Bayesian Inference for Inverse Problems with Hyperparameters Estimation of the Field Covariance Function

Nadège Polette[1,2], Pierre Sochala[1], Alexandrine Gesret[2], Olivier Le Maître[3]

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

PFE - November '22 - March '23, ENPC supervisor: Julien Reygner [CERMICS]
February, 23rd 2023

Context

Detection and analysis of seismic events

Global scale

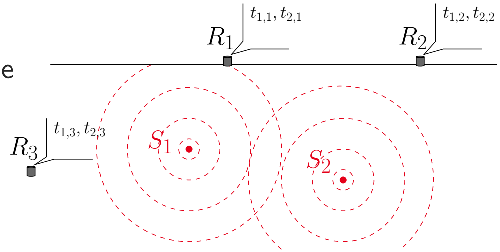
- International treaties (CTBT, NTP)
- Environment monitoring (IMS)

Regional scale

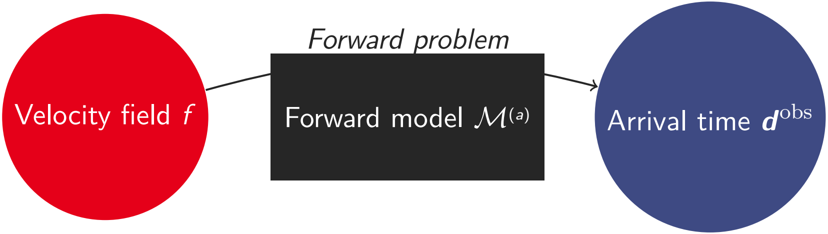
- Tsunami and seism alerts
- Risk prevention

Local scale

- Knowledge of subsurface
- Exploitation

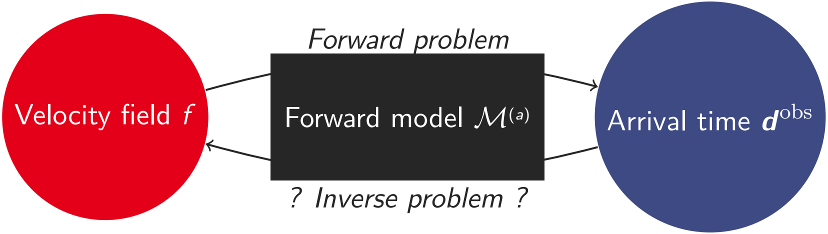


Context: seismic tomography



[a] Eikonal solver [Noble et al., 2011]

Context: seismic tomography



[a] Eikonal solver [Noble et al., 2011]

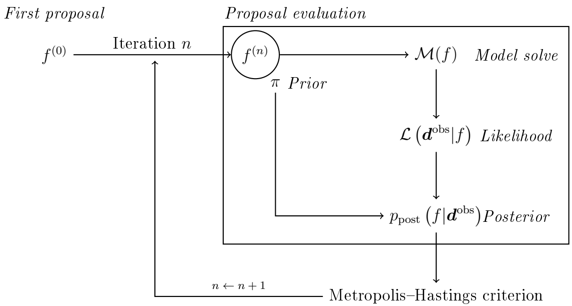
Objective: Estimation of a field (*i*) accurate, (*ii*) with uncertainties, (*iii*) fast

Table of contents

Bayes formulation

Bayes rule: $$p_{\text{post}}(f|\boldsymbol{d}^{\text{obs}}) \propto \mathcal{L}(\boldsymbol{d}^{\text{obs}}|f)\pi_{\mathcal{F}}(f).$$

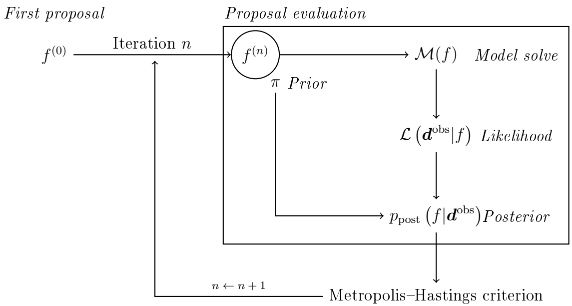
Markov Chain Monte–Carlo algorithm:



Bayes formulation

Bayes rule: $$p_{\text{post}}(f|\boldsymbol{d}^{\text{obs}}) \propto \mathcal{L}(\boldsymbol{d}^{\text{obs}}|f)\pi_{\mathcal{F}}(f).$$

Markov Chain Monte–Carlo algorithm:

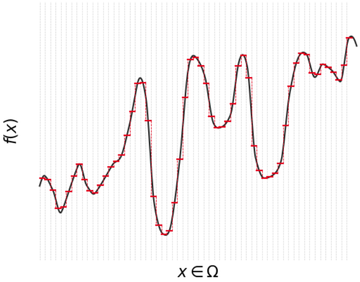


$\Rightarrow$ Evaluation of $\mathcal{M}$?

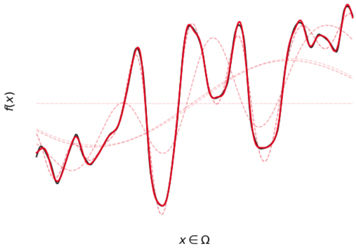
$\rightarrow$ Polynomial chaos surrogate [Marzouk et al., 2009].

Representation of f ?

Representation of the field



(a) Nodal representation



(b) Modal representation

- ✗ Large number of parameters (expensive)
- ✗ Interpolation needed
- ✓ Easy to implement

- ✓ Few number of modes
- ✓ Defined on all the spatial domain

$\Rightarrow$ Implementation ?

Karhunen–Loève decomposition

$f(\boldsymbol{x})$ is seen a particular *realization of a Gaussian process* $\mathcal{G} \sim \mathcal{N}(0, k)$, where k is the *autocovariance function* [Karhunen, Loève, 1946, 1977].

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

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

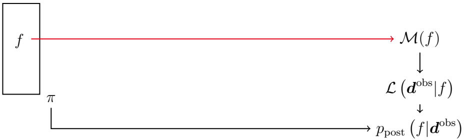
$$\langle k(\boldsymbol{x}, \cdot), u_i \rangle_\Omega := \int_\Omega k(\boldsymbol{x}, \boldsymbol{x}') u_i(\boldsymbol{x}') d\boldsymbol{x}' = \lambda_i u_i(\boldsymbol{x})$$

- *Bi-orthonormality* of the decomposition:
  - $\forall i, j \in \mathbb{N}^*$, u_i, u_j orthonormal, $\langle u_i, u_j \rangle_\Omega = \delta_{i,j}$,
  - $\boldsymbol{\eta} := (\eta_i)_{1 \leqslant i \leqslant r} \sim \mathcal{N}(0, \mathrm{I}_r)$

$\Rightarrow p_{\mathrm{post}}(f(\boldsymbol{\eta}) | \boldsymbol{d}^{\mathrm{obs}}) \propto \mathcal{L}(\boldsymbol{d}^{\mathrm{obs}} | f) \pi(\boldsymbol{\eta})$.

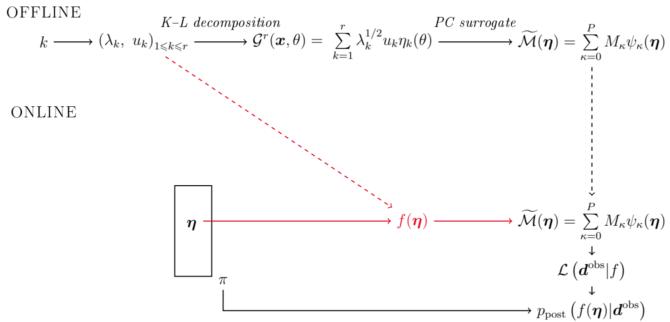
Karhunen-Loève decomposition

ONLINE



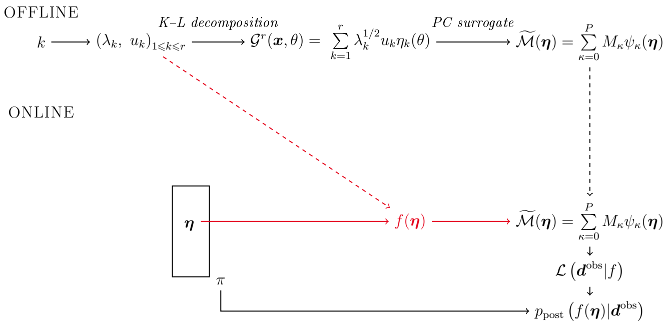
- Representation of f ?
- Evaluation cost of $\mathcal{M}$?

