Matrix product state for quantum-inspired feature extraction and compressed sensing

Wen-Jun Li
School of Physical Sciences
University of Chinese Academy of Sciences
P. O. Box 4588, Beijing 100049, China
liwenjun18@mails.ucas.edu.cn

Zheng-Zhi Sun
School of Physical Sciences
University of Chinese Academy of Sciences
P. O. Box 4588, Beijing 100049, China
sunzhengzhi16@mails.ucas.edu.cn

Ya-Ru Wang
School of Physical Sciences
University of Chinese Academy of Sciences
P. O. Box 4588, Beijing 100049, China
wangyaru18@mails.ucas.edu.cn

Shi-Ju Ran*
Department of Physics
Capital Normal University
Beijing 100048, China
sjran@cnu.edu.cn

Gang Su†
Kavli Institute for Theoretical Sciences,
and CAS Center for Excellence in Topological Quantum Computation
University of Chinese Academy of Sciences
Beijing 100190, China
School of Physical Sciences
University of Chinese Academy of Sciences
P. O. Box 4588, Beijing 100049, China
gsu@ucas.ac.cn

Abstract

Improving the interpretability and efficiency of machine learning methods are challenging and important issues. The concept of entanglement, which is a quantity from quantum information science, may contribute to these issues. In this extended abstract, we introduce two tensor-network (TN) machine learning methods, which concerns feature extraction and compressed sensing, respectively, based on entanglement. For the former [12], the entanglement obtained from matrix product state (MPS) is used as a measure of the importance of the features in the real-life datasets. An entanglement-based feature extraction algorithm is proposed, and used to improve the efficiency of TN machine learning. In the latter [13], TN states are used to realize efficient compressed sampling by entanglement-guided projective measurements. This scheme can be applied in the future to compress and communicate the real-life data by entangled qubits.

1 Introduction

Classical machine learning models such as neural networks are sometimes called "black boxes", since it is hard to understand the logic behind them. Improving the interpretability of machine learning is becoming increasingly important [1–4]. As a mathematical model, tensor network (TN) has been

*Corresponding Author
†Corresponding Author

Second Workshop on Quantum Tensor Networks in Machine Learning, 35th Conference on Neural Information Processing Systems (NeurIPS 2021)

used to effectively represent a large class of quantum many-body states that satisfy the area law of entanglement entropy [5–7]. The powerful representation ability and high interpretability of TN play important roles in the quantum-inspired and quantum machine learning [8–10].

TN can be directly used for the classification tasks as a probabilistic model, where the entanglement is shown capable to characterize the importance of features. This method can be regarded as a TN version of sensitivity analysis [11]. The feature extraction algorithm based on entanglement in Ref [12] extracts the important features from the original dataset and improves the efficiency of TN machine learning. Numerical experiments show that the number of features in MNIST dataset can be reduced to $1/10$ with little harm to the accuracy.

The TN is also used for the compressed sensing, which can reduce the complexity of quantum communication for the real-life information [13]. In recent years, the number of controllable qubits has increased rapidly [14–16], while quantum measurements and communication are still expensive and extremely challenging for the real-life data. One possible way to improve the efficiency is to combine the compression and communication in a unified process. Towards this direction, tensor network compressed sensing (TNCS) was proposed [13] based on MPS (also called Born machine) and its entanglement properties. Specifically, the receiver recovers the whole information from a small piece of it by implementing an entanglement-guided projective measurement protocol on the MPS. The numerical experiments show on MNIST and Fashion-MNIST datasets that the TNCS is accurate and efficient. Thanks to the close connections between TN and quantum states, TNCS can be readily generalized to quantum encrypted communications based on multi-qubit entangled states.

2 Feature extraction based on entanglement

The first step of the MPS-based feature extraction is to map the images to many-qubit quantum states. For instance, the n-th pixel x_n of the image $\mathbf{x}$ is mapped to a one-qubit state given by a two-dimensional normalized vector

$$x_n \to |s(x_n)\rangle = \cos(x_n\pi/2)\,|0\rangle + \sin(x_n\pi/2)\,|1\rangle \tag{1}$$

