Adaptive Construction for Surrogate-based Inference

POLETTE Nadège, *CEA DAM DIF/Mines Paris PSL*

Supervisors:

SOCHALA Pierre, *CEA DAM DIF*

GESRET Alexandrine, *Mines Paris PSL*

LE MAÎTRE Olivier, *CMAP, CNRS*

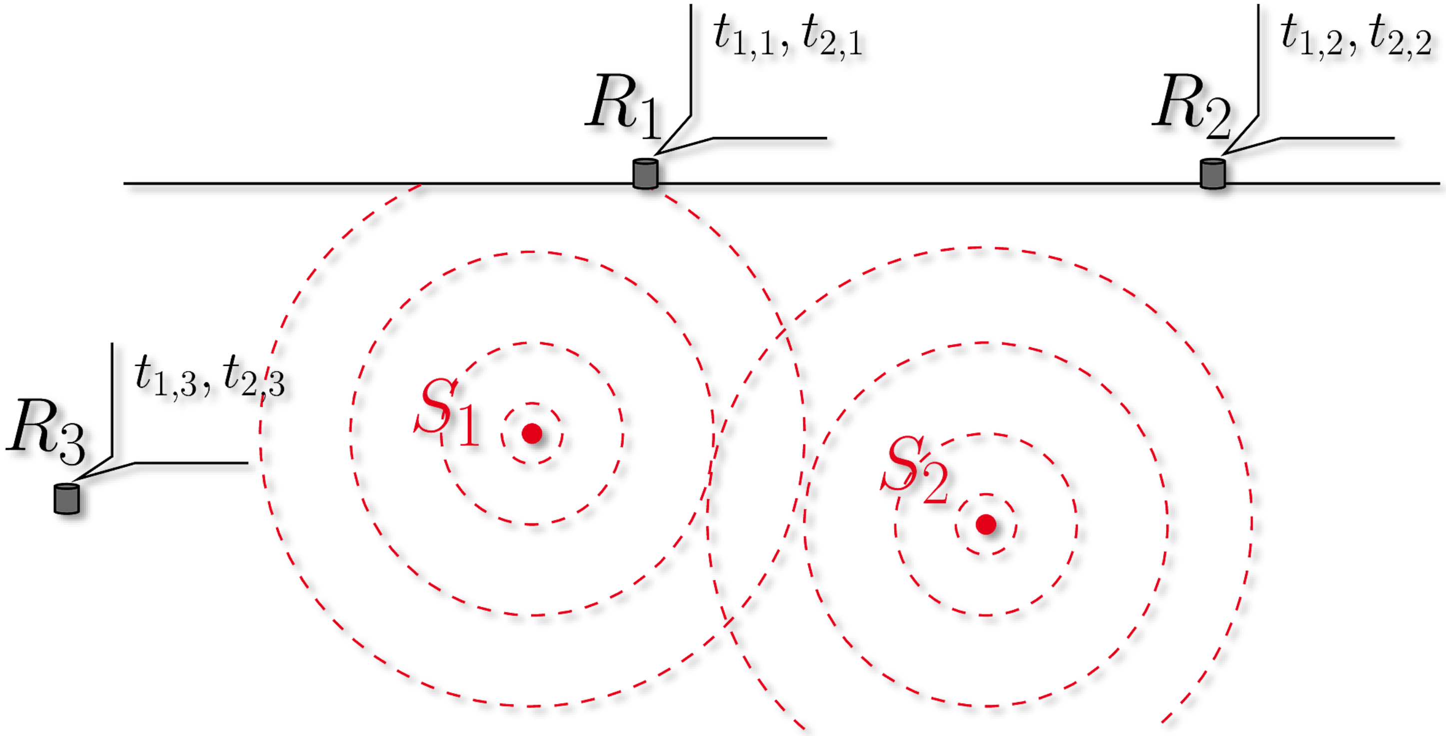
Table of contents

1. Context: Detection and analysis of seismic events

Global scale
- International treaties (CTBT, NTP)
- Environment monitoring (IMS)

Regional scale
- Tsunami and earthquake alerts
- Risk prevention

Local scale
- Subsurface knowledge
- Exploitation

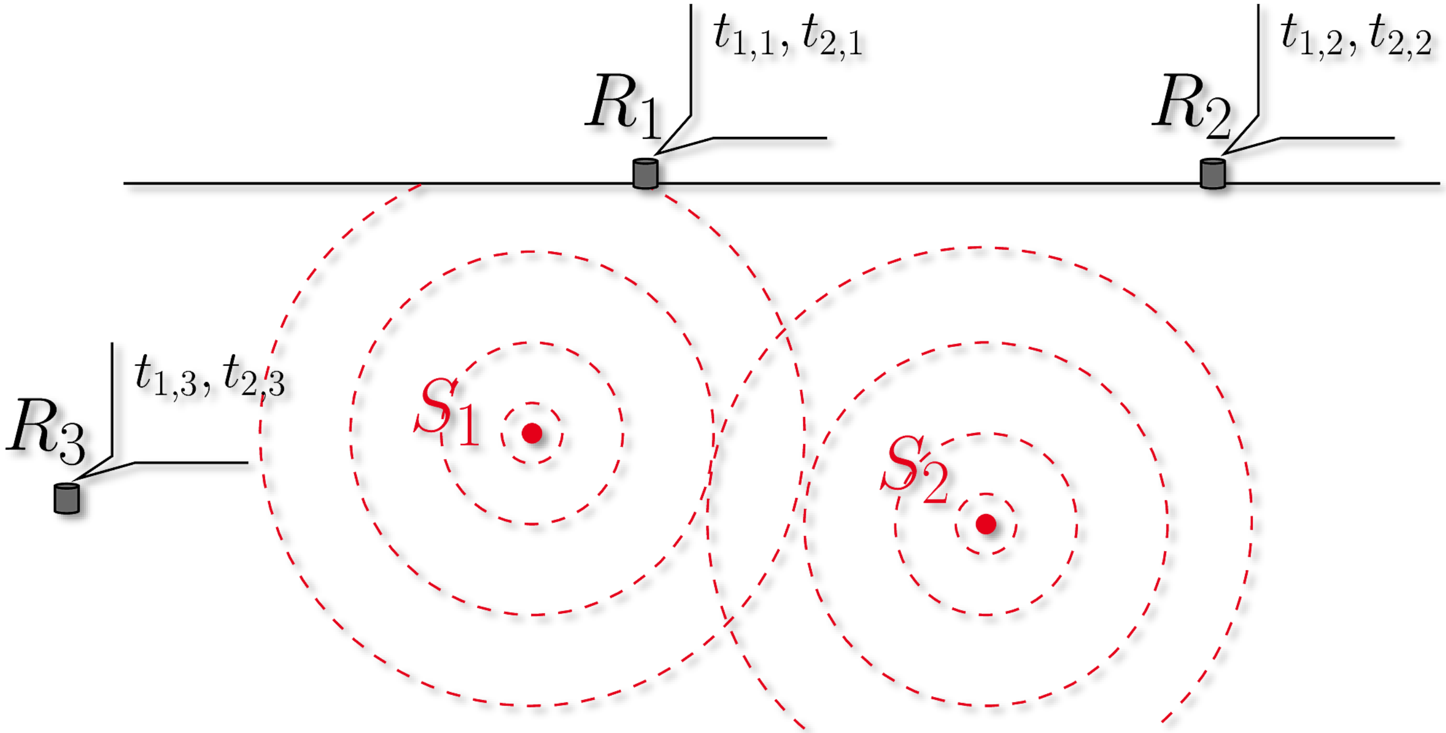


1. Context: Detection and analysis of seismic events

Global scale
- International treaties (CTBT, NTP)
- Environment monitoring (IMS)