Karhunen-Loève decomposition



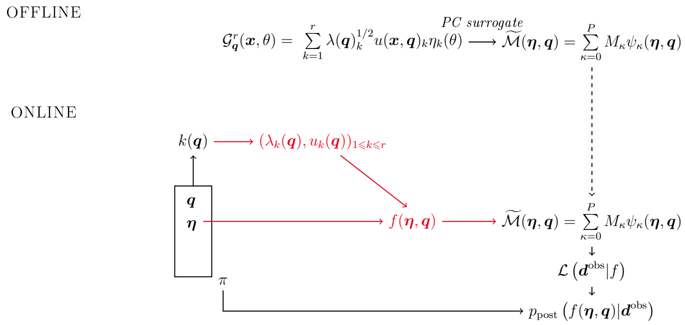
- $f(\boldsymbol{\eta}) = \sum_{i=1}^{r} \lambda_i^{1/2} u_i \eta_i$ with $\eta_i = \lambda_i^{-1/2} \langle u_i, \mathcal{G} \rangle_\Omega$
- $\widetilde{\mathcal{M}}(\boldsymbol{\eta}) = \sum_\kappa M_\kappa \psi_\kappa(\boldsymbol{\eta})$
- In reality, k depends on hyperparameters $\boldsymbol{q} \in \mathbb{H}$: $\mathcal{G} \sim \mathcal{N}(0, k(\boldsymbol{q}))$

Hyperparameters dependency



- $f(\boldsymbol{\eta}, \boldsymbol{q}) = \sum_{i=1}^{r} \lambda_i(\boldsymbol{q})^{1/2} u_i(\boldsymbol{q}) \eta_i$ with $\eta_i = \lambda_i(\boldsymbol{q})^{-1/2} \langle u_i(\boldsymbol{q}), \mathcal{G} \rangle_\Omega$
- $\widetilde{\mathcal{M}}(\boldsymbol{\eta}, \boldsymbol{q}) = \sum_\kappa M_\kappa \psi_\kappa(\boldsymbol{\eta}, \boldsymbol{q})$
- In reality, k depends on hyperparameters $\boldsymbol{q} \in \mathbb{H}$:
  $\mathcal{G} \sim \mathcal{N}(0, k(\boldsymbol{q}))$

Hyperparameters dependency



- $f(\boldsymbol{\eta},\boldsymbol{q}) = \sum_{i=1}^{r} \lambda_i(\boldsymbol{q})^{1/2} u_i(\boldsymbol{q}) \eta_i$ with $\eta_i = \lambda_i(\boldsymbol{q})^{-1/2} \langle u_i(\boldsymbol{q}), \mathcal{G}\rangle_\Omega$
- $\widetilde{\mathcal{M}}(\boldsymbol{\eta},\boldsymbol{q}) = \sum_\kappa M_\kappa \psi_\kappa(\boldsymbol{\eta},\boldsymbol{q})$
- Eigenvalue problem at each iteration + difficulties to build $\widetilde{\mathcal{M}}$

Reference basis [Sraj et al., 2016]

$$\forall \boldsymbol{x}, \boldsymbol{x}' \in \Omega, \qquad \overline{k}(\boldsymbol{x}, \boldsymbol{x}') := \mathbb{E}_{\mathbb{H}}(k(\boldsymbol{x}, \boldsymbol{x}', \cdot)) := \int_{\mathbb{H}} k(\boldsymbol{x}, \boldsymbol{x}', \boldsymbol{q}) \pi_{\mathbb{H}}(\boldsymbol{q}) d\boldsymbol{q},$$

The reference eigenelements $\{\overline{u}_i, \overline{\lambda}_i\}_{i \in \mathbb{N}^*}$, are solution of the reference eigenvalue problem:

$$\forall \boldsymbol{x} \in \Omega,\ \forall i \in \mathbb{N}^*, \qquad \int_{\Omega} \overline{k}(\boldsymbol{x}, \boldsymbol{x}') \overline{u}_i(\boldsymbol{x}') d\boldsymbol{x} = \overline{\lambda}_i \overline{u}_i(\boldsymbol{x}).$$

