

Machine learning of chaotic characteristics in classical nonlinear dynamics using variational quantum circuit

Sheng-Chen Bai(白生辰) and Shi-Ju Ran(冉仕举)
Citation: Chin. Phys. B, 2026, 35 (**2**): 020303. DOI: 10.1088/1674-1056/adeb5b
http://cpb.iphy.ac.cn; http://iopscience.iop.org/cpb

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

Preparation of digital-encoded and analog-encoded quantum states corresponding to matrix operations

Kaitian Gao(高凯天), Youlong Yang(杨有龙), and Zhenye Du(杜振叶)
Chin. Phys. B, 2026, 35 (**1**): 010202. DOI: 10.1088/1674-1056/ade3ad

Quantum algorithm for marginal Fisher analysis

Jing Li(李静), Yanqi Song(宋燕琪), Sujuan Qin(秦素娟), Wenmin Li(李文敏), and Fei Gao(高飞)
Chin. Phys. B, 2025, 34 (**12**): 120302. DOI: 10.1088/1674-1056/addeb8

Domain adaptation method inspired by quantum convolutional neural network

Chunhui Wu(武春辉), Junhao Pei(裴骏豪), Yihua Wu(吴逸华), Anqi Zhang(张安琪), and Shengmei Zhao(赵生妹)
Chin. Phys. B, 2025, 34 (**7**): 070302. DOI: 10.1088/1674-1056/adc7ed

A backdoor attack against quantum neural networks with limited information

Chen-Yi Huang(黄晨猗) and Shi-Bin Zhang(张仕斌)
Chin. Phys. B, 2023, 32 (**10**): 100306. DOI: 10.1088/1674-1056/acd8ab

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

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

Machine learning of chaotic characteristics in classical nonlinear dynamics using variational quantum circuit

Sheng-Chen Bai(白生辰) and Shi-Ju Ran(冉仕举)†

Center for Quantum Physics and Intelligent Sciences, Department of Physics, Capital Normal University, Beijing 100048, *China*

(Received 15 April 2025; revised manuscript received 12 June 2025; accepted manuscript online 3 July 2025)

Replicating the chaotic characteristics inherent in nonlinear dynamical systems via machine learning (ML) is a key challenge in this rapidly advancing interdisciplinary field. In this work, we explore the potential of variational quantum circuits (VQC) for learning the stochastic properties of classical nonlinear dynamical systems. Specifically, we focus on the one- and two-dimensional logistic maps, which, while simple, remain under-explored in the context of learning dynamical characteristics. Our findings reveal that, even for such simple dynamical systems, accurately replicating long-term characteristics is hindered by a pronounced sensitivity to overfitting. While increasing the parameter complexity of the ML model typically enhances short-term prediction accuracy, it also leads to a degradation in the model's ability to replicate long-term characteristics, primarily due to the detrimental effects of overfitting on generalization power. By comparing the VQC with two widely recognized classical ML techniques, which are long short-term memory (LSTM) networks for time-series processing and reservoir computing, we demonstrate that VQC outperforms these methods in terms of replicating long-term characteristics. Our results suggest that for the ML of dynamics, it is demanded to develop more compact and efficient models (such as VQC) rather than more complicated and large-scale ones.

Keywords: variational quantum circuit, machine learning, chaos

PACS: 03.67.Ac, 07.05.Mh, 05.45.–a

DOI: 10.1088/1674-1056/adeb5b **CSTR:** 32038.14.CPB.adeb5b

1. Introduction

To what extent can a machine learning (ML) model effectively learn from the dynamical data, such as trajectories, of a classical chaotic system? A defining feature of nonlinear systems is chaos, which refers to the deterministic yet seemingly random behavior of dynamic systems.[1,2] Chaos is ubiquitous in a variety of real-world systems that exhibit strong nonlinearity, including population dynamics[3] and atmospheric phenomena.[4] Chaotic systems are characterized by several unique properties, such as extreme sensitivity to initial conditions and perturbations.[1,2] These characteristics render long-term time series prediction (TSP) exceedingly challenging, as even small uncertainties can lead to significantly divergent outcomes over time.

