Appendix A

Theoretical Results: Proofs

The proofs of a number of theoretical results from Chapter 3 are presented in this Appendix.

A.1 Theorem 1

In order to prove Theorem 1, this section first establishes a series of supporting lemmas.

Lemma 6 shows that a bound on the local information hallucinated during the batch implies a relationship between batch and sequential confidence intervals:

Lemma 6. *Suppose that at some round t, there exists a bound, C, uniform over D, on the conditional mutual information with respect to $f(\boldsymbol{x})$ which will be acquired by the set of actions initiated since the last observation at round $fb[t]$, where this bound is of form*

$$I(f(\boldsymbol{x}); \boldsymbol{y}_{fb[t]+1:t-1} \mid \boldsymbol{y}_{1:fb[t]}) \leq C, \ \forall \boldsymbol{x} \in D, \tag{A.1}$$

for some constant $C > 0$. Choose

$$\beta_t = \exp(2C)\alpha_{fb[t]} \tag{A.2}$$

where Equation (3.6) relates sequential confidence intervals $C^{seq}_{fb[t]+1}(\boldsymbol{x})$ with the parameter $\alpha_{fb[t]+1}$ and Equation (3.8) relates batch confidence intervals $C^{batch}_t(\boldsymbol{x})$ with the parameter β_t. Then, conditional on the event that for all $\boldsymbol{x} \in D$, $f(\boldsymbol{x}) \in C^{seq}_{fb[t]+1}(\boldsymbol{x})$, it holds that $f(\boldsymbol{x}) \in C^{batch}_{t'}(\boldsymbol{x})$ for all $\boldsymbol{x} \in D$ and all t' such that $fb[t] + 1 \leq t' \leq t$.

Proof. Noting that the confidence intervals $C^{\text{seq}}_{\text{fb}[t]+1}(\boldsymbol{x})$ and $C^{\text{batch}}_t(\boldsymbol{x})$ are both centered on $\mu_{\text{fb}[t]}(\boldsymbol{x})$,

$$C^{\text{seq}}_{\text{fb}[t]+1}(\boldsymbol{x}) \subseteq C^{\text{batch}}_t(\boldsymbol{x}) \ \forall \boldsymbol{x} \in D \iff \alpha^{1/2}_{\text{fb}[t]}\sigma_{\text{fb}[t]}(\boldsymbol{x}) \leq \beta^{1/2}_t \sigma_{t-1}(\boldsymbol{x}) \ \forall \boldsymbol{x} \in D.$$

By the definition of the conditional mutual information with respect to $f(\boldsymbol{x})$, and by employing

Equation (A.1), Equation (3.14) follows. Choosing β_t as in Equation (A.2), it follows that

$$\alpha_{\text{fb}[t]}^{1/2}\sigma_{\text{fb}[t]}(\boldsymbol{x}) = \beta_t^{1/2} \exp(-C) \cdot \sigma_{\text{fb}[t]}(\boldsymbol{x}) \leq \beta_t^{1/2}\sigma_{t-1}(\boldsymbol{x}),$$

where the inequality is based on Equation (3.14), thus implying $C_{\text{fb}[t]+1}^{\text{seq}}(\boldsymbol{x}) \subseteq C_t^{\text{batch}}(\boldsymbol{x})\, \forall \boldsymbol{x} \in D$. In turn, if $f(\boldsymbol{x}) \in C_{\text{fb}[t]+1}^{\text{seq}}(\boldsymbol{x})$, then $f(\boldsymbol{x}) \in C_t^{\text{batch}}(\boldsymbol{x})$. Further, since Equation (A.2) relates β_t to $\alpha_{\text{fb}[t]}$, then $\beta_{t'} = \beta_t$ for all $t' \in \{\text{fb}[t]+1, \ldots, t\}$. Since $\sigma_{t'}$ is non-increasing, $C_{t'}^{\text{batch}}(\boldsymbol{x}) \supseteq C_t^{\text{batch}}(\boldsymbol{x})$ for all such t', completing the proof. □

With a bound C on the conditional mutual information gain with respect to $f(\boldsymbol{x})$ for any $\boldsymbol{x} \in D$, as in Equation (A.1), Lemma 6 links the confidence intervals and GP-BUCB decision rule at time t with the GP posterior after observation fb$[t]$. Lemma 7 extends this link to all $t \geq 1$ and all $\boldsymbol{x} \in D$, given a high-probability guarantee of confidence interval correctness at the beginning of all batches. This step is required for the regret bound of Theorem 1.

Lemma 7. *Suppose there exist a constant $C > 0$, a sequence of actions $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{t-1}\}$, and a feedback mapping fb$[t]$ such that for all $\boldsymbol{x} \in D$*

$$C \geq I(f(\boldsymbol{x}); \boldsymbol{y}_{fb[t]+1:t-1} \mid \boldsymbol{y}_{1:fb[t]}),\ \forall t \geq 1.$$

Then, if $\beta_t = \exp(2C)\alpha_{fb[t]}$, $\forall t \geq 1$,

$$P(f(\boldsymbol{x}) \in C_{fb[t]+1}^{seq}(\boldsymbol{x})\, \forall \boldsymbol{x} \in D, \forall t \geq 1) \geq 1 - \delta$$
$$\implies P(f(x) \in C_t^{batch}(\boldsymbol{x})\, \forall \boldsymbol{x} \in D, \forall t \geq 1) \geq 1 - \delta.$$

Proof. If β_t is chosen as specified, then for any t and τ such that $\tau = \text{fb}[t]$ and $f(x) \in C_{\tau+1}^{\text{seq}}(\boldsymbol{x})$, Lemma 6 implies that $f(x) \in C_t^{\text{batch}}(\boldsymbol{x})$. If there exists a set $\mathcal{S} = \{\tau_1, \tau_2, \ldots\}$ such that fb$[t] \in \mathcal{S}$ for all $t \geq 1$, and additionally $f(x) \in C_{\tau+1}^{\text{seq}}(\boldsymbol{x})$ for all $\boldsymbol{x} \in D$ and $\tau \in \mathcal{S}$, then $f(\boldsymbol{x}) \in C_t^{\text{batch}}(\boldsymbol{x})$ for all $\boldsymbol{x} \in D$ and all $t \geq 1$. The event $f(\boldsymbol{x}) \in C_{\text{fb}[t]+1}^{\text{seq}}(\boldsymbol{x})$, $\forall \boldsymbol{x} \in D$, $\forall t \geq 1$ satisfies these conditions directly. The lemma follows because for logical propositions A and B, $[A \implies B] \implies [P(A) \leq P(B)]$ and thus if $P(A) \geq 1 - \delta \implies P(B) \geq 1 - \delta$. □

The high-probabilty confidence intervals are next be related to the instantaneous regret and thence to the cumulative regret.

Lemma 8. *Conditional on the event $f(\boldsymbol{x}) \in C_t^{batch}(\boldsymbol{x})$, $\forall \boldsymbol{x} \in D$, $\forall t \geq 1$, and given that actions $\boldsymbol{x}_t$, $\forall t \geq 1$ are selected using Equation (3.7), it holds that*

$$R_T \leq \sqrt{TC_1\beta_T\gamma_T},$$

where $C_1 = 8/\log(1+\sigma_n^{-2})$, γ_T *is defined in Equation (3.3), and* β_t *is defined in Equation (A.2).*

Proof. Given $f(\boldsymbol{x}) \in C_t^{\text{batch}}(\boldsymbol{x})$, $\forall \boldsymbol{x} \in D$, $\forall t \geq 1$, using the GP-BUCB decision rule, Equation (3.7), Lemma 5.2 in Srinivas et al. (2010), and our assumptions about $C_t^{\text{batch}}(\boldsymbol{x})$, it follows that the instantaneous regret r_t is bounded as $r_t \leq 2\beta_t^{1/2}\sigma_{t-1}(\boldsymbol{x}_t)$, $\forall t \geq 1$. From Lemma 5.3 and 5.4 in Srinivas et al. (2010) it follows that $\sum_{t=1}^{T} r_t^2 \leq C_1 \beta_T \gamma_T$. The claim then follows as a consequence of the Cauchy Schwarz inequality, since $R_T^2 \leq T \sum_{t=1}^{T} r_t^2$. □

Proof of Theorem 1. Taken together, Lemmas 6 through 8, a bound C satisfying Equation (A.1), and a high-probability guarantee that some set of sequential confidence intervals always contain the values of f allow us to construct a batch algorithm with high-probability regret bounds. Srinivas et al. (2010) develop choices of the exploration-exploitation tradeoff parameter α_t such that guarantees of the form $P(f(\boldsymbol{x}) \in C_t^{\text{seq}}(\boldsymbol{x}) \, \forall \boldsymbol{x} \in D, \forall t \geq 1) \geq 1 - \delta$ are realizable for arbitrarily small δ. Employing Lemma 8, as well as Lemmas 5.1 and 5.8 of Srinivas et al. (2010) (for assumptions 1 and 2) and Theorem 6 of Srinivas et al. (2010) (for assumption 3), Theorem 1 follows as an immediate corollary. □

A.2 Theorem 4: Initialization Set Size Bounds

Thorough initialization of GP-BUCB can drive down the constant C, which bounds the information which can be hallucinated in the course of post-initialization batches and also governs the asymptotic scaling of the regret bound with batch size B. First, we connect the information which can be gained in post-initialization batches with the amount of information being gained in the initialization, through Lemma 3, the formal statement of which is in Section 3.3.5, and the proof of which is presented here.

Proof of Lemma 3. Since the initialization procedure is greedy and information gain is submodular (See Section 3.2.3), the information gain from adding the last element of the initialization set, $I(f; \boldsymbol{x}_{T^{\text{init}}}^{\text{init}} \mid D_{T^{\text{init}}-1}^{\text{init}})$, must be the smallest marginal information gain in the initialization process, and thus is no greater than the mean of the marginal gains, i.e,

$$I(f; \boldsymbol{x}_{T^{\text{init}}}^{\text{init}} \mid D_{T^{\text{init}}-1}^{\text{init}}) \leq I(f; D_{T^{\text{init}}}^{\text{init}})/T^{\text{init}}.$$

Further, again because information gain is submodular and the initializaiton set was constructed greedily, no subsequent decision can yield information gain greater than $I(f; \boldsymbol{x}_{T^{\text{init}}}^{\text{init}} \mid D_{T^{\text{init}}-1}^{\text{init}})$, and thus $\gamma_{B-1}^{\text{init}} \leq (B-1) \cdot I(f; \boldsymbol{x}_{T^{\text{init}}}^{\text{init}} \mid D_{T^{\text{init}}-1}^{\text{init}})$. Combining these two inequalities with the definition of $\gamma_{T^{\text{init}}}$ yields the result. □

A.2.1 Initialization Set Size: Linear Kernel

For the linear kernel, there exists a logarithmic bound on the maximum information gain of a set of queries, precisely, $\exists \eta \geq 0 : \gamma_t \leq \eta d \log(t+1)$ (Srinivas et al., 2010). We attempt to initialize GP-BUCB with a set D^{init} of size T^{init}, where, motivated by this bound and the form of Inequality (3.16), we assume T^{init} is of the form

$$T^{\text{init}} = k\eta d(B-1)\log B. \tag{A.3}$$

We must show that there exists a k of finite size for which an initialization set of size T^{init} as in Equation (A.3) implies that any subsequent set $\mathcal{S}$, $|\mathcal{S}| = B - 1$, produces a conditional information gain with respect to $\mathbf{f}$ of no more than C. This requires showing that the inequality $\frac{B-1}{T^{\text{init}}}\gamma_{T^{\text{init}}} \leq C$ holds for this choice of k and thus T^{init}. Since we consider non-trivial batches, i.e., $B-1 \geq 1$, if k is sufficiently large such that $k\eta d(B-1) \geq 1$,

$$\log(\log(B) + 1/(k\eta d(B-1))) \leq \log(\log(B) + 1) \leq \log B.$$

Using Equation (A.3) and the bound for $\gamma_{T^{\text{init}}}$, and following algebraic rearrangement, this inequality implies that if $k\eta d(B-1) \geq 1$,

$$\frac{B-1}{T^{\text{init}}}\gamma_{T^{\text{init}}} \leq C \Longleftarrow \frac{\log k}{k \log B} + \frac{\log \eta + \log d}{k \log B} + \frac{2}{k} \leq C.$$

