A. Proof of Lemma 2

Lemma 2. Similar to the definition of $f(\mathbf{x}, \mathbf{Z})$, let $l_{t-1}(\mathbf{x}, \mathbf{Z})$ and $u_{t-1}(\mathbf{x}, \mathbf{Z})$ denote the random function over $\mathbf{x}$ where the randomness comes from the random variable $\mathbf{Z}$; l_{t-1} and u_{t-1} are defined in (3). Then, $\forall \mathbf{x} \in \mathcal{D}_{\mathbf{x}}, t \geq 1$, $\alpha \in (0,1)$,

$$\begin{aligned} V_\alpha(f(\mathbf{x},\mathbf{Z})) \in\ & I_{t-1}[V_\alpha(f(\mathbf{x},\mathbf{Z}))] \\ & \triangleq [V_\alpha(l_{t-1}(\mathbf{x},\mathbf{Z})), V_\alpha(u_{t-1}(\mathbf{x},\mathbf{Z}))] \end{aligned}$$

holds with probability $\geq 1-\delta$ for δ in Lemma 1, where $V_\alpha(l_{t-1}(\mathbf{x}, \mathbf{Z}))$ and $V_\alpha(u_{t-1}(\mathbf{x}, \mathbf{Z}))$ are defined as (1).

Proof. Conditioned on the event $f(\mathbf{x}, \mathbf{z}) \in I_{t-1}[f(\mathbf{x}, \mathbf{z})] \triangleq [l_{t-1}(\mathbf{x}, \mathbf{z}), u_{t-1}(\mathbf{x}, \mathbf{z})]$ for all $\mathbf{x} \in \mathcal{D}_{\mathbf{x}}$, $\mathbf{z} \in \mathcal{D}_{\mathbf{z}}$, $t \geq 1$ which occurs with probability $\geq 1-\delta$ for δ in Lemma 1, we will prove that $V_\alpha(l_{t-1}(\mathbf{x}, \mathbf{Z})) \leq V_\alpha(f(\mathbf{x}, \mathbf{Z}))$. The proof of $V_\alpha(f(\mathbf{x}, \mathbf{Z})) \leq V_\alpha(u_{t-1}(\mathbf{x}, \mathbf{Z}))$ can be done in a similar manner.

From $f(\mathbf{x}, \mathbf{z}) \in I_{t-1}[f(\mathbf{x}, \mathbf{z})] \triangleq [l_{t-1}(\mathbf{x}, \mathbf{z}), u_{t-1}(\mathbf{x}, \mathbf{z})]$ for all $\mathbf{x} \in \mathcal{D}_{\mathbf{x}}$, $\mathbf{z} \in \mathcal{D}_{\mathbf{z}}$, $t \geq 1$ we have $\forall \mathbf{x} \in \mathcal{D}_{\mathbf{x}}, \mathbf{z} \in \mathcal{D}_{\mathbf{z}}, t \geq 1$,

$$f(\mathbf{x},\mathbf{z}) \geq l_{t-1}(\mathbf{x},\mathbf{z}) \ .$$

Therefore, for all $\omega \in \mathbb{R}$, $\mathbf{x} \in \mathcal{D}_{\mathbf{x}}$, $\mathbf{z} \in \mathcal{D}_{\mathbf{z}}$, $t \geq 1$,

$$\begin{aligned} f(\mathbf{x},\mathbf{z}) \leq \omega &\Rightarrow l_{t-1}(\mathbf{x},\mathbf{z}) \leq \omega \\ P(f(\mathbf{x},\mathbf{Z}) \leq \omega) &\leq P(l_{t-1}(\mathbf{x},\mathbf{Z}) \leq \omega) \ . \end{aligned}$$

So, for all $\omega \in \mathbb{R}$, $\alpha \in (0,1)$, $\mathbf{x} \in \mathcal{D}_{\mathbf{x}}$, $t \geq 1$,

$$P(f(\mathbf{x},\mathbf{Z}) \leq \omega) \geq \alpha \Rightarrow P(l_{t-1}(\mathbf{x},\mathbf{Z}) \leq \omega) \geq \alpha \ .$$

Hence, the set $\{\omega : P(f(\mathbf{x}, \mathbf{Z}) \leq \omega) \geq \alpha\}$ is a subset of $\{\omega : P(l_{t-1}(\mathbf{x}, \mathbf{Z}) \leq \omega) \geq \alpha\}$ for all $\alpha \in (0, 1)$, $\mathbf{x} \in \mathcal{D}_{\mathbf{x}}$, $t \geq 1$, which implies that $\inf\{\omega : P(f(\mathbf{x}, \mathbf{Z}) \leq \omega) \geq \alpha\} \geq \inf\{\omega : P(l_{t-1}(\mathbf{x}, \mathbf{Z}) \leq \omega) \geq \alpha\}$, i.e., $V_\alpha(l_{t-1}(\mathbf{x}, \mathbf{Z})) \leq V_\alpha(f(\mathbf{x}, \mathbf{Z}))$ for all $\alpha \in (0, 1)$, $\mathbf{x} \in \mathcal{D}_{\mathbf{x}}$, $t \geq 1$. □

B. Proof of (4)

We prove that

$$r(\mathbf{x}_t) \leq V_\alpha(u_{t-1}(\mathbf{x}_t,\mathbf{Z})) - V_\alpha(l_{t-1}(\mathbf{x}_t,\mathbf{Z}))\ \forall t \geq 1$$

which holds with probability $\geq 1-\delta$ for δ in Lemma 1.

