# Quantum Kernel Methods for Malware beaconing Detection


Maissa BEJI

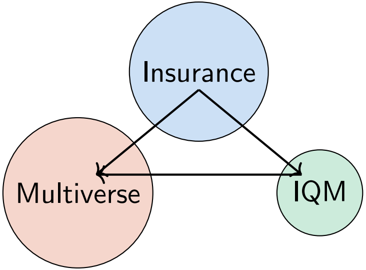
Industrial Supervisors: Luc ANDREA, Llorenç ESPINOSA, Multiverse Computing
Academic Supervisor: Peter BROWN, Télécom Paris

# Outline

# Context: AQACYB Project

**Quantum Advantage for Cyber Threat Analysis**

- Focus: Quantum-enhanced anomaly detection
- Target: Malware beaconing detection
- Hardware: IQM's 20-qubit quantum processor



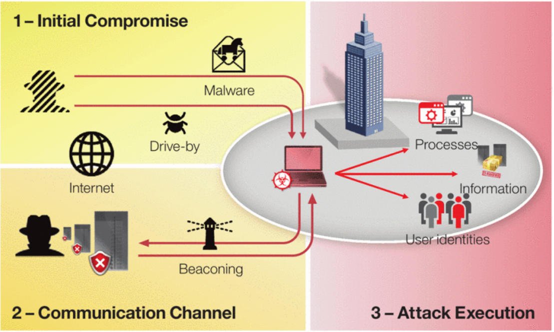

## Challenge

$\Rightarrow$ Malware beaconing hides in encrypted periodic traffic to sustain long-term access.

# Problem Statement

**How Attackers Control Infected Machines:**

- Malware opens **stealth channel** to C&C server
- Connection **starts from inside**
- Regular "calls home" for:
  - Presence announcement
  - New instructions
- This repeated pattern = **Beaconing**




*Source: Hu et al., IEEE, 2016.*

# Problem Statement

## Key Challenges in Detecting Beaconing

### Traffic Looks Normal

- Blends with regular encrypted traffic

### False Positives

- Legitimate apps appear beacon-like

### Scalability

- Millions of daily connections

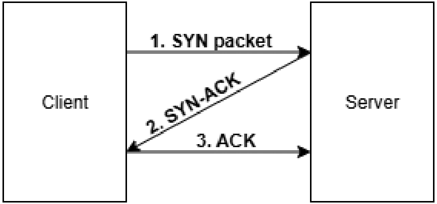
### Imbalanced Data

- Malicious traffic is rare

# Dataset Evolution

## Original Dataset

- 721,899 normal events, 2,523 attack events
- Highly imbalanced (99.65% normal)

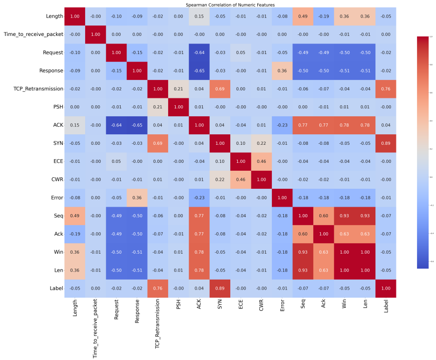



*TCP three-way handshake*

# Dataset Evolution

## Original Dataset

- Data leakage issues with SYN and TCP flags
- No time dependency captured in baseline features




*Spearman correlation matrix*

# Dataset Evolution

## Feature Engineering Approach

For each source-destination IP pair:

- Mean and std of time intervals between connections
- Number of new connections
- Mean and std of packet sizes
- Number of unique source ports
- Number of unique destination ports

**119 Non-attacks**
**1 Attack**

**Insufficient for evaluation**

# Final Dataset Construction

### Data Sources

- IoT traffic
- Normal logs
- Client malware
- CTU-13 dataset
- Laptop traffic

### Dataset Size

**6,459**

Training logs

**1,741**

Testing logs

**126**

Labeled attacks

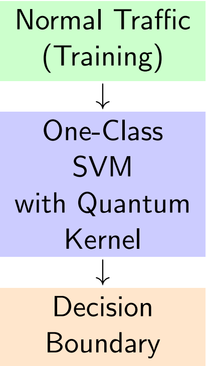
### Flow Definition

- Source IP
- Dest. IP
- Protocol
- Ext. port

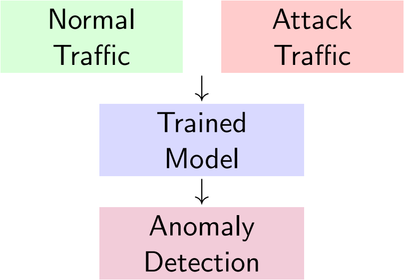
**Note:** IP-only insufficient for periodicity detection

