Can Small Islands Protect Nearby Coasts from Tsunamis? An active experimental design approach

Themistoklis Stefanakis[1], *Emile Contal* [1], Nicolas Vayatis[1], Frederic Dias[2], Costas Synolakis[3]

[1] CMLA, ENS Cachan, France
[2] University College Dublin, Ireland
[3] University of Southern California, USA

LITTORAL 12-10-2015, ENS Cachan

Run-up Amplification: Stefanakis et al. (2011)

PRL **107,** 124502 (2011) PHYSICAL REVIEW LETTERS week ending 16 SEPTEMBER 2011

Local Run-Up Amplification by Resonant Wave Interactions

Themistoklis S. Stefanakis[1,2] and Frédéric Dias[2,1,*]
[1]*CMLA, ENS Cachan, 61 avenue du Président Wilson, 94235 Cachan cedex, France*
[2]*School of Mathematical Sciences, University College of Dublin, Dublin, Ireland*

Denys Dutykh[3]
[3]*Université de Savoie-CNRS Laboratoire de Mathématiques LAMA - UMR 5127 Campus Scientifique 73376 Le Bourget-du-Lac, France*
(Received 1 July 2011; published 16 September 2011)

⋮

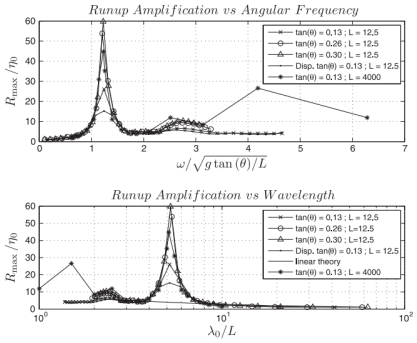


FIG. 2. Maximum run-up amplification ratio as a function of nondimensional angular frequency (top) and nondimensional wavelength (bottom) for two beach lengths, namely $L = 12.5$ m and 4000 m.

more light on the resonant mechanism ($\omega = 0.4$ s^{-1}, $\tan\theta = 0.13$). The first snapshot is taken at the instant

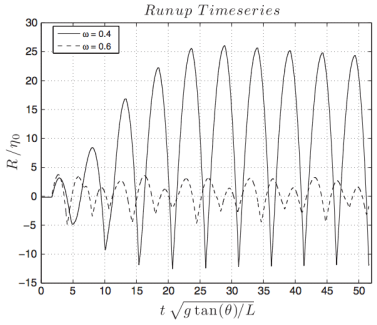
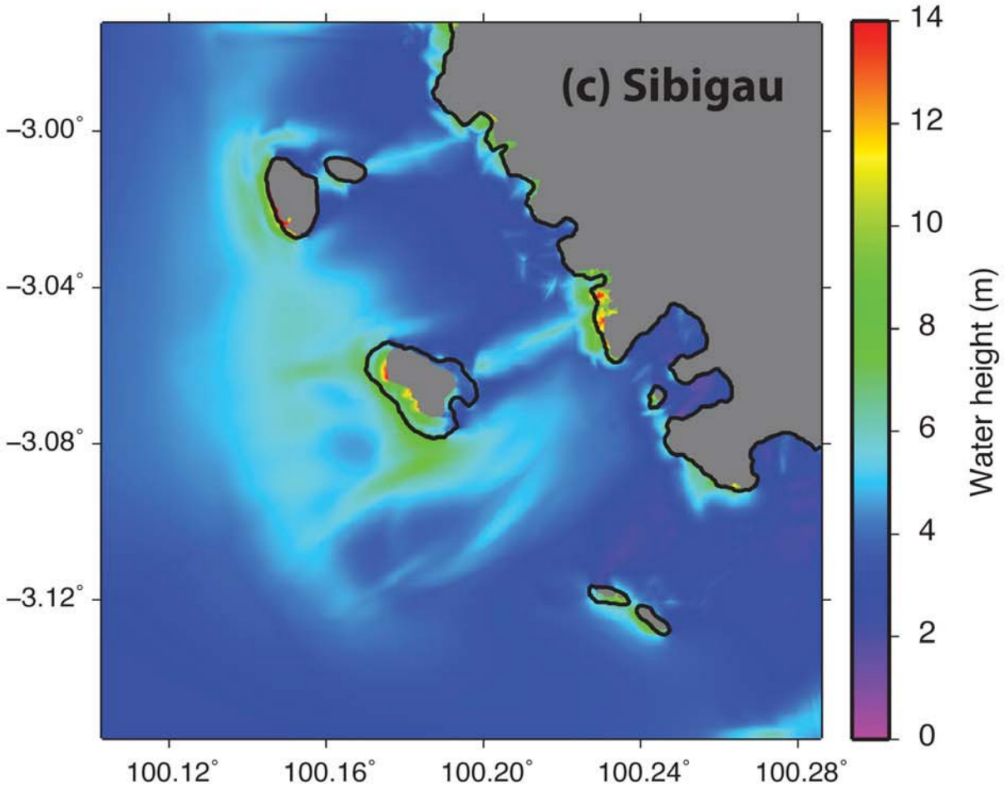


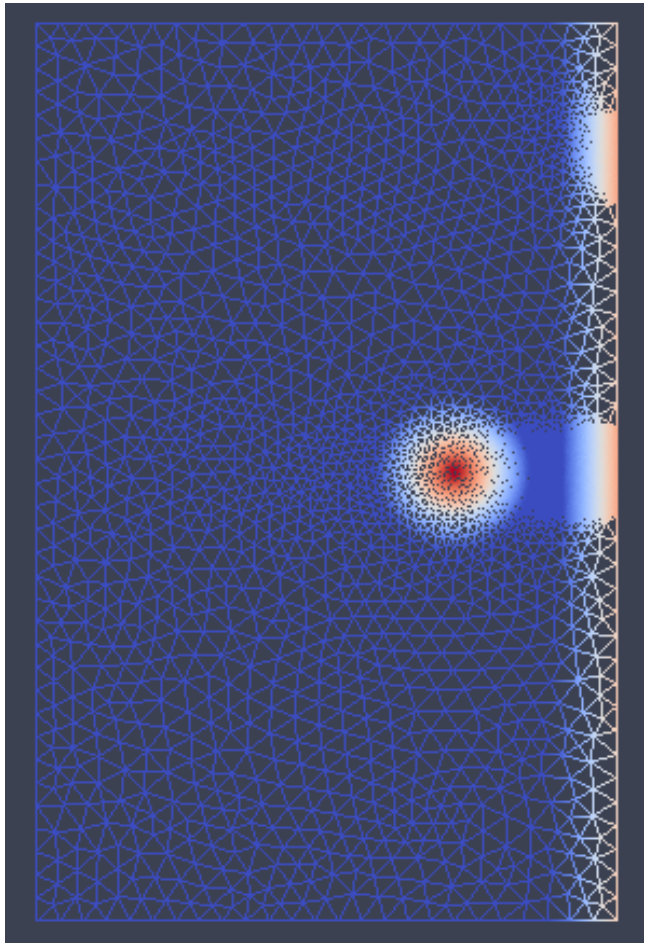
FIG. 3. Run-up time series for two different angular frequencies: $\omega = 0.4$ s^{-1}, which is the resonant frequency for $\tan\theta = 0.13$, and $\omega = 0.6$ s^{-1}, which is a nonresonant frequency for the same slope ($L = 12.5$ m).

