Supplementary Material of Bayesian Optimization with Abstract Properties (BOAP)

7 Kernels and Reproducing Kernel Hilbert Spaces

The kernel functions $k(\mathbf{x}, \mathbf{x}') : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ used in Gaussian Process (GP) uniquely define an associated Reproducing Kernel Hilbert Space (RKHS) [1]. Formally:

Definition 1. *Let $\mathcal{H}_k$ be a Hilbert space of real valued functions $f : \mathcal{X} \to \mathbb{R}$ on a non-empty set $\mathcal{X}$. A function $k : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ is a reproducing kernel of $\mathcal{H}_k$, and $\mathcal{H}_k$ a **Reproducing Kernel Hilbert Space (RKHS)**, if*

- $\forall \mathbf{x}, \mathbf{x}' \in \mathcal{X}$, $k(\mathbf{x}, \mathbf{x}') = \langle k(\cdot, \mathbf{x}), k(\cdot, \mathbf{x}') \rangle_{\mathcal{H}_k}$,
- *k spans $\mathcal{H}_k$ i.e. $\forall \mathbf{x} \in \mathcal{X}$, $k(\ \cdot\ , \mathbf{x}) \in \mathcal{H}_k$,*
- $\forall \mathbf{x} \in \mathcal{X}, \forall f \in \mathcal{H}_k, \langle f(\cdot), k(\cdot, \mathbf{x}) \rangle_{\mathcal{H}_k} = f(\mathbf{x})$ *(the reproducing property).*

There exists varieties of kernels that can be used in fitting a GP surrogate model. A kernel that depends only on the distance between two given points i.e. $k = k(\mathbf{x} - \mathbf{x}')$ is called as stationary kernels. Stationary kernels are also called as translation-invariant kernels. Some of the popular kernel functions are listed below.

7.1 Matérn Kernel

Matérn kernel is a kernel that is commonly used in numerous machine learning applications. There are two variants of Matérn kernels that differ in their smoothness coefficient (ν) as shown below.

$$k_{\mathrm{MAT}}(\mathbf{x}, \mathbf{x}')_{\nu=\frac{3}{2}} = \left(1 + \frac{\sqrt{3}}{l}\|\mathbf{x} - \mathbf{x}'\|\right) \exp\left(-\frac{\sqrt{3}}{l}\|\mathbf{x} - \mathbf{x}'\|\right) \tag{11}$$

$$k_{\mathrm{MAT}}(\mathbf{x}, \mathbf{x}')_{\nu=\frac{5}{2}} = \left(1 + \frac{\sqrt{5}}{l}\|\mathbf{x} - \mathbf{x}'\| + \frac{\sqrt{5}}{3l^2}\|\mathbf{x} - \mathbf{x}'\|^2\right) \exp\left(-\frac{\sqrt{5}}{l}\|\mathbf{x} - \mathbf{x}'\|\right) \tag{12}$$

where l is the lengthscale hyperparameter of Matérn kernel.

7.2 Squared Exponential Kernel

Squared Exponential (SE) kernel or Radial Basis Function (RBF) or Gaussian kernel function is the popular stationary kernel function. The closed-form formulation of the SE kernel is represented as shown below.

$$k_{\mathrm{SE}}(\mathbf{x}, \mathbf{x}') = \sigma_f^2 \exp\left(-\frac{1}{2l^2}\|\mathbf{x} - \mathbf{x}'\|^2\right) \tag{13}$$

where, σ_f^2 and l corresponds to the signal variance and lengthscale hyperparameter, respectively, collectively represented as $\Theta_k = \{l, \sigma_f^2\}$.

7.3 Linear Kernel

Linear kernel is a commonly used non-stationary kernels defined as the inner product of the input data points. The mathematical formulation of a linear kernel is given by:

$$k_{\text{LIN}}(\mathbf{x}, \mathbf{x}') = \mathbf{x}'\mathbf{x}^{\intercal} + c \tag{14}$$

where c is the bias hyperparameter of linear kernels.

7.4 Multi-kernel Learning

Multi-kernel is a non-stationary kernel function defined as a linear combination of stationary and non-stationary kernels. For instance, multi-kernels can be constructed as:

$$k_{\text{MKL}}(\mathbf{x}, \mathbf{x}') = w_1\, k_{\text{SE}}(\mathbf{x}, \mathbf{x}') + w_2\, k_{\text{MAT}}(\mathbf{x}, \mathbf{x}') + w_3\, k_{\text{LIN}}(\mathbf{x}, \mathbf{x}') \tag{15}$$

where $\mathbf{w} = [w_1\ w_2\ w_3]$ corresponds to the kernel weights, that are usually tuned by maximizing GP log-likelihood.

8 Bayesian Optimization

The central idea of Bayesian Optimization (BO) strategy is to define a prior distribution over all the possible set of objective functions and then refine the model sequentially with the observed samples. BO is built on top of the Bayes theorem that incorporate prior belief about the black-box objective function under consideration. According to Bayes theorem, given a model $\mathcal{M}$ and data $\mathcal{D}$, the posterior probability of the model conditioned on data *i.e.* $\mathcal{P}(\mathcal{M}|\mathcal{D})$ is directly proportional to the likelihood of data $\mathcal{D}$ conditioned on model $\mathcal{M}$ *i.e.* $\mathcal{P}(\mathcal{D}|\mathcal{M})$, multiplied by the prior probability of model $\mathcal{P}(\mathcal{M})$,

$$\mathcal{P}(\mathcal{M}|\mathcal{D}) \ \propto\ \mathcal{P}(\mathcal{D}|\mathcal{M})\, \mathcal{P}(\mathcal{M}) \tag{16}$$

The observation model of BO is collected as $\mathcal{D}_{1:t} = \{\mathbf{x}_{1:t}, \mathbf{y}_{1:t}\}$, where $y_t = f(\mathbf{x}_t) + \eta_t$ is a noisy observation of the black-box objective function f evaluated at input location $\mathbf{x}_t$ corrupted with a white Gaussian noise $\eta_t \sim \mathcal{GP}(0, \sigma_\eta^2)$. In BO, we compute the posterior distribution $\mathcal{P}(f\,|\mathcal{D})$ by combining the prior $\mathcal{P}(f)$ with the likelihood $\mathcal{P}(\mathcal{D}|\,f)$ represented as,

$$\mathcal{P}(f\,|\mathcal{D}) \ \propto\ \mathcal{P}(\mathcal{D}|\,f)\, \mathcal{P}(f) \tag{17}$$

The posterior distribution $\mathcal{P}(f\,|\mathcal{D})$ computed captures our updated belief about the black-box objective function. BO can be perceived as a two step sequential strategy. First step focuses on defining the prior distribution that capture our prior beliefs. Usually GPs are used in placing prior distributions. Second step focuses on determining the best candidate that can be evaluated

next. Acquisition functions are used to find the next candidate point with the high promise of finding the optima. An algorithm for the standard Bayesian optimization procedure is provided in Algorithm 2.

Algorithm 2 Standard Bayesian Optimization

Input: Set of observations $\mathcal{D}_{1:t'} = \{\mathbf{x}_{1:t'}, \mathbf{y}_{1:t'}\}$, Sampling Budget T

1. **for** $t = t', \ldots, T$ iterations **do**
2. optimize $\Theta^\star = \underset{\Theta}{\operatorname{argmax}} \log \mathcal{L}$
3. update GP model with optimal kernel hyperparameters $\Theta^\star$
4. find the next query point $\mathbf{x}_{t+1} = \underset{\mathbf{x}\in\mathcal{X}}{\operatorname{argmax}}\, u(\mathbf{x})$
5. query $f(\mathbf{x})$ at $\mathbf{x}_{t+1}$ as $y_{t+1} = f(\mathbf{x}_{t+1}) + \eta_{t+1}$
6. augment the data as $\mathcal{D}_{1:t+1} = \mathcal{D}_{1:t} \cup (\mathbf{x}_{t+1}, y_{t+1})$
7. update GP model
8. **end for**

8.1 Acquisition Functions

The acquisition function guides the optimization by balancing the trade-off between exploration and exploitation. [10] proposed Expected Improvement (EI) acquisition function to guide the search by taking into account both the probability and magnitude of improvement over the current known optima. The next candidate point is obtained by maximizing the acquisition function given as:

$$u_{\text{EI}}(\mathbf{x}) = \begin{cases} (\mu(\mathbf{x}) - f(\mathbf{x}^+))\, \Phi(Z) + \sigma(\mathbf{x})\, \phi(Z) & \text{if } \sigma(\mathbf{x}) > 0 \\ 0 & \text{if } \sigma(\mathbf{x}) = 0 \end{cases} \tag{18}$$

$$\mathcal{Z} = \frac{\mu(\mathbf{x}) - f(\mathbf{x}^+)}{\sigma(\mathbf{x})}$$

where $\Phi(\mathcal{Z})$ and $\phi(\mathcal{Z})$ represents the Cumulative Distribution Function (CDF) and Probability Distribution Function (PDF) of the standard normal distribution, respectively and $f(\mathbf{x}^+)$ is the best value observed so far in the optimization.

Gaussian Process-Upper Confidence Bound (GP-UCB) acquisition function is another popular acquisition function defined based on confidence bounds criteria. GP-UCB acquisition function is given as:

$$u_{\text{GP-UCB}}(\mathbf{x}) = \mu(\mathbf{x}) + \sqrt{\beta_t}\, \sigma(\mathbf{x})$$

where β_t is a hyperparameter that balances the exploration-exploitation at iteration t. [16] discussed in detail the possible values for trade-off parameter β_t. Following [3], we set the value for trade-off parameter (β_t) at iteration t as:

$$\beta_t = 2 \log \left(\frac{t^{\frac{d}{2}+2} \pi^2}{3\delta'} \right)$$

where $\delta' \in (0, 1)$, d is the number of input dimensions.

8.2 Thompson Sampling based Bayesian Optimization

Thompson sampling [17] is a randomized selection strategy to select the next candidate for function evaluation by maximizing a random sample drawn from the posterior distribution. There have been significant advancements [9, 15, 2, 5] in Thompson sampling literature that demonstrate the theoretical guarantees of Thompson sampling. [13] provided a Bayesian regret bound for Thompson sampling using the notion of eluder dimension [11]. A complete algorithm of Thompson sampling based Bayesian optimization is provided in Algorithm 3.