# Unsupervised Learning Setup

**Training Phase:**




**Testing Phase:**



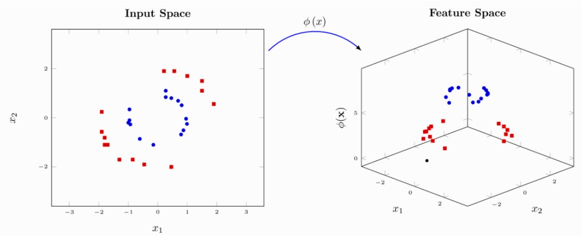

- Evaluation using F1-score

# Quantum Feature Maps

## Data Encoding

Classical data $x \in \mathbb{R}^n$ encoded into quantum states:

$$|\phi(x)\rangle = U(x)|0\rangle^{\otimes n}$$

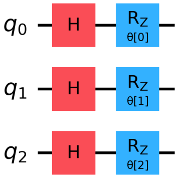



*Source: squLEARN - Quantum Kernel Methods Example*

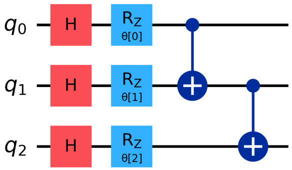
# Quantum Feature Maps

## Tested Feature Maps:

- **Z Feature Map**: Separable rotations



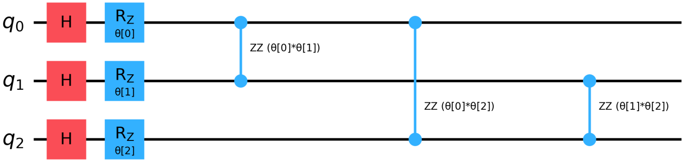

- **CNOT Feature Map**: Entangling gates



# Quantum Feature Maps

## Tested Feature Maps:

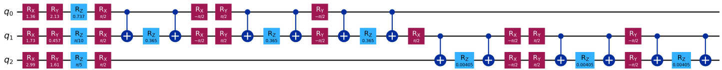
- **IQP-like Circuit**: Commuting gates



# Quantum Feature Maps

## Tested Feature Maps:

- **Hamiltonian Evolution**: Many-body inspired



## Mathematical Definition

$$|x_i\rangle = \left( \prod_{j=1}^{n} \exp\left( -i \frac{t}{T} x_{ij} H_j^{XYZ} \right) \right)^T \bigotimes_{j=1}^{n+1} |w_j\rangle$$

where: $H_j^{XYZ} = X_j X_{j+1} + Y_j Y_{j+1} + Z_j Z_{j+1}$

# Quantum Kernel Computation

## Fidelity Kernel

$$\kappa(x_i, x_j) = |\langle\phi(x_i)|\phi(x_j)\rangle|^2$$

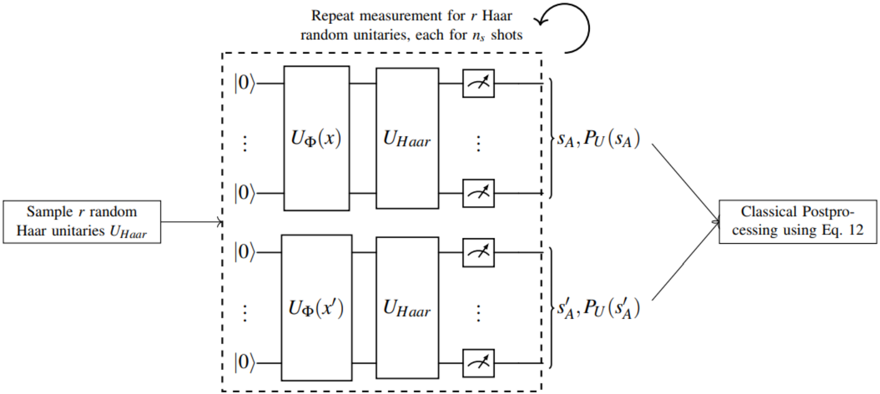
Direct measure of state similarity

## Main Challenge

- **Scalability**: Quadratic growth with dataset size

# Approach 1: Randomized Measurements

## Protocol Overview




Source: *Kölle et al., arXiv:2407.20753 (2024)*

# Approach 1: Randomized Measurement

## Mathematical Framework

Estimate kernel entries using random basis measurements:

$$K(x_i, x_j) = 2^N \sum_{s_A, s'_A} (-2)^{-H(s_A, s'_A)} \overline{p_U^{(i)}(s_A) p_U^{(j)}(s'_A)}$$

