Continuous-variable stochastic gradient descent methods applied for a photonic quantum neural network

Kareem H. El-Safty[1], Dániel Nagy[1,2], Péter Hága[2], Zsófia Kallus[2], Zoltán Zimborás[1]
1: Wigner Research Centre for Physics, 2: Ericsson Research; Contact: kareem.elsafty@wigner.hu, zimboras.zoltan@wigner.hu

Abstract

Quantum machine learning is a research subject that is becoming increasingly relevant within near-term quantum computation. Quantum neural networks are a well-known example for learning tasks in both qubit and qumode-based systems. Stochastic Gradient Descent (SGD) methods have been introduced in recent work for various variational quantum circuits but only in qubit-based architectures. In this work, we present different variants of the SGD method in the continuous-variable framework via the example application of photonic quantum neural networks for regression models. Using only a low number of shots, we see a nicely converging training process using simulated circuits on StrawberryFields [3] photonic simulator and its GPU enabled TensorFlow backend. Regression results are found best on 2-qumode circuits for the presented test function approximations and the effect of further stochastically sampling with mini-batches during the training process.

Objectives

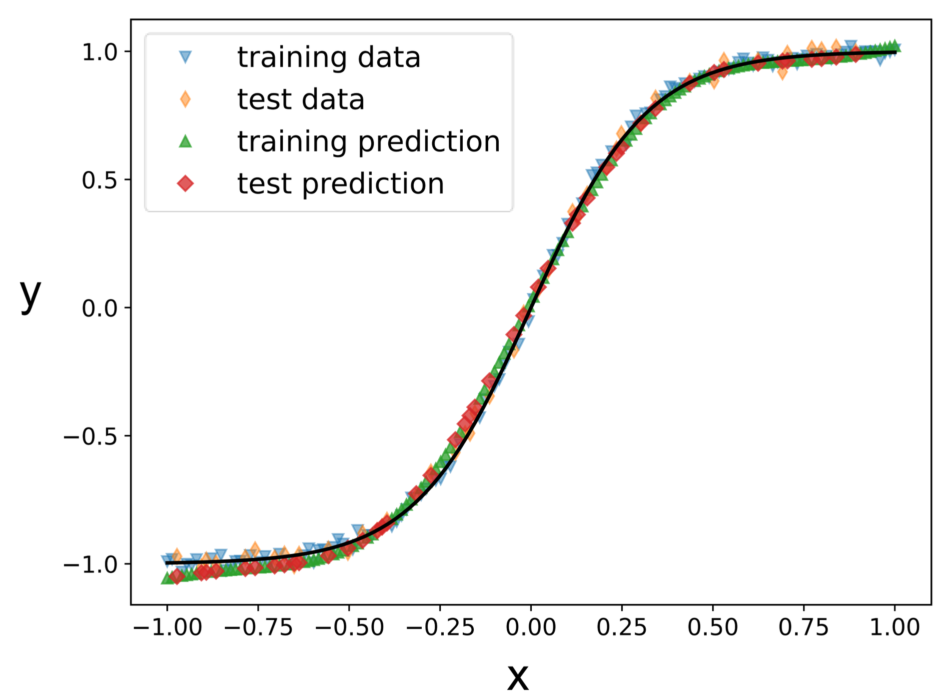
In this work, we present continuous variable quantum neural networks [1] trained with different variants of the SGD method. We chose a regression model to demonstrate the performance of the stochastic approach in approximating the tanh function. First, we investigate the number of epochs necessary for the convergence of the iterative training process. Second, we test the doubly stochastic version, by varying the size of the mini-batches while using a fixed shot number.

Regression with Photonic QNN

Continuous-Variable Quantum neural networks which contain active gates should always have some regularization to keep the trace of the state close to 1. Therefore, we used L2 regularizations on the active weights:

$$\mathcal{L} = \frac{1}{m}\sum_{i=0}^{m}(y_i - \langle\psi|\hat{x}|\psi\rangle)^2 + \frac{\lambda}{m}\sum_{S,D}\|W_i\|_2$$

A final inference plot for the tanh function is shown on the figure below. We added a small noise to the training data to avoid overfitting.

