Highlights

Magnitude-Dependent Quantum Advantage in Forbush Decrease Detection: A Quantum Kernel SVM Benchmark

D. Sierra-Porta

- Quantum kernel SVMs outperform classical RBF above a 4% FD magnitude threshold.
- Quantum advantage reaches 100% of configurations at Magn $\geq$ 5% (ΔAUC = +0.074).
- FD magnitude is a key predictor of quantum classification performance in space weather.

Magnitude-Dependent Quantum Advantage in Forbush Decrease Detection: A Quantum Kernel SVM Benchmark

D. Sierra-Porta[a]

[a] *Universidad Tecnológica de Bolívar. Grupo de Investigación Gravitación y Matemática Aplicada - GIGMA & Grupo de Investigación Física Aplicada y Procesamiento de Imágenes y Señales - FAPIS., Parque Industrial y Tecnológico Carlos Vélez Pombo Km 1 Vía Turbaco, Cartagena de Indias, 130010, Bolívar, Colombia*

Abstract

Forbush decreases (FDs)—transient reductions in galactic cosmic ray intensity driven by interplanetary coronal mass ejections—are key observables in space weather monitoring, yet their automated detection from multivariate solar wind and neutron monitor time series remains a challenging classification problem. Here we report a systematic benchmark of quantum kernel support vector machines (QKSVM) for FD detection, in which the FD magnitude threshold emerges as the governing factor separating two distinct classification regimes. Using 2,971 confirmed events from the Forbush Effects and Interplanetary Disturbances (FEID) catalog, combined with hourly OMNI solar wind parameters—including interplanetary magnetic field (IMF) components, solar wind speed, proton density, proton temperature, and the Kp and Dst geomagnetic indices—and galactic cosmic ray count rates from the Jungfraujoch neutron monitor station (JUNG, NMDB), we construct a balanced FD versus quiet-time classification dataset and extract 121 statistical features across eleven physical channels. A `ZZFeatureMap` quantum kernel with 4–8 qubits is benchmarked against a classical radial basis function (RBF) SVM across 180 experimental configurations spanning FD magnitude thresholds of 0–7%, circuit depths of 1–3 repetitions, and quantum training sizes of 50–250 samples. We find that below a magnitude threshold of 4%, the classical kernel consistently outperforms the quantum alternative (mean ΔAUC = -0.100 at `min_magn` = 0%). Above this threshold, the relationship inverts: at `min_magn` $\geq$ 4% the quantum kernel achieves positive mean ΔAUC in 72% of configurations, rising to 100% of configurations at `min_magn` $\geq$ 5% (mean ΔAUC = $+0.074$, peak AUC = 1.000 with 4–8 qubits), indicating that the entanglement structure of the `ZZFeatureMap` captures non-linear correlations between IMF dynamics and cosmic ray modulation that the RBF kernel cannot represent. The magnitude threshold of 4% thus constitutes a physically interpretable boundary between a noise-dominated regime where classical methods suffice and a signal-rich regime where quantum kernels provide measurable and statistically significant advantage ($p < 0.001$, Wilcoxon signed-rank test). These results establish FD magnitude as a key predictor of quantum classification performance, and suggest that near-term quantum machine learning applications in heliophysics should preferentially target high-amplitude space weather events.

Keywords: Forbush decrease, quantum kernels, machine learning, classification task, geomagnetic indices
2020 MSC: 81P68, 68T05, 85A35, 62H30, 68Q12

1. Introduction

Forbush decreases (FDs) are among the most well-characterized transient phenomena in the heliosphere, manifesting as sudden reductions in the flux of galactic cosmic rays (GCRs) observed at ground level by neutron monitor networks (Forbush, 1938). These events are predominantly associated with the passage of interplanetary coronal mass ejections (ICMEs) and their preceding shocks, which modulate the GCR flux through enhanced magnetic field turbulence and the shielding effect of ICME magnetic structures (Cane, 2000; Richardson and Cane, 2011). Because FDs encode information about the magnetic topology and dynamical state of interplanetary disturbances, their detection and characterization constitute a central problem in space weather monitoring and GCR transport physics (Belov, 2008).

The automated identification of FDs from continuous neutron monitor time series is, however, a non-trivial classification task. FD amplitudes span nearly two orders of magnitude—from sub-percent modulations indistinguishable from diurnal variations to decreases exceeding 20% in major events—and their onset morphology ranges from abrupt shock-associated drops to gradual, diffusion-driven depressions (Melkumyan et al., 2023, 2024; Richardson and Cane, 2011; Belov et al., 2001). Classical threshold-based detection methods (Lockwood, 1971; Dumbović et al., 2024) and rule-based algorithms struggle with weak events and with the high background variability characteristic of solar maximum periods. The principal difficulty is that weak FDs occupy a regime where the signal-to-noise ratio in individual station data is insufficient for reliable automated discrimination, even when multivariate solar wind context is available. These limitations have motivated a growing body of work applying supervised machine learning to FD detection and space weather forecasting more broadly (Campo-

*Corresponding author:
Email address: `dporta@utb.edu.co` (D. Sierra-Porta)

reale, 2019; Bobra and Couvidat, 2015; Ye et al., 2025). Support vector machines (SVMs), random forests, and deep neural networks have all been applied to classification problems in heliophysics, generally achieving strong performance on well-defined event catalogs.

Despite this progress, the fundamental question of whether quantum machine learning (QML) methods can offer any advantage over classical approaches for space weather classification problems has not been addressed. Quantum kernel methods, in particular, have attracted considerable theoretical and experimental interest as a potential route to quantum advantage in supervised learning (Schuld and Killoran, 2019; Havlíček et al., 2019). In this framework, classical data vectors $\mathbf{x}$ are encoded into quantum states $|\phi(\mathbf{x})\rangle$ via a parameterized quantum circuit—the feature map—using *angle encoding*, in which each feature component x_j is mapped to a rotation angle within the circuit gates, so that the full feature vector is embedded into the amplitudes of a multi-qubit quantum state—and the inner product $|\langle\phi(\mathbf{x})|\phi(\mathbf{x}')\rangle|^2$ defines a kernel function that replaces the classical RBF or polynomial kernels in a standard SVM (Liu et al., 2021). The expressibility of such kernels grows exponentially with the number of qubits, raising the possibility that quantum feature maps can capture correlations in physical data that are inaccessible to polynomial-time classical kernels (Schuld, 2021; Huang et al., 2021). However, empirical demonstrations of genuine quantum advantage on real-world datasets remain scarce and contested (Kübler et al., 2021; Thanasilp et al., 2024), underscoring the need for careful, domain-specific benchmarks. A self-contained introduction to the quantum computing concepts underlying this framework is provided in Appendix Appendix A.

1.1. Related Work

1.1.1. Machine Learning for Space Weather and Forbush Decrease Detection

The application of supervised machine learning to space weather problems has grown substantially over the past decade, driven by the increasing availability of multi-instrument observational datasets and the computational maturity of modern ML frameworks (Camporeale, 2019). Within solar physics, SVMs and random forests have been applied to solar flare prediction from photospheric magnetograms (Bobra and Couvidat, 2015), while deep learning architectures have been employed for solar wind forecasting (Upendran et al., 2020) and geomagnetic storm classification (Ye et al., 2025). A comprehensive review of these developments is provided by Camporeale (2019), who identifies classification under class imbalance and the physical interpretability of learned features as the two central open challenges in the field.

Despite this progress, the automated detection of Forbush decreases has received comparatively limited attention from the machine learning community. Traditional FD identification relies on visual inspection of neutron monitor time series or on threshold-based algorithms applied to the global survey method output (Melkumyan et al., 2024, 2023), both of which are labor-intensive and prone to inconsistency for weak events. Okike (2020) proposed a Fourier-based preprocessing pipeline to disentangle FD signals from diurnal cosmic ray anisotropies prior to automated detection, highlighting the signal complexity that complicates direct ML approaches. While prior work by the present author has explored complexity-based and graph-theoretic descriptors for FD characterization (Sierra-Porta et al., 2024; Perez-Navarro and Sierra-Porta, 2024), the application of kernel-based supervised classifiers to the binary FD versus quiet-time detection problem using multivariate solar wind and neutron monitor features remains, to our knowledge, largely unexplored. The present study aims to fill this gap with a systematic experimental benchmark.

1.1.2. Quantum Kernel Methods: Theory and Empirical Benchmarks

A self-contained introduction to the quantum computing concepts used throughout this paper is provided in Appendix Appendix A, which readers may find useful as background for the material presented in this and subsequent sections.

Quantum kernel methods were placed on a rigorous theoretical footing by Schuld and Killoran (2019) and Havlíček et al. (2019), who demonstrated that quantum feature maps can define kernel functions whose evaluation is classically intractable for sufficiently expressive circuits. The `ZZFeatureMap` circuit used in the present work was introduced in Havlíček et al. (2019) as a concrete instantiation of this framework, and subsequently shown by Liu et al. (2021) to satisfy conditions sufficient for a provable quantum speedup over classical kernel methods on specific problem classes (see Section 3 for the explicit circuit definition and encoding scheme). The broader theoretical landscape of quantum machine learning is reviewed in Biamonte et al. (2017) and Cerezo et al. (2022).

Empirical demonstrations of quantum kernel advantage on real-world datasets have, however, proven elusive. Huang et al. (2021) established that the relative performance of quantum and classical kernels is governed by the geometry of the data distribution in feature space, implying that quantum advantage is dataset-dependent rather than universal. Systematic benchmarking studies—including Alvarez-Estevez (2025), who evaluated quantum kernel training across a range of standard classification datasets—consistently report mixed results: quantum kernels outperform classical counterparts on some tasks while underperforming on others, with the outcome depending sensitively on the choice of feature map and hyperparameter configuration. These findings underscore the importance of domain-specific benchmarks, which motivates the systematic experimental design of the present study.

1.1.3. Quantum Machine Learning for Earth and Atmospheric Sciences

The application of QML methods to geophysical and Earth observation problems is an emerging research direction. In the domain of remote sensing, quantum kernel SVMs have been applied to satellite image classification (Miroszewski et al., 2023) and hyperspectral image analysis (Otgonbaatar and Datcu, 2022), with results suggesting that quantum kernels can match or marginally exceed classical SVM performance on spectrally

complex datasets with limited training samples. Jaderberg et al. (2024) demonstrated that parameterized quantum circuits can be trained to reproduce global atmospheric stream function dynamics, providing an early proof of concept for quantum ML in weather modeling. To the best of our knowledge, no prior work has applied quantum kernel methods to space weather data or to cosmic ray time series, and no study has investigated the dependence of quantum kernel performance on the physical amplitude of the geophysical events being classified. Addressing both of these gaps is the primary contribution of the present paper.

1.2. Aims and Scope

In this work we present a systematic benchmark of quantum kernel SVMs for FD detection using a dataset constructed from the Forbush Effects and Interplanetary Disturbances (FEID) catalog (Belov et al., 2021; Okoye et al., 2024; Jerry-Okafor et al., 2024), hourly OMNI solar wind data (King and Papitashvili, 2005), and cosmic ray count rates from the Jungfraujoch neutron monitor (JUNG) of the Neutron Monitor Database (NMDB) (Mavromichalaki et al., 2011). Our central finding is that quantum kernel performance is strongly regime-dependent, governed by the magnitude of the FD events included in the classification problem: the classical RBF-SVM dominates in the noise-limited low-amplitude regime, while the quantum kernel achieves consistent and statistically significant advantage above a physically interpretable magnitude threshold. The entanglement structure of the `ZZFeatureMap` circuit (Havlíček et al., 2019) appears to capture non-linear IMF–GCR correlations that are inaccessible to the Euclidean-distance-based RBF kernel, but only when those correlations are sufficiently strong and structured to constitute a genuine learning signal. Detailed results and their physical interpretation are provided in Sections 4 and 5.

