

Fourier Symmetrization for Geometric Quantum Machine Learning

Letao Wang¹, Abdel Lisser^{1,2}, Sreejith Sreekumar¹, and Zeno Toffano¹

¹Laboratory of Signals and Systems, CentraleSupélec, CNRS, Paris-Saclay University, 3 Rue Joliot Curie, Gif-sur-Yvette, 91190, Île-de-France, France

²Fédération de Mathématiques de CentraleSupélec, CentraleSupélec, CNRS, Paris-Saclay University, 3 Rue Joliot Curie, Gif-sur-Yvette, 91190, Île-de-France, France

Symmetry is a structural prior in scientific learning, and encoding it into quantum models can improve data efficiency, generalization, and trainability. This idea has motivated geometric quantum machine learning (GQML) [1–3]. We bring this viewpoint to data re-uploading models through the quantum Fourier model (QFM) formalism [4, 5]. For L encoding layers, the QFM unitary operation is

$$U(\mathbf{x}, \boldsymbol{\theta}) = W^{(L+1)}(\boldsymbol{\theta}) \prod_{l=1}^L \left(E^{(l)}(\mathbf{x}) W^{(l)}(\boldsymbol{\theta}) \right), \quad (1)$$

This unitary alternates data encoding layers $E(\mathbf{x})$ and trainable layers $W(\boldsymbol{\theta})$. Given the initial state $|0\rangle$ and observable $\hat{O}$, the induced hypothesis function is

$$f(\mathbf{x}, \boldsymbol{\theta}) = \langle 0 | U^\dagger(\mathbf{x}, \boldsymbol{\theta}) \hat{O} U(\mathbf{x}, \boldsymbol{\theta}) | 0 \rangle = \sum_{\boldsymbol{\omega} \in \Omega} c_{\boldsymbol{\omega}}(\boldsymbol{\theta}) e^{-i\mathbf{x} \cdot \boldsymbol{\omega}}. \quad (2)$$

Here $\mathbf{x} \in \mathbb{R}^m$ is the m -dimensional input, $\boldsymbol{\theta}$ denotes trainable parameters, $\Omega \subset \mathbb{R}^m$ is the finite frequency spectrum, and $c_{\boldsymbol{\omega}}(\boldsymbol{\theta})$ is the coefficient of $\boldsymbol{\omega} \in \Omega$. Fourier coefficient variances can decay exponentially with the number of qubits, a phenomenon known as *vanishing expressivity* [5]. We show that symmetry can help mitigate this phenomenon by transferring input invariance to the Fourier coefficient level, where symmetry-related coefficients are aggregated.

Twirling is a standard way to enforce symmetry in GQML by averaging gates, operators, or channels over a group representation [2, 3]. We use the same averaging idea at the level of the QFM hypothesis function. Let $\mathcal{S}$ be a finite group and let $V : \mathcal{S} \rightarrow O(m)$, $s \mapsto V_s$, be an orthogonal group representation, where $O(m)$ is the orthogonal group on $\mathbb{R}^m$. Define the *averaging operator* $\mathcal{A}$ by

$$(\mathcal{A}f)(\mathbf{x}, \boldsymbol{\theta}) := \frac{1}{|\mathcal{S}|} \sum_{s \in \mathcal{S}} f(V_s[\mathbf{x}], \boldsymbol{\theta}). \quad (3)$$

Here $|\mathcal{S}|$ is the number of elements in $\mathcal{S}$. An $\mathcal{S}$ -invariant hypothesis function satisfies $f(V_s[\mathbf{x}], \boldsymbol{\theta}) = f(\mathbf{x}, \boldsymbol{\theta})$ for all $s \in \mathcal{S}$. The identity $V_s[\mathbf{x}] \cdot \boldsymbol{\omega} = \mathbf{x} \cdot V_{s^{-1}}[\boldsymbol{\omega}]$ lets the same representation organize the frequency spectrum. Assume that $V_s[\boldsymbol{\omega}] \in \Omega$ for all $s \in \mathcal{S}$. Frequencies related by V form an orbit. From each orbit choose one representative frequency, called an *orbit frequency* $\boldsymbol{\nu}$, and write the orbit as

$$[\boldsymbol{\nu}] = \{V_s[\boldsymbol{\nu}] \mid s \in \mathcal{S}\}. \quad (4)$$

The set of all orbit frequencies is the orbit frequency spectrum $\tilde{\Omega}$. Figure 1 illustrates this notation.