Tsunamis Amplification Phenomenon



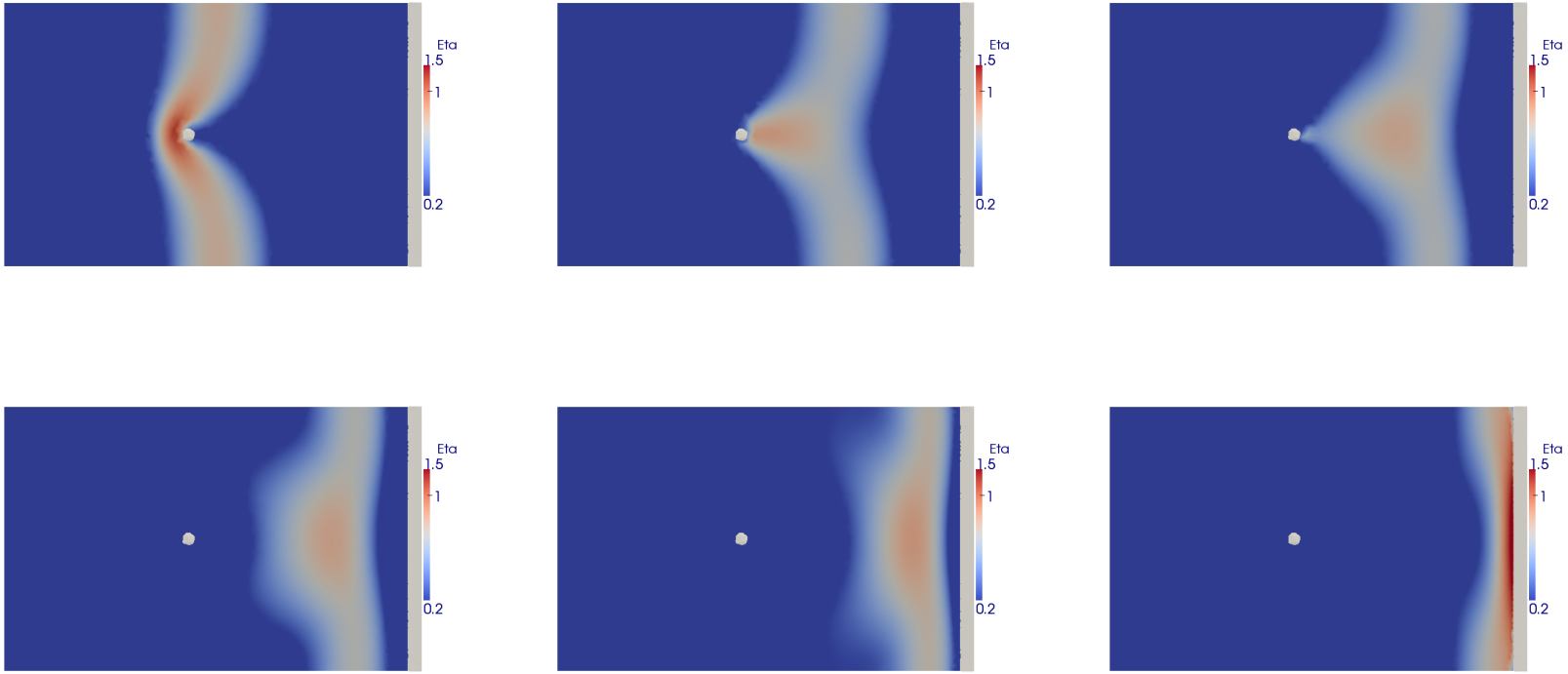
2010 Sumatra tsunami and the Mentawai Islands (Hill et al., 2012)

Numerical Simulations



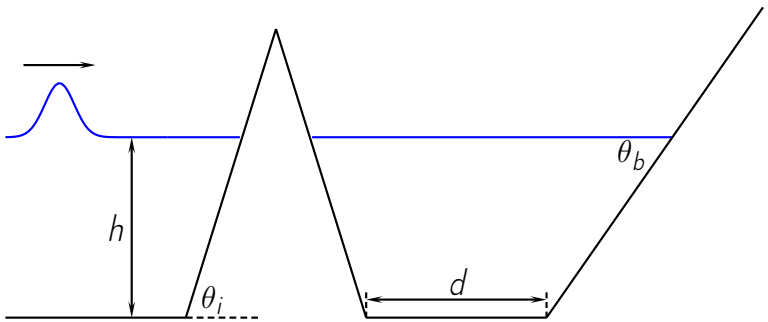
Adaptive mesh grid of the VOLNA solver

Tsunamis Amplification Phenomena

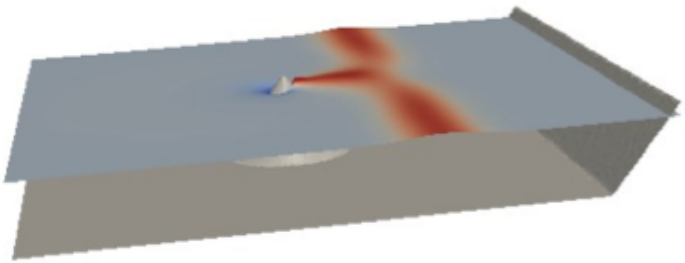


Numerical simulations of a tsunami amplification generated by a conical island

Model Parameters



$\tan\theta_i$	$0.05 - 0.2$
$\tan\theta_b$	$0.05 - 0.2$
d	$0 - 5000$m
h	$100 - 1000$m
ω	$0.01 - 0.1$rad/s



Five parameters modelling the geometry

Problem Statement

Setup

- d real parameters denoted by d-dimensional vectors $x \in \mathcal{X}$
- $\mathcal{X} \subseteq \mathbb{R}^d$ compact and convex
- Unknown objective function $f(x) \in \mathbb{R}$ for all $x \in \mathcal{X}$
- Noisy measurement $y = f(x) + \epsilon$, where $\epsilon \overset{\text{iid}}{\sim} \mathcal{N}(0, \eta^2)$