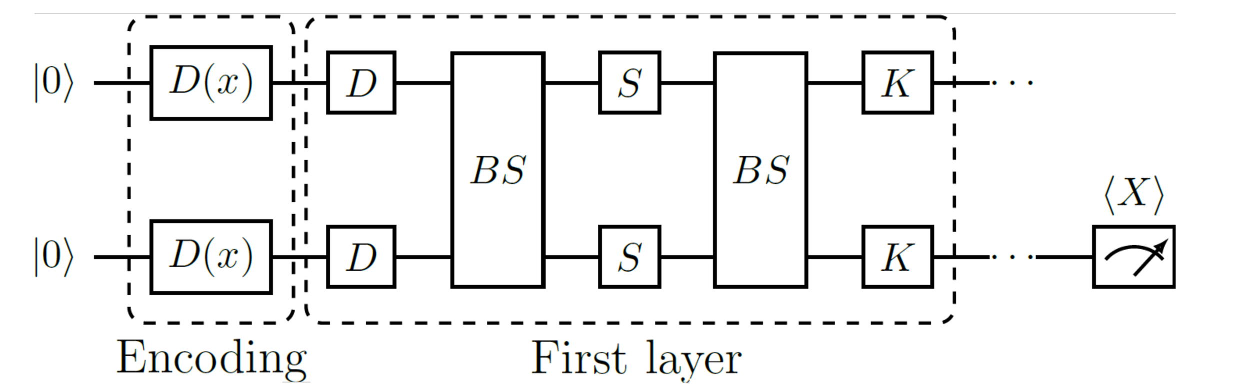


Stochastic methods

Here we present the adaptation of two stochasticity versions. First, by using only a finite number of shots per measurement, we show that an estimation by a finite number of shots gives acceptable convergence even if we use relatively few shots. We can further reduce the total number of shots by sampling mini-batches from the dataset.

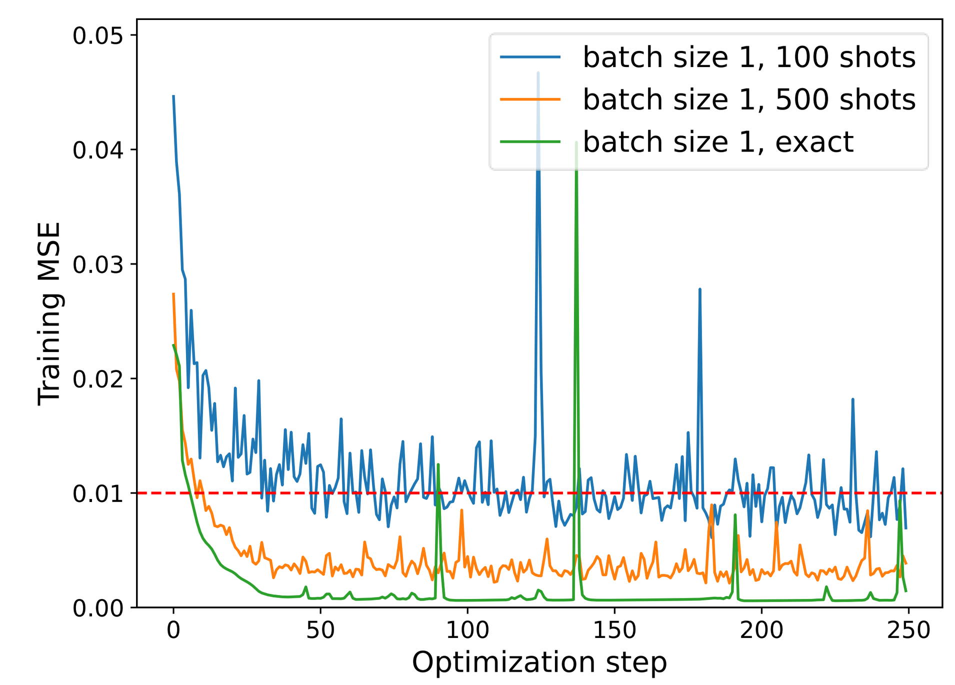
Ansatz circuit

We implemented a generalized ansatz circuit composed of passive and active photonic gates. Each layer is made of a series of Displacement, Beam-Splitters, Squeezing, Kerr, and Displacement gates. We used six consecutive layers as shown below. Then we use the last qumode for the measurement process.