Proposition 1 (Fourier Orbit Decomposition). *Given $f(\mathbf{x}, \boldsymbol{\theta})$ in Eq. (2), applying the averaging operator $\mathcal{A}$ yields an $\mathcal{S}$ -invariant function:*

$$(\mathcal{A}f)(\mathbf{x}, \boldsymbol{\theta}) = \sum_{\boldsymbol{\nu} \in \tilde{\Omega}} a_{\boldsymbol{\nu}}(\boldsymbol{\theta}) \phi_{[\boldsymbol{\nu}]}(\mathbf{x}), \quad (5)$$

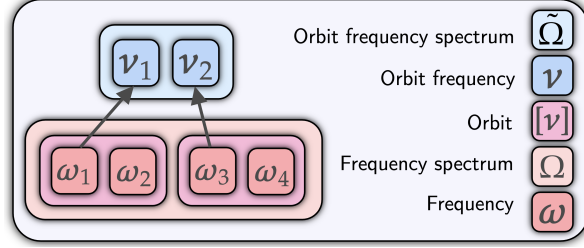


Figure 1: Illustration of the orbit notation. ω_1 and ω_2 belong to one orbit, while ω_3 and ω_4 belong to another orbit. We choose ν_1 and ν_2 as orbit frequencies representing ω_1 and ω_3 , respectively, so the two orbits are denoted by $[\nu_1]$ and $[\nu_2]$. The orbit frequencies ν_1 and ν_2 together form the orbit frequency spectrum $\tilde{\Omega}$.

where the symmetrized coefficient is

$$a_{\nu}(\theta) := \sum_{\omega \in [\nu]} c_{\omega}(\theta), \quad (6)$$

and the normalized orbit basis is

$$\phi_{[\nu]}(\mathbf{x}) = \frac{1}{|[\nu]|} \sum_{\tilde{\omega} \in [\nu]} e^{-i\mathbf{x} \cdot \tilde{\omega}}. \quad (7)$$

Proposition 1 shows that invariance groups the original frequencies into orbit basis functions and aggregates the coefficients in each orbit into a_{ν} . If $f(\mathbf{x}, \theta)$ is already $\mathcal{S}$ -invariant, then $f(\mathbf{x}, \theta) = \frac{1}{|\mathcal{S}|} \sum_{s \in \mathcal{S}} f(V_s[\mathbf{x}], \theta)$. Therefore, Proposition 1 also shows that any $\mathcal{S}$ -invariant QFM under this orthogonality assumption on V admits a Fourier orbit decomposition. Now we show that symmetry can help mitigate this phenomenon by increasing coefficient variance.

Theorem 1 (Symmetrized Coefficient Variance under 2-designs). *Consider the QFM in Eq. (1), whose hypothesis function has the Fourier coefficients $c_{\omega}(\theta)$ in Eq. (2), and let $a_{\nu}(\theta)$ be the symmetrized coefficients defined in Eq. (6). If each trainable layer forms an independent unitary 2-design, then for any orbit frequency ν in the orbit frequency spectrum $\tilde{\Omega}$,*

$$\text{Var}_{\theta}(a_{\nu}) = \sum_{\omega \in [\nu]} \text{Var}_{\theta}(c_{\omega}) = |[\nu]| \cdot \overline{\text{Var}_{\theta}(c_{\omega})}, \quad (8)$$

where $\overline{\text{Var}_{\theta}(c_{\omega})}$ denotes the mean variance of the frequencies within the orbit $[\nu]$.

Figure 2 illustrates this coefficient-level effect. The pink curve for $\text{Var}(a_{\nu})$ coincides with the purple curve for $\sum_{\omega \in [\nu]} \text{Var}(c_{\omega})$. The overlap between the pink and purple curves numerically confirms Theorem 1.

For realistic ansatzes, the trainable layers may only approximate 2-designs. We state the bound for the single-uploading setting $L = 1$. The circuit then has two trainable layers, one on each side of $E(\mathbf{x})$. Write $E(\mathbf{x})|j\rangle = e^{-i\mathbf{x} \cdot \boldsymbol{\lambda}_j} |j\rangle$ for $j \in [1, d]$, where d is the Hilbert-space dimension, $[1, d] = \{1, \dots, d\}$, $|j\rangle$ is the j -th basis state, and $\boldsymbol{\lambda}_j \in \mathbb{R}^m$ is the corresponding eigenvalue vector. For an orbit $[\nu]$, define

$$\mathfrak{R}_{[\nu]} := \{(j, j') \in [1, d]^2 : \boldsymbol{\lambda}_j - \boldsymbol{\lambda}_{j'} \in [\nu]\}, \quad R := |\mathfrak{R}_{[\nu]}|. \quad (9)$$

We call R the total redundancy of the orbit $[\nu]$, following the frequency-generator notation of Mhiri et al. [5]. The single-frequency generator $\mathfrak{R}(\omega)$ is obtained by replacing $[\nu]$ in Eq. (9) with $\{\omega\}$.

