Balancing Expressivity and Learnability in Quantum Kernel Bandit Optimization

Yuqi Huang[1] **Vincent Y. F. Tan**[1,2] **Sharu Theresa Jose**[3]

[1]Department of Mathematics, National University of Singapore, Singapore
[2]Department of Electrical and Computer Engineering, National University of Singapore, Singapore
[3]School of Computer Science, University of Birmingham, Birmingham, United Kingdom

Abstract

We investigate Gaussian process (GP) bandit optimization with quantum kernels, assuming the mean reward function lies in the reproducing kernel Hilbert space (RKHS) induced by the quantum kernel. This setting is motivated by NISQ-era tasks such as quantum control, state preparation and variational quantum algorithms. While quantum kernels can offer a 'quantum advantage' via domain-specific inductive biases, naïvely using full, high-dimensional kernels increases model complexity and information gain, leading to higher cumulative regret and poor learnability. To address this, we propose projected quantum kernels and classical kernel approximation techniques that reduce feature dimensionality while preserving key quantum properties. Using these approximate kernels, we develop misspecified GP bandit algorithms and derive regret bounds that characterize the trade-off between approximation error and information gain. The regret bounds provide principled guidance for selecting the optimal model complexity. Empirically, our methods outperform full quantum kernels in sample efficiency, while substantially reducing computational overhead, enabling scalable GP optimization for quantum-native applications.

1 INTRODUCTION

Recent advances in quantum hardware have ushered in the era of *noisy intermediate-scale quantum* (NISQ) computing, characterized by quantum devices with limited qubits and high noise levels. Within this regime, quantum machine learning has emerged as a promising application, aiming to leverage these NISQ-era devices to accelerate learning from classical or quantum data. This includes discriminative and generative learning approaches based on quantum neural networks Schuld [2021] and quantum kernel methods Blank et al. [2020], Rebentrost et al. [2014], as well as sequential decision making problems.

Recently, there has been growing interest in developing quantum algorithms for classical bandit learning problems, aiming to achieve speedups in sequential decision-making. These include quantum algorithms for best arm identification [Casalé et al., 2020, Wang et al., 2025, Buchholz et al., 2025, Wang et al., 2021a], and for addressing exploration-exploitation tradeoff under both linear [Lumbreras et al., 2022, Wan et al., 2023, Wu et al., 2023] and non-linear reward models [Hikima et al., 2024, Dai et al., 2023]. However, these approaches assume access to a quantum reward oracle, which simplifies quantum algorithm design, but limits applicability in realistic settings where the classical reward distribution must be encoded into a quantum state.

In this work, we depart from this oracle-based paradigm, and study the *quantum kernel bandit* problem. Here, the learner interacts sequentially with an unknown reward function through noisy evaluations. We assume that the unknown mean reward function $f^* : \mathcal{X} \to \mathbb{R}$ lies in the RKHS $\mathcal{H}_{\kappa_Q}$ induced by a quantum kernel $\kappa_Q(\cdot,\cdot)$. At each round t, the learner selects an action $\mathbf{x}_t \in \mathcal{X}$ and observes a noisy reward $y_t = f^*(\mathbf{x}_t) + \eta_t$, where η_t is zero mean noise. The objective of the learner is to minimize the *cumulative regret* over T rounds, $R_T = \sum_{t=1}^{T}(f^*(\mathbf{x}^*) - f^*(\mathbf{x}_t))$, where $\mathbf{x}^* = \arg\max_{\mathbf{x}\in\mathcal{X}} f^*(\mathbf{x})$ denotes the optimal action in hindsight.

This formulation naturally captures a range of NISQ-era optimization tasks, including state preparation in quantum sensing [Schuff et al., 2020], experimental quantum control [Bukov et al., 2018], and the optimization of variational quantum algorithms [Wanner et al., 2025]. Here, actions $\mathbf{x}$ correspond to continuous parameters (e.g., gate angles or pulse amplitudes), and a fixed quantum embedding map $\boldsymbol{\rho}(\mathbf{x})$ induces a corresponding quantum kernel. The resulting quantum RKHS provides a natural hypothesis space for modelling the mean reward function $f^*(\mathbf{x})$, while reward feedback is obtained from noisy finite-shot measurements.

Proceedings of the 42nd Conference on Uncertainty in Artificial Intelligence (UAI 2026), PMLR 337:2274–2313.

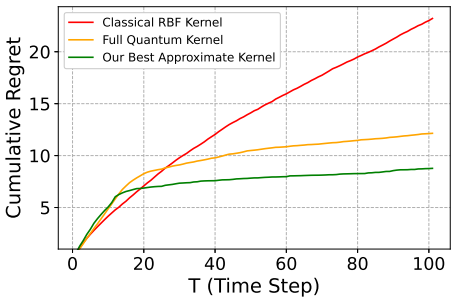


Figure 1: Cumulative regrets over $T = 100$ rounds (using the EC-GP-UCB algorithm) with three different modelling kernels on a synthetic task where the unknown reward function lies in a quantum RKHS. While the full quantum kernel beats the classical RBF kernel, our proposed *projected quantum kernel-based approximation* achieves an even smaller regret. See subsection 4.1 for experimental details.

We approach this problem using quantum Gaussian process (GP) models. Leveraging the duality between RKHS and GPs [Kanagawa et al., 2018], we encode the inductive bias (i.e., $f^* \in \mathcal{H}_{\kappa_Q}$) by placing a GP prior with the kernel κ_Q. While such models can theoretically outperform classical GPs with standard kernels such as RBF or Matérn (see Figure 1) [Smith et al., 2023], they face two unique challenges:

- The expressivity of quantum kernels on n-qubit systems scales exponentially with n, leading to high *information gain* and, consequently, excessive cumulative regret. This hinders the learnability of the quantum GP models.
- As the quantum system size n grows, kernel values concentrate exponentially, requiring exponentially many quantum measurements for accurate estimation [Thanasilp et al., 2024].

To address these challenges, we propose new GP algorithms that *approximate the full quantum RKHS* using either lower-dimensional quantum subspaces or quantum-inspired classical approximations. These methods reduce the model complexity, and thus the information gain, at the cost of *kernel misspecification*, which can impact the regret [Bogunovic and Krause, 2021]. Our main contributions are:

- **Quantum-Specific Learnability Barrier**: We identify the exponential scaling of the maximum information gain with the number of qubits n as the fundamental learnability barrier for GP bandits with fidelity quantum kernels. This statistical barrier, together with kernel concentration as n grows, motivates the use of approximate quantum-kernel bandit algorithms.
- **Algorithmic Framework**: We develop a new framework for approximate GP optimization in quantum kernel bandits by combining quantum kernel approximation techniques with GP (and linear) bandit algorithms. Our approximation approaches include: linear projected quantum kernels (LPQKs), operating on reduced quantum subspaces, and quantum-inspired classical approximations – Random Fourier Features (RFF) and Newton basis expansions.
- **Quantum-Structured Trade-off Analysis with Regret Bounds**: We then derive regret bounds that quantify the trade-off between information gain and kernel misspecification. In regimes where the full quantum kernel is excessively high-dimensional, we show that a suitably chosen approximate kernel achieves *strictly lower regret* than the full model (Figure 1). The bounds also guide the choice of approximation parameters, such as the number of qubits traced out in projected quantum kernels, or the number of random features used in RFF. Although some approximation and misspecified-bandit tools used in this work are kernel-general, our analysis specializes them to quantum kernels by exploiting the structure of quantum feature maps. In particular, LPQKs correspond to Pauli-weight projections of the n-qubit density-matrix feature space. For quantum encodings that admit a Fourier form, the frequencies are determined by eigenvalue gaps of the encoding generators. Thus, our results connect circuit and observable structure directly to the information-gain/misspecification trade-off in regret.
- **Computational Efficiency**: While naive GP optimization with a full quantum kernel incurs $\mathcal{O}(T^3)$ time complexity and costly circuit evaluations, our approximation-based methods leverage a D-dimensional feature map ($D \ll T$) to reduce the time complexity to $\mathcal{O}(TD^2 + D^3)$; or rely on smaller qubit subsystems, thereby enhancing the scalability.
- **Experimental Validation**: We evaluate our methods on synthetic and real quantum tasks (e.g., phase classification, variational quantum eigensolver optimization), showing that our approximations often *outperform the full quantum kernel* both in sample efficiency and computational cost, and that theoretical insights into approximation error can guide dimension selection, yielding near-optimal performance in practice.

Related Works: A common approach to optimizing the NISQ-era objectives is to use variational quantum algorithms (VQAs), which directly update the parameters of a parameterized quantum circuit [Cerezo et al., 2021]. These methods often rely on gradients estimated by finite differences or the parameter-shift rule, such gradient estimates can be measurement-intensive, and may suffer from barren plateaus [McClean et al., 2018, Wang et al., 2021b]. We take a different, gradient-free surrogate-optimization viewpoint: the energy or reward is treated as a noisy black-box function, and a quantum-kernel GP surrogate is used to select queries to balance exploration and exploitation.

Kernelized bandits (including GP bandits) traditionally as-

sume *realizability*: the unknown reward function f^* either lies in the RKHS associated with the kernel $\kappa(\mathbf{x}, \mathbf{x}')$, or is drawn from a mean-zero GP prior $\mathrm{GP}(0, \kappa(\mathbf{x}, \mathbf{x}'))$ [Valko et al., 2013, Srinivas et al., 2009]. However, in practice, the true RKHS or the exact kernel function is often unknown, making the realizability assumption too restrictive. Our work belongs to the *misspecified kernelized bandit* setting, where the learner optimizes using an approximate kernel rather than the ground truth. This paradigm was first formalized for linear bandits by Ghosh et al. [2017], who established a regret lower bound of $\Omega(\varepsilon\sqrt{dT})$ for ε-perturbed additive reward models with d-dimensional vectors $\mathbf{x}$. Subsequent research has deepened the understanding of misspecification in linear bandits Zanette et al. [2020], Neu and Olkhovskaya [2020]. More recently, Bogunovic and Krause [2021] and Camilleri et al. [2021] extended these insights to kernelized settings, demonstrating that misspecification introduces an additive regret term of $\mathcal{O}(\varepsilon\sqrt{\gamma_T T})$, where ε is the misspecification error and γ_T is the maximum information gain of the kernel.

While quantum kernels have been widely used in supervised and generative learning settings, their integration into probabilistic models like GPs remains relatively underexplored. Smith et al. [2023] pioneered the use of GP surrogates with quantum kernels for variational quantum eigensolver (VQE) optimization, showing that encoding physical knowledge into the kernel can outperform GPs using classical kernels (e.g., RBF, Matérn). Nicoli et al. [2023] proposed a quantum-inspired classical kernel, termed the VQE-kernel, for Bayesian optimization of VQE problems. Other recent works, such as Rapp and Roth [2024] and Guo et al. [2024], explore quantum GP regression with NISQ-compatible quantum kernels, highlighting the promise of quantum probabilistic models in practical settings.

In contrast to prior work, we study GP bandit optimization with quantum kernels under kernel approximation, bridging the misspecified kernel bandit literature with quantum kernel methods. Our analysis explicitly characterizes how approximating high-dimensional quantum kernels impacts information gain, regret, and computational complexity.

2 PROBLEM SETTING

2.1 QUANTUM KERNEL BANDIT LEARNING

We consider a sequential online learning problem where at each round t, the learner chooses an action $\mathbf{x}_t \in \mathcal{X}$ from the finite action set $\mathcal{X}$. Corresponding to the chosen action, the learner observes a noisy reward

$$y_t = f^*(\mathbf{x}_t) + \eta_t, \tag{1}$$

where $f^* : \mathcal{X} \to \mathbb{R}$ is the true unknown reward function and η_t is the i.i.d. zero-mean noise. The goal of the learner is to minimize the *cumulative regret* over T rounds,

$$R_T = \sum_{t=1}^{T} f^*(\mathbf{x}^*) - \sum_{t=1}^{T} f^*(\mathbf{x}_t), \tag{2}$$

with respect to the optimal $\mathbf{x}^* = \arg\max_{\mathbf{x}\in\mathcal{X}} f^*(\mathbf{x})$. We assume that the f^* belongs to the RKHS $\mathcal{H}_{\kappa_Q}(\mathcal{X})$ defined by a quantum kernel $\kappa_Q(\mathbf{x}, \mathbf{x}')$, with bounded RKHS norm, i.e., $\|f^*\|_{\mathcal{H}_{\kappa_Q}} \leq B$ for some $B < \infty$.

Specifically, we consider the *global fidelity quantum kernel*, defined via a *quantum feature map* $\boldsymbol{\rho} : \mathcal{X} \to \mathcal{H}$. Here, $\mathcal{H}$ denotes the Hilbert space of $2^n \times 2^n$ Hermitian matrices with the inner product $\langle \mathbf{A}, \mathbf{B}\rangle_{\mathcal{H}} = \mathrm{Tr}(\mathbf{A}\mathbf{B}^\dagger)$. The feature map $\boldsymbol{\rho}$ is realized through a quantum circuit $\mathbf{S}(\mathbf{x})$, which maps $\mathbf{x} \in \mathcal{X}$ to an n-qubit quantum state $\boldsymbol{\rho}(\mathbf{x})$ by operating on the initial state $|\mathbf{0}\rangle$ as

$$\boldsymbol{\rho}(\mathbf{x}) = \mathbf{S}(\mathbf{x})|\mathbf{0}\rangle\langle\mathbf{0}|\mathbf{S}(\mathbf{x})^\dagger. \tag{3}$$

The resulting quantum state $\boldsymbol{\rho}(\mathbf{x})$ is represented as a $2^n \times 2^n$-dimensional *density matrix* (i.e., a positive semi-definite, Hermitian matrix with unit trace).

For $\mathbf{x} = (x_1, \ldots, x_d) \in \mathbb{R}^d$, the quantum circuit $\mathbf{S}(\mathbf{x})$ can be constructed by interleaving fixed unitary evolutions $\mathbf{W}$ with data dependent encoding gates $e^{-ix_j\mathbf{G}_j}, 1 \leq j \leq d$ as [Schuld and Petruccione, 2021]:

$$\mathbf{S}(\mathbf{x}) = \mathbf{W}^{(d+1)} e^{-ix_d\mathbf{G}_d}\mathbf{W}^{(d)} \cdots e^{-ix_1\mathbf{G}_1}\mathbf{W}^{(1)}. \tag{4}$$

Here, each $\mathbf{G}_j$ is a $d_{\mathbf{G}_j} \leq 2^n$-dimensional Hermitian operator serving as the generator for the encoding gate, while $\mathbf{W}^{(1)}, \ldots, \mathbf{W}^{(d+1)}$ denote arbitrary unitary evolutions before and after the encoding gates. This construction naturally generalizes to multiple repeated layers of $\mathbf{S}(\mathbf{x})$ [Schuld et al., 2021].

The quantum feature map induces the *global fidelity quantum kernel* [Huang et al., 2021],

$$\kappa_Q(\mathbf{x}, \mathbf{x}') = \mathrm{Tr}(\boldsymbol{\rho}(\mathbf{x})\boldsymbol{\rho}(\mathbf{x}')^\dagger). \tag{5}$$

The kernel evaluates the similarity between $\mathbf{x}, \mathbf{x}'$ using global evaluation on the n-qubit quantum state. The associated quantum RKHS $\mathcal{H}_{\kappa_Q}(\mathcal{X})$ is the Hilbert-space completion of $\mathrm{span}\{\kappa_Q(\cdot, \mathbf{x}) : \mathbf{x} \in \mathcal{X}\}$ with inner product satisfying $\langle \kappa_Q(\cdot, \mathbf{x}), \kappa_Q(\cdot, \mathbf{x}')\rangle_{\mathcal{H}_{\kappa_Q}} = \kappa_Q(\mathbf{x}, \mathbf{x}')$. Consequently, every $f \in \mathcal{H}_{\kappa_Q}(\mathcal{X})$ satisfies the reproducing property $f(\mathbf{x}) = \langle f, \kappa_Q(\cdot, \mathbf{x})\rangle_{\mathcal{H}_{\kappa_Q}}$.

For the global fidelity kernel, this RKHS corresponds to reward models that are linear in the quantum feature space, i.e.,

$$f^*(\mathbf{x}) = \mathrm{Tr}(\mathbf{H}\boldsymbol{\rho}(\mathbf{x}))$$

for some Hermitian operator $\mathbf{H} \in \mathcal{H}$ [Schuld and Petruccione, 2021]. Such models arise in many quantum machine learning applications. For instance, the variational

quantum eigensolver (VQE) aims to find the parameterized quantum state $\boldsymbol{\rho}(\mathbf{x}^*)$ that approximates the ground state of a quantum Hamiltonian $\mathbf{H}$, based on noisy observations $y_t = f^*(\mathbf{x}_t) + \eta_t$ of the energy function $f^*(\mathbf{x}) = \mathrm{Tr}(\mathbf{H}\boldsymbol{\rho}(\mathbf{x}))$, where the noise stems from measurement uncertainty [Nicoli et al., 2023].

2.2 QUANTUM GP BANDIT OPTIMIZATION

We approach the quantum kernelized bandit problem using quantum GP models. Leveraging the equivalence between RKHS and GPs, we encode the inductive bias $f^* \in \mathcal{H}_{\kappa_Q}$ by placing a GP prior over f^* with the quantum kernel, i.e., $f^* \sim \mathrm{GP}(0, \kappa_Q(\cdot,\cdot))$. Assuming Gaussian noise $\eta_t \sim \mathcal{N}(0, \sigma^2)$, the posterior over $f^*(\mathbf{x})$ is also a GP, with mean $\mu_T(\mathbf{x})$ and variance $\sigma_T(\mathbf{x},\mathbf{x})$ given by standard formulas (see Appendix A).

In this setting, a natural bandit algorithm is GP-UCB [Srinivas et al., 2010] applied with the quantum GP model. At each round t, GP-UCB selects an action $\mathbf{x}_t$ that maximizes an upper confidence bound (UCB) that contains $f^*(\mathbf{x})$ with high probability over the time horizon:

$$\mathbf{x}_t \;=\; \arg\max_{\mathbf{x}\in\mathcal{X}} \big[\mu_{t-1}(\mathbf{x}) + \sqrt{\beta_t}\,\sigma_{t-1}(\mathbf{x})\big],$$

where β_t is an exploration parameter and $\sigma_{t-1}(\mathbf{x})$ is an estimate of the standard deviation of the process at point $\mathbf{x}$.

Theorem 1 ((Simplified) [Srinivas et al., 2010])**.** *Under mild regularity conditions on the kernel $\kappa_Q(\cdot,\cdot)$, there exists a suitable sequence $\{\beta_t\}$ such that quantum GP-UCB satisfies, with high probability,*

$$R_T \;=\; \mathcal{O}\big(\sqrt{T\,\beta_T\,\gamma_T}\big),$$

where

$$\gamma_T \;=\; \max_{A\subset\mathcal{X},|A|=T} I\big(\mathbf{f}_A;\,\mathbf{y}_A\big)$$

is the maximum information gain, with $\mathbf{f}_A = [f^*(\mathbf{x})]_{\mathbf{x}\in A}$ and $\mathbf{y}_A$ the corresponding noisy observations.

Challenges of Quantum GP Bandit Optimization: Theorem 1 shows that the cumulative regret scales with the maximum information gain γ_T, a kernel-dependent quantity that measures how informative the noisy observations are about the unknown f^*. In particular, GP models with large γ_T are harder to learn and typically incur higher regret.

For GPs equipped with an n-qubit fidelity quantum kernel as in (5), the information gain scales as $\gamma_T = \mathcal{O}(4^n \log T)$ since the model is equivalent to a linear model with 4^n dimensional features (see Appendix B). This scaling is exponential in the number of qubits n, and is *tight* for many practical data-encoding circuits [Schuld and Killoran, 2019, Havlicek et al., 2019].

As an illustrative example, consider the quantum feature map in (3)-(4) with $\mathbf{W} = \mathbf{I}$ and generators $\mathbf{G}_i \in \{\mathbf{X}, \mathbf{Y}, \mathbf{Z}\}$ given by single-qubit Pauli operators. This feature map, called the *rotation encoding*, generates a feature space whose dimension grows exponentially in n. For example, even without entangling gates, a product of single-qubit fixed-axis rotations, e.g., $R_y(x_i)$, has a three-dimensional local span generated by $\{\mathbf{I}, \mathbf{X}, \mathbf{Z}\}$, and therefore gives feature dimension 3^n when applied independently to n qubits. Richer encodings, such as two noncommuting Pauli rotations or data re-uploading, can yield a feature dimension of *exactly* 4^n [Kübler et al., 2021]. We empirically validate this exponential dependence in Figure 2, where we plot the information gain γ_T as a function of the number of data samples T for different values of n (left); and as a function of n for fixed $T = 1000$ (right). The latter clearly illustrates the exponential dependence of γ_T on n. We use Pauli-Y and Pauli-Z encoding, similar to our experiments in Section 4.

Consequently, applying naïve GP-UCB algorithm with high-qubit quantum kernels can lead to prohibitively large regret. Moreover, as n increases, kernel values for different data pairs can concentrate exponentially around a fixed value, requiring an exponential number of quantum measurements to accurately estimate the kernel [Thanasilp et al., 2024].

To overcome these challenges, we propose surrogate GP models that use a lower-dimensional RKHS $\mathcal{H}_{\kappa_{\mathrm{app}}}$, induced by a suitably chosen kernel κ_{app}. In particular, we assume that there exists an approximation

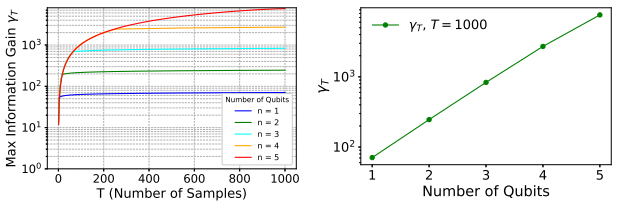
$$\tilde{f} = \arg\min_{f\in\mathcal{H}_{\kappa_{\mathrm{app}}}(\mathcal{X})} \|f - f^*\|_\infty$$

that has small uniform error, i.e., $\|\tilde{f} - f^*\|_\infty \le \varepsilon$ for small $\varepsilon > 0$. The surrogate GP then models $\tilde{f} \sim \mathrm{GP}(0, \kappa_{\mathrm{app}})$.

While such approximations can substantially reduce information gain, they typically introduce *kernel misspecification* relative to the full quantum model. This misspecification can adversely affect cumulative regret [Bogunovic and Krause, 2021]. This motivates designing approximate RKHS for surrogate GP models that balance information gain with misspecification error to achieve more favorable regret.

3 BALANCING INFORMATION GAIN AND MODEL MISSPECIFICATION

This section proposes quantum GP surrogate models based on approximate kernels κ_{app}, designed to balance information gain and model misspecification error. A key challenge in designing effective approximations is to preserve the quantum properties of the full kernel while avoiding exponential model complexity. We consider three low-rank approximation approaches: (a) projected quantum kernels via reduced quantum circuits; (b) classical Random Fourier Features (RFFs) for efficient kernel approximation; and (c)



(a) γ_T against T for diff. n (b) γ_T at $T = 1000$ against n

Figure 2: Plot of γ_T as a function of (a) the iterations T for varying system qubit size n, with Pauli-Y and Z rotation encoding feature maps, and (b) the number n of qubits with $T = 1000$, verifying the exponential increase of γ_T with n.

P-greedy kernel approximations via data-dependent subset selection. While projected quantum kernels still rely on quantum hardware, RFF and P-greedy are fully classical, offering scalability and implementation flexibility.

3.1 LINEAR PROJECTED QUANTUM KERNELS-BASED EC-GP-UCB ALGORITHM

Linear projected quantum kernels (LPQK) [Huang et al., 2021, Kübler et al., 2021] compare quantum states at the level of individual sub-systems, by projecting the global quantum state onto a reduced subspace. Specifically, for a sub-system $s \subseteq \{1, \ldots, n\}$ (indexed by a subset of qubit labels), the LPQK is defined as

$$\kappa_Q^s(\mathbf{x}, \mathbf{x}') = \mathrm{Tr}\big(\boldsymbol{\rho}^s(\mathbf{x})\boldsymbol{\rho}^s(\mathbf{x}')\big), \tag{6}$$

where $\boldsymbol{\rho}^s(\mathbf{x}) = \mathrm{Tr}_{s^c}\big(\boldsymbol{\rho}(\mathbf{x})\big)$ is the reduced density matrix obtained by tracing out the complement sub-system s^c. The resulting state $\boldsymbol{\rho}^s(\mathbf{x})$ lies in a Hilbert space of dimension $2^{|s|} \times 2^{|s|}$, which is exponentially smaller than the full state dimension $2^n \times 2^n$.

While projection onto one sub-system may discard significant information about the global state $\boldsymbol{\rho}(\mathbf{x})$, this loss can be mitigated by aggregating information across multiple sub-systems. To this end, we define a summed LPQK over a collection $\mathbb{S}_b = \{s_1, \ldots, s_w\}$ of all sub-systems satisfying $|s_i| \leq b$, for $i = 1, \ldots, w$, as $\kappa_{\mathbb{S}_b}(\mathbf{x}, \mathbf{x}') = \sum_{s \in \mathbb{S}_b} \frac{1}{|\mathbb{S}_b|} \kappa_Q^s(\mathbf{x}, \mathbf{x}')$ [Gan et al., 2023]. The resulting kernel has effective dimensionality upper bounded by the sum of the dimensions of the individual projected subspaces.

LPQK with Misspecified Bandit Algorithm. We use the summed LPQK $\kappa_{\mathbb{S}_b}$ as the approximate kernel in a surrogate GP model, with $f \sim \mathrm{GP}(0, \kappa_{\mathbb{S}_b}(\cdot, \cdot))$. The next theorem shows that the RKHS $\mathcal{H}_{\kappa_{\mathbb{S}_b}}(\mathcal{X})$, defined by the summed LPQKs for $1 \leq b < n$, is a natural surrogate for approximating the full quantum RKHS $\mathcal{H}_{\kappa_Q}(\mathcal{X})$, and provides an explicit bound on the misspecification error.

