Supplementary Material for "Time-Varying Gaussian Process Bandit Optimization"

(AISTATS 2016; Ilija Bogunovic, Volkan Cevher, Jonathan Scarlett)

Note that all citations here are to the bibliography in the main document, and similarly for many of the cross-references.

A Posterior Updates

Here we derive the posterior update rules for the time-varying setting via a suitable adaptation of the derivation for the time-invariant setting [6]. Observe from (3)–(4) that each function f_t depends only on the functions g_i for $i \le t$. By a simple recursion, we readily obtain for all t, j and x that

$$\mathrm{Cov}[f_t(x)f_{t+j}(x')] = (1-\epsilon)^{j/2}\mathbb{E}[f_t(x)f_t(x')] = (1-\epsilon)^{j/2}k(x,x'). \tag{22}$$

Hence, and since each output y_t equals the corresponding function sample $f_t(x_t)$ plus additive Gaussian noise z_t with variance σ^2, the joint distribution between the previous outputs $\mathbf{y}_t = (y_1, \dots, y_t)$ (corresponding to the points $\mathbf{x}_t = (x_1, \dots, x_t)$) and the next function value $f_{t+1}(x)$ is

$$\begin{bmatrix} \mathbf{y} \\ f_{t+1}(x) \end{bmatrix} \sim \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} \widetilde{\mathbf{K}}_t + \sigma^2\mathbf{I}_t & \widetilde{\mathbf{k}}_t(x) \\ \widetilde{\mathbf{k}}_t(x)^T & k(x,x) \end{bmatrix}\right) \tag{23}$$

using the definitions in the proposition statement. Using the formula for the conditional distribution associated with a jointly Gaussian random vector [6, App. A], we find that $f_{t+1}(x)$ is conditionally Gaussian with mean $\tilde{\mu}_t(x)$ and variance $\tilde{\sigma}_t(x)^2$, as was to be shown.

B Learning ϵ via Maximum-Likelihood

In this section, we provide an overview of how ϵ can be learned from training data in a principled manner; the details can be found in [20, Section 4.3] and [6, Section 5]. Throughout this appendix, we assume that the kernel matrix is parametrized by a set of hyperparameters θ (e.g., $\theta = (\nu, l)$ for the Mátern kernel), σ and ϵ.

Let $\bar{\mathbf{y}}$ be a vector of observations such that the i-th entry is observed at time t_i as a result of sampling the function f_{t_i} at location x_i. Note that there will typically be many indices i sharing common values of t_i, since in the training data we often have multiple samples at each time. Under our time-varying GP model, the marginal log-likelihood of $\bar{\mathbf{y}}$ given the hyperparameters is

$$\log p(\bar{\mathbf{y}}|\theta,\sigma,\epsilon) = -\frac{1}{2}\bar{\mathbf{y}}^T(\bar{\mathbf{K}}\circ\bar{\mathbf{D}} + \sigma^2\mathbf{I})^{-1}\bar{\mathbf{y}} - \frac{1}{2}\log|\bar{\mathbf{K}}\circ\bar{\mathbf{D}} + \sigma^2\mathbf{I}| - \frac{n}{2}\log(2\pi) \tag{24}$$

where $(\bar{\mathbf{K}})_{ij} = k(x_i, x_j)$ and $(\bar{\mathbf{D}})_{ij} = (1-\epsilon)^{|t_i - t_j|/2}$. To set the hyperparameters by maximizing the marginal likelihood, we can use the partial derivatives with respect to the hyperparameters. In particular, we have

$$\frac{\partial \log p(\bar{\mathbf{y}}|\theta,\sigma,\epsilon)}{\partial\epsilon} = \frac{1}{2}\mathrm{tr}\big((\boldsymbol{\alpha}\boldsymbol{\alpha}^T - (\bar{\mathbf{K}}\circ\bar{\mathbf{D}} + \sigma^2\mathbf{I})^{-1})(\bar{\mathbf{K}}\circ\bar{\mathbf{D}}')\big), \tag{25}$$

where $\boldsymbol{\alpha} = (\bar{\mathbf{K}}\circ\bar{\mathbf{D}} + \sigma^2\mathbf{I})^{-1}\bar{\mathbf{y}}$, and $(\bar{\mathbf{D}}')_{ij} = -v(1-\epsilon)^{v-1}$ with $v = |t_i - t_j|/2$.

We can now fit ϵ and the other hyperparameters by optimizing the marginal likelihood on the training data, e.g. by using an optimization algorithm from the family of quasi-Newton methods.

C Analysis of TV-GP-UCB (Theorem 4.3)

We recall the following alternative form for the mutual information (see (14)) from [1, Lemma 5.3], which extends immediately to the time-varying setting:

$$\tilde{I}(\mathbf{f}_T; \mathbf{y}_T) = \frac{1}{2}\sum_{t=1}^{T}\log\big(1 + \sigma^{-2}\tilde{\sigma}_{t-1}^2(x_t)\big). \tag{26}$$

C.1 Proof of (20)

The initial steps of the proof follow similar ideas to [1], but with suitable modifications to handle the fact that we have a different function f_t at each time instant. A key difficulty is in subsequently bounding the maximum mutual information in the presence of time variations, which is done in the following subsection.

We first fix a discretization $D_t \subset D \subseteq [0,r]^d$ of size $(\tau_t)^d$ satisfying

$$\|x - [x]_t\|_1 \le rd/\tau_t, \quad \forall x \in D, \tag{27}$$

where $[x]_t$ denotes the closest point in D_t to x. For example, a uniformly-spaced grid suffices to ensure that this holds.

We now fix a constant $\delta > 0$ and an increasing sequence of positive constants $\{\pi_t\}_{t=1}^{\infty}$ satisfying $\sum_{t\ge1} \pi_t^{-1} = 1$ (e.g. $\pi_t = \pi^2 t^2/6$), and condition on three high-probability events:

1. We first claim that if $\beta_t \ge 2\log\frac{3\pi_t}{\delta}$ then the selected points $\{x_t\}_{t=1}^{T}$ satisfy the confidence bounds

$$|f_t(x_t) - \tilde{\mu}_{t-1}(x_t)| \le \beta_t^{1/2}\tilde{\sigma}_{t-1}(x_t), \quad \forall t \tag{28}$$

   with probability at least $1 - \frac{\delta}{3}$. To see this, we note that conditioned on the outputs $(y_1, \dots, y_{t-1})$, the sampled points $(x_1, \dots, x_t)$ are deterministic, and $f_t(x_t) \sim \mathcal{N}(\tilde{\mu}_{t-1}(x_t), \tilde{\sigma}_{t-1}^2(x_t))$. An $\mathcal{N}(\mu, \sigma^2)$ random variable is within $\sqrt{\beta}\sigma$ of μ with probability at least $1 - e^{-\beta/2}$, and hence our choice of β_t ensures that the event corresponding to time t in (28) occurs with probability at least $\frac{\delta}{3\pi_t}$. Taking the union bound over t establishes the claim.

