A Table of Notation

We provide a table of notation so that the proof section can be easily followed.

Symbol	Explanation
m	Variable to denote a base arm
G	Variable to denote a group
S	Variable to denote a super arm
S_t	Variable to denote the chosen super arm at round t
$\mathcal{M}_t$	Set of base arms that are available at round t
M_t	Number of base arms that become available at round t
$\mathcal{G}_t$	Set of feasible groups in round t
$\mathcal{G}_{t,\text{good}}$	Set of groups whose expected rewards are above their thresholds
$\mathcal{S}_t$	Set of feasible super arms in round t
$\mathcal{S}'_t$	Set of super arms whose corresponding groups satisfy their thresholds
$\hat{\mathcal{S}}'_t$	Set of feasible super arms whose corresponding groups' reward indices are above their thresholds
$\mathcal{S}$	Overall feasible set of super arms
$\mathcal{X}$	Context set
$\mathcal{X}_t$	Set of available contexts in round t
$x_{t,m}$	Context associated with base arm m
$\boldsymbol{x}_{t,G}$	Context vector of base arms in G
$\boldsymbol{x}_{t,S}$	Context vector of base arms in S
$\boldsymbol{r}(x)$	Random outcome of base arm with context x
$\boldsymbol{f}(x)$	Expected outcome of base arm with context x
$\boldsymbol{\eta}$	$\mathcal{N}(\mathbf{0}, \sigma^2\boldsymbol{I})$ independent observation noise
$U(S, r_1(\boldsymbol{x}_{t,S}))$	Random super arm reward
$u(S, f_1(\boldsymbol{x}_{t,S}))$	Expected super arm reward function
$V_G(r_2(\boldsymbol{x}_{t,G\cap S_t}))$	Random group reward
$v_G(f_2(\boldsymbol{x}_{t,G\cap S_t}))$	Expected group reward function
$\gamma_{t,G}$	Threshold for group G at round t
ζ	Trade-off parameter
$i_t(x_{t,m})$	Index of base arm m at round t which is given to $\text{Oracle}_{\text{spr}}$ (reward index)
$i'_t(x_{t,m})$	Index of base arm m at round t which is given to $\text{Oracle}_{\text{grp}}$ (satisfying index)
$\overline{\gamma}_T$	Maximum information gain which is associated with the context arrivals $\mathcal{X}_1, \ldots, \mathcal{X}_T$
γ_N	Maximum information gain given N base arm outcome observations
K	Maximum possible number of base arms in a super arm

When we state that $a \geq b$ with a and b being vectors, we mean that every component of a is greater than or equal to the corresponding component of b. This holds for other comparison operators as well. Also, we omit G from v_G and V_G; and S from $U(S, \ldots)$ and $u(S, \ldots)$ when it is obvious from the context. In the definition of $\mathcal{S}'_t$ and $\hat{\mathcal{S}}'_t$, the term corresponding groups means the groups that contain the base arms of the chosen super arm.

B Additional Experimental Results

B.1 Changing Trade-off Parameter (ζ)

In this simulation we showcase the behavior of our algorithm for changing trade-off parameter (ζ) values.

B.1.1 Setup

We use a synthetic setup where we generate the arm and group data needed for the simulation. Similar to the main paper simulations, in each round t, we first sample the number of groups, $|\mathcal{G}_t|$, from a Poisson distribution with mean 50. Then, for each group we generate the contexts of the base arms in that group, where the number of base arms is sampled from a Poisson distribution with mean 5. Each base arm has a two-dimensional (2D) context that is sampled uniformly from $[0, 1]^2$. Then, we sample the expected outcome of each base arm of each group from a GP with zero mean and two squared exponential kernels, each given by

