MULTIVERSE BAYESIAN OPTIMIZATION WITH JOINT FINITE-FEATURE SURROGATES

Anonymous authors
Paper under double-blind review

ABSTRACT

Bayesian optimization (BO) typically relies on a Gaussian process (GP) surrogate whose performance depends on the choice of kernel. When an objective contains several patterns, committing to one kernel can miss useful structure and direct acquisition toward uninformative regions. We propose Multiverse Bayesian Optimization (MVBO), a novel BO framework that represents candidate kernels as complementary worlds and learns their contributions jointly in a compact feature space. By retaining features from each world before combining them, the surrogate can preserve structure that compressing the kernel sum would discard. The resulting Bayesian linear regression provides a Gaussian posterior that captures dependence between worlds, supports standard acquisition functions, and requires $O(r^2)$ time per update for a fixed total feature dimension r. For upper confidence bound (UCB) acquisition with fixed features, we establish best-query regret $\tilde{O}(\sqrt{r/T})$ under a well-specified prior on a finite domain without repeated queries, and extend this guarantee to account for the error introduced by the finite-feature approximation. We also study when preserving each world's features helps and quantify the accuracy and computational trade-offs of the proposed representation. Across three emerging scientific and AI design tasks, including molecular conformer search, virtual screening for drug discovery, and LLM data-mixture optimization, MVBO-UCB reduces final normalized simple regret by approximately 42%, 32%, and 38%, respectively, relative to the strongest competing BO baseline.

1 INTRODUCTION

Bayesian optimization (BO) is a sample-efficient framework for optimizing expensive black-box functions in science and engineering (Shahriari et al., 2016; Garnett, 2023). It maintains a probabilistic surrogate of the objective, typically a Gaussian process (GP) (Rasmussen & Williams, 2006), and selects queries by maximizing an acquisition function that balances exploration and exploitation. Both the flexibility of GPs and the quality of their uncertainty estimates ultimately depend on how well the chosen kernel captures the structure of the target function.

In standard BO practice, the kernel is often chosen up front, using a squared exponential or Matérn kernel by default, and remains fixed for the entire optimization process (Snoek et al., 2012). This approach assumes that the objective behaves similarly throughout the domain. However, real-world objectives often exhibit heterogeneous or multi-regime structure. For instance, molecular energy landscapes exhibit discrete rotameric basins with incommensurable curvature (Hawkins et al., 2010; Riniker & Landrum, 2015), while drug-discovery objectives fragment across scaffolds with distinct structure–activity relationships (Gómez-Bombarelli et al., 2018; Griffiths & Hernández-Lobato, 2020; Korovina et al., 2020). A more recent example is optimizing large language model (LLM) training data mixtures for multiple skills, which produces distinct regions in the objective landscape, corresponding to different regimes with unique behavior (Chew et al., 2026; Lambert et al., 2025). In such settings, a single kernel may be insufficient since transitions between regimes can be blurred or misinterpreted as noise (see Appendix B.2.6). Even sophisticated non-stationary kernels (Paciorek & Schervish, 2004; Damianou & Lawrence, 2013; Wilson et al., 2016) often oversmooth across boundaries, failing to represent sharp transitions between distinct regions.

Addressing this limitation requires surrogates that can adapt to different regimes across the input space, rather than committing to a single global representation. Existing attempts to do so include

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

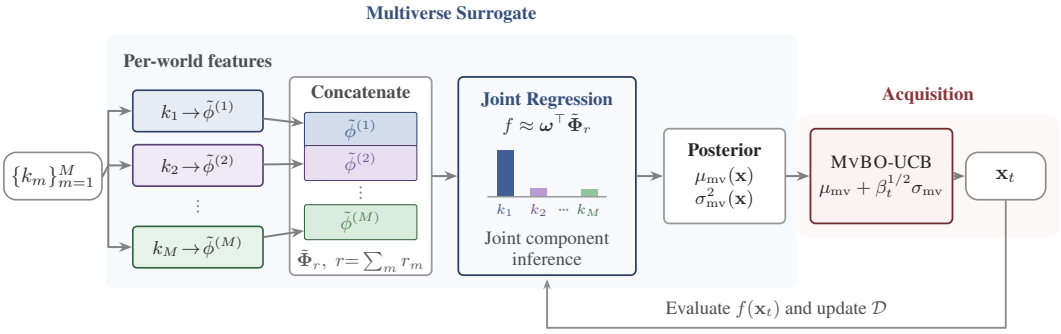


Figure 1: Overview of MvBO. Per-world feature maps are concatenated into a single fixed representation. Joint Bayesian regression on this representation yields a posterior that guides acquisition through its mean, variance and cross-world covariance.

automated kernel search (Duvenaud et al., 2013; Malkomes & Garnett, 2018; Suwandi et al., 2025), multiple kernel learning (Gönen & Alpaydın, 2011; Yin et al., 2020), and GP ensembles (Lu et al., 2023; Candelieri et al., 2025), but each settles on one kernel, one weighted mixture, or one averaged prediction applied uniformly across the domain. To close this gap, we propose Multiverse Bayesian Optimization (MvBO), which treats each kernel as a "world" representing a distinct structural hypothesis (Figure 1). Our main contributions are:

- We introduce MvBO, a joint surrogate that retains per-world features and learns their contributions together so acquisition accounts for uncertainty in their sum. With total feature dimension fixed at r, each posterior update costs $O(r^2)$ independent of the number of observations (Section 3).
- We prove a best-query regret bound for upper confidence bound (UCB) acquisition with fixed features and extend the analysis to account for approximation error when the true function is not fully captured by the finite-feature model (Section 4).
- We evaluate MvBO on molecular conformer search, virtual screening for drug discovery, and LLM data-mixture optimization, and analyze when retaining per-world features helps (Section 6).

2 Preliminaries

Bayesian optimization. Let $f : \mathcal{X} \to \mathbb{R}$ be an objective on a compact domain. BO seeks $\mathbf{x}^* \in \arg\max_{\mathbf{x}\in\mathcal{X}} f(\mathbf{x})$ through observations $y_i = f(\mathbf{x}_i) + \varepsilon_i$. A GP prior with kernel k and independent noise $\varepsilon_i \sim \mathcal{N}(0, \sigma_\varepsilon^2)$ yields a Gaussian posterior with latent mean $\mu_t(\mathbf{x})$ and variance $\sigma_t^2(\mathbf{x})$ after t observations (Rasmussen & Williams, 2006). Acquisition functions then use this posterior to decide where to query next. For example, expected improvement measures the expected gain over an incumbent (Jones et al., 1998), while Thompson sampling optimizes a function drawn from the posterior (Russo et al., 2018). Here, we focus on the upper confidence bound (UCB),

$$\alpha_{\mathrm{UCB}}(\mathbf{x}) = \mu_t(\mathbf{x}) + \beta_t^{1/2}\sigma_t(\mathbf{x}), \tag{1}$$

which scores each candidate by its posterior mean plus an exploration bonus and has well-studied regret guarantees (Srinivas et al., 2010; Chowdhury & Gopalan, 2017).

Information gain. The convergence rate of BO methods depends on how quickly the surrogate model's uncertainty can be reduced, which is quantified by the maximum information gain of the kernel k:

$$\gamma_T(k) = \max_{(\mathbf{x}_1,\ldots,\mathbf{x}_T)\in\mathcal{X}^T} \frac{1}{2}\log\det\left(\mathbf{I}_T + \sigma_\varepsilon^{-2}\mathbf{K}_T\right), \tag{2}$$

where $[\mathbf{K}_T]_{ij} = k(\mathbf{x}_i, \mathbf{x}_j)$ denotes the kernel matrix entries. For kernels of rank at most r, such as those induced by r features (Section 3.1), with bounded feature norms, the maximum information gain scales as $\gamma_T(k) = O(r \log T)$ (Srinivas et al., 2010; Vakili et al., 2021). Since regret bounds for algorithms like GP-UCB scale with $\sqrt{\gamma_T}$, fewer features imply faster uncertainty reduction and tighter regret bounds of order $\tilde{O}(\sqrt{r/T})$, as derived in Section 4.2.

Finite-dimensional features. A finite-dimensional feature map $\tilde{\phi}_r : \mathcal{X} \to \mathbb{R}^r$ approximates a kernel via the inner product $\tilde{k}(\mathbf{x}, \mathbf{x}') = \tilde{\phi}_r(\mathbf{x})^\top \tilde{\phi}_r(\mathbf{x}')$, where placing a Gaussian prior on the coefficients yields a finite-rank Gaussian Process (Rasmussen & Williams, 2006, Ch. 2). Common constructions include Nyström methods, which use kernel evaluations at landmark points (Williams & Seeger, 2001), and Random Fourier Features, which sample frequencies for stationary kernels (Rahimi & Recht, 2007). The choice of construction method primarily determines the approximation quality and computational cost, while the kernel's behavior is governed by the feature count r, which sets the representation size and model capacity, as well as the feature scaling, which controls the kernel amplitude or signal variance.

3 Multiverse Bayesian Optimization

We construct features for each kernel and fit their coefficients jointly, obtaining a posterior whose predictions and uncertainty guide the next evaluation.

3.1 Kernel Worlds and Their Joint Representation

For candidate kernels $\{k_m\}_{m=1}^M$, assign each world a feature map $\tilde{\phi}_{r_m}^{(m)} : \mathcal{X} \to \mathbb{R}^{r_m}$ that approximates its kernel as $\tilde{k}_m(\mathbf{x}, \mathbf{x}') = \tilde{\phi}_{r_m}^{(m)}(\mathbf{x})^\top \tilde{\phi}_{r_m}^{(m)}(\mathbf{x}')$. Concatenating these maps gives the joint representation,

$$\tilde{\mathbf{\Phi}}_r(\mathbf{x}) = \left[\tilde{\phi}_{r_1}^{(1)\top}(\mathbf{x}), \ldots, \tilde{\phi}_{r_M}^{(M)\top}(\mathbf{x})\right]^\top, \qquad r = \sum_m r_m. \tag{3}$$

The objective model is $f = \sum_m f_m$, where $f_m(\mathbf{x}) = \boldsymbol{\omega}^{(m)\top} \tilde{\phi}_{r_m}^{(m)}(\mathbf{x})$. Several worlds can therefore contribute at the same input, with potentially overlapping features that make the effective dimension smaller than r.

Spectral implementation. For landmarks $L = \{\ell_i\}_{i=1}^q$, let $\mathbf{K}_m = [k_m(\ell_i, \ell_j)]_{i,j} = \mathbf{V}_m \mathbf{\Lambda}_m \mathbf{V}_m^\top$ and $k_m(L, \mathbf{x}) = [k_m(\ell_i, \mathbf{x})]_i$. Retaining the r_m leading positive eigenpairs gives

$$\tilde{\phi}_{r_m}^{(m)}(\mathbf{x}) = \mathbf{\Lambda}_{m,r_m}^{-1/2} \mathbf{V}_{m,r_m}^\top k_m(L, \mathbf{x}). \tag{4}$$

Our experiments use this spectral Nyström map (Williams & Seeger, 2001). By default, L is a uniform sample from the candidate domain, drawn without using responses and then frozen (Appendix B.2.7). Although Eq. (3) accepts other finite maps, including random Fourier features for compatible kernels, each construction needs its own approximation-error bound (Appendix B.1.1). Once initialization is complete, kernel parameters, features, and response normalization remain fixed, and new observations update only the joint coefficient posterior.

3.2 Closed-Form Posterior and the Multiverse Kernel

We place the prior $\boldsymbol{\omega} \sim \mathcal{N}(\mathbf{0}, \sigma_\omega^2 \mathbf{I}_r)$ on the concatenated coefficients, with $f(\mathbf{x}) = \boldsymbol{\omega}^\top \tilde{\mathbf{\Phi}}_r(\mathbf{x})$ and observations $y = f(\mathbf{x}) + \varepsilon$, $\varepsilon \sim \mathcal{N}(0, \sigma_\varepsilon^2)$. For observations $\mathcal{D}$, conjugacy gives the Bayesian linear regression posterior (Rasmussen & Williams, 2006, Ch. 2) of $\boldsymbol{\omega} \mid \mathcal{D} \sim \mathcal{N}(\boldsymbol{\mu}_\omega, \mathbf{\Sigma}_\omega)$, with

$$\mathbf{\Sigma}_\omega = \left(\sigma_\omega^{-2} \mathbf{I}_r + \sigma_\varepsilon^{-2} \sum_{(\mathbf{x}_i, y_i) \in \mathcal{D}} \tilde{\mathbf{\Phi}}_r(\mathbf{x}_i) \tilde{\mathbf{\Phi}}_r(\mathbf{x}_i)^\top \right)^{-1}, \tag{5}$$

$$\boldsymbol{\mu}_\omega = \sigma_\varepsilon^{-2} \mathbf{\Sigma}_\omega \sum_{(\mathbf{x}_i, y_i) \in \mathcal{D}} \tilde{\mathbf{\Phi}}_r(\mathbf{x}_i) y_i. \tag{6}$$

Because the worlds explain the observed sum jointly, one component can account for structure left unexplained by another, whereas in an independent ensemble each model is fit separately to the same observations y_i before their predictions are combined (Appendix A.1.4). The resulting predictive distribution is $p(f(\mathbf{x}) \mid \mathcal{D}) = \mathcal{N}(\mu_{\mathrm{mv}}(\mathbf{x}), \sigma_{\mathrm{mv}}^2(\mathbf{x}))$, with

$$\mu_{\mathrm{mv}}(\mathbf{x}) = \boldsymbol{\mu}_\omega^\top \tilde{\mathbf{\Phi}}_r(\mathbf{x}), \tag{7}$$

$$\sigma_{\mathrm{mv}}^2(\mathbf{x}) = \tilde{\mathbf{\Phi}}_r(\mathbf{x})^\top \mathbf{\Sigma}_\omega \tilde{\mathbf{\Phi}}_r(\mathbf{x}). \tag{8}$$

Adding σ_ε^2 gives the predictive variance of a noisy observation, while partitioning the mean reveals each world's contribution:

$$\mu_{\mathrm{mv}}(\mathbf{x}) = \sum_m c_m(\mathbf{x}), \qquad c_m(\mathbf{x}) = \boldsymbol{\mu}_\omega^{(m)\top} \tilde{\phi}_{r_m}^{(m)}(\mathbf{x}), \tag{9}$$

World m's fitted coefficients $\boldsymbol{\mu}_\omega^{(m)}$ do not depend on $\mathbf{x}$, so $c_m(\mathbf{x})$ varies across the input only because its features $\tilde{\phi}_{r_m}^{(m)}(\mathbf{x})$ do. We use $w_m^{\mathrm{attr}}(\mathbf{x}) = |c_m(\mathbf{x})|$ to measure this contribution's size at each point, subject to the usual caveats for additive-model attribution scores (Appendices A.1.4 and B.2.2).

Writing $\tilde{k}_m(\mathbf{x}, \mathbf{x}') = \tilde{\phi}_{r_m}^{(m)}(\mathbf{x})^\top \tilde{\phi}_{r_m}^{(m)}(\mathbf{x}')$, the prior induces

$$\mathrm{MvK}(\mathbf{x}, \mathbf{x}') = \sigma_\omega^2 \tilde{\boldsymbol{\Phi}}_r(\mathbf{x})^\top \tilde{\boldsymbol{\Phi}}_r(\mathbf{x}') = \sigma_\omega^2 \sum_m \tilde{k}_m(\mathbf{x}, \mathbf{x}'). \tag{10}$$

Equation (10) shows that MvBO's prior is exactly the GP defined by this finite-rank sum kernel, so its RKHS $\mathcal{H}_{\mathrm{MvK}}$ contains every sum of functions drawn from the worlds' individual RKHSs (Proposition 1), a property shared by any positively weighted sum of kernels. MvBO's distinguishing feature is instead its explicit finite representation, which fixes each world's feature dimension r_m and updates all worlds' coefficients in a single joint fit, with c_m retaining its standard additive-GP interpretation despite the truncation (Appendix A.1.4).

3.3 Query Selection with Multiverse Acquisition

We use the joint posterior in UCB acquisition, balancing each candidate's predicted value against uncertainty in the combined objective:

$$\alpha_{\mathrm{UCB},t}(\mathbf{x} \mid \mathcal{D}_{t-1}) = \mu_{\mathrm{mv},t-1}(\mathbf{x}) + \sqrt{\beta_t}\, \sigma_{\mathrm{mv},t-1}(\mathbf{x}). \tag{11}$$

For $M = 1$, this reduces to GP-UCB with the corresponding finite-feature kernel (Proposition 4). With multiple worlds, partitioning $\boldsymbol{\Sigma}_\omega$ into blocks $\boldsymbol{\Sigma}_{m\ell}$ and writing $\tilde{\phi}_m = \tilde{\phi}_{r_m}^{(m)}$ reveals the covariance terms in the acquisition variance:

$$\sigma_{\mathrm{mv}}^2(\mathbf{x}) = \sum_{m=1}^{M} \sum_{\ell=1}^{M} \tilde{\phi}_m(\mathbf{x})^\top \boldsymbol{\Sigma}_{m\ell} \tilde{\phi}_\ell(\mathbf{x}). \tag{12}$$

Observing the sum couples components whose priors were independent, producing off-diagonal covariance terms that a model fitting each component separately would omit (Appendix A.1.4). These terms can have either sign, so joint conditioning can increase or decrease the exploration bonus at a given input. With features and noise variance fixed, the posterior covariance $\boldsymbol{\Sigma}_\omega$ depends only on the query locations, while the coefficient means $\boldsymbol{\mu}_\omega^{(m)}$ also depend on the observed values.

3.4 The Full MvBO Loop

By fixing r features, we cap each posterior update at $O(r^2)$ time and storage, instead of the $O(n^2)$ cost an exact dense sum-kernel GP incurs as observations accumulate (Appendix B.1.5). This saving is possible because Eqs. (5)–(6) reduce to two running sums over $r \times r$ matrices and r-vectors, from the inverse of $\boldsymbol{\Sigma}_\omega$ and the vector it multiplies to produce $\boldsymbol{\mu}_\omega$. Each new observation adds one rank-one term to each sum, so updating these sums costs $O(r^2)$ regardless of how many observations came before. After $n_0 + T$ calls, the loop recommends the queried maximizer of the posterior mean (Appendix A.2.3).

4 Regret Analysis

We analyze best-query simple regret, the gap between the optimum and the best true value among evaluated points, first assuming fixed features and an objective drawn from the resulting GP model. We then extend the analysis to objectives containing structure outside this representation, so the guarantee accounts for both learning and approximation error.

