Supplementary Material

Corruption-Tolerant Gaussian Process Bandit Optimization

Ilija Bogunovic, Andreas Krause, Jonathan Scarlett (AISTATS 2020)

All citations below are to the reference list in the main document.

A Proof of Lemma 2 (Corrupted vs. Non-Corrupted Posterior Mean)

Our analysis uses techniques from [Chowdhury and Gopalan, 2017, Appendix C]. Let $\boldsymbol{x}$ be any point in D, and fix a time index $t \geq 1$. From the definitions of $\tilde{\mu}_t(\cdot), \mu_t(\cdot)$ and $\tilde{y}_t$ (Eq. (4), (9) and (1)), we have

$$\begin{aligned}\tilde{\mu}_t(\boldsymbol{x}) &= \boldsymbol{k}_t(\boldsymbol{x})^T(\boldsymbol{K}_t + \lambda \boldsymbol{I}_t)^{-1}\tilde{\boldsymbol{y}}_t && (37)\\ &= \boldsymbol{k}_t(\boldsymbol{x})^T(\boldsymbol{K}_t + \lambda \boldsymbol{I}_t)^{-1}\boldsymbol{y}_t + \boldsymbol{k}_t(\boldsymbol{x})^T(\boldsymbol{K}_t + \lambda \boldsymbol{I}_t)^{-1}\boldsymbol{c}_t && (38)\\ &= \mu_t(\boldsymbol{x}) + \boldsymbol{k}_t(\boldsymbol{x})^T(\boldsymbol{K}_t + \lambda \boldsymbol{I}_t)^{-1}\boldsymbol{c}_t, && (39)\end{aligned}$$

where $\tilde{\boldsymbol{y}}_t = [\tilde{y}_1, \dots, \tilde{y}_t]^T$ and $\boldsymbol{c}_t = [c_1(\boldsymbol{x}_1), \dots, c_t(\boldsymbol{x}_t)]^T$. We proceed by upper bounding the absolute difference between $\tilde{\mu}_t(\boldsymbol{x})$ and $\mu_t(\boldsymbol{x})$, i.e, $|\boldsymbol{k}_t(\boldsymbol{x})^T(\boldsymbol{K}_t + \lambda \boldsymbol{I}_t)^{-1}\boldsymbol{c}_t|$.

Let $\mathcal{H}_k(D)$ denote the RKHS associated with the kernel k and domain D. We define $\phi(\boldsymbol{x}) := k(\boldsymbol{x}, \cdot)$, where $\phi : D \to \mathcal{H}_k(D)$ maps $\boldsymbol{x} \in D$ to the RKHS associated with the kernel. For any two functions f_1, f_2 in $\mathcal{H}_k(D)$, we write $f_1^T f_2$ to denote the kernel inner product $\langle f_1, f_2 \rangle_k$, which implies that $\|f\|_k = \sqrt{f^T f}$. By the RKHS reproducing property, i.e., $f(\boldsymbol{x}) = \langle f, k(\boldsymbol{x}, \cdot)\rangle_k$ for all $\boldsymbol{x} \in D$, and the fact that $k(\boldsymbol{x}, \cdot) \in \mathcal{H}_k(D)$ for all $\boldsymbol{x} \in D$, we can write

$$k(\boldsymbol{x}, \boldsymbol{x}') = \langle k(\boldsymbol{x}, \cdot), k(\boldsymbol{x}', \cdot)\rangle_k = \langle \phi(\boldsymbol{x}), \phi(\boldsymbol{x}')\rangle_k = \phi(\boldsymbol{x})^T\phi(\boldsymbol{x}')$$

for all $\boldsymbol{x}, \boldsymbol{x}' \in D$. It also follows that $\boldsymbol{K}_t = \Phi_t\Phi_t^T$ where $\Phi_t = [\phi(\boldsymbol{x}_1), \dots, \phi(\boldsymbol{x}_t)]^T$, and $k_t(\boldsymbol{x}) = \Phi_t\phi(\boldsymbol{x})$. (Here and subsequently, the notation $f_1^T f_2 = \langle f_1, f_2\rangle_k$ similarly extends to matrix multiplication operations.)

Using these properties, we can characterize the second term of (39) as follows:

$$\begin{aligned}&|\boldsymbol{k}_t(\boldsymbol{x})^T(\boldsymbol{K}_t + \lambda \boldsymbol{I}_t)^{-1}\boldsymbol{c}_t|\\ &\quad= |\phi(\boldsymbol{x})^T\Phi_t^T(\Phi_t\Phi_t^T + \lambda \boldsymbol{I}_t)^{-1}\boldsymbol{c}_t| && (40)\\ &\quad= |\langle \phi(\boldsymbol{x})^T(\Phi_t^T\Phi_t + \lambda \boldsymbol{I}_{\mathcal{H}_k})^{-1}, \Phi_t^T\boldsymbol{c}_t\rangle_k| && (41)\\ &\quad\leq \|(\Phi_t^T\Phi_t + \lambda \boldsymbol{I}_{\mathcal{H}_k})^{-1/2}\phi(\boldsymbol{x})\|_k\|(\Phi_t^T\Phi_t + \lambda \boldsymbol{I}_{\mathcal{H}_k})^{-1/2}\Phi_t^T\boldsymbol{c}_t\|_k && (42)\\ &\quad= \sqrt{\phi(\boldsymbol{x})^T(\Phi_t^T\Phi_t + \lambda \boldsymbol{I}_{\mathcal{H}_k})^{-1}\phi(\boldsymbol{x})}\sqrt{(\Phi_t^T\boldsymbol{c}_t)^T(\Phi_t^T\Phi_t + \lambda \boldsymbol{I}_{\mathcal{H}_k})^{-1}\Phi_t^T\boldsymbol{c}_t} && (43)\\ &\quad= \lambda^{-1/2}\sigma_t(\boldsymbol{x})\sqrt{\boldsymbol{c}_t^T\Phi_t\Phi_t^T(\Phi_t\Phi_t^T + \lambda \boldsymbol{I}_t)^{-1}\boldsymbol{c}_t} && (44)\\ &\quad= \lambda^{-1/2}\sigma_t(\boldsymbol{x})\sqrt{\boldsymbol{c}_t^T\boldsymbol{K}_t(\boldsymbol{K}_t + \lambda \boldsymbol{I}_t)^{-1}\boldsymbol{c}_t} && (45)\\ &\quad\leq \lambda^{-1/2}\sigma_t(\boldsymbol{x})\sqrt{\lambda_{\max}\left(\boldsymbol{K}_t(\boldsymbol{K}_t + \lambda \boldsymbol{I}_t)^{-1}\right)\|\boldsymbol{c}_t\|_2^2} && (46)\\ &\quad\leq \lambda^{-1/2}\sigma_t(\boldsymbol{x})C\sqrt{\lambda_{\max}\left(\boldsymbol{K}_t(\boldsymbol{K}_t + \lambda \boldsymbol{I}_t)^{-1}\right)} && (47)\\ &\quad\leq C\lambda^{-1/2}\sigma_t(\boldsymbol{x}), && (48)\end{aligned}$$

where:

- Eq. (41) follows from the standard identity (see, e.g., [Chowdhury and Gopalan, 2017, Eq. (12)])

$$\Phi_t^T(\Phi_t\Phi_t^T + \lambda \boldsymbol{I}_t)^{-1} = (\Phi_t^T\Phi_t + \lambda \boldsymbol{I}_{\mathcal{H}_k})^{-1}\Phi_t^T. \tag{49}$$

- Eq. (42) is by Cauchy-Schwartz.
- The first term $\lambda^{-1/2}\sigma_t(\boldsymbol{x})$ in (44) follows from the following identity:

$$\sigma_t^2(\boldsymbol{x}) = \lambda\phi(\boldsymbol{x})^T(\Phi_t^T\Phi_t + \lambda \boldsymbol{I}_{\mathcal{H}_k})^{-1}\phi(\boldsymbol{x}). \tag{50}$$

To prove (50), we first claim the following:

$$\phi(\boldsymbol{x}) = \Phi_t^T \left(\Phi_t \Phi_t^T + \lambda \boldsymbol{I}_t\right)^{-1} \Phi_t \phi(\boldsymbol{x}) + \lambda \left(\Phi_t^T \Phi_t + \lambda \boldsymbol{I}_{\mathcal{H}_k}\right)^{-1} \phi(\boldsymbol{x}). \tag{51}$$