2. By the same reasoning with an additional union bound over $x \in D_t$, if $\beta_t \ge 2\log\frac{3|D_t|\pi_t}{\delta}$, then

$$|f_t(x) - \tilde{\mu}_{t-1}(x)| \le \beta_t^{1/2}\tilde{\sigma}_{t-1}(x), \quad \forall t, x \in D_t \tag{29}$$

   with probability at least $1 - \frac{\delta}{3}$.

3. Finally, we claim that setting $L_t = b\sqrt{\log(3da\pi_t/\delta)}$ yields

$$|f_t(x) - f_t(x')| \le L_t\|x - x'\|_1, \quad \forall t, x \in D, x' \in D \tag{30}$$

   with probability at least $1-\frac{\delta}{3}$. To see this, we note that by the assumption in (10) and the union bound over $j = 1, \dots, d$, the event corresponding to time t in (30) holds with probability at least $1 - dae^{-L_t^2/b^2} = \frac{\delta}{3\pi_t}$. Taking the union bound over t establishes the claim.

Again applying the union bound, all three of (28)–(30) hold with probability at least $1 - \delta$. We henceforth condition on each of them occurring.

Combining (27) with (30) yields for all x that

$$|f_t(x) - f_t([x]_t)| \le L_t rd/\tau_t \tag{31}$$

$$= b\sqrt{\log(3da\pi_t/\delta)}rd/\tau_t, \tag{32}$$

and hence choosing $\tau_t = rdbt^2\sqrt{\log(2da\pi_t/\delta)}$ yields

$$|f_t(x) - f_t([x]_t)| \le 1/t^2. \tag{33}$$

Note that this choice of τ_t yields $|D_t| = (\tau_t)^d = (rdbt^2\sqrt{\log(3da\pi_t/\delta)})^d$. In order to satisfy both lower bounds on β_t stated before (28) and (29), it suffices to take higher of the two (i.e., the second), yielding

$$\beta_t = 2\log(3\pi_t/\delta) + 2d\log(rdbt^2\sqrt{\log(3da\pi_t/\delta)}). \tag{34}$$

This coincides with (19) upon setting $\pi_t = \pi^2t^2/6$.

Substituting (33) into (29) and applying the triangle inequality, we find the maximizing point x_t^* at time t satisfies

$$|f_t(x_t^*) - \tilde{\mu}_{t-1}([x_t^*]_t)| \le \beta_t^{1/2} \tilde{\sigma}_{t-1}([x_t^*]_t) + 1/t^2. \tag{35}$$

Thus, we can bound the instantaneous regret as follows:

$$\begin{aligned} r_t &= f_t(x_t^*) - f_t(x_t) & (36)\\ &\le \tilde{\mu}_{t-1}([x_t^*]_t) + \beta_t^{1/2}\tilde{\sigma}_{t-1}([x_t^*]_t) + 1/t^2 - f_t(x_t) & (37)\\ &\le \tilde{\mu}_{t-1}(x_t) + \beta_t^{1/2}\tilde{\sigma}_{t-1}(x_t) + 1/t^2 - f_t(x_t) & (38)\\ &\le 2\beta_t^{1/2}\tilde{\sigma}_{t-1}(x_t) + 1/t^2, & (39) \end{aligned}$$

where in (37) we used (35), (38) follows since the function $\tilde{\mu}_{t-1}(x) + \beta_t^{1/2}\tilde{\sigma}_{t-1}(x)$ is maximized at x_t by the definition of the algorithm, and (39) follows from (28).

Finally, we bound the cumulative regret as

$$\begin{aligned} R_T = \sum_{t=1}^T r_t &\le \sum_{t=1}^T \left(2\beta_t^{1/2}\tilde{\sigma}_{t-1}(x_t) + 1/t^2\right) & (40)\\ &\le \sqrt{T\sum_{t=1}^T 4\beta_t \tilde{\sigma}_{t-1}^2(x_t)} + 2 & (41)\\ &\le \sqrt{C_1 T \beta_T \tilde{\gamma}_T} + 2, & (42) \end{aligned}$$

where (41) follows using $\sum_{t=1}^\infty 1/t^2 = \pi^2/6 \le 2$ the fact that $\|z\|_1 \le \sqrt{T}\|z\|_2$ for any vector $z \in \mathbb{R}^T$. Equation (42) is proved using following steps from [1, Lemma 5.4], which we include for completeness:

$$\begin{aligned} \sum_{t=1}^T 4\beta_t \tilde{\sigma}_{t-1}^2(x_t) &\le 4\beta_T\sigma^2 \sum_{t=1}^T \sigma^{-2}\tilde{\sigma}_{t-1}^2(x_t) & (43)\\ &\le 4\beta_T\sigma^2 \sum_{t=1}^T C_2 \log\left(1 + \sigma^{-2}\tilde{\sigma}_{t-1}^2(x_t)\right) & (44)\\ &\le C_1\beta_T\tilde{\gamma}_T, & (45) \end{aligned}$$

where (43) follows since β_t is increasing in T, (44) holds with $C_2 = \sigma^{-2}/\log(1+\sigma^{-2})$ using the identity $z^2 \le C_2 \log(1+z^2)$ for $z^2 \in [0, \sigma^{-2}]$ (note also that $\sigma^{-2}\tilde{\sigma}_{t-1}^2(\mathbf{x}_t) \le \sigma^{-2}k(\mathbf{x}_t, \mathbf{x}_t) \le \sigma^{-2}$), and (45) follows from the definitions of C_1 and $\tilde{\gamma}_T$, along with the alternative form for the mutual information in (26).