$$k(x, x') = \exp\left(-\frac{1}{2l^2}\|x - x'\|\right),$$

where we set $l = 0.5$. Note that given that we run the simulation for $T = 100$ rounds, there will be an expected number of 25000 arms and to sample a GP function with that many points we will need to compute the Cholesky decomposition of a 625 million element matrix, which would be very resource-heavy. Instead, we first sample 6000 2D contexts from $[0, 1]^2$ and then sample the GP function at those points. Then, during our simulation, we sample each base arm's context x and corresponding expected outcome $\boldsymbol{f}(x)$ from the generated sets. Finally, we set $\boldsymbol{r}(x) = \boldsymbol{f}(x) + \boldsymbol{\eta}$, where $\boldsymbol{\eta} \sim \mathcal{N}(\boldsymbol{0}, 0.1^2 \boldsymbol{I}_2)$. We set the group reward, v, to be the variance of the outcomes in the group and we set the threshold to be the 80% percentile of the group rewards of all groups in all rounds of the simulation. We use a high percentile value to increase the difficulty of the group thresholding, so that minimizing super arm regret does not necessarily yield minimizing group regret. Finally, the super arm reward is the linear sum of the base arms.

B.1.2 Results

We run our algorithm using five different values of ζ, linearly spaced between 0.001 and 0.999. Figure 5 shows the final super arm and group regret of each ζ run (i.e., the cumulative regrets at the end of the simulation). First, notice a trade-off between super arm and group regret. As one increases, the other decreases. This is expected because to minimize group regret, groups with arms whose outcomes are spread out and have high variance must be picked, but to minimize super arm regret, groups with high outcomes must be picked. Second, as ζ increases, the super arm regret increases while the group regret decreases. This is because the variance of the indices given to the first oracle (i'), which determines which groups pass their thresholds, decreases as ζ approaches 1. Thus, the larger ζ is, the stricter the first oracle is. Conversely, the smaller ζ is the laxer the first oracle and the stricter the second oracle, which determines which super arm to play.

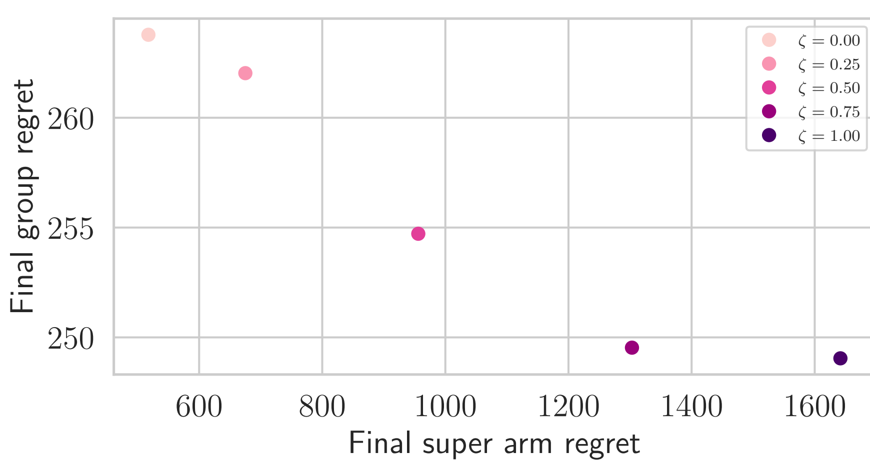

Figure 5: Final super arm and group regret for different trade-off parameter, ζ, values

C Proofs

C.1 Auxiliary Proofs

In this part, we will prove results that will be used in the proofs of the theorems in the paper.

Throughout this section, we take $j \in \{1, 2\}$. We denote by $\boldsymbol{r}_{j\,[\![t-1]\!]}$ the vector of base arm outcome observations made until the beginning of round t, where

$$\boldsymbol{r}_{j\,[\![t-1]\!]} = [r_j^T(\boldsymbol{x}_{1,S_1}), \ldots, r_j^T(\boldsymbol{x}_{t-1,S_{t-1}})]^T .$$

For any $t \geq 1$, the posterior distribution of $f_j(x)$ given the observation vector $\boldsymbol{r}_{j\,[\![t-1]\!]}$ is $\mathcal{N}(\mu_{j\,[\![t-1]\!]}(x), (\sigma_{j\,[\![t-1]\!]}(x))^2)$, for any $x \in \mathcal{X}_t$. In our analysis, we will resort to the following Gaussian tail bound

$$\mathbb{P}\left(|f_j(x) - \mu_{j\,[\![t-1]\!]}(x)| > (\sqrt{\beta_t})\sigma_{j\,[\![t-1]\!]}(x) \Big| \boldsymbol{r}_{j\,[\![t-1]\!]}\right) \leq 2\exp\left(\frac{-\beta_t}{2}\right) \text{ for } \beta_t \geq 0 \ . \tag{3}$$

The following lemma shows that the base arm indices upper bound the expected outcomes with high probability.

