

Spectral Born machines: Classically trainable quantum generative models for discrete data

Austin Huang, William Maxwell, Vasilis Belis, Evan Peters, Jason Pye, Soran Jahangiri,
and Joseph Bowles

Xanadu Quantum Technologies Inc., Toronto, Ontario, M5G 2C8, Canada

Paper summary

We present *spectral Born machines*, a class of quantum generative models that results from viewing and generalizing the class of IQP Born machines through the lens of group Fourier analysis. These quantum models exploit the quantum Fourier transform in order to induce smoothness priors that make them naturally suited to learning integer-structured data, while remaining classically hard to sample from in general. Similar to IQP Born machines, spectral Born machines can be trained efficiently at scale on classical hardware via a maximum mean discrepancy loss based on graph spectral analysis, which we make available in a new *tcdq* module of the PennyLane software platform. In numerical experiments, we show how the spectral bias of the model leads to significantly reduced parameter counts compared to unstructured approaches, and demonstrate the scalability of the software by training a 190-qubit model with over 1 million parameters to successfully learn a distribution of 93 nucleotide-long ribosomal RNA. Our results suggest that highly over-parameterized spectral Born machines may be immune to overfitting, even in strongly data-scarce regimes.

Extended abstract

The ability to use interference to construct and sample from high-dimensional probability distributions is what enables quantum computers to surpass the limits of classical computation. This is closely tied to the task of generative machine learning, where one aims to construct—or learn—a probability distribution given access to a finite set of empirical data. Viewed from this angle, generative machine learning is a natural area in which quantum computers may provide advantages for machine learning.

Modern machine learning models invariably make use of a large number of trainable parameters [1]. This is beneficial since it allows us to represent the belief that the ground truth distribution is a complicated, unknown function, while simultaneously adding a soft bias towards ‘simpler’ distributions that generalise better due to the benefits of scale [2, 3, 4]. It is therefore highly likely that successful quantum generative models will require at least modestly large parameter counts. However, even with an optimistic view on the clock speeds of future quantum computers [5], it is now well understood that traditional approaches to quantum circuit gradient estimation are unable to scale beyond relatively small numbers of parameters. New approaches to training are thus in desperate need [6, 7, 8, 9, 10, 11].

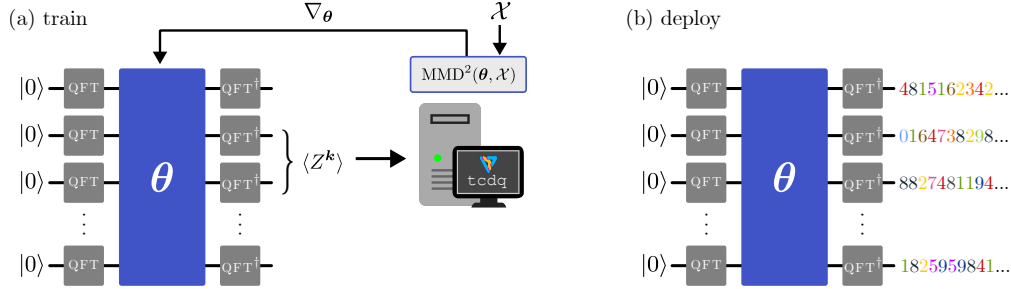
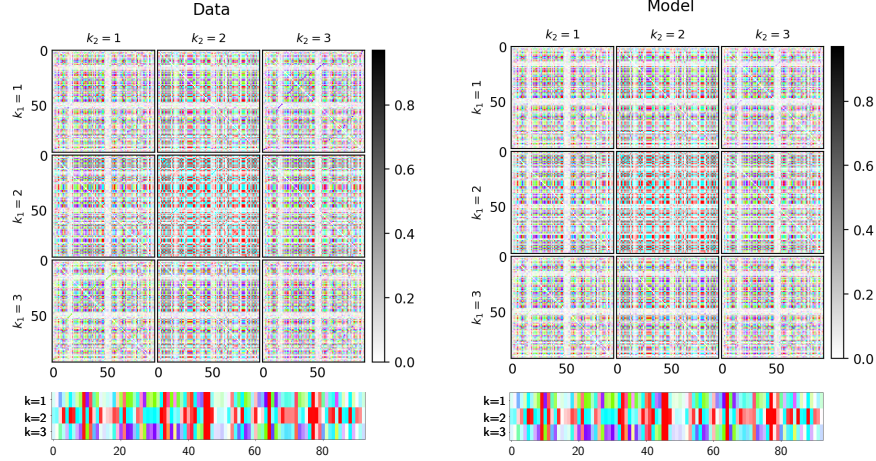