The remainder of this paper is organized as follows. Section 2 describes the datasets and preprocessing pipeline. Section 3 introduces the quantum kernel framework, the feature extraction procedure, and the experimental design. Results are presented in Section 4, including the magnitude-dependent performance analysis, learning curves, and circuit architecture sensitivity. Section 5 discusses the physical interpretation of the magnitude threshold and the implications for near-term quantum machine learning in heliophysics. Conclusions are drawn in Section 6. A self-contained primer on quantum computing concepts is provided in Appendix Appendix A.

2. Data

2.1. Forbush Decrease Catalog

Confirmed FD events were drawn from the Forbush Effects and Interplanetary Disturbances (FEID) catalog, maintained by the Pushkov Institute of Terrestrial Magnetism, Ionosphere and Radio Wave Propagation (IZMIRAN) and publicly accessible at `https://tools.izmiran.ru/w/feid` (Belov et al., 2021; Okoye et al., 2024; Jerry-Okafor et al., 2024). The catalog provides onset timestamps, FD magnitude estimates derived from global neutron monitor network data using the global survey method (GSM), and a suite of associated solar wind and geomagnetic parameters for each event. The version used in this study spans the period from 1995 January 1 to 2020 April 30, comprising 2,981 catalogued events. The primary classification target is the FD magnitude parameter `Magn`, defined as the maximal range of cosmic ray density variations for particles with 10 GV rigidity as derived from the GSM analysis of the neutron monitor network (Belov, 2008). Events were filtered by `Magn` threshold in the range 0–7% to define distinct experimental subsets, as described in Section 3.4.

2.2. OMNI Solar Wind Data

Hourly-resolution solar wind and geomagnetic index data were obtained from the OMNI database, compiled and distributed by the NASA Space Physics Data Facility (SPDF) and accessible via the OMNIWeb interface at `https://omniweb.gsfc.nasa.gov/form/dx1.html` (King and Papitashvili, 2005). The OMNI dataset covers the period from 1995 January 1 to 2021 December 31, comprising 236,688 hourly records. Ten parameters were selected as input features for the classification pipeline: the interplanetary magnetic field (IMF) total field strength (`Scalar B`), the GSE components of the IMF vector (`Bx`, `By`, `Bz`), solar wind proton temperature (`SW Temperature`), solar wind proton density, solar wind plasma speed, the Kp geomagnetic activity index, the Dst ring current index, and the F10.7 solar radio flux index used as a proxy for solar activity level. The proton temperature is a particularly informative discriminator for ICME identification, as ejecta material is systematically cooler than the ambient solar wind at equivalent speeds (Richardson and Cane, 2010). Missing values in the OMNI record, encoded as fill values following the OMNI data format convention, were replaced by `NaN` prior to any further processing.

2.3. Neutron Monitor Data

Cosmic ray count rates were obtained from the Neutron Monitor Database (NMDB), accessible at `https://www.nmdb.eu/nest/` (Mavromichalaki et al., 2011). Specifically, one-hour validated count rates from the Jungfraujoch station (JUNG, geographic coordinates: 46.55°N, 7.98°E; altitude: 3,475 m a.s.l.; vertical cutoff rigidity: ∼4.49 GV) were used, with pressure and efficiency corrections applied as provided by NMDB (`1HCOR_E_JUNG`). The JUNG record spans 1995 January 1 to 2021 December 31, comprising 235,277 hourly records. Jungfraujoch was selected on the basis of its long and continuous operational history, its intermediate cutoff rigidity—making it sensitive to the moderate-energy GCR population most strongly modulated during FDs—and its established use in previous FD studies (Melkumyan et al., 2024, 2023). Count rate values equal to or below zero, indicative of instrumental dropout or data transmission errors, were flagged and treated as missing. Outlier rejection was applied using a 5σ threshold relative to the station median, removing sporadic high-count excursions associated with solar energetic particle events or detector glitches.

2.4. Dataset Construction

The effective temporal coverage of the dataset is determined by the intersection of the three sources: 1995 January 1 to 2020 April 30. For each FD onset timestamp in the catalog, a multivariate time series window of 73 hourly steps was extracted, spanning 24 hours before the onset to 48 hours after it. The asymmetric design of this window reflects the physical structure of ICME-driven FDs: the 24-hour pre-onset segment is sufficient to characterise the ambient solar wind state and capture the arrival of the preceding interplanetary shock, while the 48-hour post-onset segment covers the full suppression phase and the early recovery, which typically unfolds over one to two days for moderate events (Cane, 2000). A symmetric window of equal pre- and post-onset duration would either truncate the recovery phase or introduce excessive quiet-time data prior to the shock arrival, in both cases degrading the discriminability of the FD signature.

Gap-filling was applied to each extracted window in three sequential steps: (1) linear interpolation across internal `NaN` segments of up to 6 consecutive hours; (2) forward-fill for leading missing values at the start of the window; and (3) backward-fill for trailing missing values at the end of the window. Steps (2) and (3) were applied only after step (1) to avoid propagating boundary values into the interior of the window. No imputation was performed for gaps longer than 6 consecutive hours in the interior; windows containing more than 20% missing values across all channels after step (1) were discarded entirely. After this procedure, and applying the `Magn` $\geq 0\%$ filter, 2,971 valid FD windows were retained, representing a loss of less than 1% of the catalog due to data gaps.

To construct a balanced binary classification dataset, an equal number of negative examples (non-FD, label = 0) were generated by randomly sampling quiet-time windows from the OMNI and JUNG records. A candidate quiet-time window was accepted only if its centre timestamp was separated by more than 72 hours from any catalogued FD onset, ensuring that negative examples were free of FD contamination. The same gap-filling and quality criteria applied to positive examples were enforced for negative examples. The resulting dataset comprises 5,942 labelled windows (2,971 FD and 2,971 non-FD), each represented as a matrix of shape 73×11, where the eleven channels correspond to the ten OMNI parameters and the JUNG cosmic ray count rate. Table 1 summarises the dataset composition.

3. Methods

3.1. Feature Extraction

Each labelled window of shape 73×11 was transformed into a fixed-length feature vector prior to classification. For each of the eleven input channels, two sets of descriptors were computed. The first set comprises eight global statistical features: mean, standard deviation, minimum, maximum, peak-to-peak range, skewness, excess kurtosis, and linear trend slope estimated by ordinary least squares. The second set comprises three onset-relative features designed to capture the characteristic morphology of FD events: the channel value at the onset hour, the maximum drop from onset to the post-onset minimum, and the normalized time elapsed from onset to the post-onset minimum. This yields $11 \times (8 + 3) = 121$ features per window. Windows containing residual `NaN` values after the gap-filling procedure described in Section 2.4—arising from constant or near-constant segments where higher-order moments are numerically ill-defined—were imputed using the column mean computed over the training partition only, to prevent data leakage.

Feature selection was then applied to reduce the dimensionality to a number compatible with near-term quantum hardware simulation. The k features with the highest mutual information (MI) with the binary label were retained (Cover and Thomas, 2006), where k equals the number of qubits in the quantum circuit (Section 3.3). MI-based selection was preferred over variance-based or correlation-based alternatives because it captures non-linear statistical dependencies between features and the target, making it consistent with the non-linear kernel approach used downstream. Selection was performed exclusively on the training partition at each experimental run to prevent data leakage.

3.2. Classical Baseline: RBF-SVM

The classical baseline is a support vector machine (SVM) with a radial basis function (RBF) kernel (Cortes and Vapnik, 1995), implemented via `scikit-learn` (Pedregosa et al., 2011). Given two feature vectors $\mathbf{x}, \mathbf{x}' \in \mathbb{R}^k$, the RBF kernel is defined as

$$K_{\mathrm{RBF}}(\mathbf{x}, \mathbf{x}') = \exp\left(-\gamma \|\mathbf{x} - \mathbf{x}'\|^2\right), \tag{1}$$

where γ is set by the `scale` heuristic ($\gamma = 1/(k \cdot \sigma_X^2)$, with σ_X^2 the variance of the training data). The regularisation parameter was fixed at $C = 1.0$ throughout all experiments. Two variants of the RBF-SVM are reported: *RBF-SVM (full)*, trained on the complete training partition comprising 4,456 samples, and *RBF-SVM (sub)*, trained on the same stratified subsample of size n_q used to train the quantum kernel. The latter provides a fair comparison against the quantum model, controlling for the effect of training set size and isolating the contribution of the kernel function itself.

3.3. Quantum Kernel SVM

3.3.1. Background

Quantum kernel methods extend the classical kernel framework to the quantum domain by exploiting the exponentially large Hilbert space of a k-qubit system as a feature space (Biamonte et al., 2017; Schuld and Killoran, 2019). The central idea is to encode classical data vectors $\mathbf{x} \in \mathbb{R}^k$ into quantum states via a parameterized unitary circuit $U_\phi(\mathbf{x})$, the *quantum feature map*, which prepares the state $|\phi(\mathbf{x})\rangle = U_\phi(\mathbf{x})|0\rangle^{\otimes k}$. The quantum kernel is then defined as the fidelity between two encoded states,

$$K_Q(\mathbf{x}, \mathbf{x}') = \left|\langle \phi(\mathbf{x}) | \phi(\mathbf{x}')\rangle\right|^2 = \left|\langle 0|^{\otimes k}\, U_\phi^\dagger(\mathbf{x})\, U_\phi(\mathbf{x}')\, |0\rangle^{\otimes k}\right|^2, \tag{2}$$

Table 1: Summary of the classification dataset. Note that the missing value fraction (∼9.5%) refers to the proportion of individual hourly entries that contained `NaN` values across all windows and channels *before* gap-filling, and is distinct from the window discard rate (<1%), which reflects the fraction of FD windows excluded due to excessive gaps after interpolation.

Property	**Value**
Temporal coverage	1995 Jan 1 – 2020 Apr 30
Total labelled windows	5,942
FD windows (label = 1)	2,971
Non-FD windows (label = 0)	2,971
Window length	73 hours
Pre-onset window	24 hours
Post-onset window	48 hours
Number of input channels	11
OMNI channels	10
NMDB channels	1 (JUNG)
Missing value fraction (pre-imputation)	∼9.5% of hourly entries
Window discard rate (post-imputation)	<1% of catalog

which can be estimated as the probability of measuring the all-zeros bitstring after applying the combined circuit $U_\phi^\dagger(\mathbf{x})\,U_\phi(\mathbf{x}')$ (Havlíček et al., 2019). Because K_Q is a valid positive semi-definite kernel, it can be used as a drop-in replacement for the RBF kernel in a standard SVM, inheriting all convergence and generalization guarantees of the classical SVM framework (Rebentrost et al., 2014; Liu et al., 2021). The expressibility of K_Q grows with the circuit depth and the number of qubits, raising the possibility of representing kernel functions that are classically intractable (Huang et al., 2021).

