Statistical finite elements for synthesis of observation data and model predictions

Fehmi Cirak

University of Cambridge

UKACM-SEMNI Autumn School

Data-Centric Engineering in Computational Mechanics

16–19 September 2024

Today's Schedule

- 09:00–09:50 Bayesian Inverse Problems
  - 10:00–11:00 Hands-on Session
- 11:15–12:10 Gaussian Process Surrogates
  - 13:30–14:30 Hands-on Session

- 14:45–15:35 Statistical Finite Elements
  - 15:45–16:45 Hands-on Session
- 16:45–17:00 Summary & Discussion

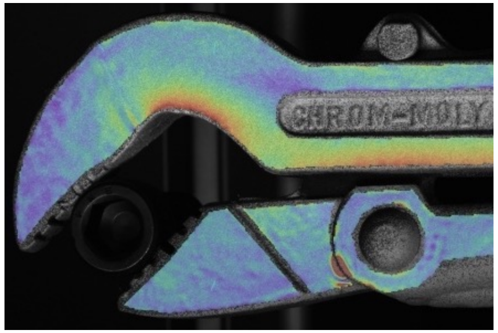
Asking questions by typing in Q&A or raising hand both are fine

Motivation

- Engineering systems are designed using inherently misspecified finite element models
  - Material properties, geometric dimensions, operating conditions etc. only partially known
  - In current practice uncertainties are considered through codified safety factors

- In-situ monitoring of engineering systems is becoming increasingly common providing abundant operational data

Digital image correlation

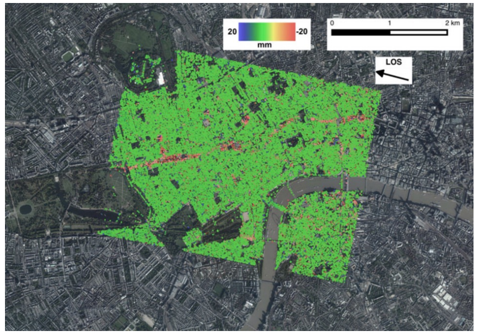


(idisc.org)

Fibre optic sensors

(Butler et al.)

Satellite remote sensing

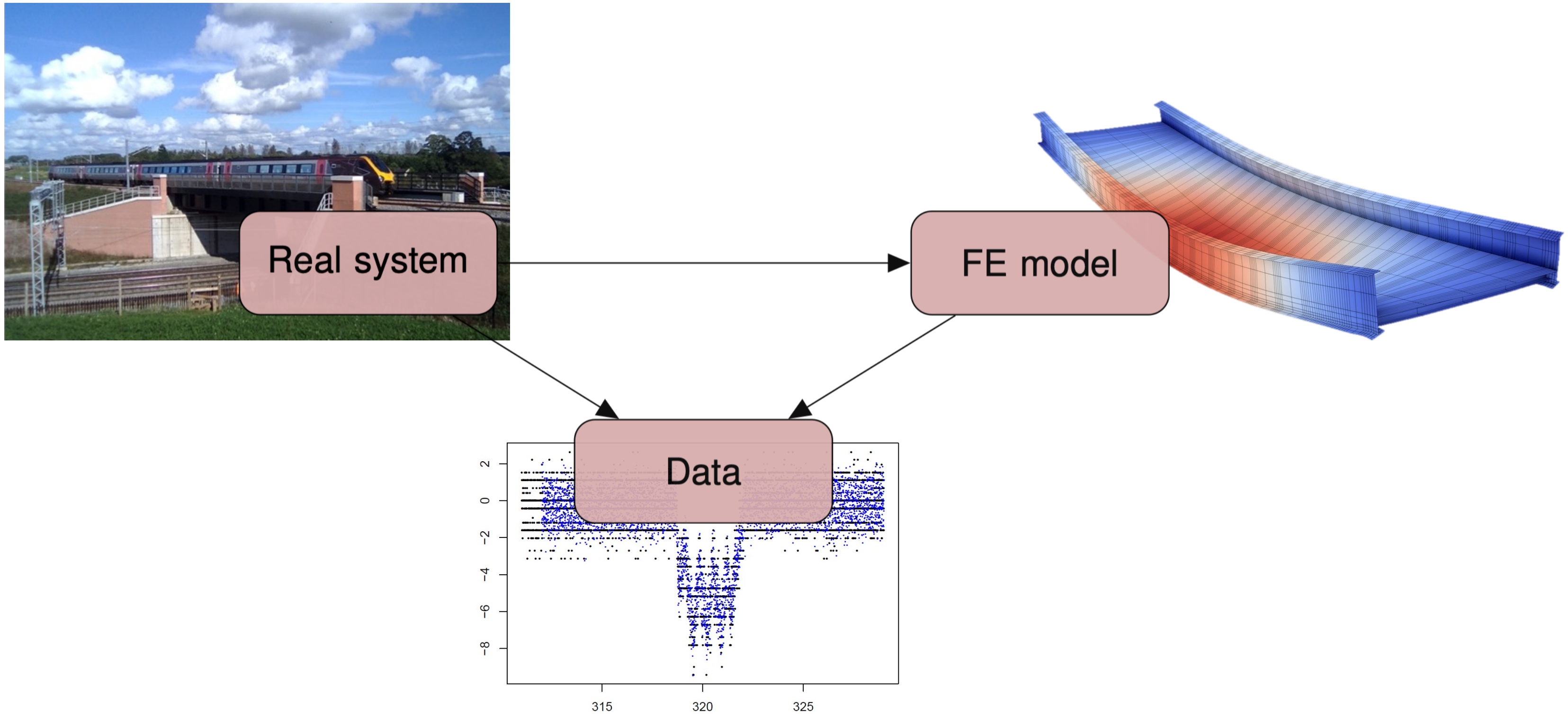


(Giardina et al.)

- Digital twinning of self-sensing systems with FE models can provide more accurate predictions
  - Crucial for condition assessment, more efficient future designs, …

Objective and Challenges

- **Research Objective**
  - To develop a statistical framework for coherent synthesis of data and finite element predictions



- **Challenges**
  - FE models are misspecified due to various modelling assumptions
  - In engineering data is relatively scarce in contrast to machine learning (ML)

Bayesian Modelling (Statistical ML)

- Uncertainties in data, model parameters and the FE model itself are represented as probabilities
  - Results in a set of random variables with corresponding probability density functions

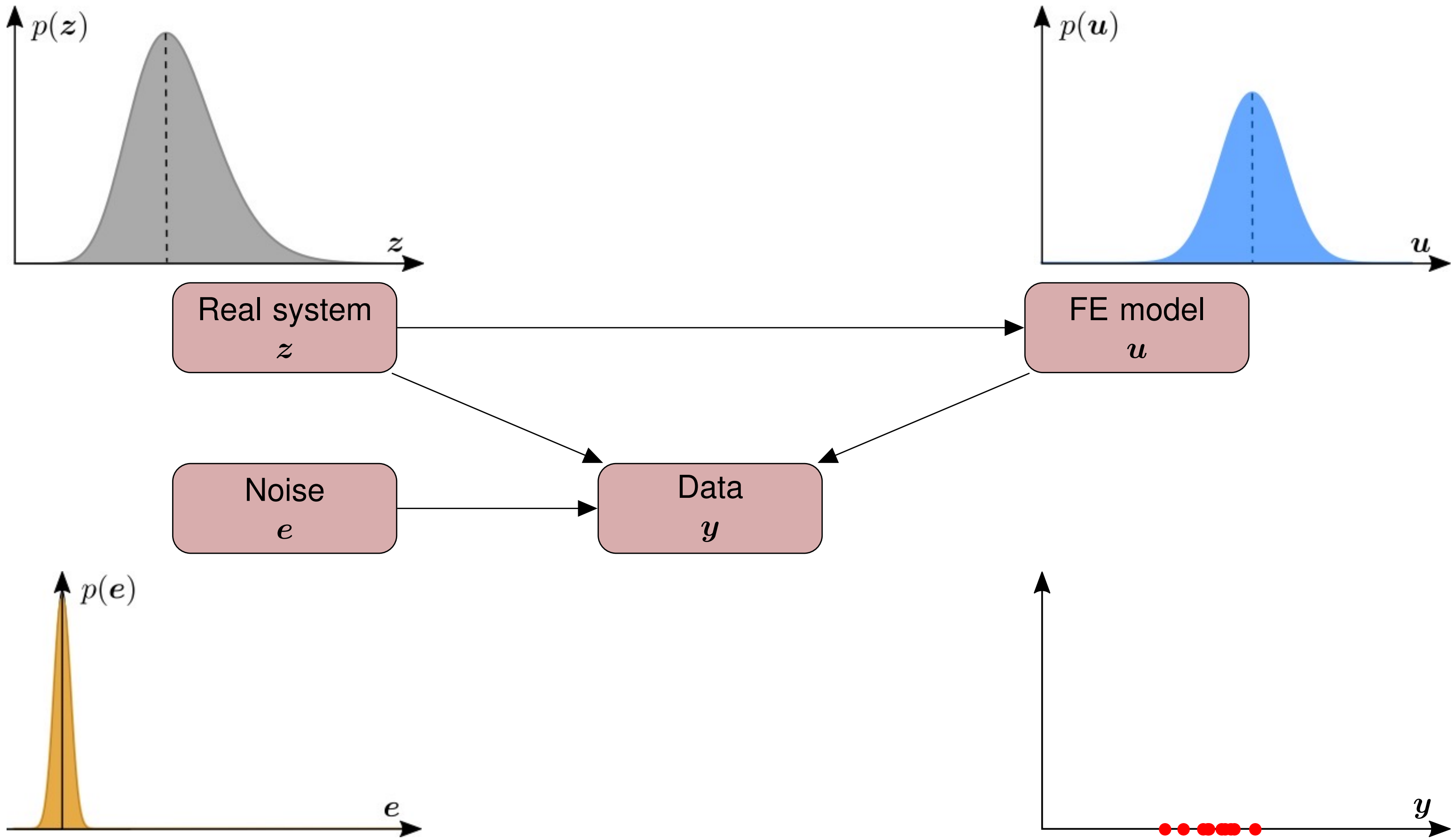


Illustration for a system with only one degree of freedom

Outline

- Statistical data generating model
- Prior probability densities
  - Finite element solution
  - Model misspecification error
- Posterior probability densities
  - Hyperparameter learning
- Examples
  - 1D, 2D and thin-shell verification examples
  - Railway bridge