# Approach 1: Randomized Measurement

## Implementation Steps

1. Prepare the target quantum state $|\phi\rangle$
2. Sample $r$ random Haar unitaries $\{U_{\text{Haar}}\}$
3. For each unitary: apply $U_\Phi(x)$, then $U_{\text{Haar}}$, then measure in computational basis
4. Repeat each measurement $n_s$ times ($n_{\text{shots}}$ per unitary)
5. Collect measurement statistics across all $r \times n_s$ runs
6. Estimate fidelity using cross-correlations between outcome probabilities

- **Error Scaling:** $\Delta K \propto \frac{1}{n_s\sqrt{r}}$
- **Error mitigation:**

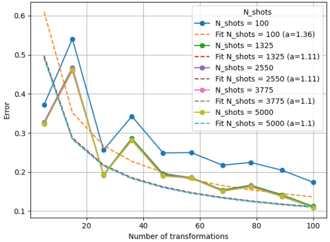
$$K_m(x_i, x_j) = \frac{\mathrm{Tr}(\rho_i \rho_j)}{\sqrt{\mathrm{Tr}(\rho_i^2)\mathrm{Tr}(\rho_j^2)}}$$

- **Complexity:** quantum $n \cdot r_s \cdot n_s$, classical post-processing $n^2$

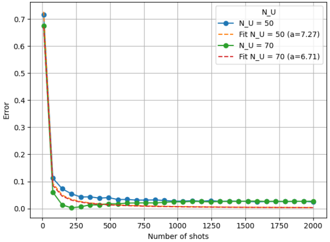
# Approach 1: Randomized Measurement

## Experimental results

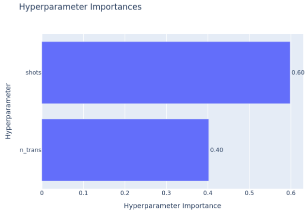
- The statistical error $\Delta K$ is the absolute deviation between the estimated purity and its ideal value.




*Statistical error vs. number of random transformations*




*Statistical error vs. number of shots*

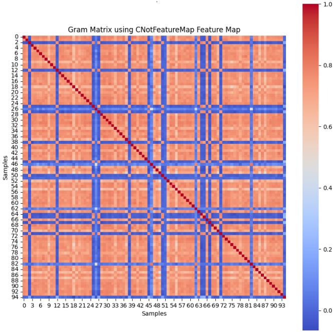



*Optuna hyperparameter importance scores*

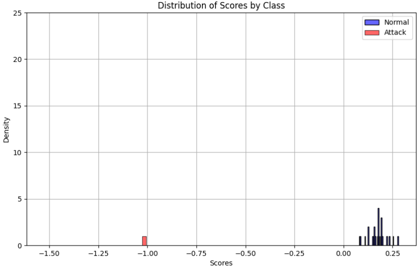
# Approach 1: Randomized Measurement

## Experimental results

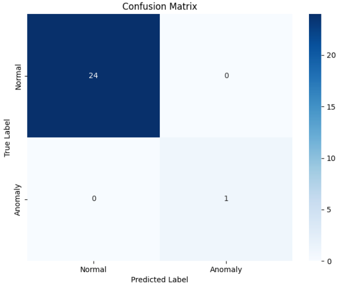
- Evaluation on the malware dataset.




*Kernel Matrix Visualization*




*Test Score Distribution*

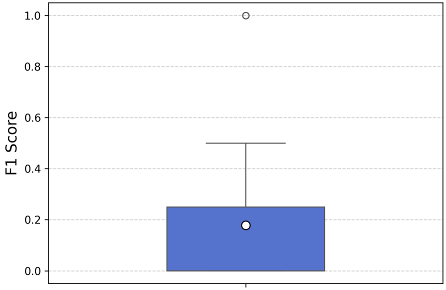



*Confusion matrix*

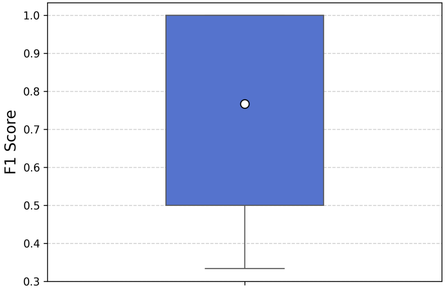
# Approach 1: Randomized Measurement

## Experimental results

- F1 Score Variability Across Experimental Repetitions




*30 transformations, 400 shots, 1 layer CNOT (29 experiments)*




*10 transformations, 180 shots, 1 layer CNOT (9 experiments)*

# Approach 1: Randomized Measurement

**Conclusions:**

1. **Error Scaling:** $\Delta K \propto \frac{1}{n_s \sqrt{r}}$
2. Strong results on malware dataset with 7 qubits.
3. High variance in performance

