Gaussian Process Optimization with Mutual Information

Emile Contal CONTAL@CMLA.ENS-CACHAN.FR
CMLA, UMR CNRS 8536, ENS Cachan, France

Vianney Perchet VIANNEY.PERCHET@NORMALESUP.ORG
LPMA, Université Paris Diderot, France

Nicolas Vayatis VAYATIS@CMLA.ENS-CACHAN.FR
CMLA, UMR CNRS 8536, ENS Cachan, France

Abstract

In this paper, we analyze a generic algorithm scheme for sequential global optimization using Gaussian processes. The upper bounds we derive on the cumulative regret for this generic algorithm improve by an exponential factor the previously known bounds for algorithms like GP-UCB. We also introduce the novel Gaussian Process Mutual Information algorithm (GP-MI), which significantly improves further these upper bounds for the cumulative regret. We confirm the efficiency of this algorithm on synthetic and real tasks against the natural competitor, GP-UCB, and also the Expected Improvement heuristic.

1. Introduction

Stochastic optimization problems are encountered in numerous real world domains including engineering design (Wang & Shan, 2007), finance (Ziemba & Vickson, 2006), natural sciences (Floudas & Pardalos, 2000), or in machine learning for selecting models by tuning the parameters of learning algorithms (Snoek et al., 2012). We aim at finding the input of a given system which optimizes the output (or reward). In this view, an iterative procedure uses the previously acquired measures to choose the next query predicted to be the most useful. The goal is to maximize the sum of the rewards received at each iteration, that is to minimize the cumulative regret by balancing exploration (gathering information by favoring locations with high uncertainty) and exploitation (focusing on the optimum by favoring locations with high predicted reward). This optimization task becomes challenging when the dimension of the search space is high and the evaluations are noisy and expensive. Efficient algorithms have been studied to tackle this challenge such as multiarmed bandit (Auer et al., 2002; Kleinberg, 2004; Bubeck et al., 2011; Audibert et al., 2011), active learning (Carpentier et al., 2011; Chen & Krause, 2013) or Bayesian optimization (Mockus, 1989; Grunewalder et al., 2010; Srinivas et al., 2012; de Freitas et al., 2012). The theoretical analysis of such optimization procedures requires some prior assumptions on the underlying function f. Modeling f as a function distributed from a Gaussian process (GP) enforces near-by locations to have close associated values, and allows to control the general smoothness of f with high probability according to the kernel of the GP (Rasmussen & Williams, 2006). Our main contribution is twofold: we propose a generic algorithm scheme for Gaussian process optimization and we prove sharp upper bounds on its cumulative regret. The theoretical analysis has a direct impact on strategies built with the GP-UCB algorithm (Srinivas et al., 2012) such as (Krause & Ong, 2011; Desautels et al., 2012; Contal et al., 2013). We suggest an alternative policy which achieves an exponential speed up with respect to the cumulative regret. We also introduce a novel algorithm, the Gaussian Process Mutual Information algorithm (GP-MI), which improves furthermore upper bounds for the cumulative regret from $\mathcal{O}(\sqrt{T(\log T)^{d+1}})$ for GP-UCB, the current state of the art, to the spectacular $\mathcal{O}(\sqrt{(\log T)^{d+1}})$, where T is the number of iterations, d is the dimension of the input space and the kernel function is Gaussian. The remainder of this article is organized as follows. We first introduce the setup and notations. We define the GP-MI algorithm in Section 2. Main results on the cumulative regret bounds are presented in Section 3. We then provide technical details in Section 4. We finally confirm the performances of GP-MI on real and synthetic tasks compared to the state of the art of GP optimization and some heuristics used in practice.

Proceedings of the 31^{st} International Conference on Machine Learning, Beijing, China, 2014. JMLR: W&CP volume 32. Copyright 2014 by the author(s).

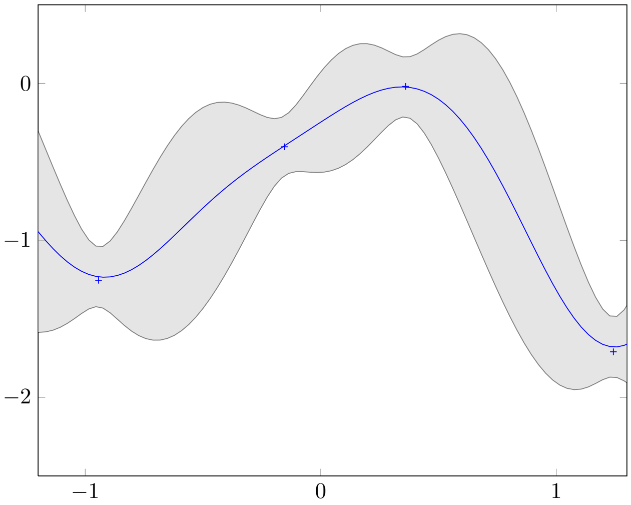


Figure 1. One dimensional Gaussian process inference of the posterior mean μ (blue line) and posterior deviation σ (half of the height of the gray envelope) with squared exponential kernel, based on four observations (blue crosses).

2. Gaussian process optimization and the GP-MI algorithm

2.1. Sequential optimization and cumulative regret

Let $f : \mathcal{X} \to \mathbb{R}$, where $\mathcal{X} \subset \mathbb{R}^d$ is a compact and convex set, be the unknown function modeling the system we want to be optimized. We consider the problem of finding the maximum of f denoted by:

$$f(x^\star) = \max_{x \in \mathcal{X}} f(x)\,,$$

via successive queries $x_1, x_2, \dots \in \mathcal{X}$. At iteration $T+1$, the choice of the next query x_{T+1} depends on the previous noisy observations, $\mathbf{Y}_T = \{y_1, \dots, y_T\}$ at locations $X_T = \{x_1, \dots, x_T\}$ where $y_t = f(x_t) + \epsilon_t$ for all $t \leq T$, and the noise variables $\epsilon_1, \dots, \epsilon_T$ are independently distributed as a Gaussian random variable $\mathcal{N}(0, \sigma^2)$ with zero mean and variance σ^2. The efficiency of a policy and its ability to address the exploration/exploitation trade-off is measured via the cumulative regret R_T, defined as the sum of the instantaneous regret r_t, the gaps between the value of the maximum and the values at the sample locations,

$$r_t = f(x^\star) - f(x_t) \text{ for } t \leq T$$
$$\text{and } R_T = \sum_{t=1}^{T} f(x^\star) - f(x_t)\,.$$

Our aim is to obtain upper bounds on the cumulative regret R_T with high probability.