Algorithm 3 Thompson Sampling based Bayesian Optimization

Input: Set of observations $\mathcal{D}_{1:t'} = \{\mathbf{x}_{1:t'}, \mathbf{y}_{1:t'}\}$, Sampling budget T

1. **for** $t = t', \cdots, T$ iterations **do**
2. optimize $\Theta^* = \underset{\Theta}{\operatorname{argmax}} \log \mathcal{L}$
3. update GP model with Θ^*
4. draw a random sample $\mathfrak{g}_{t+1}$ from the updated GP.
5. find the next query point $\mathbf{x}_{t+1} = \underset{\mathbf{x}\in\mathcal{X}}{\operatorname{argmax}}\, \mathfrak{g}_{t+1}(\mathbf{x})$
6. query $f(\mathbf{x})$ at $\mathbf{x}_{t+1}$ as $y_{t+1} = f(\mathbf{x}_{t+1}) + \eta_{t+1}$
7. augment the data as $\mathcal{D}_{1:t+1} = \mathcal{D}_{1:t} \cup (\mathbf{x}_{t+1}, y_{t+1})$
8. update GP posterior model
9. **end for**

At each iteration $t+1$, Thompson sampling strategy selects a point $\mathbf{x}_t$ that is highly likely to be the optimum according to the posterior distribution *i.e.,* $\mathbf{x}_t$ is drawn from the posterior distribution $\mathcal{P}_{\mathbf{x}^\star}(\cdot|\mathcal{D}_{1:t-1})$ conditioned on data $\mathcal{D}_{1:t-1}$. Thompson sampling strategy simplifies if Gaussian Processes (GPs) are used for prior and posterior distributions. For GPs, if $\mathfrak{g}$ is a sample drawn from $\mathcal{GP}(\mu_{\mathcal{D}_{1:t-1}}, k_{\mathcal{D}_{1:t-1}})$ we have:

$$\mathcal{P}_{\mathbf{x}^\star}(\mathbf{x}|\mathcal{D}_{1:t-1}) = \int \mathcal{P}_{\mathbf{x}^\star}(\mathbf{x}|\mathfrak{g})\mathcal{P}(\mathfrak{g}|\mathcal{D}_{1:t-1})\, d\mathfrak{g}$$

The probability $\mathcal{P}_{\mathbf{x}^\star}(\mathbf{x}|\mathfrak{g})$ has its mass at the maximizer:

$$\underset{\mathbf{x}\in\mathcal{X}}{\operatorname{argmax}}\, \mathfrak{g}(\mathbf{x}) \tag{19}$$

Using this utility, at each iteration $t+1$, we draw a random sample $\mathfrak{g}_{t+1}$ from $\mathcal{GP}(\mu_{\mathcal{D}_{1:t-1}}, k_{\mathcal{D}_{1:t-1}})$ and then find its maxima as per Eq. (19). The obtained maxima is used as the next candidate for function evaluation.

9 Additional Details of BOAP Framework

9.1 Abstract Properties

In numerous scenarios, human experts often reason about a system based on their approximate or qualitative knowledge on the system at different abstraction levels. The abstract properties can be simple physical properties or even a combination of multiple physical properties that an expert uses to reason about why one design is better than another. A well-chosen abstract property captures the input space of a system sufficiently enough that enables the domain expert to better reason about the output of the system in terms of such higher-order abstract properties. The main objective of augmenting abstract properties into the input space is to capture the inherent transformations of the function space $\mathcal{F}$ that significantly reduce the complexity and thereby simplify the task of GP modeling, thus a better overall optimization performance.

Furthermore, the human expert feedback is not available always, thus we cannot simply use the expert derived feedback as input always. However we model this (intermittent) expert feedback using a GP, and that model will act as a proxy for the expert in experiments where no expert feedback is given.

9.2 Construction of Preference Data P

Let $X = \{\mathbf{x}_i \mid \forall i \in \mathbb{N}_n\}$ be a set of n training instances. Let $\omega = \omega(\mathbf{x})$ be the immeasurable latent preference function values associated with each of the instances $\mathbf{x} \in X$. We assume that an expert compares and prefers an instance over another based on the utility ω. Let P be the set of p pairwise preferences between instances in X, defined as $P^{\omega_i} = \{(\mathbf{x} \succ \mathbf{x}') \quad \text{if } \omega_i(\mathbf{x}) > \omega_i(\mathbf{x}'), \mathbf{x} \in X, \forall j \in \mathbb{N}_P\}$. We generate preference list P^{ω_i} for each high-level feature of the designs by comparing its utility ω. We initially start with $\binom{t'}{2}$ preferences in P. For each candidate suggested at the end of i^{th} iteration, preference relations are updated accordingly. Here, we randomly choose a subset of datapoints containing $n/2 = (t' + i)/2$ datapoints to compare with the new candidate suggested for function evaluation. With the updated observation model $\bar{D} = \{\mathbf{x}_{1:n}, P = \{(\mathbf{x} \succ \mathbf{x}\prime)_j | \forall \mathbf{x}, \mathbf{x}' \in X, j \in \mathbb{N}_P\}\}$, we tune the hyperparameters (θ_h) of the kernel and refit $\mathcal{GP}_h$.

9.3 Frequency and cost of Expert Preference Data

BOAP allows an expert to intervene only when they are willing to provide any preferential input thereby allowing them to evolve their knowledge over the course of optimization. Therefore, we have not considered any human query costs in the current version of our research work. The preferential feedback is provided by experts at their discretion *i.e.,* when they feel they can make a clear judgment or that the algorithm has gone astray. How the human expert intervenes in giving the preferential feedback is not controlled by the algorithm. In our empirical studies with simulated human experts (zero query cost), at each iteration we

randomly choose a subset of datapoints containing $n/2$ datapoints to compare with the new candidate suggested for function evaluation.

As discussed in the main paper, our proposed framework uses a model selection based decision making on whether to choose augmented GP ($\mathcal{GP}_h$) built on the expert preferential knowledge on abstract properties or the standard GP ($\mathcal{GP}_h$) for suggesting the next candidate for function evaluation. The arm containing the standard GP model is denoted as Arm-$\mathfrak{f}$ and the arm containing the augmented GP model is termed as Arm-$\mathfrak{h}$.

9.4 2-Arm Scheme based Model Selection

In BOAP, Arm-$\mathfrak{f}$ directly models the given objective function $f(\mathbf{x})$, whereas Arm-$\mathfrak{h}$ models $f(\mathbf{x})$ via a human objective function ($h(\tilde{\mathbf{x}})$) in a search space comprising of inputs that are augmented with the latent abstract properties computed using rank GPs. As discussed in the main paper, the human objective function $h(\tilde{\mathbf{x}})$ incorporates the additional expert feedback from experts, and thus more-informed and a simplified version of $f(\mathbf{x})$. Therefore, at iteration t, if Arm-$\mathfrak{h}$ is selected and $\tilde{\mathbf{x}}_t = [\mathbf{x}_t, \mu_{\omega_1}(\mathbf{x}_t), \cdots, \mu_{\omega_m}(\mathbf{x}_t)]$ is the candidate suggested, then we observe y_t as:

$$y_t = h([\mathbf{x}_t, \mu_{\omega_1}(\mathbf{x}_t), \cdots, \mu_{\omega_m}(\mathbf{x}_t)]) \approx f(\mathbf{x}_t)$$

A graphical representation of BOAP framework and its components are shown in Figure 6. Nodes highlighted in blue color corresponds to inputs or outputs of a Gaussian process. Rectangular boxes shaded in Grey correspond to the nodes representing rank GPs, whereas the rectangular boxes shaded in orange correspond to the nodes representing conventional GPs. The estimated parameters and the latent variables are highlighted in yellow and green color, respectively.

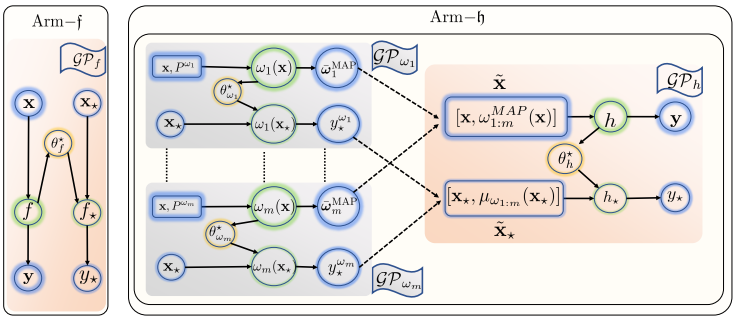

Fig. 6: Components of Arm−$\mathfrak{f}$ and Arm−$\mathfrak{h}$ in BOAP framework. Nodes highlighted in blue color corresponds to inputs or outputs of a Gaussian process. The estimated parameters and the latent variables are highlighted in yellow and green color, respectively. Rectangular boxes shaded in Grey and Orange correspond to the nodes representing rank GPs and conventional GPs, respectively.

9.5 Derivatives of the Likelihood Function

In rank GP distributions the MAP estimates are computed using Eq. (4) mentioned in the main paper. However, the Newton-Raphson recursion needs access to the first and second order derivatives of the loss function $L = -\ln \Phi\big(z(\mathbf{x}, \mathbf{x}')\big)$ with respect to latent function values $\boldsymbol{\omega}$. The analytical formulation of the first order and second order derivatives are given as:

$$\frac{\partial L}{\partial \omega(\mathbf{x}_i)} = \frac{\Gamma(\mathbf{x}_i)}{\sqrt{2\sigma_\eta^2}} \frac{\phi(z)}{\Phi(z)} \tag{20}$$

$$\frac{\partial^2 L}{\partial \omega(\mathbf{x}_i)\partial \omega(\mathbf{x}_j)} = \frac{\Gamma(\mathbf{x}_i)\ \Gamma(\mathbf{x}_j)}{2\sigma_\eta^2} \left(\frac{\phi^2(z)}{\Phi^2(z)} + z\frac{\phi(z)}{\Phi(z)} \right) \tag{21}$$

$$\text{where} \qquad \Gamma_{(\mathbf{x},\mathbf{x}')}(\mathbf{x}_i) = \begin{cases} 1 & \text{if } \mathbf{x} = \mathbf{x}_i \\ -1 & \text{if } \mathbf{x}' = \mathbf{x}_i \\ 0 & \text{Otherwise} \end{cases}$$

10 Convergence Analysis

We begin with the kernel for the augmented model:

$$K(\mathbf{x}, \mathbf{x}') = \kappa(\mathbf{x}, \mathbf{x}') \exp\left(- \sum_{d=1}^{D} \frac{\left(x_d - x'_d\right)^2}{2l_d^2} - \sum_{i=1}^{m} \frac{\left(\mu_{\omega_i}(\mathbf{x}) - \mu_{\omega_i}(\mathbf{x}')\right)^2}{2l_{\omega_i}(\mathbf{x},\mathbf{x}')^2} \right)$$

where:

$$\kappa(\mathbf{x}, \mathbf{x}') = \kappa(\mathbf{x}', \mathbf{x}) = \prod_{i=1}^{m} \sqrt{\frac{2\tilde{\sigma}_{\omega_i}(\mathbf{x})\tilde{\sigma}_{\omega_i}(\mathbf{x}')}{\tilde{\sigma}^2_{\omega_i}(\mathbf{x}) + \tilde{\sigma}^2_{\omega_i}(\mathbf{x}')}} \leq 1 \text{ (equality if } \mathbf{x} = \mathbf{x}')$$
$$l_{\omega_i}(\mathbf{x}, \mathbf{x}') = l_{\omega_i}(\mathbf{x}', \mathbf{x}) = \frac{\alpha_{\omega_i}}{\sqrt{2}} \sqrt{\tilde{\sigma}^2_{\omega_i}(\mathbf{x}) + \tilde{\sigma}^2_{\omega_i}(\mathbf{x}')}$$

and we implicitly let μ_{ω_i} and $\tilde{\sigma}_{\omega_i}$ be part of the kernel definition. Recall that:

$$\tilde{\sigma}_{\omega_i}(\mathbf{x}) = s_{\omega_i}\sigma_{\omega_i}(\mathbf{x}) + b_{\omega_i}$$

such that $\tilde{\sigma}_{\omega_i}(\mathbf{x}) \in (0, 1]\ \forall \mathbf{x} \in \mathbb{X}$. We wish to bound the maximum information gain γ_T characterizing this kernel, specifically:

$$\gamma_T \leq \frac{\frac{1}{2}}{1 - \frac{1}{e}} \max_{(m_t : \sum_t m_t = T)} \sum_{t=1}^{|D|} \log\left(1 + \sigma^{-2} m_t \lambda_t\right)$$

where $\lambda_1 \geq \lambda_2 \geq \ldots$ are the eigenvalues of $\mathbf{K}_T$ for observations $\{(\mathbf{x}_t, y_t) : 1 \leq t \leq T\}$ of the objective.