Lemma 1. *(Lemma 1 of (Nika et al., 2021)) Fix $\delta \in (0, 1)$, and set $\beta_t := 2\log(M\pi^2 t^2/3\delta)$. Let $\mathcal{F}_j := \{\forall t \geq 1, \forall x \in \mathcal{X}_t : |f_j(x) - \mu_{j\,[\![t-1]\!]}(x)| \leq (\sqrt{\beta_t})\sigma_{j\,[\![t-1]\!]}(x)\}$. We have $\mathbb{P}(\mathcal{F}_j) \geq 1 - \delta$.*

Now, we state that the modified indices upper bound the expected base arm outcomes with high probability under the events $\mathcal{F}_j$.

Lemma 2. *The following arguments hold under $\mathcal{F}_1 \cap \mathcal{F}_2$.*

$$\begin{aligned} i_t(x) &\geq f_1(x), \forall t \geq 1, \ \forall x \in \mathcal{X}_t \\ i'_t(x) &\geq f_2(x), \forall t \geq 1, \ \forall x \in \mathcal{X}_t \ . \end{aligned} \tag{4}$$

Proof. Fix $t \geq 1$ and $x \in \mathcal{X}_t$. Under event $\mathcal{F}_1$, we have the following chain of inequalities

$$f_1(x) - \mu_{1\,[\![t-1]\!]}(x) \leq (\sqrt{\beta_t})\sigma_{1\,[\![t-1]\!]}(x)$$

$$f_1(x) - \mu_{1\,[\![t-1]\!]}(x) \leq \frac{1}{1-\zeta}(\sqrt{\beta_t})\sigma_{1\,[\![t-1]\!]}(x) \tag{5}$$

$$f_1(x) \leq \mu_{1\,[\![t-1]\!]}(x) + \frac{1}{1-\zeta}(\sqrt{\beta_t})\sigma_{1\,[\![t-1]\!]}(x) = i_t(x) \ .$$

Proceeding in the same fashion for event $\mathcal{F}_2$, we obtain

$$f_2(x) - \mu_{2\,[\![t-1]\!]}(x) \leq \frac{1}{\zeta}(\sqrt{\beta_t})\sigma_{2\,[\![t-1]\!]}(x) \tag{6}$$

$$f_2(x) \leq \mu_{2\,[\![t-1]\!]}(x) + \frac{1}{\zeta}(\sqrt{\beta_t})\sigma_{2\,[\![t-1]\!]}(x) = i'_t(x) \ ,$$

where equation 5 and equation 6 follow from the fact that $\frac{1}{1-\zeta} \geq 1$ and $\frac{1}{\zeta} \geq 1$ when $\zeta \in [0, 1]$.

□

Next, we show that the set of feasible super arms whose corresponding groups satisfy their thresholds is a subset of the set of feasible super arms whose groups' reward indices are above their thresholds. We later use this result in Lemma 5.

