



HAL
open science

Gaussian Processes: from knowledge-informed Machine Learning to optimization

Didier Rullière, Rodolphe Le Riche, Xavier Bay, Soumyodeep Mukhopadhyay, Victor Trappler, Hassan Maatouk, Tanguy Appriou, Laurent Genest, David Gaudrie, Marc Grossouvre, et al.

► To cite this version:

Didier Rullière, Rodolphe Le Riche, Xavier Bay, Soumyodeep Mukhopadhyay, Victor Trappler, et al.. Gaussian Processes: from knowledge-informed Machine Learning to optimization. HCERES LIMOS visit 2025, Oct 2025, Clermont -Ferrand, France. 2025. <hal-05324906>

HAL Id: hal-05324906

<https://hal.science/hal-05324906v1>

Submitted on 21 Oct 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY 4.0 - Attribution - International License

Why learning with Gaussian Processes ?

Many situations lead to sparse data or uncertain data. Gaussian processes (GPs) offer a mathematically founded framework for learning, optimizing, and other decision-making activities.

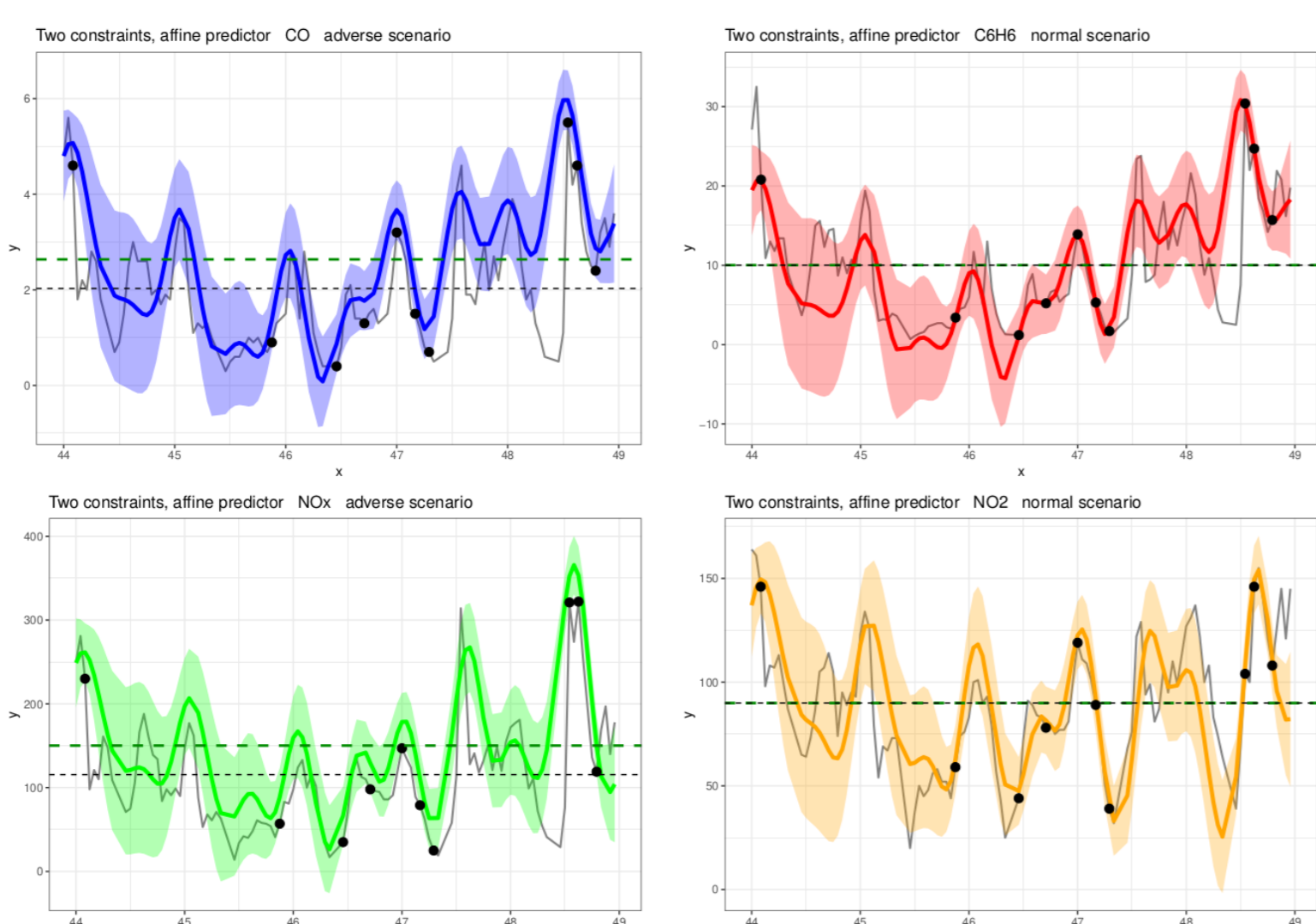
Gaussian process regression framework

- **Problem:** Predict a phenomenon based on few observations
- ▷ Express the GP as a best linear unbiased predictor