To see this, we apply (49) to the first term to obtain the equivalent expression

$$\phi(\boldsymbol{x}) = \left(\Phi_t^T \Phi_t + \lambda \boldsymbol{I}_{\mathcal{H}_k}\right)^{-1} \Phi_t^T \Phi_t \phi(\boldsymbol{x}) + \lambda \left(\Phi_t^T \Phi_t + \lambda \boldsymbol{I}_{\mathcal{H}_k}\right)^{-1} \phi(\boldsymbol{x}).$$

Multiplying from the left by $\left(\Phi_t^T \Phi_t + \lambda \boldsymbol{I}_{\mathcal{H}_k}\right)$, we find that this is in turn equivalent to

$$\left(\Phi_t^T \Phi_t + \lambda \boldsymbol{I}_{\mathcal{H}_k}\right) \phi(\boldsymbol{x}) = \Phi_t^T \Phi_t \phi(\boldsymbol{x}) + \lambda \phi(\boldsymbol{x}),$$

which trivially holds. Then, note by the definition of $\sigma_t^2(\boldsymbol{x})$ and (51) that

$$\begin{aligned}
\sigma_t^2(\boldsymbol{x}) &= k(\boldsymbol{x}, \boldsymbol{x}) - k_t(\boldsymbol{x})^T \left(\boldsymbol{K}_t + \lambda \boldsymbol{I}_t\right)^{-1} k_t(\boldsymbol{x}) \\
&= \phi(\boldsymbol{x})^T \phi(\boldsymbol{x}) - \phi(\boldsymbol{x})^T \Phi_t^T \left(\Phi_t \Phi_t^T + \lambda \boldsymbol{I}_t\right)^{-1} \Phi_t \phi(\boldsymbol{x}) \\
&\overset{(51)}{=} \phi(\boldsymbol{x})^T \Phi_t^T \left(\Phi_t \Phi_t^T + \lambda \boldsymbol{I}_t\right)^{-1} \Phi_t \phi(\boldsymbol{x}) + \lambda \phi(\boldsymbol{x})^T (\Phi_t^T \Phi_t + \lambda \boldsymbol{I}_{\mathcal{H}_k})^{-1} \phi(\boldsymbol{x}) \\
&\quad - \phi(\boldsymbol{x})^T \Phi_t^T \left(\Phi_t \Phi_t^T + \lambda \boldsymbol{I}_t\right)^{-1} \Phi_t \phi(\boldsymbol{x}) \\
&= \lambda \phi(\boldsymbol{x})^T (\Phi_t^T \Phi_t + \lambda \boldsymbol{I}_{\mathcal{H}_k})^{-1} \phi(\boldsymbol{x}),
\end{aligned}$$

yielding (50). The second term in (44) (i.e., the square root) follows by again applying (49).

- In (46), $\lambda_{\max}\left(\boldsymbol{K}_t(\boldsymbol{K}_t + \lambda \boldsymbol{I}_t)^{-1}\right)$ denotes the largest eigenvalue of $\boldsymbol{K}_t(\boldsymbol{K}_t + \lambda \boldsymbol{I}_t)^{-1}$.
- Eq. (47) follows since $\|\boldsymbol{c}_t\|_1 \le C$ (see (2)), and since the ℓ_1 norm is always an upper bound on the ℓ_2-norm.
- Eq. (48) follows since
$$\lambda_{\max}(\boldsymbol{K}_t(\boldsymbol{K}_t + \lambda \boldsymbol{I}_t)^{-1}) \le 1.$$
This follows since all eigenvectors of $\boldsymbol{K}_t$ are also eigenvectors of $(\boldsymbol{K}_t + \lambda \boldsymbol{I}_t)^{-1}$, and hence, the eigenvalues of $\boldsymbol{K}_t(\boldsymbol{K}_t + \lambda \boldsymbol{I}_t)^{-1}$ are of the form $\frac{\lambda(\boldsymbol{K}_t)}{\lambda(\boldsymbol{K}_t)+\lambda}$. Since, $\lambda(\boldsymbol{K}_t) \ge 0$ and $\lambda > 0$, all the eigenvalues of $\boldsymbol{K}_t(\boldsymbol{K}_t + \lambda \boldsymbol{I}_t)^{-1}$ are bounded by 1.

B Proof of Lemma 4 (Regret Bound with Known Corruption)

Conditioned on the confidence bounds (10) being valid according to Lemma 3, we have

$$\begin{aligned}
& f(\boldsymbol{x}^*) - f(\boldsymbol{x}_t) \\
& \le f(\boldsymbol{x}^*) - \tilde{\mu}_{t-1}(\boldsymbol{x}_t) + \beta_t \sigma_{t-1}(\boldsymbol{x}_t) + \lambda^{-1/2} C \sigma_{t-1}(\boldsymbol{x}_t) & (52) \\
& \le \tilde{\mu}_{t-1}(\boldsymbol{x}^*) + \lambda^{-1/2} C \sigma_{t-1}(\boldsymbol{x}^*) + \beta_t \sigma_{t-1}(\boldsymbol{x}^*) - \tilde{\mu}_{t-1}(\boldsymbol{x}_t) + \beta_t \sigma_{t-1}(\boldsymbol{x}_t) + \lambda^{-1/2} C \sigma_{t-1}(\boldsymbol{x}_t) & (53) \\
& \le \tilde{\mu}_{t-1}(\boldsymbol{x}_t) + \lambda^{-1/2} C \sigma_{t-1}(\boldsymbol{x}_t) + \beta_t \sigma_{t-1}(\boldsymbol{x}_t) - \tilde{\mu}_{t-1}(\boldsymbol{x}_t) + \beta_t \sigma_{t-1}(\boldsymbol{x}_t) + \lambda^{-1/2} C \sigma_{t-1}(\boldsymbol{x}_t) & (54) \\
& = 2(\lambda^{-1/2} C + \beta_t) \sigma_{t-1}(\boldsymbol{x}_t). & (55)
\end{aligned}$$

where (52) uses the lower confidence bound from (10), (53) uses the upper confidence bound from (10), and (54) uses the selection rule in (13).

When $\lambda \ge 1$, we have from [Chowdhury and Gopalan, 2017, Lemma 4] that[3]

$$\sum_{t=1}^T \sigma_{t-1}(\boldsymbol{x}_t) \le \sqrt{4T\lambda\gamma_T}. \tag{56}$$

This is a variant of a more widely-used upper bound on $\sum_{t=1}^T \sigma_{t-1}(\boldsymbol{x}_t)$ in terms of γ_T from [Srinivas et al., 2010].

[3] The statement of [Chowdhury and Gopalan, 2017, Lemma 4] uses $\lambda = 1 + 2/T$, but the proof states the result for general $\lambda \ge 1$.

We set $\lambda = 1$ in accordance with the lemma statement, and sum over the time steps:

$$\begin{aligned} R_T &= \sum_{t=1}^{T} \big(f(\boldsymbol{x}^*) - f(\boldsymbol{x}_t) \big) && (57) \\ &\le (2C + 2\beta_T) \sum_{t=1}^{T} \sigma_{t-1}(\boldsymbol{x}_t) && (58) \\ &\le (2C + 2\beta_T) \sqrt{4T\gamma_T} && (59) \\ &\le \left(2C + 2B + 2\sigma \sqrt{2\left(\gamma_T + \ln(\tfrac{1}{\delta})\right)} \right) \sqrt{4T\gamma_T}, && (60) \end{aligned}$$

where (58) uses (55) and the monotonicity of β_t, (59) uses (56), and (60) substitutes the choice of β_t in (7) and applies $\gamma_{T-1} \le \gamma_T$. Hence, we have $R_T = \mathcal{O}\big((B + C + \sqrt{\ln(1/\delta)})\sqrt{\gamma_T T} + \gamma_T \sqrt{T}\big)$, which establishes the lemma.

C Bounding the Simple Regret

While we have focused exclusively on the cumulative regret in our exposition, we can easily adapt our analysis to handle the simple regret similarly to the idea used in the proof of [Bogunovic et al., 2018a, Theorem 1]. We outline this procedure for Theorem 5, since all of the other results can be adapted in the same manner.