By noting that the maximum of $\frac{\log k}{k}$ over $k \in (0, \infty)$ is $1/e$ and choosing for convenience $C = 2/e$, we obtain for $k \geq 1/(\eta d(B-1))$:

$$\frac{B-1}{T^{\text{init}}}\gamma_{T^{\text{init}}} \leq \frac{2}{e} \Longleftarrow \frac{1}{e\log B} + \frac{1}{k}\left(\frac{\log \eta + \log d + 2\log B}{\log B}\right) \leq \frac{2}{e},$$

or equivalently, choosing k to satisfy both constraints simultaneously,

$$\frac{B-1}{T^{\text{init}}}\gamma_{T^{\text{init}}} \leq \frac{2}{e} \Longleftarrow k \geq \max\left[\frac{1}{\eta d(B-1)}, \frac{e(\log \eta + \log d + 2\log B)}{2\log(B) - 1}\right].$$

Thus, for a linear kernel and such a k, an initialization set D^{init} of size T^{init}, where $T^{\text{init}} \geq k\eta d(B-1)\log(B)$, ensures that the hallucinated conditional information in any future batch of size B is $\leq \frac{2}{e}$.

A.2.2 Initialization Set Size: Matérn Kernel

For the Matérn kernel, $\gamma_t \leq \nu t^\epsilon$, $\epsilon \in (0,1)$ for some $\nu > 0$ (Srinivas et al., 2010). Hence:

$$\begin{aligned}\frac{(B-1)}{T^{\text{init}}}\gamma_{T^{\text{init}}} \leq C \Longleftarrow & \frac{\nu(B-1)(T^{\text{init}})^\epsilon}{T^{\text{init}}} \leq C \\ \Longleftrightarrow & \ \nu(B-1)(T^{\text{init}})^{\epsilon-1} \leq C \\ \Longleftrightarrow & \ T^{\text{init}} \geq \left(\frac{\nu(B-1)}{C}\right)^{1/(1-\epsilon)}.\end{aligned}$$

Thus, for a Matérn kernel, an initialization set D^{init} of size $T^{\text{init}} \geq \left(\frac{\nu(B-1)}{C}\right)^{1/(1-\epsilon)}$ implies that the conditional information gain of any future batch is $\leq C$. Choosing $C = 1$, we obtain the results presented in the corresponding row of Table 3.1.

A.2.3 Initialization Set Size: Squared Exponential (RBF) Kernel

For the RBF kernel, the information gain is bounded by an expression similar to that of the linear kernel, $\gamma_t \leq \eta(\log{(t+1)})^{d+1}$ (Srinivas et al., 2010). Again, motivated by Inequality (3.16), one reasonable choice for an initialization set size is $T^{\text{init}} = k\eta(B-1)(\log B)^{d+1}$. It is necessary to show that there exists a finite k such that the conditional information gain of any post-initialization batch is $\leq C$. By a similar parallel argument to that for the linear kernel (Appendix A.2.1), and assuming that $B \geq 2$ and $k\eta(B-1) \geq 1$, it follows that

$$\begin{aligned}&\frac{B-1}{T^{\text{init}}}\gamma_{T^{\text{init}}} \leq C \\ &\Longleftarrow \frac{\log k + \log\eta + \log{(B-1)}}{k^{1/(d+1)}(\log B)} \frac{\log{[(\log B)^{d+1}+1]}}{k^{1/(d+1)}(\log B)} \leq C^{1/(d+1)} \\ &\Longleftarrow \frac{\log k}{k^{1/(d+1)}(\log B)} + \frac{\log\eta}{k^{1/(d+1)}(\log B)} + \frac{(d+2)}{k^{1/(d+1)}} \leq C^{1/(d+1)},\end{aligned}$$

where the last implication follows because for $a \geq 0, b \geq 1, (a^b+1) \leq (a+1)^b$.

By noting that the maximum of $k^{-1/(d+1)}\log k$ over $k \in (0,\infty)$ is $(d+1)/e$ and choosing $C = (2(d+1)/e)^{d+1}$, we obtain for $k \geq 1/(\eta(B-1))$:

$$\frac{B-1}{T^{\text{init}}}\gamma_{T^{\text{init}}} \leq \left(\frac{2d+2}{e}\right)^{d+1} \Longleftarrow \frac{d+1}{e\log B} + \frac{1}{k^{1/(d+1)}}\left(\frac{\log\eta + (d+2)\log B}{\log B}\right) \leq \frac{2d+2}{e},$$

or equivalently, incorporating the constraint $k \geq 1/(\eta(B-1))$ explicitly,

$$\frac{B-1}{T^{\text{init}}}\gamma_{T^{\text{init}}} \leq \left(\frac{2d+2}{e}\right)^{d+1} \Longleftarrow k \geq \max\left[\frac{1}{\eta(B-1)}, \left(\frac{e(\log\eta + (d+2)\log B)}{(d+1)(2\log{(B)}-1)}\right)^{d+1}\right].$$

Thus, for a Squared Exponential kernel and such a k, an initialization set D^{init} of size T^{init}, where $T^{\text{init}} \geq k\eta(B-1)(\log(B))^{d+1}$, ensures that the hallucinated conditional information in any future batch of size B is no more than $\left(\frac{2d+2}{e}\right)^{d+1}$.

A.3 GP-AUCB: Finite Batch Size

In the absence of an explicitly specified maximum batch size, it is interesting to consider the scaling of batche sizes produced by GP-AUCB for large T. We are concerned with the case where actions are chosen when much is known about the structure of the reward function: many actions could be selected with little "danger" of choosing poorly, but also little information gain. In such a case, a great deal of regret could be accumulated between observations if the posterior mean fails to correctly order the available actions in D with respect to their reward.

The set of size T which gains the least information with respect to f, conditioned on observations $\boldsymbol{y}(\mathcal{S})$, is one which queries $x_* = \operatorname{argmin}_{x \in D} \sigma^2(x|\boldsymbol{y}(\mathcal{S}))$ T times. These samples gain information $1/2\ \log(1 + T\sigma_n^{-2}\sigma_{\mathcal{S}}^2(x_*))$, where $\sigma^2(x|\boldsymbol{y}_{\mathcal{S}})) = \sigma_{\mathcal{S}}^2(x)$ is the posterior variance, conditioned on the observations $\boldsymbol{y}_{\mathcal{S}}$. Using this observation, if a batch is terminated when a threshold C for hallucinated conditional information with respect to f is exceeded, as in the GP-AUCB algorithm, the maximum possible length of a batch, B_{max}, can be bounded as follows:

$$\begin{aligned} C \geq & 1/2\ \log(1 + (B_{max} - 1)\sigma_n^{-2}\sigma_{\mathbf{S}}^2(x_*)) \\ \implies & \left\lceil \frac{\sigma_n^2}{\sigma_{\mathbf{S}}^2(x_*)}[\exp(2C) - 1] \right\rceil + 1 \geq B_{max}. \end{aligned} \tag{A.4}$$

Thus, if there does not exist any $x \in D$ such that $\sigma^2(x) = 0$, which is the case under the GP model for any finite noise, this upper limit on B_{max} is finite for any finite C and any previous sampling history; the batch sizes of GP-AUCB do not diverge to infinity in a finite number of rounds.

Bounding the rate at which the batch length B_{max} can grow is of interest, however. Consider cases where time is indexed by action number t or by batch number N. In the case of iteration number, by rearrangement of Equation (3.14) and using Inequalities (3.11) and (3.13), we have

$$\sigma_t^2(\boldsymbol{x}) \geq \sigma_0^2(\boldsymbol{x}) \exp\left(-2I(f; \boldsymbol{y}_{1:t-1})\right) \geq \sigma_0^2(\boldsymbol{x}) \exp\left(-2\gamma_{t-1}\right)\ \forall t \in \mathbb{N}.$$

At time t, using this result and Inequality (A.4), the maximum length of the batch which can be constructed under GP-AUCB (or any sampling procedure such that the batch terminates when the information gain threshold C is exceeded) is bounded as

$$B_{max} \leq \left\lceil \frac{\sigma_n^2}{\min_{x \in D}(\sigma_0^2(x))}[\exp(2C) - 1][\exp\left(2\gamma_{t-1}\right)] \right\rceil + 1.$$

This bound is $O(\exp(tC))$, since γ_t is no more than linear in t.

A similar bounding result may be obtained for the Nth batch. After $N-1$ batches, the posterior variance of $f(\boldsymbol{x})$, $\sigma^2_{N-1}(\boldsymbol{x})$, may be bounded as follows, for any $\boldsymbol{x} \in D$, via Equation (3.14) and Inequalities (3.11) and (3.13):

$$\sigma^2_{N-1}(\boldsymbol{x}) \geq \sigma^2_0(\boldsymbol{x}) \exp\left(-2(N-1)C_B\right) \forall t \in \mathbb{N}.$$

Here, C_B is an upper bound on the information which is obtained when the observations corresponding to the batch are made. C_B is greater than C, since the batch terminates only when the information which would be hallucinated in order to select the next action exceeds the threshold C. One useful bound is $C_B \leq C + 1/2 \log\left(1 + \sigma_n^{-2} \max_{\boldsymbol{x} \in D} \sigma^2_0(\boldsymbol{x})\right)$, which follows because the termination condition is checked every round and mutual information is submodular. Using Equation (A.4), the length of the Nth batch is thus bounded as

$$B_{max} \leq \left[\frac{\sigma^2_n}{\min_{x \in D}(\sigma^2_0(x))}[\exp(2C) - 1][\exp\left(2(N-1)C_B\right)]\right] + 1,$$

which is $O(\exp(NC))$, but is bounded for finite batch number.

Appendix B

Tabulated Computational Results

B.1 Tables of Results from Experiments

These experiments are described in detail in Section 3.6. Tables of numerical results are presented here; these include the regret (or elapsed time) with the standard error. Each table presents the results of each data set and algorithm combination tested for a particular experimental setting, averaged over 200 runs. Minimum regret of zero indicates that the optimal set was visited by every run.