Theorem 2. *Let $\mathcal{H}_{\kappa_{\mathbb{S}_b}}(\mathcal{X})$ be the RKHS induced by the summed LPQKs $\kappa_{\mathbb{S}_b}$, and let $\mathcal{H}_{\kappa_Q}(\mathcal{X})$ be the RKHS induced by the full quantum kernel in* (5)*. Then,*

(i) $\mathcal{H}_{\kappa_{\mathbb{S}_b}}(\mathcal{X}) \subseteq \mathcal{H}_{\kappa_Q}(\mathcal{X})$,

(ii) If the true function $f^*(\mathbf{x}) = \mathrm{Tr}[\mathbf{H}\boldsymbol{\rho}(\mathbf{x})]$ is generated by an observable $\mathbf{H}$ drawn uniformly at random from the sphere of operators with Hilbert–Schmidt norm B (equivalent to normalized Gaussian, see Appendix C), then with probability at least $1-\delta$, the misspecification error is bounded as

$$\varepsilon_b := \inf_{f \in \mathcal{H}_{\kappa_{\mathbb{S}_b}}(\mathcal{X})} \|f - f^*\|_\infty \leq B\sqrt{\left(1 - \frac{\sum_{w=0}^{b}\binom{n}{w}3^w}{4^n}\right) + \sqrt{\frac{\ln(1/\delta)}{4^n}}}.$$

Moreover, if $(b+1)/n \geq 3/4$, then

$$\varepsilon_b \leq B\sqrt{e^{-\frac{8n}{3}\left(\frac{b+1}{n} - \frac{3}{4}\right)^2} + \sqrt{\frac{\ln(1/\delta)}{4^n}}}. \tag{7}$$

See Appendix C for the proof. Part (i) shows that the RKHS induced by LPQKs is a *subspace* of the full quantum RKHS. Part (ii) should be interpreted as a conservative baseline: for a dense uniformly random observable, the error scales with the fraction of Pauli-weight components discarded by the projection. Thus aggressive dimension reduction is not guaranteed to have small error in this worst-case model. Instead, the practical benefit of LPQK arises when the target observable or circuit has additional structure, such as locality or a rapidly decaying high-weight Pauli tail; see Appendix C.5 for more details.

For this misspecified setting, where the true reward function $f^* \notin \mathcal{H}_{\kappa_{\mathbb{S}_b}}(\mathcal{X})$, we can leverage the EC-GP-UCB algorithm of Bogunovic and Krause [2021], which yields the following cumulative regret with probability at least $1 - \delta$,

$$R_T = \mathcal{O}\big(B\sqrt{\gamma_T T} + \sqrt{(\ln(1/\delta) + \gamma_T)\gamma_T T} + \varepsilon_b T\sqrt{\gamma_T}\big), \tag{8}$$

where $B < \infty$ is the RKHS-norm of the approximating function f, i.e., $\|f\|_{\mathcal{H}_{\kappa_Q}} \leq B$. Compared to the regret achieved in the realizable setting (Theorem 1), there is an additional $\mathcal{O}(\varepsilon_b T\sqrt{\gamma_T})$ term in the regret that depends on the misspecification error ε_b. For the summed LPQK kernel, this misspecification error depends on the sub-system size b as shown in (7). Additionally, as shown in Appendix C, the information gain of the approximate kernel $\kappa_{\mathbb{S}_b}$ scales as $\gamma_T = \mathcal{O}\big((\sum_{w=0}^{b}\binom{n}{w}3^w)\log T\big)$ with $\sum_{w=0}^{b}\binom{n}{w}3^w \ll 4^n$, significantly smaller than using the n-qubit full fidelity kernel. Consequently, the cumulative regret of EC-GP-UCB algorithm with summed LPQK kernel can be optimized by effectively balancing the information gain γ_T and the misspecification error ε_b by choosing appropriate size b of the sub-system. For structured observables with small or rapidly decaying high-weight Pauli components (cf. Appendix C.5), much smaller values of b can substantially reduce the effective dimension while keeping ε_b small.

3.2 RFF-BASED SQUARECB ALGORITHM

We now investigate RFF-based approximation of the quantum fidelity kernel $\kappa_Q(\mathbf{x}, \mathbf{x}')$ [Landman et al., 2022], leveraging the Fourier representation of quantum kernels [Schuld and Petruccione, 2021]. For exposition, we focus on the simple circuit family in (4), where $\mathbf{G}_i = \mathbf{G}$ for all i with an eigenspectrum $\{\lambda_1, \ldots, \lambda_{d_\mathbf{G}}\}$, and $d_\mathbf{G} \leq 2^n$. Extensions to repeated and multi-layer data encodings admit analogous Fourier expansions and can be derived similarly [Schuld et al., 2021]. Under this encoding, κ_Q admits the Fourier expansion $\kappa_Q(\mathbf{x}, \mathbf{x}') = \sum_{\mathbf{s},\mathbf{t}\in\Omega} c_{\mathbf{s},\mathbf{t}} \exp(i\,\mathbf{s}\cdot\mathbf{x}) \exp(i\,\mathbf{t}\cdot\mathbf{x}')$, where the frequency set Ω contains vectors of the form $\{\mathbf{\Lambda}_j - \mathbf{\Lambda}_k\}$, with $\mathbf{\Lambda}_j = (\lambda_{j_1}, \ldots \lambda_{j_d})$ for indices $j_1, \ldots, j_d \in \{1, \ldots, d_\mathbf{G}\}$, and the Fourier coefficients $c_{\mathbf{s},\mathbf{t}} \in \mathbb{C}$ satisfy $c_{\mathbf{s},\mathbf{t}} = c^*_{-\mathbf{s},-\mathbf{t}}$. The frequency spectrum depends solely on the eigenspectrum of the generator $\mathbf{G}$.

When the quantum kernel $\kappa_Q(\mathbf{x}, \mathbf{x}')$ is *shift-invariant*, i.e., $\kappa_Q(\mathbf{x}, \mathbf{x}') = g(\mathbf{x} - \mathbf{x}')$ [Schuld and Petruccione, 2021] for some function $g : \mathbb{R}^d \to \mathbb{R}$, Bochner's theorem [Rudin, 2017] ensures that κ_Q admits a Fourier integral representation: $\kappa_Q(\mathbf{x}, \mathbf{x}') = \int_\Omega p(\boldsymbol{\omega}) e^{i\boldsymbol{\omega}^\top(\mathbf{x}-\mathbf{x}')} d\boldsymbol{\omega}$, where $p(\boldsymbol{\omega})$ is a probability distribution over $\boldsymbol{\omega} \in \Omega$ (assuming κ_Q is normalized). For the circuit family above, shift-invariance arises when each generator $\mathbf{G}_i = \mathbf{G}$ encodes input x_i into distinct, non-overlapping qubits.

RFF approximates this kernel as [Rahimi and Recht, 2007]

$$\kappa_Q(\mathbf{x}, \mathbf{x}') \;\approx\; \boldsymbol{\phi}_{\mathrm{RFF}}(\mathbf{x})^\top \boldsymbol{\phi}_{\mathrm{RFF}}(\mathbf{x}') =: \kappa_{\mathrm{RFF}}(\mathbf{x}, \mathbf{x}'), \quad (9)$$

where $\boldsymbol{\phi}_{\mathrm{RFF}} : \mathcal{X} \to \mathbb{R}^D$ is the *feature map* obtained by drawing D frequency samples $\{\boldsymbol{\omega}_j\}$ from $p(\boldsymbol{\omega})$ and random phases $\{b_j\}$ from the uniform distribution over $[0, 2\pi]$:

$$\boldsymbol{\phi}_{\mathrm{RFF}}(\mathbf{x}) = \sqrt{\tfrac{2}{D}} \left[\cos\big(\boldsymbol{\omega}_1\cdot\mathbf{x}+b_1\big) \;\ldots\; \cos\big(\boldsymbol{\omega}_D\cdot\mathbf{x}+b_D\big)\right]^\top.$$

For many standard quantum circuits, the frequency set Ω and coefficients $\{c_{\mathbf{s},\mathbf{t}}\}$ can be derived explicitly, allowing one to directly sample frequencies from Ω for RFF [Schuld and Petruccione, 2021]. See Appendix D for more details.

Misspecified Linear Bandit via RFF. The RFF kernel (9) induces the following RKHS,

$$\mathcal{H}_{\kappa_{\mathrm{RFF}}}(\mathcal{X}) = \{f(\cdot) = \mathbf{w}^\top \boldsymbol{\phi}_{\mathrm{RFF}}(\cdot) : \mathbf{w} \in \mathbb{R}^D\}, \quad (10)$$

consisting of functions linear in $\boldsymbol{\phi}_{\mathrm{RFF}}(\cdot)$. While κ_{RFF} can define a GP model, its finite dimensionality also allows us to treat the problem as a *misspecified linear bandit* with the RFF feature. Concretely, we model the unknown mean reward function as $\tilde{f}(\cdot) = \mathbf{w}^\top \boldsymbol{\phi}_{\mathrm{RFF}}(\cdot) \in \mathcal{H}_{\kappa_{\mathrm{RFF}}}(\mathcal{X})$ such that the true reward function $f^*(\mathbf{x}) = \tilde{f}(\mathbf{x}) + \xi(\mathbf{x})$ with $\sup_\mathbf{x} |\xi(\mathbf{x})| \leq \varepsilon_D$. In this setting, the SquareCB algorithm Foster and Rakhlin [2020] achieves an expected regret of $\mathbb{E}[R_T] = \tilde{\mathcal{O}}(\sqrt{DKT} + \varepsilon_D\sqrt{KT})$, where the expectation is over the randomness of action policy and $K = |\mathcal{X}|$ is the number of actions.

Theorem 3. *Let $f^* \in \mathcal{H}_{\kappa_Q}(\mathcal{X})$ with $\|f^*\|_{\mathcal{H}_{\kappa_Q}} \leq B < \infty$, where $\kappa_Q(\cdot,\cdot)$ is an n-qubit shift-invariant quantum kernel. Consider $\mathcal{H}_{\kappa_{\mathrm{RFF}}}(\mathcal{X})$ as in* (10) *obtained by sampling D Fourier features of κ_Q. Then, with probability at least $1-\delta$ over the sampled Fourier features, there exists $\tilde{f} \in \mathcal{H}_{\kappa_{\mathrm{RFF}}}(\mathcal{X})$ such that the misspecification error satisfies*

$$\sup_{\mathbf{x}\in\mathcal{X}} |f^*(\mathbf{x}) - \tilde{f}(\mathbf{x})| < \varepsilon_D \quad \textit{where} \quad (11)$$

$$\varepsilon_D := \frac{4B|\Omega|}{\sqrt{D}}\Big(\sqrt{\log\frac{1}{\delta}} + 4B_\mathcal{X}\,\omega_{\max}\Big),$$

and where $B_\mathcal{X} := \sup_{\mathbf{x}\in\mathcal{X}} \|\mathbf{x}\|_2$ and $\omega_{\max} := \max_{\boldsymbol{\omega}\in\Omega} \|\boldsymbol{\omega}\|_2$ are the domain and frequency radii respectively. Using SquareCB with $\mathcal{H}_{\kappa_{RFF}}(\mathcal{X})$ yields an expected cumulative regret of

$$\mathbb{E}[R_T] \leq \min\Big\{2\sqrt{4^n KT\log\big(1+\tfrac{T}{4^n}\big)}, 2\sqrt{K}T^{3/4}\sqrt{\log T} + 20\sqrt{K}T^{3/4}B|\Omega|\Big(\sqrt{\log T} + 4B_\mathcal{X}\,\omega_{\max}\Big)\Big\}, \quad (12)$$

where the expectation is over SquareCB's random action sampling, the stochastic rewards, as well as the sampled Fourier features.

See Appendix E for proof. In the regret bound in (12), the first term corresponds to applying SquareCB with the full quantum model (no misspecification), scaling as $\tilde{\mathcal{O}}(\sqrt{4^n KT})$. The second term arises from using SquareCB in the misspecified setting, with the misspecification error in (11) optimized with the choice of $D = \sqrt{T}$, giving a regret of $\tilde{\mathcal{O}}(\sqrt{K}T^{3/4})$. This term is smaller than the $\tilde{\mathcal{O}}(\sqrt{4^n KT})$ regret when $4^n \gtrsim \sqrt{T}$, i.e., for $n = \Omega(\log T)$. This analysis demonstrates that RFF-based surrogate models outperform full quantum fidelity kernel methods in regimes where the kernel complexity grows faster than the required approximation capacity. Furthermore, the analysis provides a theoretical prescription for selecting the optimal dimension D to balance approximation error and model complexity.

We remark that, if the horizon T is not known in advance, it can be chosen in an anytime manner via a standard doubling trick (cf. Besson and Kaufmann [2018]), which preserves the regret order up to logarithmic factors in T.

Alternative Sampling and Possible Faster Error Decay. In practice, instead of sampling $\boldsymbol{\omega} \sim p(\boldsymbol{\omega})$, one can use structured sampling (e.g., prioritizing frequencies with the largest Fourier coefficients). When the Fourier expansion of the kernel is explicitly known, this can tighten misspecification bounds [Schuld, 2021], particularly if the coefficients decay rapidly—such as polynomially: $\varepsilon_D = \mathcal{O}(1/D)$ or exponentially: $\varepsilon_D = \exp(-\Omega(D))$. In these regimes, both SquareCB and EC-GP-UCB benefit from faster convergence and improved regret guarantees.

Computational Complexity. GP or kernelized regression typically require matrix inversion operations with $\mathcal{O}(T^3)$

cost. In contrast, approximating κ_Q with a D-dimensional feature map yields a linear model, solvable by ridge regression in $\mathcal{O}(TD^2 + D^3)$. As long as $D < T$, the RFF approach provides substantial computational savings. Additionally, for quantum kernels involving a large number of qubits n, kernel evaluations can become expensive; in such cases, RFF-based methods are more practical and scalable (more details are provided in Appendix F).

3.3 P-GREEDY KERNEL APPROXIMATION

While RFF is effective for *shift-invariant* kernels with accessible feature maps, many quantum kernels lack these properties or are challenging to analyze. As an alternative, P-greedy [Marchi et al., 2005, Takemori and Sato, 2021] requires only access to kernel evaluations $\kappa_Q(\mathbf{x}, \mathbf{x}')$.

The P-greedy algorithm constructs a low-dimensional subspace $V(X_D) = \operatorname{span}\{\kappa_Q(\cdot, \boldsymbol{\xi}_1), \ldots, \kappa_Q(\cdot, \boldsymbol{\xi}_D)\} \subset \mathcal{H}_{\kappa_Q}(\mathcal{X})$, spanned by kernel evaluated at a chosen subset of points $X_D = \{\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_D\} \subset \mathcal{X}$. This subspace is selected iteratively: At each iteration k, the algorithm greedily selects the point $\boldsymbol{\xi}_k \in \mathcal{X}$ that maximizes the power function:

$$P_{V(X_{k-1})}(\mathbf{x}) := \sup_{f \in \mathcal{H}_{\kappa_Q}(\mathcal{X}) \setminus \{0\}} \frac{|f(\mathbf{x}) - (\Pi_{V(X_{k-1})} f)(\mathbf{x})|}{\|f\|_{\mathcal{H}_{\kappa_Q}}},$$

which identifies where the current subspace approximates the quantum RKHS worst. Here, $\Pi_{V(X)} : \mathcal{H}_{\kappa_Q}(\mathcal{X}) \to V(X)$ denotes the orthogonal projection onto $V(X)$. After D iterations, the algorithm returns X_D, defining a subspace $V(X_D)$ that approximates the full quantum RKHS. The final feature map corresponds to the Newton basis formed via Gram–Schmidt orthonormalization of basis $\{\kappa_Q(\cdot, \boldsymbol{\xi}_1), \ldots, \kappa_Q(\cdot, \boldsymbol{\xi}_D)\}$, yielding an efficient low-rank approximation of the quantum kernel.

Performance and Spectral Decay. The effectiveness of the P-greedy algorithm is governed by the spectral decay of the target quantum kernel [Jordão and Menegatto, 2018]. For kernels with rapid eigenvalue decay, the algorithm may admit polynomial or exponential improvement with each greedily chosen point [DeVore et al., 2013]. Nevertheless, P-greedy typically achieves near-optimal performance: The misspecification error using $2D$ points satisfies [DeVore et al., 2013] $\varepsilon_{2D} \leq B/\gamma \cdot \sqrt{2d_D}$ for some $0 < \gamma < 1$, where d_D is the D-dimensional Kolmogorov width, representing the best possible D-dimensional approximation of the quantum RKHS $\mathcal{H}_{\kappa_Q}(\mathcal{X})$.

As an example, consider quantum kernels κ_Q with a Fourier expansion,

$$\kappa_Q(\mathbf{x}, \mathbf{x}') = \sum_{\boldsymbol{\omega} \in \Omega} c_{\boldsymbol{\omega}}\, e^{i\langle \boldsymbol{\omega}, \mathbf{x} - \mathbf{x}' \rangle}.$$

The Kolmogorov width d_D is controlled by the spectral tail with

$$d_D^2 = \sum_{j > D} c_j = \sum_{\boldsymbol{\omega} \notin \Omega_D} c_{\boldsymbol{\omega}},$$

where $\{c_j\}_{j \geq 1}$ are the eigenvalues in non-increasing order and Ω_D indexes the top-D eigenvalues (see Appendix H for a proof). Without additional circuit structure, fast spectral decay is generally not guaranteed. However, there exist many circuit families that produce non-uniform spectra, often leading to exponentially decaying d_D, and consequently, ε_D (see Appendix G for examples).

Once an instance-specific (or circuit-specific) bound ε_D is available for the P-greedy surrogate class, it can be directly integrated into the regret bounds of misspecified bandit algorithms to obtain a theoretically principled choice of the approximation dimension D.

4 NUMERICAL EXPERIMENTS

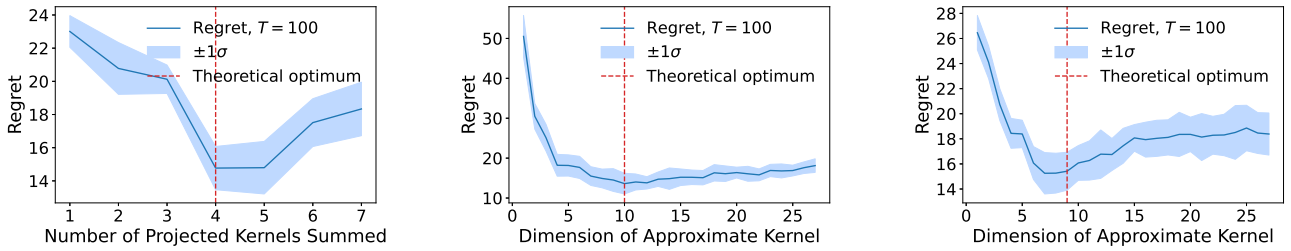
We demonstrate the effectiveness of our kernel approximation approach on a variety of tasks using classical bandit algorithms such as EC-GP-UCB [Bogunovic and Krause, 2021] or SquareCB [Foster and Rakhlin, 2020]. Further implementation details are in Appendix J.

4.1 GP OPTIMIZATION WITH SYNTHETIC QUANTUM FUNCTIONS

Our first synthetic task uses a true reward function f^* sampled from a three-qubit quantum kernel (extension to larger $n = 6$ qubit systems can be found in Appendix K). We construct a parameterized quantum circuit in PennyLane with Pauli-X data-encoding gates followed by random rotation and entangling gates. This circuit defines a global fidelity quantum kernel κ_Q, which we use to generate random functions f^* over the domain $[0, 2\pi]^3$. Observed rewards are corrupted with i.i.d. Gaussian noise (variance $\sigma^2 = 0.01$) and normalized to $[0, 1]$.

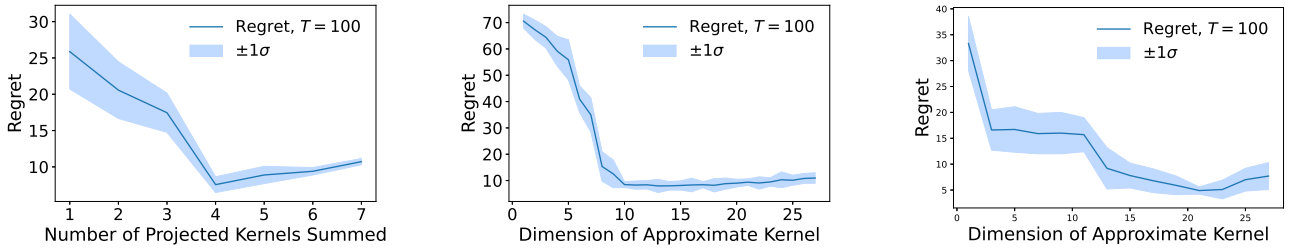
We evaluate three kernel-approximation strategies: projected quantum kernels, RFF, and P-greedy, each paired with *SquareCB* and *EC-GP-UCB*. Model complexity is controlled by the number of summed LPQKs, the RFF dimension D, or the number of basis elements in P-greedy. Figures 3 and 4 plot the cumulative regret as a function of model complexity for the three approximation methods, using SquareCB and EC-GP-UCB respectively. The rightmost point in each plot corresponds to the full quantum kernel (i.e., maximum model complexity). Each algorithm was run for $T = 100$ rounds, with results averaged over 30 trials.

Across all settings, we observe the characteristic "U-shaped" trend in regret. At low model complexity, the surrogate kernel is too coarse to capture f^* accurately, leading to underfitting and high regret. As model complexity increases, the approximation error decreases, improving performance.



(a) Projected Quantum Kernels (b) Random Fourier Features (c) P-greedy

Figure 3: Cumulative regret of SquareCB against approximation model complexity for (a) projected quantum kernels, (b) RFF and (c) P-greedy approximation.



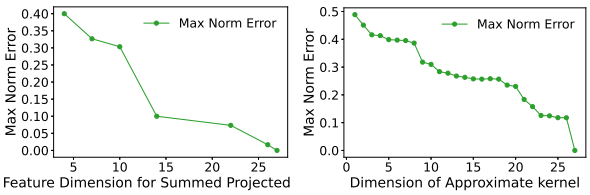
(a) Projected Quantum Kernels (b) Random Fourier Features (c) P-greedy

Figure 4: Cumulative regret of EC-GP-UCB against approximation model complexity for (a) projected quantum kernels, (b) RFF and (c) P-greedy approximation.

However, beyond an optimal point, further increase in complexity leads to higher regret due to increased information gain, mirroring overfitting in classical learning.

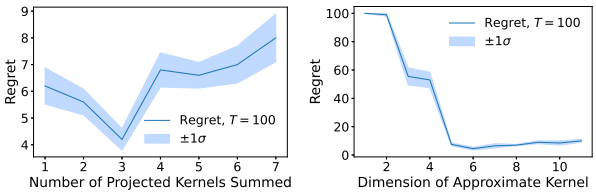
Choosing the approximate kernel dimension. Theorem 3 suggests setting the RFF dimension $D = \sqrt{T}$ to minimize the regret bound. For $T = 100$, this yields $D = 10$, which is highlighted by the red vertical line in Figure 3(b). Note that our theoretical choice of D aligns perfectly with the best empirical performance.

For P-greedy, the relationship between the approximation dimension D and the misspecification error ε is circuit-dependent. One can analyze the circuit structure—empirically or analytically—to estimate how ε varies with D. Plugging this estimate into misspecified bandit regret bounds (e.g., SquareCB) then allows us to select the optimal D. Figure 5 illustrates the empirical relationship between ε and D for both the projected kernel and P-greedy methods. Substituting these estimates into SquareCB's regret bounds yields a theoretically optimal D, marked in red in Figures 3(a) and (c). These values closely match the observed performance peaks, validating the approach. The same principle applies to EC-GP-UCB.



(a) Projected Quantum Kernels (b) P-greedy

Figure 5: Uniform (max-norm) approximation error against the kernel dimension (or effective dimension) for (a) projected quantum kernels and (b) P-greedy algorithm.



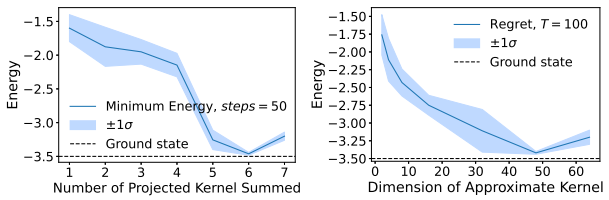
(a) Projected Quantum Kernels (b) P-greedy

Figure 6: Cumulative regrets of EC-GP-UCB after $T = 100$ iterations as a function of the model complexity for (a) projected quantum kernels and (b) P-greedy approximation for quantum phase classification.

4.2 PHASE CLASSIFICATION

We next evaluate our approach on a *quantum phase classification* task inspired by Caro et al. [2022]. We consider the generalized cluster Hamiltonian

$$\mathbf{H}_C(J_1, J_2) = \sum_{j=1}^{n} (\mathbf{Z}_j - J_1 \mathbf{X}_j \mathbf{X}_{j+1} - J_2 \mathbf{X}_{j-1} \mathbf{Z}_j \mathbf{X}_{j+1}),$$



(a) Projected Quantum Kernels (b) P-greedy

Figure 7: Bayesian optimization of VQE with different kernel approximation approaches, $n = 3$

where $\mathbf{X}_j$ and $\mathbf{Z}_j$ are Pauli-X and Pauli-Z operators, respectively, acting on qubit j. By varying the coupling coefficients J_1 and J_2, the ground state of the Hamiltonian can fall into one of four distinct phases: symmetry-protected topological (I), ferromagnetic (II), anti-ferromagnetic (III), or trivial (IV). Our task is to identify parameters (J_1, J_2) corresponding to the *ferromagnetic phase* (II).

In each round, we sample a point $\mathbf{x} = (J_1, J_2)$, compute the ground state $|\psi_{\mathbf{x}}\rangle$ of $\mathbf{H}_C$, and observe $y = f^*(\mathbf{x}) + \eta_t$ where $\eta_t \sim \mathcal{N}(0, 0.01)$. Crucially, for this physical system, the reward function $f^*(\mathbf{x})$ can be represented within a quantum RKHS via the normalized total magnetization squared,

$$\mathbf{H} = \frac{1}{n^2} \sum_{j,k=1}^{n} \mathbf{X}_j \mathbf{X}_k,$$

as $f^*(\mathbf{x}) = \mathrm{Tr}(\mathbf{H}|\psi_{\mathbf{x}}\rangle\langle\psi_{\mathbf{x}}|)$, which is approximately 1 within the ferromagnetic phase and 0 elsewhere. The objective is to minimize regret for choosing points not in phase II.