We claim that under the setup of Theorem 5, for a given $\Delta > 0$, Algorithm 1 achieves $f(\boldsymbol{x}^*) - f(\boldsymbol{x}^{(T)}) \le \Delta$ after $T = \mathcal{O}\left(\frac{\gamma_T(\beta_T + C)^2}{\Delta^2}\right)$ rounds, where the reported point $\boldsymbol{x}^{(T)}$ is defined as

$$\boldsymbol{x}^{(T)} = \boldsymbol{x}_{t^*}, \quad \text{with} \quad t^* = \operatorname*{arg\,max}_{1,\dots,T} \left\{ \tilde{\mu}_{t-1}(\boldsymbol{x}_t) - (C + \beta_t)\sigma_{t-1}(\boldsymbol{x}_t) \right\}. \tag{61}$$

To prove this claim, we continue from the end of Appendix B. We set $\lambda = 1$ as before, and define

$$\bar{r}(\boldsymbol{x}_t) := f(\boldsymbol{x}^*) - \tilde{\mu}_{t-1}(\boldsymbol{x}_t) + (C + \beta_t)\sigma_{t-1}(\boldsymbol{x}_t).$$

Using (52), we have $f(\boldsymbol{x}^*) - f(\boldsymbol{x}) = r(\boldsymbol{x}_t) \le \bar{r}(\boldsymbol{x}_t)$ for each $t \ge 1$. From the definition of the reported point $\boldsymbol{x}^{(T)}$ in (61), we have that t^* is the time index with the smallest value of $\bar{r}(\boldsymbol{x}_t)$. It follows that

$$\begin{aligned} \bar{r}(\boldsymbol{x}^{(T)}) &\le \frac{1}{T} \sum_{t=1}^{T} 2(C + \beta_t)\sigma_{t-1}(\boldsymbol{x}_t) && (62) \\ &\le \frac{2(C + \beta_T)}{T} \sum_{t=1}^{T} \sigma_{t-1}(\boldsymbol{x}_t) && (63) \\ &\le \frac{2(C + \beta_T)}{T} \sqrt{4T\gamma_T}, && (64) \end{aligned}$$

where (62) upper bounds the minimum by the average, (63) uses the monotonicity of β_t, and (64) uses (56) with $\lambda = 1$.

Re-arranging (64), we find that after $T = \mathcal{O}\big(\frac{\gamma_T(\beta_T + C)^2}{\Delta^2}\big)$ time steps, $\bar{r}(\boldsymbol{x}^{(T)}) \le \Delta$, which further implies that $r(\boldsymbol{x}^{(T)}) \le \Delta$.

D Proof of Lemma 6 (Total Corruption Observed by S)

We follow the proof of [Lykouris et al., 2018, Lemma 3.3], making use of the following martingale concentration inequality.[4]

[4] This result is presented in [Beygelzimer et al., 2011] for the filtration $\mathcal{F}_t$ generated by $M_1, \dots, M_T$ itself, but the proof applies in the general case. To prove Lemma 6, we could in fact resort to the classical martingale concentration bound of Freedman [Freedman et al., 1975], but we found the form given in [Beygelzimer et al., 2011] to be more convenient.

Lemma 11. [Beygelzimer et al., 2011, Lemma 1] *Let $M_1, \dots, M_T$ be a sequence of real-valued random variables forming a martingale with respect to a filtration $\{\mathcal{F}_t\}$, i.e., $\mathbb{E}[M_t|\mathcal{F}_{t-1}] = 0$, and suppose that $M_t \leq R$ almost surely. Then for any $\delta > 0$, the following holds:*

$$\mathbb{P}\left[\sum_{t=1}^{T} M_t \leq \frac{V}{R}(e-2) + R\ln(1/\delta)\right] \geq 1-\delta,$$

where $V = \sum_{t=1}^{T} \mathbb{E}[M_t^2|\mathcal{F}_{t-1}]$.

Let $\boldsymbol{x}_t^{(S)}$ be the point that would be selected at time t if instance S were chosen. We let $C_t = |c_t(\boldsymbol{x}_t^{(S)})|\mathbb{1}\{A_t = S\}$ denote the amount of corruption observed by instance S at time t in Algorithm 2.

Let $\mathcal{H}_{t-1}$ denote the history (i.e., all selected instances $A_i \in \{F, S\}$, inputs $\boldsymbol{x}_i \in D$, and observations $\tilde{y}_i \in \mathbb{R}$) prior to round t. Noting that $\boldsymbol{x}_t^{(S)}$ is deterministic given $\mathcal{H}_{t-1}$, we find that C_t is a random variable equaling $|c_t(\boldsymbol{x}_t^{(S)})|$ with probability $\rho := \min\{1, C^{-1}\}$ and 0 otherwise. As a result, we can define the following martingale sequence:

$$M_t = C_t - \mathbb{E}[C_t|\mathcal{H}_{t-1}],$$

where $\mathbb{E}[C_t|\mathcal{H}_{t-1}] = \rho|c_t(\boldsymbol{x}_t^{(S)})|$ as stated above. Since $c_t(\boldsymbol{x}) \in [-B_0, B_0]$ for all t and $\boldsymbol{x} \in D$ (see Section 2), we have $M_t \leq B_0$ for all t. Hence, we can set $R = B_0$ in Lemma 11.

Next, we note the following:

$$\begin{aligned}\mathbb{E}[M_t^2|\mathcal{H}_{t-1}] &= \rho\left(|c_t(\boldsymbol{x}_t^{(S)})| - \rho|c_t(\boldsymbol{x}_t^{(S)})|\right)^2 + (1-\rho)\left(\rho|c_t(\boldsymbol{x}_t^{(S)})|\right)^2 \\ &= \rho c_t(\boldsymbol{x}_t^{(S)})^2(1-\rho)^2 + (1-\rho)(\rho c_t(\boldsymbol{x}_t^{(S)}))^2 \\ &\leq \rho c_t(\boldsymbol{x}_t^{(S)})^2 + \rho c_t(\boldsymbol{x}_t^{(S)})^2 \\ &= 2\rho c_t(\boldsymbol{x}_t^{(S)})^2 \\ &\leq 2\rho B_0|c_t(\boldsymbol{x}_t^{(S)})|.\end{aligned}$$

where the two inequalities use $\rho \in [0, 1]$ and $c_t(\boldsymbol{x}_t^{(S)}) \leq B_0$ respectively. By summing over all the rounds and using the definition of C in (2), we obtain

$$V = \sum_{t=1}^{T} \mathbb{E}[M_t^2|\mathcal{H}_{t-1}] \leq 2B_0\rho\sum_{t=1}^{T}|c_t(\boldsymbol{x}_t^{(S)})| \leq 2B_0\rho C \leq 2B_0,$$

since $\rho \leq C^{-1}$. Applying Lemma 11, we have with probability at least $1-\delta$ that

$$\sum_{t=1}^{T} M_t \leq \frac{2B_0}{B_0}(e-2) + B_0\ln(1/\delta) \leq 2 + B_0\ln(1/\delta). \tag{65}$$

Finally, we complete the proof of Lemma 6 by adding the total expected corruption:

$$\begin{aligned}\sum_{t=1}^{T} C_t &= \sum_{t=1}^{T} M_t + \sum_{t=1}^{T} \mathbb{E}\left[C_t|\mathcal{H}_{t-1}\right] \\ &\leq 3 + B_0\ln(1/\delta),\end{aligned}$$

where we have used (65) and $\sum_{t=1}^{T}\mathbb{E}\left[C_t|\mathcal{H}_{t-1}\right] = \rho\sum_{t=1}^{T}|c_t(\boldsymbol{x}_t^{(S)})| \leq \rho C \leq 1$.

