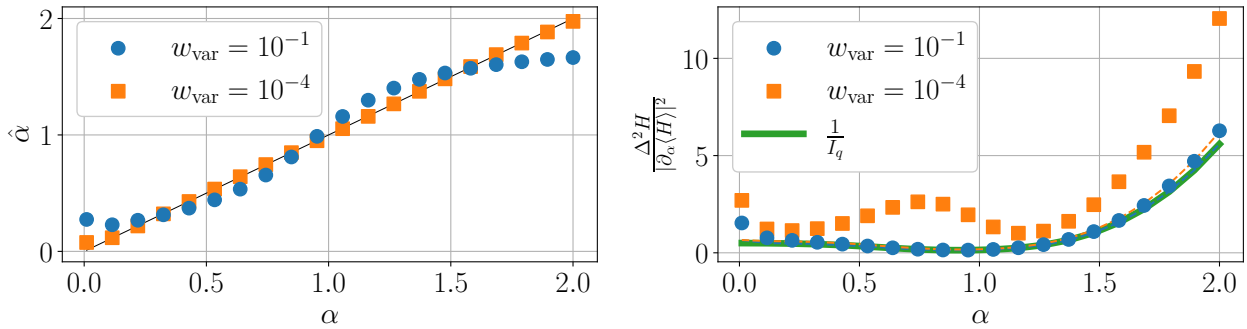


Figure 1: Left: predicted $\hat{\alpha}$ vs. true α transverse field of the 8-qubit Ising Hamiltonian (5). Right: error propagation and CRB (4) vs. α. The observable H is trained on a set $\mathcal{T} = \big\{|\psi_j\rangle, \alpha_j\big\}_{j=1}^{20}$ with $|\psi_j\rangle$ being the ground states of (5) with randomly generated fields α_j. In the right panel, the dashed lines of the corresponding colors indicate the achieved classical CRB, while the solid green line stands for the quantum CRB.

As a second test case, we train our model to predict the label α of the states of the form

$$\rho_\alpha = e^{-i\frac{\alpha}{2}\sigma_z} |+\rangle\langle+| e^{i\frac{\alpha}{2}\sigma_z} \tag{6}$$

For this particular case, there is no observable H which gives $\operatorname{Tr} H\rho_\alpha = \alpha$, evaluating therefore, instead, to some function $f(\alpha)$. However, we can simultaneously process c copies of the state ρ_α, which introduces non-linearity [10, 11] and may improve the performance of QML for some tasks [12, 4]. Therefore, our estimator (2) becomes $\hat{f}(\boldsymbol{x}, \boldsymbol{\theta}, c) = \operatorname{Tr} H(\boldsymbol{x}, \boldsymbol{\theta})\rho_\alpha^{\otimes c}$. In Figure 2, we show the results for the observables trained to predict the rotation angle α in (6) with different number of simultaneously processed copies c. As can be seen, with $c = 2$ copies, we obtain very accurate predictions for α. However, in contrast to $c = 1$, we do not saturate the CRB (4).

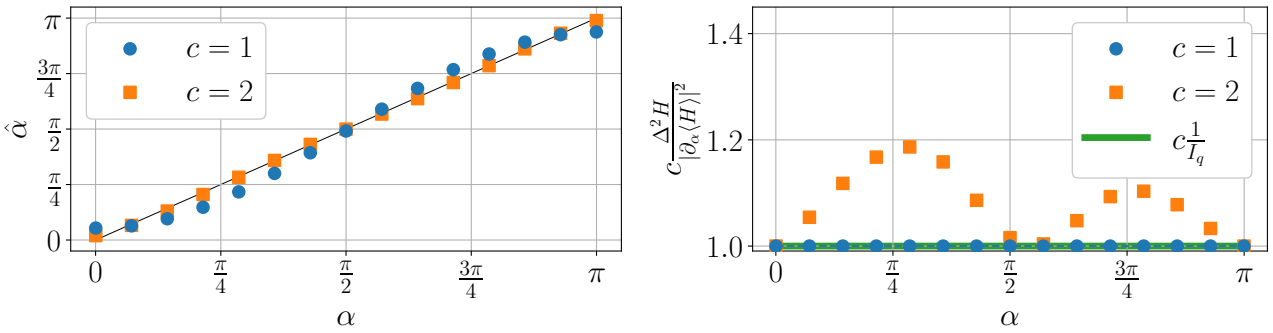


Figure 2: Left: predicted $\hat{\alpha}$ vs. true α rotation angle of the state (6). Right: error propagation and CRB (4). The observables H are trained on a set $\mathcal{T}_c = \{\rho_j^{\otimes c}, \alpha_j\}_{j=1}^{10}$ with randomly generated α_j and $c = 1, 2$ simultaneously processed copies.