Figure 1: (a) A spectral Born machine consists of a diagonal parameterized unitary sandwiched between two layers of quantum Fourier transforms acting on qudits. For these circuits, expectation values of Heisenberg-Weyl observables can be classically estimated in order to train the maximum mean discrepancy loss function between the output distribution of the circuit and the empirical data distribution, which we make available via the PennyLane module `tcdq`. (b) Once training is complete, the trained circuit can be deployed on quantum hardware and used to sample new data, where genuine computational advantages may be present.

One solution to this problem was recently proposed in the form of a ‘train classical, deploy quantum’ (TCDQ) approach [12, 13]. Here, the basic idea is to construct a highly parameterized generative model based on a class of quantum circuits that are believed to be hard to classically sample from, while also identifying classically simulable observables that can be used to train the model. Since the observables are classically simulable, training can be performed on classical hardware. A quantum computer is therefore only required at inference time in order to sample from the model, where genuine computational advantages may be present. This enables scalable training, dramatically reducing the amount of quantum resources needed in the overall pipeline, while avoiding quantum data-loading bottlenecks since the model only sees data during the classical training loop. The possibility of training quantum generative models at scale in this way was first shown for instantaneous quantum polynomial (IQP) Born machines [12], and later for fermionic [14], bosonic [15, 16, 17] and correlation-based [18] circuits, and has attracted considerable attention as a promising route to quantum advantage [19, 20, 21, 22, 23].

All the existing models are however natively binary, meaning they are designed to learn distributions over bitstrings. This presents a problem, since most datasets of interest are not binary in nature, and simply encoding higher dimensional data via a binary encoding is not ideal due to the mismatch between the natural structure of the integers and bitstrings. The main drive of this work is therefore to construct a TCDQ-type generative model that is designed specifically with the discrete structure of integers in mind. To do this, we look to group theory and Fourier analysis, which provide the mathematical tools to reason about both structure and smoothness for such data. In particular, since we are working with data consisting of vectors of discrete values, the group we consider is $\mathbb{Z}_d^n$, the n -fold product group of the group $\mathbb{Z}_d$ that describes the integers $\{0, \dots, d-1\}$.

We show that the original IQP Born machines can be understood as models that parameterize the phase structure of the quantum state in Fourier space. Since IQP Born machines are based on qubits, the relevant Fourier transform is the group Fourier transform over $\mathbb{Z}_2^n$, which is realised by a layer of Hadamard unitaries. If we replace the Hadamard unitaries in an IQP Born machine by quantum Fourier transforms over the group $\mathbb{Z}_d$, and define an analogous parameterized diagonal phase layer, we obtain a classically trainable model that ties directly to the group-Fourier structure of $\mathbb{Z}_d^n$. Importantly for machine learning, this structure allows us to control inductive biases of the generative model distribution that can be expected to aid generalization. On the one hand, we can



directly bias towards distributional smoothness [24, 25, 26] with respect to the ordering of integers by selecting trainable gates that control only the low-frequency content of the phase unitary. On the other hand, we can bias the training of the model to match the structure of either ordinal or categorical discrete variables, which we show by formulating two versions of the maximum mean discrepancy loss function based on graph spectral analysis.

Like IQP Born machines, we show how training can be efficiently implemented using basic linear algebra operations, which allows the method to scale to large parameter and qudit counts. To facilitate future research, we implement the method in a dedicated `tcdq` module [27] of the PennyLane software platform [28], which we release alongside this work. The software is written in JAX [29], making algorithm construction flexible, differentiable, and scalable. Using `tcdq`, we first train a spectral Born machine on synthetic data generated via the Potts model. This distribution has a bias towards smoothness over the integers, making it a natural testbed for our model class. We show that spectral Born machines with gate sets that are engineered to match this bias are able to successfully learn from the data using much fewer parameters than unstructured equivalents. Second, in order to demonstrate the software’s scalability, we train a spectral Born machine with over 1 million parameters on a scarce distribution of 93-nucleotide ribosomal RNA. The figure above shows the values all Fourier coefficients of order one (panel below) and two (panel above) Fourier coefficients for the train plus test data (left) versus the trained model (right), where hue represents magnitude and colour represents phase. The very close resemblance of the two plots essentially means the model is able to accurately capture the low body information present in the data, showing that the model is able to learn well at scale. Further to this, we also give evidence that despite extreme overparameterization, spectral Born machines may be immune to overfitting, even in strongly data-scarce regimes. Overall, we hope this work ignites further interest in classically trainable generative models, enables an exciting new era of empirical quantum machine learning, and ultimately leads to the discovery of groundbreaking new applications of quantum machine learning.