We apply EC-GP-UCB using approximate kernels (projected kernels and P-greedy) derived from a three-qubit *quantum kernel* implemented via a parameterized circuit with random single-qubit rotations and entangling gates acting on the initial ground state. Figure 6 shows the regret of EC-GP-UCB over $T = 100$ iterations against model complexity. Both approximation methods significantly outperform the full quantum kernel (rightmost points) at their optimal complexity. This suggests that lower-dimensional kernels suffice to distinguish phase II, while avoiding the higher information-gain penalty of the full kernel.

4.3 VARIATIONAL QUANTUM EIGENSOLVER

In our final experiment, we follow the setup in Nicoli et al. [2023] to approximate the ground state of the XYZ Heisenberg Hamiltonian:

$$\mathbf{H} = -\sum_{j=1}^{n-1} (J_X \mathbf{X}_j \mathbf{X}_{j+1} + J_Y \mathbf{Y}_j \mathbf{Y}_{j+1} + J_Z \mathbf{Z}_j \mathbf{Z}_{j+1}) - \sum_{j=1}^{n} (h_X \mathbf{X}_j + h_Y \mathbf{Y}_j + h_Z \mathbf{Z}_j)$$

We employ a 3-qubit "Efficient SU(2)" circuit [Nicoli et al., 2023] as the parameterized ansatz $\mathbf{U}(\mathbf{x})$, acting on an initial state $|\psi_0\rangle$. Our goal is to find the parameters $\mathbf{x}$ that minimize the energy function

$$f^*(\mathbf{x}) = \langle\psi_0| \mathbf{U}(\mathbf{x})^\dagger \mathbf{H} \mathbf{U}(\mathbf{x}) |\psi_0\rangle .$$

We use a GP model with kernel from the SU(2) circuit, performing Bayesian optimization with the Expected Improvement (EI) acquisition function, optimized via L-BFGS. Each run consists of 50 optimization steps, averaged over 30 trials.

Figure 7 illustrates the energy landscape as a function of model complexity for the projected quantum kernel and P-greedy approximations. The full kernel appears excessively complex for this task, leading to slower convergence. In contrast, simpler approximations achieve faster convergence. While this is an optimization task (minimizing energy) rather than a regret-minimization bandit problem, the results still show that choosing an appropriate kernel complexity is crucial for efficiency. This demonstrates that the balance between expressivity and sample efficiency extends beyond standard bandit optimization contexts.

5 CONCLUSION

We developed approximate GP algorithms for quantum kernel bandits, addressing the challenges of scalability and sample efficiency in NISQ-era learning tasks. By exploiting reduced quantum subspaces and classical approximations, our methods achieve a practical trade-off between information gain and model misspecification. Theoretical regret bounds, supported by empirical results on both synthetic and quantum tasks, demonstrate the effectiveness of our approach. Finally, the techniques discussed here are applicable to high dimensional classical kernels, offering lower regret and reduced computational complexity (see Appendix I). For limitations and future work, see Appendix L.

Acknowledgements

V. Y. F. Tan is supported by a Singapore Ministry of Education (MOE) AcRF Tier 1 grant (A-8002934-00-00) and an AcRF Tier 2 grant (A-8004062-00-00). S. T. Jose is supported through Royal Society International Travel Grant (Reference no: IES\R2\242104) and through EPSRC Quantum Technologies Career Acceleration Fellowship (UKRI1218).

References

Alice Barthe and Adrián Pérez-Salinas. Gradients and frequency profiles of quantum re-uploading models. *Quantum*, 8:1523, November 2024. ISSN 2521-327X. doi: 10.22331/q-2024-11-14-1523. URL https://doi.org/10.22331/q-2024-11-14-1523.

Lilian Besson and Emilie Kaufmann. What doubling tricks can and can't do for multi-armed bandits, 2018. URL https://arxiv.org/abs/1803.06971.

Carsten Blank, Daniel K Park, June-Koo Kevin Rhee, and Francesco Petruccione. Quantum classifier with tailored quantum kernel. *npj Quantum Information*, 6(1):41, 2020.

Ilija Bogunovic and Andreas Krause. Misspecified gaussian process bandit optimization. *Advances in neural information processing systems*, 34:3004–3015, 2021.

Simon Buchholz, Jonas M. Kübler, and Bernhard Schölkopf. Multi-Armed Bandits and Quantum Channel Oracles. *Quantum*, 9:1672, March 2025. ISSN 2521-327X. doi: 10.22331/q-2025-03-25-1672. URL https://doi.org/10.22331/q-2025-03-25-1672.

Marin Bukov, Alexandre GR Day, Dries Sels, Phillip Weinberg, Anatoli Polkovnikov, and Pankaj Mehta. Reinforcement learning in different phases of quantum control. *Physical Review X*, 8(3):031086, 2018.

Romain Camilleri, Kevin Jamieson, and Julian Katz-Samuels. High-dimensional experimental design and kernel bandits. In *International Conference on Machine Learning*, pages 1227–1237. PMLR, 2021.

Matthias C. Caro, Hsin-Yuan Huang, M. Cerezo, Kunal Sharma, Andrew Sornborger, Lukasz Cincio, and Patrick J. Coles. Generalization in quantum machine learning from few training data. *Nature Communications*, 13(1), August 2022. ISSN 2041-1723. doi: 10.1038/s41467-022-32550-3. URL http://dx.doi.org/10.1038/s41467-022-32550-3.

Balthazar Casalé, Giuseppe Di Molfetta, Hachem Kadri, and Liva Ralaivola. Quantum bandits. *Quantum Machine Intelligence*, 2:1–7, 2020.

Marco Cerezo, Andrew Arrasmith, Ryan Babbush, Simon C Benjamin, Suguru Endo, Keisuke Fujii, Jarrod R McClean, Kosuke Mitarai, Xiao Yuan, Lukasz Cincio, et al. Variational quantum algorithms. *Nature Reviews Physics*, 3(9):625–644, 2021.

John B. Conway. *A Course in Functional Analysis*. Springer, 2007.

Zhongxiang Dai, Gregory Kang Ruey Lau, Arun Verma, Yao Shu, Bryan Kian Hsiang Low, and Patrick Jaillet. Quantum bayesian optimization. *Advances in Neural Information Processing Systems*, 36:20179–20207, 2023.

Ronald DeVore, Guergana Petrova, and Przemyslaw Wojtaszczyk. Greedy algorithms for reduced bases in banach spaces. *Constructive Approximation*, pages 455–466, 2013.

Dylan J. Foster and Alexander Rakhlin. Beyond ucb: optimal and efficient contextual bandits with regression oracles. In *Proceedings of the 37th International Conference on Machine Learning*. JMLR.org, 2020.

Beng Yee Gan, Daniel Leykam, and Supanut Thanasilp. A unified framework for trace-induced quantum kernels. *arXiv preprint arXiv:2311.13552*, 2023.

Avishek Ghosh, Sayak Ray Chowdhury, and Aditya Gopalan. Misspecified linear bandits. In *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*, pages 3761–3767, 2017.

Xuyang Guo, Jun Dai, and Roman V Krems. Benchmarking of quantum fidelity kernels for gaussian process regression. *Machine Learning: Science and Technology*, 5(3): 035081, 2024.

Vojtech Havlicek, Antonio D. Corcoles, Kristan Temme, Aram W. Harrow, Abhinav Kandala, Jerry M. Chow, and Jay M. Gambetta. Supervised learning with quantum-enhanced feature spaces. *Nature*, 567(7747): 209–212, March 2019. ISSN 1476-4687. doi: 10.1038/s41586-019-0980-2. URL http://dx.doi.org/10.1038/s41586-019-0980-2.

Yasunari Hikima, Kazunori Murao, Sho Takemori, and Yuhei Umeda. Quantum kernelized bandits. In *The 40th Conference on Uncertainty in Artificial Intelligence*, 2024.

Hsin-Yuan Huang, Michael Broughton, Masoud Mohseni, Ryan Babbush, Sergio Boixo, Hartmut Neven, and Jarrod R McClean. Power of data in quantum machine learning. *Nature communications*, 12(1):2631, 2021.

T. Jordão and V.A. Menegatto. Kolmogorov widths on the sphere via eigenvalue estimates for hölderian integral operators, 2018. URL https://arxiv.org/abs/1805.08476.

Motonobu Kanagawa, Philipp Hennig, Dino Sejdinovic, and Bharath K Sriperumbudur. Gaussian processes and kernel methods: A review on connections and equivalences. *arXiv preprint arXiv:1807.02582*, 2018.

Jonas Kübler, Simon Buchholz, and Bernhard Schölkopf. The inductive bias of quantum kernels. *Advances in Neural Information Processing Systems*, 34:12661–12673, 2021.

Jonas Landman, Slimane Thabet, Constantin Dalyac, Hela Mhiri, and Elham Kashefi. Classically approximating variational quantum machine learning with random

fourier features, 2022. URL https://arxiv.org/abs/2210.13200.

Josep Lumbreras, Erkka Haapasalo, and Marco Tomamichel. Multi-armed quantum bandits: Exploration versus exploitation when learning properties of quantum states. *Quantum*, 6:749, 2022.

Olivier Marchal and Julyan Arbel. On the sub-gaussianity of the beta and dirichlet distributions. *Electronic Communications in Probability*, 22(none), January 2017. ISSN 1083-589X. doi: 10.1214/17-ecp92. URL http://dx.doi.org/10.1214/17-ECP92.

Stefano De Marchi, Robert Schaback, and Holger Wendland. Near-optimal data-independent point locations for radial basis function interpolation. *Advances in Computational Mathematics*, 23:317–330, 2005.

Jarrod R McClean, Sergio Boixo, Vadim N Smelyanskiy, Ryan Babbush, and Hartmut Neven. Barren plateaus in quantum neural network training landscapes. *Nature communications*, 9(1):4812, 2018.

Gergely Neu and Julia Olkhovskaya. Efficient and robust algorithms for adversarial linear contextual bandits. In *Conference on Learning Theory*, pages 3049–3068. PMLR, 2020.

Kim Nicoli, Christopher J Anders, Lena Funcke, Tobias Hartung, Karl Jansen, Stefan Kühn, Klaus-Robert Müller, Paolo Stornati, Pan Kessel, and Shinichi Nakajima. Physics-informed bayesian optimization of variational quantum circuits. *Advances in Neural Information Processing Systems*, 36:18341–18376, 2023.

Ali Rahimi and Benjamin Recht. Random features for large-scale kernel machines. In *Proceedings of the 21st International Conference on Neural Information Processing Systems*, pages 1177–1184. Curran Associates Inc., 2007.

Ali Rahimi and Benjamin Recht. Uniform approximation of functions with random bases. In *2008 46th Annual Allerton Conference on Communication, Control, and Computing*, pages 555–561, 2008. doi: 10.1109/ALLERTON.2008.4797607.

Frederic Rapp and Marco Roth. Quantum gaussian process regression for bayesian optimization. *Quantum Machine Intelligence*, 6(1):5, 2024.

Patrick Rebentrost, Masoud Mohseni, and Seth Lloyd. Quantum support vector machine for big data classification. *Physical review letters*, 113(13):130503, 2014.

Walter Rudin. *Fourier analysis on groups*. Courier Dover Publications, 2017.

Jonas Schuff, Lukas J Fiderer, and Daniel Braun. Improving the dynamics of quantum sensors with reinforcement learning. *New Journal of Physics*, 22(3):035001, 2020.

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

Maria Schuld and Nathan Killoran. Quantum machine learning in feature hilbert spaces. *Physical Review Letters*, 122(4), February 2019. ISSN 1079-7114. doi: 10.1103/physrevlett.122.040504. URL http://dx.doi.org/10.1103/PhysRevLett.122.040504.

Maria Schuld and Francesco Petruccione. Quantum models as kernel methods. *Machine Learning with Quantum Computers*, pages 217–245, 2021.

Maria Schuld, Ryan Sweke, and Johannes Jakob Meyer. Effect of data encoding on the expressive power of variational quantum-machine-learning models. *Physical Review A*, 103(3), March 2021. ISSN 2469-9934. doi: 10.1103/physreva.103.032430. URL http://dx.doi.org/10.1103/PhysRevA.103.032430.

Alistair WR Smith, AJ Paige, and MS Kim. Faster variational quantum algorithms with quantum kernel-based surrogate models. *Quantum Science and Technology*, 8 (4):045016, 2023.

Niranjan Srinivas, Andreas Krause, Sham M Kakade, and Matthias Seeger. Gaussian process optimization in the bandit setting: No regret and experimental design. *arXiv preprint arXiv:0912.3995*, 2009.

Niranjan Srinivas, Andreas Krause, Sham M Kakade, and Matthias Seeger. Gaussian process optimization in the bandit setting: No regret and experimental design. *In International Conference on Machine Learning (ICML), pages 1015–1022*, 2010.

Sho Takemori and Masahiro Sato. Approximation theory based methods for rkhs bandits. In *International Conference on Machine Learning*, pages 10076–10085. PMLR, 2021.

Supanut Thanasilp, Samson Wang, Marco Cerezo, and Zoë Holmes. Exponential concentration in quantum kernel methods. *Nature communications*, 15(1):5200, 2024.

Michal Valko, Nathaniel Korda, Rémi Munos, Ilias Flaounas, and Nelo Cristianini. Finite-time analysis of kernelised contextual bandits. *arXiv preprint arXiv:1309.6869*, 2013.

Zongqi Wan, Zhijie Zhang, Tongyang Li, Jialin Zhang, and Xiaoming Sun. Quantum multi-armed bandits and stochastic linear bandits enjoy logarithmic regrets. In *Proceedings of the Thirty-Seventh AAAI Conference on Artificial Intelligence and Thirty-Fifth Conference on Innovative Applications of Artificial Intelligence and Thirteenth Symposium on Educational Advances in Artificial Intelligence*, AAAI'23/IAAI'23/EAAI'23. AAAI Press,

2023. ISBN 978-1-57735-880-0. doi: 10.1609/aaai.v37i8.26202. URL https://doi.org/10.1609/aaai.v37i8.26202.

Daochen Wang, Xuchen You, Tongyang Li, and Andrew M. Childs. Quantum exploration algorithms for multi-armed bandits. *Proceedings of the AAAI Conference on Artificial Intelligence*, 35(11):10102–10110, May 2021a. ISSN 2159-5399. doi: 10.1609/aaai.v35i11.17212. URL http://dx.doi.org/10.1609/aaai.v35i11.17212.

Samson Wang, Enrico Fontana, Marco Cerezo, Kunal Sharma, Akira Sone, Lukasz Cincio, and Patrick J Coles. Noise-induced barren plateaus in variational quantum algorithms. *Nature communications*, 12(1):6961, 2021b.

Xuchuang Wang, Yu-Zhen Janice Chen, Matheus Guedes de Andrade, Jonathan Allcock, Mohammad Hajiesmaili, John C.S. Lui, and Don Towsley. Best arm identification with quantum oracles. In *Proceedings of the Thirty-Ninth AAAI Conference on Artificial Intelligence and Thirty-Seventh Conference on Innovative Applications of Artificial Intelligence and Fifteenth Symposium on Educational Advances in Artificial Intelligence*, AAAI'25/IAAI'25/EAAI'25. AAAI Press, 2025. ISBN 978-1-57735-897-8. doi: 10.1609/aaai.v39i20.35432. URL https://doi.org/10.1609/aaai.v39i20.35432.

Marc Wanner, Johan Jonasson, Emil Carlsson, and Devdatt Dubhashi. Variational quantum optimization with continuous bandits. *arXiv preprint arXiv:2502.04021*, 2025.

Yulian Wu, Chaowen Guan, Vaneet Aggarwal, and Di Wang. Quantum heavy-tailed bandits. *arXiv preprint arXiv:2301.09680*, 2023.

Andrea Zanette, Alessandro Lazaric, Mykel Kochenderfer, and Emma Brunskill. Learning near optimal policies with low inherent bellman error. In *International Conference on Machine Learning*, pages 10978–10989. PMLR, 2020.

Balancing Expressivity and Learnability in Quantum Kernel Bandit Optimization (Supplementary Material)

Yuqi Huang[1] **Vincent Y. F. Tan**[1,2] **Sharu Theresa Jose**[3]

[1]Department of Mathematics, National University of Singapore, Singapore
[2]Department of Electrical and Computer Engineering, National University of Singapore, Singapore
[3]School of Computer Science, University of Birmingham, Birmingham, United Kingdom

A DETAILS ON GAUSSIAN PROCESSES

Suppose we have a set of noisy samples

$$\mathbf{y}_T \;=\; \big[y_1,\dots,y_T\big]^\top$$

at the points

$$\mathbf{A}_T \;=\; \big[\mathbf{x}_1,\dots,\mathbf{x}_T\big],$$

where $y_t = f(\mathbf{x}_t) + \eta_t$ with $\eta_t \sim \mathcal{N}\big(0,\sigma^2\big)$ i.i.d. Gaussian noise. Then the posterior on f is again a GP distribution with mean $\mu_T(\mathbf{x})$ and variance $\sigma_T^2(\mathbf{x})$ given by the standard formulas:

$$\mu_T(\mathbf{x}) \;=\; \mathbf{k}_T(\mathbf{x})^\top\big(\mathbf{K}_T+\sigma^2\mathbf{I}\big)^{-1}\mathbf{y}_T, \tag{13}$$

$$\sigma_T^2(\mathbf{x}) \;=\; \kappa(\mathbf{x},\mathbf{x}) \;-\; \mathbf{k}_T(\mathbf{x})^\top\big(\mathbf{K}_T+\sigma^2\mathbf{I}\big)^{-1}\mathbf{k}_T(\mathbf{x}), \tag{14}$$

where $\mathbf{k}_T(\mathbf{x})$ is the length-T-vector $[\,\kappa(\mathbf{x}_1,\mathbf{x}),\dots,\kappa(\mathbf{x}_T,\mathbf{x})\,]^\top$, and $\mathbf{K}_T$ is the positive semidefinite kernel matrix $\big[\kappa(\mathbf{x},\mathbf{x}')\big]_{\mathbf{x},\mathbf{x}'\in\mathbf{A}_T}$.

B INFORMATION GAIN OF QUANTUM FIDELITY KERNEL

Consider a n-qubit system, where each classical input $\mathbf{x}$ produces a density matrix $\boldsymbol{\rho}(\mathbf{x}) \in \mathbb{C}^{2^n\times 2^n}$. Let us define the quantum fidelity kernel as

$$\kappa_Q(\mathbf{x},\mathbf{x}') \;=\; \mathrm{tr}\big(\boldsymbol{\rho}(\mathbf{x})^\dagger\boldsymbol{\rho}(\mathbf{x}')\big).$$

Note that the associated RKHS has dimension at most 4^n [Kübler et al., 2021].

It is well known that for any D-dimensional linear kernel, the information gain γ_T scales as $\mathcal{O}(D\log T)$ [Srinivas et al., 2010]. A similar argument reveals that γ_T of κ_Q is $\mathcal{O}(4^n\log T)$. Concretely, we can define the feature map

$$\boldsymbol{\phi}(\mathbf{x}) \;=\; \mathrm{vec}\big(\boldsymbol{\rho}(\mathbf{x})\big) \;\in\; \mathbb{C}^{4^n},$$

satisfying $\|\boldsymbol{\phi}(\mathbf{x})\| \le 1$ [Schuld, 2021], where $\mathrm{vec}(\cdot)$ flattens a $2^n\times 2^n$ matrix into a vector of dimension 4^n. Subsequently

$$\kappa_Q(\mathbf{x},\mathbf{x}') \;=\; \mathrm{Tr}\big(\boldsymbol{\rho}(\mathbf{x})^\dagger\boldsymbol{\rho}(\mathbf{x}')\big) = \boldsymbol{\phi}(\mathbf{x})^\dagger\boldsymbol{\phi}(\mathbf{x}').$$

Corresponding to the observed T data points $\{\mathbf{x}_1,\dots,\mathbf{x}_T\}$, we then collect the feature vectors into the following matrix:

$$\mathbf{X}_T \;=\; \begin{bmatrix} | & | & \cdots & | \\ \boldsymbol{\phi}(\mathbf{x}_1) & \boldsymbol{\phi}(\mathbf{x}_2) & \cdots & \boldsymbol{\phi}(\mathbf{x}_T) \\ | & | & \cdots & | \end{bmatrix} \;\in\; \mathbb{C}^{4^n\times T}.$$

Recall that, for reward function $\mathbf{f}_A = [f^*(\mathbf{x})]_{\mathbf{x}\in A}$ defined on domain $A \subset \mathcal{X}$, the *information gain* on observing T points $\{\mathbf{x}_1, ..., \mathbf{x}_T\}$ with observations $\mathbf{y}_A$ is $\gamma_T = \max_{A\subset\mathcal{X}:|A|=T} I(\mathbf{y}_A; \mathbf{f}_A)$.

Furthermore,

$$\begin{aligned} I(\mathbf{y}_A; \mathbf{f}_A) &= \frac{1}{2}\log|\mathbf{I} + \sigma^{-2}\mathbf{K}_A| \\ &= \frac{1}{2}\log|\mathbf{I} + \sigma^{-2}\mathbf{X}_T^\dagger\mathbf{X}_T| \\ &= \frac{1}{2}\log|\mathbf{I} + \sigma^{-2}\mathbf{X}_T\mathbf{X}_T^\dagger| \\ &= \frac{1}{2}\log\prod_{i=1}^{4^n}(1+\sigma^{-2}\lambda_i) \\ &\leq \frac{1}{2}4^n\log\big(1+\sigma^{-2}\cdot O(T)\big) \end{aligned}$$

where λ_i's are the eigenvalues of $\mathbf{X}_T\mathbf{X}_T^\dagger$. The last inequality follows since the largest eigenvalue of $\mathbf{X}_T\mathbf{X}_T^\dagger$ is $O(T)$, since all $\|\boldsymbol{\phi}(\mathbf{x})\| \leq 1$. Hence, γ_T for the quantum kernel is upper bounded by $\mathcal{O}(4^n \log T)$.

C PROOF AND DISCUSSION ON THEOREM 2

C.1 PROOF OF THEOREM 2 (I)

We will show that the RKHS associated with a (summed) projected kernel is a subspace of the full kernel's RKHS.

Proof. By standard arguments (e.g., the Moore–Aronszajn theorem or see Kübler et al. [2021]), the n-qubit quantum kernel

$$\kappa_Q(\mathbf{x}, \mathbf{x}') = \mathrm{Tr}[\boldsymbol{\rho}(\mathbf{x})\boldsymbol{\rho}(\mathbf{x}')^\dagger]$$

defines the RKHS

$$\mathcal{H}_{\kappa_Q}(\mathcal{X}) = \{f : f(\cdot) = \mathrm{Tr}[\boldsymbol{\rho}(\cdot)\mathbf{H}], \mathbf{H} \in \mathbb{C}^{2^n\times 2^n}, \mathbf{H} = \mathbf{H}^\dagger\},$$

of real-valued functions $f : \mathcal{X} \to \mathbb{R}$.

Likewise, for the projected quantum kernel onto the set of qubits $s \subseteq \{1, ..., n\}$

$$\kappa_Q^s(\mathbf{x}, \mathbf{x}') = \mathrm{Tr}(\boldsymbol{\rho}^s(\mathbf{x})\boldsymbol{\rho}^s(\mathbf{x}')),$$

where the reduced density matrix $\boldsymbol{\rho}^s(\mathbf{x})$ is obtained from the full $\boldsymbol{\rho}(\mathbf{x})$ by tracing out the complementary qubits s^c as follows

$$\boldsymbol{\rho}^s(\mathbf{x}) = \mathrm{Tr}_{s^c}[\boldsymbol{\rho}(\mathbf{x})].$$

The associated RKHS $\mathcal{H}^s_{\kappa_Q}(\mathcal{X})$ is comprised of functions f^s of the form

$$\begin{aligned} \mathcal{H}^s_{\kappa_Q}(\mathcal{X}) = \{f^s : f^s(\cdot) = \mathrm{Tr}[\boldsymbol{\rho}^s(\cdot)\mathbf{H}'], \\ \mathbf{H}' \in \mathbb{C}^{2^{|s|}\times 2^{|s|}}, \mathbf{H}' = (\mathbf{H}')^\dagger\}. \end{aligned}$$

The definition of partial trace implies that, for any operator $\mathbf{H}' \in \mathbb{C}^{2^{|s|}\times 2^{|s|}}$,

$$f^s(\mathbf{x}) = \mathrm{Tr}[\boldsymbol{\rho}^s(\mathbf{x})\mathbf{H}'] = \mathrm{Tr}\big[\boldsymbol{\rho}(\mathbf{x})\,(\mathbf{H}'\otimes\mathbf{I}_{s^c})\big].$$

Now note $\mathbf{H} = \mathbf{H}' \otimes \mathbf{I}_{s^c}$ is an operator in $\mathbb{C}^{2^n\times 2^n}$. Hence any $f^s \in \mathcal{H}^s_{\kappa_Q}(\mathcal{X})$ must also belong to the full RKHS $\mathcal{H}_{\kappa_Q}(\mathcal{X})$. Therefore, we have $\mathcal{H}^s_{\kappa_Q}(\mathcal{X}) \subseteq \mathcal{H}_{\kappa_Q}(\mathcal{X})$. Subsequently, since $\kappa_{\mathbb{S}_b} = \frac{1}{|\mathbb{S}_b|}\sum_{s\in\mathbb{S}_b}\kappa_Q^s$, we have $\mathcal{H}_{\kappa_{\mathbb{S}_b}}(\mathcal{X}) \subseteq \mathcal{H}_{\kappa_Q}(\mathcal{X})$. □

C.2 EXTENSION OF THEOREM 2 (I)

We showed that the RKHS of a projected kernel is a subspace of the full kernel's RKHS. Here, we present a stronger result: this subspace relationship arises from a simple orthogonal projection in the underlying feature space:

Proposition 1. *Suppose $\boldsymbol{\phi}_Q$ is the feature map corresponding to κ_Q such that $\kappa_Q(\mathbf{x}, \mathbf{x}') = \langle \boldsymbol{\phi}_Q(\mathbf{x}), \boldsymbol{\phi}_Q(\mathbf{x}') \rangle$. Then for each $s \subseteq \{1, ..., n\}$, there exists a unique self-adjoint idempotent operator P such that*

$$\kappa_Q^s(\mathbf{x}, \mathbf{x}') = \langle P\boldsymbol{\phi}_Q(\mathbf{x}), P\boldsymbol{\phi}_Q(\mathbf{x}') \rangle.$$

This allows us to define a feature map of κ_Q^s as $\boldsymbol{\phi}_s := P\boldsymbol{\phi}_Q$.