3.3.2. ZZFeatureMap Circuit

The quantum feature map $U_\phi(\mathbf{x})$ is instantiated as the `ZZFeatureMap` (Havlíček et al., 2019), a two-local circuit that alternates layers of single-qubit Hadamard and R_z rotation gates with layers of entangling ZZ-interaction gates. For a single repetition, the circuit implements

$$U_\phi(\mathbf{x}) = \exp\left(i \sum_{j<l} \phi_{jl}(\mathbf{x})\, Z_j Z_l\right) \exp\left(i \sum_j x_j Z_j\right) H^{\otimes k}, \qquad (3)$$

where $\phi_{jl}(\mathbf{x}) = (\pi - x_j)(\pi - x_l)$ encodes pairwise feature correlations via entangling two-qubit rotations, and $H^{\otimes k}$ denotes the tensor product of Hadamard gates acting on all k qubits. This constitutes a form of *angle encoding*: each selected feature x_j, after rescaling to $[-\pi, \pi]$, is mapped directly to a rotation angle in both the single-qubit R_z layer and the two-qubit ZZ entangling layer. The key distinction from classical kernel methods is that the entangling layer encodes *products* of feature values $\phi_{jl}(\mathbf{x}) = (\pi - x_j)(\pi - x_l)$ into two-qubit interactions, introducing cross-feature correlations that are absent from any kernel that acts on features independently. The circuit is repeated for $r \in \{1, 2, 3\}$ layers (the `reps` parameter) to control the expressibility of the feature map, with linear entanglement topology to limit circuit depth. Prior to encoding, all selected features were rescaled to the interval $[-\pi, \pi]$ using a min-max scaler fitted on the training partition, ensuring that the rotation angles in Equation (3) span the full periodicity of the quantum gates (Schuld, 2021).

Figure 1 shows the explicit gate-level diagram of the `ZZFeatureMap` for $k = 4$ qubits and reps = 1, as implemented in Qiskit 1.4. Each data point $\mathbf{x} \in \mathbb{R}^4$ is encoded in a single pass through the Hadamard layer, the single-qubit P-gate layer (equivalent to R_z up to global phase), and the two-qubit entangling C_P-gate layer. The fidelity kernel is evaluated by appending the conjugate circuit $U_\phi^\dagger(\mathbf{x}')$ and measuring the probability of the all-zeros outcome.

3.3.3. Kernel Matrix Computation and SVM Training

For a training set of n_q samples, the quantum kernel matrix $\mathbf{K}_{\text{train}} \in \mathbb{R}^{n_q \times n_q}$ is computed by evaluating Equation (2) for all $n_q(n_q + 1)/2$ unique pairs, exploiting the symmetry of the kernel. The full training partition of 4,456 samples is not used directly for quantum kernel training because the resulting $\mathbf{K}_{\text{train}}$ would require $4{,}456 \times 4{,}457/2 \approx 9.9 \times 10^6$ statevector simulations per configuration — computationally prohibitive under local simulation on the hardware used. Instead, stratified subsamples of size $n_q \in \{50, 100, 150, 200, 250\}$ are drawn from the training partition, reducing the number of kernel evaluations to at most $250 \times 251/2 = 31{,}375$ for the training matrix. The test kernel matrix $\mathbf{K}_{\text{test}} \in \mathbb{R}^{n_{\text{test}} \times n_q}$ is computed analogously, with $n_{\text{test}} = 1{,}486$ in all experiments. To avoid out-of-memory failures under local simulation, the test matrix was computed in sequential batches of 50 rows. All kernel evaluations were performed using exact statevector simulation via Qiskit's `ComputeUncompute` fidelity protocol (Qiskit contributors, 2024), which provides shot-noise-free estimates of K_Q. Qiskit version 1.4 was used throughout, rather than the more recent version 2.x, because the `ComputeUncompute` fidelity estimator and the `Sampler` V1 primitive required for its evaluation are fully supported in version 1.4 and have undergone interface-breaking changes in version 2.x that introduce incompatibilities with the `FidelityQuantumKernel` class used here. The SVM classifier uses $\mathbf{K}_{\text{train}}$ as a precomputed kernel matrix with $C = 1.0$, implemented via `scikit-learn` (Pedregosa et al., 2011).

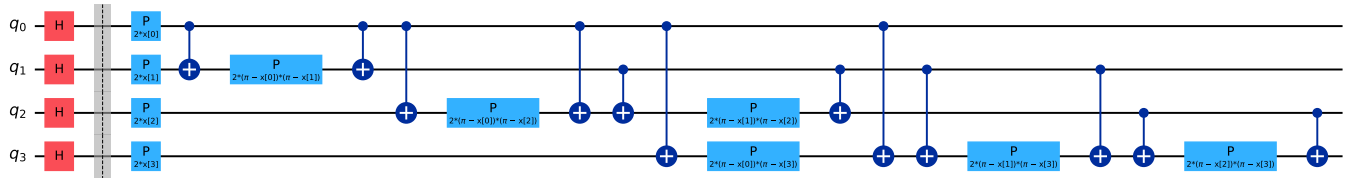

Figure 1: Gate-level diagram of the `ZZFeatureMap` circuit for $k = 4$ qubits and reps $= 1$. The circuit begins with a layer of Hadamard gates (H) that place all qubits in equal superposition, followed by single-qubit phase gates (P) that encode each feature x_j as a rotation angle, and two-qubit controlled-phase gates (C_P) that encode pairwise products $\phi_{jl}(\mathbf{x}) = (\pi - x_j)(\pi - x_l)$ via ZZ interactions with linear entanglement topology. To evaluate the quantum kernel $K_Q(\mathbf{x}, \mathbf{x}')$, the conjugate circuit $U_\phi^\dagger(\mathbf{x}')$ is appended and the probability of measuring the all-zeros bitstring is recorded.

3.4. Experimental Design

A systematic grid of 270 experimental configurations was evaluated, varying four independent parameters:

- **FD magnitude threshold** (`min_magn`): $\{0, 3, 4, 5, 6, 7\}\%$, controlling which FD events enter the dataset and defining the signal-to-noise regime of the classification problem.
- **Number of qubits** (k): $\{4, 6, 8\}$, simultaneously setting the number of MI-selected features and the dimension of the quantum feature space.
- **Circuit depth** (`reps`): $\{1, 2, 3\}$, controlling the expressibility of the `ZZFeatureMap`.
- **Quantum training size** (n_q): $\{50, 100, 150, 200, 250\}$, the number of stratified samples used to train both the quantum kernel and the RBF-SVM (sub) baseline.

For each configuration, the full dataset was split into training (75%) and test (25%) partitions using stratified sampling to preserve class balance, yielding a fixed test set of 1,486 samples. A stratified subsample of size n_q was drawn from the training partition to train both the quantum kernel and the RBF-SVM (sub). Stratified sampling was used for the subsample—rather than simple random sampling—to guarantee that both classes remain equally represented at all values of n_q, which is essential for obtaining unbiased AUC estimates when n_q is small relative to the full training partition. Without stratification, random subsamples at $n_q = 50$ would have a non-negligible probability of class imbalance severe enough to distort the comparison. The RBF-SVM (full) was trained on the complete training partition. All three models were evaluated on the same held-out test set, ensuring that performance comparisons are not confounded by differences in the evaluation data. A fixed random seed of 42 was used for all train–test splits, subsample draws, and SVM solvers throughout the study, ensuring full reproducibility of all reported results.

The primary comparison metric is

$$\Delta\text{AUC} = \text{AUC}_{\text{QSVM}} - \text{AUC}_{\text{RBF-sub}}, \tag{4}$$

where AUC denotes the area under the receiver operating characteristic (ROC) curve (Fawcett, 2006). A positive ΔAUC indicates that the quantum kernel outperforms its classical counterpart under identical data conditions. Secondary metrics include classification accuracy and the confusion matrix. Five-fold cross-validation on the full dataset was additionally performed for the RBF-SVM (full) to provide a bias-corrected estimate of classical performance.

Statistical significance of ΔAUC within each magnitude threshold group was assessed using the Wilcoxon signed-rank test (Wilcoxon, 1945), a non-parametric test that evaluates whether the median of a paired difference—here, $\text{AUC}_{\text{QSVM}} - \text{AUC}_{\text{RBF-sub}}$ across all configurations within a group—differs significantly from zero. The one-sided variant (alternative: $\Delta\text{AUC} > 0$) was used to test the directional hypothesis that quantum kernels outperform the classical baseline. To assess whether the distribution of ΔAUC differs significantly across the six magnitude threshold groups, the Kruskal–Wallis test (Kruskal and Wallis, 1952) was applied — a non-parametric one-way analysis of variance that tests for stochastic dominance among $g \geq 2$ independent groups without assuming normality. Effect sizes are reported as rank-biserial correlation coefficients r alongside each Wilcoxon test statistic, and bootstrap 95% confidence intervals (9,999 resamples, seed = 42) are reported for all mean ΔAUC estimates.

Individual experiment durations ranged from approximately 0.5 minutes (high-magnitude thresholds with small datasets and $n_q = 50$) to approximately 51 minutes (`min_magn` = 0%, $k = 8$ qubits, reps $= 3$, $n_q = 250$), with a median of approximately 2.4 minutes per configuration. The complete benchmark of 270 configurations required approximately 2–3 days of wall-clock time on an Intel Core i7-8565U workstation (4 physical cores, 16 GiB RAM, Ubuntu 26.04 LTS) running 4 parallel processes, reflecting both the $O(n_q^2)$ scaling of the quantum kernel matrix and the thermal throttling inherent to sustained parallel load on mobile hardware. By contrast, the classical RBF-SVM required less than 2 seconds per configuration on the same machine.

All experiments were implemented in Python 3.13 using Qiskit 1.4 (Qiskit contributors, 2024) and scikit-learn 1.6 (Pedregosa et al., 2011), and executed using GNU Parallel (Tange, 2011) to distribute the 270 configurations across four concurrent processes.

4. Results

4.1. Dataset and Experimental Summary

The final classification dataset constructed following the procedure described in Section 2.4 comprises 5,942 labelled win-

dows (2,971 FD and 2,971 non-FD), with a training partition of 4,456 samples and a held-out test set of 1,486 samples for the unfiltered case (`min_magn` = 0%). As the magnitude threshold is raised, the number of available FD events decreases and the dataset is rebalanced accordingly, with the most restrictive filter (`min_magn` = 7%) yielding 34 FD events, a training partition of 51 samples, and a test set of 17 samples. Table 2 summarises the dataset composition at each threshold. The progressive reduction in dataset size as the threshold increases is an inherent consequence of the FD magnitude distribution in the FEID catalog, and constitutes the primary source of increased statistical uncertainty at higher thresholds.

Table 2: Dataset composition by FD magnitude threshold. FD and non-FD windows are balanced at all thresholds. Training and test partitions are obtained by stratified 75/25 splitting of the total window count.

Threshold	FD	Non-FD	Total	Train	Test
Magn≥0%	2,971	2,971	5,942	4,456	1,486
Magn≥3%	190	190	380	285	95
Magn≥4%	107	107	214	160	54
Magn≥5%	65	65	130	97	33
Magn≥6%	47	47	94	70	24
Magn≥7%	34	34	68	51	17

Across all 180 valid experimental configurations (6 magnitude thresholds × 3 qubit counts × 3 circuit depths × 5 training sizes), the quantum kernel achieved higher AUC than the classical RBF-SVM (sub) in 73 cases (40.6%), with 13 cases (7.2%) in parity ($|\Delta\text{AUC}| \leq 0.01$) and 94 cases (52.2%) where the classical kernel outperformed. The global mean $\Delta\text{AUC} = +0.0021$ ($\sigma = 0.083$) is statistically indistinguishable from zero when all configurations are pooled, consistent with the regime-dependent behaviour described below.