2.2. The Gaussian process framework

Prior assumption on f. In order to control the smoothness of the underlying function, we assume that f is sampled from a Gaussian process $\mathcal{GP}(m, k)$ with mean function $m : \mathcal{X} \to \mathbb{R}$ and kernel function $k : \mathcal{X} \times \mathcal{X} \to \mathbb{R}^+$. We formalize in this manner the prior assumption that high local variations of f have low probability. The prior mean function is considered without loss of generality to be zero, as the kernel k can completely define the GP (Rasmussen & Williams, 2006). We consider the normalized and dimensionless framework introduced by (Srinivas et al., 2010) where the variance is assumed to be bounded, that is $k(x, x) \leq 1$ for all $x \in \mathcal{X}$.

Bayesian inference. At iteration $T+1$, given the previously observed noisy values $\mathbf{Y}_T$ at locations X_T, we use Bayesian inference to compute the current posterior distribution (Rasmussen & Williams, 2006), which is a GP of mean μ_{T+1} and variance σ^2_{T+1} given at any $x \in \mathcal{X}$ by,

$$\mu_{T+1}(x) = \mathbf{k}_T(x)^\top \mathbf{C}_T^{-1} \mathbf{Y}_T \qquad (1)$$
$$\text{and } \sigma^2_{T+1}(x) = k(x, x) - \mathbf{k}_T(x)^\top \mathbf{C}_T^{-1} \mathbf{k}_T(x)\,, \qquad (2)$$

where $\mathbf{k}_T(x) = [k(x_t, x)]_{x_t \in X_T}$ is the vector of covariances between x and the query points at time T, and $\mathbf{C}_T = \mathbf{K}_T + \sigma^2 \mathbf{I}$ with $\mathbf{K}_T = [k(x_t, x_{t'})]_{x_t, x_{t'} \in X_T}$ the kernel matrix, σ^2 is the variance of the noise and $\mathbf{I}$ stands for the identity matrix. The Bayesian inference is illustrated on Figure 1 in a sample problem in dimension one, where the posteriors are based on four observations of a Gaussian Process with squared exponential kernel. The height of the gray area represents two posterior standard deviations at each point.

2.3. The Gaussian Process Mutual Information algorithm

The Gaussian Process Mutual Information algorithm (GP-MI) is presented as Algorithm 1. The key statement is the choice of the query, $x_t = \operatorname{argmax}_{x \in \mathcal{X}} \mu_t(x) + \phi_t(x)$. The exploitation ability of the procedure is driven by μ_t, while the exploration is governed by $\phi_t : \mathcal{X} \to \mathbb{R}$, which is an increasing function of $\sigma_t^2(x)$. The novelty in the GP-MI algorithm is that ϕ_t is empirically controlled by the amount of exploration that has been already done, that is, the more the algorithm has gathered information on f, the more it will focus on the optimum. In the GP-UCB algorithm from (Srinivas et al., 2012) the exploration coefficient is a $\mathcal{O}(\log t)$ and therefore tends to infinity. The parameter α in Algorithm 1 governs the trade-off between precision and confidence, as shown in Theorem 2. The efficiency of the algorithm is robust to the choice of its value. We confirm empirically this property and provide further discussion on the calibration of α in Section 5.

Algorithm 1 GP-MI

```
$\widehat{\gamma}_0 \leftarrow 0$
for $t = 1, 2, \ldots$ do
    // Bayesian inference (Eq. 1, 2)
    Compute $\mu_t$ and $\sigma_t^2$
    // Definition of $\phi_t(x)$ for all $x \in \mathcal{X}$
    $\phi_t(x) \leftarrow \sqrt{\alpha}\left(\sqrt{\sigma_t^2(x) + \widehat{\gamma}_{t-1}} - \sqrt{\widehat{\gamma}_{t-1}}\right)$
    // Selection of the next query location
    $x_t \leftarrow \operatorname{argmax}_{x \in \mathcal{X}} \mu_t(x) + \phi_t(x)$
    // Update $\widehat{\gamma}_t$
    $\widehat{\gamma}_t \leftarrow \widehat{\gamma}_{t-1} + \sigma_t^2(x_t)$
    // Query
    Sample $x_t$ and observe $y_t$
end for
```

Mutual information. The quantity $\widehat{\gamma}_T$ controlling the exploration in our algorithm forms a lower bound on the information acquired on f by the query points X_T. The information on f is formally defined by $I_T(X_T)$, the mutual information between f and the noisy observations $\mathbf{Y}_T$ at X_T, hence the name of the GP-MI algorithm. For a Gaussian process distribution $I_T(X_T) = \frac{1}{2}\log\det(\mathbf{I} + \sigma^{-2}\mathbf{K}_T)$ where $\mathbf{K}_T$ is the kernel matrix $[k(x_i, x_j)]_{x_i, x_j \in X_T}$. We refer to (Cover & Thomas, 1991) for further reading on mutual information. We denote by $\gamma_T = \max_{X_T \subset \mathcal{X} : |X_T| = T} I_T(X_T)$ the maximum mutual information obtainable by a sequence of T query points. In the case of Gaussian processes with bounded variance, the following inequality is satisfied (Lemma 5.4 in (Srinivas et al., 2012)):

$$\widehat{\gamma}_T = \sum_{t=1}^{T} \sigma_t^2(x_t) \leq C_1 \gamma_T \tag{3}$$

where $C_1 = \frac{2}{\log(1+\sigma^{-2})}$ and σ^2 is the noise variance. The upper bounds on the cumulative regret we derive in the next section depend mainly on this key quantity.