Machine learning (ML) has proven to be a powerful approach for TSP. Notably, neural networks, such as the well-established long short-term memory (LSTM) and transformer models,[5–7] have demonstrated exceptional performance in a variety of TSP-related tasks, including weather forecasting.[8,9] ML-based approaches for TSP in chaotic systems involve two primary issues: (i) short-term prediction of the system's states,[10–12] and (ii) long-term replication of the system's statistical or ergodic dynamical behaviors.[13–17]

Various methods have been proposed targeting at the second issue. One example is the reservoir computing with neural networks. The idea is to freeze most of the variational parameters and only train a small subset (e.g., those of the output neural layer) in order to suppress overfitting and improve computational efficiency.[18–20]

Another challenge in the ML of nonlinear dynamics is the fact that the stochastic properties often vary with the dynamical parameter(s). A common approach is to train separate models for each distinct value of the dynamical parameter(s),[21–23] a strategy we refer to as specific learning [see Fig. 1(a)]. More recently, some studies have proposed using a single ML model to universally capture the system's characteristics across a non-trivial range of dynamical parameters, which we here dub as universal learning.[24–26] We hypothesize that universal learning may allow the model to better learn the underlying dynamics, rather than simply learning the properties associated with specific values of the dynamical parameters.

In this work, we propose the universal learning of nonlinear dynamical systems using variational quantum circuit (VQC),[27–29] which is trained to predict the state one-step-ahead from M latest historic states and show excellent ability on replicating the long-term chaotic characteristics. Universal leaning is efficiently realized by introducing a preprocessing quantum feature map that is tuned by the dynamical parameter(s). We particularly consider two simplest nonlinear systems, namely, one- and two-dimensional (2D) logistic maps.[3,30,31] Our results surprisingly show that even for

†Corresponding author. E-mail: sjran@cnu.edu.cn

© 2026 Chinese Physical Society and IOP Publishing Ltd. All rights, including for text and data mining, AI training, and similar technologies, are reserved.

such simple nonlinear systems, the replication of their chaotic statistics is challenging, with high sensitivity to the generalization power of the ML model.

We further suggest the divergence in the trends of accuracies on short-term state prediction and long-term characteristics replication as a signature on the lack of generalization in the ML of nonlinear dynamics. We compare VQC with LSTM[5,6] and echo state network (ESN)[26] [Note: ESN serves as a canonical model in reservoir computing] by simulating the bifurcation diagram and Lyapunov exponents. VQC well suppresses the divergence of the two trends, and exhibits remarkable accuracy on replicating the long-term statistics. We speculate the main reason for the advantages of VQC on learning nonlinear dynamics is its block-wise (or tensor-wise) structure. Such a structure is generally known as tensor network (TN),[32–36] which has demonstrated not only remarkable efficiency in addressing various problems in quantum many-body physics,[37–39] but also significant success in tackling ML problems.[40–44] Typical application cases include quantum neural networks[45,46] and matrix product states,[36,47] which have further advanced the development of TNs in the interdisciplinary field of quantum computing and artificial intelligence.[48–50] Notably, its applications to the ML of nonlinear dynamics are for the moment sparse.[51,52]

2. Basics on the machine learning for time series prediction

A dynamical system can be generally written as a map $x_{t+1} = f(x_t)$, with x_t called the state of system at the time t. Taking the one-dimensional (1D) logistic chaotic map as an example, which is a fundamental nonlinear dynamical system [see the top-left panel of Fig. 1(a)], with wide applications in understanding chaotic phenomena and generating random numbers.[3] The map is written as

$$f(x_t;\mu) = \mu x_t(1-x_t), \tag{1}$$

with $0 \le \mu \le 4$ the dynamical parameter.

An ML model is also essentially a map. Considering the prediction of the state one-step-ahead from the latest M historic states (called a trajectory), the ML map can be formally written as

$$\tilde{y} = F(x_1, x_2, \ldots, x_M; \boldsymbol{W}), \tag{2}$$

with $\boldsymbol{W}$ denoting all the variational parameters of the ML model. One should not confuse the parameters of the dynamical system (say the dynamical parameter μ) with the variational parameters of the ML model (say $\boldsymbol{W}$). The time series generated by the dynamical system f is used to train the ML model, i.e., to update $\boldsymbol{W}$. Accordingly, ML samples for training are trajectories, and the labels are the ground truth of the states one-step-ahead. Note a state in a sample is also called a feature in the language of ML.

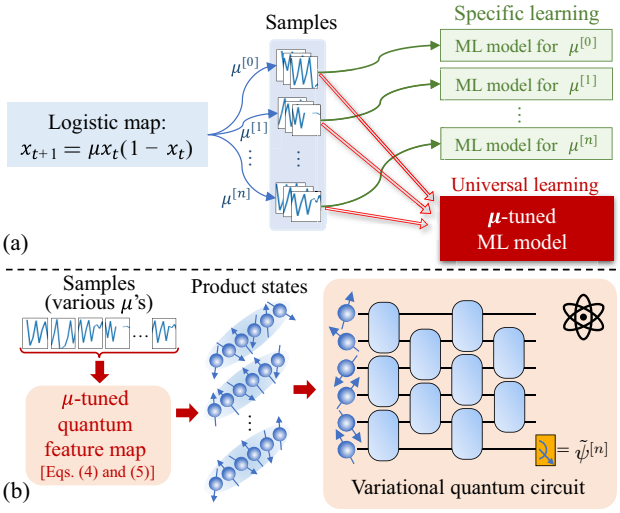


Fig. 1. (a) The illustration of specific learning (independently training one ML model for each value of the dynamical parameter μ) and universal learning (training a single ML model for various values of μ). (b) The main processes of the ML using VQC, where the μ-tuned quantum feature map [see Eqs. (4) and (5)] is introduced to realized universal learning.

