A study on the expressibility and learnability of quantum circuit learning

CHIH-CHIEH CHEN[1,a)] MASAYA WATABE[2] KODAI SHIBA[1,2] MASARU SOGABE[1]
KATSUYOSHI SAKAMOTO[2,3] TOMAH SOGABE[2,3,1,b)]

Abstract: Using quantum circuits for supervised machine learning is one potential way to harness quantum advantage for Noisy Intermediate-Scale Quantum hardware. Many implementations and algorithms have been proposed and studied for quantum circuit learning, but the learnability and generalization ability of quantum circuit ansatz is not well-understood yet. In this work, we study the relation between the circuit ansatz, the expressive power, and the PAC-learnability of quantum circuit learning. The model complexity and generalization ability are studied using a KL-divergence based measure, a VC-dimension upper bound, and various numerical simulation. Our result provides a way to understand the learnability of quantum circuit.

Keywords: Machine learning, quantum circuit

1. Introduction

The quest for artificial intelligence and the study of machine learning started with a quesiton asked by Alan Turing: "Can machines think?" [1]. Recent advance of machine learning [2] shows that classical computers can outperform human players in a computationally hard game GO [3], and recent demonstration of Noisy Intermediate-Scale Quantum (NISQ) [4] hardware shows that quantum information processors could outperform classical processors [5]. At this point, it is reasonable to ask a question: "Can quantum machines learn?"

We are interested in the learning capability of near-term quantum devices on the classical machine learning tasks [6], [7]. For NISQ hardware, applying variational quantum circuit [8], [9] for hybrid quantum-classical supervised machine learning leads to the quantum circuit learning (QCL) method [10]. There are recent efforts toward the understanding of expressitivity and model complexity of QCL [11], [12], [13], [14], [15], [16]. The learnability of machines can be understood with the Probably Approximately Correct (PAC) model proposed by Valiant [17]. If a learning machine has too much model complexity, it is possible to overfit the data and the generalization ability would be bad. For classiciation tasks, Vapnik–Chervonenkis (VC) theory [18] can be used to establish the generalization ability by using the VC dimension [19]. For regression tasks, the pseudo-dimension [20] and the fat-shattering dimension [21], [22] could be used.

1 Grid Inc., 107-0061 Tokyo, Japan
2 Engineering department, The University of Electro-Communications, 182-8585 Tokyo, Japan
3 i-PERC, The University of Electro-Communications, 182-8585 Tokyo, Japan
a) chen.chih.chieh@gridsolar.jp
b) sogabe@uec.ac.jp

Our recent study gives a VC dimesnion upper bound in terms of circuit width and circuit depth, and provides a recipe to control the model complexity of QCL [23], [24]. In this report, we present a summary and some further discussions and investigations related to the expressibility and learnability. Previous learnability results for quantum machine learning are based on fat-shattering dimension [25], pseudo-dimension [26], or quantum sampling complexity [27]. Another VC-dimension upper bound, which is different from our result, is recently proposed in Ref. [28].

2. Method

2.1 Quantum circuit learning

In supervised binary classification learning problem, there is an unknown target function $f : X \mapsto Y = \{-1, 1\}$, and we are given some training dataset $\{(\vec{x}_i, y_i = f(\vec{x}_i)) | \vec{x}_i \in X\}$ where $\vec{x}_i$ is drawn from some unknown distribution $\vec{x}_i \sim P(\vec{x}_i)$. (In the setting that the training data is noisy, the dataset is drawn from some unknown joint distribution $(\vec{x}_i, y_i) \sim P(\vec{x}, y)$.) The goal of learning is to obtain a hypothesis $h : X \mapsto Y$ such that the prediction error (out-of-sample error) $E_{out} = \mathbb{P}_{\vec{x}\sim P(\vec{x})}[h(\vec{x}) \neq f(\vec{x})]$ (or $E_{out} = \mathbb{P}_{(\vec{x},y)\sim P(\vec{x},y)}[h(\vec{x}) \neq y]$ for the noisy case) is small.