3. Main Results

3.1. Generic Optimization Scheme

We first consider the generic optimization scheme defined in Algorithm 2, where we let ϕ_t as a generic function viewed as a parameter of the algorithm. We only require ϕ_t to be measurable with respect to $\mathbf{Y}_{t-1}$, the observations available at iteration $t-1$. The theoretical analysis of Algorithm 2 can be used as a plug-in theorem for existing algorithms. For example the GP-UCB algorithm with parameter $\beta_t = \mathcal{O}(\log t)$ is obtained with $\phi_t(x) = \sqrt{\beta_t \sigma_t^2(x)}$. A generic analysis of Algorithm 2 leads to the following upper bounds on the cumulative regret with high probability.

Algorithm 2 Generic Optimization Scheme (ϕ_t)

```
for $t = 1, 2, \ldots$ do
    Compute $\mu_t$ and $\phi_t$
    $x_t \leftarrow \operatorname{argmax}_{x \in \mathcal{X}} \mu_t(x) + \phi_t(x)$
    Sample $x_t$ and observe $y_t$
end for
```

Theorem 1 (Regret bounds for the generic algorithm)**.** *For all $\delta > 0$ and $T > 0$, the regret R_T incurred by Algorithm 2 on f distributed as a GP perturbed by independent Gaussian noise with variance σ^2 satisfies the following bound with high probability, with $C_1 = \frac{2}{\log(1+\sigma^{-2})}$ and $\alpha = \log\frac{2}{\delta}$:*

$$\Pr\left[R_T \leq \sum_{t=1}^{T}\left(\phi_t(x_t) - \phi_t(x^\star)\right) + 4\sqrt{\alpha(C_1\gamma_T + 1)} + \frac{\sqrt{\alpha}}{2}\frac{\sum_{t=1}^{T}\sigma_t^2(x^\star)}{\sqrt{C_1\gamma_T + 1}}\right] \geq 1 - \delta\,.$$

The proof of Theorem 1, relies on concentration guarantees for Gaussian processes (see Section 4.1). Theorem 1 provides an intermediate result used for the calibration of ϕ_t to face the exploration/exploitation trade-off. For example by choosing $\phi_t(x) = \frac{\sqrt{\alpha}}{2}\sigma_t^2(x)$ (where the dimensional constant is hidden), Algorithm 2 becomes a variant of the GP-UCB algorithm where in particular the exploration parameter $\sqrt{\beta_t}$ is fixed to $\frac{\sqrt{\alpha}}{2}$ instead of being an increasing function of t. The upper bounds on the cumulative regret with this definition of ϕ_t are of the form $R_T = \mathcal{O}(\gamma_T)$, as stated in Corollary 1. We then also consider the case where the kernel k of the Gaussian process is under the form of a squared exponential (RBF) kernel, $k(x_1, x_2) = \exp\left(-\frac{\|x_1 - x_2\|^2}{2l^2}\right)$, for all $x_1, x_2 \in \mathcal{X}$ and length scale $l \in \mathbb{R}$. In this setting, the maximum mutual information γ_T satisfies the upper bound $\gamma_T = \mathcal{O}((\log T)^{d+1})$, where d is the dimension of the input space (Srinivas et al., 2012).

Corollary 1. *Consider the Algorithm 2 where we set $\phi_t(x) = \frac{\sqrt{\alpha}}{2}\sigma_t^2(x)$. Under the assumptions of Theorem 1, we have that the cumulative regret for Algorithm 2 satisfies the following upper bounds with high probability:*

- *For f sampled from a GP with general kernel:* $R_T = \mathcal{O}(\gamma_T)$.
- *For f sampled from a GP with RBF kernel:* $R_T = \mathcal{O}((\log T)^{d+1})$.

To prove Corollary 1 we apply Theorem 1 with the given definition of ϕ_t and then Equation 3, which leads to $\Pr\left[R_T \leq \frac{\sqrt{\alpha}}{2}C_1\gamma_T + 4\sqrt{\alpha(C_1\gamma_T + 1)}\right] \geq 1 - \delta$. The previously known upper bounds on the cumulative regret for the GP-UCB algorithm are of the form $R_T = \mathcal{O}(\sqrt{T\beta_T\gamma_T})$ where $\beta_T = \mathcal{O}(\log\frac{T}{\delta})$. The improvement of the generic Algorithm 2 with $\phi_t(x) = \frac{\sqrt{\alpha}}{2}\sigma_t^2(x)$ over the GP-UCB algorithm with respect to the cumulative regret is then exponential in the case of Gaussian processes with RBF kernel. For f sampled from a GP with linear kernel, corresponding to $f(x) = w^T x$ with $w \sim \mathcal{N}(0, \mathbf{I})$, we obtain $R_T = \mathcal{O}(d \log T)$. We remark that the GP assumption with linear kernel is more restrictive than the linear bandit framework, as it implies a Gaussian prior over the linear coefficients w. Hence there is no contradiction with the lower bounds stated for linear bandit like those of (Dani et al., 2008). We refer to (Srinivas et al., 2012) for the analysis of γ_T with other kernels widely used in practice.

3.2. Regret bounds for the GP-MI algorithm

We present here the main result of the paper, the upper bound on the cumulative regret for the GP-MI algorithm.

Theorem 2 (Regret bounds for the GP-MI algorithm)**.** *For all $\delta > 0$ and $T > 1$, the regret R_T incurred by Algorithm 1 on f distributed as a GP perturbed by independent Gaussian noise with variance σ^2 satisfies the following bound with high probability, with $C_1 = \frac{2}{\log(1+\sigma^{-2})}$ and $\alpha = \log\frac{2}{\delta}$:*

$$\Pr\left[R_T \leq 5\sqrt{\alpha C_1 \gamma_T} + 4\sqrt{\alpha}\right] \geq 1 - \delta\,.$$

The proof for Theorem 2 is provided in Section 4.2, where we analyze the properties of the exploration functions ϕ_t. Corollary 2 describes the case with RBF kernel for the GP-MI algorithm.