Note. Another work [30] appeared on the arxiv at a similar time which generalizes IQP Born machines in a similar manner. Ref. [30] emphasizes an application to high-energy physics, with symmetry-adapted circuits, hardware compilation, and small-scale sampling-based validation, whereas this work develops the group-Fourier formulation over $\mathbb{Z}_d^n$, using general Heisenberg-Weyl observables and a graph-spectral MMD, discusses the associated spectral inductive bias, and releases scalable software that demonstrates the method scales to large models.

References

- [1] Jared Kaplan, Sam McCandlish, Tom Henighan, Tom B Brown, Benjamin Chess, Rewon Child, Scott Gray, Alec Radford, Jeffrey Wu, and Dario Amodei. Scaling laws for neural language models. *arXiv preprint arXiv:2001.08361*, 2020.
- [2] Andrew Gordon Wilson. Deep learning is not so mysterious or different. *arXiv preprint arXiv:2503.02113*, 2025.
- [3] Andrew G Wilson and Pavel Izmailov. Bayesian deep learning and a probabilistic perspective of generalization. *Advances in neural information processing systems*, 33:4697–4708, 2020.
- [4] Chaoyue Liu, Libin Zhu, and Mikhail Belkin. Loss landscapes and optimization in over-parameterized non-linear systems and neural networks. *Applied and Computational Harmonic Analysis*, 59:85–116, 2022.
- [5] Torsten Hoefer, Thomas Häner, and Matthias Troyer. Disentangling hype from practicality: On realistically achieving quantum advantage. *Communications of the ACM*, 66(5):82–87, 2023.
- [6] Joseph Bowles, David Wierichs, and Chae-Yeun Park. Backpropagation scaling in parameterised quantum circuits. *Quantum*, 9:1873, 2025.
- [7] Amira Abbas, Robbie King, Hsin-Yuan Huang, William J Huggins, Ramis Movassagh, Dar Gilboa, and Jarrod McClean. On quantum backpropagation, information reuse, and cheating measurement collapse. *Advances in Neural Information Processing Systems*, 36:44792–44819, 2023.
- [8] Brian Coyle, Snehal Raj, Natansh Mathur, El Amine Cherrat, Nishant Jain, Skander Kazdaghi, and Iordanis Kerenidis. Training-efficient density quantum machine learning. *npj Quantum Information*, 11(1):172, 2025.
- [9] Koki Chinzei, Shinichiro Yamano, Quoc Hoan Tran, Yasuhiro Endo, and Hirotaka Oshima. Trade-off between gradient measurement efficiency and expressivity in deep quantum neural networks. *npj Quantum Information*, 11(1):79, 2025.
- [10] Natansh Mathur, Panagiotis Kl Barkoutsos, Masako Yamada, Martin Roetteler, and Iordanis Kerenidis. Scalable on-hardware training of quantum neural networks and application to clinical data imputation. *arXiv preprint arXiv:2606.03517*, 2026.
- [11] Brian Coyle, Snehal Raj, Virag Umathe, El Amine Cherrat, and Elham Kashefi. Adaptive directional gradients for parameterised quantum circuits. *arXiv preprint arXiv:2606.09734*, 2026.
- [12] Erik Recio-Armengol, Shah Nawaz Ahmed, and Joseph Bowles. Train on classical, deploy on quantum: scaling generative quantum machine learning to a thousand qubits. *arXiv preprint arXiv:2503.02934*, 2025.
- [13] Sachin Kasture, Oleksandr Kyriienko, and Vincent E Elfving. Protocols for classically training quantum generative models on probability distributions. *Physical Review A*, 108(4):042406, 2023.