Regional scale
- Tsunami and earthquake alerts
- Risk prevention

Local scale
- Subsurface knowledge
- Exploitation

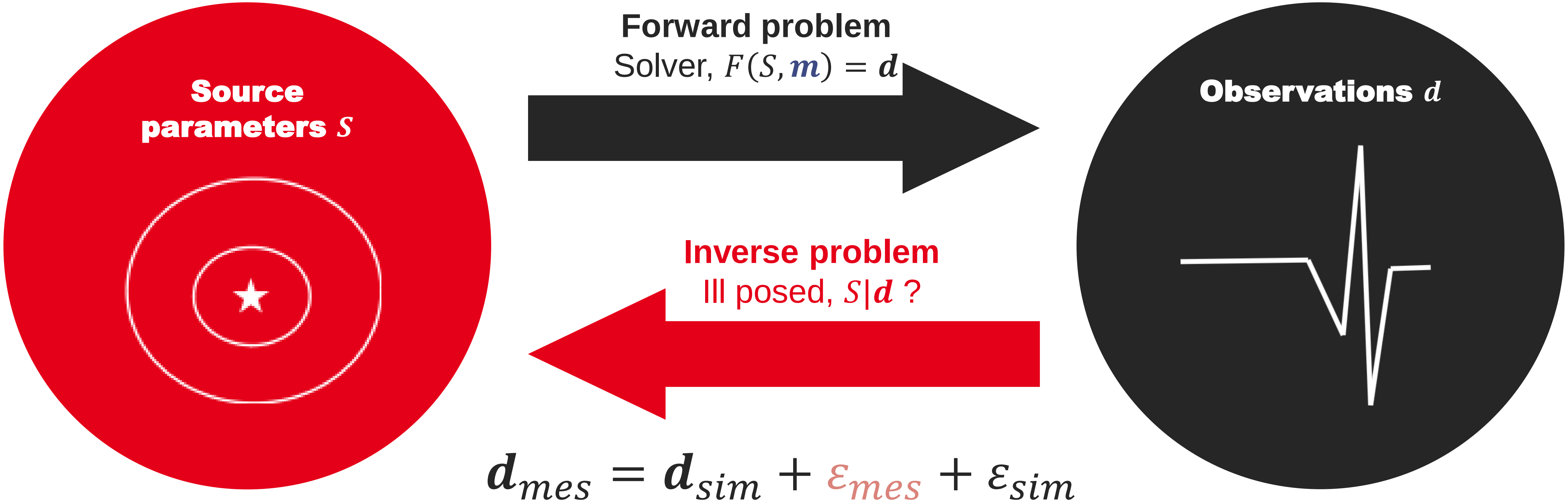


$$F(S) = d$$

F: forward model
S: source parameters
d: data

Objective: retrieve S from d
- fast
- with accuracy
- with uncertainties

1. Context: Inverse problem



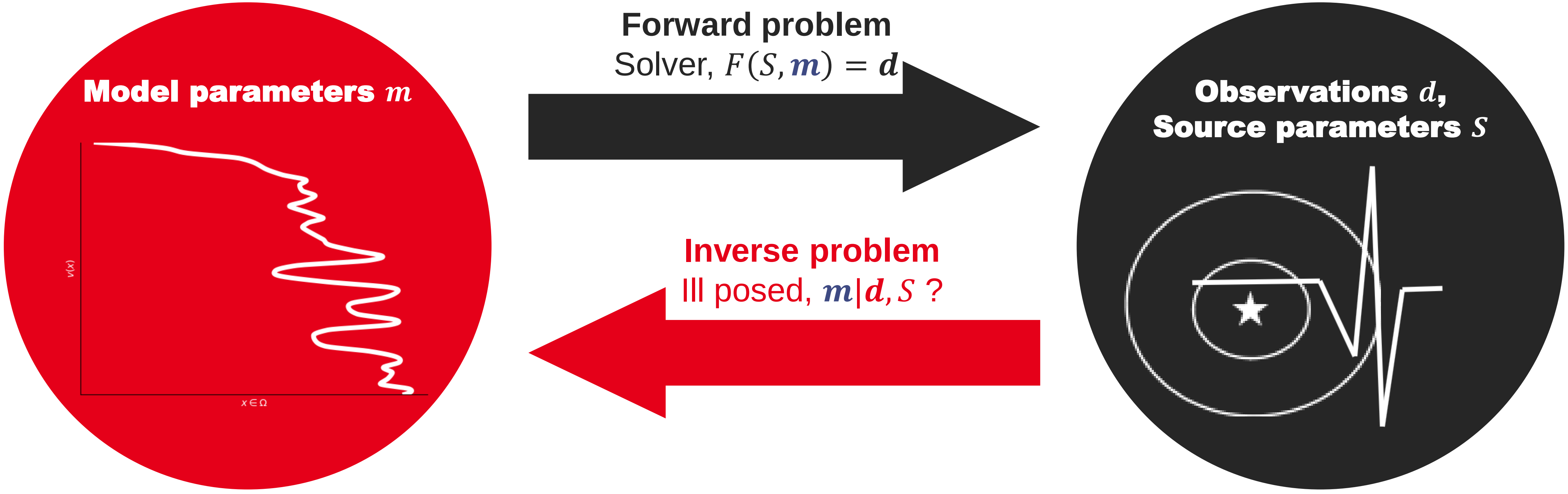
$$\boldsymbol{d}_{mes} = \boldsymbol{d}_{sim} + \varepsilon_{mes} + \underbrace{\varepsilon_{sim}}_{\varepsilon_{num} + \varepsilon_{mod} + \varepsilon_{par}}$$

Uncertainty sources: **observations**, physical model, **model parameters**, ...

Objective: improve **uncertainty quantification** of model parameters

A. Tarantola, *Inverse Problem Theory and Methods for Model Parameter Estimation*, SIAM 2005

1. Context: Inverse problem



Objective: **to characterize the velocity field m and its uncertainty** from indirect observations $\boldsymbol{d}$

$\Rightarrow$ to find the probability distribution of the field knowing the observations $\boldsymbol{P(m|d_{mes})}$

A. Tarantola, *Inverse Problem Theory and Methods for Model Parameter Estimation*, SIAM 2005

1. Context: Bayesian inference

Proposition $\boldsymbol{m}^{(n)}$

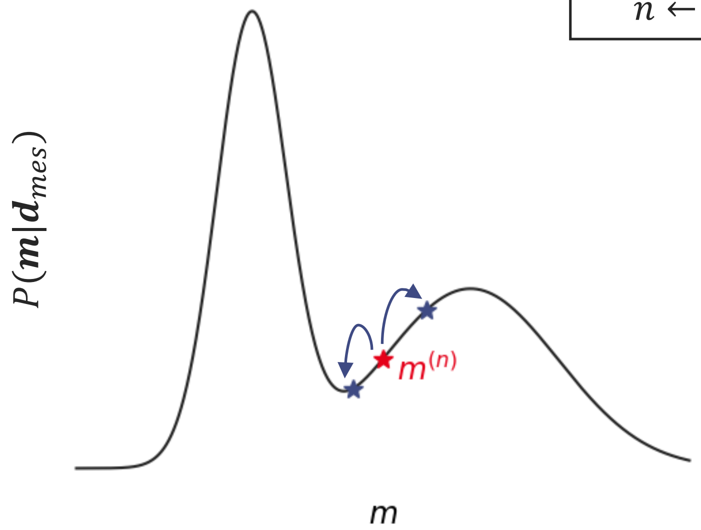
Forward model
Solver $F(\boldsymbol{m}; S) = \boldsymbol{d}$

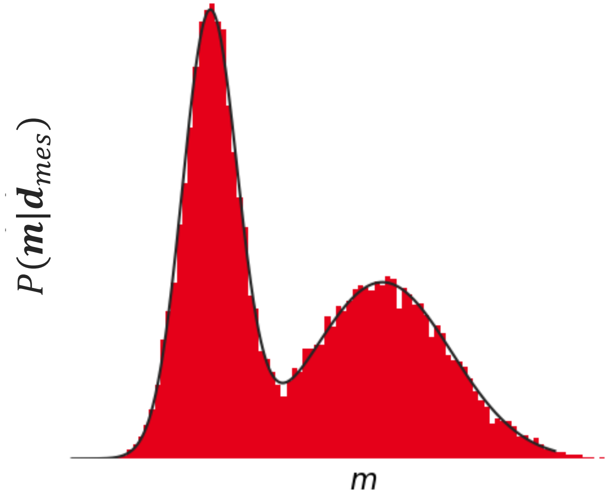
Comparison
$\boldsymbol{d}$ vs $\boldsymbol{d}_{mes}$ (data)

Bayes' rule : $P(\boldsymbol{m}|\boldsymbol{d}_{mes}) \propto L(\boldsymbol{d}_{mes}|\boldsymbol{m})P(\boldsymbol{m})$

