ENM 540: Data-driven modeling and probabilistic scientific computing

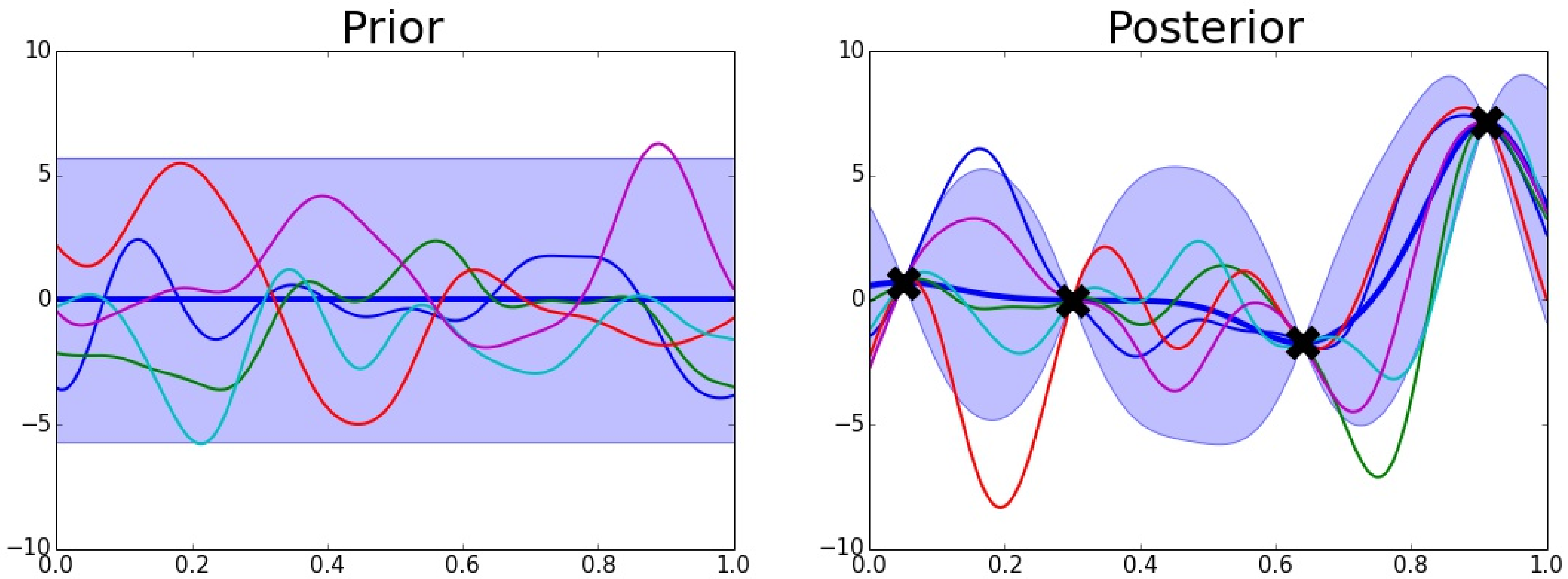
Multi-output Gaussian processes and multi-fidelity modeling

Paris Perdikaris
March 1st, 2018

Penn
University *of* Pennsylvania

Data-driven modeling with Gaussian processes

$$y = f(\boldsymbol{x}) + \epsilon \qquad f \sim \mathcal{GP}(0, k(\boldsymbol{x}, \boldsymbol{x}'; \boldsymbol{\theta}))$$



Training via maximizing the marginal likelihood

$$\log p(\boldsymbol{y}|\boldsymbol{X}, \boldsymbol{\theta}) = -\frac{1}{2}\log|\boldsymbol{K} + \sigma_\epsilon^2 \boldsymbol{I}| - \frac{1}{2}\boldsymbol{y}^T(\boldsymbol{K} + \sigma_\epsilon^2\boldsymbol{I})^{-1}\boldsymbol{y} - \frac{N}{2}\log 2\pi$$

Prediction via conditioning on available data

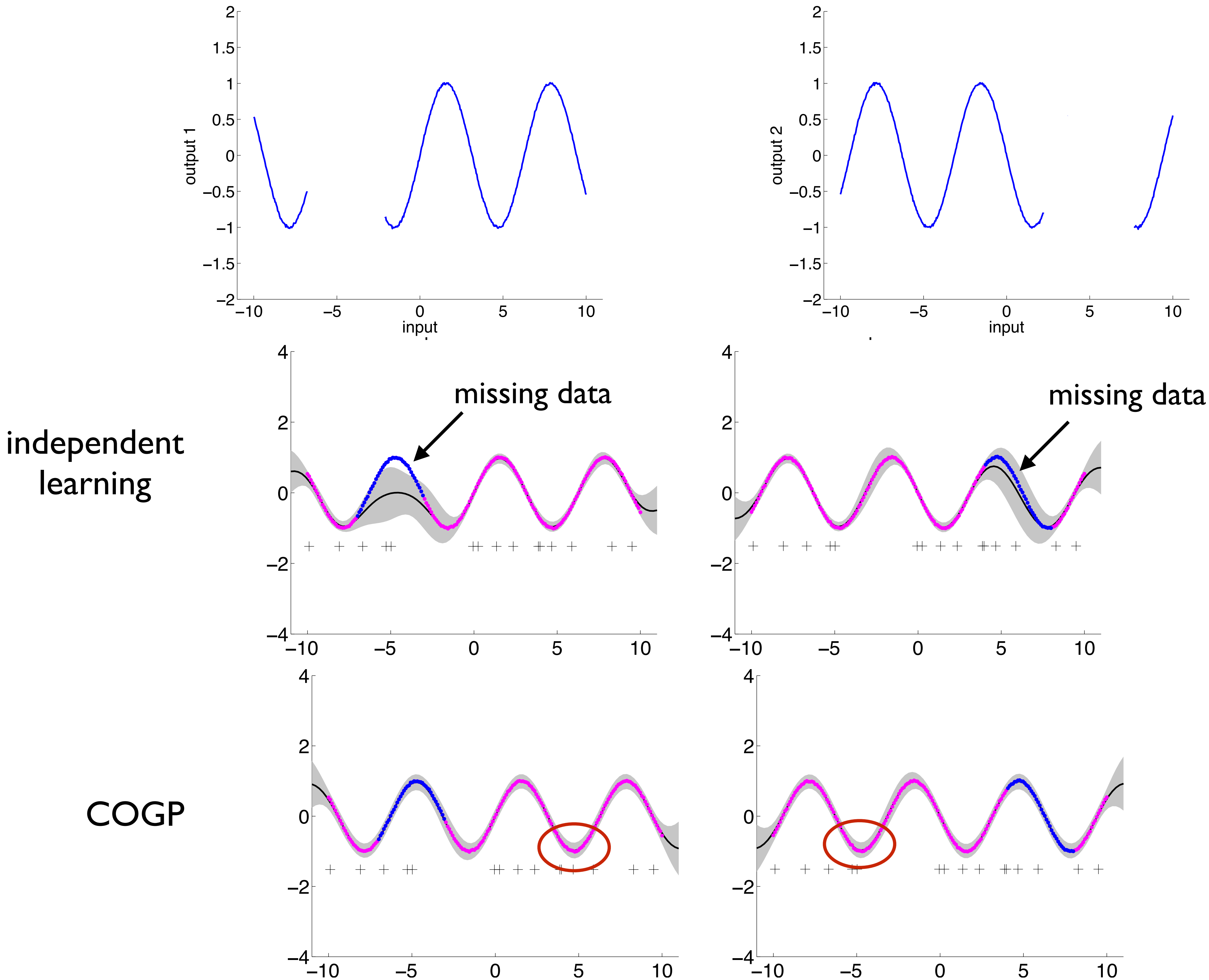
$$\begin{aligned} p(f_*|\boldsymbol{y}, \boldsymbol{X}, \boldsymbol{x}_*) &= \mathcal{N}(f_*|\mu_*, \sigma_*^2), \\ \mu_*(\boldsymbol{x}_*) &= \boldsymbol{k}_{*N}(\boldsymbol{K} + \sigma_\epsilon^2\boldsymbol{I})^{-1}\boldsymbol{y}, \\ \sigma_*^2(\boldsymbol{x}_*) &= \boldsymbol{k}_{**} - \boldsymbol{k}_{*N}(\boldsymbol{K} + \sigma_\epsilon^2\boldsymbol{I})^{-1}\boldsymbol{k}_{N*}, \end{aligned}$$

Demo code: https://github.com/PredictiveIntelligenceLab/GPTutorial

Multi-output Gaussian process regression

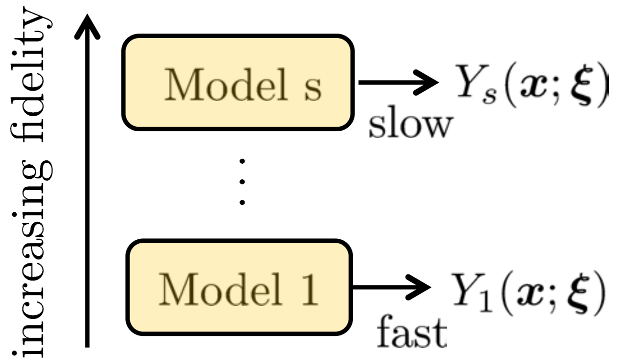
Learn two correlated tasks (outputs) with lots of data + missing data

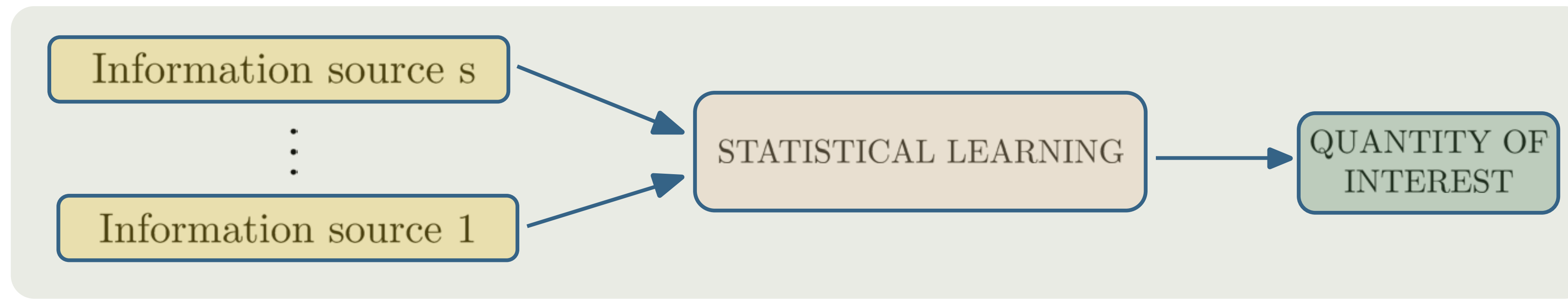
output 2 = - output 1



Nguyen, T. V., & Bonilla, E. V. (2014, July). Collaborative Multi-output Gaussian Processes. In UAI (pp. 643-652).