Theorem 2 (Symmetrized Coefficient Variance under Approximate 2-designs). *Consider the n -qubit QFM in Eq. (1) with $L = 1$, whose hypothesis function has the Fourier series in Eq. (2). Suppose the two trainable layers form independent ε -approximate 2-designs in spectral norm, with approximation error $\varepsilon > 0$. For any frequency orbit $[\nu]$ excluding the zero frequency $\mathbf{0}$, let R be its total redundancy as defined in Eq. (9). Let $\hat{O}$ be the observable, let $\|\cdot\|_F$ denote the Frobenius norm, let $\text{Var}_{2\text{-design}}$ denote*

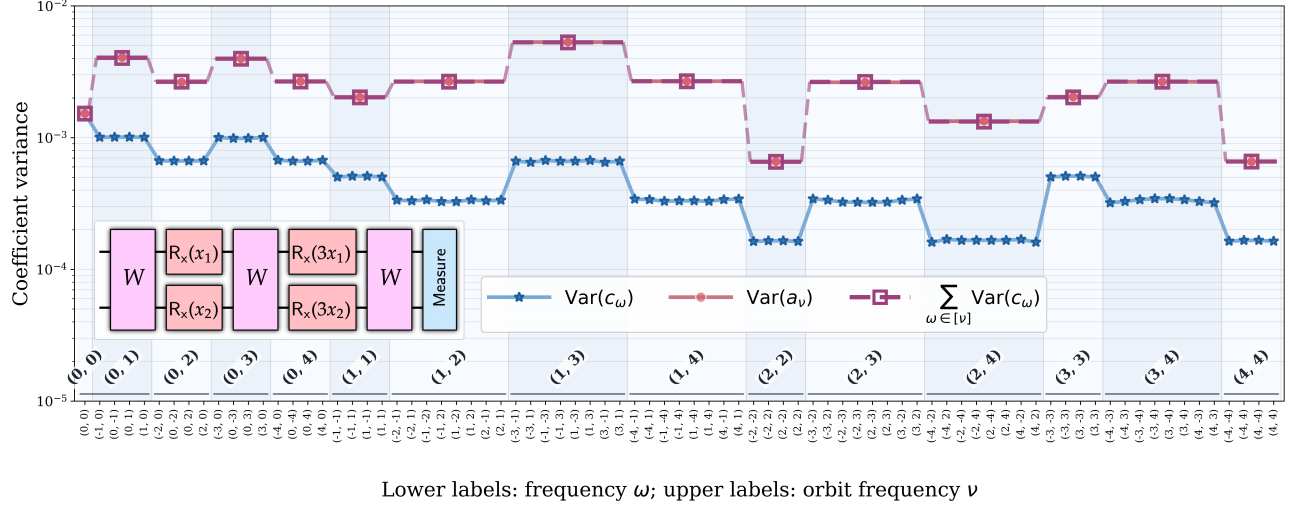


Figure 2: Fourier coefficient variances reorganized into frequency orbits with respect to the hyperoctahedral group under the 2-design assumption. The x-axis uses two rows of labels: the lower row lists the original frequencies ω , and the upper row labels each orbit by its orbit frequency ν . For example, under the coordinate permutations and independent sign flips in the hyperoctahedral group, the orbit frequency $(0,1)$ represents the orbit containing frequencies $(-1,0)$, $(0,-1)$, $(0,1)$, and $(1,0)$. The blue curve shows the variance of each coefficient c_ω , the pink curve shows the variance of each symmetrized coefficient a_ν , and the purple curve shows the sum $\sum_{\omega \in [\nu]} \text{Var}(c_\omega)$ over the same orbit. The circuit used in this numerical illustration is shown in the lower-left inset, where each trainable layer W denotes five repetitions of the Strongly Entangling Ansatz and the measurement uses the global observable.

the variance obtained by replacing the approximate trainable layers with exact unitary 2-designs, and let $\mathcal{O}(\cdot)$ denote asymptotic upper-bound notation. Then

$$|\text{Var}_\theta(a_\nu) - \text{Var}_{2\text{-design}}(a_\nu)| \in \mathcal{O}\left(\varepsilon \|\hat{O}\|_F^2 \left(\frac{\sqrt{R}}{4^n} + \varepsilon\right)\right). \quad (10)$$