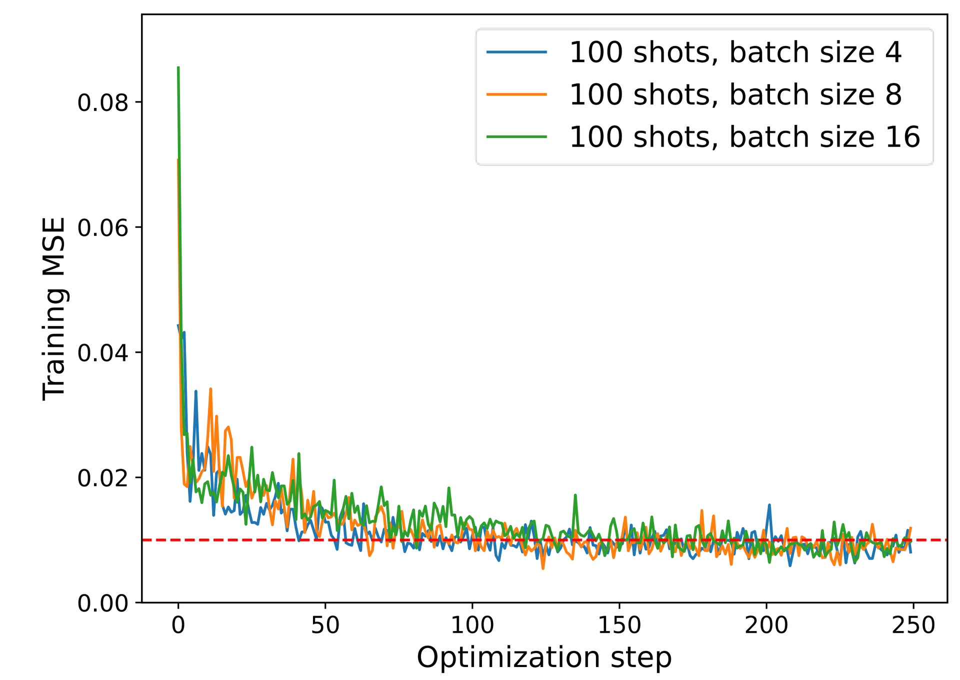
Stochastic results

Using SGD and a varying number of shots, an acceptable convergence is reached for 500 shots within 10 steps, and using only 100 shots, within 50 steps. Hence, the stochastic approach introduced for photonic circuits can effectively be used with a significant gain in costly measurements.



Doubly stochastic results

Introducing double stochasticity by using minibatch training, one can significantly reduce the number necessary gradient evaluations and by finding optimal batch size, the convergence process can become more stable. Here, using minibatch sizes below 20, the 100 shot simulations converge without significant slowdown, and at the same time, the large fluctuations disappeared.



Based on these results, we find that using the doubly stochastic method might be a very good choice in continuous-variable quantum neural networks.

Conclusion and future work

- We found that the optimal number of qumodes was two, and the optimal number of layers in 6.
- The doubly stochastic method takes more iterations to converge.
- In the future we want to study the tread-off between the number of shots and iteration steps.
- We will benchmark this method against a real world measured dataset.
- We will investigate the effect of encoding the dataset on the convergence.

References

1. Nathan Killoran, Thomas R. Bromley, Juan Miguel Arrazola, Maria Schuld, Nicolás Quesada, Seth Lloyd, Continuous-variable quantum neural networks, Phys. Rev. Research 1, 033063 (2019),Eprint: arXiv:1806.06871
2. Ryan Sweke, Frederik Wilde, Johannes Meyer, MariaSchuld, Paul K. Faehrmann, Barthélémy Meynard-Piganeau, and Jens Eisert. Stochastic gradient descent for hybrid quantum-classical optimization. Quantum,4, 314 (2020). Eprint: arXiv:1910.01155v3
3. Nathan Killoran, Josh Izaac, Nicolás Quesada, VilleBergholm, Matthew Amy, and Christian Weedbrook. Strawberry Fields: A Software Platform for Photonic Quantum Computing. Quantum, 3, 129 (2019). Eprint: arXiv:1804.03159

Acknowledgement

- Special thanks to the Wigner RCP - GPU-Lab for providing their computational infrastructure.