On correlation and budget constraints in model-based bandit optimization

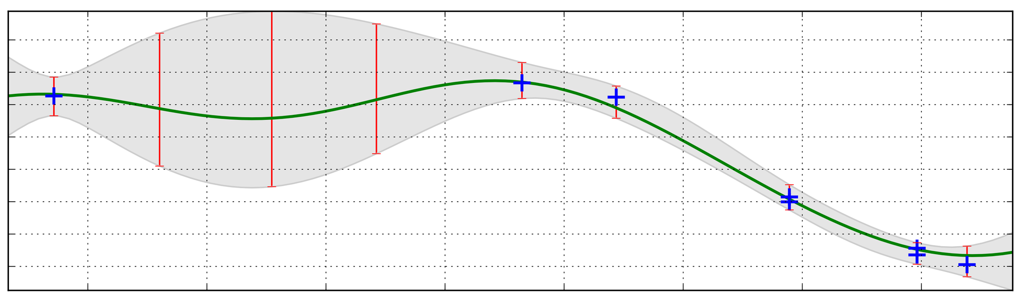
with application to automatic machine learning

Matthew W. Hoffman[†], Bobak Shahriari, Nando de Freitas[‡]

University Of British Columbia; and University of Cambridge[†], University of Oxford[‡]

Motivation

We address the problem of finding the **maximizer** of a **nonlinear smooth function** at **discrete points**



To properly do so we take into account:
- the **correlation** between different points;
- that we are only interested in finding an ϵ-**best point**;
- a limited query **budget**.

Our contribution

- first Bayesian approach for 3 criteria—better performance when all 3 matter!
- comprehensive (and first) comparison of many bandit (including best arm) and Bayesian optimization methods.

Problem formulation

We will consider a discrete set of *arms* $\mathcal{A} = \{1, \dots, k\}$ and horizon T.
- when arm a_t is selected at time t we observe $Y_t \sim \nu_k(\cdot)$;
- each arm has mean $\mu_k = \mathbb{E}_{\nu_k}[Y]$ and regret $R_k = \mu^* - \mu_k$;
- at time T we will make a single recommendation $\Omega_T \in \mathcal{A}$.

Goal: define $a_{1:T}$ and Ω_T such that with high probability Ω_T is the best arm

Bayesian bandits

We will let the arm models depend on unknown parameters, i.e. $\nu_k(\cdot|\theta)$; after $t-1$ rounds we have a **posterior** $\pi_t(\theta)$ which depends on the data seen. This induces a **posterior over means** $\rho_{kt}(\mu_k)$ which will guide our exploration.

Gaussian models

For known x_k and σ, consider the likelihood $\nu_k(y|\theta) = \mathcal{N}(y; x_k^T\theta, \sigma^2)$ and a zero-mean prior of variance η^2.

After t arm pulls the posterior can be written as,

$$\pi_t(\theta) = \mathcal{N}(\theta; \hat\theta_t, \hat\Sigma_t) \quad \text{with} \quad \hat\Sigma_t^{-1} = \sigma^{-2}X_t^TX_t + \eta^{-2}I, \quad \hat\theta_t = \sigma^{-2}\hat\Sigma_tX_t^TY_t,$$

and the mean for each arm is marginally

$$\rho_{kt}(\mu_k) = \mathcal{N}(\mu_k; \hat\mu_{kt}, \hat\sigma_{kt}^2) \quad \text{with} \quad \hat\mu_{kt} = x_k^T\hat\theta_t, \quad \hat\sigma_{kt}^2 = x_k^T\hat\Sigma_tx_k.$$

Gaussian Process (GP) models

The above linear observation model is not restrictive; for discrete arms it includes the GP setting. Given a kernel matrix $G \in \mathbb{R}^{K\times K}$

$$X = VD^{\frac{1}{2}} \quad \text{for } G = VDV^T$$

defines the vectors x_k for each arm.

Bayesian gap-based exploration

Let $U_k(t)$ and $L_k(t)$ be upper/lower bounds on μ_k at time t given by $\hat\mu_{kt} \pm \beta\hat\sigma_{kt}$. Write the following **gap** quantity:

$$B_k(t) = \max_{i\neq k} U_i(t) - L_k(t),$$

which upper bounds the simple regret.

Strategy: explore by choosing between the arm $J(t)$ minimizing the gap and the best "alternative" arm $j(t)$—pick the one with highest uncertainty

Choose β in order to guarantee high probability that best arm is selected.