Finally, we apply our method for learning to predict the entanglement of two-qubit random mixed states. As a measure of entanglement, we chose the negativity $N(\rho_{AB}) = \left\|\rho_{AB}^{T_B}\right\|_1 - 1$, where $\|\cdot\|_1$ is the trace norm, and $\rho_{AB}^{T_B} \equiv (\mathbb{1} \otimes T)[\rho]$ is a state ρ_{AB} of two subsystems A and B transposed with respect to the subsystem B [13, 14]. This time, we allow our model to process $c = 4$ copies of the labeled states. The results of the prediction of the negativity with the trained observable are shown in Figure 3. As can be seen, with our method, one is able to predict the entanglement of two-qubit states with good accuracy. Additionally, the performance is dependent on the purity $P(\rho) = \operatorname{Tr}\rho^2$ of a given state ρ: the greater is the purity, the more accurate is the estimation of the negativity $N(\rho)$, lower is the variance.

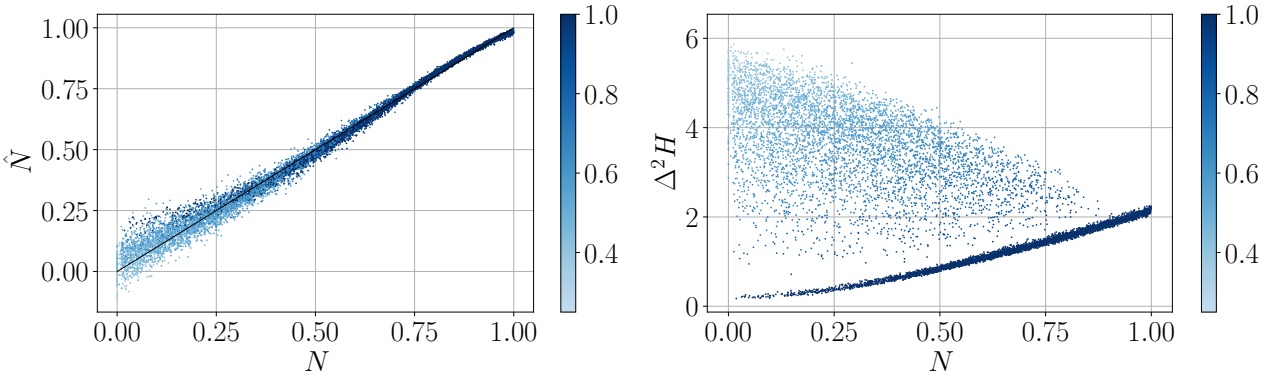


Figure 3: Predicted negativity $\hat{N}$ of 10^4 random mixed states (left) and variance of the trained observable H (right) vs. the true negativity N. The color of points indicates the purity of the corresponding states. The model is trained on a set $\mathcal{T} = \{\rho_j^{\otimes 4}, N_j\}_{j=1}^{1000}$, where the states ρ_j are generated such that their negativities are distributed approximately evenly on $[0, 1]$.

References

[1] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd. Quantum machine learning. *Nature*, 549(7671):195–202, 2017.

[2] M. Schuld and F. Petruccione. *Machine Learning with Quantum Computers*. Springer, 2021.

[3] Daniel K Park, Carsten Blank, and Francesco Petruccione. The theory of the quantum kernel-based binary classifier. *Physics Letters A*, 384(21):126422, 2020.

[4] Leonardo Banchi, Jason Pereira, and Stefano Pirandola. Generalization in quantum machine learning: A quantum information standpoint. *PRX Quantum*, 2(4):040321, 2021.

[5] Maria Schuld, Ilya Sinayskiy, and Francesco Petruccione. Prediction by linear regression on a quantum computer. *Physical Review A*, 94(2):022342, 2016.

[6] Pranath Reddy and Aranya B Bhattacherjee. A hybrid quantum regression model for the prediction of molecular atomization energies. *Machine Learning: Science and Technology*, 2(2):025019, 2021.

[7] Marcello Benedetti, Erika Lloyd, Stefan Sack, and Mattia Fiorentini. Parameterized quantum circuits as machine learning models. *Quantum Science and Technology*, 4(4):043001, 2019.

[8] Jasminder S Sidhu and Pieter Kok. Geometric perspective on quantum parameter estimation. *AVS Quantum Science*, 2(1), 2020.

[9] A. Kandala, A. Mezzacapo, K. Temme, M. Takita, M. Brink, J.M. Chow, and J.M. Gambetta. Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets. *Nature*, 549:242, Sep 2017.

[10] Yadong Wu, Juan Yao, Pengfei Zhang, and Hui Zhai. Expressivity of quantum neural networks. *Physical Review Research*, 3(3):L032049, 2021.

[11] Zoë Holmes, Nolan J Coble, Andrew T Sornborger, and Yiğit Subaşı. Nonlinear transformations in quantum computation. *Physical Review Research*, 5(1):013105, 2023.

[12] M. Schuld, A. Bocharov, K. Svore, and N. Wiebe. Circuit-centric quantum classifiers. *Physical Review A*, 101(3):032308, 2020.

[13] K. Życzkowski, P. Horodecki, A. Sanpera, and M. Lewenstein. Volume of the set of separable states. *Physical Review A*, 58:883, 1998.

[14] G. Vidal and R. F. Werner. Computable measure of entanglement. *Physical Review A*, 65:032314, 2002.