Data Set	Algorithm	AR, Query 100	MR, Query 100	AR, Query 200	MR, Query 200
Matérn GP	GP-UCB	0.1268 ± 0.0076	0.0285 ± 0.0075	0.0845 ± 0.0073	0.0243 ± 0.0069
	GP-BUCB	0.1434 ± 0.0040	0.0113 ± 0.0032	0.0855 ± 0.0035	0.0107 ± 0.0032
	SM-UCB	0.1479 ± 0.0055	0.0089 ± 0.0052	0.0849 ± 0.0037	0.0035 ± 0.0011
	SM-MEI	0.1549 ± 0.0048	0.0147 ± 0.0031	0.0937 ± 0.0036	0.0099 ± 0.0026
SE GP	GP-UCB	0.0513 ± 0.0038	0.0054 ± 0.0033	0.0322 ± 0.0033	0.0021 ± 0.0012
	GP-BUCB	0.0577 ± 0.0014	0.0005 ± 0.0002	0.0329 ± 0.0008	0.0003 ± 0.0001
	SM-UCB	0.0612 ± 0.0018	0.0016 ± 0.0011	0.0349 ± 0.0011	0.0004 ± 0.0002
	SM-MEI	0.0593 ± 0.0017	0.0016 ± 0.0007	0.0338 ± 0.0011	0.0006 ± 0.0002
Rosenbrock	GP-UCB	0.0571 ± 0.0005	0.0000 ± 0.0000	0.0353 ± 0.0003	0.0000 ± 0.0000
	GP-BUCB	0.0579 ± 0.0005	0.0000 ± 0.0000	0.0359 ± 0.0003	0.0000 ± 0.0000
	SM-UCB	0.0598 ± 0.0004	0.0000 ± 0.0000	0.0366 ± 0.0003	0.0000 ± 0.0000
	SM-MEI	0.0560 ± 0.0005	0.0000 ± 0.0000	0.0340 ± 0.0003	0.0000 ± 0.0000
Cosines	GP-UCB	0.2109 ± 0.0013	0.0009 ± 0.0002	0.1152 ± 0.0007	0.0001 ± 0.0000
	GP-BUCB	0.2110 ± 0.0013	0.0010 ± 0.0002	0.1158 ± 0.0008	0.0002 ± 0.0001
	SM-UCB	0.2195 ± 0.0012	0.0010 ± 0.0002	0.1213 ± 0.0008	0.0003 ± 0.0001
	SM-MEI	0.2092 ± 0.0013	0.0019 ± 0.0004	0.1173 ± 0.0011	0.0010 ± 0.0003
Vaccine Design	GP-UCB	0.8147 ± 0.0402	0.3465 ± 0.0346	0.6009 ± 0.0354	0.2987 ± 0.0304
	GP-BUCB	0.8605 ± 0.0374	0.2834 ± 0.0291	0.6013 ± 0.0314	0.2326 ± 0.0269
	SM-UCB	0.8149 ± 0.0321	0.1521 ± 0.0212	0.5261 ± 0.0264	0.1446 ± 0.0207
	SM-MEI	0.7750 ± 0.0337	0.1525 ± 0.0214	0.5125 ± 0.0266	0.1066 ± 0.0171
SCI	GP-UCB	0.3099 ± 0.0142	0.1540 ± 0.0129	0.2329 ± 0.0127	0.1345 ± 0.0121
	GP-BUCB	0.2965 ± 0.0102	0.0666 ± 0.0085	0.1920 ± 0.0085	0.0544 ± 0.0076
	SM-UCB	0.3016 ± 0.0092	0.0398 ± 0.0061	0.1813 ± 0.0069	0.0303 ± 0.0052
	SM-MEI	0.3622 ± 0.0085	0.0146 ± 0.0019	0.2280 ± 0.0049	0.0096 ± 0.0008

Table B.1: Average (AR) and Minimum regret (MR) for fixed batch size B = 5.

Data Set	Algorithm	AR, Round 100	MR, Round 100	AR, Round 200	MR, Round 200
Matérn GP	GP-UCB	0.1307 ± 0.0054	0.0167 ± 0.0052	0.0811 ± 0.0050	0.0095 ± 0.0030
	GP-BUCB	0.1601 ± 0.0046	0.0096 ± 0.0039	0.0925 ± 0.0041	0.0079 ± 0.0038
	GP-AUCB	0.1527 ± 0.0052	0.0105 ± 0.0041	0.0898 ± 0.0047	0.0101 ± 0.0041
SE GP	GP-UCB	0.0482 ± 0.0012	0.0013 ± 0.0005	0.0290 ± 0.0010	0.0009 ± 0.0005
	GP-BUCB	0.0656 ± 0.0014	0.0003 ± 0.0001	0.0369 ± 0.0008	0.0002 ± 0.0001
	GP-AUCB	0.0597 ± 0.0015	0.0013 ± 0.0005	0.0343 ± 0.0010	0.0009 ± 0.0005
Rosenbrock	GP-UCB	0.0598 ± 0.0004	0.0000 ± 0.0000	0.0369 ± 0.0003	0.0000 ± 0.0000
	GP-BUCB	0.0601 ± 0.0004	0.0000 ± 0.0000	0.0376 ± 0.0003	0.0000 ± 0.0000
	GP-AUCB	0.0635 ± 0.0005	0.0000 ± 0.0000	0.0382 ± 0.0003	0.0000 ± 0.0000
Cosines	GP-UCB	0.2224 ± 0.0013	0.0013 ± 0.0003	0.1224 ± 0.0008	0.0004 ± 0.0002
	GP-BUCB	0.2199 ± 0.0013	0.0012 ± 0.0003	0.1214 ± 0.0009	0.0001 ± 0.0000
	GP-AUCB	0.2693 ± 0.0013	0.0024 ± 0.0005	0.1352 ± 0.0010	0.0002 ± 0.0000
Vaccine Design	GP-UCB	0.8217 ± 0.0371	0.3058 ± 0.0317	0.5834 ± 0.0316	0.2555 ± 0.0286
	GP-BUCB	0.9650 ± 0.0337	0.2501 ± 0.0279	0.6453 ± 0.0277	0.2100 ± 0.0248
	GP-AUCB	0.9653 ± 0.0355	0.2031 ± 0.0252	0.6153 ± 0.0281	0.1783 ± 0.0243
SCI	GP-UCB	0.3092 ± 0.0131	0.0920 ± 0.0100	0.2108 ± 0.0106	0.0718 ± 0.0091
	GP-BUCB	0.3558 ± 0.0095	0.0339 ± 0.0060	0.2068 ± 0.0067	0.0237 ± 0.0050
	GP-AUCB	0.3155 ± 0.0122	0.0455 ± 0.0072	0.1880 ± 0.0084	0.0317 ± 0.0058

Table B.2: Average (AR) and Minimum regret (MR) for fixed delay length B = 5.

Data Set	Algorithm	AR, Query 100	MR, Query 100	AR, Query 200	MR, Query 200
Matérn GP	GP-BUCB, $B = 5$	0.1405 ± 0.0033	0.0080 ± 0.0024	0.0827 ± 0.0028	0.0076 ± 0.0024
	GP-BUCB, $B = 10$	0.1751 ± 0.0029	0.0068 ± 0.0016	0.0980 ± 0.0020	0.0060 ± 0.0016
	GP-BUCB, $B = 20$	0.2843 ± 0.0047	0.0038 ± 0.0009	0.1513 ± 0.0024	0.0029 ± 0.0008
	SM-UCB, $B = 5$	0.1509 ± 0.0048	0.0117 ± 0.0043	0.0889 ± 0.0045	0.0110 ± 0.0043
	SM-UCB, $B = 10$	0.1891 ± 0.0028	0.0029 ± 0.0009	0.1036 ± 0.0017	0.0025 ± 0.0009
	SM-UCB, $B = 20$	0.3022 ± 0.0051	0.0025 ± 0.0008	0.1597 ± 0.0026	0.0005 ± 0.0002
	SM-MEI, $B = 5$	0.1524 ± 0.0047	0.0141 ± 0.0041	0.0905 ± 0.0040	0.0099 ± 0.0036
	SM-MEI, $B = 10$	0.1897 ± 0.0037	0.0076 ± 0.0025	0.1064 ± 0.0028	0.0068 ± 0.0023
	SM-MEI, $B = 20$	0.2978 ± 0.0047	0.0081 ± 0.0019	0.1609 ± 0.0029	0.0063 ± 0.0015
SE GP	GP-BUCB, $B = 5$	0.0600 ± 0.0014	0.0005 ± 0.0001	0.0344 ± 0.0008	0.0002 ± 0.0001
	GP-BUCB, $B = 10$	0.0937 ± 0.0024	0.0005 ± 0.0001	0.0515 ± 0.0014	0.0004 ± 0.0001
	GP-BUCB, $B = 20$	0.1653 ± 0.0045	0.0010 ± 0.0002	0.0864 ± 0.0023	0.0004 ± 0.0002
	SM-UCB, $B = 5$	0.0607 ± 0.0016	0.0006 ± 0.0002	0.0349 ± 0.0011	0.0004 ± 0.0002
	SM-UCB, $B = 10$	0.0920 ± 0.0024	0.0004 ± 0.0002	0.0501 ± 0.0013	0.0001 ± 0.0000
	SM-UCB, $B = 20$	0.1660 ± 0.0048	0.0006 ± 0.0001	0.0869 ± 0.0024	0.0003 ± 0.0001
	SM-MEI, $B = 5$	0.0606 ± 0.0017	0.0014 ± 0.0004	0.0349 ± 0.0011	0.0011 ± 0.0003
	SM-MEI, $B = 10$	0.0920 ± 0.0025	0.0019 ± 0.0006	0.0501 ± 0.0014	0.0013 ± 0.0005
	SM-MEI, $B = 20$	0.1639 ± 0.0049	0.0013 ± 0.0002	0.0853 ± 0.0024	0.0009 ± 0.0002
Rosenbrock	GP-BUCB, $B = 5$	0.0576 ± 0.0004	0.0000 ± 0.0000	0.0356 ± 0.0003	0.0000 ± 0.0000
	GP-BUCB, $B = 10$	0.0573 ± 0.0004	0.0000 ± 0.0000	0.0353 ± 0.0003	0.0000 ± 0.0000
	GP-BUCB, $B = 20$	0.0771 ± 0.0004	0.0000 ± 0.0000	0.0453 ± 0.0003	0.0000 ± 0.0000
	SM-UCB, $B = 5$	0.0590 ± 0.0005	0.0000 ± 0.0000	0.0368 ± 0.0003	0.0000 ± 0.0000
	SM-UCB, $B = 10$	0.0639 ± 0.0005	0.0000 ± 0.0000	0.0386 ± 0.0003	0.0000 ± 0.0000
	SM-UCB, $B = 20$	0.0828 ± 0.0008	0.0000 ± 0.0000	0.0483 ± 0.0004	0.0000 ± 0.0000
	SM-MEI, $B = 5$	0.0558 ± 0.0005	0.0000 ± 0.0000	0.0340 ± 0.0003	0.0000 ± 0.0000
	SM-MEI, $B = 10$	0.0626 ± 0.0006	0.0000 ± 0.0000	0.0374 ± 0.0004	0.0000 ± 0.0000
	SM-MEI, $B = 20$	0.0806 ± 0.0007	0.0000 ± 0.0000	0.0464 ± 0.0004	0.0000 ± 0.0000
Cosines	GP-BUCB, $B = 5$	0.2107 ± 0.0013	0.0014 ± 0.0003	0.1158 ± 0.0008	0.0003 ± 0.0001
	GP-BUCB, $B = 10$	0.2066 ± 0.0013	0.0009 ± 0.0002	0.1131 ± 0.0009	0.0002 ± 0.0001
	GP-BUCB, $B = 20$	0.2136 ± 0.0017	0.0023 ± 0.0007	0.1186 ± 0.0011	0.0006 ± 0.0004
	SM-UCB, $B = 5$	0.2211 ± 0.0014	0.0012 ± 0.0003	0.1210 ± 0.0009	0.0003 ± 0.0001
	SM-UCB, $B = 10$	0.2330 ± 0.0014	0.0013 ± 0.0004	0.1278 ± 0.0008	0.0003 ± 0.0001
	SM-UCB, $B = 20$	0.2729 ± 0.0015	0.0019 ± 0.0004	0.1505 ± 0.0011	0.0001 ± 0.0000
	SM-MEI, $B = 5$	0.2106 ± 0.0016	0.0033 ± 0.0006	0.1184 ± 0.0012	0.0015 ± 0.0005
	SM-MEI, $B = 10$	0.2253 ± 0.0016	0.0027 ± 0.0005	0.1257 ± 0.0010	0.0011 ± 0.0002
	SM-MEI, $B = 20$	0.2631 ± 0.0016	0.0041 ± 0.0006	0.1454 ± 0.0010	0.0011 ± 0.0003
Vaccine Design	GP-BUCB, $B = 5$	0.9413 ± 0.0406	0.3302 ± 0.0340	0.6615 ± 0.0348	0.2775 ± 0.0293
	GP-BUCB, $B = 10$	1.0379 ± 0.0349	0.1839 ± 0.0278	0.6540 ± 0.0299	0.1711 ± 0.0265
	GP-BUCB, $B = 20$	1.4637 ± 0.0323	0.1024 ± 0.0176	0.8327 ± 0.0247	0.0951 ± 0.0169
	SM-UCB, $B = 5$	0.8531 ± 0.0366	0.1790 ± 0.0245	0.5428 ± 0.0279	0.1444 ± 0.0215
	SM-UCB, $B = 10$	1.0513 ± 0.0275	0.0906 ± 0.0174	0.6170 ± 0.0219	0.0866 ± 0.0168
	SM-UCB, $B = 20$	1.5212 ± 0.0278	0.0349 ± 0.0113	0.8341 ± 0.0190	0.0349 ± 0.0113
	SM-MEI, $B = 5$	0.8239 ± 0.0325	0.1667 ± 0.0229	0.5418 ± 0.0256	0.1383 ± 0.0214
	SM-MEI, $B = 10$	1.0751 ± 0.0330	0.1202 ± 0.0231	0.6557 ± 0.0249	0.0801 ± 0.0158
	SM-MEI, $B = 20$	1.5440 ± 0.0270	0.0277 ± 0.0098	0.8590 ± 0.0195	0.0271 ± 0.0098
SCI	GP-BUCB, $B = 5$	0.2748 ± 0.0103	0.0492 ± 0.0076	0.1728 ± 0.0082	0.0433 ± 0.0071
	GP-BUCB, $B = 10$	0.3884 ± 0.0091	0.0440 ± 0.0065	0.2275 ± 0.0069	0.0391 ± 0.0060
	GP-BUCB, $B = 20$	0.6031 ± 0.0093	0.0427 ± 0.0063	0.3349 ± 0.0070	0.0298 ± 0.0052
	SM-UCB, $B = 5$	0.3075 ± 0.0094	0.0445 ± 0.0066	0.1894 ± 0.0072	0.0392 ± 0.0063
	SM-UCB, $B = 10$	0.4162 ± 0.0078	0.0190 ± 0.0030	0.2290 ± 0.0049	0.0152 ± 0.0025
	SM-UCB, $B = 20$	0.6608 ± 0.0120	0.0236 ± 0.0045	0.3571 ± 0.0072	0.0213 ± 0.0043
	SM-MEI, $B = 5$	0.3734 ± 0.0089	0.0170 ± 0.0026	0.2379 ± 0.0050	0.0109 ± 0.0013
	SM-MEI, $B = 10$	0.4838 ± 0.0078	0.0132 ± 0.0022	0.2981 ± 0.0048	0.0087 ± 0.0015
	SM-MEI, $B = 20$	0.7177 ± 0.0099	0.0124 ± 0.0022	0.4086 ± 0.0060	0.0072 ± 0.0013