In QCL [10] we use quantum circuit as the hypothesis set. The number of qubits n is also called circuit width. The circuit starts with an input layer where the input data $\vec{x}$ is encoded by $|\psi_{in}(\vec{x})\rangle = \otimes_{i=0}^{n-1} U(\theta_{i,in}(\vec{x}))|0\rangle$. After the input layer, the trainable part $U_\theta = (\prod_{j=1}^{L} U_{loc}^{(j)}(\theta_j) U_{ent}) U_{loc}^{(0)}(\theta_j)$ consists of alternating layers of trainable single qubit rotations $U_{loc}^{(j)}(\theta_j) = \otimes_{i=0}^{n-1} U(\theta_{i,j})$ and fixed (not tranable) entangling layer. The local rotation could include any subset of $\{R_X, R_Y, R_Z\}$. The number of layers is a hyper-parameter, also known as circuit depth L. The measurement result is used to compute the Z expectation values of i-th qubit

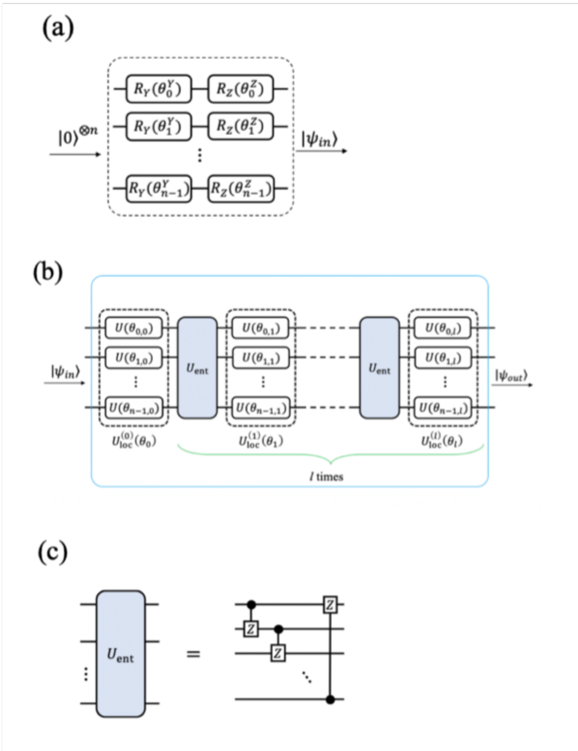

Fig. 1 The learning circuit used in this work. Figure is taken from Ref. [24].

$$\langle Z_i(\theta, \vec{x})\rangle \quad = \mathrm{Tr}(\rho_{in}(\vec{x}) U_\theta Z_i U_\theta^\dagger), \tag{1}$$

where $\rho_{in} = |\psi_{in}(\vec{x})\rangle\langle\psi_{in}(\vec{x})|$. The expectation value is then thresholded by some value t for the binary classification, leading to a hypothesis set $\{\mathrm{sgn}(\langle Z_i(\theta, \vec{x})\rangle - t)\}$. Multi-class classification is done by the usual softmax method. Since we are interested in NISQ applications, we use the Hardware Efficient Ansatz (HEA) where the entangler consists of 2-qubit quantum gate acting on nearest neighbor qubits for some lattice. The learning circuit studied in this work is depicted in Fig. 1, where the entangling lattice is an 1D ring with periodic boundary condition, and the entangling gate is the controlled-Z (CZ) gate.

2.2 Information theory: KL expressibility

A KL-divergence expressibility measure for quantum circuit is defined as [11]

$$D_{KL}(P_C(F)||P_{Haar}(F)) = \int_0^1 P_C(F)\log(\frac{P_C(F)}{P_{Haar}(F)})dF. \tag{2}$$

where $F = |\langle\psi_\theta|\psi_\phi\rangle|^2$ is the fidelity. The quantity can be numerically evaluated by random uniform sampling over the circuit parameters θ and ϕ. The circuit simulation is done by using Qiskit [29]. Notice that lower value of D_{KL} means higher model complexity. In general we are more interested in the relative value of D_{KL}. Hence we define re-scaled $D^*_{KL} = aD_{KL} + b$ for some real number a and b such that higher D^*_{KL} means higher expressibility.

2.3 Statistical learning theory: VC theory

We use the definition that the generalization error is $E_{out} - E_{in}$, where E_{out} is the out-of-sample error (prediction error) and $E_{in} = \frac{1}{N}\sum_{i=1}^N [\![h(\vec{x}_i) \neq f(\vec{x}_i)]\!]$ is the in-sample-error (training error). The VC generalization error bound is [18]

$$E_{out} - E_{in} \leq \sqrt{\frac{8}{N}\log(\frac{4m_H(2N)}{\delta})}, \tag{3}$$

with probability $\geq 1-\delta$. N is the sample size. δ is the confidence interval. The quantity $m_H(N)$ is a function which could be bounded by $m_H(N) \leq N^{d_{VC}} + 1$ for finite VC-dimension d_{VC}.[31] VC dimension is the maximum number of points that can be shattered by the hypothesis set. In general, d_{VC} could be infinite for an infinite hypothesis set. If d_{VC} is finite, then the generalization ability is guaranteed by the VC bound. There are two advantages of VC theory [30]: (1) VC bound is independent of the input distribution. (2) VC bound is non-asymptotic, and it can be applied when the size of training data set is small.