New iteration
$n \leftarrow n+1$

Accept/Reject criterion





D. Sivia, J. Skilling, *Data Analysis: A Bayesian Tutorial*, Oxford, sc. Pub. 2006
A. Doucet et al., *Sequential Monte Carlo Methods in Practice*, ISS 2013

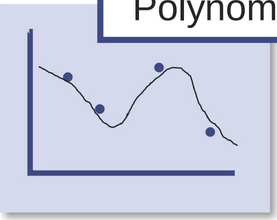
1. Context: Bayesian inference

Proposition $\boldsymbol{m}^{(n)}$

$\boldsymbol{m}$ infinite dimensional
→ dimension reduction

- ❑ 'small' dimension
- ❑ fixed dimension
- ❑ allowing for various shapes

Surrogate model
Polynomial chaos $\tilde{F}(\boldsymbol{m}) = \sum_k f_k \Psi_k(\boldsymbol{m}) = \boldsymbol{d}$

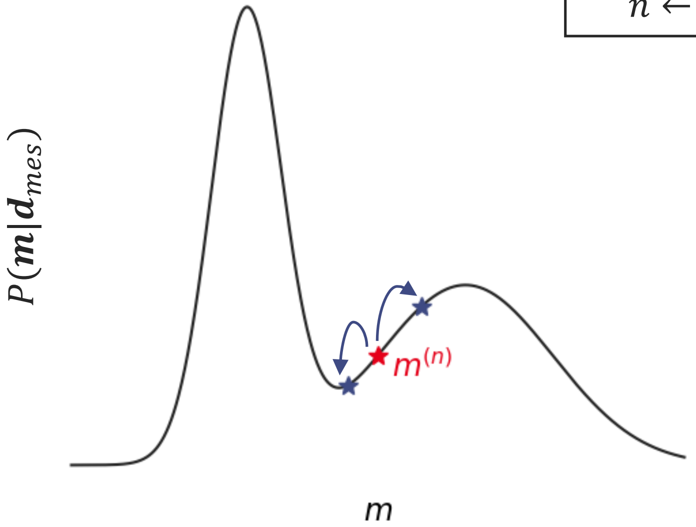


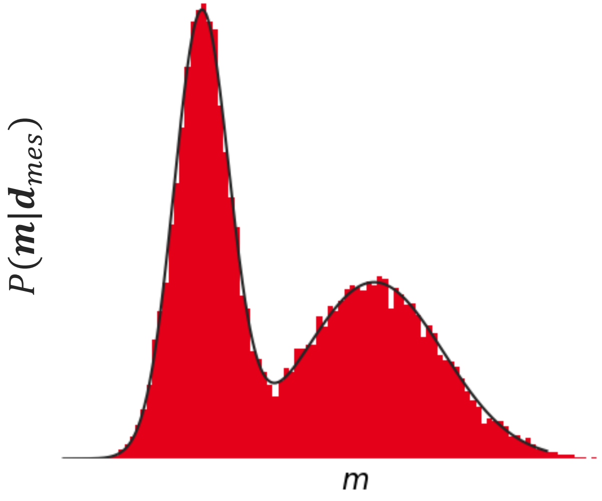
Comparison
$\boldsymbol{d}$ vs $\boldsymbol{d}_{mes}$ (data)

Bayes' rule : $P(\boldsymbol{m}|\boldsymbol{d}_{mes}) \propto L(\boldsymbol{d}_{mes}|\boldsymbol{m})P(\boldsymbol{m})$

Accept/Reject criterion

New iteration
$n \leftarrow n+1$





D. Sivia, J. Skilling, *Data Analysis: A Bayesian Tutorial*, Oxford, sc. Pub. 2006
A. Doucet et al., *Sequential Monte Carlo Methods in Practice*, ISS 2013

1. Context: Bayesian inference

Proposition $m^{(n)}$

m infinite dimensional
→ dimension reduction

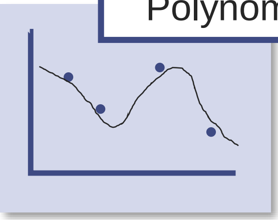
- ☑ 'small' dimension
- ☑ fixed dimension
- ☐ allowing for various shapes
- ▸ Modal decomposition with a change of measure method

$$m(x) = \sum_{i=1}^{r} U_i(x)\eta_i$$

Parameters: $\eta = \{\eta_i\}_{1\leq i\leq r}$

m

Surrogate model
Polynomial chaos $\tilde{F}(\eta) = \sum_k f_k \Psi_k(\eta) = d$

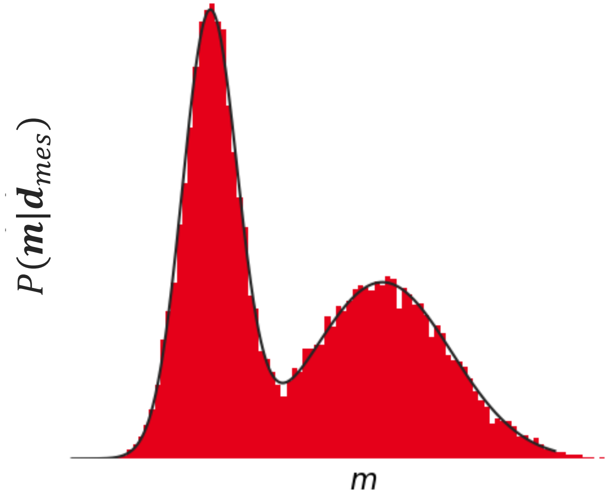


Comparison
d vs d_{mes} (data)

Bayes' rule : $P(\eta|d_{mes}) \propto L(d_{mes}|\eta)P(\eta)$

New iteration
$n \leftarrow n+1$

Accept/Reject criterion



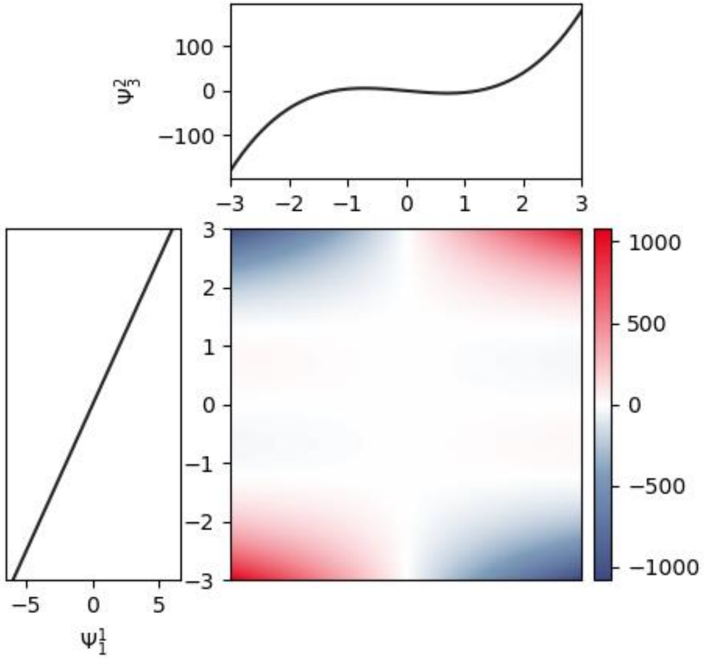
D. Sivia, J. Skilling, *Data Analysis: A Bayesian Tutorial*, Oxford, sc. Pub. 2006
A. Doucet et al., *Sequential Monte Carlo Methods in Practice*, ISS 2013
N. Polette, O. Le Maître, P. Sochala, A. Gesret, *Change of Measure for Bayesian Field Inversion with Hierarchical Hyperparameters Sampling*, In Rev

2. Initial surrogate construction: Polynomials

Surrogate model

Polynomial chaos $\tilde{F}(\boldsymbol{\eta}) = \sum_{a \in A} f_a \psi_a(\boldsymbol{\eta}) = \boldsymbol{d}$

- A: set of multi-indices $\{(0,1,0); (2,0,0); (1,0,1)\}$
- Ψ_a: product of orthonormal univariate polynomials:
  $$\Psi_{a=(a_1,\ldots,a_r)}(\boldsymbol{\eta}) = \prod_{i=1}^{r} \psi_{a_i}^{i}(\eta_i)$$
- f_a: coefficients to compute