Table B.3: Average (AR) and Minimum regret (MR) for batch sizes $B = 5$, 10, and 20, non-adaptive algorithms.

Data Set	Algorithm	AR, Query 100	MR, Query 100	AR, Query 200	MR, Query 200
Matérn GP	GP-AUCB, $B_{max} = 5$	0.1303 ± 0.0057	0.0182 ± 0.0053	0.0819 ± 0.0053	0.0166 ± 0.0052
	GP-AUCB, $B_{max} = 10$	0.1293 ± 0.0060	0.0185 ± 0.0058	0.0816 ± 0.0057	0.0138 ± 0.0042
	GP-AUCB, $B_{max} = 20$	0.1326 ± 0.0061	0.0197 ± 0.0059	0.0835 ± 0.0059	0.0187 ± 0.0059
	GP-AUCB Local, $B_{max} = 5$	0.1290 ± 0.0042	0.0112 ± 0.0036	0.0774 ± 0.0038	0.0108 ± 0.0036
	GP-AUCB Local, $B_{max} = 10$	0.1667 ± 0.0071	0.0238 ± 0.0069	0.1020 ± 0.0068	0.0190 ± 0.0057
	GP-AUCB Local, $B_{max} = 20$	0.1952 ± 0.0081	0.0173 ± 0.0055	0.1148 ± 0.0062	0.0153 ± 0.0055
	HBBO UCB, $B_{max} = 5$	0.1455 ± 0.0070	0.0179 ± 0.0064	0.0895 ± 0.0065	0.0163 ± 0.0063
	HBBO UCB, $B_{max} = 10$	0.1459 ± 0.0047	0.0136 ± 0.0038	0.0873 ± 0.0040	0.0121 ± 0.0037
	HBBO UCB, $B_{max} = 20$	0.1587 ± 0.0055	0.0150 ± 0.0044	0.0950 ± 0.0047	0.0137 ± 0.0044
	HBBO MEI, $B_{max} = 5$	0.1515 ± 0.0059	0.0197 ± 0.0051	0.0935 ± 0.0052	0.0169 ± 0.0050
	HBBO MEI, $B_{max} = 10$	0.1644 ± 0.0063	0.0224 ± 0.0053	0.1017 ± 0.0054	0.0162 ± 0.0047
	HBBO MEI, $B_{max} = 20$	0.1771 ± 0.0069	0.0131 ± 0.0040	0.1049 ± 0.0050	0.0111 ± 0.0040
SE GP	GP-AUCB, $B_{max} = 5$	0.0489 ± 0.0015	0.0005 ± 0.0002	0.0286 ± 0.0009	0.0004 ± 0.0002
	GP-AUCB, $B_{max} = 10$	0.0505 ± 0.0016	0.0017 ± 0.0006	0.0296 ± 0.0011	0.0010 ± 0.0003
	GP-AUCB, $B_{max} = 20$	0.0590 ± 0.0041	0.0061 ± 0.0035	0.0362 ± 0.0035	0.0026 ± 0.0012
	GP-AUCB Local, $B_{max} = 5$	0.0540 ± 0.0035	0.0044 ± 0.0033	0.0323 ± 0.0026	0.0007 ± 0.0005
	GP-AUCB Local, $B_{max} = 10$	0.0591 ± 0.0021	0.0022 ± 0.0012	0.0340 ± 0.0015	0.0020 ± 0.0012
	GP-AUCB Local, $B_{max} = 20$	0.0683 ± 0.0027	0.0012 ± 0.0005	0.0382 ± 0.0015	0.0008 ± 0.0005
	HBBO UCB, $B_{max} = 5$	0.0547 ± 0.0020	0.0040 ± 0.0016	0.0331 ± 0.0017	0.0021 ± 0.0013
	HBBO UCB, $B_{max} = 10$	0.0554 ± 0.0017	0.0010 ± 0.0004	0.0326 ± 0.0011	0.0003 ± 0.0001
	HBBO UCB, $B_{max} = 20$	0.0610 ± 0.0023	0.0015 ± 0.0005	0.0343 ± 0.0013	0.0006 ± 0.0003
	HBBO MEI, $B_{max} = 5$	0.0533 ± 0.0022	0.0017 ± 0.0006	0.0315 ± 0.0014	0.0013 ± 0.0005
	HBBO MEI, $B_{max} = 10$	0.0601 ± 0.0023	0.0021 ± 0.0006	0.0346 ± 0.0014	0.0014 ± 0.0005
	HBBO MEI, $B_{max} = 20$	0.0640 ± 0.0036	0.0032 ± 0.0017	0.0361 ± 0.0021	0.0010 ± 0.0002
Rosenbrock	GP-AUCB, $B_{max} = 5$	0.0572 ± 0.0005	0.0000 ± 0.0000	0.0353 ± 0.0003	0.0000 ± 0.0000
	GP-AUCB, $B_{max} = 10$	0.0580 ± 0.0005	0.0000 ± 0.0000	0.0359 ± 0.0004	0.0000 ± 0.0000
	GP-AUCB, $B_{max} = 20$	0.0577 ± 0.0005	0.0000 ± 0.0000	0.0359 ± 0.0003	0.0000 ± 0.0000
	GP-AUCB Local, $B_{max} = 5$	0.0574 ± 0.0005	0.0000 ± 0.0000	0.0356 ± 0.0003	0.0000 ± 0.0000
	GP-AUCB Local, $B_{max} = 10$	0.0579 ± 0.0005	0.0000 ± 0.0000	0.0360 ± 0.0003	0.0000 ± 0.0000
	GP-AUCB Local, $B_{max} = 20$	0.0602 ± 0.0006	0.0000 ± 0.0000	0.0368 ± 0.0004	0.0000 ± 0.0000
	HBBO UCB, $B_{max} = 5$	0.0579 ± 0.0005	0.0000 ± 0.0000	0.0362 ± 0.0003	0.0000 ± 0.0000
	HBBO UCB, $B_{max} = 10$	0.0578 ± 0.0005	0.0000 ± 0.0000	0.0356 ± 0.0003	0.0000 ± 0.0000
	HBBO UCB, $B_{max} = 20$	0.0580 ± 0.0006	0.0000 ± 0.0000	0.0362 ± 0.0004	0.0000 ± 0.0000
	HBBO MEI, $B_{max} = 5$	0.0540 ± 0.0005	0.0000 ± 0.0000	0.0332 ± 0.0003	0.0000 ± 0.0000
	HBBO MEI, $B_{max} = 10$	0.0545 ± 0.0005	0.0000 ± 0.0000	0.0334 ± 0.0004	0.0000 ± 0.0000
	HBBO MEI, $B_{max} = 20$	0.0550 ± 0.0005	0.0000 ± 0.0000	0.0343 ± 0.0004	0.0000 ± 0.0000
Cosines	GP-AUCB, $B_{max} = 5$	0.2168 ± 0.0012	0.0009 ± 0.0002	0.1191 ± 0.0009	0.0002 ± 0.0001
	GP-AUCB, $B_{max} = 10$	0.2183 ± 0.0015	0.0014 ± 0.0003	0.1182 ± 0.0009	0.0002 ± 0.0001
	GP-AUCB, $B_{max} = 20$	0.2156 ± 0.0014	0.0020 ± 0.0004	0.1186 ± 0.0010	0.0005 ± 0.0001
	GP-AUCB Local, $B_{max} = 5$	0.2118 ± 0.0014	0.0009 ± 0.0002	0.1162 ± 0.0008	0.0002 ± 0.0000
	GP-AUCB Local, $B_{max} = 10$	0.2108 ± 0.0015	0.0016 ± 0.0004	0.1160 ± 0.0011	0.0005 ± 0.0002
	GP-AUCB Local, $B_{max} = 20$	0.2187 ± 0.0015	0.0016 ± 0.0003	0.1218 ± 0.0011	0.0005 ± 0.0002
	HBBO UCB, $B_{max} = 5$	0.2122 ± 0.0012	0.0010 ± 0.0002	0.1160 ± 0.0009	0.0002 ± 0.0001
	HBBO UCB, $B_{max} = 10$	0.2129 ± 0.0013	0.0009 ± 0.0003	0.1155 ± 0.0008	0.0001 ± 0.0000
	HBBO UCB, $B_{max} = 20$	0.2168 ± 0.0017	0.0019 ± 0.0004	0.1204 ± 0.0012	0.0004 ± 0.0001
	HBBO MEI, $B_{max} = 5$	0.2018 ± 0.0014	0.0031 ± 0.0008	0.1126 ± 0.0010	0.0017 ± 0.0006
	HBBO MEI, $B_{max} = 10$	0.2031 ± 0.0014	0.0030 ± 0.0006	0.1138 ± 0.0010	0.0008 ± 0.0002
	HBBO MEI, $B_{max} = 20$	0.2074 ± 0.0016	0.0033 ± 0.0005	0.1177 ± 0.0012	0.0009 ± 0.0002
Vaccine Design	GP-AUCB, $B_{max} = 5$	0.8811 ± 0.0451	0.3048 ± 0.0341	0.6217 ± 0.0373	0.2714 ± 0.0316
	GP-AUCB, $B_{max} = 10$	0.8410 ± 0.0402	0.3016 ± 0.0325	0.6048 ± 0.0345	0.2710 ± 0.0305
	GP-AUCB, $B_{max} = 20$	0.8455 ± 0.0387	0.2963 ± 0.0315	0.6136 ± 0.0353	0.2725 ± 0.0301
	GP-AUCB Local, $B_{max} = 5$	0.9044 ± 0.0367	0.2729 ± 0.0293	0.6174 ± 0.0314	0.2580 ± 0.0287
	GP-AUCB Local, $B_{max} = 10$	0.9564 ± 0.0375	0.2197 ± 0.0265	0.6202 ± 0.0301	0.1911 ± 0.0248
	GP-AUCB Local, $B_{max} = 20$	1.0741 ± 0.0367	0.1562 ± 0.0228	0.6656 ± 0.0277	0.1436 ± 0.0211
	HBBO UCB, $B_{max} = 5$	0.8472 ± 0.0413	0.3141 ± 0.0326	0.6021 ± 0.0346	0.2767 ± 0.0315
	HBBO UCB, $B_{max} = 10$	0.8628 ± 0.0381	0.3456 ± 0.0324	0.6343 ± 0.0326	0.3153 ± 0.0305
	HBBO UCB, $B_{max} = 20$	0.8606 ± 0.0420	0.3188 ± 0.0344	0.6154 ± 0.0368	0.2862 ± 0.0327
	HBBO MEI, $B_{max} = 5$	0.8574 ± 0.0403	0.3030 ± 0.0310	0.6020 ± 0.0316	0.2134 ± 0.0248
	HBBO MEI, $B_{max} = 10$	0.8712 ± 0.0378	0.3135 ± 0.0324	0.6299 ± 0.0325	0.2366 ± 0.0290
	HBBO MEI, $B_{max} = 20$	0.8675 ± 0.0370	0.2934 ± 0.0316	0.6105 ± 0.0308	0.2168 ± 0.0275
SCI	GP-AUCB, $B_{max} = 5$	0.3152 ± 0.0132	0.1444 ± 0.0121	0.2339 ± 0.0120	0.1287 ± 0.0115
	GP-AUCB, $B_{max} = 10$	0.3242 ± 0.0138	0.1252 ± 0.0110	0.2315 ± 0.0118	0.1124 ± 0.0107
	GP-AUCB, $B_{max} = 20$	0.3226 ± 0.0125	0.1403 ± 0.0111	0.2354 ± 0.0112	0.1326 ± 0.0111
	GP-AUCB Local, $B_{max} = 5$	0.3114 ± 0.0134	0.0961 ± 0.0106	0.2128 ± 0.0113	0.0810 ± 0.0100
	GP-AUCB Local, $B_{max} = 10$	0.3084 ± 0.0128	0.0727 ± 0.0088	0.1962 ± 0.0099	0.0606 ± 0.0080
	GP-AUCB Local, $B_{max} = 20$	0.3766 ± 0.0128	0.0739 ± 0.0089	0.2306 ± 0.0095	0.0627 ± 0.0081
	HBBO UCB, $B_{max} = 5$	0.2751 ± 0.0100	0.0557 ± 0.0081	0.1745 ± 0.0087	0.0491 ± 0.0078
	HBBO UCB, $B_{max} = 10$	0.3060 ± 0.0093	0.0516 ± 0.0078	0.1862 ± 0.0081	0.0448 ± 0.0073
	HBBO UCB, $B_{max} = 20$	0.3186 ± 0.0106	0.0613 ± 0.0090	0.1966 ± 0.0094	0.0559 ± 0.0087
	HBBO MEI, $B_{max} = 5$	0.3206 ± 0.0099	0.0292 ± 0.0055	0.1989 ± 0.0058	0.0104 ± 0.0008
	HBBO MEI, $B_{max} = 10$	0.3362 ± 0.0086	0.0201 ± 0.0038	0.2080 ± 0.0053	0.0087 ± 0.0008
	HBBO MEI, $B_{max} = 20$	0.3527 ± 0.0087	0.0274 ± 0.0049	0.2221 ± 0.0057	0.0093 ± 0.0008

