

Circuit Harmonic Matrices: A Spectral Framework for Quantum Machine Learning

Paper at <https://arxiv.org/abs/2604.04292>

Kyle J. S. Campbell¹

Luigi Del Debbio²

Petros Wallden¹

¹Quantum Software Lab, School of Informatics, University of Edinburgh, United Kingdom

²School of Physics and Astronomy, University of Edinburgh, United Kingdom

Overview. This extended abstract summarises the full paper *Circuit Harmonic Matrices: A Spectral Framework for Quantum Machine Learning* [1]. Parametrised quantum circuits (PQCs) are a common modelling tool in near-term quantum machine learning, but the effect of architectural choices such as the encoder, ansatz, entangling pattern, observable, and input state on a model’s expressivity and training behaviour is difficult to determine before training. Our central contribution is the introduction and analysis of the *circuit harmonic matrix* C , a data-independent architecture matrix for quantum Fourier models (QFMs) with commuting phase encoders and Pauli-rotation trainable blocks [1, 2, 3, 4]. In this representation the model output has the joint input–parameter harmonic expansion

$$f(x; \theta) = \sum_{\omega \in \Omega} a_{\omega}(\theta) e^{i\omega \cdot x} = \sum_{\omega \in \Omega} \sum_{k \in K} C_{\omega k} e^{i\omega \cdot x} e^{ik \cdot \theta}, \quad a(\theta) = C\psi(\theta), \quad \psi_k(\theta) = e^{ik \cdot \theta}. \quad (1)$$

The rows of C are indexed by encoder-accessible input frequencies $\omega \in \Omega$, its columns by parameter harmonics $k \in K$ generated by the trainable circuit, and its entries give the amplitude with which each parameter harmonic contributes to each output frequency. We show that the coefficient variances, frequency–frequency covariance and correlation, and quantum neural tangent kernel (QNTK) are quadratic forms in C . This makes architecture-induced coupling between learnable Fourier modes explicit before choosing a dataset or optimisation trajectory, supporting architecture comparison, problem–model matching, and more resource-efficient QML model design.

1 Problem and contributions

Existing QFM expressivity theory characterises which input frequencies Ω a circuit can represent, while QNTK theory describes local gradient-flow geometry on a specified dataset [5, 6, 7]. These perspectives do not directly identify which Fourier modes share the same trainable degrees of freedom, or how circuit architecture determines those shared degrees before data projection. We introduce the circuit harmonic matrix C to fill this gap: a circuit-derived map from parameter harmonics to input harmonics that makes architectural coupling visible before training.

This work makes the following contributions:

- We introduce and analyse the circuit harmonic matrix C , the joint Fourier coefficient matrix of the model on the input–parameter torus (Equation (1)). C is fixed once the encoder, ansatz, observable, and input state are chosen, and is independent of the dataset and optimisation trajectory.
- We give a constructive interpretation of C using Heisenberg-picture Pauli propagation: each non-commuting gate–observable encounter produces a trigonometric branch, which becomes a parameter-harmonic label k after conversion to torus characters [8, 9, 10].
- We derive closed-form coefficient statistics from C : variances are row norm squares, covariances are row Gram matrices, and correlations identify frequencies coupled through shared parameter harmonics.
- We show that the QNTK evaluated on any finite input set factorises through the same matrix C , separating into an architecture factor (C), a data factor, and a circuit-independent parameter factor, so that coefficient statistics and tangent-kernel geometry are controlled by C .

- We validate these identities numerically against independent Monte Carlo estimates. Since the number of parameter harmonics $|K|$ can grow exponentially with circuit size, making C itself intractable to construct, we evaluate $\text{Cov}_\theta[a] = CC^\dagger$ exactly by carrying out the parameter-torus average analytically over pairs of Pauli propagation branches.

2 Results: covariance, correlation, and kernels

The factorisation $a(\theta) = C\psi(\theta)$ expresses the Fourier coefficients as a fixed linear image of the parameter-side features $\psi_k(\theta) = e^{ik\cdot\theta}$. Under uniform sampling over the parameter torus $\theta \sim \text{Unif}(\mathbb{T}^m)$, character orthogonality reduces the second-order coefficient statistics to row Gram matrices of C . Excluding the constant harmonic $k = 0$ (which carries the mean $\mathbb{E}_\theta[a] = C_{\bullet 0}$) and writing C hereafter for the resulting *centred* matrix,

$$\text{Cov}_\theta[a] = CC^\dagger, \quad \text{Corr}_\theta[a] = D^{-1/2} CC^\dagger D^{-1/2}, \quad \text{Var}_\theta[a_\omega] = \sum_{k \neq 0} |C_{\omega k}|^2 = \|C_{\omega,:}\|_2^2, \quad (2)$$

with $D = \text{diag}(CC^\dagger)$. Each coefficient variance is the squared norm of a row of C (its row energy), and two frequencies are correlated when their rows share a non-zero parameter harmonic, i.e. when the same trainable degree of freedom contributes to both. $CC^\dagger$ is therefore a frequency-coupling matrix determined by the architecture; it gives a closed form for the ‘‘Fourier fingerprint’’ used empirically to compare ansatzes [11], and it distinguishes access to a frequency from trainable coupling between frequencies. In the limit of strong parameter mixing it reproduces the redundancy–variance scaling of [12].