Example with Hermite polynomials and $a = (1,1,0)$

2. Initial surrogate construction: Coefficients

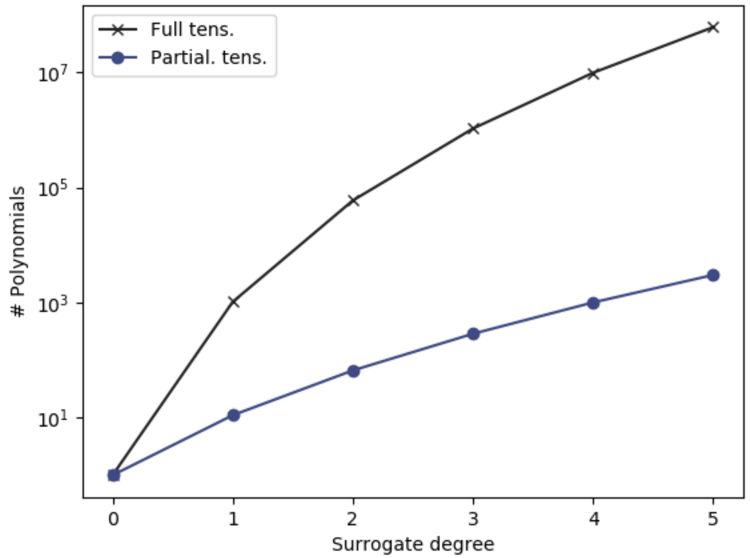
Surrogate model

Polynomial chaos $\tilde{F}(\boldsymbol{\eta}) = \sum_{a \in A} f_a \psi_a(\boldsymbol{\eta}) = \boldsymbol{d}$

Non-intrusive ordinary least squares approach

- $\{\boldsymbol{\eta}^{(n)}\}_{1 \leq n \leq N} \sim P(\boldsymbol{\eta})$, training set $N \gg K = |\mathcal{A}|$
- $\boldsymbol{U} = \left(F\left(\boldsymbol{\eta}^{(1)}\right), \dots, F\left(\boldsymbol{\eta}^{(N)}\right)\right)^T$, training evaluations
- $\boldsymbol{\Psi} \in \mathbb{R}^{N \times K}$, polynomial evaluations at training points $\Psi_{ij} = \psi_j(\eta^{(i)})$
- $\boldsymbol{f} = (f_1, \dots, f_K)^T$, vector of coefficients

$$\boldsymbol{f} = (\boldsymbol{\Psi}^T \boldsymbol{\Psi})^{-1} \boldsymbol{\Psi}^T \boldsymbol{U}$$



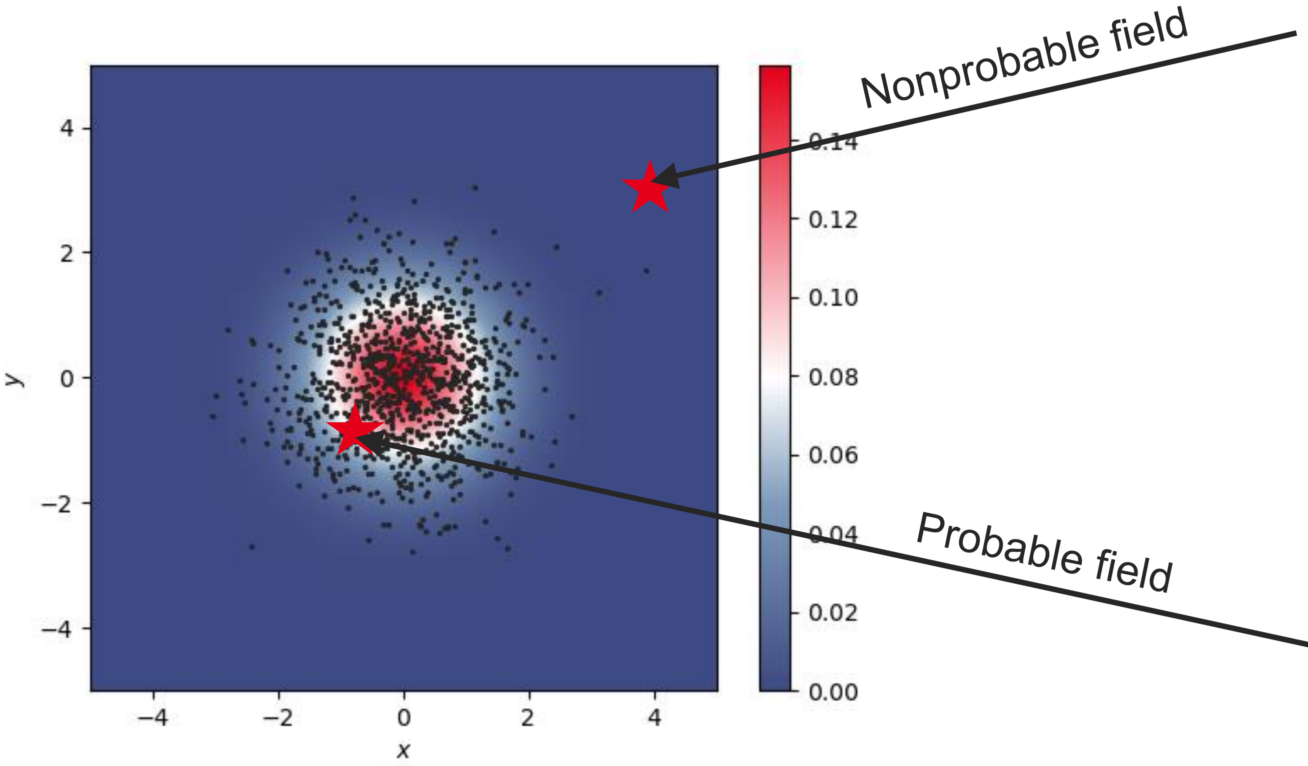
Number of polynomials according to surrogate degree n_o $(r = 10)$

Full tensorization: $\max(n_{o,i}) \leq n_o$

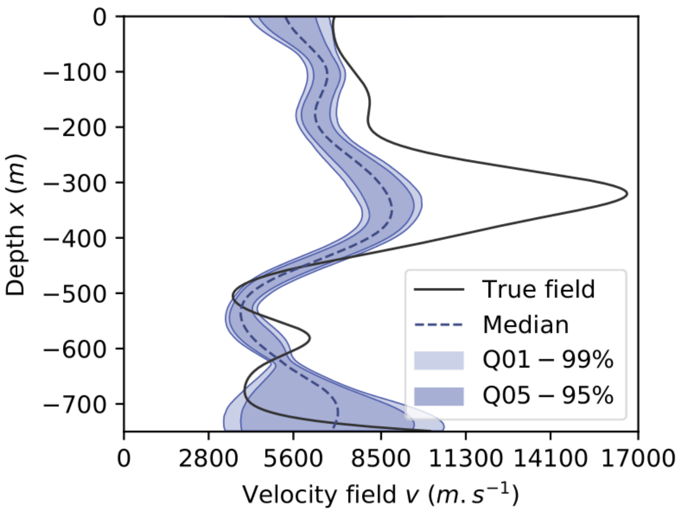
Partial tensorisation: $\sum n_{o,i} \leq n_o$

2. Initial surrogate construction: Illustration

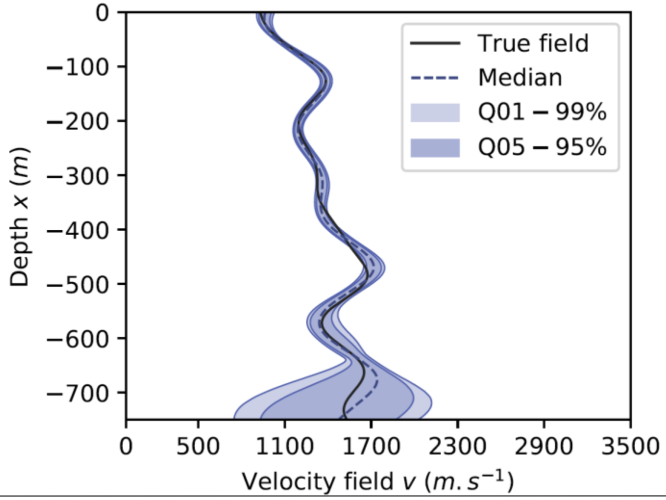
$$\boldsymbol{P(\eta|d_{mes}) \propto L(d_{mes}|\eta)P(\eta)}$$
with $\boldsymbol{d_{mes}} = F(\boldsymbol{\eta}) + \varepsilon \simeq \tilde{F}(\boldsymbol{\eta}) + \varepsilon$