Multi-fidelity modeling

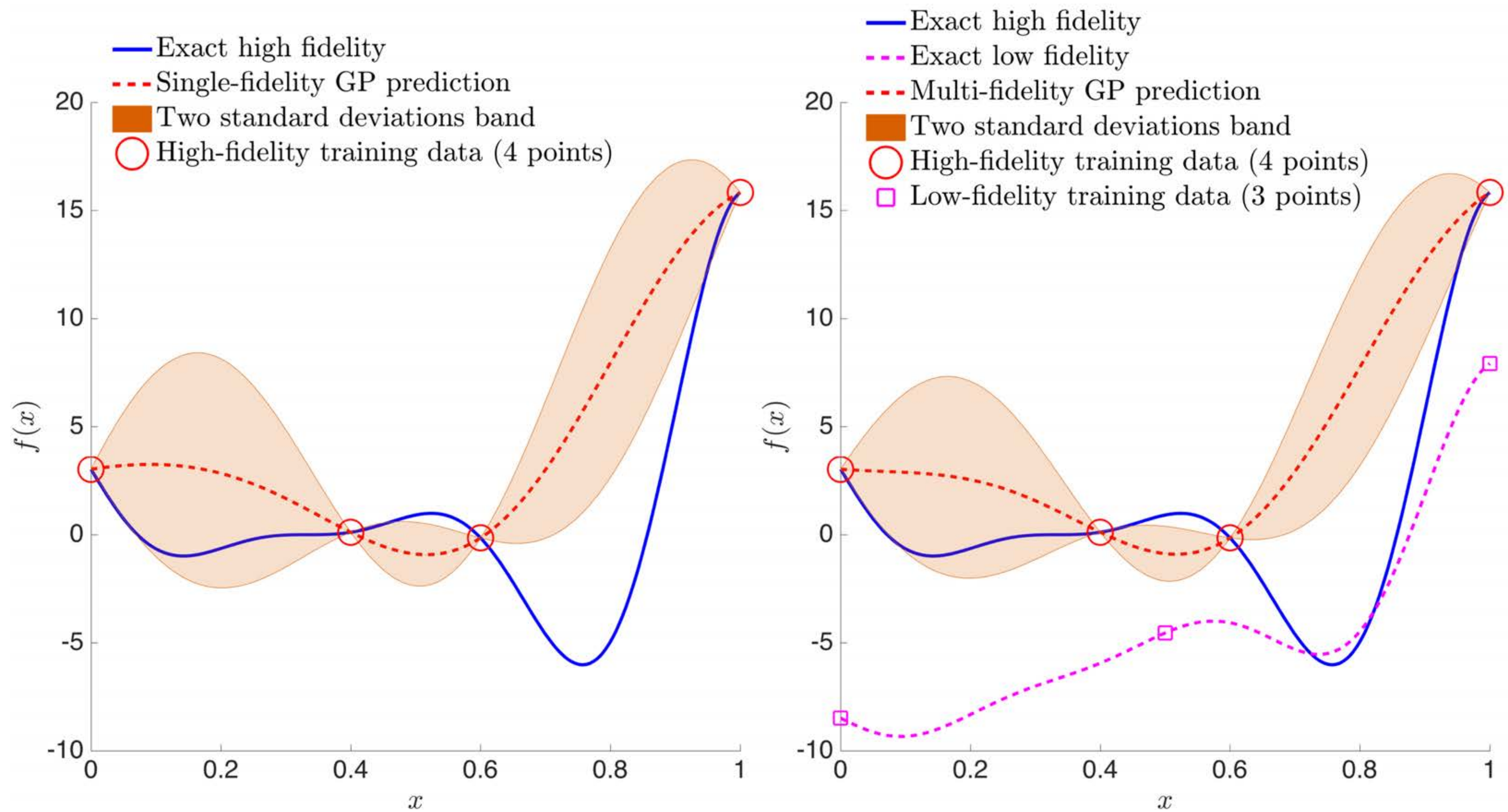




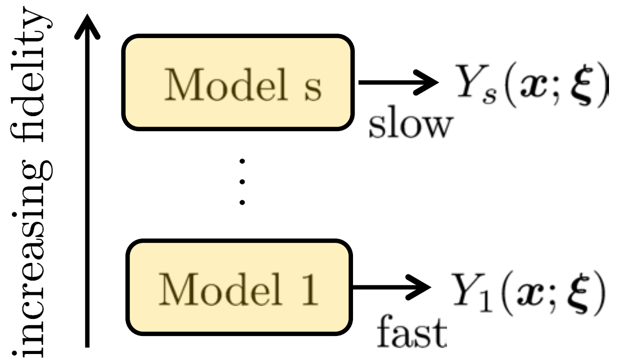
Number of runs is limited by time and computational resources

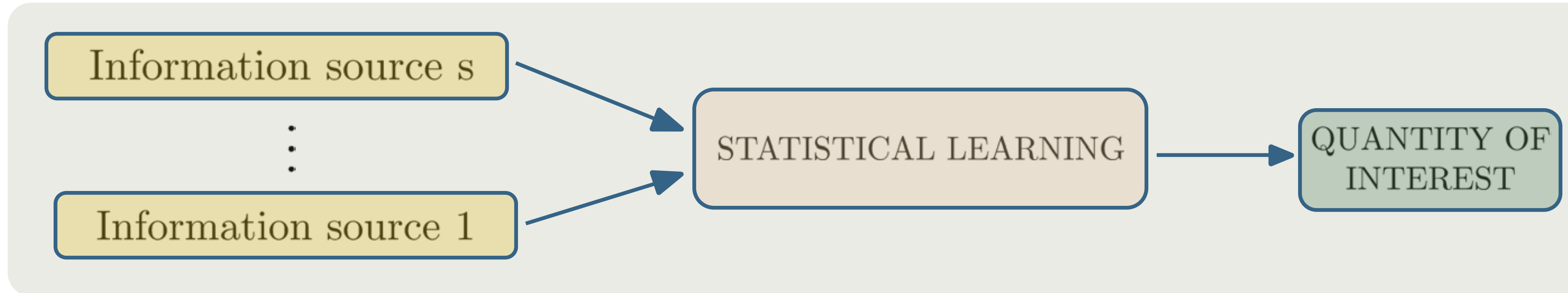
We cannot compute at all $(\boldsymbol{x};\boldsymbol{\xi})$

Prediction of $Z_i(\boldsymbol{x}) = \mathbb{E}[f(Y_i(\boldsymbol{x};\boldsymbol{\xi}))]$ is a problem of statistical inference