In Figure 2a, the RBF-SVM (full) and RBF-SVM (sub) curves are nearly identical across all thresholds. This is expected: both models use the same 121-dimensional feature representation and the same MI-selected feature subset; the only difference is training set size. The near-identical performance indicates that the RBF classifier saturates its AUC well within the subsampled training set sizes tested ($n_q \leq 250$), so that the additional samples in the full partition provide negligible benefit. This saturation confirms that the RBF-SVM is not data-starved in the subsampled regime, making the fair comparison (QSVM vs. RBF-sub) a genuine test of kernel expressibility rather than a confound of sample size.

4.2. Magnitude-Dependent Performance

The dominant result of this study is the strong dependence of quantum kernel advantage on the FD magnitude threshold, illustrated in Figure 2 and summarised in Table 5. For `min_magn` = 0%—including all 2,971 FD events regardless of amplitude—the quantum kernel is uniformly outperformed by the classical RBF-SVM (sub), with a mean $\Delta\text{AUC} = -0.100$ (95% CI: [−0.112, −0.088]), and not a single configuration showing quantum advantage (0/45). This regime is dominated by weak FDs whose time-series signatures are largely indistinguishable from quiet-time fluctuations in the solar wind and neutron monitor records, producing a classification problem that is fundamentally noise-limited.

Raising the threshold to `min_magn` = 3% (190 events) reduces the classical advantage but does not reverse it: mean $\Delta\text{AUC} = -0.025$ (95% CI: [−0.030, −0.020]), with only 4 out of 45 configurations showing marginal quantum advantage (8.9%). The transition occurs sharply between 3% and 4% (Figure 3). At `min_magn` = 4% (107 events), the mean ΔAUC becomes strongly positive (+0.070, 95% CI: [+0.048, +0.090]), with 72.2% of configurations showing quantum advantage. The transition is statistically confirmed: a Wilcoxon signed-rank test (testing whether $\Delta\text{AUC} > 0$) yields $p = 1.9 \times 10^{-5}$, rank-biserial $r = 0.787$ for `min_magn` ≥ 4%, compared to $p = 1.0$ for the two lower thresholds. A Kruskal–Wallis test across all six threshold groups confirms that the distribution of ΔAUC differs significantly across groups ($H = 138.2$, $p = 4.3 \times 10^{-28}$). Table 3 reports the complete statistical results including bootstrap confidence intervals and rank-biserial effect sizes for all threshold groups.

At `min_magn` = 5% and 6% (65 and 47 events, respectively), quantum advantage reaches its maximum consistency: 100% of tested configurations yield $\Delta\text{AUC} > 0$, with mean values of +0.074 (95% CI: [+0.062, +0.086], $r = 1.000$) and +0.087 (95% CI: [+0.067, +0.103], $r = 1.000$) respectively. The mean AUC of the quantum kernel in these regimes exceeds 0.95, reaching a peak of 1.000 in isolated configurations (Table 5). At `min_magn` = 7% (34 events), quantum advantage persists on average (+0.031, 95% CI: [+0.012, +0.052]) but becomes less consistent (55.6% of configurations), likely reflecting the increased statistical variability arising from the small dataset size at this threshold (Table 2). The Wilcoxon test remains significant at this threshold ($p = 0.007$, $r = 0.754$, **).

Table 3: Statistical summary of ΔAUC by FD magnitude threshold. The Wilcoxon signed-rank test (one-sided, H_1: $\Delta\text{AUC} > 0$) is reported alongside the rank-biserial correlation r as an effect size measure ($|r| > 0.1$ small, > 0.3 medium, > 0.5 large). Bootstrap 95% confidence intervals on the mean ΔAUC are based on 9,999 resamples (seed = 42). Significance codes: $^{**}p < 0.01$; $^{***}p < 0.001$; ns: not significant. For Magn≥0% and Magn≥3%, r reflects strong classical dominance; the test is one-sided (H_1: $\Delta\text{AUC} > 0$), so ns indicates failure to reject in the quantum-advantage direction.

Threshold	n	Mean ΔAUC	95% CI	r	Sig.
Magn≥0%	45	−0.100	[−0.112, −0.088]	1.000	ns
Magn≥3%	45	−0.025	[−0.030, −0.020]	0.951	ns
Magn≥4%	36	+0.070	[+0.048, +0.090]	0.787	***
Magn≥5%	18	+0.074	[+0.062, +0.086]	1.000	***
Magn≥6%	18	+0.087	[+0.067, +0.103]	1.000	***
Magn≥7%	18	+0.031	[+0.012, +0.052]	0.754	**
Kruskal–Wallis: $H = 138.2$, $p = 4.3 \times 10^{-28}$					

4.3. Effect of Circuit Architecture

Figure 4 shows ΔAUC averaged over all training sizes for each combination of magnitude threshold, qubit count, and circuit depth. The most striking feature is the strong interaction between `min_magn` and `n_qubits`: for `min_magn` ≥ 4%, four-qubit circuits consistently achieve the largest ΔAUC, with a

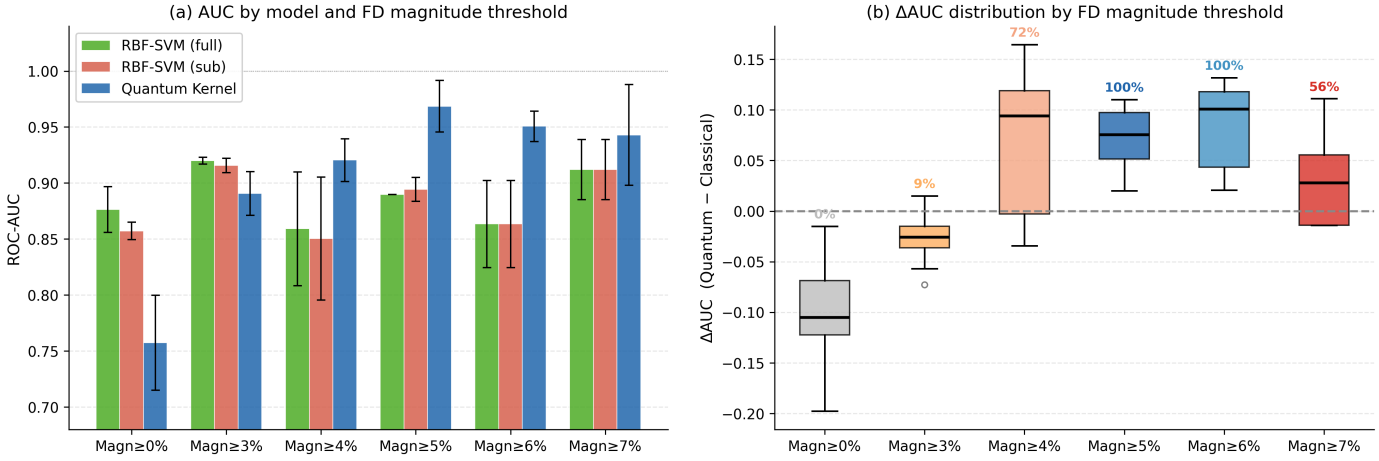

Figure 2: (a) Mean ROC-AUC for each of the three classifiers—RBF-SVM (full), RBF-SVM (sub), and Quantum Kernel SVM—as a function of FD magnitude threshold. Error bars indicate ±1 standard deviation across experimental configurations. (b) Distribution of $\Delta\text{AUC} = \text{AUC}_{\text{QSVM}} - \text{AUC}_{\text{RBF-sub}}$ as a function of threshold, shown as boxplots. Percentages above each box indicate the fraction of configurations in which the quantum kernel outperforms the classical baseline. The dashed line marks the parity boundary ($\Delta\text{AUC} = 0$). The near-identical performance of RBF-SVM (full) and RBF-SVM (sub) in panel (a) reflects AUC saturation of the classical kernel within the tested subsample sizes ($n_q \leq 250$); see Section 4.1.

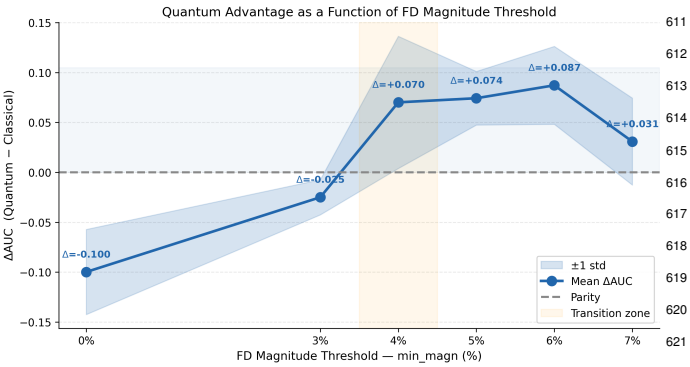

Figure 3: Mean ΔAUC as a function of FD magnitude threshold (`min_magn`), averaged over all 180 experimental configurations. The shaded band indicates ±1 standard deviation. The transition from classical to quantum advantage occurs between 3% and 4%, marked by the orange transition zone. Annotated values correspond to the mean ΔAUC at each threshold.

peak of +0.135 at (`min_magn` = 4%, $k = 4$). Six-qubit circuits follow closely (+0.091 at the same threshold). Eight-qubit circuits, by contrast, show near-zero or negative ΔAUC at `min_magn` ≤ 4%, though they recover positive advantage at higher thresholds (+0.088 at `min_magn` = 5%, +0.108 at `min_magn` = 6%). We note that the non-monotonic pattern with qubit count is based on only three values ($k \in \{4, 6, 8\}$), which limits the strength of any claim about the functional form of this dependence. The observed pattern—in which intermediate qubit counts do not consistently outperform smaller ones—is nevertheless qualitatively consistent with the concentration phenomenon described by Thanasilp et al. (2024): deeper circuits with more qubits produce kernel matrices whose entries become exponentially concentrated around their mean, effectively reducing the discriminative power of the kernel for the training sizes used here.

Figure 6a summarises these trends across all qubit counts and magnitude thresholds; panel (b) shows the corresponding QSVM AUC as a function of training size in the winning regime (`min_magn` ≥ 5%).

The effect of circuit depth (`reps`) is secondary but notable. For `min_magn` ≥ 4%, deeper circuits (`reps` = 3) tend to produce higher ΔAUC at thresholds of 5–6%, achieving a maximum of +0.094 at (`min_magn` = 5%, `reps` = 3). For lower thresholds, additional circuit depth does not help and can slightly hurt performance. The panel (c) of Figure 4 confirms that, when averaged across all magnitude thresholds, the qubit–depth interaction is modest, with four and six qubit circuits in parity or slight advantage across all depth settings.

4.4. Learning Curves and Sample Efficiency

Figure 5a shows ΔAUC as a function of the quantum training size n_q for each magnitude threshold. For the low-threshold regimes (`min_magn` = 0% and 3%), ΔAUC remains negative across all training sizes, indicating that additional quantum training data does not recover competitiveness. For `min_magn` ≥ 5%, ΔAUC is positive and broadly stable across $n_q \in \{50, 100, 150, 200, 250\}$, with no systematic improvement as n_q increases. This robustness to training size is a practically important property: quantum advantage at high FD magnitudes is achievable even with the smallest training sets tested ($n_q = 50$), suggesting that the kernel structure of the `ZZFeatureMap` efficiently leverages the available data in the signal-rich regime.