Algorithm 1: MvBO: Multiverse Bayesian Optimization

Input: Additional query budget T, initial size n_0, kernels $\{k_m\}_{m=1}^{M}$, feature budgets $\{r_m\}$, landmark set L (default: uniform on the candidate domain)
Output: Recommended queried point $\mathbf{x}_{\mathrm{rec}}$

Evaluate n_0 distinct random points and form $\mathcal{D}_0$
for $m = 1$ **to** M **do**
  Construct and freeze $\tilde{\phi}_{r_m}^{(m)}$ from L via Eq. (4)
Concatenate the maps and initialize the posterior on $\mathcal{D}_0$
for $t = 1$ **to** T **do**
  Compute $\mu_{\mathrm{mv},t-1}$ and $\sigma^2_{\mathrm{mv},t-1}$ via Eqs. (7)–(8)
  Select an unqueried $\mathbf{x}_t$ by the acquisition, e.g. Eq. (11)
  Observe y_t and set $\mathcal{D}_t = \mathcal{D}_{t-1} \cup \{(\mathbf{x}_t, y_t)\}$
  Update Eqs. (5)–(6) incrementally
Recompute $\mu_{\mathrm{mv},T}$ on $\mathcal{D}_T$ via Eq. (7)
$\mathbf{x}_{\mathrm{rec}} \leftarrow \arg\max_{(\mathbf{x},y)\in\mathcal{D}_T} \mu_{\mathrm{mv},T}(\mathbf{x})$
return $\mathbf{x}_{\mathrm{rec}}$

4.1 Assumptions

Assumption 1 (Domain and queries)**.** *$\mathcal{X}$ is finite. The initial input set S_0 contains n_0 distinct points. Subsequent UCB queries maximize over all unqueried points, with $1 \le T \le |\mathcal{X}| - n_0$.*

Assumption 2 (Fixed, well-specified prior)**.** *Kernel parameters, landmarks, and sampled frequencies are chosen without using observed responses. Conditional on these choices, the feature map is fixed and $f(\mathbf{x}) = \boldsymbol{\omega}^\top \tilde{\boldsymbol{\Phi}}_r(\mathbf{x})$ with $\boldsymbol{\omega} \sim \mathcal{N}(\mathbf{0}, \sigma_\omega^2 \mathbf{I}_r)$. The posterior conditions this same model on all observations.*

Assumption 3 (Noise)**.** *Observation noise is i.i.d. $\mathcal{N}(0, \sigma_\varepsilon^2)$ and independent of the prior coefficients $\boldsymbol{\omega}$, with $\sigma_\varepsilon^2 > 0$.*

Assumption 4 (Bounded features)**.** *$s_{\max} = \sup_{\mathbf{x}\in\mathcal{X}} \|\tilde{\boldsymbol{\Phi}}_r(\mathbf{x})\|^2 < \infty$ and $\kappa^2 := \sigma_\omega^2 s_{\max} > 0$.*

These assumptions distinguish the analyzed policy from practical variations. Fitting kernels to observed responses requires the alternative analysis in Appendices A.2.4 and A.2.7, while candidate subsampling and continuous local refinement fall outside Assumption 1's requirement to maximize UCB over every unqueried point in the fixed domain.

4.2 Information Gain and Simple Regret

Lemma 1 (Finite-rank information gain)**.** *For* MvK *in Eq.* (10)*,*

$$\gamma_T := \gamma_T(\mathrm{MvK}) \le \frac{r}{2} \log\left(1 + \frac{\kappa^2 T}{r\sigma_\varepsilon^2}\right). \tag{13}$$

The bound also holds with r replaced by the dimension of the span of $\{\tilde{\boldsymbol{\Phi}}_r(\mathbf{x}) : \mathbf{x} \in \mathcal{X}\}$.

Theorem 1 (Best-query guarantee without replacement)**.** *Under Assumptions 1–4, use $\beta_t = 2\log(|\mathcal{X}|t^2\pi^2/(6\delta))$, $\delta \in (0,1)$. With probability at least $1-\delta$, for all admissible T,*

$$s_T := f(\mathbf{x}^*) - \max_{\mathbf{x}\in S_0 \cup \{\mathbf{x}_1,\dots,\mathbf{x}_T\}} f(\mathbf{x}) \le \sqrt{\frac{C_1 \beta_T \gamma_T}{T}}, \quad C_1 = \frac{8\kappa^2}{\log(1+\kappa^2/\sigma_\varepsilon^2)}. \tag{14}$$

Proof sketch. A union bound gives simultaneous Gaussian confidence intervals (Srinivas et al., 2010). If an optimum has been queried, $s_T = 0$, while an unqueried optimum remains eligible and UCB gives $f(\mathbf{x}^*) - f(\mathbf{x}_t) \le 2\sqrt{\beta_t}\sigma_{t-1}(\mathbf{x}_t)$. Bounding the variance sum by information gain and applying the Cauchy–Schwarz inequality gives a cumulative-regret bound. Since the minimum per-step regret cannot exceed the average, dividing by T proves the result. Appendix A.2 gives both full proofs.

Corollary 1 (Dependence on feature dimension)**.** *The Bayesian best-query rate is $\tilde{O}(\sqrt{r/T})$, suppressing logarithmic factors, under the stated fixed-domain assumptions. With a fixed function of RKHS norm at most B and the confidence width in Appendix A.2.4, it becomes $\tilde{O}((B\sqrt{r}+r)/\sqrt{T})$ for fixed noise scales and initialization.*

The number of worlds enters through the total feature count r without an additional factor of M, although feature norms, noise, and domain size also affect the bound (Appendix A.2). This guarantee concerns the best true value evaluated. Algorithm 1 recommends the queried posterior-mean maximizer (Appendix A.2.3). Appendix A.2.5 gives a cumulative-regret bound when repeated queries are permitted.

Accounting for approximation error. To account for structure the features cannot represent, write $f = g + h$ with $g \in \mathcal{H}_{\mathrm{MvK}}$, $\|g\|_{\mathcal{H}_{\mathrm{MvK}}} \le B$, and $\|h\|_\infty \le \epsilon$. Under the noise assumptions in Appendix A.2.6, enlarging the UCB confidence width to include this residual gives

$$s_T = \tilde{O}\left(\frac{B\sqrt{r}+r}{\sqrt{T}} + \epsilon\sqrt{r} + \epsilon\right). \tag{15}$$

This extension of error-aware GP-UCB (Bogunovic & Krause, 2021) accounts for both pointwise approximation error and its effect on fitting the surrogate to observations of f. Obtaining the guarantee requires valid bounds on B and ϵ in the enlarged confidence width, so it does not automatically apply to unmodified Bayesian UCB when the objective lies outside the model.

For the spectral features in Eq. (4), we can bound ϵ using the full sum kernel $k_+ = \sigma_\omega^2 \sum_m k_m$. If $\|f\|_{\mathcal{H}_{k_+}} \le B$, there is $g \in \mathcal{H}_{\mathrm{MvK}}$ with $\|g\|_{\mathcal{H}_{\mathrm{MvK}}} \le B$ and

$$\|f - g\|_\infty \le \epsilon, \qquad \epsilon = B \sup_{\mathbf{x}\in\mathcal{X}} \sqrt{k_+(\mathbf{x},\mathbf{x}) - \mathrm{MvK}(\mathbf{x},\mathbf{x})}. \tag{16}$$

Proposition 3 establishes this bound throughout the domain, where approximation error can remain even after every positive landmark eigenvalue is retained. Other feature constructions can use Eq. (15) with their own error bounds: the nonnegativity of $k_+(\mathbf{x},\mathbf{x}) - \mathrm{MvK}(\mathbf{x},\mathbf{x})$ in Eq. (16) follows from spectral projection specifically, and need not hold for other constructions, so ϵ must be bounded by other means in that case.

5 Related Work

Kernel design and composition. Several lines of work adapt the kernel itself to capture structure a single stationary kernel would miss. Multiple kernel learning fits a weighted sum of kernels (Gönen & Alpaydın, 2011), spectral mixture kernels parameterize the kernel's spectral density directly (Wilson & Adams, 2013), and deep kernel learning replaces the input representation with a learned embedding (Wilson et al., 2016). More recently, CAKE searches over kernel expressions with an LLM (Suwandi et al., 2025). Each of these methods ultimately fits or selects a single combined kernel, so the resulting surrogate exposes one composite view of the objective rather than each structural hypothesis's separate contribution. Additive GPs keep components separate and condition them jointly on their sum (Duvenaud et al., 2011), which is closest in spirit to our approach, but there the components index subsets of input dimensions rather than compete as alternative kernel hypotheses over the whole domain. MvBO generalizes this idea to handle multiple, possibly overlapping kernel hypotheses such that each world keeps its own finite-feature representation, and only the fitted predictions are combined in the joint posterior.

Ensembles and mixture models. A second line of work keeps each kernel's model separate but combines their predictions after fitting. EGP-TS fits one GP per kernel to the same observations, maintains a posterior weight over each, and selects queries by Thompson sampling over the ensemble (Lu et al., 2023). WBGP-BO instead averages the per-kernel posteriors through a Wasserstein barycenter (Candelieri et al., 2025). IMoGPE and RAMBO take a complementary approach, partitioning the input space into regimes, whether through input-dependent expert assignments (Rasmussen & Ghahramani, 2001) or a Dirichlet process mixture of GPs (Zhang et al., 2026), and assigning each region to a locally responsible model. In both families, each model is fit to explain the full observed response on its own, and only the fitted predictions or region assignments are combined afterward.

MVBO instead fits every world's coefficients jointly to the same observed sum, so no single world needs to explain the whole objective. One world can account for structure another leaves unexplained, and covariance between worlds emerges directly from this joint conditioning rather than from an explicit weighting or regime-assignment rule.

Finite features and BO guarantees. Our finite-feature implementation builds on a specific line of prior work, including the spectral Nyström construction used in kernel adaptive filtering (Li & Príncipe, 2026; Li et al., 2026), alongside random Fourier features as a compatible alternative map (Williams & Seeger, 2001; Rahimi & Recht, 2007). On the theoretical side, existing finite-feature BO analyses study the approximation of a single kernel, whether through quadrature Fourier features (Mutný & Krause, 2018) or uncertainty-preserving adaptive sketches (Calandriello et al., 2019). These guarantees cover one fixed finite-dimensional surrogate, while MVBO extends the argument to a sum of such surrogates, one per world, and shows that the total feature count enters the bound without an additional factor for the number of worlds.

6 EXPERIMENTS

We evaluate MVBO on three emerging scientific and AI applications including molecular conformer search, virtual screening for drug discovery, and LLM data-mixture optimization. For each, we compare MVBO against the baselines described below and examine each world's fitted contribution. A separate set of controlled experiments then isolates the effect of compressing each kernel separately versus compressing their sum, using the same spectral features throughout (Appendices B.2 and B.2.7).

6.1 APPLICATIONS

Molecular conformer optimization ($d = 12$). Following Zhang et al. (2026), we minimize the MMFF94 energy (Halgren, 1996) of pentadecane over twelve torsion angles, fixing the queried angles while relaxing the remaining geometry. Rotatable bonds admit competing conformations separated by steric barriers (Hawkins et al., 2010; Riniker & Landrum, 2015), making the search multimodal. Finding a favorable basin and refining its geometry require both broad exploration and local resolution, motivating kernels with different scales and periodic structure.

Virtual screening for drug discovery ($d = 50$). We use the Cancer-6T2W benchmark of Zhang et al. (2026), minimizing precomputed docking scores over one million molecules from ZINC15 (Sterling & Irwin, 2015) via DrugImprover (Liu et al., 2023b). Queries select library molecules represented by 50-dimensional fingerprint PCA coordinates. The challenge is to find favorable scores with few queries across chemical scaffolds whose structure–activity relationships can differ (Gómez-Bombarelli et al., 2018; Griffiths & Hernández-Lobato, 2020; Korovina et al., 2020).

LLM data-mixture optimization (DMO, $d = 6$). To study how data composition shapes LLM post-training performance (Lambert et al., 2025), we use BoLT (Chew et al., 2026) to optimize the proportions of instruction following (IF), mathematics, and code data in two training stages. The six mixture coordinates have four free parameters under simplex constraints, and maximizing the mean skill score requires balancing both skills and stages. BoLT's emulator evaluates each mixture without retraining an LLM, making this a tractable test of skill-aligned kernel assumptions.

6.2 BASELINES

We compare MVBO with single-kernel and local surrogates, kernel ensembles and learned sums, kernel search, and mixtures of GP experts. For single-kernel BO, GP-UCB and GP-UCB-M52 use squared exponential and Matérn-$5/2$ kernels, respectively (Srinivas et al., 2010). TuRBO provides a local GP baseline with adaptive trust regions (Eriksson et al., 2019). Among methods using the same set of kernels, the ensembles EGP-TS (Lu et al., 2023) and WBGP-BO (Candelieri et al., 2025) combine models through Thompson sampling with updated model probabilities and a Wasserstein barycenter, respectively. MKL-BO is a learned kernel-sum baseline with global combination weights optimized alongside GP hyperparameters (Gönen & Alpaydın, 2011). We also include CAKE for LLM-guided kernel search (Suwandi et al., 2025). The mixture-of-experts baselines are IMoGPE,

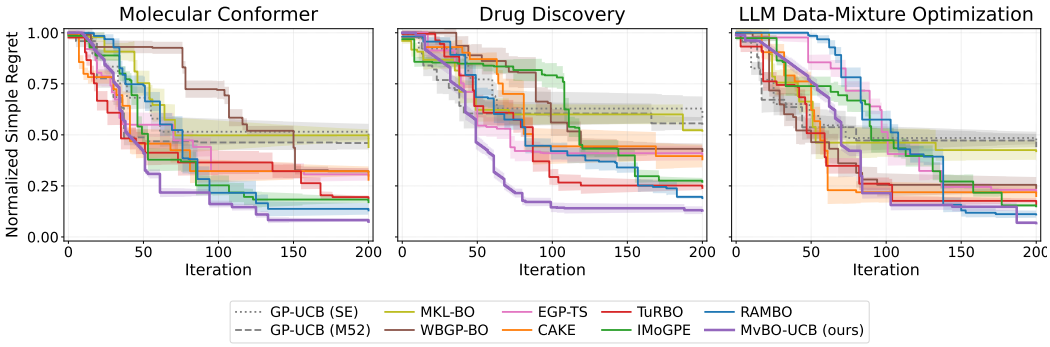


Figure 2: Application optimization trajectories, showing mean normalized simple regret and standard deviation over ten seeds (lower is better). Each task uses its corresponding objective and reference values as defined in Section 6.3 and Appendix B.2.4. Appendix B.2.4 gives the protocols.

with input-dependent expert assignments (Rasmussen & Ghahramani, 2001), and RAMBO, with a Dirichlet process mixture of GPs (Zhang et al., 2026).

6.3 Setup

The objectives for the tasks mentioned above are negative molecular energy, negative docking score, and mean DMO skill score, respectively. For all three tasks, we report normalized simple regret $\rho_j = [f_{\text{ref}} - \max_{i \le j} f(x_i)]/[f_{\text{ref}} - f_{\text{low}}]$, where $f_{\text{ref}} > f_{\text{low}}$ are fixed task-specific reference maximum and minimum values, shared across methods and seeds. Lower regret values indicate better performance. Appendix B.2.4 gives the definitions for each objective. Across ten paired seeds, each run uses $N = 200$ total iterations.

We choose kernel worlds to cover complementary structural assumptions. Molecular search uses squared exponential (SE), Matérn-$5/2$, SE $\times$ periodic kernel, and rational quadratic (RQ) kernels to represent different smoothness, periodicity, and scales. Screening uses automatic relevance determination (ARD) variants of SE, Matérn-$5/2$, Matérn-$3/2$, and RQ, giving each world a separate lengthscale per PCA coordinate. DMO combines SE and Matérn-$5/2$ with three anisotropic SE worlds, each using shorter lengthscales along one skill's mixture coordinates in both training stages. With $r_m = 10$ features per world, the total dimensions are 40, 40, and 50, respectively (Appendix B.2.3).

6.4 Results

Application results. Across all three applications, MvBO reaches the lowest displayed endpoint, but the shape of its advantage differs by task (Figure 2). On molecular conformer search ($d = 12$), MvBO pulls ahead around 50 evaluations, while single-kernel GP-UCB and MKL-BO flatten earlier at higher values and RAMBO continues improving later in the run, so the separation reflects both faster progress within the budget and a better final value. On virtual screening for drug discovery ($d = 50$), MvBO opens its largest lead between 40 and 100 evaluations, after which TuRBO closes the gap quickly before plateauing and RAMBO narrows it late, yet MvBO still ends with the lowest final normalized simple regret. On LLM data-mixture optimization ($d = 6$), the ordering changes over the course of search: CAKE and TuRBO improve faster initially and RAMBO remains competitive late, but a late improvement gives MvBO the lowest displayed endpoint, so its final advantage here does not imply uniformly faster progress.

Local component contributions. Figure 3 shows that each specialized world contributes more to the final prediction when its corresponding skill is more common in the training data. For example, the IF world has the strongest influence where IF data dominates, while the Math and Code worlds behave similarly for their respective skills. This demonstrates that MvBO automatically learns to rely on the right expert for the right task, rather than spreading contribution evenly or arbitrarily across all worlds (Appendix B.2.2).

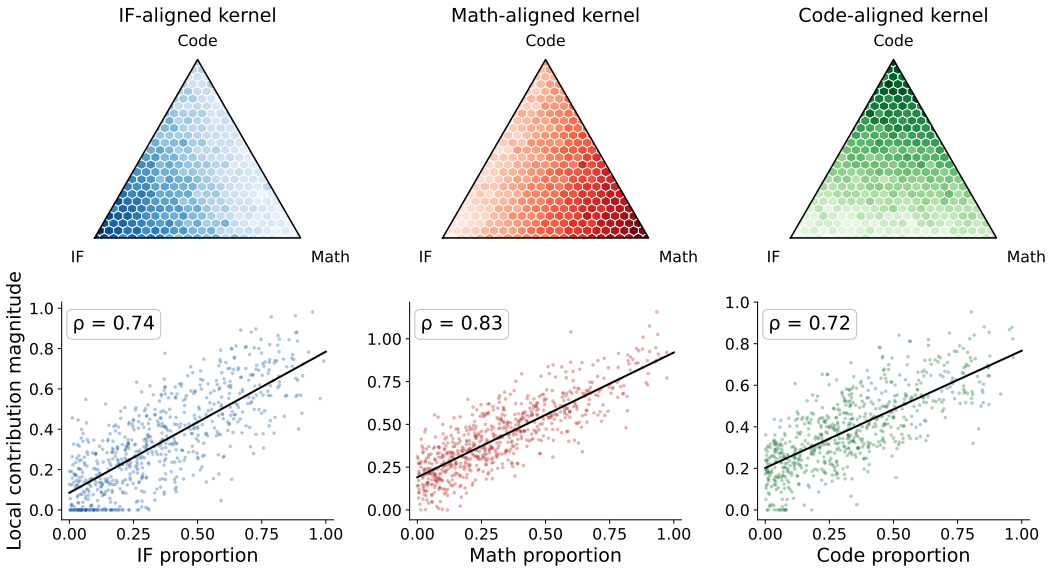


Figure 3: Local component magnitudes on DMO. The top panels show variation across a stage-1 mixture slice, while the bottom panels relate each magnitude to its skill proportion, with linear trends and Spearman correlations (IF 0.74, Math 0.83, Code 0.72). Each axis corresponds to a skill-aligned kernel specified in the model.