Data Generating Model

- Data vector $\boldsymbol{y}$, e.g., displacements or strains, is composed of the true response $\boldsymbol{z}$ and measurement noise $\boldsymbol{e}$

$$\boldsymbol{y} = \boldsymbol{z} + \boldsymbol{e}$$

  - In turn, true response $\boldsymbol{z}$ is composed of FE response $\boldsymbol{u}$ and model-reality mismatch $\boldsymbol{d}$

$$\boldsymbol{y} = \boldsymbol{z} + \boldsymbol{e} = \rho \boldsymbol{P} \boldsymbol{u} + \boldsymbol{d} + \boldsymbol{e}$$

    - $\boldsymbol{P}$ is a known observation matrix and ρ a scalar unknown hyperparameter

- We assume that each random vector is <u>normally</u> distributed
  - Finite element prior $\boldsymbol{u} \sim p(\boldsymbol{u}) = \mathcal{N}(\overline{\boldsymbol{u}}, \, \boldsymbol{C}_u)$
  - Misspecification prior $\boldsymbol{d} \sim p(\boldsymbol{d}) = \mathcal{N}(\boldsymbol{0}, \, \boldsymbol{C}_d)$
  - Measurement noise $\boldsymbol{e} \sim p(\boldsymbol{e}) = \mathcal{N}(\boldsymbol{0}, \, \boldsymbol{C}_e)$

- Note that:
  - Each random vector has a mean vector and a covariance matrix
  - Priors consolidate knowledge at hand before taking any measurements

FE Prior: Probabilistic Forward Problem

- Illustrative governing equation (e.g., heat diffusion)

$$
\begin{aligned}
-\boldsymbol{\nabla} \cdot (\mu(\boldsymbol{x}) \boldsymbol{\nabla} u(\boldsymbol{x})) = f(\boldsymbol{x}) & \qquad \text{in } \Omega \\
u(\boldsymbol{x}) = 0 & \qquad \text{on } \partial\Omega
\end{aligned}
$$

  - Random diffusivity $\mu(\boldsymbol{x}) = \exp\big(\kappa(\boldsymbol{x})\big)$ with Gaussian field $\kappa(\boldsymbol{x})$ and random Gaussian forcing $f(\boldsymbol{x})$

- Finite element discretisation

$$
u(x) = \boldsymbol{\phi}(\boldsymbol{x}) \cdot \boldsymbol{u} \qquad \kappa(x) = \boldsymbol{\psi}(\boldsymbol{x}) \cdot \boldsymbol{\kappa} \qquad f(x) = \boldsymbol{\phi}(\boldsymbol{x}) \cdot \boldsymbol{f}
$$

  - Governing equation

$$
\boldsymbol{A}(\boldsymbol{\kappa})\boldsymbol{u} = \boldsymbol{f}
$$

  - Prescribed diffusivity and forcing densities (multivariate normal)

$$
\boldsymbol{\kappa} \sim p(\boldsymbol{\kappa}) = \mathcal{N}\left(\overline{\boldsymbol{\kappa}},\, \boldsymbol{C}_\kappa\right) \qquad \boldsymbol{f} \sim p(\boldsymbol{f}) = \mathcal{N}\left(\overline{\boldsymbol{f}},\, \boldsymbol{C}_f\right)
$$

  - Random solution

$$
\boldsymbol{u}(\boldsymbol{\kappa},\, \boldsymbol{f}) = \boldsymbol{A}(\boldsymbol{\kappa})^{-1} \boldsymbol{f}
$$

    - Non-Gaussian because depends nonlinearly on the diffusivity

FE Prior: Approximate Forward Problem

- Any of the conventional approaches, including Monte Carlo, perturbation, polynomial chaos, etc. can be used to determine density of the FE solution

- We use a first-order perturbation of the random FE solution

$$\boldsymbol{u}(\boldsymbol{\kappa},\, \boldsymbol{f}) = \boldsymbol{u}(\overline{\boldsymbol{\kappa}},\, \boldsymbol{f}) + \sum_{e=1}^{n_e} \frac{\partial \boldsymbol{u}(\overline{\boldsymbol{\kappa}},\, \boldsymbol{f})}{\partial \kappa_e}(\kappa_e - \overline{\kappa}_e) + \dots = \boldsymbol{u}^{(0)} + \sum_{e=1}^{n_e} \boldsymbol{u}_e^{(1)} \lambda_e + \dots$$

  - Mean $\overline{\boldsymbol{u}} = \mathbb{E}\left[\boldsymbol{u}^{(0)} + \sum_e \boldsymbol{u}_e^{(1)} \lambda_e\right] = \boldsymbol{A}(\overline{\boldsymbol{\kappa}})^{-1} \overline{\boldsymbol{f}}$

  - Covariance $\boldsymbol{C}_u = \mathbb{E}\left[\left(\boldsymbol{u}^{(0)} + \sum_e \boldsymbol{u}_e^{(1)} \lambda_e\right) \otimes \left(\boldsymbol{u}^{(0)} + \sum_e \boldsymbol{u}_e^{(1)} \lambda_e\right)\right] - \overline{\boldsymbol{u}} \otimes \overline{\boldsymbol{u}}$

    - Covariance has a (lengthy) closed-form solution

  - Approximate FE density $p(\boldsymbol{u}) = \mathcal{N}(\overline{\boldsymbol{u}},\, \boldsymbol{C}_u)$

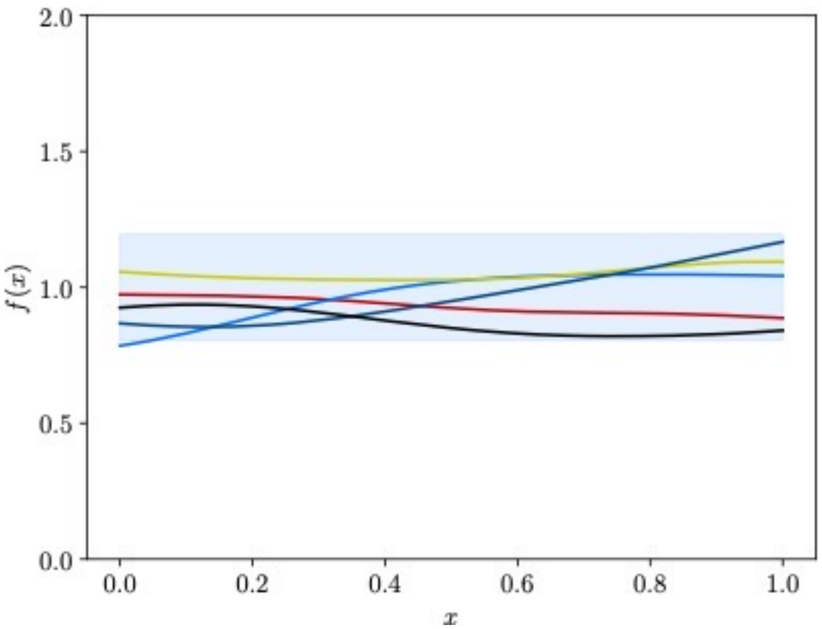
Illustrative Example – Random Forcing

- Governing equation in 1D

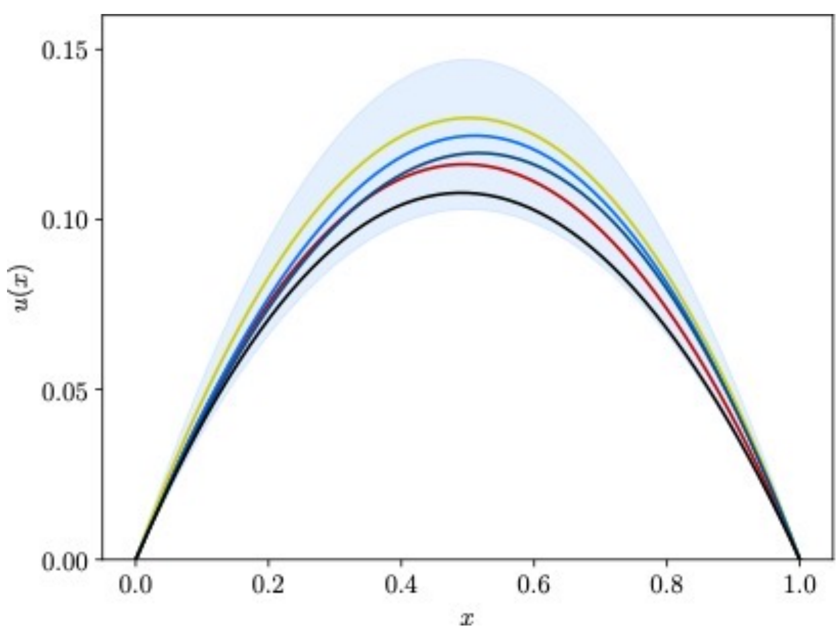
$$\frac{\mathrm{d}^2 u}{\mathrm{d}x^2} = f(x) \qquad \text{with} \quad u(0) = u(1) = 0$$

- Normally distributed random loading with a mean $\bar{f}(x) = 1$, standard deviation $\sigma_f = 0.1$ and correlation length $\ell_f = 0.4$

Random forcing



Random FE solution



Lines indicate samples from the respective densities and shaded areas the 95% confidence regions