For `min_magn` = 4%, the learning curve shows a slight improvement from $n_q = 50$ to $n_q = 100$, followed by stabilisation. The intermediate threshold (`min_magn` = 7%) shows the most variability, consistent with the limited number of available FD events at this threshold. Figure 5b provides a complementary view as an AUC scatter plot, in which each point represents one (k, `reps`, n_q) configuration coloured by magnitude threshold: configurations with `min_magn` ≥ 4% cluster above

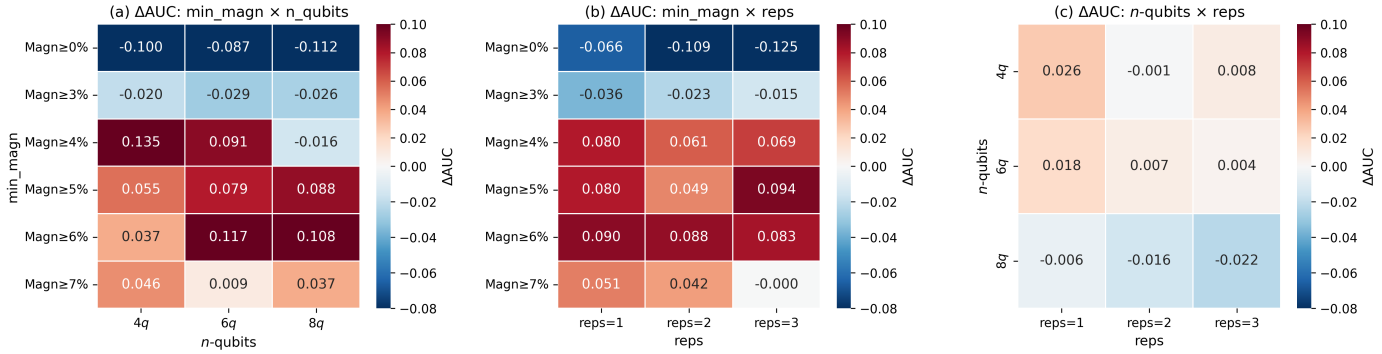

Figure 4: Heatmaps of mean ΔAUC for all parameter combinations. (a) ΔAUC as a function of FD magnitude threshold and qubit count. (b) ΔAUC as a function of FD magnitude threshold and circuit depth (`reps`). (c) ΔAUC as a function of qubit count and circuit depth, averaged across all magnitude thresholds. Red cells indicate quantum advantage; blue cells indicate classical advantage.

the parity diagonal, confirming quantum advantage across the full range of classical AUC values achieved in those regimes. The dense vertical clustering of points at certain $\text{AUC}_{\text{RBF-sub}}$ values reflects the fact that the classical baseline is fixed for a given (`min_magn`, n_q) combination regardless of the quantum circuit parameters, so multiple quantum configurations share the same x-coordinate.

4.5. Best Configurations and Computational Cost

The fifteen highest-ΔAUC configurations are dominated by `min_magn` = 4% with $k = 4$ qubits, achieving ΔAUC up to +0.165 with AUC_{QSVM} reaching 0.964. The second cluster corresponds to `min_magn` = 6% with $k = 6$ qubits, with ΔAUC up to +0.132. In the winning region defined as `min_magn` $\geq$ 5% and $k \in \{4, 6\}$, comprising 36 configurations, the quantum kernel achieves a mean AUC of 0.949 ± 0.032 against a classical baseline of 0.892 ± 0.033, with $\Delta\text{AUC} = +0.057 \pm 0.044$ and quantum advantage in 30 of 36 configurations (83%).

Table 4 quantifies the computational cost of the quantum kernel relative to the classical baseline. The number of fidelity evaluations required for the training kernel matrix grows as $O(n_q^2)$, reaching 31,375 for $n_q = 250$, while the test matrix requires $n_{\text{test}} \times n_q$ evaluations. For the largest tested configuration (`min_magn` = 0%, $k = 8$, $n_q = 250$), this amounts to 402,875 statevector simulations per experiment, compared to less than 2 seconds total for the classical RBF-SVM on the same hardware. This difference of three to four orders of magnitude in wall-clock time (see Section 3.4) underscores that quantum kernel methods are currently competitive in accuracy only in the signal-rich regime where the advantage is large enough to justify the cost.

Table 4: Theoretical quantum kernel evaluation counts by training size (n_q) and qubit count (k). K_{train} counts fidelity evaluations for the upper-triangular training kernel matrix; K_{test} counts evaluations against the held-out test set (n_{test} = 1,486, worst case `min_magn` = 0%). The classical RBF-SVM requires less than 2 s per configuration on the same hardware for all (k, n_q) combinations shown.

k	n_q	K_{train}	K_{test}	Total	RBF time
4	50	1,275	74,300	75,575	< 2 s
4	100	5,050	148,600	153,650	< 2 s
4	250	31,375	371,500	402,875	< 2 s
6	50	1,275	74,300	75,575	< 2 s
6	100	5,050	148,600	153,650	< 2 s
6	250	31,375	371,500	402,875	< 2 s
8	50	1,275	74,300	75,575	< 2 s
8	100	5,050	148,600	153,650	< 2 s
8	250	31,375	371,500	402,875	< 2 s

5. Discussion

5.1. Physical Interpretation of the Magnitude Threshold

The central empirical finding of this study—that quantum kernel advantage emerges sharply above a FD magnitude threshold of 4%—has a natural physical interpretation rooted in the structure of ICME-driven cosmic ray modulation. Large-amplitude FDs are predominantly associated with well-developed ICME magnetic structures: organised flux ropes with strong, ordered magnetic fields that produce coherent and sustained shielding of the GCR flux (Petukhova et al., 2020; Gutierrez et al., 2024). In these events, the modulation of the JUNG count rate is tightly correlated with the IMF dynamics—specifically, the temporal evolution of the Bz component drives the ring current response (Dst), which in turn correlates with the depth of the GCR suppression. The ZZFeatureMap encodes pairwise feature correlations through the entangling term $\phi_{jl}(\mathbf{x}) = (\pi - x_j)(\pi - x_l)$, and it is precisely these cross-channel correlations between IMF components (`Bx`, `By`, `Bz`), geomagnetic indices, and the JUNG count rate that characterise strong FD events. The RBF kernel, by contrast, measures only the Euclidean distance between feature vectors and cannot represent such structured cross-correlations explicitly.

Weak FDs (`Magn` < 4%), on the other hand, arise from a heterogeneous mixture of physical drivers: partial ICME encounters, corotating interaction regions, and flux ropes without well-developed sheaths (Gutierrez et al., 2024). In these cases, the IMF–GCR coupling is weaker and more variable, producing feature vectors that are difficult to separate from quiet-time windows regardless of the kernel used. The classification problem in this regime is fundamentally noise-limited, and the addi-

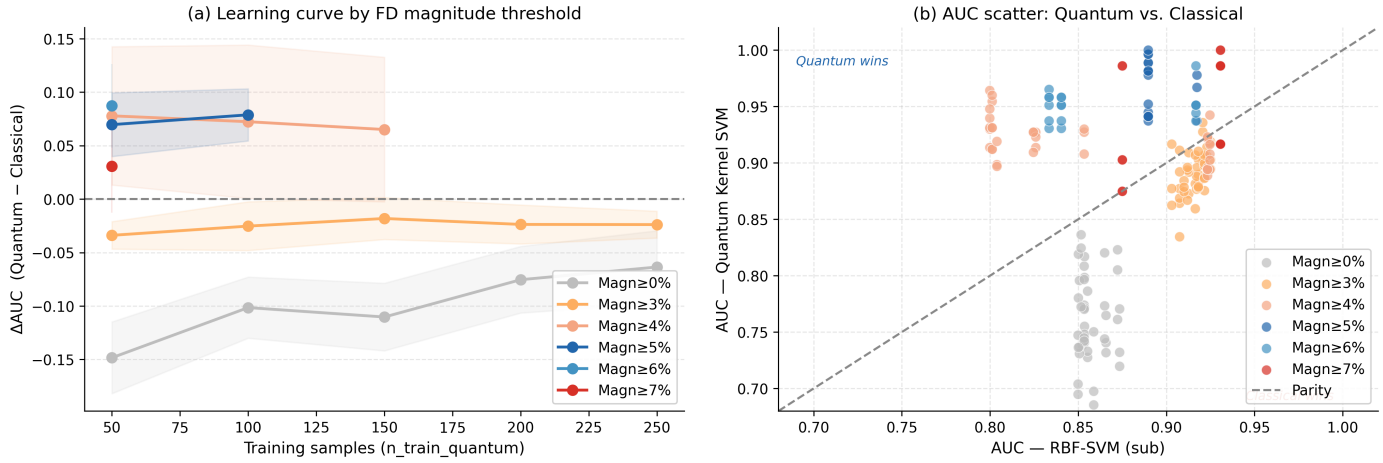

Figure 5: (a) ΔAUC as a function of quantum training size n_q for each FD magnitude threshold. Shaded bands indicate ±1 standard deviation across qubit count and circuit depth configurations. (b) Scatter plot of AUC values: each point represents one experimental configuration (k, `reps`, n_q), coloured by magnitude threshold. Points above the dashed diagonal (parity) indicate quantum advantage. Vertical clustering of points at fixed x-values reflects configurations sharing the same classical baseline (RBF-sub) AUC but different quantum circuit parameters.

tional representational capacity of the quantum kernel provides no benefit—and indeed can hurt performance, as the expressible function class is too large relative to the available training signal (Thanasilp et al., 2024).

The transition at 4% is therefore not merely a statistical artifact of the experimental design but reflects a genuine physical boundary: below it, the FD catalogue is dominated by events where the IMF–GCR connection is too noisy for quantum feature correlations to be useful; above it, the structured IMF dynamics of well-developed ICMEs create a feature geometry that the `ZZFeatureMap` can exploit more effectively than the RBF kernel.

5.2. *Interpretation of the Qubit Count Effect*

The non-monotonic behaviour of ΔAUC with qubit count—four and six qubits outperforming eight qubits in the low-threshold regime—is consistent with the exponential concentration phenomenon described by Thanasilp et al. (2024) and studied theoretically by Wang et al. (2021). We emphasise, however, that this pattern is observed across only three qubit values ($k \in \{4, 6, 8\}$), and should be regarded as a qualitative tendency rather than a quantitatively established functional relationship. Testing additional values (e.g., $k = 2$, $k = 10$) would be required to characterise the qubit-count dependence more precisely; $k = 10$ was not included in the present grid due to the severe kernel concentration observed at $k = 8$ in the low-magnitude regime, which suggested that larger circuits would be unlikely to provide further benefit. As the number of qubits and circuit depth increase, the entries of the kernel matrix $\mathbf{K}$ become exponentially concentrated around a constant value, reducing the effective rank of the matrix and limiting the discriminative power of the SVM. This effect is particularly pronounced at low magnitude thresholds where the classification signal is weak, and the high expressibility of larger circuits effectively washes out the relevant structure.

