

Harmoniq: Efficient Quantum Data Augmentation Inspired by Quantum Harmonic Analysis

Extended abstract for QTML 2026

1 Motivation

Recent years have witnessed growing interest in quantum machine learning [1, 2, 3], motivated by the ability of quantum systems to process information in high-dimensional Hilbert spaces. This has led to the development of various QML paradigms, including quantum kernel methods [4, 5] and variational quantum neural networks [6, 7], often inspired by successful classical machine learning architectures [8, 9, 10]. However, these approaches face significant challenges related to trainability and scalability [11, 7]. Moreover, comparatively little attention has been paid to data-centric techniques such as data augmentation [12, 13], which play a crucial role in classical ML.

In this work, we draw inspiration from the field of Quantum Harmonic Analysis (QHA) [14, 15, 16] to propose *Harmoniq*: a novel quantum data augmentation protocol, which is non-variational, efficiently implementable on a quantum computer, and modular. It relies on the stochastic implementation of the Weyl-Heisenberg channel, which smooths the covariance operator of the dataset, thus allowing for better learning from the data. We demonstrate the approach in the context of signal denoising and show that it is particularly effective in the small-sample regime.

2 Background

Quantum harmonic analysis. Classical harmonic analysis studies continuous signals in $L^2(\mathbb{R})$ using time- and frequency-shift operators. Their composition yields the time-frequency (TF) shift operator $W(x, z)$ for $(x, z) \in \mathbb{R}^2$: $W(x, z)g(t) = e^{2\pi i z t} g(t - x)$.

Quantum Harmonic Analysis (QHA) [14] lifts these concepts to the operator level. TF shifts act on a trace-class operator O via conjugation, allowing us to define its convolution with an integrable weight function λ :

$$\lambda \star O = \int_{\mathbb{R}^2} \lambda(x, z) W_{(x, z)} O W_{(x, z)}^\dagger dx dz. \quad (1)$$

Crucially, this convolution acts as a smoothing operation that suppresses unstructured noise. When O represents a dataset’s covariance matrix, $\lambda \star O$ yields the covariance of an augmented dataset [16]. This induced regularity forces a faster eigenvalue decay, producing a sharper spectral cutoff for Principal Component Analysis.

Weyl–Heisenberg matrices. On n qubits ($d = 2^n$), the finite-dimensional counterparts of the TF shifts are the Weyl–Heisenberg (generalized Pauli) [17, 18, 19] matrices given by

$$W_{(x, z)} = e^{-i\pi x z / d} Z_d^z X_d^x, \quad (x, z) \in \{0, \dots, d-1\}^{\times 2}, \quad (2)$$

with clock and shift operators given by $Z_d = \sum_{k=0}^{d-1} \omega_d^k |k\rangle\langle k|$ and $X_d = \sum_{k=0}^{d-1} |k+1\rangle\langle k|$ where $\omega_d = \exp(2\pi i / d)$ is the d -th root of unity. With this, the operator convolution in discrete dimensions becomes