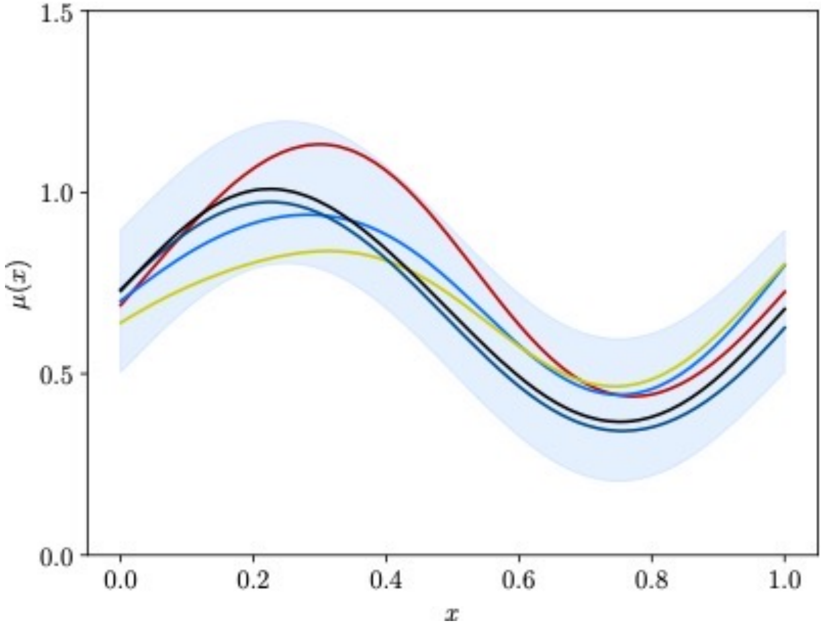
Illustrative Example – Random Diffusion

- Governing equation in 1D

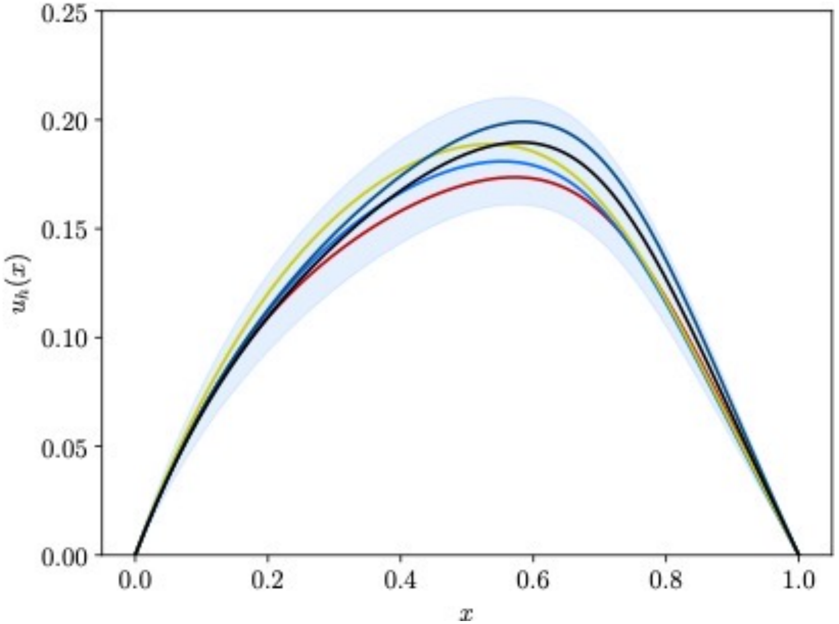
$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\mu(x)\frac{\mathrm{d}u}{\mathrm{d}x}\right) = 1 \qquad \text{with} \quad u(0) = u(1) = 0$$

- Normally distributed diffusivity with standard deviation $\sigma_\kappa = 0.1$, correlation length $\ell_\kappa = 0.25$ and mean $\bar{\kappa}(x) = \ln(0.7 + 0.3\sin(2\pi x))$

Random diffusion coefficient



Random FE solution



Lines indicate samples from the respective densities and shaded areas the 95% confidence regions

Modell Misspecification Prior –1–

- FE models contain many simplifying assumptions resulting in irreducible model misspecification errors

- Unknown misspecification errors are assumed to have the probability density

$$\boldsymbol{d} \sim p(\boldsymbol{d} \mid \sigma,\, \nu,\, \ell) = \mathcal{N}\left(\mathbf{0},\, \boldsymbol{C}_d(\sigma,\, \nu,\, \ell)\right)$$

  - Covariance matrix is obtained by evaluating the chosen Matérn kernel at the observation locations

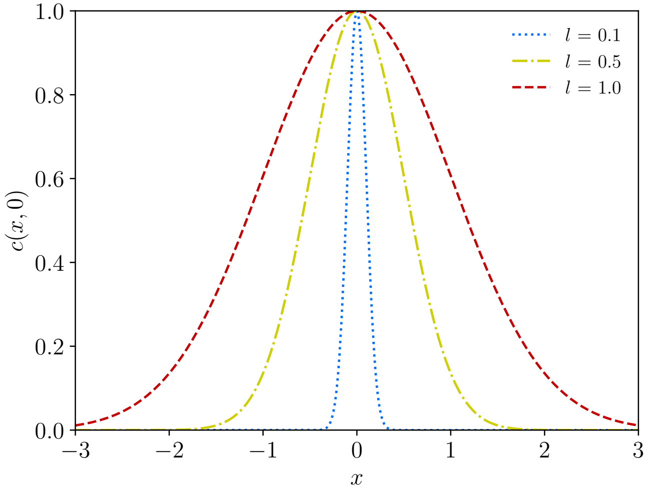
$$c(\boldsymbol{x},\, \boldsymbol{x}') = \frac{\sigma^2}{2^{\nu-1}\Gamma(\nu)} \left(\frac{\sqrt{2\nu}}{\ell} \|\boldsymbol{x} - \boldsymbol{x}'\| \right)^{\nu} K_\nu \left(\frac{\sqrt{2\nu}}{\ell} \|\boldsymbol{x} - \boldsymbol{x}'\| \right)$$

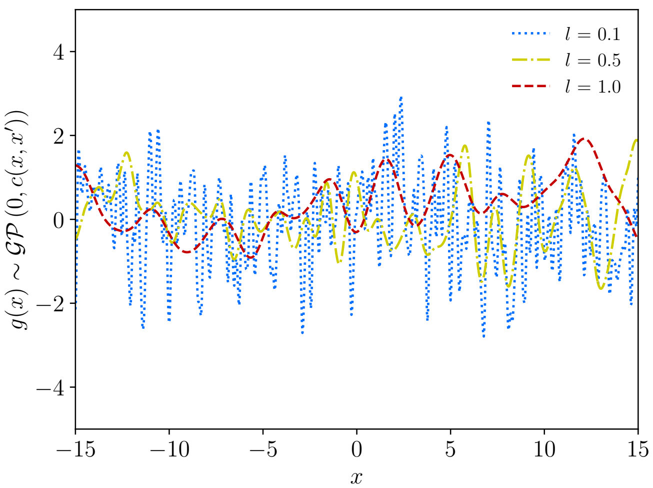
    - Kernel gives the correlation between the values $d(\boldsymbol{x})$ and $d(\boldsymbol{x}')$
    - Standard deviation σ, smoothness ν and ℓ are hyperparameters, which are estimated from the data

  - Other kernels, like square-exponential, polynomial, periodic, etc. or their combinations are possible

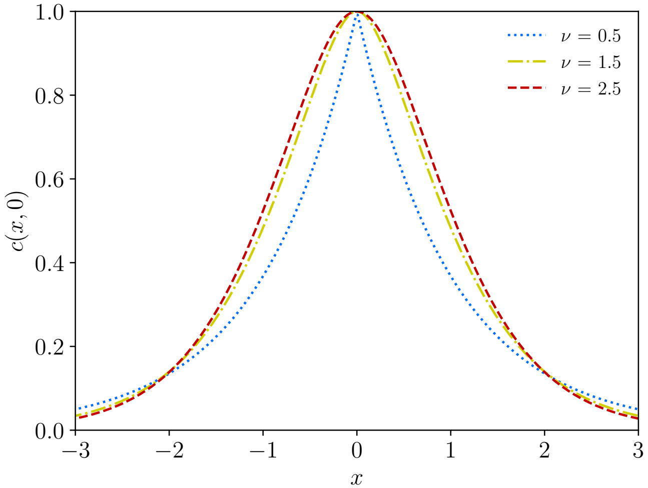
Modell Misspecification Prior –2–

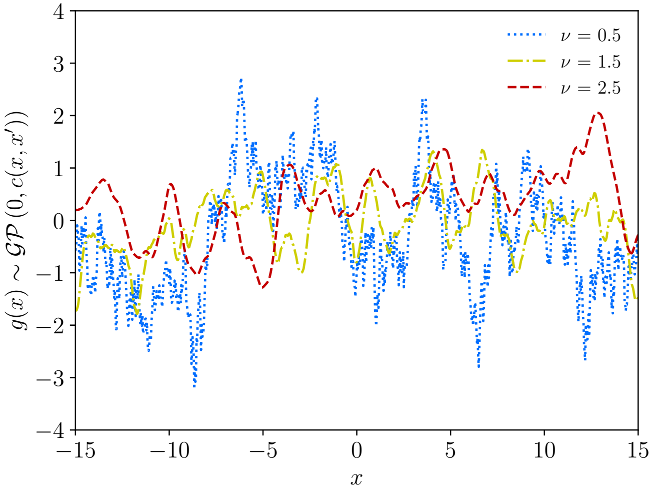
- A kernel defines a parameterised family of random functions
  - Matérn kernel and random functions for three different length-space parameters