Our approach is in two phases. First, recall the inequality of Weyl, and assuming ordered eigenvalues with largest first for positive definite $\mathbf{K}$, $\mathbf{K}'$:

$$|\lambda_i(\mathbf{K}) - \lambda_i(\mathbf{K}')| \leq \|\mathbf{K} - \mathbf{K}'\|_2$$

where the spectral norm is on the right. Starting from the augmented kernel we apply two approximations to (a) replace the variable lengthscale with a discontinuous first-order approximation and, subsequently, (b) drop irrelevant dimensions from the kernel. Having suitably lowered the dimension, we draw on existing results to give convergence for γ_T in terms of the (reduced) dimension.

We assume throughout that:

1. Bounded inputs: $\mathbf{x} \in [-1, 1]^D$.
2. All kernels (objective model and augmentation models) are Lipschitz.

Before proceeding we have the following preliminary result

Lemma 1 *Consider the posterior given observations* $\{(\mathbf{x}_t, y_t) : t \leq T\}$. *Define* $K_t(\mathbf{x}) = K(\mathbf{x}, \mathbf{x}_t)$, $D_{T,t}(\mathbf{x}) = K(\mathbf{x}, \mathbf{x}) - |K_t(\mathbf{x})|$ *and* $D_{T,\downarrow}(\mathbf{x}) = \min_{t \leq T}\{D_{T,t}(\mathbf{x})\}$, $D_{T,\uparrow}(\mathbf{x}) = \max_{t \leq T}\{D_{T,t}(\mathbf{x})\}$. *If* $K(\mathbf{x}, \mathbf{x}') > 0$ *and* $K(\mathbf{x}, \mathbf{x}) = 1$. *If* $K(\mathbf{x}, \mathbf{x}') > 0$ *and* $K(\mathbf{x}, \mathbf{x}) = 1$ *then the posterior variance* σ_T *is bounded as:*

$$\sigma_{T,\downarrow}(\mathbf{x}) \leq \sigma_T(\mathbf{x}) \leq \sigma_{T,\uparrow}(\mathbf{x})$$

where:

$$\begin{aligned}
\sigma^2_{T,\downarrow}(\mathbf{x}) &= \tfrac{\sigma^2}{\sigma^2+T} \sim \Omega\left(\tfrac{1}{T}\right) \\
\sigma^2_{T,\uparrow}(\mathbf{x}) &= \tfrac{\sigma^2+TD_{T,\uparrow}(\mathbf{x})(2-D_{T,\uparrow}(\mathbf{x}))}{\sigma^2+T} \sim \mathcal{O}\left(D_{T,\uparrow}(\mathbf{x})\left(2 - D_{T,\uparrow}(\mathbf{x})\right)\right)
\end{aligned}$$

If $D_{T,\uparrow} \sim \mathcal{O}(g(T))$, $g(T) > 0$, $g(T) \to 0$, *then* $\sigma^2_T(\mathbf{x}) \sim \mathcal{O}(g(T))$.

Proof. Using the definition of GP posterior variance and K bounds we can minimize the posterior variance within the constraints given as:

$$\begin{aligned}
\sigma^2_T(\mathbf{x}) &\geq K(\mathbf{x}, \mathbf{x}) - K^2_\uparrow(\mathbf{x})\, \mathbf{1}_T^{\mathrm{T}} \left(K_{\uparrow,\uparrow}\mathbf{1}_T\mathbf{1}_T^{\mathrm{T}} + \sigma^2\mathbf{I}_T\right)^{-1} \mathbf{1}_T \\
&= K(\mathbf{x}, \mathbf{x}) \left(1 - \tfrac{K^2_\uparrow(\mathbf{x})}{K(\mathbf{x},\mathbf{x})} \tfrac{T}{\sigma^2+TK_{\uparrow,\uparrow}}\right) \\
&= K(\mathbf{x}, \mathbf{x}) \left(1 - \tfrac{TK^2_\uparrow(\mathbf{x})}{\sigma^2 K(\mathbf{x},\mathbf{x})+TK_{\uparrow,\uparrow}K(\mathbf{x},\mathbf{x})}\right) \\
&= K(\mathbf{x}, \mathbf{x}) \left(\tfrac{\sigma^2 K(\mathbf{x},\mathbf{x})+TK_{\uparrow,\uparrow}K(\mathbf{x},\mathbf{x})-TK^2_\uparrow(\mathbf{x})}{\sigma^2 K(\mathbf{x},\mathbf{x})+TK_{\uparrow,\uparrow}K(\mathbf{x},\mathbf{x})}\right) \\
&= K(\mathbf{x}, \mathbf{x}) \left(\tfrac{\sigma^2 K(\mathbf{x},\mathbf{x})+TK_{\uparrow,\uparrow}K(\mathbf{x},\mathbf{x})-TK^2(\mathbf{x},\mathbf{x})+2TK(\mathbf{x},\mathbf{x})D_{T,\downarrow}(\mathbf{x})-TD^2_{T,\downarrow}(\mathbf{x})}{\sigma^2 K(\mathbf{x},\mathbf{x})+TK_{\uparrow,\uparrow}K(\mathbf{x},\mathbf{x})}\right) \\
&\geq K(\mathbf{x}, \mathbf{x}) \left(\tfrac{\sigma^2+2TD_{T,\downarrow}(\mathbf{x})-T\frac{1}{K(\mathbf{x},\mathbf{x})}D^2_{T,\downarrow}(\mathbf{x})}{\sigma^2+TK_{\uparrow,\uparrow}}\right)
\end{aligned}$$

where $K_\uparrow(\mathbf{x}) = \max_t\{|K(\mathbf{x}, \mathbf{x}_t)|\}$ and $K_{\uparrow,\uparrow} = \max_t\{|K(\mathbf{x}_t, \mathbf{x}_t)|\}$. For the upper bound, we pessimise using the maximum eigenvalue:

$$\sigma^2_T(\mathbf{x}) \leq K(\mathbf{x}, \mathbf{x}) - \frac{\sum_{t \leq T} K^2_t(\mathbf{x})}{\lambda_{\max}(\mathbf{K}_T)+\sigma^2}$$

and by Gershgorin's circle theorem $\lambda_{\max}(\mathbf{K}_T) \leq \sum_{t\leq T} K_{t,t}$, so:

$$\begin{aligned}
\sigma_T^2(\mathbf{x}) &\leq K(\mathbf{x},\mathbf{x}) - \frac{\sum_{t\leq T} K_t^2(\mathbf{x})}{\sigma^2+\sum_{t\leq T} K_{t,t}} \\
&= \frac{K(\mathbf{x},\mathbf{x})}{\sigma^2 K(\mathbf{x},\mathbf{x})+\sum_{t\leq T} K_{t,t}K(\mathbf{x},\mathbf{x})}(\sigma^2 K(\mathbf{x},\mathbf{x})+\sum_{t\leq T} K_{t,t}K(\mathbf{x},\mathbf{x})-\sum_{t\leq T} K_t^2(\mathbf{x})) \\
&= \frac{K(\mathbf{x},\mathbf{x})}{\sigma^2 K(\mathbf{x},\mathbf{x})+\sum_{t\leq T} K_{t,t}K(\mathbf{x},\mathbf{x})}\big(\sigma^2 K(\mathbf{x},\mathbf{x})+\sum_{t\leq T} K_{t,t}K(\mathbf{x},\mathbf{x})-TK^2(\mathbf{x},\mathbf{x})+... \\
&\qquad ...2K(\mathbf{x},\mathbf{x})\sum_{t\leq T} D_{T,t}(\mathbf{x})-\sum_{t\leq T} D_{T,t}^2(\mathbf{x})) \\
&\leq \frac{K(\mathbf{x},\mathbf{x})}{\sigma^2 K(\mathbf{x},\mathbf{x})+\sum_{t\leq T} K^2(\mathbf{x},\mathbf{x})}\big(\sigma^2 K(\mathbf{x},\mathbf{x})+\sum_{t\leq T} K^2(\mathbf{x},\mathbf{x})-TK^2(\mathbf{x},\mathbf{x})+... \\
&\qquad ...2K(\mathbf{x},\mathbf{x})\sum_{t\leq T} D_{T,t}(\mathbf{x})-\sum_{t\leq T} D_{T,t}^2(\mathbf{x})\big) \\
&\leq \frac{K(\mathbf{x},\mathbf{x})}{\sigma^2 K(\mathbf{x},\mathbf{x})+TK^2(\mathbf{x},\mathbf{x})}\big(\sigma^2 K(\mathbf{x},\mathbf{x})+2K(\mathbf{x},\mathbf{x})\sum_{t\leq T} D_{T,t}(\mathbf{x})-\sum_{t\leq T} D_{T,t}^2(\mathbf{x})\big) \\
&= \frac{\sigma^2+2\sum_{t\leq T} D_{T,t}(\mathbf{x})-\frac{1}{K(\mathbf{x},\mathbf{x})}\sum_{t\leq T} D_{T,t}^2(\mathbf{x})}{\sigma^2+TK(\mathbf{x},\mathbf{x})}K(\mathbf{x},\mathbf{x})
\end{aligned}$$

The final results follow by substitution.

So the convergence of the augmented mean follows the convergence of the coverage of input space by the expert, as measured by the worst-case similarity measure implied by the kernel.

10.1 Step 1: Local-Linear Kernel Approximation

Let $\{\mathbb{X}_n \subset \mathbb{X} : n \leq M\}$ for $\mathbb{X}$ be a finite ($M \in \mathbb{N}$), disjoint ($\mathbb{X}_n \cap \mathbb{X}_{n'} = \emptyset$), cover ($\cup_{n\leq M}\mathbb{X}_n = \mathbb{X}$) for $\mathbb{X}$. For some $f : \mathbb{X} \to \mathbb{R}$, define:

$$\begin{aligned}
&\bar{f}_{(n)} = f(\mathbf{x}_n) \\
&\text{where} : \bar{\mathbf{x}}_n = \frac{\int_{\mathbf{x}\in\mathbb{X}_n} \mathbf{x}d\mathbf{x}}{\int_{\mathbf{x}\in\mathbb{X}_n} d\mathbf{x}}
\end{aligned}$$

This is a zeroth-order approximation of f on $\mathbb{X}_n$. We also define an analogous first-order approximation of f on $\mathbb{X}_n$ using:

$$\begin{aligned}
&\hat{f}_{(n)}(\mathbf{x}) = \bar{f}_{(n)} + \mathbf{f}_{(n)}^{\mathrm{T}}(\mathbf{x}-\mathbf{x}_n) \\
&\text{where} : \mathbf{f}_{(n)} = \operatorname*{arginf}_{\mathbf{f}_{(n)}} \left\{ \left| \bar{f}_{(n)} + \mathbf{f}_{(n)}^{\mathrm{T}}(\mathbf{x}-\mathbf{x}_n) - f(\mathbf{x}) \right| \forall \mathbf{x} \in \mathbb{X}_n \right\}
\end{aligned}$$

We extend these to all $\mathbb{X}$ using:

$$\begin{aligned}
\bar{f}(\mathbf{x}) &= \bar{f}_{(\{\mathbf{x}\})} \\
\hat{f}(\mathbf{x}) &= \hat{f}_{(\{\mathbf{x}\})}(\mathbf{x})
\end{aligned}$$

where $\{\mathbf{x}\} = n|\mathbf{x} \in \mathbb{X}_n$. We will use $\hat{\mu}_{\omega_i}, \hat{\bar{\sigma}}_{\omega_i}$ to approximate the (normalized) posterior mean and (normalized) posterior standard deviation of the augmented model ω_i. Note that these models are piecewise linear with discontinuities at the cover boundaries, and moreover that the difference between these and the (normalized) posterior mean and variance are bounded. It follows from our

Lipschitz assumption on the kernel that there exist $\epsilon, \gamma \in (0,1)$ such that $|\mu_{\omega_i}(\mathbf{x}) - \hat{\mu}_{\omega_i}(\mathbf{x})| \leq \epsilon$ and $|\tilde{\sigma}^2_{\omega_i}(\mathbf{x}) - \bar{\tilde{\sigma}}^2_{\omega_i}(\mathbf{x})| \leq \epsilon^2 < \gamma^2 \tilde{\sigma}^2_{\omega_i}(\mathbf{x})\ \forall \mathbf{x} \in \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_T\},\ \forall i, n$.

