Field uncertainties estimation through (hyper)parameters sampling using Bayesian inference

PhD student: Nadège Polette\textsuperscript{1,2},
Supervisors: Pierre Sochala\textsuperscript{1}, Alexandrine Gesret\textsuperscript{2}, Olivier Le Maître\textsuperscript{3}

\textsuperscript{1} CEA, DAM, DIF, F-91297 Arpajon, France
\textsuperscript{2} Mines Paris PSL, centre de Géosciences, Fontainebleau, France
\textsuperscript{3} CMAP, CNRS, Inria, École Polytechnique, IPP, Palaiseau, France
Context

Detection and analysis of seismic events

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

Local scale
- Knowledge of subsurface
- Exploitation

Regional scale
- Tsunami and seism alerts
- Risk prevention
Context: seismic tomography

Quantity of interest

Velocity field $f$

Forward problem

Forward model\(^{(a)}\)

$M$

Indirect observations

Arrival time

$\mathbf{d}^{\text{obs}} = M(f) + \varepsilon$

(a) Eikonal solver [Noble and Gesret, 2011]
Context: seismic tomography

Quantity of interest

Velocity field $f$

Forward problem

Forward model\(^{(a)}\)

$\mathcal{M}$

Indirect observations

Arrival time

$\mathbf{d}_{\text{obs}} = \mathcal{M}(f) + \varepsilon$

? Inverse problem ?

(a) Eikonal solver [Noble and Gesret, 2011]

**Objective:** Estimation of a field

(i) accurate,

(ii) with uncertainties,

(iii) fast
Bayes formulation

Bayes rule:  
\[ p_{\text{post}}(f|d^{\text{obs}}) \propto \mathcal{L}(d^{\text{obs}}|f)\pi_F(f) \]

Markov Chain Monte–Carlo algorithm:

First proposal  
First proposal  
\[ f^{(0)} \]

Proposal evaluation  
\[ f^{(n)} \]

\[ \pi \quad \text{Prior} \]

\[ \mathcal{L}(d^{\text{obs}}|f) \quad \text{Likelihood} \]

\[ p_{\text{post}}(f|d^{\text{obs}}) \quad \text{Posterior} \]

\[ n \leftarrow n + 1 \]

\[ \mathcal{M}(f) \quad \text{Model solve} \]

\[ \text{Metropolis–Hastings criterion} \]
Bayes formulation

**Bayes rule:**
\[ p_{\text{post}}(f|d^{\text{obs}}) \propto \mathcal{L}(d^{\text{obs}}|f)\pi_F(f) \]

**Markov Chain Monte–Carlo algorithm:**

\[
\begin{align*}
\text{First proposal} \quad & f^{(0)} \quad \text{Iteration } n \\
\text{Proposal evaluation} \quad & f^{(n)} \quad \mathcal{M}(f) \quad \text{Model solve} \\
\quad & \pi \quad \text{Prior} \\
& L \left( d^{\text{obs}}|f \right) \quad \text{Likelihood} \\
& p_{\text{post}} \left( f|d^{\text{obs}} \right) \quad \text{Posterior} \\
& n \leftarrow n + 1 \quad \text{Metropolis–Hastings criterion}
\end{align*}
\]

⇒ Representation of \( f \)?
Evaluation of \( \mathcal{M} \)?
→ Polynomial chaos surrogate [Marzouk and Najm, 2009]

N. Polette - ETICS2023 - Cap Féret

2023/10/12
Field representation

$f(x)$ is seen as a particular realization of a random (Gaussian) process $G \sim N(0, k)$, where $k$ is the autocovariance function.