Goal

Find the parameters x maximizing $f(x)$

Constraints

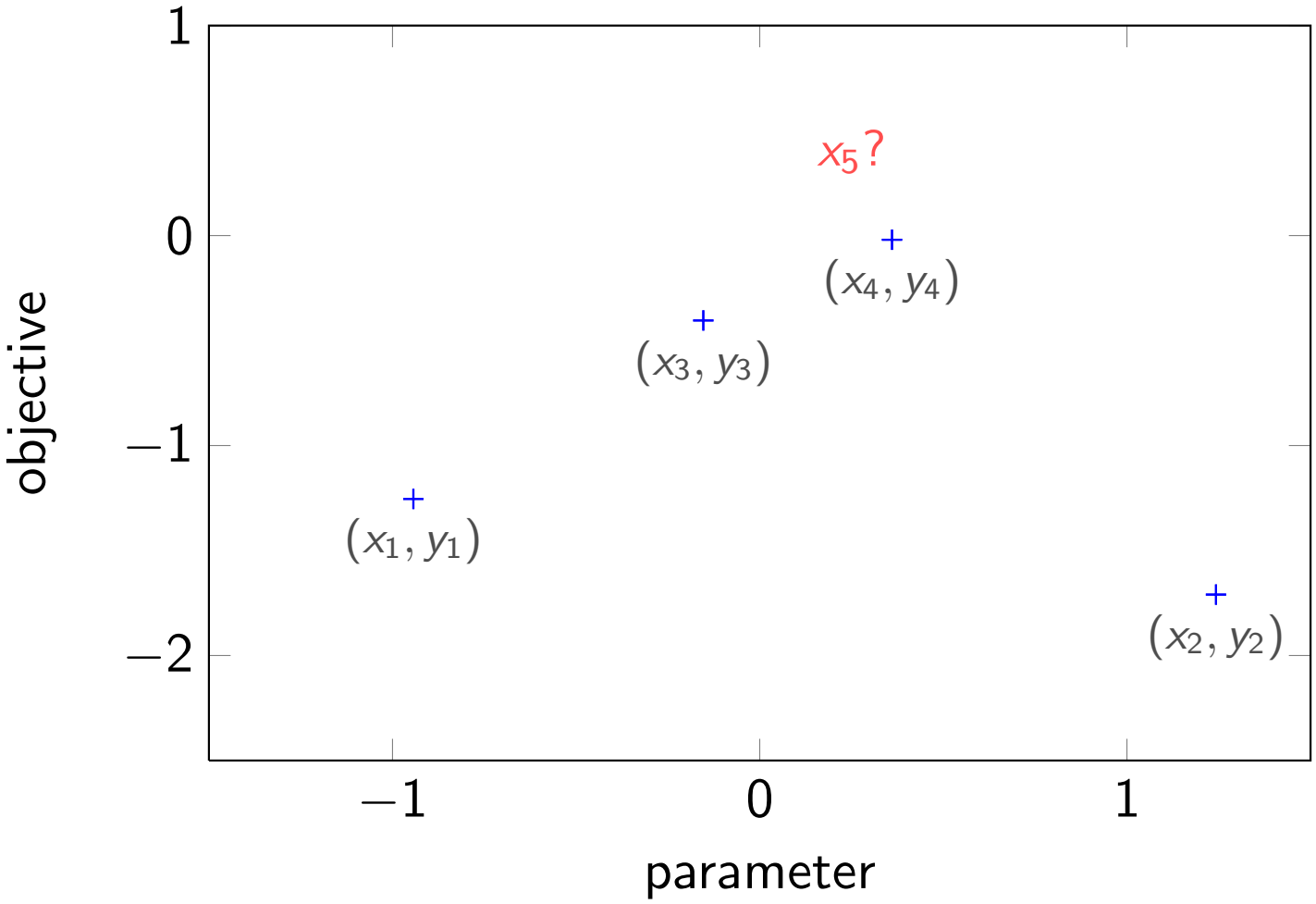
Challenges

- Expensive evaluations
- Joint optimization of several parameters

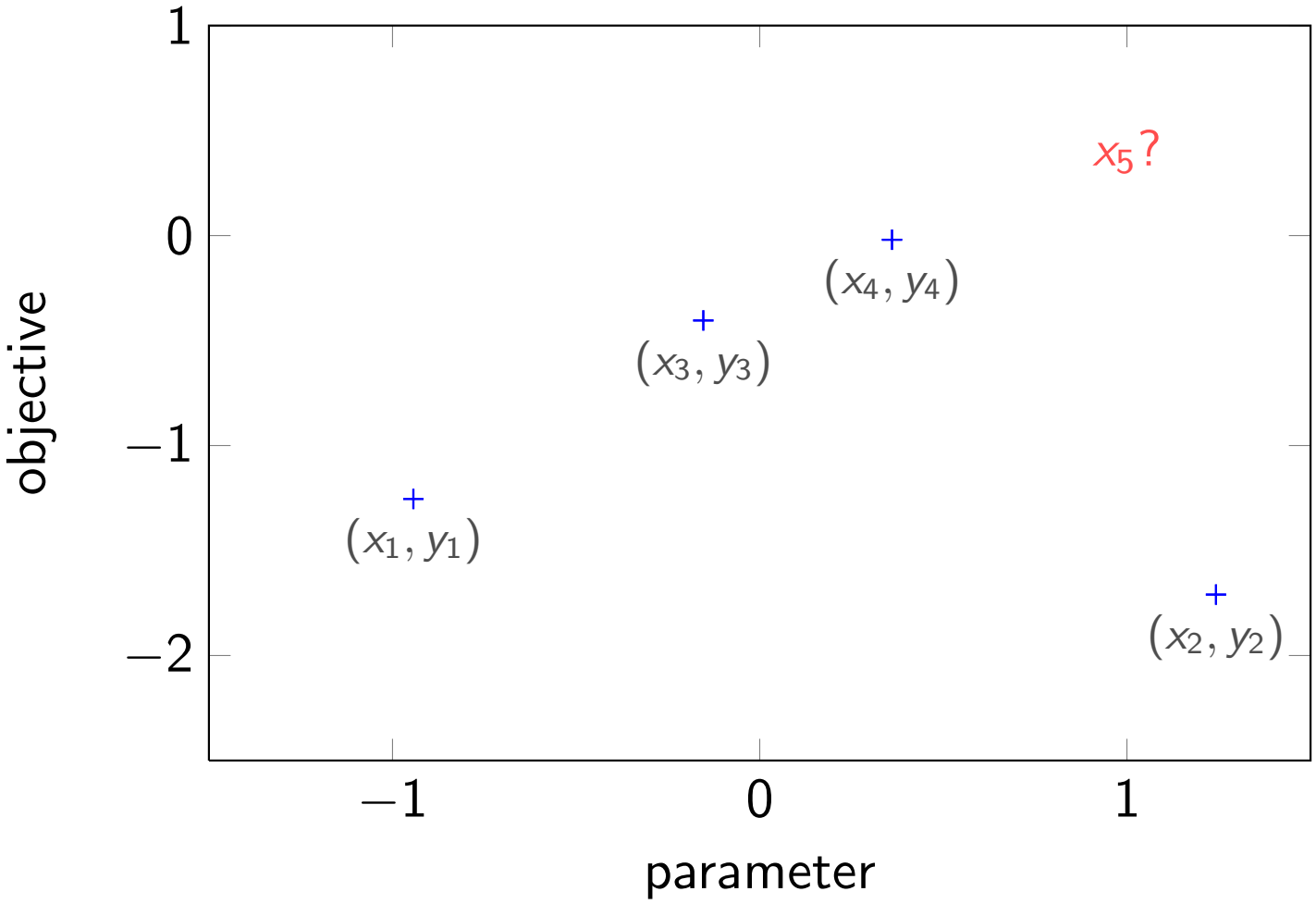
Case: run-up

- 5 parameters
- Each simulation takes 2 hours of computation
- A regular grid with 10 values per parameters needs 10^5 points
- The standard approach would take 23 years of computation

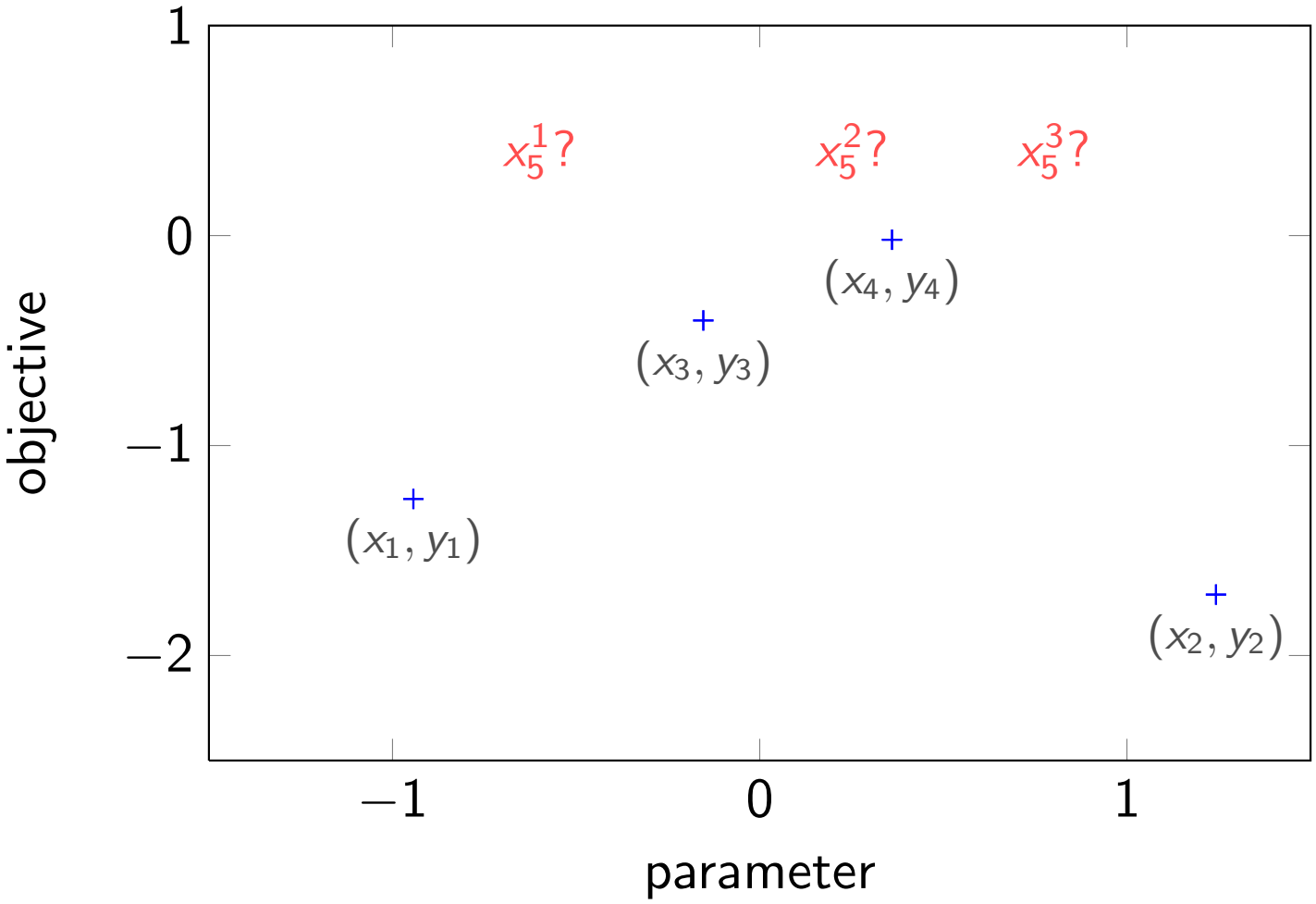
Motivating Example: Sequential Optimization