C.2 Proof of (21)

It remains to show that

$$\tilde{\gamma}_T \le \left(\frac{T}{\tilde{N}} + 1\right)\left(\gamma_{\tilde{N}} + \tilde{N}^3\epsilon\right) \tag{46}$$

under the definitions in (14)–(15). Recall that $\mathbf{x}_T = (x_1, \dots, x_T)$ are the points of interest, $\mathbf{f}_T = (f_1(x_1), \dots, f_T(x_T))$ are the corresponding function values, and $\mathbf{y}_T = (y_1, \dots, y_T)$ contains the corresponding noisy observations with $y_i = f_i(x_i) + z_i$.

At a high level, we bound the mutual information with time variations in terms of the corresponding quantity for the time-invariant case [1] by splitting the time steps $\{1, \dots, T\}$ into $\frac{T}{\tilde{N}}$ blocks of length $\tilde{N}$, such that within each block the function f_i does not vary significantly. We assume for the time being that $T/\tilde{N}$ is an integer, and then handle the general case.

Using the chain rule for mutual information and the fact that the noise sequence $\{z_i\}$ is independent, we have [25, Lemma 7.9.2]

$$\tilde{I}(\mathbf{f}_T; \mathbf{y}_T) \le \sum_{i=1}^{T/\tilde{N}} \tilde{I}(\mathbf{f}_{\tilde{N}}^{(i)}; \mathbf{y}_{\tilde{N}}^{(i)}), \tag{47}$$

where $\mathbf{y}_{\tilde{N}}^{(i)} = (y_{\tilde{N}(i-1)+1}, \dots, y_{\tilde{N}i})$ contains the measurements in the i-th block, and $\mathbf{f}_{\tilde{N}}^{(i)}$ is defined analogously. Maximizing both sides over $(x_1, \dots, x_T$, we obtain

$$\tilde{\gamma}_T \le \frac{T}{\tilde{N}} \tilde{\gamma}_{\tilde{N}}. \tag{48}$$

We are left to bound $\tilde{\gamma}_{\tilde{N}}$. To this end, we write the relevant covariance matrix as

$$\widetilde{\mathbf{K}}_{\tilde{N}} = \mathbf{K}_{\tilde{N}} \circ \mathbf{D}_{\tilde{N}} = \mathbf{K}_{\tilde{N}} + \mathbf{A}_{\tilde{N}}, \tag{49}$$

where

$$\mathbf{A}_{\tilde{N}} := \mathbf{K}_{\tilde{N}} \circ \mathbf{D}_{\tilde{N}} - \mathbf{K}_{\tilde{N}} \tag{50}$$

$$= \mathbf{K}_{\tilde{N}} \circ (\mathbf{D}_{\tilde{N}} - \mathbf{1}_{\tilde{N}}) \tag{51}$$

and $\mathbf{1}_{\tilde{N}}$ is the $\tilde{N} \times \tilde{N}$ matrix of ones. Observe that the (i,j)-th entry of $\mathbf{D}_{\tilde{N}} - \mathbf{1}_{\tilde{N}}$ has absolute value $1 - (1-\epsilon)^{\frac{|i-j|}{2}}$, which is upper bounded for all $\epsilon \in [0,1]$ by $\epsilon|i-j|$.[4] Hence, and using the fact that each entry of $\mathbf{K}_{\tilde{N}}$ lies in the range $[0,1]$, we obtain the following bound on the Frobenius norm:

$$\|\mathbf{A}_{\tilde{N}}\|_F^2 \le \sum_{i,j} (i-j)^2 \epsilon^2 \tag{52}$$

$$= \frac{1}{6} \tilde{N}^2 (\tilde{N}^2 - 1) \epsilon^2 \tag{53}$$

$$\le \tilde{N}^4 \epsilon^2, \tag{54}$$

where (53) is a standard double summation formula. We will use this inequality to bound $\tilde{\gamma}_N$ via Mirsky's theorem, which is given as follows.

Lemma C.1. (Mirsky's theorem [26, Cor. 7.4.9.3]) *For any matrices $\mathbf{U}_{\tilde{N}}$ and $\mathbf{V}_{\tilde{N}}$, and any unitarily invariant norm $|||\cdot|||$, we have*

$$|||\mathrm{diag}(\lambda_1(\mathbf{U}_{\tilde{N}}), \dots, \lambda_{\tilde{N}}(\mathbf{U}_{\tilde{N}})) - \mathrm{diag}(\lambda_1(\mathbf{V}_{\tilde{N}}), \dots, \lambda_{\tilde{N}}(\mathbf{V}_{\tilde{N}}))||| \le |||\mathbf{U}_{\tilde{N}} - \mathbf{V}_{\tilde{N}}|||, \tag{55}$$

where $\lambda_i(\cdot)$ is the i-th largest eigenvalue.

Using this lemma with $\mathbf{U}_{\tilde{N}} = \mathbf{K}_{\tilde{N}} + \mathbf{A}_{\tilde{N}}$, $\mathbf{V}_{\tilde{N}} = \mathbf{K}_{\tilde{N}}$, and $|||\cdot||| = \|\cdot\|_F$, and making use of (54), we find that $\lambda_i(\mathbf{K}_{\tilde{N}} + \mathbf{A}_{\tilde{N}}) = \lambda_i(\mathbf{K}_{\tilde{N}}) + \Delta_i$ for some $\{\Delta_i\}_{i=1}^{\tilde{N}}$ satisfying $\sum_{i=1}^{\tilde{N}} \Delta_i^2 \le \tilde{N}^4 \epsilon^2$. We thus have

$$\tilde{\gamma}_{\tilde{N}} = \sum_{i=1}^{\tilde{N}} \log\big(1 + \lambda_i(\mathbf{K}_{\tilde{N}} + \mathbf{A}_{\tilde{N}})\big) \tag{56}$$

$$= \sum_{i=1}^{\tilde{N}} \log\big(1 + \lambda_i(\mathbf{K}_{\tilde{N}}) + \Delta_i\big) \tag{57}$$

$$\le \gamma_{\tilde{N}} + \sum_{i=1}^{\tilde{N}} \log(1 + \Delta_i) \tag{58}$$

$$\le \gamma_{\tilde{N}} + \tilde{N} \log(1 + \tilde{N}^2 \epsilon) \tag{59}$$

$$\le \gamma_{\tilde{N}} + \tilde{N}^3 \epsilon, \tag{60}$$

where (58) follows from the inequality $\log(1+a+b) \le \log(1+a) + \log(1+b)$ for non-negative a and b (and the definition in (12)), (59) follows since a simple analysis of the optimality conditions of

$$\text{maximize} \quad \sum_{i=1}^{\tilde{N}} \log(1+\Delta_i) \qquad \text{subject to} \qquad \sum_{i=1}^{\tilde{N}} \Delta_i^2 \le \tilde{N}^4 \epsilon^2 \tag{61}$$