Karhunen–Loève decomposition

\[
f(x) = G(x, \theta) \simeq \sum_{i=1}^{r} \lambda_i^{1/2} u_i(x) \eta_i(\theta), \text{ with } \eta_i = \lambda_i^{-1/2} \langle u_i, G \rangle_{\Omega}
\]

- $(u_i, \lambda_i)_{i \in \mathbb{N}^*}$ eigenelements of $k$:

\[
\langle k(x, \cdot), u_i \rangle_{\Omega} := \int_{\Omega} k(x, x') u_i(x') dx' = \lambda_i u_i(x)
\]

- Bi-orthonormality of the decomposition:
  - $\forall i, j \in \mathbb{N}^*, u_i, u_j$ orthonormal, $\langle u_i, u_j \rangle_{\Omega} = \delta_{i,j}$,
  - $\eta := (\eta_i)_{1 \leq i \leq r} \sim N(0, I_r)$ (if Gaussian process)

\[
\Rightarrow p_{\text{post}}(\eta | d^{\text{obs}}) \propto L(d^{\text{obs}} | \eta)\pi(\eta)
\]

[Karhunen, 1946][Loève, 1977]
Dependency on hyperparameters

In fact, $G \sim \mathcal{N}(0, k(q))$,

$$f(x) \simeq \sum_{i=1}^{r} \lambda_i^{1/2}(q) u_i(x, q) \eta_i,$$

with $\eta_i = \lambda_i^{-1/2}(q) \langle u_i(\cdot, q), f \rangle_{\Omega}$

Squared exponential kernel: $k(x, y, q := \{A, \ell\}) = A \exp\left(-\frac{\|x-y\|^2}{2\ell^2}\right)$

$$q^{(1)} \sim \eta_1 + \eta_2 + \ldots + \eta_r$$

$$q^{(2)} \sim \eta_1 + \eta_2 + \ldots + \eta_r$$

$$\Rightarrow p_{\text{post}}(\eta, q | d^{\text{obs}}) \propto \mathcal{L}(d^{\text{obs}} | \eta, q) \pi(\eta, q)$$
Change of measure method

Change of measure:

- Reference kernel $\overline{k}$ and associated basis $(\overline{\lambda}_i, \overline{u}_i)_{i \in \mathbb{N}^*}$ [Sraj et al., 2016]

- Sample $(\xi, q)$: the q-dependency is transferred to the coordinates law, [NP/Sochala/Le Maître/Gesret, in prep.]

$$f(x) \sim \overline{G}^r(x, \theta) := \sum_{i=1}^{r} \lambda_i^{1/2} \overline{u}_i(x) \xi_i(\theta) \text{ with } \xi \sim \mathcal{N}(0, \Sigma(q))$$

and $p_{\text{post}}(\xi, q | d^{\text{obs}}) \propto \mathcal{L}(d^{\text{obs}} | \xi) \pi(\xi | q) \pi(q)$

The covariance matrix $\Sigma(q)$ writes

$$\forall 1 \leq i, j \leq r, \forall q \in \mathbb{H}, \quad \Sigma_{ij}(q) := (\overline{\lambda}_i \overline{\lambda}_j)^{-1/2} \langle \langle k(\cdot, \cdot, q), \overline{u}_j \rangle_{\Omega}, \overline{u}_i \rangle_{\Omega}$$
\[
\xi 
\xrightarrow{\text{Sampling}} q 
\xrightarrow{\Sigma(q)} \xi \sim \mathcal{N}(0, \Sigma(q)) 
\xrightarrow{\mathcal{L}(d^{\text{obs}} | \xi)} p_{\text{post}}(\xi, q | d^{\text{obs}})
\]

\[
\Sigma_{ij}(q) = (\bar{\lambda}_i \bar{\lambda}_j)^{-1/2} \left\langle \left\langle k(\cdot, \cdot, q), \bar{u}_j \right\rangle_{\Omega} , \bar{u}_i \right\rangle_{\Omega}
\]
Sampling problem

Hierarchical formulation: \( p_{\text{post}}(\xi, q|d^{\text{obs}}) \propto \mathcal{L}(d^{\text{obs}}|\xi)p(\xi|q)p(q), \)
\( \xi \sim \mathcal{N}(0, \Sigma(q))): \text{the prior distribution of } \xi \text{ is highly sensitive to } q \)

Projection of \( \xi \) priors for two different \( q \)
Sampling strategy

Introduction of an auxiliary variable $\bar{\xi}$ whose prior law does not depend on the hyperparameters (e.g. $\bar{\xi} \sim \mathcal{N}(0, \Sigma_{\bar{\xi}})$)

Algorithm 1  Sampling step for CoM method.