Proof. By the representer theorem, there is a Hilbert space $\mathcal{F}_Q$ (feature space) and a map $\boldsymbol{\phi}_Q : \mathcal{X} \to \mathcal{F}_Q$ (feature map) such that

$$\kappa_Q(\mathbf{x}, \mathbf{x}') \;=\; \langle \boldsymbol{\phi}_Q(\mathbf{x}),\, \boldsymbol{\phi}_Q(\mathbf{x}') \rangle_{\mathcal{F}_Q}.$$

It is also standard that the RKHS $\mathcal{H}_{\kappa_Q}(\mathcal{X})$ is isometrically isomorphic to the closed linear span of $\{\boldsymbol{\phi}_Q(\mathbf{x}) : \mathbf{x} \in \mathcal{X}\}$ in $\mathcal{F}_Q$. Specifically, one can define an isomorphism $\Phi : \overline{\operatorname{span}\{\boldsymbol{\phi}_Q(\mathbf{x}) : \mathbf{x} \in \mathcal{X}\}} \to \mathcal{H}_{\kappa_Q}(\mathcal{X})$ by

$$\Phi\left(\sum_i \alpha_i\, \boldsymbol{\phi}_Q(\mathbf{x}_i)\right) = \sum_i \alpha_i\, \kappa_Q(\cdot, \mathbf{x}_i).$$

Because $\mathcal{H}^s_{\kappa_Q}(\mathcal{X}) \subseteq \mathcal{H}_{\kappa_Q}(\mathcal{X})$ is a closed subspace, we define, with the above Φ,

$$\mathcal{V} \;=\; \Phi^{-1}\big(\mathcal{H}^s_{\kappa_Q}(\mathcal{X})\big) \;\subseteq\; \overline{\operatorname{span}\{\boldsymbol{\phi}_Q(\mathbf{x})\}} \;\subseteq\; \mathcal{F}_Q.$$

Then $\mathcal{V}$ is indeed a closed subspace of $\mathcal{F}_Q$. By the Hilbert projection theorem [Conway, 2007], there exists a unique linear orthogonal projector $P : \mathcal{F}_Q \to \mathcal{F}_Q$, such that $P^2 = P$, $P = P^\dagger$ and $\operatorname{Range}(P) = \mathcal{V}$. We can then use this P to define a new feature map

$$\tilde{\boldsymbol{\phi}}^s_Q(\mathbf{x}) \;=\; P\,\boldsymbol{\phi}_Q(\mathbf{x}),$$

and the corresponding kernel,

$$\tilde{\kappa}^s_Q(\mathbf{x}, \mathbf{x}') \;=\; \langle \tilde{\boldsymbol{\phi}}^s_Q(\mathbf{x}),\, \tilde{\boldsymbol{\phi}}^s_Q(\mathbf{x}') \rangle_{\mathcal{F}_Q} \;=\; \langle P\,\boldsymbol{\phi}_Q(\mathbf{x}),\, P\,\boldsymbol{\phi}_Q(\mathbf{x}') \rangle_{\mathcal{F}_Q}.$$

Since P is an orthogonal projector, $\tilde{\kappa}^s_Q$ is positive semidefinite. Let $\tilde{\mathcal{H}}^s_{\kappa_Q}(\mathcal{X})$ be its associated RKHS. We now show that $\tilde{\mathcal{H}}^s_{\kappa_Q}(\mathcal{X}) = \mathcal{H}^s_{\kappa_Q}(\mathcal{X})$.

(i) $\tilde{\mathcal{H}}^s_{\kappa_Q}(\mathcal{X}) \subseteq \mathcal{H}^s_{\kappa_Q}(\mathcal{X})$.

Any $f \in \tilde{\mathcal{H}}^s_{\kappa_Q}(\mathcal{X})$ has the form

$$\begin{aligned} f(\cdot) &= \sum_{i=1}^N \alpha_i\, \tilde{\kappa}^s_Q(\cdot, \mathbf{x}_i) \\ &= \sum_{i=1}^N \alpha_i\, \big\langle \tilde{\boldsymbol{\phi}}^s_Q(\cdot),\, \tilde{\boldsymbol{\phi}}^s_Q(\mathbf{x}_i) \big\rangle_{\mathcal{F}_Q} \\ &= \Big\langle \sum_{i=1}^N \alpha_i\, \tilde{\boldsymbol{\phi}}^s_Q(\mathbf{x}_i),\; \tilde{\boldsymbol{\phi}}^s_Q(\cdot) \Big\rangle_{\mathcal{F}_Q}. \end{aligned}$$

Define $\mathbf{v} \;:=\; \sum_{i=1}^N \alpha_i\, \tilde{\boldsymbol{\phi}}^s_Q(\mathbf{x}_i) \;\in\; \mathcal{V}$. Then

$$f_{\mathbf{v}}(\cdot) \;:=\; \langle \mathbf{v},\; \tilde{\boldsymbol{\phi}}^s_Q(\cdot) \rangle_{\mathcal{F}_Q} \;=\; \langle \mathbf{v},\; P\boldsymbol{\phi}_Q(\cdot) \rangle_{\mathcal{F}_Q}.$$

But since $\mathbf{v} \in \operatorname{Range}(P)$, $\mathbf{v} = P\,\mathbf{v}$, and since P is self-adjoint,

$$f_{\mathbf{v}}(\cdot) = \langle \mathbf{v},\; P\boldsymbol{\phi}_Q(\cdot) \rangle_{\mathcal{F}_Q} = \langle \mathbf{v},\; \boldsymbol{\phi}_Q(\cdot) \rangle_{\mathcal{F}_Q}$$

By the isomorphism property of Φ, $f_{\mathbf{v}}$ is exactly $\Phi(\mathbf{v}) \in \mathcal{H}_{\kappa_Q}(\mathcal{X})$. Furthermore, since $\mathbf{v} \in \mathcal{V}$, we have $\Phi(\mathbf{v}) \in \mathcal{H}^s_{\kappa_Q}(\mathcal{X})$, hence $f \in \mathcal{H}^s_{\kappa_Q}(\mathcal{X})$.

(ii) $\mathcal{H}^s_{\kappa_Q}(\mathcal{X}) \subseteq \tilde{\mathcal{H}}^s_{\kappa_Q}(\mathcal{X})$.

Similarly, each $f \in \mathcal{H}^s_{\kappa_Q}(\mathcal{X})$ can be written as $\Phi(\mathbf{v})$ with some vector $\mathbf{v} \in \mathcal{V}$, i.e.,

$$f(\cdot) \;=\; \Phi(\mathbf{v}) \;=\; \langle \mathbf{v},\, \boldsymbol{\phi}_Q(\cdot)\rangle_{\mathcal{F}_Q}.$$

We have $\mathbf{v} = P\mathbf{v}$ since $\mathbf{v} \in \mathrm{Range}(P)$, thus

$$f(\cdot) \;=\; \langle \mathbf{v},\, \boldsymbol{\phi}_Q(\cdot)\rangle \;=\; \langle \mathbf{v},\, P\,\boldsymbol{\phi}_Q(\cdot)\rangle \;=\; \langle \mathbf{v},\, \tilde{\boldsymbol{\phi}}^s_Q(\cdot)\rangle,$$

which, by definition, belongs to $\tilde{\mathcal{H}}^s_{\kappa_Q}(\mathcal{X})$.

The relationships **(i)** and **(ii)** together imply that $\mathcal{H}^s_{\kappa_Q}(\mathcal{X}) = \tilde{\mathcal{H}}^s_{\kappa_Q}(\mathcal{X})$. Finally, as RKHS is uniquely determined by its kernel,

$$\kappa^s_Q(\mathbf{x}, \mathbf{x}') \;=\; \tilde{\kappa}^s_Q(\mathbf{x}, \mathbf{x}'), \quad \forall\, \mathbf{x}, \mathbf{x}' \in \mathcal{X}.$$

This completes the proof that κ^s_Q is realized exactly by the projected feature map $\tilde{\boldsymbol{\phi}}^s_Q = P\,\boldsymbol{\phi}_Q$. □

C.3 PROOF OF THEOREM 2 (II)

This proof establishes a high-probability bound on the misspecification error, ε, when approximating a global function $f^*(\mathbf{x}) = \mathrm{Tr}[\mathbf{O}\boldsymbol{\rho}(\mathbf{x})]$ with one constructed from a "low-weight" or "local" observable $\mathbf{H}$.

From Function Approximation to Operator Distance. To understand the misspecification error, it's helpful to start with a simple illustrative case. Suppose our approximate model is restricted to information from a single b-qubit subsystem, denoted by the index set s. The function $f(\mathbf{x})$ in this model then depends only on the reduced density matrix $\boldsymbol{\rho}^s(\mathbf{x}) = \mathrm{Tr}_{s^c}(\boldsymbol{\rho}(\mathbf{x}))$, i.e., $f(\mathbf{x}) = \mathrm{Tr}(\mathbf{H}\boldsymbol{\rho}^s(\mathbf{x}))$, where $\mathbf{H}$ is a b-qubit Hermitian observable. This can be rewritten on the full Hilbert space by letting $\mathbf{H}' = \mathbf{H} \otimes \mathbf{I}_{s^c}$. Then $f(\mathbf{x}) = \mathrm{Tr}(\mathbf{H}'\boldsymbol{\rho}(\mathbf{x}))$. The error of the best possible approximation for a given $\mathbf{O}$ is:

$$\begin{aligned}
\varepsilon_s(\mathbf{O}) &= \inf_f \|f - f^*\|_\infty \\
&= \inf_{\mathbf{H}'} \sup_{\boldsymbol{\rho}(\mathbf{x})} |\mathrm{Tr}(\mathbf{O}\boldsymbol{\rho}(\mathbf{x})) - \mathrm{Tr}(\mathbf{H}'\boldsymbol{\rho}(\mathbf{x}))| \\
&= \inf_{\mathbf{H}'} \sup_{\boldsymbol{\rho}(\mathbf{x})} |\mathrm{Tr}[(\mathbf{O} - \mathbf{H}')\boldsymbol{\rho}(\mathbf{x})]| \\
&\le \inf_{\mathbf{H}'} \|\mathbf{O} - \mathbf{H}'\|_{op}
\end{aligned}$$

Generalization to Low-Weight Observables. Now we generalize the previous setup, to approximate with a combination of subsystems. Let $\mathcal{H}_n = (\mathbb{C}^2)^{\otimes n}$ be the Hilbert space of an n-qubit system, with dimension $= 2^n$. Let $\mathcal{A}_n$ be the vector space of Hermitian operators on $\mathcal{H}_n$, with dimension $\dim(\mathcal{A}_n) = 4^n$. The Hilbert-Schmidt inner product on this space is $\langle \mathbf{A}, \mathbf{B}\rangle = \mathrm{Tr}(\mathbf{A}^\dagger\mathbf{B})$.

In a quantum system with n qubits, any observable $\mathbf{O} \in \mathcal{A}_n$ can be uniquely decomposed in the Pauli basis: $\mathbf{O} = \sum_P c_P\mathbf{P}$, where $\mathbf{P}$ are n-qubit Pauli operators (i.e., P is a tensor product of n single-qubit matrices. $\mathbf{P} = \mathbf{P}_1 \otimes \mathbf{P}_2 \otimes \cdots \otimes \mathbf{P}_n$, $\mathbf{P}_i \in \{\mathbf{I}, \mathbf{X}, \mathbf{Y}, \mathbf{Z}\}$, each being a 2×2 matrix). The **Pauli weight** of a Pauli operator $\mathbf{P}$, denoted $\mathrm{wt}(\mathbf{P})$, is the number of qubits on which it acts non-trivially (i.e., not as the identity) (recall if $\mathbf{H}' = \mathbf{H} \otimes \mathbf{I}_{s^c}$ is from a b-qubit observable $\mathbf{H}$, $\mathbf{H}'$ acts as identity on qubits outside b).

Our goal is to approximate f^* with (sum of) projected kernels. For a fully expressive (4^n-dimensional) full quantum kernel κ_Q, this is equivalent to approximating a global observable $\mathbf{O}$ using a combination of such "local" (low Pauli weight) operators. We define the subspace of low-weight operators, $\mathcal{A}_{\le b}$, as the span of all Pauli operators with weight at most b:

$$\mathcal{A}_{\le b} = \mathrm{span}\{\mathbf{P} \in \mathcal{P}_n \mid \mathrm{wt}(\mathbf{P}) \le b\}$$

where $\mathcal{P}_n$ is the set of n-qubit Pauli operators.

Now the error of the best possible approximation of $\mathbf{O}$ using an operator $\mathbf{H}' \in \mathcal{A}_{\leq b}$ is given by

$$\varepsilon_{\leq b}(\mathbf{O}) = \inf_{\mathbf{H}' \in \mathcal{A}_{\leq b}} \|\mathbf{O} - \mathbf{H}'\|_{op}$$

Further, the operator norm is upper-bounded by the Hilbert-Schmidt norm, $\|\mathbf{A}\|_{op} \leq \|\mathbf{A}\|_2$. Therefore, we can find an upper bound by analyzing the Hilbert-Schmidt distance:

$$\varepsilon_{\leq b}(\mathbf{O}) \leq \inf_{\mathbf{H}' \in \mathcal{A}_{\leq b}} \|\mathbf{O} - \mathbf{H}'\|_2 \tag{15}$$

The problem is now to find the minimum of $\|\mathbf{O} - \mathbf{H}'\|_2$ over all $\mathbf{H}' \in \mathcal{A}_{\leq b}$. In a Hilbert space, the point in a subspace that is closest to an external point is its orthogonal projection.

Let $P_{\leq b} : \mathcal{A}_n \to \mathcal{A}_{\leq b}$ be the orthogonal projection operator. The infimum is achieved when $\mathbf{H}' = P_{\leq b}(\mathbf{O})$.

$$\inf_{\mathbf{H}' \in \mathcal{A}_{\leq b}} \|\mathbf{O} - \mathbf{H}'\|_2 = \|\mathbf{O} - P_{\leq b}(\mathbf{O})\|_2 \tag{16}$$

By the property of orthogonal projections, we have:

$$\|\mathbf{O} - P_{\leq b}(\mathbf{O})\|_2^2 = \|\mathbf{O}\|_2^2 - \|P_{\leq b}(\mathbf{O})\|_2^2 \tag{17}$$

Expected Error for Spherical Operators. Here we assume $\mathbf{O}$ is drawn uniformly from the sphere of operators with $\|\mathbf{O}\|_2 = 1$ (For simplicity in this step, we'll assume the norm is one, i.e., $||\mathbf{O}||_2 = 1$. This analysis can be trivially extended to any arbitrary norm $||\mathbf{O}||_2 = B$ by scaling the final result by B). Then $\|\mathbf{O} - P_{\leq b}(\mathbf{O})\|_2^2 = \|\mathbf{O}\|_2^2 - \|P_{\leq b}(\mathbf{O})\|_2^2 = 1 - \|P_{\leq b}(\mathbf{O})\|_2^2$. Taking the expectation:

$$\mathbb{E}\left[\|\mathbf{O} - P_{\leq b}(\mathbf{O})\|_2^2\right] = 1 - \mathbb{E}\left[\|P_{\leq b}(\mathbf{O})\|_2^2\right]$$

We use the following result for random projections.

Lemma 1. *Let $\mathbf{O}$ be a random vector chosen uniformly from the unit sphere in a D-dimensional space V. Let P_S be the orthogonal projection onto a d-dimensional subspace S. Then, $\mathbb{E}[\|P_S(\mathbf{O})\|_2^2] = d/D$.*

Proof. Let $\{\mathbf{e}_1, \ldots, \mathbf{e}_D\}$ be an orthonormal basis for V such that $\{\mathbf{e}_1, \ldots, \mathbf{e}_d\}$ is an orthonormal basis for S. Any vector $\mathbf{O}$ on the unit sphere can be written as $\mathbf{O} = \sum_{i=1}^{D} c_i \mathbf{e}_i$ with $\sum_i c_i^2 = 1$. The projection is $P_S(\mathbf{O}) = \sum_{i=1}^{d} c_i \mathbf{e}_i$, and its squared norm is $\|P_S(\mathbf{O})\|_2^2 = \sum_{i=1}^{d} c_i^2$. By linearity of expectation, $\mathbb{E}\left[\|P_S(\mathbf{O})\|_2^2\right] = \sum_{i=1}^{d} \mathbb{E}[c_i^2]$. Because the distribution of $\mathbf{O}$ is spherically symmetric, the expected value of its squared component along any basis vector is the same, i.e., $\mathbb{E}[c_i^2] = C$ for all i for some constant C. We know $\mathbb{E}[\|\mathbf{O}\|_2^2] = \mathbb{E}[1] = 1$. Also, $\mathbb{E}[\|\mathbf{O}\|_2^2] = \mathbb{E}[\sum_{i=1}^{D} c_i^2] = \sum_{i=1}^{D} \mathbb{E}[c_i^2] = D \cdot C$. Thus, $D \cdot C = 1 \implies C = 1/D$. Finally, $\mathbb{E}\left[\|P_S(\mathbf{O})\|_2^2\right] = \sum_{i=1}^{d} C = d \cdot C = d/D$. □

High-Probability Bound for Spherical Operators. Next we derive a high-probability bound on the error by determining the probability distribution of the squared norm of the projection, $\|P_{\leq b}(\mathbf{O})\|_2^2$.

Lemma 2. *Let $\mathbf{O}$ be a random vector chosen uniformly from the unit sphere in a D-dimensional space V. Let P_S be the orthogonal projection onto a d-dimensional subspace S. The random variable $X = \|P_S(\mathbf{O})\|_2^2$ follows a Beta distribution with parameters $\alpha = d/2$ and $\beta = (D-d)/2$.*

$$\|P_S(\mathbf{O})\|_2^2 \sim \text{Beta}\left(\frac{d}{2}, \frac{D-d}{2}\right)$$

Proof of Lemma. A random vector $\mathbf{O}$ on the unit sphere can be generated by taking a standard D-dimensional Gaussian vector $\mathbf{G} = (g_1, \ldots, g_D)$ where each $g_i \sim \mathcal{N}(0,1)$ are i.i.d., and normalizing it: $\mathbf{O} = \mathbf{G}/\|\mathbf{G}\|_2$. Let $\{\mathbf{e}_1, \ldots, \mathbf{e}_D\}$ be an orthonormal basis for V where $\{\mathbf{e}_1, \ldots, \mathbf{e}_d\}$ is a basis for S. The coefficients of $\mathbf{O}$ are $c_i = \langle \mathbf{O}, \mathbf{e}_i \rangle = g_i/\|\mathbf{G}\|_2$. The squared norm of the projection is:

$$\|P_S(\mathbf{O})\|_2^2 = \sum_{i=1}^{d} c_i^2 = \frac{\sum_{i=1}^{d} g_i^2}{\sum_{j=1}^{D} g_j^2}$$

Let $U = \sum_{i=1}^{d} g_i^2$ and $V = \sum_{j=d+1}^{D} g_j^2$. By definition of the chi-squared distribution, $U \sim \chi^2(d)$ and $V \sim \chi^2(D-d)$. Since the g_i are independent, U and V are independent. The random variable is $X = U/(U+V)$. A random variable constructed as the ratio of a chi-squared variable to the sum of itself and another independent chi-squared variable is known to follow the Beta distribution. Specifically, if $U \sim \chi^2(\nu_1)$ and $V \sim \chi^2(\nu_2)$, then $U/(U+V) \sim \text{Beta}(\nu_1/2, \nu_2/2)$. Substituting $\nu_1 = d$ and $\nu_2 = D - d$ proves the lemma. □

The Beta distribution is known to be strongly concentrated around its mean. In fact, it is $\frac{1}{4(\alpha+\beta)}$-sub-gaussian [Marchal and Arbel, 2017]. Hence we can use Hoeffding's inequality to obtain a concentration inequality: For a random variable $X \sim \text{Beta}(\alpha, \beta)$ with mean $\mu = \alpha/(\alpha+\beta)$, a tail bound is given by:

$$\mathbb{P}(|X - \mu| \geq t) \leq 2\exp\left(-\frac{t^2}{2\sigma^2}\right) \leq 2e^{-2(\alpha+\beta)t^2}$$

In our case, $\mu = \frac{d/2}{D/2} = d/D$ and $\alpha + \beta = D/2$. We are interested in a lower bound on X to get an upper bound on the error.

$$\mathbb{P}\left(X \leq \frac{d}{D} - t\right) \leq e^{-Dt^2}$$

Setting the failure probability to $\delta = e^{-Dt^2}$ gives $t = \sqrt{\ln(1/\delta)/D}$. This means with probability at least $1 - \delta$, we have $X \geq d/D - \sqrt{\ln(1/\delta)/D}$.

Corollary 1 (Spherical Error Bound)**.** *With probability at least $1 - \delta$, the squared Hilbert-Schmidt norm of the error is bounded by:*

$$\|\mathbf{O} - P_{\leq b}(\mathbf{O})\|_2^2 \leq \left(1 - \frac{\sum_{w=0}^{b}\binom{n}{w}3^w}{4^n}\right) + \sqrt{\frac{\ln(1/\delta)}{4^n}}.$$

Proof. From the Pythagorean identity, $\|\mathbf{O} - P_{\leq b}(\mathbf{O})\|_2^2 = 1 - \|P_{\leq b}(\mathbf{O})\|_2^2$. Let $d = \dim(\mathcal{A}_{\leq b})$, using the high-probability lower bound on $\|P_{\leq b}(\mathbf{O})\|_2^2$ gives:

$$\begin{aligned}\|\mathbf{O} - P_{\leq b}(\mathbf{O})\|_2^2 &\leq 1 - \left(\frac{d}{D} - \sqrt{\frac{\ln(1/\delta)}{D}}\right) \\ &= \left(1 - \frac{d}{D}\right) + \sqrt{\frac{\ln(1/\delta)}{D}}\end{aligned}$$

Next we show that $D = 4^n$ and $d = \sum_{w=0}^{b}\binom{n}{w}3^w$ in our case.

Recall any operator $\mathbf{O} \in \mathcal{A}_n$ can be uniquely decomposed in the Pauli basis: $\mathbf{O} = \sum_{\mathbf{P}} c_P \mathbf{P}$, where $\mathbf{P}$ are n-qubit Pauli operators. And it follows directly from our assumption that $c_{\mathbf{P}} \neq 0$ for all $c_{\mathbf{P}}$ with probability 1. The n-qubit Pauli operators form a basis of $\mathcal{A}_n$, hence $D = 4^n$.

d is the dimension of the subspace $\mathcal{A}_{\leq b}$, the number of linearly independent Pauli operators with weight at most b. Note the following:

- The number of ways to choose w qubits to act on from a total of n is $\binom{n}{w}$.
- For each of the chosen w qubits, there are 3 non-identity Pauli choices ($\sigma_x, \sigma_y, \sigma_z$).
- The total number of Pauli strings with weight exactly w is $\binom{n}{w}3^w$.

The dimension d is the sum over all allowed weights (from 0 to b):

$$d = \dim(\mathcal{A}_{\leq b}) = \sum_{w=0}^{b}\binom{n}{w}3^w$$

Substitute the expression for d and D completes the proof.

□

Now it only remains for us to show that

$$p := 1 - \frac{\sum_{w=0}^{b} \binom{n}{w} 3^w}{4^n} \leq e^{-\frac{8n}{3}\left(\frac{b+1}{n} - \frac{3}{4}\right)^2}.$$

For this purpose we note that

$$\begin{aligned} p &= \frac{1}{4^n} \sum_{w=b+1}^{n} \binom{n}{w} 3^w \\ &= \sum_{w=b+1}^{n} \binom{n}{w} \Big(\frac{3}{4}\Big)^w \Big(\frac{1}{4}\Big)^{n-w} \\ &= \Pr\Big(\frac{1}{n}\sum_{i=1}^{n} Z_i \geq \frac{b+1}{n}\Big), \end{aligned}$$

where Z_i are independent Bernoulli random variables satisfying $\Pr(Z_i = 1) = \frac{3}{4}$. If $\frac{b+1}{n} \geq \frac{3}{4}$ which is the regime of interest, by the Chernoff bound,

$$p \leq \exp\Big(-nD_{\mathrm{KL}}\Big(\frac{b+1}{n}\Big\|\frac{3}{4}\Big)\Big),$$

where $D_{\mathrm{KL}}(q\|p)$ denotes the KL divergence between Bernoulli distributions with parameters q and p. Furthermore, it holds that $D_{\mathrm{KL}}(q+x\|q) \geq \frac{x^2}{2q(1-q)}$. Thus,

$$D_{\mathrm{KL}}\Big(\frac{b+1}{n}\Big\|\frac{3}{4}\Big) \geq \frac{(\frac{b+1}{n} - \frac{3}{4})^2}{2 \cdot \frac{3}{4} \cdot \frac{1}{4}} = \frac{8}{3}\Big(\frac{b+1}{n} - \frac{3}{4}\Big)^2.$$

We conclude that

$$p \leq \exp\Big(-\frac{8n}{3}\Big(\frac{b+1}{n} - \frac{3}{4}\Big)^2\Big).$$

C.4 JUSTIFICATION OF THE RANDOM OPERATOR ASSUMPTION

The proof of Theorem 2(ii) relies on the assumption that the global observable $\mathbf{O}$ is drawn uniformly at random from the sphere of operators with a fixed norm. While this may seem abstract and restrictive, it can be motivated from an isotropic Gaussian prior on the observable $\mathbf{O}$ (which induces a GP prior on $f^*(\mathbf{x}) = \mathrm{Tr}[\mathbf{O}\boldsymbol{\rho}(\mathbf{x})]$). This assumption corresponds to conditioning that Gaussian prior on $\|\mathbf{O}\|_2$ (equivalently, normalizing a Gaussian draw).

Here's the connection:

From Gaussian Processes to Operator Priors. A function f^* drawn from a GP is defined by a prior distribution over a function space. In our setting, where functions are of the form $f^*(\mathbf{x}) = \mathrm{Tr}[\mathbf{O}\boldsymbol{\rho}(\mathbf{x})]$, this prior on functions induces a prior on the operator $\mathbf{O}$. We can represent any observable $\mathbf{O}$ by expanding it in an orthonormal basis of Hermitian operators, $\{\mathbf{B}_i\}$ (e.g., the normalized Pauli basis):

$$\mathbf{O} = \sum_{i=1}^{4^n} c_i \mathbf{B}_i$$

A simple and common case of GP prior is one where the coefficients c_i are assumed to be independent and identically distributed (i.i.d.) Gaussian random variables, i.e., $c_i \sim \mathcal{N}(0, \sigma^2)$. This is equivalent to placing a white-noise Gaussian prior on the operator $\mathbf{O}$.

