Mitigating Overconfidence in Bayesian Field Inversion thanks to Hyperparameters Sampling

UQSAY #77

PhD student: Nadège Polette[1,2],
Supervisors:
Alexandrine Gesret[1], Pierre Sochala[2], Olivier Le Maître[3]

[1] Mines Paris PSL, Geosciences center, Fontainebleau, France
[2] CEA, DAM, DIF, F-91297 Arpajon, France
[3] CNRS, CMAP, IPP École Polytechnique, Palaiseau, France

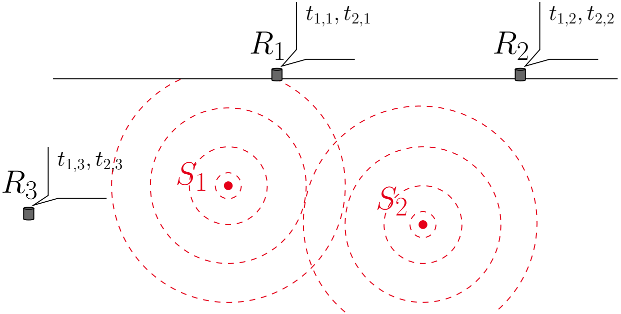
Table of contents

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

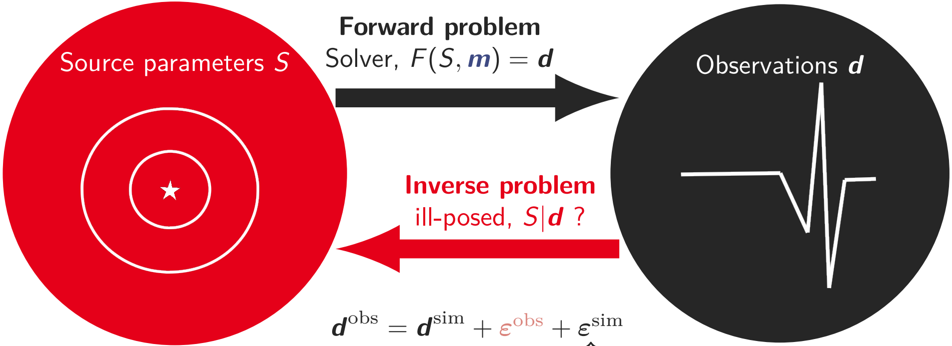


$$F(S) = \boldsymbol{d}$$

F: forward model
S: source parameters
$\boldsymbol{d}$: data

Objective: retrieve S from $\boldsymbol{d}$
- fast
- accurately
- with uncertainties

1. Context: Inverse problem



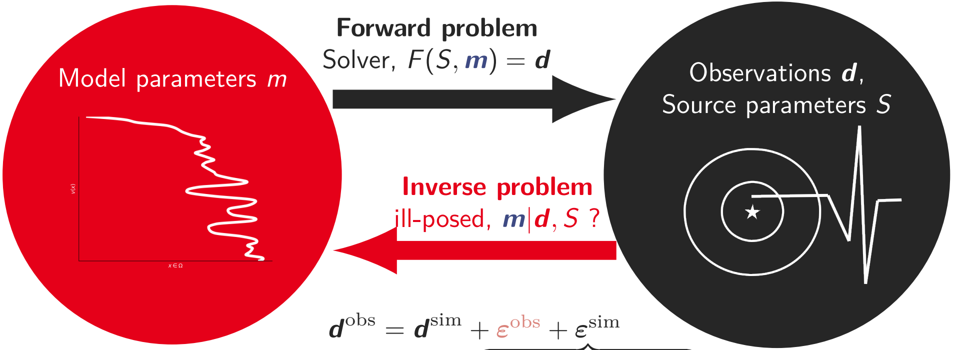
$$\boldsymbol{d}^{\text{obs}} = \boldsymbol{d}^{\text{sim}} + \varepsilon^{\text{obs}} + \underbrace{\varepsilon^{\text{sim}}}_{\varepsilon^{\text{num}} + \varepsilon^{\text{mod}} + \varepsilon^{\text{par}}}$$

Uncertainty sources: **observations**, physical model, **model parameters**,. . .
Objective: improve **uncertainty quantification** of model parameters

---

Tarantola, *SIAM*, 2005; **Noble et al.**, *GJI*, 2014

1. Context: Inverse problem



$$\boldsymbol{d}^{\text{obs}} = \boldsymbol{d}^{\text{sim}} + \varepsilon^{\text{obs}} + \underbrace{\varepsilon^{\text{sim}}}_{\varepsilon^{\text{num}} + \varepsilon^{\text{mod}} + \varepsilon^{\text{par}}}$$

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 $\pi_{\text{post}}(m|\boldsymbol{d}^{\text{obs}})$

Tarantola, *SIAM*, 2005; **Noble et al.**, *GJI*, 2014

1. Context: Bayesian inference and Markov chain Monte Carlo

Proposition $m^{(i)}$

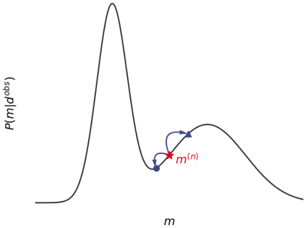
Forward model
Solver $F(m; S) = \boldsymbol{d}$

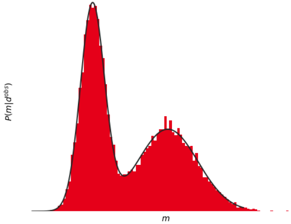
Comparison
$\boldsymbol{d}$ versus $\boldsymbol{d}^{\text{obs}}$ (data)