Input: $Y^{(n)} = (\bar{\xi}^{(n)}, \xi^{(n)}, q^{(n)})$  ▶ Current state

1: Propose $(\bar{\xi}^*, q^*) \sim \mathcal{N}((\bar{\xi}^{(n)}, q^{(n)}), \widehat{C}^{(n)})$  ▶ Adaptive random walk
2: Set $\xi^* = \Sigma(q^{(n)})^{1/2} \Sigma_{\bar{\xi}}^{-1/2} \bar{\xi}^*$  ▶ Change of variable
3: Accept/Reject according to MH criterion on $p_{\text{post}}(\bar{\xi}^*, q^*)$.

- the ratio of the transition probabilities becomes

$$\frac{p(\xi^{(n)}, q^{(n)}|\bar{\xi}^*, q^*)}{p(\bar{\xi}^*, q^*|\xi^{(n)}, q^{(n)})} = \left( \frac{\det \Sigma(q^*)}{\det \Sigma(q^{(n)})} \right)^{1/2}$$
Surrogate quantities

Evaluating forward model $\mathcal{M}$ and derived quantities of $\Sigma(q)$ at each step is expensive $\Rightarrow$ use of *Polynomial Chaos (PC) expansions* [Wiener, 1938][Ghanem and Spanos, 1991][Xiu and Karniadakis, 2002]

$$m(\zeta) \approx \tilde{m}(\zeta) = \sum_{a \in \mathcal{A}} m_a \Psi_a(\zeta),$$

- Forward model: $\mathcal{M} \circ f(\xi) = \tilde{\mathcal{M}}(\xi)$
- $\Sigma(q)^{-1} = \tilde{\Sigma}^{-1/2}(q)\tilde{\Sigma}^{-1/2}(q)$
- $\log \det [\Sigma(q)] = (\log \det \Sigma)(q)$
- $\Sigma(q)^{1/2} = \tilde{\Sigma}^{1/2}(q)$

$\Sigma^{\pm 1/2}$ is obtained by eigendecomposition of $\Sigma$

$$\Sigma^{\pm 1/2} = U\Lambda^{\pm 1/2}U^\top,$$

where $(U, \Lambda)$ are the eigenelements of $\Sigma$
Application to seismic tomography

Application case: 1D section of Amoco model [O’Brien and Regone, 1994] and location of stations

\( \mathbf{d}^{\text{obs}} \): time of arrival, with noise level \( \alpha = 0.002 \text{s} \)

\( r = 20, \quad \mathbf{q} = \{A, \ell\} \)
Inference of a field (SW)

(a) Proposed method

Comparison of inference results for different bases - small wavelength field

(b) $\ell = 10$
Inference of a field (SW)

(a) Proposed method

(b) $\ell = 34$

Comparison of inference results for different bases - small wavelength field
Inference of a field (SW)

(a) Proposed method

(b) $\ell = 80$

Comparison of inference results for different bases - small wavelength field
Inference of a field (LW)

(a) Proposed method

(b) $\ell = 10$

Comparison of inference results for different bases - large wavelength field
Inference of a field (LW)

(a) Proposed method

(b) $\ell = 34$

Comparison of inference results for different bases - large wavelength field
Inference of a field (LW)

(a) Proposed method

Comparison of inference results for different bases - large wavelength field

(b) $\ell = 80$
Hyperparameters distribution (LW)

(a) Field posterior distribution
(b) Correlation length posterior distribution

Posterior distributions - large wavelength field
Hyperparameters distribution (SW)

(a) Field posterior distribution

(b) Correlation length posterior distribution

Posterior distributions - small wavelength field
Conclusion and perspectives

- Change of measure
  - Dimension reduction of the field
  - Enlarge *a priori* parametrization \(\Rightarrow\) uncertainties are less ruled by the model selection
  - Without large computational cost increase

Thank you!
Conclusion and perspectives

- Change of measure
  - Dimension reduction of the field
  - Enlarge *a priori* parametrization $\Rightarrow$ uncertainties are less ruled by the model selection
  - Without large computational cost increase
- Development of adaptive methods
- Use of the posterior quantity to propagate uncertainty to other parameters (e.g. seismic location)
- *How to quantify the relevance of the uncertainties?*
Conclusion and perspectives

- Change of measure
  - Dimension reduction of the field
  - Enlarge *a priori* parametrization $\Rightarrow$ uncertainties are less ruled by the model selection
  - Without large computational cost increase

- Development of adaptive methods

- Use of the posterior quantity to propagate uncertainty to other parameters (e.g. seismic location)

- *How to quantify the relevance of the uncertainties?*

Thank you!