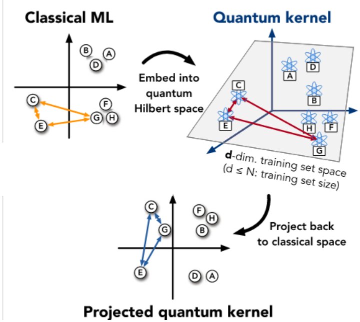
**Challenges:**

- Statistical error plateaus
- Concentration effects persist
- Requires error mitigation

# Approach 2: Projected Quantum Kernels

## Core Concept

Extract classical features from reduced density matrices




*Source: Huang et al., Nature Communications (2021)*

# Approach 2: Projected Quantum Kernels

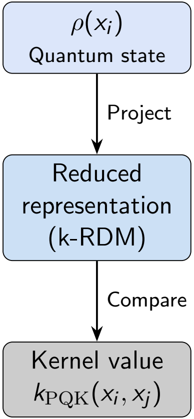
## Core Concept

- **k-particle reduced density matrix (k-RDM)** approach:

$$\rho_K(x_i) = \mathrm{Tr}_{\bar{K}}[\rho(x_i)]$$

**Kernel Types:**

- **RBF:** $k(x_i, x_j) = \exp(-\gamma \| f(x_i) - f(x_j) \|^2)$
- **Sigmoid:** $k(x_i, x_j) = \tanh(\alpha \langle f(x_i), f(x_j) \rangle + c)$

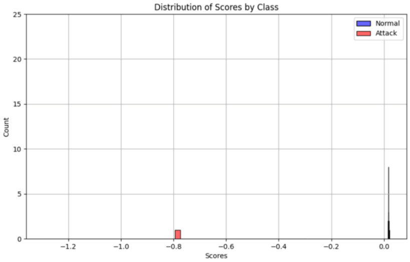



Projection $\Rightarrow$ Classical feature space

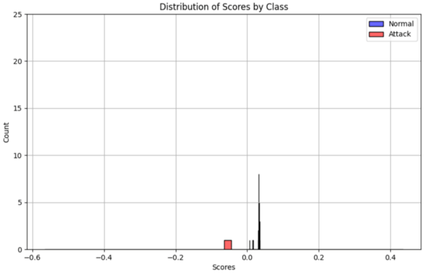
# Approach 2: Projected Quantum Kernels

## Evaluation on Malware Dataset

- **Sigmoid Kernel**




*Test scores with Hamiltonian feature map*

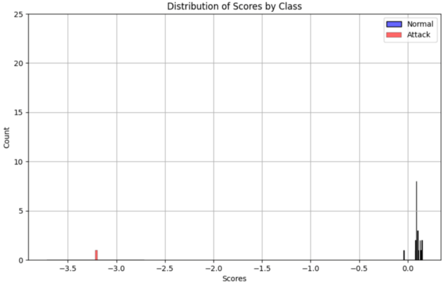



*Test scores with CNOT feature map*

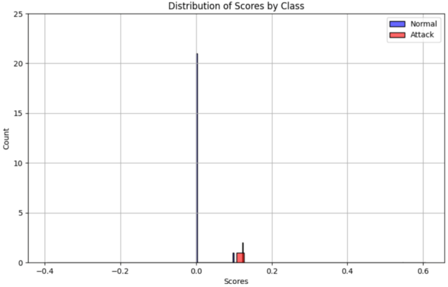
# Approach 2: Projected Quantum Kernels

## Evaluation on Malware Dataset

- **RBF Kernel**




*Test scores with Hamiltonian feature map*




*Test scores with CNOT feature map*

# Approach 2: Projected Quantum Kernels

## Evaluation on Open-Source Network Dataset

- **Performance comparison for different training set sizes**

| **Data size** | **PQKs** | | | | **Classical** | | | |
|---|---|---|---|---|---|---|---|---|
| | F1 | Recall | Precision | Acc. | F1 | Recall | Precision | Acc. |
| 200 | 0.2308 | 0.1313 | 0.9524 | 0.9524 | 0.7467 | 0.6936 | 0.8087 | 0.6140 |
| 1000 | 0.1193 | 0.0634 | 1 | 0.2316 | 0.2368 | 0.1422 | 0.7065 | 0.2477 |

**Takeaway:** PQKs perform consistently worse than the classical kernel, with F1-scores dropping further as data size increases.