Multi-fidelity modeling

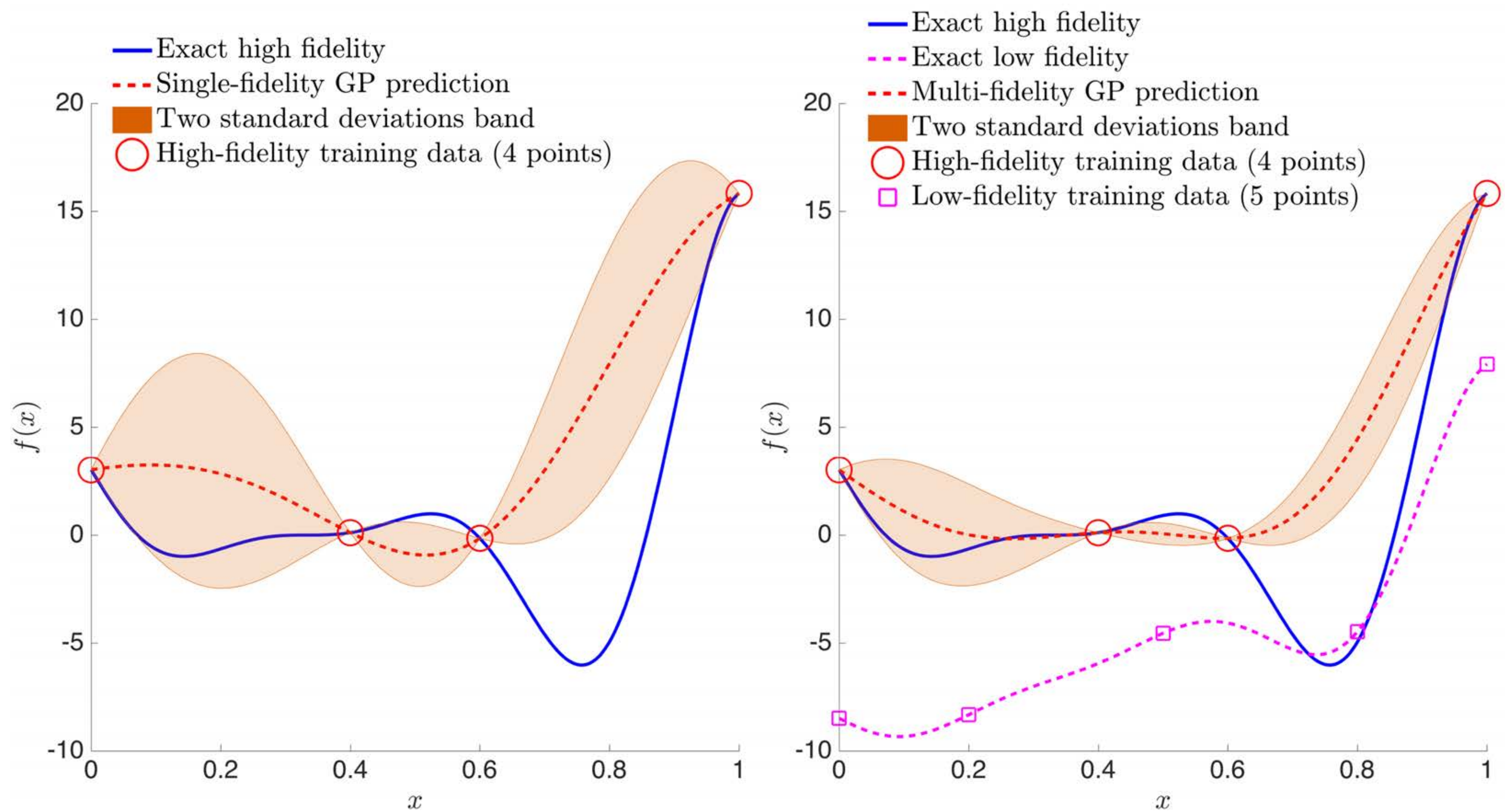




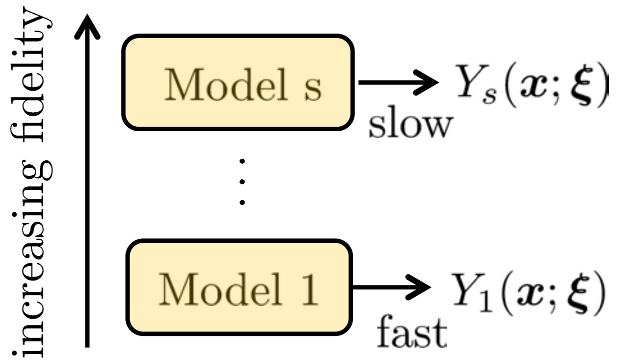
Number of runs is limited by time
and computational resources

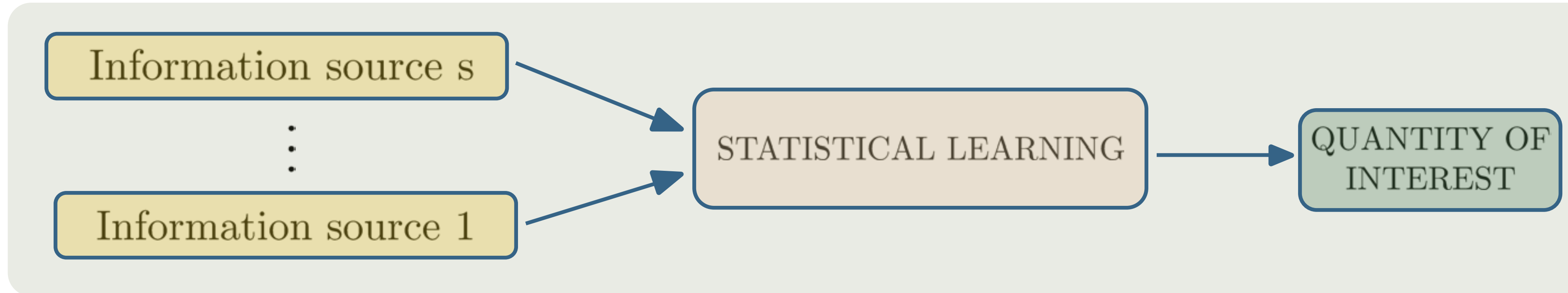
We cannot compute at all $(\boldsymbol{x};\boldsymbol{\xi})$

Prediction of $Z_i(\boldsymbol{x}) = \mathbb{E}[f(Y_i(\boldsymbol{x};\boldsymbol{\xi}))]$ is a
problem of statistical inference



Multi-fidelity modeling

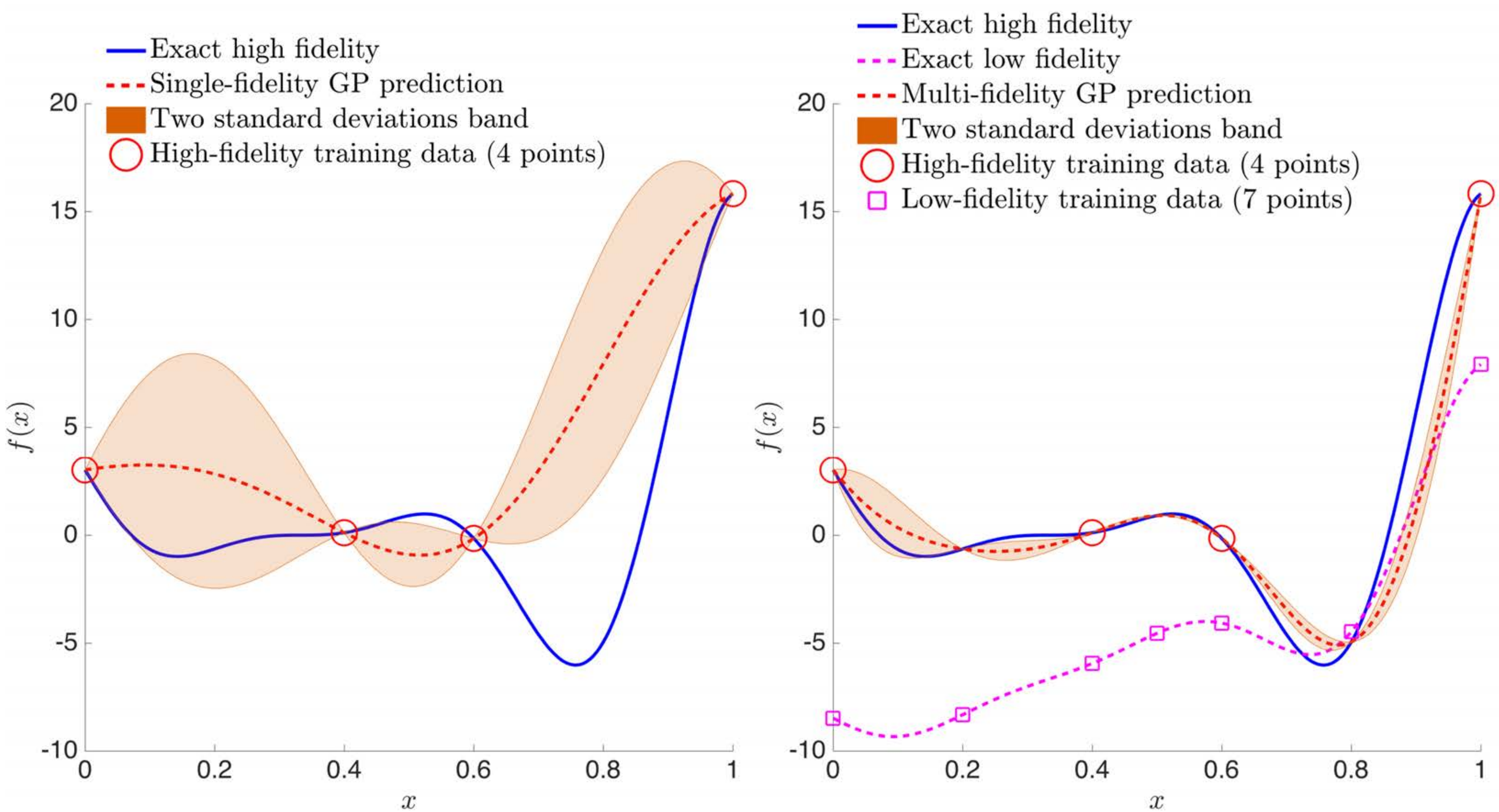




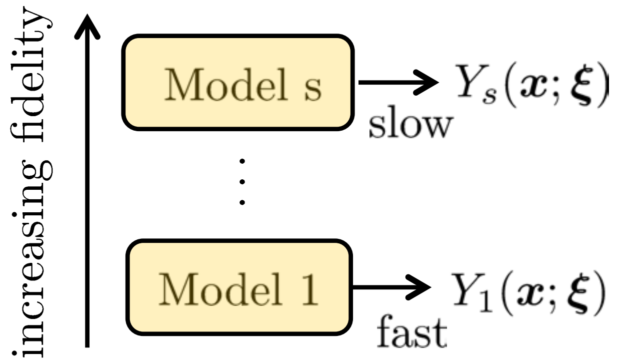
Number of runs is limited by time and computational resources