Proof. Conditioned on the event $V_\alpha(f(\mathbf{x}, \mathbf{Z})) \in I_{t-1}[V_\alpha(f(\mathbf{x}, \mathbf{Z}))] \triangleq [V_\alpha(l_{t-1}(\mathbf{x}, \mathbf{Z})), V_\alpha(u_{t-1}(\mathbf{x}, \mathbf{Z}))]$ for all $\alpha \in (0, 1)$, $\mathbf{x} \in \mathcal{D}_{\mathbf{x}}$, $t \geq 1$, which occurs with probability $\geq 1-\delta$ in Lemma 2,

$$\begin{aligned} V_\alpha(f(\mathbf{x}_*,\mathbf{Z})) &\leq V_\alpha(u_{t-1}(\mathbf{x}_*,\mathbf{Z})) \\ V_\alpha(f(\mathbf{x}_t,\mathbf{Z})) &\geq V_\alpha(l_{t-1}(\mathbf{x}_t,\mathbf{Z})) \ . \end{aligned}$$

Hence,

$$\begin{aligned} r(\mathbf{x}_t) &\triangleq V_\alpha(f(\mathbf{x}_*,\mathbf{Z})) - V_\alpha(f(\mathbf{x}_t,\mathbf{Z})) \\ &\leq V_\alpha(u_{t-1}(\mathbf{x}_*,\mathbf{Z})) - V_\alpha(l_{t-1}(\mathbf{x}_t,\mathbf{Z})) \ . \end{aligned} \quad (9)$$

Since $\mathbf{x}_t$ is selected as $\operatorname{argmax}_{\mathbf{x} \in \mathcal{D}_{\mathbf{x}}} V_\alpha(u_{t-1}(\mathbf{x}, \mathbf{Z}))$,

$$V_\alpha(u_{t-1}(\mathbf{x}_*,\mathbf{Z})) \leq V_\alpha(u_{t-1}(\mathbf{x}_t,\mathbf{Z})) \ ,$$

equivalently, $V_\alpha(u_{t-1}(\mathbf{x}_*, \mathbf{Z})) - V_\alpha(l_{t-1}(\mathbf{x}_t, \mathbf{Z})) \leq V_\alpha(u_{t-1}(\mathbf{x}_t, \mathbf{Z})) - V_\alpha(l_{t-1}(\mathbf{x}_t, \mathbf{Z}))$. Hence, from (9), $r(\mathbf{x}_t) \leq V_\alpha(u_{t-1}(\mathbf{x}_t, \mathbf{Z})) - V_\alpha(l_{t-1}(\mathbf{x}_t, \mathbf{Z}))$ for all $\alpha \in (0, 1)$ and $t \geq 1$. □

C. Proof of Theorem 1

Theorem 1. $\forall \alpha \in (0, 1)$, $\forall \mathbf{x} \in \mathcal{D}_{\mathbf{x}}$, $\forall t \geq 1$, there exists a lacing value in $\mathcal{D}_{\mathbf{z}}$ with respect to $\mathbf{x}$ and t.

Proof. Recall that

$$\mathcal{Z}_l^{\leq} \triangleq \{\mathbf{z} \in \mathcal{D}_{\mathbf{z}} : l_{t-1}(\mathbf{x},\mathbf{z}) \leq V_\alpha(l_{t-1}(\mathbf{x},\mathbf{Z}))\} \ .$$

From to the definition of $\mathcal{Z}_l^{\leq}$ and $V_\alpha(l_{t-1}(\mathbf{x}, \mathbf{Z}))$, we have

$$P(\mathbf{Z} \in \mathcal{Z}_l^{\leq}) \geq \alpha \ . \quad (10)$$

Since $\alpha \in (0, 1)$, $\mathcal{Z}_l^{\leq} \neq \emptyset$. We prove the existence of LV by contradiction: (a) assuming that $\exists \mathbf{x} \in \mathcal{D}_{\mathbf{x}}, \exists t \geq 1, \forall \mathbf{z} \in \mathcal{Z}_l^{\leq}, u_{t-1}(\mathbf{x}, \mathbf{z}) < V_\alpha(u_{t-1}(\mathbf{x}, \mathbf{Z}))$ and then, (b) proving that $V_\alpha(u_{t-1}(\mathbf{x}, \mathbf{Z}))$ is not a lower bound of $\{\omega : P(u_{t-1}(\mathbf{x}, \mathbf{Z}) \leq \omega) \geq \alpha\}$ which is a contradiction.

Since the GP posterior mean μ_{t-1} and posterior standard deviation σ_{t-1} are continuous functions in $\mathcal{D}_{\mathbf{x}} \times \mathcal{D}_{\mathbf{z}}$, l_{t-1} and u_{t-1} are continuous functions in the closed $\mathcal{D}_{\mathbf{z}} \subset \mathbb{R}^{d_z}$ ($\mathbf{x}$ and t are given and remain fixed in this proof). We will prove that $\mathcal{Z}_l^{\leq}$ is closed in $\mathbb{R}^{d_z}$ by contradiction.

If $\mathcal{Z}_l^{\leq}$ is not closed in $\mathbb{R}^{d_z}$, there exists a limit point $\mathbf{z}_p$ of $\mathcal{Z}_l^{\leq}$ such that $\mathbf{z}_p \notin \mathcal{Z}_l^{\leq}$. Since $\mathcal{Z}_l^{\leq} \subset \mathcal{D}_{\mathbf{z}}$ and $\mathcal{D}_{\mathbf{z}}$ is closed in $\mathbb{R}^{d_z}$, $\mathbf{z}_p \in \mathcal{D}_{\mathbf{z}}$. Thus, for $\mathbf{z}_p \notin \mathcal{Z}_l^{\leq}$, $l_{t-1}(\mathbf{x}, \mathbf{z}_p) > V_\alpha(l_{t-1}(\mathbf{x}, \mathbf{Z}))$ (from the definition of $\mathcal{Z}_l^{\leq}$). Then, there exists $\epsilon_0 > 0$ such that $l_{t-1}(\mathbf{x}, \mathbf{z}_p) > V_\alpha(l_{t-1}(\mathbf{x}, \mathbf{Z})) + \epsilon_0$. The pre-image of the open interval $I_0 = (l_{t-1}(\mathbf{x}, \mathbf{z}_p) - \epsilon_0/2, l_{t-1}(\mathbf{x}, \mathbf{z}_p) + \epsilon_0/2)$ under l_{t-1} is also an open set $\mathcal{V}$ containing $\mathbf{z}_p$ (because l_{t-1} is a continuous function). Since $\mathbf{z}_p$ is a limit point of $\mathcal{Z}_l^{\leq}$, there exists an $\mathbf{z}' \in \mathcal{Z}_l^{\leq} \cap \mathcal{V}$. Then, $l_{t-1}(\mathbf{x}, \mathbf{z}') \in I_0$, so $l_{t-1}(\mathbf{x}, \mathbf{z}') \geq l_{t-1}(\mathbf{x}, \mathbf{z}_p) - \epsilon_0/2 \geq V_\alpha(l_{t-1}(\mathbf{x}, \mathbf{Z})) + \epsilon_0 - \epsilon_0/2 = V_\alpha(l_{t-1}(\mathbf{x}, \mathbf{Z})) + \epsilon_0/2$. It contradicts $\mathbf{z}' \in \mathcal{Z}_l^{\leq}$.

