Multi-fidelity Gaussian processes for noisy outputs and non-nested experiment designs: a comparison between the recursive and non-recursive formulations

ENBIS 2026, Florence, Italy, 8 September

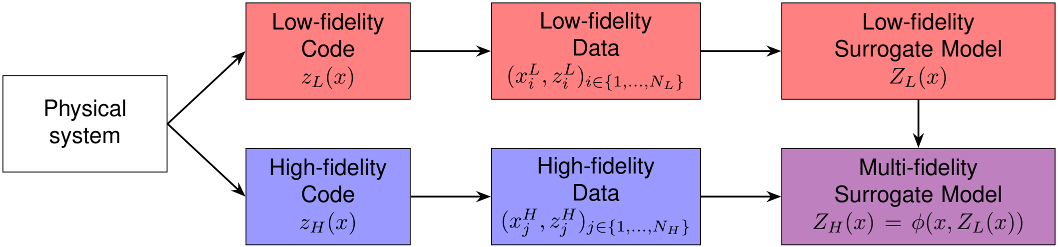
Nils Baillie[1,2], Baptiste Kerleguer[1], Cyril Feau[1], Josselin Garnier[2]

[1] CEA
[2] Ecole Polytechnique, CMAP

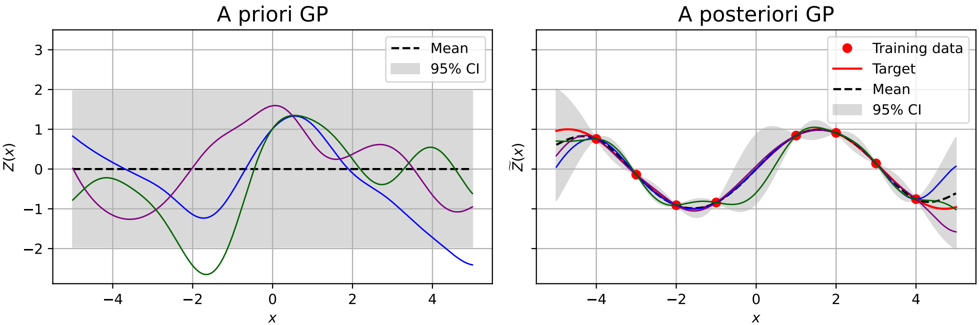
Outline

Outline

Multi-fidelity surrogate model



Gaussian processes



- **Conditioning:** $\begin{cases} Z(x) \sim \mathcal{N}(\mu(x), k(x,x)) \\ \widetilde{Z}(x) = (Z(x) \,|\, \forall i, Z(x_i) = z_i) \sim \mathcal{N}(m(x), V(x,x)) \end{cases}$
- **GP model:** $\begin{cases} \mu_\beta(x) = f(x)^\top \beta \\ k_{\sigma^2,\theta}(x,x') = \sigma^2 \exp\left(-0.5 \cdot |x - x'|^2 / \theta^2\right) \end{cases}$

Outline

Multi-fidelity Gaussian processes

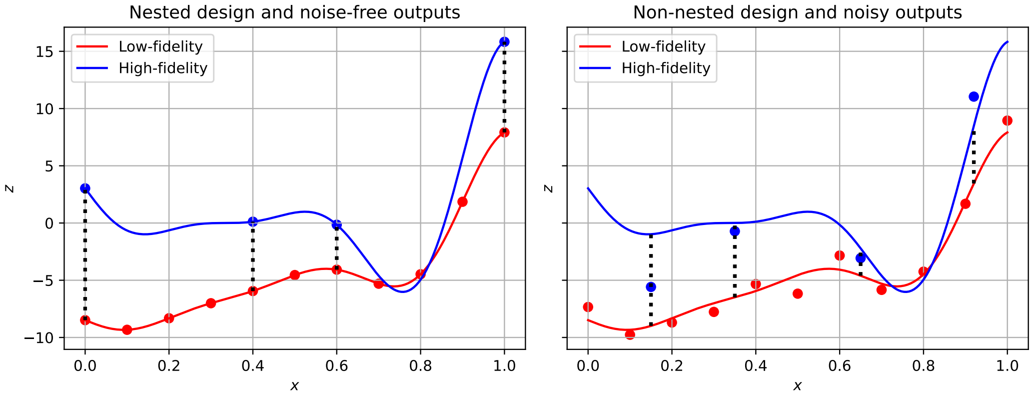
- Computer codes z_L, z_H, observations: $\begin{cases} z_i^L = z_L(x_i^L), & i \in \{1, \dots, N_L\} \\ z_j^H = z_H(x_j^H), & j \in \{1, \dots, N_H\} \end{cases}$
- Auto-regressive model AR(1) (Kennedy and O'Hagan 2000):

