

Uncertainty quantification (UQ) in geotechnical engineering

Ningxin Yang

Civil and Environmental Engineering, Imperial College London

n.yang23@imperial.ac.uk

May 20, 2024

Supervisors: Dr Truong Le; Prof. Lidija Zdravković

Presentation Overview

- ① UQ and uncertainty types
- ② UQ framework
- ③ How to consider uncertainties into UQ framework
- ④ Final goal of UQ
- ⑤ Current work and future considerations

- ① UQ and uncertainty types
- ② UQ framework
- ③ How to consider uncertainties into UQ framework
- ④ Final goal of UQ
- ⑤ Current work and future considerations

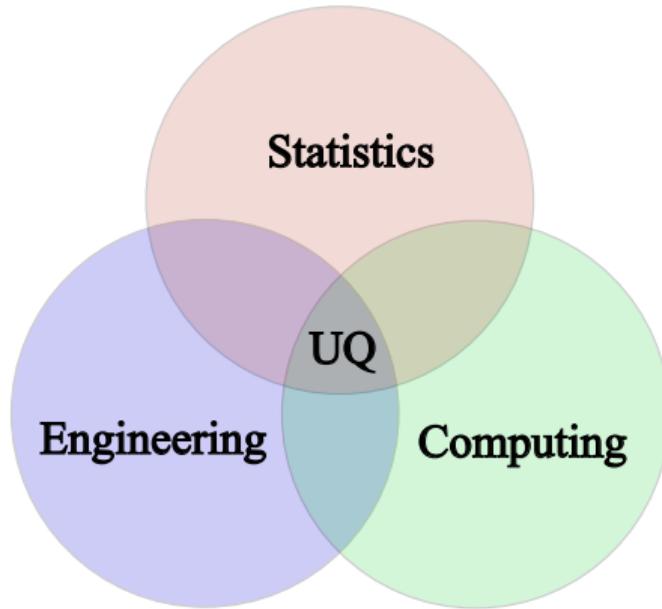
What is Uncertainty Quantification (UQ)?

All models are wrong, but some are useful.
-George E.P. Box

How inaccurate might the models be?
When are they useful in engineering problems?
How much confidence can we have in model's predictions?

UQ provides a framework answering
these questions and making model useful.

What is Uncertainty Quantification (UQ)?



*UQ is the science of quantitative **characterization** and **reduction** of uncertainties in both computational and real world applications¹*

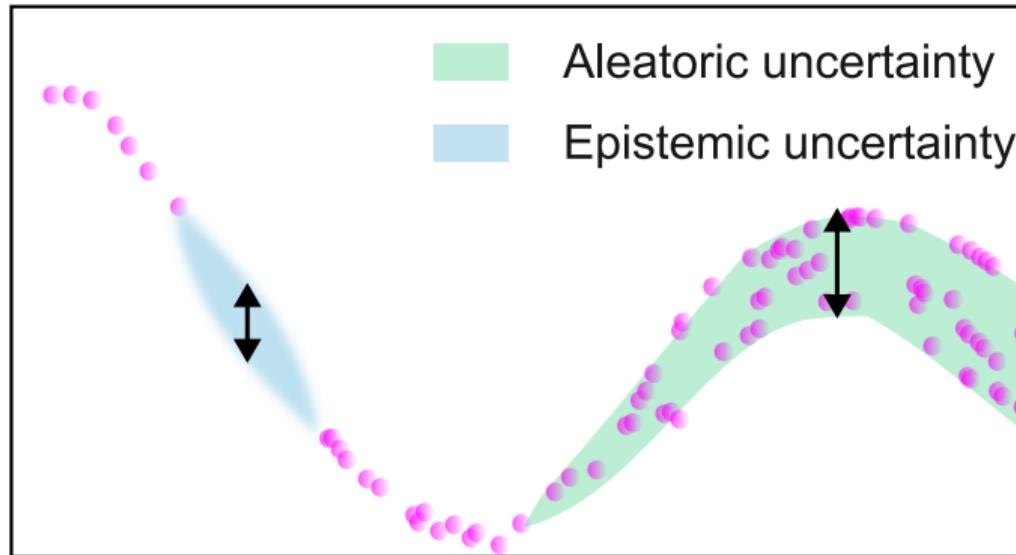
¹Victor E. Saouma and M. Amin Hariri-Ardebili. "Uncertainty Quantification". In: *Aging, Shaking, and Cracking of Infrastructures: From Mechanics to Concrete Dams and Nuclear Structures*. Cham: Springer International Publishing, 2021, pp. 423–454.

Unforecast exposures in geotechnics



Types of two uncertainties

Aleatoric vs epistemic



Aleatoric uncertainty vs Epistemic uncertainty

Distinguish:

Aleatoric uncertainty

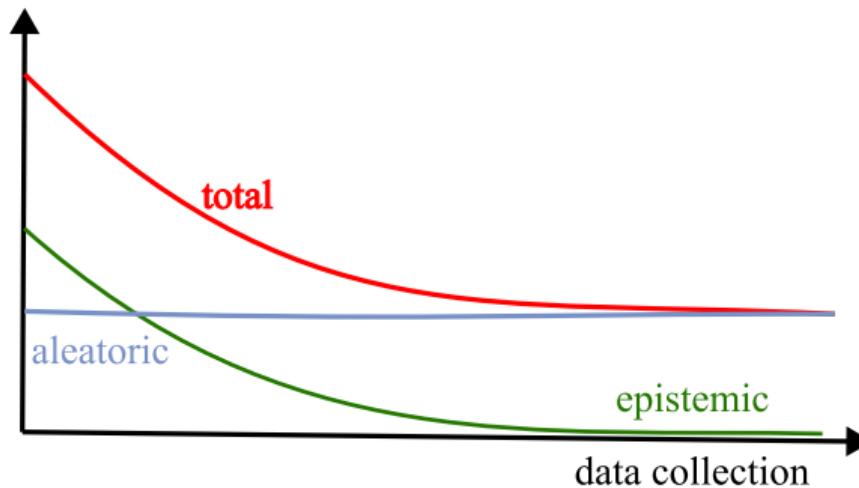
statistical variability,
inherently random
effects (**irreducible**)

Epistemic uncertainty

model uncertainty, a lack
of knowledge (**reducible**)

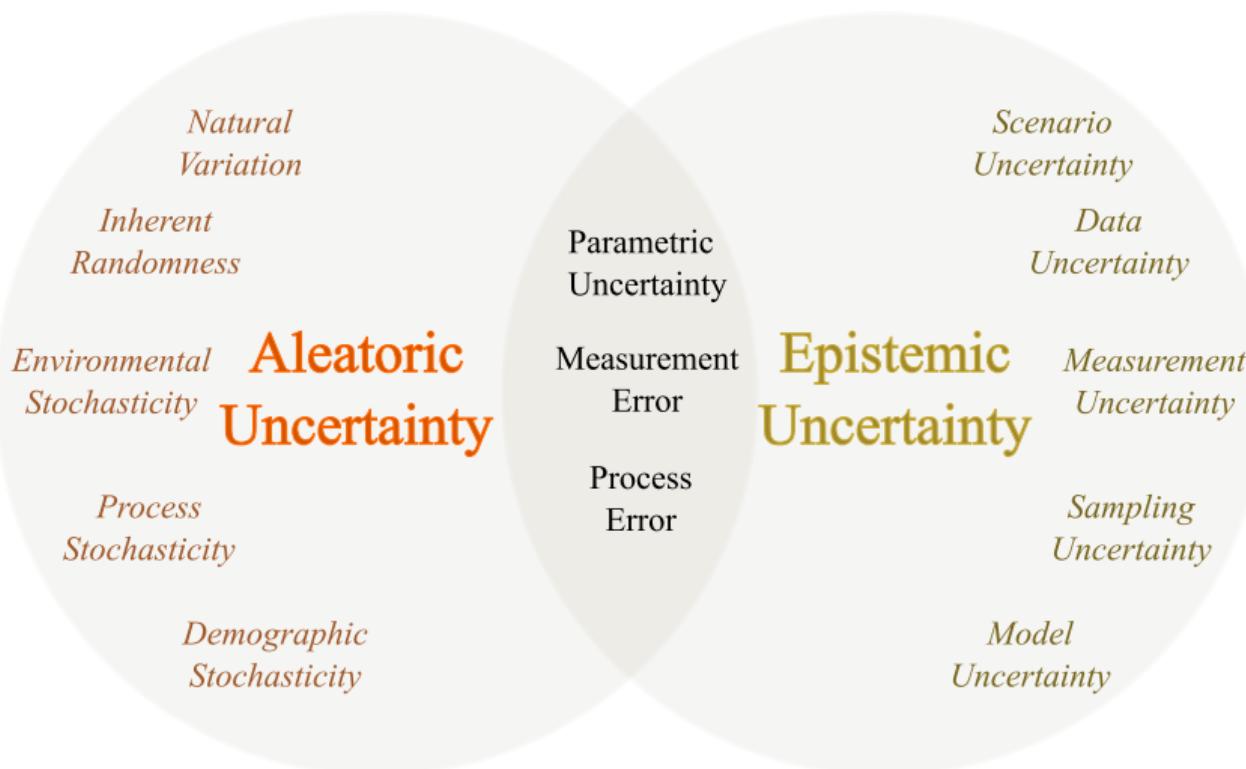
Uncertainty components:

Total uncertainty \approx aleatoric uncertainty + epistemic uncertainty

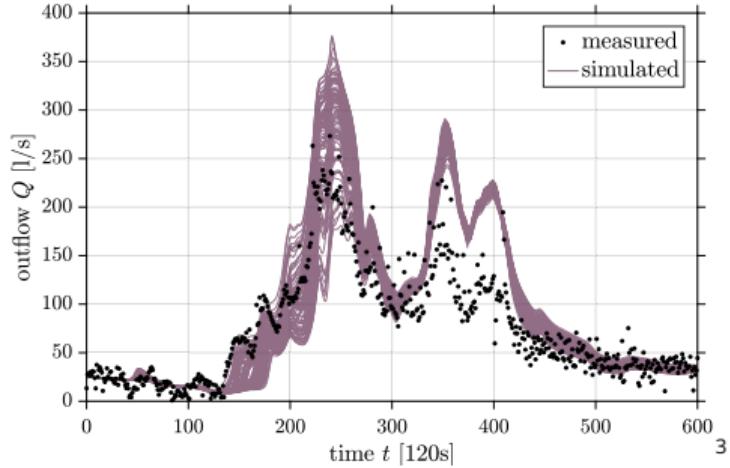
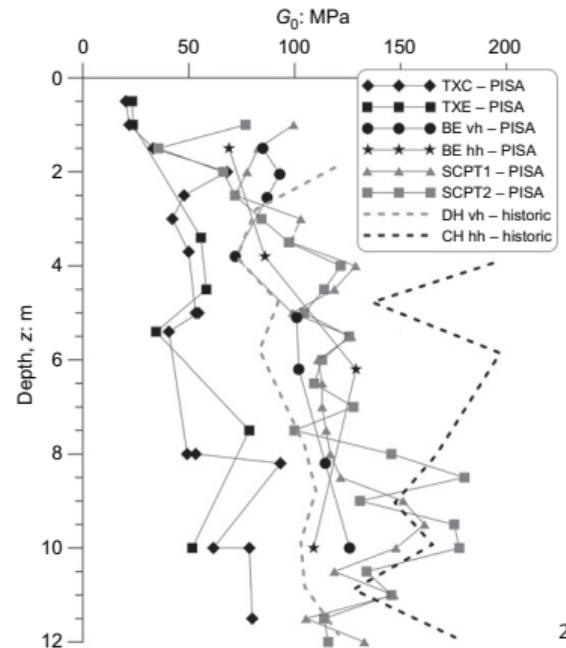


Not simple as it is

Mix of aleatoric and epistemic: Involve too much subjectivity



Examples in geotechnics

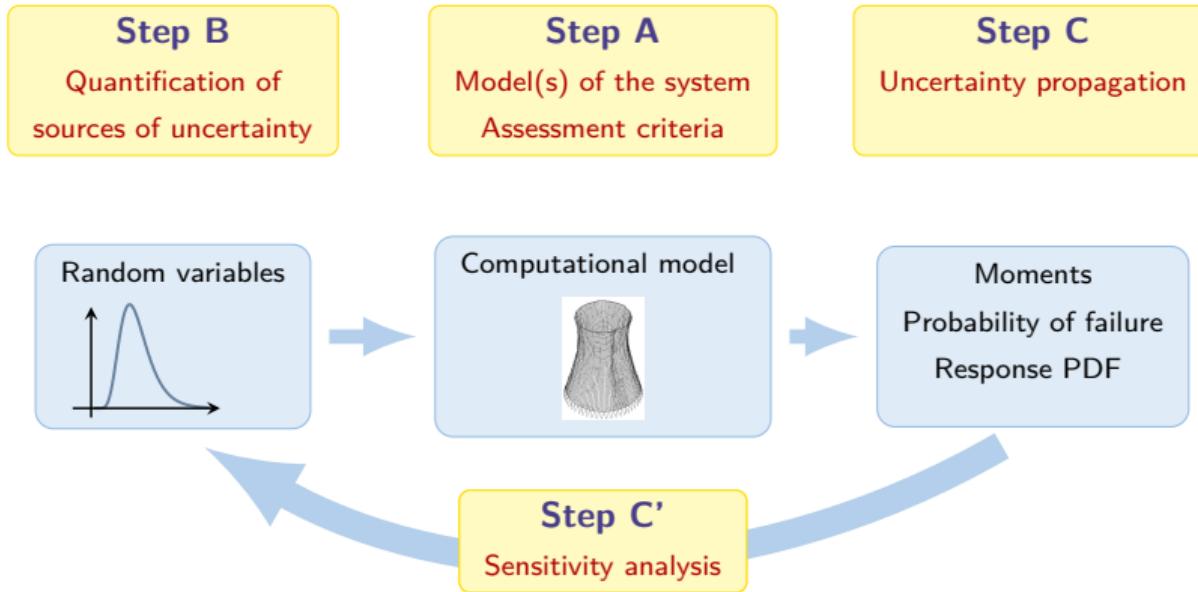


²Lidija Zdravković et al. "Ground characterisation for PISA pile testing and analysis". In: *Géotechnique* 70.11 (2020), pp. 945–960.

³Joseph B Nagel, Jörg Rieckermann, and Bruno Sudret. "Principal component analysis and sparse polynomial chaos expansions for global sensitivity analysis and model calibration: Application to urban drainage simulation". In: *Reliability Engineering & System Safety* 195 (2020) p. 106737.

- ① UQ and uncertainty types
- ② UQ framework
- ③ How to consider uncertainties into UQ framework
- ④ Final goal of UQ
- ⑤ Current work and future considerations

UQ framework and components

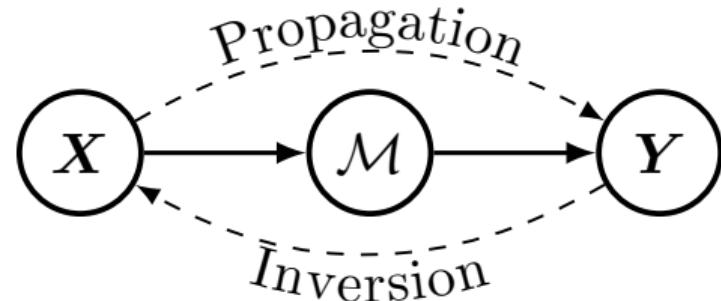


4

⁴Bruno Sudret. "Uncertainty quantification in the simulation of complex systems". In: *1st International Conference on Infrastructure Resilience*. ETH Zurich. 2018.

UQ framework and components

Two UQ types of problems:



Four components:

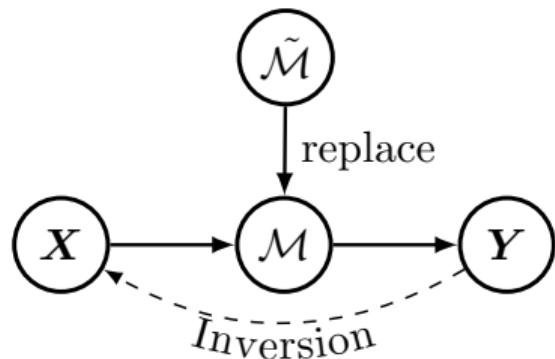
- ① Model assessment
- ② Uncertainty propagation
- ③ Model calibration
- ④ Sensitivity analysis

Definition

A computational model \mathcal{M} should contain:

- a **mathematical description** of the physics
 - may be seen as a **black box** to compute the QoI
-
- Analytical formula
 - Empirical formula
 - Packaged programs based on FEM/FDM ...
 - ...

UQ component one-Model assessment



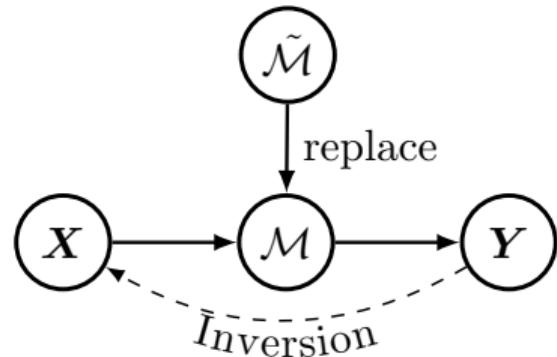
Usage

$$\mathcal{M}(\mathbf{x}) \approx \tilde{\mathcal{M}}(\mathbf{x})$$

Hours to run seconds to run

- It is built from a **limited** set of runs of the original model \mathcal{M} called the **experimental design** $\mathcal{X} = \{\mathbf{x}^{(i)}\}, i = 1, \dots, N\}$
- It assume some regularity of the model \mathcal{M} and some general functional shape

UQ component one-Model assessment



Some popular choices for \tilde{M} :

Name	Shape	Parameters
Polynomial chaos expansions	$\tilde{M}(\mathbf{x}) = \sum_{\alpha \in \mathcal{A}} \mathbf{y}_\alpha \Psi_\alpha(\mathbf{x})$	\mathbf{y}_α
Low-rank tensor approximations	$\tilde{M}(\mathbf{x}) = \sum_{l=1}^R b_l \left(\prod_{i=1}^M v_l^i x_i \right)$	$b_l, z_{k,l}^i$
Kriging (a.k.a Gaussian processss)	$\tilde{M}(\mathbf{x}) = \boldsymbol{\beta}^T \cdot \mathbf{f}(\mathbf{x}) + Z(\mathbf{x}, \omega)$	$\boldsymbol{\beta}, \sigma_Z^2, \theta$
Support vector machines	$\tilde{M}(\mathbf{x}) = \sum_{i=1}^m a_i K(\mathbf{x}_i, \mathbf{x}) + b$	a, b
Neural networks	$\tilde{M}(\mathbf{x}) = f_n(\cdots f_2(b_2 + f_1(b_1 + \mathbf{w}_1 \cdot \mathbf{x}) \cdot \mathbf{w}_2))$	\mathbf{w}, \mathbf{b}

UQ component two-Model calibration

Choice for UQ inversion

Choice for the UQ method is totally based on the **quantity** of accessible data:

- **Lack or no** data available, model can be solely based on expert judgement
- **Substantial** volume data available, model can fully use statistical inference (e.g., the methods of moments)
- **Combination** of two above: Bayesian methods

$$\pi(\mathbf{x}|\mathcal{Y}) = \frac{\mathcal{L}(\mathbf{x}|\mathcal{Y}) \cdot \pi(\mathbf{x})}{\pi(\mathcal{Y})}$$

Bayesian Data Analysis

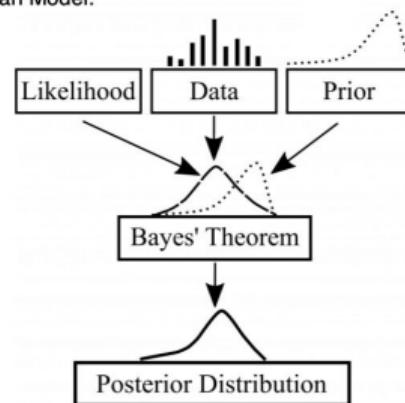
Formula:

$$f(\text{model} | \text{data}) = \frac{f(\text{data} | \text{model}) \times f(\text{model})}{f(\text{data})}$$