An ML model can be trained by minimizing the prediction error on the so-called training samples. Here, we take μ as 50 discretized values in the range of $2 < \mu \le 4$. For each value of μ, we generate 2000 samples as the training set, where each sample is obtained by iteratively implementing the dynamical map on a random initial state for $(M-1)$ times. The variational parameters are optimized by minimizing the root mean-square error (RMSE)

$$L = \sqrt{\frac{1}{N}\sum_n \left(\tilde{y}^{[n]} - y^{[n]}\right)^2}, \tag{3}$$

where $\tilde{y}^{[n]}$ denotes the prediction by the ML model for the n-th sample, $y^{[n]}$ denotes the ground truth (the label of this sample), and the summation over n goes through the whole training set. The variational parameters are updated using an improved gradient-descent method AdamW.[53] AdamW modifies the update rule and adapts the learning rate (gradient step) by decoupling the weight-decay term from the gradient update, thereby suppressing over-fitting and improving generalization performance. In gradient descent, the variational parameters $\boldsymbol{W}$ are updated as $\boldsymbol{W} \to \boldsymbol{W} - \eta \frac{\partial L}{\partial \boldsymbol{W}}$, where η is a small positive constant known as the gradient step or learning rate.

3. Universal machine learning of nonlinear dynamics by variational quantum circuit

To replicate the stochastic characteristics of dynamical system with different values of μ by a VQC, we propose a μ-tuned trainable quantum feature map (QFM). Our idea is to

map each feature (which is a scalar) to vector by a μ-tuned trainable map, say $x_t^{[n]} \to \boldsymbol{v}^{[n,t]} = (v_1^{[n,t]}, v_2^{[n,t]}, \ldots, v_d^{[n,t]})$, with $x_t^{[n]}$ the t-th feature of the n-th sample and $d = \dim(\boldsymbol{v}^{[n,t]})$ a preset dimension. In this way, a sample is mapped to a set of vectors. Note that both LSTM and VQC can take a set of vectors as input. In the terms of quantum physics, QFM maps one feature a d-level quantum state.

To enhance the learning ability, the QFM can depend on not just μ but also some variational parameters that will be optimized in the training stage. Here, we introduce a $(d \times d \times d)$-dimensional tensor $\boldsymbol{T}$ as the variational parameters of QFM and define it as

$$v_k^{[n,t]} = \sum_{i,j} \xi_i(x_t^{[n]}; \theta)\xi_j(\mu^{[n]}; \theta)T_{ijk}. \tag{4}$$

We will update $\boldsymbol{T}$ when training the ML model. The vectors $\boldsymbol{\xi}(x_t^{[n]}; \theta)$ and $\boldsymbol{\xi}(\mu^{[n]}; \theta)$, which are both d-dimensional, are obtained by the following map that transforms a scalar to a normalized vector.[40] For a given feature (say a), the j-th elements of the resulting vector $\boldsymbol{\xi}(a; \theta)$ satisfies

$$\xi_j(a; \theta) = \sqrt{\binom{d-1}{j-1}} \cos\left(\frac{\theta\pi}{2}a\right)^{d-j} \sin\left(\frac{\theta\pi}{2}a\right)^{j-1}. \tag{5}$$

The parameter θ will also be optimized (independently on μ and samples) in the training stage. Note since the training samples are obtained by taking the dynamical parameter μ to be different values, one should take $\mu^{[n]}$ in Eq. (4) to be the corresponding value of μ for the n-th sample.

The QFM can be defined flexibly. It can also be generalized to the dynamical system that contains multiple variables and dynamical parameters. We can always use Eq. (5) and map these variables (and dynamical parameters) to multiple d-dimensional vectors.