Therefore, $\mathcal{Z}_l^{\leq}$ is a closed set in $\mathbb{R}^{d_z}$. Besides, since $\{u_{t-1}(\mathbf{x}, \mathbf{z}) : \mathbf{z} \in \mathcal{Z}_l^{\leq}\}$ is upper bounded by $V_\alpha(u_{t-1}(\mathbf{x}, \mathbf{Z}))$ (due to our assumption), so

$\sup\{u_{t-1}(\mathbf{x},\mathbf{z}) : \mathbf{z} \in \mathcal{Z}_l^{\leq}\}$ exists. Let $\mathbf{z}_l^+$ be such that $u_{t-1}(\mathbf{x},\mathbf{z}_l^+) \triangleq \sup\{u_{t-1}(\mathbf{x},\mathbf{z}) : \mathbf{z} \in \mathcal{Z}_l^{\leq}\}$. Then, $\mathbf{z}_l^+ \in \mathcal{Z}_l^{\leq}$ because $\mathcal{Z}_l^{\leq}$ is closed.

Moreover, from our assumption that $\forall \mathbf{z} \in \mathcal{Z}_l^{\leq}, u_{t-1}(\mathbf{x},\mathbf{z}) < V_\alpha(u_{t-1}(\mathbf{x},\mathbf{Z}))$, we have $u_{t-1}(\mathbf{x},\mathbf{z}_l^+) < V_\alpha(u_{t-1}(\mathbf{x},\mathbf{Z}))$. Furthermore,

$$P(u_{t-1}(\mathbf{x},\mathbf{Z}) \leq u_{t-1}(\mathbf{x},\mathbf{z}_l^+)) \geq P(\mathbf{Z} \in \mathcal{Z}_l^{\leq}) \geq \alpha\ .$$

where the first inequality is because $u_{t-1}(\mathbf{x},\mathbf{z}_l^+) = \sup\{u_{t-1}(\mathbf{x},\mathbf{z}) : \mathbf{z} \in \mathcal{Z}_l^{\leq}\}$ and the last inequality is from (10). Hence, $V_\alpha(u_{t-1}(\mathbf{x},\mathbf{Z})$ is not a lower bound of $\{\omega : P(u_{t-1}(\mathbf{x},\mathbf{Z}) \leq \omega) \geq \alpha\}$. □

D. Proof of Lemma 3

Lemma 3. By selecting $\mathbf{x}_t$ as a maximizer of $V_\alpha(u_{t-1}(\mathbf{x},\mathbf{Z}))$ and selecting $\mathbf{z}_t$ as an LV w.r.t $\mathbf{x}_t$, the instantaneous regret is upper-bounded by:

$$r(\mathbf{x}_t) \leq 2\beta_t^{1/2}\sigma_{t-1}(\mathbf{x}_t,\mathbf{z}_t)\ \forall t \geq 1$$

with probability $\geq 1-\delta$ for δ in Lemma 1.

Proof. Conditioned on the event $f(\mathbf{x},\mathbf{z}) \in I_{t-1}[f(\mathbf{x},\mathbf{z})] \triangleq [l_{t-1}(\mathbf{x},\mathbf{z}), u_{t-1}(\mathbf{x},\mathbf{z})]$ for all $\mathbf{x} \in \mathcal{D}_{\mathbf{x}}$, $\mathbf{z} \in \mathcal{D}_{\mathbf{z}}$, $t \geq 1$ which occurs with probability $\geq 1-\delta$ in Lemma 1, it follows that $V_\alpha(f(\mathbf{x},\mathbf{Z})) \in I_{t-1}[V_\alpha(f(\mathbf{x},\mathbf{Z}))] \triangleq [V_\alpha(l_{t-1}(\mathbf{x},\mathbf{Z})), V_\alpha(u_{t-1}(\mathbf{x},\mathbf{Z}))]$ for all $\alpha \in (0,1)$, $\mathbf{x} \in \mathcal{D}_{\mathbf{x}}$, and $t \geq 1$ in Lemma 2.