Theorem 2 makes Theorem 1 stable under approximate 2-design errors. Compared with the approximate 2-design bound of Mhiri et al. [5, Theorem 7], this result gives a tighter error estimate. In the single-frequency case, the previous redundancy-dependent term $\varepsilon^2 \|\hat{O}\|_F^2 |\Re(\omega)|$ is replaced by $(\varepsilon^2 + \varepsilon^4) \|\hat{O}\|_F^2$. For symmetry groups that collect exponentially many frequencies into the same orbit, the resulting advantage can be exponentially large.

The averaging operator can also reshape the hypothesis space. Chebyshev polynomials provide useful bases for solving partial differential equations, but direct Chebyshev feature maps as in [6] can introduce unwanted non-polynomial terms. Averaging removes these terms and yields a pure multivariate Chebyshev polynomial basis, which can lead to better PDE solvers.

For implementation, conventional symmetry constructions often use gate symmetrization [2] or circuit-based twirling with the induced quantum representation [3], which can require expensive offline construction, additional ancilla qubits, or extra representation-circuit depth. We propose randomized encoding: instead of implementing the induced quantum representation on the quantum device, it samples the encoding layer $E(V_s[\mathbf{x}])$, where E is the original encoding layer. The estimator is unbiased for $(\mathcal{A}f)(\mathbf{x}, \theta)$ and adds no qubits or quantum depth.

We further perform two-dimensional quantum physics-informed neural network [6, 7] experiments for the Poisson equation and the Hamilton–Jacobi equation. These experiments show that the symmetry-induced Chebyshev construction can provide a clear training advantage for solving partial differential equations.

References

- [1] Martín Larocca, Frédéric Sauvage, Faris M. Sbahi, Guillaume Verdon, Patrick J. Coles, and M. Cerezo. Group-invariant quantum machine learning. *PRX Quantum*, 3:030341, Sep 2022. doi: 10.1103/PRXQuantum.3.030341. URL <https://link.aps.org/doi/10.1103/PRXQuantum.3.030341>.
- [2] Johannes Jakob Meyer, Marian Mularski, Elies Gil-Fuster, Antonio Anna Mele, Francesco Arzani, Alissa Wilms, and Jens Eisert. Exploiting symmetry in variational quantum machine learning. *PRX Quantum*, 4:010328, Mar 2023. doi: 10.1103/PRXQuantum.4.010328. URL <https://link.aps.org/doi/10.1103/PRXQuantum.4.010328>.
- [3] Quynh T. Nguyen, Louis Schatzki, Paolo Braccia, Michael Ragone, Patrick J. Coles, Frédéric Sauvage, Martín Larocca, and M. Cerezo. Theory for equivariant quantum neural networks. *PRX Quantum*, 5:020328, May 2024. doi: 10.1103/PRXQuantum.5.020328. URL <https://link.aps.org/doi/10.1103/PRXQuantum.5.020328>.
- [4] Maria Schuld, Ryan Sweke, and Johannes Jakob Meyer. Effect of data encoding on the expressive power of variational quantum-machine-learning models. *Phys. Rev. A*, 103:032430, Mar 2021. doi: 10.1103/PhysRevA.103.032430. URL <https://link.aps.org/doi/10.1103/PhysRevA.103.032430>.
- [5] Hela Mhiri, Leo Monbroussou, Mario Herrero-Gonzalez, Slimane Thabet, Elham Kashefi, and Jonas Landman. Constrained and Vanishing Expressivity of Quantum Fourier Models. *Quantum*, 9:1847, September 2025. ISSN 2521-327X. doi: 10.22331/q-2025-09-03-1847. URL <https://doi.org/10.22331/q-2025-09-03-1847>.
- [6] Oleksandr Kyriienko, Annie E. Paine, and Vincent E. Elfving. Solving nonlinear differential equations with differentiable quantum circuits. *Phys. Rev. A*, 103:052416, May 2021. doi: 10.1103/PhysRevA.103.052416. URL <https://link.aps.org/doi/10.1103/PhysRevA.103.052416>.
- [7] Letao Wang, Abdel Lisser, Sreejith Sreekumar, and Zeno Toffano. Learning PDEs for portfolio optimization with quantum physics-informed neural networks, 2026. URL <https://arxiv.org/abs/2604.03346>.