[4] For $|i-j| \ge 2$, this follows since the function of interest is concave, passes through the origin, and has derivative $\frac{|i-j|}{2} \le |i-j|$ there. For $k=1$, the statement follows by observing that equality holds for $\epsilon \in \{0,1\}$, and noting that the function of interest is convex.

reveals that the maximum is achieved when all of the Δ_i are equal to $\tilde{N}^2\epsilon$, and (60) follows from the inequality $\log(1+a) \le a$.

Recalling that we are considering the case that $T/\tilde{N}$ is an integer, we obtain (21) by combining (48) and (60). In the general case, we simply use the fact that $\tilde{\gamma}_T$ is increasing in T by definition, hence leading to the addition of one in (21).

D Analysis of R-GP-UCB (Theorem 4.2)

Parts of the proof of Theorem 4.2 overlap with that of Theorem 4.3; we focus primarily on the key differences. First, overloading the notation from the TV-GP-UCB analysis, we let $\tilde{\mu}_t(x)$ and $\tilde{\sigma}_t(x)$ be defined as in (7)–(8), but using *only the samples since the previous reset in the R-GP-UCB algorithm*, and similarly for $\mathbf{k}_t$, $\widetilde{\mathbf{k}}_t$, $\mathbf{d}_t$, and so on. Thus, for example, the dimension of $\mathbf{k}_t$ is at most the length N between resets, and the entries of $\mathbf{D}_t$ are no smaller than $(1-\epsilon)^{N/2}$. Note that the time-invariant counterparts $\mu_t(x)$ and $\sigma_t(x)$ (computed using $\mathbf{k}$ and $\mathbf{K}$ in place of $\widetilde{\mathbf{k}}$ and $\widetilde{\mathbf{K}}$) are used in the algorithm, thus creating a mismatch that must be properly handled.

Recall the definitions of the discretization D_t (whose cardinality is again set to τ_t^d for some τ_t), the corresponding quantization function $[x]_t$, and the constants π_t. We now condition on four (rather than three) high probability events:

- Setting $\beta_t = 2\log\frac{4|D_t|\pi_t}{\delta}$, the same arguments as those leading to (28)–(29) reveal that

$$|f_t(x_t) - \tilde{\mu}_{t-1}(x_t)| \le \beta_t^{1/2}\tilde{\sigma}_{t-1}(x_t) \quad \forall t \ge 1 \tag{62}$$

$$|f_t(x) - \tilde{\mu}_{t-1}(x)| \le \beta_t^{1/2}\tilde{\sigma}_{t-1}(x) \quad \forall t \ge 1, x \in D_t \tag{63}$$

  with probability at least $1-\frac{\delta}{2}$. Note that in proving these claims we only condition on the observations since the last reset, rather than all of the points since $t=1$.

- Using the same argument as (30), the assumption in (10) implies that

$$\Big|\frac{\partial f_t(x)}{\partial x^{(j)}}\Big| \le L_t := b_1\sqrt{\log\frac{4da_1\pi_t}{\delta}} \quad \forall t \ge 1, x \in D, j \in \{1,\dots,d\} \tag{64}$$

  with probability at least $1-\frac{\delta}{4}$.

- We claim that the assumption in (9) similarly implies that

$$|y_t| \le \tilde{L}_t := (2+b_0)\sqrt{\log\frac{4(1+a_0)\pi_t}{\delta}} \quad \forall t \ge 1, x \in D \tag{65}$$

  with probability at least $1-\frac{\delta}{4}$. To see this, we first note that $\Pr\big(|z_t| \le L\big) \le e^{-L^2/2}$ since $z_t \sim \mathcal{N}(0,1)$, and by a standard bound on the standard normal tail probability. Combining this with (9) and noting that $|y_t| \le |f_t(x_t)| + |z_t|$, we find that $\Pr\big(|y_t| > 2L\big)$ is upper bounded by $e^{-L^2/2} + a_0 e^{-(L/b_0)^2}$, which in turn is upper bounded by $(1+a_0)e^{-(L/(2+b_0))^2}$. Choosing $L = \tilde{L}_t/2$ and applying the union bound and some simple manipulations, we obtain (65).

By the union bound, all four of (62)–(65) hold with probability at least $1-\delta$.

As in the TV-GP-UCB proof, we set $\tau_t = rdt^2L_t$, thus ensuring that $|f_t(x) - f_t([x]_t)| \le \frac{1}{t^2}$ for all $x \in D$. Defining

$$\Delta_t^{(\mu)} := \sup_{x\in D} |\tilde{\mu}_t(x) - \mu_t(x)| \tag{66}$$

$$\Delta_t^{(\sigma)} := \sup_{x\in D} |\tilde{\sigma}_t(x) - \sigma_t(x)| \tag{67}$$

to be the maximal errors between the true and the mismatched posterior updates, we have the following:

$$r_t = f_t(x_t^*) - f_t(x_t) \tag{68}$$
$$\leq f_t([x_t^*]_t) - f_t(x_t) + \frac{1}{t^2} \tag{69}$$
$$\leq \tilde{\mu}_{t-1}([x_t^*]_t) + \beta_t^{1/2}\tilde{\sigma}_{t-1}([x_t^*]_t) - \tilde{\mu}_{t-1}(x_t) + \beta_t^{1/2}\tilde{\sigma}_{t-1}(x_t) + \frac{1}{t^2} \tag{70}$$
$$\leq \mu_{t-1}([x_t^*]_t) + \beta_t^{1/2}\sigma_{t-1}([x_t^*]_t) - \mu_{t-1}(x_t) + \beta_t^{1/2}\sigma_{t-1}(x_t) + 2\Delta_t^{(\mu)} + 2\beta_t^{1/2}\Delta_t^{(\sigma)} + \frac{1}{t^2} \tag{71}$$
$$\leq 2\beta_t^{1/2}\sigma_{t-1}(x_t) + 2\Delta_t^{(\mu)} + 2\beta_t^{1/2}\Delta_t^{(\sigma)} + \frac{1}{t^2}, \tag{72}$$

where (69) follows in the same way as (37), (70) follows from (62)–(63), (71) follows from the definitions in (66)–(67), and (72) follows from the choice of x_t in the algorithm.

