Detecting Underdetermination in Parametrized Quantum Circuits

Marie Kempkes[1, 2], Jakob Spiegelberg[2], Evert van Nieuwenburg[1] and Vedran Dunjko[1]

[1] Leiden University, Niels Bohrweg 1, 2333 CA Leiden, Netherlands
[2] Volkswagen Group Innovation, Berliner Ring 2, 38440 Wolfsburg, Germany

Scientific progress and enhanced computing capacity enabled AI methods to acquire increasing influence in the past few years, and it can be assumed that this trend will continue. With the application of machine learning models in safety-related domains, the significance of making them safe should not be overlooked. Safe AI is however still an underexplored research area that receives only a fraction of resources compared to performance improvements. At the same time, the field of quantum machine leaning (QML) continues to evolve and scientists work persistently at identifying application fields in which QML promises advantages over classical methods. We argue that while seeking potential benefits of QML methods, it is just as crucial to devote energy to making them safe.

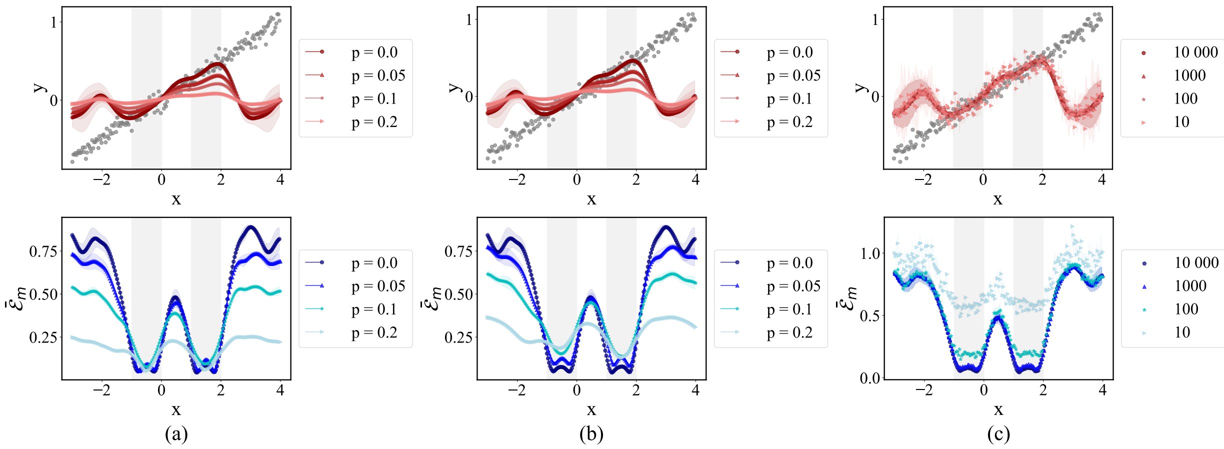

Figure 1: **Underdetermination detection in noisy quantum circuits.** The method for detecting underdetermination employed in this work is robust against different sources of hardware noise. Plot shows predictions and average extrapolation score $\bar{\mathcal{E}}_m$ for varying levels of (a) depolarization noise, (b) bitflip noise and (c) shot noise. Training data only originates from grey shaded intervals, which is why underdetermination is expected to occur elsewhere.

In this work, besides providing an overview over important concepts on Safe AI in the classical literature, we employ a method using second-order information based on the Hessian matrix [1] on PQCs to detect underdetermination, which is a major contributor to predictive uncertainty. Not only do we show that it is possible to reliably detect underdetermination, but we also analyze how the method behaves under different sources of hardware noise in numerical experiments. Compared to ensembling techniques for underdetermination detection, the method we utilize is computationally light-weight during training as well as in inference. Our work extends on the scarce literature in the field of Safe QML and aims at emphasizing the importance of doing more research in this field.

[1] David Madras, James Atwood, and Alexander D'Amour. Detecting Extrapolation with Local Ensembles. In *International Conference on Learning Representations*, 2020.