For all $\mathbf{x} \in \mathbb{X}_n$, $\mathbf{x}' \in \mathbb{X}_{n'}$, define the (approximated) kernel on $\mathbb{X}_n \times \mathbb{X}_{n'}$ as:

$$\begin{aligned}\hat{K}_{(n,n')}(\mathbf{x}, \mathbf{x}') &= \kappa_{(n,n')} \exp\left(- \sum_{d=1}^{D} \frac{\left(x_d - x'_d\right)^2}{2l_d^2} - \sum_{i=1}^{M} \frac{\left(\hat{\mu}_{\omega_i(n)}(\mathbf{x}) - \hat{\mu}_{\omega_i(n')}(\mathbf{x}')\right)^2}{2l_{\omega_i(n,n')}} \right) \\ &= \kappa_{(n,n')} \exp\Bigg(\ldots \\ &\quad \ldots - \sum_{d=1}^{D} \frac{\left(x_d - x'_d\right)^2}{2l_d^2} - \sum_{i=1}^{M} \frac{\left(\left(\bar{\mu}_{\omega_i(n)} + \boldsymbol{\mu}^{\mathrm{T}}_{\omega_i(n)}\mathbf{x}\right) - \left(\bar{\mu}_{\omega_i(n')} + \boldsymbol{\mu}^{\mathrm{T}}_{\omega_i(n')}\mathbf{x}'\right)\right)^2}{2l^2_{\omega_i(n,n')}} \Bigg)\end{aligned}$$

where, using that the posterior variances are normalized:

$$\begin{aligned}\kappa_{(n,n')} &= \kappa_{(n',n)} = \prod_{i=1}^{M} \sqrt{\frac{2\bar{\sigma}_{\omega_i(n)}\bar{\sigma}_{\omega_i(n')}}{\bar{\sigma}^2_{\omega_i(n)} + \bar{\sigma}^2_{\omega_i(n')}}} \leq 1 \text{ (equality if } n = n') \\ l^2_{\omega_i(n,n')} &= l^2_{\omega_i(n',n)} = \tfrac{1}{2}\alpha^2_{\omega_i}\left(\bar{\sigma}^2_{\omega_i(n)} + \bar{\sigma}^2_{\omega_i(n')}\right) \leq \alpha^2_{\omega_i}\end{aligned}$$

Subsequently we extend the approximated kernel to the whole of $\mathbb{X}$ using:

$$\hat{K}(\mathbf{x}, \mathbf{x}') = \hat{K}_{(\{\mathbf{x}\},\{\mathbf{x}'\})}(\mathbf{x}, \mathbf{x}')$$

We wish to bound the difference between $\hat{K}$ and K on $\mathbb{X} \times \mathbb{X}$. Observing the fact that that $\kappa(\mathbf{x}, \mathbf{x}'), \kappa_{(n,n')}, \exp(-x^2) \leq 1$, it follows that:

$$\begin{aligned}&\left|K(\mathbf{x}_t, \mathbf{x}_{t'}) - \hat{K}(\mathbf{x}_t, \mathbf{x}_{t'})\right| \ldots \\ &\quad \ldots = \kappa(\mathbf{x}_t, \mathbf{x}_{t'}) \exp\left(- \sum_{d=1}^{D} \frac{\left(x_{t,d} - x_{t',d}\right)^2}{2l_d^2} - \sum_{i=1}^{M} \frac{\left(\mu_{\omega_i}(\mathbf{x}_t) - \mu_{\omega_i}(\mathbf{x}_{t'})\right)^2}{2l^2_{\omega_i}(\mathbf{x}_t, \mathbf{x}_{t'})} \right) \ldots \\ &\qquad - \kappa_{(\{\mathbf{x}_t\},\{\mathbf{x}_{t'}\})} \exp\Bigg(- \sum_{d=1}^{D} \frac{\left(x_{t,d} - x_{t',d}\right)^2}{2l_d^2} \ldots \\ &\qquad - \sum_{i=1}^{M} \frac{\left(\left(\bar{\mu}_{\omega_i(\{\mathbf{x}_t\})} + \boldsymbol{\mu}^{\mathrm{T}}_{\omega_i(\{\mathbf{x}_t\})}\mathbf{x}_t\right) - \left(\bar{\mu}_{\omega_i(\{\mathbf{x}_{t'}\})} + \boldsymbol{\mu}^{\mathrm{T}}_{\omega_i(\{\mathbf{x}_{t'}\})}\mathbf{x}_{t'}\right)\right)^2}{\ldots 2l^2_{\omega_i(\{\mathbf{x}_t\},\{\mathbf{x}_{t'}\})}} \Bigg) \\ &\quad \leq \left|\kappa(\mathbf{x}_t, \mathbf{x}_{t'}) - \kappa_{(\{\mathbf{x}_t\},\{\mathbf{x}_{t'}\})}\right| + \Bigg| \exp\Bigg(- \sum_{d=1}^{D} \frac{\left(x_{t,d} - x_{t',d}\right)^2}{2l_d^2} \ldots \\ &\qquad \ldots - \sum_{i=1}^{M} \frac{\left(\mu_{\omega_i}(\mathbf{x}_t) - \mu_{\omega_i}(\mathbf{x}_{t'})\right)^2}{2l^2_{\omega_i}(\mathbf{x}_t, \mathbf{x}_{t'})} \Bigg) - \exp\Bigg(- \sum_{d=1}^{D} \frac{\left(x_{t,d} - x_{t',d}\right)^2}{2l_d^2} \ldots \\ &\qquad \ldots - \sum_{i=1}^{M} \frac{\left(\left(\bar{\mu}_{\omega_i(\{\mathbf{x}_t\})} + \boldsymbol{\mu}^{\mathrm{T}}_{\omega_i(\{\mathbf{x}_t\})}\mathbf{x}_t\right) - \left(\bar{\mu}_{\omega_i(\{\mathbf{x}_{t'}\})} + \boldsymbol{\mu}^{\mathrm{T}}_{\omega_i(\{\mathbf{x}_{t'}\})}\mathbf{x}_{t'}\right)\right)^2}{2l^2_{\omega_i(\{\mathbf{x}_t\},\{\mathbf{x}_{t'}\})}} \Bigg)\Bigg|\end{aligned}$$

and also:

$$\begin{aligned}
\left|\kappa\left(\mathbf{x}_t,\mathbf{x}_{t'}\right)-\kappa_{(\{\mathbf{x}_t\},\{\mathbf{x}_{t'}\})}\right| &= \left|\prod_{i=1}^{M}\sqrt{\frac{2\tilde{\sigma}_{\omega_i}(\mathbf{x}_t)\tilde{\sigma}_{\omega_i}(\mathbf{x}_{t'})}{\tilde{\sigma}^2_{\omega_i}(\mathbf{x}_t)+\tilde{\sigma}^2_{\omega_i}(\mathbf{x}_{t'})}}-\prod_{i=1}^{M}\sqrt{\frac{2\bar{\sigma}_{\omega_i(n)}\bar{\sigma}_{\omega_i(n')}}{\bar{\sigma}^2_{\omega_i(n)}+\bar{\sigma}^2_{\omega_i(n')}}}\right| \\
&\leq \prod_{i=1}^{M}\sqrt{\frac{2\left(\tilde{\sigma}_{\omega_i}(\mathbf{x}_t)+\epsilon\right)\left(\tilde{\sigma}_{\omega_i}(\mathbf{x}_{t'})+\epsilon\right)}{\left(\tilde{\sigma}_{\omega_i}(\mathbf{x}_t)-\epsilon\right)^2+\left(\tilde{\sigma}_{\omega_i}(\mathbf{x}_{t'})-\epsilon\right)^2}}-\prod_{i=1}^{M}\sqrt{\frac{2\tilde{\sigma}_{\omega_i}(\mathbf{x}_t)\tilde{\sigma}_{\omega_i}(\mathbf{x}_{t'})}{\tilde{\sigma}^2_{\omega_i}(\mathbf{x}_t)+\tilde{\sigma}^2_{\omega_i}(\mathbf{x}_{t'})}} \\
&\leq \sqrt{\prod_{i=1}^{M}\frac{2\left(\tilde{\sigma}_{\omega_i}(\mathbf{x}_t)+\epsilon\right)\left(\tilde{\sigma}_{\omega_i}(\mathbf{x}_{t'})+\epsilon\right)}{\left(\tilde{\sigma}_{\omega_i}(\mathbf{x}_t)-\epsilon\right)^2+\left(\tilde{\sigma}_{\omega_i}(\mathbf{x}_{t'})-\epsilon\right)^2}-\prod_{i=1}^{M}\frac{2\tilde{\sigma}_{\omega_i}(\mathbf{x}_t)\tilde{\sigma}_{\omega_i}(\mathbf{x}_{t'})}{\tilde{\sigma}^2_{\omega_i}(\mathbf{x}_t)+\tilde{\sigma}^2_{\omega_i}(\mathbf{x}_{t'})}} \\
&\leq \sqrt{\left(\prod_{i=1}^{M}\frac{2\tilde{\sigma}_{\omega_i}(\mathbf{x}_t)\tilde{\sigma}_{\omega_i}(\mathbf{x}_{t'})}{\tilde{\sigma}^2_{\omega_i}(\mathbf{x}_t)+\tilde{\sigma}^2_{\omega_i}(\mathbf{x}_{t'})}\right)\left(\prod_{i=1}^{M}\frac{2(1+\gamma)(1+\gamma)}{\frac{(\tilde{\sigma}_{\omega_i}(\mathbf{x}_t)-\epsilon)^2+(\tilde{\sigma}_{\omega_i}(\mathbf{x}_{t'})-\epsilon)^2}{\tilde{\sigma}^2_{\omega_i}(\mathbf{x}_t)+\tilde{\sigma}^2_{\omega_i}(\mathbf{x}_{t'})}}-1\right)} \\
&\leq \sqrt{\left(\prod_{i=1}^{M}\frac{2\tilde{\sigma}_{\omega_i}(\mathbf{x}_t)\tilde{\sigma}_{\omega_i}(\mathbf{x}_{t'})}{\tilde{\sigma}^2_{\omega_i}(\mathbf{x}_t)+\tilde{\sigma}^2_{\omega_i}(\mathbf{x}_{t'})}\right)\left(\prod_{i=1}^{M}\frac{(1+\gamma)^2}{(1-\gamma)^2}-1\right)} \\
&= \kappa\left(\mathbf{x}_t,\mathbf{x}_{t'}\right)\sqrt{\left(\frac{1+\gamma}{1-\gamma}\right)^{2M}-1} \\
&\leq \sqrt{\left(\frac{1+\gamma}{1-\gamma}\right)^{2M}-1}
\end{aligned}$$