The key remaining step is to characterize $\Delta_t^{(\mu)}$ and $\Delta_t^{(\sigma)}$. Our findings are summarized in the following lemma.

Lemma D.1. *Conditioned on the event in* (65)*, we have* $\Delta_t^{(\mu)} \leq \big(\sigma^{-2}+\sigma^{-4}\big)N^3\epsilon\tilde{L}_t$ *and* $\Delta_t^{(\sigma)} \leq \big(3\sigma^{-2}+\sigma^{-4}\big)N^3\epsilon$ *almost surely.*

This lemma implies Theorem 4.2 upon substitution into (72), setting $\pi_t = \pi^2t^2/6$, and following the steps from (40) onwards. In the remainder of the section, we prove the lemma. The claims on $\Delta_t^{(\mu)}$ and $\Delta_t^{(\sigma)}$ are proved similarly; we focus primarily on the latter since it is the (slightly) more difficult of the two.

The subsequent analysis applies for arbitrary values of t and x, so we use the shorthands $\mathbf{k} := \mathbf{k}_t(x)$, $\mathbf{K} := \mathbf{K}_t(x)$, $\widetilde{\mathbf{k}} := \widetilde{\mathbf{k}}_t(x)$, $\widetilde{\mathbf{K}} := \widetilde{\mathbf{K}}_t$ and $\mathbf{I} := \mathbf{I}_t$. We first use the definition in (8) and the triangle inequality to write

$$|\tilde{\sigma}_t(x)^2 - \sigma_t(x)^2|$$
$$= \big|\widetilde{\mathbf{k}}^T(\widetilde{\mathbf{K}} + \sigma^2\mathbf{I})^{-1}\widetilde{\mathbf{k}} - \mathbf{k}^T(\mathbf{K} + \sigma^2\mathbf{I})^{-1}\mathbf{k}\big| \tag{73}$$
$$\leq \big|\widetilde{\mathbf{k}}^T(\widetilde{\mathbf{K}} + \sigma^2\mathbf{I})^{-1}\widetilde{\mathbf{k}} - \widetilde{\mathbf{k}}^T(\mathbf{K} + \sigma^2\mathbf{I})^{-1}\widetilde{\mathbf{k}}\big| + \big|\widetilde{\mathbf{k}}^T(\mathbf{K} + \sigma^2\mathbf{I})^{-1}\widetilde{\mathbf{k}} - \mathbf{k}^T(\mathbf{K} + \sigma^2\mathbf{I})^{-1}\mathbf{k}\big| \tag{74}$$
$$:= T_1 + T_2. \tag{75}$$

We proceed by bounding T_1 and T_2 separately, starting with the latter.

Set $\mathbf{M} := (\mathbf{K} + \sigma^2\mathbf{I})^{-1}$ for brevity. By expanding the quadratic function $(\widetilde{\mathbf{k}} - \mathbf{k})^T\mathbf{M}(\widetilde{\mathbf{k}} - \mathbf{k})^T$, grouping the terms appearing in T_2, and applying the triangle inequality, we obtain

$$T_2 \leq 2|\mathbf{k}^T\mathbf{M}(\widetilde{\mathbf{k}} - \mathbf{k})| + |(\widetilde{\mathbf{k}} - \mathbf{k})^T\mathbf{M}(\widetilde{\mathbf{k}} - \mathbf{k})|. \tag{76}$$

We upper bound each of these terms of the form $\mathbf{a}^T\mathbf{M}\mathbf{b}$ by $\|\mathbf{a}\|_2\|\mathbf{M}\|_{2\to2}\|\mathbf{b}\|_2$, where $\|\mathbf{M}\|_{2\to2}$ is the spectral norm. By definition, λ is an eigenvalue of $\mathbf{K}$ if and only if $\frac{1}{\lambda+\sigma^2}$ is an eigenvalue of $\mathbf{M}$; since $\mathbf{K}$ is positive semi-definite, it follows that $\|\mathbf{M}\|_{2\to2} \leq \frac{1}{\sigma^2}$. We also have $\|\mathbf{k}\|_2^2 \leq N$ since the entries of $\mathbf{k}$ lies in $[0,1]$, and $\|\widetilde{\mathbf{k}} - \mathbf{k}\|_2^2 \leq N^3\epsilon^2$ since the absolute values of the entries of $\widetilde{\mathbf{k}} - \mathbf{k}$ are upper bounded by $N\epsilon$ by the argument following (51). Combining these, we obtain

$$T_2 \leq 2\sigma^{-2}N^2\epsilon + \sigma^{-2}N^3\epsilon^2. \tag{77}$$

To bound T_1, we use the following inequality for positive definite matrices $\mathbf{U}, \mathbf{V}$ and any unitarily invariant norm $|||\cdot|||$ [27, Lemma X.1.4]:

$$\big|\big|\big|(\mathbf{U} + \mathbf{I})^{-1} - (\mathbf{U} + \mathbf{V} + \mathbf{I})^{-1}\big|\big|\big| \leq \big|\big|\big|\mathbf{I} - (\mathbf{V} + \mathbf{I})^{-1}\big|\big|\big|. \tag{78}$$

Specializing to the spectral norm, multiplying through by $\frac{1}{\sigma^2}$, and choosing $\mathbf{U} = \frac{1}{\sigma^2}\mathbf{K}$ and $\mathbf{V} = \frac{1}{\sigma^2}(\widetilde{\mathbf{K}} - \mathbf{K})$, we obtain

$$\big\|(\mathbf{K} + \sigma^2\mathbf{I})^{-1} - (\widetilde{\mathbf{K}} + \sigma^2\mathbf{I})^{-1}\big\|_{2\to2} \leq \big\|\sigma^{-2}\mathbf{I} - (\widetilde{\mathbf{K}} - \mathbf{K} + \sigma^2\mathbf{I})^{-1}\big\|_{2\to2}. \tag{79}$$