- [14] Bence Bakó, Zoltán Kolarovszki, and Zoltán Zimborás. Fermionic born machines: Classical training of quantum generative models based on fermion sampling. *arXiv preprint arXiv:2511.13844*, 2025.
- [15] Felix Gottlieb, Rawad Mezher, Brian Ventura, Shane Mansfield, and Alexia Salavrakos. Efficient training of photonic quantum generative models. *arXiv preprint arXiv:2603.08793*, 2026.
- [16] Zoltán Kolarovszki, Bence Bakó, Michal Oszmaniec, Changhun Oh, and Zoltán Zimborás. Generative modeling with gaussian boson sampling: classically trainable bosonic born machines. *arXiv preprint arXiv:2603.11195*, 2026.
- [17] Andrii Kurkin, Ulysse Chabaud, Zoltán Kolarovszki, Bence Bakó, Zoltán Zimborás, and Vedran Dunjko. Universality of classically trainable, quantum-deployed boson-sampling generative models. *arXiv preprint arXiv:2603.11014*, 2026.
- [18] Cenk Tüysüz, Oleksandr Kyriienko, and Michele Grossi. Quantum fourier generative models trainable at large scale. *arXiv preprint arXiv:2606.28483*, 2026.
- [19] Mario Herrero-Gonzalez, Brian Coyle, Kieran McDowall, Ross Grassie, Sjoerd Beentjes, Ava Khamseh, and Elham Kashefi. The born ultimatum: Conditions for classical surrogation of quantum generative models with correlators. *arXiv preprint arXiv:2511.01845*, 2025.
- [20] Oriol Balló-Gimbernát, Marcos Arroyo-Sánchez, Paula García-Molina, Adan Garriga, and Fernando Vilariño. Shallow instantaneous quantum polynomial-time circuits for generative modeling on noisy intermediate-scale quantum hardware. *Physical Review A*, 113(4):042617, 2026.
- [21] Sabrina Herbst, Ivona Brandić, and Adrián Pérez-Salinas. Limits of quantum generative models with classical sampling hardness. *arXiv preprint arXiv:2512.24801*, 2025.
- [22] Jamal Slim, Saverio Monaco, Florian Rehm, Dirk Krücker, and Kerstin Borrás. An iqp born machine for calorimeter image generation at 64 qubits with compiled-iqp deployment. *arXiv preprint arXiv:2605.27735*, 2026.
- [23] Markus Baumann, Daniel Hein, Steffen Udluft, Tobias Rohe, Claudia Linnhoff-Popien, and Jonas Stein. Parity supervision as a driver of generalization in quantum generative modeling. *arXiv preprint arXiv:2605.10258*, 2026.
- [24] Mihaela Rosca, Theophane Weber, Arthur Gretton, and Shakir Mohamed. A case for new neural network smoothness constraints. *PMLR*, 137:21–32, 2020.
- [25] Vasilis Belis, Joseph Bowles, Rishabh Gupta, Evan Peters, and Maria Schuld. Spectral methods: crucial for machine learning, natural for quantum computers? *arXiv preprint arXiv:2603.24654*, 2026.
- [26] Benoit Dherin, Michael Munn, Mihaela Rosca, and David Barrett. Why neural networks find simple solutions: The many regularizers of geometric complexity. *Advances in Neural Information Processing Systems*, 35:2333–2349, 2022.
- [27] tcdq is currently under development but available as part of PennyLane labs. See <https://pennylane.ai/blog/2025/01/cutting-edge-research-with-labs/>.

- [28] Ville Bergholm, Josh Izaac, Maria Schuld, Christian Gogolin, Shahnawaz Ahmed, Vishnu Ajith, M Sohaib Alam, Guillermo Alonso-Linaje, B AkashNarayanan, Ali Asadi, et al. PennyLane: Automatic differentiation of hybrid quantum-classical computations. *arXiv preprint arXiv:1811.04968*, 2018.
- [29] James Bradbury, Roy Frostig, Peter Hawkins, Matthew James Johnson, Yash Katariya, Chris Leary, Dougal Maclaurin, George Necula, Adam Paszke, Jake VanderPlas, Skye Wanderman-Milne, and Qiao Zhang. JAX: composable transformations of Python+NumPy programs, 2018.
- [30] Robert J Banks, Arianna Crippa, Matthias Traube, Josua Unger, Christian Ertler, and Wolfgang Lechner. Qudit extension of parameterized iqp circuits: A generative quantum machine learning approach to integer data. *arXiv preprint arXiv:2606.28236*, 2026.