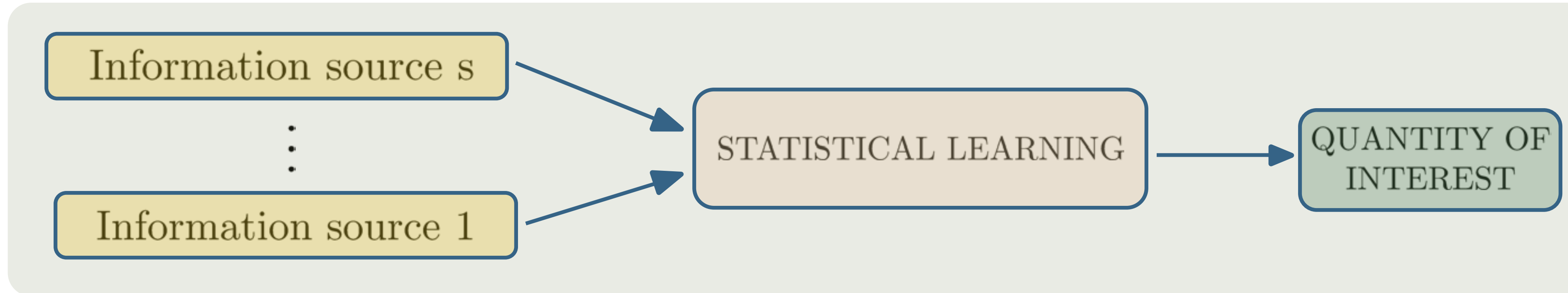
We cannot compute at all $(\boldsymbol{x};\boldsymbol{\xi})$

Prediction of $Z_i(\boldsymbol{x}) = \mathbb{E}[f(Y_i(\boldsymbol{x};\boldsymbol{\xi}))]$ is a problem of statistical inference



Multi-fidelity modeling

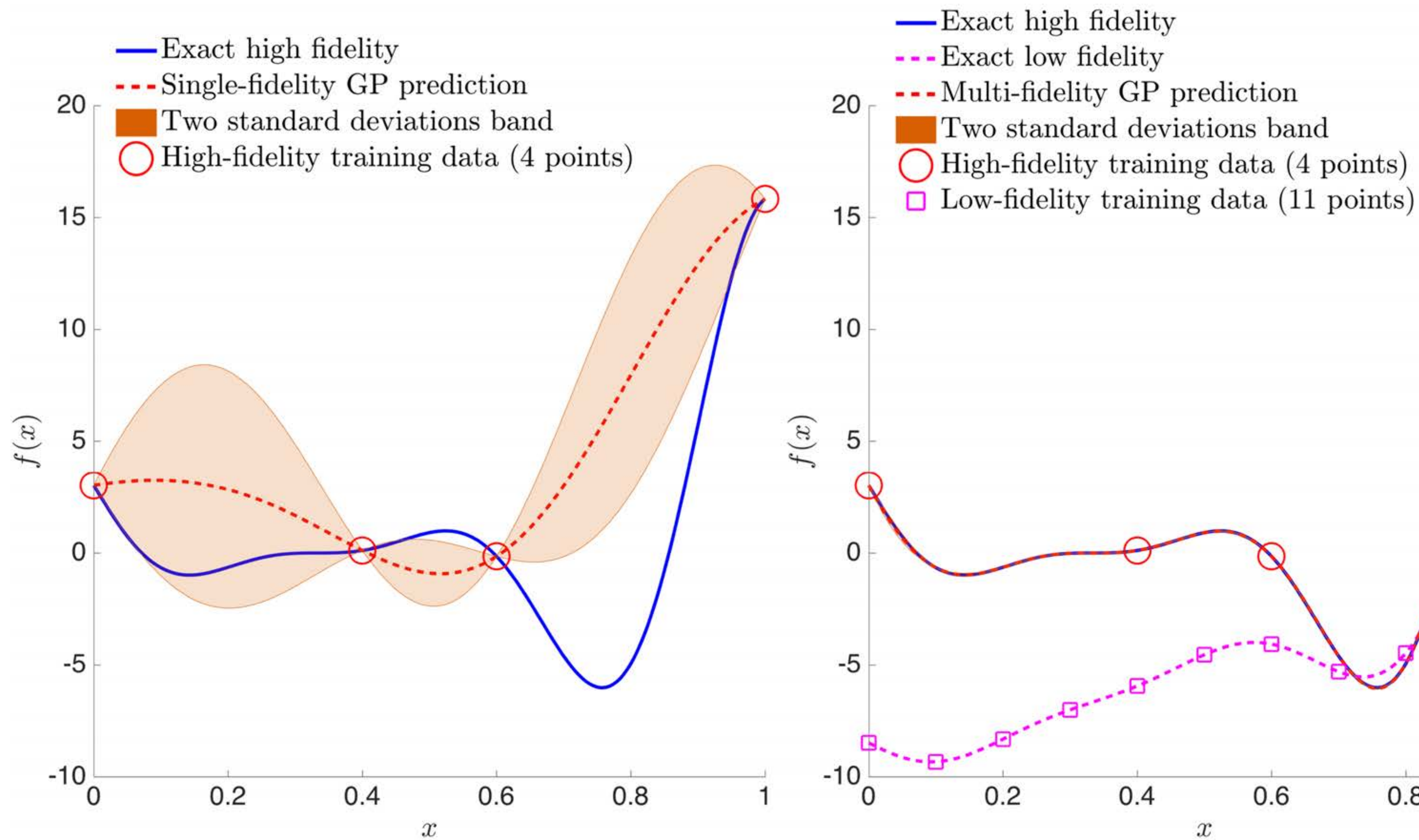




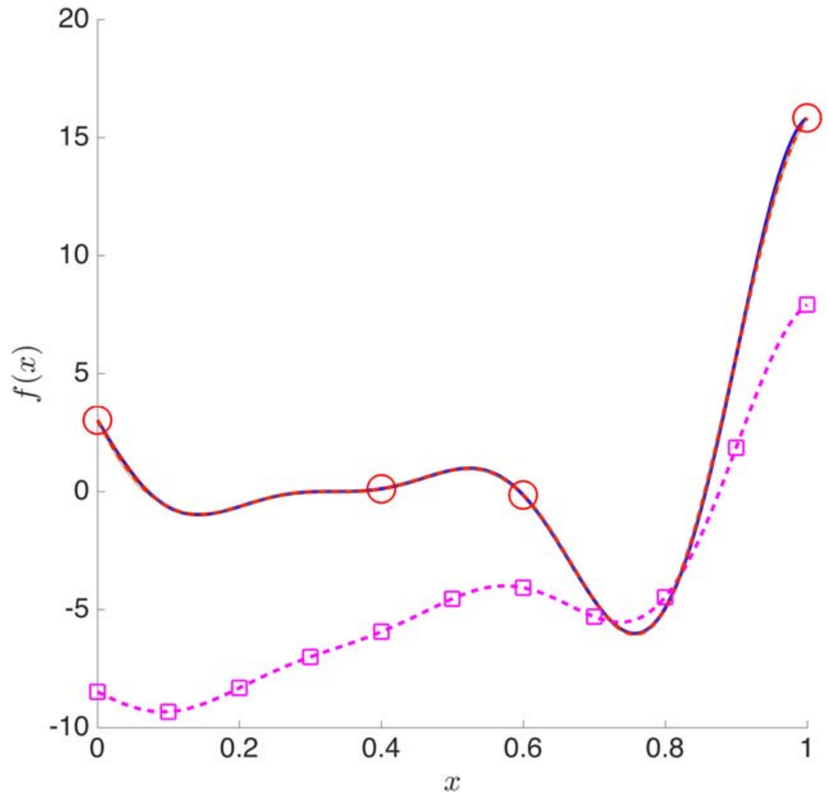
Number of runs is limited by time
and computational resources

We cannot compute at all $(\boldsymbol{x};\boldsymbol{\xi})$

Prediction of $Z_i(\boldsymbol{x}) = \mathbb{E}[f(Y_i(\boldsymbol{x};\boldsymbol{\xi}))]$ is a
problem of statistical inference



Multi-fidelity modeling with GPs



Multi-fidelity observations:

$$\boldsymbol{y}_L = f_L(\boldsymbol{x}_L) + \boldsymbol{\epsilon}_L$$
$$\boldsymbol{y}_H = f_H(\boldsymbol{x}_H) + \boldsymbol{\epsilon}_H$$

Probabilistic model:

$$f_H(\boldsymbol{x}) = \rho f_L(\boldsymbol{x}) + \delta(\boldsymbol{x})$$
$$f_L(\boldsymbol{x}) \sim \mathcal{GP}(0, k_L(\boldsymbol{x}, \boldsymbol{x}'; \theta_L))$$
$$\delta(\boldsymbol{x}) \sim \mathcal{GP}(0, k_H(\boldsymbol{x}, \boldsymbol{x}'; \theta_H))$$
$$\epsilon_L \sim \mathcal{N}(0, \sigma^2_{\epsilon_L} \boldsymbol{I})$$
$$\epsilon_H \sim \mathcal{N}(0, \sigma^2_{\epsilon_H} \boldsymbol{I})$$

Training:

$$\boldsymbol{y} = \begin{bmatrix} \boldsymbol{y}_L \\ \boldsymbol{y}_H \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} k_L(\boldsymbol{x}_L, \boldsymbol{x}'_L; \theta_L) + \sigma^2_{\epsilon_L} \boldsymbol{I} & \rho k_L(\boldsymbol{x}_L, \boldsymbol{x}'_H; \theta_L) \\ \rho k_L(\boldsymbol{x}_H, \boldsymbol{x}'_L; \theta_L) & \rho^2 k_L(\boldsymbol{x}_H, \boldsymbol{x}'_H; \theta_L) + k_H(\boldsymbol{x}_H, \boldsymbol{x}'_H; \theta_H) + \sigma^2_{\epsilon_H} \boldsymbol{I} \end{bmatrix} \right)$$

$$\boldsymbol{X} = \begin{bmatrix} \boldsymbol{x}_L \\ \boldsymbol{x}_H \end{bmatrix} \qquad -\log p(\boldsymbol{y} | \boldsymbol{X}, \theta_L, \theta_H, \rho, \sigma^2_{\epsilon_L}, \sigma^2_{\epsilon_H}) = \frac{1}{2} \log |\boldsymbol{K}| + \frac{1}{2} \boldsymbol{y}^T \boldsymbol{K}^{-1} \boldsymbol{y} - \frac{N_L + N_H}{2} \log 2\pi$$