2.4 Tensor networks: Light cone limitation

Due to the locality and unitarity of 2-qubit entangling gate used in HEA, the light cone limitation in the tensor network can be applied [32]. The tensor contraction outside of the light cone gives identities, hence only the qubits covered by the light cone contribute to the hypothesis set. This limitation leads to the limitation to the degree of trignometric polynomials and hence limits the VC dimension as well as the KL-expressibility [23].

3. Results and Discussions

3.1 Theoretical results

Under the assumptions of using the encoding schemes in Ref. [10], we derive a VC dimension upper bound for QCL hypothesis set. The d-dimensional input vector is $\vec{x} \in [-1, 1]^d$, and we assume that d divides n. Each qubit encodes one input dimension with map $R_z(\arccos(x_i^2))R_y(\arcsin(x_i))$, so the input layer is $|\psi_{in}(\vec{x})\rangle = (\otimes_{i=0}^{d-1} R_z(\phi_{i,in})R_y(\theta_{i,in})|0\rangle)^{\otimes\frac{n}{d}}$ where $\phi_{i,in} = \arccos(x_i^2)$ and $\theta_{i,in} = \arcsin(x_i)$. Since the hypothesis set is a real trignometric polynomial of variables $\{\phi_{i,in}, \theta_{i,in}| i = 0, ..., d-1\}$, we obtain the upper bound

$$d_{VC} \leq (2\frac{n}{d} + 1)^{2d}. \tag{4}$$

For 1D periodic boundary condition HEA, we may apply the light cone limitation to improve the upper bound

$$d_{VC} \leq (2\min(\frac{n}{d}, \lfloor\frac{2L+1}{d}\rfloor + 1) + 1)^{2d}. \tag{5}$$

Notice that for fixed d, the bound grows polynomially with respect to depth L, and saturates at a critical value $\frac{n}{d} = \lfloor\frac{2L^*+1}{d}\rfloor + 1$. This is just the result for one special case HEA. For other HEA entangler topology, one could obtain a similar bound polynomial in L by counting the number of qubits covered by the corresponding light cone. Also notice that unlike classical neural networks (number of edges=$|E|$, number of vertices=$|V|$) where the VC dimension grows asymptotically $O(|E|\log(|E|))$ (for sign activation function) or $O(|V|^2|E|^2)$ (for sigmoid activation function) [33], [34], the VC dimension saturates for deep circuit in QCL. This difference reminds us that QCL and neural networks are two different learning models.

We may also obtain analytical results regarding the saturation of expressibility. The KL-expressibility for light cone limited Haar distribution can be evaluated to be