$$\hat{Y}(x) := \sum_{i=1}^n \alpha_i(x) Y(x_i)$$

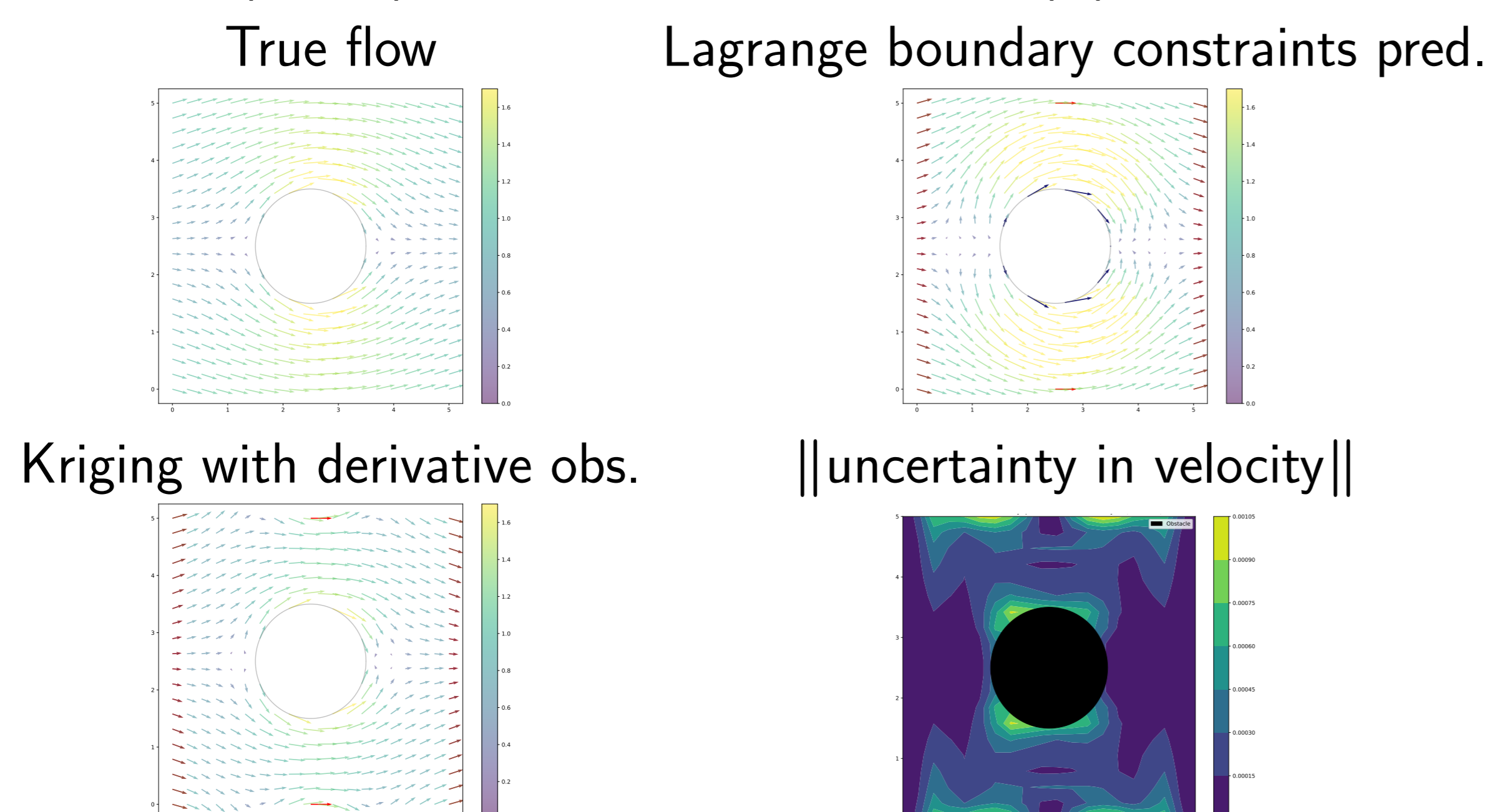
$$\text{with } (\alpha_1(x), \dots, \alpha_n(x)) = \arg \min_{\alpha} \mathbb{E} \left[\left(\hat{Y}(x) - Y(x) \right)^2 \right]$$