# Approach 2: Projected Quantum Kernels

## Evaluation on Open-Source Network Dataset

- **Varying k in the k-reduced density matrices**

| k | F1 | Recall | Precision | Acc. |
|---|---|---|---|---|
| 1 | 0.2308 | 0.1313 | 0.9524 | 0.2819 |
| 2 | 0.2031 | 0.1138 | 0.9454 | 0.2675 |
| 3 | 0.1996 | 0.1116 | 0.9444 | 0.2657 |
| 4 | 0.2136 | 0.1203 | 0.9482 | 0.2728 |

**Takeaway:** Performance is nearly flat across $k$; best and most efficient choice is $k = 1$.

# Approach 2: Projected Quantum Kernels

**Conclusion:**

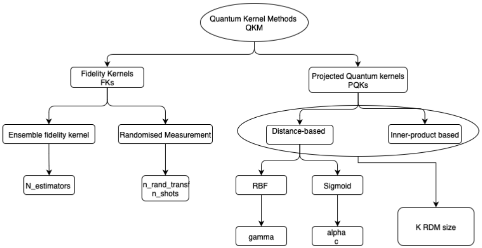
- Hamiltonian mapping more robust than CNOT.
- Classical kernels often outperform PQKs; RBF ($\gamma = 1$) limits gains.
- Minimal impact; $k = 1$ optimal.
- PQKs highly dependent on feature map and kernel choice.

**Advantages:**

- Avoids concentration effects
- Computationally efficient
- No fidelity estimation needed
- Stable across runs

# Study Design: Quantum Kernels and Feature Maps

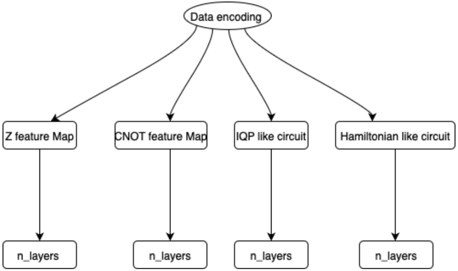
- **Focus**: Hyperparameter benchmarking for quantum anomaly detection
- **Retained methods**:
  - *EFK*: Scalable fidelity-based
  - *PQK*: Distance-based with $k$-RDMs
- **Excluded**: Randomized Measurement (scaling issues)




*Quantum kernels and hyperparameters.*

# Study Design: Quantum Kernels and Feature Maps

- **Focus**: Hyperparameter benchmarking for quantum anomaly detection
- **Retained methods**:
  - *EFK*: Scalable fidelity-based
  - *PQK*: Distance-based with $k$-RDMs
- **Excluded**: Randomized Measurement (scaling issues)




*Quantum data encoding methods.*

# Experimental setup:

## Hyperparameter search space:

- **Feature map:** {Z feature map, CNOT feature map, Hamiltonian-like circuit, IQP-like circuit}, number of layers $\in [1, 5]$
- **EFK:** number of estimators $\in [1, 5]$
- **PQK_RBF:** $k \in [1, 3]$, $\gamma \in [10^{-3}, 10^{3}]$
- **PQK_Sigmoid:** $k \in [1, 5]$, $\alpha \in [-1, 1]$, $c_{\text{sigmoid}} \in [10^{-3}, 10^{3}]$