Polynomial Chaos expansions

Evaluating forward model and derived quantities of $\Sigma(q)$ at each step is expensive $\Rightarrow$ use of *Polynomial Chaos (PC) expansions*


$m(\zeta) \simeq \tilde{m}(\zeta) = \sum_{a \in \mathcal{A}} m_a \psi_a(\zeta),$

- $\{\psi_a(\zeta)\}_{a \in \mathcal{A}}$: orthogonal polynomials in $L^2(\zeta)$
- $\{m_a\}_{a \in \mathcal{A}}$: PC coefficients
- $\mathcal{A}$: set of multi-indexes

Building framework:
- Exact evaluations at quadrature points (Gauss–Legendre or sparse grids)
- Computations of PC coefficients by matrix-vector product
Surrogate quantities

Evaluating forward model and derived quantities of $\Sigma(q)$ at each step is expensive $\Rightarrow$ use of *Polynomial Chaos (PC) expansions* 


$$m(\zeta) \simeq \tilde{m}(\zeta) = \sum_{a \in A} m_a \Psi_a(\zeta),$$

- **Forward model:** $\mathcal{M} \circ f(\xi) = \tilde{\mathcal{M}}(\xi)$
- $\Sigma(q)^{-1} = \Sigma^{-1/2}(q)\Sigma^{-1/2}(q)$
- $\log \det [\Sigma(q)] = (\log \det \Sigma)(q)$
- $\Sigma(q)^{1/2} = \tilde{\Sigma}^{1/2}(q)$

$\Sigma^{\pm 1/2}$ is obtained by eigendecomposition of $\Sigma$

$$\Sigma^{\pm 1/2} = U \Lambda^{\pm 1/2} U^\top,$$

where $(U, \Lambda)$ are the eigenelements of $\Sigma$
Projection of a discrete field

Velocity field $v$ (m. s$^{-1}$)

Depth $x$ (m)

$V_{\text{true}}$  $V_{\text{KL}}$
Inference of a discrete field

The diagram illustrates the velocity field \( v \) (m/s) as a function of depth \( x \) (m). The true field is depicted by a solid black line, with median values shown as a dashed line. The 99% quantile (Q01) and 95% quantile (Q05) are indicated by shaded areas, and the Maximum A Posteriori (MAP) estimate is represented by a dotted line.
Projection of a non-stationary field
Inference of a non-stationary field

Velocity field $v$ (m $s^{-1}$)

Depth $x$ (m)

- True field
- Median
- Q01 – 99%
- Q05 – 95%
- MAP
Sampling problem

Hierarchical formulation: \( p_{\text{post}}(\xi, q|d_{\text{obs}}) \propto \mathcal{L}(d_{\text{obs}}|\xi)\pi(\xi|q)\pi(q) \), \( \xi \sim \mathcal{N}(0, \Sigma(q)) \): the prior distribution of \( \xi \) is highly sensitive to \( q \)

Projection of \( \xi \) priors for two different \( q \)
Sampling problem

Hierarchical formulation: \( p_{\text{post}}(\xi, q|d^{\text{obs}}) \propto \mathcal{L}(d^{\text{obs}}|\xi)\pi(\xi|q)\pi(q) \), \( \xi \sim \mathcal{N}(0, \Sigma(q)) \): *the prior distribution of \( \xi \) is highly sensitive to \( q \)*

![Graph showing projection of \( \xi \) priors for two different \( q \)]

Initial sampling: Metropolis–Hastings (MH) random walk with adapted covariance proposal

\[
Y^* := (\xi^*, q^*) \sim \mathcal{N}(Y^{(n)}, \hat{C}^{(n)}), \text{ with } \hat{C}^{(n)} \propto \text{Cov} \left( Y^{(1)}, \ldots, Y^{(n)} \right)
\]