$$\begin{cases} Z_H(x) = \rho(x) Z_L(x) + \Delta_H(x) \\ Z_L(x) \perp \Delta_H(x) \end{cases} \implies \left(Z_H(x) \,\middle|\, \begin{cases} \forall i, Z_L(x_i^L) = z_i^L, \\ \forall j, Z_H(x_j^H) = z_j^H \end{cases} \right)$$

- Recursive formulation (Le Gratiet and Garnier 2014):

$$\begin{cases} Z_H(x) = \rho(x) \widetilde{Z}_L(x) + \Delta_H(x) \\ \widetilde{Z}_L(x) \perp \Delta_H(x) \\ \widetilde{Z}_L(x) = \left(Z_L(x) \,|\, \forall i, Z_L(x_i^L) = z_i^L \right) \end{cases} \implies \left(Z_H(x) \,|\, \forall j, Z_H(x_j^H) = z_j^H \right)$$

Noisy and non-nested data



➤ Generalization to non-nested and noisy data (Baillie et al. 2026, pre-print)

Noisy model

- Computer codes z_L, z_H are tainted by noise:

$$\begin{cases} z_L(x) = y_L(x) + \varepsilon_L(x), & \varepsilon_L(x) \sim \mathcal{N}(0, \sigma^2_{\varepsilon,L}) \\ z_H(x) = y_H(x) + \varepsilon_H(x), & \varepsilon_H(x) \sim \mathcal{N}(0, \sigma^2_{\varepsilon,H}) \end{cases}$$

- Noisy recursive AR(1) model:

$$\begin{cases} Y_H(x) = \rho(x)\widetilde{Y}_L(x) + \Delta_H(x) \\ \widetilde{Y}_L(x) \perp \Delta_H(x) \\ \widetilde{Y}_L(x) = \left(Y_L(x) \,|\, \forall i, Z_L(x_i^L) = z_i^L\right) \end{cases} \implies \widetilde{Y}_H(x) = \left(Y_H(x) \,|\, \forall j, Z_H(x_j^H) = z_j^H\right)$$

- Explicit distribution: $\widetilde{Y}_H(x) \sim \mathcal{N}(m_{Y_H}(x), V_{Y_H}(x,x))$

Parameter estimation

A priori GPs: $\begin{cases} Y_L(x) \sim \mathcal{N}(\mu_L(x), k_L(x,x)) \\ \Delta_H(x) \sim \mathcal{N}(\mu_H(x), k_H(x,x)), \end{cases}$ with:

$$\begin{cases} \mu_L(x) = \beta_L \\ \mu_H(x) = \beta_H \end{cases} \quad \text{and} \quad \begin{cases} k_L(x,x') = \sigma_L^2 \exp\left(-0.5 \cdot |x-x'|^2 / \theta_L^2\right) \\ k_H(x,x') = \sigma_H^2 \exp\left(-0.5 \cdot |x-x'|^2 / \theta_H^2\right) \end{cases}$$

Maximum Likelihood Estimation (MLE)

- Non-recursive formulation: Joint optimization, high-dimensional problem

	β_L	σ_L^2	θ_L	$\sigma_{\varepsilon,L}^2$
ρ	β_H	σ_H^2	θ_H	$\sigma_{\varepsilon,H}^2$