$$D_{KL}(P_{Harr}^{LightCone}(F)||P_{Haar}(F)) \tag{6}$$

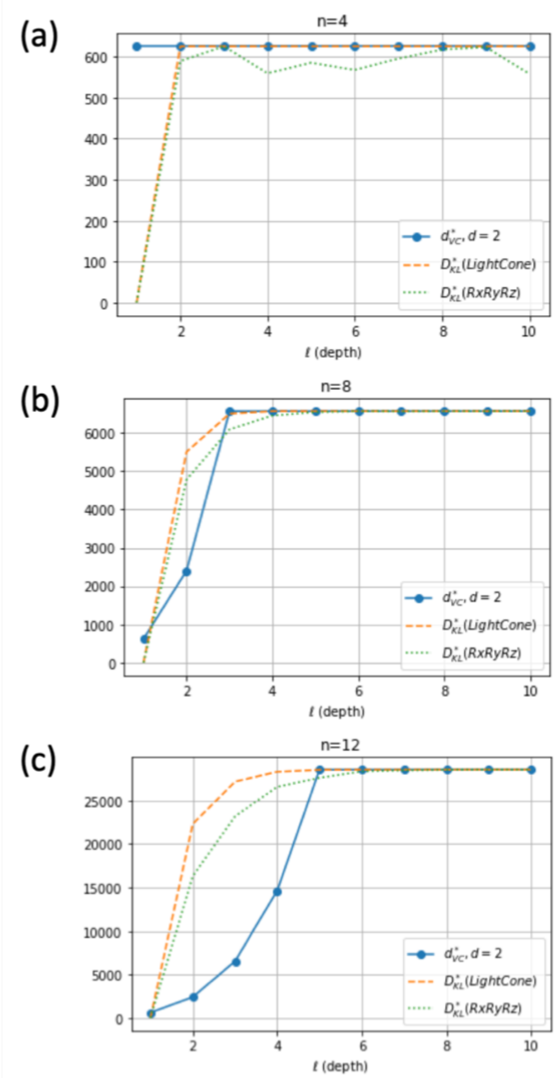

Fig. 2 Re-scaled KL-expressibility and VC dimension upper bounds plotted against circuit depth for different circuit widths. Blue solid line with circle marker is the VC dimension upper bound for feature space dimension $d = 2$. Green dotted line is the rescaled KL-expressibility for RxRyRz local unitaries ansatz. Orange dashed line is the light cone restricted KL expressibility in Eq. 6. (a) $n = 4$. (b) $n = 8$ (c) $n = 12$.

$$= \frac{M-1}{M-2}\log(\frac{M-1}{N-1}) - \frac{M-N}{M-1}, \tag{7}$$

where $N = 2^n$. M is the effective Hilbert space dimension covered by the light cone, and in the case of 1D periodic boundary condition HEA, it is $M = 2^{\min(2L+1,n)}$. Fig. 2 shows a comparison of the saturation behavior against circuit depth L for d^*_{VC}, D^*_{KL} with $R_xR_yR_z$ HEA sampling [29], and $D^*_{KL}(LightCone) \equiv D^*_{KL}(P^{LightCone}_{Harr}(F)\|P_{Haar}(F))$ with light cone limitation Eq. 6 . One can observe that the scaling of light-cone limited expressibility is similar to the expressibility obtained by random circuit sampling. The saturation behavior is also similar to that of the VC dimension upper bound.

3.2 Experimental results

Fig. 3 shows numerical simulation result for $n = 4$ QCL using a simulator and back-propagation algorithm built in-house [24]. The ansatz is CZ-HEA with $RyRz$ local rotations. The input data is Scikit-learn "make moon" data [35], and the input data noise

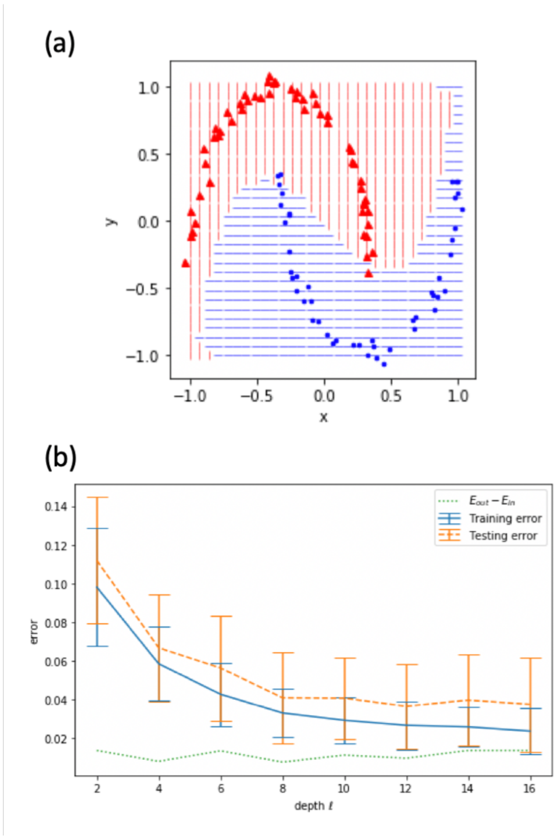

Fig. 3 (a) One example for the training data and prediction results for Scikit-learn "make moon" binary classification data. Blue and red color denotes two classes. Triangle and circle markers denote training data for two classes. Vertical and horizontal lines are the test results for two classes. (b) In-sample error, out-of-sample error, and generalization error are plotted against circuit depth for $n = 4$. Average of 50 independent runs. Error bar denotes one sigma fluctuations.

is 0.04. The size of training data set is 100 and the size of test data set is 100. We observe both the in-sample-error and out-of-sample error are decreasing for deeper circuit. We do not see significant sign of overfitting. This observation supports the theory of saturation of model complexity. We would point out that this saturation behavior can not be explained by a counting argument which is comparing the number of trainable parameters $2n(L+1)$ with the real degrees of freedom of unitary group 2^{2n}. This counting argument would give a critical depth $L^* = \frac{2^{2n-1}}{n} - 1 = 31$ for $n = 4$, which is much higher than the observed value. More experimental results for various parameter settings could be found in Ref. [23].