After implementing QFM, a sample is mapped to M vectors (recall that M is the length of a sample), which will be subsequently input to the VQC. To proceed, one may construct a product state as $\prod_{\otimes m=1}^{M} \boldsymbol{v}^{[n,t]}$ as the initial state of VQC. By acting the VQC on such a state following the standard evolution process,[28,46] the prediction can be calculated with the final state. See more details about the QFM and the evolution by VQC in the supplemental material.[54]

4. Numerical results and discussion

Figure 2 demonstrates the ability of LSTM and VQC on the prediction of the states with different values of μ. In panel (a), we show the bifurcation diagrams obtained by VQC and LSTM, which are consistent with the one given by the logistic map (see the inset). Each data point in the diagrams is obtained by iterating the map (the logistic map or those of the ML models) for more than 200 times (which is sufficiently large), starting from a random state that does not appear in the training set. We take 500 such random states for each value of μ to generate the testing set. Since these samples are not used to train the ML model, the testing set reveal the so-called generalization ability, which refers to the power of the ML model on dealing with the unlearned samples.

The dense of the data points is indicated by the darkness of the colors. The similarity between two bifurcation diagrams can be characterized by the peak signal-to-noise ratios (PSNR) r_{P}, which is widely used in the field of computer vision to characterize the similarity between two images.[55] We have $r_{\mathrm{P}} = 43.58$ and 43.40 by comparing the ground-truth diagram by logistic map with the ones by LSTM and VQC, respectively. Such a consistency between the bifurcation diagrams does not require to accurately predict the states, but requires the validity of accurately replicating the distribution of the states after long-term evolutions.

In Fig. 2(b), we show the relative error $\varepsilon_{\mathrm{R}} = \frac{1}{N}\sum_{n=1}^{N}\left|(\tilde{y}^{[n]} - y^{[n]})/y^{[n]}\right|$ by VQC versus the discrete time t for different values of μ. Note the ML models are always trained to predict the state one-step-ahead from M historic states, and thus the training process just tries to minimize the error on the short-time predictions. The summation in ε_{R} is over the testing test. Different values of μ are taken so that the system is in the converged ($\mu = 2.2$), bifurcation ($\mu = 3.2$), quadrifurcation ($\mu = 3.4$), and chaotic ($\mu = 3.92$) regions. The values of Lyapunov exponent (LE)[56,57] are given in the legend, which is calculated as

$$\lambda = \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \ln\left|\frac{\mathrm{d}f(x;\mu)}{\mathrm{d}x}\right|_{x=x_t}. \tag{6}$$

In our numerical simulations, we take $T > 200$, which is sufficiently large.

For the state prediction in the converged and bifurcation regions (with negative LE), VQC can well predict the system's state even after a long-time evolution, with the relative error $\varepsilon_{\mathrm{R}} \sim O(10^{-3})$. When the system is in a chaotic region (with positive LE), previous work suggests that the state prediction is valid in a short duration characterized usually by the Lyapunov time $T_{\mathrm{LE}} \equiv \lambda^{-1}$.[58] Our results reveal an exponential scaling law, where ε_{R} grows exponentially with t before it converges to $O(1)$ as

$$\varepsilon_{\mathrm{R}} \sim \mathrm{e}^{\eta t}. \tag{7}$$

The table in Fig. 2(b) shows the exponential index for $3.57 < \mu < 4$ to be about $\eta \approx 0.4 + O(10^{-1})$. In such a range of μ, the logistic map exhibits complex dynamical behavior characterized by the coexistence of periodic orbits and aperiodic trajectories that manifest chaotic dynamics.[59]

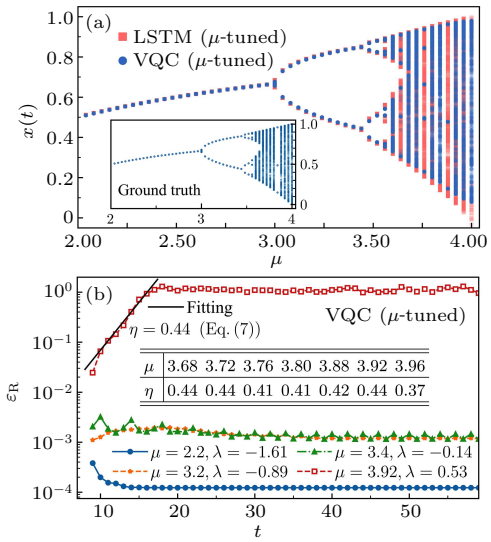


Fig. 2. (a) The bifurcation diagrams obtained by VQC, LSTM, and logistic map (see the inset). The dense of the data points is indicated by the darkness. (b) The semi-logarithmic plot of the relative error ε_R versus time t in different regions of the bifurcation diagram. We take $\mu = 2.2, 3.2, 3.4$, and 3.92, which give three negative and one positive values for the Lyapunov exponent. The exponential growth of ε_R for $\mu = 3.92$ is fitted by Eq. (7) with the exponential index $\eta = 0.44$ (see the black solid line). The table in panel (b) gives η for different values of μ in the chaotic region.