Motivating Example: Sequential Optimization



Motivating Example: Batch Optimization



Objective

Setting

After t iterations of the sequential query procedure,

- Query the batch $x_{t+1}^1, \dots, x_{t+1}^K$ using the information acquired during the previous iterations
- Observe the respective noisy evaluations $y_{t+1}^1, \dots, y_{t+1}^K$

Target

- We want $\max\limits_{\substack{1 \leq t \leq T \\ 1 \leq k \leq K}} f(x_t^k) \xrightarrow[T \to \infty]{} \max\limits_{x \in \mathcal{X}} f(x)$ as fast as possible.
- Exploration vs Exploitation tradeoff

Gaussian Processes: Framework

Definition

$f \sim \mathcal{GP}(m, k)$, with mean function $m : \mathcal{X} \to \mathbb{R}$ and covariance function $k : \mathcal{X} \times \mathcal{X} \to \mathbb{R}^+$, when for all $x_1, \ldots, x_n$,

$$\big(f(x_1), \ldots, f(x_n)\big) \sim \mathcal{N}(\boldsymbol{\mu}, \mathbf{K}) \ ,$$

$$\text{with } \boldsymbol{\mu}[x_i] = m(x_i) \text{ and } \mathbf{K}[x_i, x_j] = k(x_i, x_j) \ .$$

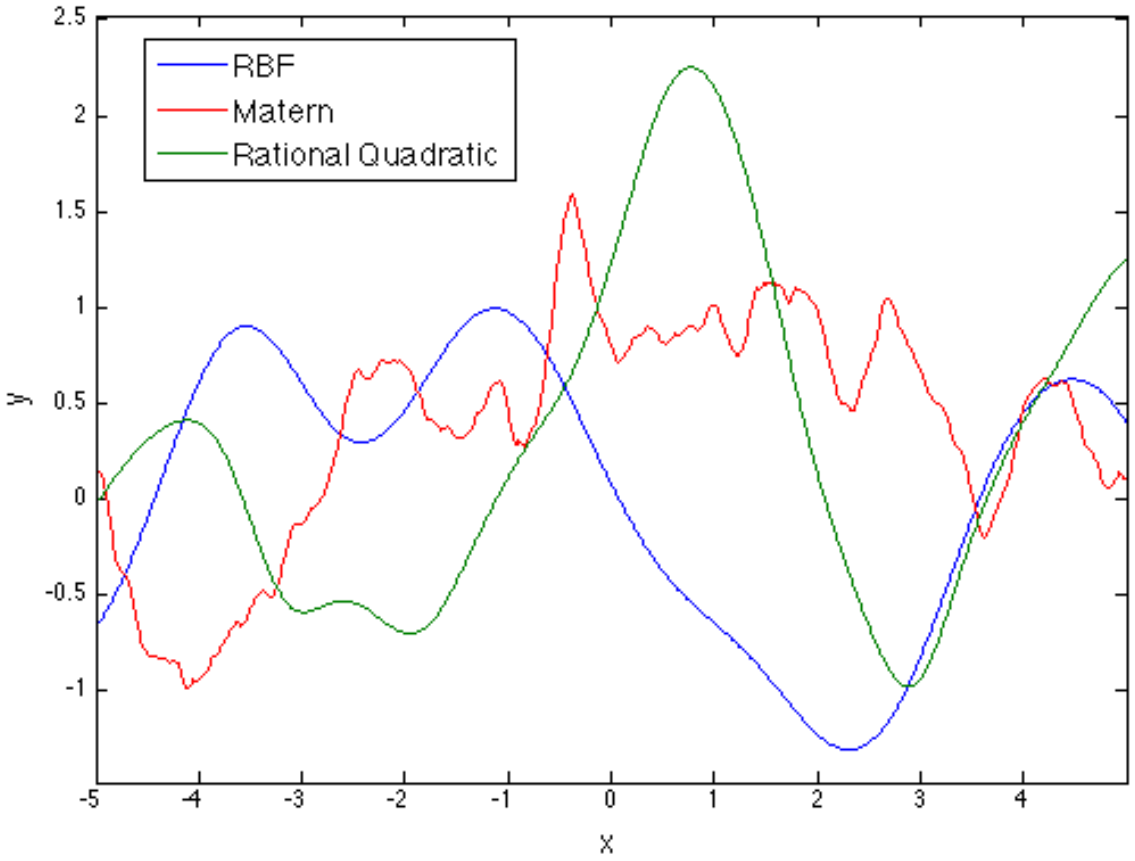
Probabilistic smoothness assumption

- Nearby locations are highly correlated
- Large local variations have low probability

Example of covariance function

- Squared Exponential RBF: $k(x, y) = \exp(-\frac{\|x-y\|_2^2}{2\ell^2})$
- Rational Quadratic: $k(x, y) = \big(1 + \frac{\|x-y\|_2^2}{2\alpha\ell^2}\big)^{-\alpha}$

Gaussian Processes: Examples



1D Gaussian Processes with different covariance functions

Gaussian Process Interpolation

Bayesian Inference (Rasmussen and Williams, 2006)

At iteration t, with observations $\mathbf{Y}_t$ for the query points X_t, the posterior mean and variances are given at all point x in the search space by:

$$\mu_t(x) := \mathbb{E}[f(x) \mid X_t, \mathbf{Y}_t] = \mathbf{k}_t(x)^\top \mathbf{C}_t^{-1} \mathbf{Y}_t \tag{1}$$

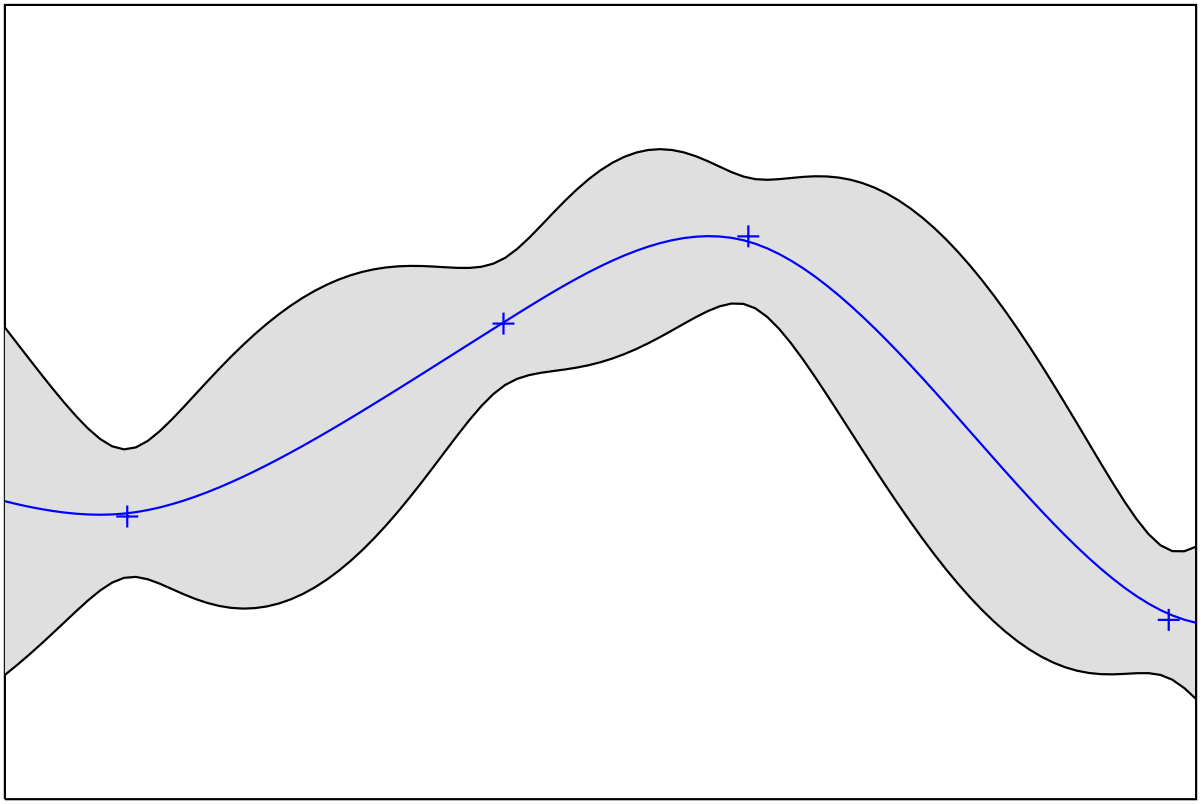
$$\sigma_t^2(x) := \mathbb{V}[f(x) \mid X_t, \mathbf{Y}_t] = k(x,x) - \mathbf{k}_t(x)^\top \mathbf{C}_t^{-1} \mathbf{k}_t(x) \ , \tag{2}$$

where $\mathbf{C}_t = \mathbf{K}_t + \eta^2 \mathbf{I}$ and $\mathbf{K}_t = [k(x_t, x_{t'})]_{x_t, x_{t'} \in X_t}$.