Bayesian Model:

```
graph TD; subgraph BM [Bayesian Model]; L[Likelihood]; D[Data]; P[Prior]; end; L --> BT[Bayes' Theorem]; D --> BT; P --> BT; BT --> PD[Posterior Distribution];
```

Bayesian Model:



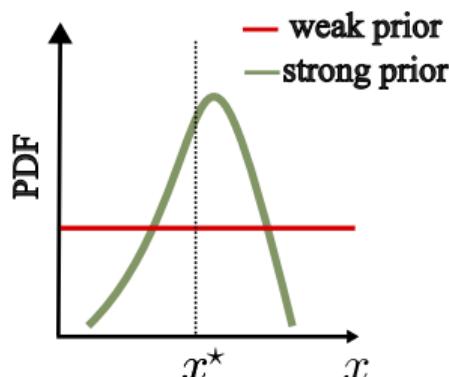
From Matt Dancho

UQ component two-Model calibration

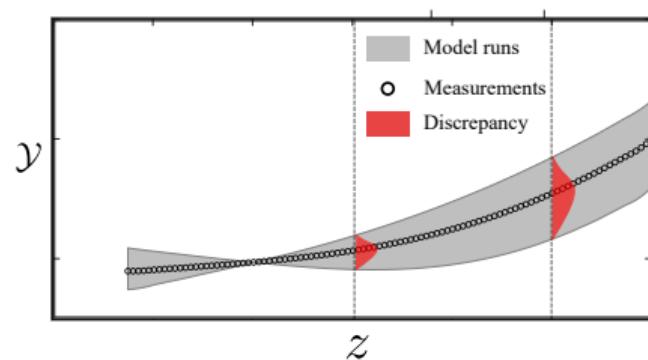
Bayesian methods

Expert guess + Limited data → Distribution

Prior:



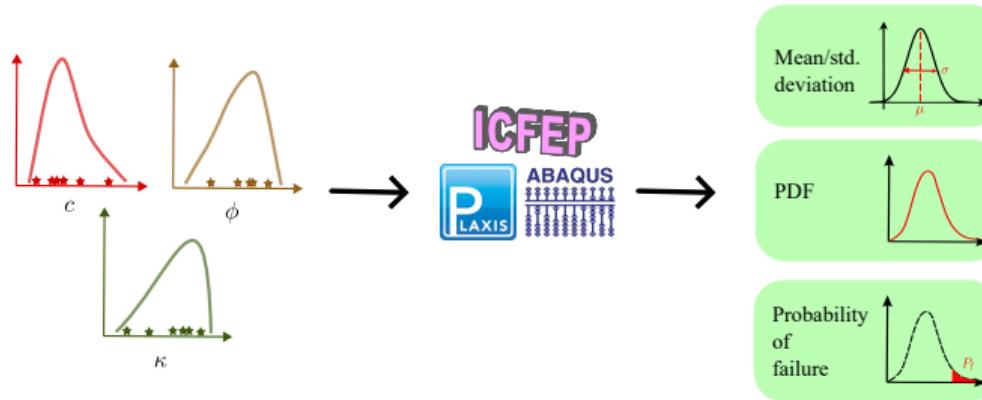
Likelihood:



Notable caveats:

- Prior-Requires specific expertise
- Likelihood-Computationally expensive

UQ component three-Uncertainty propagation



- **Output statistics**, i.e., mean, standard deviation. etc.

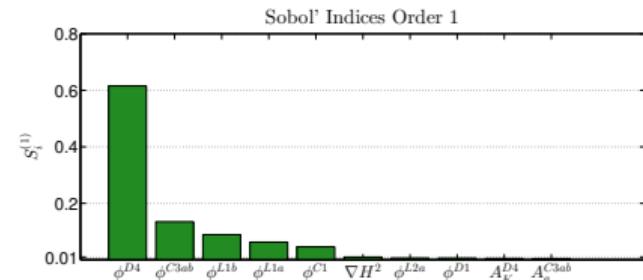
$$\mu_{\mathbf{Y}} = \mathbb{E}_{\mathbf{X}}[\mathcal{M}(\mathbf{X})]; \sigma_{\mathbf{Y}}^2 = \mathbb{E}_{\mathbf{X}}[(\mathcal{M}(\mathbf{X}) - \mu_{\mathbf{Y}})^2]$$

- **Distribution** of the QoI
- **Probability** of exceeding an admissible threshold y_{adm} following
$$P_f = \mathbb{P}(\mathbf{Y} \geq y_{adm})$$

UQ component four-Sensitivity analysis

Sensitivity analysis-Determine what are the input parameters whose uncertainty explains the variability of the QoI

- detect input parameters whose uncertainty has **no impact** on the output variability
- detect input parameters which allow one to best **decrease the output variability** when set to a deterministic value
- detect **interactions** between parameters



Variance decomposition (Sobol' indices)^a

^aBruno Sudret. "Uncertainty quantification in the simulation of complex systems". In: *1st International Conference on Infrastructure Resilience*. ETH Zurich. 2018.

UQ component four-Sensitivity analysis

Total variance:

$$D \equiv \text{Var}[\mathcal{M}(\mathbf{X})] = \text{Var}\left[\sum_{u \subset \{1, \dots, M\}} \mathcal{M}_u(\mathbf{X}_u)\right] = \sum_{u \subset \{1, \dots, M\}} \text{Var}[\mathcal{M}_u(\mathbf{X}_u)]$$

- Sobol's indice:

$$S_u \stackrel{\text{def}}{=} \frac{\text{Var}[\mathcal{M}_u(\mathbf{X}_u)]}{D}$$

- First-order Sobol's indice:

$$S_i = \frac{D_i}{D} = \frac{\text{Var}[\mathcal{M}_i(\mathbf{X}_i)]}{D}$$

Quantify the effect of each input parameter **separately**

- Total Sobol's indice:

$$S_i^T \stackrel{\text{def}}{=} \sum_{u \supset i} S_u$$

Quantify the **total effect** of x_i , including **interactions** with other variables

Uncertainty quantification for engineering problems

Research topics

- Uncertainty modelling for engineering systems
- Bayesian model calibration
- Structural reliability analysis
- Surrogate models (low dimensions/high dimensions)
- Stochastic inverse problem
- Global sensitivity analysis
- Reliability-based design optimization
- ...

- ① UQ and uncertainty types
- ② UQ framework
- ③ How to consider uncertainties into UQ framework
- ④ Final goal of UQ
- ⑤ Current work and future considerations

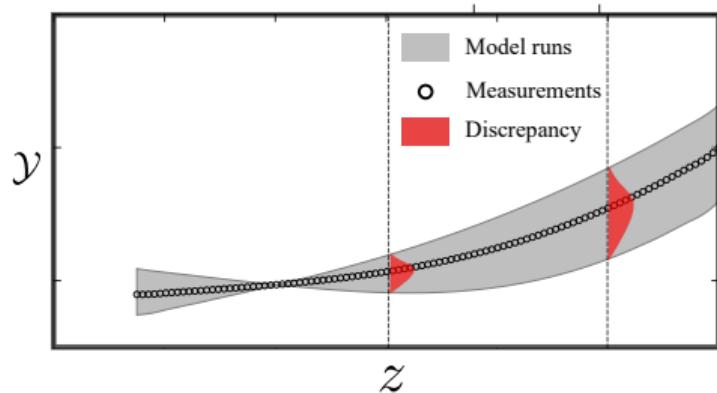
Connecting physics with FE modelling

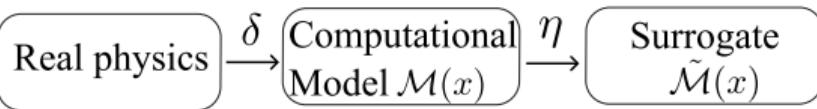
Through an additive observation error:

$$\mathbf{y} = \mathcal{M}(\mathbf{x}) + \boldsymbol{\varepsilon} \quad (1)$$

The rationale behind the equation:

- Assume observation error $\boldsymbol{\varepsilon}$ is additive
- Assume $\boldsymbol{\varepsilon}$ Gaussian form





Other uncertainties?

- Observation error ε
- Model discrepancy $\delta(\mathbf{x})$
- Numerical/truncation error $\eta(\mathbf{x})$

Revised:

$$\begin{aligned}\mathbf{y} &= \mathcal{M}(\mathbf{x}) + \delta(\mathbf{x}) + \varepsilon \\ &= \tilde{\mathcal{M}}(\mathbf{x}) + \eta(\mathbf{x}) + \delta(\mathbf{x}) + \varepsilon\end{aligned}$$

Question: where should the uncertainties go to?

$$\pi(\mathbf{x}|\mathcal{Y}) = \frac{\mathcal{L}(\mathbf{x}|\mathcal{Y}) \cdot \pi(\mathbf{x})}{\pi(\mathcal{Y})} = \frac{\mathcal{L}(\mathbf{x}|\mathcal{Y}) \cdot \pi(\mathbf{x})}{\int_{\mathcal{D}_X} \pi(\mathbf{x}) \pi(\mathcal{Y}|\mathbf{x}) d\mathbf{x}}$$

Incorporate uncertainties into $\mathcal{L}(\mathbf{x}|\mathcal{Y})$

If only consider observations error ϵ

$$\mathbf{y}_i = \tilde{\mathcal{M}}(\mathbf{x}) + \boldsymbol{\epsilon}, i = 1, \dots, k; \boldsymbol{\epsilon} \in \mathcal{N}(\boldsymbol{\epsilon}|\mathbf{0}, \boldsymbol{\Sigma})$$

$$\begin{aligned}\mathcal{L}(\mathbf{x}|\mathcal{Y}) &= \prod_{i=1}^k N(\mathbf{y}_i | \tilde{\mathcal{M}}(\mathbf{x}), \boldsymbol{\Sigma}) \\ &= \prod_{i=1}^k \frac{1}{\sqrt{(2\pi)^N \det(\boldsymbol{\Sigma})}} \exp \left(-\frac{1}{2} (\mathbf{y}_i - \tilde{\mathcal{M}}(\mathbf{x}))^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y}_i - \tilde{\mathcal{M}}(\mathbf{x})) \right)\end{aligned}$$

Incorporate all uncertainties into $\boldsymbol{\Sigma}$

$$\mathbf{y} = \tilde{\mathcal{M}}(\mathbf{x}) + \eta(\mathbf{x}) + \delta(\mathbf{x}) + \boldsymbol{\epsilon}$$

Numeric/
Truncation
Model
discrepancy
Observation
error



Difficulty with calculating evidence $\pi(\mathcal{Y})$

$$\pi(\mathbf{x}|\mathcal{Y}) = \frac{\mathcal{L}(\mathbf{x}|\mathcal{Y}) \cdot \pi(\mathbf{x})}{\pi(\mathcal{Y})}$$

Computational methods:

$$\pi(\mathbf{x}|\mathcal{Y}) \approx \mathcal{L}(\mathbf{x}|\mathcal{Y}) \cdot \pi(\mathbf{x})$$

Samples from the posterior can be obtained through *Sampling methods* or
Optimization methods: MCMC, variational methods,...