E Proof of Lemma 7 (Characterizing the Points Not Sampled by F)

Consider any round $t \in \{1, \dots, T\}$ and any point $\boldsymbol{x} \in \mathcal{S}_t$ (see (23)). We wish to show that F never selects $\boldsymbol{x}$, i.e., $\boldsymbol{x}_t \neq \boldsymbol{x}$. To establish this, it suffices to prove that

$$\min_{A\in\{F,S\}} \overline{\mathrm{ucb}}_{t_A-1}^{(A)}(\boldsymbol{x}; 1) < \min_{A\in\{F,S\}} \overline{\mathrm{ucb}}_{t_A-1}^{(A)}(\boldsymbol{x}'; 1). \tag{66}$$

for some $\boldsymbol{x}' \in D$; this means that $\boldsymbol{x}'$ is favored over $\boldsymbol{x}$ according to the selection rule of F.

To show (66), we first trivially write

$$\min_{A \in \{F,S\}} \overline{\mathrm{ucb}}^{(A)}_{t_A - 1}(\boldsymbol{x}; 1) \le \overline{\mathrm{ucb}}^{(S)}_{t_S - 1}(\boldsymbol{x}; 1). \tag{67}$$

Since $\boldsymbol{x} \in \mathcal{S}_t$, by the definition of $\mathcal{S}_t$ in (23), there exists $\boldsymbol{x}' \in D$ such that

$$\overline{\mathrm{ucb}}^{(S)}_{t_S - 1}(\boldsymbol{x}; 1) < \overline{\mathrm{lcb}}^{(S)}_{t_S - 1}(\boldsymbol{x}'; 1). \tag{68}$$

Moreover, the following two equations provide upper bounds on $\overline{\mathrm{lcb}}^{(S)}_{t_S - 1}(\boldsymbol{x}'; 1)$:

$$\overline{\mathrm{lcb}}^{(S)}_{t_S - 1}(\boldsymbol{x}'; 1) \le \overline{\mathrm{ucb}}^{(S)}_{t_S - 1}(\boldsymbol{x}'; 1) \tag{69}$$
$$\overline{\mathrm{lcb}}^{(S)}_{t_S - 1}(\boldsymbol{x}'; 1) \le \overline{\mathrm{ucb}}^{(F)}_{t_F - 1}(\boldsymbol{x}'; 1), \tag{70}$$

where (69) follows from the validity of the confidence bounds (see (21)), and (70) is due to $A_t = F$, which means that the condition (19) used in Fast-Slow GP-UCB (Line 12) is not satisfied and thus it cannot hold that $\overline{\mathrm{lcb}}^{(S)}_{t_S - 1}(\boldsymbol{x}'; 1) > \overline{\mathrm{ucb}}^{(F)}_{t_F - 1}(\boldsymbol{x}'; 1)$.

From (69) and (70) we have $\overline{\mathrm{lcb}}^{(S)}_{t_S - 1}(\boldsymbol{x}'; 1) \le \min_{\{F,S\}} \overline{\mathrm{ucb}}^{(A)}_{t_A - 1}(\boldsymbol{x}'; 1)$, and from (67) and (68) we have $\overline{\mathrm{lcb}}^{(S)}_{t_S - 1}(\boldsymbol{x}'; 1) > \min_{A \in \{F,S\}} \overline{\mathrm{ucb}}^{(A)}_{t_A - 1}(\boldsymbol{x}; 1)$, which together prove that (66) holds.

F Proof of Lemma 8 (Characterizing the Points Ruled Out via S)

Although we consider the S instance run with $\alpha = 2$, we are interested in how long it takes before the following (corresponding to $\alpha = 1$) is observed for the given suboptimal $\boldsymbol{x}$ and some $\boldsymbol{x}' \in D$:

$$\mathrm{ucb}^{(S)}_{t_S - 1}(\boldsymbol{x}; 1) < \mathrm{lcb}^{(S)}_{t_S - 1}(\boldsymbol{x}'; 1). \tag{71}$$

Since $\overline{\mathrm{ucb}}^{(S)}_{t_S - 1}$ and $\overline{\mathrm{lcb}}^{(S)}_{t_S - 1}$ are tighter confidence bounds than $\mathrm{ucb}^{(S)}_{t_S - 1}$ and $\mathrm{lcb}^{(S)}_{t_S - 1}$, (71) holding implies that

$$\overline{\mathrm{ucb}}^{(S)}_{t_S - 1}(\boldsymbol{x}; 1) < \overline{\mathrm{lcb}}^{(S)}_{t_S - 1}(\boldsymbol{x}'; 1), \tag{72}$$

meaning that $\boldsymbol{x} \in \mathcal{S}_{t_S}$ (see (23)). Since $\overline{\mathrm{ucb}}^{(S)}_{t_S - 1}$ and $\overline{\mathrm{lcb}}^{(S)}_{t_S - 1}$ are monotone, (72) holding for some t_S means that it continues to hold for all $t_S' > t_S$. Hence, to establish the lemma, it suffices to show that after t_S rounds (with t_S given in (24)), there exists a point $\boldsymbol{x}' \in D$ such that (71) holds.

Since this proof only concerns points selected by S, we abuse notation slightly and let $\boldsymbol{x}_i$ denote the i-th point queried by S. We use the fact that the instant regret incurred by the S instance satisfies

$$r(\boldsymbol{x}_i) = f(\boldsymbol{x}^*) - f(\boldsymbol{x}_i) \le 2\alpha \beta^{(S)}_i \sigma^{(S)}_{i-1}(\boldsymbol{x}_i) \tag{73}$$

(via an identical argument[5] to (55)), and the sum of posterior standard deviations satisfies

$$\frac{1}{t_S} \sum_{i=1}^{t_S} \sigma^{(S)}_{i-1}(\boldsymbol{x}_i) \le \sqrt{\frac{4\gamma_{t_S}}{t_S}} \tag{74}$$

when we set $\lambda = 1$ (by a direct application of (56)). Combining these gives

$$\frac{1}{t_S} \sum_{i=1}^{t_S} r(\boldsymbol{x}_i) \le \frac{1}{t_S} \sum_{i=1}^{t_S} 2\alpha \beta^{(S)}_i \sigma^{(S)}_{i-1}(\boldsymbol{x}_i) \le \sqrt{\frac{C_1 (\beta^{(S)}_{t_S})^2 \gamma_{t_S}}{t_S}}, \tag{75}$$

[5] See also (92) in Appendix G.

where $C_1 = 16\alpha^2$. It is useful to "invert" the right-hand side of (75); to do this, we define the function

$$\tau(\Delta) = \min\left\{\tau \,:\, \sqrt{\frac{C_1(\beta_\tau^{(S)})^2\gamma_\tau}{\tau}} \le \Delta\right\}. \tag{76}$$

Since (75) and (76) state that the "average" value of $2\alpha\beta_i^{(S)}\sigma_{i-1}^{(S)}(\boldsymbol{x}_i)$ by time $\tau(\Delta)$ is at most Δ, we deduce that

$$\forall\Delta > 0, \exists i \le \tau(\Delta) \text{ such that } 2\alpha\beta_i^{(S)}\sigma_{i-1}^{(S)}(\boldsymbol{x}_i) \le \Delta. \tag{77}$$

That is, at least one time index i yields a value less than or equal to the average.

Now consider the given $\boldsymbol{x} \in D$ with instant regret satisfying $r(\boldsymbol{x}) \ge \Delta_0 > 0$ in accordance with the lemma statement. Setting the parameter $\Delta = \frac{\Delta_0}{10}$ in (77) gives

$$\exists i \le \tau(\Delta_0/10) \text{ such that } 2\alpha\beta_i^{(S)}\sigma_{i-1}^{(S)}(\boldsymbol{x}_i) \le \frac{\Delta_0}{10} \tag{78}$$

$$\text{and hence } r(\boldsymbol{x}_i) \le \frac{\Delta_0}{10}, \tag{79}$$

where (79) follows from (73). This means that $\boldsymbol{x}_i$ is much closer to optimal than $\boldsymbol{x}$ is. The properties in (78) and (79) allow us to characterize the confidence bounds of $\boldsymbol{x}_i$:

$$\begin{aligned}
\mathrm{ucb}_{i-1}^{(S)}(\boldsymbol{x}_i;\alpha) &= \mathrm{lcb}_{i-1}^{(S)}(\boldsymbol{x}_i;\alpha) + 2\alpha\beta_i^{(S)}\sigma_{i-1}^{(S)}(\boldsymbol{x}_i) && (80)\\
&\le f(\boldsymbol{x}_i) + 2\alpha\beta_i^{(S)}\sigma_{i-1}^{(S)}(\boldsymbol{x}_i) && (81)\\
&\le f(\boldsymbol{x}^*) + \frac{\Delta_0}{10}, && (82)
\end{aligned}$$