  - Matérn kernel and random functions for three different smoothness parameters

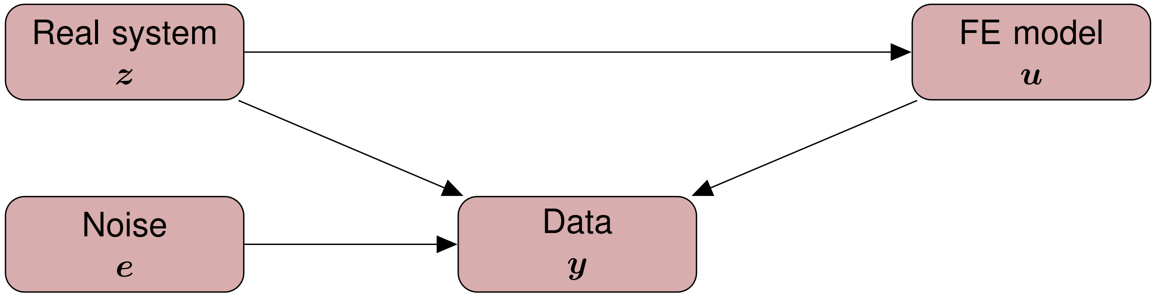




Summary So Far (No Data Yet)

- Data is decomposed into three (independent) random components

$$\boldsymbol{y} = \boldsymbol{z} + \boldsymbol{e} = \rho \boldsymbol{P} \boldsymbol{u} + \boldsymbol{d} + \boldsymbol{e}$$



- Each random variable has a prior density
  - Finite element solution $\boldsymbol{u} \sim p(\boldsymbol{u}) = \mathcal{N}(\overline{\boldsymbol{u}}, \boldsymbol{C}_u)$
  - Model misspecification $\boldsymbol{d} \sim p(\boldsymbol{d} \mid \sigma, \nu, \ell) = \mathcal{N}\left(\boldsymbol{0}, \boldsymbol{C}_d(\sigma, \nu, \ell)\right)$
  - Measurement noise $\boldsymbol{e} \sim p(\boldsymbol{e}) = \mathcal{N}(\boldsymbol{0}, \boldsymbol{C}_e)$
- Hyperparameters of the statistical data generating model
  - Scaling parameter ρ
  - Covariance parameters σ, ν, ℓ

Bayesian Inference –1–

- Bayes rule

Posterior Data likelihood FE prior

$$p(\boldsymbol{u} \mid \boldsymbol{y},\, \boldsymbol{w}) = \frac{p(\boldsymbol{y} \mid \boldsymbol{u},\, \boldsymbol{w})\, p(\boldsymbol{u})}{p(\boldsymbol{y} | \boldsymbol{w})}$$

Marginal likelihood

  - Vector of hyperparameters $\boldsymbol{w} = (\rho \quad \sigma \quad \nu \quad \ell)^{\mathsf{T}}$

- Data likelihood and FE prior are Gaussians so that the posterior and marginal likelihood are Gaussians as well

Bayesian Inference –2 –

Posterior density (Gaussian)

$$p(\boldsymbol{u} \mid \boldsymbol{y},\, \boldsymbol{w}) = \mathcal{N}\left(\overline{\boldsymbol{u}}_{|y}(\boldsymbol{w}) \mid \boldsymbol{C}_{u|y}(\boldsymbol{w})\right)$$

- Posterior mean

$$\overline{\boldsymbol{u}}_{|y}(\boldsymbol{w}) = \overline{\boldsymbol{u}} + \rho \boldsymbol{C}_u \boldsymbol{P}^{\mathsf{T}} \left(\rho^2 \boldsymbol{P} \boldsymbol{C}_u \boldsymbol{P}^{\mathsf{T}} + \boldsymbol{C}_d(\boldsymbol{w}) + \boldsymbol{C}_e\right)^{-1} (\boldsymbol{y} - \rho \boldsymbol{P} \overline{\boldsymbol{u}})$$

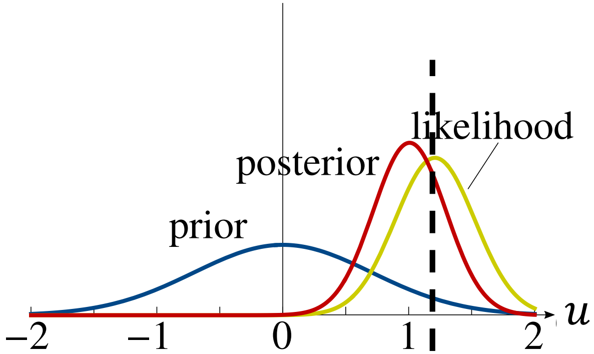
  - Posterior mean is interpolated between data $\boldsymbol{y}$ and prior mean $\overline{\boldsymbol{u}}$

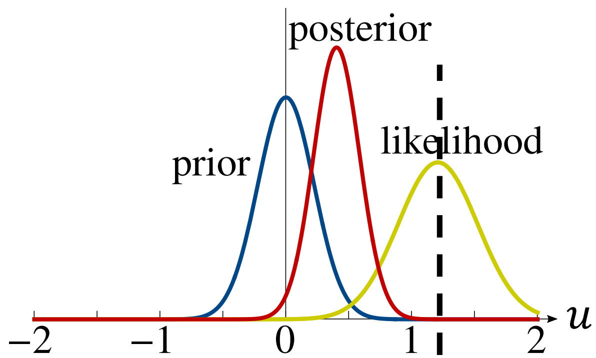
- Posterior covariance

$$\boldsymbol{C}_{u|y}(\boldsymbol{w}) = \boldsymbol{C}_u - \rho^2 \boldsymbol{C}_u \boldsymbol{P}^{\mathsf{T}} \left(\rho^2 \boldsymbol{P} \boldsymbol{C}_u \boldsymbol{P}^{\mathsf{T}} + \boldsymbol{C}_d(\boldsymbol{w}) + \boldsymbol{C}_e\right)^{-1} \boldsymbol{P} \boldsymbol{C}_u$$

  - Measurements lead to a reduction in the prior FE covariance

- One-dimensional illustration





Hyperparameter ($\boldsymbol{w}$) Learning

- Marginal likelihood is the probability of observing the data $\boldsymbol{y}$ over all possible FE solutions $\boldsymbol{u}$

$$p(\boldsymbol{y} \mid \boldsymbol{w}) = \mathcal{N}\left(\rho \boldsymbol{P}\overline{u},\, \rho^2 \boldsymbol{P}\boldsymbol{C}_u \boldsymbol{P}^{\mathsf{T}} + \boldsymbol{C}_d(\boldsymbol{w}) + \boldsymbol{C}_e\right)$$

- Point estimates for the hyperparameter values
  - Approach 1: maximum of the marginal likelihood

$$\boldsymbol{w}^* = \arg\max_{\boldsymbol{w}} p(\boldsymbol{y} \mid \boldsymbol{w})$$

  - Approach 2: empirical mean of the marginal likelihood sampled with Markov-chain Monte Carlo

$$\boldsymbol{w}^* = \frac{1}{N} \sum_{i=1}^{N} p(\boldsymbol{y} \mid \boldsymbol{w}^{(i)})$$

- Posterior density

$$p(\boldsymbol{u} \mid \boldsymbol{y}, \boldsymbol{w}^*) = \mathcal{N}\left(\overline{\boldsymbol{u}}_{|y}(\boldsymbol{w}^*) \mid \boldsymbol{C}_{u|y}(\boldsymbol{w}^*)\right)$$

Examples

One-Dimensional Problem

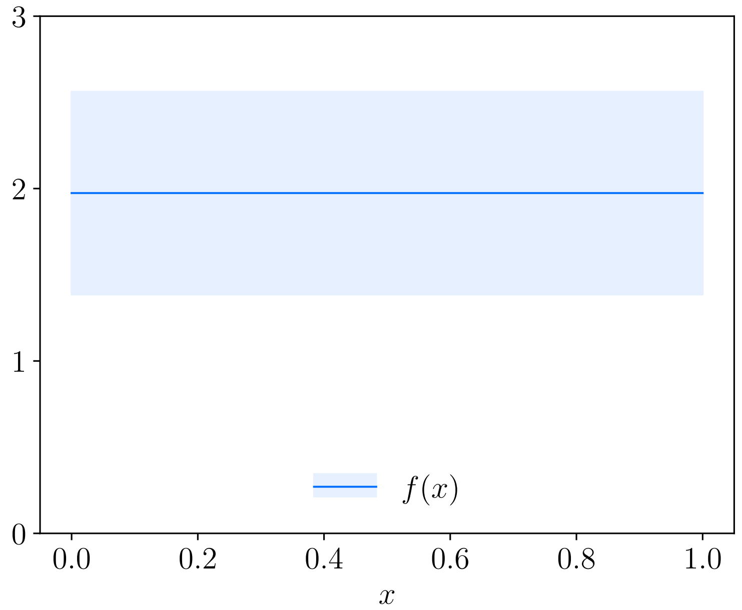
- Poisson-Dirichlet problem with random forcing

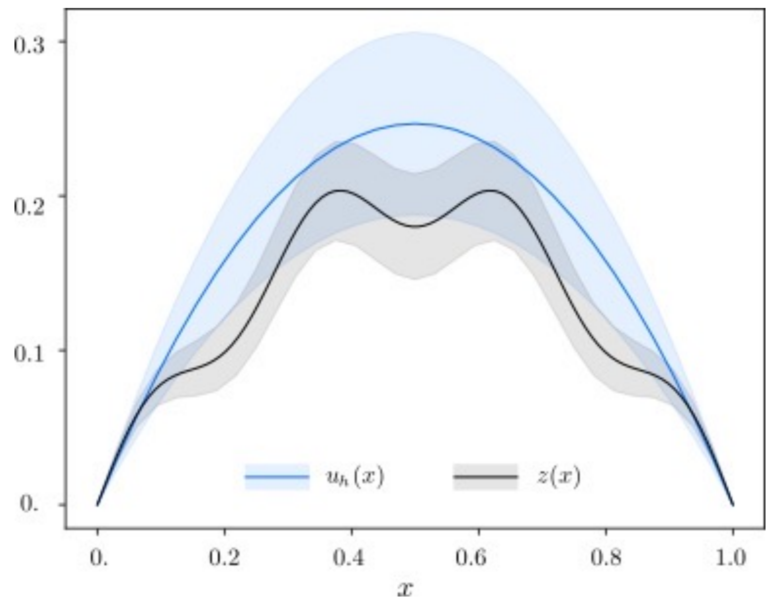
$$\frac{\mathrm{d}^2 u}{\mathrm{d}x^2} = f(x) \qquad \text{in } \Omega = (0,\, 1)$$

$$u(x) = 0 \qquad \text{on } x = 0 \text{ and } x = 1$$

  - Random forcing with density $f(x) \sim \mathcal{GP}\left(1.97,\, 0.3^2 \exp\left(-\frac{\|x - x'\|^2}{2 \cdot 0.25^2}\right)\right)$