BayesGap pseudo-code

1: **for** $t = 1, \dots, T$ **do**
2: set $J(t) = \arg\min_{k\in\mathcal{A}} B_k(t)$
3: set $j(t) = \arg\max_{k\neq J(t)} U_k(t)$
4: select arm $a_t = \arg\max_{k\in\{j(t),J(t)\}} \hat\sigma_{kt}$
5: observe $y_t \sim \nu_{a_t}(\cdot)$; update posterior and bounds
6: **end for**
7: **return** $\Omega_T = J\big(\arg\min_{t\leq T} B_{J(t)}(t)\big)$

Theorem (Simple Regret of BayesGap)

Consider a K-armed Gaussian bandit problem, horizon T, and upper and lower bounds defined as above. For $\epsilon > 0$ and

$$\beta^2 = \left(\frac{T-K}{\sigma^2} + \frac{\sum_k \|x_k\|^{-2}}{\eta^2}\right) \Big/ \left(4\textstyle\sum_k H_{k\epsilon}^{-2}\right)$$

then BayesGap attains simple regret satisfying

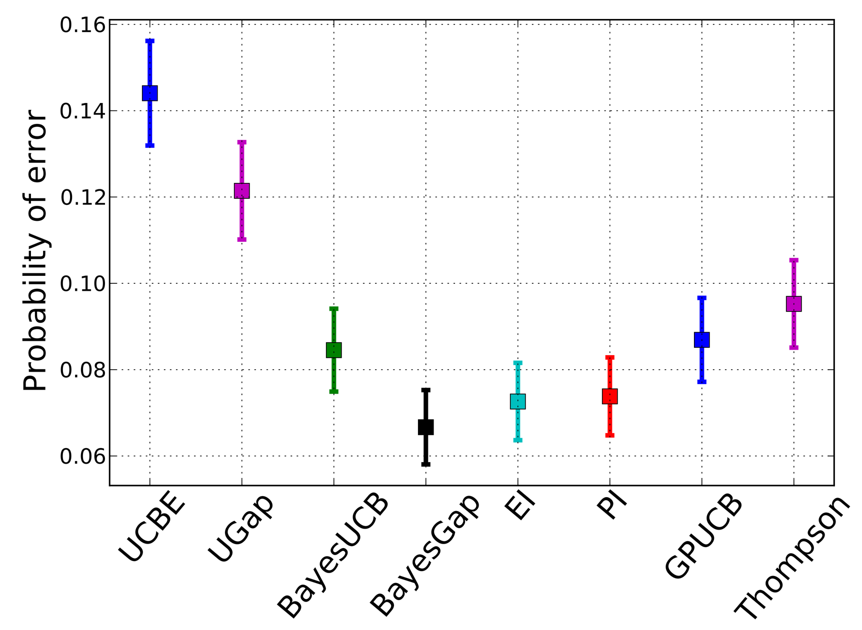
$$\Pr(R_{\Omega_T} \leq \epsilon) \geq 1 - KTe^{-\beta^2/2}.$$

Example 1: optimization in sensor networks

Given data taken from $K = 357$ traffic speed sensors on highway I-880 South in California, identify the single location with the highest expected speed

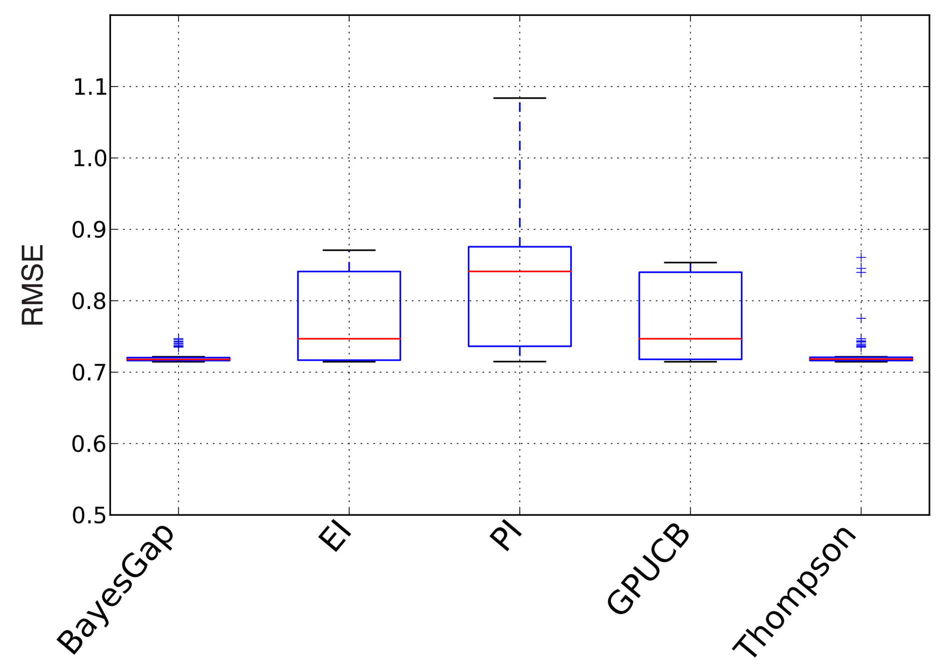
Probability of error while finding the traffic speed sensor with lowest congestion.

BayesGap provides considerable improvements over both cumulative and simple regret alternatives.