4. Conclusion

Experimental study for QCL shows that the generalization error does not significantly increase for increasing circuit depth in many cases. In this work, we provide theoretical explanation of the phenomena by using tools from statistical learning theory, information theory, and tensor networks. The light cone limitation causes the saturation of both VC-dimension upper bound and also KL-expressibility for deep variational quantum circuit. This observation provides both theoretical understanding of the learnability of QCL and also useful recipe to control the generalization ability of QCL. Some future efforts are required to further understand the learnability of QCL, including the scaling behavior of

VC-dimension lower bound, and other possible quantum circuit architectures.

References

[1] TURING, A. M.: I.—COMPUTING MACHINERY AND INTELLIGENCE, *Mind*, Vol. LIX, No. 236, pp. 433–460 (online), DOI: 10.1093/mind/LIX.236.433 (1950).

[2] Silver, D., Schrittwieser, J., Simonyan, K., Antonoglou, I., Huang, A., Guez, A., Hubert, T., Baker, L., Lai, M., Bolton, A., Chen, Y., Lillicrap, T., Hui, F., Sifre, L., van den Driessche, G., Graepel, T. and Hassabis, D.: Mastering the game of Go without human knowledge, *Nature*, Vol. 550, No. 7676, pp. 354–359 (online), DOI: 10.1038/nature24270 (2017).

[3] Lichtenstein, D. and Sipser, M.: GO Is Polynomial-Space Hard, *J. ACM*, Vol. 27, No. 2, p. 393–401 (online), DOI: 10.1145/322186.322201 (1980).

[4] Preskill, J.: Quantum Computing in the NISQ era and beyond, *Quantum*, Vol. 2, p. 79 (online), DOI: 10.22331/q-2018-08-06-79 (2018).

[5] Arute, F., Arya, K., Babbush, R., Bacon, D., Bardin, J. C., Barends, R., Biswas, R., Boixo, S., Brandao, F. G. S. L., Buell, D. A., Burkett, B., Chen, Y., Chen, Z., Chiaro, B., Collins, R., Courtney, W., Dunsworth, A., Farhi, E., Foxen, B., Fowler, A., Gidney, C., Giustina, M., Graff, R., Guerin, K., Habegger, S., Harrigan, M. P., Hartmann, M. J., Ho, A., Hoffmann, M., Huang, T., Humble, T. S., Isakov, S. V., Jeffrey, E., Jiang, Z., Kafri, D., Kechedzhi, K., Kelly, J., Klimov, P. V., Knysh, S., Korotkov, A., Kostritsa, F., Landhuis, D., Lindmark, M., Lucero, E., Lyakh, D., Mandrà, S., McClean, J. R., McEwen, M., Megrant, A., Mi, X., Michielsen, K., Mohseni, M., Mutus, J., Naaman, O., Neeley, M., Neill, C., Niu, M. Y., Ostby, E., Petukhov, A., Platt, J. C., Quintana, C., Rieffel, E. G., Roushan, P., Rubin, N. C., Sank, D., Satzinger, K. J., Smelyanskiy, V., Sung, K. J., Trevithick, M. D., Vainsencher, A., Villalonga, B., White, T., Yao, Z. J., Yeh, P., Zalcman, A., Neven, H. and Martinis, J. M.: Quantum supremacy using a programmable superconducting processor, *Nature*, Vol. 574, No. 7779, pp. 505–510 (online), DOI: 10.1038/s41586-019-1666-5 (2019).

[6] Havlíček, V., Córcoles, A. D., Temme, K., Harrow, A. W., Kandala, A., Chow, J. M. and Gambetta, J. M.: Supervised learning with quantum-enhanced feature spaces, *Nature*, Vol. 567, No. 7747, pp. 209–212 (online), DOI: 10.1038/s41586-019-0980-2 (2019).

[7] Bishop, C. M.: *Pattern Recognition and Machine Learning (Information Science and Statistics)*, Springer-Verlag, Berlin, Heidelberg (2006).