- **Contribution [1]:** Application to multi-output, classification.
- ▷ Add unbiased constraints, linear constraints (e.g. adverse scenarios)
- ▷ Extend to vector interpolation, to classification



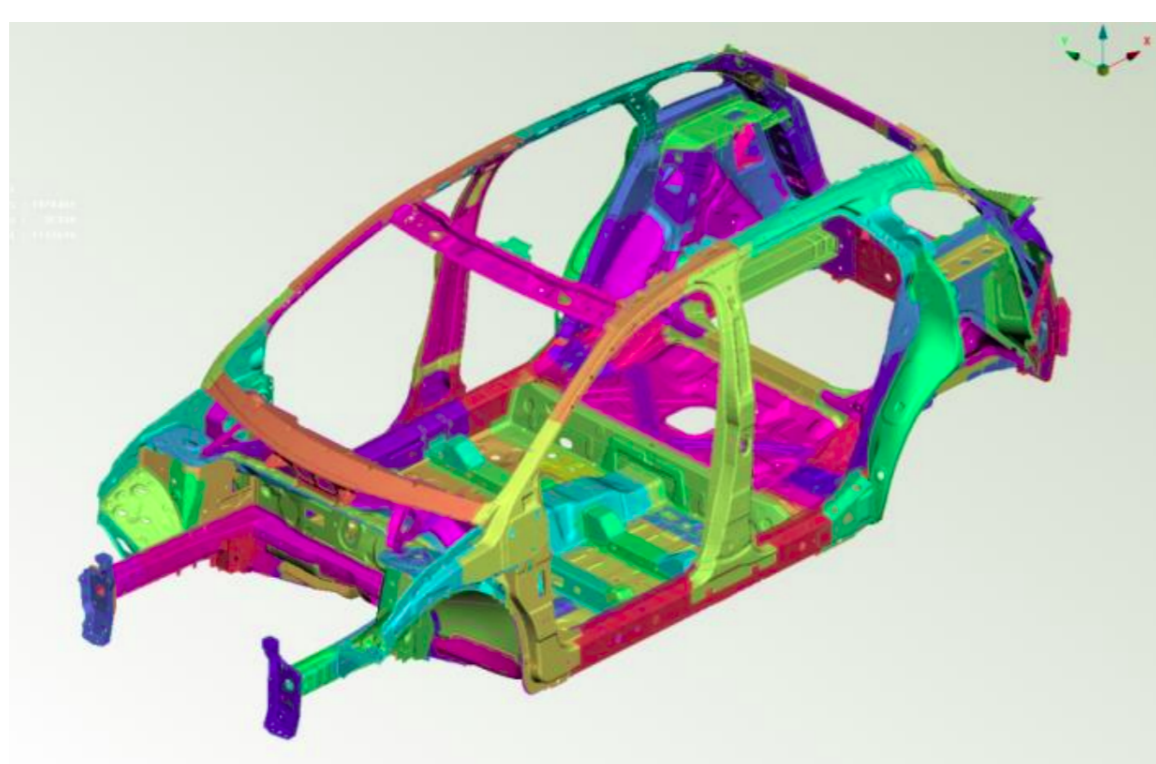
GPs and physics-informed predictions

- **Problem:** Incorporate physical information.
- **Contribution [2]:** Impose linear partial differential equations in a kriging model either as observations or as constraints in the model loss minimization (BLUE). For a potential flow, $\nabla \cdot v(x) = 0$.



GPs for optimization in high dimension

- **Problem:** Optimize in high dimension.
- GPs are a key asset in the optimization of expensive functions: they can partially replace the true function. But GPs tend to fail in high-dimension, even more so in the context of optimization.