where $|0\rangle$ and $|1\rangle$ represent the spin-up and spin-down states, respectively. An image that contains L pixels is mapped to a L-qubit product state $\mathbf{v} = \prod_{\otimes n=1}^{L} |s(x_n)\rangle$. An MPS is trained to map a given product state to the prediction of its classification, named as MPS classifier. As MPS represents essentially a 1D state, the 1D path to cover the 2D image by the MPS is flexible but critical. Ref. [12] uses two kinds of entanglement entropy to optimize the path. Single-site entanglement entropy (SEE) describes the amount of entanglement between a specific qubit and the rest, and bipartite entanglement entropy (BEE) describes that between the qubits on the two halves of the 1D path. Known from the quantum information theories, the qubits with small entanglement in MPS have less quantum correlations with the rest of system. In accordance, the numerical results given below show that the features whose corresponding qubits are largely entangled in the MPS classifier contribute more information to the classification results, and they should be retained. The entanglement-based feature extraction algorithm is proposed by optimizing the path based on the SEE to concentrate the important features in the middle and cutting the tails of the MPS according to the BEE to retain only $\widetilde{L}$ features with largest BEE.

The results from the "1-7" binary classifier (for images with labels "1" and "7") on MNIST are given as an example. Fig. 1 (a) shows the BEE and SEE of the MPS classifiers. The darkness of each dot indicates the strength of SEE, and the thickness of each bond indicates the BEE. An overlapped image of "1" and "7" emerges from the important features with dark dots connected with strong bonds. These dots give the features that will be extracted. Fig. 1 (b) demonstrates the same quantities, where the MPS is trained by the images after implementing the discrete-cosine transformation (DCT) [17–19]. In accordance with the experience in the computer vision that the lower-frequency parts are usually more important, the dark dots connected with strong bonds appear near the top-right corner. In Fig. 1 (c), the path is optimized by arranging the qubit with larger SEE closer to the middle of the path, and the features arranged at the non-entangled tails of the path are discarded according to BEE. This respects the quantum correlations indicated by the entanglement entropies, and leads to a better extraction.

For the ten-class classification on MNIST dataset, Fig. 2 shows the testing accuracy with different numbers of the extracted features $\tilde{L}$ using the feature extraction explained above. The feature

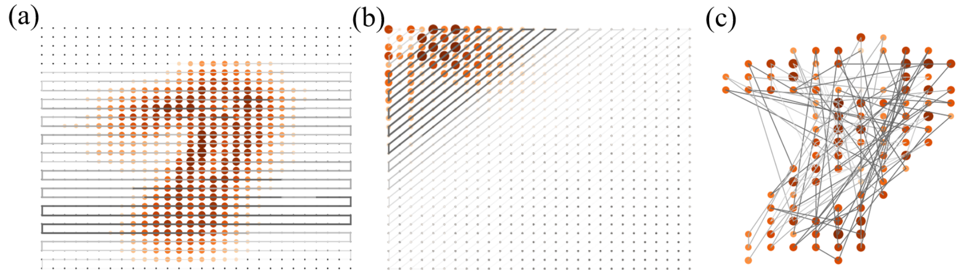

Figure 1: The entanglement properties of the ''1–7'' binary MPS classifier using different paths (which show how the MPS covers the 2D images). Each point corresponds to one feature. The size and darkness of nodes represent the SEE's strength on each site. The thickness of each bond represents the strength of the BEE. See more details in the main text. These figures are reused from Ref. [12].

extraction is tested on both the original data and frequency components after DCT. The results show that both DCT and path optimization can improve the accuracy with a fixed number of the retained features. In particular when both strategies are applied, the accuracy is 88.69% with only 40 features. Note these 40 features are simply selected from the original features. Even using DCT, the combination (linear transformation) of the features is independent to the samples. This means the importance of the original features are directly evaluated. It is different from some other feature extraction schemes such as principle component analysis (PCA) [20], where the combination (linear transformation) of the features depends on the samples. Thus, one cannot directly evaluate the importance of the pixels in images by PCA.

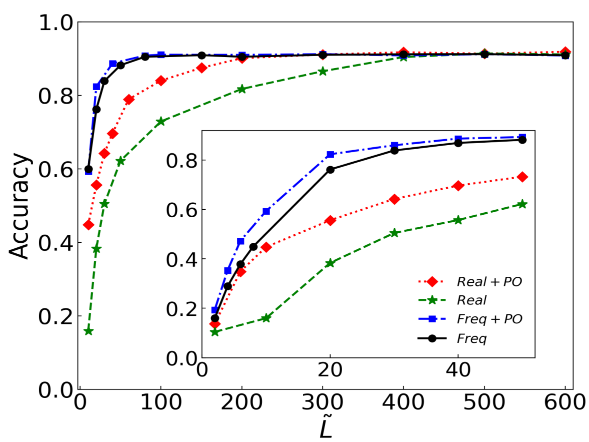

Figure 2: The classification accuracy on the test dataset versus the number of the retained extracted features $\widetilde{L}$. The path optimization (PO) on frequency space gives the best results. This figure is reused from Ref. [12].