This result may appear to contradict the common intuition that greater circuit expressibility implies better classification performance. The apparent paradox is resolved by noting that expressibility and kernel utility are distinct properties (Thanasilp et al., 2024): a highly expressive circuit can access a rich function class, but if the kernel matrix entries concentrate exponentially toward a constant, the SVM receives effectively no discriminative information regardless of the functional richness of the feature map. In other words, more qubits expand the representable hypothesis class while simultaneously collapsing the kernel matrix toward a scalar multiple of the identity—a phenomenon that is fundamentally different from overfitting and cannot be corrected by regularisation alone.

At higher magnitude thresholds (`min_magn` $\geq$ 5%), the strong physical signal overcomes this concentration tendency, and all three qubit counts achieve positive ΔAUC, with the eight-qubit configuration recovering competitive performance (+0.088 at `min_magn` = 5%). This suggests that the concentration effect is signal-dependent: when the data contains strong, structured feature correlations, even larger circuits can exploit them effectively. The practical implication for near-term quantum hardware is that modest circuit sizes (k = 4–6 qubits) represent the most reliable operating point across a range of signal conditions, consistent with recommendations in the QML literature for NISQ devices (Incudini et al., 2025; Cerezo et al., 2022).

5.3. *Sample Efficiency and the Small-Data Regime*

A noteworthy aspect of the results is the robustness of quantum advantage to training set size in the signal-rich regime. For `min_magn` $\geq$ 5%, the mean ΔAUC is positive and stable across $n_q \in \{50, 100, 150, 200, 250\}$, indicating that the quantum kernel achieves competitive classification with as few as 50 training samples. This is consistent with theoretical results showing that quantum kernels can achieve lower generalisation error

Table 5: Summary of results by FD magnitude threshold and qubit count. Columns report the mean AUC of the full and subsampled RBF-SVM baselines, the mean AUC of the Quantum Kernel SVM, the mean and standard deviation of ΔAUC, and the fraction of configurations where the quantum kernel outperforms the classical baseline (Q>C). Results are averaged over all circuit depths (`reps` ∈ {1, 2, 3}) and training sizes ($n_q \in \{50, 100, 150, 200, 250\}$).

Threshold	k	n	**AUC$_{\text{RBF}}$(full)**	**AUC$_{\text{RBF}}$(sub)**	**AUC$_{\text{QSVM}}$**	**ΔAUC**	$\pm\sigma$	**Q>C**
Magn≥0%	4	15	0.861	0.854	0.753	−0.100	0.045	0/15 (0%)
	6	15	0.863	0.851	0.764	−0.087	0.043	0/15 (0%)
	8	15	0.905	0.867	0.755	−0.112	0.038	0/15 (0%)
Magn≥3%	4	15	0.924	0.919	0.899	−0.020	0.017	2/15 (13%)
	6	15	0.918	0.918	0.889	−0.029	0.014	0/15 (0%)
	8	15	0.918	0.910	0.884	−0.026	0.021	2/15 (13%)
Magn≥4%	4	12	0.801	0.800	0.936	+0.135	0.018	12/12 (100%)
	6	12	0.853	0.827	0.917	+0.091	0.016	12/12 (100%)
	8	12	0.923	0.924	0.908	−0.016	0.016	2/12 (17%)
Magn≥5%	4	6	0.890	0.904	0.959	+0.056	0.026	6/6 (100%)
	6	6	0.890	0.890	0.969	+0.079	0.022	6/6 (100%)
	8	6	0.890	0.890	0.978	+0.088	0.025	6/6 (100%)
Magn≥6%	4	6	0.917	0.917	0.954	+0.037	0.017	6/6 (100%)
	6	6	0.833	0.833	0.950	+0.117	0.014	6/6 (100%)
	8	6	0.840	0.840	0.948	+0.108	0.011	6/6 (100%)
Magn≥7%	4	6	0.875	0.875	0.921	+0.046	0.052	4/6 (67%)
	6	6	0.931	0.931	0.940	+0.009	0.036	2/6 (33%)
	8	6	0.931	0.931	0.968	+0.037	0.040	4/6 (67%)

than classical methods with fewer samples when the data geometry aligns with the kernel structure (Huang et al., 2021). In the context of FD classification, the practical relevance is significant: strong FDs are rare events. The FEID catalog contains only 34–65 events above the 5–7% magnitude thresholds tested here, and the entire 25-year observational record provides limited training data at these amplitudes. The ability of the quantum kernel to operate effectively in this small-data regime is therefore not a limitation but a natural alignment between the method and the application domain.

Nevertheless, the small absolute event counts at high thresholds (Table 2) introduce non-negligible statistical uncertainty into the performance estimates. Although the datasets are perfectly balanced (equal FD and non-FD windows at every threshold), the total sample sizes at `min_magn` ≥ 5%, 6%, and 7% are only 130, 94, and 68 windows respectively, with test sets of 33, 24, and 17 samples. AUC estimates computed on such small test sets carry wide confidence intervals, and isolated perfect-AUC configurations (e.g., AUC$_{\text{QSVM}}$ = 1.000 reported at `min_magn` = 5%) should be interpreted with corresponding caution. The bootstrap confidence intervals in Table 3 reflect this uncertainty and show that the mean ΔAUC at `min_magn` = 7% has the widest interval ([+0.012, +0.052]), consistent with the higher variability expected at this threshold.

5.4. Comparison with Prior QML Benchmarks

The regime-dependent quantum advantage observed here is broadly consistent with the emerging consensus from QML benchmarking studies. Alvarez-Estevez (2025) found that quantum kernels outperform classical counterparts on ad-hoc datasets but show mixed results on standard benchmarks, attributing the difference to whether the data geometry aligns with the inductive bias of the quantum feature map. Huang et al. (2021) established theoretically that quantum advantage is governed by the *geometric difference* between the quantum and classical kernel matrices, a quantity that is data-dependent. Our results provide an empirical demonstration of this principle in a physical science domain: the geometric difference between the `ZZFeatureMap` kernel and the RBF kernel is beneficial precisely in the regime where the data contains structured cross-channel correlations (strong FDs), and detrimental where it does not (weak FDs). The magnitude threshold at which this transition occurs—approximately 4%—is therefore a property of the FD data distribution, not of the quantum circuit architecture.

Compared to related applications in Earth observation, where quantum kernels have shown modest but consistent improvements for hyperspectral classification (Miroszewski et al., 2023; Otgonbaatar and Datcu, 2022), the advantage magnitudes reported here (mean ΔAUC = +0.087 at `min_magn` = 6%) are among the largest reported in any domain-specific QML benchmark to date, suggesting that strong FD events represent a particularly favourable problem geometry for quantum kernel methods.

5.5. Limitations and Future Directions

Several limitations of the present study should be acknowledged. First, all kernel evaluations were performed using exact statevector simulation, which provides shot-noise-free fidelity estimates but does not model the decoherence and gate errors of real quantum hardware. In practice, quantum hardware estimates each kernel entry $K_Q(\mathbf{x}, \mathbf{x}')$ by measuring the all-zeros bitstring probability over a finite number of circuit executions (shots). This introduces statistical shot noise into every entry of the kernel matrix, with variance scaling as $O(1/N_{\text{shots}})$ per entry. For a kernel matrix of size $n_q \times n_q$, errors in individual entries propagate to the SVM decision boundary and can substantially degrade classification performance, particularly for

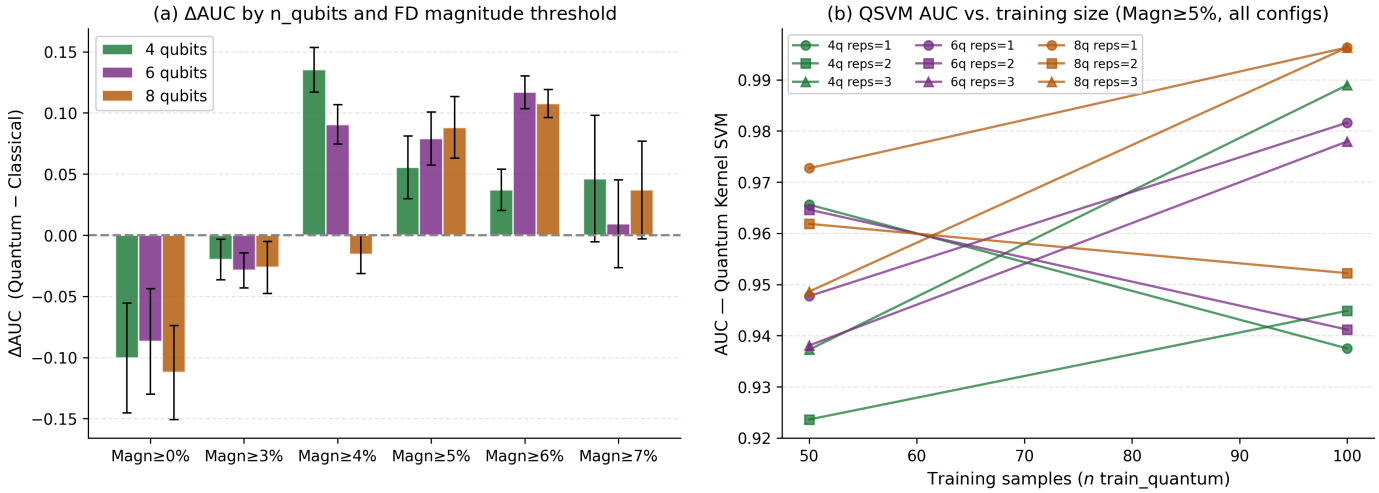

Figure 6: (a) ΔAUC as a function of FD magnitude threshold for each qubit count ($k \in \{4, 6, 8\}$). Error bars indicate ±1 standard deviation across circuit depths and training sizes. The apparent non-monotonic trend across qubit counts should be interpreted cautiously, as it is based on only three values ($k \in \{4, 6, 8\}$). (b) $\mathrm{AUC}_{\mathrm{QSVM}}$ as a function of training size for all nine (k, `reps`) combinations in the winning regime (`min_magn` ≥ 5%). The figure shows only two training size values ($n_q = 50$ and $n_q = 100$) because the total training partition at this threshold comprises only 97 samples, making $n_q > 100$ infeasible. The crossing and interleaving of curves at each n_q value reflects configuration-dependent sensitivity to training size rather than a systematic trend, consistent with the AUC saturation observed in the learning curves of Figure 5a.

deeper circuits where the target probability K_Q is already small and concentrated. Achieving the AUC values reported here on real hardware would require either a large number of shots per kernel entry (typically 10^3–10^4) or error mitigation techniques, both of which further increase the computational cost. Prior work has shown that NISQ-era noise can degrade quantum kernel performance significantly for deeper circuits (Wang et al., 2021; Incudini et al., 2025), and the advantage margins reported here may be reduced or eliminated in a noisy hardware setting. Experiments on actual quantum processors would be required to assess the robustness of our results to realistic noise, and constitute the most immediate next step for this research.

Second, the dataset is limited to a single neutron monitor station (JUNG). The global survey method combines data from the worldwide NMDB network to derive the `Magn` parameter used as the classification target, but the feature vectors used for classification include only the JUNG count rate as the GCR proxy. Including additional stations with different cutoff rigidities—such as Moscow (MOSC, 2.43 GV) or Oulu (3.7 GV)—would provide complementary spectral information about the GCR modulation and may further improve classification performance, particularly for weak events (Belov, 2008).

Third, the magnitude threshold of 4% is empirically derived from the FEID catalog and the specific feature representation used here. Whether this threshold generalises to other classification formulations (e.g., multiclass FD strength, onset-type classification) or other cosmic ray datasets is an open question. Future work should also explore whether projected quantum kernels (Huang et al., 2021)—which project the quantum state onto a lower-dimensional subspace before computing the kernel—can mitigate the concentration effect at larger qubit counts and potentially extend quantum advantage to the weak FD regime.