- **Contribution [3]:** Build a GP as a linear combination of GPs with fixed length-scales: more robust, avoid repeated minimizations of the loss. Analytical model. E.g. Optimization of a vehicle structure, 130 variables.
- **Contribution [4]:** Optimize with a GP within an adapted Trust Region. TREGO (Trust Region Efficient Global Optimization)

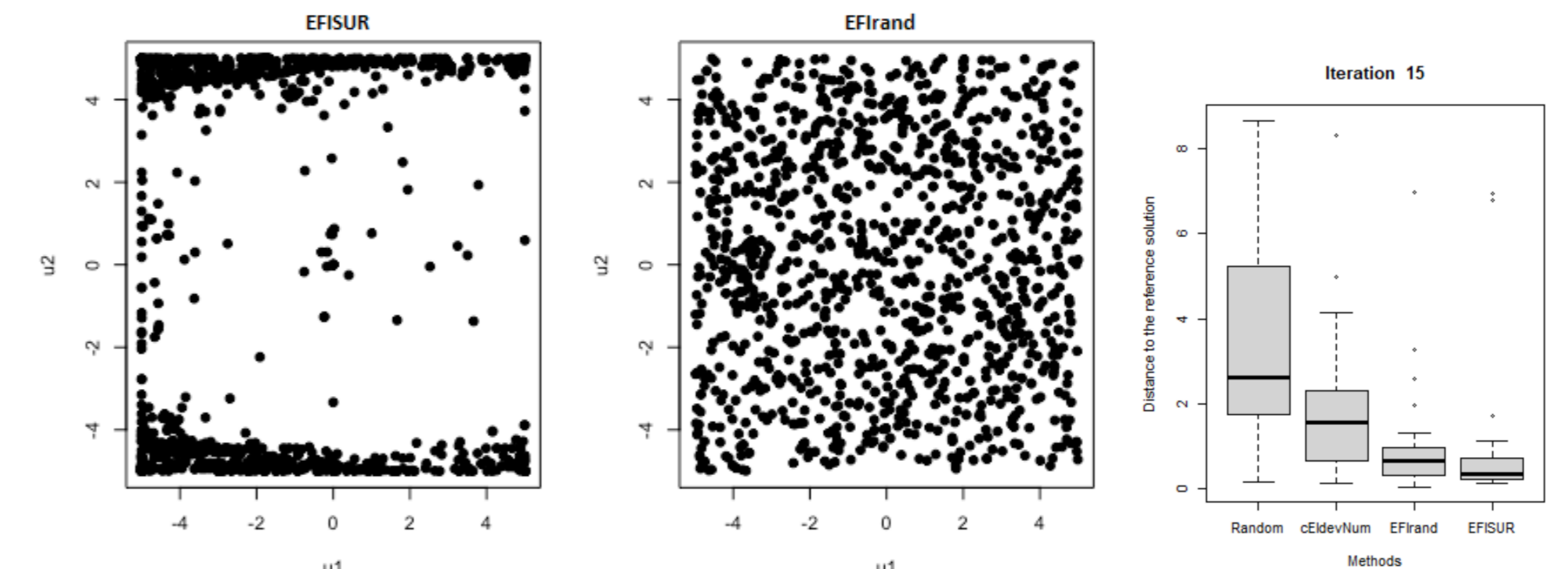
GPs for optimization with uncertainties

- **Problem:** Optimize with uncertain parameters
- $$x^* = \arg \min_{x \in \mathcal{S}_x} \mathbb{E}[f(x, U)] \text{ s.t. } \mathbb{P}(g(x, U) \leq 0) \geq 1 - \alpha,$$
- where U has a known probability density.

- **Contribution [5]:** EFISUR (Efficient Feasible Improvement with Stepwise Uncertainty Reduction) method alternates between

$$\begin{cases} x^{\text{next}} = \arg \max_{x \in \mathcal{S}_x} \text{ExpectedFeasibleImprovement}(x) \\ u^{\text{next}} = \arg \min_{u \in \mathcal{S}_u} \text{VarianceNextStep}(x^{\text{next}}, u) \end{cases}$$