Next, λ is an eigenvalue of $\widetilde{\mathbf{K}} - \mathbf{K}$ if and only if $\sigma^{-2} - \frac{1}{\lambda+\sigma^2}$ is an eigenvalue of $\sigma^{-2}\mathbf{I} - (\widetilde{\mathbf{K}} - \mathbf{K} + \sigma^2\mathbf{I})^{-1}$. Writing $\sigma^{-2} - \frac{1}{\lambda+\sigma^2} = \sigma^{-2}\big(1 - \frac{1}{\lambda/\sigma^2+1}\big) \leq \sigma^{-4}\lambda$, it follows that the right-hand side of (79) is upper bounded by

$\sigma^{-4}\|\widetilde{\mathbf{K}}-\mathbf{K}\|_{2\to 2}$. Using (54) (observe that $\mathbf{A}_N = \widetilde{\mathbf{K}}-\mathbf{K}$) and the fact that the spectral norm is upper bounded by the Frobenius norm, we obtain $\|\widetilde{\mathbf{K}}-\mathbf{K}\|_{2\to 2} \le N^2\epsilon$. Substituting into (79), we conclude that the matrix $\mathbf{M}' := (\mathbf{K}+\sigma^2\mathbf{I})^{-1} - (\widetilde{\mathbf{K}}+\sigma^2\mathbf{I})^{-1}$ has a spectral norm which is upper bounded by $\sigma^{-4}N^2\epsilon$. Finally, T_1 can be written as $\widetilde{\mathbf{k}}^T\mathbf{M}'\widetilde{\mathbf{k}}$, and since $\|\widetilde{\mathbf{k}}\|_2^2 \le N$ (since each entry of $\widetilde{\mathbf{k}}$ lies in $[0,1]$), we obtain

$$T_1 \le \sigma^{-4}N^3\epsilon. \tag{80}$$

Combining (77) and (80) and crudely writing $N^2\epsilon \le N^3\epsilon$ and $N^3\epsilon^2 \le N^3\epsilon$, we obtain

$$|\tilde{\sigma}_t(x)^2 - \sigma_t(x)^2| \le \big(3\sigma^{-2}+\sigma^{-4}\big)N^3\epsilon, \tag{81}$$

and hence, applying the inequality $(a-b)^2 \le |a^2-b^2|$, we obtain

$$\Delta_t^{(\sigma)} \le \sqrt{\big(3\sigma^{-2}+\sigma^{-4}\big)N^3\epsilon}. \tag{82}$$

To characterize $\Delta_t^{(\mu)}$, we write the following analog of (74):

$$\begin{aligned}
&|\tilde{\mu}_t(x)^2 - \mu_t(x)^2| \\
&\quad\le \big|\widetilde{\mathbf{k}}^T(\widetilde{\mathbf{K}}+\sigma^2\mathbf{I})^{-1}\mathbf{y} - \widetilde{\mathbf{k}}^T(\mathbf{K}+\sigma^2\mathbf{I})^{-1}\mathbf{y}\big| + \big|\widetilde{\mathbf{k}}^T(\mathbf{K}+\sigma^2\mathbf{I})^{-1}\mathbf{y} - \mathbf{k}^T(\mathbf{K}+\sigma^2\mathbf{I})^{-1}\mathbf{y}\big| &(83)\\
&\quad:= T_1 + T_2. &(84)
\end{aligned}$$

Following the same arguments as those above, and noting from (65) that $\|\mathbf{y}\|_2^2 \le N\tilde{L}_N^2$, we obtain

$$T_1 \le \sigma^{-2}N^2\epsilon\tilde{L}_N \tag{85}$$
$$T_2 \le \sigma^{-4}N^3\epsilon\tilde{L}_N, \tag{86}$$

and hence

$$\Delta_t^{(\mu)} \le \big(\sigma^{-2}+\sigma^{-4}\big)N^3\epsilon\tilde{L}_N. \tag{87}$$

E Applications to Specific Kernels (Corollary 4.1)

Throughout this section, we let $\mathcal{I}_T(z)$ denote the integer in $\{1,\dots,T\}$ which is closest to $z \in \mathbb{R}$. We focus primarily on the proof for TV-GP-UCB, since the proof for R-GP-UCB is essentially identical.

We begin with the squared exponential kernel. From [1, Thm. 5], we have $\gamma_{\tilde{N}} = \mathcal{O}(d\log\tilde{N}) = \tilde{\mathcal{O}}(1)$, and we thus obtain

$$\Big(\frac{T}{\tilde{N}}+1\Big)\Big(\gamma_{\tilde{N}} + N^3\epsilon\Big) = \tilde{\mathcal{O}}\bigg(\Big(\frac{T}{\tilde{N}}+1\Big)(1+\tilde{N}^3\epsilon)\bigg). \tag{88}$$

Setting $\tilde{N} = \mathcal{I}_T(\epsilon^{-1/3})$, we find that this behaves as $\tilde{\mathcal{O}}(T\epsilon^{1/3})$ when $\epsilon \ge \frac{1}{T^3}$, and as $\tilde{\mathcal{O}}(1)$ when $\epsilon < \frac{1}{T^3}$ (and hence $\tilde{N}=T$). Substitution into Theorem 4.3 yields the desired result.

For the Matérn kernel, we have from [1, Thm. 5] that $\gamma_{\tilde{N}} = \mathcal{O}(\tilde{N}^c\log\tilde{N}) = \tilde{\mathcal{O}}(\tilde{N}^c)$ with $c = \frac{d(d+1)}{2\nu+d(d+1)}$, and we thus obtain

$$\Big(\frac{T}{\tilde{N}}+1\Big)\Big(\gamma_{\tilde{N}} + \tilde{N}^3\epsilon\Big) = \tilde{\mathcal{O}}\bigg(\Big(\frac{T}{\tilde{N}}+1\Big)(\tilde{N}^c+\tilde{N}^3\epsilon)\bigg). \tag{89}$$

Setting $\tilde{N} = \mathcal{I}_T(\epsilon^{-\frac{1}{3-c}})$, we find that this behaves as $\tilde{\mathcal{O}}(T\epsilon^{\frac{1-c}{3-c}})$ when $\epsilon \ge \frac{1}{T^{3-c}}$, and as $\tilde{\mathcal{O}}(T^c)$ when $\epsilon < \frac{1}{T^{3-c}}$ (and hence $\tilde{N}=T$). Substitution into Theorem 4.3 yields the desired result.