Parameter estimation

A priori GPs: $\begin{cases} Y_L(x) \sim \mathcal{N}(\mu_L(x), k_L(x,x)) \\ \Delta_H(x) \sim \mathcal{N}(\mu_H(x), k_H(x,x)), \end{cases}$ with:

$$\begin{cases} \mu_L(x) = \beta_L \\ \mu_H(x) = \beta_H \end{cases} \quad \text{and} \quad \begin{cases} k_L(x,x') = \sigma_L^2 \exp\big(-0.5 \cdot |x-x'|^2 / \theta_L^2\big) \\ k_H(x,x') = \sigma_H^2 \exp\big(-0.5 \cdot |x-x'|^2 / \theta_H^2\big) \end{cases}$$

- Recursive formulation: Decoupled optimization
  - Low-fidelity: Simple GP MLE

β_L	σ_L^2	θ_L	$\sigma_{\varepsilon,L}^2$

  - High-fidelity: Expectation-Maximization (EM) algorithm (Zertuche 2015)

ρ	β_H	σ_H^2	θ_H	$\sigma_{\varepsilon,H}^2$

Outline

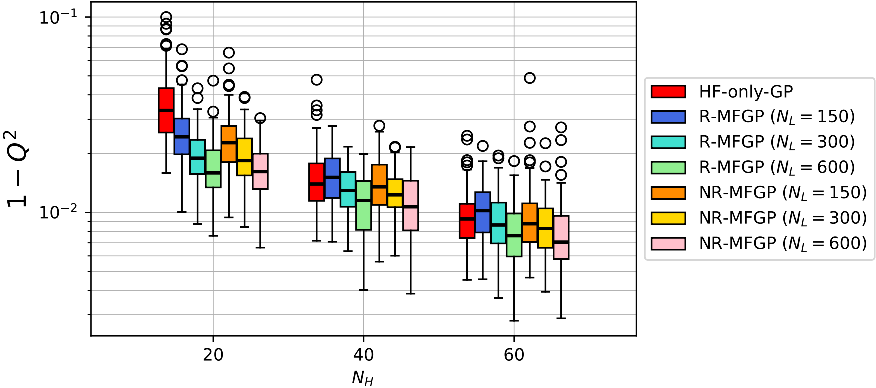
Analytic test case

- Park function (Cox et al. 2001):

$$\begin{cases} y_H(x_1,x_2,x_3,x_4) = \frac{x_1}{2}\left(\sqrt{1+(x_2+x_3^2)\frac{x_4}{x_1^2}}-1\right) + (x_1+3x_4)\exp(1+\sin(x_3)) \\ y_L(x_1,x_2,x_3,x_4) = (1+0.1\sin(x_1))\, y_H(x_1,x_2,x_3,x_4) - 2x_1 + x_2^2 + x_3^2 + 0.5 \end{cases}$$

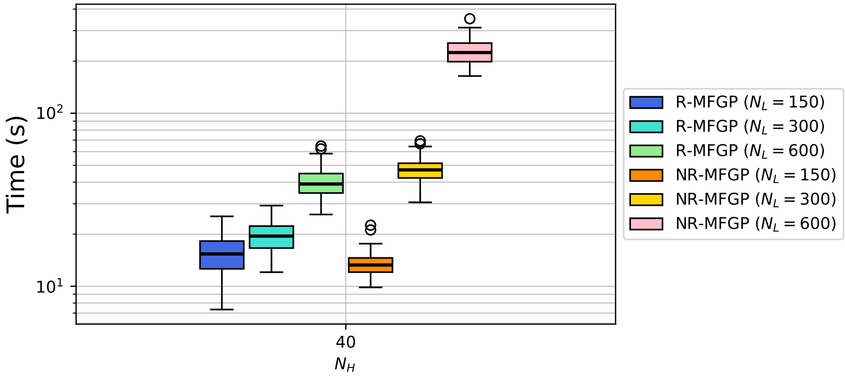
- Data: $N_L \in \{150, 300, 600\}$, $N_H \in \{20, 40, 60\}$, $\sigma_{\varepsilon,L} = 2.5$, $\sigma_{\varepsilon,H} = 0.5$
- Metrics: $1 - Q^2 = \frac{\sum_{t=1}^{N_{test}}(y(x_t)-m(x_t))^2}{\sum_{t=1}^{N_{test}}(y(x_t)-\bar{y})^2}$ and training time
- Comparison of the models on 100 repetitions:
  - Simple GP on HF data (HF-only-GP)
  - Recursive (R-MFGP): decoupled optimization
  - Non-recursive (NR-MFGP): joint optimization