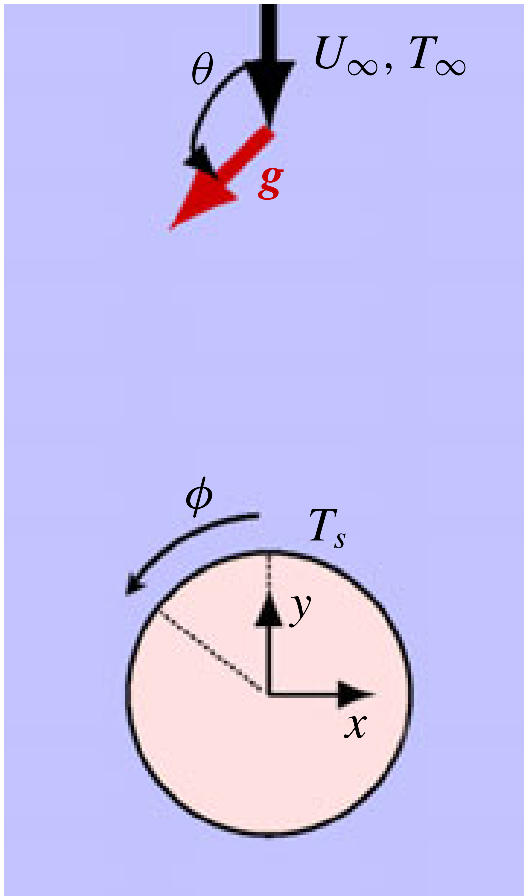
Prediction:

$$p(f(\boldsymbol{x}^*) | \boldsymbol{y}, \boldsymbol{X}, \boldsymbol{x}^*) \sim \mathcal{N}(f(\boldsymbol{x}^*) | \mu(\boldsymbol{x}^*), \sigma^2(\boldsymbol{x}^*))$$
$$\mu(\boldsymbol{x}^*) = \boldsymbol{k}(\boldsymbol{x}^*, \boldsymbol{X}) \boldsymbol{K}^{-1} \boldsymbol{y}$$
$$\sigma(\boldsymbol{x}^*) = \boldsymbol{k}(\boldsymbol{x}^*, \boldsymbol{x}^*) - \boldsymbol{k}(\boldsymbol{x}^*, \boldsymbol{X}) \boldsymbol{K}^{-1} \boldsymbol{k}(\boldsymbol{X}, \boldsymbol{x}^*)$$

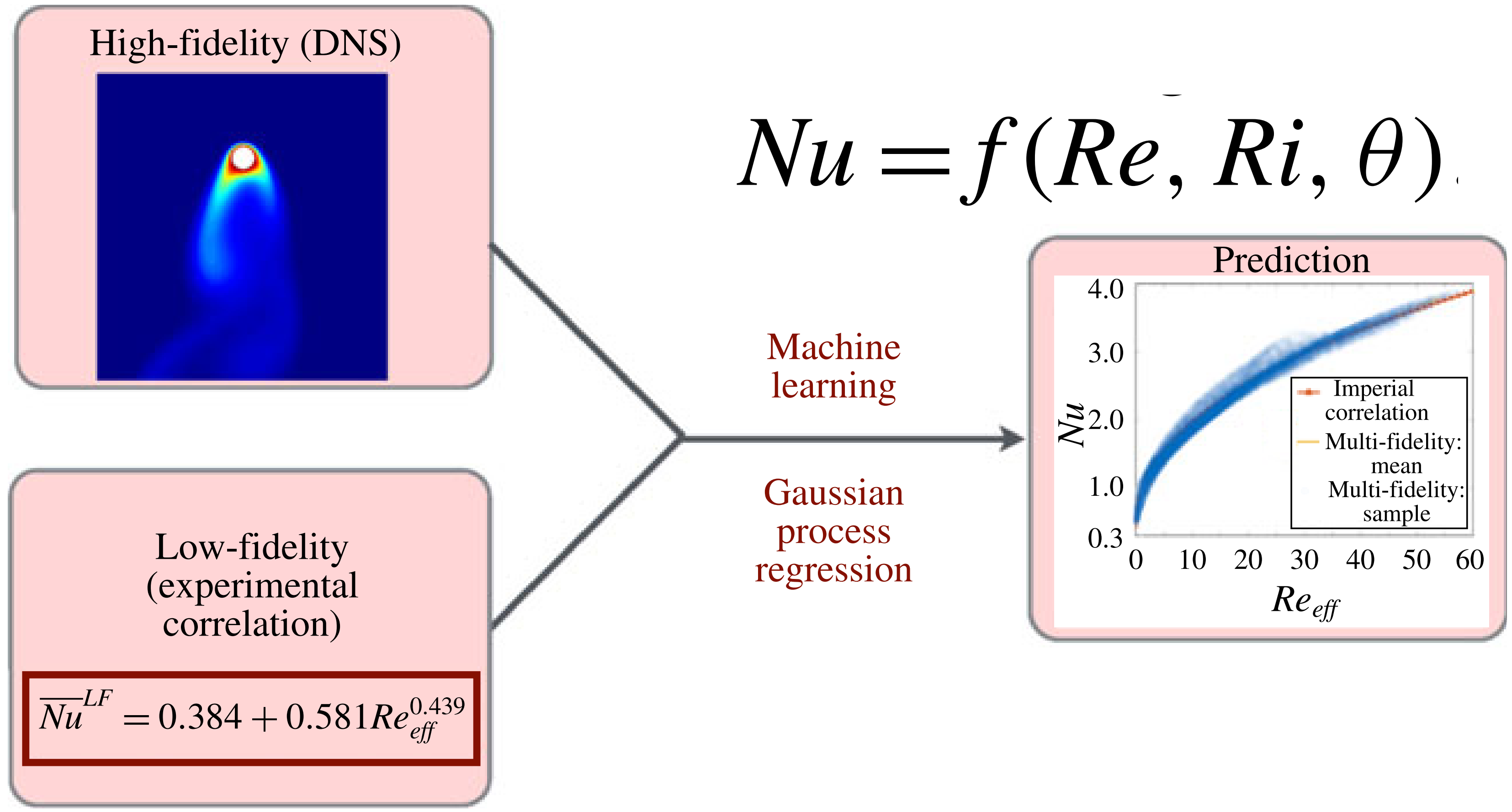
M.C Kennedy, and A. O'Hagan. *Predicting the output from a complex computer code when fast approximations are available*, 2000.

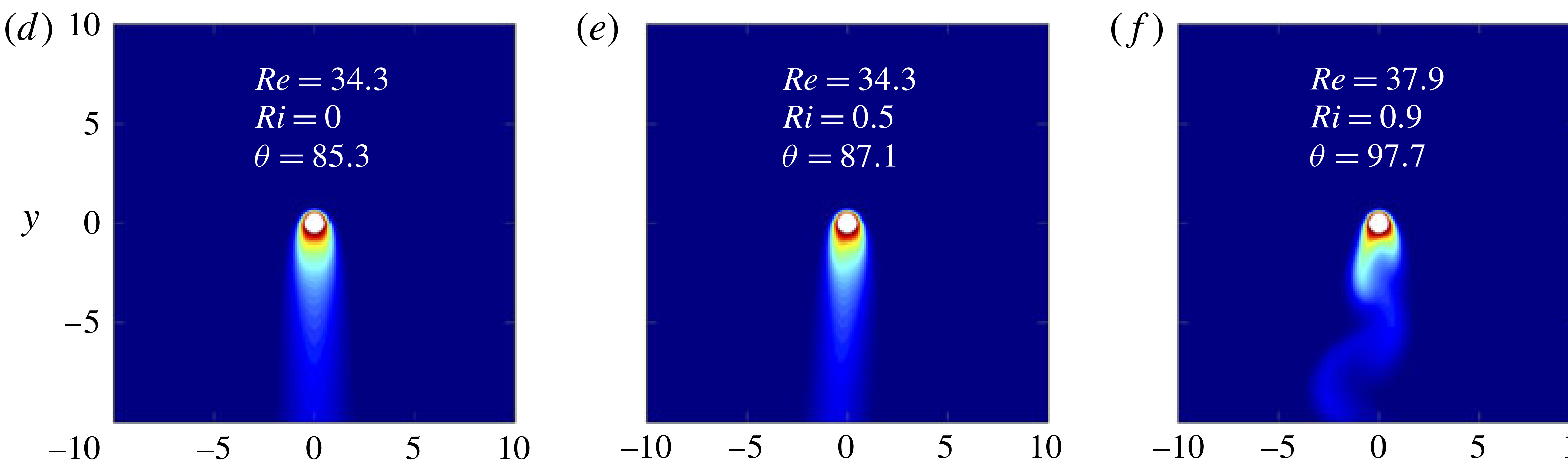
Demo code: https://github.com/PredictiveIntelligenceLab/GPTutorial