[8] Peruzzo, A., McClean, J., Shadbolt, P., Yung, M.-H., Zhou, X.-Q., Love, P. J., Aspuru-Guzik, A. and O'Brien, J. L.: A variational eigenvalue solver on a photonic quantum processor, *Nature Communications*, Vol. 5, p. 4213 (online), DOI: 10.1038/ncomms5213 (2014).

[9] McClean, J. R., Romero, J., Babbush, R. and Aspuru-Guzik, A.: The theory of variational hybrid quantum-classical algorithms, *New Journal of Physics*, Vol. 18, No. 2, p. 023023 (online), DOI: 10.1088/1367-2630/18/2/023023 (2016).

[10] Mitarai, K., Negoro, M., Kitagawa, M. and Fujii, K.: Quantum circuit learning, *Phys. Rev. A*, Vol. 98, p. 032309 (online), DOI: 10.1103/PhysRevA.98.032309 (2018).

[11] Sim, S., Johnson, P. D. and Aspuru-Guzik, A.: Expressibility and Entangling Capability of Parameterized Quantum Circuits for Hybrid Quantum-Classical Algorithms, *Advanced Quantum Technologies*, Vol. 2, No. 12, p. 1900070 (online), DOI: https://doi.org/10.1002/qute.201900070 (2019).

[12] Nakaji, K. and Yamamoto, N.: Expressibility of the alternating layered ansatz for quantum computation, *Quantum*, Vol. 5, p. 434 (online), DOI: 10.22331/q-2021-04-19-434 (2021).

[13] Schuld, M., Sweke, R. and Meyer, J. J.: Effect of data encoding on the expressive power of variational quantum-machine-learning models, *Phys. Rev. A*, Vol. 103, p. 032430 (online), DOI: 10.1103/PhysRevA.103.032430 (2021).

[14] Hubregtsen, T., Pichlmeier, J., Stecher, P. and Bertels, K.: Evaluation of Parameterized Quantum Circuits: on the relation between classification accuracy, expressibility and entangling capability, *arXiv e-prints*, p. arXiv:2003.09887 (2020).

[15] Goto, T., Hoan Tran, Q. and Nakajima, K.: Universal Approximation Property of Quantum Feature Map, *arXiv e-prints*, p. arXiv:2009.00298 (2020).

[16] Abbas, A., Sutter, D., Zoufal, C., Lucchi, A., Figalli, A. and Woerner, S.: The power of quantum neural networks, *arXiv e-prints*, p. arXiv:2011.00027 (2020).

[17] Valiant, L. G.: A Theory of the Learnable, *Commun. ACM*, Vol. 27, No. 11, p. 1134–1142 (online), DOI: 10.1145/1968.1972 (1984).

[18] Vapnik, V. N. and Chervonenkis, A. Y.: On the Uniform Convergence of Relative Frequencies of Events to Their Probabilities, *Theory of Probability & Its Applications*, Vol. 16, No. 2, pp. 264–280 (online), DOI: 10.1137/1116025 (1971).

[19] Blumer, A., Ehrenfeucht, A., Haussler, D. and Warmuth, M. K.: Learnability and the Vapnik-Chervonenkis Dimension, *J. ACM*, Vol. 36, No. 4, p. 929–965 (online), DOI: 10.1145/76359.76371 (1989).

[20] Pollard, D.: *Convergence of Stochastic Processes*, Springer (1984).

[21] Kearns, M. J. and Schapire, R. E.: Efficient distribution-free learning of probabilistic concepts, *Journal of Computer and System Sciences*, Vol. 48, No. 3, pp. 464–497 (online), DOI: https://doi.org/10.1016/S0022-0000(05)80062-5 (1994).

[22] Bartlett, P. L., Long, P. M. and Williamson, R. C.: Fat-Shattering and the Learnability of Real-Valued Functions, *Journal of Computer and System Sciences*, Vol. 52, No. 3, pp. 434–452 (online), DOI: https://doi.org/10.1006/jcss.1996.0033 (1996).

[23] Chen, C.-C., Watabe, M., Shiba, K., Sogabe, M., Sakamoto, K. and Sogabe, T.: On the expressibility and overfitting of quantum circuit learning, *To be published in ACM Transactions on Quantum Computing*.

[24] Watabe, M., Shiba, K., Sogabe, M., Sakamoto, K. and Sogabe, T.: Quantum Circuit Parameters Learning with Gradient Descent Using Backpropagation, *arXiv e-prints*, p. arXiv:1910.14266 (2019).

[25] Aaronson, S.: The learnability of quantum states, *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, Vol. 463, No. 2088, pp. 3089–3114 (online), DOI: 10.1098/rspa.2007.0113 (2007).

