Parallel Gaussian Process Optimization with Upper Confidence Bound and Pure Exploration

Emile Contal   David Buffoni   Alexandre Robicquet   Nicolas Vayatis
CMLA, ENS Cachan, CNRS, 61 Avenue du Président Wilson, F-94230 Cachan

Motivations

Examples of optimization problems with batches of size K.

- Tuning the parameters of a numerical experiment with a cluster of K cores
- Choose the locations of K sensors to find out the maximum response
- Design a marketing strategy with batches of K customers

Related Work.

- Experimental design
- Bayesian optimization
- Active learning
- Multiarmed bandit

Problem Statement

Sequential Batch Optimization.
Let $f : \mathcal{X} \to \mathbb{R}$ be our unknown objective with compact and convex $\mathcal{X} \subset \mathbb{R}^d$, we want to find the maximum $f(x^\star)$ via sequential evaluations. At iteration t, we choose a batch of K queries $\{x_t^1, \dots x_t^K\}$, and then observe simultaneously $y_t^k = f(x_t^k) + \epsilon_t^k$, where $\epsilon_t^k \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma^2)$. Assuming that evaluating f is expensive, we have to deal with the exploration/exploitation tradeoff.

Objective.
The cost incurred by a strategy which selects $x_1^1, \dots, x_T^K$ is given by the cumulative regret. Depending on the application, we may be interested in two forms of regrets.

$$\text{Batch cumulative regret: } R_T^K = \sum_{t<T} f(x^\star) - \max_{k<K} f(x_t^k) .$$

$$\text{Full cumulative regret: } R_{TK} = \sum_{t<T} \sum_{k<K} f(x^\star) - f(x_t^k) .$$

Gaussian Processes

Definition.
We say that f is distributed as a GP with mean function $m : \mathcal{X} \to \mathbb{R}$ and kernel function $k : \mathcal{X} \times \mathcal{X} \to \mathbb{R}^+$, written $f \sim \mathcal{GP}(m, k)$, when for all $x_1, \dots, x_n$ the values $\big(f(x_1), \dots, f(x_n)\big)$ are distributed as a multivariate Gaussian with mean and variance given by m and k.

Bayesian Inference.
At iteration t, with observations $\mathbf{Y}_{X_t}$ at $X_t = \{x_1^1, \dots, x_t^K\}$, the posterior distribution $\Pr[f \mid \mathbf{Y}_{X_t}]$ is a Gaussian process with mean and covariance given by the Bayesian inference. At each point $x \in \mathcal{X}$, we can compute a prediction with $\mu_{t+1}(x)$ and uncertainty with $\sigma_{t+1}^2(x)$.

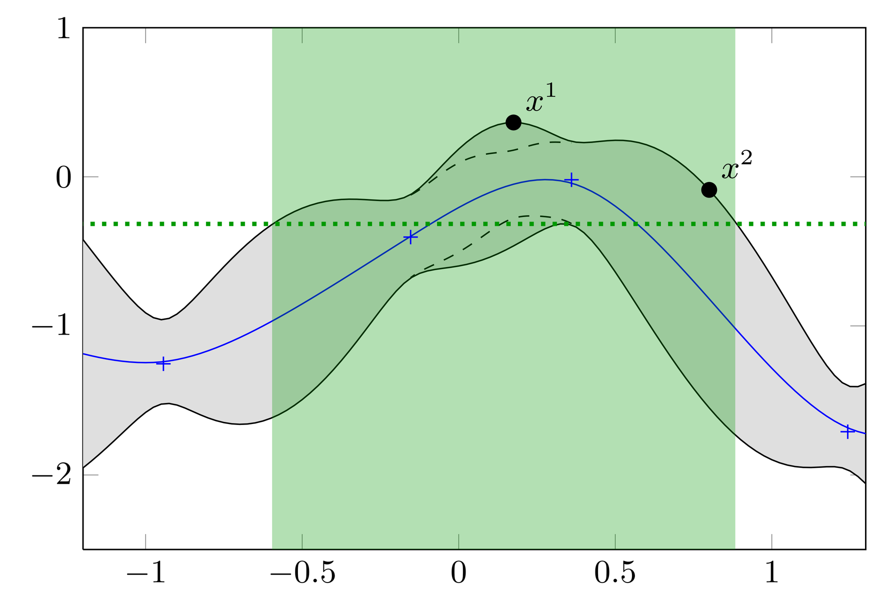


Figure : Two queries of GP-UCB-PE based on four observations of a GP in dimension 1.