where (80) uses the definition of the confidence bounds in (15)–(16), (81) uses the validity of the confidence bounds in (21), and (82) uses (78). Similarly,

$$\begin{aligned}
\mathrm{lcb}_{i-1}^{(S)}(\boldsymbol{x}_i;\alpha) &= \mathrm{ucb}_{i-1}^{(S)}(\boldsymbol{x}_i;\alpha) - 2\alpha\beta_i^{(S)}\sigma_{i-1}^{(S)}(\boldsymbol{x}_i) && (83)\\
&\ge \mathrm{ucb}_{i-1}^{(S)}(\boldsymbol{x}^*;\alpha) - 2\alpha\beta_i^{(S)}\sigma_{i-1}^{(S)}(\boldsymbol{x}_i) && (84)\\
&\ge f(\boldsymbol{x}^*) - 2\alpha\beta_i^{(S)}\sigma_{i-1}^{(S)}(\boldsymbol{x}_i) && (85)\\
&\ge f(\boldsymbol{x}^*) - \frac{\Delta_0}{10}, && (86)
\end{aligned}$$

where (83) is the same as (80), (84) uses the UCB selection rule, (85) uses the validity of the confidence bounds, and (86) uses (78). Combining (82) and (86), we find that the confidence interval $[\mathrm{lcb}_{i-1}^{(S)}(\boldsymbol{x}_i;\alpha), \mathrm{ucb}_{i-1}^{(S)}(\boldsymbol{x}_i;\alpha)]$ is within the range

$$\mathcal{I} = \left[f(\boldsymbol{x}^*) - \frac{\Delta_0}{10}, f(\boldsymbol{x}^*) + \frac{\Delta_0}{10}\right]. \tag{87}$$

It also holds that

$$\mathrm{ucb}_{i-1}^{(S)}(\boldsymbol{x};\alpha) \le \mathrm{ucb}_{i-1}^{(S)}(\boldsymbol{x}_i;\alpha) \tag{88}$$

by the UCB rule used in the S instance. For this fixed $i \le \tau(\Delta_0/10)$ and $\boldsymbol{x}_i$, there are then two possible cases that we need to consider:

1. If it also holds that $\mathrm{ucb}_{i-1}^{(S)}(\boldsymbol{x};\alpha) < \mathrm{lcb}_{i-1}^{(S)}(\boldsymbol{x}_i;\alpha)$, then we immediately obtain

   $$\mathrm{ucb}_{i-1}^{(S)}(\boldsymbol{x};1) < \mathrm{lcb}_{i-1}^{(S)}(\boldsymbol{x}_i;1)$$

   because we chose $\alpha = 2$, and decreasing α only makes $\mathrm{ucb}_{i-1}^{(S)}(\cdot;\alpha)$ decrease and $\mathrm{lcb}_{i-1}^{(S)}(\cdot;\alpha)$ increase (see (15)–(16)). Hence, the condition in (71) holds as required.

2. Otherwise, by (88), we must have

$$\mathrm{lcb}_{i-1}^{(S)}(\boldsymbol{x}_i, \alpha) \le \mathrm{ucb}_{i-1}^{(S)}(\boldsymbol{x}; \alpha) \le \mathrm{ucb}_{i-1}^{(S)}(\boldsymbol{x}_i; \alpha).$$

   By (82) and (86), this means that $\mathrm{ucb}_{i-1}^{(S)}(\boldsymbol{x}; \alpha)$ lies in the interval $\mathcal{I}$ given in (87).

   Since the confidence bounds (21) are valid and $f(\boldsymbol{x}) \le f(\boldsymbol{x}^*) - \Delta_0$ (i.e., $r(\boldsymbol{x}) \ge \Delta_0$), we must also have $\mathrm{lcb}_{i-1}^{(S)}(\boldsymbol{x}; \alpha) \le f(\boldsymbol{x}^*) - \Delta_0$. Comparing this with $\mathcal{I}$ above, we notice a gap of at least $\frac{9\Delta_0}{10}$ between the upper and lower confidence bounds at $\boldsymbol{x}$. Let this gap be denoted by $\mathrm{Gap}(\alpha) \ge \frac{9\Delta_0}{10}$.

   The confidence bounds $\mathrm{ucb}_{i-1}^{(S)}(\boldsymbol{x}; \alpha)$ and $\mathrm{lcb}_{i-1}^{(S)}(\boldsymbol{x}; \alpha)$ are equal to $\tilde{\mu} \pm \frac{1}{2}\mathrm{Gap}(\alpha)$, where $\tilde{\mu}$ is shorthand for the corrupted posterior mean. When we compare to $\mathrm{ucb}_{i-1}^{(S)}(\boldsymbol{x}; 1)$ and $\mathrm{lcb}_{i-1}^{(S)}(\boldsymbol{x}; 1)$, the value $\tilde{\mu}$ remains unchanged, but we have $\mathrm{Gap}(1) = \frac{1}{\alpha}\mathrm{Gap}(\alpha)$; see (15)–(16). Therefore, we have

$$\begin{aligned} \mathrm{ucb}_{i-1}^{(S)}(\boldsymbol{x}; 1) &= \mathrm{ucb}_{i-1}^{(S)}(\boldsymbol{x}; \alpha) - \frac{1}{2}\Big(1 - \frac{1}{\alpha}\Big)\mathrm{Gap}(\alpha) \\ &\le \mathrm{ucb}_{i-1}^{(S)}(\boldsymbol{x}; \alpha) - \frac{1}{2}\Big(1 - \frac{1}{\alpha}\Big)\frac{9\Delta_0}{10} \end{aligned}$$

   since $\mathrm{Gap}(\alpha) \ge \frac{9\Delta_0}{10}$. Substituting $\alpha = 2$ gives $\mathrm{ucb}_{i-1}^{(S)}(\boldsymbol{x}; 1) \le \mathrm{ucb}_{i-1}^{(S)}(\boldsymbol{x}; 2) - \frac{9\Delta_0}{40}$. Since the width of the interval $\mathcal{I}$ (in which $\mathrm{ucb}_{i-1}^{(S)}(\boldsymbol{x}; 2)$ lies) is only $\frac{2\Delta_0}{10} = \frac{8\Delta_0}{40}$, we conclude that $\mathrm{ucb}_{i-1}^{(S)}(\boldsymbol{x}; 1)$ lies strictly below $\mathcal{I}$.

   On the other hand, using (82) and (86), we see that the entire confidence interval for $\boldsymbol{x}_i$ lies within $\mathcal{I}$ (recall that replacing $\alpha > 1$ by $\alpha = 1$ only shrinks this interval). Hence, $\mathrm{ucb}_{i-1}^{(S)}(\boldsymbol{x}; 1) < \mathrm{lcb}_{i-1}^{(S)}(\boldsymbol{x}_i; 1)$, as required.

Recall that the above findings all correspond to some time index $i \le \tau(\Delta_0/10)$. Hence, (24) follows by setting $t_S = \tau(\Delta_0/10)$.

G Proof of Theorem 9 (Regret Bound in the Known-or-Zero Setting)

Throughout the proof, we condition on the events (20)–(22) that simultaneously hold with probability at least $1 - \frac{4\delta}{5}$.

G.1 Non-corrupted case

Recall that at time t, the chosen instance and input are denoted by A_t and $\boldsymbol{x}_t$, respectively, and we use t_A to denote the number of times an instance $A \in \{F, S\}$ has been chosen up to time t.

In the non-corrupted case, the condition (19) cannot hold (conditioned on the events (20) and (21)), since the confidence bounds for both S and F are valid and hence $\mathrm{ucb}_{t_F}^{(F)}(\boldsymbol{x}; 1)$ can never be smaller than $\mathrm{lcb}_{t_S}^{(S)}(\boldsymbol{x}; 1)$. Consequently, Algorithm 2 selects only S or F, and never switches permanently to Algorithm 1.