From Gaussian Vectors to Uniform Spherical Vectors. The vector of coefficients, $\mathbf{c} = (c_1, c_2, \ldots, c_{4^n})$, is therefore a vector whose components are i.i.d. Gaussians. A fundamental property of multivariate Gaussian distributions is that, taking such a vector and normalize it by its length, the resulting direction vector is uniformly distributed on the surface of the unit sphere. So, if we normalize our coefficient vector to get $\mathbf{c}' = \mathbf{c}/\|\mathbf{c}\|_2$, the vector $\mathbf{c}'$ will be uniformly random on the unit sphere in $\mathbb{R}^{4^n}$.

Hence, the operator constructed from these normalized coefficients, $\mathbf{O}' = \sum_i c'_i \mathbf{B}_i$, will have a unit Hilbert-Schmidt norm ($\|\mathbf{O}'\|_2 = \|\mathbf{c}'\|_2 = 1$) and its "direction" in the space of operators will be uniformly random. This is precisely the assumption made in our proof.

C.5 STRUCTURED OBSERVABLES: LOCALITY, PAULI-WEIGHT TAILS, AND REGRET

The bound in Theorem 2(ii) provides a useful worst-case baseline, but it is pessimistic for many physical tasks. In many quantum optimization problems, the observable $\mathbf{O}$ has additional structure, such as locality or a rapidly decaying high-weight Pauli tail. In this subsection, we show how such structure directly improves the LPQK misspecification error and the resulting regret bound.

Recall in a quantum system with n qubits, any observable $\mathbf{O} \in \mathcal{A}_n$ can be uniquely decomposed in the normalized Pauli basis: $\mathbf{O} = \sum_P c_P \mathbf{P}$, where $\mathbf{P}$ are n-qubit Pauli operators. And the **Pauli weight** of a Pauli operator $\mathbf{P}$, denoted $\mathrm{wt}(\mathbf{P})$, is the number of qubits on which it acts non-trivially (i.e., not as the identity).

Define the low-weight and high-weight components

$$\mathbf{O}_{\le b} := \sum_{\mathrm{wt}(\mathbf{P}) \le b} c_{\mathbf{P}} \mathbf{P}, \qquad \mathbf{O}_{>b} := \sum_{\mathrm{wt}(\mathbf{P}) > b} c_{\mathbf{P}} \mathbf{P}.$$

The summed LPQK over all subsystems $|s| \le b$ retains exactly the observable components of Pauli weight at most b. Therefore, for $f^*(\mathbf{x}) = \mathrm{Tr}[\mathbf{O}\boldsymbol{\rho}(\mathbf{x})]$, the best LPQK approximation satisfies

$$\varepsilon_b := \inf_{f \in \mathcal{H}_{\kappa_{\mathbb{S}_b}}(\mathcal{X})} \|f - f^*\|_\infty \le \sup_{\mathbf{x} \in \mathcal{X}} |\mathrm{Tr}[\mathbf{O}_{>b}\boldsymbol{\rho}(\mathbf{x})]| \le \|\mathbf{O}_{>b}\|_{\mathrm{op}} \le \|\mathbf{O}_{>b}\|_{\mathrm{HS}} = \left(\sum_{\mathrm{wt}(\mathbf{P}) > b} c_{\mathbf{P}}^2 \right)^{1/2}. \tag{18}$$

That is, the LPQK misspecification error is controlled by the high-weight Pauli tail of the target observable.

On the other hand, the effective dimension of the summed LPQK is $S_b := \sum_{w=0}^{b} \binom{n}{w} 3^w$, and hence its information gain satisfies

$$\gamma_T(b) = \mathcal{O}(S_b \log T).$$

Substituting this into the EC-GP-UCB bound in Eq. (8) gives, suppressing logarithmic factors and constants,

$$R_T(b) = \widetilde{\mathcal{O}}\left((B\sqrt{S_b} + S_b)\sqrt{T} + \varepsilon_b T \sqrt{S_b} \right). \tag{19}$$

Example 1: k-local observables. Suppose $\mathbf{O}$ is k-*local*, meaning that its Pauli expansion contains no terms of weight larger than k:

$$c_{\mathbf{P}} = 0 \qquad \text{for all } \mathrm{wt}(\mathbf{P}) > k.$$

Then $\mathbf{O}_{>b} = 0$ for all $b \ge k$, and therefore

$$\varepsilon_b = 0 \qquad \text{for all } b \ge k.$$

Thus, choosing $b = k$ makes the LPQK surrogate realizable while using only

$$S_k = \sum_{w=0}^{k} \binom{n}{w} 3^w = \mathcal{O}(n^k)$$

features for fixed k, instead of the full 4^n-dimensional quantum feature space. The regret bound becomes

$$R_T(k) = \widetilde{\mathcal{O}}\left((B\sqrt{S_k} + S_k)\sqrt{T} \right) = \widetilde{\mathcal{O}}\left(n^k \sqrt{T} \right)$$

for fixed k under the EC-GP-UCB bound. Therefore, for local physical observables, LPQK can replace the exponential full-kernel dimension 4^n by the polynomial dimension $S_k = \mathcal{O}(n^k)$.

Example 2: exponentially decaying high-weight Pauli tail. More generally, suppose the total squared Pauli mass at weight w decays exponentially:

$$E_w := \sum_{\mathrm{wt}(\mathbf{P}) = w} c_{\mathbf{P}}^2 \le C_0 e^{-2\alpha w} \qquad \text{for some } C_0, \alpha > 0.$$

Then

$$\varepsilon_b \le \|\mathbf{O}_{>b}\|_{\mathrm{HS}} = \left(\sum_{w=b+1}^{n} E_w\right)^{1/2} \le \left(C_0 \sum_{w=b+1}^{\infty} e^{-2\alpha w}\right)^{1/2} = C_\alpha e^{-\alpha(b+1)}, \tag{20}$$

where $C_\alpha := \sqrt{C_0/(1-e^{-2\alpha})}$. Hence the misspecification error decreases exponentially in the retained subsystem size b, while the information gain increases through

$$\gamma_T(b) = \mathcal{O}(S_b \log T), \qquad S_b = \sum_{w=0}^{b} \binom{n}{w} 3^w.$$

Combining (19) and (20) gives

$$R_T(b) = \widetilde{\mathcal{O}}\left((B\sqrt{S_b} + S_b)\sqrt{T} + C_\alpha e^{-\alpha b} T\sqrt{S_b}\right).$$

Thus, an optimal subsystem size can be selected by

$$b^\star \in \operatorname*{arg\,min}_{0\le b\le n} \left\{(B\sqrt{S_b} + S_b)\sqrt{T} + C_\alpha e^{-\alpha b} T\sqrt{S_b}\right\}.$$

Since for small fixed b, $S_b = \mathcal{O}(n^b)$, balancing the two terms gives the heuristic scaling $b^\star = \widetilde{\mathcal{O}}\left(\frac{\log T}{2\alpha + \log n}\right)$, and eventually yielding

$$R_T = \widetilde{\mathcal{O}}\left(T^{\frac{1}{2} + \frac{\log n}{2\alpha + \log n}}\right).$$

Thus, faster Pauli-tail decay, i.e., larger α, permits a smaller subsystem size and regret closer to the realizable $\sqrt{T}$-type rate.

Overall, LPQK is more beneficial when the target observable has most of its Pauli mass concentrated on low-weight terms. In the exact k-local case, the misspecification error vanishes at $b = k$ and the exponential full-kernel dimension 4^n is replaced by the polynomial dimension $\mathcal{O}(n^k)$. For observables with decaying high-weight tails, the optimal b is determined by balancing the information-gain term S_b against the tail error ε_b.

Remark on Effective Dimensionality It is important to note that the dimensions used in proof of Theorem 2(ii), $D = 4^n$ and $d = \sum_{w=0}^{b} \binom{n}{w} 3^w$, are based on the assumption that the quantum feature map is sufficiently expressive to span the entire space of n-qubit Hermitian operators. While this is often the case for complex, generic circuits, it may not hold for all feature maps.

In many practical scenarios, a quantum circuit may have specific symmetries or a limited number of parameters, constraining the reachable states $\{\boldsymbol{\rho}(\mathbf{x})\}$ to a lower-dimensional subspace. In such cases, the actual dimension of the full space D would be less than 4^n, and the actual dimension of the low-weight subspace, d, would likewise be smaller than $\sum_{w=0}^{b} \binom{n}{w} 3^w$.

D FOURIER FORM OF QUANTUM KERNELS AND RFF

D.1 FOURIER FORM OF QUANTUM KERNELS

Theorem 4 (Fourier Representation of Quantum Kernels, Theorem 1 of Schuld [2021])**.** *Let $\mathcal{X} = \mathbb{R}^N$ and let*

$$\mathbf{S}(\mathbf{x}) \;=\; \mathbf{W}^{(N+1)}\, e^{-ix_N \mathbf{G}}\, \mathbf{W}^{(N)} \;\cdots\; \mathbf{W}^{(2)}\, e^{-ix_1 \mathbf{G}}\, \mathbf{W}^{(1)}$$

be a data-encoding circuit, where each $\mathbf{G}$ is a diagonal $(d \times d)$ Hermitian operator with eigenvalues $\{\lambda_1, \ldots, \lambda_d\}$, and $\mathbf{W}^{(1)}, \ldots, \mathbf{W}^{(N+1)}$ denote fixed unitary operations. Define the quantum state $|\boldsymbol{\phi}(\mathbf{x})\rangle = \mathbf{S}(\mathbf{x})\,|0\rangle$ and let $\kappa_Q(\mathbf{x}, \mathbf{x}')$ be the resulting quantum kernel, i.e.,

$$\kappa_Q(\mathbf{x}, \mathbf{x}') \;=\; \left|\langle \boldsymbol{\phi}(\mathbf{x}) | \boldsymbol{\phi}(\mathbf{x}')\rangle\right|^2.$$

Then there exists a finite index set $\Omega \subseteq \mathbb{R}^N$ (derived from the eigenvalues $\{\lambda_j\}$) and coefficients $\{c_{\mathbf{s},\mathbf{t}}\}_{\mathbf{s},\mathbf{t}\in\Omega}$ such that

$$\kappa_Q(\mathbf{x}, \mathbf{x}') \;=\; \sum_{\mathbf{s},\mathbf{t}\in\Omega} c_{\mathbf{s},\mathbf{t}} \exp\big(i\,\mathbf{s}\cdot\mathbf{x}\big)\, \exp\big(i\,\mathbf{t}\cdot\mathbf{x}'\big) \tag{21}$$

Proof. Assume without loss of generality that the encoding generator $\mathbf{G}$ is diagonal because any Hermitian operator can be diagonalized as $\mathbf{G} = \mathbf{V}\mathbf{\Sigma}\mathbf{V}^\dagger$, where $\mathbf{\Sigma}$ is diagonal. Consequently, an encoding gate $e^{-ix_i\mathbf{G}}$ can be written as

$$e^{-ix_i\mathbf{G}} \;=\; \mathbf{V}\left(e^{-ix_i\mathbf{\Sigma}}\right)\mathbf{V}^\dagger,$$

and thus we may "absorb" $\mathbf{V}$ and $\mathbf{V}^\dagger$ into the preceding and subsequent arbitrary unitaries in the circuit. Hence, we have

$$e^{-ix_i\mathbf{\Sigma}} \;=\; \begin{pmatrix} e^{-ix_i\lambda_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & e^{-ix_i\lambda_d} \end{pmatrix},$$

where $\{\lambda_1, \ldots, \lambda_d\}$ are the eigenvalues of $\mathbf{G}$.

The quantum kernel can be written as

$$\begin{aligned}
\kappa_Q(\mathbf{x}, \mathbf{x}') &= \left|\langle \boldsymbol{\phi}(\mathbf{x}) | \boldsymbol{\phi}(\mathbf{x}')\rangle\right|^2 \\
&= \left|\langle 0|\ \left(\mathbf{W}^{(1)}\right)^\dagger \left(e^{-\,i\,x_1'\,\mathbf{\Sigma}}\right)^\dagger \cdots \left(e^{-\,i\,x_N'\,\mathbf{\Sigma}}\right)^\dagger \left(\mathbf{W}^{(N+1)}\right)^\dagger \times \mathbf{W}^{(N+1)}\, e^{-\,i\,x_N\,\mathbf{\Sigma}} \cdots e^{-\,i\,x_1\,\mathbf{\Sigma}}\, \mathbf{W}^{(1)}\ |0\rangle\right|^2 \\
&= \left|\langle 0|\ \left(\mathbf{W}^{(1)}\right)^\dagger \left(e^{-\,i\,x_1'\,\mathbf{\Sigma}}\right)^\dagger \cdots \left(e^{-\,i\,x_N'\,\mathbf{\Sigma}}\right)^\dagger \times e^{-\,i\,x_N\,\mathbf{\Sigma}} \cdots e^{-\,i\,x_1\,\mathbf{\Sigma}}\, \mathbf{W}^{(1)}\ |0\rangle\right|^2 \\
&= \Big| \sum_{j_1,\ldots,j_N=1}^{d} \sum_{k_1,\ldots,k_N=1}^{d} \exp\Big(-i\big(\lambda_{j_1}x_1 - \lambda_{k_1}x_1' + \cdots + \lambda_{j_N}x_N - \lambda_{k_N}x_N'\big)\Big) \\
&\qquad \times \left(\mathbf{W}^{(1)}_{1\,k_1} \cdots \mathbf{W}^{(N)}_{k_{N-1}\,k_N}\right)^* \times \left(\mathbf{W}^{(N)}_{j_N\,j_{N-1}} \cdots \mathbf{W}^{(1)}_{j_1\,1}\right)\Big|^2 \\
&= \left| \sum_{\mathbf{j}} \sum_{\mathbf{k}} \exp\big(-\,i\,(\mathbf{\Lambda_j}\cdot\mathbf{x} - \mathbf{\Lambda_k}\cdot\mathbf{x}')\big)\,\left(w_\mathbf{k}\right)^*\left(w_\mathbf{j}\right)\right|^2 \\
&= \sum_{\mathbf{j}}\sum_{\mathbf{k}}\sum_{\mathbf{h}}\sum_{\mathbf{l}} \exp\big(-\,i\,(\mathbf{\Lambda_j} - \mathbf{\Lambda_l})\cdot\mathbf{x}\big)\,\exp\big(i\,(\mathbf{\Lambda_k} - \mathbf{\Lambda_h})\cdot\mathbf{x}'\big) \times \left(w_\mathbf{k}\,w_\mathbf{h}\right)^*\left(w_\mathbf{j}\,w_\mathbf{l}\right).
\end{aligned}$$

Here, the scalars $W^{(i)}_{ab}$ for $i = 1, \ldots, N$ refer to the element $\langle a|\,\mathbf{W}^{(i)}\,|b\rangle$ of the unitary operator $\mathbf{W}^{(i)}$. The bold multi-index $\mathbf{j}$ summarizes the set $(j_1, \ldots, j_N)$ where $j_i \in \{1, \ldots, d\}$, and $\mathbf{\Lambda_j}$ is a vector containing the eigenvalues selected by the multi-index (and similarly for $\mathbf{k}, \mathbf{h}, \mathbf{l}$).

We can now group together all terms where $\mathbf{\Lambda_j} - \mathbf{\Lambda_l} = \mathbf{s}$ and $\mathbf{\Lambda_k} - \mathbf{\Lambda_h} = \mathbf{t}$; in other words, where the differences of eigenvalues amount to the same vectors $\mathbf{s}, \mathbf{t}$. Then we have

$$\begin{aligned}
\kappa_Q(\mathbf{x}, \mathbf{x}') &= \sum_{\mathbf{s},\mathbf{t}\in\Omega} e^{-i\mathbf{s}\cdot\mathbf{x}}\, e^{i\mathbf{t}\cdot\mathbf{x}'} \left(\sum_{\substack{\mathbf{j},\mathbf{l}: \\ \mathbf{\Lambda_j}-\mathbf{\Lambda_l}=\mathbf{s}}} \sum_{\substack{\mathbf{k},\mathbf{h}: \\ \mathbf{\Lambda_k}-\mathbf{\Lambda_h}=\mathbf{t}}} w_\mathbf{j} w_\mathbf{l} \big(w_\mathbf{k} w_\mathbf{h}\big)^* \right) \\
&= \sum_{\mathbf{s},\mathbf{t}\in\Omega} e^{-i\mathbf{s}\cdot\mathbf{x}}\, e^{i\mathbf{t}\cdot\mathbf{x}'}\, c_{\mathbf{s},\mathbf{t}}
\end{aligned}$$

The frequency set Ω contains all vectors of the form $\{\mathbf{\Lambda_j} - \mathbf{\Lambda_k}\}$ with $\mathbf{\Lambda_j} = (\lambda_{j_1}, \ldots, \lambda_{j_N})$, $j_1, \ldots, j_N \in \{1, \ldots, d\}$. □

This result demonstrates that a *quantum* kernel can be interpreted as a sum of complex exponentials, which are determined by the spectra of the encoding generator $\mathbf{G}$ and the unitaries interleaved within the circuit. Although, in principle, this sum may encompass *exponentially many* terms—stemming from the combinatorial complexity of multi-qubit gates—the specific structure of $\mathbf{G}$ and the unitaries $\mathbf{W}$ often introduce patterns or symmetries. These patterns can be leveraged to simplify analysis. More importantly, once the circuit structure is fixed, the coefficients in this sum can be computed explicitly using a formula, as outlined earlier in the proof.

Remark (extension to re-uploading / repeated encodings). Theorem 4 is stated for an encoding sequence in which each scalar input x_j appears once. If a circuit re-uploads the same input $\mathbf{x} \in \mathbb{R}^d$ across multiple layers (so that some coordinates are used multiple times), the resulting kernel still admits a Fourier representation of the above form [Schuld et al., 2021].

Next, we will examine how this perspective facilitates the analysis of the feature map corresponding to quantum kernels.

D.2 RFF WITH PAULI ENCODING QUANTUM KERNEL

In this section, we consider the special case where the quantum kernel $\kappa_Q(\mathbf{x}, \mathbf{x}')$ satisfies two simplifying properties:

- The kernel is translation-invariant—i.e., it depends only on the difference $\mathbf{x} - \mathbf{x}'$.
- The frequencies Ω in the Fourier representation (cf. Equation 21) are integer-valued.

These assumptions hold in many common settings, e.g., when data encoding is implemented by Pauli rotations [Landman et al., 2022]. In this case, one may write

$$\kappa_Q(\mathbf{x}, \mathbf{x}') = \sum_{\boldsymbol{\omega} \in \Omega} c_{\boldsymbol{\omega}} \exp\big(i\, \boldsymbol{\omega}^\top (\mathbf{x} - \mathbf{x}')\big) = \sum_{\boldsymbol{\omega} \in \Omega} c_{\boldsymbol{\omega}} \cos\big(\boldsymbol{\omega}^\top (\mathbf{x} - \mathbf{x}')\big),$$

where $\Omega \subset \mathbb{Z}^d$ is a finite set of frequencies and $c_{\boldsymbol{\omega}} \geq 0$ are real Fourier coefficients.

Expressing κ in a Real Feature Space From standard trigonometric identities, each cosine term $\cos(\boldsymbol{\omega}^\top \mathbf{x} - \boldsymbol{\omega}^\top \mathbf{x}')$ can be represented as a dot product of some 2 dimensional real feature. Specifically,

$$\begin{aligned} C\cos(\alpha - \beta) &= C\cos\alpha\cos\beta + C\sin\alpha\sin\beta \\ &= \left\langle \begin{bmatrix} \sqrt{C}\cos\alpha \\ \sqrt{C}\sin\alpha \end{bmatrix}, \begin{bmatrix} \sqrt{C}\cos\beta \\ \sqrt{C}\sin\beta \end{bmatrix} \right\rangle \end{aligned}$$

For example, let $\Delta_i = x_i - x_i'$,

$$\begin{aligned} &\cos(\Delta_1 - \Delta_2 + \Delta_3) \\ &= \cos\left((x_1 - x_1') - (x_2 - x_2') + (x_3 - x_3')\right) \\ &= \cos\left((x_1 - x_2 + x_3) - (x_1' - x_2' + x_3')\right) \\ &= \left\langle \begin{bmatrix} \cos(x_1 - x_2 + x_3) \\ \sin(x_1 - x_2 + x_3) \end{bmatrix}, \begin{bmatrix} \cos(x_1' - x_2' + x_3') \\ \sin(x_1' - x_2' + x_3') \end{bmatrix} \right\rangle. \end{aligned}$$

Hence, if $\kappa_Q(\mathbf{x}, \mathbf{x}') = \sum_{\boldsymbol{\omega} \in \Omega} c_{\boldsymbol{\omega}} \cos\big(\boldsymbol{\omega}^\top (\mathbf{x} - \mathbf{x}')\big)$, then we can rewrite κ_Q as $\langle \boldsymbol{\phi}(\mathbf{x}), \boldsymbol{\phi}(\mathbf{x}') \rangle$ for some (potentially high-dimensional) $\boldsymbol{\phi}(\mathbf{x})$ built from the features. Consequently, for a translation-invariant quantum kernel κ_Q with a finite Fourier decomposition, κ_Q corresponds to the real feature set

$$\Big\{ \sqrt{c_{\boldsymbol{\omega}}} \cos\big(\boldsymbol{\omega} \cdot \mathbf{x}\big), \sqrt{c_{\boldsymbol{\omega}}} \sin\big(\boldsymbol{\omega} \cdot \mathbf{x}\big) \Big\}_{\boldsymbol{\omega} \in \Omega}.$$

Approximation via Random Fourier Features Following the approach in [Landman et al., 2022], one can sample D frequencies $(\boldsymbol{\omega}_1, ..., \boldsymbol{\omega}_D)$ uniformly from Ω, and then construct the approximate kernel

$$\kappa_{\text{RFF}}(\mathbf{x}, \mathbf{x}') := \boldsymbol{\phi}_{\text{RFF}}(\mathbf{x})^T \boldsymbol{\phi}_{\text{RFF}}(\mathbf{x}'),$$

where

$$\boldsymbol{\phi}_{\text{RFF}}(\mathbf{x}) \;=\; \sqrt{\tfrac{1}{D}}\, \left[\cos\big(\boldsymbol{\omega}_1 \cdot \mathbf{x}\big), \ldots, \sin\big(\boldsymbol{\omega}_D \cdot \mathbf{x}\big)\right]^\top.$$

Uniform approximation of the target function via random bases. To control the misspecification error induced by replacing the full feature $\boldsymbol{\phi}_Q$ with a random low-dimensional feature map, we use the following uniform approximation theorem of Rahimi and Recht [2008].

Theorem 5 (Uniform approximation with random bases [Rahimi and Recht, 2008, Theorem 3.2]). *Let $\mathcal{X} \subset \mathbb{R}^d$ be compact and let $\varphi(\mathbf{x};\boldsymbol{\theta}) = \varphi(\boldsymbol{\theta}^\top \mathbf{x})$, where $\varphi : \mathbb{R} \to \mathbb{R}$ is L-Lipschitz, satisfies $\varphi(0) = 0$, and $|\varphi| \leq 1$. Let p be a distribution over $\boldsymbol{\theta}$ with finite second moment. Define the mixture class*

$$\mathcal{F}(\mathcal{X}, \Theta, \varphi, p) = \left\{ f(\mathbf{x}) = \int_\Theta \alpha(\boldsymbol{\theta})\varphi(\mathbf{x};\boldsymbol{\theta})\, d\boldsymbol{\theta} \ : \ \|f\|_p := \sup_{\boldsymbol{\theta}} \left| \frac{\alpha(\boldsymbol{\theta})}{p(\boldsymbol{\theta})} \right| < \infty \right\}.$$

Fix any $f \in \mathcal{F}$ and let $B_\mathcal{X} = \sup_{\mathbf{x}\in\mathcal{X}} \|\mathbf{x}\|_2$. Draw $\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_D \overset{iid}{\sim} p$. Then for any $\delta > 0$, with probability at least $1 - \delta$, there exist coefficients $c_1, \ldots, c_D$ such that

$$\hat{f}(\mathbf{x}) = \sum_{k=1}^{D} c_k \varphi(\boldsymbol{\theta}_k^\top \mathbf{x})$$

satisfies the uniform bound

$$\|\hat{f} - f\|_\infty < \frac{\|f\|_p}{\sqrt{D}} \left(\sqrt{\log \frac{1}{\delta}} + 4LB_\mathcal{X} \sqrt{\mathbb{E}_{\boldsymbol{\theta}\sim p} \|\boldsymbol{\theta}\|_2^2} \right).$$

This theorem yields a *uniform* $O(1/\sqrt{D})$ approximation rate on the entire compact domain $\mathcal{X}$. In [Rahimi and Recht, 2008], the authors discuss that Fourier features of the form $\varphi(\mathbf{x};\boldsymbol{\theta}) = \cos(\boldsymbol{\omega}^\top \mathbf{x} + b)$ satisfy the conditions of Theorem 5.

D.3 RFF WITH GENERAL ENCODING QUANTUM KERNEL

When one or more encoding Hamiltonians in a variational quantum circuit (VQC) are not trivially diagonalizable, directly identifying and sampling from their associated frequency set can be challenging. To address this, Landman et al. [2022] propose a more flexible grid-based sampling approach to approximate the quantum kernel. The central idea involves sampling from a grid of frequencies evenly spaced between zero and an upper bound, $\boldsymbol{\omega}_{\max}$, along each dimension. Here, $\boldsymbol{\omega}_{\max}$ is chosen as the maximum possible frequency that can appear in the VQC—such as the one inferred via the Shannon criterion or derived from known bounds on the largest eigenvalues of the encoding Hamiltonians.