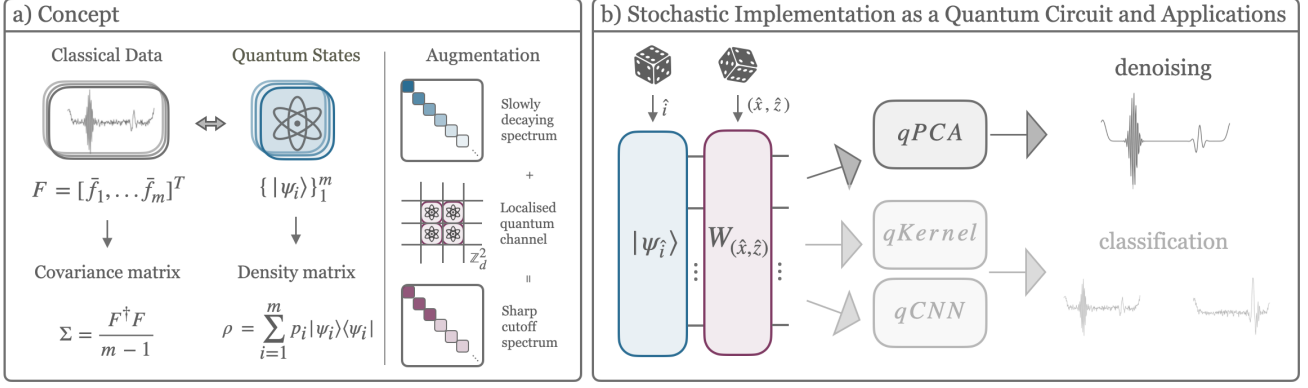


Figure 1: *Concept and Implementation of Harmoniq.* (a) Classical data signals are normalized into pure quantum states, allowing the covariance matrix to be treated as a density matrix. The augmentation via the localised Weyl-Heisenberg quantum channel results in an eigenvalue spectrum with a sharper cutoff. (b) The augmentation can be implemented stochastically by drawing pure quantum states and Weyl-Heisenberg matrices. While applied here together with qPCA for denoising, our modular framework supports diverse methods and applications.

a *quantum channel* on a density matrix ρ ,

$$\mathcal{P}_\Lambda(\rho) = \sum_{(x,z) \in \Omega} \lambda_{(x,z)} W_{(x,z)} \rho W_{(x,z)}^\dagger, \quad \lambda_{(x,z)} \geq 0 \quad \text{and} \quad \sum_{(x,z) \in \Omega} \lambda_{(x,z)} = 1, \quad (3)$$

Operationally, this channel corresponds to a probabilistic mixture of d^2 unitary evolutions, where each unitary is a Weyl-Heisenberg matrix $W_{(x,z)}$ and the associated probability weight is $\lambda_{(x,z)}$.

3 The framework

We introduce an efficient quantum data augmentation protocol. Inspired by quantum harmonic analysis, we frame augmentation as a localized quantum channel acting on the data's covariance matrix.

Let $F \in \mathbb{C}^{m \times d}$ denote the data matrix whose (conjugated) rows are the centered data vectors $\bar{f}_i \in \mathbb{C}^d$. Studying the eigenspectrum of the data covariance matrix $\Sigma = \frac{1}{m-1} F^\dagger F$ reveals the underlying structure of the data. The normalized data covariance matrix $\Sigma/\text{tr}(\Sigma)$ can be encoded as a mixed d -dimensional quantum state [20] $\rho = \sum_{i=1}^m p_i |\psi_i\rangle\langle\psi_i|$ where $|\psi_i\rangle$ is the quantum state corresponding to an amplitude encoding of the normalized $\bar{f}_i$, $|\psi_i\rangle = \bar{f}_i / \|\bar{f}_i\|_2$ and every vector enters with a probability proportional to its norm, $p_i = \|\bar{f}_i\|_2^2 / \sum_{j=1}^m \|\bar{f}_j\|_2^2$. Operationally, the density matrix admits a stochastic realization: it can be viewed as the ensemble average over pure states obtained by randomly sampling data points according to their energy and encoding them as quantum states, $|\hat{\psi}\rangle = \bar{f}_i / \|\bar{f}_i\|_2$ with $\Pr[\hat{i} = i] = p_i$.

To augment the data, we pass this mixed state through the Weyl-Heisenberg channel, Eq. (3), which applies phase-space displacements sampled from a specified distribution. In the language of QHA, this is an operator-function convolution, Eq. (1). In machine learning terms, it acts as a covariance-level data augmentation. If we sampled these displacements uniformly across all possible unitaries, we would destroy the data structure by twirling ρ into the fully mixed state. Instead, we restrict the shifts to a tightly bounded *augmentation window* $\Omega \subseteq \mathbb{Z}_d^2$ ($|\Omega| \ll d^2$), directly applying the principle of *localized augmentation* [15, 16].