Interpretation

- posterior mean μ_t: prediction
- posterior variance σ_t^2: uncertainty

Gaussian Process Interpolation: Example



Upper and Lower Confidence Bounds

Definition

Fix $0 < \delta < 1$,

$$f_t^+(x) = \mu_t(x) + \sqrt{\beta_t \sigma_t^2(x)}$$
$$f_t^-(x) = \mu_t(x) - \sqrt{\beta_t \sigma_t^2(x)}$$

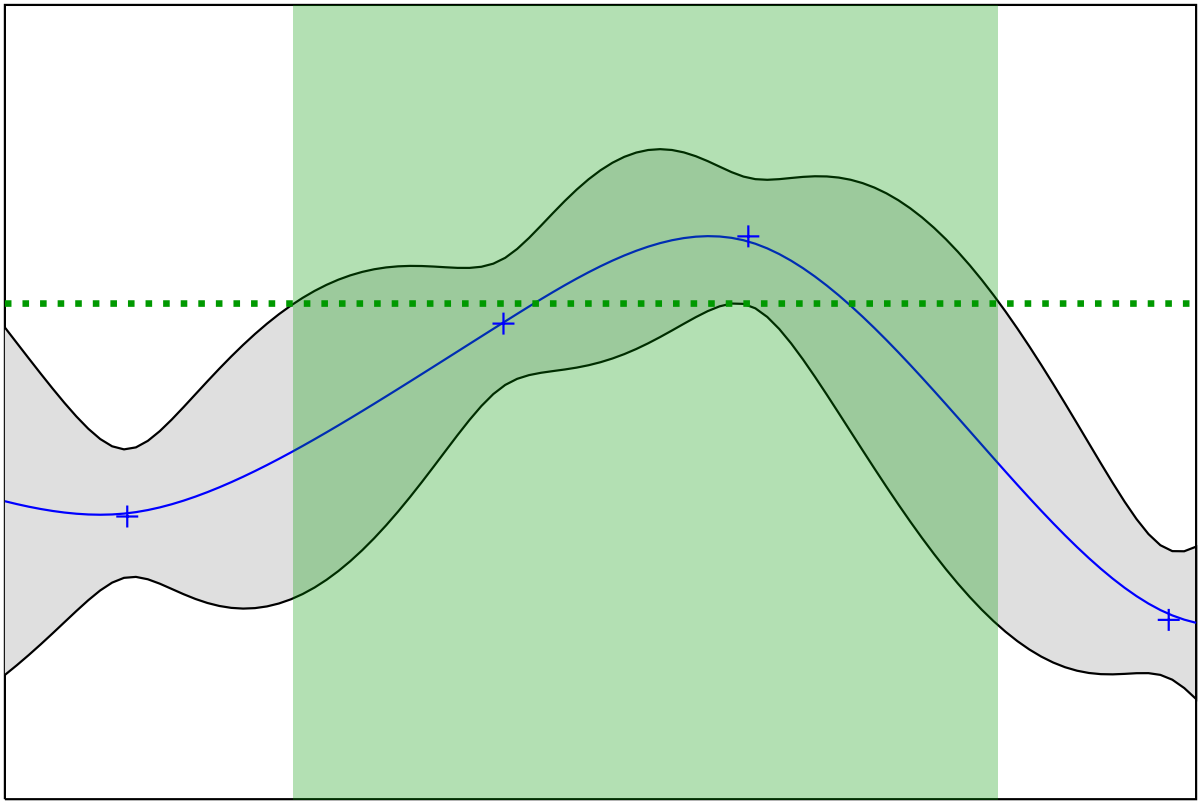
with $\beta_t = \mathcal{O}(\log \frac{t}{\delta})$ defined in Srinivas et al. (2012)

Property

$$\forall x \in \mathcal{X}, \forall t \geq 1,$$

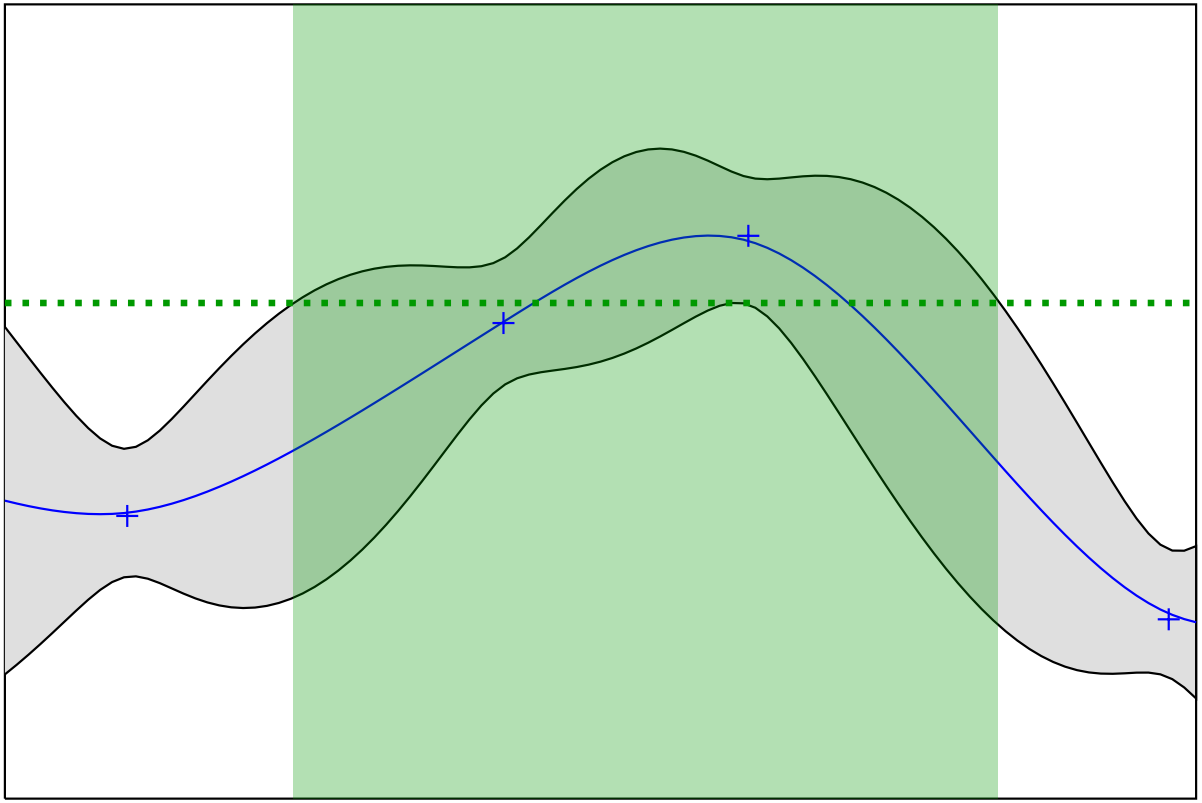
$$f(x) \in \left[f_t^-(x), f_t^+(x)\right] \textit{ holds with probability at least } 1 - \delta$$