Corollary 2 (RBF kernels)**.** *The cumulative regret R_T incurred by Algorithm 1 on f sampled from a GP with RBF kernel satisfies with high probability,*

$$R_T = \mathcal{O}\Big((\log T)^{\frac{d+1}{2}}\Big)\,.$$

The GP-MI algorithm significantly improves the upper bounds for the cumulative regret over the GP-UCB algorithm and the alternative policy of Corollary 1.

4. Theoretical analysis

In this section, we provide the proofs for Theorem 1 and Theorem 2. The approach presented here to study the cumulative regret incurred by Gaussian process optimization strategies is general and can be used further for other algorithms.

4.1. Analysis of the general algorithm

The theoretical analysis of Theorem 1 uses a similar approach to the Azuma-Hoeffding inequality adapted for Gaussian processes. Let $r_t = f(x^\star) - f(x_t)$ for all $t \leq T$. We define M_T, which is shown later to be a martingale with respect to $\mathbf{Y}_{T-1}$,

$$M_T = \sum_{t=1}^{T}\Big(r_t - \big(\mu_t(x^\star) - \mu_t(x_t)\big)\Big)\,, \tag{4}$$

for $T \geq 1$ and $M_0 = 0$. Let Y_t be defined as the martingale difference sequence with respect to M_T, that is the difference between the instantaneous regret and the gap between the posterior mean for the optimum and the one for the point queried,

$$Y_t = M_t - M_{t-1} = r_t - \big(\mu_t(x^\star) - \mu_t(x_t)\big) \text{ for } t \geq 1\,.$$

Lemma 1. *The sequence M_T is a martingale with respect to $\mathbf{Y}_{T-1}$ and for all $t \leq T$, given $\mathbf{Y}_{t-1}$, the random variable Y_t is distributed as a Gaussian $\mathcal{N}(0, \ell_t^2)$ with zero mean and variance ℓ_t^2, where:*

$$\ell_t^2 = \sigma_t^2(x^\star) + \sigma_t^2(x_t) - 2k(x^\star, x_t)\,. \tag{5}$$

Proof. From the GP assumption, we know that given $\mathbf{Y}_{t-1}$, the distribution of $f(x)$ is Gaussian $\mathcal{N}(\mu_t(x), \sigma_t^2(x))$ for all $x \in \mathcal{X}$, and r_t is a projection of a Gaussian random vector, that is r_t is distributed as a Gaussian $\mathcal{N}(\mu_t(x^\star) - \mu_t(x_t), \ell_t^2)$ and Y_t is distributed as Gaussian $\mathcal{N}(0, \ell_t^2)$, with $\ell_t^2 = \sigma_t^2(x^\star) + \sigma_t^2(x_t) - 2k(x^\star, x_t)$, hence M_T is a Gaussian martingale. □

We now give a concentration result for M_T using inequalities for self-normalized martingales.

Lemma 2. *For all $\delta > 0$ and $T > 1$, the martingale M_T normalized by the predictable quadratic variation $\sum_{t=1}^T \ell_t^2$ satisfies the following concentration inequality with $\alpha = \log\frac{2}{\delta}$ and $y = 8(C_1\gamma_T + 1)$:*

$$\Pr\left[M_T \leq \sqrt{2\alpha y} + \sqrt{\frac{2\alpha}{y}}\sum_{t=1}^T \sigma_t^2(x^\star)\right] \geq 1 - \delta\,.$$

Proof. Let $y = 8(C_1\gamma_T + 1)$. We introduce the notation $\Pr_>[A]$ for $\Pr[A \wedge \sum_{t=1}^T \ell_t^2 > y]$ and $\Pr_\leq[A]$ for $\Pr[A \wedge \sum_{t=1}^T \ell_t^2 \leq y]$. Given that M_t is a Gaussian martingale, using Theorem 4.2 and Remark 4.2 from (Bercu & Touati, 2008) with $\langle M\rangle_T = \sum_{t=1}^T \ell_t^2$ and $a = 0$ and $b = 1$ we obtain for all $x > 0$:

$$\Pr_>\left[\frac{M_T}{\sum_{t=1}^T \ell_t^2} > x\right] < \exp\left(-\frac{x^2 y}{2}\right)\,.$$

With $x = \sqrt{\frac{2\alpha}{y}}$ where $\alpha = \log\frac{2}{\delta}$, we have:

$$\Pr_>\left[M_T > \sqrt{\frac{2\alpha}{y}}\sum_{t=1}^T \ell_t^2\right] < \frac{\delta}{2}\,.$$

By definition of ℓ_t in Eq. 5 and with $k(x^\star, x_t) \geq 0$, we have for all $t \geq 1$ that $\ell_t^2 \leq \sigma_t^2(x_t) + \sigma_t^2(x^\star)$. Using Equation 3 we have $\widehat{\gamma}_T \leq \frac{y}{8}$, we finally get:

$$\Pr_{>}\left[M_T > \frac{\sqrt{2\alpha y}}{8} + \sqrt{\frac{2\alpha}{y}} \sum_{t=1}^T \sigma_t^2(x^\star)\right] < \frac{\delta}{2}\,. \tag{6}$$

Now, using Theorem 4.1 and Remark 4.2 from (Bercu & Touati, 2008) the following inequality is satisfied for all $x > 0$:

$$\Pr_{\leq}\left[M_T > x\right] < \exp\left(-\frac{x^2}{2y}\right)\,.$$

With $x = \sqrt{2\alpha y}$ we have:

$$\Pr_{\leq}\left[M_T > \sqrt{2\alpha y}\right] < \frac{\delta}{2}\,. \tag{7}$$

Combining Equations 6 and 7 leads to,

$$\Pr\left[M_T > \sqrt{2\alpha y} + \sqrt{\frac{2\alpha}{y}} \sum_{t=1}^T \sigma_t^2(x^\star)\right] < \delta\,,$$

proving Lemma 2. □

The following lemma concludes the proof of Theorem 1 using the previous concentration result and the properties of the generic Algorithm 2.