Note that:

$$\begin{aligned}
\left|l^2_{\omega_i}\left(\mathbf{x},\mathbf{x}'\right)-l^2_{\omega_i(\{\mathbf{x}\},\{\mathbf{x}'\})}\right| &= \Big|\tfrac{1}{2}\alpha^2_{\omega_i}\left(\tilde{\sigma}^2_{\omega_i}\left(\mathbf{x}\right)+\tilde{\sigma}^2_{\omega_i}\left(\mathbf{x}'\right)\right)-\ldots \\
&\quad \ldots\tfrac{1}{2}\alpha^2_{\omega_i}\left(\bar{\sigma}^2_{\omega_i(\{\mathbf{x}\})}+\bar{\sigma}^2_{\omega_i(\{\mathbf{x}'\})}\right)\Big| \\
&\leq \tfrac{1}{2}\alpha^2_{\omega_i}\Big(\left|\tilde{\sigma}^2_{\omega_i}\left(\mathbf{x}\right)-\bar{\sigma}^2_{\omega_i(\{\mathbf{x}\})}\right|+\ldots \\
&\quad \ldots\left|\tilde{\sigma}^2_{\omega_i}\left(\mathbf{x}'\right)-\bar{\sigma}^2_{\omega_i(\{\mathbf{x}'\})}\right|\Big) \\
&\leq \alpha^2_{\omega_i}\epsilon^2
\end{aligned}$$

Furthermore, $\exp(-x^2)$ is $L = \frac{2\sqrt{2}}{e^2}$-Lipchitz,[1], so:

$$\begin{aligned}
&\left|\exp\left(-\sum_{d=1}^{D}\frac{\left(x_d-x'_d\right)^2}{2l_d^2}-\sum_{i=1}^{M}\frac{\left(\left(\bar{\mu}_{\omega_i(\{\mathbf{x}\})}+\boldsymbol{\mu}^{\mathrm{T}}_{(\{\mathbf{x}\})}\mathbf{x}\right)-\left(\bar{\mu}_{\omega_i(\{\mathbf{x}'\})}+\boldsymbol{\mu}^{\mathrm{T}}_{(\{\mathbf{x}'\})}\mathbf{x}'\right)\right)^2}{2l^2_{\omega_i(\{\mathbf{x}\},\{\mathbf{x}'\})}}\right)\dots\right.\\
&\qquad\left.\dots-\exp\left(-\sum_{d=1}^{D}\frac{\left(x_d-x'_d\right)^2}{2l_d^2}-\sum_{i=1}^{M}\frac{\left(\mu_{\omega_i}(\mathbf{x})-\mu_{\omega_i}\left(\mathbf{x}'\right)\right)^2}{2l^2_{\omega_i}(\mathbf{x},\mathbf{x}')}\right)\right|\dots\\
&\dots\leq\left|\exp\left(-\sum_{d=1}^{D}\frac{\left(x_d-x'_d\right)^2}{2l_d^2}-\sum_{i=1}^{M}\frac{\left(\left(\bar{\mu}_{\omega_i(\{\mathbf{x}\})}+\boldsymbol{\mu}^{\mathrm{T}}_{(\{\mathbf{x}\})}\mathbf{x}\right)-\left(\bar{\mu}_{\omega_i(\{\mathbf{x}'\})}+\boldsymbol{\mu}^{\mathrm{T}}_{(\{\mathbf{x}'\})}\mathbf{x}'\right)\right)^2}{2l^2_{\omega_i(\{\mathbf{x}\},\{\mathbf{x}'\})}}\right)\dots\right.\\
&\qquad\left.\dots-\exp\left(-\sum_{d=1}^{D}\frac{\left(x_d-x'_d\right)^2}{2l_d^2}-\sum_{i=1}^{M}\frac{\left(\mu_{\omega_i}(\mathbf{x})-\mu_{\omega_i}\left(\mathbf{x}'\right)\right)^2}{2l^2_{\omega_i(\{\mathbf{x}\},\{\mathbf{x}'\})}}\right)\right|\dots\\
&\qquad\dots+\left|\exp\left(-\sum_{d=1}^{D}\frac{\left(x_d-x'_d\right)^2}{2l_d^2}-\sum_{i=1}^{M}\frac{\left(\mu_{\omega_i}(\mathbf{x})-\mu_{\omega_i}\left(\mathbf{x}'\right)\right)^2}{2l^2_{\omega_i(\{\mathbf{x}\},\{\mathbf{x}'\})}}\right)-\dots\right.\\
&\qquad\left.\dots\exp\left(-\sum_{d=1}^{D}\frac{\left(x_d-x'_d\right)^2}{2l_d^2}-\sum_{i=1}^{M}\frac{\left(\mu_{\omega_i}(\mathbf{x})-\mu_{\omega_i}\left(\mathbf{x}'\right)\right)^2}{2l^2_{\omega_i}(\mathbf{x},\mathbf{x}')}\right)\right|\dots\\
&\dots\leq\prod_{i=1}^{M}\frac{L}{\sqrt{2}}\left|\frac{\left(\bar{\mu}_{\omega_i(\{\mathbf{x}\})}+\boldsymbol{\mu}^{\mathrm{T}}_{(\{\mathbf{x}\})}\mathbf{x}\right)-\left(\bar{\mu}_{\omega_i(\{\mathbf{x}'\})}+\boldsymbol{\mu}^{\mathrm{T}}_{(\{\mathbf{x}'\})}\mathbf{x}'\right)}{l_{\omega_i(\{\mathbf{x}\},\{\mathbf{x}'\})}}-\frac{\mu_{\omega_i}(\mathbf{x})-\mu_{\omega_i}\left(\mathbf{x}'\right)}{l_{\omega_i(\{\mathbf{x}\},\{\mathbf{x}'\})}}\right|\dots\\
&\qquad\dots+\prod_{i=1}^{M}\frac{L}{\sqrt{2}}\left|\frac{\mu_{\omega_i}(\mathbf{x})-\mu_{\omega_i}\left(\mathbf{x}'\right)}{l_{\omega_i(\{\mathbf{x}\},\{\mathbf{x}'\})}}-\frac{\mu_{\omega_i}(\mathbf{x})-\mu_{\omega_i}\left(\mathbf{x}'\right)}{l_{\omega_i}(\mathbf{x},\mathbf{x}')}\right|\dots\\
&\dots\leq\prod_{i=1}^{M}\frac{2L\epsilon}{\sqrt{2}l_{\omega_i(\{\mathbf{x}\},\{\mathbf{x}'\})}}+\prod_{i=1}^{M}\frac{L\left|\mu_{\omega_i}(\mathbf{x})-\mu_{\omega_i}\left(\mathbf{x}'\right)\right|}{\sqrt{2}}\left|\frac{1}{l_{\omega_i(\{\mathbf{x}\},\{\mathbf{x}'\})}}-\frac{1}{l_{\omega_i(\{\mathbf{x}\},\{\mathbf{x}'\})}+\alpha_{\omega_i}\epsilon}\right|\dots\\
&\dots=\prod_{i=1}^{M}\frac{2L\epsilon}{\sqrt{2}l_{\omega_i(\{\mathbf{x}\},\{\mathbf{x}'\})}}+\prod_{i=1}^{M}\frac{L\left|\mu_{\omega_i}(\mathbf{x})-\mu_{\omega_i}\left(\mathbf{x}'\right)\right|}{\sqrt{2}l_{\omega_i(\{\mathbf{x}\},\{\mathbf{x}'\})}}\left|\frac{\alpha_{\omega_i}\epsilon}{l_{\omega_i(\{\mathbf{x}\},\{\mathbf{x}'\})}+\alpha_{\omega_i}\epsilon}\right|\dots\\
&\dots\leq\prod_{i=1}^{M}\frac{L2\epsilon}{\sqrt{2}l_{\omega_i(\{\mathbf{x}\},\{\mathbf{x}'\})}}+\prod_{i=1}^{M}\frac{L\alpha_{\omega_i}\epsilon\left|\mu_{\omega_i}(\mathbf{x})-\mu_{\omega_i}\left(\mathbf{x}'\right)\right|}{\sqrt{2}l_{\omega_i(\{\mathbf{x}\},\{\mathbf{x}'\})}}
\end{aligned}$$

Using that the augmentation is normalized in mean:

$$\begin{aligned}
&\left|\exp\left(-\sum_{d=1}^{D}\frac{\left(x_d-x'_d\right)^2}{2l_d^2}-\sum_{i=1}^{M}\frac{\left(\left(\bar{\mu}_{\omega_i(\{\mathbf{x}\})}+\boldsymbol{\mu}^{\mathrm{T}}_{(\{\mathbf{x}\})}\mathbf{x}\right)-\left(\bar{\mu}_{\omega_i(\{\mathbf{x}'\})}+\boldsymbol{\mu}^{\mathrm{T}}_{(\{\mathbf{x}'\})}\mathbf{x}'\right)\right)^2}{2l^2_{\omega_i(\{\mathbf{x}\},\{\mathbf{x}'\})}}\right)\dots\right.\\
&\left.\dots-\exp\left(-\sum_{d=1}^{D}\frac{\left(x_d-x'_d\right)^2}{2l_d^2}-\sum_{i=1}^{M}\frac{\left(\mu_{\omega_i}(\mathbf{x})-\mu_{\omega_i}\left(\mathbf{x}'\right)\right)^2}{2l^2_{\omega_i}(\mathbf{x},\mathbf{x}')}\right)\right|\dots\\
&\dots\leq\prod_{i=1}^{M}\frac{\sqrt{2}L\left(\epsilon+\alpha_{\omega_i}\epsilon\right)}{l_{\omega_i(\{\mathbf{x}\},\{\mathbf{x}'\})}}\leq\prod_{i=1}^{M}\frac{\sqrt{2}L\epsilon\left(1+\alpha_{\omega_i}\right)}{\sqrt{\frac{\alpha^2_{\omega_i}\tilde{\sigma}^2}{\tilde{\sigma}^2+T}}}\leq\left(L\epsilon\left(1+\alpha_{\min}\right)\sqrt{2\frac{\tilde{\sigma}^2+T}{\alpha^2_{\min}\tilde{\sigma}^2}}\right)^M\leq\dots\\
&\qquad\dots\left(\frac{4}{e^2}\sqrt{\frac{\tilde{\sigma}^2+T}{\tilde{\sigma}^2}}\frac{1+\alpha_{\min}}{\alpha_{\min}}\epsilon\right)^M
\end{aligned}$$

where by Lemma 1 we have used:

$$l^2_{\omega_i(n,n')}=\tfrac{1}{2}\alpha^2_{\min}\left(\bar{\sigma}^2_{\omega_i(n)}+\bar{\sigma}^2_{\omega_i(n')}\right)\geq\frac{\alpha^2_{\min}\tilde{\sigma}^2}{\tilde{\sigma}^2+T}$$

and thus:

$$\left|K\left(\mathbf{x},\mathbf{x}'\right)-\hat{K}\left(\mathbf{x},\mathbf{x}'\right)\right|\leq\hat{K}_{\Delta}=\sqrt{\left(\frac{1+\gamma}{1-\gamma}\right)^{2M}-1}+\left(\frac{4}{e^2}\sqrt{\frac{\tilde{\sigma}^2+T}{\tilde{\sigma}^2}}\frac{1+\alpha_{\min}}{\alpha_{\min}}\epsilon\right)^M$$

[1] Here $\frac{d}{dx}\exp(-x^2)=-2x\exp(-x^2)$, $\frac{d^2}{dx^2}\exp(-x^2)=(4x^2-2)\exp(-x^2)$, so the gradient maximum occurs at $x=\sqrt{2}$ with value $-\frac{2\sqrt{2}}{e^2}$.

