Variational Quantum Support Vector Machine With Inference Transfer

Siheon Park[1], Daniel K. Park[2], and June-Koo Kevin Rhee*[1,3]

[1]School of Electrical Engineering and KAIST Institute for IT Convergence, KAIST, Daejeon, Korea
[2]Dept. of Applied Statistics & Dept. of Statistics and Data Science, Yonsei University, Seoul, Korea
[3]Qunova Computing, Inc., Deajeon, S. Korea

October 5, 2022

Supplementary Information

A Proof of Eqs. (15) and (16)

The proof of Eqs. (15) is as follows:

$$\begin{aligned}
|\psi\rangle &= \sum_{i=0}^{M-1} \sqrt{\alpha_i}\,|0, i, \phi(\mathbf{x}_i), y_i, \phi(\hat{\mathbf{x}})\rangle \\
&\Longrightarrow \frac{1}{\sqrt{2}} \sum_{i=0}^{M-1} \sqrt{\alpha_i}\,|0, i, \phi(\mathbf{x}_i), y_i, \phi(\hat{\mathbf{x}})\rangle + \frac{1}{\sqrt{2}} \sum_{i=0}^{M-1} \sqrt{\alpha_i}\,|1, i, \phi(\mathbf{x}_i), y_i, \phi(\hat{\mathbf{x}})\rangle \\
&\Longrightarrow \frac{1}{\sqrt{2}} \sum_{i=0}^{M-1} \sqrt{\alpha_i}\,|0, i, \phi(\mathbf{x}_i), y_i, \phi(\hat{\mathbf{x}})\rangle + \frac{1}{\sqrt{2}} \sum_{i=0}^{M-1} \sqrt{\alpha_i}\,|1, i, \phi(\hat{\mathbf{x}}), y_i, \phi(\mathbf{x}_i)\rangle \\
&\Longrightarrow \frac{1}{2} \sum_{i=0}^{M-1} \sqrt{\alpha_i}\left(|0, i, \phi(\mathbf{x}_i), y_i, \phi(\hat{\mathbf{x}})\rangle + |0, i, \phi(\hat{\mathbf{x}}), y_i, \phi(\mathbf{x}_i)\rangle\right) \\
&\quad + \frac{1}{2} \sum_{i=0}^{M-1} \sqrt{\alpha_i}\left(|1, i, \phi(\mathbf{x}_i), y_i, \phi(\hat{\mathbf{x}})\rangle - |1, i, \phi(\hat{\mathbf{x}}), y_i, \phi(\mathbf{x}_i)\rangle\right) \\
&\xrightarrow{\langle Z_a Z_y\rangle} \sum_{i=0}^{M-1} \alpha_i y_i |\langle\phi(\mathbf{x}_i)|\phi(\hat{\mathbf{x}})\rangle|^2, \ \xrightarrow{\langle I_a Z_y\rangle} \sum_{i=0}^{M-1} \alpha_i y_i
\end{aligned} \tag{1}$$

The proof of Eqs. (16) is as follows:

$$
\begin{aligned}
|\psi\rangle &= \sum_{i,j=0}^{M-1} \sqrt{\alpha_i}\sqrt{\alpha_j}\,|0, i, \phi(\mathbf{x}_i), y_i, j, \phi(\mathbf{x}_j), y_j\rangle \\
&\Longrightarrow \frac{1}{\sqrt{2}} \sum_{i,j=0}^{M-1} \sqrt{\alpha_i}\sqrt{\alpha_j}\,|0, i, \phi(\mathbf{x}_i), y_i, j, \phi(\mathbf{x}_j), y_j\rangle + \frac{1}{\sqrt{2}} \sum_{i,j=0}^{M-1} \sqrt{\alpha_i}\sqrt{\alpha_j}\,|1, i, \phi(\mathbf{x}_i), y_i, j, \phi(\mathbf{x}_j), y_j\rangle \\
&\Longrightarrow \frac{1}{\sqrt{2}} \sum_{i,j=0}^{M-1} \sqrt{\alpha_i}\sqrt{\alpha_j}\,|0, i, \phi(\mathbf{x}_i), y_i, j, \phi(\mathbf{x}_j), y_j\rangle + \frac{1}{\sqrt{2}} \sum_{i,j=0}^{M-1} \sqrt{\alpha_i}\sqrt{\alpha_j}\,|1, i, \phi(\mathbf{x}_j), y_i, j, \phi(\mathbf{x}_i), y_j\rangle \\
&\Longrightarrow \frac{1}{2} \sum_{i,j=0}^{M-1} \sqrt{\alpha_i}\sqrt{\alpha_j}\,(|0, i, \phi(\mathbf{x}_i), y_i, j, \phi(\mathbf{x}_j), y_j\rangle + |0, i, \phi(\mathbf{x}_j), y_i, j, \phi(\mathbf{x}_i), y_j\rangle) \\