Training set and true coordinates (η_1, η_2) *(red stars)*



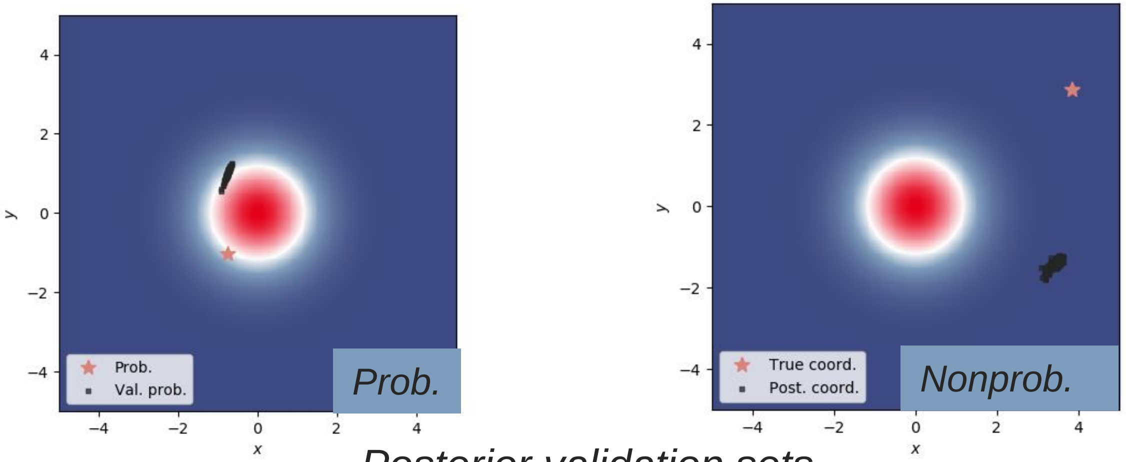
Posterior field distributions



2. Initial surrogate construction: Illustration

- Initial training set $\{\boldsymbol{\eta}^{(n)}\}_{1\leq n\leq N} \sim P(\boldsymbol{\eta})$
- Initial surrogate $\tilde{F}(\boldsymbol{\eta})$
- Monte-Carlo sampling $P(\boldsymbol{\eta}|\boldsymbol{d}_{mes}) \propto \tilde{L}(\boldsymbol{d}_{mes}|\boldsymbol{\eta})P(\boldsymbol{\eta})$
- Validation

$$RMSRE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\frac{\tilde{L}(\boldsymbol{d}_{mes}|\boldsymbol{\eta}^{(i)}) - L(\boldsymbol{d}_{mes}|\boldsymbol{\eta}^{(i)})}{L(\boldsymbol{d}_{mes}|\boldsymbol{\eta}^{(i)})}\right)^2}$$



Posterior validation sets

Case	Valid. set	$n_o = 1$	$n_o = 2$	$n_o = 3$
Probable	Prior Posterior	47.4 97.9	10.3 84.6	2.57 3.80
Nonprobable	Prior Posterior	26.2 99.8	7.00 81.0	2.11 95.5

Surrogate error RMSRE (%)

This construction does not ensure that the error on the posterior subspace is bounded.

Objective: to improve the surrogate by *minimizing the error on the final quantity of interest* $E_{P(\cdot|\boldsymbol{d}_{mes})}\left(\|\boldsymbol{L}(\boldsymbol{\eta}) - \tilde{\boldsymbol{L}}(\boldsymbol{\eta})\|^2\right)$

3. Adaptive construction: Training set

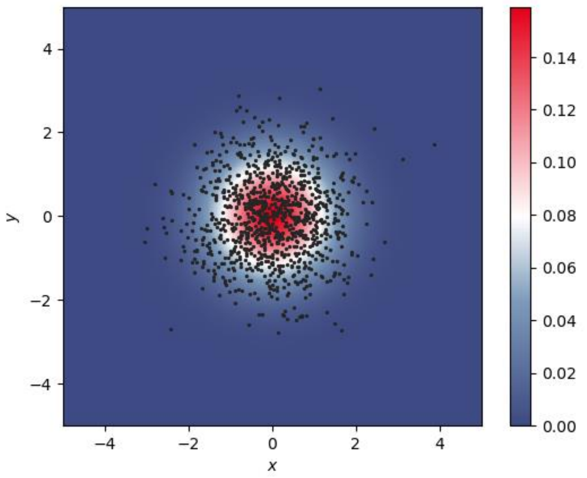
Adaptive workflow

- Initial surrogate $\tilde{L}^{(0)}$ with $X^{(0)} = \{\boldsymbol{\eta}^{(n)}\}_{1\leq n\leq N} \sim P(\boldsymbol{\eta})$
- **While** **general convergence** not achieved
  - MCMC with $\tilde{P}^{(i)}(\cdot\,|\boldsymbol{d}_{mes})$
  - $i \leftarrow i+1$
  - Update training set and surrogate:
    $X^{(i)} = X^{(i-1)} \backslash X^{(i-1)}_{1\leq k\leq n_r} \cup \{\boldsymbol{\eta}^{(n)}\}_{1\leq n\leq n_a} \sim \tilde{P}^{(i-1)}(\cdot\,|\boldsymbol{d}_{mes})$

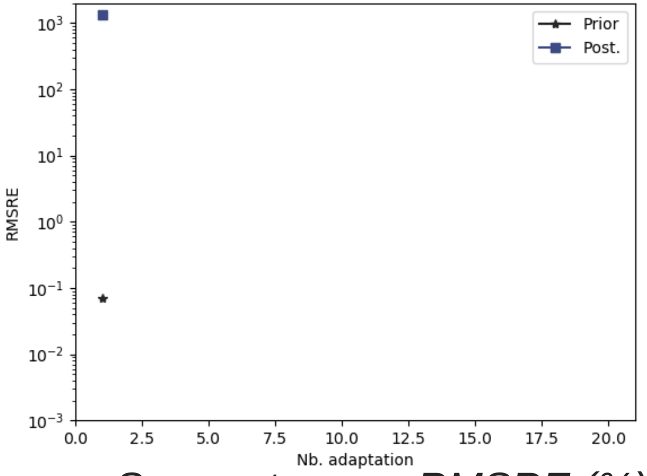
General convergence: surrogate quality

STOP if $R^2\left(\tilde{L}^{(i-1)}, L|X^{(i)}\right) > R^2_{target}$

$$with\ \ R^2(u,v|X) = 1 - \frac{\sum_{k=1}^{n}(u(X_k)-v(X_k))^2}{\sum_{k=1}^{n}\left(v(X_k) - \frac{1}{n}\sum_{j=1}^{n} v(X_j)\right)^2}$$



Training set on prior pdf (initialization)



Surrogate error RMSRE (%)

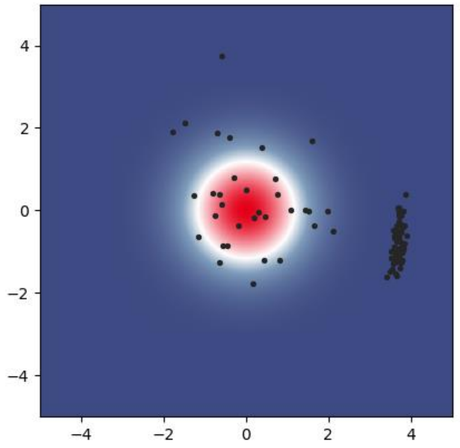
Lucor D., Le Maître O., *Cardiovascular Modeling With Adapted Parametric Inference*, ESAIM Proceedings, 2018

3. Adaptive construction: Training set

Adaptive workflow

- Initial surrogate $\tilde{L}^{(0)}$ with $X^{(0)} = \{\boldsymbol{\eta}^{(n)}\}_{1\le n\le N} \sim P(\boldsymbol{\eta})$
- **While general convergence** not achieved
  - MCMC with $\tilde{P}^{(i)}(\cdot\,|\boldsymbol{d}_{mes})$
  - $i \leftarrow i+1$
  - Update training set and surrogate:

$$X^{(i)} = X^{(i-1)} \backslash X^{(i-1)}_{1\le k\le n_r} \cup \{\boldsymbol{\eta}^{(n)}\}_{1\le n\le n_a} \sim \tilde{P}^{(i-1)}(\cdot\,|\boldsymbol{d}_{mes})$$