Bayes' rule: $\pi_{\text{post}}(m|\boldsymbol{d}^{\text{obs}}) \propto \overbrace{\mathcal{L}(\boldsymbol{d}^{\text{obs}}|m)}\, \pi_{\text{prior}}(m)$

Accept/Reject criterion

New iteration
$i \leftarrow i+1$





Sivia and Skilling, *Oxford*, 2006; **Doucet et al.**, *Springer NY*, 2013

1. Context: Bayesian inference and Markov chain Monte Carlo

Proposition $m^{(i)}$

m infinite dimensional

$\rightarrow$ dimension reduction

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

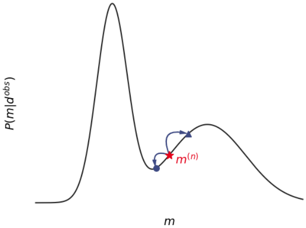
Forward model
Solver $F(m; S) = \boldsymbol{d}$

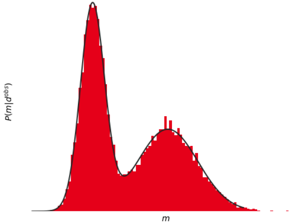
Comparison
$\boldsymbol{d}$ versus $\boldsymbol{d}^{\text{obs}}$ (data)

Bayes' rule: $\pi_{\text{post}}(m|\boldsymbol{d}^{\text{obs}}) \propto \overbrace{\mathcal{L}(\boldsymbol{d}^{\text{obs}}|m)}\,\pi_{\text{prior}}(m)$

Accept/Reject criterion

New iteration
$i \leftarrow i + 1$





Sivia and Skilling, *Oxford*, 2006; **Doucet et al.**, *Springer NY*, 2013

1. Context: Bayesian inference and Markov chain Monte Carlo

Proposition $m^{(i)}$

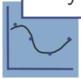
m infinite dimensional

→ dimension reduction

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

Surrogate model

Polynomial chaos $\widetilde{F}(m) = \sum_{a \in \mathcal{A}} f_a \Psi_a(m) = \boldsymbol{d}$



Comparison

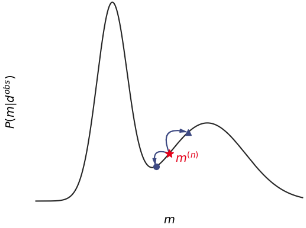
$\boldsymbol{d}$ versus $\boldsymbol{d}^{\text{obs}}$ (data)

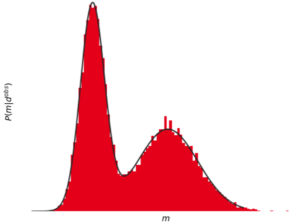
Bayes' rule: $\pi_{\text{post}}(m|\boldsymbol{d}^{\text{obs}}) \propto \underbrace{\mathcal{L}(\boldsymbol{d}^{\text{obs}}|m)}\, \pi_{\text{prior}}(m)$

Accept/Reject criterion

New iteration

$i \leftarrow i + 1$

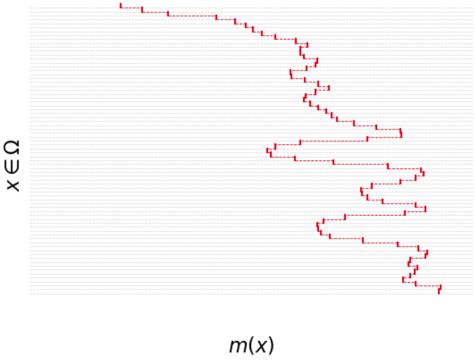




Sivia and Skilling, *Oxford*, 2006; **Doucet et al.**, *Springer NY*, 2013

Table of contents

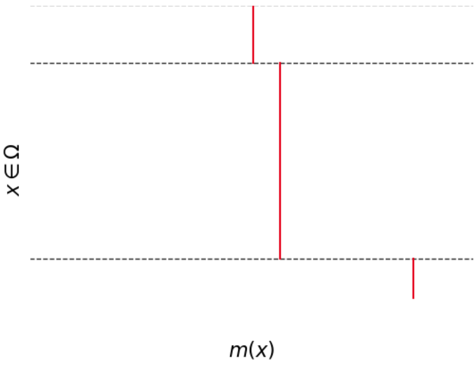
2. Field parametrization: Spatial mesh



$\forall x \in [x_i, x_{i+1}],\ m(x) = m_i$
Parameters: $\{m_i\}_{1 \leqslant i \leqslant N_{\text{meshes}}}$

- ✗ "small" dimension
- ✓ fixed dimension
- ✓ allowing for various shapes

2. Field parametrization: Layered velocity model

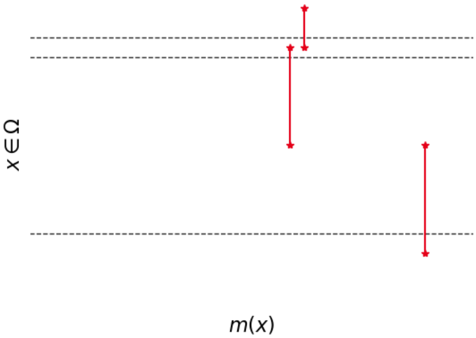


$\forall x \in [x_i, x_{i+1}],\ m(x) = m_i$
Parameters: $\{m_i, x_i\}_{1 \leqslant i \leqslant N_{\text{layers}}}$

- ☑ "small" dimension
- ☑ fixed dimension
- ✗ allowing for various shapes

Sochala et al., *GEM*, 2021

2. Field parametrization: Voronoï tesselation



$\forall x \in V(x_i),\ m(x) = m_i$

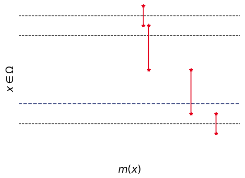
Parameters:

$\{m_i, x_i\}_{1 \leqslant i \leqslant N_{\text{cells}}} \cup N_{\text{cells}}$

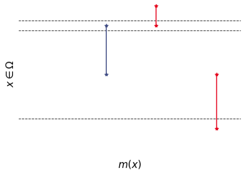
- ☑ "small" dimension
- ✗ fixed dimension
- ☑ allowing for various shapes

Bodin et al., *GJI*, 2012; **Piana Agostinetti et al.**, *GJI*, 2015; **Belhadj et al.**, *Inverse Problems*, 2018

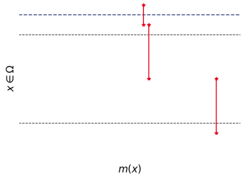
2. Field parametrization: Voronoï tesselation



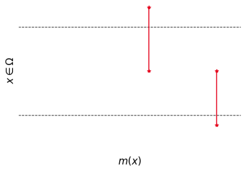
(a) Add a layer



(b) Change a value



(c) Change a depth



(d) Remove a layer

$\forall x \in V(x_i),\ m(x) = m_i$

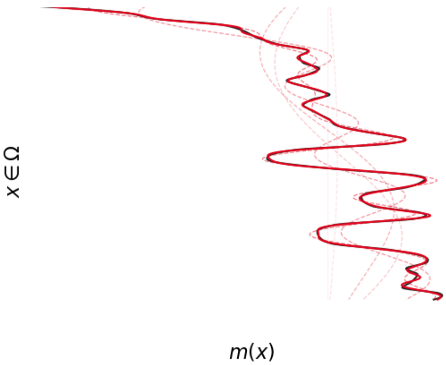
Parameters:

$\{m_i, x_i\}_{1 \leqslant i \leqslant N_{\text{cells}}} \cup N_{\text{cells}}$

- ☑ "small" dimension
- ✗ fixed dimension
- ☑ allowing for various shapes

Bodin et al., *GJI*, 2012; **Piana Agostinetti et al.**, *GJI*, 2015; **Belhadj et al.**, *Inverse Problems*, 2018

2. Field parametrization: Modal representation



$$\forall x,\ m(x) = \sum_{i=1}^{r} U_i(x) w_i$$

Parameters: $\{w_i\}_{1 \leqslant i \leqslant r}$

- ☑ "small" dimension
- ☑ fixed dimension
- ☑ allowing for various shapes

Marzouk and Najm, *JCP*, 2009

Assuming m is the realization of a **random process** with autocovariance function k, $m \sim \mathrm{GP}(0, k)$

$$m(x) = \sum_{i=1}^{r} \sqrt{\lambda_i} u_i(x) \eta_i$$

$(\lambda_i, u_i)_{1 \leqslant i \leqslant r}$ **eigenelements of** k: $\langle k(x, \cdot),\ u_i \rangle = \int_\Omega k(x,y) u_i(y) dy = \lambda_i u_i(x)$

The decomposition is **bi-orthonormal**:

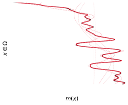
- $\langle u_i,\ u_j \rangle = \delta_{i,j}$
- $\mathbb{E}(\eta_i) = 0$ and $\mathbb{E}(\eta_i \eta_j) = \delta_{i,j}$

In the case of a Gaussian process, $\boldsymbol{\eta} \sim \mathcal{N}(0, \mathrm{I}_r)$

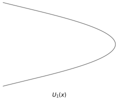
Karhunen, *Ann. Acad. Sci. Fenn.*, 1946; **Loève**, *Springer NY*, 1977; **Marzouk and Najm**, *JCP*, 2009

Assuming m is the realization of a **random process** with autocovariance function k,
$m \sim \mathrm{GP}(0, k)$

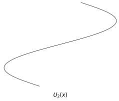
$$m(x) = \sum_{i=1}^{r} \sqrt{\lambda_i} u_i(x) \eta_i$$



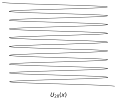
$= \eta_1$



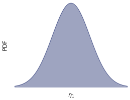
$+ \eta_2$

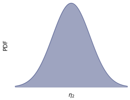


$+ \cdots + \eta_r$



$\boldsymbol{\eta} \sim \mathcal{N}(0, \mathrm{I}_r)$





Karhunen, *Ann. Acad. Sci. Fenn.*, 1946; **Loève**, *Springer NY*, 1977; **Marzouk and Najm**, *JCP*, 2009

2. Field parametrization: Karhunen-loève decomposition

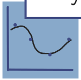
Proposition $\eta^{(i)}$

$m \sim \mathrm{GP}(0, k)$

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

Surrogate model

Polynomial chaos $\widetilde{F}(\boldsymbol{\eta}) = \sum_{a \in \mathcal{A}} f_a \Psi_a(\boldsymbol{\eta}) = \boldsymbol{d}$



Comparison

$\boldsymbol{d}$ versus $\boldsymbol{d}^{\mathrm{obs}}$ (data)

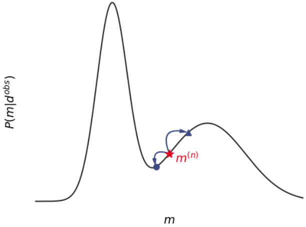
Bayes' rule: $\pi_{\mathrm{post}}(\boldsymbol{\eta}|\boldsymbol{d}^{\mathrm{obs}}) \propto \mathcal{L}(\boldsymbol{d}^{\mathrm{obs}}|\boldsymbol{\eta})\, \pi_{\mathrm{prior}}(\boldsymbol{\eta})$

$\boldsymbol{\eta} \sim \mathcal{N}(0, \mathrm{I}_r)$

New iteration

$i \leftarrow i + 1$

Accept/Reject criterion



Problem: how to choose k ?