- ① UQ and uncertainty types
- ② UQ framework
- ③ How to consider uncertainties into UQ framework
- ④ Final goal of UQ
- ⑤ Current work and future considerations

A unified and scalable digital twin

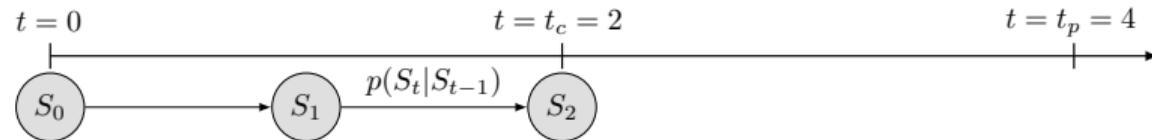
Digital twin is used to:

- Visualize the calibration and prediction
- Make prompt actions

A unified and scalable digital twin

Digital twin is used to:

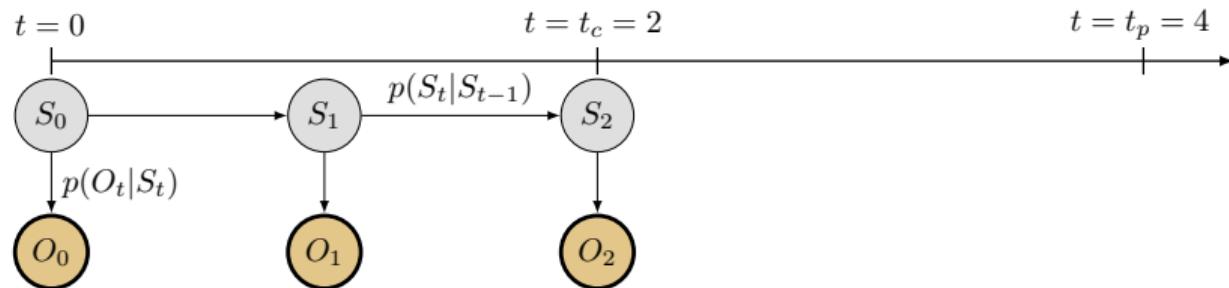
- Visualize the calibration and prediction
- Make prompt actions



A unified and scalable digital twin

Digital twin is used to:

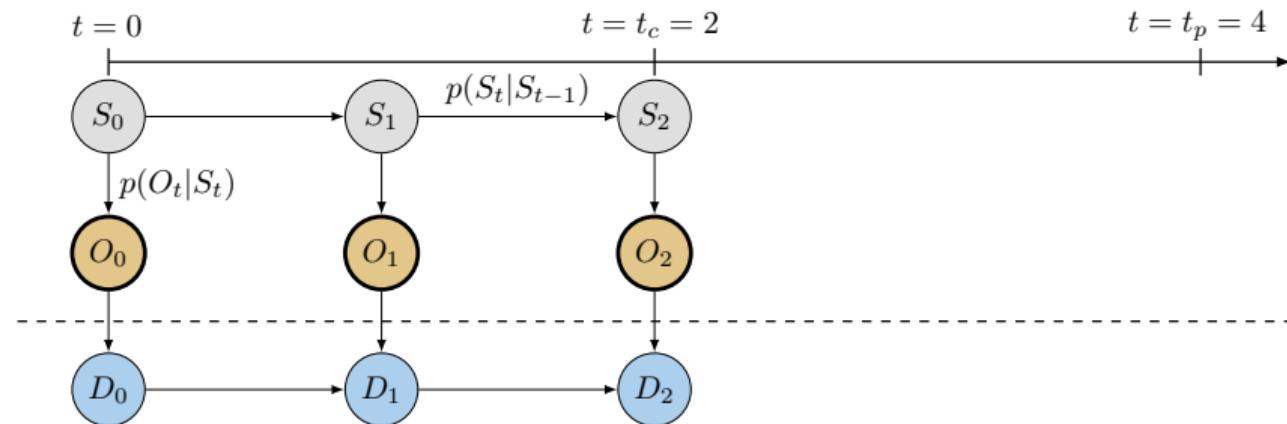
- Visualize the calibration and prediction
- Make prompt actions



A unified and scalable digital twin

Digital twin is used to:

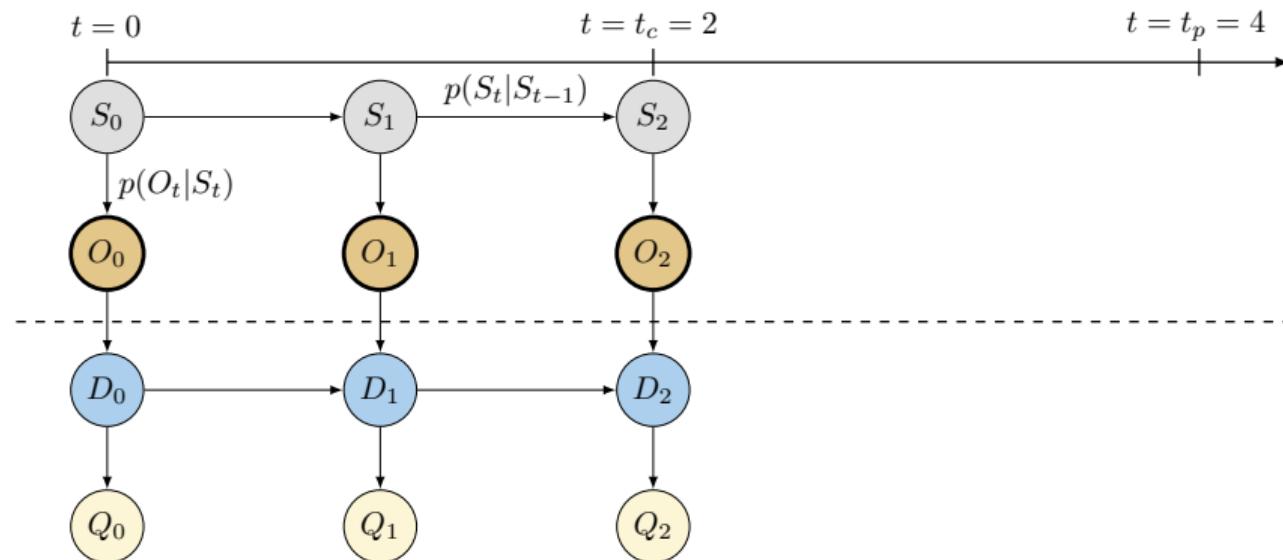
- Visualize the calibration and prediction
- Make prompt actions



A unified and scalable digital twin

Digital twin is used to:

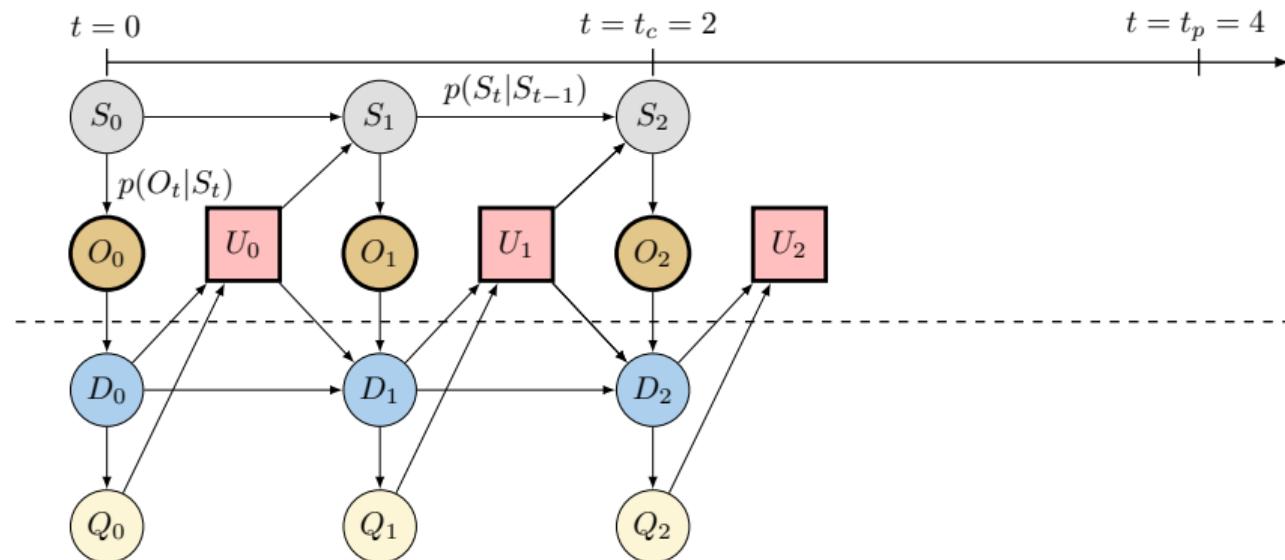
- Visualize the calibration and prediction
- Make prompt actions



A unified and scalable digital twin

Digital twin is used to:

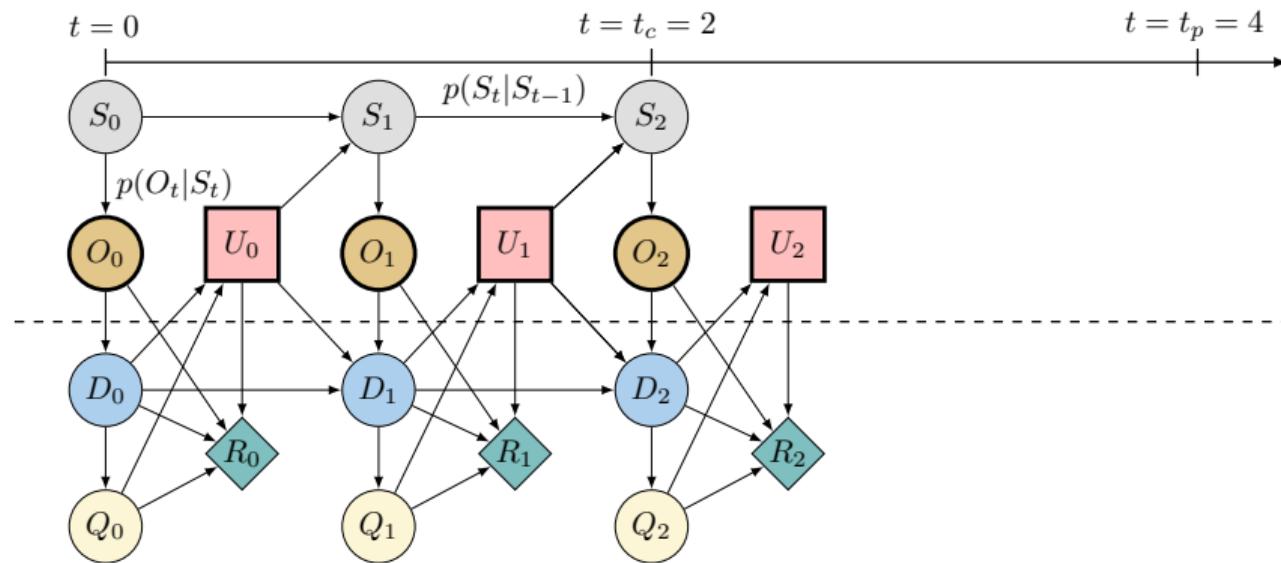
- Visualize the calibration and prediction
- Make prompt actions



A unified and scalable digital twin

Digital twin is used to:

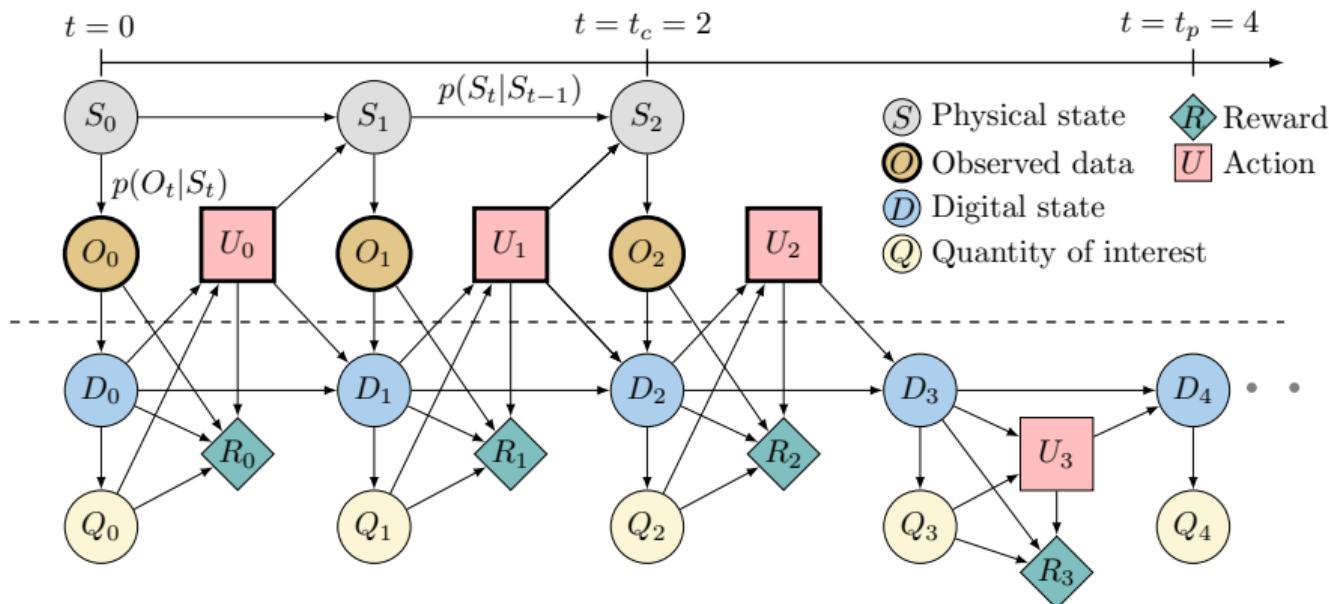
- Visualize the calibration and prediction
- Make prompt actions



A unified and scalable digital twin

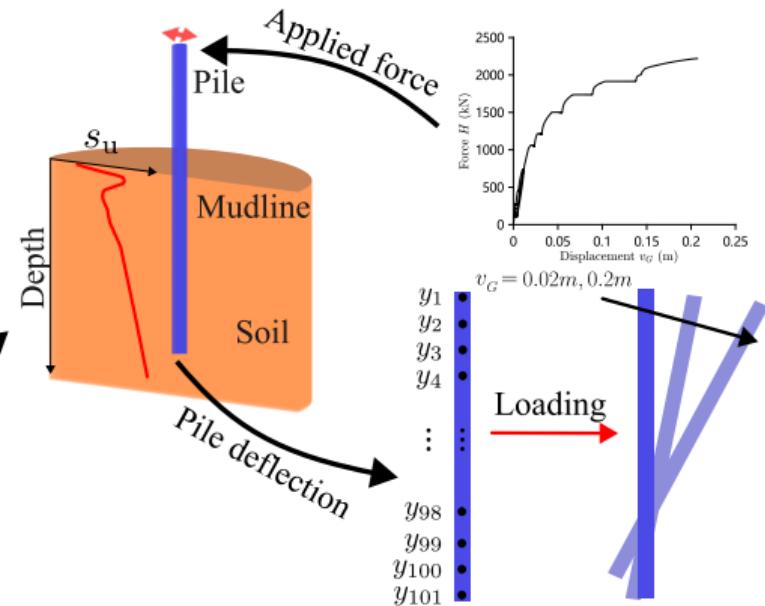
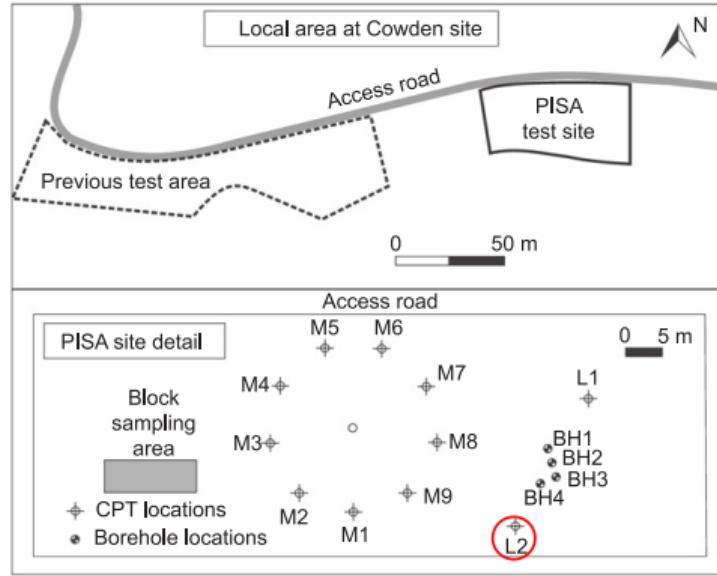
Digital twin is used to:

- Visualize the calibration and prediction
- Make prompt actions



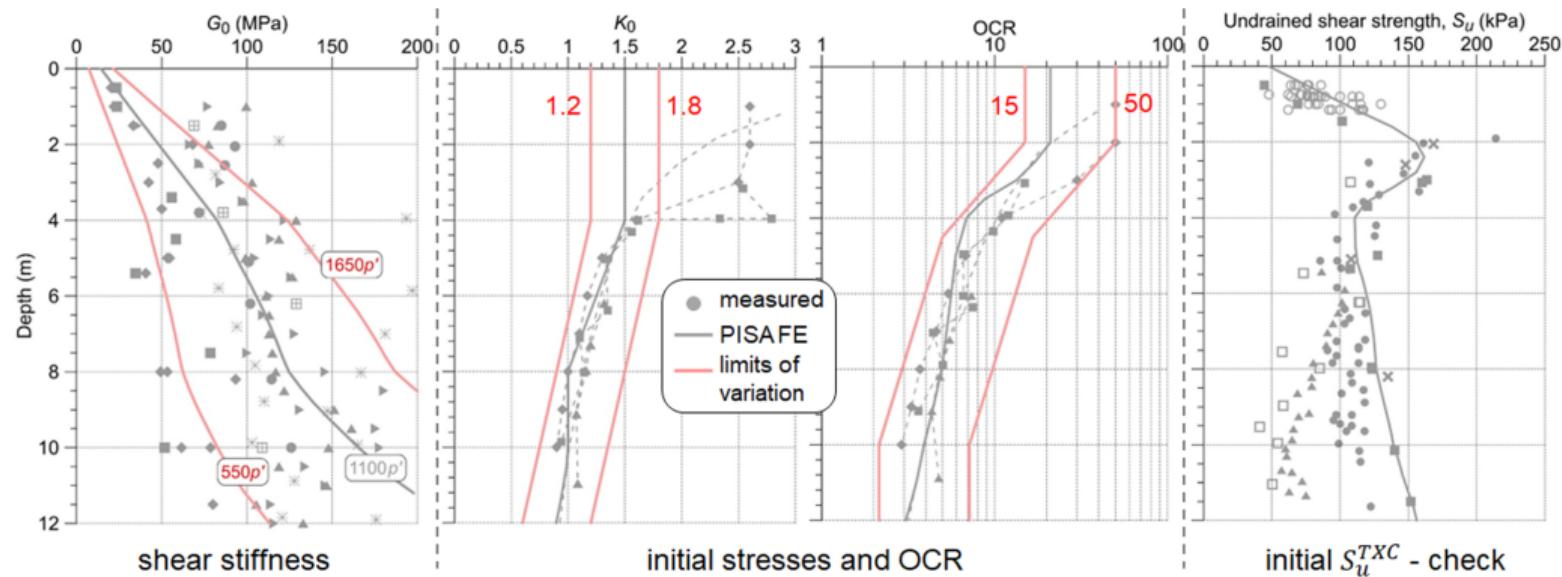
- ① UQ and uncertainty types
- ② UQ framework
- ③ How to consider uncertainties into UQ framework
- ④ Final goal of UQ
- ⑤ Current work and future considerations