Lemma 3. *The cumulative regret for Algorithm 2 on f sampled from a GP satisfies the following bound for all $\delta > 0$ and α and y defined in Lemma 2:*

$$\Pr\Bigg[R_T \leq \sum_{t=1}^T \big(\phi_t(x_t) - \phi_t(x^\star)\big) + \sqrt{2\alpha y} + \sqrt{\frac{2\alpha}{y}} \sum_{t=1}^T \sigma_t^2(x^\star)\Bigg] \geq 1 - \delta\,.$$

Proof. By construction of the generic Algorithm 2, we have $x_t = \operatorname{argmax}_{x \in \mathcal{X}} \mu_t(x) + \phi_t(x)$, which guarantees for all $t \geq 1$ that $\mu_t(x^\star) - \mu_t(x_t) \leq \phi_t(x_t) - \phi_t(x^\star)$. Replacing M_T by its definition in Eq. 4 and using the previous property in Lemma 2 proves Lemma 3. □

4.2. Analysis of the GP-MI algorithm

In order to bound the cumulative regret for the GP-MI algorithm, we focus on an alternative definition of the exploration functions ϕ_t where the last term is modified inductively so as to simplify the sum $\sum_{t=1}^T \phi_t(x_t)$ for all $T > 0$. Being a constant term for a fixed $t > 0$, Algorithm 1 remains unchanged. Let ϕ_t be defined as,

$$\phi_t(x) = \sqrt{\alpha\big(\sigma_t^2(x) + \widehat{\gamma}_{t-1}\big)} - \sum_{i=1}^{t-1} \phi_i(x_i)\,,$$

where x_t is the point selected by Algorithm 1 at iteration t. We have for all $T > 1$,

$$\begin{aligned}\sum_{t=1}^T \phi_t(x_t) &= \sqrt{\alpha \widehat{\gamma}_T} - \sum_{t=1}^{T-1} \phi_t(x_t) + \sum_{t=1}^{T-1} \phi_t(x_t)\\ &= \sqrt{\alpha \widehat{\gamma}_T}\,.\end{aligned} \tag{8}$$

We can now derive upper bounds for $\sum_{t=1}^T \big(\phi_t(x_t) - \phi_t(x^\star)\big)$ which will be plugged in Theorem 1 in order to cancel out the terms involving $x^\star$. In this manner we can calibrate sharply the exploration/exploitation trade-off by optimizing the remaining terms.

Lemma 4. *For the GP-MI algorithm, the exploration term in the equation of Theorem 1 satisfies the following inequality:*

$$\sum_{t=1}^T \big(\phi_t(x_t) - \phi_t(x^\star)\big) \leq \sqrt{\alpha \widehat{\gamma}_T} - \frac{\sqrt{\alpha}}{2} \frac{\sum_{t=1}^T \sigma_t^2(x^\star)}{\sqrt{\widehat{\gamma}_T + 1}}\,.$$

Proof. Using our alternative definition of ϕ_t which gives the equality stated in Equation 8, we know that,

$$\sum_{t=1}^T \big(\phi_t(x_t) - \phi_t(x^\star)\big) = \sqrt{\alpha}\left(\sqrt{\widehat{\gamma}_T} + \sum_{t=1}^T \left(\sqrt{\widehat{\gamma}_{t-1}} - \sqrt{\widehat{\gamma}_{t-1} + \sigma_t^2(x^\star)}\right)\right)\,.$$

By concavity of the square root, we have for all $a \geq -b$ that $\sqrt{a+b} - \sqrt{a} \leq \frac{b}{2\sqrt{a}}$. Introducing the notations $a_t = \widehat{\gamma}_{t-1} + \sigma_t^2(x^\star)$ and $b_t = -\sigma_t^2(x^\star)$, we obtain,

$$\sum_{t=1}^T \big(\phi_t(x_t) - \phi_t(x^\star)\big) \leq \sqrt{\alpha \widehat{\gamma}_T} + \frac{\sqrt{\alpha}}{2} \sum_{t=1}^T \frac{b_t}{\sqrt{a_t}}\,.$$

Moreover, with $0 \leq \sigma_t^2(x) \leq 1$ for all $x \in \mathcal{X}$, we have $a_t \leq \widehat{\gamma}_T + 1$ and $b_t \leq 0$ for all $t \leq T$ which gives,

$$\sum_{t=1}^T \frac{b_t}{\sqrt{a_t}} \leq -\frac{\sum_{t=1}^T \sigma_t^2(x^\star)}{\sqrt{\widehat{\gamma}_T + 1}}\,,$$

leading to the inequality of Lemma 4. □

The following lemma combines the results from Theorem 1 and Lemma 4 to derive upper bounds on the cumulative regret for the GP-MI algorithm with high probability.

Lemma 5. *The cumulative regret for Algorithm 1 on f sampled from a GP satisfies the following bound for all $\delta > 0$ and α defined in Lemma 2,*

$$\Pr\left[R_T \leq 5\sqrt{\alpha C_1 \gamma_T} + 4\sqrt{\alpha}\right] \geq 1 - \delta\,.$$

Proof. Considering Theorem 1 in the case of the GP-MI algorithm and bounding $\sum_{t=1}^{T}\left(\phi_t(x_t) - \phi_t(x^\star)\right)$ with Lemma 4, we obtain the following bound on the cumulative regret incurred by GP-MI:

$$\Pr\Bigg[R_T \leq \sqrt{\alpha\widehat{\gamma}_T} - \frac{\sqrt{\alpha}}{2}\frac{\sum_{t=1}^{T}\sigma_t^2(x^\star)}{\sqrt{\widehat{\gamma}_T+1}} + 4\sqrt{\alpha(C_1\gamma_T+1)} + \frac{\sqrt{\alpha}}{2}\frac{\sum_{t=1}^{T}\sigma_t^2(x^\star)}{\sqrt{C_1\gamma_T+1}}\Bigg] \geq 1-\delta\,,$$

which simplifies to the inequality of Lemma 5 using Equation 3, and thus proves Theorem 2. □