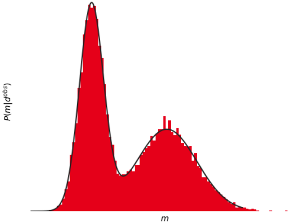
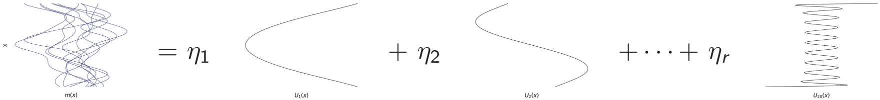


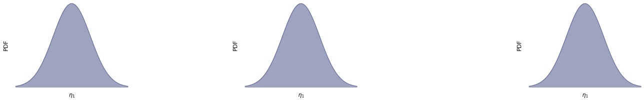
Table of contents

- **Squared exponential** autocovariance function: $k(x,y) = A\exp\left(-\|x-y\|^2/(2l^2)\right)$
- **Hyperparameters** $\boldsymbol{q} = \{A, l\} \Rightarrow m(x) = \sum_{i=1}^{r} \sqrt{\lambda_i(\boldsymbol{q})}\, u_i(x, \boldsymbol{q})\, \eta_i$

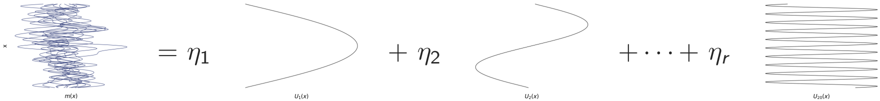
Large l



$\boldsymbol{\eta} \sim \mathcal{N}(0, \mathrm{I}_r)$



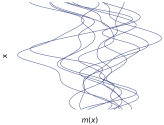
Small l



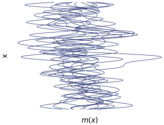
3. Change of measure: Objective

- **Squared exponential** autocovariance function: $k(x,y) = A\exp\left(-\|x-y\|^2/(2l^2)\right)$
- **Hyperparameters** $\boldsymbol{q} = \{A, l\} \Rightarrow m(x) = \sum_{i=1}^{r} \sqrt{\lambda_i(\boldsymbol{q})}\, u_i(x,\boldsymbol{q})\eta_i$

Field realizations



Large l



Small l

- Hyperparameters are determined *a priori*
  $\rightarrow$ expert judgement, MSE, LOOCV...
  Overconfidence risk
- Hyperparameters are inferred during the procedure
  $\rightarrow$ Bayes' rule:
  $\pi_{\text{post}}(\boldsymbol{\eta}, \boldsymbol{q} | \boldsymbol{d}^{\text{obs}}) \propto \mathcal{L}(\boldsymbol{d}^{\text{obs}} | \boldsymbol{\eta}, \boldsymbol{q})\pi_{\text{prior}}(\boldsymbol{\eta}, \boldsymbol{q})$
  Expensive

Objective: develop a cheap method to take into account hyperparameters

Rasmussen and Williams, *The MIT Press*, 2005; **Pion and Vazquez**, *LOD 2024*, 2024
Tagade and Choi, *IPSE*, 2014; **Sraj et al.**, *CMAME*, 2016

Reference kernel

$$\overline{k} = \int_{\mathbb{H}} k(\cdot,\cdot,\boldsymbol{q})\pi_{\mathrm{prior}}(\boldsymbol{q})d\boldsymbol{q}$$

Reference basis $(\overline{\lambda}_i, \overline{u}_i)_{1\leqslant i\leqslant r}$ $\rightarrow$ eigenelements of $\overline{k}$: $\int_{\Omega} \overline{k}(x,y)\overline{u}_i(y)dy = \overline{\lambda}_i\overline{u}_i(x)$

Field decomposition

$$m(x) = \sum_{i=1}^{r} \sqrt{\overline{\lambda}_i}\,\overline{u}_i(x)\xi_i$$

Hierarchical Bayes formulation

$$\pi_{\mathrm{post}}(\boldsymbol{\xi},\boldsymbol{q}|\boldsymbol{d}^{\mathrm{obs}}) \propto \mathcal{L}(\boldsymbol{d}^{\mathrm{obs}}|\boldsymbol{\xi})\pi_{\mathrm{prior}}(\boldsymbol{\xi},\boldsymbol{q}) = \mathcal{L}(\boldsymbol{d}^{\mathrm{obs}}|\boldsymbol{\xi})\pi_{\mathrm{prior}}(\boldsymbol{\xi}|\boldsymbol{q})\pi_{\mathrm{prior}}(\boldsymbol{q})$$

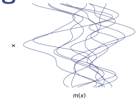
$\Rightarrow$ The $\boldsymbol{q}$-dependency is transferred to the prior law of the coordinates $\boldsymbol{\xi}$

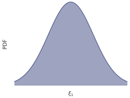
Sraj et al., *CMAME*, 2016; **Polette et al.**, *in Rev.*, 2024

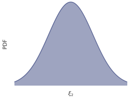
3. Change of measure: Reference basis

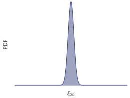
The field decomposition writes $m(x) = \sum_{i=1}^{r} \sqrt{\lambda_i}\,\overline{u}_i(x)\xi_i,\ \boldsymbol{\xi} \sim \pi_{\text{prior}}(\cdot|\boldsymbol{q})$