Relevant Region $\mathcal{R}_t$



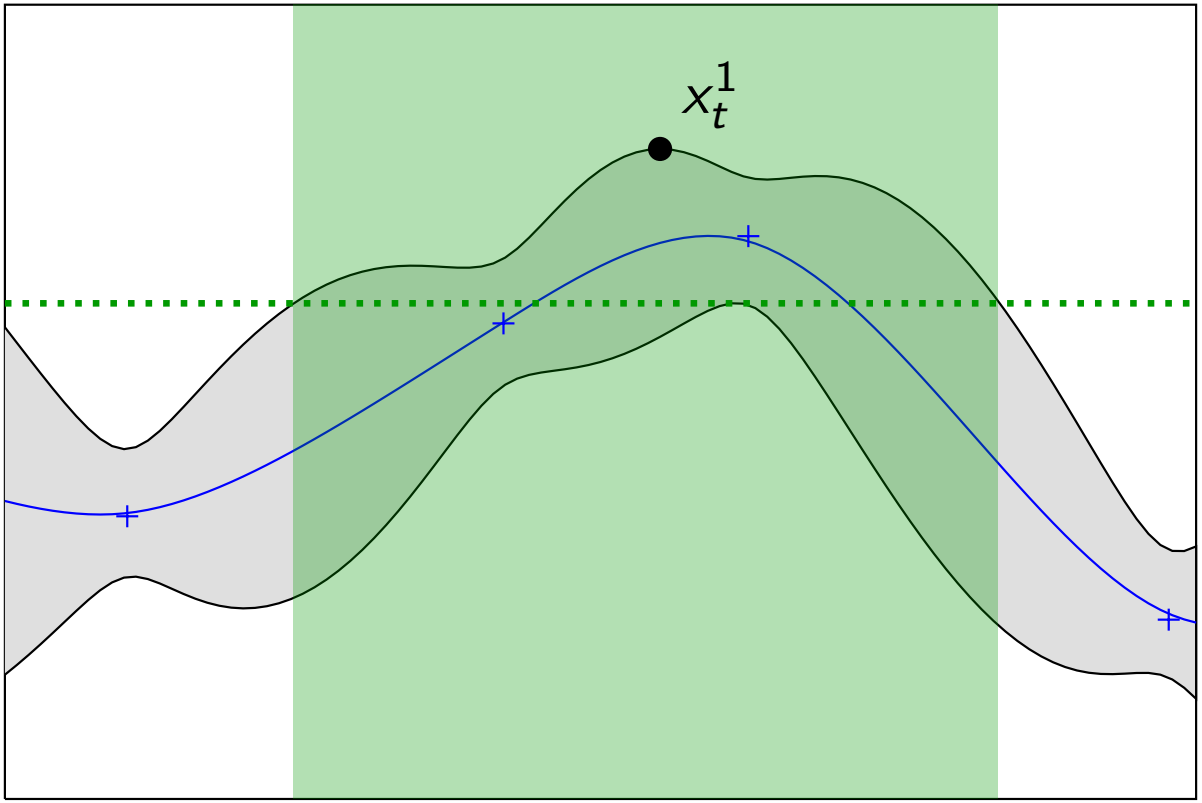
The GP-UCB-PE algorithm (Contal et al., 2013)

$$x_t^1 = ?$$
$$x_t^2 = ?$$
$$x_t^3 = ?$$



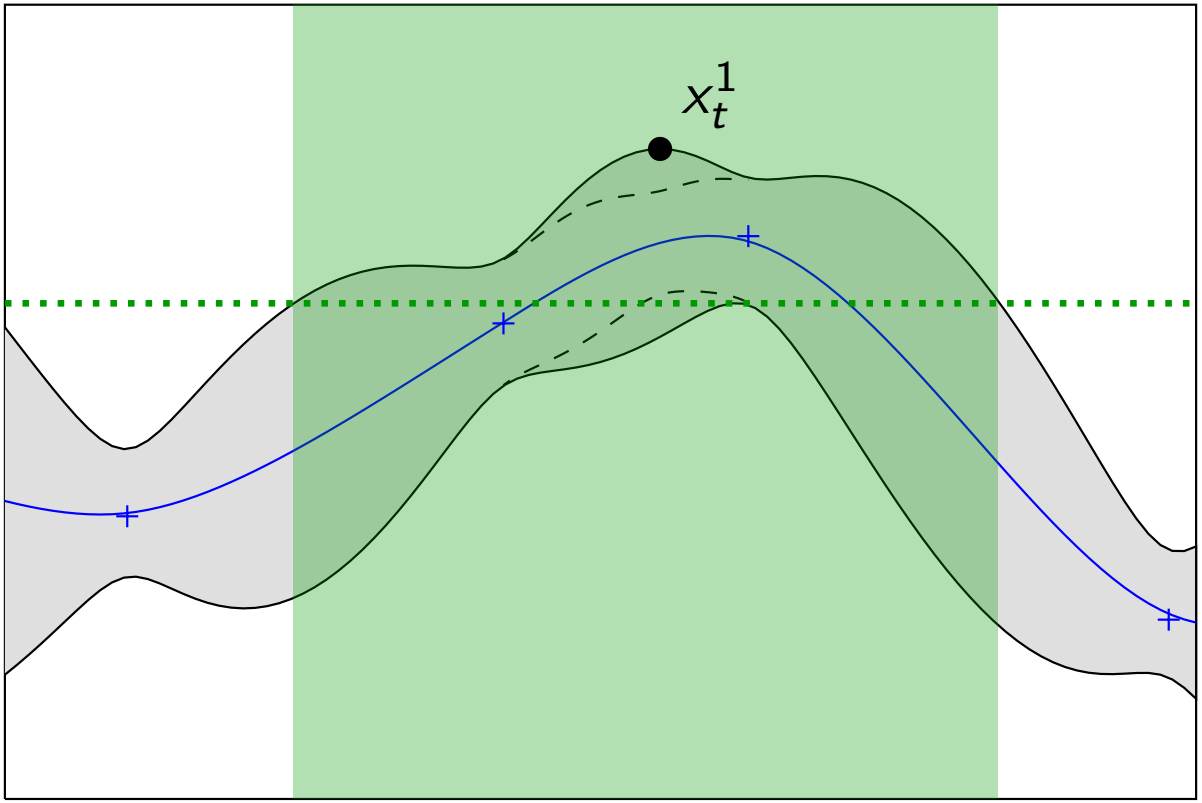
The GP-UCB-PE algorithm (Contal et al., 2013)

$$x_t^1 = \text{argmax}_{x \in \mathcal{R}_t} f_t^+(x)$$
$$x_t^2 = ?$$
$$x_t^3 = ?$$



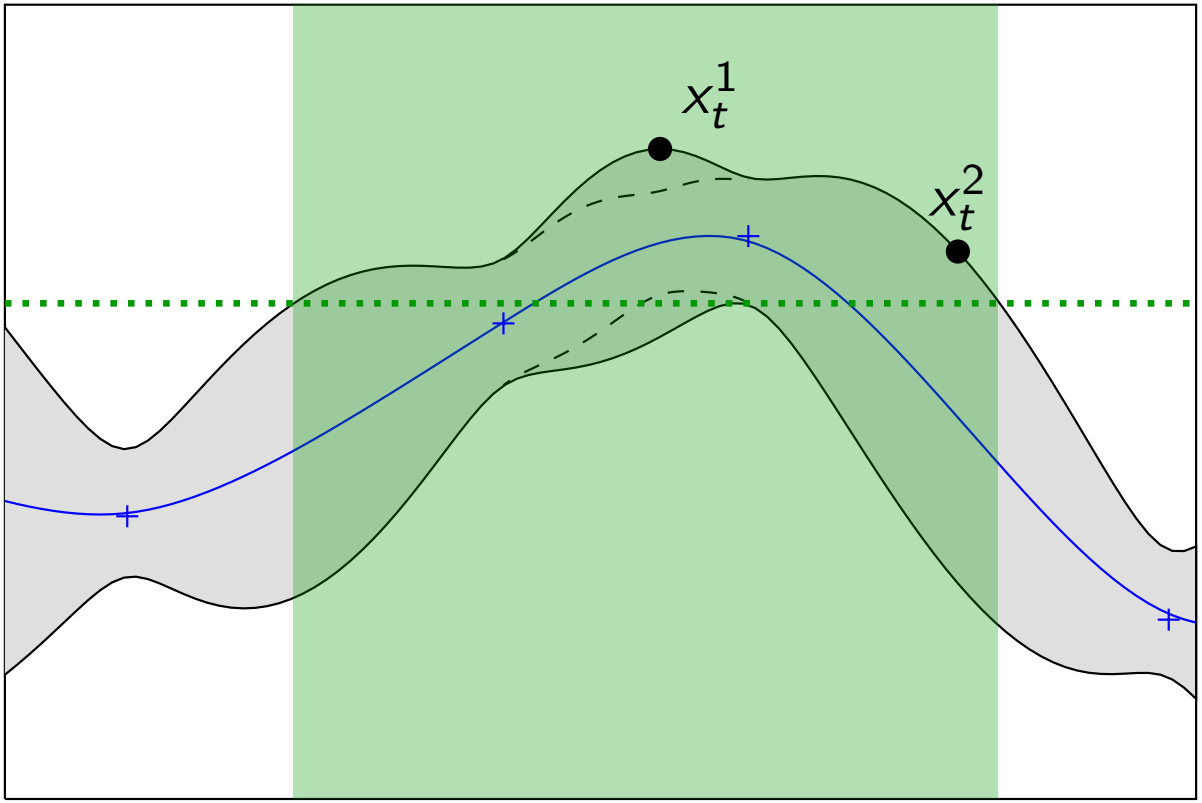
The GP-UCB-PE algorithm (Contal et al., 2013)

$$x_t^1 = \text{argmax}_{x \in \mathcal{R}_t} f_t^+(x)$$
$$x_t^2 = ?$$
$$x_t^3 = ?$$