Lemma 3. *Fix $\delta \in (0, 1)$. The following argument holds under the event $\mathcal{F}_2$ when the group reward function v satisfies the monotonicity assumption given in Assumption 1.*

$$\mathcal{S}'_t \subseteq \hat{\mathcal{S}}'_t, \ \forall t \geq 1.$$

Proof. For any S_t and for all G we have:

$$\begin{aligned} S_t \in \mathcal{S}'_t &\iff v(f_2(\boldsymbol{x}_{t,G\cap S_t})) \geq \gamma_{t,G} && (7)\\ &\iff v(f_2(\boldsymbol{x}_{t,G\cap S_t})) - \gamma_{t,G} \geq 0 \\ &\implies v(i'_t(\boldsymbol{x}_{t,G\cap S_t})) - \gamma_{t,G} \geq 0 && (8)\\ &\iff S_t \in \hat{\mathcal{S}}'_t && (9) \end{aligned}$$

where equation 7 follows from the definition of $\mathcal{S}'_t$, equation 8 follows from the inequality that $v(i'_t(\boldsymbol{x}_{t,G\cap S_t})) \geq v(f_2(\boldsymbol{x}_{t,G\cap S_t}))$. Since $i'_t(\boldsymbol{x}_{t,G\cap S_t}) \geq f_2(\boldsymbol{x}_{t,G\cap S_t})$ under the event $\mathcal{F}_2$ and v is monotone by assumption, this inequality is valid. Lastly, equation 9 follows from the fact that $\text{Oracle}_{\text{grp}}$ is an exact oracle and will return the groups where $v(i'_t(\boldsymbol{x}_{t,G\cap S_t})) > \gamma_{t,G}$. As this reasoning is true for any S_t, we indicate that $\mathcal{S}'_t \subseteq \hat{\mathcal{S}}'_t$.

Detail:

$$\begin{aligned} \mathcal{S}'_t &:= \{S \in \mathcal{S}_t : \forall (G \in \mathcal{G}_t), v_G(f_2(\boldsymbol{x}_{t,G\cap S})) \geq \gamma_{t,G}\}. \\ \hat{\mathcal{S}}'_t &:= \{S \in \mathcal{S}_t : \forall G \in \mathcal{G}_t,\ S \cap G \in \mathcal{G}_{t,\text{good}}\} \\ \mathcal{G}_{t,\text{good}} &:= \{G' \subseteq G \mid G \in \mathcal{G}_t \text{ and } v_G(i'_t(\boldsymbol{x}_{t,G'})) \geq \gamma_{t,G'}\} \\ i'_t(x_{t,m}) &:= \mu_{2[\![t-1]\!]}(x_{t,m}) + \frac{1}{\zeta}(\sqrt{\beta_t})\sigma_{2[\![t-1]\!]}(x_{t,m}). \end{aligned}$$

□

Next, we upper bound the group regret in terms of the posterior variances of base arms. Note that group regret is incurred when a selected group's expected reward is below its threshold whereas its index is above. Therefore, whenever a group G incurs group regret in round t, then $v(f_2(\boldsymbol{x}_{t,G\cap S_t})) < \gamma_{t,G} < v(i'_t(\boldsymbol{x}_{t,G\cap S_t}))$ must happen. This observation plays a key role in the analysis of the next lemma. Moreover, we use the notation $\tilde{x}_{t,k}$ to denote the context of the kth base arm in $G \cap S_t$ at round t for convenience, unless otherwise stated.

Lemma 4. *Fix $t \geq 1$, and consider $G \in \mathcal{G}_t$. The following argument holds under the event $\mathcal{F}_2$:*

$$[\gamma_{t,G} - v(f_2(\boldsymbol{x}_{t,G\cap S_t}))]_+ \leq \left(\frac{\zeta+1}{\zeta}\right) B\sqrt{\beta_t} \sum_{k=1}^{|G\cap S_t|} |\sigma_{2[\![t-1]\!]}(\tilde{x}_{t,k})| \ . \tag{10}$$

Proof. $[\gamma_{t,G} - v(f_2(\boldsymbol{x}_{t,G\cap S_t}))]_+ > 0$ implies that $v(f_2(\boldsymbol{x}_{t,G\cap S_t})) < \gamma_{t,G}$ and $v(i'_t(\boldsymbol{x}_{t,G\cap S_t})) \geq \gamma_{t,G}$. Therefore, whenever G incurs group regret it holds that $v(f_2(\boldsymbol{x}_{t,G\cap S_t})) < \gamma_{t,G} \leq v(i'_t(\boldsymbol{x}_{t,G\cap S_t}))$.