The above results demonstrate the ability of the ML models on dynamical characteristics replication rather than state prediction. This is consistent with the previous works, where ML models possess certain ability on reproducing the bifurcation diagrams.[60] Below, we further characterize quantitatively the ability on replicating the characteristics by comparing the LE, as shown in Fig. 3(a). The LE of the ML models is obtained by replacing the differential term in Eq. (6) by $\mathrm{d}F(x_{t-M+1},\ldots,x;\boldsymbol{W})/\mathrm{d}x$ at $x = x_t$.

The LE's with varying μ given by the LSTM and VQC with the μ-tuned pre-processing (green and blue lines with symbols) are consistent with those given by the logistic map (the ground truth shown by red solid line). As a comparison, much worse consistence is reached by the standard LSTM without the μ-tuned pre-processing (black dash–dot line). This is a demonstration of the improvement brought by our μ-tuned pre-processing. The two horizontal stripes indicate whether the μ-tuned LSTM and VQC give the LE with correct (green) or incorrect (red) sign. Generally, the accuracy on replicating the LE's sign is high. The incorrect signs appeal in the chaotic region.

Similar results are obtained for the 2D logistic map [see Fig. 3(b)], which represents a more intricate nonlinear system with richer dynamical behaviors by incorporating a linear coupling term.[30,31] The two variables $[x_t, x'_t]$ are mapped as

$$\begin{aligned} x_{t+1} &= 4\mu_1 x_t(1-x_t) + \beta x'_t, \\ x'_{t+1} &= 4\mu_2 x'_t(1-x'_t) + \beta x_t. \end{aligned} \tag{8}$$

We take $\mu_1 = \mu_2 = \mu$ and fix $\beta = 0.1$ for simplicity. Two LE's are defined. Obviously, replicating the LE's of the 2D logistic map is much more challenging than the 1D case. Consequently, the LSTM without the μ-tuned pre-processing becomes almost invalid. With the μ-tuned pre-processing, VQC achieves higher accuracy than LSTM.

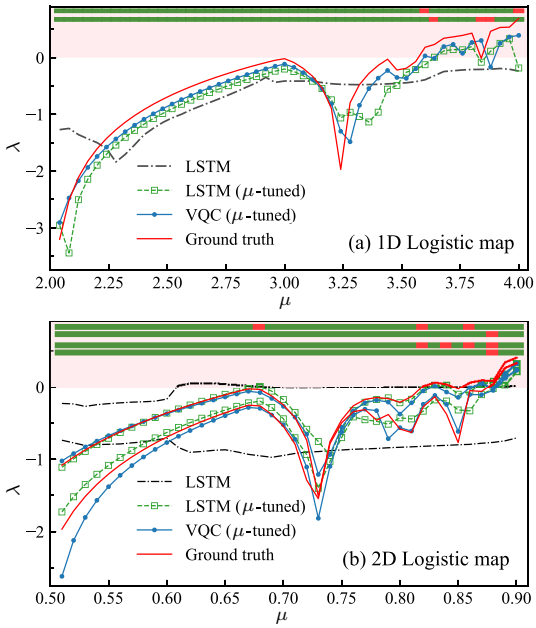


Fig. 3. The LE's of the (a) 1D and (b) 2D logistic maps and that of the corresponding ML models (LSTM and VQC). Each point is the average of five independent simulations. At the top of panel (a), the two horizontal stripes show whether the LSTM and VQC correctly (green) or incorrectly (red) give the sign of LE. In panel (b), the first two stripes at the top show the accuracy on giving the sign of the two LE's by LSTM, and the last two stripes give that by VQC.

Normally, a valid way to improve the performance of an ML model is to increase its complexity. However, such a way possibly fails in our case. For the dynamical system containing just a handful of hyper-parameters (say one parameter μ in the 1D logistic map), learning its dynamical characteristics is a big challenge. The key factor on improving the performance of an ML model on replicating the long-term characteristics should not be its parameter complexity. First of all, the model cannot be too simple, or otherwise necessary priori information on the dynamics (such as its mathematical form) is required to ensure the accuracy. In other words, proper complexity is needed so that the ML model would possess certain generalization ability to universally replicate the characteristics of dynamical systems without any priori knowledge on their mathematical forms. However, if we simply increase the model complexity, over-fitting emerges instead of any guaranteed improvement on the accuracy.

Our results support the above expectation. Firstly, increasing the parameter complexity can generally reduce the error on the short-term state prediction. See the RMSE L against the model complexity for the 1D and 2D logistic maps using

LSTM, VQC, and ESN in Fig. 4. Secondly, the error on replicating long-term characteristics $L_{\rm LE}$ does not decrease with the model complexity. For the LSTM, the results show obvious opposition on the trends of L and $L_{\rm LE}$ while increasing the hidden dimension $D_{\rm h}$. For ESN, it achieves the best short-term accuracy among the three (with $L \sim O(10^{-2})$–$O(10^{-4})$ that is slightly higher than those from the other two models) but worst long-term accuracy (with $L_{\rm LE} \sim O(1)$). For VQC, its short-term accuracy is comparable to LSTM (with $L \sim O(10^{-3})$), and remarkably, its long-term accuracy is stably maintained (which is the highest among the three models) while increasing the number of circuit layers.