From (4), by selecting $\mathbf{z}_t$ as an LV, for all $t \geq 1$,

$$\begin{aligned}
r(\mathbf{x}_t) &\leq V_\alpha(u_{t-1}(\mathbf{x}_t,\mathbf{Z})) - V_\alpha(l_{t-1}(\mathbf{x}_t,\mathbf{Z})) \\
&\leq u_{t-1}(\mathbf{x}_t,\mathbf{z}_t) - l_{t-1}(\mathbf{x}_t,\mathbf{z}_t) \text{ (since } \mathbf{z}_t \text{ is an LV)} \\
&\leq \mu_{t-1}(\mathbf{x}_t,\mathbf{z}_t) + \beta_t^{1/2}\sigma_{t-1}(\mathbf{x}_t,\mathbf{z}_t) \\
&\quad - \mu_{t-1}(\mathbf{x}_t,\mathbf{z}_t) + \beta_t^{1/2}\sigma_{t-1}(\mathbf{x}_t,\mathbf{z}_t) \\
&= 2\beta_t^{1/2}\sigma_{t-1}(\mathbf{x}_t,\mathbf{z}_t)\ .
\end{aligned}$$

□

E. Bound on the Average Cumulative Regret

Conditioned on the event $f(\mathbf{x},\mathbf{z}) \in I_{t-1}[f(\mathbf{x},\mathbf{z})] \triangleq [l_{t-1}(\mathbf{x},\mathbf{z}), u_{t-1}(\mathbf{x},\mathbf{z})]$ for all $\mathbf{x} \in \mathcal{D}_{\mathbf{x}}$, $\mathbf{z} \in \mathcal{D}_{\mathbf{z}}$, $t \geq 1$ which occurs with probability $\geq 1-\delta$ in Lemma 1, it follows that $r(\mathbf{x}_t) \leq 2\beta_t^{1/2}\sigma_{t-1}(\mathbf{x}_t,\mathbf{z}_t)\ \forall t \geq 1$ in Lemma 3. Therefore,

$$\begin{aligned}
R_T &\triangleq \sum_{t=1}^{T} r(\mathbf{x}_t) \leq \sum_{t=1}^{T} 2\beta_t^{1/2}\sigma_{t-1}(\mathbf{x}_t,\mathbf{z}_t) \\
&\leq 2\beta_T^{1/2}\sum_{t=1}^{T}\sigma_{t-1}(\mathbf{x}_t,\mathbf{z}_t)
\end{aligned}$$

since β_t is a non-decreasing sequence.

From Lemma 5.4 and the Cauchy-Schwarz inequality in (Srinivas et al., 2010), we have

$$2\sum_{t=1}^{T}\sigma_{t-1}(\mathbf{x}_t,\mathbf{z}_t) \leq \sqrt{C_1 T\gamma_T} \tag{11}$$

where $C_1 = 8/\log(1+\sigma_n^{-2})$. Hence,

$$R_T \leq \sqrt{C_1 T\beta_T\gamma_T}\ .$$

Equivalently, $R_T/T \leq \sqrt{C_1\beta_T\gamma_T/T}$. Since γ_T is shown to be bounded for several common kernels in (Srinivas et al., 2010), the above implies that $\lim_{T\to\infty} R_T/T = 0$, i.e., the algorithm is no-regret.

F. Bound on $r(\mathbf{x}_{t_*(T)})$

Conditioned on the event $f(\mathbf{x},\mathbf{z}) \in I_{t-1}[f(\mathbf{x},\mathbf{z})] \triangleq [l_{t-1}(\mathbf{x},\mathbf{z}), u_{t-1}(\mathbf{x},\mathbf{z})]$ for all $\mathbf{x} \in \mathcal{D}_{\mathbf{x}}$, $\mathbf{z} \in \mathcal{D}_{\mathbf{z}}$, $t \geq 1$, which occurs with probability $\geq 1-\delta$ in Lemma 1, it follows that $V_\alpha(f(\mathbf{x},\mathbf{Z})) \in I_{t-1}[V_\alpha(f(\mathbf{x},\mathbf{Z}))] \triangleq [V_\alpha(l_{t-1}(\mathbf{x},\mathbf{Z})), V_\alpha(u_{t-1}(\mathbf{x},\mathbf{Z}))]$ for all $\alpha \in (0,1), \mathbf{x} \in \mathcal{D}_{\mathbf{x}}, t \geq 1$ in Lemma 2. Furthermore, we select $\mathbf{z}_t$ as an LV, so $l_{t-1}(\mathbf{x}_t,\mathbf{z}_t) \leq V_\alpha(l_{t-1}(\mathbf{x}_t,\mathbf{Z})) \leq V_\alpha(u_{t-1}(\mathbf{x}_t,\mathbf{Z})) \leq u_{t-1}(\mathbf{x}_t,\mathbf{z}_t)$ according to the Definition 1.

At T-th iteration, by recommending $\mathbf{x}_{t_*(T)}$ as an estimate of $\mathbf{x}_*$ where $t_*(T) \triangleq \text{argmax}_{t\in\{1,\ldots,T\}} V_\alpha(l_{t-1}(\mathbf{x}_t,\mathbf{Z}))$, we have

$$\begin{aligned}
V_\alpha(l_{t_*(T)-1}(\mathbf{x}_{t_*(T)},\mathbf{Z})) &= \max_{t\in\{1,\ldots,T\}} V_\alpha(l_{t-1}(\mathbf{x}_t,\mathbf{Z})) \\
&\geq \frac{1}{T}\sum_{t=1}^{T} V_\alpha(l_{t-1}(\mathbf{x}_t,\mathbf{Z}))\ .
\end{aligned}$$

Therefore,

$$\begin{aligned}
r(\mathbf{x}_{t_*(T)}) &= V_\alpha(f(\mathbf{x}_*,\mathbf{Z})) - V_\alpha(f(\mathbf{x}_{t_*(T)},\mathbf{Z})) \\
&\leq V_\alpha(f(\mathbf{x}_*,\mathbf{Z})) - V_\alpha(l_{t_*(T)-1}(\mathbf{x}_{t_*(T)},\mathbf{Z})) \\
&\leq \frac{1}{T}\sum_{t=1}^{T}\left(V_\alpha(f(\mathbf{x}_*,\mathbf{Z})) - V_\alpha(l_{t-1}(\mathbf{x}_t,\mathbf{Z}))\right)\ .
\end{aligned}$$