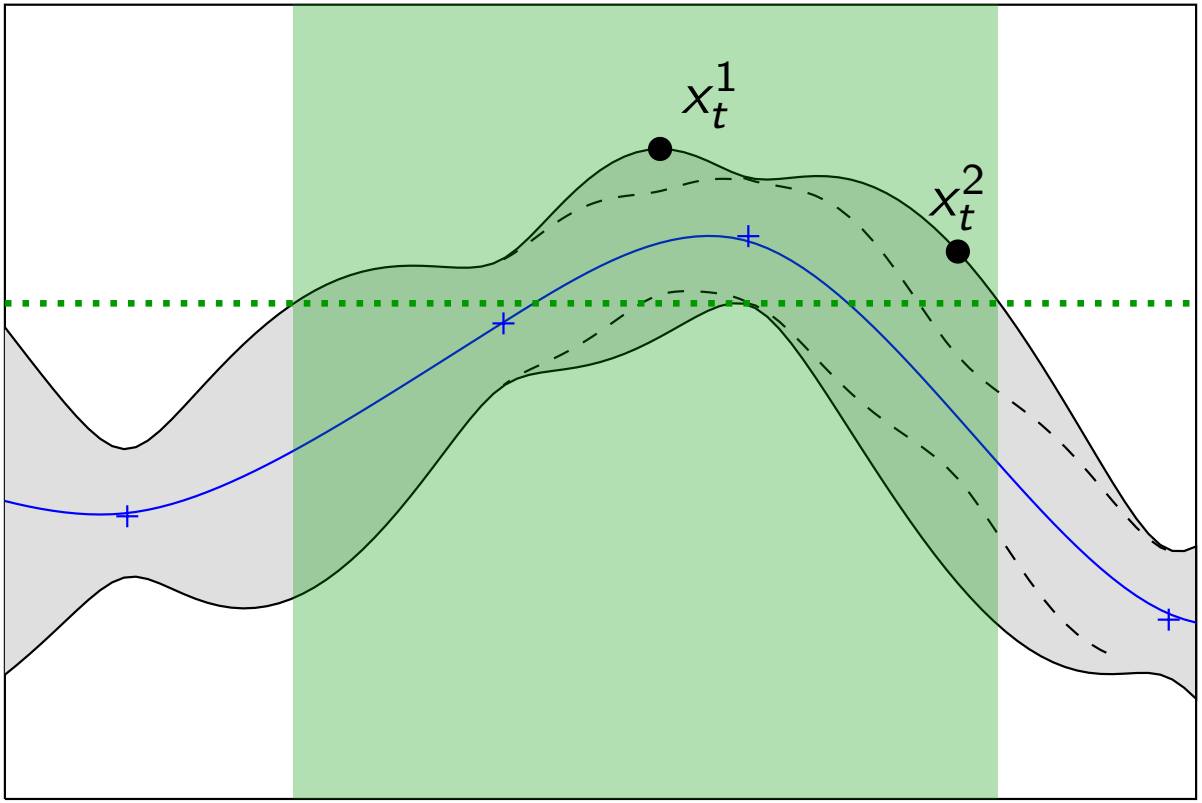
The GP-UCB-PE algorithm (Contal et al., 2013)

$$x_t^1 = \text{argmax}_{x \in \mathcal{R}_t} f_t^+(x)$$
$$x_t^2 = \text{argmax}_{x \in \mathcal{R}_t} \sigma_t^{(1)}(x)$$
$$x_t^3 = ?$$



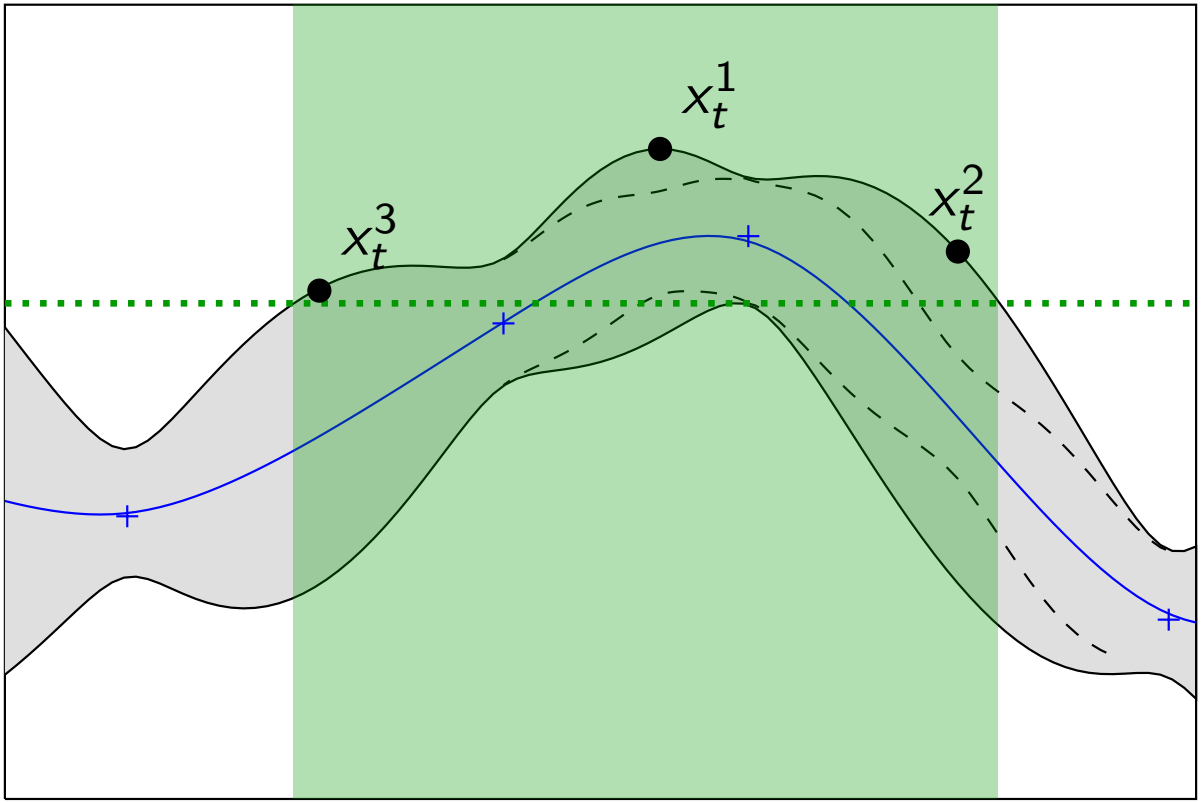
The GP-UCB-PE algorithm (Contal et al., 2013)

$$x_t^1 = \text{argmax}_{x \in \mathcal{R}_t} f_t^+(x)$$
$$x_t^2 = \text{argmax}_{x \in \mathcal{R}_t} \sigma_t^{(1)}(x)$$
$$x_t^3 = ?$$



The GP-UCB-PE algorithm (Contal et al., 2013)

$$x_t^1 = \text{argmax}_{x \in \mathcal{R}_t} f_t^+(x)$$
$$x_t^2 = \text{argmax}_{x \in \mathcal{R}_t} \sigma_t^{(1)}(x)$$
$$x_t^3 = \text{argmax}_{x \in \mathcal{R}_t} \sigma_t^{(2)}(x)$$



Theoretical Analysis

Theorem (Contal et al. (2013))

With $f \sim \mathcal{GP}(0, k)$, with probability at least $1 - \delta$:

$$\max_{x \in \mathcal{X}} f(x) - \max_{\substack{1 \leq t \leq T \\ 1 \leq k \leq K}} f(x_t^k) = \mathcal{O}\left(\sqrt{\frac{\gamma_{TK}}{TK}}\right)$$

Information theory

γ_{TK} is the maximum information gain about f obtainable by a sequence of TK queries.