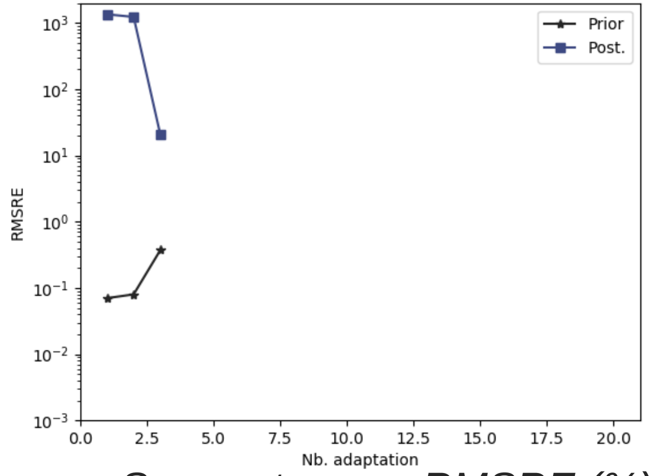


Training set (adaptation nb. 1)

General convergence: surrogate quality

STOP if $R^2\left(\tilde{L}^{(i-1)}, L|X^{(i)}\right) > R^2_{target}$

$$with\ \ R^2(u,v|X) = 1 - \frac{\sum_{k=1}^{n}\left(u(X_k) - v(X_k)\right)^2}{\sum_{k=1}^{n}\left(v(X_k) - \frac{1}{n}\sum_{j=1}^{n} v(X_j)\right)^2}$$



Surrogate error RMSRE (%)

Lucor D., Le Maître O., *Cardiovascular Modeling With Adapted Parametric Inference*, ESAIM Proceedings, 2018

3. Adaptive construction: Training set

Adaptive workflow

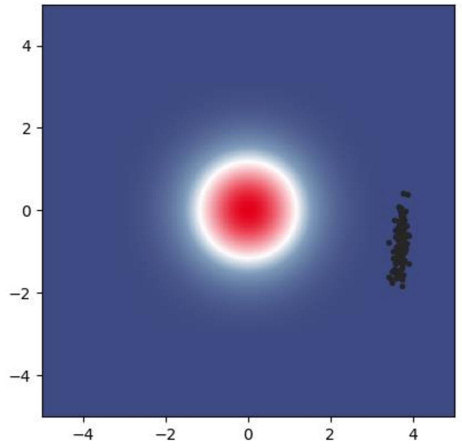
- Initial surrogate $\tilde{L}^{(0)}$ with $X^{(0)} = \{\boldsymbol{\eta}^{(n)}\}_{1 \le n \le N} \sim P(\boldsymbol{\eta})$
- **While general convergence not achieved**
  - MCMC with $\tilde{P}^{(i)}(\cdot\,|\boldsymbol{d}_{mes})$
  - $i \leftarrow i + 1$
  - Update training set and surrogate:

    $X^{(i)} = X^{(i-1)} \backslash X^{(i-1)}_{1 \le k \le n_r} \cup \{\boldsymbol{\eta}^{(n)}\}_{1 \le n \le n_a} \sim \tilde{P}^{(i-1)}(\cdot\,|\boldsymbol{d}_{mes})$

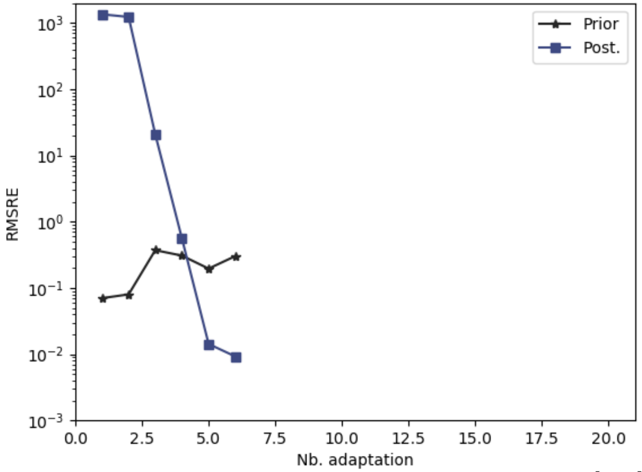
General convergence: surrogate quality

STOP if $R^2\left(\tilde{L}^{(i-1)}, L | X^{(i)}\right) > R^2_{target}$

$$with\ \ R^2(u, v|X) = 1 - \frac{\sum_{k=1}^{n}(u(X_k) - v(X_k))^2}{\sum_{k=1}^{n}\left(v(X_k) - \frac{1}{n}\sum_{j=1}^{n} v(X_j)\right)^2}$$



Training set (adaptation nb. 6)



Surrogate error RMSRE (%)

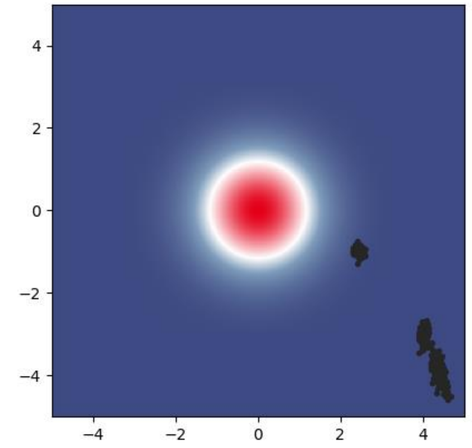
Lucor D., Le Maître O., *Cardiovascular Modeling With Adapted Parametric Inference*, ESAIM Proceedings, 2018

3. Adaptive construction: Training set

Adaptive workflow

- Initial surrogate $\tilde{L}^{(0)}$ with $X^{(0)} = \{\boldsymbol{\eta}^{(n)}\}_{1\leq n\leq N} \sim P(\boldsymbol{\eta})$
- **While general convergence not achieved**
  - MCMC with $\tilde{P}^{(i)}(\cdot\,|\boldsymbol{d}_{mes})$
  - $i \leftarrow i+1$
  - Update training set and surrogate:

$$X^{(i)} = X^{(i-1)} \backslash X^{(i-1)}_{1\leq k\leq n_r} \cup \{\boldsymbol{\eta}^{(n)}\}_{1\leq n\leq n_a} \sim \tilde{P}^{(i-1)}(\cdot\,|\boldsymbol{d}_{mes})$$

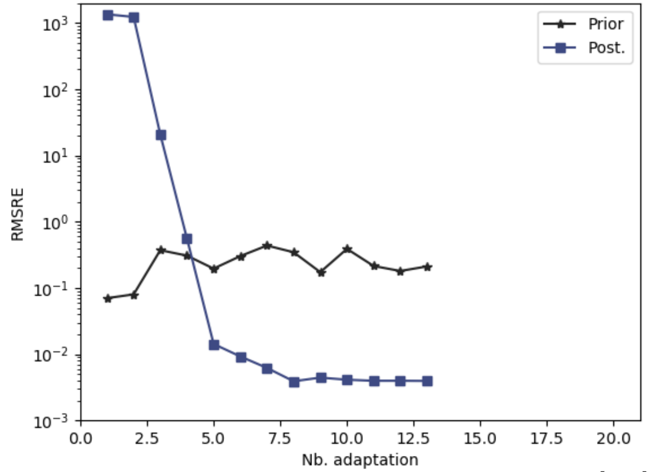


Training set (adaptation nb. 13)

General convergence: surrogate quality

STOP if $R^2\left(\tilde{L}^{(i-1)}, L|X^{(i)}\right) > R^2_{target}$

$$with\ \ R^2(u,v|X) = 1 - \frac{\sum_{k=1}^{n}(u(X_k) - v(X_k))^2}{\sum_{k=1}^{n}\left(v(X_k) - \frac{1}{n}\sum_{j=1}^{n} v(X_j)\right)^2}$$



Surrogate error RMSRE (%)

Lucor D., Le Maître O., *Cardiovascular Modeling With Adapted Parametric Inference*, ESAIM Proceedings, 2018