Application on PISA pile



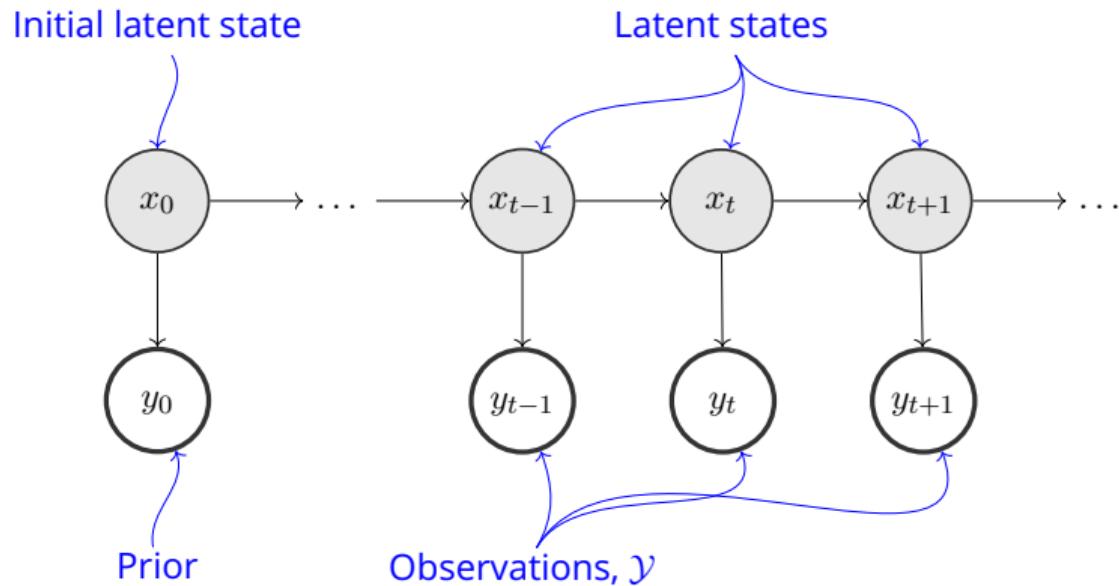
Application on PISA pile

Range of Parameters

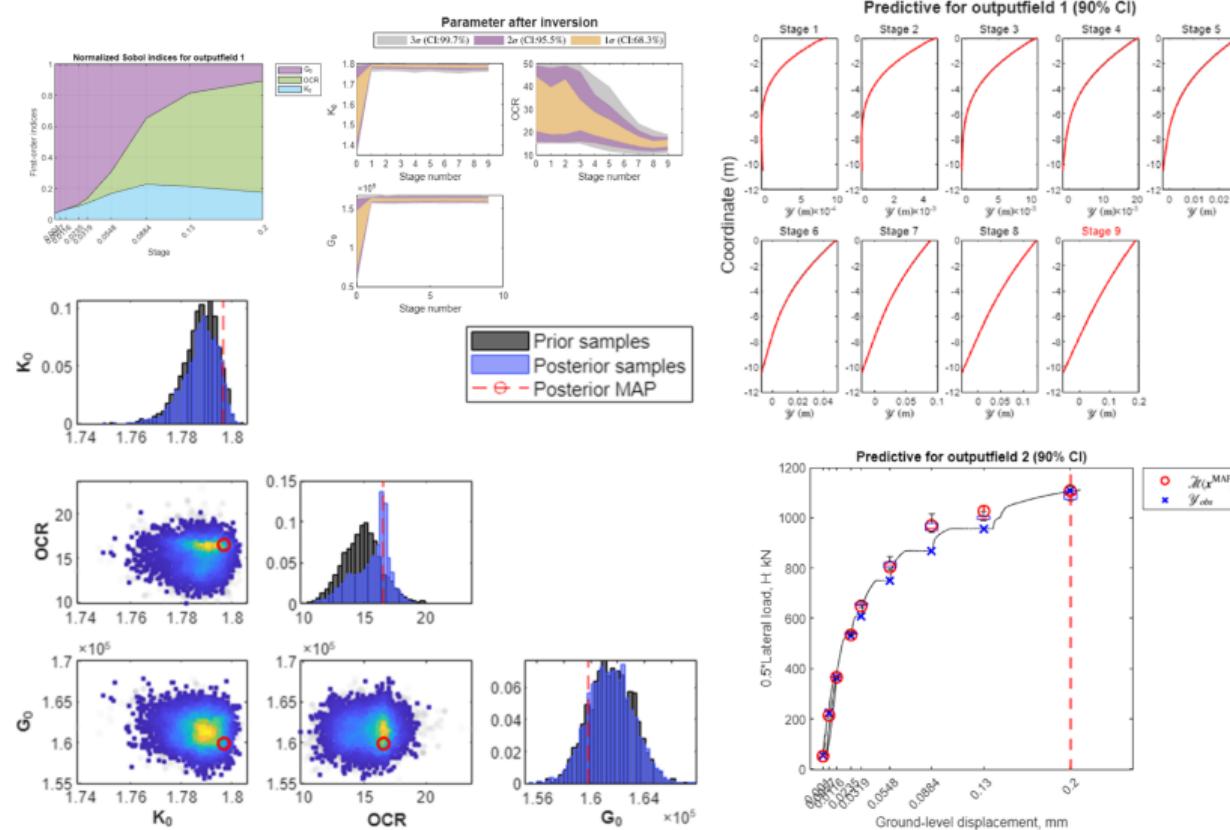


⁴Lidija Zdravković. "Geotechnical Engineering for a Sustainable Society". In: Rankine lecture (2024).

Application on PISA pile



Application on PISA pile



Future considerations:

- ① Careful with aleatoric and epistemic types of uncertainties
- ② Four UQ components are equally important
- ③ Consider different types of uncertainty into Bayesian inference required linked with likelihood-(**currently doing**)
- ④ Consider taking actions in digital twin-(**future plan**)

References I

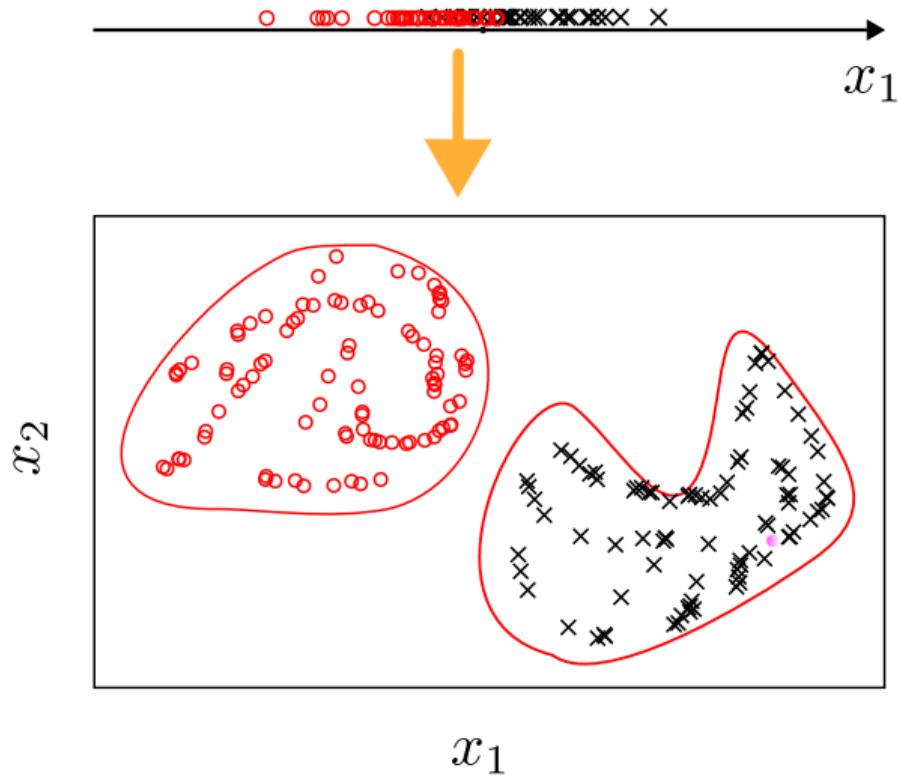
-  Nagel, Joseph B, Jörg Rieckermann, and Bruno Sudret. "Principal component analysis and sparse polynomial chaos expansions for global sensitivity analysis and model calibration: Application to urban drainage simulation". In: *Reliability Engineering & System Safety* 195 (2020), p. 106737.
-  Saouma, Victor E. and M. Amin Hariri-Ardebili. "Uncertainty Quantification". In: *Aging, Shaking, and Cracking of Infrastructures: From Mechanics to Concrete Dams and Nuclear Structures*. Cham: Springer International Publishing, 2021, pp. 423–454.
-  Sudret, Bruno. "Uncertainty quantification in the simulation of complex systems". In: *1st International Conference on Infrastructure Resilience*. ETH Zurich. 2018.
-  Zdravković, Lidija. "Geotechnical Engineering for a Sustainable Society". In: *Rankine lecture* (2024).
-  Zdravković, Lidija et al. "Ground characterisation for PISA pile testing and analysis". In: *Géotechnique* 70.11 (2020), pp. 945–960.