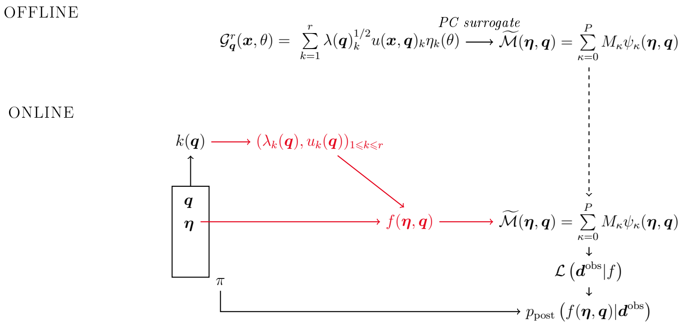
The representation basis does not depends on $\boldsymbol{q}$ *anymore.*

New field representation is obtained by *coordinates transformation*:

$$f(\boldsymbol{x}) \simeq \hat{f}^r(\boldsymbol{x}) := \sum_{j=1}^{r} \overline{\lambda}_j^{1/2} \overline{u}_j(\boldsymbol{x}) \hat{\eta}_j(\boldsymbol{q}, \theta),$$

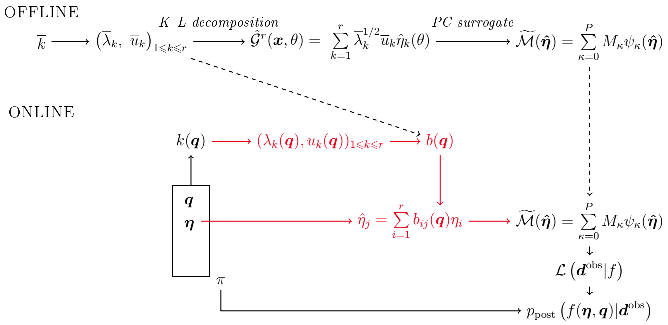
$$\text{where } \hat{\eta}_j(\boldsymbol{q}, \theta) := \sum_{i=1}^{r} \underbrace{\overline{\lambda}_j^{-1/2} \left\langle \lambda_i(\boldsymbol{q})^{1/2} u_i(\cdot, \boldsymbol{q}), \overline{u}_j \right\rangle_{\Omega}}_{:= b_{ij}(\boldsymbol{q})} \eta_i(\theta)$$

Reference basis



- $f(\boldsymbol{\eta},\boldsymbol{q}) = \sum_{i=1}^{r} \lambda_i(\boldsymbol{q})^{1/2} u_i(\boldsymbol{q})\eta_i$ with $\eta_i = \lambda_i(\boldsymbol{q})^{-1/2} \langle u_i(\boldsymbol{q}), \mathcal{G}\rangle_\Omega$
- $\widetilde{\mathcal{M}}(\boldsymbol{\eta},\boldsymbol{q}) = \sum_\kappa M_\kappa \psi_\kappa(\boldsymbol{\eta},\boldsymbol{q})$
- Eigenvalue problem at each iteration + difficulties to build $\widetilde{\mathcal{M}}$

Reference basis



- $f(\boldsymbol{\eta}, \boldsymbol{q}) = \sum_{i=1}^{r} \overline{\lambda}_i^{1/2} \overline{u}_i \hat{\eta}_i$ with $\hat{\eta}_j(\boldsymbol{q}, \theta) := \sum_{i=1}^{r} b_{ij}(\boldsymbol{q}) \eta_i(\theta)$
- $\widetilde{\mathcal{M}}(\hat{\boldsymbol{\eta}}) = \sum_{\kappa} M_{\kappa} \psi_{\kappa}(\hat{\boldsymbol{\eta}})$
- $\widetilde{b}$ difficult to build + physical sense [Sraj et al., Siripatana et al., 2016, 2020]