- For linear covariance, $\gamma_{TK} = \mathcal{O}(d \log TK)$
- For Squared Exponential covariance, $\gamma_{TK} = \mathcal{O}\big((\log TK)^{d+1}\big)$

Batch vs Sequential

Complexity

If n is the number of training points,

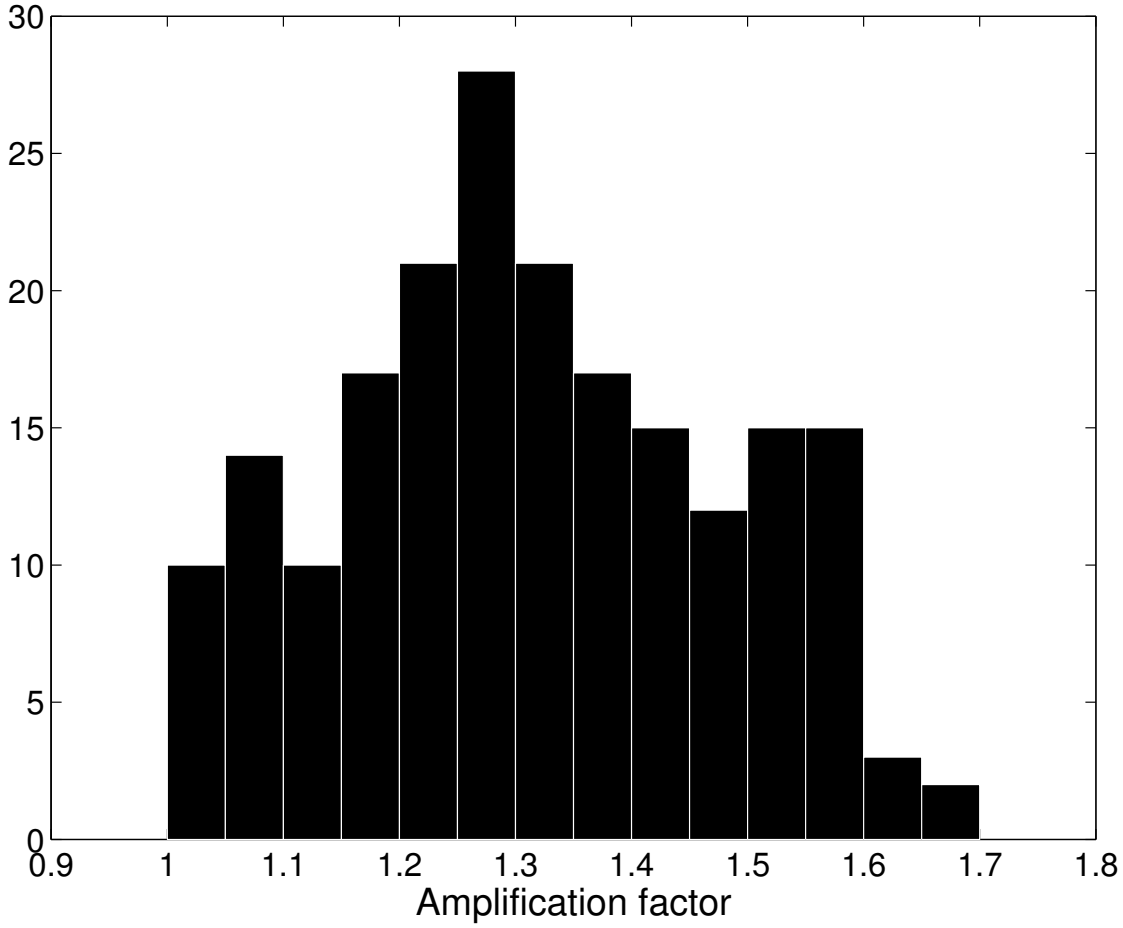
$$\text{Sequential: } n\,\text{Cost}(f) + n\,\text{Cost}(\text{GP})$$
$$\text{Batch: } \frac{n}{K}\,\text{Cost}(f) + n\,\text{Cost}(\text{GP})$$

With exact inference, $\text{Cost}(\text{GP}) = \mathcal{O}(n^2)$.

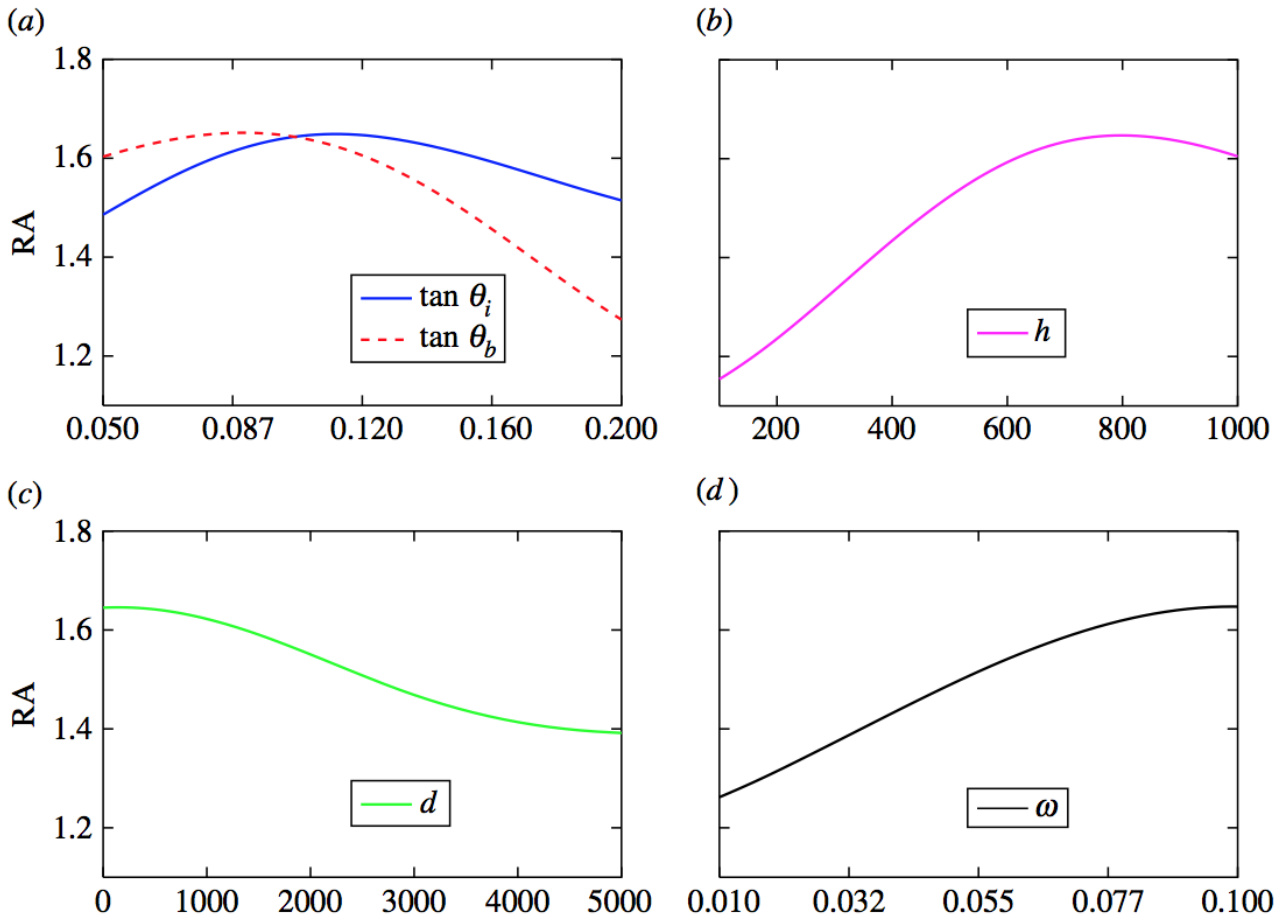
Impact on the convergence speed

Take $K \ll T$, then for equivalent $\text{Cost}(f)$ the improvement of the parallel strategy over the sequential one is $\sqrt{K}$ with respect to the convergence speed.

Histogram of the run-up amplification



Local sensitivity of the maximum run-up amplification



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

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

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

Stefanakis, T. S., Contal, E., Vayatis, N., Dias, F., and Synolakis, C. (2014). Can small islands protect nearby coasts from tsunamis? an active experimental design approach. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 470(2172).

Stefanakis, T. S., Dias, F., and Dutykh, D. (2011). Local run-up amplification by resonant wave interactions. *Phys. Rev. Lett.*, 107:124502.