## Metrics:

- F1-score

$$F_1 = \frac{2 \cdot TP}{2 \cdot TP + FP + FN}$$

- Geometric difference

$$g_{C \to Q} = \| \sqrt{K_Q} (K_C)^{-1} \sqrt{K_Q} \|_\infty$$

# Dataset Configuration

**Training Dataset:**

- 200 datapoints with 5 selected features
- Results in 5-qubit quantum system

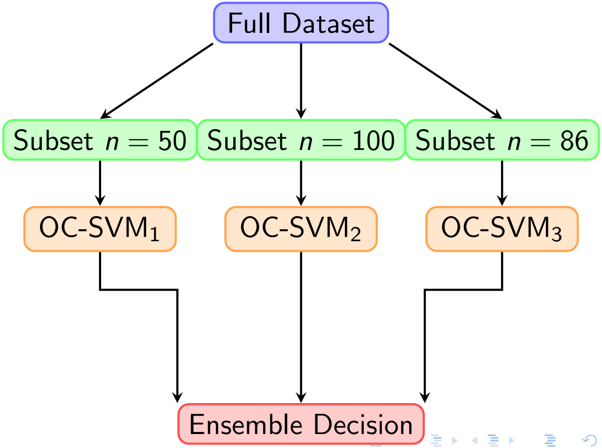
**Test Dataset:**

- 100 non-attack samples
- 457 attack samples
- Same 5-feature selection as training

# Variable Subsampling Ensembles with Inversion Test Kernels (Ensemble Fidelity Kernel)

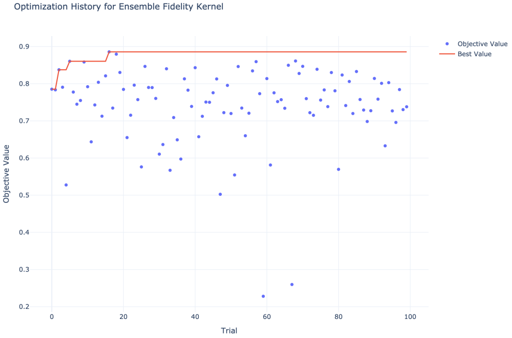
**Key Idea**

- Train multiple OC-SVMs on subsets of different sizes ($n_i$).
- Each subset $\to$ different decision boundary.
- Aggregate predictions (average $\to$ reduce variance, max $\to$ reduce bias).
- Scalable: training complexity $\sim \lfloor \frac{n}{100} \rfloor \times \left(\frac{50+100}{2}\right)^2$.