$$\begin{aligned} 0 < \gamma_{t,G} - v(f_2(\boldsymbol{x}_{t,G\cap S_t})) &< v(i'_t(\boldsymbol{x}_{t,G\cap S_t})) - v(f_2(\boldsymbol{x}_{t,G\cap S_t})) \\ 0 < [\gamma_{t,G} - v(f_2(\boldsymbol{x}_{t,G\cap S_t}))]_+ &< v(i'_t(\boldsymbol{x}_{t,G\cap S_t})) - v(f_2(\boldsymbol{x}_{t,G\cap S_t})) \\ &\leq B_{G\cap S_t} \sum_{k=1}^{|G\cap S_t|} |i'_t(\tilde{x}_{t,k}) - f_2(\tilde{x}_{t,k})| && (11)\\ &\leq B_{G\cap S_t} \sum_{k=1}^{|G\cap S_t|} |\mu_{2[\![t-1]\!]}(\tilde{x}_{t,k}) - f_2(\tilde{x}_{t,k})| + B_{G\cap S_t} \sum_{k=1}^{|G\cap S_t|} \left|\frac{1}{\zeta}\sqrt{\beta_t}\sigma_{2[\![t-1]\!]}(\tilde{x}_{t,k})\right| && (12)\\ &\leq \left(\frac{\zeta+1}{\zeta}\right) B(\sqrt{\beta_t}) \sum_{k=1}^{|G\cap S_t|} |\sigma_{2[\![t-1]\!]}(\tilde{x}_{t,k})| \ , && (13) \end{aligned}$$

where equation 11 follows from the Lipschitz continuity of v; equation 12 follows from the definition of index and the triangle inequality; for equation 13 we use Lemma 1 and the definition of B.

□

Next, we upper bound the gap of a selected super arm in round t (aka simple regret) in terms of the posterior variances of base arms. Hereafter, we use the notation $\overline{x}_{t,k}$ to denote the context $x_{t,s_{t,k}}$ of the kth selected base arm $s_{t,k}$ at round t for convenience, unless otherwise stated. Note that $S_t^* \in \arg\max_{S \in \mathcal{S}'_t} u(f_1(\boldsymbol{x}_{t,S}))$ is the optimal super arm in round t.

Lemma 5. *Given round $t \geq 1$, let $S_t^* = \{s_{t,1}^*, \ldots, s_{t,|S_t^*|}^*\}$ denote the optimal super arm in round t. Then, the following argument holds under the event $\mathcal{F}_1$:*

$$\alpha \cdot u(f_1(\boldsymbol{x}_{t,S_t^*})) - u(f_1(\boldsymbol{x}_{t,S_t})) \leq \left(\frac{2-\zeta}{1-\zeta}\right) B' \sqrt{\beta_t} \sum_{k=1}^{|S_t|} |\sigma_{1\,[\![t-1]\!]}(\overline{x}_{t,k})| \tag{14}$$

Proof. We define $H_t = \arg\max_{S \in \hat{\mathcal{S}}'_t} u(i_t(\boldsymbol{x}_{t,S}))$. Given that event $\mathcal{F}_1$ holds, we have:

$$\begin{aligned}
\alpha \cdot u(f_1(\boldsymbol{x}_{t,S_t^*})) - u(f_1(\boldsymbol{x}_{t,S_t})) &\leq \alpha \cdot u(i_t(\boldsymbol{x}_{t,S_t^*})) - u(f_1(\boldsymbol{x}_{t,S_t})) &(15)\\
&\leq \alpha \cdot u(i_t(\boldsymbol{x}_{t,H_t})) - u(f_1(\boldsymbol{x}_{t,S_t})) &(16)\\
&\leq u(i_t(\boldsymbol{x}_{t,S_t})) - u(f_1(\boldsymbol{x}_{t,S_t})) &(17)\\
&\leq B' \sum_{k=1}^{|S_t|} |i_t(\overline{x}_{t,k}) - f_1(\overline{x}_{t,k})| &(18)\\
&\leq B' \sum_{k=1}^{|S_t|} |\mu_{1\,[\![t-1]\!]}(\overline{x}_{t,k}) - f_1(\overline{x}_{t,k})| + B' \sum_{k=1}^{|S_t|} \left| \frac{1}{1-\zeta} (\sqrt{\beta_t}) \sigma_{1\,[\![t-1]\!]}(\overline{x}_{t,k}) \right| &(19)\\
&\leq \left(\frac{2-\zeta}{1-\zeta}\right) B' \sqrt{\beta_t} \sum_{k=1}^{|S_t|} |\sigma_{1\,[\![t-1]\!]}(\overline{x}_{t,k})| \ , &(20)
\end{aligned}$$