&\quad + \frac{1}{2} \sum_{i,j=0}^{M-1} \sqrt{\alpha_i}\sqrt{\alpha_j}\,(|0, i, \phi(\mathbf{x}_i), y_i, j, \phi(\mathbf{x}_j), y_j\rangle - |0, i, \phi(\mathbf{x}_j), y_i, j, \phi(\mathbf{x}_i), y_j\rangle) \\
&\quad \xrightarrow{\left\langle Z_a Z_{y_0} Z_{y_1}\right\rangle} \sum_{i,j=0}^{M-1} \alpha_i\alpha_j y_i y_j |\langle\phi(\mathbf{x}_i)|\phi(\mathbf{x}_j)\rangle|^2, \ \xrightarrow{\left\langle I_a Z_{y_0} Z_{y_1}\right\rangle} \sum_{i,j=0}^{M-1} \alpha_i\alpha_j y_i y_j = \left(\sum_{i=0}^{M-1} \alpha_i y_i\right)^2
\end{aligned} \tag{2}
$$

For more detail, see.[1,2]

B Supplementary Figures

In this section, we provide supplementary figures mentioned in the main text.

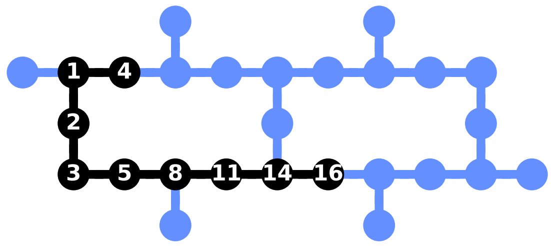


Figure 1: *ibm_montreal* qubits configurations. We used qubits colored in black for QASVM with toy data set

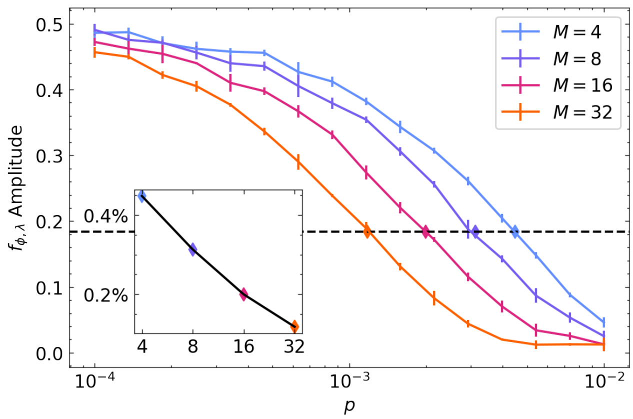
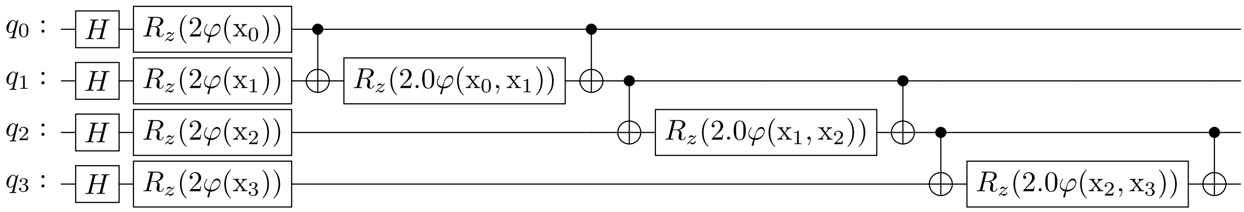
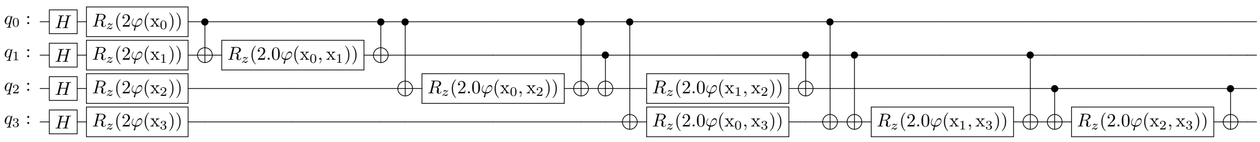