For R-GP-UCB, the arguments are analogous using Theorem 4.2 in place of Theorem 4.3, with N playing the role of $\tilde{N}$. We set $N = \mathcal{I}_T(\epsilon^{-1/4})$ for the squared exponential kernel and $N = \mathcal{I}_T(\epsilon^{-\frac{1}{4-c}})$ for the Matérn kernel.

F Lower Bound (Theorem 4.1)

We obtain a lower bound on the regret of *any* algorithm by considering the optimal algorithm for a genie-aided setting. Specifically, suppose that at time t, the entire function f_{t-1} is known perfectly. We claim that the optimal strategy, in the sense of minimizing the expected regret, is to choose x_t to be any maximizer of f_{t-1}. This can be seen by noting that minimizing the regret $r_t = f_t(x_t^*) - f_t(x_t)$ is equivalent to maximizing the function value $f_t(x_t)$, since $f_t(x_t^*)$ is unaffected by the choice of x_t. Then, conditioned on the entire function f_{t-1}, the next value $f_t(x)$ is distributed as $\mathcal{N}(\sqrt{1-\epsilon}f_{t-1}(x), \epsilon)$, and clearly the optimal strategy is to choose the point that maximizes the mean.

We proceed by lower bounding the regret incurred by such a scheme. Recall that for each t, both f_t and g_t are distributed as $\mathcal{GP}(0, k)$. Thus, (10) and (11) hold for all such functions.

We let ∇f denote the gradient vector of a function f, and let $\nabla^2 f$ denote the Hessian matrix. For the time being, we condition on the previous function f_{t-1}, the selected point x_t (i.e., the maximizer of f_{t-1}) and the innovation function g_t satisfying the following events for some positive constants L and η:

$$\mathcal{A}_1 := \bigg\{ \Big| \frac{\partial^2 f_{t-1}(x)}{\partial x^{(j_1)} \partial x^{(j_2)}} \Big| \le L, \quad \forall j_1, j_2, x \bigg\} \tag{90}$$

$$\mathcal{A}_2 := \bigg\{ \Big| \frac{\partial^2 g_t(x)}{\partial x^{(j_1)} \partial x^{(j_2)}} \Big| \le L, \quad \forall j_1, j_2, x \bigg\} \tag{91}$$

$$\mathcal{A}_3 := \bigg\{ \frac{\sqrt{\epsilon}\big|\frac{\partial g_t(x)}{\partial x^{(j)}}\big|}{2L\sqrt{d}} \le \eta, \quad \forall j \bigg\} \tag{92}$$

$$\mathcal{A}_4 := \big\{ d(x_t, B) \ge \eta \big\}, \tag{93}$$

where $d(x_t, B) := \min_{x \in B} \|x_t - x\|_2$ is the distance of x_t to the closest point on the boundary B of the compact domain D. Observe that for any fixed η, $\Pr[\mathcal{A}_i]$ can be made arbitrarily close to one for $i = 1, 2, 3$ by choosing L sufficiently large. Moreover, we have $\Pr[\mathcal{A}_4] > 0$ for sufficiently small η, since otherwise the maximum of $f \sim \mathcal{GP}(0, k)$ would be on the boundary of the domain D with probability one. Applying the union bound, we conclude that the event $\mathcal{A} := \mathcal{A}_1 \cap \mathcal{A}_2 \cap \mathcal{A}_3 \cap \mathcal{A}_4$ occurs with strictly positive probability for suitable chosen η and L.

We fix an arbitrary vector v with $\|v\|_2 = 1$ and a constant $\delta > 0$, and note that the regret r_t at time t an be lower bounded as follows provided that $x_t + v\delta \in D$:

$$r_t = \max_x f_t(x) - f_t(x_t) \tag{94}$$

$$\ge f_t(x_t + v\delta) - f_t(x_t) \tag{95}$$

$$= \sqrt{1-\epsilon}\big(f_{t-1}(x_t + v\delta) - f_{t-1}(x_t)\big) + \sqrt{\epsilon}\big(g_t(x_t + v\delta) - g_t(x_t)\big) \tag{96}$$

$$= \sqrt{1-\epsilon}\,\frac{1}{2}\delta^2 v^T \big[\nabla^2 f_{t-1}(x_t + v\delta_f)\big] v + \sqrt{\epsilon}\big(\delta v^T \nabla g_t(x_t) + \frac{1}{2}\delta^2 v^T \big[\nabla^2 g_t(x_t + v\delta_g)\big] v\big), \tag{97}$$

where (96) follows by substituting the update equations in (3)–(4), and (97) holds for some $\delta_f \in [0, \delta]$ and $\delta_g \in [0, \delta]$ by a second-order Taylor expansion; note that $\nabla f_{t-1}(x_t) = 0$ since x_t maximizes f_{t-1}, whose peak is away from the boundary of the domain by (93).

We choose the unit vector v to have the same direction as $\nabla g_t(x_t)$, so that $\delta v^T \nabla g_t(x_t) = \delta \|\nabla g_t(x_t)\|_2$. By (90)–(91), the entries of $\nabla^2 f_{t-1}(x_t + v\delta_f)$ and $\nabla^2 g_t(x_t + v\delta_g)$ are upper bounded by L, and thus a standard inequality between the entry-wise ℓ_∞ norm and the spectral norm reveals that the latter is upper bounded by Ld. This, in turn, implies that $v^T \big[\nabla^2 f_{t-1}(x_t + v\delta_f)\big] v$ and $v^T \big[\nabla^2 g_t(x_t + v\delta_g)\big] v$ are upper bounded by Ld, and hence

$$r_t \ge \sqrt{\epsilon}\delta \|\nabla g_t(x_t)\|_2 - \frac{1}{2} L d \delta^2 \big(\sqrt{1+\epsilon} + \sqrt{\epsilon}\big) \tag{98}$$