Applications: Mixed convection



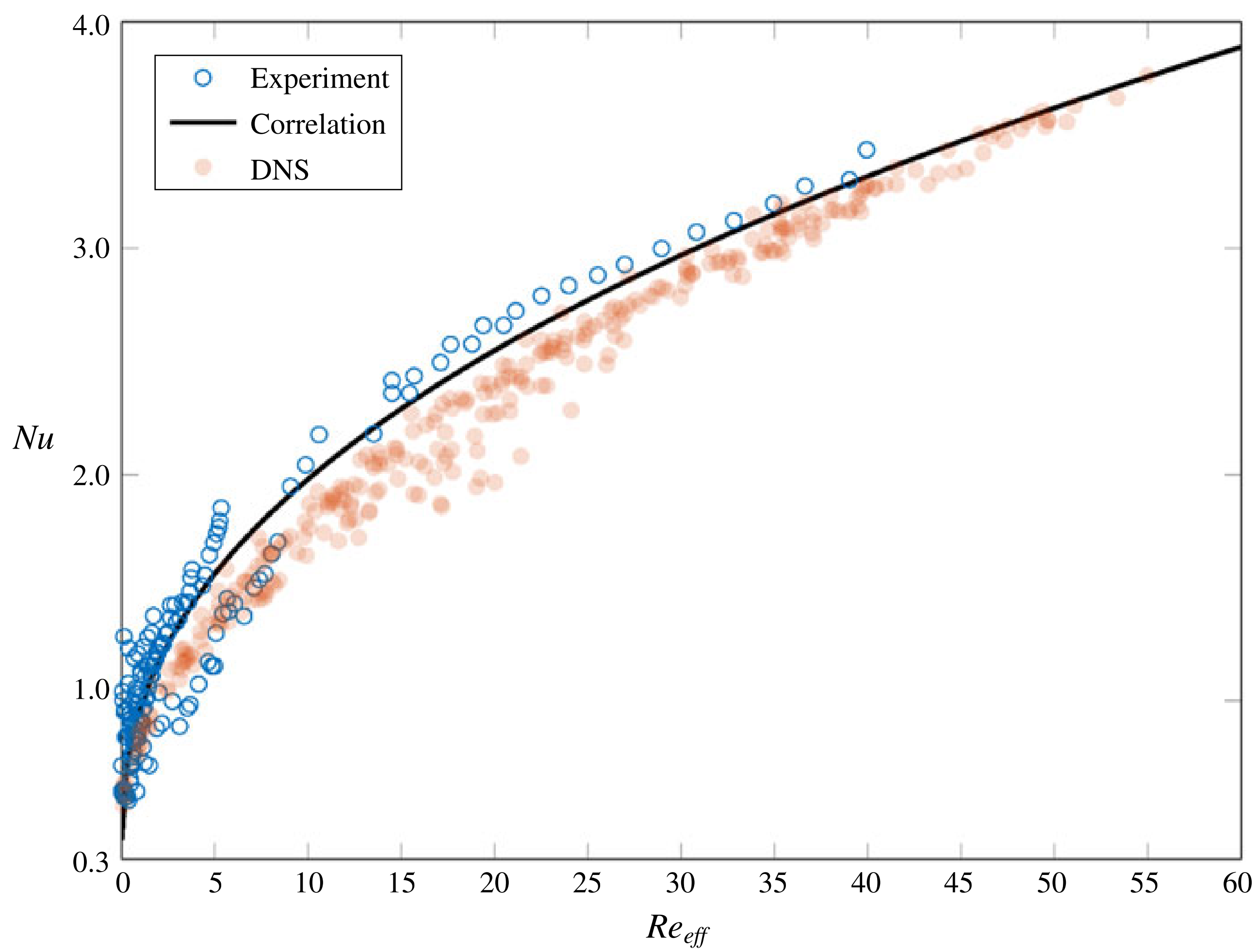
$$Nu = f(Re, Ri, \theta)$$

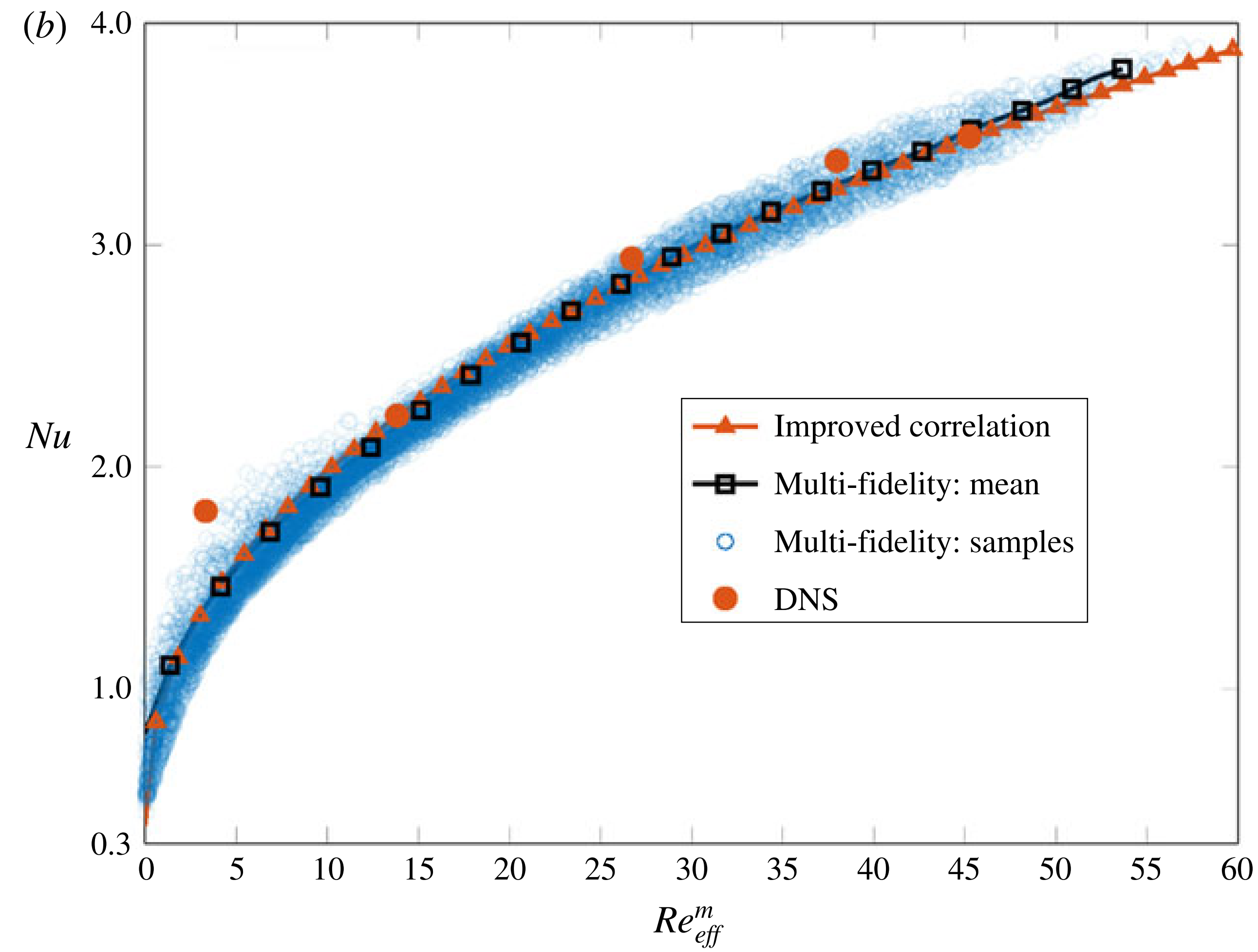




Babaee, H., Perdikaris, P., Chryssostomidis, C., & Karniadakis, G. E. (2016). Multi-fidelity modelling of mixed convection based on experimental correlations and numerical simulations. Journal of Fluid Mechanics, 809, 895-917.