Specifically, suppose each dimension of $\mathbf{x}$ could contain frequencies in $[0, \boldsymbol{\omega}_{\max}]$. We partition this interval into regularly spaced bins of size $s > 0$, giving $\lceil \boldsymbol{\omega}_{\max}/s \rceil^d$ possible frequency values per dimension. Over d dimensions, this yields a regular grid of size $[0, \boldsymbol{\omega}_{\max}]^d$. One then sample D frequencies $(\boldsymbol{\omega}_1, \ldots, \boldsymbol{\omega}_D)$ from the grid, and similarly construct the approximate kernel $k_{\text{RFF}}(\mathbf{x}, \mathbf{x}') = \boldsymbol{\phi}_{\text{RFF}}(\mathbf{x})^\top \boldsymbol{\phi}_{\text{RFF}}(\mathbf{x}')$ where, as before,

$$\boldsymbol{\phi}_{\text{RFF}}(\mathbf{x}) \ = \ \sqrt{\tfrac{1}{D}} \ \left[\cos\big(\boldsymbol{\omega}_1 \cdot \mathbf{x}\big) \quad \ldots \quad \cos\big(\boldsymbol{\omega}_D \cdot \mathbf{x}\big) \right]^\top.$$

Then we have [Landman et al., 2022]

Theorem 6 (RFF for Grid Sampling). *Let $\mathcal{X}$ be a compact subset of $\mathbb{R}^d$, and $\varepsilon > 0$. Consider a training set $\{(\mathbf{x}_i, y_i)\}_{i=1}^M$. Let f^* be a VQC model with any Hamiltonian encoding, with a maximum individual frequency $\omega_{\max}$ and full freedom on the associated frequency coefficients, trained with a regularization λ. Let $\sigma_y^2 = \frac{1}{M}\sum_{i=1}^M y_i^2$ and $|\mathcal{X}|$ be the diameter of $\mathcal{X}$. Let $\tilde{f}$ be the RFF model with D samples in the grid strategy, trained on the same dataset and the same regularization. Let $C = \|f^*\|_\infty |\mathcal{X}|$ and s be the sampling rate as in the grid strategy. Then, for $0 < s < \frac{1}{C}$, we can guarantee $\sup_{\mathbf{x}\in\mathcal{X}} |f^*(\mathbf{x}) - \tilde{f}(\mathbf{x})| \leq \varepsilon$ with probability $1 - \delta$, provided the number D of samples satisfies*

$$D = \Omega\Bigg(\frac{dC_1(1+\lambda)}{\lambda^4(\varepsilon - sC)^2} \Big[\log\big(\boldsymbol{\omega}_{\max} |\mathcal{X}|\big) + \log\Big(\frac{C_2(1+\lambda)}{\lambda^2(\varepsilon - sC)}\Big) - \log\delta \Big] \Bigg),$$

where C_1 and C_2 are constants depending on σ_y and $\mathcal{X}$. Note that f^* can be equivalently defined to be the Kernel Ridge Regression (KRR) predictor derived from the full quantum kernel κ, and this theorem can be used similarly as Theorem 9

D.4 A DIRECT TRUNCATED KERNEL APPROXIMATION APPROACH

Recall that for a translation invariant quantum kernel, we can express its feature map as

$$\Big\{ \sqrt{c_\mathbf{s}} \cos\big(\mathbf{s} \cdot \mathbf{x}\big), \sqrt{c_\mathbf{s}} \sin\big(\mathbf{s} \cdot \mathbf{x}\big) \Big\}_{\mathbf{s}\in\Omega}.$$

Truncated Features for Kernel Approximation To construct a smaller feature map (or kernel) from κ_Q, a natural idea is to pick only the largest-magnitude coefficients $\{c_{\mathbf{s}}\}$ and retain those frequencies $\mathbf{s}$ in the approximation while discarding others. Concretely, let $S \subset \Omega$ be a subset of frequencies corresponding to the largest values of $\{c_{\mathbf{s}}\}$ and define

$$\kappa_{\text{small}}(\mathbf{x}, \mathbf{x}') = \sum_{\mathbf{s} \in S \subseteq \Omega} c_{\mathbf{s}} \Big[\cos\big(\mathbf{s} \cdot \mathbf{x}\big) \cos\big(\mathbf{s} \cdot \mathbf{x}'\big) + \sin\big(\mathbf{s} \cdot \mathbf{x}\big) \sin\big(\mathbf{s} \cdot \mathbf{x}'\big) \Big].$$

This kernel κ_{small} is essentially a truncated version of κ_Q: it retains only frequencies in S. We now show that a function f drawn from $\mathrm{GP}(0, \kappa_Q)$ can be uniformly approximated by a function in the RKHS of κ_{small}, with an error bounded by the sum of the left-out coefficients.

Theorem 7 (Approximation by Truncating Fourier Series)**.** *Suppose a quantum kernel admits a Fourier series*

$$\kappa_Q(\mathbf{x}, \mathbf{x}') = \sum_{\mathbf{s} \in \Omega} c_{\mathbf{s}} \, \cos\big(\mathbf{s} \cdot (\mathbf{x} - \mathbf{x}')\big),$$

and let f^* be a function sampled from $\mathrm{GP}(0, \kappa_Q)$. Consider the truncated kernel

$$\kappa_{\textit{small}}(\mathbf{x}, \mathbf{x}') = \sum_{\mathbf{s} \in S \subset \Omega} c_{\mathbf{s}} \Big[\cos\big(\mathbf{s} \cdot \mathbf{x}\big) \cos\big(\mathbf{s} \cdot \mathbf{x}'\big) + \sin\big(\mathbf{s} \cdot \mathbf{x}\big) \sin\big(\mathbf{s} \cdot \mathbf{x}'\big) \Big],$$

with the associated RKHS $\mathcal{H}_{\kappa_{\textit{small}}}(\mathcal{X})$. Then, the expected uniform approximation error satisfies

$$\mathbb{E}\Big[\min_{f \in \mathcal{H}_{\kappa_{\textit{small}}}(\mathcal{X})} \|f - f^*\|_\infty \Big] \leq \sum_{\mathbf{s} \in \Omega \setminus S} 2\sqrt{c_{\mathbf{s}}}.$$

Proof. We outline the proof in the following three steps.

1. Sampling f from $\mathrm{GP}(0, \kappa_Q)$:

By standard Gaussian process theory, a function f drawn from $\mathrm{GP}(0, \kappa_Q)$ admits a representation in terms of feature basis corresponding to κ_Q. Concretely, one may write

$$f(\mathbf{x}) = \sum_{\boldsymbol{\omega} \in \Omega} \alpha_{\boldsymbol{\omega}} \, \psi_{\boldsymbol{\omega}}(\mathbf{x}),$$

where $\alpha_{\boldsymbol{\omega}} \sim \mathcal{N}(0, 1)$, and $\{\psi_{\boldsymbol{\omega}}\}$ are orthonormal basis functions scaled by $\sqrt{c_{\boldsymbol{\omega}}}$. In our case, we can take $\psi_{\boldsymbol{\omega}} = (\sqrt{c_{\boldsymbol{\omega}}} \cos(\boldsymbol{\omega} \cdot \mathbf{x}) + \sqrt{c_{\boldsymbol{\omega}}} \sin(\boldsymbol{\omega} \cdot \mathbf{x}))$. For notational convenience, we may redefine $\alpha_{\boldsymbol{\omega}} \leftarrow \sqrt{c_{\boldsymbol{\omega}}} \alpha_{\boldsymbol{\omega}}$ so that $\psi_{\boldsymbol{\omega}}$ is just $(\cos(\boldsymbol{\omega} \cdot \mathbf{x}) + \sin(\boldsymbol{\omega} \cdot \mathbf{x}))$.

2. Restricting to a Smaller Kernel k_{small}:

Suppose we now try to approximate f^* by a function g lying in the RKHS associated with the truncated kernel κ_{small}. This space corresponds to a smaller set of frequencies $S \subset \Omega$. Consequently, any $g(\mathbf{x}) \in \mathcal{H}_{\kappa_{\text{small}}}(\mathcal{X})$ can be written as

$$g(\mathbf{x}) = \sum_{\boldsymbol{\omega} \in S} \beta_{\boldsymbol{\omega}} \psi_{\boldsymbol{\omega}}(\mathbf{x}),$$

where $\{\beta_{\boldsymbol{\omega}}\}$ are real coefficients determined by the regression procedure between f^* and g.

Since the functions $\{\psi_{\boldsymbol{\omega}}\}$ are orthogonal on $[0, 2\pi]^d$ (i.e., $\int_{[0,2\pi]^d} \psi_{\boldsymbol{\omega}}(\mathbf{x}) \psi_{\boldsymbol{\omega}'}(\mathbf{x}) \, d\mathbf{x} = 0$ for all $\boldsymbol{\omega} \neq \boldsymbol{\omega}'$), the best fit to f^* in the mean-square sense is achieved by matching coefficients on those frequencies in S. In other words, for each $\boldsymbol{\omega} \in S \cap \Omega$, the optimal $\beta_{\boldsymbol{\omega}}$ is $\alpha_{\boldsymbol{\omega}}$; frequencies outside $S \cap \Omega$ cannot be matched, leaving those $\beta_{\boldsymbol{\omega}} = 0$.

This can be seen by expanding the objective

$$\begin{aligned} &\int_{[0,2\pi]^d} \big[f(\mathbf{x}) - g(\mathbf{x})\big]^2 \, d\mathbf{x} \\ &= \int_{[0,2\pi]^d} f(\mathbf{x})^2 - 2 f(\mathbf{x}) g(\mathbf{x}) + g(\mathbf{x})^2 \, d\mathbf{x} \end{aligned}$$

and noting that in

$$\int_{[0,2\pi]^d} f(\mathbf{x})g(\mathbf{x})\,d\mathbf{x}$$
$$= \int_{[0,2\pi]^N} \sum_{\boldsymbol{\omega}} \alpha_{\boldsymbol{\omega}}\beta_{\boldsymbol{\omega}}\psi_{\boldsymbol{\omega}}(\mathbf{x}) + \sum_{\mathbf{s},\mathbf{t},\mathbf{s}\neq\mathbf{t}} \alpha_{\mathbf{s}}\beta_{\mathbf{t}}\psi_{\mathbf{s}}(\mathbf{x})\psi_{\mathbf{t}}(\mathbf{x})\,d\mathbf{x},$$

the second summand is zero by the orthogonality of $\{\psi_{\boldsymbol{\omega}}\}$. Thus, the solution matching $\beta_{\boldsymbol{\omega}} = \alpha_{\boldsymbol{\omega}}$ for $\boldsymbol{\omega} \in S \cap \Omega$ and $\beta_{\boldsymbol{\omega}} = 0$ otherwise minimizes the integral, and hence yields the optimal least-squares approximation.

3. Uniform Approximation Error:

Finally, consider the uniform approximation error, $\mathbb{E}\Big[\min_{f\,\in\,\mathcal{H}_{\kappa_{\text{small}}}(\mathcal{X})} \|f - f^*\|_\infty\Big]$. Since the optimal $f \in \mathcal{H}_{\kappa_{\text{small}}}(\mathcal{X})$ contains coefficients that match the corresponding ones in f^*, we obtain

$$\min_{f\in\mathcal{H}_{\kappa_{\text{small}}}(\mathcal{X})} \|f - f^*\|_\infty$$
$$= \left| \sum_{\mathbf{s}\in\Omega\setminus S} \sqrt{c_{\mathbf{s}}}\,(\cos(\mathbf{s}\cdot\mathbf{x}) + \sin(\mathbf{s}\cdot\mathbf{x})) \right|$$
$$\leq \sum_{\mathbf{s}\in\Omega\setminus S} 2\sqrt{c_{\mathbf{s}}},$$

which is the statement to be proved. □

E PROOF OF THEOREM 3

E.1 INTRODUCTION TO SQUARECB

Algorithm 1 SquareCB

Parameters: Learning rate $\gamma > 0$, exploration parameter $\mu > 0$, online regression oracle SqAlg.

for $t = 1, \ldots, T$ **do**
  Receive context $\mathbf{x}_t$.
  For each action $\mathbf{a} \in \mathcal{A}$, compute $\hat{y}_{t,\mathbf{a}} := \hat{y}_t(\mathbf{x}_t, \mathbf{a})$
  Let $\mathbf{b}_t = \arg\min_{\mathbf{a}\in\mathcal{A}} \hat{y}_{t,\mathbf{a}}$
  For each $\mathbf{a} \neq \mathbf{b}_t$, define $p_{t,\mathbf{a}} = \frac{1}{\mu+\gamma(\hat{y}_{t,\mathbf{a}} - \hat{y}_{t,\mathbf{b}_t})}$ and let $p_{t,\mathbf{b}_t} = 1 - \sum_{\mathbf{a}\neq\mathbf{b}_t} p_{t,\mathbf{a}}$
  Sample action $\mathbf{a}_t \sim p_t$ and observe loss $\ell_t(\mathbf{a}_t)$.
  Update SqAlg with example $\big((\mathbf{x}_t, \mathbf{a}_t), \ell_t(\mathbf{a}_t)\big)$.
end for

Foster and Rakhlin [2020] considered the following contextual bandit setting: Over T rounds:

- Nature or an adversary picks a context $\mathbf{x}_t \in \mathcal{X}$ and a loss function $\ell_t(\cdot)\colon \mathcal{A} \to [0,1]$.
- Learner observes $\mathbf{x}_t$, selects $\mathbf{a}_t \in \mathcal{A}$.
- Learner incurs and observes loss $\ell_t(\mathbf{a}_t)$.

Note that we are considering the linear bandit setting in this work, which is simply a special case of their setting (with constant $\mathbf{x}_t$).

The learner has access to a function class $\mathcal{F} \subseteq \{f : \mathcal{X} \times \mathcal{A} \to [0,1]\}$. Under *realizability*, there is an $f \in \mathcal{F}$ so that $\mathbb{E}[\ell_t(\mathbf{a}) \mid \mathbf{x}_t] = f(\mathbf{x}_t, \mathbf{a})$. In a *misspecified* setting, one relaxes this equality to

$$\big|\mathbb{E}[\ell_t(\mathbf{a}) \mid \mathbf{x}_t] - f(\mathbf{x}_t, \mathbf{a})\big| \leq \varepsilon, \quad \forall t, \mathbf{a}$$

for some $\varepsilon > 0$. Note this agrees with our definition for misspecified setting in the main text, namely that there exists $f \in \mathcal{F}$ such that $|f(\mathbf{x}_t, \mathbf{a}) - f^*(\mathbf{x}_t, \mathbf{a})| \le \varepsilon$. Here $f^*(\mathbf{x}_t, \mathbf{a}) = \mathbb{E}[\ell_t(\mathbf{a}) \mid \mathbf{x}_t]$.

Foster and Rakhlin [2020] introduced SquareCB, a contextual linear bandit algorithm. See Algorithm 1 for details.

SquareCB uses an online algorithm called SqAlg that provides predictions $\hat{y}_t(\mathbf{x}_t, \mathbf{a}) \in [0, 1]$, and receives the true label (loss) $y_t \in [0, 1]$. Over T rounds, SqAlg aims to do as well as the best function $f \in \mathcal{F}$ in terms of the square loss, i.e., it satisfies a bound of the form

$$\sum_{t=1}^{T} \big(\widehat{y}_t - y_t\big)^2 - \inf_{f \in \mathcal{F}} \sum_{t=1}^{T} \big(f(\mathbf{x}_t, \mathbf{a}_t) - y_t\big)^2 \le \mathrm{Reg}_{\mathrm{Sq}}(T),$$

where $\mathrm{Reg}_{\mathrm{Sq}}(T)$ is typically called the square-loss regret of the online forecaster. If $\mathcal{F}$ is a simple class (e.g., linear in our case), there are known efficient algorithms that achieve small $\mathrm{Reg}_{\mathrm{Sq}}(T)$.

In the misspecified setting, Foster and Rakhlin [2020] showed the following:

Theorem 8 (Pseudoregret Bound for SquareCB)**.** *Under the misspecified setting, SquareCB with $\mu = K := |\mathcal{A}|$ and $\gamma = \sqrt{8KT/(\mathrm{Reg}_{\mathrm{Sq}}(T) + \varepsilon^2 T)}$ ensures that*

$$\sup_{\pi} \mathbb{E}\left[\sum_{t=1}^{T} \ell_t(\mathbf{a}_t) - \sum_{t=1}^{T} \ell_t\big(\pi(\mathbf{x}_t)\big)\right] \le 2\sqrt{KT\,\mathrm{Reg}_{\mathrm{Sq}}(T)} + 5\varepsilon\sqrt{K}T,$$

where $\mathbf{a}_t$ is the action chosen with SquareCB at time t, and $\sup_\pi$ ranges over all policies $\pi : \mathcal{X} \to \mathcal{A}$. Then, by instantiating SquareCB with the Vovk-Azoury-Warmuth forecaster, which has $\mathrm{Reg}_{\mathrm{Sq}}(T) \le D\log(1 + T/D)$, one gets an efficient algorithm with regret

$$R_T = \mathbb{E}\left[\sum_{t=1}^{T} \ell_t(\mathbf{a}_t) - \sum_{t=1}^{T} \ell_t\big(\pi^*(\mathbf{x}_t)\big)\right] \le 2\sqrt{KTD\log(1 + T/D)} + 5\varepsilon\sqrt{K}T.$$

with π^* being the optimal policy.

E.2 CHOOSING OPTIMAL D AND THE PROOF OF THEOREM 3

We now apply the SquareCB algorithm, combined with the Distinct-Sampling RFF technique, to the linear bandit setting where, at each round t, the following hold:

- The context $\mathbf{x}_t$ is now constant.
- The reward (or equivalently, the loss function) is linear in the quantum feature map $\boldsymbol{\rho}(\mathbf{a})$. Specifically, $\mathbb{E}[\ell_t(\mathbf{a})|\mathbf{x}_t] = \mathrm{Tr}[\mathbf{H}\boldsymbol{\rho}(\mathbf{a})]$ for some observable $\mathbf{H}$ and known quantum feature map $\boldsymbol{\rho}(\mathbf{a}) \in \mathbb{C}^{2^n \times 2^n}$. In other words, $f^*(\mathbf{a}) := \mathbb{E}[\ell_t(\mathbf{a})|\mathbf{x}_t]$ belongs to the RKHS induced by the quantum kernel κ_Q.

By construction, f^* belongs to the RKHS induced by the quantum kernel κ_Q.

We now instantiate Theorem 5 for the Pauli-encoding kernel κ_Q above. Define the parameter set

$$\Theta := \Omega \times \{c, s\}, \qquad p(\boldsymbol{\omega}, u) := \frac{1}{2|\Omega|} \quad \text{(uniform over } \Theta),$$

and define the real basis functions

$$\varphi(\mathbf{x}; (\boldsymbol{\omega}, c)) := \frac{1}{2}\cos(\boldsymbol{\omega}^\top \mathbf{x}) - \frac{1}{2}, \qquad \varphi(\mathbf{x}; (\boldsymbol{\omega}, s)) := \sin(\boldsymbol{\omega}^\top \mathbf{x}).$$

Note that any function $f \in \mathcal{H}_{\kappa_Q}$ can be written as a finite mixture of these bases:

$$f(\mathbf{x}) = c_0 + \sum_{\boldsymbol{\omega} \in \Omega} \Big(a_{\boldsymbol{\omega}} \cos(\boldsymbol{\omega}^\top \mathbf{x}) + b_{\boldsymbol{\omega}} \sin(\boldsymbol{\omega}^\top \mathbf{x})\Big),$$

which corresponds to a discrete mixture representation in the class $\mathcal{F}(\mathcal{X}, \Theta, \varphi, p)$. The scaling by $\frac{1}{2}$ in $\varphi(\mathbf{x}; (\boldsymbol{\omega}, c))$ is merely to match the requirement for Theorem 5 and does not affect the conclusion since it just shifts the result by a fixed constant.

We obtain the following corollary, which we will use as the misspecification guarantee for SquareCB.

Theorem 9 (Uniform misspecification bound of RFF approximation of Pauli-induced quantum f^*)**.** *Let κ_Q be a shift-invariant Pauli-encoding quantum kernel on a compact domain $\mathcal{X} \subset \mathbb{R}^d$ with finite spectrum $\Omega \subset \mathbb{Z}^d$, and let $f^* \in \mathcal{H}_{\kappa_Q}$. Fix $\delta \in (0,1)$ and draw $(\boldsymbol{\omega}_1, u_1), \ldots, (\boldsymbol{\omega}_D, u_D) \overset{iid}{\sim} p$ uniformly from $\Theta = \Omega \times \{c, s\}$. Define the random feature map*

$$\boldsymbol{\phi}_{\mathrm{RFF}}(\mathbf{x}) = \frac{1}{\sqrt{D}}\big[1, \varphi(\mathbf{x}; (\boldsymbol{\omega}_1, u_1)), \ldots, \varphi(\mathbf{x}; (\boldsymbol{\omega}_D, u_D))\big]^\top \in \mathbb{R}^{D+1}.$$

Then with probability at least $1 - \delta$, there exists a linear predictor $\tilde{f}(\mathbf{x}) = \langle \mathbf{w}, \boldsymbol{\phi}_{\mathrm{RFF}}(\mathbf{x}) \rangle$ such that

$$\varepsilon_D := \sup_{\mathbf{x} \in \mathcal{X}} |f^*(\mathbf{x}) - \tilde{f}(\mathbf{x})| < \frac{\|f^*\|_p}{\sqrt{D}} \left(\sqrt{\log \frac{1}{\delta}} + 4B_{\mathcal{X}}\, \omega_{\max} \right), \tag{22}$$

where $B_{\mathcal{X}} = \sup_{\mathbf{x} \in \mathcal{X}} \|\mathbf{x}\|_2$ and $\omega_{\max} := \max_{\boldsymbol{\omega} \in \Omega} \|\boldsymbol{\omega}\|_2$.

Note that since Ω is finite in any fixed-n qubit system, we have assumed $\|f^*\|_{\mathcal{H}_{\kappa_Q}}$ is bounded by B and p is uniform, we have that $\omega_{\max}$ and $\|f^*\|_p$ are all finite constants.

By Foster and Rakhlin [2020], under the misspecified setting, SquareCB with $\mu = K := |\mathcal{A}|$ and $\gamma = \sqrt{8KT/(\mathrm{Reg}_{\mathrm{Sq}}(T) + \varepsilon^2 T)}$ guarantees the expected pseudoregret bound

$$\sup_{\pi} \mathbb{E}\left[\sum_{t=1}^{T} \ell_t(\mathbf{a}_t) - \sum_{t=1}^{T} \ell_t(\pi(\mathbf{x}_t)) \right] \leq 2\sqrt{KT\, \mathrm{Reg}_{\mathrm{Sq}}(T)} + 5\varepsilon\sqrt{K}T,$$

where $\mathrm{Reg}_{\mathrm{Sq}}(T)$ is the square-loss regret of the regression oracle.

We instantiate the regression oracle by the Vovk–Azoury–Warmuth forecaster. For a d-dimensional linear class, this gives $\mathrm{Reg}_{\mathrm{Sq}}(T) \leq d \log(1 + T/d)$, hence

$$\mathbb{E}[R_T] \leq 2\sqrt{KT \cdot d \log(1 + T/d)} + 5\varepsilon\sqrt{K}T \tag{23}$$

Case 1: Exact quantum model. Using the exact quantum feature map with ambient dimension $d = 4^n$ yields a realizable model (so $\varepsilon = 0$), and therefore

$$\mathbb{E}[R_T] \leq 2\sqrt{KT \cdot 4^n \log\Big(1 + \tfrac{T}{4^n}\Big)}.$$

Case 2: RFF approximate model. Using a D-dimensional RFF surrogate gives a misspecification level $\varepsilon = \varepsilon_D$. Set $\delta = 1/T$ in 22, then with probability at least $1 - 1/T$ (over the sampled Fourier features),

$$\varepsilon_D < \frac{\|f^*\|_p}{\sqrt{D}} \left(\sqrt{\log T} + 4B_{\mathcal{X}}\, \omega_{\max} \right).$$

On the (probability $\geq 1 - 1/T$) event above, we substitute the explicit bound on ε_D as in 22; on the complement event we use the trivial bound $R_T \leq T$. Taking total expectation gives

$$\mathbb{E}[R_T] \leq 2\sqrt{KT \cdot D \log\Big(1 + \tfrac{T}{D}\Big)} + 5\sqrt{K}T\, \frac{\|f^*\|_p}{\sqrt{D}} \left(\sqrt{\log T} + 4B_{\mathcal{X}}\, \omega_{\max} \right) + 1.$$

Choosing $D = \lceil \sqrt{T} \rceil$ yields

$$\mathbb{E}[R_T] \leq 2\sqrt{K}T^{3/4}\sqrt{\log T} + 5\sqrt{K}T^{3/4}\|f^*\|_p \Big(\sqrt{\log T} + 4B_{\mathcal{X}}\, \omega_{\max} \Big) + 1 = \tilde{\mathcal{O}}\big(\sqrt{K}T^{3/4}\big)$$

Finally, taking the minimum of the exact-model bound and the RFF-surrogate bound yields the result in Theorem 3.

E.3 BOUND ON $\|f^*\|_p$

Assume the quantum kernel is shift-invariant with a finite Fourier expansion

$$\kappa_Q(\mathbf{x}, \mathbf{x}') \;=\; \sum_{\boldsymbol{\omega}\in\Omega} c_{\boldsymbol{\omega}} \cos\big(\boldsymbol{\omega}^\top(\mathbf{x}-\mathbf{x}')\big).$$

Let $f^* \in \mathcal{H}_{\kappa_Q}$ with $\|f^*\|_{\mathcal{H}_{\kappa_Q}} \le B$. Then we show that

$$\|f^*\|_p \;\le\; 4|\Omega|B.$$

This, together with the derivation in Appendix E.2, yields the result in Theorem 3.