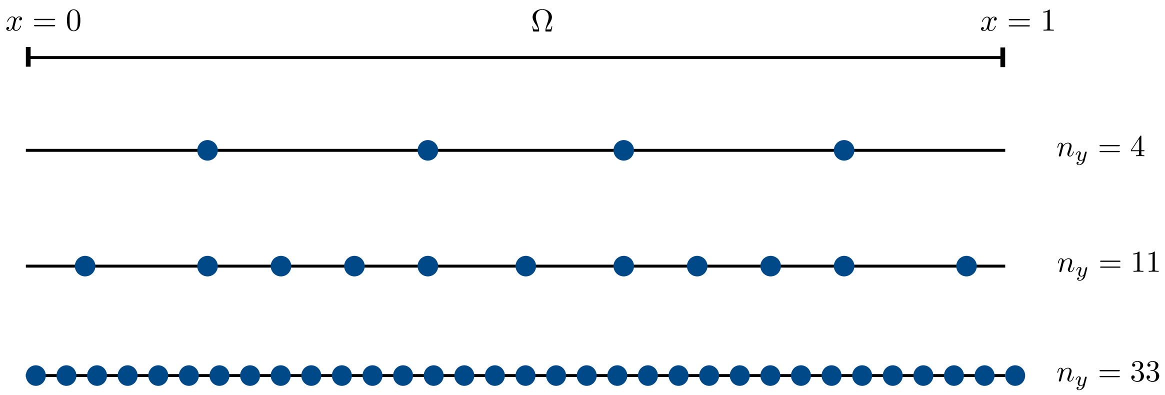
  - Assumed true system density $z(x) \sim \mathcal{GP}(\overline{z}(x),\, c_z(x, x'))$





Synthetic Observation Data

- Variable number observation locations



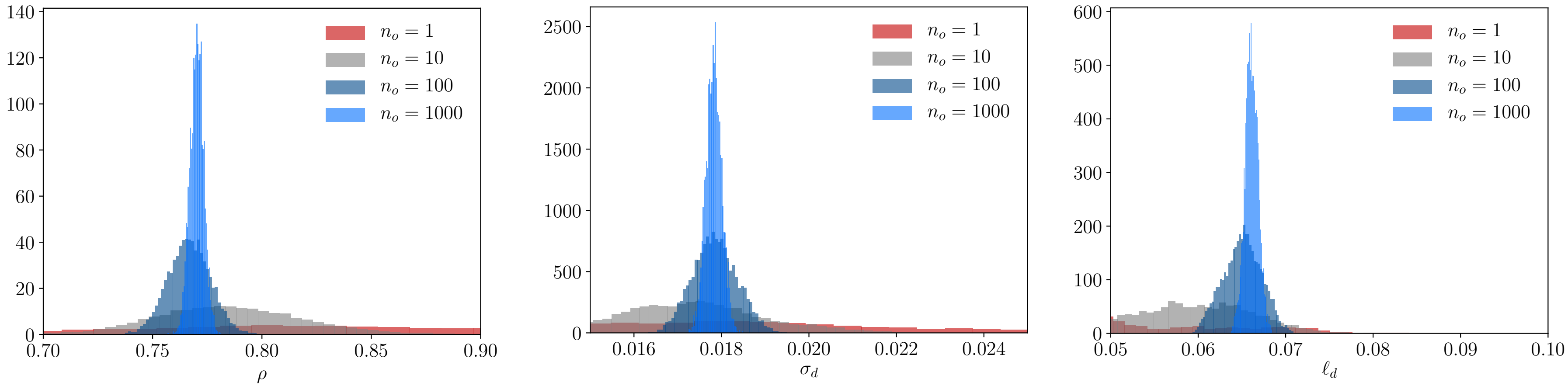
- We generate synthetic observations by sampling from

$$y(x) \sim \mathcal{GP}(\overline{z}(x),\, c_z(x, x') + 2.5 \cdot 10^{-5} \delta_{xx'})$$

  - At the observation locations repeated readings n_o are generated from the assumed true system density
    - $n_o = \{1, 10, 100, 1000\}$

Hyperparameter ($\boldsymbol{w}$) Learning

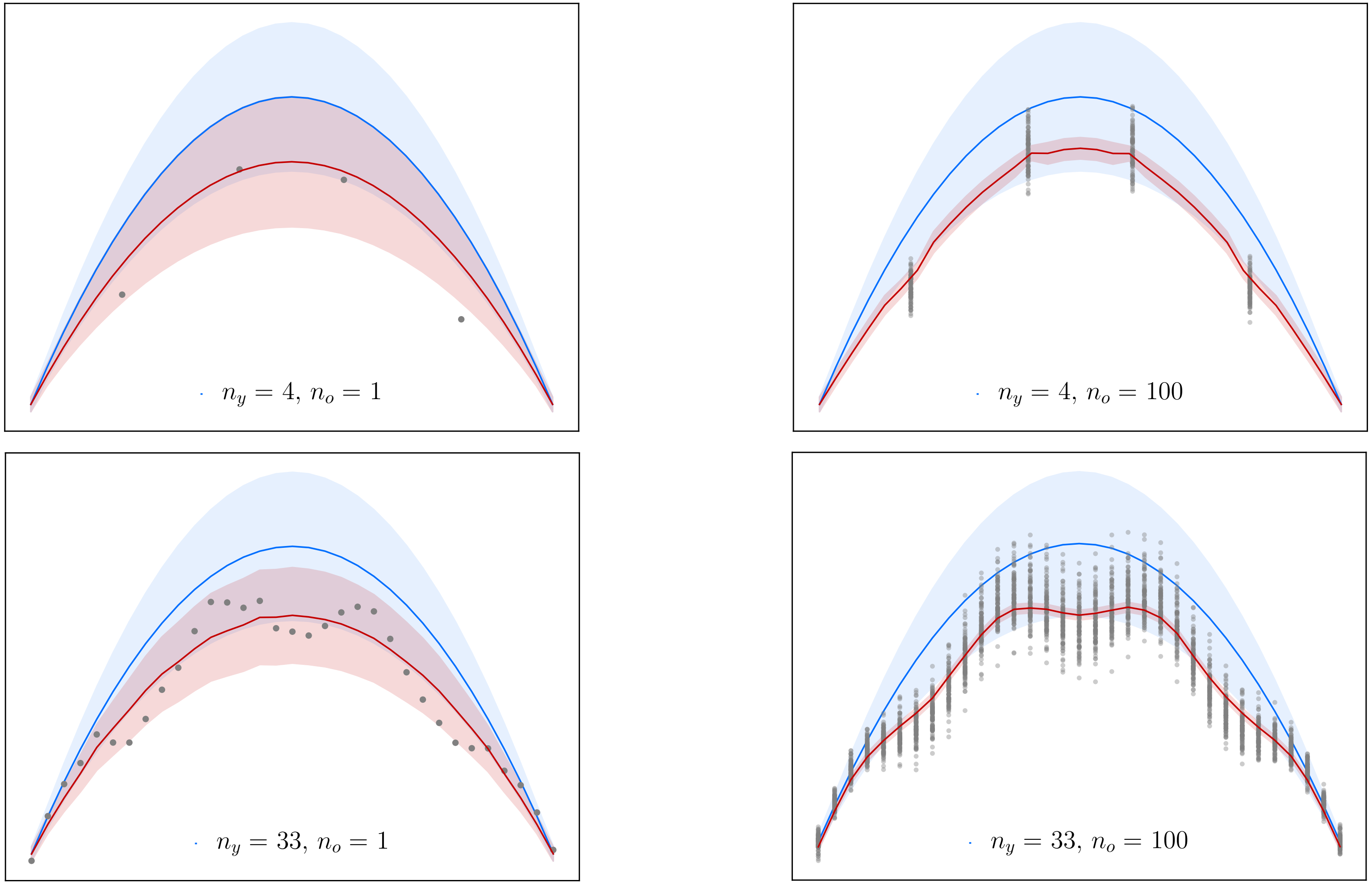
- Marginal likelihood sampled with the Markov-chain Monte Carlo method
  - $n_y = 11$ observation locations
  - $n_o = \{1, 10, 100, 1000\}$ repeated readings at each location



  - Increasing the number of readings n_o leads to convergence of the mean and decrease of variance
  - Empirical mean of the samples is used as a point estimate $(\rho^*, \sigma^*, \ell^*)$