6. Conclusions

We have presented a systematic benchmark of quantum kernel support vector machines for the classification of Forbush decreases, using a 25-year observational dataset constructed from the FEID catalog, OMNI solar wind data, and JUNG neutron monitor count rates. A `ZZFeatureMap` quantum kernel with 4–8 qubits was evaluated across 180 experimental configurations and benchmarked against a classical RBF-SVM baseline under identical data conditions. The principal conclusions are as follows.

1. **Quantum advantage is regime-dependent, not universal.** The quantum kernel consistently outperforms the classical baseline only above a FD magnitude threshold of 4%. Below this threshold, the classical RBF-SVM is superior across all tested configurations. This result is consistent with theoretical predictions that quantum advantage requires data distributions whose geometry aligns with the inductive bias of the quantum feature map (Huang et al., 2021).
2. **The transition at 4% is statistically robust.** A Wilcoxon signed-rank test confirms that $\Delta\mathrm{AUC} > 0$ is statistically significant ($p = 1.9 \times 10^{-5}$, rank-biserial $r = 0.787$) for `min_magn` = 4%, with increasing effect sizes at higher thresholds ($r = 1.000$ at `min_magn` ≥ 5%). The Kruskal–Wallis test confirms that the six threshold groups exhibit significantly different ΔAUC distributions ($H = 138.2$, $p = 4.3 \times 10^{-28}$). At `min_magn` ≥ 5%, 100% of tested configurations show quantum advantage, with mean $\Delta\mathrm{AUC} = +0.074$ (95% CI: [+0.062, +0.086]) and peak $\mathrm{AUC}_{\mathrm{QSVM}} = 1.000$ (noting that this maximum is observed on a test set

of 33 samples and should be interpreted with appropriate caution).

3. **The magnitude threshold has a physical interpretation.** Large-amplitude FDs are driven by well-developed ICME magnetic structures (flux ropes) whose passage produces structured, correlated dynamics across the IMF components, geomagnetic indices, and cosmic ray count rates. The `ZZFeatureMap` encodes pairwise feature correlations via its entangling layer, providing a representational advantage over the Euclidean-distance-based RBF kernel precisely in this structured regime. Weak FDs, by contrast, arise from a heterogeneous mixture of drivers and exhibit noise-dominated feature distributions that neither kernel can separate effectively.
4. **Modest circuit architectures are optimal for near-term hardware.** Circuits with 4–6 qubits achieve the most consistent quantum advantage across magnitude regimes, while 8-qubit circuits suffer from kernel concentration effects at lower magnitude thresholds. This trend is based on three qubit-count values ($k \in \{4, 6, 8\}$) and should be regarded as a qualitative tendency pending evaluation at additional circuit sizes. Circuit depth (`reps`) has a secondary effect, with deeper circuits providing marginal benefits in the signal-rich regime.
5. **Quantum advantage is robust to small training sets.** In the winning regime (`min_magn` $\geq 5\%$, $k \in \{4, 6\}$), mean $\Delta\text{AUC} = +0.057$ is achieved with as few as 50 training samples, consistent with the theoretical expectation that quantum kernels are sample-efficient when the data geometry favours their inductive bias. This property is particularly relevant for FD classification, where strong events are rare across the historical record.

These results establish FD magnitude as a principled criterion for predicting when quantum kernel methods will outperform classical alternatives in heliophysical classification tasks. More broadly, they suggest a general strategy for deploying near-term QML in physical sciences: identify the physical regime where structured, correlated multi-channel dynamics are present, and restrict quantum kernel methods to that regime. The present study is limited to statevector simulation; extension to real quantum hardware, additional neutron monitor stations, and projected quantum kernels represent natural directions for future work that will further test the robustness and scalability of these findings.

Data Availability

The FEID catalog is publicly available at `https://tools.izmiran.ru/w/feid`. OMNI data were obtained from the NASA OMNIWeb service at `https://omniweb.gsfc.nasa.gov`. Neutron monitor data were obtained from the NMDB at `https://www.nmdb.eu/nest/`. The Python code implementing the full experimental pipeline, including data preprocessing, feature extraction, quantum kernel computation, and analysis notebooks, is openly available at `https://github.com/sierraporta/quantum-kernel-forbush`.

Acknowledgements

The author thanks the IZMIRAN team for maintaining the FEID catalog, the NASA Space Physics Data Facility for the OMNI database, and the NMDB team and station operators for providing open access to neutron monitor data. This work was carried out within the Escuela de Transformación Digital at Universidad Tecnológica de Bolívar (UTB), Cartagena de Indias, Colombia. Quantum kernel computations were performed using the open-source Qiskit framework (Qiskit contributors, 2024) on an Intel Core i7-8565U workstation.

Appendix A. Quantum Computing Primer

This appendix provides a self-contained introduction to the quantum computing concepts used in this paper. It is intended for readers with a background in classical machine learning who are unfamiliar with quantum mechanics. A comprehensive treatment can be found in Nielsen and Chuang (2010).

Appendix A.1. Qubits and Quantum States

The fundamental unit of quantum computation is the *qubit*, the quantum analogue of a classical bit. Whereas a classical bit is always in state 0 or 1, a qubit can exist in a *superposition* of both:

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle, \qquad \alpha, \beta \in \mathbb{C}, \qquad |\alpha|^2 + |\beta|^2 = 1. \quad \text{(A.1)}$$

Here $|0\rangle$ and $|1\rangle$ are the computational basis states (analogous to the unit vectors $\mathbf{e}_1$ and $\mathbf{e}_2$), and $|\alpha|^2$, $|\beta|^2$ are the probabilities of measuring the qubit in state 0 or 1 respectively. The state $|\psi\rangle$ is a unit vector in a two-dimensional complex Hilbert space $\mathcal{H} \cong \mathbb{C}^2$.

A system of k qubits occupies a 2^k-dimensional Hilbert space $\mathcal{H}^{\otimes k}$, with basis states $|b_1 b_2 \cdots b_k\rangle$ for all binary strings $b_i \in \{0, 1\}$. The general state of a k-qubit system is

$$|\Psi\rangle = \sum_{b \in \{0,1\}^k} c_b \, |b\rangle, \qquad \sum_b |c_b|^2 = 1, \quad \text{(A.2)}$$

where the 2^k complex amplitudes c_b encode a probability distribution over all possible measurement outcomes. This exponential growth of the state space with k is the core resource that quantum computing exploits.

Appendix A.2. Quantum Gates and Circuits

Quantum computation proceeds by applying *quantum gates* — unitary transformations U acting on the Hilbert space — to an initial state (typically $|0\rangle^{\otimes k}$). The unitarity condition $UU^\dagger = I$ ensures that quantum evolution is reversible and probability-preserving. Single-qubit gates act on one qubit at a time; two-qubit gates introduce interactions between pairs of qubits. The gates relevant to the `ZZFeatureMap` used in this paper are:

- **Hadamard gate** (H): places a qubit in equal superposition,

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \qquad H|0\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}}, \quad H|1\rangle = \frac{|0\rangle - |1\rangle}{\sqrt{2}}. \quad \text{(A.3)}$$

- **Phase gate** ($P(\theta)$, equivalent to $R_z(\theta)$ up to global phase): rotates the phase of the $|1\rangle$ component by angle θ,

$$P(\theta) = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\theta} \end{pmatrix}. \tag{A.4}$$

  In the `ZZFeatureMap`, $\theta = 2x_j$ encodes the j-th feature value as a rotation angle (angle encoding).

- **Controlled-phase gate** ($C_P(\theta)$): a two-qubit gate that applies a phase $e^{i\theta}$ only when both qubits are in state $|1\rangle$,

$$C_P(\theta) = \mathrm{diag}(1,\ 1,\ 1,\ e^{i\theta}). \tag{A.5}$$

  In the `ZZFeatureMap`, $\theta = 2\phi_{jl}(\mathbf{x}) = 2(\pi - x_j)(\pi - x_l)$ encodes pairwise feature products into the two-qubit interaction.

A *quantum circuit* is a sequence of gates applied to an initial state, read from left to right. Figure 1 shows the gate-level diagram of the `ZZFeatureMap` for $k = 4$ qubits and one repetition.

Appendix A.3. Measurement and Probability Estimation

At the end of a quantum circuit, the state is *measured* in the computational basis. Measurement is probabilistic: the outcome $b \in \{0, 1\}^k$ is observed with probability $|c_b|^2$, and the quantum state collapses to $|b\rangle$ upon measurement. Quantum algorithms typically require many repeated executions (*shots*) of the same circuit to estimate these probabilities.

For quantum kernel computation, the quantity of interest is the probability of the all-zeros outcome $|00\cdots0\rangle$, which equals the squared overlap between two encoded states:

$$K_Q(\mathbf{x}, \mathbf{x}') = \Pr[\text{measure } 0^k] = \left|\langle 0|^{\otimes k}\, U_\phi^\dagger(\mathbf{x})\, U_\phi(\mathbf{x}')\, |0\rangle^{\otimes k}\right|^2. \tag{A.6}$$

In this paper, K_Q is computed via *statevector simulation* — exact numerical integration of the Schrödinger equation — which provides shot-noise-free estimates. On real hardware, a finite number of shots introduces statistical noise; see Section 5.5.

Appendix A.4. Entanglement and Quantum Feature Maps

A key property of multi-qubit systems is *entanglement*: quantum states that cannot be written as a product $|\psi_1\rangle \otimes |\psi_2\rangle \otimes \cdots \otimes |\psi_k\rangle$ of individual qubit states. Entanglement is generated by two-qubit gates such as C_P. In the `ZZFeatureMap`, the entangling layer creates correlations between pairs of qubits that depend on products of input features $\phi_{jl}(\mathbf{x}) = (\pi - x_j)(\pi - x_l)$. This is the mechanism by which the quantum kernel captures cross-feature correlations that are absent from classical kernels operating on features independently.

A *quantum feature map* $U_\phi(\mathbf{x})$ is a parameterized quantum circuit that encodes a classical data vector $\mathbf{x} \in \mathbb{R}^k$ into a quantum state $|\phi(\mathbf{x})\rangle = U_\phi(\mathbf{x})|0\rangle^{\otimes k}$. The resulting *quantum kernel* $K_Q(\mathbf{x}, \mathbf{x}') = |\langle\phi(\mathbf{x})|\phi(\mathbf{x}')\rangle|^2$ is a valid positive semi-definite kernel and can be used as a drop-in replacement for classical kernels in a support vector machine (Schuld and Killoran, 2019; Havlíček et al., 2019). The expressibility of the kernel — its ability to represent complex decision boundaries — grows with the number of qubits k and the circuit depth, but so does the risk of *kernel concentration* (Section 5.2), which limits practical advantage to datasets with sufficiently structured feature geometry (Thanasilp et al., 2024; Huang et al., 2021).

References

Alvarez-Estevez, D., 2025. Benchmarking Quantum Machine Learning Kernel Training for Classification Tasks. IEEE Transactions on Quantum Engineering 6, TQE.2025. doi:doi:10.1109/TQE.2025.3541882, arXiv:2408.10274.

Belov, A., Eroshenko, E., Oleneva, V., Struminsky, A., Yanke, V., 2001. What determines the magnitude of forbush decreases? Advances in Space Research 27, 625–630. doi:doi:10.1016/S0273-1177(01)00095-3.