Thank you!

Questions? Comments?

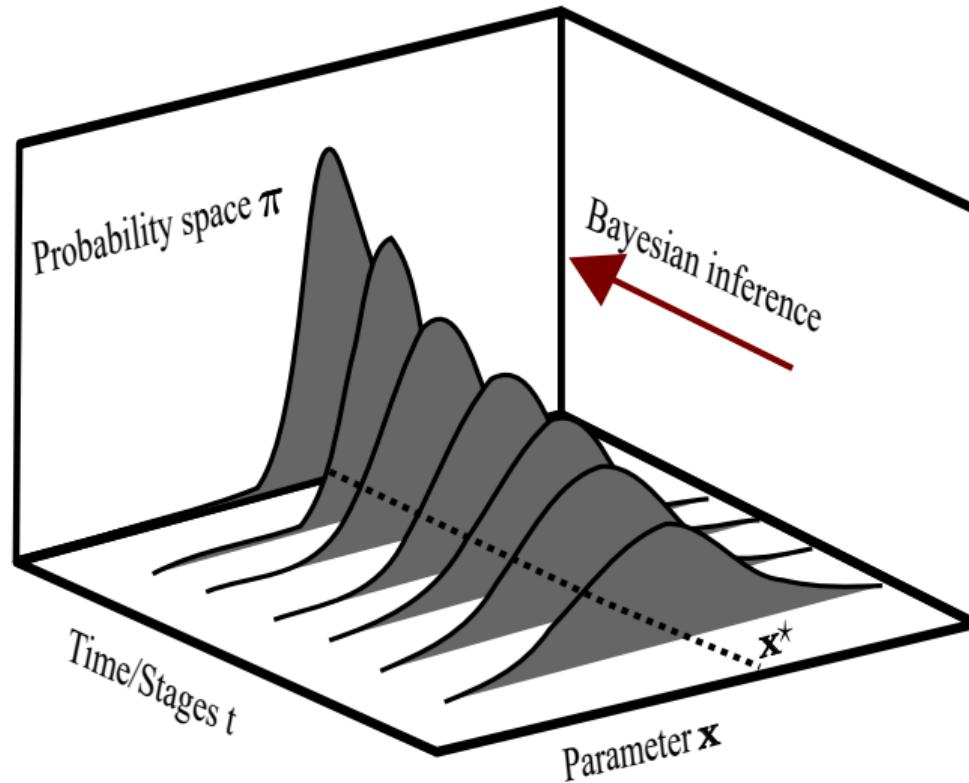
Appendix:

Aleatoric to epistemic



Appendix:

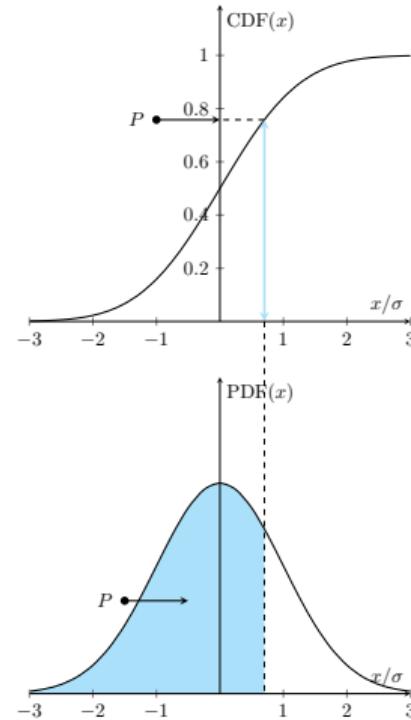
Sequential Bayesian inference



Appendix:

Inverse probability transform

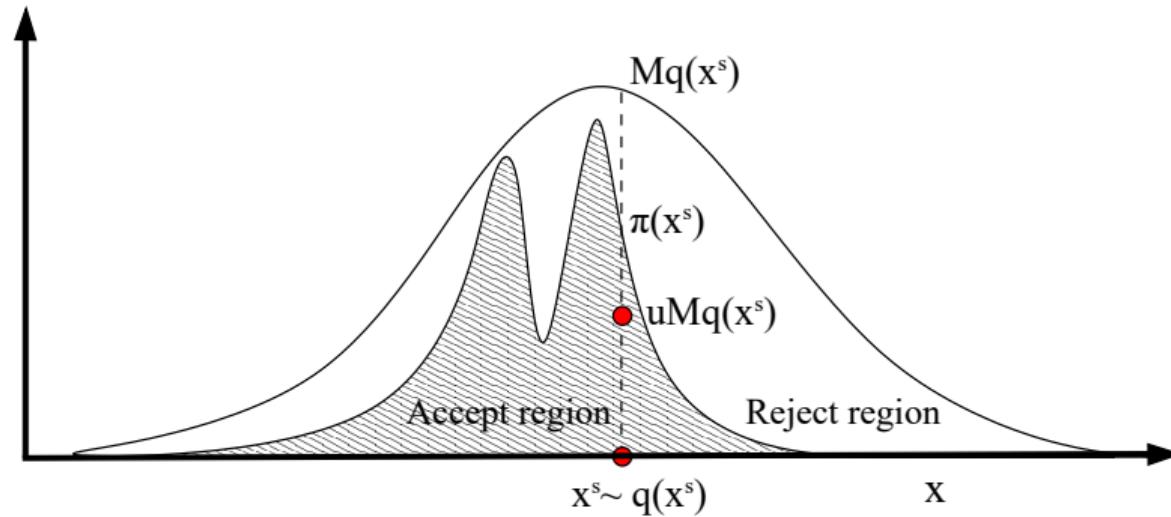
- Inverse probability transform



Appendix:

Rejection sampling

- Rejection sampling



Appendix:

importance sampling

$$E[f(\mathbf{x}_t)] = \int f(\mathbf{x}_t) \pi(\mathbf{x}_t) d\mathbf{x}_t \approx \frac{1}{S} \sum_{s=1}^S f(\mathbf{x}_t^s)$$

$$E[f(\mathbf{x}_t)] = \int f(\mathbf{x}_t) \pi(\mathbf{x}_t) d\mathbf{x}_t = \int f(\mathbf{x}_t) \frac{\pi(\mathbf{x}_t)}{q(\mathbf{x}_t)} q(\mathbf{x}_t) d\mathbf{x}_t \approx \frac{1}{S} \sum_{s=1}^S f(\mathbf{x}_t^s) \frac{\pi(\mathbf{x}_t^s)}{q(\mathbf{x}_t^s)}$$

- Require a proper proposal distribution $q(\mathbf{x}_t)$

Appendix:

Metropolish-Hasting

$$f(\mathbf{x}_t^{s+1} | \mathbf{x}_t^s) = \min(1, \alpha)$$
$$\alpha = \min\left(1, \frac{q(\mathbf{x}_t^s | \mathbf{x}_t^{s+1}) \pi(\mathbf{x}_t^{s+1} | \mathcal{Y}_t)}{q(\mathbf{x}_t^{s+1} | \mathbf{x}_t^s) \pi(\mathbf{x}_t^s | \mathcal{Y}_t)}\right)$$

Appendix:

Metropolish-Hasting

Algorithm 1: MH algorithm at t_{th} step

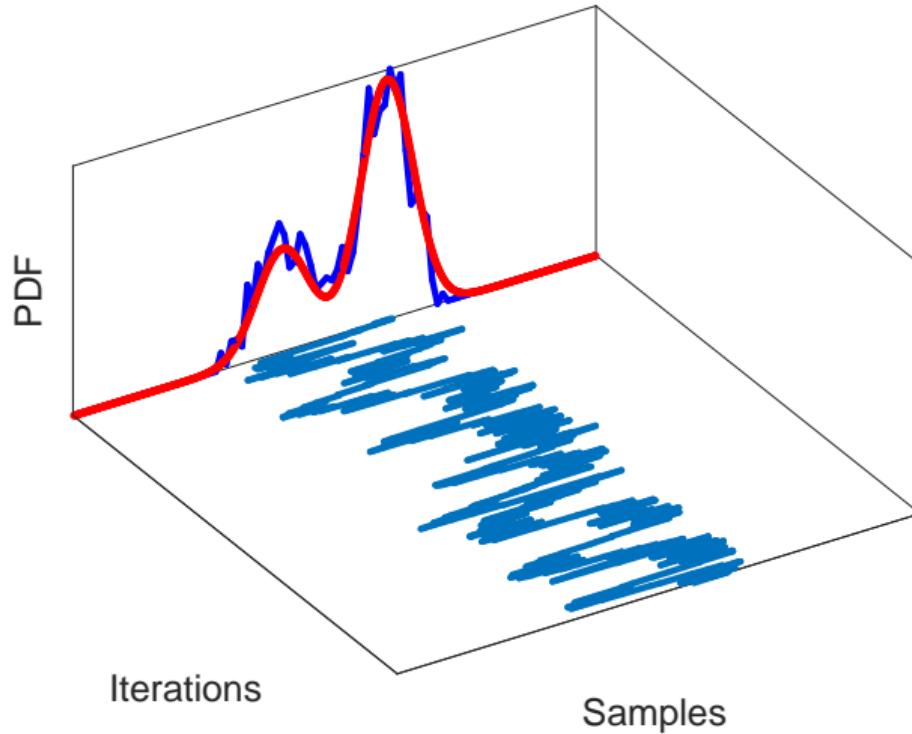
Data: $q(\mathbf{x}_t)$: Proposal distribution; $\pi(\mathbf{x}_t|\mathcal{Y})$: Target posterior.