First, we consider the case that $A_t = S$ is used to select $\boldsymbol{x}_t$ for some t. We have

$$\begin{aligned} f(\boldsymbol{x}^*) - f(\boldsymbol{x}_t) &\le \mathrm{ucb}_{t_S-1}^{(S)}(\boldsymbol{x}^*; \alpha) - f(\boldsymbol{x}_t) && (89) \\ &\le \mathrm{ucb}_{t_S-1}^{(S)}(\boldsymbol{x}_t; \alpha) - f(\boldsymbol{x}_t) && (90) \\ &\le \mathrm{ucb}_{t_S-1}^{(S)}(\boldsymbol{x}_t; \alpha) - \mathrm{lcb}_{t_S-1}^{(S)}(\boldsymbol{x}_t; \alpha) && (91) \\ &\le 2\alpha\beta_{t_S}^{(S)}\sigma_{t_S-1}^{(S)}(\boldsymbol{x}_t), && (92) \end{aligned}$$

where (89) and (91) use the validity of the confidence bounds, (90) follows from the selection rule of S, and (92) uses the definitions (15)–(16).

Next, we consider the case that $A_t = F$ is used to select $\boldsymbol{x}_t$ for some t. We have

$$f(\boldsymbol{x}^*) - f(\boldsymbol{x}_t) \le \min_{A \in \{F,S\}} \overline{\text{ucb}}_{t_A-1}^{(A)}(\boldsymbol{x}^*; 1) - f(\boldsymbol{x}_t) \tag{93}$$

$$\le \min_{A \in \{F,S\}} \overline{\text{ucb}}_{t_A-1}^{(A)}(\boldsymbol{x}_t; 1) - f(\boldsymbol{x}_t) \tag{94}$$

$$\le \overline{\text{ucb}}_{t_F-1}^{(F)}(\boldsymbol{x}_t; 1) - f(\boldsymbol{x}_t) \tag{95}$$

$$\le \overline{\text{ucb}}_{t_F-1}^{(F)}(\boldsymbol{x}_t; 1) - \overline{\text{lcb}}_{t_F-1}^{(F)}(\boldsymbol{x}_t; 1) \tag{96}$$

$$\le 2\beta_{t_F}^{(F)} \sigma_{t_F-1}^{(F)}(\boldsymbol{x}_t), \tag{97}$$

where (93) and (96) use the validity of the confidence bounds, (94) uses the selection rule of F, and (97) follows similarly to (92) by noting that the intersected confidence bounds are at least as tight as the non-intersected ones.

The regret R_T of Algorithm 2 after T rounds can be trivially bounded by the sum $R_T^{(S)} + R_T^{(F)}$, where $R_T^{(A)}$ is the regret of instance A when run for T rounds in the non-corrupted case:

$$R_T \le R_T^{(F)} + R_T^{(S)} \tag{98}$$

$$\le \sum_{t_F=1}^{T} 2\beta_{t_F}^{(F)} \sigma_{t_F-1}^{(F)}(\boldsymbol{x}_{t_F}) + \sum_{t_S=1}^{T} 2\alpha\beta_{t_S}^{(S)} \sigma_{t_S-1}^{(S)}(\boldsymbol{x}_{t_S}) \tag{99}$$

$$\le 2\beta_T^{(F)} \sum_{t_F=1}^{T} \sigma_{t_F-1}^{(F)}(\boldsymbol{x}_{t_F}) + 2\alpha\beta_T^{(S)} \sum_{t_S=1}^{T} \sigma_{t_S-1}^{(S)}(\boldsymbol{x}_{t_S}) \tag{100}$$

$$\le 2\beta_T^{(F)} \sqrt{4T\gamma_T} + 2\alpha\beta_T^{(S)} \sqrt{4T\gamma_T} \tag{101}$$

$$\le 4\alpha\beta_T^{(S)} \sqrt{4T\gamma_T}, \tag{102}$$

where (99) follows from (92) and (97), (100) follows since both $\beta_{t_S}^{(S)}$ and $\beta_{t_F}^{(F)}$ are non-decreasing in the time index, (101) follows from (56) by setting $\lambda = 1$, and (102) follows since $\alpha \ge 1$ and $\beta_T^{(S)} \ge \beta_T^{(F)}$ (see (25)–(26)).

Substituting $\beta_T^{(S)} = B + \sigma\sqrt{2\left(\gamma_{T-1} + \ln\left(\frac{5}{\delta}\right)\right)} + (3 + B_0 \ln\left(\frac{5}{\delta}\right))$ and $\alpha = 2$ in (102), we arrive at the regret bound, i.e., with probability at least $1 - \frac{4}{5}\delta \ge 1 - \delta$, the regret of Algorithm 2 after T rounds is

$$R_T = \mathcal{O}\left(\left(B + B_0 \ln(\tfrac{1}{\delta}) + \sqrt{\ln(\tfrac{1}{\delta})}\right)\sqrt{T\gamma_T} + \gamma_T\sqrt{T}\right).$$

G.2 C-corrupted case

Similarly to the non-corrupted case, we condition on (20)–(22), and we set $\lambda = 1$. We first address the two parts of Fast-Slow GP-UCB whose contributions to the cumulative regret are the simplest to handle: That from Algorithm 1, and that from the slow instance S.

Supposing that Algorithm 1 is run for $T' \le T$ rounds, we simply use the confidence bounds (22) and apply Lemma 4 (with $\frac{\delta}{5}$ in place of δ): If $\beta_t^{(A_1)} = B + \sigma\sqrt{2\left(\gamma_{t-1} + \ln\left(\frac{5}{\delta}\right)\right)} + C$, then the cumulative regret after T' rounds satisfies

$$R_{T'}^{(A_1)} = \mathcal{O}\left(\left(B + C + \sqrt{\ln(\tfrac{1}{\delta})}\right)\sqrt{\gamma_{T'}T'} + \gamma_{T'}\sqrt{T'}\right). \tag{103}$$

The regret obtained by S is analyzed in the same way via Lemma 4, but with $B_0 \ln\left(\frac{5}{\delta}\right)$ in place of C, and the confidence bounds (21) in place of (22). Lemma 4 then implies that the regret coming from S for a total of T' rounds satisfies

$$R_{T'}^{(S)} = \mathcal{O}\left(\left(B + B_0 \ln(\tfrac{1}{\delta}) + \sqrt{\ln(\tfrac{1}{\delta})}\right)\sqrt{\gamma_{T'}T'} + \gamma_{T'}\sqrt{T'}\right), \tag{104}$$

which is the same as (103) but with $B_0 \ln\left(\frac{5}{\delta}\right)$ in place of C (and possibly a different T' value). It now only remains to bound the regret of the F instance in the corrupted case.

Regret incurred by the F instance. First, we recall a few facts. The F-confidence bounds in (20) are only valid when there is no corruption, and hence they cannot be used to characterize the regret of the F instance in the corrupted case. Unlike the F-confidence bounds, the S-confidence bounds in (21) are valid even in the corrupted case, and they are useful since the F rule explicitly depends on them (Fast-Slow GP-UCB, Line 6). In Lemma 7, we have shown that no point that is suboptimal according to the S-confidence bounds is sampled by the F instance. Subsequently, in Lemma 8, we have characterized how many points need to be queried in S before this occurs. More formally, the results of Lemmas 7 and 8 (with $\alpha = 2$) state that by time

$$t_S = \min\big\{\tau \,:\, 8\beta_{\tau}^{(S)}\sqrt{\tfrac{\gamma_\tau}{\tau}} \le \tfrac{\Delta_0}{10}\big\}, \tag{105}$$

all Δ_0-suboptimal points are ruled out and are not sampled by F in the subsequent time steps. We observe that the following two statements are equivalent:

- After time $t_S = \min\big\{\tau \,:\, 8\beta_{\tau}^{(S)}\sqrt{\frac{\gamma_\tau}{\tau}} \le \frac{\Delta_0}{10}\big\}$, the instant regret of each point selected by F is at most Δ_0;
- After time t_S, the instant regret of each point selected by F is at most $80\beta_{t_S}^{(S)}\sqrt{\frac{\gamma_{t_S}}{t_S}}$.

This is by a simple inversion; if we set $\Delta_0 = 80\beta_{\tau}^{(S)}\sqrt{\frac{\gamma_\tau}{\tau}}$ in (105) then it trivially holds that $8\beta_{\tau}^{(S)}\sqrt{\frac{\gamma_\tau}{\tau}} \le \frac{\Delta_0}{10}$.