[26] Caro, M. C. and Datta, I.: Pseudo-dimension of quantum circuits, *Quantum Machine Intelligence*, Vol. 2, No. 2 (online), DOI: 10.1007/s42484-020-00027-5 (2020).

[27] Arunachalam, S. and de Wolf, R.: Optimal Quantum Sample Complexity of Learning Algorithms, *Journal of Machine Learning Research*, Vol. 19, No. 71, pp. 1–36 (online), available from ⟨http://jmlr.org/papers/v19/18-195.html⟩ (2018).

[28] Gyurik, C., van Vreumingen, D. and Dunjko, V.: Structural risk minimization for quantum linear classifiers, *arXiv e-prints*, p. arXiv:2105.05566 (2021).

[29] Abraham, H., AduOffei, Agarwal, R., Akhalwaya, I. Y., Aleksandrowicz, G., Alexander, T., Amy, M., Arbel, E., Arijit02, Asfaw, A., Avkhadiev, A., Azaustre, C., AzizNgoueya, Banerjee, A., Bansal, A., Barkoutsos, P., Barnawal, A., Barron, G., Barron, G. S., Bello, L., Ben-Haim, Y., Bevenius, D., Bhobe, A., Bishop, L. S., Blank, C., Bolos, S., Bosch, S., Brandon, Bravyi, S., Bryce-Fuller, Bucher, D., Burov, A., Cabrera, F., Calpin, P., Capelluto, L., Carballo, J., Carrascal, G., Chen, A., Chen, C.-F., Chen, E., Chen, J. C., Chen, R., Chow, J. M., Churchill, S., Claus, C., Clauss, C., Cocking, R., Correa, F., Cross, A. J., Cross, A. W., Cross, S., Cruz-Benito, J., Culver, C., Córcoles-Gonzales, A. D., Dague, S., Dandachi, T. E., Daniels, M., Dartiailh, M., DavideFrr, Davila, A. R., Dekusar, A., Ding, D., Doi, J., Drechsler, E., Drew, Dumitrescu, E., Dumon, K., Duran, I., EL-Safty, K., Eastman, E., Eberle, G., Eendebak, P., Egger, D., Everitt, M., Fernández, P. M., Ferrera, A. H., Fouilland, R., FranckChevallier, Frisch, A., Fuhrer, A., Fuller, B., GEORGE, M., Gacon, J., Gago, B. G., Gambella, C., Gambetta, J. M., Gammanpila, A., Garcia, L., Garg, T., Garion, S., Gilliam, A., Giridharan, A., Gomez-Mosquera, J., Gonzalo, de la Puente González, S., Gorzinski, J., Gould, I., Greenberg, D., Grinko, D., Guan, W., Gunnels, J. A., Haglund, M., Haide, I., Hamamura, I., Hamido, O. C., Harkins, F., Havlicek, V., Hellmers, J., Herok, Ł., Hillmich, S., Horii, H., Howington, C., Hu, S., Hu, W., Huang, J., Huisman, R., Imai, H., Imamichi, T., Ishizaki, K., Iten, R., Itoko, T., JamesSeaward, Javadi, A., Javadi-Abhari, A., Javed, W., Jessica, Jivrajani, M., Johns, K., Johnstun, S., Jonathan-Shoemaker, K, V., Kachmann, T., Kale, A., Kanazawa, N., Kang-Bae, Karazeev, A., Kassebaum, P., Kelso, J., King, S., Knabberjoe, Kobayashi, Y., Kovyrshin, A., Krishnakumar, R., Krishnan, V., Krsulich, K., Kumkar, P., Kus, G., LaRose, R., Lacal, E., Lambert, R., Lapeyre, J., Latone, J., Lawrence, S., Lee, C., Li, G., Liu, D., Liu, P., Maeng, Y., Majmudar, K., Malyshev, A., Manela, J., Marecek, J., Marques, M., Maslov, D., Mathews, D., Matsuo, A., McClure, D. T., McGarry, C., McKay, D., McPherson, D., Meesala, S., Metcalfe, T., Mevissen, M., Meyer, A., Mezzacapo, A., Midha, R., Minev, Z., Mitchell, A., Moll, N., Montanez, J., Monteiro, G., Mooring, M. D., Morales, R., Moran, N., Motta, M., MrF, Murali, P., Müggenburg, J., Nadlinger, D., Nakanishi, K., Nannicini, G., Nation, P., Navarro, E., Naveh, Y., Neagle, S. W., Neuweiler, P., Nicander, J., Niroula, P., Norlen, H., NuoWenLei, O'Riordan, L. J., Ogunbayo, O., Ollitrault, P., Otaolea, R., Oud, S., Padilha, D., Paik, H., Pal, S., Pang, Y., Pascuzzi, V. R., Perriello, S., Phan, A., Piro, F., Pistoia, M., Piveteau, C., Pocreau, P., Pozas-Kerstjens, A., Prokop, M., Prutyanov, V., Puzzuoli, D., Pérez, J., Quintiii, Rahman, R. I., Raja, A., Ramagiri, N., Rao, A., Raymond, R., Redondo, R. M.-C., Reuter, M., Rice, J., Riedemann, M., Rocca, M. L., Rodríguez, D. M., RohithKarur, Rossmannek, M., Ryu, M., SAPV, T., SamFerracin, Sandberg, M., Sandesara, H., Sapra, R., Sargsyan,