Figure 2: Noise robustness simulation result on depolarizing error. $f_{\phi,\lambda}$ amplitude represents amplitude of sine-fitted function of measured decision values. p is depolarizing error rate of single-qubit gate. We set the error rate of two-qubit gates 10 times higher than that of single-qubit gates.

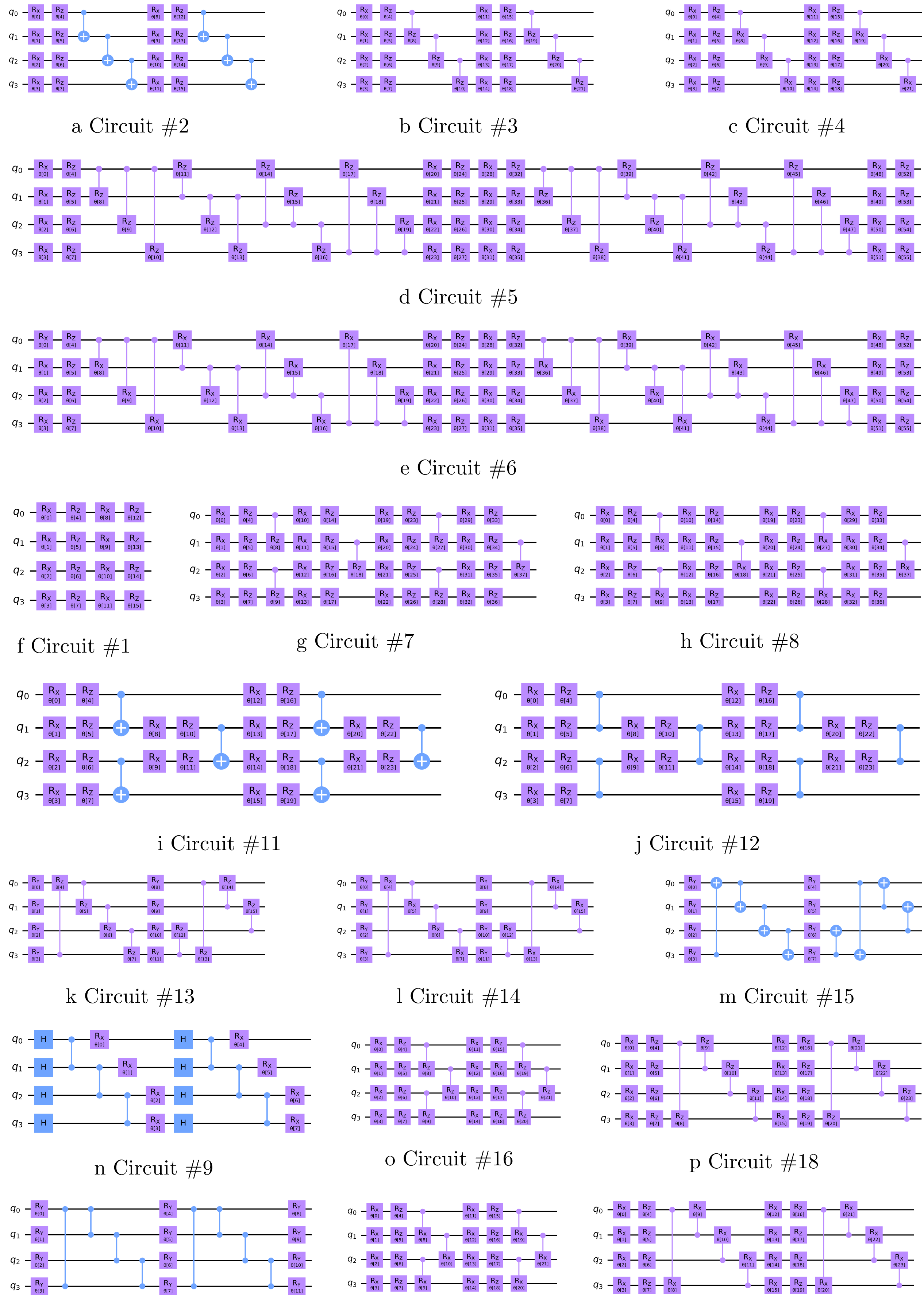


a *ZZFeatureMap* with linear entanglement



b *ZZFeatureMap* with all-to-all entanglement

Figure 3: Example quantum feature map from Ref.[3]



a Circuit #2 b Circuit #3 c Circuit #4

d Circuit #5

e Circuit #6

f Circuit #1 g Circuit #7 h Circuit #8

i Circuit #11 j Circuit #12

k Circuit #13 l Circuit #14 m Circuit #15

n Circuit #9 o Circuit #16 p Circuit #18

q Circuit #10 r Circuit #17 s Circuit #19

Figure 4: 19 sample PQC architecture provided in Ref. 4. Each circuit has 4 qubits and 2 layers.