Table of contents

Change of measure

Change of coordinates: Sample $(\boldsymbol{\eta}, \boldsymbol{q})$

$$f(x) \simeq \hat{\mathcal{G}}^r(\boldsymbol{x}, \theta) = \sum_{i=1}^{r} \overline{\lambda}_i^{1/2} \overline{u}_i(\boldsymbol{x}) \hat{\eta}_i(\theta) \text{ with } \hat{\eta}_j(\boldsymbol{q}, \theta) = \sum_{i=1}^{r} b_{ij}(\boldsymbol{q}) \eta_i(\theta)$$

$$p_{\mathrm{post}}(f(\boldsymbol{\eta}, \boldsymbol{q}) | \boldsymbol{d}^{\mathrm{obs}}) \propto \mathcal{L}(\boldsymbol{d}^{\mathrm{obs}} | f(\boldsymbol{\eta}, \boldsymbol{q})) \pi(\boldsymbol{\eta}) \pi(\boldsymbol{q}).$$

Change of measure

Change of coordinates: Sample $(\boldsymbol{\eta}, \boldsymbol{q})$

$$f(x) \simeq \hat{\mathcal{G}}^r(\boldsymbol{x}, \theta) = \sum_{i=1}^{r} \overline{\lambda}_i^{1/2} \overline{u}_i(\boldsymbol{x}) \hat{\eta}_i(\theta) \text{ with } \hat{\eta}_j(\boldsymbol{q}, \theta) = \sum_{i=1}^{r} b_{ij}(\boldsymbol{q}) \eta_i(\theta)$$

$$p_{\text{post}}(f(\boldsymbol{\eta}, \boldsymbol{q}) | \boldsymbol{d}^{\text{obs}}) \propto \mathcal{L}(\boldsymbol{d}^{\text{obs}} | f(\boldsymbol{\eta}, \boldsymbol{q})) \pi(\boldsymbol{\eta}) \pi(\boldsymbol{q}).$$

Change of measure: Sample $(\boldsymbol{\xi}, \boldsymbol{q})$

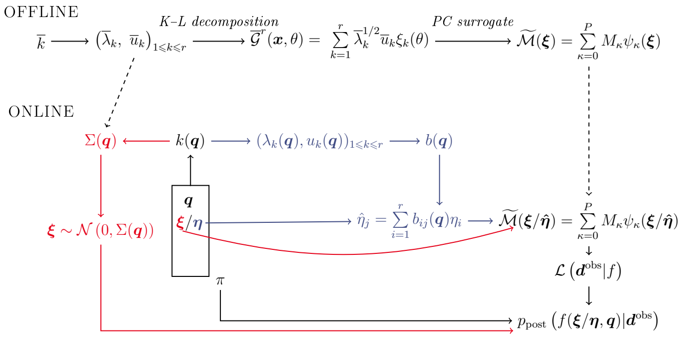
$$f(x) \simeq \overline{\mathcal{G}}^r(\boldsymbol{x}, \theta) := \sum_{i=1}^{r} \overline{\lambda}_i^{1/2} \overline{u}_i(\boldsymbol{x}) \xi_i(\theta) \text{ with } \boldsymbol{\xi} \sim \mathcal{N}(0, \Sigma(\boldsymbol{q}))$$

$$p_{\text{post}}(f(\boldsymbol{\xi}) | \boldsymbol{d}^{\text{obs}}) \propto \mathcal{L}(\boldsymbol{d}^{\text{obs}} | f(\boldsymbol{\xi})) \pi(\boldsymbol{\xi} | \boldsymbol{q}) \pi(\boldsymbol{q}).$$

The $\boldsymbol{q}$-dependency is transferred to the coordinates law. The covariance matrix $\Sigma(\boldsymbol{q})$ writes:

$$\forall 1 \leqslant i, j \leqslant r,\ \forall \boldsymbol{q} \in \mathbb{H}, \qquad \Sigma_{ij}(\boldsymbol{q}) := (\overline{\lambda}_i \overline{\lambda}_j)^{-1/2} \left\langle \left\langle k(\cdot, \cdot, \boldsymbol{q}), \overline{u}_j \right\rangle_\Omega, \overline{u}_i \right\rangle_\Omega .$$

Workflow



(CoC)

- $f(\boldsymbol{\eta},\boldsymbol{q})=\sum_{i=1}^{r}\overline{\lambda}_i^{1/2}\overline{u}_i\hat{\eta}_i$
- $\hat{\eta}_j(\boldsymbol{q},\theta)=\sum_{i=1}^{r}b_{ij}(\boldsymbol{q})\eta_i(\theta)$
- $b(\boldsymbol{q})$ ambiguous along $\boldsymbol{q}$