Proof. Recall that if we defined the feature map

$$\boldsymbol{\phi}(\mathbf{x}) := \big(\sqrt{c_{\boldsymbol{\omega}}}\cos(\boldsymbol{\omega}^\top\mathbf{x}),\ \sqrt{c_{\boldsymbol{\omega}}}\sin(\boldsymbol{\omega}^\top\mathbf{x})\big)_{\boldsymbol{\omega}\in\Omega} \in \mathbb{R}^{2|\Omega|},$$

then $\kappa_Q(\mathbf{x}, \mathbf{x}') = \langle\boldsymbol{\phi}(\mathbf{x}), \boldsymbol{\phi}(\mathbf{x}')\rangle$, and every $f^* \in \mathcal{H}_{\kappa_Q}$ can be written as $f^*(\mathbf{x}) = \langle\mathbf{w}, \boldsymbol{\phi}(\mathbf{x})\rangle$ for some $\mathbf{w} \in \mathbb{R}^{2|\Omega|}$ with $\|f^*\|_{\mathcal{H}_{\kappa_Q}} = \|\mathbf{w}\|_2$. Write $\mathbf{w} = (u_{\boldsymbol{\omega}}, v_{\boldsymbol{\omega}})_{\boldsymbol{\omega}\in\Omega}$ such that

$$f^*(\mathbf{x}) = \sum_{\boldsymbol{\omega}\in\Omega}\Big(u_{\boldsymbol{\omega}}\sqrt{c_{\boldsymbol{\omega}}}\cos(\boldsymbol{\omega}^\top\mathbf{x}) + v_{\boldsymbol{\omega}}\sqrt{c_{\boldsymbol{\omega}}}\sin(\boldsymbol{\omega}^\top\mathbf{x})\Big) = \sum_{\boldsymbol{\omega}\in\Omega}\Big(a_{\boldsymbol{\omega}}\cos(\boldsymbol{\omega}^\top\mathbf{x}) + b_{\boldsymbol{\omega}}\sin(\boldsymbol{\omega}^\top\mathbf{x})\Big),$$

where $a_{\boldsymbol{\omega}} := u_{\boldsymbol{\omega}}\sqrt{c_{\boldsymbol{\omega}}}$ and $b_{\boldsymbol{\omega}} := v_{\boldsymbol{\omega}}\sqrt{c_{\boldsymbol{\omega}}}$. Since $\kappa_Q(\mathbf{x}, \mathbf{x}) = \sum_{\boldsymbol{\omega}\in\Omega} c_{\boldsymbol{\omega}} \le 1$, we have $c_{\boldsymbol{\omega}} \le 1$ for all $\boldsymbol{\omega}$, hence

$$|a_{\boldsymbol{\omega}}| \le |u_{\boldsymbol{\omega}}|\sqrt{c_{\boldsymbol{\omega}}} \le \|\mathbf{w}\|_2\sqrt{c_{\boldsymbol{\omega}}} \le \|\mathbf{w}\|_2 \le B, \quad \text{similarly,} \quad |b_{\boldsymbol{\omega}}| \le B.$$

Now note that $\varphi(0; (\boldsymbol{\omega}, c)) = 0$, so the centered function $g(\mathbf{x}) := f^*(\mathbf{x}) - f^*(0)$ admits the exact mixture representation

$$g(\mathbf{x}) = \sum_{\boldsymbol{\omega}\in\Omega} a_{\boldsymbol{\omega}}\big(\cos(\boldsymbol{\omega}^\top\mathbf{x}) - 1\big) + \sum_{\boldsymbol{\omega}\in\Omega} b_{\boldsymbol{\omega}}\sin(\boldsymbol{\omega}^\top\mathbf{x}) = \sum_{\boldsymbol{\omega}\in\Omega}\big(2a_{\boldsymbol{\omega}}\big)\,\varphi(\mathbf{x}; (\boldsymbol{\omega}, c)) + \sum_{\boldsymbol{\omega}\in\Omega} b_{\boldsymbol{\omega}}\,\varphi(\mathbf{x}; (\boldsymbol{\omega}, s)).$$

Thus we may take $\alpha(\boldsymbol{\omega}, c) = 2a_{\boldsymbol{\omega}}$ and $\alpha(\boldsymbol{\omega}, s) = b_{\boldsymbol{\omega}}$. Also recall that we sampled the Random Features uniformly with $p(\boldsymbol{\omega}, u) = \frac{1}{2|\Omega|}$, hence

$$\|g\|_p = \sup_{(\boldsymbol{\omega},u)\in\Theta}\left|\frac{\alpha(\boldsymbol{\omega}, u)}{p(\boldsymbol{\omega}, u)}\right| = 2|\Omega|\cdot\max\Big\{\max_{\boldsymbol{\omega}}|2a_{\boldsymbol{\omega}}|, \max_{\boldsymbol{\omega}}|b_{\boldsymbol{\omega}}|\Big\} \le 2|\Omega|\cdot\max\{2B, B\} = 4|\Omega|B.$$

□

E.4 BOUND ON $\omega_{\max}$

Recall $\omega_{\max} := \max_{\boldsymbol{\omega}\in\Omega}\|\boldsymbol{\omega}\|_2$, where Ω is the frequency set in the Fourier/Bochner representation of a shift-invariant quantum kernel $\kappa_Q(\mathbf{x}, \mathbf{x}') = g(\mathbf{x} - \mathbf{x}')$, appeared in Theorem 3. We hereby give a short discussion on $\omega_{\max}$ to show that it is moderate and usually negligible.

A simple upper bound. Consider any data-encoding circuit in which each input coordinate x_j appears m_j times via gates $\exp(-ix_j\mathbf{G}_{j,\ell})$, $\ell = 1, \ldots, m_j$. Let $\Delta_{j,\ell} := \lambda_{\max}(\mathbf{G}_{j,\ell}) - \lambda_{\min}(\mathbf{G}_{j,\ell})$ be the spectral diameter of the generator. Then every frequency vector $\boldsymbol{\omega} \in \Omega$ satisfies

$$|\omega_j| \;\le\; \sum_{\ell=1}^{m_j}\Delta_{j,\ell} \qquad (j = 1, \ldots, d), \tag{24}$$

and hence

$$\omega_{\max} \;\le\; \left\|\Big(\sum_{\ell=1}^{m_1}\Delta_{1,\ell}, \ldots, \sum_{\ell=1}^{m_d}\Delta_{d,\ell}\Big)\right\|_2 < \infty. \tag{25}$$

We next give a concrete bound on $\omega_{\max}$ for some example cases.

Example 1. Suppose $d = n$ and each x_j is encoded once on one qubit using $\mathbf{G}_j \in \{\mathbf{X}, \mathbf{Y}, \mathbf{Z}\}$, so $m_j = 1$. Since $\mathbf{X}, \mathbf{Y}, \mathbf{Z}$ have eigenvalues ± 1, we have $\Delta_{j,1} = 2$ and thus $|\omega_j| \leq 2$ for all j. Therefore

$$\omega_{\max} \;\leq\; \sqrt{\sum_{j=1}^{n} 2^2} \;=\; 2\sqrt{n} \;=\; \mathcal{O}\left(\sqrt{n}\right).$$

Example 2. Let $d = 1$ and encode the same scalar x on each of n qubits once, e.g., $|\psi(x)\rangle = (e^{-ix\mathbf{X}/2}|0\rangle)^{\otimes n}$. Then the fidelity kernel is

$$\kappa(x, x') = \left|\langle\psi(x)|\psi(x')\rangle\right|^2 = \cos^{2n}\Big(\frac{x - x'}{2}\Big).$$

Using $\cos(\theta) = \frac{1}{2}(e^{i\theta} + e^{-i\theta})$,

$$\cos^{2n}\Big(\frac{\Delta}{2}\Big) = 2^{-2n}\sum_{j=0}^{2n}\binom{2n}{j}e^{i(n-j)\Delta} = \sum_{\ell=-n}^{n} c_\ell e^{i\ell\Delta}, \qquad \Delta := x - x'.$$

Hence $\Omega = \{-n, -n+1, \ldots, n\}$ and

$$\omega_{\max} = n.$$

Example 3. Suppose $d = n$ and each coordinate x_j is re-uploaded L times with local Pauli generators (so $m_j = L$ and each $\Delta_{j,\ell} = 2$). Then by (25),

$$\omega_{\max} \;\leq\; \sqrt{\sum_{j=1}^{n}(2L)^2} \;=\; 2L\sqrt{n} \;=\; \mathcal{O}\left(L\sqrt{n}\right)$$

F DISCUSSION ON RUNNING TIME

Algorithm 2 GP-UCB

Input: Input (Action) space $\mathcal{A}$; GP Prior $\mu_0 = 0, \sigma_0, \kappa$.

1: **for** $t = 1, \ldots, T$ **do**
2: Choose $\mathbf{x}_t = \arg\max_{\mathbf{x}\in\mathcal{A}} \left[\mu_{t-1}(\mathbf{x}) + \sqrt{\beta_t}\sigma_{t-1}(\mathbf{x})\right]$, where $\beta_t = 2\log(|\mathcal{A}|t^2\pi^2/6\delta)$
3: Sample $y_t := f^*(\mathbf{x}_t) + \eta_t$, where η_t is the noise
4: Update posterior mean and variance μ_t and σ_t
5: **end for**

Algorithm 3 EC-GP-UCB (Enlarged Confidence GP-UCB)

Input: Kernel function $\kappa(\cdot,\cdot)$, domain $\mathcal{X}$, misspecification ε, and parameters B, λ, σ.

1: Set $\mu_0(\mathbf{x}) = 0$ and $\sigma_0(\mathbf{x}) = \kappa(\mathbf{x}, \mathbf{x})$ for all $\mathbf{x} \in \mathcal{X}$.
2: **for** $t = 1 \ldots T$ **do**
3: Choose

$$\mathbf{x}_t \in \arg\max_{\mathbf{x}\in\mathcal{X}} \left[\mu_{t-1}(\mathbf{x}) + \left(\beta_t + \frac{\varepsilon\sqrt{t}}{\sqrt{\lambda}}\right)\sigma_{t-1}(\mathbf{x})\right].$$

4: Observe reward $y_t^* = f^*(\mathbf{x}_t) + \eta_t$.
5: Update posterior mean and variance $\mu_t(\cdot)$ and $\sigma_t(\cdot)$
6: **end for**

In this work, we focus on the *quantum kernel bandit* problem, where the unknown mean-reward function f^* resides in the RKHS induced by a quantum kernel. Our objective is to minimize the cumulative regret $R_T = \sum_{t=1}^{T}(f^*(\mathbf{x}^*) - f^*(\mathbf{x}_t))$ where $\mathbf{x}^* = \arg\max_{\mathbf{x}\in\mathcal{X}} f^*(\mathbf{x})$ denotes the optimal action in hindsight.

A common approach, assuming the true (quantum) kernel is given, is to apply Gaussian process (GP) based methods such as GP-UCB [Srinivas et al., 2010]; see Algorithm 2. However, the computational burden of GP inference is typically high. Specifically, at round t, updating the GP posterior naively (as in Appendix A for line 4 in the algorithm) requires $\mathcal{O}(t^3)$ for an exact matrix inversion; or $\mathcal{O}(t^2)$ per round if one employs Sherman-Morrison rank-1 updates. Moreover, computing the next action $\mathbf{x}_t$ within the action set $\mathcal{A} = \mathcal{X}$ requires $\mathcal{O}(t^2|\mathcal{A}|)$ for posterior mean/variance queries. Summing over T rounds yields a total time of order

$$\mathcal{O}(T \times (T^2 + T^2|\mathcal{A}|)) = \mathcal{O}(T^3 + T^3|\mathcal{A}|).$$

Such a cubic dependence on T becomes intractable for large-scale bandit problems.

In contrast, our method considers a D-dimensional approximate feature map $\boldsymbol{\phi}_{\mathrm{RFF}}$ for the quantum kernel. When using SquareCB with the Vovk-Azoury-Warmuth forecaster in this D-dimensional feature space, the per-round update cost is only $\mathcal{O}(D^2)$. And inference (deciding $\mathbf{x}_t$) over the set $\mathcal{A}$ takes $\mathcal{O}(D \times |\mathcal{A}|)$. Consequently, the total cost across T rounds is

$$\mathcal{O}(TD^2 + TD|\mathcal{A}|),$$

which grows linearly in T.

Alternatively, one may run EC-GP-UCB [Bogunovic and Krause, 2021] (see Algorithm 3) using the same $\boldsymbol{\phi}_{\mathrm{RFF}}$. Rather than maintaining a full GP posterior, one performs standard Bayesian linear regression in the D-dimensional space. As above, each round's update can be done in $\mathcal{O}(D^2)$ using rank-1 matrix updates (Sherman-Morrison), and the inference cost is $\mathcal{O}(D \times |\mathcal{A}|)$. Hence, this approach also achieves a total cost of

$$\mathcal{O}(TD^2 + TD|\mathcal{A}|).$$

Thus, by replacing the exact (and computationally heavy) GP posterior inference with a D-dimensional approximation via random Fourier features (or other finite-dimensional embeddings), one obtains linear dependence on T, which is far more tractable in large-scale bandit settings.

G ABOUT P-GREEDY

Algorithm 4 Construction of Newton basis with P-greedy algorithm

Input: Kernel K, admissible error $e > 0$, a subset of points $\hat{\mathcal{X}} \subseteq \mathcal{X}$
Output: A subset of points $X_m \subseteq \hat{\mathcal{X}}$ and Newton basis $N_1, \ldots, N_m$ of $V(X_m)$

$\boldsymbol{\xi}_1 := \arg\max_{\mathbf{x} \in \hat{\mathcal{X}}} K(\mathbf{x}, \mathbf{x})$
$N_1(\mathbf{x}) := \frac{K(\mathbf{x}, \boldsymbol{\xi}_1)}{\sqrt{K(\boldsymbol{\xi}_1, \boldsymbol{\xi}_1)}}$
for $m = 1, 2, 3, \ldots$ **do**
  $P_m^2(\mathbf{x}) := K(\mathbf{x}, \mathbf{x}) - \sum_{k=1}^{m} \big(N_k(\mathbf{x})\big)^2$
  if $\max_{\mathbf{x} \in \hat{\mathcal{X}}} P_m^2(\mathbf{x}) < e^2$ **then**
    return $\{\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_m\}$ and $\{N_1, \ldots, N_m\}$
  end if
  $\boldsymbol{\xi}_{m+1} := \arg\max_{\mathbf{x} \in \hat{\mathcal{X}}} P_m^2(\mathbf{x})$
  $u(\mathbf{x}) := K(\mathbf{x}, \boldsymbol{\xi}_{m+1}) - \sum_{k=1}^{m} N_k(\boldsymbol{\xi}_{m+1})\, N_k(\mathbf{x})$
  $N_{m+1}(\mathbf{x}) := \frac{u(\mathbf{x})}{\sqrt{P_m^2(\boldsymbol{\xi}_{m+1})}}$
end for

We provide a more detailed introduction to P-greedy methods along with some established results to facilitate discussion in Appendix I.

Let $\mathcal{X}$ be a non-empty subset of a Euclidean space $\mathbb{R}^d$, and let $K : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ be a symmetric, positive definite kernel defined on $\mathcal{X}$. For a given discretization $\hat{\mathcal{X}} \subset \mathcal{X}$ and a function $f \in \mathcal{H}_K(\mathcal{X})$, the goal is to approximate f using the P-greedy procedure. This method iteratively selects a small set $X_n = \{\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_n\} \subset \hat{\mathcal{X}}$ of n representative points from the candidate set $\hat{\mathcal{X}}$, which often serves as a discrete approximation of $\mathcal{X}$. These points define a finite-dimensional subspace $V(X_n)$ that approximates the entire RKHS $\mathcal{H}_K(\mathcal{X})$.

At each iteration, P-greedy evaluates an "error surface" via the power function

$$P_{V(X)}(\mathbf{x}) := \sup_{f \in \mathcal{H}_K \setminus \{0\}} \frac{|f(\mathbf{x}) - (\Pi_{V(X)} f)(\mathbf{x})|}{\|f\|_{\mathcal{H}_K(\mathcal{X})}},$$

which quantifies the current approximation's looseness at each point $\mathbf{x}$. The method identifies the point where this error is the largest and adds it to the set X to reduce the worst-case error in subsequent iterations. Specifically, the next point is chosen as

$$\boldsymbol{\xi}_m = \arg\max_{\mathbf{x} \in \hat{\mathcal{X}}} P_{V(X_{m-1})}(\mathbf{x}).$$

Repeating this process for m iterations results in a set of points $X_m = \{\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_m\}$. These points generate the finite-dimensional subspace

$$V(X_m) = \operatorname{span}\{K(\cdot, \boldsymbol{\xi}_1), \ldots, K(\cdot, \boldsymbol{\xi}_m)\},$$

which approximates $\mathcal{H}_K(\mathcal{X})$. The resulting approximation effectively yields an approximate kernel whose feature map corresponds to the Newton basis $\{N_1, \ldots, N_m\}$, obtained via Gram–Schmidt orthonormalization of the basis $\{K(\cdot, \boldsymbol{\xi}_1), \ldots, K(\cdot, \boldsymbol{\xi}_m)\}$.

We provide the pseudo-code for the P-greedy algorithm, adapted from Takemori and Sato [2021], in Algorithm 4.

Furthermore, the following is known from Takemori and Sato [2021].

Theorem 10. *Let $\alpha, q > 0$ be parameters, d be the dimension of input data, and denote by $D = D_{q,\alpha}(T)$ the number of points returned by the P-greedy algorithm with error $e = \alpha / T^q$. Then:*

(i) Suppose K has finite smoothness *(e.g., Matérn kernels) with smoothness parameter $\nu > 0$. Then $D_{q,\alpha}(T) = O\big(\alpha^{-\frac{d}{\nu}} T^{\frac{dq}{\nu}}\big)$.*

(ii) Suppose K has infinite smoothness *(e.g., Gaussian kernel). Then $D_{q,\alpha}(T) = O\big((q \log T - \log(\alpha))^d\big)$.*

H P-GREEDY MISSPECIFICATION ERROR

In this section, we provide some ways to understand and bound the misspecification error from using the P-greedy algorithm.

H.1 SETUP

Let $\mathcal{X}$ be compact and let $\kappa : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ be continuous, symmetric, and positive definite with RKHS $\mathcal{H}_\kappa$ and norm $\|\cdot\|_{\mathcal{H}_\kappa}$. For each $\mathbf{x} \in \mathcal{X}$ define the kernel section $\kappa_{\mathbf{x}} := \kappa(\cdot, \mathbf{x}) \in \mathcal{H}_\kappa$ and the compact set

$$\mathcal{F} \;:=\; \{\kappa_{\mathbf{x}} :\; \mathbf{x} \in \mathcal{X}\} \subset \mathcal{H}_\kappa.$$

Given points $\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_D \in \mathcal{X}$ selected by P-greedy, define the associated approximation space

$$V_D \;:=\; \operatorname{span}\{\kappa_{\boldsymbol{\xi}_1}, \ldots, \kappa_{\boldsymbol{\xi}_D}\} \subset \mathcal{H}_\kappa, \tag{26}$$

and let $\Pi_{V_D} : \mathcal{H}_\kappa \to V_D$ denote the $\mathcal{H}_\kappa$-orthogonal projector.

We are interested in the misspecification between $f^* \in \mathcal{H}_\kappa$, and best possible approximation $f_D \in V_D$:

$$\varepsilon_D := \varepsilon_D(f^*) := \inf_{f_D \in V_D} \|f^* - f_D\|_\infty, \qquad \|g\|_\infty := \sup_{\mathbf{x} \in \mathcal{X}} |g(\mathbf{x})|.$$

H.2 BOUNDING MISSPECIFICATION ERROR BY THE KOLMOGOROV WIDTH

ε_D is controlled by the greedy error $\sigma_D(\mathcal{F})$. We first recall DeVore et al. [2013]'s definition of greedy error for the compact set $\mathcal{F} \subset \mathcal{H}_\kappa$:

$$\sigma_D(\mathcal{F}) \;:=\; \sup_{k^* \in \mathcal{F}} \inf_{k_D \in V_D} \|k^* - k_D\|_{\mathcal{H}_\kappa}. \tag{27}$$

Lemma 3 (Greedy error equals worst-case power function)**.** *For V_D defined in* (26)*, one has*

$$\sigma_D(\mathcal{F}) = \sup_{\mathbf{x}\in\mathcal{X}} \|\kappa_{\mathbf{x}} - \Pi_{V_D}\kappa_{\mathbf{x}}\|_{\mathcal{H}_\kappa}.$$

Proof. Fix $\mathbf{x} \in \mathcal{X}$. Since $\Pi_{V_D}\kappa_{\mathbf{x}} \in V_D$ and Π_{V_D} is the best approximation in Hilbert norm,

$$\inf_{f_D\in V_D} \|\kappa_{\mathbf{x}} - f_D\|_{\mathcal{H}_\kappa} = \|\kappa_{\mathbf{x}} - \Pi_{V_D}\kappa_{\mathbf{x}}\|_{\mathcal{H}_\kappa}.$$

Taking $\sup_{f\in\mathcal{F}}$, which is equivalent to $\sup_{\mathbf{x}\in\mathcal{X}}$, yields the claim. □

Lemma 4 (Misspecification bounded by greedy error)**.** *For any $f^* \in \mathcal{H}_\kappa$,*

$$\varepsilon_D(f^*) \;\le\; \|f^*\|_{\mathcal{H}_\kappa}\, \sigma_D(\mathcal{F}).$$

In particular, since we assumed $\|f^*\|_{\mathcal{H}_\kappa} \le B$, we have $\varepsilon_D(f^*) \le B\,\sigma_D(\mathcal{F})$.

Proof. Let $f_D := \Pi_{V_D} f^* \in V_D$. Then $\varepsilon_D(f^*) \le \|f^* - f_D\|_\infty$. Fix $\mathbf{x} \in \mathcal{X}$. By the reproducing property,

$$(f^* - f_D)(\mathbf{x}) = \langle f^* - f_D,\ \kappa_{\mathbf{x}}\rangle_{\mathcal{H}_\kappa}.$$

Since Π_{V_D} is an orthogonal projector, $\langle f^* - f_D,\ \Pi_{V_D}\kappa_{\mathbf{x}}\rangle_{\mathcal{H}_\kappa} = 0$ and therefore

$$(f^* - f_D)(\mathbf{x}) = \langle f^* - f_D,\ \kappa_{\mathbf{x}} - \Pi_{V_D}\kappa_{\mathbf{x}}\rangle_{\mathcal{H}_\kappa}.$$

By Cauchy–Schwarz and properties of orthogonal projection,

$$|(f^* - f_D)(\mathbf{x})| \le \|f^* - f_D\|_{\mathcal{H}_\kappa}\, \|\kappa_{\mathbf{x}} - \Pi_{V_D}\kappa_{\mathbf{x}}\|_{\mathcal{H}_\kappa} \le \|f^*\|_{\mathcal{H}_\kappa}\, \|\kappa_{\mathbf{x}} - \Pi_{V_D}\kappa_{\mathbf{x}}\|_{\mathcal{H}_\kappa}.$$

Taking the supremum over $\mathbf{x} \in \mathcal{X}$ and invoking Lemma 3 yields

$$\|f^* - f_D\|_\infty \le \|f^*\|_{\mathcal{H}_\kappa}\, \sigma_D(\mathcal{F}).$$

□

$\sigma_{2D}(\mathcal{F})$ is controlled by Kolmogorov width $d_D(\mathcal{F})$. Define the Kolmogorov D-width of $\mathcal{F}$ in $\mathcal{H}_\kappa$:

$$d_D(\mathcal{F}) \;:=\; \inf_{\substack{Y_D\subset\mathcal{H}_\kappa \\ \dim(Y_D)=D}} \sup_{k^*\in\mathcal{F}} \inf_{y_D\in Y_D} \|k^* - y_D\|_{\mathcal{H}_\kappa}. \tag{28}$$

It measures the best possible D-dimensional approximates to the full space.

We use the following result from DeVore et al. [2013], which asserts that P-greedy typically achieves near-optimal performance: the error σ_{2D} using $2D$ points satisfies $\sigma_{2D} \le \gamma^{-1}\sqrt{2d_D}$ for some constant $0 < \gamma < 1$:

Lemma 5 (greedy error bounded by Kolmogorov width)**.** *(Corollary 3.3 of DeVore et al. [2013])*

$$\sigma_{2D}(\mathcal{F}) \;\le\; \frac{\sqrt{2}}{\gamma}\,\sqrt{d_D(\mathcal{F})}, \tag{29}$$

for some constant γ.

Combining Lemma 4 with (29) gives the key implication:

$$\varepsilon_{2D}(f^*) \;\le\; B\,\sigma_{2D}(\mathcal{F}) \;\le\; \frac{\sqrt{2}B}{\gamma}\,\sqrt{d_D(\mathcal{F})}. \tag{30}$$

H.3 BOUNDING $d_D(\mathcal{F})$ FOR FINITE FOURIER-TYPE KERNELS

Further assume the kernel is translation invariant and admits the finite Fourier-type expansion

$$\kappa(\mathbf{x}, \mathbf{x}') = \sum_{\boldsymbol{\omega} \in \Omega} c_{\boldsymbol{\omega}}\, e^{i\langle \boldsymbol{\omega},\, \mathbf{x}-\mathbf{x}' \rangle}, \qquad c_{\boldsymbol{\omega}} \geq 0, \tag{31}$$

where Ω is finite. Let $\Omega_D \subset \Omega$ index the D largest coefficients $\{c_{\boldsymbol{\omega}}\}$.