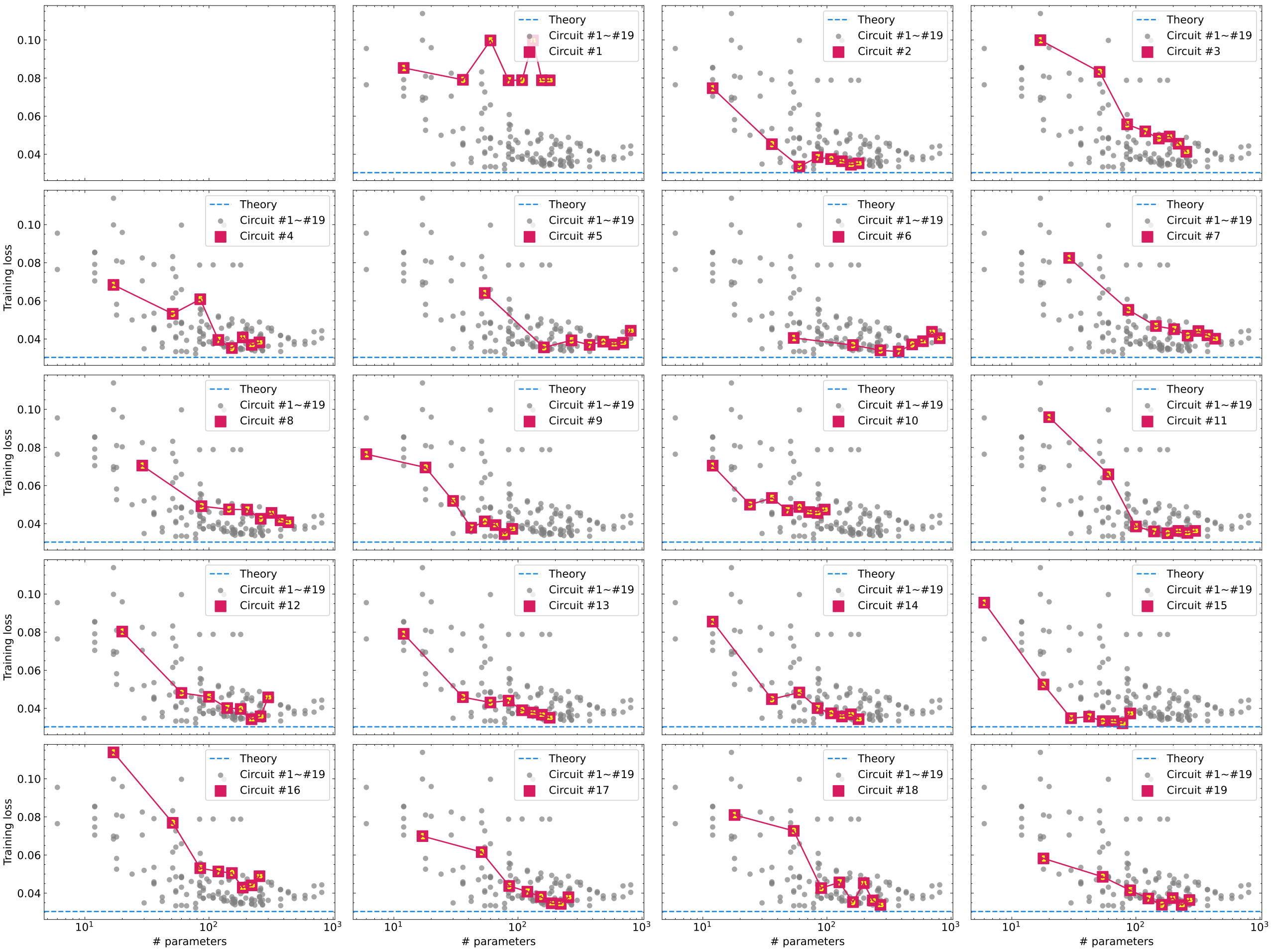


Figure 5: Training results on iris data set ($M = 64$)