Table B.4: Average (AR) and Minimum regret (MR) for maximum adaptive batch sizes $B_{max} = 5$, 10, and 20.

Data Set	Algorithm	AR, Round 100	MR, Round 100	AR, Round 200	MR, Round 200
Matérn GP	GP-BUCB, $B = 5$	0.1530 ± 0.0029	0.0037 ± 0.0013	0.0857 ± 0.0020	0.0033 ± 0.0013
	GP-BUCB, $B = 10$	0.2089 ± 0.0032	0.0036 ± 0.0014	0.1138 ± 0.0019	0.0033 ± 0.0014
	GP-BUCB, $B = 20$	0.3314 ± 0.0053	0.0033 ± 0.0012	0.1758 ± 0.0028	0.0022 ± 0.0012
	GP-AUCB, $B_{max} = 5$	0.1501 ± 0.0056	0.0130 ± 0.0052	0.0883 ± 0.0053	0.0120 ± 0.0052
	GP-AUCB, $B_{max} = 10$	0.1742 ± 0.0038	0.0062 ± 0.0018	0.0904 ± 0.0022	0.0029 ± 0.0013
	GP-AUCB, $B_{max} = 20$	0.3144 ± 0.0095	0.0217 ± 0.0040	0.1220 ± 0.0042	0.0087 ± 0.0029
	GP-AUCB Local, $B_{max} = 5$	0.1578 ± 0.0028	0.0057 ± 0.0017	0.0891 ± 0.0020	0.0050 ± 0.0016
	GP-AUCB Local, $B_{max} = 10$	0.2089 ± 0.0035	0.0022 ± 0.0007	0.1138 ± 0.0021	0.0014 ± 0.0007
	GP-AUCB Local, $B_{max} = 20$	0.3287 ± 0.0052	0.0017 ± 0.0007	0.1746 ± 0.0027	0.0013 ± 0.0007
SE GP	GP-BUCB, $B = 5$	0.0663 ± 0.0015	0.0007 ± 0.0002	0.0369 ± 0.0008	0.0004 ± 0.0002
	GP-BUCB, $B = 10$	0.1027 ± 0.0024	0.0005 ± 0.0002	0.0553 ± 0.0012	0.0004 ± 0.0002
	GP-BUCB, $B = 20$	0.1784 ± 0.0047	0.0008 ± 0.0005	0.0931 ± 0.0024	0.0002 ± 0.0001
	GP-AUCB, $B_{max} = 5$	0.0591 ± 0.0015	0.0015 ± 0.0009	0.0338 ± 0.0011	0.0013 ± 0.0009
	GP-AUCB, $B_{max} = 10$	0.0673 ± 0.0027	0.0009 ± 0.0004	0.0361 ± 0.0015	0.0003 ± 0.0001
	GP-AUCB, $B_{max} = 20$	0.0885 ± 0.0025	0.0024 ± 0.0012	0.0422 ± 0.0013	0.0010 ± 0.0007
	GP-AUCB Local, $B_{max} = 5$	0.0683 ± 0.0015	0.0011 ± 0.0003	0.0387 ± 0.0010	0.0006 ± 0.0003
	GP-AUCB Local, $B_{max} = 10$	0.0941 ± 0.0022	0.0004 ± 0.0001	0.0506 ± 0.0011	0.0002 ± 0.0001
	GP-AUCB Local, $B_{max} = 20$	0.1074 ± 0.0028	0.0004 ± 0.0001	0.0549 ± 0.0014	0.0001 ± 0.0001
Rosenbrock	GP-BUCB, $B = 5$	0.0594 ± 0.0005	0.0000 ± 0.0000	0.0371 ± 0.0003	0.0000 ± 0.0000
	GP-BUCB, $B = 10$	0.0602 ± 0.0005	0.0000 ± 0.0000	0.0375 ± 0.0003	0.0000 ± 0.0000
	GP-BUCB, $B = 20$	0.0794 ± 0.0005	0.0000 ± 0.0000	0.0468 ± 0.0003	0.0000 ± 0.0000
	GP-AUCB, $B_{max} = 5$	0.0638 ± 0.0005	0.0000 ± 0.0000	0.0379 ± 0.0003	0.0000 ± 0.0000
	GP-AUCB, $B_{max} = 10$	0.0809 ± 0.0007	0.0000 ± 0.0000	0.0423 ± 0.0003	0.0000 ± 0.0000
	GP-AUCB, $B_{max} = 20$	0.2001 ± 0.0019	0.0004 ± 0.0001	0.0570 ± 0.0005	0.0000 ± 0.0000
	GP-AUCB Local, $B_{max} = 5$	0.0598 ± 0.0004	0.0000 ± 0.0000	0.0369 ± 0.0003	0.0000 ± 0.0000
	GP-AUCB Local, $B_{max} = 10$	0.0602 ± 0.0004	0.0000 ± 0.0000	0.0373 ± 0.0003	0.0000 ± 0.0000
	GP-AUCB Local, $B_{max} = 20$	0.0806 ± 0.0004	0.0000 ± 0.0000	0.0474 ± 0.0003	0.0000 ± 0.0000
Cosines	GP-BUCB, $B = 5$	0.2199 ± 0.0012	0.0010 ± 0.0003	0.1216 ± 0.0008	0.0002 ± 0.0001
	GP-BUCB, $B = 10$	0.2265 ± 0.0015	0.0019 ± 0.0005	0.1255 ± 0.0010	0.0003 ± 0.0001
	GP-BUCB, $B = 20$	0.2401 ± 0.0015	0.0030 ± 0.0005	0.1358 ± 0.0012	0.0003 ± 0.0001
	GP-AUCB, $B_{max} = 5$	0.2719 ± 0.0014	0.0023 ± 0.0004	0.1356 ± 0.0010	0.0003 ± 0.0001
	GP-AUCB, $B_{max} = 10$	0.3930 ± 0.0007	0.0696 ± 0.0032	0.2034 ± 0.0014	0.0006 ± 0.0002
	GP-AUCB, $B_{max} = 20$	0.4205 ± 0.0015	0.1032 ± 0.0037	0.3933 ± 0.0008	0.0726 ± 0.0032
	GP-AUCB Local, $B_{max} = 5$	0.2220 ± 0.0014	0.0013 ± 0.0004	0.1217 ± 0.0010	0.0003 ± 0.0001
	GP-AUCB Local, $B_{max} = 10$	0.2251 ± 0.0015	0.0020 ± 0.0007	0.1247 ± 0.0010	0.0004 ± 0.0002
	GP-AUCB Local, $B_{max} = 20$	0.2383 ± 0.0014	0.0020 ± 0.0004	0.1340 ± 0.0010	0.0002 ± 0.0001
Vaccine Design	GP-BUCB, $B = 5$	0.9260 ± 0.0380	0.1995 ± 0.0235	0.5953 ± 0.0289	0.1796 ± 0.0230
	GP-BUCB, $B = 10$	1.2659 ± 0.0345	0.1252 ± 0.0215	0.7446 ± 0.0266	0.1200 ± 0.0214
	GP-BUCB, $B = 20$	1.8475 ± 0.0307	0.0490 ± 0.0124	1.0281 ± 0.0208	0.0345 ± 0.0097
	GP-AUCB, $B_{max} = 5$	0.9702 ± 0.0394	0.2391 ± 0.0292	0.6358 ± 0.0325	0.2149 ± 0.0280
	GP-AUCB, $B_{max} = 10$	1.1655 ± 0.0446	0.2131 ± 0.0288	0.6466 ± 0.0293	0.1515 ± 0.0222
	GP-AUCB, $B_{max} = 20$	2.1901 ± 0.0789	0.6204 ± 0.0627	1.0715 ± 0.0521	0.1958 ± 0.0258
	GP-AUCB Local, $B_{max} = 5$	1.0020 ± 0.0342	0.2335 ± 0.0287	0.6645 ± 0.0301	0.2074 ± 0.0280
	GP-AUCB Local, $B_{max} = 10$	1.1721 ± 0.0278	0.1076 ± 0.0163	0.6812 ± 0.0201	0.0825 ± 0.0138
	GP-AUCB Local, $B_{max} = 20$	1.8291 ± 0.0306	0.0226 ± 0.0059	1.0094 ± 0.0201	0.0189 ± 0.0050
SCI	GP-BUCB, $B = 5$	0.3614 ± 0.0100	0.0386 ± 0.0065	0.2140 ± 0.0071	0.0201 ± 0.0040
	GP-BUCB, $B = 10$	0.5019 ± 0.0086	0.0200 ± 0.0037	0.2757 ± 0.0052	0.0094 ± 0.0008
	GP-BUCB, $B = 20$	0.7114 ± 0.0075	0.0045 ± 0.0013	0.3775 ± 0.0041	0.0033 ± 0.0006
	GP-AUCB, $B_{max} = 5$	0.3641 ± 0.0136	0.0648 ± 0.0091	0.2203 ± 0.0100	0.0455 ± 0.0076
	GP-AUCB, $B_{max} = 10$	0.4735 ± 0.0215	0.0747 ± 0.0093	0.2548 ± 0.0114	0.0434 ± 0.0073
	GP-AUCB, $B_{max} = 20$	0.7793 ± 0.0353	0.1353 ± 0.0173	0.3831 ± 0.0197	0.0543 ± 0.0073
	GP-AUCB Local, $B_{max} = 5$	0.3701 ± 0.0109	0.0434 ± 0.0069	0.2192 ± 0.0076	0.0235 ± 0.0049
	GP-AUCB Local, $B_{max} = 10$	0.4893 ± 0.0083	0.0199 ± 0.0040	0.2674 ± 0.0050	0.0103 ± 0.0021
	GP-AUCB Local, $B_{max} = 20$	0.7197 ± 0.0070	0.0032 ± 0.0010	0.3849 ± 0.0042	0.0023 ± 0.0005