Applications: Mixed convection

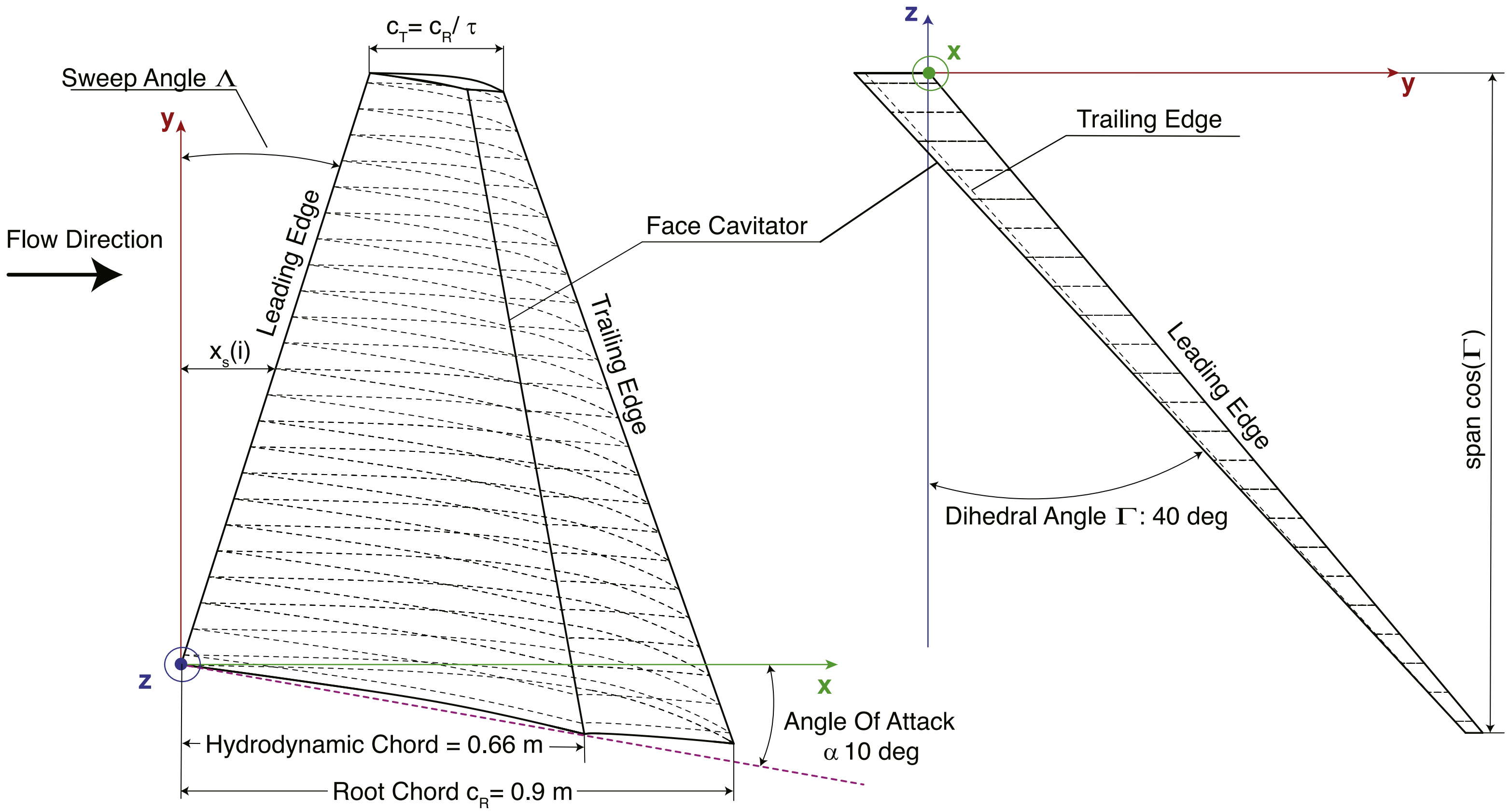




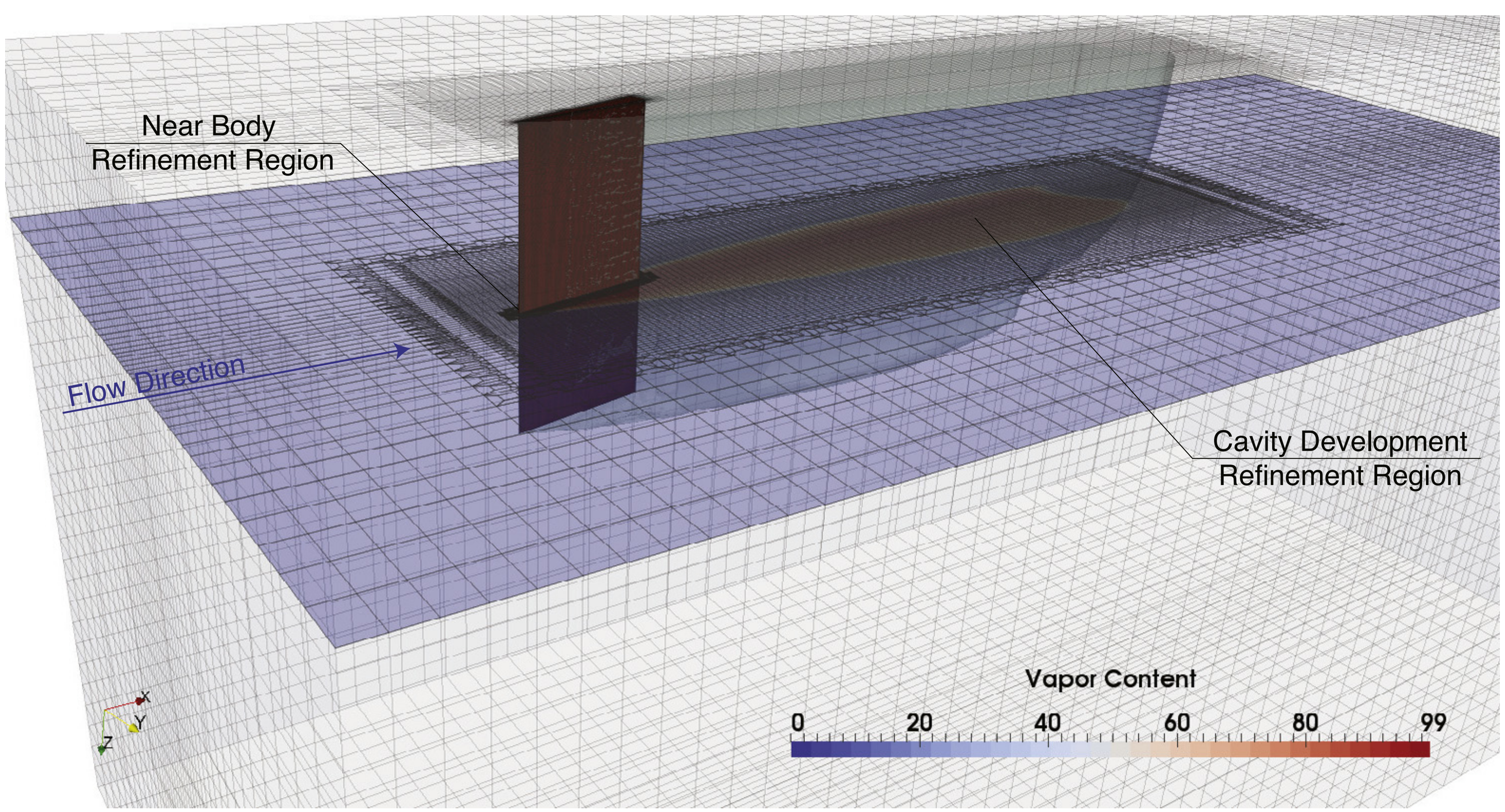
$g(\theta, Ri) = 1 - 0.38 Ri \sin(\theta)$

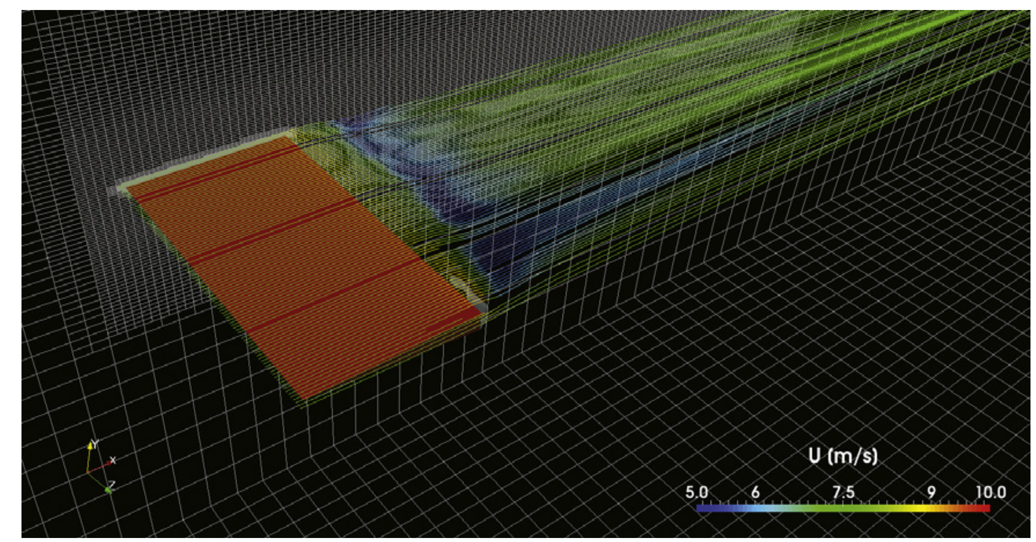
Babaee, H., Perdikaris, P., Chryssostomidis, C., & Karniadakis, G. E. (2016). Multi-fidelity modelling of mixed convection based on experimental correlations and numerical simulations. Journal of Fluid Mechanics, 809, 895-917.

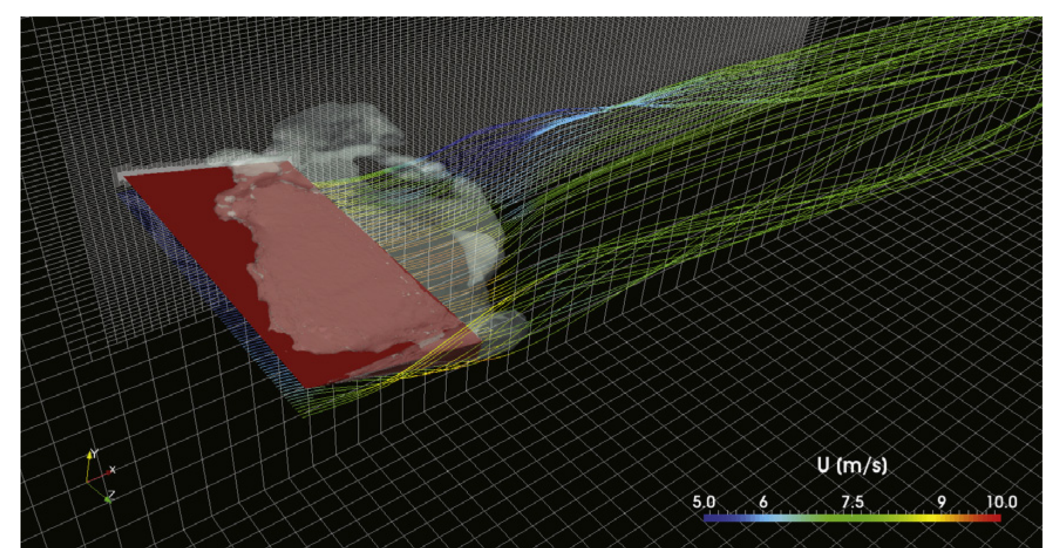
Applications: Design optimization



Geometry obtained through the combination of 17 design variables.