Results: Prediction error



- Multi-fidelity is relevant for few HF samples
- Similar performance for R-MFGP and NR-MFGP

Results: Training time



- The recursive approach yields faster training times for large LF datasets
- The influence of N_H on the training time is negligible

Outline

Conclusion and Prospects

- Auto-regressive MFGP model : Generalization of the recursive formulation in the case of non-nested designs and noisy outputs
- Improve the estimation of the high-fidelity noise variance $\sigma^2_{\varepsilon,H}$
- Heteroscedastic model: non-constant noise variances
- Sequential design of experiments

References I

Nils Baillie, Baptiste Kerleguer, Cyril Feau, and Josselin Garnier (2026). *Multi-fidelity Gaussian process regression for noisy outputs and non-nested experimental designs: a comparison between the recursive and non-recursive formulations*. URL: `https://arxiv.org/abs/2511.20183v2`.

Dennis D Cox, Jeong-Soo Park, and Clifford E Singer (2001). “A statistical method for tuning a computer code to a data base”. In: *Computational Statistics & Data Analysis* 37.1, pp. 77–92. ISSN: 0167-9473. DOI: `10.1016/S0167-9473(00)00057-8`.

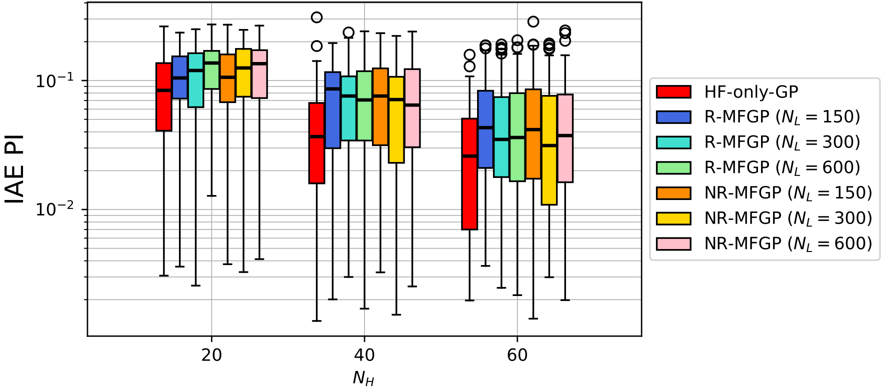
Marc Kennedy and Anthony O’Hagan (Oct. 2000). “Predicting the Output from a Complex Computer Code When Fast Approximations Are Available”. In: *Biometrika* 87. DOI: `10.1093/biomet/87.1.1`.

Loic Le Gratiet and Josselin Garnier (2014). “Recursive co-kriging model for Design of Computer experiments with multiple levels of fidelity”. In: *International Journal for Uncertainty Quantification* 4.5, pp. 365–386. DOI: `10.1615/Int.J.UncertaintyQuantification.2014006914`.

Federico Zertuche (Oct. 2015). “Assessment of uncertainty in computer experiments when working with multifidelity simulators.”. Theses. Université Grenoble Alpes.

Thank you for your attention !

Results: Prediction uncertainty



- Simple GP tends to be more reliable when adding HF data
- Similar performance for R-MFGP and NR-MFGP

Results: Prediction interval width