Table B.5: Average (AR) and Minimum regret (MR) for delay lengths $B = 5$, 10, and 20.

Data Set	Algorithm	Query 40	Query 100	Query 200
Matern GP	GP-UCB	0.5992 ± 0.0010	1.9532 ± 0.0037	6.5840 ± 0.0040
	GP-UCB Lazy	0.1764 ± 0.0026	0.2357 ± 0.0033	0.3824 ± 0.0044
	GP-BUCB	0.5947 ± 0.0006	1.9363 ± 0.0027	6.5302 ± 0.0053
	GP-BUCB Lazy	0.2957 ± 0.0018	0.3481 ± 0.0026	0.4592 ± 0.0035
	SM-UCB	2.9618 ± 0.0078	10.8404 ± 0.0123	44.5631 ± 0.0137
	SM-UCB Lazy	6.3990 ± 0.0186	16.0585 ± 0.0370	37.0588 ± 0.0703
	SM-MEI	3.0716 ± 0.0010	11.1101 ± 0.0035	45.1232 ± 0.0072
	SM-MEI Lazy	15.8325 ± 0.0273	38.2692 ± 0.0358	80.2965 ± 0.0446
	HBBO UCB	0.5594 ± 0.0016	1.8075 ± 0.0044	6.2063 ± 0.0064
	HBBO MEI	0.5658 ± 0.0007	1.8250 ± 0.0008	6.2486 ± 0.0009
	GP-AUCB	0.6699 ± 0.0001	1.8984 ± 0.0003	6.2417 ± 0.0020
	GP-AUCB Lazy	0.3527 ± 0.0021	0.4040 ± 0.0029	0.5125 ± 0.0039
	GP-AUCB Local	0.4708 ± 0.0008	1.5885 ± 0.0030	5.6950 ± 0.0071
	GP-AUCB Lazy Local	0.2823 ± 0.0022	0.3338 ± 0.0029	0.4420 ± 0.0040
SE GP	GP-UCB	0.5993 ± 0.0001	1.9520 ± 0.0004	6.5182 ± 0.0011
	GP-UCB Lazy	0.2891 ± 0.0061	0.4265 ± 0.0109	0.6406 ± 0.0170
	GP-BUCB	0.6011 ± 0.0001	1.9462 ± 0.0003	6.5005 ± 0.0008
	GP-BUCB Lazy	0.3703 ± 0.0053	0.4982 ± 0.0100	0.7039 ± 0.0165
	SM-UCB	2.9105 ± 0.0009	10.6819 ± 0.0028	44.0896 ± 0.0062
	SM-UCB Lazy	7.3562 ± 0.0410	17.7201 ± 0.0773	39.3720 ± 0.1229
	SM-MEI	3.0133 ± 0.0010	10.9425 ± 0.0028	44.6246 ± 0.0056
	SM-MEI Lazy	17.6639 ± 0.0782	41.0165 ± 0.1256	83.6221 ± 0.1674
	HBBO UCB	0.5549 ± 0.0006	1.7936 ± 0.0008	6.1503 ± 0.0013
	HBBO MEI	0.5630 ± 0.0006	1.8161 ± 0.0007	6.1973 ± 0.0011
	GP-AUCB	0.6749 ± 0.0001	1.8999 ± 0.0003	6.2147 ± 0.0007
	GP-AUCB Lazy	0.4348 ± 0.0057	0.5588 ± 0.0103	0.7555 ± 0.0162
	GP-AUCB Local	0.4684 ± 0.0001	1.5820 ± 0.0003	5.6564 ± 0.0010
	GP-AUCB Lazy Local	0.3206 ± 0.0050	0.4442 ± 0.0095	0.6404 ± 0.0155
Rosenbrock	GP-UCB	0.5916 ± 0.0030	1.9470 ± 0.0091	6.4193 ± 0.0172
	GP-UCB Lazy	0.3356 ± 0.0013	0.4521 ± 0.0020	0.6437 ± 0.0034
	GP-BUCB	0.5841 ± 0.0009	1.9074 ± 0.0032	6.3437 ± 0.0071
	GP-BUCB Lazy	0.3787 ± 0.0012	0.4900 ± 0.0020	0.6732 ± 0.0034
	SM-UCB	2.8155 ± 0.0010	10.4979 ± 0.0023	43.1365 ± 0.0113
	SM-UCB Lazy	6.4663 ± 0.0075	15.9830 ± 0.0109	36.1740 ± 0.0158
	SM-MEI	2.9311 ± 0.0077	10.7848 ± 0.0181	43.6449 ± 0.0199
	SM-MEI Lazy	15.5865 ± 0.0180	36.8226 ± 0.0204	76.0677 ± 0.0253
	HBBO UCB	0.6033 ± 0.0012	1.8202 ± 0.0017	6.0460 ± 0.0019
	HBBO MEI	0.6076 ± 0.0010	1.8366 ± 0.0011	6.0840 ± 0.0014
	GP-AUCB	0.6510 ± 0.0001	1.8495 ± 0.0004	6.0354 ± 0.0008
	GP-AUCB Lazy	0.4372 ± 0.0010	0.5463 ± 0.0018	0.7317 ± 0.0031
	GP-AUCB Local	0.4582 ± 0.0001	1.5449 ± 0.0002	5.4983 ± 0.0039
	GP-AUCB Lazy Local	0.3569 ± 0.0011	0.4609 ± 0.0016	0.6349 ± 0.0029
Cosines	GP-UCB	0.5829 ± 0.0001	1.9113 ± 0.0004	6.3865 ± 0.0013
	GP-UCB Lazy	0.2750 ± 0.0007	0.3654 ± 0.0010	0.4913 ± 0.0015
	GP-BUCB	0.5810 ± 0.0001	1.9094 ± 0.0004	6.3497 ± 0.0011
	GP-BUCB Lazy	0.3452 ± 0.0009	0.4267 ± 0.0012	0.5242 ± 0.0013
	SM-UCB	2.8365 ± 0.0014	10.6211 ± 0.0046	43.3563 ± 0.0134
	SM-UCB Lazy	5.9717 ± 0.0038	15.0289 ± 0.0059	34.3330 ± 0.0123
	SM-MEI	2.9458 ± 0.0010	10.9126 ± 0.0049	43.9570 ± 0.0134
	SM-MEI Lazy	14.8944 ± 0.0084	35.7111 ± 0.0094	74.2792 ± 0.0243
	HBBO UCB	0.6536 ± 0.0012	1.8741 ± 0.0013	6.1179 ± 0.0015
	HBBO MEI	0.6654 ± 0.0014	1.9006 ± 0.0015	6.1691 ± 0.0017
	GP-AUCB	0.6493 ± 0.0002	1.8570 ± 0.0003	6.0636 ± 0.0008
	GP-AUCB Lazy	0.3962 ± 0.0006	0.4772 ± 0.0010	0.5727 ± 0.0011
	GP-AUCB Local	0.4602 ± 0.0001	1.5595 ± 0.0004	5.5394 ± 0.0008
	GP-AUCB Lazy Local	0.4940 ± 0.0009	0.5946 ± 0.0022	0.6863 ± 0.0023
Vaccine Design	GP-UCB	1.8238 ± 0.0005	6.0469 ± 0.0019	20.3408 ± 0.0047
	GP-UCB Lazy	0.6347 ± 0.0094	0.7021 ± 0.0094	0.9267 ± 0.0094
	GP-BUCB	1.8252 ± 0.0004	5.9950 ± 0.0016	20.1145 ± 0.0039
	GP-BUCB Lazy	1.1121 ± 0.0024	1.1472 ± 0.0024	1.2584 ± 0.0024
	SM-UCB	8.3346 ± 0.0032	32.0995 ± 0.0167	134.3169 ± 0.0458
	SM-UCB Lazy	22.5270 ± 0.2438	46.4903 ± 0.7422	99.7192 ± 1.6901
	SM-MEI	8.5207 ± 0.0023	32.6054 ± 0.0135	135.4947 ± 0.0391
	SM-MEI Lazy	49.0936 ± 0.2575	115.2695 ± 0.7228	243.0290 ± 1.2733
	HBBO UCB	2.3149 ± 0.0128	6.2659 ± 0.0234	19.8061 ± 0.0377
	HBBO MEI	2.2665 ± 0.0121	6.2584 ± 0.0211	19.9313 ± 0.0376
	GP-AUCB	2.2982 ± 0.0003	6.1147 ± 0.0014	19.5122 ± 0.0041
	GP-AUCB Lazy	1.2533 ± 0.0049	1.2877 ± 0.0050	1.3968 ± 0.0049
	GP-AUCB Local	1.4302 ± 0.0003	4.8850 ± 0.0013	17.4855 ± 0.0039
	GP-AUCB Lazy Local	1.0196 ± 0.0083	1.0676 ± 0.0090	1.1843 ± 0.0095
Spinal Cord Therapy	GP-UCB	0.0721 ± 0.0001	0.2404 ± 0.0006	0.8161 ± 0.0021
	GP-UCB Lazy	0.0138 ± 0.0001	0.0323 ± 0.0002	0.0917 ± 0.0003
	GP-BUCB	0.0721 ± 0.0000	0.2395 ± 0.0003	0.8134 ± 0.0014
	GP-BUCB Lazy	0.0236 ± 0.0001	0.0421 ± 0.0001	0.1042 ± 0.0004
	SM-UCB	0.5920 ± 0.0005	1.9495 ± 0.0020	6.7668 ± 0.0024
	SM-UCB Lazy	0.7619 ± 0.0009	2.0412 ± 0.0039	4.8634 ± 0.0116
	SM-MEI	0.6547 ± 0.0002	2.1107 ± 0.0019	7.0953 ± 0.0037
	SM-MEI Lazy	1.8187 ± 0.0018	4.4822 ± 0.0096	9.3589 ± 0.0268
	HBBO UCB	0.0702 ± 0.0003	0.2333 ± 0.0008	0.8007 ± 0.0019
	HBBO MEI	0.0759 ± 0.0003	0.2482 ± 0.0009	0.8317 ± 0.0021
	GP-AUCB	0.0742 ± 0.0001	0.2268 ± 0.0004	0.7722 ± 0.0019
	GP-AUCB Lazy	0.0266 ± 0.0001	0.0442 ± 0.0002	0.1050 ± 0.0004
	GP-AUCB Local	0.0585 ± 0.0003	0.1996 ± 0.0007	0.7152 ± 0.0018
	GP-AUCB Lazy Local	0.0342 ± 0.0004	0.0533 ± 0.0006	0.1186 ± 0.0012

Table B.6: Mean wall-clock execution times and standard deviations of estimate (S).