Result: MCMC samples at t_{th} stage: $\mathcal{X}_t = \{\mathbf{x}_t^1, \dots, \mathbf{x}_t^{N_X}\}$

- 1 Initialization $\mathbf{x}_t^1 \in \mathcal{D}_{\mathbf{X}}$;
 - 2 **for** $s \leftarrow 2$ to N_X **do**
 - 3 Sample $\mathbf{x}_t^{s+1} \sim q(\mathbf{x}_t^{s+1} | \mathbf{x}_t^s)$;
 - 4 Compute acceptance probability α ;
 - 5 Compute $f(\mathbf{x}_t^{s+1} | \mathbf{x}_t^s) = \min(1, \alpha)$;
 - 6 Sample $u \sim \mathcal{U}(0, 1)$;
 - 7 Set candidate sample $\mathbf{x}_t^{(*)}$ to \mathbf{x}_t^{s+1} with probability α ;
 - 8 **end for**
-

Appendix:

Metropolish-Hasting



Appendix:

AIES

$$\boldsymbol{x}_t^{(\star)} = \boldsymbol{x}_{t_i}^{(s)} + z \cdot (\boldsymbol{x}_{t_j}^{(\tilde{s})} - \boldsymbol{x}_{t_i}^{(s)})$$

$$p(z|a) = \begin{cases} \frac{1}{\sqrt{z}(2\sqrt{a} - \frac{2}{\sqrt{a}})} & \text{if } z \in [1/a, a] \\ 0 & \text{otherwise} \end{cases}$$

$$\alpha = \min(1, z^{M-1} \frac{\pi(x_t^{(\star)} | \mathcal{Y})}{\pi(x_{t_i}^{(s)} | \mathcal{Y})})$$

Appendix:

AIES

Algorithm 2: AIES algorithm at t_{th} step

Data: $\pi(x_t | \mathcal{Y})$: Target posterior; tuning parameter a

Result: MCMC samples at t_{th} stage: $\mathcal{X}_t = \{\mathcal{X}_{t-1}, \dots, \mathcal{X}_{t-N_{chain}}\}$, with

$$\mathcal{X}_{t-i} = \{x_{t-i}^1, \dots, x_{t-i}^{N_{\mathcal{X}}}\}$$

- 1 Initialization N_{chain} samples $\{x_{t-1}^1, \dots, x_{t-N_{chain}}^1\}$, with $x_{t-i}^1 \in \mathcal{D}_X$
 - 2 **for** $s \leftarrow 2$ to $N_{\mathcal{X}}$ **do**
 - 3 **for** $i \in \{1, \dots, N_{chain}\}$ **do**
 - 4 Pick random j from $\{1, \dots, N_{chain}\} \setminus i$;
 - 5 Propose $x_t^{(\star)}$ with ??;
 - 6 Set $x_{t-i}^s = x_t^{(\star)}$ with probability α (see ??);
 - 7 **end for**
 - 8 **end for**
-

Appendix:

Sequential Monte Carlo

$$\boldsymbol{x}_t = g(\boldsymbol{x}_{t-1}) + \boldsymbol{v} \quad (\text{state equation})$$

$$\mathcal{Y}_t = m(\boldsymbol{x}_t) + \boldsymbol{w} \quad (\text{observation equation})$$

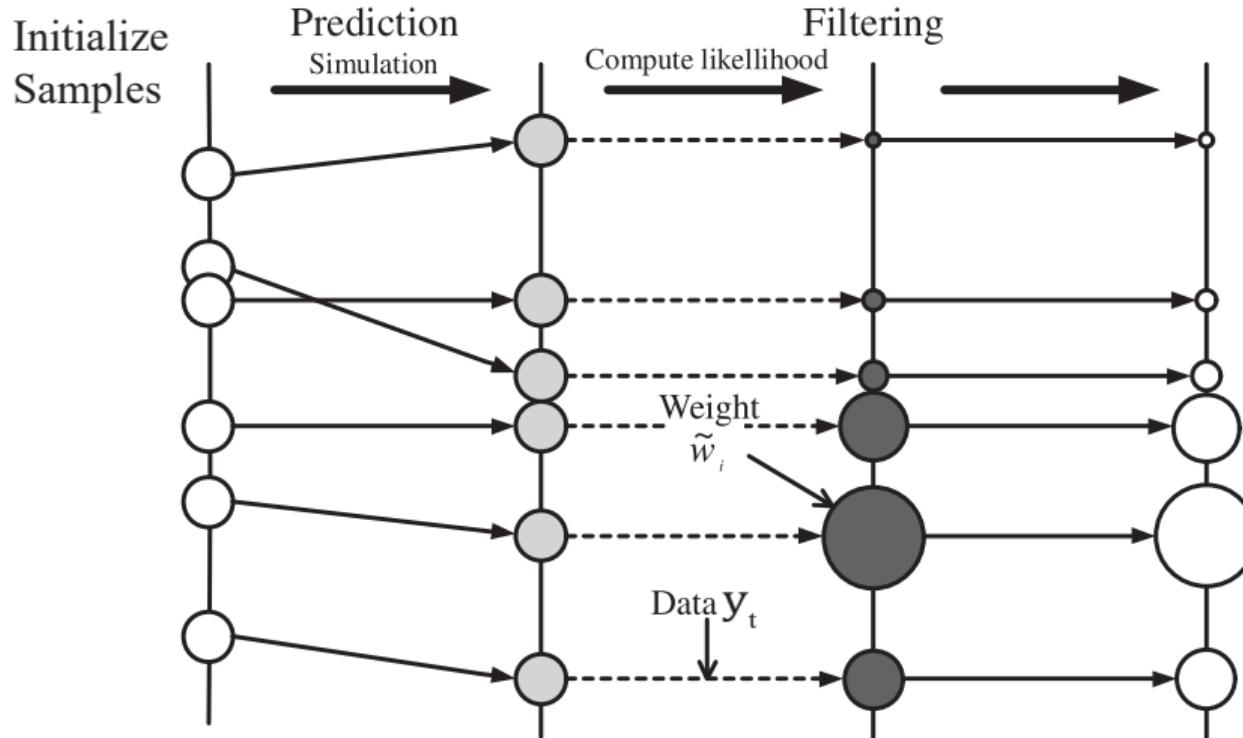
$$\pi(\boldsymbol{x}_{1:t} | \mathcal{Y}_{1:t}) \approx \sum_{s=1}^S \tilde{w}_t^s \delta_{\boldsymbol{x}_{1:t}^s}(\boldsymbol{x}_{1:t})$$

$$\tilde{w}_t^s = \frac{w_t^s}{\sum_{s=1}^S (w_t^s)}$$

$$\pi(\boldsymbol{x}_{1:t} | \mathcal{Y}_{1:t}) \propto \pi(\mathcal{Y}_t | \boldsymbol{x}_t) \pi(\boldsymbol{x}_t | \boldsymbol{x}_{t-1}) \pi(\boldsymbol{x}_{t-1} | \mathcal{Y}_{t-1})$$

Appendix:

Sequential Monte Carlo



Appendix:

Sequential Monte Carlo with resampling

Algorithm 3: SISR algorithm at t_{th} step

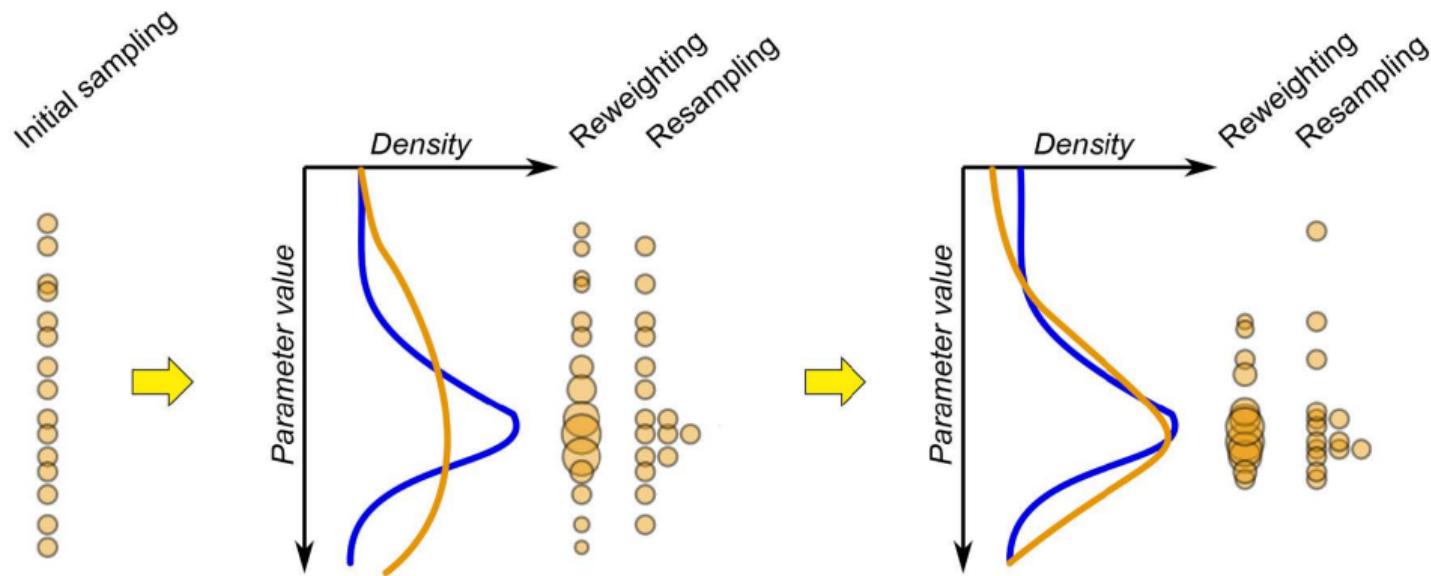
Data: Samples \mathbf{x}_{t-1}^s with weights w_{t-1}^s , $s = \{1, \dots, N\}$; observation \mathcal{Y}_t at t_{th} stage

Result: SMC samples with normalized weights \tilde{w}_t^s at t_{th} stage: $\mathbf{x}_t^{(*)} = \{\mathbf{x}_t^1, \dots, \mathbf{x}_t^N\}$)

```
1 for  $s \leftarrow 1$  to  $N$  do
2   | Sample from proposal distribution  $\mathbf{x}_t^s \sim q(\mathbf{x}_t^s | \mathbf{x}_{t-1}^s, \mathcal{Y}_t)$ ;
3   | Compute weight using ??;
4 end for
5 Normalized weights;
6 Calculate degeneracy measure using ??;
7 if  $\hat{S}_{eff} < S$  then
8   | Resample;
9 end if
```

Appendix:

Sequential Monte Carlo with resampling



Appendix:

Visualized PCE