(CoM)

- $f(\boldsymbol{\xi})=\sum_{i=1}^{r}\overline{\lambda}_i^{1/2}\overline{u}_i\xi_i$
- $\boldsymbol{\xi}\sim\mathcal{N}(0,\Sigma(\boldsymbol{q}))$
- Σ is smooth along $\boldsymbol{q}$

Polynomial chaos surrogates

At each iteration, computation of $\mathcal{M}(f)$ and

$$\log p_{\text{post}}(f|\boldsymbol{d}^{\text{obs}}) \propto \log \mathcal{L}(\boldsymbol{d}^{\text{obs}}|f) + \log \pi_{\mathbb{H}}(\boldsymbol{q}) + \log \pi(\xi|\boldsymbol{q})$$

$\mathcal{L}(\boldsymbol{d}^{\text{obs}}|f)$ depends on $\mathcal{M}(f)$
$\xi \sim \mathcal{N}(0, \Sigma(\boldsymbol{q}))$ depends on $\Sigma^{-1/2}$ and logdet_Σ

Polynomial chaos surrogates: $Q(\zeta) \simeq \widetilde{Q}(\zeta) = \sum_{a \in \mathcal{A}} c_a P_a(\zeta).$

Quality assessment:

- Accuracy: use of *RRMSE*

$$\mathrm{RRMSE}(Q, \widetilde{Q}) = \sqrt{\frac{\sum_{i=1}^N \|Q^{(i)} - \widetilde{Q}^{(i)}\|^2}{\sum_{i=1}^N \|Q^{(i)}\|^2}}.$$

- Cost reduction: speed (*speed-up factor*) and number of exact evaluations

Table of contents

Transient diffusion equation (WIP)

$$\frac{\partial U}{\partial t} = \frac{\partial}{\partial \boldsymbol{x}} \left(f \frac{\partial U}{\partial x} \right).$$

Objective: comparison with (CoC) method

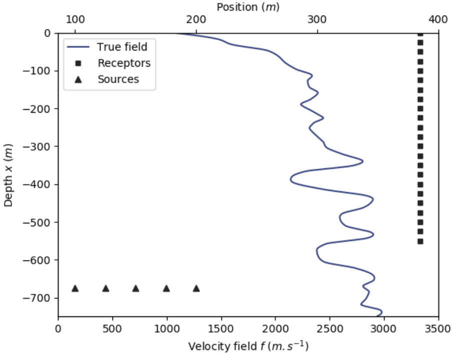
Finite numerical inaccuracy

For some $\boldsymbol{q}$, the magnitude of eigenmodes quickly decays.

(CoC) $b_{ij}(\boldsymbol{q})$ set to 0 if $\overline{\lambda}_r / \overline{\lambda_1} < \kappa \rightsquigarrow r = 15$, 7 modes are really inferred.

(CoM) choose r such that $\min_{\boldsymbol{q} \in \mathbb{H}} \lambda_r(\boldsymbol{q}) / \lambda_1(\boldsymbol{q}) > \kappa \rightsquigarrow r = 7$ sufficient to explain more than 99.8% of the field variance.

Application to seismic tomography

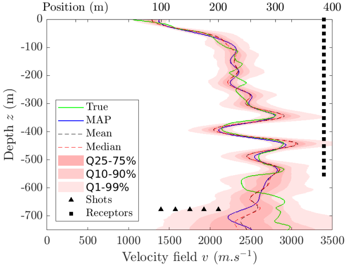


Application case: 1D section of Amoco model [O'Brien et al., 1994] and location of stations

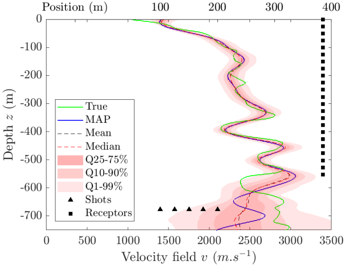
$\boldsymbol{d}^{\mathrm{obs}}$: time of arrival, with noise level $\alpha = 0.001s$