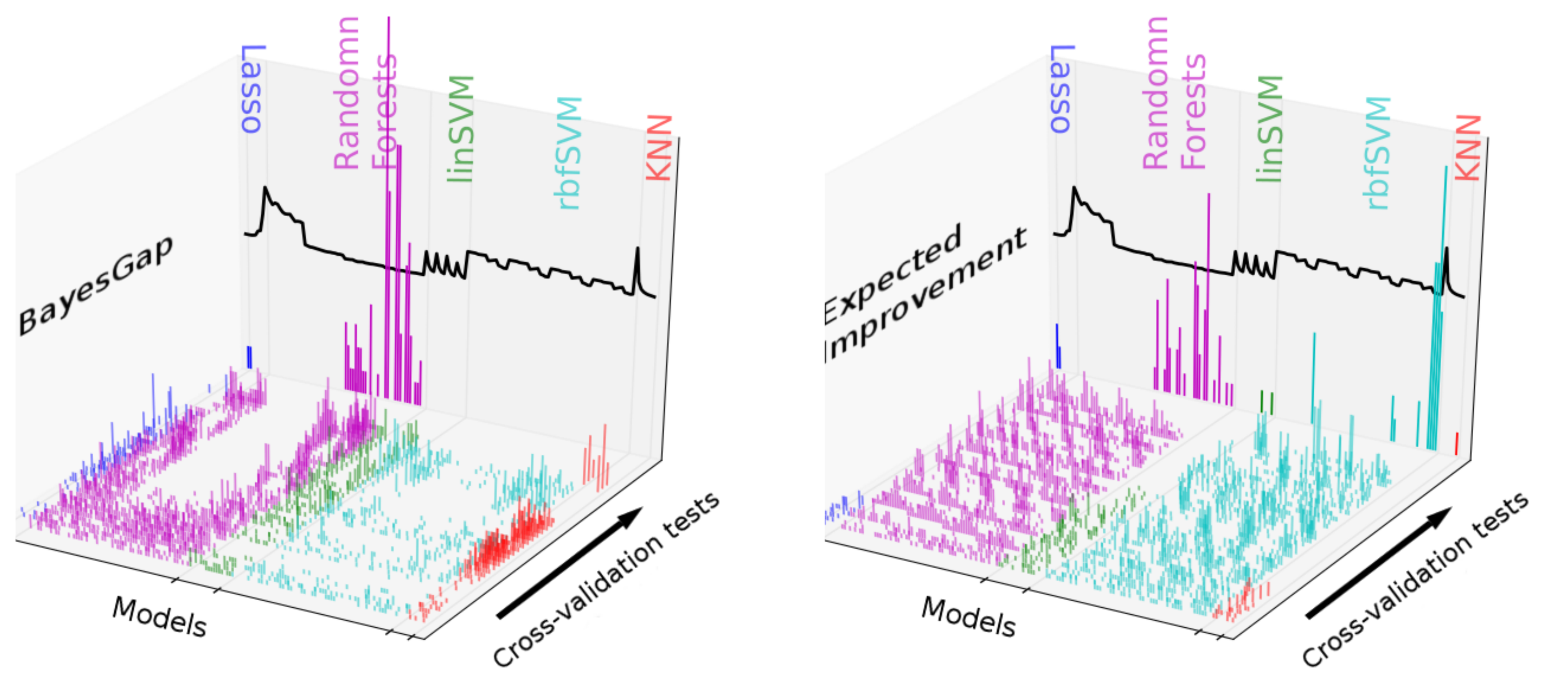
Example 2: automatic machine learning (model selection)

Automating model selection for regression with `scikit-learn`.
- goal: minimize the cost of cross-validation with large amounts of data
- cannot afford to try all models in all cross-validation tests
- better estimation of the generalization error than single test/train split.



Performance over 100 runs at a budget of $T = 40$ with 160 models (i.e., more arms than possible observations).

Allocations and recommendations of BayesGap and EI



Histogram of the arms pulled at each round; far wall shows the final arm recommendation over 100 different runs; the solid black line on the far wall shows the estimated "ground truth".

Key references

J.-Y. Audibert, S. Bubeck, and R. Munos. *Best arm identification in multi-armed bandits.* In *CoLT*, 2010.

E. Brochu, V. Cora, and N. de Freitas. *A tutorial on Bayesian optimization of expensive cost functions.* Technical Report arXiv:1012.2599, 2010.

V. Gabillon, M. Ghavamzadeh, and A. Lazaric. *Best arm identification: A unified approach to fixed budget and fixed confidence.* In *NIPS*, 2012.

M. Hoffman, B. Shahriari, and N. de Freitas. *Exploiting correlation and budget constraints in bayesian multi-armed bandit optimization.* Technical Report arXiv:1303.6746, 2013.

E. Kaufmann, O. Cappé, and A. Garivier. *On Bayesian upper conf. bounds for bandit problems.* In *AIStats*, 2012.

N. Srinivas, A. Krause, S. M. Kakade, and M. Seeger. *Gaussian process optimization in the bandit setting: No regret and experimental design.* In *ICML*, 2010.