Using the bound on the kernel approximation we obtain the spectral norm bound:

$$\left\|\mathbf{K}_T - \hat{\mathbf{K}}_T\right\|_2 = \sup_{\|\mathbf{x}\|_2=1}\left\{\left\|\left(\mathbf{K}_T - \hat{\mathbf{K}}_T\right)\mathbf{x}\right\|_2\right\} \leq \sup_{\|\mathbf{x}\|_2=1}\left\{\left\|\hat{K}_\Delta \mathbf{1}_T \mathbf{1}_T^{\mathrm{T}}\mathbf{x}\right\|_2\right\} = \hat{K}_\Delta T$$
$$\leq T\sqrt{\left(\tfrac{1+\gamma}{1-\gamma}\right)^{2M} - 1} + T\left(\tfrac{4}{e^2}\sqrt{\tfrac{\tilde{\sigma}^2+T}{\tilde{\sigma}^2}}\tfrac{1+\alpha_{\min}}{\alpha_{\min}}\epsilon\right)^M$$

which can be made arbitrarily small by letting $\epsilon, \gamma \to 0$. Recall our assumption $\epsilon^2 < \gamma^2\tilde{\sigma}^2_{\omega_i}(\mathbf{x})\ \forall \mathbf{x} \in \mathbb{X}, \forall i, n$. It suffices that $\epsilon^2 = \gamma^2\frac{\tilde{\sigma}^2}{\tilde{\sigma}^2+T}$, so:

$$\left\|\mathbf{K}_T - \hat{\mathbf{K}}_T\right\|_2 \leq T\sqrt{\left(\tfrac{1+\gamma}{1-\gamma}\right)^{2M} - 1} + T\left(\tfrac{4}{e^2}\tfrac{1+\alpha_{\min}}{\alpha_{\min}}\gamma\right)^M$$

Finally for this approximation, defining $\gamma = \frac{\hat{\gamma}}{T^{\frac{1}{M}}}$ (where $0 < \hat{\gamma} < 1$):[2].

$$\left\|\mathbf{K}_T - \hat{\mathbf{K}}_T\right\|_2 \leq \left(T\sqrt{\left(\left(\frac{1+\frac{\hat{\gamma}}{T^{1/M}}}{1-\frac{\hat{\gamma}}{T^{1/M}}}\right)^{2M} - 1\right)} + \left(\tfrac{4}{e^2}\tfrac{1+\alpha_{\min}}{\alpha_{\min}}\hat{\gamma}\right)^M\right) \sim \mathcal{O}(1)$$

10.2 Step 2: Dimensional Pruning Kernel Approximation

Our next step is to study the effect of removing features - real or augmented - from the kernel. To this end, define:

$$\mathbb{D} \subseteq \mathbb{N}_D$$
$$\mathbb{I} \subseteq \mathbb{N}_M$$

and $d_{\text{eff}} = \#(\mathbb{D}) + \#(\mathbb{I})$. Assume that these are the features (primary or augmented) that will dominate (in the sense that the corresponding lengthscale becomes "large") as $T \to \infty$. Define:

$$\breve{K}_{(n,n')}(\mathbf{x}, \mathbf{x}') = \hat{K}_{(n,n')}(\mathbf{x}, \mathbf{x}')\exp\left(\sum_{d\notin\mathbb{D}}\frac{\left(x_d - x'_d\right)^2}{2l_d^2} + \ldots\right.$$
$$\left.\ldots\sum_{i\notin\mathbb{I}}\frac{\left(\left(\bar{\mu}_{\omega_i(n)} + \boldsymbol{\mu}^{\mathrm{T}}_{(n)}\mathbf{x}\right) - \left(\bar{\mu}_{\omega_i(n')} + \boldsymbol{\mu}^{\mathrm{T}}_{(n')}\mathbf{x}'\right)\right)^2}{2l^2_{\omega_i(n,n')}}\right)$$

Using the Lipschitz properties of $\exp(-x^2)$ discussed previously:

$$\exp\left(\sum_{d\notin\mathbb{D}}\frac{\left(x_d - x'_d\right)^2}{2l_d^2} + \sum_{i\notin\mathbb{I}}\frac{\left(\left(\bar{\mu}_{\omega_i(n)} + \boldsymbol{\mu}^{\mathrm{T}}_{(n)}\mathbf{x}\right) - \left(\bar{\mu}_{\omega_i(n')} + \boldsymbol{\mu}^{\mathrm{T}}_{(n')}\mathbf{x}'\right)\right)^2}{2l^2_{\omega_i(n,n')}}\right)\ldots$$
$$\leq \kappa_{(n,n')}\tfrac{2\sqrt{2}}{e^2}\left(\sum_{d\notin\mathbb{D}}\frac{\left(x_d - x'_d\right)^2}{2l_d^2} + \sum_{i\notin\mathbb{I}}\frac{\left(\left(\bar{\mu}_{\omega_i(n)} + \boldsymbol{\mu}^{\mathrm{T}}_{(n)}\mathbf{x}\right) - \left(\bar{\mu}_{\omega_i(n')} + \boldsymbol{\mu}^{\mathrm{T}}_{(n')}\mathbf{x}'\right)\right)^2}{2l^2_{\omega_i(n,n')}}\right)$$
$$\leq \tfrac{4\sqrt{2}}{e^2}\left(\sum_{d\notin\mathbb{D}}\frac{1}{l_d^2} + \sum_{i\notin\mathbb{I}}\frac{(1+\epsilon)^2}{l^2_{\omega_i(n,n')}}\right)$$
$$\leq \tfrac{4\sqrt{2}}{e^2}(1+\epsilon)^2\left(\sum_{d\notin\mathbb{D}}\frac{1}{l_d^2} + \sum_{i\notin\mathbb{I}}\frac{1}{l^2_{\omega_i(n,n')}}\right)$$

[2] Taylor expanding, $\left(\frac{1+a\gamma}{1-a\gamma}\right)^{\frac{2}{M}} = 1 + \frac{4}{M}a\gamma + \ldots o((a\gamma)^2)$. We neglect the potential influence of M dependence on T in the asymptote.

and hence, using $\epsilon^2 = \gamma^2 \frac{\tilde{\sigma}^2}{\tilde{\sigma}^2+T}$:

$$\exp\left(\sum_{d\notin\mathbb{D}} \frac{\left(x_d - x'_d\right)^2}{2l_d^2} + \sum_{i\notin\mathbb{I}} \frac{\left(\left(\bar{\mu}_{\omega_i(n)} + \boldsymbol{\mu}^{\mathrm{T}}_{(n)}\mathbf{x}\right) - \left(\bar{\mu}_{\omega_i(n')} + \boldsymbol{\mu}^{\mathrm{T}}_{(n')}\mathbf{x}'\right)\right)^2}{2l^2_{\omega_i(n,n')}}\right)\dots$$
$$\leq \breve{K}_{\Delta(n,n')} = d_{\text{eff}} \frac{4\sqrt{2}}{e^2}\left(1 + \frac{\hat{\gamma}}{T^{\frac{1}{M}}}\sqrt{\frac{\tilde{\sigma}^2}{\tilde{\sigma}^2+T}}\right)^2 \left(\frac{1}{d_{\text{eff}}}\left(\sum_{d\notin\mathbb{D}}\frac{1}{l_d^2} + \sum_{i\notin\mathbb{I}}\frac{1}{l^2_{\omega_i(n,n')}}\right)\right)$$
$$\leq \breve{K}_{\Delta(n,n')} = \frac{4\sqrt{2}}{e^2}\left(1 + \frac{\hat{\gamma}}{T^{\frac{1}{M}}}\sqrt{\frac{\tilde{\sigma}^2}{\tilde{\sigma}^2+T}}\right)^2 \frac{d_{\text{eff}}}{l^2_{\max}}$$

where:

$$l_{\max} = \min\left\{\min_{d\notin\mathbb{D}}\left\{l_d\right\}, \min_{i\notin\mathbb{I}}\left\{l_{\omega_i(n,n')}\right\}\right\}$$

Recall that implicitly here $\mu_{\omega_i,T}$, $\sigma_{\omega_i,T}$ are the posterior mean and variance for the underlying augmenting feature ω_i. We construct the normalized posterior variance $\tilde{\sigma}_{\omega_i,T}(\mathbf{x}) = s_{\omega_i,T}\sigma_{\omega_i,T}(\mathbf{x}) + b_{\omega_i,T}$, and the lengthscale function is $l^2_{i,T}(\mathbf{x},\mathbf{x}') = \frac{1}{2}\alpha_{\omega_i T}(\tilde{\sigma}^2_{\omega_i,T}(\mathbf{x}) + \tilde{\sigma}^2_{\omega_i,T}(\mathbf{x}'))$, where $\alpha_{\omega_i T}$ is selected (along with other kernel parameters) to maximize log-marginal likelihood. We have from Lemma 1 that $\tilde{\sigma}^2_{\omega_i,T}(\mathbf{x}) \sim \mathcal{O}(g(T))$, where $g(T)$ is a decreasing surrogate of the rate of human expert feedback (this also governs the rate at which $\mu_{\omega_i} \to \omega_i$). Thus the normalized posterior variance $\tilde{\sigma}_{\omega_i,T}(\mathbf{x}) \to b_{\omega_i,T}$ as $T \to \infty$, with the difference converging to zero at $\mathcal{O}(g(T))$. So $l_{d,T} \to l_d\ \forall d$, $\alpha_{\omega_i T} \to \alpha_{\omega_i}$, $b_{\omega_i,T} \to b_{\omega_i}$, $s_{\omega_i,T} \to s_{\omega_i}$ as $T \to \infty$; and subsequently $l_{\omega_i,T}(\mathbf{x},\mathbf{x}')^2 \to l^2_{\omega_i} = \alpha_{\omega_i T} b^2_{\omega_i}$. Subsequently:

$$\exp\left(\sum_{d\notin\mathbb{D}} \frac{\left(x_d - x'_d\right)^2}{2l_d^2} + \sum_{i\notin\mathbb{I}} \frac{\left(\left(\bar{\mu}_{\omega_i(\{\mathbf{x}\})} + \boldsymbol{\mu}^{\mathrm{T}}_{(\{\mathbf{x}\})}\mathbf{x}\right) - \left(\bar{\mu}_{\omega_i(\{\mathbf{x}'\})} + \boldsymbol{\mu}^{\mathrm{T}}_{(\{\mathbf{x}'\})}\mathbf{x}'\right)\right)^2}{2l^2_{\omega_i(\{\mathbf{x}\},\{\mathbf{x}'\})}}\right)\dots$$
$$\overset{T\to\infty}{\leq} \frac{4\sqrt{2}}{e^2}\left(1 + \gamma\sqrt{\frac{\tilde{\sigma}^2}{\tilde{\sigma}^2+T}}\right)\frac{d_{\text{eff}}}{l^2_{\max}}$$

with convergence governed by $g(T)$ - ie. the rate at which the human expert is able to provide feedback to build the augmenting model. Subsequently:

$$\left\|\mathbf{K}_T - \breve{\mathbf{K}}_T\right\|_2 \overset{T\to\infty}{\leq} \left(\left(\frac{8}{M}\right)^M + \left(\frac{8}{e^2}\frac{1+\alpha_{\min}}{\alpha_{\min}}\right)^M\right)\left(\frac{\hat{\gamma}}{2}\right)^M \frac{4\sqrt{2}}{e^2}\frac{d_{\text{eff}}}{l^2_{\max}}$$

Cleaning up, if we let $\hat{\gamma} = \frac{4\sqrt{2}}{e^2}((\frac{8}{M})^M + (\frac{8}{e^2}\frac{1+\alpha_{\min}}{\alpha_{\min}})^M)^{\frac{1}{M}} 2\tilde{\gamma}^{\frac{1}{M}}$, where $\tilde{\gamma} > 0$ then:

$$\left\|\mathbf{K}_T - \breve{\mathbf{K}}_T\right\|_2 \overset{T\to\infty}{\leq} \tilde{\gamma}\frac{d_{\text{eff}}}{l^2_{\max}}$$

Note that the bound does not go to zero as M - the size of the cover - grows as ϵ decreases, and ϵ is a function of $\tilde{\gamma}$.