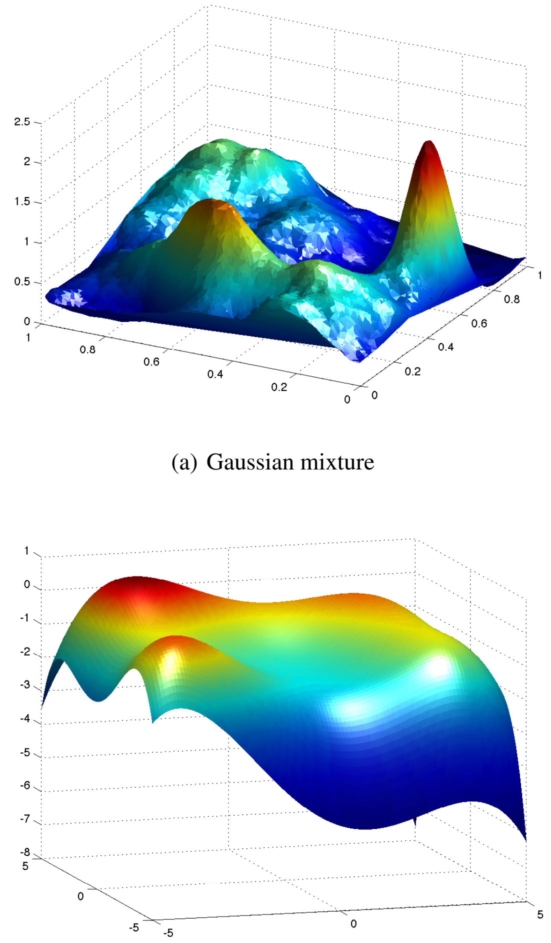
5. Practical considerations and experiments

5.1. Numerical experiments

Protocol. We compare the empirical performances of our algorithm against the state-of-the-art of GP optimization, the GP-UCB algorithm (Srinivas et al., 2012), and a commonly used heuristic, the Expected Improvement (EI) algorithm with GP (Jones et al., 1998). The tasks used for assessment come from two real applications and five synthetic problems described here. For all data sets and algorithms the learners were initialized with a random subset of 10 observations $\{(x_i, y_i)\}_{i\leq 10}$. When the prior distribution of the underlying function was not known, the Bayesian inference was made using a squared exponential kernel. We first picked the half of the data set to estimate the hyperparameters of the kernel via cross validation in this subset. In this way, each algorithm was running with the same prior information. The value of the parameter δ for the GP-MI and the GP-UCB algorithms was fixed to $\delta = 10^{-6}$ for all these experimental tasks. Modifying this value by several orders of magnitude is insignificant with respect to the empirical mean cumulative regret incurred by the algorithms, as discussed in Section 5.2. The results are provided in Figure 3. The curves show the evolution of the average regret $\frac{R_T}{T}$ in term of iteration T. We report the mean value with the confidence interval over a hundred experiments.

Description of the data sets. We describe briefly all the data sets used for assessment.

- *Generated GP.* The generated Gaussian process functions are random GPs drawn from an isotropic Matérn kernel in dimension 2 and 4, with the kernel bandwidth set to 1 for dimension 2, and 16 for dimension 4. The Matérn parameter was set to $\nu = 3$ and the noise standard deviation to 1% of the signal standard deviation.
- *Gaussian Mixture.* This synthetic function comes from the addition of three 2-D Gaussian functions. We then perturb these Gaussian functions with smooth variations generated from a Gaussian Process with isotropic Matérn Kernel and 1% of noise. It is shown on Figure 2(a). The highest peak being thin, the sequential search for the maximum of this function is quite challenging.



(a) Gaussian mixture

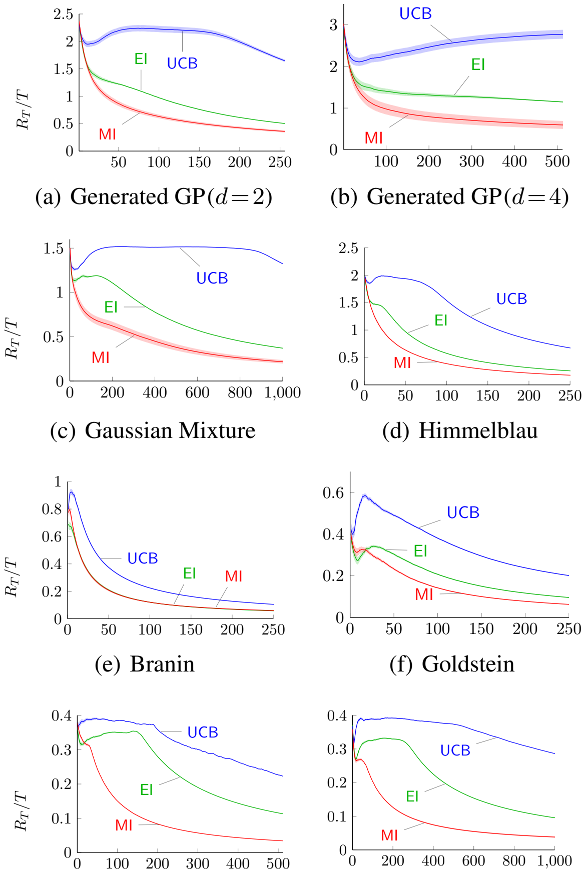
(b) Himmelblau

Figure 2. Visualization of the synthetic functions used for assessment

- *Himmelblau.* This task is another synthetic function in dimension 2. We compute a slightly tilted version of the Himmelblau's function with the addition of a linear function, and take the opposite to match the challenge of finding its maximum. This function presents four peaks but only one global maximum. It gives a practical way to test the ability of a strategy to manage exploration/exploitation trade-offs. It is represented in Figure 2(b).