3. Adaptive construction: Training set

Adaptive workflow

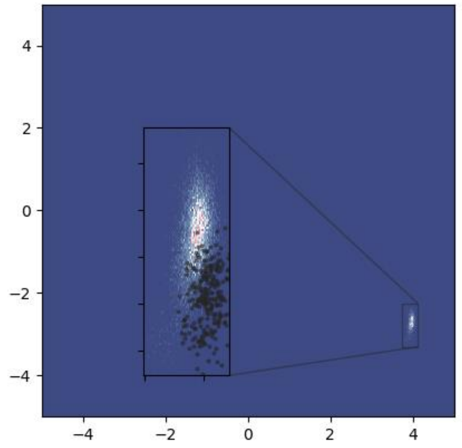
- Initial surrogate $\tilde{L}^{(0)}$ with $X^{(0)} = \{\boldsymbol{\eta}^{(n)}\}_{1\leq n\leq N} \sim P(\boldsymbol{\eta})$
- **While** general convergence not achieved
  - MCMC with $\tilde{P}^{(i)}(\cdot\,|\boldsymbol{d}_{mes})$
  - $i \leftarrow i+1$
  - Update training set and surrogate:

$$X^{(i)} = X^{(i-1)} \backslash X^{(i-1)}_{1\leq k\leq n_r} \cup \{\boldsymbol{\eta}^{(n)}\}_{1\leq n\leq n_a} \sim \tilde{P}^{(i-1)}(\cdot\,|\boldsymbol{d}_{mes})$$

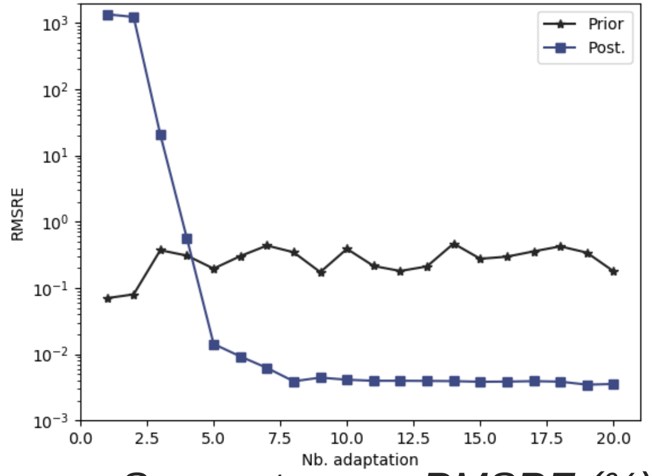
General convergence: surrogate quality

STOP if $R^2\left(\tilde{L}^{(i-1)}, L|X^{(i)}\right) > R^2_{target}$

$$with\ \ R^2(u,v|X) = 1 - \frac{\sum_{k=1}^{n}(u(X_k) - v(X_k))^2}{\sum_{k=1}^{n}\left(v(X_k) - \frac{1}{n}\sum_{j=1}^{n} v(X_j)\right)^2}$$



Training set on posterior pdf (final adaptation)



Surrogate error RMSRE (%)

Lucor D., Le Maître O., *Cardiovascular Modeling With Adapted Parametric Inference*, ESAIM Proceedings, 2018

3. Adaptive construction: Training set

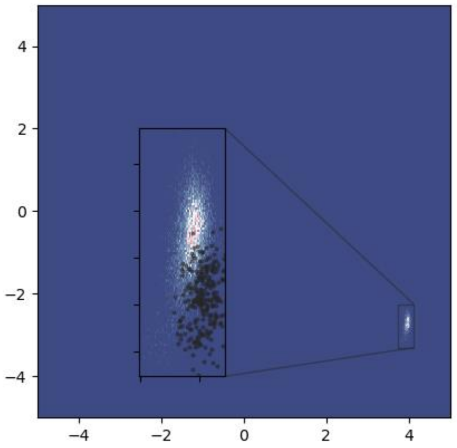
Adaptive workflow

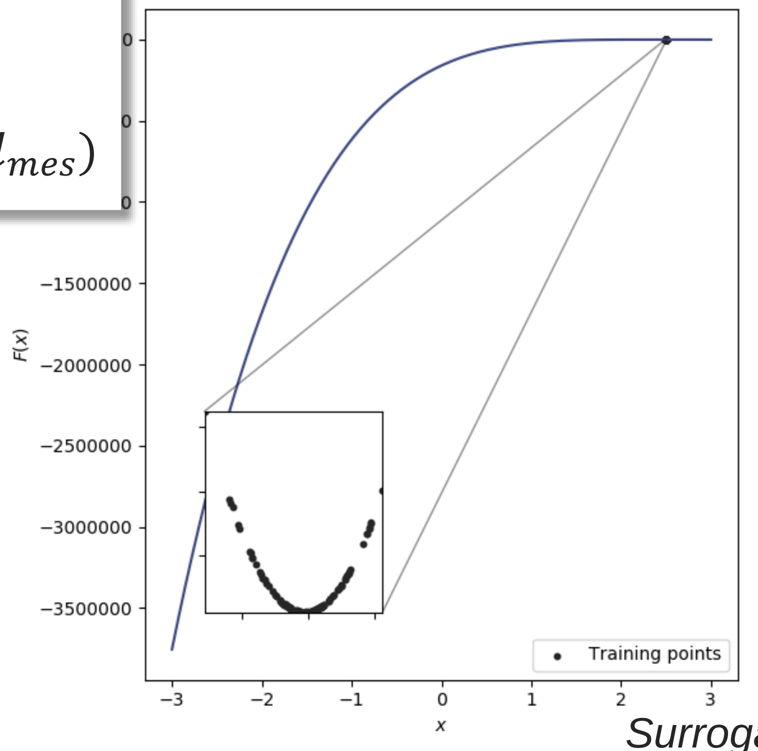
- Initial surrogate $\tilde{L}^{(0)}$ with $X^{(0)} = \{\boldsymbol{\eta}^{(n)}\}_{1\leq n\leq N} \sim P(\boldsymbol{\eta})$
- **While** general convergence not achieved
  - MCMC with $\tilde{P}^{(i)}(\cdot\,|\boldsymbol{d}_{mes})$
  - $i \leftarrow i+1$
  - Update training set and surrogate:

$$X^{(i)} = X^{(i-1)} \backslash X^{(i-1)}_{1\leq k\leq n_r} \cup \{\boldsymbol{\eta}^{(n)}\}_{1\leq n\leq n_a} \sim \tilde{P}^{(i-1)}(\cdot\,|\boldsymbol{d}_{mes})$$

All the training points **are concentrated on a small subspace** of the prior density → surrogate not pertinent

Training set on posterior pdf (final adaptation)





Surrogate approximation

Lucor D., Le Maître O., *Cardiovascular Modeling With Adapted Parametric Inference*, ESAIM Proceedings, 2018

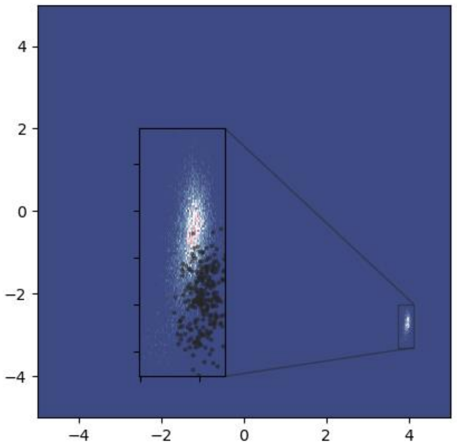
3. Adaptive construction: Training set

Adaptive workflow

- Initial surrogate $\tilde{L}^{(0)}$ with $X^{(0)} = \{\boldsymbol{\eta}^{(n)}\}_{1 \leq n \leq N} \sim P(\boldsymbol{\eta})$
- **While** general convergence not achieved
  - MCMC with $\tilde{P}^{(i)}(\cdot\,|\boldsymbol{d}_{mes})$
  - $i \leftarrow i + 1$
  - Update training set and surrogate:

$$X^{(i)} = X^{(i-1)} \backslash X^{(i-1)}_{1 \leq k \leq n_r} \cup \{\boldsymbol{\eta}^{(n)}\}_{1 \leq n \leq n_a} \sim \tilde{P}^{(i-1)}(\cdot\,|\boldsymbol{d}_{mes})$$