where equation 15 follows from monotonicity of u and the fact that $f_1(x_{t,s_{t,k}^*}) \leq i_t(x_{t,s_{t,k}^*})$, for $k \leq |S_t^*|$ (Lemma 2); equation 16 follows from the definition of H_t and the fact that $\mathcal{S}'_t \subseteq \hat{\mathcal{S}}'_t$ on event $\mathcal{F}_2$ (Lemma 3), in other words, $\max_{S \in \hat{\mathcal{S}}'_t} u(i_t(\boldsymbol{x}_{t,S})) \geq \max_{S \in \mathcal{S}'_t} u(i_t(\boldsymbol{x}_{t,S}))$ since $\mathcal{S}'_t \subseteq \hat{\mathcal{S}}'_t$; equation 17 holds since S_t is the super arm chosen by the α-approximation oracle; equation 18 follows from the Lipschitz continuity of u; equation 19 follows from the definition of index and the triangle inequality; for equation 20 we use Lemma 1.

Before proving our theorems, we prove our last lemma which enables us to have our regrets bounds in terms of the information gain. We note that both of our oracles are deterministic and our algorithm doesn't give any extra randomization.

Lemma 6. *(Lemma 3 of (Nika et al., 2021)) Let $\boldsymbol{z}_t := \boldsymbol{x}_{t,S_t}$ be the vector of selected contexts at time $t \geq 1$. Given $T \geq 1$, we have:*

$$I_j\left(r_j(\boldsymbol{z}_{[T]}); f_j(\boldsymbol{z}_{[T]})\right) \geq \frac{1}{2(\sigma^{-2}\lambda^*(K) + 1)} \sum_{t=1}^{T} \sum_{k=1}^{|S_t|} \sigma^{-2} \sigma^2_{j_{[\![t-1]\!]}}(\overline{x}_{t,k}) \ ,$$

where $\boldsymbol{z}_{[T]} = [\boldsymbol{z}_1, \ldots, \boldsymbol{z}_T]^T$ is the vector of all selected contexts until round T and λ^* is the maximum eigenvalue of matrix $(\Sigma_{[\![t-1]\!]}(\boldsymbol{z}_{[T]}))_{t=1}^T$.

C.2 Proof of Theorem 1

From Lemma 4 we have:

$$\begin{aligned}
R_g(T) &= \sum_{t=1}^{T} \sum_{G \in \tilde{\mathcal{G}}_t} [\gamma_{t,G} - v(f_2(\boldsymbol{x}_{t,G\cap S_t}))]_+ \\
&\leq \left(\frac{\zeta+1}{\zeta}\right) B\sqrt{\beta_T} \sum_{t=1}^{T} \sum_{G \in \mathcal{G}_t} \sum_{k=1}^{|G\cap S_t|} |\sigma_{2\,[\![t-1]\!]}(\tilde{x}_{t,k})| \\
&\leq \left(\frac{\zeta+1}{\zeta}\right) B\sqrt{\beta_T} \sum_{t=1}^{T} \sum_{k=1}^{|S_t|} |\sigma_{2\,[\![t-1]\!]}(\overline{x}_{t,k})| \ , \qquad (21)
\end{aligned}$$

using the fact that $\sqrt{\beta_t}$ is monotonically increasing in t. Also, we changed the notation of $\tilde{x}_{t,k}$ with $\overline{x}_{t,k}$ as we are summing through all the base arms in S_t in equation 21. We have:

$$\begin{aligned}
R_g^2(T) &\leq \left(\frac{\zeta+1}{\zeta}\right)^2 B^2\beta_T \left(\sum_{t=1}^{T} \sum_{k=1}^{|S_t|} |\sigma_{2\,[\![t-1]\!]}(\overline{x}_{t,k})|\right)^2 \qquad (22) \\
&\leq \left(\frac{\zeta+1}{\zeta}\right)^2 B^2\beta_T T \sum_{t=1}^{T} \left(\sum_{k=1}^{|S_t|} |\sigma_{2\,[\![t-1]\!]}(\overline{x}_{t,k})|\right)^2 \qquad (23) \\
&\leq \left(\frac{\zeta+1}{\zeta}\right)^2 B^2\beta_T T \sum_{t=1}^{T} |S_t| \sum_{k=1}^{|S_t|} \left(\sigma_{2\,[\![t-1]\!]}(\overline{x}_{t,k})\right)^2 \qquad (24) \\
&\leq \left(\frac{\zeta+1}{\zeta}\right)^2 B^2\beta_T TK \sum_{t=1}^{T} \sum_{k=1}^{|S_t|} \left(\sigma_{2\,[\![t-1]\!]}(\overline{x}_{t,k})\right)^2 \\
&\leq \left(\frac{\zeta+1}{\zeta}\right)^2 B^2\beta_T TK\sigma^2 \sum_{t=1}^{T} \sum_{k=1}^{|S_t|} \sigma^{-2} \left(\sigma_{2\,[\![t-1]\!]}(\overline{x}_{t,k})\right)^2 \qquad (25) \\
&\leq 2(\sigma^{-2}\lambda^*(K)+1) \left(\frac{\zeta+1}{\zeta}\right)^2 B^2\beta_T TK\sigma^2 I\left(r_2(\boldsymbol{z}_{[T]}); f_2(\boldsymbol{z}_{[T]})\right) \qquad (26) \\
&\leq C_1(K)K\beta_T T\overline{\gamma_{2}}_T \ , \qquad (27)
\end{aligned}$$

where for equation 23 and equation 24 we have used the Cauchy-Schwarz inequality twice; in equation 25 we just multiply by σ^2 and σ^{-2}; equation 26 follows from Lemma 6 and for equation 27 we use the definition of $\overline{\gamma_2}_T$. Taking the square root of both sides we obtain our desired result.

From Lemma 5 we have:

$$\begin{aligned}
R_s(T) &= \alpha \sum_{t=1}^{T} \mathrm{opt}(f_t) - \sum_{t=1}^{T} u(f_1(\boldsymbol{x}_{t,S_t})) \\
&\leq \left(\frac{2-\zeta}{1-\zeta}\right) B'\sqrt{\beta_T} \sum_{t=1}^{T} \sum_{k=1}^{|S_t|} |\sigma_{1\,[\![t-1]\!]}(\overline{x}_{t,k})| \ , \qquad (28)
\end{aligned}$$

using the fact that $\sqrt{\beta_t}$ is monotonically increasing in t. We have:

$$R_s^2(T) \leq \left(\frac{2-\zeta}{1-\zeta}\right)^2 (B')^2\beta_T \left(\sum_{t=1}^{T} \sum_{k=1}^{|S_t|} |\sigma_{1\,[\![t-1]\!]}(\overline{x}_{t,k})|\right)^2 \qquad (29)$$

The middle steps are the same as equation 23-equation 26 except that we have $\left(\frac{2-\zeta}{1-\zeta}\right)^2 (B')^2\beta_T$ instead of $\left(\frac{\zeta+1}{\zeta}\right)^2 B^2\beta_T$ as the constant multiplier. Also, we use σ_1 instead of σ_2. Hence, the last steps are modified as:

$$\begin{aligned}R_s^2(T) &\leq 2(\sigma^{-2}\lambda^*(K)+1)\left(\frac{2-\zeta}{1-\zeta}\right)^2 (B')^2\beta_T T K\sigma^2 I\left(r_1(\boldsymbol{z}_{[T]}); f_1(\boldsymbol{z}_{[T]})\right)\\ &\leq C_2(K)K\beta_T T\overline{\gamma_1}_T \ ,\end{aligned}$$