3 Compressed sensing based on matrix product state and entanglement

As the number of usable qubits rises fast, it becomes increasingly important to explore algorithms with many-qubit states. Towards this aim, TNCS is proposed as a quantum-inspired compressed sensing scheme, and can be generalized to quantum set-ups for the encrypted communications of the real-life data.

The main idea of TNCS is to compress the information by implementing designed projective measurements on an trained MPS $|\psi\rangle$. Let us consider the following scenario. Alice wants to send an image of hand-written digit "3" to Bob, while she intends to only send a small number of pixels publicly

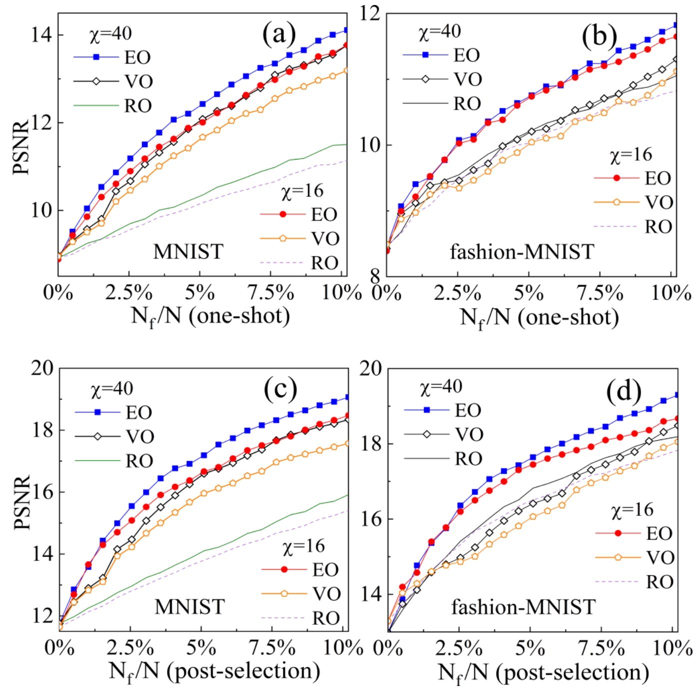

Figure 3: Average peak signal-to-noise ratio (PSNR) of the constructed images in the testing dataset of the handwriting digits "3" in MNIST and the dresses in fashion-MNIST. Choosing features with high entanglement (i.e., EO) and decompressing the data by post-selection gives the best result. These figures are reused from Ref. [13].

or through an unsafe channel. To make the communication safe, Alice provides an generative MPS $|\psi\rangle$ [21] to Bob. This MPS is trained to capture the probability distribution of the training set of many "3" images. Then, Bob can acquire a partially projected MPS by collapsing the pixels sent from Alice. The rest pixels that Bob wants to recover can be obtained from a separable state that has the minimum distance to the projected MPS. Bob can safely decode the unsent pixels from the projected state (the separable state) by measuring every qubit (one-shot) or estimating the dominant eigenstates of the one-qubit reduced density matrices (post-selection).

In Fig. 3, we present the numerical results of recovering images in MNIST and Fashion-MNIST using TNCS, where the performance is characterized by the average peak signal-to-noise ratio (PSNR). Different methods to encode and decode the information using TNCS with different virtual bond dimensions χ of the MPS are tested, where χ determines the complexity of the MPS. The numerical results on both datasets show that measuring the qubits from higher entangled to lower entangled qubits (namely entanglement ordering, EO in short) is better than other ways, such as the order according to the variance (VO) or random orders (RO). Besides, the information decoded by post-selections shows higher accuracy than that by one-shot measurements. With EO and post-selections, TNCS permits to compress and communicate real-life information more accurately and efficiently.

4 Summary

In this extended abstract, we introduce two quantum-inspired algorithms for feature extraction and compressed sensing. These two schemes achieve impressive performance in the numerical experiments on the real-life datasets. These works shed light on improving the interpretability of machine learning by incorporating with TN and entanglement. With sufficient controllable qubits and the techniques on state preparation in the future, TNCS can be readily applied for encrypted communications of real-life messages by preparing the required MPS's with entangled qubits. In the future, these two entanglement-based methods can also be extended to TN's with more sophisticated architecture and be verified with more complex datasets.