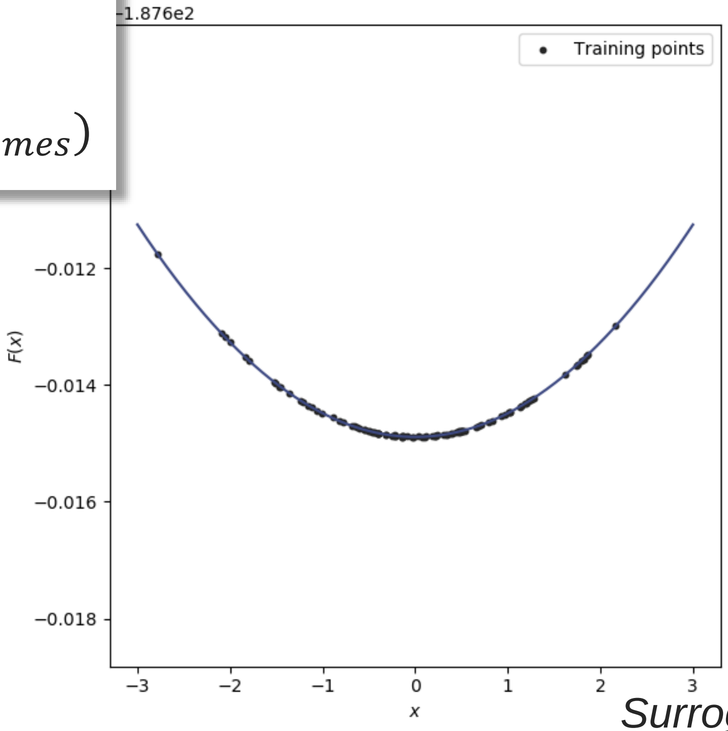
Training set on posterior pdf (final adaptation)



All the training points **are concentrated on a small subspace** of the prior density → surrogate not pertinent

Rescaling: using mean $\overline{\boldsymbol{\eta}}$ and $\boldsymbol{C}$ the mean and covariance matrix of $X^{(i)}$

$$\tilde{F}^{(i)}(\boldsymbol{\eta}) = \sum_{a \in A} f_a \psi_a\left(\boldsymbol{C}^{-\frac{1}{2}}(\boldsymbol{\eta} - \overline{\boldsymbol{\eta}})\right)$$

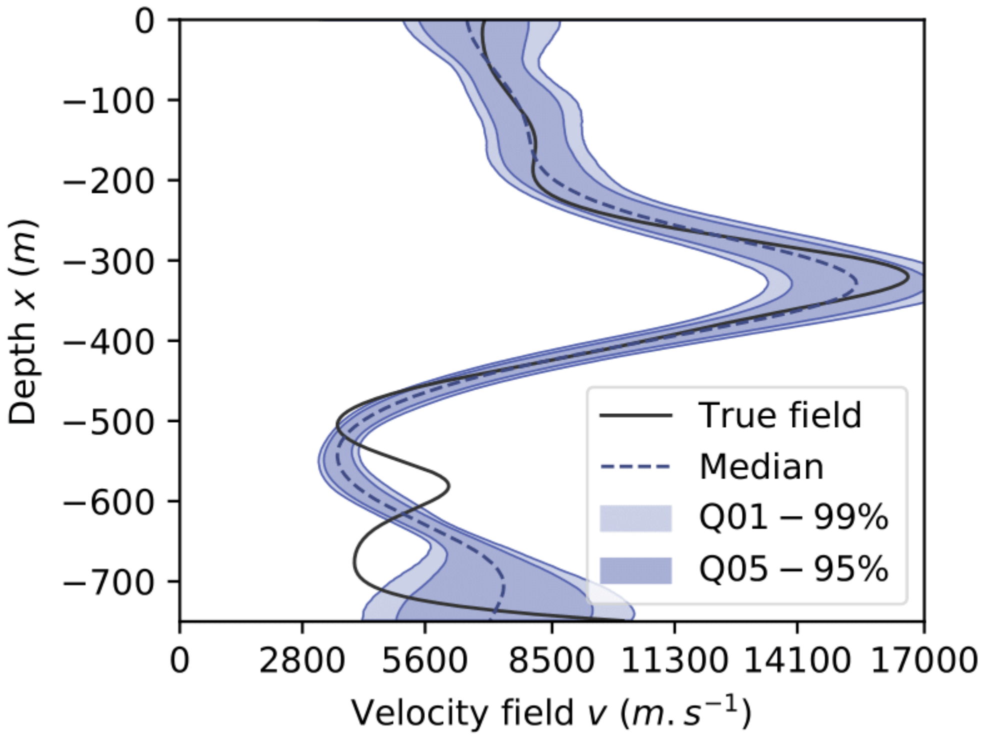


Surrogate approximation

3. Adaptive construction: Polynomial order

Adaptive workflow

- Initial surrogate $\tilde{L}^{(0)}$with $X^{(0)} = \{\boldsymbol{\eta}^{(n)}\}_{1 \leq n \leq N} \sim P(\boldsymbol{\eta})$
- **While** general convergence not achieved
  - MCMC with $\tilde{P}^{(i)}(\cdot\,|\boldsymbol{d}_{mes})$
  - $i \leftarrow i + 1$
  - Update training set and surrogate:
    $X^{(i)} = X^{(i-1)} \backslash X^{(i-1)}_{1 \leq k \leq n_r} \cup \{\boldsymbol{\eta}^{(n)}\}_{1 \leq n \leq n_a} \sim \tilde{P}^{(i-1)}(\cdot\,|\boldsymbol{d}_{mes})$
  - **If** $|X^{(i)}| > 5 \times N_{PC}(n_o + 1)$**:** $n_o \leftarrow n_o + 1$



Posterior field distribution (nonprobable case)

3. Adaptive construction: State-of-the-art

Review (ED enrichment, sparse constructions)

→ **Teixeira R. et al**, *Adaptive Approaches in Metamodel-based Reliability Analysis: A Review*, Structural Safety, 2021

Adaptive training sets

→ **Li, J. and Marzouk Y.**, *Adaptive Construction of Surrogates for the Bayesian Solution of Inverse Problems*, SIAM Journal on Scientific Computing, 2014
→ **Fu S. et al.**, *An Adaptive Kriging Method for Solving Nonlinear Inverse Statistical Problems,* Environmetrics, 2017
→ **Lucor D., Le Maître O.**, *Cardiovascular Modeling With Adapted Parametric Inference*, ESAIM Proceedings, 2018

Adaptive polynomial basis (PCE)

→ **Blatman G. and Sudret B.**, *An Adaptive Algorithm to build up Sparse Polynomial Chaos Expansions for Stochastic Finite Element Analysis*, Probabilistic Engineering Mechanics, 2010
→ **Blatman G. and Sudret B.**, *Adaptive Sparse Polynomial Chaos Expansion Based on Least Angle Regression*, JCP, 2011
→ **Poëtte G., Lucor D.**, *Non Intrusive Iterative Stochastic Spectral Representation with Application to Compressible Gas Dynamics*, JCP, 2012
→ **Zhou Y. and al.**, *Adaboost-based Ensemble of Polynomial Chaos Expansion with Adaptive Sampling*, CMAME, 2022

Dimension reduction with surrogate models

→ **Lieberman C. and al.**, *Parameter and State Model Reduction for Large-Scale Statistical Inverse Problems*, SIAM Journal on Scientific Computing, 2010
→ **Vohra M. and al.**, *Fast Surrogate Modeling using Dimensionality Reduction in Model Inputs and Field Output: Application to Additive Manufacturing*, Reliability Engineering & System Safety, 2020

Conclusion

- Surrogate models require a training set
- The prior space can be very different from the posterior space
- Adaptive construction allow to improve surrogate on the space of interest while mitigating costs

Thank you !

nadege.polette@cea.fr

Keywords: inverse problem, (hierarchical) Bayesian inference, surrogate models (polynomial chaos), Markov Chain Monte Carlo, Dimension reduction, Karhunen-Loève decomposition

Bibliography

- **N. Polette, O. Le Maître, P. Sochala, A. Gesret,** *Change of Measure for Bayesian Field Inversion with Hierarchical Hyperparameters Sampling*, In Rev
- **A. Tarantola**, *Inverse Problem Theory and Methods for Model Parameter Estimation*, SIAM 2005
- **D. Sivia, J. Skilling**, *Data Analysis: A Bayesian Tutorial*, Oxford, sc. Pub. 2006
- **A. Doucet et al.**, *Sequential Monte Carlo Methods in Practice*, ISS 2013
- **Lucor D., Le Maître O.**, *Cardiovascular Modeling With Adapted Parametric Inference*, ESAIM Proceedings, 2018