Belov, A., Papaioannou, A., Abunina, M., Dumbovic, M., Richardson, I.G., Heber, B., Kuhl, P., Herbst, K., Anastasiadis, A., Vourlidas, A., et al., 2021. On the rigidity spectrum of cosmic-ray variations within propagating interplanetary disturbances: Neutron monitor and soho/ephin observations at 1–10 gv. The Astrophysical Journal 908, 5. doi:doi:10.3847/1538-4357/abd724.

Belov, A.V., 2008. Forbush effects and their connection with solar, interplanetary and geomagnetic phenomena. Proceedings of the International Astronomical Union 4, 439–450. doi:doi:10.1017/S1743921309029676.

Biamonte, J., Wittek, P., Pancotti, N., Rebentrost, P., Wiebe, N., Lloyd, S., 2017. Quantum machine learning. Nature 549, 195–202. doi:doi:10.1038/nature23474.

Bobra, M.G., Couvidat, S., 2015. Solar flare prediction using SDO/HMI vector magnetic field data with a machine-learning algorithm. The Astrophysical Journal 798, 135. doi:doi:10.1088/0004-637X/798/2/135.

Camporeale, E., 2019. The challenge of machine learning in space weather: Nowcasting and forecasting. Space Weather 17, 1166–1207. doi:doi:10.1029/2018SW002061.

Cane, H.V., 2000. Coronal mass ejections and forbush decreases. Space Science Reviews 93, 55–77. doi:doi:10.1023/A:1026532125747.

Cerezo, M., Verdon, G., Huang, H.Y., Cincio, L., Coles, P.J., 2022. Challenges and opportunities in quantum machine learning. Nature Computational Science 2, 567–576. doi:doi:10.1038/s43588-022-00311-3.

Cortes, C., Vapnik, V., 1995. Support-vector networks. Machine Learning 20, 273–297. doi:doi:10.1007/BF00994018.

Cover, T.M., Thomas, J.A., 2006. Elements of Information Theory. 2nd ed., Wiley, Hoboken, NJ. doi:doi:10.1002/047174882X.

Dumbović, M., Kramarić, L., Benko, I., Heber, B., Vršnak, B., 2024. A new method of measuring forbush decreases. Astronomy & astrophysics 683, A168. doi:doi:10.1051/0004-6361/202346969.

Fawcett, T., 2006. An introduction to ROC analysis. Pattern Recognition Letters 27, 861–874. doi:doi:10.1016/j.patrec.2005.10.010.

Forbush, S.E., 1938. On world-wide changes in cosmic-ray intensity. Physical Review 54, 975–988. doi:doi:10.1103/PhysRev.54.975.

Gutierrez, C., Dasso, S., Démoulin, P., Janvier, M., 2024. A comparison of forbush decreases driven by icmes and sirs. Journal of Atmospheric and Solar-Terrestrial Physics 259, 106232. doi:doi:10.1016/j.jastp.2024.106232.

Havlíček, V., Córcoles, A.D., Temme, K., Harrow, A.W., Kandala, A., Chow, J.M., Gambetta, J.M., 2019. Supervised learning with quantum-enhanced feature spaces. Nature 567, 209–212. doi:doi:10.1038/s41586-019-0980-2.

Huang, H.Y., Broughton, M., Mohseni, M., Babbush, R., Boixo, S., Neven, H., McClean, J.R., 2021. Power of data in quantum machine learning. Nature Communications 12, 2631. doi:doi:10.1038/s41467-021-22539-9.

Incudini, M., Martini, F., Pierro, A.D., 2025. Toward useful quantum kernels. Advanced Quantum Technologies 8, 2300298. doi:doi:10.1002/qute.202300298.

Jaderberg, B., Gentile, A.A., Ghosh, A., Elfving, V.E., Jones, C., Vodola, D., Manobianco, J., Weiss, H., 2024. Potential of quantum scientific machine learning applied to weather modeling. Physical Review A 110, 052423. doi:doi:10.1103/PhysRevA.110.052423.

Jerry-Okafor, O., Onuchukwu, C., Okoye, V., Okoli, L., 2024. Statistical study and analysis of the parameters in forbush effects and interplanetary disturbances (feid) during solar cycles 23 and 24. Asian Research Journal of Current Science 6, 191–218. URL: https://hal.science/hal-05086121/.

King, J.H., Papitashvili, N.E., 2005. Solar wind spatial scales in and comparisons of hourly wind and ACE plasma and magnetic field data. Journal of Geophysical Research: Space Physics 110, A02104. doi:doi:10.1029/2004JA010649.

Kruskal, W.H., Wallis, W.A., 1952. Use of ranks in one-criterion variance analysis. Journal of the American statistical Association 47, 583–621. doi:doi:10.1080/01621459.1952.10483441.

Kübler, J.M., Buchheim, S., Sweke, R., Eisert, J., 2021. The inductive bias of quantum kernels. Advances in Neural Information Processing Systems 34, 12137–12148. URL: https://proceedings.neurips.cc/paper_files/paper/2021/hash/69adc1e107f7f7d035d7baf04342e1ca-Abstract.html.

Liu, Y., Arunachalam, S., Temme, K., 2021. A rigorous and robust quantum speed-up in supervised machine learning. Nature Physics 17, 1013–1017. doi:doi:10.1038/s41567-021-01287-z.

Lockwood, J.A., 1971. Forbush decreases in the cosmic radiation. Space Science Reviews 12, 658–715. doi:doi:10.1007/BF00173346.

Mavromichalaki, H., Papaioannou, A., Plainaki, C., Sarlanis, C., Souvatzoglou, G., Gerontidou, M., Papailiou, M., Eroshenko, E., Belov, A., Yanke, V., et al., 2011. Applications and usage of the real-time neutron monitor database. Advances in Space Research 47, 2210–2222. doi:doi:10.1016/j.asr.2010.02.019.

Melkumyan, A., Belov, A., Shlyk, N., Abunina, M., Abunin, A., Oleneva, V., Yanke, V., 2023. Forbush decreases and geomagnetic disturbances: 1. events associated with different types of solar and interplanetary sources. Geomagnetism and Aeronomy 63, 686–700. doi:doi:10.1134/S0016793223600650.

Melkumyan, A., Belov, A., Shlyk, N., Abunina, M., Abunin, A., Oleneva, V., Yanke, V., 2024. Forbush decreases and geomagnetic disturbances: 2. comparison of solar cycles 23–24 and events with sudden and gradual commencement. Geomagnetism and Aeronomy 64, 32–43. doi:doi:10.1134/S0016793223600911.

Miroszewski, A., Mielczarek, J., Czelusta, G., Szczepanek, F., Grabowski, B., Le Saux, B., Nalepa, J., 2023. Detecting clouds in multispectral satellite images using quantum-kernel support vector machines. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 16, 7601–7613. doi:doi:10.1109/JSTARS.2023.3304122.

Nielsen, M.A., Chuang, I.L., 2010. Quantum computation and quantum information. Cambridge university press.

Okike, O., 2020. Automated detection of simultaneous/non-simultaneous forbush decreases and the associated cosmic ray phenomena. Journal of Atmospheric and Solar-Terrestrial Physics 211, 105460. doi:doi:10.1016/j.jastp.2020.105460.

Okoye, V., Onuchukwu, C., Okoli, L., Jerry-Okafor, O., 2024. Statistical study of forbush effects and interplanetary disturbances (feid) from 1957-2019. Asian Basic and Applied Research Journal 6, 116–131. URL: https://hal.science/hal-05083994/.

Otgonbaatar, S., Datcu, M., 2022. Classification of remote sensing images with parameterized quantum gates. IEEE Geoscience and Remote Sensing Letters 19, 1–5. doi:doi:10.1109/LGRS.2021.3108014.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D.,

Brucher, M., Perrot, M., Duchesnay, E., 2011. Scikit-learn: Machine learning in Python. Journal of Machine Learning Research 12, 2825–2830. URL: https://jmlr.org/papers/v12/pedregosa11a.html.

Perez-Navarro, J.D., Sierra-Porta, D., 2024. Derivative-aligned anticipation of forbush decreases from entropy and fractal markers. The Open Journal of Astrophysics 9. doi:doi:10.33232/001c.157585.

Petukhova, A., Petukhov, S., Starodubtsev, S., 2020. Forbush decrease characteristics in a magnetic cloud. Space Weather 18, e2020SW002616. doi:doi:10.1029/2020SW002616.

Qiskit contributors, 2024. Qiskit: An open-source framework for quantum computing. Zenodo URL: https://doi.org/10.5281/zenodo.2573505, doi:doi:10.5281/zenodo.2573505.

Rebentrost, P., Mohseni, M., Lloyd, S., 2014. Quantum support vector machine for big data classification. Physical Review Letters 113, 130503. doi:doi:10.1103/PhysRevLett.113.130503.

Richardson, I.G., Cane, H.V., 2010. Near-earth interplanetary coronal mass ejections during solar cycle 23 (1996–2009): Catalog and summary of properties. Solar Physics 264, 189–237. doi:doi:10.1007/s11207-010-9568-6.

Richardson, I.G., Cane, H.V., 2011. Galactic cosmic ray intensity response to interplanetary coronal mass ejections/magnetic clouds in 1995–2009. Solar Physics 270, 609–627. doi:doi:10.1007/s11207-011-9774-x.

Schuld, M., 2021. Supervised quantum machine learning models are kernel methods. arXiv preprint arXiv:2101.11020. URL: https://arxiv.org/abs/2101.11020.

Schuld, M., Killoran, N., 2019. Quantum machine learning in feature Hilbert spaces. Physical Review Letters 122, 040504. doi:doi:10.1103/PhysRevLett.122.040504.

Sierra-Porta, D., Petro-Ramos, J., Ruiz-Morales, D., Herrera-Acevedo, D., García-Teheran, A., Alvarado, M.T., 2024. Machine learning models for predicting geomagnetic storms across five solar cycles using dst index and heliospheric variables. Advances in Space Research 74, 3483–3495. doi:doi:10.1016/j.asr.2024.08.031.

Tange, O., 2011. GNU parallel: The command-line power tool. ;login: The USENIX Magazine 36, 42–47. URL: https://www.gnu.org/software/parallel/, doi:doi:10.5281/zenodo.7347980.

Thanasilp, S., Wang, S., Cerezo, M., Holmes, Z., 2024. Exponential concentration in quantum kernel methods. Nature communications 15, 5200. doi:doi:10.1038/s41467-024-49287-w.

Upendran, V., Cheung, M.C.M., Hanasoge, S., Krishnamurthi, G., 2020. Solar wind prediction using deep learning. Space Weather 18, e2020SW002478. doi:doi:10.1029/2020SW002478.

Wang, X., Du, Y., Luo, Y., Tao, D., 2021. Towards understanding the power of quantum kernels in the nisq era. Quantum 5, 531. doi:doi:10.22331/q-2021-08-30-531.

Wilcoxon, F., 1945. Individual comparisons by ranking methods. Biometrics bulletin 1, 80–83. doi:doi:10.2307/3001968.

Ye, Y., Liu, J., Hao, Y., Cui, J., Feng, X., 2025. Assessing the geoeffectiveness of stream interaction regions through physically interpretable machine learning. The Astrophysical Journal 993, 10. doi:doi:10.3847/1538-4357/ae0454.