Large /

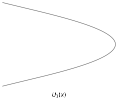




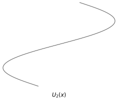




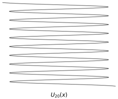
Reference decomposition: ξ_1



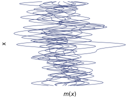
$+\ \xi_2$

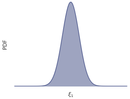


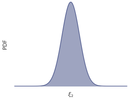
$+ \cdots + \xi_r$

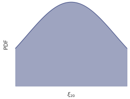


Small /









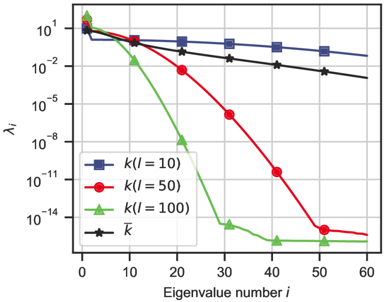
How to define $\pi_{\text{prior}}(\boldsymbol{\xi}|\boldsymbol{q})$?

3. Change of measure: Formulation

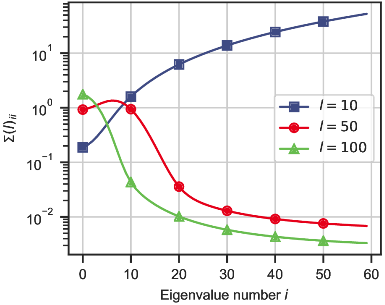
Prior law of the reference coordinates according to the hyperparameters

$$m(x) = \sum_{i=1}^{r} \sqrt{\overline{\lambda}_i}\,\overline{u}_i(x)\xi_i, \quad \boldsymbol{\xi} \sim \mathcal{N}\left(0, \Sigma(\boldsymbol{q})\right) \text{ with } \Sigma(\boldsymbol{q})_{ij} = \left(\overline{\lambda}_i \overline{\lambda}_j\right)^{-1/2} \langle\langle k(\cdot,\cdot,\boldsymbol{q}),\ \overline{u}_i\rangle,\ \overline{u}_j\rangle$$

$\Sigma(\boldsymbol{q})$ is the double projection of the $\boldsymbol{q}$-dependent kernel on the reference basis



(a) Eigenvalues decay according to the basis



(b) Variance of ξ according to l

Polette et al., *in Rev.*, 2024

3. Change of measure: Formulation

Objective: Find $\pi_{\text{prior}}(\boldsymbol{\xi}|\boldsymbol{q})$ such that

$$\sum_{i=1}^{r} \sqrt{\overline{\lambda}_i}\,\overline{u}_i \xi_i \simeq \sum_{j=1}^{+\infty} \sqrt{\lambda_j(\boldsymbol{q})}\, u_j(\boldsymbol{q}) \eta_j$$

Projection on the reference modes:

$$\sqrt{\overline{\lambda}_i}\,\xi_i = \sum_{j=1}^{+\infty} \sqrt{\lambda_j(\boldsymbol{q})}\, \langle u_j(\boldsymbol{q}),\ \overline{u}_i \rangle \eta_j$$

Since $\boldsymbol{\eta} \sim \mathcal{N}(0, \mathrm{I})$, $\boldsymbol{\xi}$ is Gaussian with $\mathbb{E}(\boldsymbol{\xi}) = 0$ and, using Mercer's theorem,

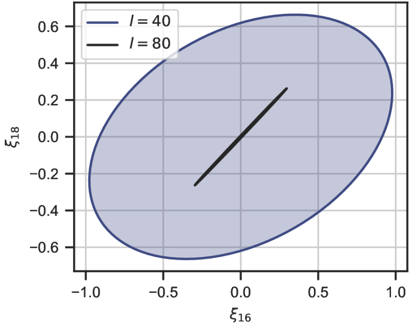
$$\Sigma(\boldsymbol{q})_{ij} = \mathbb{E}(\xi_i \xi_j) = \left(\overline{\lambda}_i \overline{\lambda}_j\right)^{-1/2} \left\langle \left\langle \sum_{k=1}^{+\infty} \lambda_k(\boldsymbol{q}) u_k(x, \boldsymbol{q}) u_k(y, \boldsymbol{q}),\ \overline{u}_i \right\rangle,\ \overline{u}_j \right\rangle$$

$$= \left(\overline{\lambda}_i \overline{\lambda}_j\right)^{-1/2} \langle\langle k(\cdot, \cdot, \boldsymbol{q}),\ \overline{u}_i \rangle,\ \overline{u}_j \rangle$$

Polette et al., *in Rev.*, 2024

3. Change of measure: Sampling

Hierarchical sampling: $\xi \sim \mathcal{N}(0, \Sigma(\boldsymbol{q}))$, the prior distribution of ξ can be highly sensitive to $\boldsymbol{q}$



$\Sigma(\boldsymbol{q})$ covariance projected on (ξ_{16}, ξ_{18})

$\Rightarrow$ Introduction of an **auxiliary variable** ζ whose prior law does not depend on hyperparameters

- Sample $\zeta \sim \mathcal{N}(0, \mathrm{I}_r)$, $\boldsymbol{q}$
- Compute $\xi \sim \mathcal{N}(0, \Sigma(\boldsymbol{q}))$ from $(\zeta, \boldsymbol{q})$,
$$\xi = \Sigma(\boldsymbol{q})^{1/2}\zeta$$

The proposition is not symmetric anymore
Ratio of the transition probabilities:

$$\frac{p_t\left(\xi^{(n)}, \boldsymbol{q}^{(n)} | \xi^*, \boldsymbol{q}^*\right)}{p_t\left(\xi^*, \boldsymbol{q}^* | \xi^{(n)}, \boldsymbol{q}^{(n)}\right)} = \left(\frac{\det\left(\Sigma(\boldsymbol{q}^*)\right)}{\det\left(\Sigma(\boldsymbol{q}^{(n)})\right)}\right)^{1/2}$$