$$\ge \sqrt{\epsilon}\delta \|\nabla g_t(x_t)\|_2 - L d \delta^2, \tag{99}$$

where we have used $\sqrt{1+\epsilon} + \sqrt{\epsilon} \le 2$. By differentiating with respect to δ, it is easily verified that the right-hand side is maximized by $\delta = \frac{\sqrt{\epsilon}\|\nabla g_t(x_t)\|_2}{2Ld}$. This choice is seen to be valid (i.e., it yields $x_t + v\delta$ still in the domain) by (92)–(93) and the fact that $\|z\|_2 \le \sqrt{d}\|z\|_\infty$ for $z \in \mathbb{R}^d$, and we obtain

$$r_t \ge \frac{\epsilon \|\nabla g_t(x_t)\|_2^2}{4Ld}. \tag{100}$$

It follows that the expectation of r_t is lower bounded by

$$\mathbb{E}[r_t] \ge \Pr[\mathcal{A}]\mathbb{E}[r_t|\mathcal{A}] \tag{101}$$

$$\ge \Pr[\mathcal{A}] \frac{\epsilon\, \mathbb{E}\big[\|\nabla g_t(x_t)\|_2^2 \,|\, \mathcal{A}\big]}{4Ld} \tag{102}$$

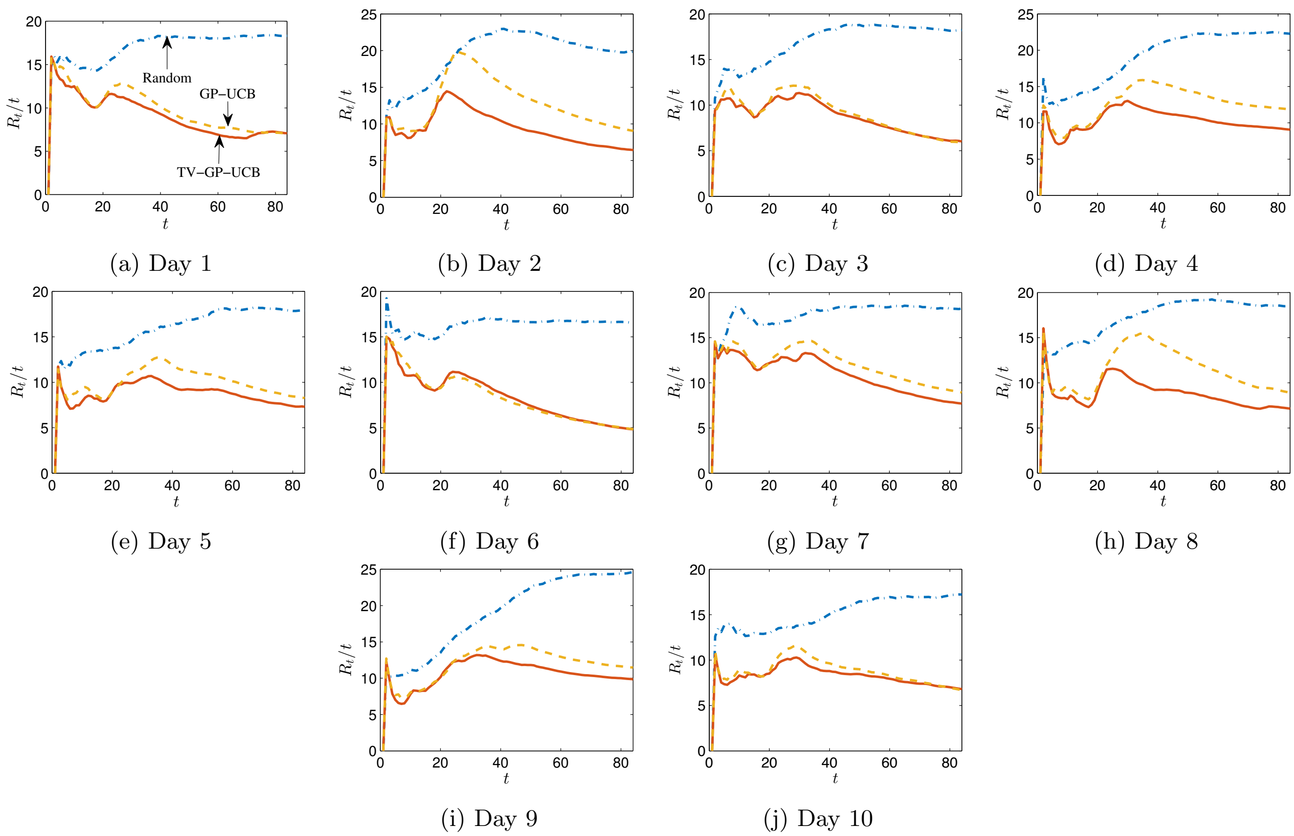
$$= \Theta(\epsilon), \tag{103}$$

where (101) follows since $r_t \ge 0$ almost surely, and (103) follows since $\mathbb{E}\big[\|\nabla g_t(x_t)\|_2^2 \,|\, \mathcal{A}\big] > 0$ by a simple proof by contradiction: The expectation can only equal zero if its (non-negative) argument is zero almost surely, but if that were the case then the unconditional distribution of $\|\nabla g_t(x_t)\|_2^2$ would satisfy $\Pr[\|\nabla g_t(x_t)\|_2^2 = 0] > \Pr[\mathcal{A}]$, which is impossible since the entries of $\nabla g_t(x_t)$ are Gaussian [6, Sec. 9.4] and hence have zero probability of being exactly zero.

Finally, using (103), the average cumulative regret satisfies $\mathbb{E}[R_T] = \sum_{i=1}^T \mathbb{E}[r_T] = \Omega(T\epsilon)$.

G Further results for traffic speed data

In Figure 4, we outline the complete results on all the testing days for the experiment described in Section 5.2. The sensors used in the experiment have the following IDs: [0, 54, 69, 77, 169, 131, 262, 216, 34, 320, 308, 177, 130, 221, 290, 348, 25, 157, 252, 83, 163, 149, 294, 21, 246, 45, 98, 74, 274, 237, 322, 29, 120, 44, 49, 241, 286, 99, 247, 297, 96, 234, 236, 205, 329, 214, 28, 175, 65, 220].



(a) Day 1 (b) Day 2 (c) Day 3 (d) Day 4

(e) Day 5 (f) Day 6 (g) Day 7 (h) Day 8

(i) Day 9 (j) Day 10

Figure 4: Numerical performance of upper confidence bound algorithms on traffic speed dataset.