Problem-informed Graphical Quantum Generative Learning

Bence Bakó, Dániel T. R. Nagy, Péter Hága, Zsófia Kallus, Zoltán Zimborás

Introduction

Quantum resources, due to their inherent probabilistic nature, can be used to efficiently draw samples from probability distributions of high complexity [1]. This makes generative learning, the machine learning paradigm aiming to capture the underlying distribution of data in order to generate realistic samples, a natural pathway towards harnessing the potential of quantum computers. Recent advances in quantum machine learning (QML) led to the adaptation of several successful classical generative models. In this work, we concentrate on quantum circuit Born machines (QCBMs) [2, 3]. This paradigmatic generative QML model naturally inherits the Born rule and thus can be used to generate tunable and discrete probability distributions that approximate a target distribution.

While their high expressivity makes general purpose QML models very powerful, they also pose several challenges. Contrary to classical neural networks, variational quantum circuits are much more affected by trainability issues, such as barren plateaus and poor local minima. Furthermore, the no-free-lunch theorem also translates to QML, suggesting, that these problem-agnostic models, like hardware-efficient Ansätze, have poor average performance. The reason behind these barriers can be seen as the lack of sufficient inductive bias, i.e., assumptions about the data that could be encoded into the learning algorithm. Consequently, a potential way of dealing with them is by constructing problem-informed models, that can be trained more efficiently for structured problems.

Probabilistic graphical models

Probabilistic graphical models (PGMs) provide a mathematical framework for representing structure in generative learning problems defined over random variables [4], and as such, can be exploited to construct problem-informed QML models. PGMs use a graph representation to compactly encode a complex distribution of interacting random variables.

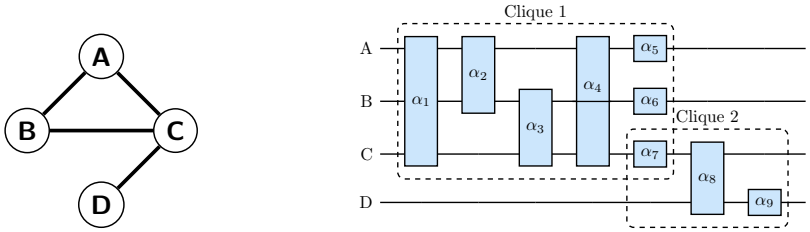
The two main classes of PGMs are Bayesian networks (BNs) and Markov networks (MNs) each having their merits in different domains. BNs use directed acyclic graphs to represent the conditional dependencies between random variables. As opposed to BNs, where the factors are represented by edges, and can be interpreted as conditional probabilities, for MNs the factors of the joint probability distribution are given by the cliques of an undirected graph and describe the compatibilities between the corresponding random variables.

While there are several excellent works concerning the quantum circuit implementation of BNs [5], MNs are not well-studied in the context of generative QML. In this work, we investigate the applicability of the framework provided by MNs to generative QML with classical data.

Markov networks for generative QML

We propose a problem-informed QCBM Ansatz for learning the joint probability distribution of binary random variables, where the independence relations are efficiently represented by a MN. We start by constructing a parametrized many-body Ising Hamiltonian, that is inspired by the log-linear model of MNs, and consequently depends on the clique-structure of the MN of interest. This Hamiltonian takes the form

$$H'(\boldsymbol{\beta}) = \sum_{C\in\mathcal{C}} \bigotimes_{v\in C} \beta_{C,v}(I + Z_v),$$



(a) (b)

Figure 1: (a) Markov network with two maximal cliques: Clique 1 = $\{A, B, C\}$, Clique 2 = $\{C, D\}$. (b) Decomposition of the corresponding $U_Z(\boldsymbol{\alpha}) = e^{-iH(\boldsymbol{\alpha})}$ operator, where parametrized gates represent k-local $e^{-i\alpha Z \ldots Z}$ operators.

where Z_v is the Pauli-Z operator acting on qubit v, $\mathcal{C}$ refers to the set of cliques and $\boldsymbol{\beta}$ is the parameter vector. Usually some of the MN cliques overlap in nonzero subsets, thus there will be reoccurring terms. Since all terms commute, we can reparametrize the Hamiltonian such that each term only appears once (and identities are excluded): $H'(\boldsymbol{\beta}) \to H(\boldsymbol{\alpha})$. For example, for the two cliques in the Markov network depicted in Fig. 1a, we would have

$$H(\boldsymbol{\alpha}) = \alpha_1 Z_A Z_B Z_C + \alpha_2 Z_A Z_B + \alpha_3 Z_B Z_C + \alpha_4 Z_A Z_C + \alpha_5 Z_C Z_D + \alpha_6 Z_A + \alpha_7 Z_B + \alpha_8 Z_C + \alpha_9 Z_D.$$

Having this parametrized Hamiltonian specific to the MN of interest, we can construct a quantum circuit model similar to the quantum circuit Ising Born machine, introduced in [3]. However, instead of a generic 2-local Hamiltonian with all-to-all connectivity, we implement the unitary generated by our higher-order Ising Hamiltonian $H(\boldsymbol{\alpha})$. The Ansatz starts in the $|+\rangle^{\otimes n}$ state, then applies the unitary $U_Z(\boldsymbol{\alpha}) = e^{-iH(\boldsymbol{\alpha})}$ (see Fig. 1) and finally implements general one-qubit gates and takes computational basis measurements. We call these problem-informed QCBM models quantum circuit Markov random fields (QCMRFs).