Additional results. We additionally test whether preserving each world's features helps when the objective's signal is concentrated in a low-amplitude kernel component, and find that separate compression substantially reduces error in that setting but not when the signal instead lies in the dominant component (Appendices B.2.7–B.2.10). We also measure how each method's posterior update time scales with the number of observations (Appendix B.1.7).

7 Conclusion

We introduce MVBO, a novel BO framework that integrates complementary kernel worlds into a single surrogate model. By retaining per-world features before jointly learning their combination, MVBO preserves independent representations for each structural hypothesis while allowing their coefficients to be learned simultaneously. This enables the acquisition function to account for dependencies between worlds and uncertainty in their aggregate prediction. Furthermore, posterior updates conditioned on these features remain computationally efficient, independent of the number of observations. Our theoretical analysis connects this representation to optimization performance, establishing a best-query regret bound under fixed feature maps. We also quantified the impact of structural approximations away from landmark points. Empirically, we demonstrate that preserving even weakly performing worlds can help locate the optimum when different regions of the search space favor different structures. However, we also find that allocating features equally across worlds is inefficient when the objective function is dominated by a few key components. More broadly, this work suggests significant value in treating kernel choice as a set of competing hypotheses, particularly for heterogeneous, multi-regime objectives where a single kernel is insufficient.

Limitations and Future Work

Our regret analysis conditions on a feature map that is fixed before any responses are observed. Extending the guarantee to features or landmarks that are chosen from data, or updated during search, remains open. Equal allocation of features across worlds is also a limitation of the current implementation. When a few components dominate, this wastes rank on weak worlds, which motivates allocating capacity by estimated world relevance rather than splitting r evenly. A separate direction is to adapt the kernel population itself as evidence accumulates, for example by using LLM-guided evolution (Suwandi et al., 2025) as an outer loop over worlds.

AI USE STATEMENT

In this work, we used generative AI tools for editing the paper to improve readability, formatting references, and summarizing, searching for, and identifying relevant literature. This manuscript assistance did not include experimental design or execution, or development of the method's mathematical formulation. We have reviewed all AI-assisted work: literature suggested by AI was verified against primary sources, readability edits were manually checked to preserve technical meaning, and reference formatting was validated. We take responsibility for the final content of this work, including text, claims, and artifacts produced with the aid of generative AI.

ETHICS STATEMENT

All authors have read and adhere to the ICLR Code of Ethics and Code of Conduct. This work does not involve human subjects, personally identifiable data, or sensitive data releases. Experiments use public benchmarks and do not present foreseeable harmful applications, dual-use risks, or privacy/security concerns beyond standard BO for scientific design. No conflicts of interest or sponsorship constraints affect the reported results, and research integrity practices (documentation of protocols, seeds, and code in the appendix/supplement) were followed.

REPRODUCIBILITY STATEMENT

We discuss reproducibility at the end of the main text and reference materials that enable replication. Theoretical assumptions and proofs for the finite-feature regret bounds, including approximation-error extensions, are given in Section 4 and Appendices A–A.2.7. Algorithm details, feature construction (Eqs. 3–4), hyperparameters, kernel populations, and complexity are in Appendix B.1. Experimental protocols for the three applications, baselines, and the controlled allocation experiment (metrics, seeds, candidate/landmark settings) are in Appendix B.2–B.2.7 with additional diagnostics and compute details in Appendices B.2.9–B.2.11. Anonymous code and reproduction scripts are provided as supplementary material.

REFERENCES

Yasin Abbasi-Yadkori, Dávid Pál, and Csaba Szepesvári. Improved algorithms for linear stochastic bandits. In *Advances in Neural Information Processing Systems*, volume 24, 2011.

Jacob Austin, Augustus Odena, Maxwell Nye, Maarten Bosma, Henryk Michalewski, David Dohan, Ellen Jiang, Carrie Cai, Michael Terry, Quoc Le, and Charles Sutton. Program synthesis with large language models. *arXiv preprint arXiv:2108.07732*, 2021.

Alain Berlinet and Christine Thomas-Agnan. *Reproducing Kernel Hilbert Spaces in Probability and Statistics*. Springer, 2004.

Ilija Bogunovic and Andreas Krause. Misspecified Gaussian process bandit optimization. In *Advances in Neural Information Processing Systems*, volume 34, 2021. URL https://papers.nips.cc/paper/2021/hash/177db6acfe388526a4c7bff88e1feb15-Abstract.html.

Daniele Calandriello, Luigi Carratino, Alessandro Lazaric, Michal Valko, and Lorenzo Rosasco. Gaussian process optimization with adaptive sketching: Scalable and no regret. In *Proceedings of the Conference on Learning Theory*, volume 99 of *Proceedings of Machine Learning Research*, pp. 533–557, 2019. URL https://proceedings.mlr.press/v99/calandriello19a.html.

Antonio Candelieri, Andrea Ponti, and Francesco Archetti. Wasserstein barycenter Gaussian process based Bayesian optimization. *arXiv preprint arXiv:2505.12471*, 2025.

Ruth Wan Theng Chew, Zhiliang Chen, Apivich Hemachandra, and Bryan Kian Hsiang Low. BoLT: A benchmark to democratize black-box optimization research for expensive LLM tasks. *arXiv preprint arXiv:2605.17000*, 2026.

Sayak Ray Chowdhury and Aditya Gopalan. On kernelized multi-armed bandits. In *Proceedings of the International Conference on Machine Learning*, pp. 844–853, 2017.

Andreas C. Damianou and Neil D. Lawrence. Deep Gaussian processes. In *Proceedings of the International Conference on Artificial Intelligence and Statistics*, volume 31 of *Proceedings of Machine Learning Research*, pp. 207–215, 2013.

David Duvenaud, James Lloyd, Roger Grosse, Joshua Tenenbaum, and Zoubin Ghahramani. Structure discovery in nonparametric regression through compositional kernel search. In *Proceedings of the International Conference on Machine Learning*, pp. 1166–1174. PMLR, 2013.

David K. Duvenaud, Hannes Nickisch, and Carl Edward Rasmussen. Additive Gaussian processes. In *Advances in Neural Information Processing Systems*, volume 24, 2011.

David Eriksson, Michael Pearce, Jacob Gardner, Ryan D Turner, and Matthias Poloczek. Scalable global optimization via local Bayesian optimization. *Advances in Neural Information Processing Systems*, 32, 2019.

Roman Garnett. *Bayesian optimization*. Cambridge University Press, 2023.

Rafael Gómez-Bombarelli, Jennifer N. Wei, David Duvenaud, José Miguel Hernández-Lobato, Benjamín Sánchez-Lengeling, Dennis Sheberla, Jorge Aguilera-Iparraguirre, Timothy D. Hirzel, Ryan P. Adams, and Alán Aspuru-Guzik. Automatic chemical design using a data-driven continuous representation of molecules. *ACS Central Science*, 4(2):268–276, 2018. doi: 10.1021/acscentsci.7b00572.

Mehmet Gönen and Ethem Alpaydın. Multiple kernel learning algorithms. *Journal of Machine Learning Research*, 12(64):2211–2268, 2011.

Ryan-Rhys Griffiths and José Miguel Hernández-Lobato. Constrained Bayesian optimization for automatic chemical design using variational autoencoders. *Chemical Science*, 11(2):577–586, 2020. doi: 10.1039/c9sc04026a.

Thomas A. Halgren. Merck molecular force field. I. basis, form, scope, parameterization, and performance of MMFF94. *Journal of Computational Chemistry*, 17(5-6):490–519, 1996. doi: 10.1002/(SICI)1096-987X(199604)17:5/6⟨490::AID-JCC1⟩3.0.CO;2-P.

Paul C. D. Hawkins, A. Geoffrey Skillman, Gregory L. Warren, Benjamin A. Ellingson, and Matthew T. Stahl. Conformer generation with OMEGA: Algorithm and validation using high quality structures from the protein databank and cambridge structural database. *Journal of Chemical Information and Modeling*, 50(4):572–584, 2010. doi: 10.1021/ci100031x.

Dan Hendrycks, Collin Burns, Saurav Kadavath, Akul Arora, Steven Basart, Eric Tang, Dawn Song, and Jacob Steinhardt. Measuring mathematical problem solving with the MATH dataset. In *Advances in Neural Information Processing Systems*, 2021.

Donald R. Jones, Matthias Schonlau, and William J. Welch. Efficient global optimization of expensive black-box functions. *Journal of Global Optimization*, 13(4):455–492, 1998. doi: 10.1023/A:1008306431147.

Ksenia Korovina, Sailun Xu, Kirthevasan Kandasamy, Willie Neiswanger, Barnabas Poczos, Jeff Schneider, and Eric Xing. ChemBO: Bayesian optimization of small organic molecules with synthesizable recommendations. In *Proceedings of the International Conference on Artificial Intelligence and Statistics*, volume 108 of *Proceedings of Machine Learning Research*, pp. 3393–3403. PMLR, 2020. URL https://proceedings.mlr.press/v108/korovina20a.html.

Nathan Lambert, Jacob Morrison, Valentina Pyatkin, Shengyi Huang, Hamish Ivison, Faeze Brahman, Lester James V. Miranda, Alisa Liu, Nouha Dziri, et al. Tulu 3: Pushing frontiers in open language model post-training. In *Proceedings of the Conference on Language Modeling*, 2025.

Greg Landrum et al. RDKit: Open-source cheminformatics. http://www.rdkit.org, 2013.

Kan Li and José C. Príncipe. Online filtering in kernel adaptive principal subspace. *Signal Processing*, 245:110552, 2026. doi: 10.1016/j.sigpro.2026.110552.

Kan Li, Benjamin Colburn, and José C. Príncipe. Multiverse kernel adaptive filtering. In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, 2026. doi: 10.1109/ICASSP55912.2026.11460575.

Jiawei Liu, Chunqiu Steven Xia, Yuyao Wang, and Lingming Zhang. Is your code generated by ChatGPT really correct? rigorous evaluation of large language models for code generation. In *Advances in Neural Information Processing Systems*, volume 36, pp. 21558–21572, 2023a.

Xuefeng Liu, Songhao Jiang, Archit Vasan, Alexander Brace, Ozan Gokdemir, Thomas Brettin, Fangfang Xia, Ian Foster, and Rick Stevens. DrugImprover: Utilizing reinforcement learning for multi-objective alignment in drug optimization. In *Advances in Neural Information Processing Systems Workshop on New Frontiers of AI for Drug Discovery and Development*, 2023b. URL https://openreview.net/forum?id=8RPbtOC6Sl.

Qin Lu, Konstantinos D Polyzos, Bingcong Li, and Georgios B Giannakis. Surrogate modeling for Bayesian optimization beyond a single Gaussian process. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(9):11283–11296, 2023.

Gustavo Malkomes and Roman Garnett. Automating Bayesian optimization with Bayesian optimization. *Advances in Neural Information Processing Systems*, 31, 2018.

Mark McGann. FRED and HYBRID docking performance on standardized datasets. *Journal of Computer-Aided Molecular Design*, 26(8):897–906, 2012. doi: 10.1007/s10822-012-9584-8.

Mojmír Mutný and Andreas Krause. Efficient high dimensional Bayesian optimization with additivity and quadrature Fourier features. In *Advances in Neural Information Processing Systems*, volume 31, 2018. URL https://papers.neurips.cc/paper_files/paper/2018/hash/4e5046fc8d6a97d18a5f54beaed54dea-Abstract.html.

Christopher J. Paciorek and Mark J. Schervish. Nonstationary covariance functions for Gaussian process regression. *Advances in Neural Information Processing Systems*, 16, 2004.

Ali Rahimi and Benjamin Recht. Random features for large-scale kernel machines. In *Advances in Neural Information Processing Systems*, volume 20, 2007.

Carl Edward Rasmussen and Zoubin Ghahramani. Infinite mixtures of Gaussian process experts. In *Advances in Neural Information Processing Systems*, volume 14, pp. 881–888, 2001.

Carl Edward Rasmussen and Christopher K. I. Williams. *Gaussian Processes for Machine Learning*. MIT Press, Cambridge, MA, 2006. ISBN 978-0-262-18253-9.

Sereina Riniker and Gregory A. Landrum. Better informed distance geometry: Using what we know to improve conformation generation. *Journal of Chemical Information and Modeling*, 55(12): 2562–2574, 2015. doi: 10.1021/acs.jcim.5b00654.

David Rogers and Mathew Hahn. Extended-connectivity fingerprints. *Journal of Chemical Information and Modeling*, 50(5):742–754, 2010. doi: 10.1021/ci100050t.

Daniel J. Russo, Benjamin Van Roy, Abbas Kazerouni, Ian Osband, and Zheng Wen. A tutorial on Thompson sampling. *Foundations and Trends in Machine Learning*, 11(1):1–96, 2018.

Bobak Shahriari, Kevin Swersky, Ziyu Wang, Ryan P. Adams, and Nando de Freitas. Taking the human out of the loop: A review of Bayesian optimization. *Proceedings of the IEEE*, 104(1): 148–175, 2016.

Jasper Snoek, Hugo Larochelle, and Ryan P. Adams. Practical Bayesian optimization of machine learning algorithms. In *Advances in Neural Information Processing Systems*, volume 25, 2012.

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

Teague Sterling and John J. Irwin. ZINC 15 – ligand discovery for everyone. *Journal of Chemical Information and Modeling*, 55(11):2324–2337, 2015. doi: 10.1021/acs.jcim.5b00559.

Richard Cornelius Suwandi, Feng Yin, Juntao Wang, Renjie Li, Tsung-Hui Chang, and Sergios Theodoridis. Adaptive kernel design for Bayesian optimization is a piece of CAKE with LLMs. In *Advances in Neural Information Processing Systems*, volume 38, 2025.

William D. Tap, Zev A. Wainberg, Stephen P. Anthony, Prabha N. Ibrahim, Chao Zhang, John H. Healey, Bartosz Chmielowski, Arthur P. Staddon, Allen Lee Cohn, Geoffrey I. Shapiro, et al. Structure-guided blockade of CSF1R kinase in tenosynovial giant-cell tumor. *New England Journal of Medicine*, 373(5):428–437, 2015. doi: 10.1056/NEJMoa1411366.

Sattar Vakili, Kia Khezeli, and Victor Picheny. On information gain and regret bounds in Gaussian process bandits. *Proceedings of the International Conference on Artificial Intelligence and Statistics*, pp. 82–90, 2021.

Christopher K. I. Williams and Matthias Seeger. Using the Nyström method to speed up kernel machines. In *Advances in Neural Information Processing Systems*, volume 13, 2001.

Andrew Gordon Wilson and Ryan Prescott Adams. Gaussian process kernels for pattern discovery and extrapolation. In *Proceedings of the International Conference on Machine Learning*, volume 28 of *Proceedings of Machine Learning Research*, pp. 1067–1075, 2013.

Andrew Gordon Wilson, Zhiting Hu, Ruslan Salakhutdinov, and Eric P. Xing. Deep kernel learning. In *Proceedings of the International Conference on Artificial Intelligence and Statistics*, volume 51 of *Proceedings of Machine Learning Research*, pp. 370–378, 2016.

An Yang, Anfeng Li, Baosong Yang, Beichen Zhang, Binyuan Hui, Bo Zheng, Bowen Yu, Chang Gao, Chengen Huang, Chen Lv, et al. Qwen3 technical report. *arXiv preprint arXiv:2505.09388*, 2025.

Feng Yin, Lishuo Pan, Tianshi Chen, Sergios Theodoridis, Zhi-Quan Luo, and Abdelhak M. Zoubir. Linear multiple low-rank kernel based stationary Gaussian processes regression for time series. *IEEE Transactions on Signal Processing*, 68:5260–5275, 2020. doi: 10.1109/TSP.2020.3023008.

Yan Zhang, Xuefeng Liu, Sipeng Chen, Sascha Ranftl, Chong Liu, and Shibo Li. Regime-adaptive Bayesian optimization via Dirichlet process mixtures of Gaussian processes. *arXiv preprint arXiv:2601.20043*, 2026.

Jeffrey Zhou, Tianjian Lu, Swaroop Mishra, Siddhartha Brahma, Sujoy Basu, Yi Luan, Denny Zhou, and Le Hou. Instruction-following evaluation for large language models. *arXiv preprint arXiv:2311.07911*, 2023.

A Mathematical Appendix

A.1 The Multiverse Kernel: Formal Properties

We first characterize the functions the multiverse surrogate can represent, then quantify what is lost by truncating each world's features. Write $\tilde{k}_m(\mathbf{x}, \mathbf{x}') = \tilde{\phi}_m(\mathbf{x})^\top \tilde{\phi}_m(\mathbf{x}')$, $\bar{k}_m = \sigma_\omega^2 \tilde{k}_m$, and $\mathrm{MvK} = \sum_m \bar{k}_m$. We suppress the retained-rank subscript on $\tilde{\phi}_m$ when unambiguous.

A.1.1 Validity and the RKHS Sum

Proposition 1 (Validity and sum structure)**.** *The multiverse kernel is positive semi-definite and has rank at most $r = \sum_m r_m$. If $\tilde{\mathcal{H}}_m$ is the RKHS of $\tilde{k}_m$, then*

$$\mathcal{H}_{\mathrm{MvK}} = \sum_m \tilde{\mathcal{H}}_m, \qquad \|g\|^2_{\mathcal{H}_{\mathrm{MvK}}} = \min_{g=\sum_m g_m} \sigma_\omega^{-2} \sum_m \|g_m\|^2_{\tilde{\mathcal{H}}_m}. \tag{17}$$

If the synthesis map $(g_m)_m \mapsto \sum_m g_m$ is injective, this sum is direct and every function has a unique component decomposition. Otherwise decompositions can differ, although the minimum-norm decomposition is still unique.