Furthermore, $V_\alpha(f(\mathbf{x}_*,\mathbf{Z})) \leq V_\alpha(u_{t-1}(\mathbf{x}_*,\mathbf{Z}))$ from our condition, so

$$r(\mathbf{x}_{t_*(T)}) \le \frac{1}{T}\sum_{t=1}^{T}(V_\alpha(f(\mathbf{x}_*,\mathbf{Z})) - V_\alpha(l_{t-1}(\mathbf{x}_t,\mathbf{Z})))$$
$$\le \frac{1}{T}\sum_{t=1}^{T}(V_\alpha(u_{t-1}(\mathbf{x}_*,\mathbf{Z})) - V_\alpha(l_{t-1}(\mathbf{x}_t,\mathbf{Z})))$$
$$\le \frac{1}{T}\sum_{t=1}^{T}(V_\alpha(u_{t-1}(\mathbf{x}_t,\mathbf{Z})) - V_\alpha(l_{t-1}(\mathbf{x}_t,\mathbf{Z})))$$
$$\le \frac{1}{T}\sum_{t=1}^{T}(u_{t-1}(\mathbf{x}_t,\mathbf{z}_t) - l_{t-1}(\mathbf{x}_t,\mathbf{z}_t)) \text{ (since } \mathbf{z}_t \text{ is an LV)}$$
$$\le \frac{1}{T}\sum_{t=1}^{T}2\beta_t^{1/2}\sigma_{t-1}(\mathbf{x}_t,\mathbf{z}_t)$$
$$\le \sqrt{\frac{C_1\beta_T\gamma_T}{T}} \text{ (from Appendix E)} .$$

Since γ_T is shown to be bounded for several common kernels in (Srinivas et al., 2010), the above implies that $\lim_{T\to\infty} r(\mathbf{x}_{t_*(T)}) = 0$.

G. Proof of Theorem 2 and Its Corollaries

G.1. Proof of Theorem 2

Theorem 2. Let $\mathbf{W}$ be a random variable with the support $\mathcal{D}_w \subset \mathbb{R}^{d_w}$ and dimension d_w. Let h be a continuous function mapping from $\mathbf{w} \in \mathcal{D}_w$ to $\mathbb{R}$. Then, $h(\mathbf{W})$ denotes the random variable whose realization is the function h evaluation at a realization $\mathbf{w}$ of $\mathbf{W}$. Suppose $h(\mathbf{w})$ has a minimizer $\mathbf{w}_{\min} \in \mathcal{D}_w$, then $\lim_{\alpha\to 0^+} V_\alpha(h(\mathbf{W})) = h(\mathbf{w}_{\min})$.

Recall that the support $\mathcal{D}_w$ of $\mathbf{W}$ is defined as the smallest closed subset $\mathcal{D}_w$ of $\mathbb{R}^{d_z}$ such that $P(\mathbf{W} \in \mathcal{D}_w) = 1$, and $\mathbf{w}_{\min} \in \mathcal{D}_w$ minimizes $h(\mathbf{w})$.

Lemma 4. For all $\alpha \in (0,1)$, $V_\alpha(h(\mathbf{W}))$ is a nondecreasing function, i.e.,

$$\forall\, 1 > \alpha > \alpha' > 0,\ V_\alpha(h(\mathbf{W})) \ge V_{\alpha'}(h(\mathbf{W})) .$$

Proof. Since $\alpha > \alpha'$, for all $\omega \in \mathbb{R}$,

$$P(h(\mathbf{W}) \le \omega) \ge \alpha \Rightarrow P(h(\mathbf{W}) \le \omega) \ge \alpha' .$$

Therefore, $\{\omega : P(h(\mathbf{W}) \le \omega) \ge \alpha\}$ is a subset of $\{\omega : P(h(\mathbf{W}) \le \omega) \ge \alpha'\}$. Thus,

$$\inf\{\omega : P(h(\mathbf{W}) \le \omega) \ge \alpha\}$$
$$\ge \inf\{\omega : P(h(\mathbf{W}) \le \omega) \ge \alpha'\}$$

i.e., $V_\alpha(h(\mathbf{W})) \ge V_{\alpha'}(h(\mathbf{W}))$. □

Let

$$\omega_{0^+} \triangleq \lim_{\alpha\to 0^+} V_\alpha(h(\mathbf{W})) . \quad (12)$$

Then, from Lemma 4, the following lemma follows.

Lemma 5. For all $\alpha \in (0,1)$, and ω_{0^+} defined in (12)

$$\omega_{0^+} \le V_\alpha(h(\mathbf{W})) .$$

We use Lemma 5 to prove the following lemma.

Lemma 6. For all $\mathbf{w} \in \mathcal{D}_w$, and ω_{0^+} defined in (12)

$$\omega_{0^+} \le h(\mathbf{w})$$

which implies that

$$\omega_{0^+} \le h(\mathbf{w}_{\min}) .$$

Proof. By contradiction, we assume that there exits $\mathbf{w}' \in \mathcal{D}_w$ such that $\omega_{0^+} > h(\mathbf{w}')$. Then, there exists $\epsilon_1 > 0$ such that $\omega_{0^+} > h(\mathbf{w}') + \epsilon_1$. Consider the pre-image $\mathcal{V}$ of the open interval $I_h = (h(\mathbf{w}') - \epsilon_1/2, h(\mathbf{w}') + \epsilon_1/2$. Since h is a continuous function, $\mathcal{V}$ is an open set and it contains $\mathbf{w}'$ (as I_h contains $h(\mathbf{w}')$). Then, consider the set $\mathcal{V} \cap \mathcal{D}_w \supset \{\mathbf{w}'\} \ne \emptyset$, we prove $P(\mathbf{W} \in \mathcal{V} \cap \mathcal{D}_\mathbf{z}) > 0$ by contradiction as follows.

If $P(\mathbf{W} \in \mathcal{V} \cap \mathcal{D}_w) = 0$ then the closure of $\mathcal{D}_w \setminus \mathcal{V}$ is a closed set that is smaller than $\mathcal{D}_w$ (since $\mathcal{V}$ is an open set, $\mathcal{D}_w$ is a closed set, and $\mathcal{V} \cap \mathcal{D}_w$ is not empty) and satisfies $P(\mathbf{W} \in \mathcal{D}_w \setminus \mathcal{V}) = 1$, which contradicts the definition of $\mathcal{D}_w$. Thus, $P(\mathbf{W} \in \mathcal{V} \cap \mathcal{D}_w) > 0$.