We now seek to characterize how many times F is selected in between successive selections of S. If $C \le 1$, then this is trivial, since S is always selected, so in the following we focus on $C > 1$. We will establish that with probability at least $1 - \frac{\delta}{5}$, in between any two selections of S (or prior to the first such selection), there are at most $C \ln\frac{5T}{\delta}$ selections of F with probability at least $1 - \frac{\delta}{5}$. We henceforth denote this event by $\mathcal{A}$.

To establish the preceding claim, fix an integer $N > 0$, and observe that after any given selection of S, the probability of selecting F for the next N rounds is $\big(1 - \frac{1}{C}\big)^N \le e^{-N/C}$. Hence, if $N = C\ln\frac{1}{\delta'}$, then the probability is at most δ'. The number of selections of S is trivially at most T, so taking a union bound over at most T associated events, we obtain $\mathbb{P}[\mathcal{A}] \ge 1 - \frac{\delta}{5}$ when $\delta' = \frac{\delta}{5T}$.

By the union bound, the event $\mathcal{A}$ and the events in (20)–(22) hold simultaneously with probability at least $1 - \frac{4}{5}\delta - \frac{1}{5}\delta = 1 - \delta$. Conditioned on these events, when Fast-Slow GP-UCB is run for T rounds, the cumulative regret of the points selected by F satisfies[6]

$$\begin{aligned}
R_T^{(F)} &\le 2B_0 N + N \cdot 80\beta_T^{(S)}\sqrt{\gamma_T}\sum_{t_S=1}^{\lfloor \frac{T}{N}\rfloor}\sqrt{\tfrac{1}{t_S}} && (106)\\
&\le 2B_0 N + 80N\beta_T^{(S)}\sqrt{4\gamma_T\tfrac{T}{N}} && (107)\\
&= 2B_0 N + 80\beta_T^{(S)}\sqrt{4N\gamma_T T}, && (108)
\end{aligned}$$

where:

- (106) is established using the equivalence stated after (105) and the definition of $\mathcal{A}$ as follows: First, the instant regret bound $80\beta_{t_S}^{(S)}\sqrt{\frac{\gamma_{t_S}}{t_S}}$ is upper bounded by $80\beta_T^{(S)}\sqrt{\gamma_T}\sqrt{\frac{1}{t_S}}$ because $\beta_{t_S}^{(S)}$ and γ_{t_S} are monotone. Then, when summing this weakened upper bound over all time instants, the conditioning on $\mathcal{A}$ means that the worst case (i.e., giving the highest upper bound) is that there are exactly N selections of F before each selection of S. The first such selection incurs cumulative regret at most $2B_0N$ since $f(\boldsymbol{x}) \in [-B_0, B_0]$, and the subsequent selections indexed by t_S incur at most $N \cdot 80\beta_T^{(S)}\sqrt{\gamma_T}\sqrt{\frac{1}{t_S}}$.
- (107) uses $\sum_{t=1}^{T}\frac{1}{\sqrt{t}} \le 1 + \int_{t=1}^{T}\frac{1}{\sqrt{t}}dt \le \sqrt{4T}$.

Substituting $N = C\ln(\frac{5T}{\delta})$ and $\beta_T^{(S)}$ (stated above (21)) into (108), we obtain

$$R_T^{(F)} = \mathcal{O}\bigg(\sqrt{C\ln(\tfrac{T}{\delta})}\Big(\Big(B + B_0\ln(\tfrac{1}{\delta}) + \sqrt{\ln(\tfrac{1}{\delta})}\Big)\sqrt{\gamma_T T} + \gamma_T\sqrt{T}\Big) + B_0 C\ln(\tfrac{T}{\delta})\bigg). \tag{109}$$

[6] We could slightly improve this bound by replacing γ_T by $\gamma_{\frac{T}{N}}$, but we proceed with the former since it is simpler and only slightly weaker.

Overall corrupted regret bound. The obtained regret bounds (103), (104), and (109) hold simultaneously with probability at least $1-\delta$. We obtain our final bound by noting that the cumulative regret of Fast-Slow GP-UCB after T rounds can be trivially upper bounded by the sum of the individual regrets in (103), (104), and (109), where in both (103) and (104) we upper bound T' by T. Therefore, with probability at least $1-\delta$, after T rounds, we obtain

$$R_T = \mathcal{O}\left((1+C)\ln(\tfrac{T}{\delta}) \left(\left(B + B_0 \ln(\tfrac{1}{\delta}) + \sqrt{\ln(\tfrac{1}{\delta})} \right) \sqrt{\gamma_T T} + \gamma_T \sqrt{T} \right) \right). \tag{110}$$

Note that we have weakened $\sqrt{C\ln(\frac{T}{\delta})}$ in (109) to $C\ln(\frac{T}{\delta})$ for the sake of attaining a simpler bound with fewer terms, since a $C\sqrt{T\gamma_T}$ term is already present in (103).

H Further Details on the Proof of Theorem 10 (Regret Bound with Unknown C)

As stated in Theorem 10, we set the exploration parameter for each layer ℓ as follows:

$$\beta_{t_\ell}^{(\ell)} = B + \sigma \sqrt{2\left(\gamma_{t_\ell - 1} + \ln\left(\frac{4(1+\log_2 T)}{\delta}\right)\right)} + 3 + B_0 \ln\left(\frac{4(1+\log_2 T)}{\delta}\right). \tag{111}$$

This ensures the following confidence bound for each $\ell \in \{1,\dots,\lceil \log_2 T \rceil\}$ such that $2^\ell \ge C$, with probability at least $1-\delta/2$:

$$\mathrm{lcb}_{t_\ell - 1}(\boldsymbol{x};1) \le f(\boldsymbol{x}) \le \mathrm{ucb}_{t_\ell - 1}(\boldsymbol{x};1), \quad \forall \boldsymbol{x} \in D, t_\ell \ge 1 \tag{112}$$

This follows from Lemmas 3 and 6 (with $3 + B_0 \ln\left(\frac{4(1+\log_2 T)}{\delta}\right)$ in place of C in Lemma 3, and $\lambda = 1$), by setting the corresponding failure probabilities to $\frac{\delta}{4(1+\log_2 T)}$ in both. By a union bound over the two events in the lemmas, followed by a union bound over $\ell \in \{1,\dots,\lceil \log_2 T \rceil\}$, we obtain (112). Once again, (112) remains true when $\mathrm{ucb}^{(\ell)}$ and $\mathrm{lcb}^{(\ell)}$ are replaced by $\overline{\mathrm{ucb}}^{(\ell)}$ and $\overline{\mathrm{lcb}}^{(\ell)}$.

There are at most $\lceil \log_2 T \rceil$ "corruption-tolerant" layers (i.e., layers such that $2^\ell \ge C$), and their regret is analyzed via Lemma 4, but with $3 + B_0 \ln\left(\frac{4(1+\log_2 T)}{\delta}\right)$ in place of C, and the confidence bounds (112) in place of (22). Lemma 4 then implies that the total regret coming from these layers for a total of T rounds is upper bounded according to the following analog of (104):

$$\mathcal{O}\left(\left(\left(B + B_0 \ln(\tfrac{\log T}{\delta}) + \sqrt{\ln(\tfrac{\log T}{\delta})} \right) \sqrt{\gamma_T T} + \gamma_T \sqrt{T} \right) \log T \right), \tag{113}$$

with probability at least $1-\delta/2$.

It remains to characterize the regret coming from the layers that are not corruption-tolerant, i.e., the layers ℓ such that $2^\ell < C$. By the algorithm design (i.e., by the established properties of the sets of potential maximizers) and similarly to Lemma 7, it holds that if a point $\boldsymbol{x} \in D$ becomes suboptimal at time step t according to the confidence bounds of some layer ℓ (i.e., $\boldsymbol{x} \notin M_t^{(\ell)}$), then it is not sampled by any layer $\{1,\dots,\ell\}$ in the subsequent time steps $\{t+1,\dots,T\}$. If we denote the minimum layer that is robust to corruption as

$$\begin{aligned} \ell^* &:= \min\left\{ \ell \in \{1,\dots,\lceil \log T \rceil\} \,:\, 2^\ell \ge C \right\} \qquad (114) \\ &= \lceil \log_2 C \rceil \qquad (\text{if } 1 \le C \le T), \qquad (115) \end{aligned}$$

then we can use this layer to characterize the number of queries t_{ℓ^*} made at ℓ^* before a suboptimal point becomes "eliminated" from this and all the lower layers $\{1,\dots,\ell^*-1\}$. This can be done by using Lemma 8 (where ℓ^* plays the role of the S instance), using the confidence bounds from (112) instead of (21).