Relevant Region

UCB and LCB.
Fix $0 < \delta < 1$ and $\beta_t = \mathcal{O}(\log \frac{t}{\delta})$, we define $\forall x \in \mathcal{X}, \forall t \geq 1$,

$$\text{Upper Confidence Bound: } f_t^+(x) = \mu_t(x) + \sqrt{\beta_t \sigma_t^2(x)} ,$$

$$\text{Lower Confidence Bound: } f_t^-(x) = \mu_t(x) - \sqrt{\beta_t \sigma_t^2(x)} .$$

Then, we have the following property with high probability:

$$\Pr\Big[f_t^-(x) \leq f(x) \leq f_t^+(x)\Big] \geq 1 - \delta .$$

Relevant Region.
We define the Relevant Region $\mathfrak{R}_t$ as,

$$\mathfrak{R}_t = \Big\{x \in \mathcal{X} \mid f_t^+(x) \geq y_t^\bullet\Big\} ,$$

where $y_t^\bullet = \max_{x \in \mathcal{X}} f_t^-(x)$ is the maximum of the LCB.
$\mathfrak{R}_t$ contains the location of the optimum $x^\star$ with high probability.

GP-UCB-PE

for $t = 1, 2, \dots$ **do**
  Compute μ_t and σ_t^2 with Bayesian inference on $y_1^1, \dots, y_{t-1}^K$
  $x_t^1 \leftarrow \operatorname{argmax}_{x \in \mathcal{X}} \widehat{f}_t^+(x)$ `// UCB rule`
  Compute $\mathfrak{R}_t$
  for $k = 2, \dots, K$ **do**
    Update $\widehat{\sigma}_t^{(k)}$ knowing the previous choices $x_t^1, \dots, x_t^{k-1}$
    $x_t^k \leftarrow \operatorname{argmax}_{x \in \mathfrak{R}_t^+} \widehat{\sigma}_t^{(k)}(x)$ `// Pure Exploration`
  end
  Query $\{x_t^k\}_{k<K}$
end

Mutual Information

We define γ_{TK} the maximum mutual information about f obtainable by a sequence of TK queries.

- For linear kernel, $\gamma_{TK} = \mathcal{O}(d \log TK)$
- For RBF kernel, $\gamma_{TK} = \mathcal{O}\big((\log TK)^{d+1}\big)$
- For Matérn kernel, $\gamma_{TK} = \mathcal{O}\big((TK)^\alpha \log TK\big)$, with $\alpha = \frac{d(d+1)}{2\nu + d(d+1)} \leq 1$