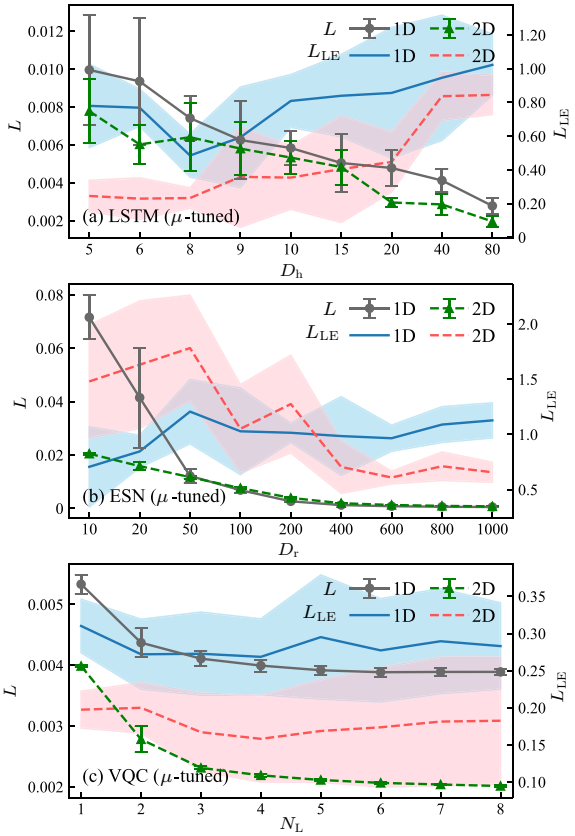


Fig. 4. The errors of predicting the state one-step-ahead and those of replicating LE for the 1D and 2D logistic maps. These two errors are characterized by L [Eq. (3)] and $L_{\rm LE}$, respectively. In panel (a), we show the L and $L_{\rm LE}$ obtained by LSTM versus its hidden dimension $D_{\rm h}$. In panel (b) we show those by ESN, whose structure is adopted from Ref. [26], versus its reservoir size $D_{\rm r}$. In panel (c), we show those by VQC versus its number of layers $N_{\rm L}$. All methods take similar μ-tuned strategy. Each data point is the average on five independent simulations, with the standard deviations demonstrated by the error bars and colored shadows.

Our results suggest the advantages of VQC on learning the long-term dynamical characteristics. We speculate that such advantages should come from its compact mathematical structure, which have been investigated in some previous works in the ML-related fields, such as image generation[41,61] and model compression.[44,62] On the other hand, our work suggests that even for the simple logistic maps, learning the long-term dynamical characteristics demands the ML model to possess high generalization power. This poses a challenge on the ML of more sophisticated dynamical systems, where VQC would be a promising candidate.

5. Summary

This work reveals the ability of variational quantum circuit (VQC) for the machine learning (ML) of long-term dynamical characteristics. We show that an ML model, which is trained to predict the state one-step-ahead from certain historic trajectories, possesses certain ability on replicating the long-term dynamical characteristics. However, such an ability is strongly hindered by the over-fitting issue, even for the simplest logistic maps. Specifically, increasing the model complexity generally improves the accuracy on the short-term state prediction but harms the accuracy of the long-term characteristics replication. By comparing with two widely-recognized ML models, VQC exhibits advantageous ability on replicating the long-term characteristics, possibly thanks to its compact tensor-network form that mitigates the over-fitting problem. Our work puts forward the strong remands for the ML of dynamics on developing more compact and efficient models (such as VQC) rather than more complicated and large-scale models.

In the future, we will investigate the applicability of VQC in high-dimensional chaotic dynamics, such as Lorenz equations, fluid dynamics models, or higher-dimensional spatiotemporal chaos models.[63] Additionally, we plan to theoretically analyze VQC's generalization capabilities, for instance, by employing quantum measures[64] to evaluate its learning capacity for chaotic dynamics. Another promising direction is integrating hybrid quantum-classical optimization strategies to enhance VQC's training efficiency in high-dimensional settings. These studies will help uncover the advantages and limitations of quantum machine learning in complex nonlinear dynamics and provide theoretical guidance for practical applications.