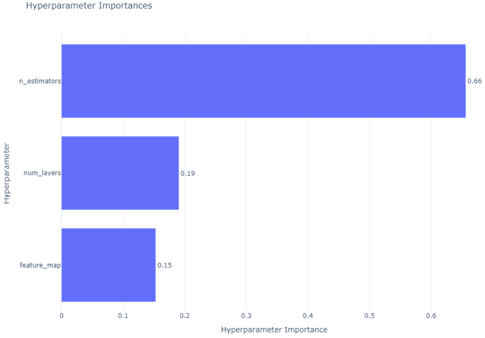


# Results: Ensemble Fidelity Kernel

- Hyperparameter optimization process



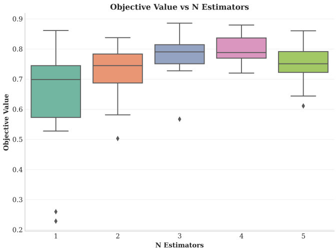

*Optimization history across 100 trials.*



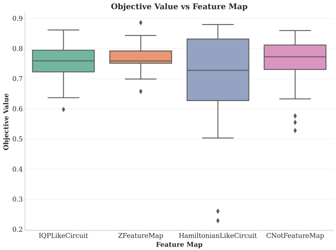

*Hyperparameter importance (Optuna analysis)*

# Results: Ensemble Fidelity Kernel

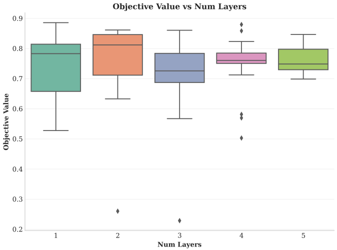
- Hyperparameter sensitivity analysis




*Number of estimators vs F1-score*

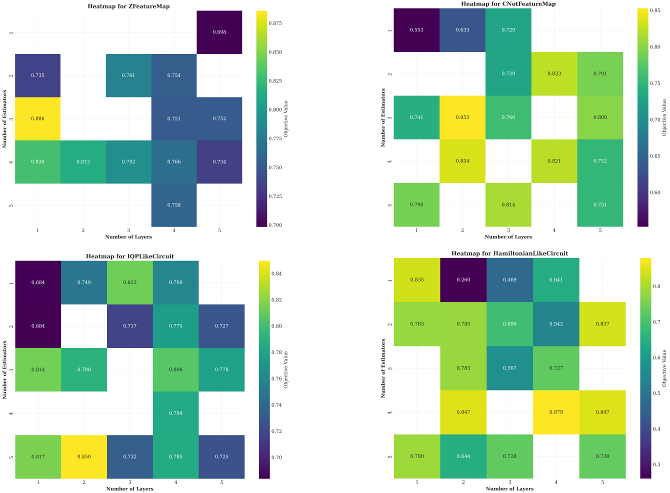



*Feature map vs F1-score*




*Number of layers vs F1-score.*

# Results: Ensemble Fidelity Kernel



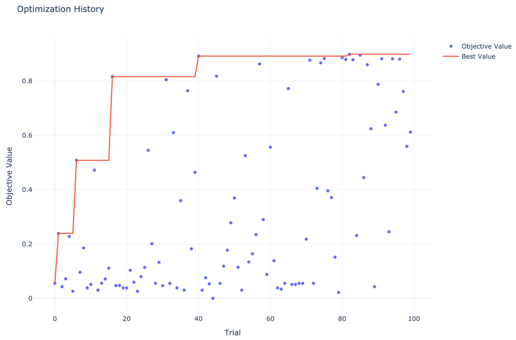
# Results: Ensemble Fidelity Kernel

**Takeaways:**

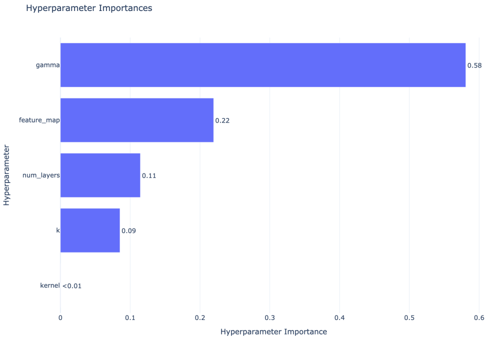
- Performance is sensitive to tuning but converges quickly within 15–20 trials
- Estimators matter most: best balance with 3–4 estimators.
- CNOT and Z maps are stable and reliable, while IQP and Hamiltonian are risky.
- Shallow circuits (1–2 layers) work best; deeper ones reduce performance.