Theorem: Regret Bounds

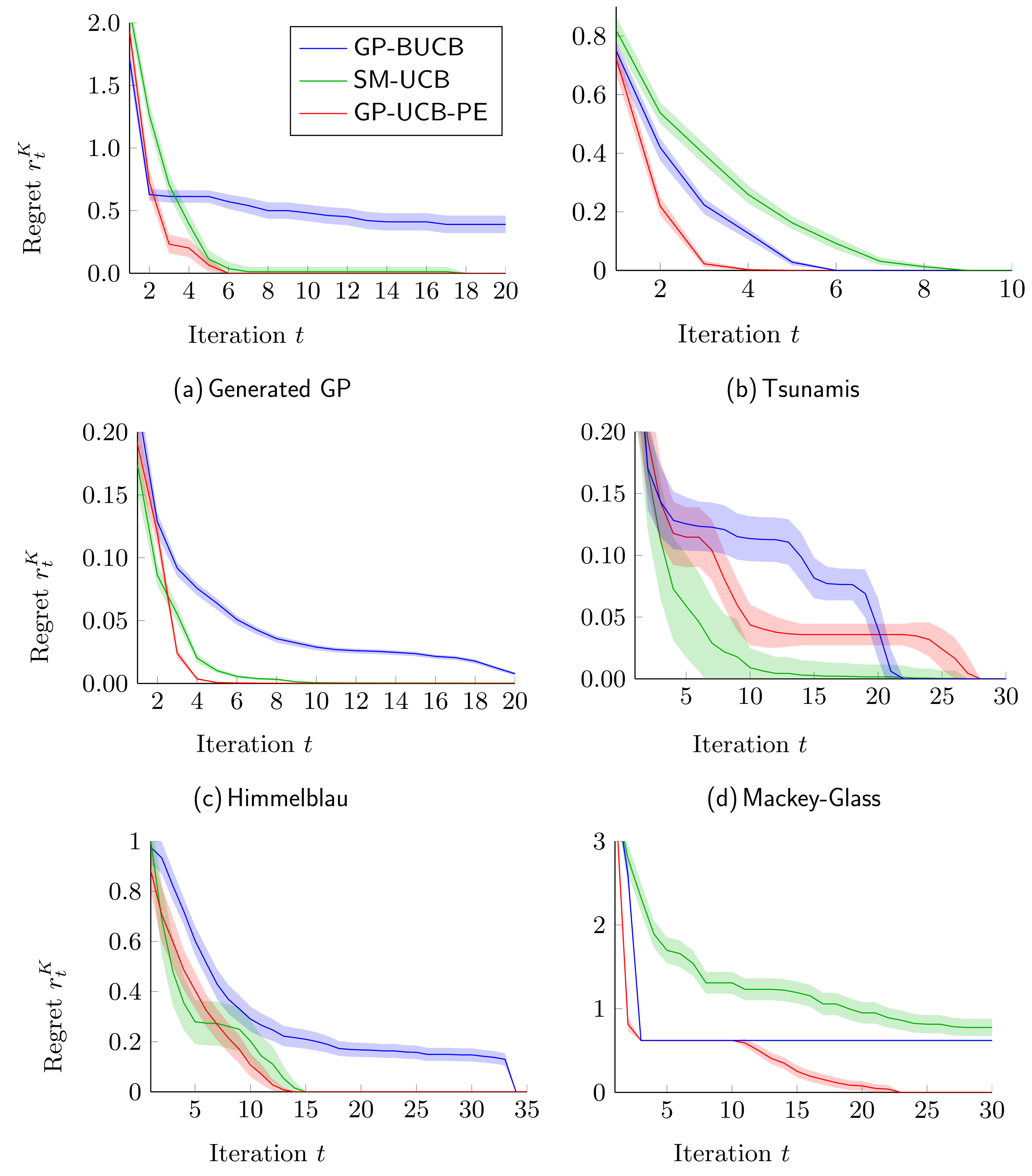
$\forall \delta > 0$, set $\beta_t = \mathcal{O}(\log \frac{t}{\delta})$ defined in Contal et al. (2013), with $f \sim \mathcal{GP}(0, k)$ and $k(x, x) \leq 1$, we have :

$$\Pr\left[R_T^K \leq \sqrt{C_1 \frac{T}{K} \beta_T \gamma_{TK} + C_2}\right] \geq 1 - \delta$$

$$\text{and } \Pr\left[R_{TK} \leq \sqrt{C_1 T K \beta_T \gamma_{TK} + C_2}\right] \geq 1 - \delta ,$$

where $C_1 = \frac{36}{\log(1+\sigma^{-2})}$ and $C_2 = \frac{\pi}{\sqrt{6}}$.

Mean batch regret and confidence interval over 64 runs.



(a) Generated GP   (b) Tsunamis
(c) Himmelblau   (d) Mackey-Glass
(e) Gaussian mixture   (f) Abalone

Discussion

- When $K \ll T$, GP-UCB-PE improves the upper bounds for R_T^K by an order of $\sqrt{K}$ compared to the sequential algorithm GP-UCB (Srinivas et al., 2012).
- The upper bounds for R_{TK} are equivalent compared to GP-UCB.
- Compared to GP-BUCB (Desautels et al., 2012), there is no need for an initialization phase. The improvement can be doubly exponential in the dimension d.

European Conference On Machine Learning And Principles And Practice Of Knowledge Discovery In Databases,   Prague, 23-27 September 2013