$r = 20$, $\boldsymbol{q} = \{A, \ell\}$

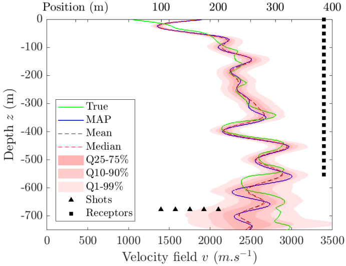
Application to seismic tomography



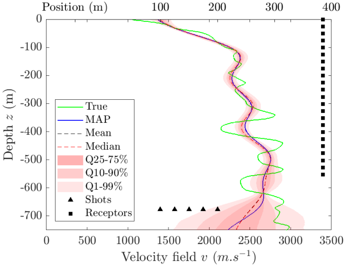
(a) Reference basis



(b) $\ell_{MAP} = 34$



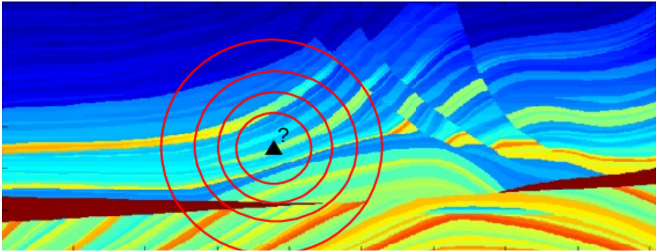
(c) $\ell = 10$



(d) $\ell = 80$

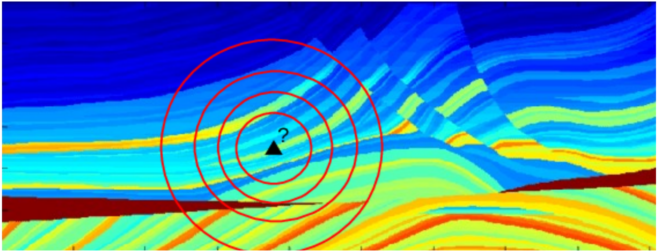
Conclusion

- Inference method allowing uncertainties estimations while remaining tractable
- WIP: comparison of results for transient diffusion equation, draft article
- Next: Reinference using a posteriori as prior; Extension to source location by using EOF



Conclusion

- Inference method allowing uncertainties estimations while remaining tractable
- WIP: comparison of results for transient diffusion equation, draft article
- Next: Reinference using a posteriori as prior; Extension to source location by using EOF



Thank you !

References I

K. Karhunen. "Zur Spektraltheorie Stochastischer Prozesse". In: *Annales Academiae Scientiarum Fennicae* (1946).

P. Lailly, F. Rocca, and R. Versteeg. "Synthesis of the Marmousi workshop". In: *The Marmousi Experience*. 1991, pp. 169–194.

M. Loève. *Probability Theory I*. Vol. 45. Graduate Texts in Mathematics. New York, NY: Springer New York, 1977. ISBN: 978-1-4684-9466-2 978-1-4684-9464-8. DOI: 10.1007/978-1-4684-9464-8.

Youssef Marzouk and Habib Najm. "Dimensionality reduction and polynomial chaos acceleration of Bayesian inference in inverse problems". In: *Journal of Computational Physics* 228 (Apr. 2009), pp. 1862–1902. DOI: 10.1016/j.jcp.2008.11.024.

N. Metropolis et al. "Equation of State Calculations by Fast Computing Machines". In: *The Journal of Chemical Physics* 21.6 (1953), pp. 1087–1092. DOI: 10.1063/1.1699114.

References II

- M. Noble and A. Gesret. *FTeik2d_3.1*. 2011.
- M. O'Brien and C. Regone. *Amoco*. 1994.
- A. Siripatana et al. "Bayesian Inference of Spatially Varying Manning's n Coefficients in an Idealized Coastal Ocean Model Using a Generalized Karhunen-Loève Expansion and Polynomial Chaos". In: *Ocean Dynamics* 70 (2020). ISSN: 1616-7228. DOI: 10.1007/s10236-020-01382-4.
- I. Sraj et al. "Coordinate Transformation and Polynomial Chaos for the Bayesian Inference of a Gaussian Process with Parametrized Prior Covariance Function". In: *Computer Methods in Applied Mechanics and Engineering* 298 (2016), pp. 205–228. ISSN: 0045-7825. DOI: 10.1016/j.cma.2015.10.002.