Proof. An inner product of finite features is positive semi-definite with rank bounded by its feature dimension, and positive scaling and summation preserve this property. The RKHS-sum theorem then gives the function space and its infimal-convolution norm (Berlinet & Thomas-Agnan, 2004, Sec. 6). Within the finite-dimensional product of component spaces, the constraint $\sum_m g_m = g$ is closed and affine, so strict convexity of the squared product norm gives a unique minimizer. When the synthesis map is also injective, the feasible decomposition itself is unique. □

Thus $\bigcup_m \tilde{\mathcal{H}}_m \subseteq \mathcal{H}_{\mathrm{MvK}}$, but the containment need not be strict: redundant or nested worlds may add no new functions. The same function space arises from any strictly positive weighted sum of these kernels. Worlds are therefore kernel components that may overlap, and their statistical interpretation follows from the joint posterior below.

A.1.2 Nyström Projection and Approximation Error

Let $\mathcal{H}_m$ denote the RKHS of the untruncated kernel k_m, and write $L = \{\ell_i\}_{i=1}^q$ for the landmarks independently of the number of evaluated pilot points. For a retained positive eigenpair (λ_{mj}, v_{mj}) of the *unjittered* landmark Gram matrix, define

$$u_{mj} = \lambda_{mj}^{-1/2} \sum_{i=1}^{q} (v_{mj})_i k_m(\ell_i, \cdot) \in \mathcal{H}_m. \tag{18}$$

These vectors are orthonormal since $\langle u_{mj}, u_{m\ell} \rangle = (\lambda_{mj}\lambda_{m\ell})^{-1/2} v_{mj}^\top \mathbf{K}_m v_{m\ell} = \mathbf{1}\{j = \ell\}$. Let P_m be their orthogonal projector. Then

$$\tilde{\phi}_m(\mathbf{x})_j = \langle u_{mj}, k_m(\mathbf{x}, \cdot) \rangle, \qquad \tilde{k}_m(\mathbf{x}, \mathbf{x}') = \langle P_m k_m(\mathbf{x}, \cdot), P_m k_m(\mathbf{x}', \cdot) \rangle. \tag{19}$$

In particular, $k_m - \tilde{k}_m$ is a positive semi-definite kernel and $0 \le \tilde{k}_m(\mathbf{x}, \mathbf{x}) \le k_m(\mathbf{x}, \mathbf{x})$ at every input.

Proposition 2 (Landmark truncation error)**.** *For landmark Gram matrices $\mathbf{K}_m$ and their spectral truncations $\widetilde{\mathbf{K}}_m$, with eigenvalues in nonincreasing order,*

$$\left\| \sigma_\omega^2 \sum_m (\mathbf{K}_m - \widetilde{\mathbf{K}}_m) \right\|_* = \sigma_\omega^2 \sum_m \sum_{j > r_m} \lambda_{mj}. \tag{20}$$

Proof. Every residual matrix is positive semi-definite. Its nuclear norm equals its trace, and traces add, yielding the identity. □

Because the spectral tail measures error only on the landmarks, it can vanish even when the approximation misses structure elsewhere. Connecting kernel approximation to optimization therefore requires a bound at new inputs, which the next proposition supplies for objectives in the full sum-kernel RKHS.

Proposition 3 (Uniform objective approximation)**.** *Let $k_+ = \sigma_\omega^2 \sum_m k_m$ and suppose $f \in \mathcal{H}_{k_+}$ with $\|f\|_{\mathcal{H}_{k_+}} \le B$. For the projected kernel* MvK *of Eq.* (19)*, there exists $g \in \mathcal{H}_{\mathrm{MvK}}$, $\|g\|_{\mathcal{H}_{\mathrm{MvK}}} \le B$, such that*

$$|f(\mathbf{x}) - g(\mathbf{x})| \le B\sqrt{d(\mathbf{x})}, \qquad d(\mathbf{x}) := k_+(\mathbf{x},\mathbf{x}) - \mathrm{MvK}(\mathbf{x},\mathbf{x}) \ge 0. \tag{21}$$

Consequently $\|f - g\|_\infty \le \epsilon := B \sup_{\mathbf{x}\in\mathcal{X}} \sqrt{d(\mathbf{x})}$.

Proof. Use the Hilbert direct-product feature representation $z(\mathbf{x}) = \sigma_\omega(k_1(\mathbf{x},\cdot),\dots,k_M(\mathbf{x},\cdot))$. The feature-space characterization of the RKHS gives a minimum-norm a with $f(\mathbf{x}) = \langle a, z(\mathbf{x})\rangle$ and $\|a\| \le B$. Let $P = \mathrm{diag}(P_1,\dots,P_M)$ and set $g(\mathbf{x}) = \langle Pa, Pz(\mathbf{x})\rangle$. Its coefficient norm is at most $\|a\|$, so $\|g\|_{\mathcal{H}_{\mathrm{MvK}}} \le B$. Orthogonality and Cauchy–Schwarz give

$$|f(\mathbf{x}) - g(\mathbf{x})| = |\langle (I-P)a, (I-P)z(\mathbf{x})\rangle| \le B\|(I-P)z(\mathbf{x})\| = B\sqrt{d(\mathbf{x})}. \tag{22}$$

□

Remark 1 (Scope and numerical implementation)**.** *Whereas Theorem 1 assumes the truncated prior, Proposition 3 starts from a function in the full sum-kernel RKHS and bounds its approximation error. Using that error in the regret analysis requires enlarging the UCB confidence width as in Appendix A.2.6. The projection identities in Eqs.* (18)–(21) *also assume exact spectral features, so jitter and eigenvalue clipping need separate treatment when applying quantitative projection bounds. These numerical changes leave the posterior and information-gain identities valid for the implemented fixed map.*

A.1.3 Reduction and Coordinate Invariance

Proposition 4 (Single-kernel reduction)**.** *For $M = 1$, the multiverse posterior equals GP regression with kernel $\sigma_\omega^2 \tilde{k}_1$. Its UCB acquisition is GP-UCB for that kernel.*

Proof. The weight-space model and GP have identical zero-mean Gaussian priors on every finite collection of function values and identical likelihoods. Gaussian conditioning therefore gives identical posterior distributions (Rasmussen & Williams, 2006, Ch. 2). □

At full landmark rank and invertible $\mathbf{K}_m$, $\tilde{k}_m(\mathbf{x},\mathbf{x}') = k_m(\mathbf{x},L)\mathbf{K}_m^{-1}k_m(L,\mathbf{x}')$. An alternative Nyström factorization with the same landmarks and regularization gives the same predictions. Comparisons between feature maps must therefore distinguish changes to the landmarks, retained space, hyperparameters, or numerical approximation from changes of coordinates within the same space.

A.1.4 Component Posterior and Interpretation

Let f_m be independent zero-mean GPs with kernels $\bar{k}_m$ before conditioning, and observe $y_i = \sum_m f_m(\mathbf{x}_i) + \varepsilon_i$. For observation inputs X, set $\mathbf{A} = \sum_m \bar{\mathbf{K}}_m(X,X) + \sigma_\varepsilon^2\mathbf{I}$. Gaussian conditioning yields

$$\mathbb{E}[f_m(\mathbf{x}) \mid \mathcal{D}] = \bar{k}_m(\mathbf{x},X)\mathbf{A}^{-1}\mathbf{y} = c_m(\mathbf{x}), \tag{23}$$

$$\mathrm{Cov}(f_m(\mathbf{x}), f_\ell(\mathbf{x}') \mid \mathcal{D}) = \mathbf{1}\{m=\ell\}\bar{k}_m(\mathbf{x},\mathbf{x}') - \bar{k}_m(\mathbf{x},X)\mathbf{A}^{-1}\bar{k}_\ell(X,\mathbf{x}'). \tag{24}$$

Summing over m,ℓ recovers the joint predictive variance and shows how conditioning introduces dependence between initially independent components. For fixed kernels and noise, this covariance depends on the observation locations but not on their values $\mathbf{y}$.

Joint regression and prediction averaging. The posterior mean also minimizes the joint regularized least-squares objective,

$$\boldsymbol{\mu}_\omega = \arg\min_{\boldsymbol{\omega}} \frac{1}{2\sigma_\varepsilon^2} \sum_{(\mathbf{x}_i,y_i)\in\mathcal{D}} \left(y_i - \sum_m \boldsymbol{\omega}^{(m)\top} \tilde{\phi}_{r_m}^{(m)}(\mathbf{x}_i) \right)^2 + \frac{1}{2\sigma_\omega^2} \sum_m \|\boldsymbol{\omega}^{(m)}\|^2. \tag{25}$$

Differentiating this quadratic gives Eqs. (5) and (6) and makes the distinction from prediction averaging explicit. Although both approaches produce sums of functions, prediction averaging fits each model to the responses under its own prior before combining predictions. Joint regression fits their sum to the observations, allowing components to explain the responses together and producing covariance terms that enter the acquisition uncertainty through Eq. (24).

Coordinate invariance. Rotating each block by an orthogonal matrix Q_m leaves its kernel unchanged, while $Q = \mathrm{diag}(Q_m)$ transforms the posterior as $\mu'_\omega = Q\mu_\omega$ and $\Sigma'_\omega = Q\Sigma_\omega Q^\top$. Both the total prediction and each $c_m(\mathbf{x})$ are therefore independent of eigenvector signs and rotations within retained spaces. Changing kernel amplitudes, duplicating worlds, or moving directions between blocks can instead redistribute explanatory credit because those operations alter the component priors. For an MKL kernel, the analogous component prediction follows by replacing $\bar{k}_m$ with $\eta_m k_m$.

Duplicating a component. Replace a component kernel $\bar{k}_m$ by two independent components with kernels $a\bar{k}_m$ and $(1-a)\bar{k}_m$, where $0 < a < 1$. Their sum leaves A and the total posterior unchanged, but Eq. (23) splits the conditional mean into ac_m and $(1-a)c_m$. Identical predictions can therefore distribute contribution magnitudes differently across world labels, making their interpretation conditional on the chosen kernels and prior amplitudes, including when retained spaces overlap.

A.1.5 Separate Truncation Versus Truncating the Sum

The order of truncation and summation determines which functions survive at a fixed feature budget. To see this directly, use all three points of a finite domain as landmarks and take

$$K_1 = \mathrm{diag}(4, 3, 0), \qquad K_2 = \mathrm{diag}(0, 0, 1), \qquad \sigma_\omega^2 = 1. \tag{26}$$

Retaining one direction per world gives $\mathrm{diag}(4, 0, 1)$, whereas retaining the leading two directions of $K_1 + K_2$ gives $\mathrm{diag}(4, 3, 0)$. Both representations have rank two. Separate allocation retains the function vector $(0, 0, 1)^\top$ and discards $(0, 1, 0)^\top$, while aggregate truncation makes the opposite choice. The discarded trace is 3 for separate allocation and 1 for aggregate truncation. Thus better spectral approximation can exclude a low-amplitude direction that matters to the objective. Neither retained space contains the other, and this example establishes no uniform optimization advantage. The useful allocation depends on the objective's alignment with the retained directions, as tested in Appendix B.2.7.

A.2 Proofs for Section 4

For the following proofs define $\psi(\mathbf{x}) = \sigma_\omega \tilde{\mathbf{\Phi}}_r(\mathbf{x})$, $\lambda = \sigma_\varepsilon^2$, and $V_n = \lambda I + \sum_{i=1}^n \psi(\mathbf{x}_i)\psi(\mathbf{x}_i)^\top$. With the first n_0 observations forming the initial design, write $n_t = n_0 + t - 1$ for the number available before query t. In this parameterization,

$$\widehat{\theta}_n = V_n^{-1} \sum_{i=1}^n \psi(\mathbf{x}_i) y_i, \quad \mu_n(\mathbf{x}) = \psi(\mathbf{x})^\top \widehat{\theta}_n, \quad \sigma_n^2(\mathbf{x}) = \lambda \psi(\mathbf{x})^\top V_n^{-1} \psi(\mathbf{x}). \tag{27}$$

Writing $J_n = \frac{1}{2}\log(\det V_n/\lambda^r)$ with $J_0 = 0$, we treat this as a realized log-determinant statistic when inputs are adaptive. The proof uses matrix identities for each realized sequence, without equating J_n to the unconditional mutual information of a random design.

A.2.1 Proof of Lemma 1

For any sequence of T inputs, let Ψ be the matrix with rows $\psi(\mathbf{x}_i)^\top$. Sylvester's identity gives

$$\frac{1}{2}\log\det(I_T + \lambda^{-1}\Psi\Psi^\top) = \frac{1}{2}\log\det(I_r + \lambda^{-1}\Psi^\top\Psi). \tag{28}$$

If $\zeta_1, \ldots, \zeta_q$ are its nonzero feature-Gram eigenvalues, then $q \le r$ and $\sum_j \zeta_j \le T\kappa^2$. Concavity gives

$$\frac{1}{2}\sum_{j=1}^q \log(1 + \zeta_j/\lambda) \le \frac{q}{2}\log\left(1 + \frac{T\kappa^2}{q\lambda}\right) \le \frac{r}{2}\log\left(1 + \frac{T\kappa^2}{r\lambda}\right). \tag{29}$$

The last inequality follows because $a\log(1 + c/a)$ increases with $a > 0$, while a zero feature-Gram matrix gives log-determinant zero directly. Maximizing over input sequences proves the result and its refinement using the effective feature dimension.

A.2.2 VARIANCE SUM AFTER INITIALIZATION

Let $v_t = \sigma^2_{n_t}(\mathbf{x}_t) \le \kappa^2$. The determinant lemma gives the pathwise identity

$$\frac{1}{2}\sum_{t=1}^{T}\log(1+v_t/\lambda) = \frac{1}{2}\log\frac{\det V_{n_0+T}}{\det V_{n_0}} \le \gamma_T. \tag{30}$$

For the inequality, put $S = \sum_{t=1}^{T}\psi(\mathbf{x}_t)\psi(\mathbf{x}_t)^\top$. Since $V_{n_0} \succeq \lambda I$, Sylvester's identity and positive semi-definite order bound the determinant ratio by $\det(I+\lambda^{-1}S)$. This is bounded by the maximum defining γ_T, even if the selected sequence is adaptive. The function $v/\log(1+v/\lambda)$ is nondecreasing on $v \ge 0$. Therefore

$$\sum_{t=1}^{T} v_t \le \frac{2\kappa^2}{\log(1+\kappa^2/\lambda)}\gamma_T = \frac{C_1}{4}\gamma_T. \tag{31}$$

Initialization may reduce uncertainty and does not require replacing γ_T by γ_{n_0+T} in this variance-sum bound. The number of worlds M enters C_1 only through the feature-norm bound κ^2 of Assumption 4, which grows as more worlds are concatenated, rather than through an explicit factor of M.

A.2.3 PROOF OF THEOREM 1

Under the stated generative model, the posterior standardized error at each fixed input is standard Gaussian conditional on the observed history. Zero posterior variance gives zero error almost surely. The Gaussian tail bound $P(|Z| > u) \le e^{-u^2/2}$ for $u \ge 0$, followed by a union bound over $\mathcal{X}$ and query rounds, yields the simultaneous event

$$|f(\mathbf{x}) - \mu_{n_t}(\mathbf{x})| \le \sqrt{\beta_t}\sigma_{n_t}(\mathbf{x}) \quad \text{for all } \mathbf{x} \in \mathcal{X} \text{ and admissible } t \tag{32}$$

with probability at least $1-\delta$, since $\sum_{t\ge1}|\mathcal{X}|e^{-\beta_t/2} = \delta$.

Conditional on this event, fix an admissible horizon T. If an optimum belongs to $S_0 \cup \{\mathbf{x}_1, \dots, \mathbf{x}_T\}$, then $s_T = 0$. Otherwise an optimum $\mathbf{x}^*$ outside this set remains eligible at every round $t \le T$, so exact UCB selection gives

$$f(\mathbf{x}^*) \le \mu_{n_t}(\mathbf{x}^*) + \sqrt{\beta_t}\sigma_{n_t}(\mathbf{x}^*) \le \mu_{n_t}(\mathbf{x}_t) + \sqrt{\beta_t}\sigma_{n_t}(\mathbf{x}_t). \tag{33}$$

Combining the confidence interval at $\mathbf{x}_t$ gives $f(\mathbf{x}^*) - f(\mathbf{x}_t) \le 2\sqrt{\beta_t}\sqrt{v_t}$. Hence

$$s_T \le \min_{t\le T}[f(\mathbf{x}^*) - f(\mathbf{x}_t)] \le \frac{2}{T}\sum_{t=1}^{T}\sqrt{\beta_t}\sqrt{v_t} \tag{34}$$

$$\le 2\sqrt{\frac{\beta_T}{T}\sum_{t=1}^{T} v_t} \le \sqrt{\frac{C_1\beta_T\gamma_T}{T}}. \tag{35}$$

Because the confidence event covers every admissible T, the result holds throughout the search and becomes exactly zero when the domain is exhausted. It does not bound cumulative regret after an optimum is excluded from future queries or assert an infinite-horizon limit on a fixed finite domain without repeated queries.

Noisy recommendation. The theorem concerns the best *latent* value among evaluated points. Algorithm 1 returns the queried maximizer of the posterior mean $\mu_{\mathrm{mv},T}$. Under noiseless evaluations, this ranking agrees with $\arg\max y_i$ whenever the fitted mean preserves the observed order, and the best observed value equals the best latent queried value. Under noise, $\arg\max y_i$ can select a positive outlier, so the posterior-mean rule is the recommendation used by the algorithm. The bound on s_T still refers to the best latent queried value, not to a separate guarantee for that recommendation.

A.2.4 FIXED-FUNCTION CONFIDENCE AND SIMPLE REGRET

Suppose instead that $f \in \mathcal{H}_{\mathrm{MvK}}$, $\|f\|_{\mathcal{H}_{\mathrm{MvK}}} \le B$, is fixed. The feature map is fixed before the noise used in regression, and each input is predictable relative to a filtration in which ε_i is conditionally

R-sub-Gaussian. The posterior formulas in Eq. (27) are used as a regularized linear estimator with $\lambda > 0$, without requiring a Bayesian draw of f. Define

$$b_n = B + \frac{R}{\sqrt{\lambda}}\sqrt{2\left(J_n + \log\frac{1}{\delta}\right)}. \tag{36}$$

Then with probability at least $1-\delta$, simultaneously for all n and all $\mathbf{x}$,

$$|f(\mathbf{x}) - \mu_n(\mathbf{x})| \le b_n \sigma_n(\mathbf{x}). \tag{37}$$

To prove this, take a minimum-norm coefficient θ for f under ψ, so $\|\theta\| \le B$. The self-normalized martingale inequality (Abbasi-Yadkori et al., 2011; Chowdhury & Gopalan, 2017) gives

$$\left\|\sum_{i=1}^{n}\psi(\mathbf{x}_i)\varepsilon_i\right\|_{V_n^{-1}} \le R\sqrt{2(J_n + \log(1/\delta))} \tag{38}$$

uniformly in n. Since $\widehat{\theta}_n - \theta = V_n^{-1}(\sum_i \psi(\mathbf{x}_i)\varepsilon_i - \lambda\theta)$ and $\|\lambda\theta\|_{V_n^{-1}} \le \sqrt{\lambda}B$, Cauchy–Schwarz with Eq. (27) proves Eq. (37). The factor $R/\sqrt{\lambda}$ accounts for the scaling of posterior standard deviation. For Gaussian noise with $R = \sigma_\varepsilon$ and $\lambda = \sigma_\varepsilon^2$, it equals one.