Appendix C

Action-matched Animal Plots

During two of the animal experiment runs (animal 5, run 1 and animal 7) a substantial number of actions performed by the human experimenter were missed or dropped. An alternate view of these experiments is presented here, where "pass" actions are inserted for those missed by either the algorithm or the human experimenter. Without inserted passes, the same action indices for the human and algorithm do not correspond to the same point in time, and visual interpretation of the regret plots is difficult; after inserted passes, this synchrony is restored.

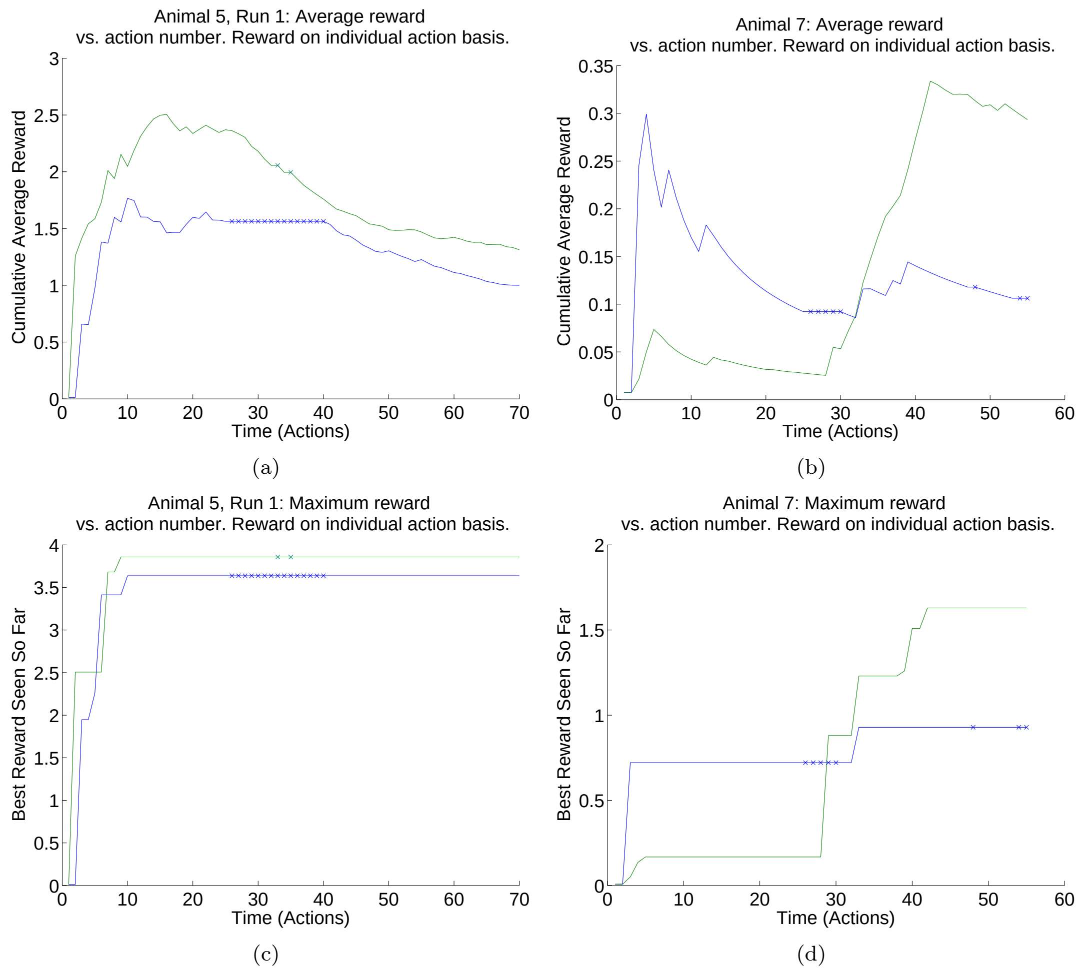

Figure C.1: Action-matched plots for animal 5, run 1 and animal 7. A missed action is treated as a "pass" to restore action synchrony between the human experimenter and the algorithm. These actions are shown with an "x" in the plots above. C.1(a) & (c): Animal 5, run 1. During this experimental run, the human experimenter missed a full day of experiments (P15). Compare these plots to Figures 4.8(b) and (d), which do not have the passes corresponding to the three missed batches on P15. (b) & (d): Animal 7. The human experimenter did not conduct a fourth and final batch on the second testing day (P12). Several actions were also missing from the third testing day, P13. These plots are action compensated versions of Figures 4.11(b) and (d).

Appendix D

Toward Human Studies: Mathematical Results

D.1 Decision-making with an Aggregated Objective

When trying to use several, possibly related GP models $f_1, \dots, f_n$ to make a decision about a known function r of those individual GPs, it is natural to attempt to apply a UCB-like approach to the problem. Unfortunately, unless r is a linear combination of these individual GPs, $r(\mathbf{f})$ is not itself a Gaussian process, nor is the posterior over $r(\mathbf{f}(\boldsymbol{x}))$ Gaussian. This problem even arises if one common and natural formulation of a reward function, that of penalized deviation from a target $\mathbf{t}$ via a weighted norm term, is used, e.g.,

$$r(\mathbf{f}(\boldsymbol{x})) = -\sqrt{(\mathbf{f}(\boldsymbol{x}) - \mathbf{t})^T W (\mathbf{f}(\boldsymbol{x}) - \mathbf{t})}, \tag{D.1}$$

where W is a symmetric, positive definite penalty matrix, such that r is -1 times a weighted 2-norm in $\mathbb{R}^n$. Such an objective function makes a great deal of sense in terms of convex optimization, and has a unique global maximum at $\mathbf{f}(\boldsymbol{x}) = \mathbf{t}$. Further, for any $\boldsymbol{x}$, the posterior over $\mathbf{f}(\boldsymbol{x})$ is $\mathbf{f}(\boldsymbol{x}) \sim \mathcal{N}(\mu(\boldsymbol{x}), \Sigma(\boldsymbol{x}))$, and as $\mu - \mathbf{t}$ becomes very large, the distribution of the weighted squared norm begins to look like the corresponding marginalization of the posterior onto the unit vector in the direction $\mu - \mathbf{t}$; this marginal distribution is a Gaussian. However, of the most interest in terms of active learning is the region near the optimum, where such an $f(\mathbf{f})$ is most strongly non-Gaussian.

Inspired by GP-UCB and GP-BUCB, it seems reasonable that it would be desirable to create a decision function of form

$$\boldsymbol{x}_t = \operatorname*{argmax}_{\boldsymbol{x} \in D} \left[\mathbb{E}[r(\mathbf{f}(\boldsymbol{x}))|\boldsymbol{y}_{1:\mathrm{fb}[t]}] + \beta_t^{1/2} \sqrt{\mathbf{Var}\left(r(\mathbf{f}(\boldsymbol{x}))|\boldsymbol{y}_{1:\mathrm{fb}[t]} \right)} \right], \tag{D.2}$$

which once again trades off exploitation, captured by the mean reward term on the left, with exploration, captured by the standard deviation term on the right. It is possible to calculate

$\mathbb{E}[r(\mathbf{f}(\boldsymbol{x}))^2|\boldsymbol{y}_{1:\text{fb}[t]}] = \mathbb{E}[(\mathbf{f}(\boldsymbol{x})-\mathbf{t})^T W(\mathbf{f}(\boldsymbol{x})-\mathbf{t})|\boldsymbol{y}_{1:\text{fb}[t]}] = (\mu_{\text{fb}[t]}(\boldsymbol{x})-\mathbf{t})^T W(\mu_{\text{fb}[t]}(\boldsymbol{x})-\mathbf{t})+\text{trace}(W\Sigma_{\text{fb}[t]})$. This leaves calculating the expected reward, $\mathbb{E}[r(\mathbf{f}(\boldsymbol{x}))|\boldsymbol{y}_{1:\text{fb}[t]}]$. Unfortunately, despite its relation to the χ distribution, I was unable to obtain a general expression for this expectation, and so I made recourse to bounding arguments. By use of Jensen's inequality, which states that for a convex function $h(\boldsymbol{x})$,

$$\mathbb{E}[h(\boldsymbol{x})] \geq h(\mathbb{E}[\boldsymbol{x}]), \tag{D.3}$$

it is possible to derive an upper bound

$$\mathbb{E}[r(\mathbf{f}(\boldsymbol{x}))|\boldsymbol{y}_{1:\text{fb}[t]}] \leq -\sqrt{(\mu_{\text{fb}[t]}(\boldsymbol{x})-\mathbf{t})^T W(\mu_{\text{fb}[t]}(\boldsymbol{x})-\mathbf{t})}$$

via the concavity of r with respect to $\mathbf{f}$, as well as a lower bound

$$\mathbb{E}[r(\mathbf{f}(\boldsymbol{x}))|\boldsymbol{y}_{1:\text{fb}[t]}] \geq \sqrt{(\mu_{\text{fb}[t]}(\boldsymbol{x})-\mathbf{t})^T W(\mu_{\text{fb}[t]}(\boldsymbol{x})-\mathbf{t}) + \text{trace}(W\Sigma_{t-1})}$$

by noting that $-\sqrt{r}$ is convex with respect to r over $r \in \mathbb{R}^+$. Using the definition of the variance in terms of the expectation of the square and the square of the expectation, and substituting in the upper bound on $\mathbb{E}[r(\mathbf{f}(\boldsymbol{x}))|\boldsymbol{y}_{1:\text{fb}[t]}]$, it can be shown that

$$\mathbf{Var}\left(r(\mathbf{f}(\boldsymbol{x}))|\boldsymbol{y}_{1:\text{fb}[t]}\right) = \mathbb{E}[r(\mathbf{f}(\boldsymbol{x}))^2|\boldsymbol{y}_{1:\text{fb}[t]}] - \mathbb{E}[r(\mathbf{f}(\boldsymbol{x}))|\boldsymbol{y}_{1:\text{fb}[t]}]^2 \leq \text{trace}(W\Sigma_{t-1}). \tag{D.4}$$