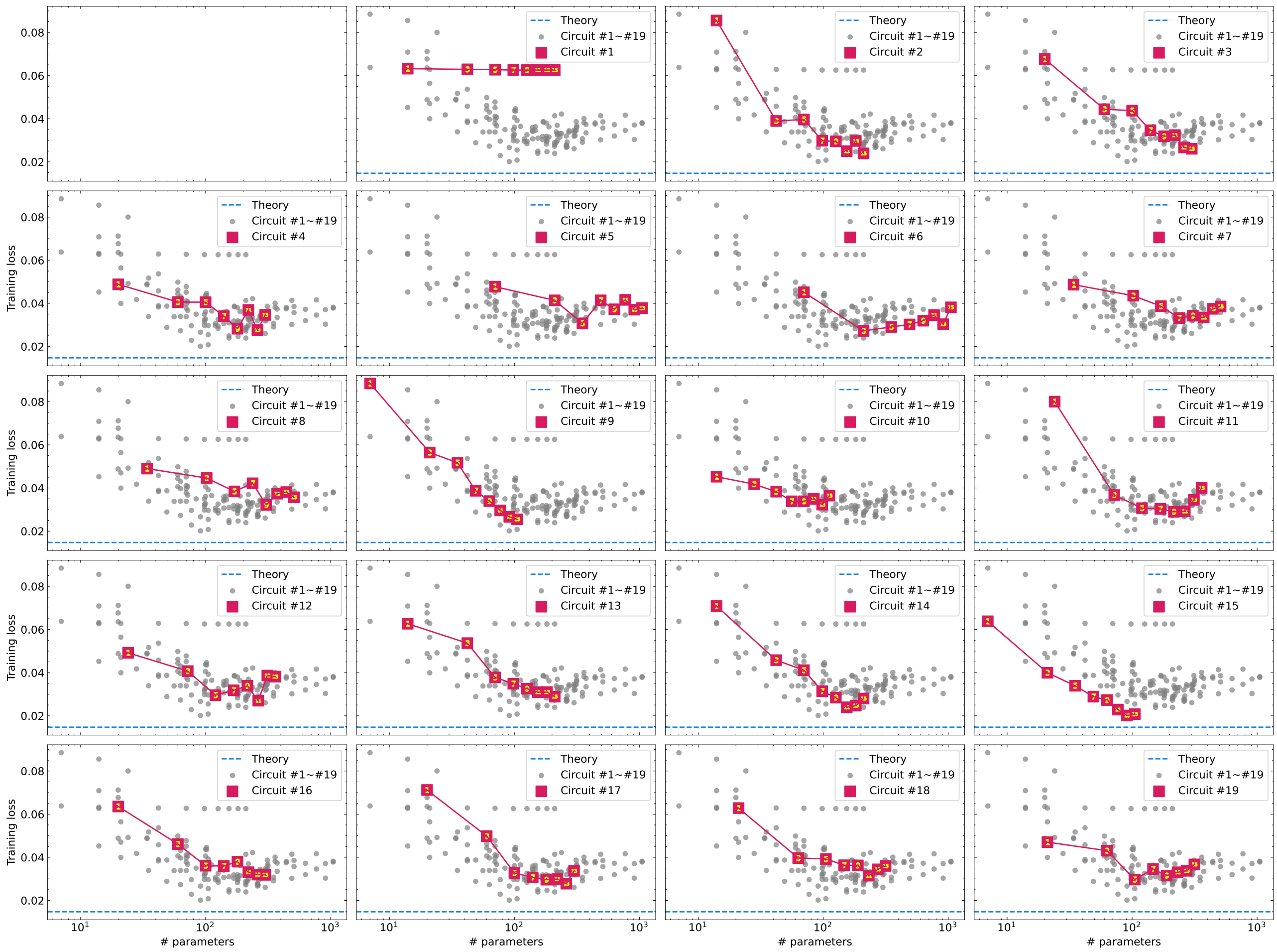
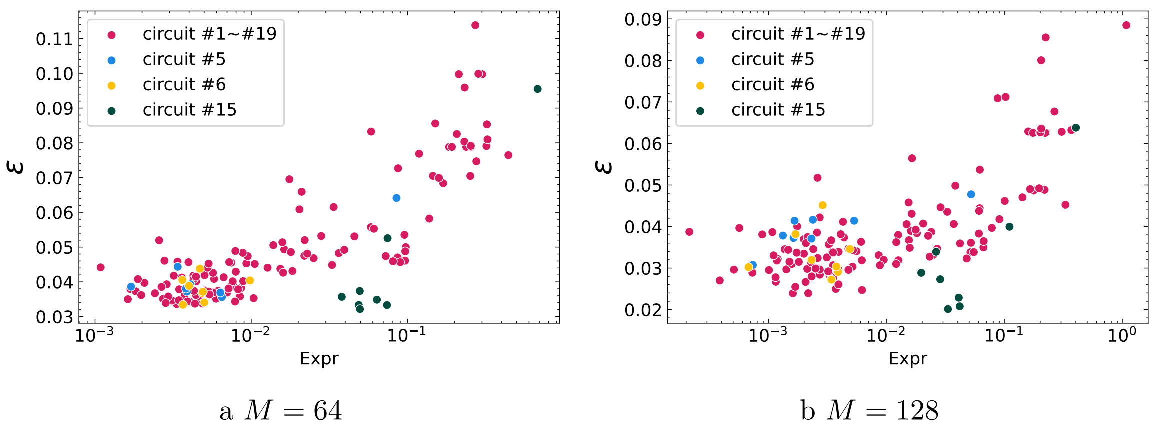


Figure 6: Training results on iris data set ($M = 128$)



a $M = 64$

b $M = 128$

Figure 7: Decision value error on iris data set. Expr refers to expressibility of PQCs.[4] Circuit #5 and #6 has $L \times \log(M)^2$ parameters.