Use $\mu_{n_t} + b_{n_t}\sigma_{n_t}$ as the UCB score. Since J_n is nondecreasing, the proof of Theorem 1 gives

$$s_T \le b_{n_0+T-1}\sqrt{\frac{C_1\gamma_T}{T}}. \tag{39}$$

Here $J_{n_0+T-1} \le \gamma_{n_0+T-1}$, so the confidence width counts initial observations even though the variance sum only needs γ_T. For fixed initialization and scales, this yields the rate in Corollary 1. Applying this result to a bounded-RKHS objective requires using b_{n_t} in the acquisition in place of the Bayesian width $\sqrt{\beta_t}$.

Fitting the representation on an initial pilot. When the initial responses are used to select the features, conditioning on those features can change the distribution of the initial noise. One way to retain valid confidence intervals is to select the map on a pilot, freeze it, and fit the regression to fresh observations only. Conditional on the pilot, the bound then holds whenever the selected representation admits the stated norm bound, using n for the fresh observations. The pilot points may remain in the evaluated set used to define s_T. Using the pilot responses in the regression instead requires a separate initial-error correction, provided in Appendix A.2.7.

Remark 2 (Coverage of distinct retained components)**.** *If the synthesis map $(g_m)_m \mapsto \sum_m g_m$ is injective, a function with at least two nonzero retained components belongs to $\mathcal{H}_{\mathrm{MvK}}$ but to no individual $\tilde{\mathcal{H}}_m$, by uniqueness of its decomposition. Under the fixed-function assumptions above, its joint norm supplies the bound in Eq.* (39)*. Injectivity is not assumed by the general construction. This observation concerns the truncated spaces and establishes neither exclusion from an untruncated RKHS nor a regret lower bound for another optimizer.*

A.2.5 Cumulative Regret When Repeated Queries Are Allowed

Consider a variant that maximizes the same UCB over all of $\mathcal{X}$ at every round, including previously queried points. Under the Bayesian assumptions, except for the no-repeat restriction, the global optimum remains feasible. Thus on the confidence event,

$$R_T := \sum_{t=1}^{T}[f(\mathbf{x}^*) - f(\mathbf{x}_t)] \le 2\sum_{t=1}^{T}\sqrt{\beta_t v_t} \le \sqrt{C_1 T\beta_T\gamma_T} = \tilde{O}(\sqrt{rT}). \tag{40}$$

Equation (2) already allows repeated inputs, so the variance-sum proof is unchanged. This variant admits arbitrary horizons on a fixed finite domain and hence an average-regret limit. Under the fixed-function assumptions, replacing $\sqrt{\beta_t}$ by b_{n_t} instead gives $R_T \le b_{n_0+T-1}\sqrt{C_1 T\gamma_T}$. Neither cumulative statement is asserted for Algorithm 1's no-repeat policy.

A.2.6 Approximation-Aware Confidence and Optimization

Let the actual objective satisfy $f = g + h$, with $g \in \mathcal{H}_{\mathrm{MvK}}$, $\|g\|_{\mathcal{H}_{\mathrm{MvK}}} \le B$, and $\|h\|_\infty \le \epsilon$. Keep the fixed-map and predictable-noise assumptions of Appendix A.2.4. Observations are of f, not of g. For the same linear estimator define

$$b_n^{\mathrm{app}} = B + \frac{R}{\sqrt{\lambda}}\sqrt{2(J_n + \log(1/\delta))} + \epsilon\sqrt{n/\lambda}. \tag{41}$$

With probability at least $1 - \delta$, uniformly in $n, \mathbf{x}$,

$$|f(\mathbf{x}) - \mu_n(\mathbf{x})| \le \epsilon + b_n^{\mathrm{app}}\sigma_n(\mathbf{x}). \tag{42}$$

Indeed, write Ψ_n for the observation feature matrix and $h_n = (h(\mathbf{x}_i))_{i=1}^n$. Relative to the well-specified fixed-function proof, the coefficient error gains $V_n^{-1}\Psi_n^\top h_n$. Since $\Psi_n V_n^{-1}\Psi_n^\top \preceq I_n$,

$$\|\Psi_n^\top h_n\|_{V_n^{-1}} \le \|h_n\|_2 \le \epsilon\sqrt{n}. \tag{43}$$

The same Cauchy–Schwarz argument adds $\epsilon\sqrt{n/\lambda}\sigma_n(\mathbf{x})$ to the confidence width for g, and $|h(\mathbf{x})| \le \epsilon$ gives Eq. (42).

Using the optimistic score $\mu_{n_t} + b_{n_t}^{\mathrm{app}}\sigma_{n_t} + \epsilon$ gives

$$s_T \le 2\epsilon + b_{n_0+T-1}^{\mathrm{app}}\sqrt{\frac{C_1\gamma_T}{T}}. \tag{44}$$

The input-independent ϵ can be omitted from the maximization. For the proof, if an optimum has been queried the regret is zero. Otherwise it stays feasible and $f(\mathbf{x}^*) - f(\mathbf{x}_t) \le 2\epsilon + 2b_{n_t}^{\mathrm{app}}\sigma_{n_t}(\mathbf{x}_t)$ at every preceding round. Average and use Eq. (31).

Proposition 3 supplies one sufficient value of ϵ when the true objective belongs to the full sum-kernel RKHS. At fixed scales and initialization, the bound is of order

$$\tilde{O}\left(\frac{B\sqrt{r} + r}{\sqrt{T}} + \epsilon\sqrt{r} + \epsilon\right). \tag{45}$$

The first term decreases as observations accumulate, while the terms involving ϵ capture error that additional observations may not resolve at fixed rank. Increasing rank can reduce this residual, at the cost of estimating more coefficients: because the irreducible term $\epsilon\sqrt{r}$ grows with r, this only shrinks the overall bound once ϵ itself decays faster than $1/\sqrt{r}$. Otherwise added capacity trades one source of error for another. Using the bound requires valid upper bounds on B and ϵ in the enlarged acquisition width. The unmodified Bayesian acquisition has no such correction, and adapting the representation during the run requires a separate analysis. Nyström variance also tends to be underestimated far from the landmarks, which can under-explore those regions under an unmodified UCB. The correction above accounts for this in the bound, and in practice we additionally rely on broad, uniform landmark coverage of the candidate domain (Appendix B.2.7) to limit how far candidates fall from the landmark set.

A.2.7 Retaining a Data-Dependent Pilot

In empirical-Bayes fitting, initial observations both select the representation and fit its coefficients, so reusing them requires an explicit bound on their noise contribution. With this contribution separated, the following result applies the martingale argument to fresh observations, whose noise remains centered conditional on the pilot.

Proposition 5 (Confidence after a data-dependent pilot)**.** *Let a pilot of n_0 observations determine a fixed feature map ψ and a fixed $\lambda > 0$. Suppose that, almost surely, the actual objective admits $f = g + h$ in the selected representation with $\|g\|_{\mathcal{H}_{\mathrm{MvK}}} \le B$ and $\|h\|_\infty \le \epsilon$, for deterministic bounds B, ϵ. Each pilot noise has an R-sub-Gaussian tail, and fresh noises are conditionally R-sub-Gaussian given the pilot and subsequent history. Fresh inputs are predictable. With Ψ_0 the pilot feature matrix, define*

$$W_0 = \lambda I + \Psi_0^\top\Psi_0, \qquad W_n = W_0 + \sum_{i=1}^n \psi(\mathbf{x}_i)\psi(\mathbf{x}_i)^\top, \tag{46}$$

$$J_{n|0} = \tfrac{1}{2}\log(\det W_n / \det W_0), \qquad Q_0 = R\sqrt{2n_0 \log(2n_0/\delta_0)}. \tag{47}$$

Set $Q_0 = 0$ when $n_0 = 0$, and choose $\delta_0, \delta_1 > 0$ with $\delta_0 + \delta_1 < 1$. Fit the regularized estimator to all pilot and fresh observations, with variance $\sigma_n^2(\mathbf{x}) = \lambda\psi(\mathbf{x})^\top W_n^{-1}\psi(\mathbf{x})$. Then, with probability at least $1 - \delta_0 - \delta_1$, simultaneously for all $n, \mathbf{x}$,

$$|f(\mathbf{x}) - \mu_n(\mathbf{x})| \le \epsilon + b_n^{\text{pilot}}\sigma_n(\mathbf{x}), \tag{48}$$

$$b_n^{\text{pilot}} = B + \frac{Q_0}{\sqrt{\lambda}} + \frac{R}{\sqrt{\lambda}}\sqrt{2(J_{n|0} + \log(1/\delta_1))} + \epsilon\sqrt{(n_0+n)/\lambda}. \tag{49}$$

Proof. A union bound on the pilot noise coordinates gives $\|\varepsilon_0\|_2 \le Q_0$ except on an event of probability δ_0. This bound does not require the selected features to be independent of those noises. Since $W_n \succeq W_0$,

$$\|\Psi_0^\top \varepsilon_0\|_{W_n^{-1}} \le \|\varepsilon_0\|_2 \le Q_0.$$

Conditional on the pilot, W_0 is fixed and positive definite. The self-normalized inequality with initial matrix W_0 bounds the fresh noise contribution uniformly in n by $R\sqrt{2(J_{n|0} + \log(1/\delta_1))}$ in the W_n^{-1} norm. The regularization contribution has norm at most $\sqrt{\lambda}B$. The residual contribution from all $n_0 + n$ observations has norm at most $\epsilon\sqrt{n_0+n}$, by the same contraction argument as in Appendix A.2.6. Adding these bounds and using $\|\psi(\mathbf{x})\|_{W_n^{-1}} = \sigma_n(\mathbf{x})/\sqrt{\lambda}$ proves the claim, including the pointwise residual $|h(\mathbf{x})| \le \epsilon$. Finally, integrate the conditional fresh-noise failure probability over the pilot and take a union bound. □

For no-repeat UCB on the fixed finite domain with this corrected width, the same best-query argument yields

$$s_T \le 2\epsilon + b_{T-1}^{\text{pilot}}\sqrt{C_1\gamma_T/T}. \tag{50}$$

Here $\kappa^2 = \sup_{\mathbf{x}} \|\psi(\mathbf{x})\|^2 > 0$ and both C_1 and γ_T use the selected regularization λ in place of σ_ε^2: specifically, $C_1 = 8\kappa^2/\log(1+\kappa^2/\lambda)$ and $\gamma_T = \max_{X_T} \frac{1}{2}\log\det(I + K_{X_T}/\lambda)$ for the selected fixed kernel, allowing repeated inputs in this maximum. The conditional variance-sum bound still holds because $W_0 \succeq \lambda I$. The result accounts for response-dependent initial fitting, but does not ensure that the selected representation satisfies a useful norm or approximation bound. It also requires the corrected acquisition width and does not justify interpreting a fitted kernel as an exogenous Bayesian prior.

B Method and Experimental Details

B.1 Method Details

B.1.1 Alternative Finite-Feature Constructions

The posterior, acquisition, and finite-rank information-gain bound only require a fixed finite map. For example, let a continuous real stationary kernel have the normalized spectral representation

$$k_m(x, x') = a_m\, \mathbb{E}_{\omega\sim p_m}[\cos(\omega^\top(x - x'))], \qquad a_m = k_m(0,0) > 0.$$

Draw $\omega_{mj} \sim p_m$ and $b_{mj} \sim \text{Unif}[0, 2\pi]$ independently and set

$$[\tilde{\phi}_{r_m}^{(m)}(x)]_j = \sqrt{2a_m/r_m}\cos(\omega_{mj}^\top x + b_{mj}), \qquad j = 1, \dots, r_m. \tag{51}$$

The expected feature kernel equals k_m (Rahimi & Recht, 2007), and freezing the draws gives a valid MvBO map with $\|\tilde{\phi}_{r_m}^{(m)}(x)\|^2 \le 2a_m$. Theorem 1 then applies under its finite-feature prior assumptions, although unbiased kernel approximation alone neither places a fixed objective in the sampled span nor bounds its uniform approximation error. Because the residual kernel need not be positive semi-definite, the spectral bound in Proposition 3 also requires a separate argument for random features. This construction shows how other features fit the method, while all numerical results use Eq. (4).

Table 1: Posterior computation for given kernels and features. n: observations, r: total feature dimension, $|C|$: candidate count. A full refit rebuilds the posterior from all observations. Incremental updates reuse the previous factorization with fixed kernels and features. Prediction costs refer to a fresh candidate batch without cached moments. Kernel/feature evaluation, basis construction, and hyperparameter optimization are additional costs.

Method	Full refit	Incremental update	Predict on C	Posterior memory
Exact GP (single or sum kernel)	$O(n^3)$	$O(n^2)$	$O(n^2\|C\|)$	$O(n^2)$
Exact-GP ensemble	$O(Mn^3)$	$O(Mn^2)$	$O(Mn^2\|C\|)$	$O(Mn^2)$
MVBO	$O(nr^2 + r^3)$	$O(r^2)$	$O(r^2\|C\|)$	$O(r^2)$

B.1.2 Rank Selection

For a world's landmark Gram matrix, define retained spectral energy

$$\rho_m(r_m) = \frac{\sum_{j \le r_m} \lambda_{mj}}{\operatorname{tr}(\mathbf{K}_m)}. \tag{52}$$

For a nonzero Gram matrix, Proposition 2 gives

$$\|\mathbf{K}_m - \widetilde{\mathbf{K}}_m\|_* = (1 - \rho_m(r_m)) \operatorname{tr}(\mathbf{K}_m). \tag{53}$$

One possible budget rule, for target $\rho \in (0, 1)$, is

$$r_m(\rho) = \min\{r : 1 \le r \le \operatorname{rank}(\mathbf{K}_m),\ \rho_m(r) \ge \rho\}. \tag{54}$$

This rule controls error on the landmarks, whereas our experiments use fixed ranks. In particular, $n_0 = r_m = 10$ retains every positive landmark eigenvector when the matrix has full rank, but zero error at those inputs does not imply a small residual elsewhere. Objective approximation therefore requires the domain-wide quantity in Eq. (21). A zero Gram matrix contributes no nonzero direction and has an undefined energy ratio.

B.1.3 Numerical Stability

When $r_m < n_0$, either a truncated eigensolver or a full eigendecomposition can be used, with numerical safeguards including jitter $\epsilon_{\text{jit}} = 10^{-6} \operatorname{tr}(\mathbf{K}_m)/q$ and clipping retained eigenvalues below 10^{-10} before inverse-square-root scaling. These operations define a slightly different finite feature map from the ideal unjittered projector. Posterior and information-gain formulas apply to that implemented map, but exact spectral-tail identities refer to the original matrix. Cholesky solves avoid explicit inversion of the posterior precision matrix.

B.1.4 Hyperparameter Fitting

The base controlled experiment uses predeclared kernel parameters. The pilot-fitting experiment jointly learns component amplitudes and lengthscales from the shared initial design, then freezes them with the eigenspaces (Appendix B.2.10). The application reproduction configuration uses the pilot-only fitting procedure specified in Appendix B.2.4. A change to kernel parameters requires recomputing the corresponding feature map. Reusing pilot responses also enters the confidence analysis because those responses selected the model. Appendix A.2.4 gives a pilot/fresh-data split, and Appendix A.2.7 gives a correction for retaining the pilot under the stated fixed-function assumptions.

B.1.5 Complexity Breakdown

With a frozen map, we maintain the posterior precision and right-hand side through rank-one updates and triangular solves, requiring $O(r^2)$ time per observation instead of the $O(nr^2 + r^3)$ cost of rebuilding from all data. An exact GP with fixed kernels can likewise update its Cholesky factor incrementally in $O(n^2)$ time, while a full dense factorization costs $O(n^3)$ whenever the Gram factor is rebuilt, such as after hyperparameter fitting. Table 1 distinguishes these full refits from incremental updates.

Table 2: Marginal (per-observation) conditioning-update cost as n grows, isolated from the rest of the BO pipeline. Mean±SD over 3 seeds, single-threaded BLAS. Ratio is exact divided by feature.

n	Exact (ms)	Feature, r=40 (ms)	Ratio
100	0.0083±0.0001	0.0096±0.0002	0.9×
200	0.020±0.0035	0.0095±0.0000	2.1×
500	0.079±0.0081	0.0094±0.0001	8.4×
1000	0.789±0.052	0.0094±0.0002	83.8×
2000	3.40±0.114	0.0091±0.0002	372.5×
4000	19.54±9.97	0.0093±0.0002	2109.5×
8000	111.6±96.09	0.0093±0.0001	11955.4×

With q landmarks, the one-time spectral cost is $O(\sum_m q^2 r_m)$ for the stated partial decomposition, or $O(Mq^3)$ for full decompositions. Evaluating features on C additionally costs $O(qr|C|)$ plus the cost of the $Mq|C|$ kernel evaluations. The maps and landmarks require $O(qr + qd)$ additional storage, with candidate-feature caching offering a further trade-off between memory and evaluation time. These spectral construction costs, together with prediction and acquisition optimization, must be included when relating the $O(r^2)$ posterior update to full BO runtime.

For comparison with feature ensembles, M independent models with D features each need $O(MD^2)$ update work. A dense joint posterior with $r = MD$ features needs $O(M^2D^2)$ work and retains cross-component covariance. Joint fitting and independent feature fitting therefore have different computational costs even at the same total coordinate budget. The fixed-feature versus exact-GP comparison above concerns how inference scales with accumulated observations.

B.1.6 Fixed Representation and Basis Refresh

We use the exact sum-kernel GP as an accuracy reference because retaining full rank on a fixed landmark set recovers the full Nyström approximation, which can still differ from the original kernel elsewhere. Refreshing the basis changes this represented space and requires rebuilding the posterior from retained observations, adding a cost that depends on refresh frequency. Such changes to the basis or kernel population fall outside Theorem 1's fixed-representation assumptions.