We can then repeat the arguments of Theorem 9 (Section G.2; Regret incurred by the F instance) and obtain the regret bounds. First, we characterize how many times layers $1,\dots,\ell^*-1$ are selected in between successive selections of ℓ^*. We can establish that with probability at least $1-\delta/2$, in between any two selections of ℓ^* (or prior to the first such selection), there are at most $N = 2C\log\frac{2T}{\delta}$ selections of layers $\{1,\dots,\ell^*-1\}$ (combined) with probability at least $1-\delta/2$. This is done via the same arguments used in the proof of Theorem 9, and the fact that layer ℓ^* is chosen with probability at least $\min\left\{1, \frac{1}{2C}\right\}$ by the definition of ℓ^*.

By taking the union bound over the previous event and the one in (112), we have that with probability at least $1-\delta$, the regret coming from the points selected by the layers $\{1,\dots,\ell^*-1\}$ is at most given by the following analog of (109):

$$\mathcal{O}\left(\sqrt{C\ln(\tfrac{T}{\delta})}\left(\left(B+B_0\ln(\tfrac{\log T}{\delta})+\sqrt{\ln(\tfrac{\log T}{\delta})}\right)\sqrt{\gamma_T T}+\gamma_T\sqrt{T}\right)+B_0 C\ln(\tfrac{T}{\delta})\right). \tag{116}$$

The following overall regret bound dominates both (113) and (116), and therefore holds for Algorithm 3 with probability at least $1-\delta$:

$$R_T=\mathcal{O}\left((1+C)\ln(\tfrac{T}{\delta})\left(\left(B+B_0\ln(\tfrac{\log T}{\delta})+\sqrt{\ln(\tfrac{\log T}{\delta})}\right)\sqrt{\gamma_T T}+\gamma_T\sqrt{T}\right)\right). \tag{117}$$

This matches the expression given in Theorem 10.

I Discussion on the Parameters λ and α

Recall that our posterior updates are done assuming a sampling noise variance $\lambda>0$ that may differ from the true variance $\sigma^2>0$. In the absence of corruptions, one may be inclined to set $\lambda=\sigma^2$, as was done (for example) in [Srinivas et al., 2010]. However, a problem with this approach in the corrupted setting is that if σ^2 is small, the posterior mean will follow the corrupted samples very closely even though they are unreliable. More generally, increasing λ generally increases robustness against corruptions, but if λ is too high then the model essentially places no trust in any of the sampled points, which prevents effective learning. In our theoretical analysis, we set $\lambda=1$ as a mathematically convenient choice controlling this trade-off, though other values may also work well in practice.

Next, we discuss the parameter $\alpha\geq 1$ in Fast-Slow GP-UCB. The idea is that if we set $\alpha=1$ everywhere, it becomes difficult or impossible to establish that suboptimal points are "ruled out" by the S instance (in the sense of Lemma 7) after a certain amount of time. This is because regardless of the suboptimality of a given point $\boldsymbol{x}$, the posterior variance may be just high enough for its upper confidence bound to be just below the maximal function value $f(\boldsymbol{x}^*)$. Then, $\boldsymbol{x}^*$ will be favored over $\boldsymbol{x}$ according to the UCB rule, and the algorithm may fail to reduce the uncertainty in $f(\boldsymbol{x})$.

In contrast, if we are using the UCB rule with $\alpha=2$ and the preceding "unlucky" scenario is encountered, then upon halving the confidence width (i.e., considering the confidence bounds with $\alpha=1$ instead of $\alpha=2$), such a point $\boldsymbol{x}$ will correctly be ruled out as suboptimal. Lemma 8 formalizes this intuition.

J Optimal Dependence on C and T

We first argue that a linear dependence on the corruption C is unavoidable in any cumulative regret bound. However, we do not make any claims of optimality regarding the *joint* dependence on (C,T).

Let the domain be the unit interval $[0,1]$, and let $f_0(x)$ and $f_1(x)$ be functions taking values in $\big[-1,1\big]$ and satisfying the RKHS norm bound, as well as the following property: Any point within $\frac{1}{2}$ of optimality for one function (e.g., $f_0(x)\geq f_0(x_0^*)-\frac{1}{2}$) is at least $\frac{1}{2}$-far away from optimality for the other function (e.g., $f_1(x)\leq f_1(x_1^*)-\frac{1}{2}$). Such functions can easily be constructed (at least when the RKHS norm B is not too small), for example, via the approach in [Scarlett et al., 2017].

Now suppose that the the true function is known to be either f_0 or f_1, but the exact one of the two is unknown. Consider an adversary that, for the first C rounds, simply perturbs the function value to zero. This can be done within the adversary's budget, since $f(x)\in\big[-1,1\big]$. Given such corruptions, the player cannot learn anything about the function, so at best can randomly guess whether the function is f_0 or f_1. However, by the property of $\frac{1}{2}$-optimality above, attaining $o(C)$ regret for one function implies incurring $\Omega(C)$ regret for the other function.

Hence, regardless of the sampling algorithm, there exist functions in the function class for which $\Omega(C)$ regret is incurred.

As for the dependence on T, we recall from (14) that when C is constant, the dependence on T matches well-known bounds from the non-corrupted setting [Srinivas et al., 2010, Chowdhury and Gopalan, 2017]. Recent lower bounds [Chowdhury and Gopalan, 2017] reveal that this dependence is near-optimal for the SE kernel, though some gaps still remain for the Matérn kernel. Closing these gaps remains a significant challenge even in the non-corrupted setting.

K Comparison to Stochastic Linear Bandits

Regret bounds for corrupted stochastic linear bandits were given in the parallel independent work of Li *et al.* [Li et al., 2019]. While the stochastic linear setting corresponds to our problem setting with a linear kernel, care should be taken in comparing our results to those of [Li et al., 2019], since the results of [Li et al., 2019] are instance-dependent (i.e., depend on certain gaps associated with the underlying function) and ours hold for an arbitrary (e.g., worst-case) instance satisfying the RKHS norm constraint.

For a polytope-shaped domain in any constant dimension, the cumulative regret bound in [Li et al., 2019] is logarithmic in T with a constant of $O\big(\frac{C}{\Delta} + \frac{1}{\Delta^2}\big)$, where Δ is the gap between the best action (necessarily a corner point of the domain) and the second-best corner point. By comparison, for fixed $B > 0$, Theorem 10 yields cumulative regret $\tilde{O}(C\sqrt{T})$, where $\tilde{O}(\cdot)$ hides $\log T$ factors. This is obtained using the fact that $\gamma_T = O(d \log T)$ for the linear kernel [Srinivas et al., 2010, Theorem 5], and the fact that we are focusing on the case $d = O(1)$ in this discussion.

Naturally, the results of [Li et al., 2019] are stronger when the gaps are constant (i.e., $\Delta = \Theta(1)$), attaining $\log T$ regret instead of $\sqrt{T}$. On the other extreme, the "worst-case" gap used to convert instance-dependent guarantees to worst-case guarantees is $\Delta = O\big(\frac{1}{\sqrt{T}}\big)$ [Abbasi-Yadkori et al., 2011], and in this case the bound of [Li et al., 2019] becomes trivial (higher than linear), whereas ours remains sublinear for $C \ll \sqrt{T}$. More generally, our bound is tighter whenever $\Delta \ll \sqrt{C}T^{1/4}$, and the bound of [Li et al., 2019] is tighter whenever $\Delta \gg T^{-1/4}$ and $C \gg 1$.

Overall, however, we believe that the main advantage of our work is the ability to handle general kernels (e.g., SE and Matérn), thereby allowing the underlying function to be highly non-linear.