Therefore, $P(h(\mathbf{W}) \in I_h) > 0$. So,

$$P(h(\mathbf{W}) \le \omega_{0^+})$$
$$\ge P(h(\mathbf{W}) \le h(\mathbf{w}') + \epsilon_1/2)$$
$$\ge P(h(\mathbf{W}) \in I_h)$$
$$> 0 .$$

Let us consider $\alpha_0 = P(h(\mathbf{W}) \le h(\mathbf{w}') + \epsilon_1/2) > 0$, the VAR at α_0 is

$$V_{\alpha_0}(h(\mathbf{W})) \triangleq \inf\{\omega : P(h(\mathbf{W}) \le \omega) \ge \alpha_0\}$$
$$\le h(\mathbf{w}') + \epsilon_1/2$$
$$< \omega_{0^+}$$

which is a contradiction to Lemma 5. □

Lemma 7. For ω_{0^+} defined in (12)

$$\omega_{0^+} \ge h(\mathbf{w}_{\min}) . \quad (13)$$

Proof. By contradiction, we assume that $\omega_{0^+} < h(\mathbf{w}_{\min})$. Then there exists $\epsilon_2 > 0$ that $\omega_{0^+} + \epsilon_2 < h(\mathbf{w}_{\min})$. Since $\omega_{0^+} \triangleq \lim_{\alpha\to 0^+} V_\alpha(h(\mathbf{W}))$ so there exits $\alpha_0 > 0$ such that $V_{\alpha_0}(h(\mathbf{W})) \in (\omega_{0^+}, \omega_{0^+} + \epsilon_2)$. However,

$$P(h(\mathbf{W}) \le V_{\alpha_0}(h(\mathbf{W})))$$
$$\le P(h(\mathbf{W}) \le \omega_{0^+} + \epsilon_2 < h(\mathbf{w}_{\min}))$$
$$= 0$$

which contradicts the fact that $\alpha_0 > 0$. Therefore, $\omega_{0^+} \geq h(\mathbf{w}_{\min})$. $\square$

From (12), Lemma 6 and Lemma 7,

$$\lim_{\alpha \to 0^+} V_\alpha(h(\mathbf{W})) = h(\mathbf{w}_{\min})$$

which directly leads to the result in Corollary 2.1 for a continuous function $f(\mathbf{x}, \mathbf{z})$ over $\mathbf{z} \in \mathcal{D}_\mathbf{z}$. While $\mathbf{Z}$ can follow any probability distribution defined on the support $\mathcal{D}_\mathbf{z}$, we can choose the distribution of $\mathbf{Z}$ as a uniform distribution over $\mathcal{D}_\mathbf{z}$.

G.2. Corollary 2.2

From Theorem 2, $\mathcal{D}_\mathbf{z}$ is a closed subset of $\mathbb{R}^{d_z}$, and $u_{t-1}(\mathbf{x}, \mathbf{z})$, $l_{t-1}(\mathbf{x}, \mathbf{z})$ are continuous functions over $\mathbf{z} \in \mathcal{D}_\mathbf{z}$, it follows that the selected $\mathbf{x}_t$ by both StableOpt (in (5)) and V-UCB are the same. Furthermore,

$$\begin{aligned}
\mathcal{Z}_l^{\leq} &\triangleq \{\mathbf{z} \in \mathcal{D}_\mathbf{z} : l_{t-1}(\mathbf{x}, \mathbf{z}) \leq V_\alpha(l_{t-1}(\mathbf{x}, \mathbf{Z}))\} \\
&= \{\mathbf{z} \in \mathcal{D}_\mathbf{z} : l_{t-1}(\mathbf{x}, \mathbf{z}) \leq \min_{\mathbf{z}' \in \mathcal{D}_\mathbf{z}} l_{t-1}(\mathbf{x}, \mathbf{z}')\} \\
&= \{\mathbf{z} \in \mathcal{D}_\mathbf{z} : l_{t-1}(\mathbf{x}, \mathbf{z}) = \min_{\mathbf{z}' \in \mathcal{D}_\mathbf{z}} l_{t-1}(\mathbf{x}, \mathbf{z}')\} \ , \\
\mathcal{Z}_u^{\geq} &\triangleq \{\mathbf{z} \in \mathcal{D}_\mathbf{z} : u_{t-1}(\mathbf{x}, \mathbf{z}) \geq V_\alpha(u_{t-1}(\mathbf{x}, \mathbf{Z}))\} \\
&= \{\mathbf{z} \in \mathcal{D}_\mathbf{z} : u_{t-1}(\mathbf{x}, \mathbf{z}) \geq \min_{\mathbf{z}' \in \mathcal{D}_\mathbf{z}} u_{t-1}(\mathbf{x}, \mathbf{z}')\} \\
&= \mathcal{D}_\mathbf{z} \ .
\end{aligned}$$

Therefore, the set of lacing values is $\mathcal{Z}_l^{\leq} \cap \mathcal{Z}_u^{\geq} = \mathcal{Z}_l^{\leq} = \{\mathbf{z} \in \mathcal{D}_\mathbf{z} : l_{t-1}(\mathbf{x}, \mathbf{z}) = \min_{\mathbf{z}' \in \mathcal{D}_\mathbf{z}} l_{t-1}(\mathbf{x}, \mathbf{z}')\}$ any of which is also the selected $\mathbf{z}_t$ in (5) by StableOpt. Thus, the selected $\mathbf{z}_t$ by both StableOpt and V-UCB are the same.