10.3 Maximum Information Gain Bound

We have from the previous section that, and by Weyl's bound:

$$\left|\lambda_i - \breve{\lambda}_i\right| \leq \tilde{\gamma}\frac{d_{\text{eff}}}{l^2_{\max}}$$

where λ_i are the eigenvalues of $\mathbf{K}_T$ and $\breve{\lambda}_i$ the eigenvalues of $\breve{\mathbf{K}}_T$. Recall that:

$$\gamma_T \leq \frac{\frac{1}{2}}{1-\frac{1}{e}} \max_{(m_t:\sum_t m_t=T)} \sum_{t=1}^{|D|} \log\left(1+\sigma^{-2} m_t \lambda_t\right)$$

The asymptotic behavior of this bound is governed by the dimension $D+M$ of the combined raw and augmented features. However, using our eigenvalue bound, we see that:

$$\begin{aligned}
\gamma_T &\overset{T\to\infty}{\leq} \frac{\frac{1}{2}}{1-\frac{1}{e}} \max_{(m_t:\sum_t m_t=T)} \sum_{t=1}^{|D|} \log\left(1+\sigma^{-2} m_t \lambda_t\right) \\
&\overset{T\to\infty}{\leq} \frac{\frac{1}{2}}{1-\frac{1}{e}} \max_{(m_t:\sum_t m_t=T)} \sum_{t=1}^{|D|} \log\left(1+\sigma^{-2} m_t \breve{\lambda}_t + m_t \tilde{\gamma} \frac{d_{\text{eff}}}{\tilde{l}^2_{\max}}\right) \\
&\overset{T\to\infty}{\leq} \frac{\frac{1}{2}}{1-\frac{1}{e}} \max_{(m_t:\sum_t m_t=T)} \sum_{t=1}^{|D|} \left(\log\left(1+\sigma^{-2} m_t \breve{\lambda}_t\right) + \frac{m_t}{1+\sigma^{-2} m_t \breve{\lambda}_t} \tilde{\gamma} \frac{d_{\text{eff}}}{\tilde{l}^2_{\max}}\right) \\
&\overset{T\to\infty}{\leq} \frac{\frac{1}{2}}{1-\frac{1}{e}} \max_{(m_t:\sum_t m_t=T)} \sum_{t=1}^{|D|} \left(\log\left(1+\sigma^{-2} m_t \breve{\lambda}_t\right) + m_t \tilde{\gamma} \frac{d_{\text{eff}}}{\tilde{l}^2_{\max}}\right) \\
&\overset{T\to\infty}{\leq} \frac{\frac{1}{2}}{1-\frac{1}{e}} \max_{(m_t:\sum_t m_t=T)} \sum_{t=1}^{|D|} \left(\log\left(1+\sigma^{-2} m_t \breve{\lambda}_t\right)\right) + |D| \tilde{\gamma} \frac{d_{\text{eff}}}{\tilde{l}^2_{\max}} \\
&\overset{T\to\infty}{=} \breve{\gamma}_T + \frac{\frac{1}{2}}{1-\frac{1}{e}} |D| \tilde{\gamma} \frac{d_{\text{eff}}}{\tilde{l}^2_{\max}}
\end{aligned}$$

whose asymptotic behavior is governed by d_{eff}. So:

$$|\gamma_T - \breve{\gamma}_T| = \mathcal{O}(d_{\text{eff}})$$

and $\breve{\gamma}_T$ is governed by d_{eff} by definition.

11 Experimental Details

11.1 Parameter Selection in BOAP framework

The kernel functions (k) used in fitting a Gaussian process model for the unknown objective function f is associated with its own hyperparameter set θ. The optimal kernel hyperparameters ($\theta^\star$) are estimated by maximizing the marginal likelihood function, given by the equation:

$$\mathcal{P}(\mathbf{y}|\mathbf{X},\theta) = \int p(\mathbf{y}|f)\; p(f|\mathbf{X},\theta)\; df \tag{22}$$

By marginalizing Eq. (22), we get the closed-form for GP log-likelihood as:

$$\mathcal{L} = \log \mathcal{P}(\mathbf{y}|\mathbf{X},\theta) = -\frac{1}{2}(\mathbf{y}^\intercal [\mathbf{K}+\sigma_\eta^2 \mathbf{I}]^{-1}\mathbf{y}) - \frac{1}{2}\log|\mathbf{K}+\sigma_\eta^2\mathbf{I}| - \frac{n}{2}\log(2\pi) \tag{23}$$

where n corresponds to the total number of training instances. For a rank GP the closed-form of the GP log-likelihood can be obtained as:

$$\bar{\mathcal{L}} = -\frac{1}{2}\boldsymbol{\omega}_{\mathbf{MAP}}^{\intercal}[\mathbf{K}+\sigma_{\eta}^{2}\mathbf{I}]^{-1}\boldsymbol{\omega}_{\mathbf{MAP}} - \frac{1}{2}\log \mid \mathbf{K}+\sigma_{\eta}^{2}\mathbf{I} \mid -\frac{n}{2}\log(2\pi) \tag{24}$$

The only difference in the formulation of log-likelihood of a rank GP and a traditional GP is that the absolute measurements ($\mathbf{y}$) of the objective function is replaced with the latent function values obtained via Maximum A Posteriori (MAP) estimates. The GP log-likelihood mentioned in Eq. (23) and Eq. (24) is now maximized to find the optimal hyperparameters $\theta^{\star}$. For instance, $\theta^{\star}$ for a traditional GP is obtained as:

$$\theta^{\star} = \underset{\theta}{\operatorname{argmax}}\, \mathcal{L}$$

Further, if we use all training instances for the computation of the log marginal likelihood, there are chances that only Control arm may get selected in majority of the rounds. Therefore to avoid this, instead of using all the training instances for computing the marginal likelihood, we use only the subset of the original training data for finding the optimal hyperparameter set and then we use the held-out instances from the original training set to compute the (predictive) likelihood.

In this paper, we implement Automatic Relevance Determination (ARD) kernel [12] based on Squared Exponential kernel mentioned in Eq. (13) at all levels of our proposed BOAP framework. ARD kernel is popular in the machine learning community due to its ability to suppress the irrelevant features. In ARD, each input dimension is assigned a different lengthscale parameter to keep track of the relevance of that dimension. Therefore, a GP fitted in a $d-$dimensional input space with ARD kernel has d lengthscale parameters *i.e.* $\boldsymbol{l} = l_{1:d}$ and optional variance hyperparameters (noise variance σ_{η}^{2} and signal variance σ_{f}^{2}).

Specifically, for the human-inspired arm (Arm-$\mathfrak{h}$) we use a spatially-varying ARD kernel where we set the lengthscales of the augmented input dimensions in proportion to the rank GP uncertainties via a parametric function of the input $\mathbf{x}$. The lengthscale function for each of the augmented input dimension (corresponding to property ω_i) is set to be $l_{\omega_i}(\mathbf{x}) = \alpha_i \sigma_{\omega_i}(\mathbf{x})$, where α_i is the scale parameter and $\sigma_{\omega_i}(\mathbf{x})$ is the normalized standard deviation predicted using the rank GP ($\mathcal{GP}_{\omega_i}$). Therefore, the hyperparameter set (θ_h) of Arm-$\mathfrak{h}$ consists of the lengthscale parameters ($l_{1:d}$) for the original (d) un-augmented dimensions and the scale factor ($\alpha_{1:m}$) from the (m) augmented input dimensions *i.e.* $\theta_h = \{l_{1:d}, \alpha_{1:m}\}$.

The overall hyperparameter set Θ of our proposed framework consists of hyperparameters from each of the m rank GPs ($\theta_{\omega_{1:m}}$) and two main GPs corresponding to the 2-arms (θ_h and θ_f) *i.e.* $\Theta = \{\theta_h, \theta_f\} = \{\{l_{1:d}, \alpha_{1:m}\}, \{l_{1:d}\}\}$. At each iteration t, we find the optimal set of hyperparameters $\Theta_t^{\star} = \{\theta_h^{\star}, \theta_f^{\star}\}$ by maximizing the GP (predictive) log-likelihoods mentioned in Eq. (23) and Eq. (24).

In all our experiments, we set the signal variance parameter $\sigma_f^2 = 1$ as we standardize the outputs of $\mathcal{GP}_h$ and $\mathcal{GP}_f$. We set the noise variance as $\eta \sim \mathcal{GP}(0, \sigma_{\eta}^2 = 0.1)$ and $\tilde{\sigma}_{\eta}^2 = 0.1$. As we normalize the input space of the GP distributions constructed in our BOAP framework, we tune each lengthscale

hyperparameter $l \in \Theta$ in the interval $[0, 1]$. Further, we normalize the outputs of each of the auxiliary rank GPs ($\mathcal{GP}_{\omega_i}$) to avoid different scaling levels in their output, that can lead to undesired structures in the augmented input space.

We run all our experiments on an Intel Xeon CPU@ 3.60GHz workstation with 16 GB of RAM capacity. We repeat our experiments with 10 different random initialization. For a $d-$dimensional problem, we consider $t' = d+3$ initial observations. The evaluation budget is set as $T = 10 \times d + 5$. For the real-world battery design experiments, due to their expensive nature, we have restricted the evaluation budget to 50 iterations even though $d \gg 5$. In all our experiments, we start with $p = \binom{t'}{2}$ preferences in P, that gets updated in every iteration of the optimization process.

11.2 Details of Real-world Experiments

As mentioned in the main paper, we evaluate the performance of BOAP framework in two real-world optimization paradigms in the Lithium-ion battery manufacturing. The advent of cheap Li-ion battery technology has significantly transformed range of industries including healthcare [14], telecommunication [4], automobiles [8], and many more due to its ability to efficiently store the electrochemical energy. However, the process for manufacturing of Li-ion batteries is very complex and expensive in nature. Thus, there is a wide scope for optimizing the battery manufacturing process to reduce its CapEx and OpEx. We now provide a brief discussion on battery manufacturing experiments considered in our main paper. For real-world experiments, we use the same set of algorithmic parameters that we used in the synthetic experiments.

Optimization of Electrode Calendering Process [7] discussed the effect of calendering process on electrode properties that significantly contribute to the underlying electrochemical performance of a battery. Authors have implemented a data-driven stochastic electrode structure generator, based on which they construct electrodes and analyze in terms of *Tortuosity* (both in solid phase τ_{sol} and liquid phase τ_{liq}), percentage of *Current Collector* (CC) surface covered by the active material and percentage of *Active Surface* (AS) covered by the electrolyte. The manufacturing process parameters considered are calendering pressure, Carbon-Binder Domain (CBD), initial electrode porosity and electrode composition. A pictorial representation of the inter dependencies between the input process variables and output electrode properties is provided in Figure 8 of [7].