Taking the square root of both sides we obtain our desired result. Finally, in order to obtain the total regret we use:

$$\begin{aligned}R(T) &= \zeta R_g(T) + (1-\zeta)R_s(T)\\ &\leq \left(\zeta\sqrt{C_1(K)} + (1-\zeta)\sqrt{C_2(K)}\right)\sqrt{\beta_T K T\overline{\gamma}_T}\end{aligned}$$

where $\overline{\gamma}_T = \max\{\overline{\gamma_1}_T, \overline{\gamma_2}_T\}$. In order to eliminate the ζ dependence we modify this expression as follows:

$$\begin{aligned}R(T) &\leq \left(\zeta\sqrt{2B^2(\frac{\zeta+1}{\zeta})^2(\lambda^*(K)+\sigma^2)} + (1-\zeta)\sqrt{2(B')^2(\frac{2-\zeta}{1-\zeta})^2(\lambda^*(K)+\sigma^2)}\right)\sqrt{\beta_T K T\overline{\gamma}_T} \quad (30)\\ &= \left(\frac{\zeta}{|\zeta|}\sqrt{2B^2(\zeta+1)^2(\lambda^*(K)+\sigma^2)} + \frac{1-\zeta}{|1-\zeta|}\sqrt{2(B')^2(2-\zeta)^2(\lambda^*(K)+\sigma^2)}\right)\sqrt{\beta_T K T\overline{\gamma}_T}\\ &= \left(\sqrt{2B^2(\zeta+1)^2(\lambda^*(K)+\sigma^2)} + \sqrt{2(B')^2(2-\zeta)^2(\lambda^*(K)+\sigma^2)}\right)\sqrt{\beta_T K T\overline{\gamma}_T} \quad (31)\\ &= \sqrt{2(\lambda^*(K)+\sigma^2)}\left(|B||\zeta+1| + |B'||2-\zeta|\right)\sqrt{\beta_T K T\overline{\gamma}_T}\\ &= \sqrt{2(\lambda^*(K)+\sigma^2)}\left(B(\zeta+1) + B'(2-\zeta)\right)\sqrt{\beta_T K T\overline{\gamma}_T} \quad (32)\\ &\leq \sqrt{2(\lambda^*(K)+\sigma^2)}\left(2B+2B'\right)\sqrt{\beta_T K T\overline{\gamma}_T} \quad (33)\\ &= \sqrt{C(K)\beta_T K T\overline{\gamma}_T} \ .\end{aligned}$$

where equation 30 follows from writing the expressions of C_1 and C_2 to the required places; equation 31 follows from $\zeta \in [0,1]$; equation 32 follows from the assumptions that $B > 0$ and $B' > 0$ and also from $\zeta \in [0,1]$ and equation 33 comes from writing the maximizing ζ values in $\zeta+1$ and $2-\zeta$.

C.3 Proof of Theorem 2

The proof is the same as that of Theorem 2 of (Nika et al., 2021).

C.4 Proof of Theorem 3

We first state an auxiliary fact that will be used in the proof (for a proof see (Williams & Vivarelli, 2000)).

Fact 1. *The predictive variance of a given context is monotonically non-increasing (i.e., given $x \in \mathcal{X}$ and the vector of selected samples $[x_1, \ldots, x_{N+1}]$, we have $\sigma^2_{N+1}(x) \leq \sigma^2_N(x)$, for any $N \geq 1$).*

The proof of Theorem 3 easily follows from Theorem 2 and this fact.

C.5 Proof of Corollary 1

This is a direct application of the explicit bounds on γ_T given in Theorem 5 of (Srinivas et al., 2012) to the bound we obtained in Theorem 3. For the Matérn kernel, we have used the tighter bounds from (Vakili et al., 2020).