Betancourt and Girolami, *arXiv*, 2013

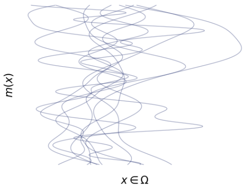
$$m(x) = \sum_{i=1}^{r} \sqrt{\overline{\lambda}_i}\,\overline{u}_i(x)\xi_i, \text{ with } \boldsymbol{\xi} \sim \mathcal{N}(0, \Sigma(\boldsymbol{q}))$$

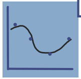
Proposition $\boldsymbol{\xi}^{(i)}, \boldsymbol{q}^{(i)}$

m infinite dimensional

→ dimension reduction

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





Surrogate model

Polynomial chaos $\widetilde{F}(\boldsymbol{\xi}) = \boldsymbol{d}$

Comparison

$\boldsymbol{d}$ versus $\boldsymbol{d}^{\text{obs}}$ (data)

Bayes' rule: $\pi_{\text{post}}(\boldsymbol{\xi}, \boldsymbol{q}|\boldsymbol{d}^{\text{obs}}) \propto \mathcal{L}(\boldsymbol{d}^{\text{obs}}|\boldsymbol{\xi})\,\pi_{\text{prior}}(\boldsymbol{\xi}|\boldsymbol{q})\pi_{\text{prior}}(\boldsymbol{q})$

Accept/Reject criterion

New iteration

$i \leftarrow i + 1$

Table of contents

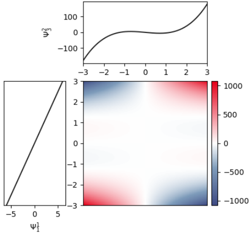
Polynomial chaos expansion (PCE)

$$\widetilde{F}(\xi) = \sum_{a \in \mathcal{A}} f_a \Psi_a(\xi) = \boldsymbol{d}$$

- $\mathcal{A}$: set of multi-indices *e.g.* $\{(0,1,0);(2,0,0);(1,0,3)\}$
- Ψ_a: product of orthonormal univariate polynomials:

$$\Psi_{a=(a_1,\ldots,a_r)}(m) = \prod_{i=1}^{r} \psi^i_{a_i}(\xi_i)$$

- f_a: coefficients to compute
- PCE also used to approximate $\log\det(\Sigma(\boldsymbol{q}))$, $\Sigma(\boldsymbol{q})^{-1/2}$, $\Sigma(\boldsymbol{q})^{1/2}$



Example with Hermite polynomials and $\alpha = \{1,3,0\}$

4. Surrogate model: Polynomial chaos expansion

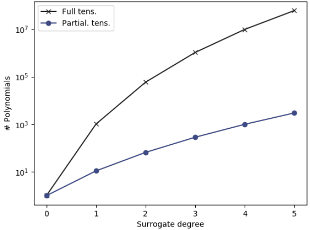
Polynomial chaos expansion (PCE)

$$\widetilde{F}(\boldsymbol{\xi}) = \sum_{a \in \mathcal{A}} f_a \Psi_a(\boldsymbol{\xi}) = \boldsymbol{d}$$

Non intrusive ordinary least squares approach

- Training set $\{\boldsymbol{\xi}^{(n)}\}_{1 \leqslant n \leqslant N} \sim \pi_{\text{prior}}(\boldsymbol{\xi})$, $N >> K = |\mathcal{A}|$
- Training evaluations $\boldsymbol{U} = \left(F(\boldsymbol{\xi}^{(1)}), \dots, F(\boldsymbol{\xi}^{(N)})\right)^\top$
- Polynomial evaluations at training points $\boldsymbol{\Psi} \in \mathbb{R}^{N \times K}$, $\Psi_{ij} = \psi_j(\boldsymbol{\xi}^{(i)})$
- Vector of coefficients $\boldsymbol{f} = (f_1, \dots, f_K)^\top$

$$\boldsymbol{f} = \left(\boldsymbol{\Psi}^\top \boldsymbol{\Psi}\right)^{-1} \boldsymbol{\Psi}^\top \boldsymbol{U}$$



Number of polynomials according to surrogate degree n_o $(r = 10)$
Full tensorization: $\max(n_{o,i}) \leqslant n_o$
Partial tensorization: $\sum_i n_{o,i} \leqslant n_o$

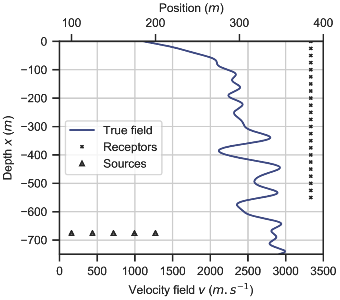
4. Surrogate model: Adaptive construction

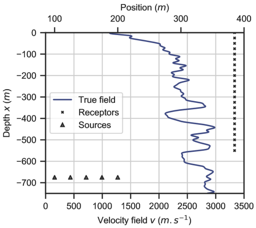
- Initial construction: does not ensure that the error on the posterior subspace is bounded
- Objective: minimize error $\mathbb{E}_{\pi_{\text{post}}}\left(\left\|\mathcal{L}(\xi) - \widetilde{\mathcal{L}}(\xi)\right\|^2\right)$
- Adaptive workflow:
  - Initial surrogate $\widetilde{\mathcal{L}}^{(0)}$ with $X^{(0)} = \{\xi^{(n)}\}_{1 \leqslant n \leqslant N} \sim \pi_{\text{prior}}$
  - While Convergence not achieved
    - MCMC with $\widetilde{\pi}^{(i)}_{\text{post}}$
    - $i \leftarrow i + 1$
    - Update training set: $X^{(i)} = X^{(i-1)} \backslash X^{(i-1)}_{1 \leqslant k \leqslant n_r} \cup \{\xi^{(n)}\}_{1 \leqslant n \leqslant n_a}$
    - Update surrogate *using centered rescaled training set*
- Convergence check: surrogate quality