The same matrix determines the training kernel. For a finite input set with design matrix $V_{i\omega} = e^{i\omega \cdot x_i}$, the QNTK Gram matrix on these inputs factorises as

$$K(\theta) = VH(\theta)V^\dagger, \quad H(\theta) = CM(\theta)C^\dagger, \quad \mathbb{E}_\theta[H] = C \text{diag}(\|k\|_2^2) C^\dagger, \quad (3)$$

where $M(\theta)$ is a character-gradient kernel on the torus that does not depend on the circuit, and the expectation is over the same uniform parameter distribution as Equation (2). We refer to $H(\theta)$, the tangent kernel expressed in the frequency basis rather than on inputs, as the *harmonic QNTK*. This separates the kernel into an architecture factor C , a data factor V , and a parameter-manifold factor $M(\theta)$. The factorisation holds pointwise in θ , so the reachable tangent directions lie in the column space of VC throughout training, not only at initialisation. The coefficient statistics and the training kernel are thus determined by the same matrix C .

3 Validation

Constructing C explicitly is intractable for large circuits, since $|K|$ can grow exponentially. For the second-order statistics this construction is not required: because $CC^\dagger$ is an average over the parameter torus, it can be evaluated analytically term by term, with matched parameter harmonics contributing their exact inner products and mismatched harmonics vanishing by character orthogonality. The covariance is thus obtained exactly, without Monte Carlo sampling over parameters and without enumerating K ; the number of harmonic pairs to evaluate can, however, grow exponentially for circuits with repeated encoding gates. For the three-qubit YZY circuit of Figure 1, this exact evaluation reproduces the Monte Carlo variance profile to linear RMSE 1.7×10^{-5} and the off-diagonal correlation structure to Frobenius error 5.2×10^{-3} (off-diagonal cosine similarity 0.998), confirming that the observed correlations are the genuine row geometry of C rather than sampling effects.

4 Outlook and conclusions

Because the coefficient statistics and training kernel are expressed through C , several architecture-level questions can be examined before training; we outline three, each developed further in the full paper [1].

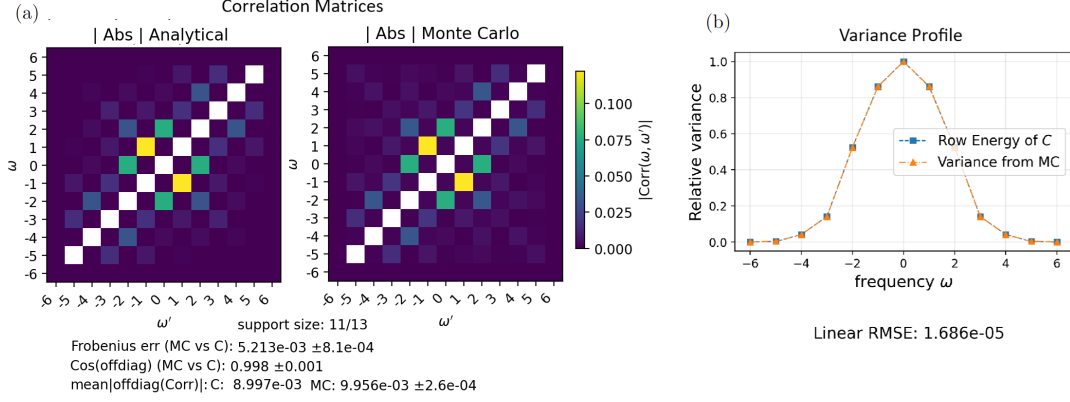


Figure 1: Architecture-derived coefficient statistics versus independent Monte Carlo (MC, 2×10^5 samples), for a three-qubit YZY ansatz with entanglers and $L = 2$ re-uploading layers [11]. (a) Absolute frequency–frequency correlations. (b) Variance profile from row energies of C and from MC estimates.

Spectral bias. The coefficient statistics of Equation (2) quantify how strongly an architecture drives each accessible frequency $\omega \in \Omega$. Weakly driven frequencies have small variance $\text{Var}_\theta[a_\omega]$, and the correlations distinguish frequencies that are fitted together through a shared parameter harmonic from frequencies that compete for the same trainable degree of freedom. Since these quantities are available in closed form before training, a circuit’s spectral bias can be assessed against the frequency content assumed present in the data and, conversely, circuits can be designed whose variance and correlation structure match a given target problem, rather than relying on generic ansatzes.