H. Local Neural Surrogate Optimization

The *local neural surrogate optimization* (LNSO) to maximize a VaR $V_\alpha(h(\mathbf{x}, \mathbf{Z}))$ is described in Algorithm 2. The algorithm can be summarized as follows:

- Whenever the current updated $\mathbf{x}^{(i)}$ is not in $\mathcal{B}(\mathbf{x}_c, r)$ (line 4), the center $\mathbf{x}_c$ of the ball $\mathcal{B}$ is updated to be $\mathbf{x}^{(i)}$ (line 6) and the surrogate function $g(\mathbf{x}, \boldsymbol{\theta})$ is re-trained (lines 7-12).
- The surrogate function $g(\mathbf{x}, \boldsymbol{\theta})$ is (re-)trained to estimate $V_\alpha(h(\mathbf{x}, \mathbf{Z}))$ well for all $\mathbf{x} \in \mathcal{B}(\mathbf{x}_c, r)$ (lines 7-12) with stochastic gradient descent by minimizing the following loss function given random mini-batches $\mathcal{Z}$ of $\mathbf{Z}$ (line 8) and $\mathcal{X}$ of $\mathbf{x} \in \mathcal{B}(\mathbf{x}_c, r)$ (line 9):

$$\mathcal{L}_g(\mathcal{X}, \mathcal{Z}) \triangleq \frac{1}{|\mathcal{X}||\mathcal{Z}|} \sum_{\mathbf{x} \in \mathcal{X}, \mathbf{z} \in \mathcal{Z}} [\rho_\alpha(h(\mathbf{x}, \mathbf{z}) - g(\mathbf{x}; \boldsymbol{\theta}))] \quad (14)$$

where ρ_α is the pinball function in Sec. 3.5.

- Instead of directly maximizing $V_\alpha(h(\mathbf{x}, \mathbf{Z}))$ whose gradient w.r.t $\mathbf{x}$ is unavailable, we find $\mathbf{x}$ that maximizes the surrogate function $g(\mathbf{x}, \boldsymbol{\theta}_s)$ (line 14) where $\boldsymbol{\theta}_s$ is the parameters trained in lines 7-12.

Algorithm 2 LNSO of $V_\alpha(h(\mathbf{x}, \mathbf{Z}))$

1: **Input:** target function h; domain $\mathcal{D}_\mathbf{x}$; initializer $\mathbf{x}^{(0)}$; α; a generator of $\mathbf{Z}$ samples `gen_Z`; radius r; no. of training iterations t_v, t_g; optimization stepsizes γ_x, γ_g
2: Randomly initialize $\boldsymbol{\theta}_s$
3: **for** $i = 1, 2, \ldots, t_v$ **do**
4:   **if** $i = 1$ or $\|\mathbf{x}^{(i)} - \mathbf{x}_c\| \geq \delta_x$ **then**
5:     Initialize $\boldsymbol{\theta}^{(0)} = \boldsymbol{\theta}_s$
6:     Update the center of $\mathcal{B}$: $\mathbf{x}_c = \mathbf{x}^{(i)}$
7:     **for** $j = 1, 2, \ldots, t_g$ **do**
8:       Draw n_z samples of $\mathbf{Z}$: $\mathcal{Z} = \texttt{gen\_Z}(n_z)$.
9:       Draw a set $\mathcal{X}$ of n_x uniformly distributed samples in $\mathcal{B}(\mathbf{x}_c, r)$.
10:       Update $\boldsymbol{\theta}^{(j)} = \boldsymbol{\theta}^{(j-1)} - \gamma_g \frac{\mathrm{d}\mathcal{L}_g(\mathcal{X}, \mathcal{Z})}{\mathrm{d}\boldsymbol{\theta}}\Big|_{\boldsymbol{\theta} = \boldsymbol{\theta}^{(j-1)}}$ where $\mathcal{L}_g(\mathcal{X}, \mathcal{Z})$ is defined in (14).
11:     **end for**
12:     $\boldsymbol{\theta}_s = \boldsymbol{\theta}_{t_g}$
13:   **end if**
14:   Update $\mathbf{x}^{(i)} = \mathbf{x}^{(i-1)} + \gamma_x \frac{\mathrm{d}g(\mathbf{x}; \boldsymbol{\theta}_s)}{\mathrm{d}\mathbf{x}}\Big|_{\mathbf{x} = \mathbf{x}^{(i-1)}}$.
15:   Project $\mathbf{x}^{(i)}$ into $\mathcal{D}_\mathbf{x}$.
16: **end for**
17: Return $\mathbf{x}^{(t_v)}$

I. Experimental Details

Regarding the construction of $\mathcal{D}_\mathbf{z}$ in optimizing the synthetic benchmark functions, the discrete $\mathcal{D}_\mathbf{z}$ is selected as equi-distant points (e.g., by dividing $[0, 1]^{d_z}$ into a grid). The probability mass of $\mathbf{Z}$ is defined as $P(\mathbf{Z} = \mathbf{z}) \propto \exp(-(\mathbf{z} - 0.5)^2 / 0.1^2)$ (the subtraction $\mathbf{z} - 0.5$ is elementwise). The continuous $\mathbf{Z}$ follows a 2-standard-deviation truncated independent Gaussian distribution with the mean of 0.5 and standard deviation 0.125. It is noted that when $\mathcal{D}_\mathbf{z}$ is discrete, there is a large region of $\mathbf{Z}$ with low probability $P(\mathbf{Z})$ in experiments with synthetic benchmark functions. This is to highlight the advantage of V-UCB Prob in exploiting $P(\mathbf{Z})$ compared with V-UCB Unif. In the robot pushing experiment, the region of $\mathbf{Z}$ with low probability is smaller than that in the experiments with synthetic benchmark functions (e.g., Hartmann-$(1, 2)$), which is illustrated in Fig. 6. Therefore, the gap in the performance between V-UCB Unif and V-UCB Prob is smaller in the robot pushing pushing experiment (Fig. 5b) than that in the experiment with Hartmann-3D-$(1, 2)$ (Fig. 3c).