[7] published a dataset reporting the input manufacturing process parameters and the output characteristics of 8800 electrode structures. Each of the manufacturing setting has been evaluated for 10 times, therefore we have averaged the results to obtain a refined dataset consisting of $n = 880$ instances with $d = 8$ process variables.

We optimize the calendering process by maximizing the *Active Surface* of an electrode (overall objective) by modeling two abstract properties: **(i)** Property 1

(ω_1): *Tortuosity in liquid phase* τ_{liq}, and **(ii)** Property 2 (ω_2): *Output Porosity* (OP). As discussed in the main paper, the abstract properties $\{\omega_{\tau_{\text{liq}}}, \omega_{\text{OP}}\}$ can only be qualitatively measured, however to simulate the expert pairwise preferential inputs $\{P^{\omega_{\tau_{\text{liq}}}}, P^{\omega_{\text{OP}}}\}$, we use the empirical measurements for τ_{liq} and OP in the published dataset.

$$P^{\omega_{\tau_{\text{liq}}}} = \{(\mathbf{x} \succ \mathbf{x}')_i \text{ if } \tau_{\text{liq}}(\mathbf{x}) > \tau_{\text{liq}}(\mathbf{x}') \mid \mathbf{x}, \mathbf{x}' \in \mathbf{x}_{1:n} \ \ \forall i \in \mathbb{N}_p\}$$

$$P^{\omega_{\text{OP}}} = \{(\mathbf{x} \succ \mathbf{x}')_i \text{ if } \text{OP}(\mathbf{x}) > \text{OP}(\mathbf{x}') \mid \mathbf{x}, \mathbf{x}' \in \mathbf{x}_{1:n} \ \ \forall i \in \mathbb{N}_p\}$$

We obtain the values $\tau_{\text{liq}}(\mathbf{x})$ and $\text{OP}(\mathbf{x})$ by referring to the dataset published. Based on these preference lists $\{P^{\omega_{\tau_{\text{liq}}}}, P^{\omega_{\text{OP}}}\}$, we fit two auxiliary rank GP distributions $\{\mathcal{GP}_{\omega_{\tau_{\text{liq}}}}, \mathcal{GP}_{\omega_{\text{OP}}}\}$. Then, we use these auxiliary rank GPs to augment the input space of the main GP surrogate $\mathcal{GP}_h$ modeling the overall objective *i.e.* active surface of the electrode.

Optimization of Electrode Manufacturing Process In a similar case study [6], the authors have analyzed the manufacturing of Lithium-ion graphite based electrodes. The main aim of [6] is to optimize the formulation and manufacturing process of Lithium-ion electrodes using machine learning. Authors have established a relationship between the process parameters at different stages of manufacturing such as mixing, coating, drying, and calendering. The published dataset records all the process parameters in manufacturing 256 coin cells, as well as the associated results showing the charge capacity of each coin cell measured after certain charge-discharge cycles. The refined dataset consists of 12 process variables. In our experiments, out of those 12 process variables we assume 2 process variables: **(i)** *Anode Thickness* (AT), and **(ii)** *Active Mass* (AM) as abstract properties that can only be qualitatively measured. The overall objective here is to maximize the battery endurance $E = \frac{D_{50}}{D_5}$, where D_{50} and D_5 are the discharge capacities of the cell at 50^{th} cycle and 5^{th} cycle, respectively.

We simulate the expert pairwise preferential inputs $\{P^{\omega_{\text{AT}}}, P^{\omega_{\text{AM}}}\}$ by comparing the empirical values recorded for these variables in the given dataset.

$$P^{\omega_{\text{AT}}} = \{(\mathbf{x} \succ \mathbf{x}')_i \text{ if } \text{AT}(\mathbf{x}) > \text{AT}(\mathbf{x}') \mid \mathbf{x}, \mathbf{x}' \in \mathbf{x}_{1:n} \ \ \forall i \in \mathbb{N}_p\}$$

$$P^{\omega_{\text{AM}}} = \{(\mathbf{x} \succ \mathbf{x}')_i \text{ if } \text{AM}(\mathbf{x}) > \text{AM}(\mathbf{x}') \mid \mathbf{x}, \mathbf{x}' \in \mathbf{x}_{1:n} \ \ \forall i \in \mathbb{N}_p\}$$

We model abstract properties $\{\omega_{\text{AT}}, \omega_{\text{AT}}\}$ by fitting rank GPs $\{\mathcal{GP}_{\omega_{\text{AM}}}, \mathcal{GP}_{\omega_{\text{AT}}}\}$ using preference lists $\{P^{\omega_{\text{AM}}}, P^{\omega_{\text{AM}}}\}$. We use rank GPs to estimate the abstract properties and then use those estimations to augment the input space of the main GP modeling the overall objective *i.e.* maximizing battery endurance E.

12 Additional Experimental Results

12.1 Frequency of Selecting Augmented/Control Arm

During the initial phase of optimization the un-augmented control arm is pulled as the augmented human arm may have insufficient data to reduce uncertainty in posteriors, making the augmented features only slightly better than noise. However, the augmented features become more accurate through the optimization process as we accumulate more data, so we observe the augmented human arm to be more accurate at the end of the optimization. To empirically evaluate we have conducted an additional experiment to record the percentage of times each arm is pulled. We have repeated the experiments for 5 times with random initialization. The results obtained are formulated in Table 3.

Table 3: Frequency (in %) of Augmented/Control Arm being pulled during the optimization.

Functions	Augmented Arm	Control Arm
Benchmark-1D	76.23±2.5	24.22±1.2
Rosenbrock-3D	71.87±4.6	25.67±2.75
Griewank-5D	63.95±3.1	35.73±2.6

Based on the empirical results we could verify that the augmented human model was chosen more than the un-augmented arm. In an ideal scenario with accurate preferential data (with accurate high-level features) we have recorded the number of times each was pulled and the results are as follows.

12.2 High-level Features as Input Dimensions

In this ablation study, we demonstrate the implications of using/knowing high-level features directly as input dimension. In practice, an expert would reason about a system only qualitatively and high-level features are not known directly. We conduct an additional experiment to directly use high-level features as additional dimensions (BOAP-HL). We found that BOAP-HL outperforms BOAP by a small margin and we believe this phenomenon is due to the fact that explicit measurements are always better than approximate qualitative measurements. We believe that the superior results of BOAP-HL is due to the additional high-level features augmented in the given input space. We have repeated the experiments for 5 times with random initialization. The obtained results are formulated in Table 4.

13 Limitations

Firstly, BOAP in its vanilla version may not be directly scalable because of the scalability issues of the underlying range of GPs used in the framework. One

Table 4: Simple regret computed after $10 \times d$ iterations for the experiments with high-level features directly used as input dimensions.

Functions	BOAP-HL	BOAP
Benchmark-1D	0.01±0.002	0.05±0.01
Rosenbrock-3D	3.16±1.19	5.32±0.12
Griewank-5D	2.15±0.8	2.74±1.1

of the well-known weaknesses of GP is that it poorly scales and suffers from a cubic time complexity $O(n^3)$ due to the inversion of the gram matrix $\mathbf{K}$. This limits the scalability of GP and thus our BOAP framework to use with large-scale datasets. In the future line of work, we aspire to overcome these limitations by employing suitable GP techniques that can be easily scaled. Secondly, the level of expertise required to conduct the experiments is very high and access to such real experts is extremely difficult. Due to the cost and access limitations of experimental facilities, we rely on numerical simulations and simulated data and thus assume null human query cost.

References

1. Aronszajn, N.: Theory of reproducing kernels. Transactions of the American mathematical society **68**(3), 337–404 (1950)
2. Bijl, H., Schön, T.B., van Wingerden, J.W., Verhaegen, M.: A sequential Monte Carlo approach to Thompson sampling for Bayesian optimization. arXiv preprint arXiv:1604.00169 (2016)
3. Brochu, E., Cora, V.M., De Freitas, N.: A tutorial on Bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning. arXiv preprint arXiv:1012.2599 (2010)
4. Brunarie, J., Billard, A.M., Lansburg, S., Belle, M.: Lithium-ion (Li-ion) battery technology evolves to serve an extended range of telecom applications. In: 2011 IEEE 33rd International Telecommunications Energy Conference (INTELEC). pp. 1–9. IEEE (2011)
5. Chowdhury, S.R., Gopalan, A.: On kernelized multi-armed bandits. In: Precup, D., Teh, Y.W. (eds.) Proceedings of the 34th International Conference on Machine Learning. Proceedings of Machine Learning Research, vol. 70, pp. 844–853. PMLR, International Convention Centre, Sydney, Australia (Aug 2017)
6. Drakopoulos, S.X., Gholamipour-Shirazi, A., MacDonald, P., Parini, R.C., Reynolds, C.D., Burnett, D.L., Pye, B., O'Regan, K.B., Wang, G., Whitehead, T.M., et al.: Formulation and manufacturing optimization of Lithium-ion graphite-based electrodes via machine learning. Cell Reports Physical Science **2**(12), 100683 (2021)
7. Duquesnoy, M., Lombardo, T., Chouchane, M., Primo, E.N., Franco, A.A.: Data-driven assessment of electrode calendering process by combining experimental results, in silico mesostructures generation and machine learning. Journal of Power Sources **480**, 229103 (2020)
8. Harper, G., Sommerville, R., Kendrick, E., Driscoll, L., Slater, P., Stolkin, R., Walton, A., Christensen, P., Heidrich, O., Lambert, S., et al.: Recycling Lithium-ion batteries from electric vehicles. nature **575**(7781), 75–86 (2019)

9. Kaufmann, E., Korda, N., Munos, R.: Thompson sampling: An asymptotically optimal finite-time analysis. In: Algorithmic Learning Theory: 23rd International Conference, ALT 2012, Lyon, France, October 29-31, 2012. Proceedings 23. pp. 199–213. Springer (2012)
10. Kushner, H.J.: A new method of locating the maximum point of an arbitrary multipeak curve in the presence of noise. Journal of Basic Engineering **86**(1), 97–106 (03 1964). https://doi.org/10.1115/1.3653121
11. Li, G., Kamath, P., Foster, D.J., Srebro, N.: Understanding the eluder dimension. In: Advances in Neural Information Processing Systems (2021)
12. Neal, R.M.: Bayesian learning for neural networks, vol. 118. Springer Science & Business Media (2012)
13. Russo, D., Van Roy, B.: Learning to optimize via posterior sampling. Mathematics of Operations Research **39**(4), 1221–1243 (2014)
14. Schmidt, C.L., Skarstad, P.M.: The future of Lithium and Lithium-ion batteries in implantable medical devices. Journal of power sources **97**, 742–746 (2001)
15. Shahriari, B., Wang, Z., Hoffman, M.W., Bouchard-Côté, A., de Freitas, N.: An entropy search portfolio for Bayesian optimization. arXiv preprint arXiv:1406.4625 (2014)
16. Srinivas, N., Krause, A., Kakade, S.M., Seeger, M.W.: Information-theoretic regret bounds for Gaussian process optimization in the bandit setting. IEEE Transactions on Information Theory **58**(5), 3250–3265 (2012)
17. Thompson, W.R.: On the likelihood that one unknown probability exceeds another in view of the evidence of two samples. Biometrika **25**(3-4), 285–294 (1933)