For $\mathbf{x} \in \mathcal{X}$, the section function satisfies

$$\kappa(\cdot, \mathbf{x}) = \sum_{\boldsymbol{\omega} \in \Omega} c_{\boldsymbol{\omega}}\, e^{-i\langle \boldsymbol{\omega}, \mathbf{x} \rangle}\, e^{i\langle \boldsymbol{\omega}, \cdot \rangle}. \tag{32}$$

(A) A general upper bound Consider the D-dimensional subspace

$$Y_D \;:=\; \operatorname{span}\{e^{i\langle \boldsymbol{\omega}, \cdot \rangle} : \boldsymbol{\omega} \in \Omega_D\} \subset \mathcal{H}_\kappa.$$

For each $\mathbf{x} \in \mathcal{X}$, define the truncation $g_D(\cdot; \mathbf{x}) \in Y_D$ by

$$g_D(\cdot; \mathbf{x}) := \sum_{\boldsymbol{\omega} \in \Omega_D} c_{\boldsymbol{\omega}}\, e^{-i\langle \boldsymbol{\omega}, \mathbf{x} \rangle}\, e^{i\langle \boldsymbol{\omega}, \cdot \rangle}.$$

Then by (32),

$$\kappa(\cdot, \mathbf{x}) - g_D(\cdot; \mathbf{x}) = \sum_{\boldsymbol{\omega} \notin \Omega_D} c_{\boldsymbol{\omega}}\, e^{-i\langle \boldsymbol{\omega}, \mathbf{x} \rangle}\, e^{i\langle \boldsymbol{\omega}, \cdot \rangle}.$$

Using the triangle inequality in $\mathcal{H}_\kappa$ and $|e^{-i\langle \boldsymbol{\omega}, \mathbf{x} \rangle}| = 1$,

$$\|\kappa(\cdot, \mathbf{x}) - g_D(\cdot; \mathbf{x})\|_{\mathcal{H}_\kappa} \leq \sum_{\boldsymbol{\omega} \notin \Omega_D} c_{\boldsymbol{\omega}}\, \|e^{i\langle \boldsymbol{\omega}, \cdot \rangle}\|_{\mathcal{H}_\kappa}.$$

In the finite-rank RKHS induced by (31), one has $\|e^{i\langle \boldsymbol{\omega}, \cdot \rangle}\|_{\mathcal{H}_\kappa} = c_{\boldsymbol{\omega}}^{-1/2}$ (whenever $c_{\boldsymbol{\omega}} > 0$), hence

$$\sup_{\mathbf{x} \in \mathcal{X}} \operatorname{dist}(\kappa(\cdot, \mathbf{x}), Y_D)_{\mathcal{H}_\kappa} \leq \sum_{\boldsymbol{\omega} \notin \Omega_D} \sqrt{c_{\boldsymbol{\omega}}}. \tag{33}$$

Since $d_D(\mathcal{F})$ is the infimum over all D-dimensional subspaces, we conclude

$$d_D(\mathcal{F}) \leq \sum_{\boldsymbol{\omega} \notin \Omega_D} \sqrt{c_{\boldsymbol{\omega}}}. \tag{34}$$

(B) An improved upper bound for integer-valued frequencies. Assume additionally that $\mathcal{X} = \mathbb{T}^d = [0, 2\pi]^d$ and that $\Omega \subset \mathbb{Z}^d$. Then the Fourier characters $\{e^{i\langle \boldsymbol{\omega}, \cdot \rangle}\}_{\boldsymbol{\omega} \in \Omega}$ yield a Mercer decomposition, and the functions $\{\sqrt{c_{\boldsymbol{\omega}}}\, e^{i\langle \boldsymbol{\omega}, \cdot \rangle}\}_{\boldsymbol{\omega} \in \Omega}$ form an orthonormal family in $\mathcal{H}_\kappa$. Consequently, the residual above is an orthogonal sum in $\mathcal{H}_\kappa$ and Pythagoras' identity gives, for all $\mathbf{x} \in \mathcal{X}$,

$$\|\kappa(\cdot, \mathbf{x}) - g_D(\cdot; \mathbf{x})\|_{\mathcal{H}_\kappa}^2 = \sum_{\boldsymbol{\omega} \notin \Omega_D} c_{\boldsymbol{\omega}}.$$

Therefore,

$$d_D(\mathcal{F}) \leq \sup_{\mathbf{x} \in \mathcal{X}} \|\kappa(\cdot, \mathbf{x}) - g_D(\cdot; \mathbf{x})\|_{\mathcal{H}_\kappa} = \Big(\sum_{\boldsymbol{\omega} \notin \Omega_D} c_{\boldsymbol{\omega}} \Big)^{1/2}. \tag{35}$$

Final misspecification bounds. Combining (30) with (34) yields the general bound

$$\varepsilon_{2D}(f^*) \leq \frac{\sqrt{2}B}{\gamma} \Big(\sum_{\boldsymbol{\omega} \notin \Omega_D} \sqrt{c_{\boldsymbol{\omega}}} \Big)^{1/2}.$$

Under the integer-frequency assumption, combining (30) with (35) yields the sharper bound

$$\varepsilon_{2D}(f^*) \leq \frac{\sqrt{2}B}{\gamma} \Big(\sum_{\boldsymbol{\omega} \notin \Omega_D} c_{\boldsymbol{\omega}} \Big)^{1/4}.$$

H.4 EXAMPLE CIRCUITS AND THEIR SPECTRAL DECAY

Example: Product single-qubit rotation encoding (exponential tail). We consider the one-dimensional periodic domain $\mathcal{X} = \mathbb{T} = [0, 2\pi]$ and the n-qubit feature map

$$|\psi(x)\rangle \;:=\; \big(R_x(x)\,|0\rangle\big)^{\otimes n}, \qquad x \in [0, 2\pi], \tag{36}$$

where $R_x(\theta) = \exp(-i\theta\mathbf{X}/2)$ is a single-qubit rotation about the Pauli-X axis. We define the corresponding quantum kernel as the squared state overlap

$$\kappa(x, x') \;:=\; \big|\langle\psi(x)|\psi(x')\rangle\big|^2, \qquad x, x' \in [0, 2\pi]. \tag{37}$$

For one qubit,

$$\langle 0|\, R_x(x)^\dagger R_x(x')\,|0\rangle = \cos\Big(\frac{x - x'}{2}\Big) = \cos\Big(\frac{\Delta}{2}\Big).$$

Taking the n-fold tensor product gives $\langle\psi(x)|\psi(x')\rangle = \cos(\Delta/2)^n$. Squaring the modulus yields, with $\Delta := x - x'$,

$$\kappa(x, x') \;=\; \cos^{2n}\Big(\frac{\Delta}{2}\Big). \tag{38}$$

Since $\kappa(x, x')$ depends only on $\Delta = x - x'$, it is shift-invariant on $\mathbb{T}$ and admits a Fourier series $\kappa(\Delta) = \sum_{\ell\in\mathbb{Z}} c_\ell e^{i\ell\Delta}$.

Moreover, the kernel (38) is a trigonometric polynomial of degree n and therefore has finite spectrum $\ell \in \{-n, \dots, n\}$. Using $\cos(\theta) = \frac{1}{2}(e^{i\theta} + e^{-i\theta})$, we obtain

$$\begin{aligned}
\cos^{2n}\Big(\frac{\Delta}{2}\Big) &= 2^{-2n}\big(e^{i\Delta/2} + e^{-i\Delta/2}\big)^{2n} \\
&= 2^{-2n}\sum_{j=0}^{2n}\binom{2n}{j}e^{i(n-j)\Delta} \\
&= \sum_{\ell=-n}^{n} c_\ell e^{i\ell\Delta},
\end{aligned}$$

where the Fourier coefficients are explicitly

$$c_\ell = 2^{-2n}\binom{2n}{n-\ell}, \qquad \ell = -n, \dots, n. \tag{39}$$

In particular, $c_\ell \geq 0$, $c_{-\ell} = c_\ell$, and the sequence is unimodal with $c_0 \geq c_1 \geq \cdots \geq c_n$.

Now, by our previous results, the (squared) Kolmogorov D-width of $\mathcal{F}$ is upper bounded by the spectral tail:

$$d_D(\mathcal{F})^2 \;\leq\; \sum_{j>D} c_{(j)} \;=\; \sum_{\ell\notin\Omega_D} c_\ell,$$

where $c_{(1)} \geq c_{(2)} \geq \cdots$ denotes the non-increasing rearrangement and Ω_D indexes the top-D coefficients. Note the top coefficients correspond to the lowest frequencies. In particular, for any integer $k \in \{0, 1, \dots, n\}$, choosing

$$\Omega_{2k+1} := \{-k, -k+1, \dots, 0, \dots, k\}$$

gives

$$d_{2k+1}(\mathcal{F})^2 \;\leq\; \sum_{|\ell|\geq k+1} c_\ell = 2\sum_{\ell=k+1}^{n} c_\ell. \tag{40}$$

Now to bound (40), let $S \sim \mathrm{Bin}(2n, \frac{1}{2})$. Then, from (39),

$$c_\ell = \mathbb{P}(S = n - \ell).$$

Hence

$$2\sum_{\ell=k+1}^{n} c_\ell = 2\,\mathbb{P}(S \le n-k-1) = 2\,\mathbb{P}(S-n \le -(k+1)) \;\le\; 2\,\mathbb{P}(|S-n| \ge k+1).$$

Applying Hoeffding's inequality to the binomial variable S yields

$$\mathbb{P}(|S-n| \ge k+1) \le 2\exp\Big(-\frac{(k+1)^2}{n}\Big),$$

and substituting into (40) gives the explicit tail bound

$$d_{2k+1}(\mathcal{F})^2 \;\le\; 4\exp\Big(-\frac{(k+1)^2}{n}\Big), \qquad k = 0,1,\dots,n-1.$$

Equivalently, for any $D \in \{1,2,\dots,2n+1\}$, letting $k = \lfloor (D-1)/2 \rfloor$,

$$d_D(\mathcal{F})^2 \;\le\; 4\exp\Big(-\frac{(\lfloor (D-1)/2 \rfloor + 1)^2}{n}\Big) \;\le\; 4\exp\Big(-\frac{D^2}{16n}\Big). \tag{41}$$

Gaussian frequency profiles in quantum re-uploading models and sub-Gaussian spectral tails. Barthe and Pérez-Salinas [2024] study quantum re-uploading models, i.e., parametrized quantum circuits in which data-encoding unitaries are repeatedly interleaved with trainable gates. They prove that such models output functions with vanishing high-frequency components and, in particular, derive uniform upper bounds on derivatives with respect to the input data (hence limiting sensitivity to fine-scale variations). Moreover, their numerical experiments indicate that when the data-uploading gates are interleaved with random trainable gates, the Fourier transform of quantum re-uploading models exhibits Gaussian profiles.

For a kernel with Fourier expansion $\kappa(\Delta) = \sum_{\ell\in\mathbb{Z}} c_\ell e^{i\ell\Delta}$, if its Fourier weights satisfy a Gaussian (sub-Gaussian) decay

$$c_\ell \;\le\; C_0 \exp(-a\ell^2), \qquad \forall\, \ell \in \mathbb{Z},$$

for some constants $C_0, a > 0$. Since the largest coefficients correspond to the lowest frequencies, choosing $\Omega_{2k+1} = \{-k,\dots,k\}$ gives

$$d_{2k+1}(\mathcal{F})^2 \;\le\; \sum_{|\ell|\ge k+1} c_\ell \;\le\; 2C_0 \sum_{\ell=k+1}^{\infty} e^{-a\ell^2} \;\le\; 2C_0 \int_k^\infty e^{-at^2}\,dt \;\le\; \frac{C_0}{ak}\,e^{-ak^2},$$

where the last inequality uses the standard Gaussian tail bound. Equivalently, for $D = 2k+1$ and $k = \lfloor (D-1)/2 \rfloor$,

$$d_D(\mathcal{F})^2 \;\le\; \frac{C_1}{D}\exp\Big(-\frac{a}{4}(D-1)^2\Big), \tag{42}$$

for some constant $C_1 > 0$ depending only on (C_0, a). Thus, a Gaussian-shaped frequency profile implies an exponential decay of the spectral tail, and hence of the Kolmogorov width.

I EXTENSION TO THE CASE WHERE THE REWARD FUNCTION LIES IN AN RKHS OF CLASSICAL KERNELS

Although we initially presented our approach in the quantum context, it applies equally well to classical kernels, whether they induce finite- or infinite-dimensional RKHS. In particular, the same analysis used for Random Fourier Features carries over to the classical setting. Moreover, for certain infinite-dimensional kernels, the P-greedy algorithm admits a rapidly decaying error bound—often polynomial or even exponential in the number of selected points—according to Theorem 10.

Consequently, in a bandit optimization scenario over a finite discretized domain, one can directly integrate P-greedy approximations into methods such as EC-GP-UCB [Bogunovic and Krause, 2021] or standard misspecified linear bandit algorithms. By substituting the resulting ε-term (representing the approximation error), it is possible to derive improved regret guarantees alongside reduced computational overhead.

J EXPERIMENTAL DETAILS

We provide more details about the experiments in Section 4.

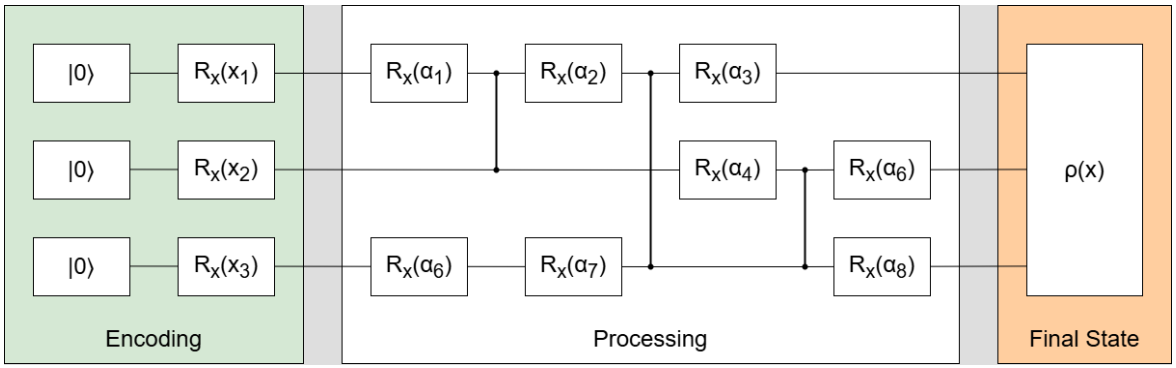


Figure 8: Example Circuit Structure for generating κ_Q, where x_i are the input data components and α_i are randomly sampled parameters for the variational layers

J.1 EXPERIMENT 1

- **Quantum Circuit Construction.** We generate the synthetic reward function f^* using a three-qubit parameterized quantum circuit composed of random layers of rotation gates and entangling operations, implemented via PennyLane. This results in a final circuit whose output density matrix defines our *global fidelity* quantum kernel $\kappa_Q(\cdot, \cdot)$.

  The structure of the parameterized quantum circuit that defines the kernel κ_Q (for this experiment) is depicted in Figure 8. The architecture begins with a data-encoding layer, by letting each component of the input vector $\mathbf{x}$ parameterizes a Pauli-X rotation gate acting on a corresponding qubit. In the notation of Eq. (4), this corresponds to $\mathbf{G}_j = \mathbf{X}_j$ for the data-encoding layer. Subsequently, the state is evolved through randomly structured layers of rotation and entangling gates, constructed using PennyLane's RandomLayers template. For a given input $\mathbf{x}$, the resulting n-qubit state is described by the density matrix $\boldsymbol{\rho}(\mathbf{x})$. The kernel is then defined as the Hilbert-Schmidt inner product between two states, $\kappa_Q(\mathbf{x}, \mathbf{x}') = \mathrm{Tr}[\boldsymbol{\rho}(\mathbf{x})\boldsymbol{\rho}(\mathbf{x}')]$, which quantifies their fidelity.

- **Action Space and Noise.** We discretize the input domain $[0, 2\pi]^3$ on a $10 \times 10 \times 10$ uniform grid, yielding 1000 possible arms (actions). At each round t, upon selecting an arm $\mathbf{x}_t$, we observe a noisy reward

$$y_t \;=\; f^*(\mathbf{x}_t) \;+\; \eta_t, \quad \eta_t \sim \mathcal{N}(0,\, 0.1^2).$$

  The observed values are then normalized to $[0, 1]$ across the entire grid.

- **Algorithms and Hyperparameters.** We compare three kernel-approximation strategies—*projected quantum kernels*, *random Fourier features* (RFF), and *P-greedy*—in combination with two bandit algorithms, *SquareCB* [Foster and Rakhlin, 2020] and *EC-GP-UCB* [Bogunovic and Krause, 2021]. We use their standard parameter settings from the respective papers, initializing the algorithms at round 1 (no special warm-up phase).

- **Projected Kernel Details.** Given our three-qubit circuit, we construct projected kernels by tracing out subsets of qubits. Specifically, we consider individual-qubit projections onto qubits $\{0\}, \{1\}, \{2\}$ as well as pairs $\{0, 1\}, \{1, 2\}, \{0, 2\}$. To create a *summed projected kernel* of size b, we sum the local projected kernels over b disjoint sub-systems. For instance, having "Number of Projected Kernels Summed" $= 3$ in the plot means summing the projected kernels on qubits 0, 1, and 2.

- **Random Fourier Features.** For RFF, we sample frequencies uniformly at random from the kernel's Fourier frequencies. We vary the RFF dimension D, as reported on the horizontal axis of Figures 3(b) and 4(b).

- **P-greedy.** For the P-greedy approach, we greedily select basis points from the same $10 \times 10 \times 10$ grid according to the power function criterion [Marchi et al., 2005], incrementally building a subspace that approximates κ_Q.

- **Trials and Metrics.** Each algorithm is run for $T = 100$ bandit rounds. We repeat the entire procedure 30 times with different random seeds. The plotted curves show the mean cumulative regret, with shaded regions or error bars indicating ± 1 standard deviation across trials. "Model complexity" on the horizontal axis is the number of summed local kernels for projected quantum kernels, the RFF dimension D for random Fourier features, or the number of basis points (kernel dimension) in P-greedy. The rightmost point in each plot always corresponds to the "Full kernel", using the unprojected 3-qubit fidelity kernel.

J.2 EXPERIMENT 2

- **Parameter Range and Actions.** We consider two coupling parameters (J_1, J_2), each discretized into 20 equally spaced points over $[-4, 4]$. This yields $20 \times 20 = 400$ total arms in the bandit problem. At each round t, we select an arm $\big(J_1, J_2\big) \in [-4, 4]^2$ and receive a reward corresponding to whether the system's ground state lies in phase II.
- **Ground-State Computation.** The generalized cluster Hamiltonian for n qubits is given by

$$\mathbf{H}_C \;=\; \sum_{j=1}^{n}\Big(\mathbf{Z}_j \;-\; J_1\,\mathbf{X}_j\mathbf{X}_{j+1} \;-\; J_2\,\mathbf{X}_{j-1}\mathbf{Z}_j\mathbf{X}_{j+1}\Big),$$

  where $\mathbf{X}_j$ and $\mathbf{Z}_j$ denote Pauli operators on qubit j. We exactly compute its ground state (i.e., the eigenvector corresponding to the lowest eigenvalue) via sparse matrix diagonalization.
- **Phase Labeling and Noise.** To label each (J_1, J_2) as "phase II" or "not phase II", we use the known boundary conditions:

$$\text{phase II if } \big(J_2 > -1 - J_1\big) \;\text{ and }\; \big(J_2 < J_1 - 1\big).$$

  Hence, if (J_1, J_2) is inside that region, the "ideal" label is 1; otherwise, it is 0. We then add *i.i.d.* Gaussian noise of variance $\sigma^2 = 0.01$, so the observed reward is

$$y \;=\; \mathbb{I}\{\text{phase II}\} \;+\; \eta, \quad \eta \sim \mathcal{N}(0, 0.01).$$

- **Regret Definition.** For each chosen (J_1, J_2), the *instantaneous regret* is 1 if that point is *outside* phase II (i.e., the ideal label is 0), and 0 if it is inside phase II. Thus, the *cumulative regret* over T rounds equals the total number of times we pick a point out of phase II in hindsight.
- **Kernel Approximations and Algorithm Setup.** We implement a three-qubit parameterized circuit (different from that of Experiment 1) with multiple layers of single-qubit rotations and entangling gates on the ground state of $\mathbf{H}_C$ as the initial state, yielding a fidelity kernel κ_Q. We then apply projected kernel and P-greedy with *EC-GP-UCB* [Bogunovic and Krause, 2021]. We run the algorithm for $T = 100$ rounds, record the cumulative regret, and then average over 30 runs. The standard deviation across these runs is shown as error bars in our plots.

J.3 EXPERIMENT 3

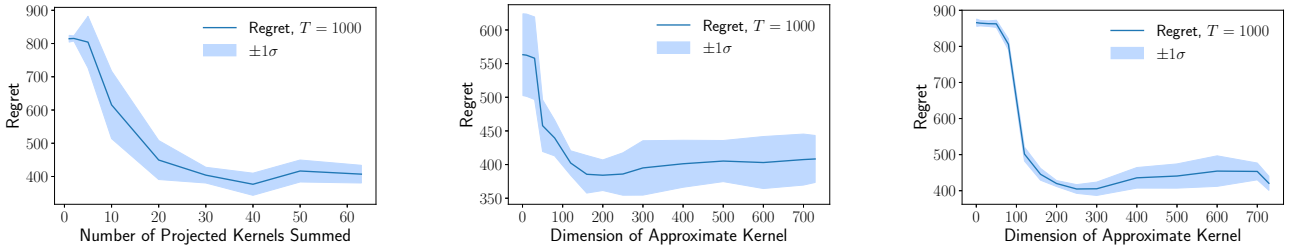
- **Hamiltonian Setup.** Specifically, we aim to approximate the ground state of the XYZ Hamiltonian, i.e., $J_X = -1$, $J_Y = J_Z = 0$, and $h_X = h_Y = 0$, $h_Z = -1$ in

$$\mathbf{H} = -\sum_{j=1}^{n-1}(J_X\mathbf{X}_j\mathbf{X}_{j+1} + J_Y\mathbf{Y}_j\mathbf{Y}_{j+1} + J_Z\mathbf{Z}_j\mathbf{Z}_{j+1}) - \sum_{j=1}^{n}(h_X\mathbf{X}_j + h_Y\mathbf{Y}_j + h_Z\mathbf{Z}_j)$$

  where $\mathbf{X}_j, \mathbf{Y}_j, \mathbf{Z}_j$ are Pauli matrices acting on the jth qubit. We fix $n = 3$ qubits for all experiments.
- **Circuit Ansatz and Initial State.** Following Nicoli et al. [2023], we employ an "Efficient SU(2)" circuit with PennyLane as a parameterized ansatz $\mathbf{U}(\mathbf{x})$. The initial state $|\psi_0\rangle$ is set to $|0\rangle^{\otimes 3}$. The circuit outputs

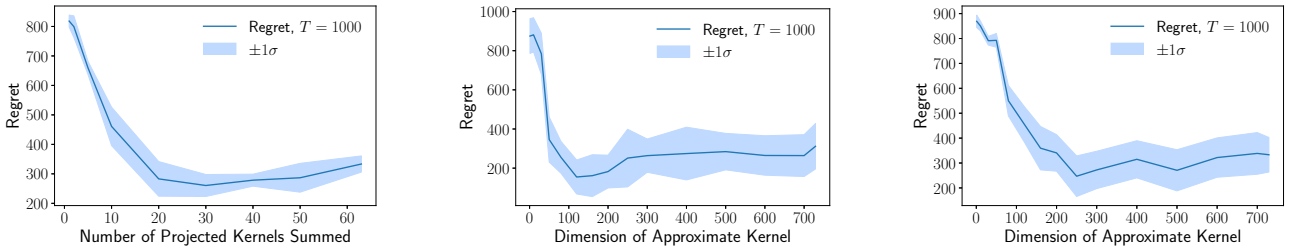
$$f^*(\mathbf{x}) \;=\; \langle\psi_0|\;\mathbf{U}(\mathbf{x})^\dagger\,\mathbf{H}\,\mathbf{U}(\mathbf{x})\;|\psi_0\rangle$$

  exactly without measurement noise.
- **Bayesian Optimization Details.** We define a Gaussian Process (GP) model with the *fidelity kernel* derived from the same 3-qubit circuit architecture, then apply Bayesian optimization using the Expected Improvement (EI) acquisition function. The EI is optimized by L-BFGS at each iteration to propose new circuit parameters. We run 50 optimization steps, repeating the procedure for 30 independent random seeds that affect parameter initialization, random draws in the approximate kernels, etc.
- **Metrics and Plots.** The optimization runs for 50 steps, we record and plot the *best* (lowest) energy found so far. The error bars (or shaded regions) depict ± 1 standard deviation across 30 trials.



(a) Projected Quantum Kernels (b) Random Fourier Features (c) P-greedy

Figure 9: Cumulative regret of SquareCB algorithm as a function of approximation dimension D for (a) projected quantum kernels, (b) RFF and (c) P-greedy approximation.



(a) Projected Quantum Kernels (b) Random Fourier Features (c) P-greedy

Figure 10: EC-GP-UCB with different kernel approximation approaches for reward functions drawn from GPs with the full quantum kernel

K BIGGER QUBIT SYSTEM FOR EXPERIMENT 1

This section provides additional results to illustrate scalability beyond the 3-qubit setting in Section 4.1. Following exactly the same protocol as Experiment 1 (same reward-generation mechanism, noise model, and evaluation procedure, though we increase T to 1000 here), we increase the number of qubits from 3 to 6, specifically, we consider a synthetic reward function f^* drawn from a GP prior induced by the full 6-qubit product kernel on the domain $[0, 2\pi]^6$. We then run the same bandit pipelines (SquareCB and EC-GP-UCB) with the same three kernel-approximation strategies: projected quantum kernels, random Fourier features (RFF), and P-greedy.

Figures 9 and 10 mirror Figures 3 and 4 in the main text, each curve reports cumulative regret as a function of the approximation dimension (number of summed projected kernels for LPQK; feature dimension D for RFF; or the number of selected basis elements for P-greedy). As in the 3-qubit experiments, we observe the characteristic "U-shaped" behavior: small approximation dimension underfits the true function (high misspecification error), while overly large dimension increases the information-gain penalty and can degrade regret.

Overall, these results confirm that the qualitative trade-off identified in Section 4.1 persists at higher qubit counts, supporting the scalability of our approximation framework to larger quantum systems.

L LIMITATIONS AND FUTURE DIRECTIONS

- **Conservativeness of Regret Bounds.** Our theoretical regret bounds provide upper estimates that may be overly conservative in practice. Consequently, using these bounds to select an optimal kernel dimension D might lead to suboptimal or unnecessarily large choices.
- **Approximation Error Uncertainty.** Although Random Fourier Features (RFF) typically yield an error rate of $\varepsilon = 1/\sqrt{D}$, this estimate can be pessimistic in many real-world scenarios. Faster decay rates—such as polynomial or exponential—are achievable but generally require problem-specific analysis. Additionally, for projected kernels and P-greedy methods, there is no universal closed-form bound for ε; each case necessitates empirical evaluation or tailored theoretical analysis.
- **Limitations of Fourier Series Representations.** Our RFF analysis relies on the assumption that the quantum kernel

admits a discrete or tractable Fourier expansion. However, some kernels may lack such a structure, making our current approach inapplicable without significant modifications.

- **Noise Model and Kernel Estimation.** We model the observed reward noise as i.i.d. Gaussian, which is a standard approximation to finite-shot measurement noise in the large-shot regime. This abstraction does not capture structured NISQ noise, such as readout errors, depolarizing noise, relaxation, drift, or cross-talk. Moreover, our approximation-error analysis assumes exact kernel values; finite-shot kernel estimation and device noise would introduce additional errors that may interact with the kernel approximation error.
- **Quantum Measurement Overhead.** The computational speedups discussed in Appendix F concern classical posterior updates and linear-algebra costs. If kernel values are estimated on quantum hardware, the total measurement cost can still be significant. Full GP-UCB requires at least $\mathcal{O}(T^2)$ pairwise kernel evaluations over T selected points, while a fixed D-dimensional approximation requires $\mathcal{O}(TD)$ feature or basis evaluations after the surrogate is fixed. These counts must be multiplied by the number of shots required per kernel or observable estimate. For LPQK, a direct tomography-based implementation over all subsystems $|s| \leq b$ may additionally require local Pauli measurements over $\sum_{w=0}^{b} \binom{n}{w} 3^w$ settings. A full measurement-aware regret and runtime analysis is left for future work.
- **Generality Beyond Quantum Kernels.** The misspecified-bandit framework applies to general classical kernels as well as quantum kernels.