Acknowledgments and Disclosure of Funding

WJL, ZZS, YRW and GS are supported by in part by the NSFC (Grant No. 11834014), the National Key RD Program of China (Grant No. 2018FYA0305804), the Strategetic Priority Research Program of the Chinese Academy of Sciences (Grant No. XDB28000000), and Beijing Municipal Science and Technology Commission (Grant No. Z190011). SJR is supported by NSFC (Grant No. 12004266 and No. 11834014), Beijing Natural Science Foundation (No. 1192005 and No. Z180013), Foundation of Beijing Education Committees (No. KM202010028013), and the key research project of Academy for Multidisciplinary Studies, Capital Normal University.

References

[1] W. Huggins, P. Patil, B. Mitchell, et al. Towards Quantum Machine Learning with Tensor Networks. *Quantum Sci. Technol.*, **4**, 024001 (2019).

[2] X. D. Cai, D. Wu, Z. E. Su, et al. Entanglement-Based Machine Learning on a Quantum Computer. *Phys. Rev. Lett.*, **114**, 110504 (2015).

[3] S. Lloyd, M. Mohseni and P. Rebentrost. Quantum algorithms for supervised and unsupervised machine learning. arXiv:1307.0411 (2013).

[4] L. Lamata. Basic Protocols in Quantum Reinforcement Learning with Superconducting Circuits. *Sci. Rep.*, **7**, 1609 (2017).

[5] R. Orús. A Practical Introduction to Tensor Networks: Matrix Product States and Projected Entangled Pair States. *Annals of Physics*, **349**, (2014) pp:117-158.

[6] J. I. Cirac and F. Verstraete. Renormalization and Tensor Product States in Spin Chains and Lattices. *J. Phys. A: Math. Theor.*, **42**, 504004 (2009).

[7] S. J. Ran, E. Tirrito, C. Peng, et al. Tensor Network Contractions. (Springer International Publishing, 2020).

[8] J. Martyn, G. Vidal, C. Roberts et al. Entanglement and tensor networks for supervised image classification. arXiv: 2007.06082 (2020).

[9] V. Pestun and Y. Vlassopoulos. Tensor network language model. arXiv: 1710.10248 (2017).

[10] D. Liu, S. J. Ran, P. Wittek, et al. Machine Learning by Unitary Tensor Network of Hierarchical Tree Structure. *New J. Phys.*, **21**, 073059 (2019).

[11] B. Iooss and P. Lemaître. Uncertainty management in Simulation-Optimization of Complex Systems: Algorithms and Applications. (Springer US, 2015) pp:101-122.

[12] Y. H. Liu, W. J. Li, X. Zhang, et al. Entanglement-Based Feature Extraction by Tensor Network Machine Learning. *Front. Appl. Math. Stat.*, **7**, 716044 (2021).

[13] S. J. Ran, Z. Z. Sun, S. M. Fei, et al. Tensor network compressed sensing with unsupervised machine learning. *Phys. Rev. Research.*, **2**, 033293 (2020).

[14] C. Song, K. Xu, H. K. Li, et al. Generation of multicomponent atomic schrödinger cat states of up to 20 qubits. *Science*, **365**, (2019) pp:574-577.

[15] A. Omran, H. Levine, A. Keesling, et al. Generation and manipulation of schrödinger cat states in rydberg atom arrays. *Science*, **365**, (2019) pp:570-574.

[16] F. Arute, K. Arya, R. Babbush, et al. Quantum supremacy using a programmable superconducting processor. *Nature*, **574**, (2019) pp:505-510.

[17] N. Ahmed, T. Natarajan and K. R. Rao. Discrete Cosine Transform. *IEEE Transactions on Computers*, **C-23**, (1974) pp:90–93.

[18] V. Holub and J. Fridrich. Low-Complexity Features for JPEG Steganalysis Using Undecimated DCT. *IEEE Transactions on Information Forensic and Security*, **10**, (2015) pp:219–228.

[19] M. A. Saad, A. C. Bovik and C. Charrier. Blind Image Quality Assessment: A Natural Scene Statistics Approach in the DCT Domain. *IEEE Transactions on Image Processing*, **21**, (2012) pp:3339–3352.

[20] I. T. Jolliffe. Principal Component Analysis. (Springer New York, 1986) pp:129-155.

[21] Z. Y. Han, J. Wang, H. Fan, et al. Unsupervised Generative Modeling using Matrix Product States. *Phys. Rev. X*, **8**, 031012 (2018).