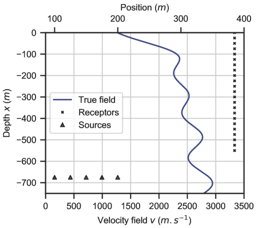
Lucor and Le Maître, *ESAIM: Proceedings and Surveys*, 2018

Table of contents

5. Results: Case presentation







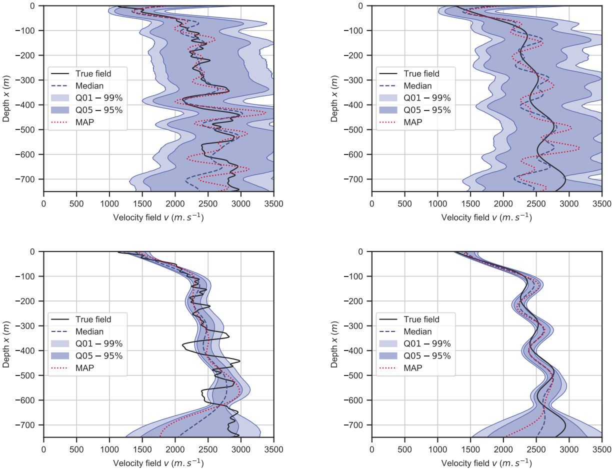
- Based on a realistic velocity model (*Amoco Tulsa Research Lab*)
- 2D velocity field, varies only along depth
- $\Omega = [0, 750]$m, 23 stations $\times$ 5 events, noise level 0.002s
- Velocity field: $v(x) = \exp\left(\mu + \sum_{i=1}^{r} \sqrt{\overline{\lambda}_i}\,\overline{u}_i(x)\xi_i\right)$
- $l \sim \mathrm{U}(10, 100)$, $A \sim \mathrm{IG}(21, 1)$, $r = 20$, $\mu \sim \mathrm{U}(6.9, 8.1)$

5. Results: with fixed hyperparameters

Small wavelength field | Large wavelength field

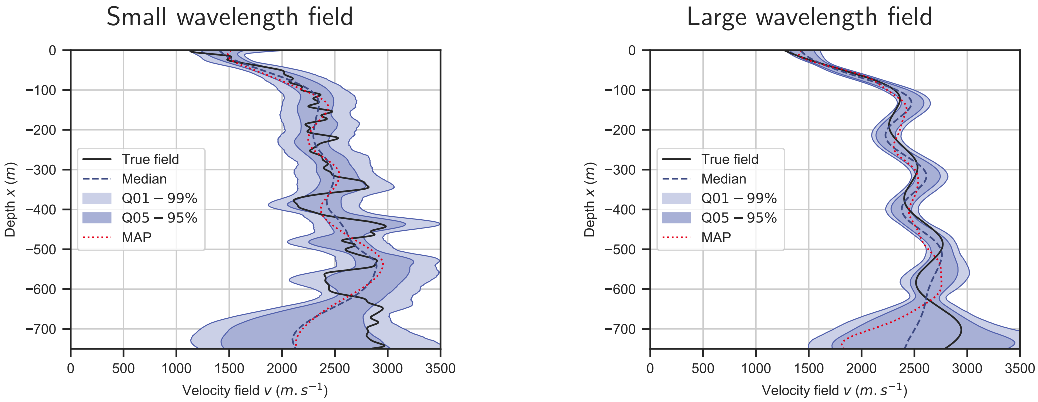
$l = 10$

$l = 80$



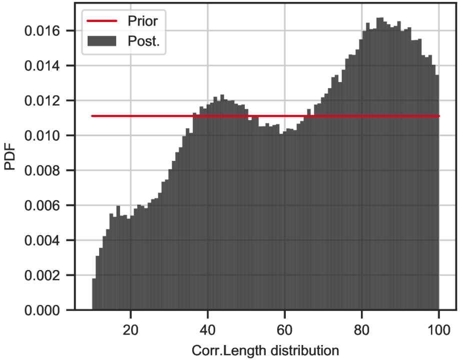
⇒ Using the same basis for both fields does not allow to distinguish them