Since constructing this entire channel deterministically on quantum hardware is impractical, we propose importance sampling to implement it stochastically. Given a pure state $|\hat{\psi}\rangle$, we apply:

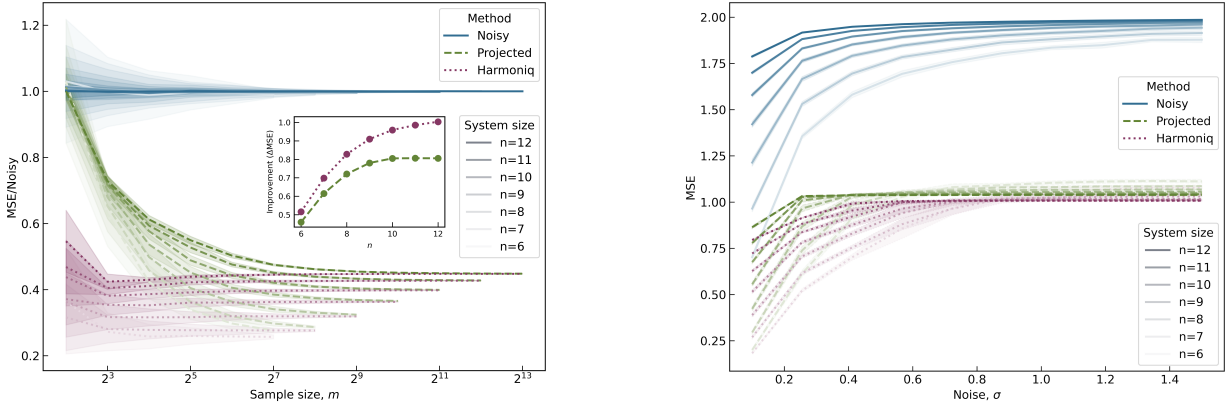
$$|\hat{\psi}\rangle \rightarrow W_{(\hat{x}, \hat{z})} |\hat{\psi}\rangle \quad \text{with} \quad \Pr[(\hat{x}, \hat{z}) = (x, z)] = \lambda_{(x,z)}. \quad (4)$$

Averaging over these individual unitary evolutions recovers the intended channel $\mathcal{P}_\Lambda(\rho)$ in expectation.

The second major contribution is demonstrating that any Weyl-Heisenberg matrix in dimension $d = 2^n$ can be efficiently implemented on a quantum computer with an $\mathcal{O}(n^2)$ deep circuit. This is due to the Z_d^z component reducing entirely to only one layer of single-qubit phase rotations. The shift X_d^x can be implemented by changing to the Fourier basis via the Quantum Fourier Transform (QFT), applying the equivalent Z_d^x rotations, and transforming back to the computational basis with the inverse QFT. Consequently, the overall circuit depth is dominated by the implementation of the QFT and inverse QFT.

4 Application

We demonstrate *Harmoniq* for signal denoising via (Quantum) Principal Component Analysis (PCA) [21, 22, 23] on synthetic, time-frequency localized signals corrupted by Gaussian noise. Simulations up to 12 qubits reveal that augmenting the data consistently improves denoising over standard PCA projection alone. This advantage is highly pronounced in the low-sample regime across diverse noise levels, proving effectively reinforces latent structures when data is scarce, see Fig. 2.



(a) Normalized Mean Squared Error (MSE) vs. sample size. *Harmoniq* significantly reduces MSE compared to standard projection, particularly in the scarce-data regime. *Inset*: Absolute MSE *improvement* over the noisy baseline scales favorably with system size, n .

(b) MSE vs. noise level at a fixed sample size ($m = 100$). *Harmoniq* consistently outperforms standard projection across all noise regimes, eventually saturating near the theoretical limit at high noise.

Figure 2: *Performance and Scalability of Harmoniq for denoising applications.* In both panels, our work is denoted by purple dotted lines, color gradients indicate system size, and all data points are averaged over 100 independent instances.

5 Conclusion and outlook

We introduced *Harmoniq*: a parameter-free quantum data augmentation protocol inspired by Quantum Harmonic Analysis that can be implemented as a stochastic channel with only quadratic circuit depth, assuming access to state preparation oracles. We demonstrated its effectiveness in a signal denoising task in the data scarcity regime. Natural next steps include task-adaptive choices of the augmentation window Ω , robustness analyses under realistic noise models, and extensions to broader classes of harmonic transforms and non-commutative signal representations. Ultimately, we see this work as paving the way to a richer theoretical foundation for data processing on quantum devices, departing from variational circuits.