By analogy to the GP-UCB and GP-BUCB decision rules (Equations 3.5 and 3.7), and using the upper bounds above to create a term capturing the reward and another term capturing the uncertainty, this suggests a decision rule of form

$$\boldsymbol{x}_t = \operatorname*{argmax}_{\boldsymbol{x}\in D}\left[-\sqrt{(\mu_{\text{fb}[t]}(\boldsymbol{x})-\mathbf{t})^T W(\mu_{\text{fb}[t]}(\boldsymbol{x})-\mathbf{t})} + \beta_t^{1/2}\sqrt{\text{trace}(W\Sigma_{t-1}(\boldsymbol{x}))}\right] \tag{D.5}$$

should have useful characteristics. Note that for the scalar case, with a very large target t, the decision rule reduces to that of GP-BUCB. Further, as we learn more and more about the function near the optima, $\Sigma_{t-1}(\boldsymbol{x})$ should decrease for these decisions, making the upper bound on $\mathbb{E}[r(\mathbf{f}(\boldsymbol{x}))|\boldsymbol{y}_{1:\text{fb}[t]}]$ tighter. Additionally, for $\mu_{t-1}(\boldsymbol{x}) - \mathbf{t}$ large, the decision rule can also be expected to closely bound the actual form in Equation D.2, allowing us to disregard these decisions as poor-performing. This leaves the poor cases as those in which $\Sigma_{t-1}(\boldsymbol{x})$ is very large and $\mu_{t-1}(\boldsymbol{x}) - \mathbf{t}$ is small; in this case, overestimating either the mean or variance should result in allocation of observations to these actions, driving down $\Sigma_{t-1}(\boldsymbol{x})$ and resolving the issue through observation. This decision rule is also practical because it is very easy to calculate; if the posterior over the values $\mathbf{f}(\boldsymbol{x})$ is available, this decision rule simply requires some linear algebraic calculations. Further, it can be shown that, under the assumption that observations can only be added to the observation set, but none can leave, the term $\sqrt{\text{trace}(W\Sigma_{t-1}(\boldsymbol{x}))}$ is non-increasing (see Appendix D.2). Because this term is non-increasing as observations are added to the observation set, the calculation of $\sqrt{\text{trace}(W\Sigma_{t-1}(\boldsymbol{x}))}$ can be done lazily, as can be done for the standard deviation in Section 3.5, enabling substantial computational savings.

D.2 Proof Multi-Muscle Uncertainty Term is Non-Increasing

A quite useful characteristic of the GP-BUCB decision rule, Equation (3.7), is that the uncertainty term (i.e., the standard deviation) cannot increase as observations are added. For computational reasons, it is important to demonstrate that Equation (D.5) also has the same property. As a first step, we define the matrix root of the weight matrix W as a symmetric, positive definite matrix $W^{1/2} = (W^{1/2})^T > 0$, such that $(W^{1/2})^2 = W$. Such a matrix can be constructed by noting that $W = VDV^T$, where $V \in \mathbb{R}^{n\times n}$ is a matrix whose columns are the set of orthonormal right eigenvectors of W and D is the diagonal matrix of the corresponding eigenvalues of W; choosing $W^{1/2} = VD^{1/2}V^T$ produces the desired properties. We then consider the uncertainty of the algorithm's estimate of $\mathbf{f}(\boldsymbol{x})$ after steps t and t' of the algorithm, where $t' > t$, the corresponding observation sets $\mathbf{y}_t$ and $\mathbf{y}_{t'}$, and the sets of past actions $\boldsymbol{X}_t$ and $\boldsymbol{X}_{t'}$. We may describe the posterior covariance function between stimuli $\boldsymbol{x}$ and $\boldsymbol{x}'$ and muscle indices i and j as $k_t((\boldsymbol{x}, i), (\boldsymbol{x}', j)) = k((\boldsymbol{x}, i), (\boldsymbol{x}', j))|\mathbf{y}_t$ and write the posterior covariance matrix at time t' for $\mathbf{f}(\boldsymbol{x})$ as

$$\Sigma_{t'}(\boldsymbol{x}) = \Sigma_t(\boldsymbol{x}) - \mathbf{k}_t(K + \sigma_n^2 I)^{-1}\mathbf{k}_t^T,$$

where $\boldsymbol{X}_{t+1;t'} = \boldsymbol{X}t' \setminus \boldsymbol{X}t$ is the set of observations occurring between times t and t', $\mathbf{k}_t = \mathbf{k}_t(\boldsymbol{x}, \boldsymbol{X}_{t+1:t'}) \in \mathbb{R}^{n\times[(t'-t)\times n]}$ is the covariance between the observations associated with $\boldsymbol{X}_{t+1;t'}$, including all n channels, and $K_t = K_t(\boldsymbol{X}_{\mathbf{t+1:t'}}, \boldsymbol{X}_{\mathbf{t+1:t'}}) \in \mathbb{R}^{[(t'-t)\times n]\times[(t'-t)\times n]}$ is the posterior covariance at time t between the noisy observations $\mathbf{y}_{t+1:t'}$ of $\mathbf{f}(\boldsymbol{X}_{t+1:t'})$ in $\boldsymbol{X}_{t+1:t'}$. Multiplying left and right by $W^{1/2}$, and then using the linearity of the trace and its invariance to circular permutations of symmetric matrices, we obtain

$$\text{trace}(W\Sigma_{t'}(\boldsymbol{x})) = \text{trace}(W\Sigma_t(\boldsymbol{x})) - \text{trace}(W^{1/2}\mathbf{k}_t(K_t + \sigma_n^2 I)^{-1}\mathbf{k}_t^T W^{1/2}). \quad \text{(D.6)}$$

Noting that $\mathbf{h} = W^{1/2}\mathbf{k}_t\mathbf{y}_{t+1:t'} \in \mathbb{R}^{n\times 1}$ is a linear combination of the $n(t'-t)$ multivariate Gaussian observations, $\mathbf{h}$ also has a multivariate Gaussian distribution, such that its covariance matrix, $W^{1/2}\mathbf{k}_t(K_t + \sigma_n^2 I)^{-1}\mathbf{k}_t^T W^{1/2}$, is positive semi-definite. Since the trace of this matrix is therefore non-negative, it follows from Equation (D.6) that $\text{trace}(W\Sigma_{t'}(\boldsymbol{x})) \leq \text{trace}(W\Sigma_t(\boldsymbol{x}))$ for $t' > t$.

D.3 Path-Based Decision Rules

As discussed in Section 5.4.2.3, it may be desirable to plan for smooth paths of length no more than B which travel from the present stimulus state through the decision set. This set of possible paths may be denoted L. While there are potentially exponentially many paths through the graph of possible stimuli, if there is a set of restrictions on path construction such that each end-point (i.e., $\boldsymbol{x} \in D$) may be reached by at most one path, these restrictions imply $|L| \leq |D|$. One reasonable idea for selecting paths from L is to extend the decision rule used by the GP-BUCB algorithm, resulting in the following equation:

$$\boldsymbol{X}_t = \underset{\boldsymbol{X} \in L}{\operatorname{argmax}} \left[\sum_{\tau=t}^{t+B-1} \left(\mu_{\mathrm{fb}[t]}(\boldsymbol{x}_\tau) + \beta_{\mathrm{fb}[t]}^{1/2} \sigma_{\tau-1}(\boldsymbol{x}_\tau) \right) \right], \tag{D.7}$$

where $\boldsymbol{X} = \{\boldsymbol{x}_t, \ldots \boldsymbol{x}_{t+B-1}\}$. This construction follows the form of the GP-BUCB decision rule, and might even be amenable to the same confidence interval analysis, at least locally. However, this decision rule may philosophically differ from the GP-BUCB rule in that the quantity which corresponds to information gained no longer maps easily to the actual information gain $I(f; \boldsymbol{y}(\boldsymbol{X}) | \boldsymbol{y}(\boldsymbol{X}_{\mathrm{fb}[t]}))$. Motivated by the transformation between $\sigma_{t-1}(\boldsymbol{x}_t)$ and $I(f; \boldsymbol{y}(\boldsymbol{x}_t) \mid \boldsymbol{y}(\boldsymbol{X}_{\mathrm{fb}[t]}))$, i.e.,

$$\sigma_{t-1}(\boldsymbol{x}_t) = \sigma_n \sqrt{-1 + \exp(I(f; \boldsymbol{y}(\boldsymbol{x}_t) \mid \boldsymbol{y}(\boldsymbol{X}_{\mathrm{fb}[t]})))}, \tag{D.8}$$

it may be reasonable to apply the same transformation to $I(f; \boldsymbol{y}(\boldsymbol{X}) | \boldsymbol{y}(\boldsymbol{X}_{\mathrm{fb}[t]}))$ to obtain a quantity $e(\boldsymbol{X})$ which corresponds to the information gain from the group of observations as follows:

$$\begin{aligned} e(\boldsymbol{X}) &= \sigma_n \sqrt{-1 + \exp(I(f; \boldsymbol{y}(\boldsymbol{X}) | \boldsymbol{y}(\boldsymbol{X}_{\mathrm{fb}[t]})))} \\ &= \sigma_n \sqrt{-1 + \prod_{\tau=t}^{t+B-1} (1 + \sigma_n^{-2} \sigma_{\tau-1}(\boldsymbol{x}_\tau))} \end{aligned} \tag{D.9}$$

where $\boldsymbol{X}$ is again $\boldsymbol{X} = \{\boldsymbol{x}_t, \ldots \boldsymbol{x}_{t+B-1}\}$. This yields a decision rule of the form

$$\boldsymbol{X}_t = \underset{\boldsymbol{X} \in L}{\operatorname{argmax}} \left[\sum_{\tau=t}^{t+B-1} [\mu_{\mathrm{fb}[t]}(\boldsymbol{x}_\tau)] + \beta_{\mathrm{fb}[t]}^{1/2} e(\boldsymbol{X}) \right]. \tag{D.10}$$

In either or both of these cases, it might be appropriate to consider the possibility that the experiment might have to be stopped during the traversal of $\boldsymbol{X}$ with some probability. Letting the uniform probability of failure of each transition be $1 - \lambda$, and assuming the reward and observation are obtained even if the individual action is a failure, the discounted version of the first decision rule is

$$\boldsymbol{X}_t = \operatorname*{argmax}_{\boldsymbol{X} \in L} \left[\sum_{\tau=t}^{t+B-1} [\lambda^{\tau-t} (\mu_{\mathrm{fb}[t]}(\boldsymbol{x}_\tau) + \beta_{\mathrm{fb}[t]}^{1/2} \sigma_{\tau-1}(\boldsymbol{x}_\tau))] \right]. \tag{D.11}$$

This decision rule has been implemented for a version of the human experimental code which is intended to search over the space of voltage and frequency parameters corresponding to a fixed set of active electrodes. Similarly, the discounted version of Equation (D.9), designated $e_\lambda(\boldsymbol{X})$, is

$$e_\lambda(\boldsymbol{X}) = \sigma_n \sqrt{-1 + \prod_{\tau=t}^{t+B-1} [(1 + \sigma_n^{-2} \sigma_{\tau-1}(\boldsymbol{x}_\tau))^{\lambda^{\tau-t}}]} \tag{D.12}$$

and the corresponding decision rule becomes

$$\boldsymbol{X}_t = \operatorname*{argmax}_{\boldsymbol{X} \in L} \left[\sum_{\tau=t}^{t+B-1} [\lambda^{\tau-t} \mu_{\mathrm{fb}[t]}(\boldsymbol{x}_\tau)] + \beta_{\mathrm{fb}[t]}^{1/2} e_\lambda(\boldsymbol{X}) \right]. \tag{D.13}$$

Either of these frameworks may make sense as a method of selecting paths through the stimulus space.

Appendix E

Code Availability

Code implementing the algorithms discussed in Chapter 3 is available at `www.its.caltech.edu/~tadesaut/`.