FE Densities

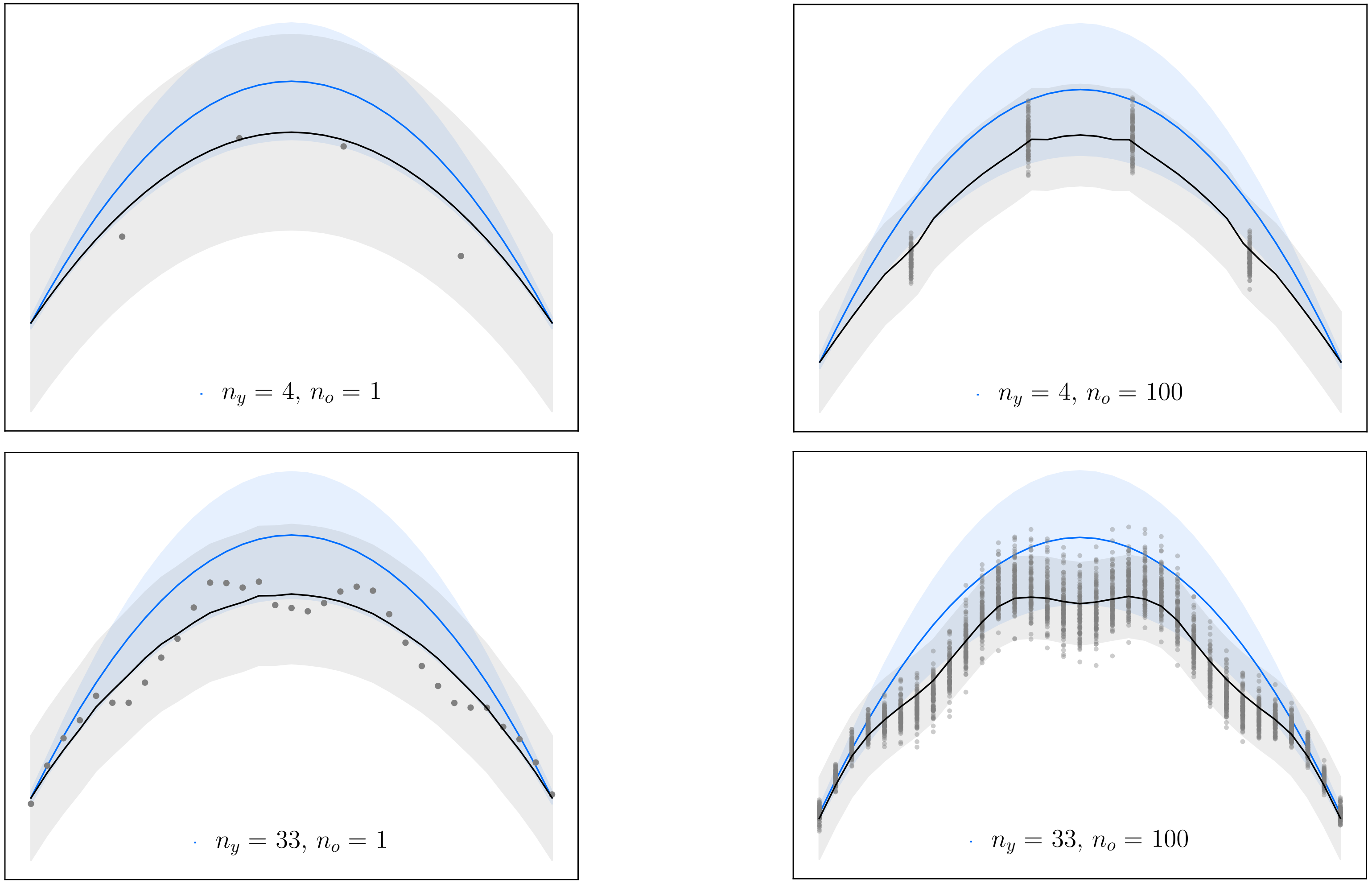
- Blue: prior FE density; red: posterior FE density; dots: observations



- Posterior mean converges to the mean of the data generating process and posterior variance converges to zero

Inferred True Solution

- Blue: prior FE density; black: inferred solution; dots: observations

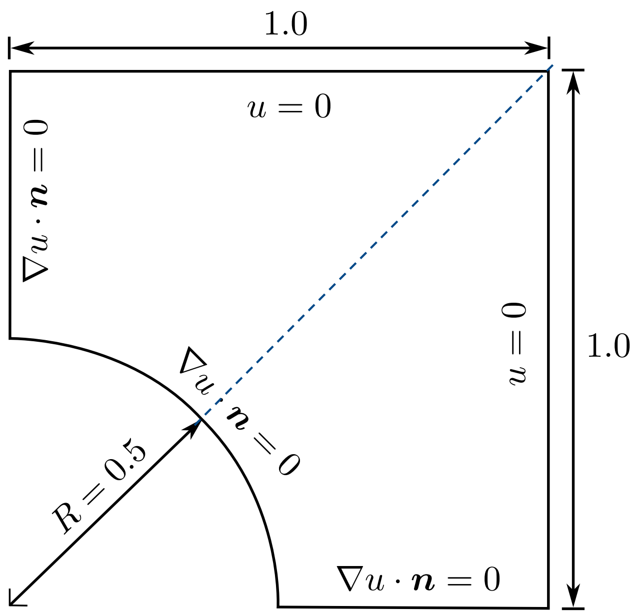


- Inferred true solution converges to the mean and variance of the data generating process

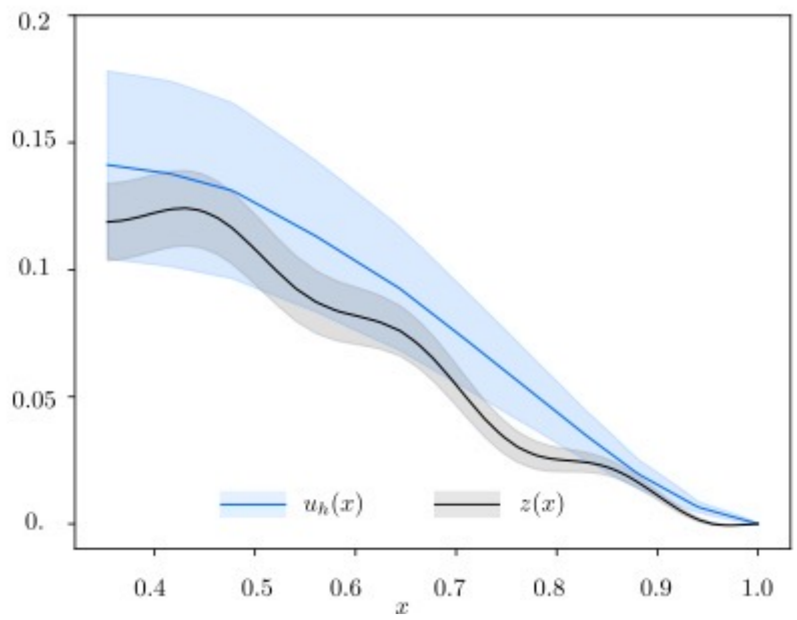
Two-Dimensional Problem

- Poisson problem with random forcing
  - Random forcing with density $f(x) \sim \mathcal{GP}\left(1.0,\ 0.3^2 \exp\left(-\frac{\|\boldsymbol{x}-\boldsymbol{x}'\|^2}{2 \cdot 0.15^2}\right)\right)$
  - Assumed true solution corresponds to FE model with a different loading and finer mesh

Geometry and BCs



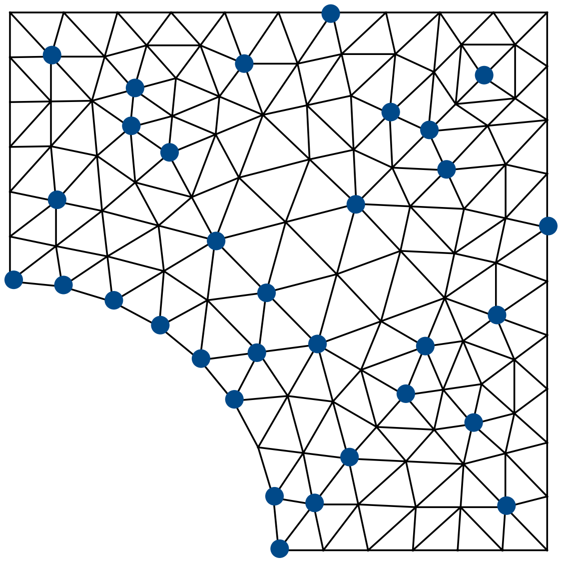
FE and assumed true solution (along diagonal)



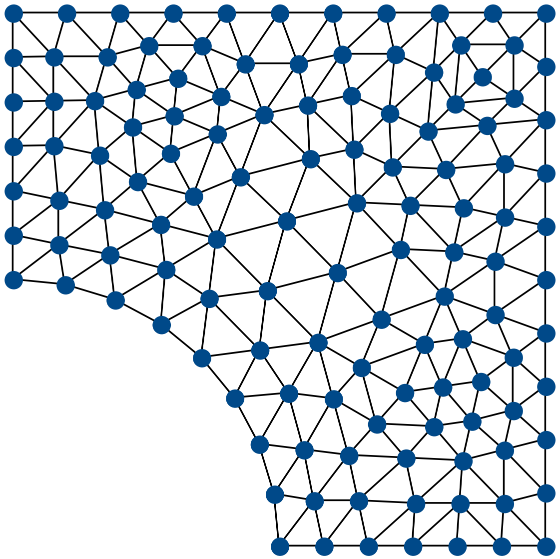
Synthetic Observation Data

- Variable number observation locations
  - Observation points chosen according to a Sobol sequence