# Results: Projected Quantum Kernel, RBF kernel

- Hyperparameter optimization process



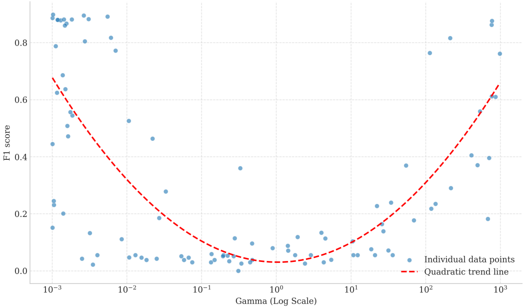

*Optimization history across 100 trials.*



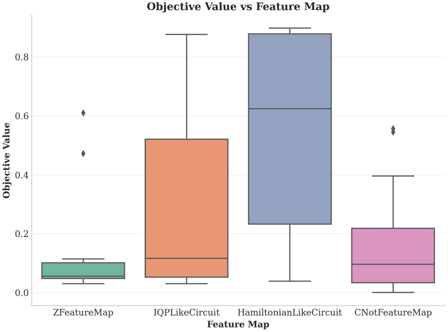

*Hyperparameter importance (Optuna analysis)*

# Results: Projected Quantum Kernel, RBF kernel

- Hyperparameter sensitivity analysis



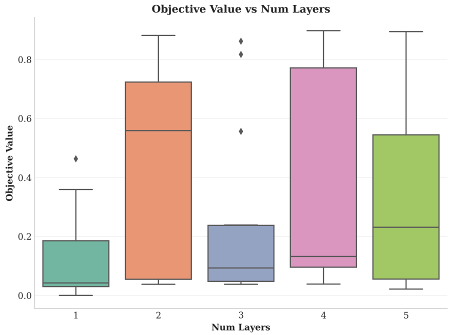

*$\gamma$ (log scale) vs F1-score*



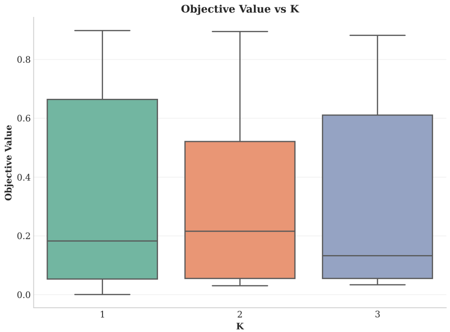

*Feature map vs F1-score.*

# Results: Projected Quantum Kernel, RBF kernel

- Hyperparameter sensitivity analysis

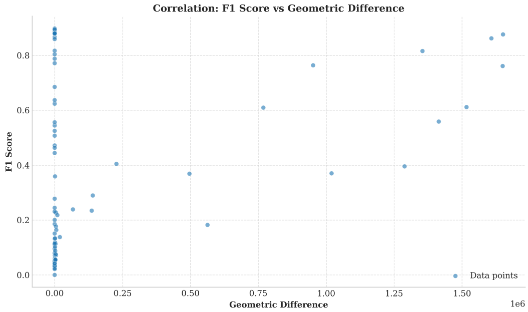



*Number of layers vs F1-score.*

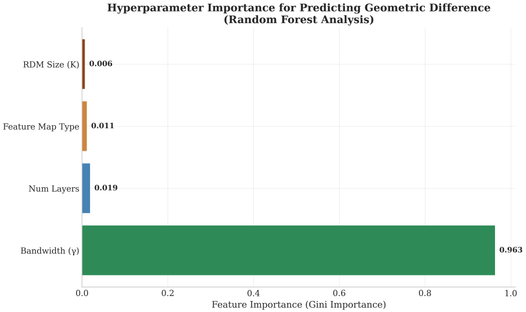



*k (K-RDM) vs F1-score.*

# Geometric difference analysis




*Correlation between F1 Score and Geometric Difference.*




*Hyperparameter importance for predicting geometric difference.*

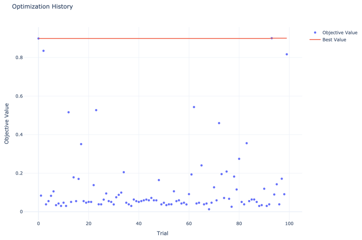
# Results: Projected Quantum Kernel, RBF kernel

## Takeaways

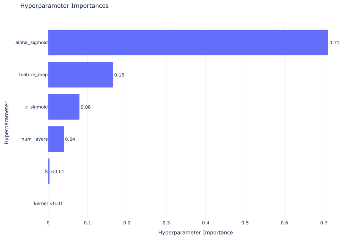
- $\gamma$ (RBF bandwidth) dominates performance; default $\gamma = 1$ is worst, extremes work best.
- Feature maps: IQP & Hamiltonian can peak but unstable; CNOT & Z more stable but weaker.
- Layers and subsystem size $k$ have little impact.
- Larger geometric difference often means lower F1, showing classical $\gamma$ overshadows quantum effects.

# Results: Projected Quantum Kernel, sigmoid kernel

- Hyperparameter optimization process




*Optimization history across 100 trials.*




*Hyperparameter importance (Optuna analysis)*

**Takeaway**

Sigmoid PQK: poor vs RBF, $\alpha$ dominates, F1 $< 0.2$.

# Conclusions

## Key Findings

- Quantum kernels viable for unsupervised anomaly detection
- Ensemble fidelity kernels show promise for scalability
- Projected quantum kernels dominated by classical parameters

## Methodological Contributions

- First comprehensive study of QKMs in unsupervised anomaly detection
- Benchmarking framework for quantum kernel comparison

# Project Impact

**Environmental**

- 105,288 CPU hours
- 47 GPU hours
- 38% CPU utilization
- Room for optimization

**Social**

- Quantum reversibility
- Enhanced interpretability
- AI transparency potential
- Ethical AI considerations

**Economic**

- Industry partnership
- Strategic cybersecurity
- Scientific publication
- Technology adoption

## Strategic Value

Real-world collaboration between academia, industry, and quantum hardware providers addressing critical cybersecurity challenges.

# Future Work

**Technical Directions:**

- Larger dataset experiments
- Noisy quantum simulators
- Hardware implementation
- Trainable kernel methods
- Quantum autoencoders comparison

**Research Questions:**

- Can quantum advantage emerge at scale?
- How to mitigate concentration effects?
- How does noise affect performance?

## Open Challenge

Finding quantum kernel constructions that provide genuine advantages over classical methods while remaining computationally tractable.

# Personal Reflection

## Technical Skills Gained

- Quantum machine learning algorithms
- Software engineering practices (Git, Docker, MLflow)
- Scientific literature review and analysis
- Hyperparameter optimization techniques

## Professional Development

- Team collaboration in research environment
- Presentation skills in consortium meetings
- Critical analysis and problem-solving
- Balancing scientific depth with practical constraints

# Thank you for your attention!

Questions?

**Contact:**
Maissa BEJI
Télécom Paris
maissa.beji@imt-atlantique.net