When the closed-form expression of the objective func-

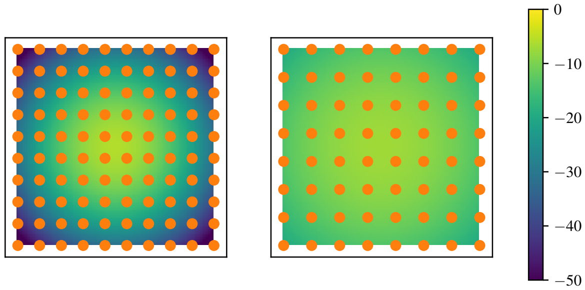


Figure 6. Plots of the log values of the un-normalized probabilities of the discrete $\mathbf{Z}$ for the Hartmann-$(1, 2)$ in the left plot and Robot pushing $(3, 2)$ in the right plot. The orange dots show the realizations of the discrete $\mathbf{Z}$.

tion is known (e.g., synthetic benchmark functions) in the evaluation of the performance metric, the maximum value $\max_{\mathbf{x}\in\mathcal{D}_{\mathbf{x}}} V_\alpha(f(\mathbf{x}, \mathbf{Z}))$ can be evaluated accurately. On the other hand, when the closed-form expression of the objective function is unknown even in the evaluation of the performance metric (e.g., the simulated robot pushing experiment), the maximum value $\max_{\mathbf{x}\in\mathcal{D}_{\mathbf{x}}} V_\alpha(f(\mathbf{x}, \mathbf{Z}))$ is estimated by $\max_{\mathbf{x}\in\mathcal{D}_T} V_\alpha(f(\mathbf{x}, \mathbf{Z})) + 0.01$ where $\mathcal{D}_T$ are input queries in the experiments with both V-UCB and ρKG^{apx}. The addition of 0.01 is to avoid $-\infty$ value in plots of the log values of the performance metric.

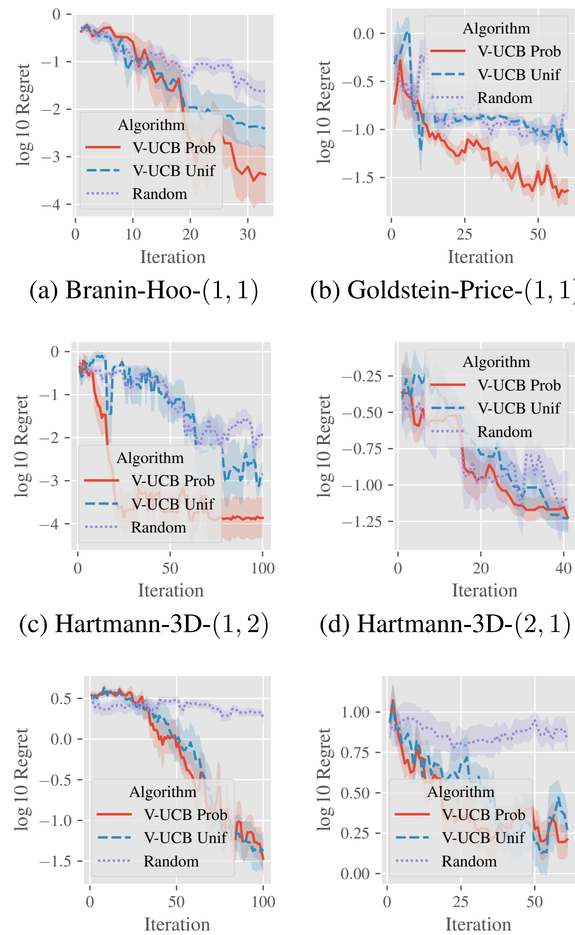
The sizes of the initial observations $\mathcal{D}_0$ are 3 for the Branin-Hoo and Goldstein-Price functions; 10 for the Hartmann-3D function; 20 for the portfolio optimization problem; and 30 for the simulated robot pushing task. The initial observations are randomly sampled for different random repetitions of the experiments, but they are the same between the same iterations in V-UCB and ρKG^{apx}.

The hyperparameters of GP (i.e., the length-scales and signal variance of the SE kernel) and the noise variance σ_n^2 are estimated by maximum likelihood estimation (Rasmussen & Williams, 2006) every 3 iterations of BO. We set a lower bound of 0.0001 for the noise variance σ_n^2 to avoid numerical errors.

To show the advantage of LNSO, we set the number of samples of $\mathbf{W}$ to be 10 for both V-UCB and ρKG^{apx}. The number of samples of $\mathbf{x}$, i.e., $|\mathcal{X}|$, in LNSO (line 9 of Algorithm 2) is 50. The radius r of the local region $\mathcal{B}$ is set to be a small value of 0.1 such that a small neural network works well: 2 hidden layers with 30 hidden neurons at each layer; the activation functions of the hidden layers and the output layer are sigmoid and linear functions, respectively.

Since the theoretical value of β_t is often considered as excessively conservative (Bogunovic et al., 2016; Srinivas et al., 2010; Bogunovic et al., 2018). We set $\beta_t = 2\log(t^2\pi^2/0.6)$ in our experiments while β_t can be tuned to achieved better exploration-exploitation trade-off (Srinivas et al., 2010) or multiple values of β_t can be used in a batch mode (Torossian et al., 2020).

Fig. 7 shows the performance advantage of our V-UCB method over a baseline that selects the input query as a random $(\mathbf{x}, \mathbf{z}) \in \mathcal{D}_{\mathbf{x}} \times \mathcal{D}_{\mathbf{z}}$, labeled as *Random* in the figure.



(a) Branin-Hoo-$(1, 1)$ (b) Goldstein-Price-$(1, 1)$

(c) Hartmann-3D-$(1, 2)$ (d) Hartmann-3D-$(2, 1)$

(e) Hartmann-6D-$(5, 1)$ (f) Portfolio optimization problem $(3, 2)$

Figure 7. Plots of the empirical performances of V-UCB and Random in optimizing synthetic functions and the portfolio optimization problem.