$n_y = 32$



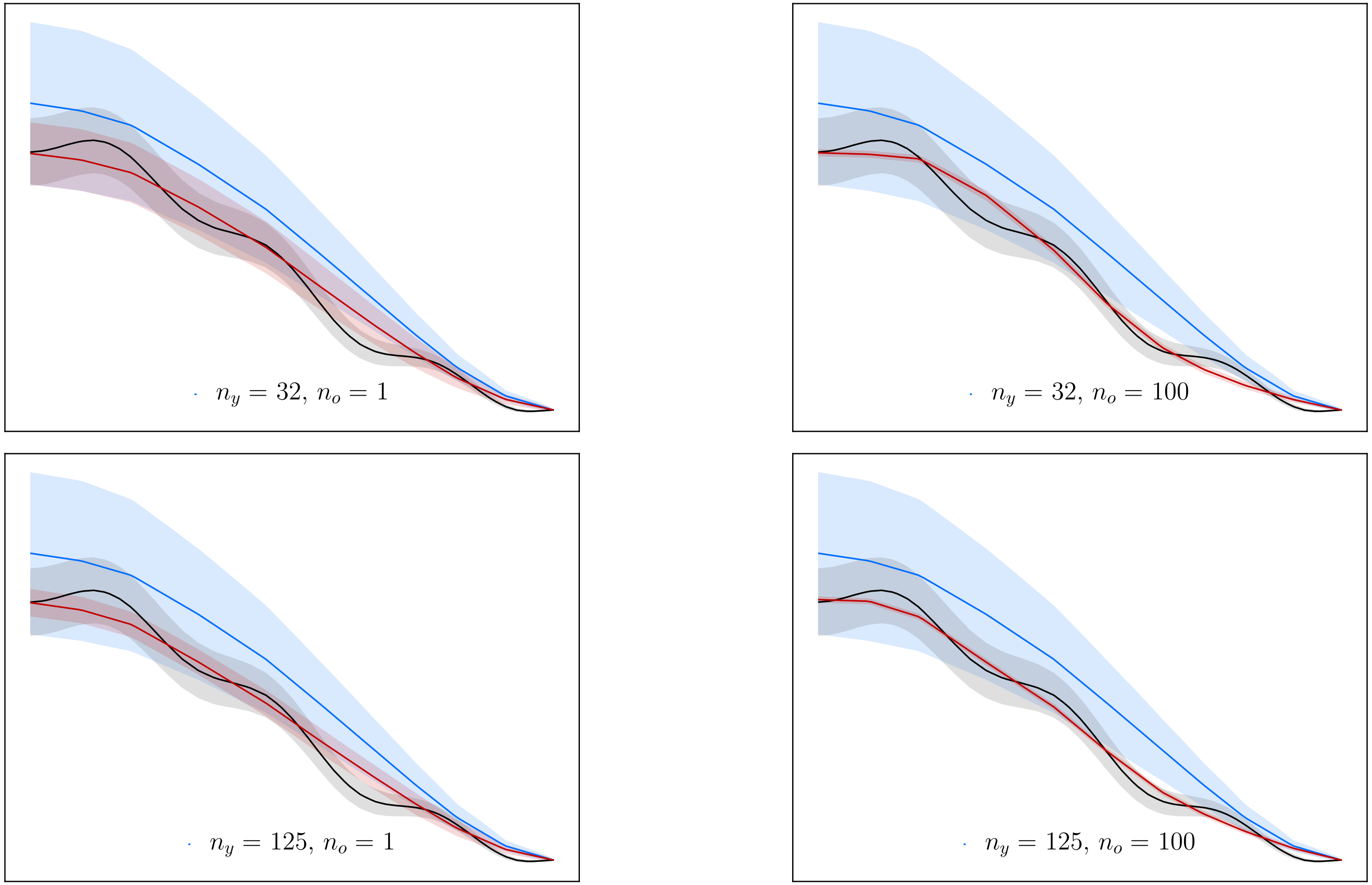
$n_y = 125$



  - At the observation locations repeated readings n_o are generated from the assumed true solution
    - $n_o = \{1, 10, 100, 1000\}$

FE Densities

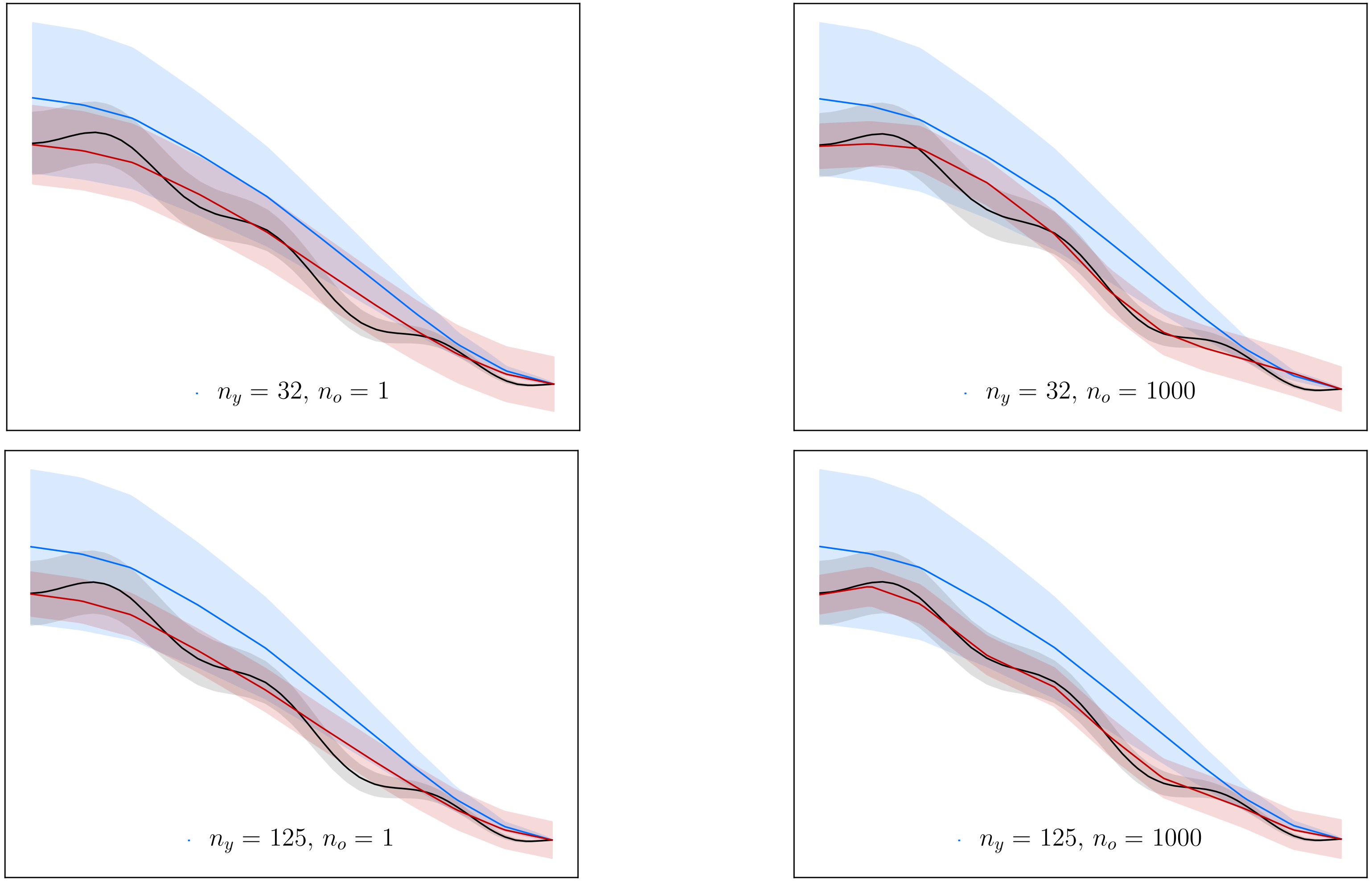
- Blue: prior FE density; red: posterior FE density; black: observations



- Posterior mean converges to the mean of the data generating process and posterior variance converges to zero

Inferred True Solution

- Blue: prior FE density; red: inferred true solution; black: observations



- Inferred true solution converges to the mean data generating process but the variance is underestimated

Summary So Far

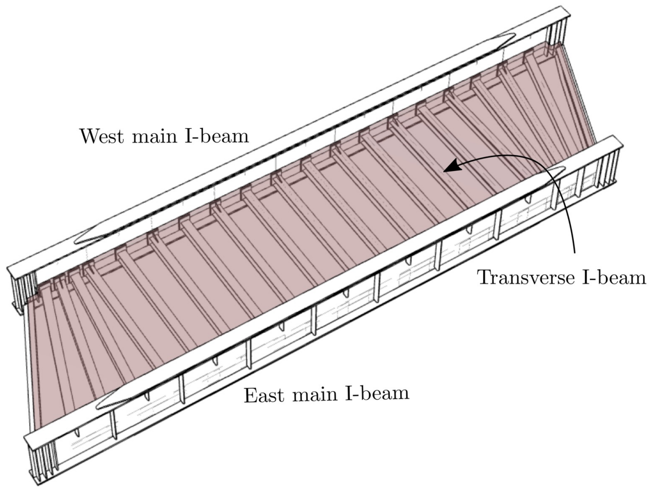
- Data and prior densities are combined to infer true system response
- Posterior mean and covariance depend on inverse of covariance matrices $\boldsymbol{C}_u$ and $\boldsymbol{C}_d$
  - Covariance matrices are dense and costly to invert
  - Implementation introduced so far only suitable for relatively small FE model
- Solution: Consider the stochastic PDE representation of random fields
  - Stiffness matrix of the stochastic PDE is the precision matrix of the random field
  - Probability densities involved can be expressed using sparse precision matrices
  - Posterior mean and variance can be evaluated using only sparse matrix operations

Railway Bridge

- Skewed steel bridge located along the UK West Coast Main Line
  - Instrumented with a fibre optic monitoring system for measuring strains
- Consists of two main longitudinal I-beams, 21 transverse I-beams, and a concrete deck
  - Span is 26.84m and width (between main I-beams) is 7.3m