- *Branin.* The Branin or Branin-Hoo function is a common benchmark function for global optimization. It presents three global optimum in the 2-D square $[-5, 10] \times [0, 15]$. This benchmark is one of the two synthetic functions used by (Srinivas et al., 2012) to evaluate the empirical performances of the GP-UCB algorithm. No noise has been added to the original signal in this experimental task.
- *Goldstein-Price.* The Goldstein & Price function is an other benchmark function for global optimization, with a single global optimum but several local optima in the 2-D square $[-2, 2] \times [-2, 2]$. This is the second synthetic benchmark used by (Srinivas et al., 2012). Like in the previous challenge, no noise has been added to the original signal.
- *Tsunamis.* Recent post-tsunami survey data as well as the numerical simulations of (Hill et al., 2012) have shown that in some cases the run-up, which is the maximum vertical extent of wave climbing on a beach, in areas which were supposed to be protected by small islands in the vicinity of coast, was significantly higher than in neighboring locations. Motivated by these observations (Stefanakis et al., 2012) investigated this phenomenon by employing numerical simulations using the VOLNA code (Dutykh et al., 2011) with the simplified geometry of a conical island sitting on a flat surface in front of a sloping beach. In the study of (Stefanakis et al., 2013) the setup was controlled by five physical parameters and the aim was to find with confidence and with the least number of simulations the parameters leading to the maximum run-up amplification.
- *Mackey-Glass function.* The Mackey-Glass delay-differential equation is a chaotic system in dimension 6, but without noise. It models real feedback systems and is used in physiological domains such as hematology, cardiology, neurology, and psychiatry. The highly chaotic behavior of this function makes it an exceptionally difficult optimization problem. It has been used as a benchmark for example by (Flake & Lawrence, 2002).

Empirical comparison of the algorithms. Figure 3 compares the empirical mean average regret $\frac{R_T}{T}$ for the three algorithms. On the easy optimization assessments like the Branin data set (Fig. 3(e)) the three strategies behave in a comparable manner, but the GP-UCB algorithm incurs a larger cumulative regret. For more difficult assessments the GP-UCB algorithm performs poorly and our algorithm always surpasses the EI heuristic. The improvement of the GP-MI algorithm against the two competitors is the most significant for exceptionally challenging optimization tasks as illustrated in Figures 3(a) to 3(d) and 3(h), where the underlying functions present several local optima. The ability of our algorithm to deal with the exploration/exploitation trade-off is emphasized by these experimental results as its average regret decreases directly after the first iterations, avoiding unwanted exploration like GP-UCB on Figures 3(a) to 3(d), or getting stuck in some local optimum like EI on Figures 3(c), 3(g) and 3(h). We further mention that the GP-MI algorithm is empirically robust against the number of dimensions of the data set (Fig. 3(b), 3(g), 3(h)).



(a) Generated GP($d = 2$)

(b) Generated GP($d = 4$)

(c) Gaussian Mixture

(d) Himmelblau

(e) Branin

(f) Goldstein

(g) Tsunamis

(h) Mackey-Glass

Figure 3. Empirical mean and confidence interval of the average regret $\frac{R_T}{T}$ in term of iteration T on real and synthetic tasks for the GP-MI and GP-UCB algorithms and the EI heuristic (lower is better).

5.2. Practical aspects

Calibration of α. The value of the parameter α is chosen following Theorem 2 as $\alpha = \log \frac{2}{\delta}$ with $0 < \delta < 1$ being a confidence parameter. The guarantees we prove in Section 4.2 on the cumulative regret for the GP-MI algorithm holds with probability at least $1 - \delta$. With α increasing linearly for δ decreasing exponentially toward 0, the algorithm is

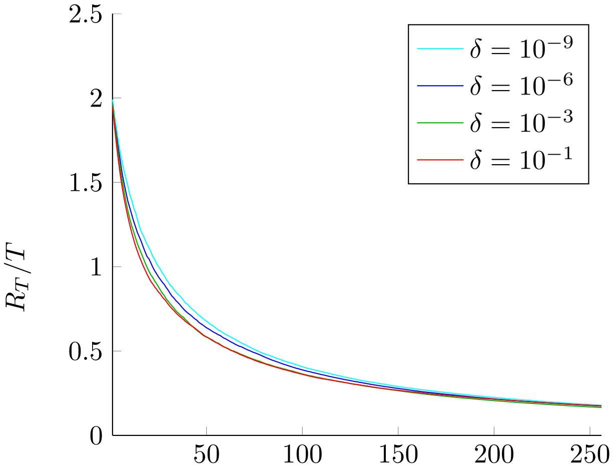


Figure 4. Small impact of the value of δ on the mean average regret of the GP-MI algorithm running on the Himmelblau data set.

robust to the choice of δ. We present on Figure 4 the small impact of δ on the average regret for four different values selected on a wide range.

Numerical Complexity. Even if the numerical cost of GP-MI is insignificant in practice compared to the cost of the evaluation of f, the complexity of the sequential Bayesian update (Osborne, 2010) is $\mathcal{O}(T^2)$ and might be prohibitive for large T. One can reduce drastically the computational time by means of Lazy Variance Calculation (Desautels et al., 2012), built on the fact that $\sigma_T^2(x)$ always decreases for increasing T and for all $x \in \mathcal{X}$. We further mention that approximated inference algorithms such as the EP approximation and MCMC sampling (Kuss et al., 2005) can be used as an alternative if the computational time is a restrictive factor.

6. Conclusion

We introduced the GP-MI algorithm for GP optimization and prove upper bounds on its cumulative regret which improve exponentially the state-of-the-art in common settings. The theoretical analysis was presented in a generic framework in order to expand its impact to other similar algorithms. The experiments we performed on real and synthetic assessments confirmed empirically the efficiency of our algorithm against both the theoretical state-of-the-art of GP optimization, the GP-UCB algorithm, and the commonly used EI heuristic.

Acknowledgements

The authors would like to thank David Buffoni and Raphaël Bonaque for fruitful discussions. The author also thank the anonymous reviewers for their detailed feedback.

References

Audibert, J-Y., Bubeck, S., and Munos, R. Bandit view on noisy optimization. In *Optimization for Machine Learning*, pp. 431–454. MIT Press, 2011.