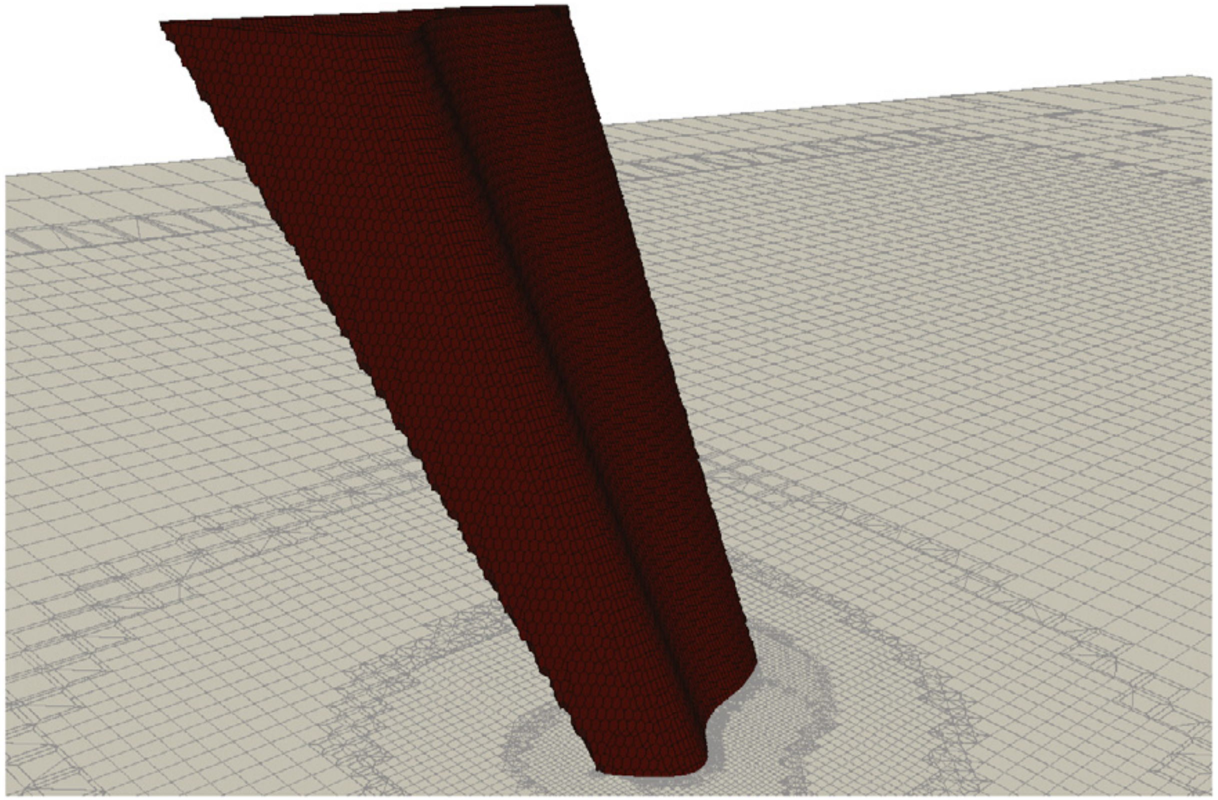




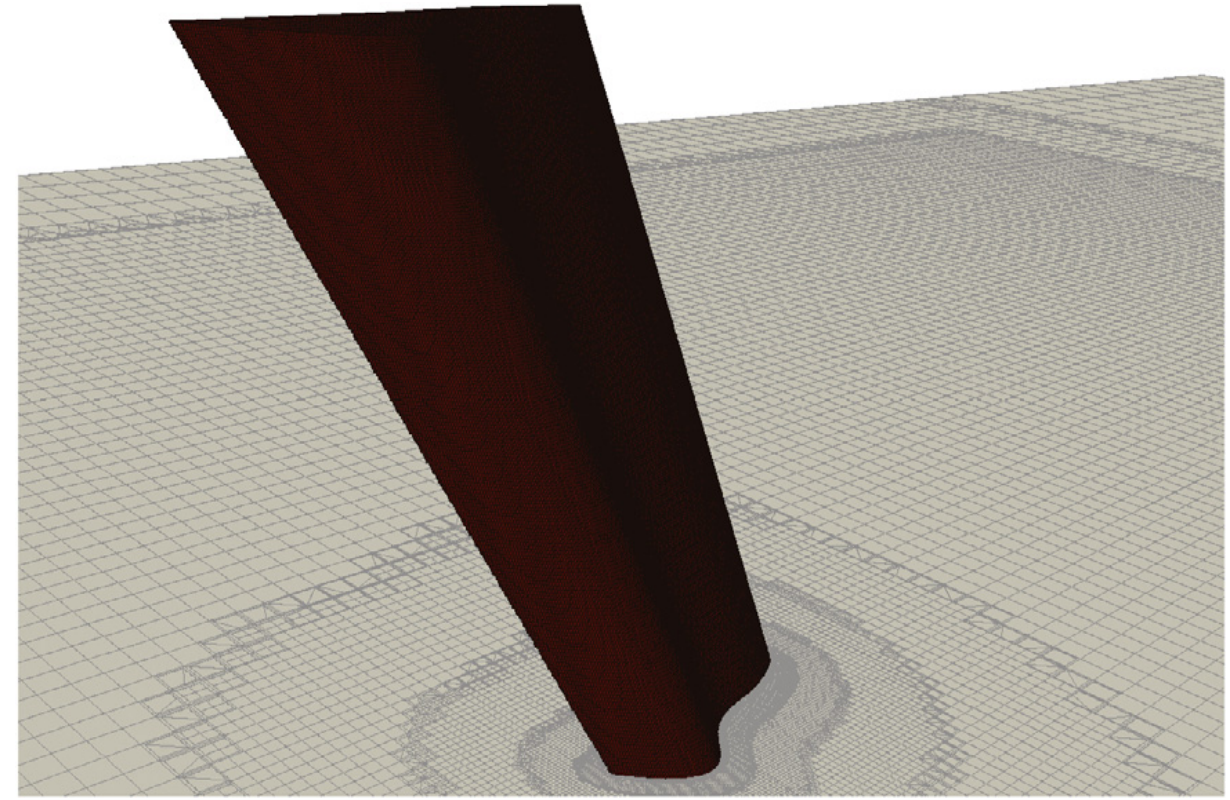


Bonfiglio, L., Perdikaris, P., Brizzolara, S., & Karniadakis, G. E. (2017). Multi-fidelity optimization of super-cavitating hydrofoils. Computer Methods in Applied Mechanics and Engineering.

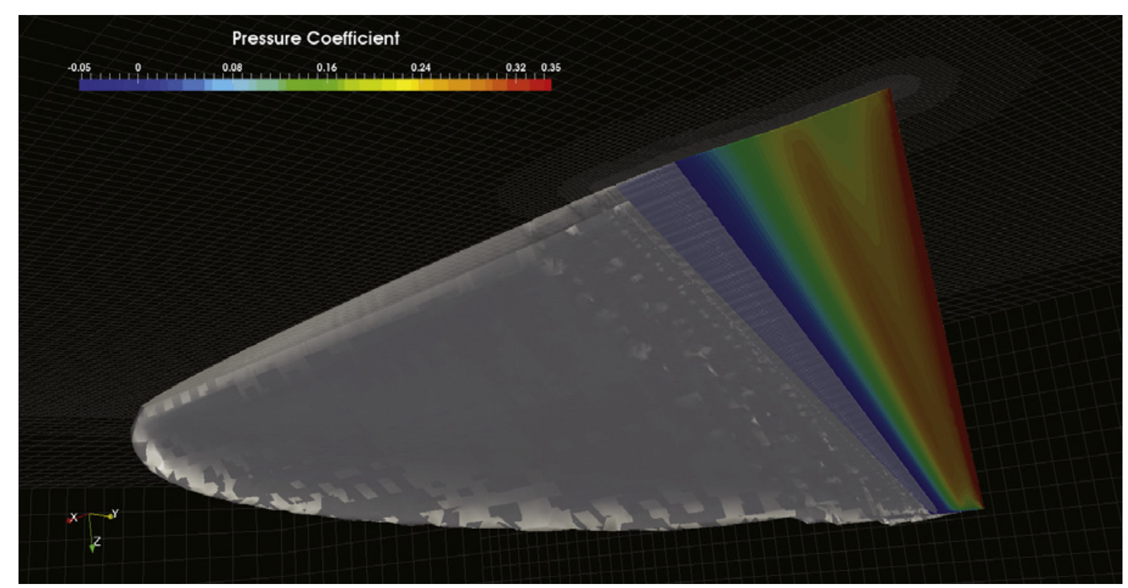
Applications: Design optimization



(a) Low fidelity resolution.



(b) High fidelity resolution.



$$
\begin{aligned}
\underset{\boldsymbol{x}\in\mathcal{X}\subset\mathbb{R}^{17}}{\text{minimize}} \quad & (C_D(\boldsymbol{x})/C_L(\boldsymbol{x}))/(C_D/C_L)_B \\
\text{subject to} \quad & 0.09 \leq C_L(\boldsymbol{x}) \leq 0.11, \\
& tk_{25\%b}(\boldsymbol{x}),\ tk_{50\%b}(\boldsymbol{x}),\ tk_{75\%b}(\boldsymbol{x}) \geq 1.32 \text{ cm.}
\end{aligned}
$$

Cost of one simulation:

High-fidelity: 3h on 256 cores.

Low-fidelity: 2h on 32 cores.

Bonfiglio, L., Perdikaris, P., Brizzolara, S., & Karniadakis, G. E. (2017). Multi-fidelity optimization of super-cavitating hydrofoils. Computer Methods in Applied Mechanics and Engineering.