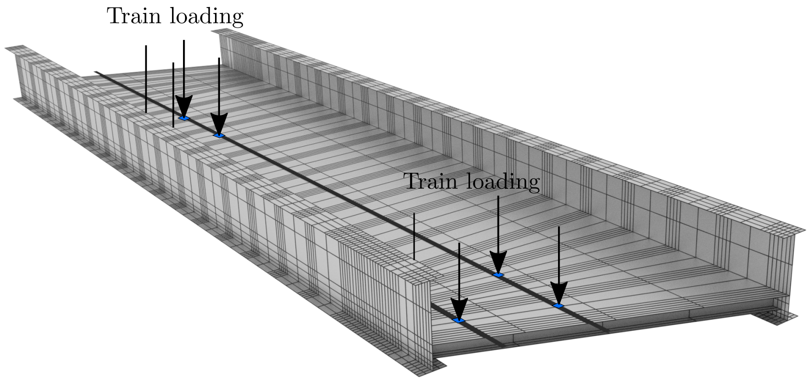
Side view of the bridge

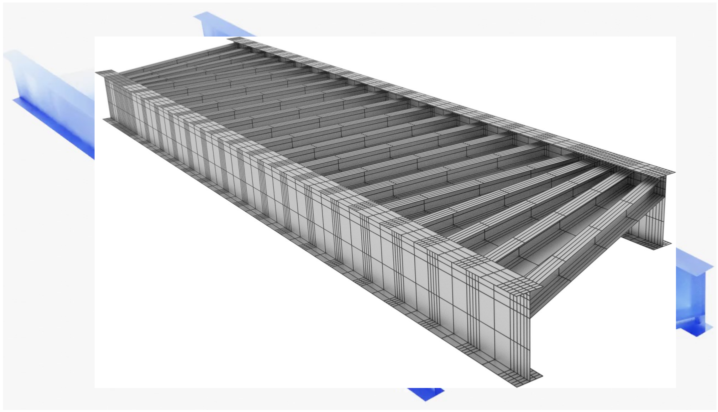
Structural layout



Finite Element Model

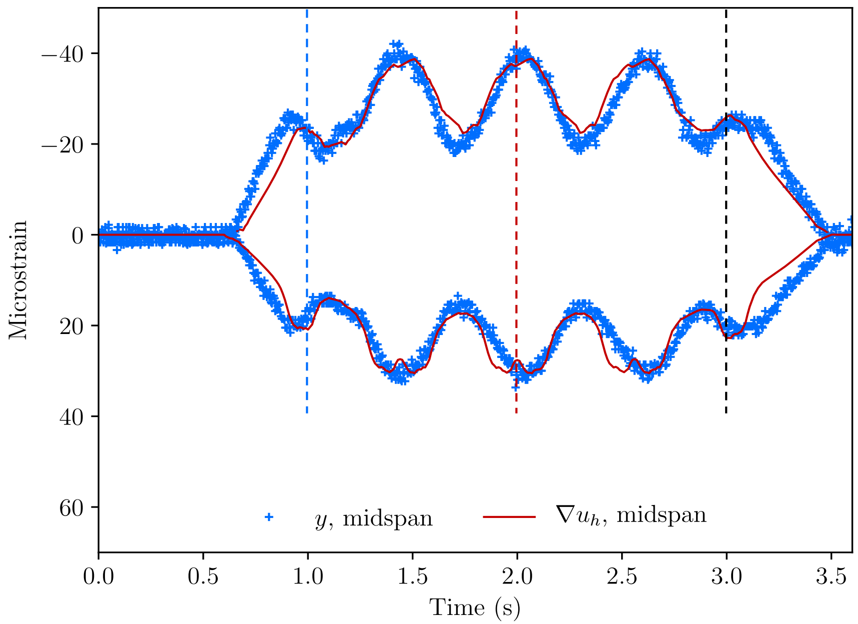
- All structural members modelled as rigidly-connected shells
  - Kirchhoff-Lovel shells discretised using B-splines / subdivision surfaces
  - Quadrilateral mesh with 4600 elements
  - Simply supported boundary conditions
- Moving point loads due to a passing train
  - Inertia effects are neglected due to the short bridge length



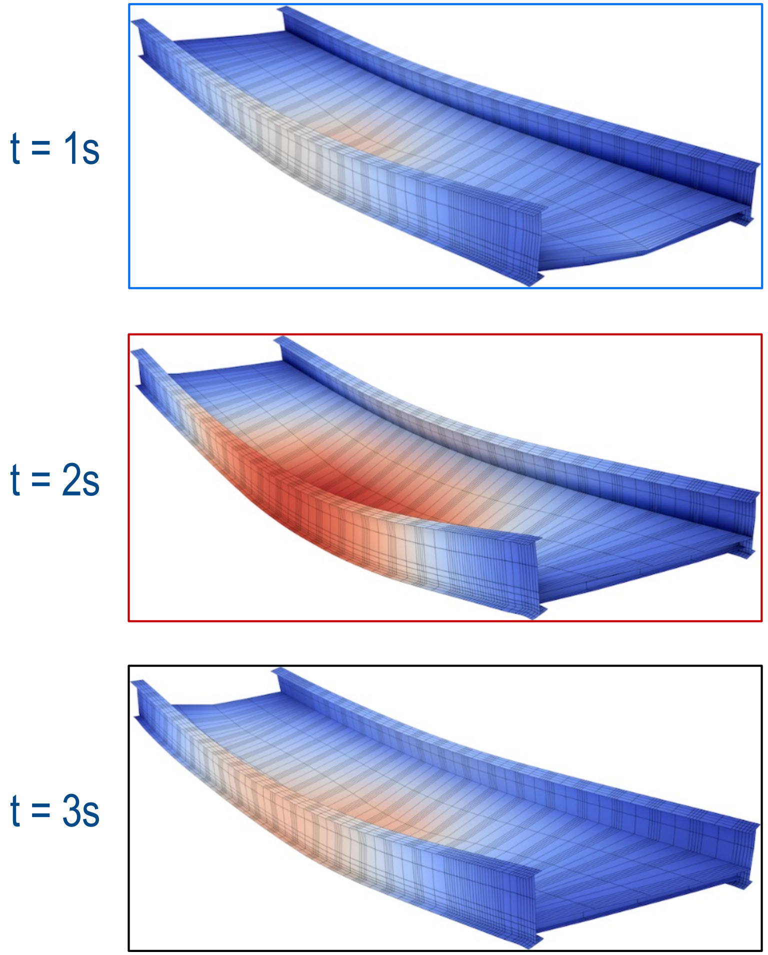


Comparison of FE and Measured Strains

- Top and bottom flange strains <u>at the mid-span</u> of the east main I-beam



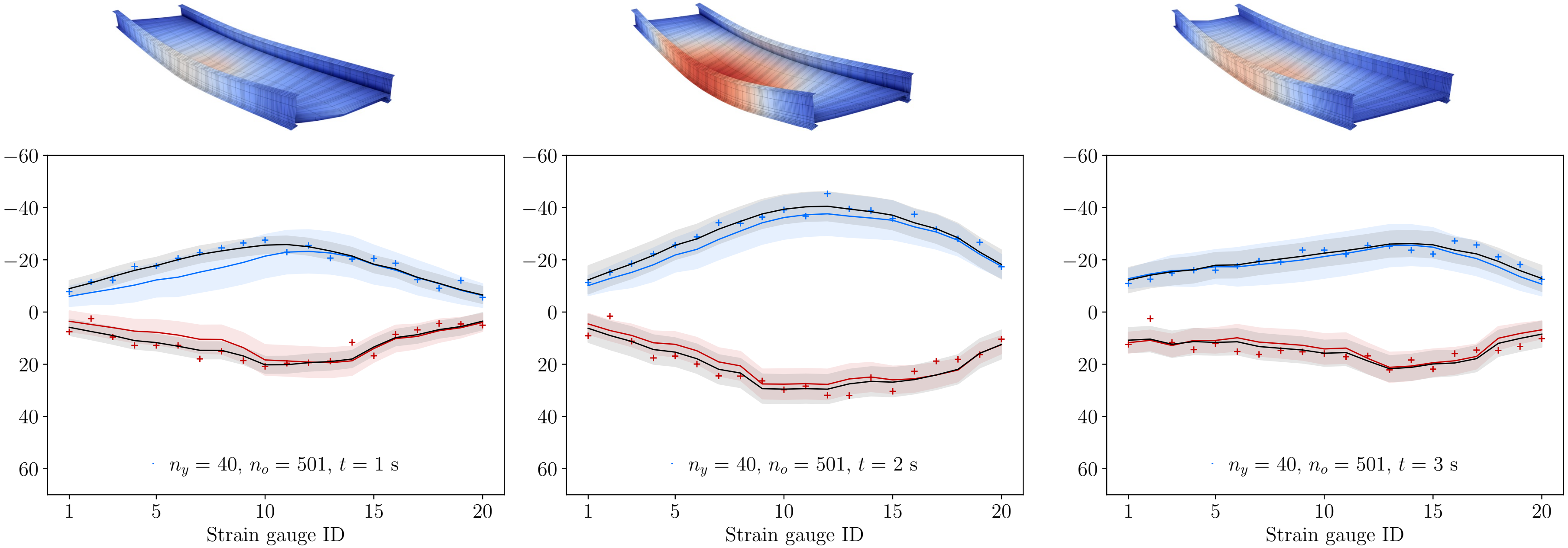
Deflected bridge



- Measured strains: +
- Red: FE strain

Inferred True Strains

- Flange strains along the east main I-beam at t = 1s, 2s, 3s



t = 1s   t = 2s   t = 3s

  - Blue and red: prior FE strains and 95% confidence regions
  - Black: inferred true flange strains and 95% confidence regions

- Inferred true strains lie between the measured strains away from the prior FE strains

Summary and Conclusions

- Introduced a statistical finite element construction for synthesis of data and FE models
  - Uncertainties in model specification, parameters and data are considered
  - Scalable to large problems using the SPDE representation of random fields
  - Limited amount of data is sufficient to estimate the true response
- Several extensions, including model selection and comparison, are available
- Finally, proposed approach can be interpreted as a physics-informed Gaussian Process regression technique with FE priors
- References:

  - Koh, Cirak, *Stochastic PDE representation of random fields for large-scale GP regression and statistical FEA,* CMAME, 2023
  - Febrianto, Butler, Girolami, Cirak, *A self-sensing digital twin of a railway bridge using the statistical finite element method,* DCE, *2022*
  - Girolami, Febrianto, Yin, Cirak, *The statistical finite element method (statFEM) for coherent synthesis of observation data and model predictions,* CMAME, 2021

Today's Schedule

- 09:00–09:50 Bayesian Inverse Problems
  - 10:00–11:00 Hands-on Session
- 11:15–12:10 Gaussian Process Surrogates
  - 13:30–14:30 Hands-on Session
- 14:45–15:35 Statistical Finite Elements
  - 15:45–16:45 Hands-on Session
- 16:45–17:00 Summary & Discussion

Asking questions by typing in Q&A or raising hand both are fine