Auer, P., Cesa-Bianchi, N., and Fischer, P. Finite-time analysis of the multiarmed bandit problem. *Machine Learning*, 47(2-3):235–256, 2002.

Bercu, B. and Touati, A. Exponential inequalities for self-normalized martingales with applications. *The Annals of Applied Probability*, 18(5):1848–1869, 2008.

Bubeck, S., Munos, R., Stoltz, G., and Szepesvári, C. X-armed bandits. *Journal of Machine Learning Research*, 12:1655–1695, 2011.

Carpentier, A., Lazaric, A., Ghavamzadeh, M., Munos, R., and Auer, P. Upper-confidence-bound algorithms for active learning in multi-armed bandits. In *Proceedings of the International Conference on Algorithmic Learning Theory*, pp. 189–203. Springer-Verlag, 2011.

Chen, Y. and Krause, A. Near-optimal batch mode active learning and adaptive submodular optimization. In *Proceedings of the International Conference on Machine Learning*. icml.cc / Omnipress, 2013.

Contal, E., Buffoni, D., Robicquet, A., and Vayatis, N. Parallel Gaussian process optimization with upper confidence bound and pure exploration. In *Machine Learning and Knowledge Discovery in Databases*, volume 8188, pp. 225–240. Springer Berlin Heidelberg, 2013.

Cover, T. M. and Thomas, J. A. *Elements of Information Theory*. Wiley-Interscience, 1991.

Dani, V., Hayes, T. P., and Kakade, S. M. Stochastic linear optimization under bandit feedback. In *Proceedings of the 21st Annual Conference on Learning Theory*, pp. 355–366, 2008.

de Freitas, N., Smola, A. J., and Zoghi, M. Exponential regret bounds for Gaussian process bandits with deterministic observations. In *Proceedings of the 29th International Conference on Machine Learning*. icml.cc / Omnipress, 2012.

Desautels, T., Krause, A., and Burdick, J.W. Parallelizing exploration-exploitation tradeoffs with Gaussian process bandit optimization. In *Proceedings of the 29th International Conference on Machine Learning*, pp. 1191–1198. icml.cc / Omnipress, 2012.

Dutykh, D., Poncet, R, and Dias, F. The VOLNA code for the numerical modelling of tsunami waves: generation, propagation and inundation. *European Journal of Mechanics B/Fluids*, 30:598–615, 2011.

Flake, G. W. and Lawrence, S. Efficient SVM regression training with SMO. *Machine Learning*, 46(1-3):271–290, 2002.

Floudas, C.A. and Pardalos, P.M. *Optimization in Computational Chemistry and Molecular Biology: Local and Global Approaches*. Nonconvex Optimization and Its Applications. Springer, 2000.

Grunewalder, S., Audibert, J-Y., Opper, M., and Shawe-Taylor, J. Regret bounds for Gaussian process bandit problems. In *Proceedings of the International Conference on Artificial Intelligence and Statistics*, pp. 273–280. MIT Press, 2010.

Hill, E. M., Borrero, J. C., Huang, Z., Qiu, Q., Banerjee, P., Natawidjaja, D. H., Elosegui, P., Fritz, H. M., Suwargadi, B. W., Pranantyo, I. R., Li, L., Macpherson, K. A., Skanavis, V., Synolakis, C. E., and Sieh, K. The 2010 Mw 7.8 Mentawai earthquake: Very shallow source of a rare tsunami earthquake determined from tsunami field survey and near-field GPS data. *J. Geophys. Res.*, 117: B06402–, 2012.

Jones, D. R., Schonlau, M., and Welch, W. J. Efficient global optimization of expensive black-box functions. *Journal of Global Optimization*, 13(4):455–492, December 1998.

Kleinberg, R. Nearly tight bounds for the continuum-armed bandit problem. In *Advances in Neural Information Processing Systems 17*, pp. 697–704. MIT Press, 2004.

Krause, A. and Ong, C.S. Contextual Gaussian process bandit optimization. In *Advances in Neural Information Processing Systems 24*, pp. 2447–2455, 2011.

Kuss, M., Pfingsten, T., Csató, L., and Rasmussen, C.E. Approximate inference for robust Gaussian process regression. *Max Planck Inst. Biological Cybern., Tubingen, GermanyTech. Rep*, 136, 2005.

Mockus, J. *Bayesian approach to global optimization*. Mathematics and its applications. Kluwer Academic, 1989.

Osborne, Michael. *Bayesian Gaussian processes for sequential prediction, optimisation and quadrature*. PhD thesis, Oxford University New College, 2010.

Rasmussen, C. E. and Williams, C. *Gaussian Processes for Machine Learning*. MIT Press, 2006.

Snoek, J., Larochelle, H., and Adams, R. P. Practical bayesian optimization of machine learning algorithms. In *Advances in Neural Information Processing Systems 25*, pp. 2960–2968, 2012.

Srinivas, N., Krause, A., Kakade, S., and Seeger, M. Gaussian process optimization in the bandit setting: No regret and experimental design. In *Proceedings of the International Conference on Machine Learning*, pp. 1015–1022. icml.cc / Omnipress, 2010.

Srinivas, N., Krause, A., Kakade, S., and Seeger, M. Information-theoretic regret bounds for Gaussian process optimization in the bandit setting. *IEEE Transactions on Information Theory*, 58(5):3250–3265, 2012.

Stefanakis, T. S., Dias, F., Vayatis, N., and Guillas, S. Long-wave runup on a plane beach behind a conical island. In *Proceedings of the World Conference on Earthquake Engineering*, 2012.

Stefanakis, T. S., Contal, E., Vayatis, N., Dias, F., and Synolakis, C. E. Can small islands protect nearby coasts from tsunamis ? An active experimental design approach. *arXiv preprint arXiv:1305.7385*, 2013.

Wang, G. and Shan, S. Review of metamodeling techniques in support of engineering design optimization. *Journal of Mechanical Design*, 129(4):370–380, 2007.

Ziemba, W. T. and Vickson, R. G. *Stochastic optimization models in finance*. World Scientific Singapore, 2006.