B.1.7 Accuracy and Computation

Bounding the posterior dimension changes the cost of a full rebuild from $O(n^3)$ for an exact dense GP to $O(nr^2 + r^3)$ for the feature model. With fixed kernels, the corresponding incremental updates cost $O(n^2)$ and $O(r^2)$, so a fixed feature dimension improves asymptotic conditioning cost as observations accumulate. Actual runtime also includes basis construction, candidate prediction, and acquisition optimization, whose costs depend on the caching strategy.

To examine this scaling, we measured per-observation posterior updates on synthetic positive semi-definite kernel matrices as n increased from 100 to 8,000, keeping $r = 40$ feature coordinates fixed. Measurements use one worker and single-threaded BLAS and report mean±SD over 3 seeds. By isolating conditioning, this benchmark measures the $O(n^2)$-versus-$O(r^2)$ step separately from the setup, candidate-moment refresh, and acquisition costs included in the full 200-query workload of Table 9. Table 2 reports the result: the feature update's marginal cost stays flat at about 0.009ms across this range, while the exact update grows from 0.008ms at $n = 100$ to 111.6ms at $n = 8{,}000$, roughly consistent with quadratic growth. The crossover in per-observation update cost falls near $n \approx 200$ in this particular microbenchmark. By $n = 8{,}000$, the feature update is about four orders of magnitude faster per observation. These measurements confirm the conditioning-step scaling in the tested implementation. They do not locate an end-to-end BO crossover, which also depends on representation setup, candidate prediction, and acquisition optimization.

Table 3: Simplex-constrained search space for DMO (Chew et al., 2026). Effective free parameters: 4.

Coord.	Meaning	Range	Constraint
x_1	IF proportion (stage 1)	$[0, 1]$	$\sum_{i=1}^{3} x_i = 1$
x_2	Math proportion (stage 1)	$[0, 1]$	$\sum_{i=1}^{3} x_i = 1$
x_3	Code proportion (stage 1)	$[0, 1]$	$\sum_{i=1}^{3} x_i = 1$
x_4	IF proportion (stage 2)	$[0, 1]$	$\sum_{j=4}^{6} x_j = 1$
x_5	Math proportion (stage 2)	$[0, 1]$	$\sum_{j=4}^{6} x_j = 1$
x_6	Code proportion (stage 2)	$[0, 1]$	$\sum_{j=4}^{6} x_j = 1$

B.2 Experimental Details

B.2.1 Application Protocols

Molecular conformer optimization (d=12). Following Zhang et al. (2026), we minimize the MMFF94 potential energy (Halgren, 1996) of a pentadecane chain ($C_{15}H_{32}$) over d=12 internal dihedral angles $\theta = (\theta_1, \ldots, \theta_{12})$, using f=−energy for BO. Each C–C bond favors three orientations, 180° (anti) and $\pm 60^\circ$ (gauche), inducing on the order of $3^{12} = 531{,}441$ conformational minima separated by high-energy steric barriers (Hawkins et al., 2010; Riniker & Landrum, 2015; Zhang et al., 2026). Each basin is locally smooth, but neighboring basins have incommensurable curvature, so the landscape is multi-regime rather than gradually non-stationary (Zhang et al., 2026). To evaluate a configuration θ, we construct the molecular backbone in RDKit (Landrum et al., 2013) and run a constrained geometry optimization that fixes the queried dihedrals while relaxing all other degrees of freedom (bond lengths, angles, and Cartesian coordinates). This yields a chemically valid energy surface rather than a rigid-rotor approximation with unrealistic steric clashes.

Virtual screening for drug discovery (d=50). We use the Cancer-6T2W docking benchmark of Zhang et al. (2026), which targets the CSF1R kinase domain (PDB 6T2W) (Tap et al., 2015). The dataset, derived from the DrugImprover framework (Liu et al., 2023b), comprises 1 million small molecules sampled from ZINC15 (Sterling & Irwin, 2015), each annotated with an OpenEye FRED docking score (McGann, 2012). The BO objective minimizes this score, where a more negative value indicates stronger predicted binding, so we maximize f=−score. Molecules are represented as 2048-bit Morgan (extended-connectivity) fingerprints (Rogers & Hahn, 2010) and projected to a continuous d=50 latent space by PCA, following standard practice in latent-space molecular optimization (Gómez-Bombarelli et al., 2018). Distinct scaffolds (different ring systems or functional groups) occupy disjoint regions of this latent space with different structure–activity relationships, so a single stationary kernel is a poor global model (Zhang et al., 2026; Gómez-Bombarelli et al., 2018; Griffiths & Hernández-Lobato, 2020; Korovina et al., 2020). Because the PCA projection is not invertible, an arbitrary point in the 50-dimensional latent space need not correspond to a real molecule, so acquisition maximization cannot use the Sobol-candidate protocol of Table 4 directly. Instead, at each round the candidate set C ($|C|$=2000) is drawn by uniform subsampling without replacement from the latent embeddings of the 1M-molecule library, excluding molecules already queried. The multiverse UCB score is evaluated on this pool, the arg max is queried, and its docking score is looked up from the precomputed library annotations. The continuous local-refinement step used on generative domains (see Appendix B.2.4) is not applied here, since gradient steps in latent space would leave the discrete manifold of valid molecules.

LLM data-mixture optimization (DMO, d=6). The base DMO task of BOLT (Chew et al., 2026) optimizes a two-stage training curriculum, allowing data proportions to change mid-training. This extends mixture optimization across supervised fine-tuning (SFT) domains (Lambert et al., 2025) by searching over both the composition and its allocation across stages. Each training mixture draws 10k samples from three TULU-3 SFT domains (Lambert et al., 2025) (IF / instruction-following, Math, and Code) at a specified mixing proportion. Chew et al. (2026) fine-tune a pretrained Qwen3-4B model (Yang et al., 2025) under a two-stage curriculum: stage 1 covers tokens 0–5M and stage 2 covers tokens 5M–10M, with mixture weights chosen independently in each stage. Models are scored on one metric per skill: instruction-level strict accuracy on IFEval (Zhou et al., 2023) for IF, 4-shot

accuracy on MATH-500 (Hendrycks et al., 2021) for Math, and pass@1 on MBPP+ (Austin et al., 2021; Liu et al., 2023a) for Code. The search domain is the two-stage mixing simplex

$$\mathcal{X}_{\text{DMO}} = \{x \in \mathbb{R}^6_{\geq 0} : \textstyle\sum_{i=1}^{3} x_i = 1,\ \sum_{j=4}^{6} x_j = 1\},$$

so $x_{1:3}$ and $x_{4:6}$ are the stage-1 and stage-2 mixture weights over IF/Math/Code (4 effective free parameters after the simplex constraints, see Table 3). Corners and edges of this product of simplices correspond to skill-dominated regimes (e.g., Math-heavy stage 1 followed by Code-heavy stage 2), giving named heterogeneous structure that classical HPO benchmarks lack. Rather than retraining for every BO query, we evaluate through BOLT's released DMO-4B emulator: a 2-layer MLP fit on 3014 real fine-tuning runs and validated on 160 held-out configurations, with test Spearman rank correlations of 0.721 (IF), 0.779 (Math), and 0.721 (Code) (Chew et al., 2026). These correlations come from BOLT's held-out validation, while our component visualizations describe the fitted surrogate of the emulator's predicted surface (Appendix B.2.2). We maximize the unweighted mean of the three skill scores, using this single-objective task throughout the main text. BOLT's additional multi-objective (DMO-MO) and heteroscedastic-noise (DMO-Het) variants provide possible extensions, with DMO-Het offering a test of the uncertainty estimate in Eq. (8).

B.2.2 Descriptive Component Contributions on DMO

A large $|c_m(x)|$ indicates a strong contribution to the fitted prediction, but its omitted sign determines whether it raises or lowers the total. Since components can offset one another, their magnitudes should be interpreted alongside their sum, as for the component means of exact additive and MKL GPs (Appendix A.1.4).

Plotting each skill-aligned magnitude against its mixture coordinate shows how the fitted prediction is distributed across the specified kernel blocks. These local predictions differ from global kernel weights and gating probabilities, so we do not rank those quantities as interchangeable attribution scores. Moreover, the stage-1 proportions sum to one and the skill directions are specified in advance, meaning coordinate correlations reflect both simplex constraints and the chosen kernels rather than establishing recovery of unknown regimes.

Emulator-to-real-data correlations measure a separate relationship and therefore impose no ceiling on component-coordinate correlations. To make the latter comparable across worlds, the contribution maps use one fitted model, shared stage-1 probe inputs, and a fixed stage-2 mixture.

Visualization configuration. The reproduction protocol uses the model from experimental seed 0 after $N = 200$ total observations. Its coefficient posterior is recomputed using all 200 observations, with kernel parameters, landmarks, feature coordinates, and pilot-defined response normalization held fixed. This includes the final observation, which arrives after the last acquisition decision.

With the stage-2 mixture fixed at $(0.13, 0.82, 0.05)$ in IF/Math/Code order, we generate 4000 stage-1 probes from $\text{Dirichlet}(1, 1, 1)$ using random seed 3. Each component prediction is evaluated once on this shared probe set, and the same magnitudes $|c_m(x)|$ feed all panels. Heatmaps display the mean magnitude in each hexagonal bin using 22 bins along the horizontal axis. The scatter panels display 700 probe points sampled without replacement using random seed 300, with the same subset used for all three components. Linear trends and Spearman correlations use all 4000 probes. Correlations are also computed separately for experimental seeds $0, \ldots, 9$ on the same probe set and summarized by their mean and standard deviation. The displayed seed is fixed in advance rather than selected by its correlation or optimization outcome. Probe evaluations use the fitted surrogate and add no objective calls or synthetic contribution noise.

B.2.3 Kernel-Population Construction

Every world in every population below is a standard stationary kernel (SE, Matérn, or RQ, as defined in Rasmussen & Williams (2006, Ch. 4)), optionally combined with a periodic factor or given anisotropic (ARD) lengthscales. The MVBO populations use these established kernels as component hypotheses in the joint finite-feature surrogate of Section 3.1. We chose the populations to cover distinct assumptions about smoothness, periodicity, scale, and per-dimension anisotropy.

Although the total coordinate count $r = \sum_m r_m$ determines the Bayesian update size, overlapping components can have a smaller effective dimension, so adding worlds need not add representable

functions (Lemma 1). Matching coordinate counts alone also does not match regret bounds, which depend on feature norms, noise, and the objective norm in Corollary 1, together with approximation error in Eq. (15). For bounded kernels, the ideal projection construction gives

$$\kappa^2 \le \sigma_\omega^2 \sum_m \sup_{\mathbf{x}\in\mathcal{X}} k_m(\mathbf{x},\mathbf{x}). \tag{55}$$

Adding an unnormalized duplicate can therefore change the prior amplitude and distribution of component contributions without expanding the function space. Population size and per-world ranks are modeling choices that jointly determine representation coverage and cost.

On the molecular conformer task ($d = 12$) the population is $M = 4$: SE, Matérn-$5/2$ (M52), a product SE×PER kernel combining a smooth trend with a periodic component (matching the torsional periodicity of Section 6.1), and a rational quadratic (RQ) kernel, which is itself a scale mixture of SE kernels over a range of lengthscales and so gives a cheap fourth world for the differing curvature across rotameric basins. On drug discovery ($d = 50$) the population is $M = 4$: ARD-SE, ARD-Matérn-$5/2$, ARD-Matérn-$3/2$ (M32), and ARD-RQ. All four are standard kernels, but giving each an independent lengthscale per PCA dimension (ARD) lets every world adapt to whichever latent axes matter for a given scaffold, while varying the smoothness order (ν=$3/2$ vs. $5/2$) and tail behavior (RQ vs. SE) keeps the four worlds structurally distinct without any non-standard geometric kernel. On DMO the population is $M = 5$: the same SE and M52 worlds, plus one anisotropic SE world per mixture axis (IF, Math, Code) with a short length scale along that axis's coordinates and a long length scale along the others, so the population can align a block with each named regime of Section 6.1. The population remains fixed within each run.

Starting values. Each world's amplitude, lengthscale, and (where applicable) periodic parameters receive an explicit starting value for the pilot fit of Appendix B.2.4. On molecular conformer search, all four amplitudes start at 0.25. Lengthscales start at 0.15 (SE), 0.30 (M52), 0.50 for the SE×PER smooth factor and 0.30 for its periodic factor, and 0.30 for RQ, whose shape parameter starts at 1. Each torsional coordinate is mapped to $[0, 1]$ over one full cycle, so the periodic factor's period is fixed at 1. On screening, all four amplitudes start at 0.25 and every ARD lengthscale starts at 0.30. The RQ shape parameter is fixed at 1 rather than fit. With four ARD kernels over d=50 dimensions, this pilot fit estimates over 200 amplitude and lengthscale parameters from the n_0=10 initial observations, which is a severely underdetermined marginal-likelihood problem. The bounded L-BFGS-B optimization (amplitudes in $[10^{-3}, 10]$, lengthscales in $[0.03, 2]$) keeps the fit near these declared starting values rather than diverging, but we do not claim the resulting per-dimension lengthscales are individually well estimated, and explicit parameter tying across dimensions, as used for DMO's skill worlds below, is a natural mitigation we did not apply here. On DMO, all five amplitudes start at 0.20, and the general SE and M52 worlds' lengthscales start at 0.30. The three skill worlds start from

$$\boldsymbol{\ell}_{\mathrm{IF}} = (.15, .50, .50, .15, .50, .50), \tag{56}$$
$$\boldsymbol{\ell}_{\mathrm{Math}} = (.50, .15, .50, .50, .15, .50), \tag{57}$$
$$\boldsymbol{\ell}_{\mathrm{Code}} = (.50, .50, .15, .50, .50, .15), \tag{58}$$

where each vector places a short lengthscale on that skill's own stage-1 and stage-2 mixture coordinate and a long lengthscale on the remaining four. Rather than fit each entry independently, the pilot fit optimizes a single positive multiplier per skill world, bounded to $[0.2, 4]$, that scales the vector while preserving its short-to-long lengthscale ratio.

LLM-guided kernel proposals, as in CAKE (Suwandi et al., 2025), offer an optional outer loop for population adaptation. Adding or replacing a world changes the prior and feature allocation and incurs search and refitting costs. These extensions fall outside the fixed-map analysis and the current allocation comparison.

B.2.4 Application Protocol and Compute

Each application uses $N = 200$ evaluations, including $n_0 = 10$ initial observations, with ten paired seeds $0, \ldots, 9$. Acquisition searches 2000 Sobol candidates followed by constrained local refinement on continuous domains, while screening selects among 2000 unqueried library molecules without continuous refinement. Because these searches use subsampling or local refinement, substituting the

Table 4: Application populations and feature budgets. These settings differ from the common four-world controlled allocation experiment.

Task	Worlds	Features/world	Total dimension
Molecular	4	10	40
Screening	4	10	40
DMO	5	10	50

candidate count into the finite-domain UCB expression is a heuristic rather than an instance of the theorem's exact finite-domain selection rule.

Kernel families are specified in Appendix B.2.3.

Application evaluation metrics. All three applications report normalized simple regret, computed from the corresponding objective before any target standardization for surrogate fitting. Each task uses a fixed reference maximum f_{ref} and minimum f_{low}, with $f_{\mathrm{ref}} > f_{\mathrm{low}}$, shared across methods and seeds. On a fully scored finite domain, these can be taken as its maximum and minimum objective values. Empirical reference values on continuous domains define a normalized reference gap and are not certificates of a global optimum.

For molecular search, let $E(x)$ denote the constrained-relaxation MMFF94 energy in kcal/mol. With reference energies $E_{\mathrm{ref}} < E_{\mathrm{high}}$, BO maximizes $f = -E$ and the reported metric is

$$\rho_j^{\mathrm{Mol}} = \frac{\min_{i \le j} E(x_i) - E_{\mathrm{ref}}}{E_{\mathrm{high}} - E_{\mathrm{ref}}}. \tag{59}$$

The numerator and denominator have the same energy units, so the reported regret is dimensionless.

For screening, let $D(x)$ denote the docking score before negation. With reference scores $D_{\mathrm{ref}} < D_{\mathrm{high}}$, BO maximizes $f = -D$ and the reported metric is

$$\rho_j^{\mathrm{Drug}} = \frac{\min_{i \le j} D(x_i) - D_{\mathrm{ref}}}{D_{\mathrm{high}} - D_{\mathrm{ref}}}. \tag{60}$$

For base DMO, $f(x)$ is the noiseless, unstandardized mean of the three emulator skill scores. Using the empirical reference $f_{\mathrm{ref}} = 0.61116$ from the released benchmark (Chew et al., 2026), the reported metric is

$$\rho_j^{\mathrm{DMO}} = \frac{0.61116 - \max_{i \le j} f(x_i)}{0.61116 - f_{\mathrm{low}}^{\mathrm{DMO}}}. \tag{61}$$

The reference applies only to base DMO. Negative gaps indicate that a query exceeds the empirical reference and must be reported as such.

These metrics are dimensionless and lower-is-better. Each curve and final table entry summarizes the per-seed normalized simple regret by its mean and standard deviation. No logarithm is applied.

Pilot fitting and observation model. The fixed-map application configuration fits the component amplitudes and lengthscales once on the ten initial observations by maximizing the exact sum-kernel marginal likelihood. Optimization uses L-BFGS-B in log-parameter space, two deterministic starts, and at most 500 iterations per start. Amplitudes are bounded to $[10^{-3}, 10]$ and lengthscales to $[0.03, 2]$ in normalized input coordinates. Periodic structure and the skill-axis assignments follow the declared task populations. The first start uses the task-specific amplitude, lengthscale, and periodic values given in the Starting Values paragraph of Appendix B.2.3. The second start rescales every amplitude from the first start by 0.8 and every lengthscale, periodic lengthscale, or DMO skill-world multiplier by 1.5, with both starts optimized in log space. For DMO, the pilot fit preserves each skill world's short-to-long lengthscale ratio by optimizing one positive vector multiplier, bounded to $[0.2, 4]$, rather than fitting its entries independently. The best converged pilot fit is frozen and shared by methods that use this kernel population: GP-UCB takes its named SE or Matérn-$5/2$ component, while ensembles, MKL-BO, the exact sum-kernel GP, and MVBO share the joint fit, including MKL weights. This matched-representation freeze isolates the surrogate from online hyperparameter search. It does

not evaluate periodically refitting GP-UCB or TuRBO. TuRBO, IMoGPE, RAMBO, and CAKE keep the settings in Appendix B.2.5. Kernel parameters, landmarks, ranks, and the numerical feature construction remain fixed throughout the subsequent search. The pilot response mean and standard deviation define the response transform, with a standard-deviation floor of 10^{-8}, and this transform is also frozen. No subsequent hyperparameter refitting or response rescaling is applied to fixed-map runs.