Acknowledgements

Sheng-Chen Bai is grateful to Qing Lv, Peng-Fei Zhou, Yong Qing, Zhang-Xu Chen, Guo-Dong Cheng, Ke Li, Rui Hong, Ying Lu, Yi-Cheng Tang, and Yu-Jia An for helpful discussions. Project supported in part by Beijing Natural Science Foundation (Grant No. 1232025), Peng Huanwu Visiting Professor Program, and Academy for Multidisciplinary Studies, Capital Normal University.

References

[1] Kantz H and Schreiber T 2004 *Nonlinear time series analysis*, Vol. 7 (Cambridge University Press)

[2] Strogatz S H 2018 *Nonlinear dynamics and chaos with student solutions manual: With applications to physics, biology, chemistry, and engineering* (CRC Press)
[3] May R M 1976 *Nature* **261** 459
[4] Lorenz E N 1963 *Journal of Atmospheric Sciences* **20** 130
[5] Hochreiter S and Schmidhuber J 1997 *Neural Computation* **9** 1735
[6] Yu Y, Si X, Hu C and Zhang J 2019 *Neural Computation* **31** 1235
[7] Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez A N, Kaiser L U and Polosukhin I 2017 *Advances in Neural Information Processing Systems*, Vol. 30, eds. Guyon I, Luxburg U V, Bengio S, Wallach H, Fergus R, Vishwanathan S and Garnett R (Curran Associates, Inc.)
[8] Bi K, Xie L, Zhang H, Chen X, Gu X and Tian Q 2023 *Nature* **619** 533
[9] Chen L, Zhong X, Zhang F, Cheng Y, Xu Y, Qi Y and Li H 2023 *npj Climate and Atmospheric Science* **6** 190
[10] Pathak J, Hunt B, Girvan M, Lu Z and Ott E 2018 *Phys. Rev. Lett.* **120** 024102
[11] Herzog S, Wörgötter F and Parlitz U 2019 *Chaos* **29** 123116
[12] Borra F, Vulpiani A and Cencini M 2020 *Phys. Rev. E* **102** 052203
[13] Pathak J, Lu Z, Hunt B R, Girvan M and Ott E 2017 *Chaos* **27** 121102
[14] Lu Z, Hunt B R and Ott E 2018 *Chaos* **28** 061104
[15] Antonik P, Gulina M, Pauwels J and Massar S 2018 *Phys. Rev. E* **98** 012215
[16] Fan H, Jiang J, Zhang C, Wang X and Lai Y C 2020 *Phys. Rev. Res.* **2** 012080
[17] Sun Y, Zhang L and Yao M 2023 *Chaos, Solitons & Fractals* **175** 113971
[18] Jaeger H 2001 *Bonn, Germany: German national research center for information technology gmd technical report* **148** 13
[19] Maass W, Natschläger T and Markram H 2002 *Neural Computation* **14** 2531
[20] Lukoševičius M 2012 *A Practical Guide to Applying Echo State Networks* (Berlin, Heidelberg: Springer Berlin Heidelberg) pp. 659–686
[21] Jaeger H and Haas H 2004 *Science* **304** 78
[22] Lukoševičius M and Jaeger H 2009 *Computer Science Review* **3** 127
[23] 2019 *Neural Networks* **115** 100
[24] Klos C, Kalle Kossio Y F, Goedeke S, Gilra A and Memmesheimer R M 2020 *Phys. Rev. Lett.* **125** 088103
[25] Kim J Z, Lu Z, Nozari E, Pappas G J and Bassett D S 2021 *Nature Machine Intelligence* **3** 316
[26] Kong L W, Fan H W, Grebogi C and Lai Y C 2021 *Phys. Rev. Res.* **3** 013090
[27] Peruzzo A, McClean J, Shadbolt P, Yung M H, Zhou X Q, Love P J, Aspuru-Guzik A and O'Brien J L 2014 *Nat. Commun.* **5** 4213
[28] Cerezo M, Arrasmith A, Babbush R, Benjamin S C, Endo S, Fujii K, McClean J R, Mitarai K, Yuan X, Cincio L and Coles P J 2021 *Nat. Rev. Phys.* **3** 625
[29] Zhou P F, Hong R and Ran S J 2021 *Phys. Rev. A* **104** 042601
[30] Yuan J M, Tung M, Feng D H and Narducci L M 1983 *Phys. Rev. A* **28** 1662
[31] Ferretti A and Rahman N 1988 *Chem. Phys.* **119** 275
[32] Cirac J I and Verstraete F 2009 *J. Phys. A: Math. Theor.* **42** 504004
[33] Orús R *The European Physical Journal B* 280
[34] Orús R 2019 *Nat. Rev. Phys.* **1** 538