5. Results: with the change of measure method

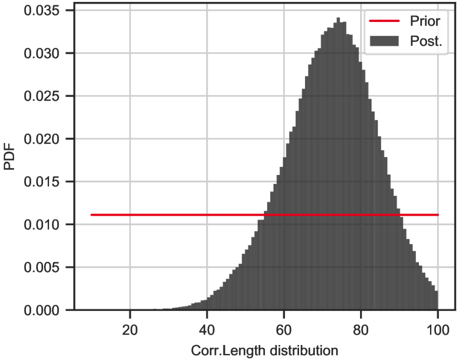


5. Results: with the change of measure method

Small wavelength field

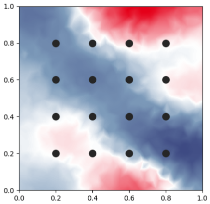


Large wavelength field

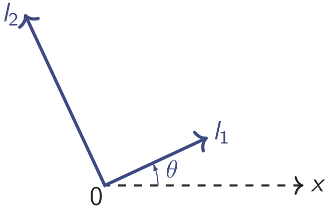


Application to a steady state diffusion equation

- **Forward model:** $-\nabla \cdot (\kappa \nabla u) = f$
- **Objective:** Find $\pi_{\text{post}}\left(\kappa | \boldsymbol{d}_{\text{obs}}\right)$, with $\boldsymbol{d}_{\text{obs}} = \{u(s_i)\}_{1 \leqslant i \leqslant N_{\text{obs}}}$
- **Prior parametrisation:**
  - $\log \kappa \sim \mathcal{N}(0, k)$
  - k anisotropic Gaussian autocovariance function $\rightarrow k(x, y) = A \exp\left(-\frac{1}{2}(x-y)^\top K(x-y)\right)$, with $K = \mathcal{R}(\theta) \begin{pmatrix} 1/l_1^2 & 0 \\ 0 & 1/l_2^2 \end{pmatrix} \mathcal{R}(\theta)^T$
  - $A \sim \mathrm{IG}(3,1)$, $l_1 \sim \mathcal{U}(0.1, 0.6)$, $l_2 \sim \mathcal{U}(0.1, 0.6)$, $\theta \sim \mathcal{U}(0, \pi/2)$



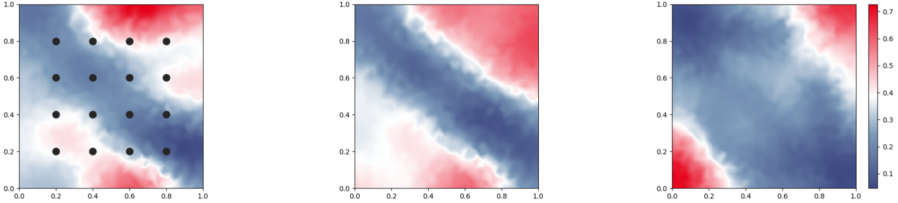
(a) True field



(b) Parametrization

Application to a steady state diffusion equation

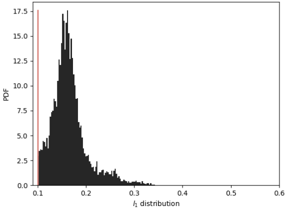
- **Forward model:** $-\nabla \cdot (\kappa \nabla u) = f$
- **Objective:** Find $\pi_{\text{post}}(\kappa | \boldsymbol{d}_{\text{obs}})$, with $\boldsymbol{d}_{\text{obs}} = \{u(s_i)\}_{1 \leqslant i \leqslant N_{\text{obs}}}$
- **Prior parametrisation:**
  - $\log \kappa \sim \mathcal{N}(0, k)$
  - k anisotropic Gaussian autocovariance function $\rightarrow k(x, y) = A \exp\left(-\frac{1}{2}(x-y)^\top K(x-y)\right)$, with $K = \mathcal{R}(\theta) \begin{pmatrix} 1/l_1^2 & 0 \\ 0 & 1/l_2^2 \end{pmatrix} \mathcal{R}(\theta)^T$
  - $A \sim \mathrm{IG}(3,1),\ l_1 \sim \mathcal{U}(0.1, 0.6),\ l_2 \sim \mathcal{U}(0.1, 0.6),\ \theta \sim \mathcal{U}(0, \pi/2)$



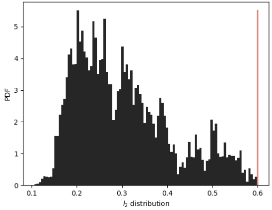
(a) True field (b) Mean posterior field (c) Std posterior field

Application to a steady state diffusion equation

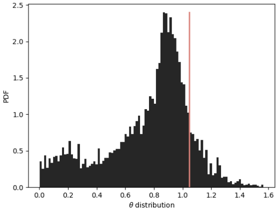
- **Forward model:** $-\nabla \cdot (\kappa \nabla u) = f$
- **Objective:** Find $\pi_{\mathrm{post}}\left(\kappa | \boldsymbol{d}_{\mathrm{obs}}\right)$, with $\boldsymbol{d}_{\mathrm{obs}} = \{u(s_i)\}_{1 \leqslant i \leqslant N_{\mathrm{obs}}}$
- **Prior parametrisation:**
  - $\log \kappa \sim \mathcal{N}(0, k)$
  - k anisotropic Gaussian autocovariance function $\rightarrow k(x, y) = A \exp\left(-\frac{1}{2}(x-y)^{\top} K (x-y)\right)$, with $K = \mathcal{R}(\theta) \begin{pmatrix} 1/l_1^2 & 0 \\ 0 & 1/l_2^2 \end{pmatrix} \mathcal{R}(\theta)^T$
  - $A \sim \mathrm{IG}(3,1)$, $l_1 \sim \mathcal{U}(0.1, 0.6)$, $l_2 \sim \mathcal{U}(0.1, 0.6)$, $\theta \sim \mathcal{U}(0, \pi/2)$



(a) l_1 marginal post. distribution



(b) l_2 marginal post. distribution



(c) θ marginal post. distribution

Conclusion

- *Change of measure*: efficient method for field inference
  - Dimension reduction
  - Flexible *a priori* parametrization → model exploration
  - Without large computational cost increase

N. Polette, O. Le Maître, P. Sochala, A. Gesret, *Change of Measure for Bayesian Field Inversion with Hierarchical Hyperparameters Sampling*, in Rev. in JCP

- Uncertainty propagation to other quantities (*e.g.* location)
- *Work in progress*: development of adaptive methods
  - Adaptive PC using posterior sampling
  - Informed modes ⊥ Non-informed modes
  - *A posteriori* dimension reduction

Thank you !

nadege.polette@minesparis.psl.eu

Keywords: inverse problem, (hierarchical) Bayesian inference, surrogate models, MCMC, dimension reduction, KL decomposition