We use coefficient-prior variance $\sigma_\omega^2 = 1$ and regression regularizer $\lambda = 0.01$ in standardized response units, while leaving the deterministic application evaluations free of added noise. Thus the regularizer stabilizes the surrogate rather than representing injected observation noise. UCB uses the latent posterior variance with $\delta = 0.1$, and the pilot-defined response standardization is kept separate from the metrics used to report performance.

Compute configuration. The application experiments use an Ubuntu Linux workstation with an AMD Ryzen 9 7950X CPU, 64 GB system RAM, and one NVIDIA GeForce RTX 4090 GPU with 24 GB VRAM. Surrogate linear algebra uses float64 precision with one experiment worker and one CPU compute thread per run. CUDA operations are synchronized at timing boundaries when GPU computation is measured. These settings are distinct from the recorded ARM64 environment of the controlled finite-domain experiments. Computational scaling of the isolated posterior update is discussed in Appendix B.1.7.

B.2.5 Baseline Hyperparameters

Single-kernel baselines use their named SE or Matérn-$5/2$ kernel, while methods with multiple kernels use the specified task families. In matched controlled experiments, kernel parameters remain fixed after construction, with the pilot-fitting variant described in Appendix B.2.10.

- **GP-UCB / GP-UCB-M52** (Srinivas et al., 2010) use a single kernel and the exploration schedule $\beta_t = 2\log(|C|t^2\pi^2/6\delta)$ with $\delta = 0.1$, where C is the finite acquisition candidate set (Appendix B.2.4).
- **TuRBO** (Eriksson et al., 2019) uses M trust regions, matching the number of worlds used by MVBO, with initial side length $L_{\text{init}} = 0.8$ and the original success/failure tolerances, restarting each region when its side length falls below $L_{\min}$. This provides a local-BO baseline with the same multi-region budget as MVBO's multi-world surrogate.
- **IMoGPE** (Rasmussen & Ghahramani, 2001) fits an infinite mixture of GP experts by variational EM, with input-dependent mixing weights that learn how expert assignments change across $\mathbf{x}$.
- **EGP-TS** (Lu et al., 2023) maintains one GP per kernel in the shared population and uses Thompson sampling over the ensemble to select evaluations.
- **WBGP-BO** (Candelieri et al., 2025) uses the same kernel population, combining GP predictions through a weighted Wasserstein barycenter.
- **MKL-BO** (Gönen & Alpaydın, 2011) forms a convex combination of the population's kernel matrices. In the fixed-map application runs, those combination weights remain at the shared pilot values rather than being refit online.
- **Exact Sum-Kernel GP** fits a single GP with $k_{\text{sum}} = \sum_m k_m$ over the same population, jointly optimizing the component hyperparameters by marginal likelihood on the shared pilot without a weight constraint across kernels (Appendix B.2.4). Full rank on a fixed landmark set recovers only a Nyström approximation to this kernel.
- **CAKE** (Suwandi et al., 2025) evolves a kernel grammar with an LLM and uses BAKER to select one kernel per iteration, following the original paper's default settings.
- **RAMBO** (Zhang et al., 2026) fits a Dirichlet process mixture of GPs using collapsed Gibbs inference with burn-in 500, an Inverse-Gamma base measure with a=2, and the Log-Sqrt concentration schedule $\alpha_t = \alpha_0\sqrt{t}/\log(t+e)$ with $\alpha_0 = 0.2$. We retain its mixture expected improvement acquisition alongside the UCB baselines, following the original paper's finding that EI is the more robust choice for this surrogate.

B.2.6 Toy Example

A kernel encodes a single inductive bias, so it is a poor model of an objective that changes structure across the input, where the geometry that fits one region is typically wrong in another. We illustrate this with the scalar function

$$f(x) = [1 - s(x)]\sin(12\pi x) + 3s(x)(x - 0.5), \qquad s(x) = \frac{1}{1 + \exp[-30(x - 0.5)]}, \tag{62}$$

which oscillates on the left half of $[0, 1]$ and follows a rising line on the right. The sigmoid $s(x)$ is used only to construct f. The surrogate sees noisy evaluations of f and is given neither regime labels nor access to s. We concatenate a locally periodic kernel, a linear kernel, and a Matérn-$5/2$ kernel and fit them jointly, as in Eqs. (9) and (10).

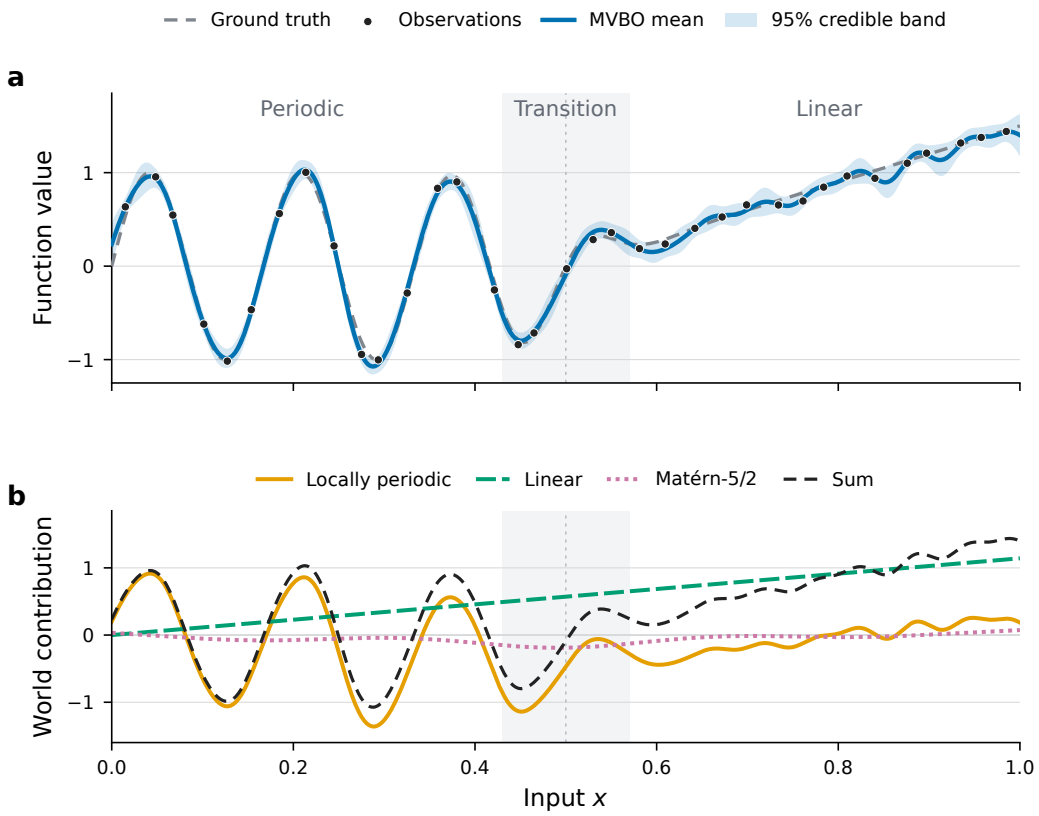


Figure 4: **A two-regime objective that a single kernel cannot represent.** (a) Joint multiverse posterior mean and pointwise 95% latent-function credible band after 50 noisy observations. The mean follows the periodic regime on the left, the linear regime on the right, and a smooth transition between them. (b) Signed contributions $c_m(x)$ of the locally periodic, linear, and Matérn-$5/2$ worlds, whose sum equals the posterior mean in (a). Each world covers only part of f, which is why a single inductive bias is insufficient.

The top panel of Figure 4 shows the joint posterior mean after these observations. It tracks the oscillations for $x < 0.5$, the rising trend for $x > 0.5$, and a smooth handoff through the shaded interval around the change in regime. Against noiseless f, the root mean squared error is 0.059 on a dense grid in $[0, 1]$. Residual ripples remain on the linear side, so the fit recovers the two-regime structure without matching Eq. (62) as an exact piecewise reconstruction.

The bottom panel shows why a single kernel cannot represent both regimes. The plotted curves are the signed contributions $c_m(x)$ from the joint fit in Eq. (9), so they should not be read as independent single-kernel GPs or as probabilities of assigning x to a regime. Their sum is the mean in Figure 4a. The locally periodic contribution accounts for most of the oscillation on the left and then drops

Table 5: Controlled allocation experiment at requested total dimension $r = 32$. Means±SD and pointwise 95% paired bootstrap intervals over 20 seeds. Negative separate-minus-aggregate differences favor separate allocation. Exact denotes the untruncated sum kernel on the same finite domain. *All Branin entries, including differences and intervals, are in units of 10^{-5}.

Task	Separate	Aggregate	Exact	Δ [95% CI]
Minority signal	0.006 ± 0.012	0.085 ± 0.053	0.003 ± 0.009	-0.079 $[-0.102, -0.054]$
Dominant signal	0.005 ± 0.014	0 ± 0	0 ± 0	0.005 $[1.79 \times 10^{-4}, 0.012]$
Branin*	2.005 ± 5.873	1.474 ± 5.593	0.266 ± 1.190	0.531 $[-0.466, 1.785]$
Hartmann-6	0.063 ± 0.078	0.043 ± 0.062	0.038 ± 0.072	0.020 $[0.002, 0.046]$
Ackley-5	0 ± 0	0 ± 0	0 ± 0	0 $[0, 0]$
Regime switch	0 ± 0	0 ± 0	0 ± 0	0 $[0, 0]$

away, rather than carrying that pattern through the linear regime. The linear contribution supplies a positive slope that becomes the dominant structure on the right, while the Matérn-$5/2$ contribution adjusts local level and smoothness, especially near the boundary. Taken alone, none of these three curves matches f: a locally periodic kernel would keep correlating neighboring oscillations past $x = 0.5$, and a linear kernel cannot represent the left-hand oscillations. Because the multiverse kernel is the sum of the world kernels, the fitted mean can change character across the transition without committing to one global geometry.

B.2.7 Controlled Allocation Protocol

To isolate feature allocation, we compare separate truncation, truncation of the summed kernel, and exact sum-kernel inference using the same kernels, landmarks, candidates, initial observations, and noise streams. In the 20-seed experiment at $r = 32$ (Table 5), separate allocation improves the constructed minority-signal control but has a higher mean gap on Hartmann-6, with small or zero final gaps on the remaining tasks. These differences indicate that preserving a component is useful when its retained directions align with the objective.

The controlled experiment uses four stationary worlds on Branin, Hartmann-6, Ackley-5, and the one-dimensional regime-switching task: SE lengthscales 0.15 and 0.5, Matérn-$5/2$ lengthscale 0.3, and RQ lengthscale 0.3 with shape one. Inputs are scaled to the unit cube, and each kernel has amplitude $1/4$ with predeclared hyperparameters, so observed responses do not influence kernel selection. Each seed supplies a scrambled Sobol domain, an independent uniform landmark subset, and a shared initial design. The evaluator uses the exact extrema of that finite domain, which serve as comparison references without certifying a continuous-domain optimum.

The two alignment controls use four orthogonal finite feature blocks with amplitudes 1, 0.01, 0.005, and 0.0025. Keeping this kernel construction shared, we switch the objective between the first and fourth blocks and vary its coefficients and domain directions by seed. These synthetic controls isolate sensitivity to informative directions with low kernel amplitude.

The default landmark set in Algorithm 1 is a uniform sample from the candidate domain, drawn without objective values and then frozen. Application runs reuse the n_0 initial inputs as landmarks ($q = n_0$). The controlled protocol uses seeds $0, \ldots, 19$, 2048 candidates, $q = 64$ uniform landmarks independent of the initial design, requested dimensions $16, 32, 64$, and 200 total queries including ten initial observations. Each world receives one quarter of the requested dimension, but retaining only positive numerical eigenvalues can give a smaller realized rank. We therefore report both requested budgets and realized dimensions, alongside projection residuals across the candidate domain. Both maps realize the requested dimension in every original run except the regime-switching task at budget 64, where separate truncation retains 58 coordinates and aggregate truncation retains 59–64. These cases share a budget ceiling.

We estimate response mean and scale from the shared pilot and hold them fixed, using $\lambda = 0.01$ to regularize the surrogate in transformed units. The default objectives are noiseless, so this regularizer does not represent noise in their evaluations. The optional noisy experiment shares a point-indexed noise stream across methods and records latent best-query and returned-incumbent quality separately. UCB uses the full fixed domain in its schedule and maximizes over all unqueried candidates. The

Table 6: Average normalized gap over the full 200-query budget at $r = 32$. Entries are mean$\pm$SD over 20 seeds. Δ is separate minus aggregate, with a pointwise 95% paired bootstrap interval. Lower values are better.

Task	Separate	Aggregate	Exact	Δ [95% CI]
Minority signal	0.078 ± 0.033	0.146 ± 0.039	0.073 ± 0.019	-0.068 $[-0.089, -0.046]$
Dominant signal	0.042 ± 0.026	0.020 ± 0.006	0.020 ± 0.006	0.021 $[0.012, 0.032]$
Branin	0.003 ± 0.003	0.003 ± 0.003	0.004 ± 0.003	-1.96×10^{-4} $[-4.99\times10^{-4}, 1.59\times10^{-4}]$
Hartmann-6	0.130 ± 0.070	0.137 ± 0.066	0.119 ± 0.072	-0.007 $[-0.028, 0.018]$
Ackley-5	0.081 ± 0.025	0.106 ± 0.051	0.084 ± 0.028	-0.025 $[-0.042, -0.008]$
Regime switch	0.007 ± 0.005	0.007 ± 0.005	0.007 ± 0.005	1.38×10^{-4} $[-3.20\times10^{-5}, 3.72\times10^{-4}]$

synthetic objectives are fixed functions, so the Bayesian schedule is an experimental policy, not a claim that these tasks follow the specified Gaussian prior.

Equivalence control. Because KAPS uses the spectral Nyström map, a conventional Nyström factorization at full landmark rank induces the same posterior when kernels, landmarks, and regularization match. We check this equivalence before the allocation comparisons to distinguish a change of coordinates from changes to landmarks or numerical regularization. The comparison therefore concerns feature allocation rather than superiority of one Nyström factorization over another.

Paired outcomes and uncertainty. For each task and budget, we computed the final gap difference $\Delta_s = \rho_{N,s}^{\text{separate}} - \rho_{N,s}^{\text{aggregate}}$ for each shared seed s. We report the mean difference and a pointwise 95% bootstrap interval obtained by resampling paired seeds 10,000 times. Negative differences favor separate allocation, while mean gap across the full budget provides a secondary measure of progress. The intervals are pointwise and have no multiplicity adjustment, so establishing equivalence would require a prespecified practical margin and paired uncertainty contained within it.

Computation. Both models update incrementally and cache candidate means and variances, so the measured conditioning step includes refreshing those moments. Exact inference uses accumulated Cholesky projections, while feature inference uses rank-one Bayesian linear updates, with acquisition selection, common kernel construction, and representation setup timed separately. Measurements use one worker and single-threaded BLAS, with randomized method order within each seed. The recorded environment was ARM64 macOS 26.5.2 with Python 3.9.6, NumPy 2.0.2, and SciPy 1.13.1. These environment details apply to the controlled experiments.

B.2.8 Search Trajectories and Rank Sensitivity

Across 120 task–seed pairs, the allocation experiment contains 840 BO trajectories covering three feature budgets and the exact reference, all of which completed 200 queries. Alongside the primary $r = 32$ endpoint comparison in Table 5, Figure 5 and Table 6 report the prespecified average normalized gap across all 200 queries, including the pilot, to measure how quickly each method improves its incumbent.

Separate truncation reduces the final gap on the minority-signal control by 0.0788 relative to aggregate truncation, with paired interval $[-0.1022, -0.0545]$. On Ackley, all three methods reach the same zero final gap at $r = 32$, but separate truncation reduces average gap relative to aggregate truncation by 0.0245 $[0.0083, 0.0420]$. Thus endpoint quality and progress throughout the budget provide complementary comparisons. The control tasks isolate alignment with kernel components, while the analytic tasks test these choices on fixed objective functions. Table 5 reports final gaps on all four analytic benchmarks, including Hartmann-6, where aggregate truncation attains a lower mean final gap.

B.2.9 Projection and Prediction Diagnostics

For each fixed representation, we evaluate $d(x) = k_+(x, x) - \mathrm{MvK}(x, x)$ on the entire 2048-point candidate domain. Figure 6 and Table 7 report the maximum deficit against feature budget, together with paired final-gap differences. Since the domain, kernels, and map remain fixed within each

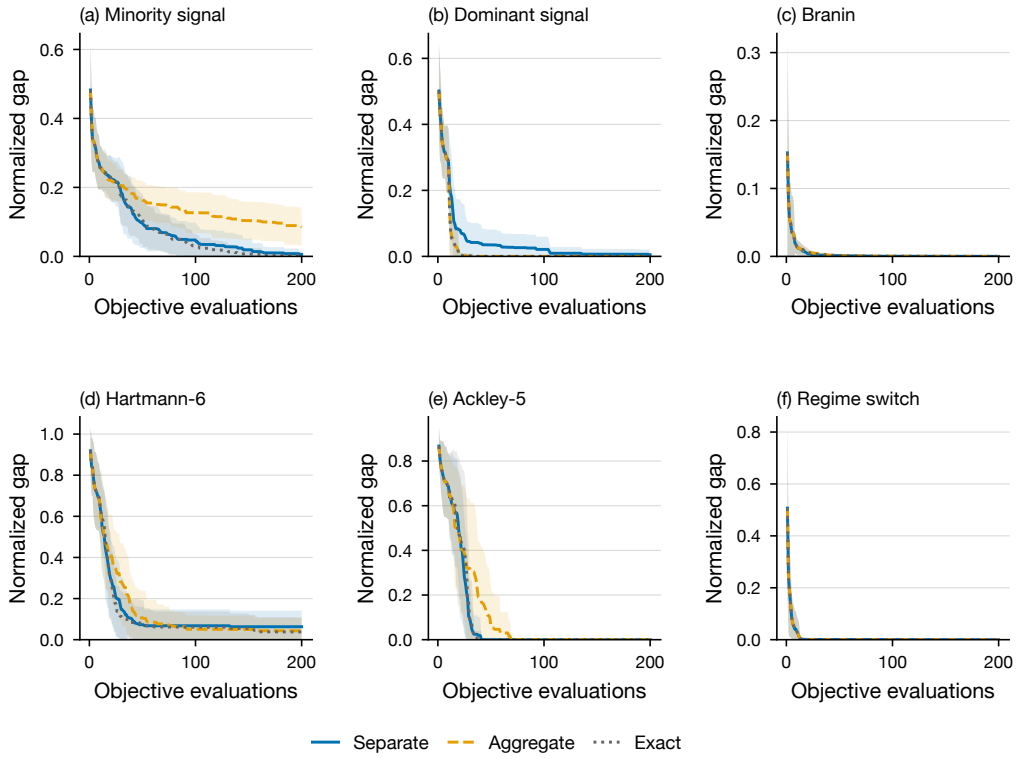


Figure 5: Allocation affects both final quality and progress during search. Curves show mean normalized best-query gap at $r = 32$ over 20 seeds. Bands show one standard deviation, clipped at zero. All 200 evaluations include the ten-query pilot. No smoothing is applied.

run, the maximum deficit changes only across representations. Increasing rank expands the retained space and, on the orthogonal controls, recovers the full generating features at rank 64, where all three methods attain identical gaps. Landmark coverage still limits the analytic tasks, with requested budgets and realized dimensions distinguished in Appendix B.2.7.