[35] Ran S J, Tirrito E, Peng C, Chen X, Tagliacozzo L, Su G and Lewenstein M 2020 *Tensor Network Contractions: Methods and Applications to Quantum Many-Body Systems* (Springer, Cham)
[36] Cirac J I, Pérez-García D, Schuch N and Verstraete F 2021 *Rev. Mod. Phys.* **93** 045003
[37] Cirac J I, Poilblanc D, Schuch N and Verstraete F 2011 *Phys. Rev. B* **83** 245134
[38] Ge Y and Eisert J 2016 *New J. Phys.* **18** 083026
[39] Verstraete F and Cirac J I 2006 *Phys. Rev. B* **73** 094423
[40] Stoudenmire E and Schwab D J 2016 *Advances in Neural Information Processing Systems* 29, eds. Lee D D, Sugiyama M, Luxburg U V, Guyon I and Garnett R (Curran Associates, Inc.) pp. 4799–4807
[41] Han Z Y, Wang J, Fan H, Wang L and Zhang P 2018 *Phys. Rev. X* **8** 031012
[42] Liu D, Ran S J, Wittek P, Peng C, García R B, Su G and Lewenstein M 2019 *New J. Phys.* **21** 073059
[43] Sun Z Z, Peng C, Liu D, Ran S J and Su G 2020 *Phys. Rev. B* **101** 075135
[44] Qing Y, Li K, Zhou P F and Ran S J 2025 *Intelligent Computing* **4** 0123
[45] Jia Z A, Yi B, Zhai R, Wu Y C, Guo G C and Guo G P 2019 *Advanced Quantum Technologies* **2** 1800077
[46] Zhao R and Wang S 2021 Preprint 2109.01840
[47] Pérez-García D, Verstraete F, Wolf M M and Cirac J I 2007 *Quantum Inf. Comput.* **7** 401
[48] Biamonte J, Wittek P, Pancotti N, Rebentrost P, Wiebe N and Lloyd S 2017 *Nature* **549** 195
[49] Cerezo M, Verdon G, Huang H Y, Cincio L and Coles P J 2022 *Nature Computational Science* **2** 567
[50] Ran S J and Su G 2023 *Intelligent Computing* **2** 0061
[51] Goeßmann A, Götte M, Roth I, Sweke R, Kutyniok G and Eisert J 2020 *NeurIPS*2020-*Tensorworkshop, December*
[52] Wang Y, Tang S and Deng M 2022 *International Journal of Robust and Nonlinear Control* **32** 7304
[53] Loshchilov I and Hutter F 2019 Preprint 1711.05101
[54] See the supplemental material for further details, including information on dataset and pre-processing, automatically differentiable quantum circuit for time series prediction, Lyapunov exponent analysis, peak signal-to-noise ratios, additional results, and model hyperparameters. Refs. [29,55–57,65–68] are included.
[55] Kotevski Z and Mitrevski P 2010 *ICT Innovations* 2009 Ed. Davcev D and Gómez J M (Berlin, Heidelberg: Springer Berlin Heidelberg) pp. 357–366
[56] Parker T S and Chua L 2012 *Practical numerical algorithms for chaotic systems* (Springer Science & Business Media)
[57] Moon F C 2008 *Chaotic and fractal dynamics: introduction for applied scientists and engineers* (John Wiley & Sons)
[58] Bezruchko B and Smirnov D 2010 *Extracting Knowledge From Time Series: An Introduction to Nonlinear Empirical Modeling* (Berlin Heidelberg: Springer)
[59] Andrecut M 1998 *Int. J. Mod. Phys. B* **12** 921
[60] Kong L W, Weng Y, Glaz B, Haile M and Lai Y C 2023 *Chaos* **33** 033111
[61] Cheng S, Wang L, Xiang T and Zhang P 2019 *Phys. Rev. B* **99** 155131
[62] Gao Z F, Cheng S, He R Q, Xie Z Y, Zhao H H, Lu Z Y and Xiang T 2020 *Phys. Rev. Res.* **2** 023300
[63] Motter A E and Campbell D K 2013 *Physics Today* **66** 27
[64] Adesso G, Bromley T R and Cianciaruso M 2016 *J. Phys. A: Math. Theor.* **49** 473001
[65] Gill S S, Kumar A, Singh H, Singh M, Kaur K, Usman M and Buyya R 2022 *Software: Practice and Experience* **52** 66
[66] Geist K, Parlitz U and Lauterborn W 1990 *Prog. Theor. Phys.* **83** 875
[67] McDonald E J and Higham D J 2001 *ETNA–Electronic Transactions on Numerical Analysis* **12** 234
[68] Beck M, Pöppel K, Spanring M, Auer A, Prudnikova O, Kopp M, Klambauer G, Brandstetter J and Hochreiter S 2024 Preprint 2405.04517