Trainability. The kernel factorisation of Equation (3) gives a pre-training diagnostic of whether the architecture supplies tangent directions relevant to the task. The effective rank of VC bounds the reachable tangent space, while the averaged kernel $C \text{diag}(\|k\|_2^2) C^\dagger$ shows where gradient capacity is concentrated. If projection to the input lattice lowers this rank, or if the kernel weight lies mostly in target-irrelevant directions, optimisation is restricted before it begins.

Circuit compression. Because the covariance and kernel both factor through C , gates or parameters whose removal changes $CC^\dagger$ or the harmonic QNTK only slightly should approximately preserve the circuit’s coefficient variances, frequency correlations, and reachable tangent directions, making them natural candidates for pruning. This suggests a route to analytical circuit compression within a chosen tolerance. On near-term hardware, reduced gate count and depth translate into shorter coherence-time requirements and less accumulated noise, so pre-training pruning methods informed by $CC^\dagger$ would have practical value.

Several extensions to this work follow naturally: multi-dimensional inputs (multi-index ω_i), parameter sharing (which enlarges K beyond $\{-1, 0, 1\}^m$ and alters the harmonic coupling), and non-commuting feature maps. A further extension is to generalise beyond QFMs to general PQCs, where the parameter harmonics decompose the circuit’s action at the channel level into fixed superoperator components, in place of input frequencies, i.e. $\Phi_\theta = \sum_k e^{i\theta} \mathcal{S}_k$ and expanding $\mathcal{S}_k$ into components indexed by α to yield a more general matrix $C_{\alpha k}$. More immediately, using this construction to explicitly improve PQC performance, either through circuit design or C -informed compression algorithms, is a planned next step.

In summary, the circuit harmonic matrix C turns architecture into an explicit spectral object: it determines coefficient statistics, frequency correlations, and training kernels for the QFM class of PQCs. This makes circuit-induced learning structure analytically accessible before training and provides a basis for comparing, selecting, analysing, and potentially compressing PQC architectures.

References

- [1] K. J. S. Campbell, L. Del Debbio, and P. Wallden, “Circuit Harmonic Matrices: A Spectral Framework for Quantum Machine Learning,” Apr. 2026. arXiv:2604.04292 [quant-ph].
- [2] A. Pérez-Salinas, A. Cervera-Lierta, E. Gil-Fuster, and J. I. Latorre, “Data re-uploading for a universal quantum classifier,” *Quantum*, vol. 4, p. 226, Feb. 2020. arXiv:1907.02085 [quant-ph].
- [3] M. Schuld, R. Sweke, and J. J. Meyer, “Effect of data encoding on the expressive power of variational quantum-machine-learning models,” *Physical Review A*, vol. 103, p. 032430, Mar. 2021.
- [4] B. Casas and A. Cervera-Lierta, “Multidimensional Fourier series with quantum circuits,” *Physical Review A*, vol. 107, p. 062612, June 2023.
- [5] J. Liu, F. Tacchino, J. R. Glick, L. Jiang, and A. Mezzacapo, “Representation Learning via Quantum Neural Tangent Kernels,” *PRX Quantum*, vol. 3, p. 030323, Aug. 2022.
- [6] E. Abedi, S. Beigi, and L. Taghavi, “Quantum Lazy Training,” *Quantum*, vol. 7, p. 989, Apr. 2023. arXiv:2202.08232 [quant-ph].
- [7] N. Shirai, K. Kubo, K. Mitarai, and K. Fujii, “Quantum tangent kernel,” *Physical Review Research*, vol. 6, Aug. 2024.
- [8] D. Gottesman, “The Heisenberg Representation of Quantum Computers,” July 1998. arXiv:quant-ph/9807006.
- [9] S. Aaronson and D. Gottesman, “Improved Simulation of Stabilizer Circuits,” *Physical Review A*, vol. 70, p. 052328, Nov. 2004. arXiv:quant-ph/0406196.
- [10] P. Rall, “Simulation of qubit quantum circuits via Pauli propagation,” *Physical Review A*, vol. 99, no. 6, 2019.
- [11] M. Strobl, M. E. Sahin, L. v. d. Horst, E. Kuehn, A. Streit, and B. Jaderberg, “Fourier Fingerprints of Ansatzes in Quantum Machine Learning,” Aug. 2025. arXiv:2508.20868 [quant-ph].
- [12] H. Mhiri, L. Monbroussou, M. Herrero-Gonzalez, S. Thabet, E. Kashefi, and J. Landman, “Constrained and Vanishing Expressivity of Quantum Fourier Models,” *Quantum*, vol. 9, p. 1847, Sept. 2025. arXiv:2403.09417 [quant-ph].