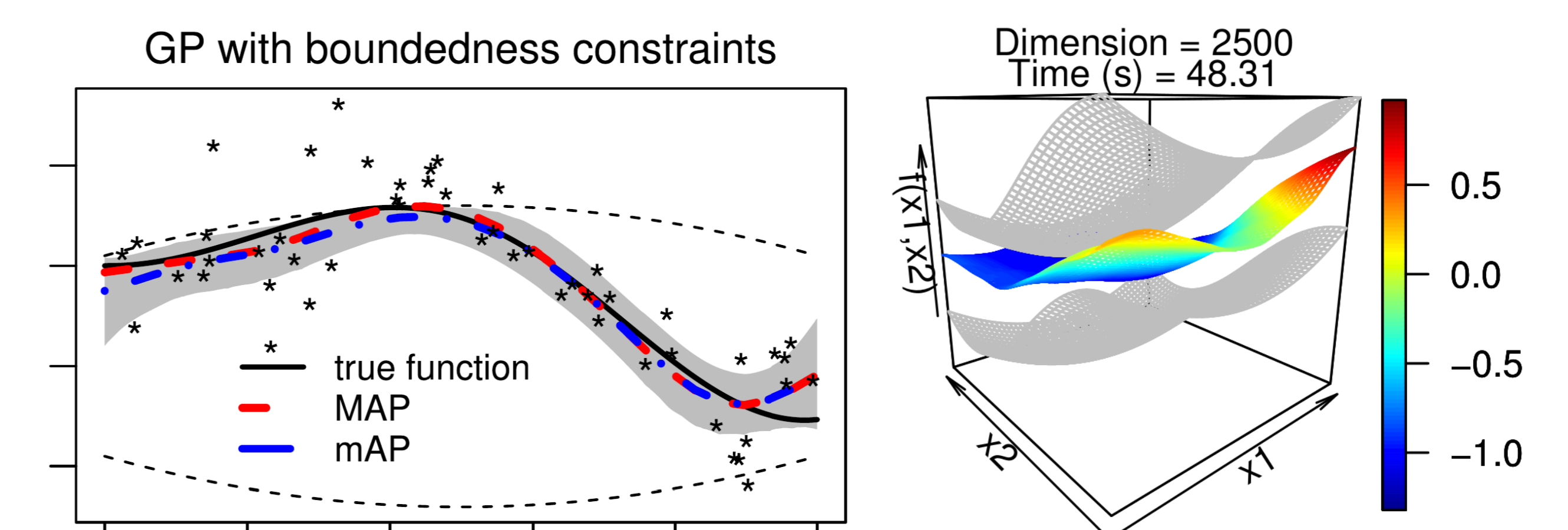
e.g. distribution of uncertain parameters by EFISUR (left) and at random (middle). Effect on the convergence accuracy (right):



- **Contributions [6, 7]:** More information, continuations.

GPs under shape constraints

- **Problem:** Incorporate constraints: monotonicity, shape constraints.
- **Contributions [8, 9]:**
- ▷ Approximate GP by finite dimensional combination of Gaussian random variables (Karhunen–Loève, hat functions...)
- ▷ Transfer constraints on these Gaussian random variables
- ▷ Sample constrained Gaussian random vectors (MCMC, ESS...)



References

- [1] Didier Rullière and Marc Grossouvre. A Joint Kriging Model with Application to Constrained Classification. *Statistics and Computing (to appear)*, September 2025.
- [2] S. Mukhopadhyay, D. Rullière, R. Le Riche, X. Bay, L. Genest, and D. Gaudrie. Collocation-based kriging with applications to the prediction of perfect flows. *CIROQUO Workshop*, May 2025.
- [3] Tanguy Appriou, Didier Rullière, and David Gaudrie. High-dimensional bayesian optimization with a combination of kriging models. *Structural and Multidisciplinary Optimization*, 67(11):196, 2024.
- [4] Youssef Diouane, Victor Picheny, Rodolphe Le Riche, and Alexandre Scotto Di Perrotolo. TREGO: a trust-region framework for efficient global optimization. *Journal of Global Optimization*, 86(1):1–23, 2023.
- [5] R. El Amri, R. Le Riche, C. Helbert, C. Blanchet-Scalliet, and S. Da Veiga. A sampling criterion for constrained bayesian optimization with uncertainties. *The SMAI Journal of computational mathematics*, 9:285–309, 2023.
- [6] Julien Pelamatti, Rodolphe Le Riche, Céline Helbert, and Christophe Blanchet-Scalliet. Coupling and selecting constraints in bayesian optimization under uncertainties. *Optimization and Engineering*, 25(1):373–412, 2024.
- [7] Victor Trappler, Céline Helbert, and Rodolphe Le Riche. Multiobjective optimization under uncertainties using conditional pareto fronts. *arXiv preprint arXiv:2504.04944*, 2025.
- [8] Hassan Maatouk, Didier Rullière, and Xavier Bay. Bayesian analysis of constrained Gaussian processes. *Bayesian Analysis (in press)*, 2024.
- [9] Hassan Maatouk, Didier Rullière, and Xavier Bay. Efficient constrained Gaussian process approximation using elliptical slice sampling. *Bayesian Analysis (to appear)*, 2025.