H., Sarkar, A., Sathaye, N., Schmitt, B., Schnabel, C., Schoenfeld, Z., Scholten, T. L., Schoute, E., Schwarm, J., Sertage, I. F., Setia, K., Shammah, N., Shi, Y., Silva, A., Simonetto, A., Singstock, N., Siraichi, Y., Sitdikov, I., Sivarajah, S., Sletfjerding, M. B., Smolin, J. A., Soeken, M., Sokolov, I. O., Sokolov, I., SooluThomas, Starfish, Steenken, D., Stypulkoski, M., Sun, S., Sung, K. J., Takahashi, H., Takawale, T., Tavernelli, I., Taylor, C., Taylour, P., Thomas, S., Tillet, M., Tod, M., Tomasik, M., de la Torre, E., Trabing, K., Treinish, M., TrishaPe, Tulsi, D., Turner, W., Vaknin, Y., Valcarce, C. R., Varchon, F., Vazquez, A. C., Villar, V., Vogt-Lee, D., Vuillot, C., Weaver, J., Weidenfeller, J., Wieczorek, R., Wildstrom, J. A., Winston, E., Woehr, J. J., Woerner, S., Woo, R., Wood, C. J., Wood, R., Wood, S., Wood, S., Wootton, J., Yeralin, D., Yonge-Mallo, D., Young, R., Yu, J., Zachow, C., Zdanski, L., Zhang, H., Zoufal, C., Zoufalc, a kapila, a matsuo, bcamorrison, brandhsn, nick bronn, brosand, chlorophyll zz, csseifms, dekel.meirom, dekelmeirom, dekool, dime10, drholmie, dtrenev, ehchen, elfrocampeador, faisaldebouni, fanizzamarco, gabrieleagl, gadial, galeinston, georgios ts, gruu, hhorii, hykavitha, jagunther, jliu45, jscott2, kanejess, klinvill, krutik2966, kurarrr, lerongil, ma5x, merav aharoni, michelle4654, ordmoj, sagar pahwa, rmoyard, saswati qiskit, scottkelso, sethmerkel, shaashwat, sternparky, strickroman, sumitpuri, tigerjack, toural, tsura crisaldo, vvilpas, welien, willhbang, yang.luh, yotamvakninibm and Čepulkovskis, M.: Qiskit: An Open-source Framework for Quantum Computing (2019).

[30] Vapnik, V. N.: *The Nature of Statistical Learning Theory*, Springer-Verlag, Berlin, Heidelberg (1995).

[31] Abu-Mostafa, Y. S., Magdon-Ismail, M. and Lin, H.-T.: *Learning From Data*, AMLBook (2012).

[32] Schollwöck, U.: The density-matrix renormalization group in the age of matrix product states, *Annals of Physics*, Vol. 326, No. 1, pp. 96–192 (online), DOI: https://doi.org/10.1016/j.aop.2010.09.012 (2011).

[33] Kearns, M. J. and Vazirani, U. V.: Computational Learning Theory, *SIGACT News*, Vol. 26, No. 1, p. 43–45 (online), DOI: 10.1145/203610.606411 (1995).

[34] Shalev-Shwartz, S. and Ben-David, S.: *Understanding Machine Learning: From Theory to Algorithms*, Cambridge University Press, USA (2014).

[35] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M. and Duchesnay, E.: Scikit-learn: Machine Learning in Python, *Journal of Machine Learning Research*, Vol. 12, pp. 2825–2830 (2011).