References

- [1] Maria Schuld, Ilya Sinayskiy, and Francesco Petruccione. An introduction to quantum machine learning. *Contemporary Physics*, 56(2):172–185, 2015.
- [2] Maria Schuld and Francesco Petruccione. Supervised learning with quantum computers. *Springer*, 2018.
- [3] Hsin-Yuan Huang, Michael Broughton, Jordan Cotler, Sitan Chen, Jerry Li, Masoud Mohseni, Hartmut Neven, Ryan Babbush, Richard Kueng, John Preskill, and Jarrod R. McClean. Quantum advantage in learning from experiments. *Science*, 376(6598):1182–1186, 2022.
- [4] Maria Schuld and Nathan Killoran. Quantum machine learning in feature hilbert spaces. *Physical Review Letters*, 122(4):040504, 2019.
- [5] Vojtěch Havlíček et al. Supervised learning with quantum-enhanced feature spaces. *Nature*, 567(7747):209–212, 2019.
- [6] Edward Farhi and Hartmut Neven. Classification with quantum neural networks on near term processors. *arXiv preprint arXiv:1802.06002*, 2018.
- [7] M. Cerezo, A. Arrasmith, R. Babbush, S. Benjamin, S. Endo, K. Fujii, J. McClean, K. Mitarai, X. Yuan, L. Cincio, and P. Coles. Variational quantum algorithms. *Nature Reviews Physics*, 3:625–644, 2021.
- [8] Patrick Rebentrost, Masoud Mohseni, and Seth Lloyd. Quantum support vector machine for big data classification. *Physical Review Letters*, 113(13):130503, 2014.
- [9] Seth Lloyd and Christian Weedbrook. Quantum generative adversarial learning. *Phys. Rev. Lett.*, 121:040502, Jul 2018.
- [10] Iris Cong, Soonwon Choi, and Mikhail D Lukin. Quantum convolutional neural networks. *Nature Physics*, 15(12):1273–1278, 2019.
- [11] Jarrod McClean, Sergio Boixo, Vadim Smelyanskiy, Ryan Babbush, and Hartmut Neven. Barren plateaus in quantum neural network training landscapes. *Nature Communications*, 9:4812, 2018.
- [12] Connor Shorten and Taghi Khoshgohfar. A survey on image data augmentation for deep learning. *Journal of Big Data*, 6:60, 2019.
- [13] Ekin Cubuk, Barret Zoph, Dandelion Mane, Vijay Vasudevan, and Quoc Le. Autoaugment: Learning augmentation policies from data. In *CVPR*, 2019.
- [14] R. Werner. Quantum harmonic analysis on phase space. *J. Math. Phys.*, 25(5):1404–1411, 1984.
- [15] Monika Dörfler, Franz Luef, and Eirik Skrettingland. Local structure and effective dimensionality of time series data sets. *Applied and Computational Harmonic Analysis*, 73:101692, 2024.
- [16] Monika Dörfler, Franz Luef, and Henry McNulty. Quantum harmonic analysis and the structure in data: Augmentation, 2025.
- [17] K. S. Gibbons, M. J. Hoffman, and W. K. Wootters. Discrete phase space based on finite fields. *Physical Review A*, 70(6):062101, 2004.

- [18] Martin Ringbauer, Michael Meth, Lukas Postler, Roman Stricker, Rainer Blatt, Philipp Schindler, and Thomas Monz. A universal qudit quantum processor with trapped ions. *Nature Physics*, 18(9):1053–1057, July 2022.
- [19] Nina Brandl, Mykyta Cherniak, Johannes Kofler, and Richard Kueng. Quickqudits: A framework for efficient simulation of noisy qudit clifford circuits via an extended stabilizer tableau formalism, 2026.
- [20] Max Hunter Gordon, M. Cerezo, Lukasz Cincio, and Patrick J. Coles. Covariance matrix preparation for quantum principal component analysis. *PRX Quantum*, 3:030334, Sep.
- [21] Ian T. Jolliffe. *Principal Component Analysis*. Springer, 2nd edition, 2002.
- [22] Seth Lloyd, Masoud Mohseni, and Patrick Rebentrost. Quantum principal component analysis. *Nature physics*, 10(9):631–633, 2014.
- [23] Tao Xin, Liangyu Che, Cheng Xi, Amandeep Singh, Xinfang Nie, Jun Li, Ying Dong, and Dawei Lu. Experimental quantum principal component analysis via parametrized quantum circuits. *Phys. Rev. Lett.*, 126:110502, Mar 2021.