Aggregate truncation can have a smaller kernel deficit while separate truncation searches more effectively. In the minority-signal control, the dominant component accounts for most prior variance, but the objective lies in the low-amplitude component. Reserving directions for that component improves search even though the remaining deficit is larger. This separates approximation of the kernel from preservation of directions informative for the objective.

We also evaluate $d(x_t)/k_+(x_t, x_t)$ at the selected points and record prediction diagnostics before each post-pilot observation. For each seed, empirical interval occupancy is the fraction of the 190 selected points satisfying $|f(x_t) - \mu_{t-1}(x_t)| \leq 1.96\sigma_{t-1}(x_t)$. The standardized squared error is the average of $(f(x_t) - \mu_{t-1}(x_t))^2/\sigma^2_{t-1}(x_t)$, with a numerical variance floor of 10^{-12} in transformed units. Table 8 summarizes these seed-level quantities at $r = 32$. The raw latent-posterior bands are evaluated here, while acquisition uses its specified UCB multiplier. These measurements characterize predictions at adaptively selected points under the fixed synthetic objectives. The projection deficit supplies the geometric term in the approximation bound, with the RKHS norm factor specified separately in Eq. (16).

B.2.10 Landmark Placement, Rank Allocation, and Pilot Fitting

We additionally varied landmark count and placement (uniform versus farthest-point traversal), compared equal against spectrum-based rank allocation across worlds, and refit kernel amplitudes and lengthscales from the pilot before freezing the representation. Each experiment holds the objective, candidate domain, and noise stream fixed across the compared settings and reports paired final-

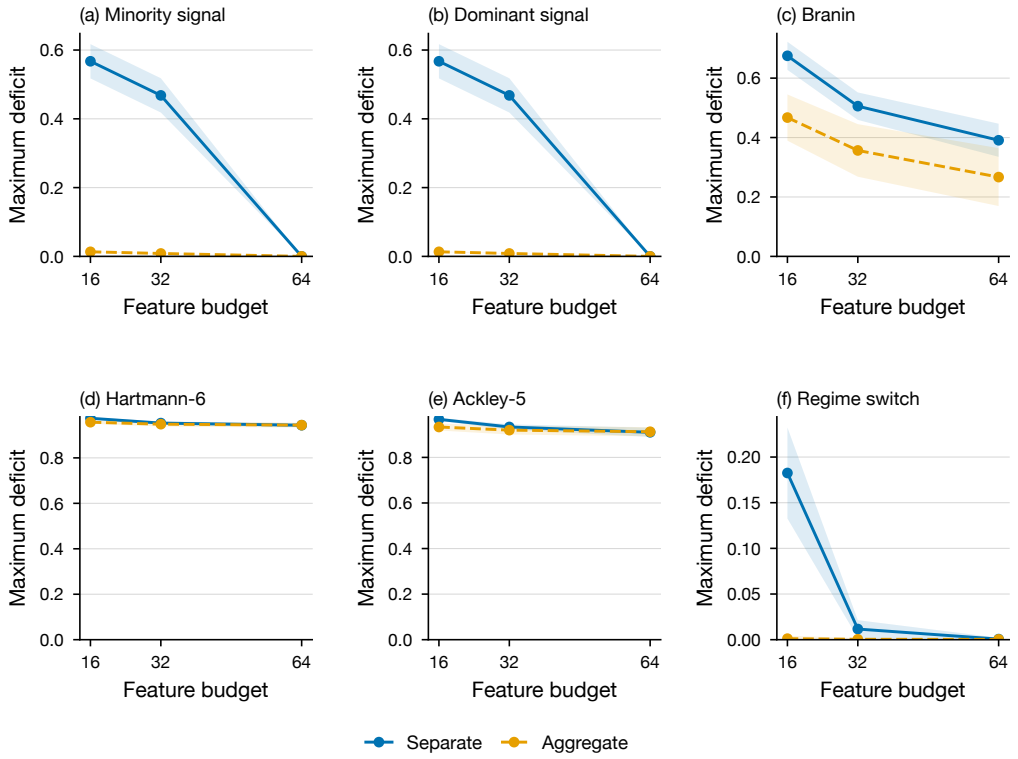


Figure 6: Increasing rank reduces the maximum projection deficit on the fixed candidate domain. Curves show mean±SD over 20 seeds for separate and aggregate truncation. Kernel parameters and 64 landmarks are shared within each seed. Exact inference has zero deficit.

Table 7: Projection and optimization across feature budgets. Maximum finite-domain deficits are mean±SD over 20 seeds. The final-gap Δ is separate minus aggregate with a pointwise 95% paired bootstrap interval. Small values are shown in scientific notation.

Task	r	Separate max. deficit	Aggregate max. deficit	Final Δ [95% CI]
Minority signal	16	0.567 ± 0.050	0.013 ± 0.001	-0.068 $[-0.090, -0.046]$
Minority signal	32	0.468 ± 0.050	$0.008 \pm 7.96\times10^{-4}$	-0.079 $[-0.102, -0.054]$
Minority signal	64	$1.50\times10^{-15} \pm 5.37\times10^{-16}$	$9.36\times10^{-12} \pm 3.93\times10^{-11}$	0 [0, 0]
Dominant signal	16	0.567 ± 0.050	0.013 ± 0.001	0.024 $[0.008, 0.044]$
Dominant signal	32	0.468 ± 0.050	$0.008 \pm 7.96\times10^{-4}$	0.005 $[1.79\times10^{-4}, 0.012]$
Dominant signal	64	$1.50\times10^{-15} \pm 5.37\times10^{-16}$	$9.36\times10^{-12} \pm 3.93\times10^{-11}$	0 [0, 0]
Branin	16	0.675 ± 0.047	0.467 ± 0.078	4.14×10^{-5} $[-2.54\times10^{-5}, 1.04\times10^{-4}]$
Branin	32	0.506 ± 0.046	0.357 ± 0.089	5.31×10^{-6} $[-4.66\times10^{-6}, 1.78\times10^{-5}]$
Branin	64	0.391 ± 0.056	0.267 ± 0.097	-1.63×10^{-5} $[-4.17\times10^{-5}, 0]$
Hartmann-6	16	0.974 ± 0.004	0.957 ± 0.007	0.035 $[0.004, 0.069]$
Hartmann-6	32	0.953 ± 0.007	0.948 ± 0.010	0.020 $[0.002, 0.046]$
Hartmann-6	64	0.944 ± 0.009	0.944 ± 0.011	0.007 $[-0.017, 0.032]$
Ackley-5	16	0.967 ± 0.007	0.933 ± 0.015	-0.032 $[-0.097, 0]$
Ackley-5	32	0.934 ± 0.012	0.920 ± 0.019	0 [0, 0]
Ackley-5	64	0.910 ± 0.020	0.913 ± 0.020	0 [0, 0]
Regime switch	16	0.183 ± 0.050	0.001 ± 0.003	0 [0, 0]
Regime switch	32	0.012 ± 0.010	$4.44\times10^{-4} \pm 0.001$	0 [0, 0]
Regime switch	64	$6.36\times10^{-4} \pm 0.001$	$3.21\times10^{-4} \pm 9.18\times10^{-4}$	0 [0, 0]

and average-gap differences with 95% bootstrap intervals over 20 seeds, following the protocol of Appendix B.2.7. These comparisons test the sensitivity of the fixed representation to landmark coverage, capacity allocation, and response-dependent kernel fitting.

Controlled pilot-fitting initialization. In the controlled pilot-fitting variant, the parameter vector contains four kernel amplitudes followed by four lengthscales, in the order SE, SE, Matérn-$5/2$, and

Table 8: Prediction diagnostics after the pilot at $r = 32$. Entries are mean±SD over 20 seeds. Query deficit is normalized by prior variance. Interval occupancy refers to latent mean±1.96SD at selected points. Standardized error is squared prediction error divided by latent posterior variance.

Task	Method	Query deficit / prior	Empirical coverage	Standardized error^2
Minority signal	Separate	0.441 ± 0.009	0.071 ± 0.029	622.214 ± 374.958
Minority signal	Aggregate	$0.011 \pm 5.68 \times 10^{-4}$	0.068 ± 0.029	588.943 ± 351.125
Minority signal	Exact	0 ± 0	0.107 ± 0.033	238.795 ± 148.817
Dominant signal	Separate	0.423 ± 0.010	0.096 ± 0.049	560.297 ± 569.881
Dominant signal	Aggregate	$0.012 \pm 4.52 \times 10^{-4}$	0.712 ± 0.291	11.193 ± 29.079
Dominant signal	Exact	0 ± 0	0.779 ± 0.263	7.085 ± 17.578
Branin	Separate	0.181 ± 0.022	0.539 ± 0.142	8.920 ± 7.476
Branin	Aggregate	0.026 ± 0.009	0.921 ± 0.079	1.301 ± 0.835
Branin	Exact	0 ± 0	0.986 ± 0.010	0.437 ± 0.283
Hartmann-6	Separate	0.721 ± 0.023	0.040 ± 0.027	$4.43 \times 10^3 \pm 7.55 \times 10^3$
Hartmann-6	Aggregate	0.609 ± 0.018	0.062 ± 0.039	$6.25 \times 10^3 \pm 1.34 \times 10^4$
Hartmann-6	Exact	0 ± 0	0.736 ± 0.207	6.851 ± 10.039
Ackley-5	Separate	0.557 ± 0.014	0.053 ± 0.030	$1.74 \times 10^3 \pm 2.10 \times 10^3$
Ackley-5	Aggregate	0.461 ± 0.019	0.092 ± 0.051	$1.94 \times 10^3 \pm 2.51 \times 10^3$
Ackley-5	Exact	0 ± 0	0.798 ± 0.178	5.475 ± 6.395
Regime switch	Separate	$7.48 \times 10^{-4} \pm 2.67 \times 10^{-4}$	0.960 ± 0.085	1.413 ± 1.690
Regime switch	Aggregate	$3.39 \times 10^{-6} \pm 7.17 \times 10^{-6}$	0.993 ± 0.006	1.323 ± 2.671
Regime switch	Exact	0 ± 0	0.993 ± 0.006	1.321 ± 2.662

RQ. The two starts in log space are

$$\theta^{(1)} = \log(0.25, 0.25, 0.25, 0.25, 0.15, 0.5, 0.3, 0.3),$$
$$\theta_i^{(2)} = \theta_i^{(1)} - 0.25 + \frac{i-1}{14}, \qquad i = 1, \ldots, 8.$$

The logarithm acts coordinatewise, and the RQ shape is fixed at one. L-BFGS-B uses at most 500 iterations per start, function tolerance 10^{-10}, and gradient tolerance 10^{-6}. Amplitude bounds are $[10^{-3}, 10]$, and lengthscale bounds are $[0.03, 2]$. We select the converged fit with the lowest negative log marginal likelihood and mark the run as failed if neither start converges. These settings apply to the controlled four-world experiment, with separate initializations defined for the application kernels.

B.2.11 Measured Computation and Search Progress

Table 9 separates the costs of the original allocation experiment at $r = 32$. Common setup includes candidate generation, cheap objective evaluation, and kernel construction. Initialization includes representation construction and posterior initialization. Updates include conditioning and refreshing candidate moments, and selection covers acquisition scoring and choosing the next query. The total includes common setup for each method. Figure 7 pairs average cumulative computation with average regret at each query index.

Common setup accounts for much of the total at this horizon, and separate eigenspace construction adds work before the $O(r^2)$ update can amortize. All timings use one worker with BLAS restricted to one thread, and method order is randomized within each seed. Candidate kernel matrices and moments are cached, so these timings characterize finite-domain search. The costs of evaluating an expensive application objective are separate.

Table 9: Measured computation in milliseconds for 200 total queries at $r = 32$. Entries are mean±SD over 20 seeds. Common setup is charged equally within each seed. Updates include candidate-moment refresh. Total includes all displayed components.

Task	Method	Common	Init.	Updates	Select.	Total
Minority signal	Separate	16.573 ± 3.595	2.992 ± 0.156	4.720 ± 0.222	1.568 ± 0.051	25.852 ± 3.564
Minority signal	Aggregate	16.573 ± 3.595	0.909 ± 0.061	4.672 ± 0.181	1.551 ± 0.036	23.704 ± 3.570
Minority signal	Exact	16.573 ± 3.595	0.017 ± 0.004	5.008 ± 0.172	1.607 ± 0.067	23.204 ± 3.571
Dominant signal	Separate	15.624 ± 0.244	3.061 ± 0.140	4.737 ± 0.156	1.576 ± 0.041	24.998 ± 0.397
Dominant signal	Aggregate	15.624 ± 0.244	0.932 ± 0.054	4.656 ± 0.117	1.578 ± 0.047	22.789 ± 0.356
Dominant signal	Exact	15.624 ± 0.244	0.018 ± 0.004	5.041 ± 0.118	1.579 ± 0.046	22.263 ± 0.303
Branin	Separate	45.689 ± 2.482	3.470 ± 0.187	4.788 ± 0.637	1.601 ± 0.150	55.549 ± 2.662
Branin	Aggregate	45.689 ± 2.482	0.909 ± 0.045	4.617 ± 0.105	1.550 ± 0.033	52.765 ± 2.543
Branin	Exact	45.689 ± 2.482	0.018 ± 0.005	5.009 ± 0.154	1.610 ± 0.080	52.326 ± 2.476
Hartmann-6	Separate	48.427 ± 1.130	3.651 ± 0.373	4.892 ± 0.587	1.609 ± 0.115	58.579 ± 1.752
Hartmann-6	Aggregate	48.427 ± 1.130	1.049 ± 0.242	4.808 ± 0.190	1.605 ± 0.090	55.889 ± 1.404
Hartmann-6	Exact	48.427 ± 1.130	0.019 ± 0.004	5.127 ± 0.416	1.661 ± 0.128	55.233 ± 1.373
Ackley-5	Separate	47.469 ± 0.431	3.510 ± 0.120	4.674 ± 0.197	1.562 ± 0.070	57.214 ± 0.421
Ackley-5	Aggregate	47.469 ± 0.431	0.918 ± 0.063	4.645 ± 0.121	1.577 ± 0.049	54.609 ± 0.462
Ackley-5	Exact	47.469 ± 0.431	0.020 ± 0.004	5.003 ± 0.119	1.595 ± 0.058	54.086 ± 0.459
Regime switch	Separate	46.551 ± 5.915	3.291 ± 0.637	4.685 ± 0.216	1.568 ± 0.086	56.094 ± 6.566
Regime switch	Aggregate	46.551 ± 5.915	0.938 ± 0.208	4.975 ± 1.469	1.608 ± 0.287	54.071 ± 6.727
Regime switch	Exact	46.551 ± 5.915	0.020 ± 0.005	5.301 ± 1.051	1.611 ± 0.203	53.482 ± 7.148

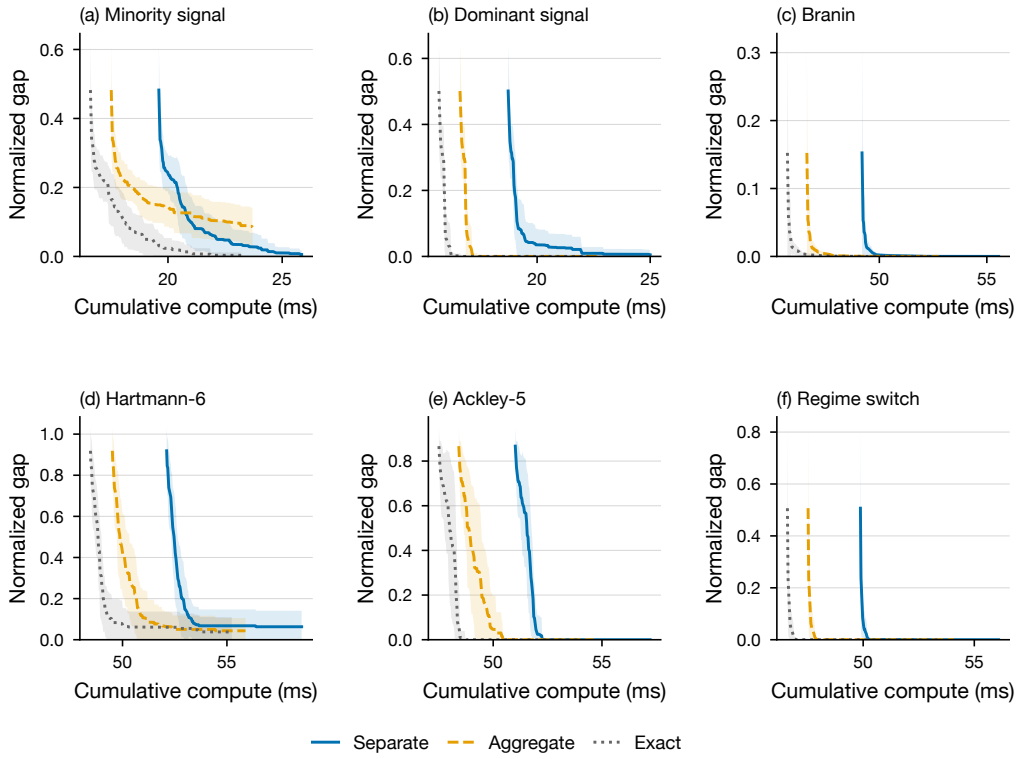


Figure 7: Search progress against measured cumulative computation. Each point pairs mean elapsed compute and mean normalized gap at the same query index over 20 seeds. Bands show one SD of gap. Setup, conditioning, and acquisition selection are included. These curves are indexed by queries and are not fixed-time-budget experiments.