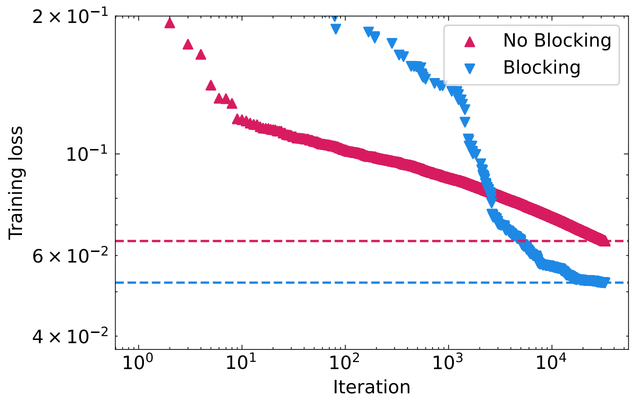
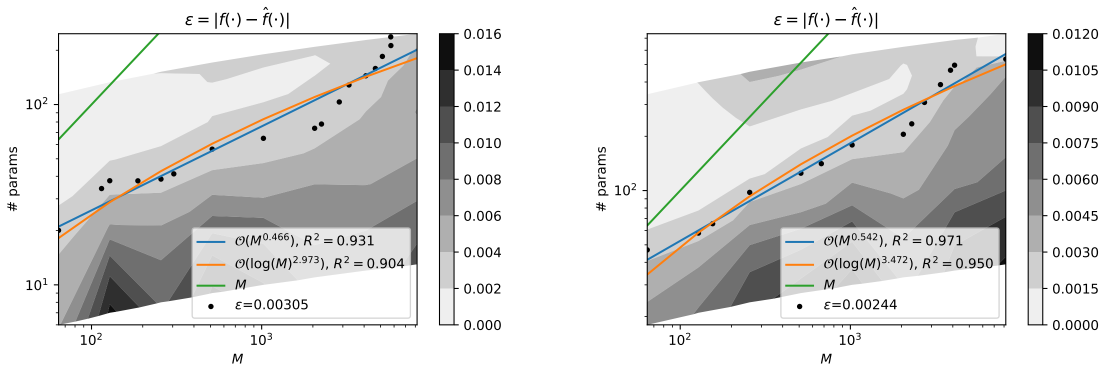


Figure 8: Blocking technique for SPSA. Training loss converges faster when blocking is used.



a *BasicEntanglerLayers* Ansatz

b *StronglyEntanglingLayers* Ansatz

Figure 9: Sublinear growth of number of parameters independent of PQC architecture.

References

[1] Blank, C., Park, D. K., Rhee, J.-K. K. & Petruccione, F. Quantum classifier with tailored quantum kernel. *npj Quantum Information* **6**, 41 (2020). URL http://www.nature.com/articles/s41534-020-0272-6.

[2] Park, D. K., Blank, C. & Petruccione, F. The theory of the quantum kernel-based binary classifier. *Physics Letters A* **384**, 126422 (2020). URL https://linkinghub.elsevier.com/retrieve/pii/S0375960120302541.

[3] Havlíček, V. *et al.* Supervised learning with quantum-enhanced feature spaces. *Nature* **567**, 209–212 (2019). URL http://www.nature.com/articles/s41586-019-0980-2.

[4] Sim, S., Johnson, P. D. & Aspuru-Guzik, A. Expressibility and Entangling Capability of Parameterized Quantum Circuits for Hybrid Quantum-Classical Algorithms. *Advanced Quantum Technologies* **2**, 1900070 (2019). URL https://onlinelibrary.wiley.com/doi/10.1002/qute.201900070.