Performance, trainability and potential quantum advantage

MNs provide a universal framework, that can represent any probability distribution. However, some cases prove to be more useful, for our QCMRF model construction, than others. Therefore, we restrict the class of MNs of interest through a series of numerical experiments and complexity theoretic arguments.

First, we compare our model to the problem-agnostic QCIBM as formulated in [3], we present simulation results based on MNs with grid-like topology, always considering the maximal clique factorization. We consider using both Kullback-Leibler (KL) divergence and the squared maximum mean discrepancy (MMD) and calculate the total variational (TV) distance analytically in each step for comparison. These results (shown in Fig. 2) demonstrate that as we increase the connectivity of the graph and the sizes of its maximal cliques, the distribution becomes harder to learn, as it is reflected in the performance of the problem-agnostic QCIBM. However, the performance of our problem-specific QCMRF model is either unaffected by this change, or it performs even better, as its complexity also increases with the underlying MN. We also argue, that this enhancement is not a consequence of a higher number of trainable parameters, but the ability of the QCMRF model to capture higher-order correlation between the random variables. These minimal experiments are complemented by other more realistic random graphs with specific community structures, where this enhancement is also present. The numerical results highlight the potential of our framework to outperform problem-agnostic models, especially when the random variables exhibit higher order correlations. Furthermore, we validate the performance of our model, comparing it to the basis-enhanced Bayesian quantum circuit (BBQCs) model introduced in [5].

Besides studying the performance, we also conduct a preliminary analysis of trainability by studying the scaling of the MMD cost variance with the number of qubits. Our investigations

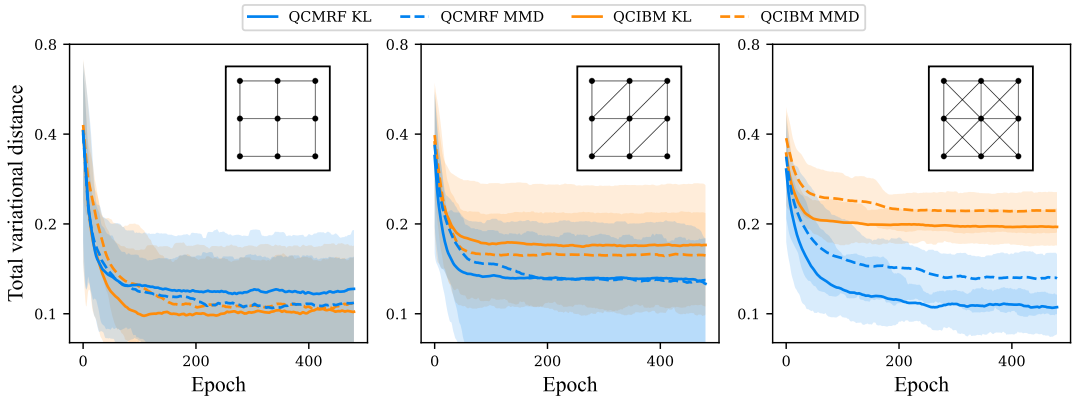


Figure 2: QCMRF benchmark results against the problem-agnostic QCIBM model. As the complexity of the problem increases with larger maximal cliques, the performance of QCIBM decreases, while QCMRFs are either unaffected by this, or they even improve.

reveal, that for complete graphs with maximal clique factorization (that have an exponential number of parameters) exhibit deterministic barren plateaus. On the other hand, for sparser graphs, the scaling is not exponential anymore. This leads to the definition of efficient MN representations, that have poly(n) degrees of freedom, and consequently the depth of the quantum circuit is also polynomial.

Lastly, we discuss the possibility that the QCMRF model provides rigorous quantum advantage relevant in the context of generative learning. We give formal definitions for two different concepts of quantum advantage (Definitions 4 and 5 in the main paper): the first concentrating on learning advantage, the other on efficiently sampling the learned distribution. We present a possible line of arguments for the second setting, highlighting the potential of our model to offer improvements over classical methods. This is based on Refs. [1, 6, 7].

The preprint of the work can be found at https://arxiv.org/abs/2405.14072.

References

[1] E. Farhi and A. W. Harrow, "Quantum supremacy through the quantum approximate optimization algorithm," *arXiv preprint arXiv:1602.07674*, 2016.

[2] M. Benedetti, D. Garcia-Pintos, O. Perdomo, V. Leyton-Ortega, Y. Nam, and A. Perdomo-Ortiz, "A generative modeling approach for benchmarking and training shallow quantum circuits," *npj Quantum Information*, vol. 5, no. 1, p. 45, 2019.

[3] B. Coyle, D. Mills, V. Danos, and E. Kashefi, "The Born supremacy: quantum advantage and training of an Ising Born machine," *npj Quantum Information*, vol. 6, no. 1, p. 60, 2020.

[4] D. Koller and N. Friedman, *Probabilistic graphical models: principles and techniques*. MIT press, 2009.

[5] X. Gao, E. R. Anschuetz, S.-T. Wang, J. I. Cirac, and M. D. Lukin, "Enhancing generative models via quantum correlations," *Physical Review X*, vol. 12, no. 2, p. 021037, 2022.

[6] A. M. Dalzell, A. W. Harrow, D. E. Koh, and R. L. La Placa, "How many qubits are needed for quantum computational supremacy?," *Quantum*, vol. 4, p. 264, 2020.

[7] H. Krovi, "Average-case hardness of estimating probabilities of random quantum circuits with a linear scaling in the error exponent," *arXiv preprint arXiv:2206.05642*, 2022.