Bayesian Field Inversion with Hyperparameters Estimation

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

MCMC2023 - June 2023

Context

Detection and analysis of seismic events

Global scale

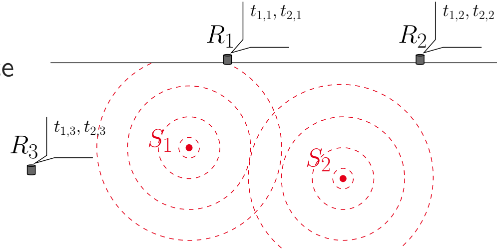
- International treaties (CTBT, NTP)
- Environment monitoring (IMS)

Regional scale

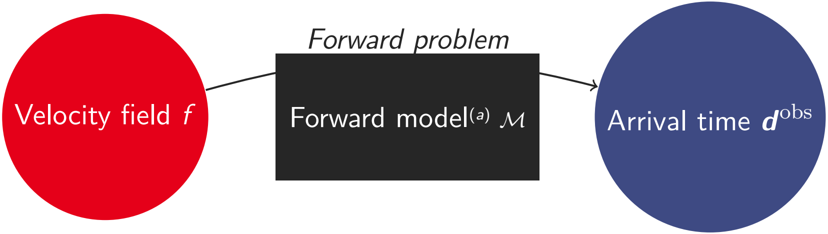
- Tsunami and seism alerts
- Risk prevention

Local scale

- Knowledge of subsurface
- Exploitation

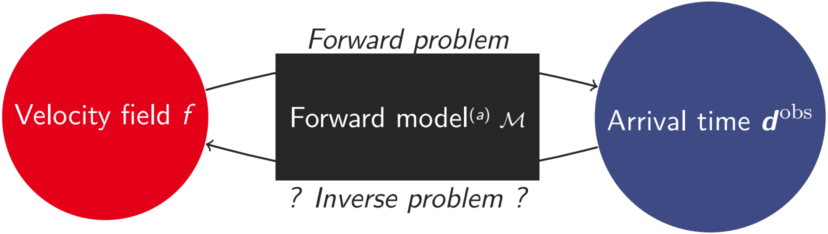


Context: seismic tomography



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

Context: seismic tomography



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

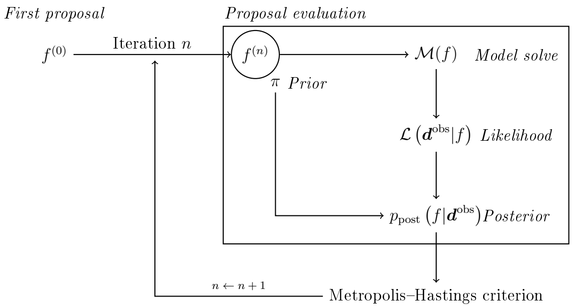
Objective: Estimation of a field (*i*) accurate,
(*ii*) with uncertainties,
(*iii*) fast.

Table of contents

Bayes formulation

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

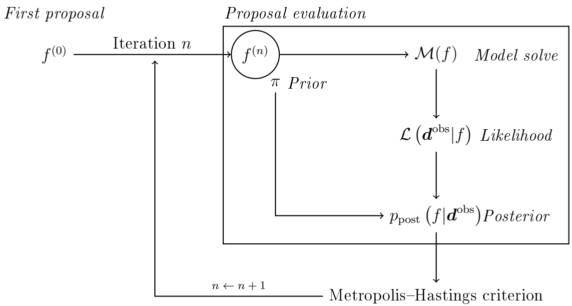
Markov Chain Monte–Carlo algorithm:



Bayes formulation

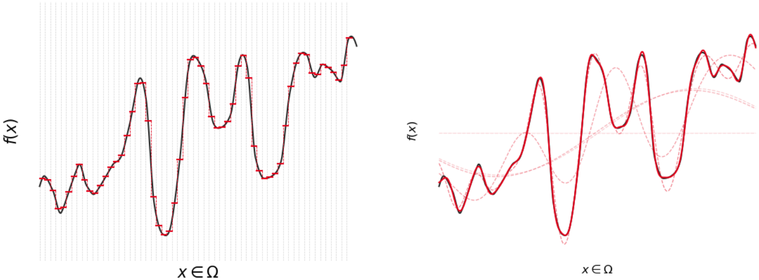
Bayes rule: $p_{\text{post}}(f|\boldsymbol{d}^{\text{obs}}) \propto \mathcal{L}(\boldsymbol{d}^{\text{obs}}|f)\pi_{\mathcal{F}}(f).$

Markov Chain Monte–Carlo algorithm:



$\Rightarrow$ Representation of f ?
Evaluation of $\mathcal{M}$?
$\rightarrow$ Polynomial chaos surrogate [Marzouk et al., 2009].

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 as 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, $\langle u_i, u_j \rangle_\Omega = \delta_{i,j}$,
  - $\boldsymbol{\eta} := (\eta_i)_{1 \leqslant i \leqslant r} \sim \mathcal{N}(0, \mathrm{I}_r)$

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

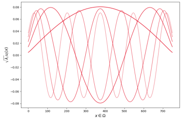
Dependency on hyperparameters

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

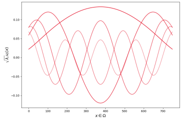
$$f(\boldsymbol{x}) \simeq \sum_{i=1}^{r} \lambda_i^{1/2}(\boldsymbol{q}) u_i(\boldsymbol{x}, \boldsymbol{q}) \eta_i(\theta), \text{ with } \eta_i = \lambda_i^{-1/2}(\boldsymbol{q}) \langle u_i(\cdot, \boldsymbol{q}), \mathcal{G} \rangle_\Omega$$

Exemple (squared exponential kernel)

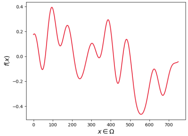
$$k(x, y, \boldsymbol{q} := \{A, \ell\}) = A \exp\left(-\frac{\|x - y\|^2}{2\ell^2}\right)$$



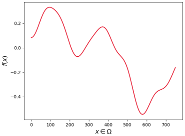
(a) Some modes for $\ell = 20$



(b) Some modes for $\ell = 60$



(c) Field obtained with $\ell = 20$



(d) Field obtained with $\ell = 60$

$\Rightarrow$ Exploration of hyperparameters space

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

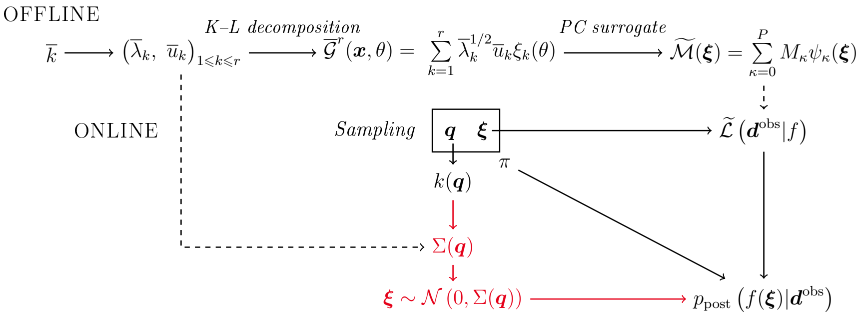
$$f(x) \simeq \overline{\mathcal{G}}^r(\boldsymbol{x}, \theta) := \sum_{i=1}^{r} \overline{\lambda}_i^{1/2} \overline{u}_i(\boldsymbol{x}) \xi_i(\theta) \text{ with } \boldsymbol{\xi} \sim \mathcal{N}(0, \Sigma(\boldsymbol{q}))$$

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

The covariance matrix $\Sigma(\boldsymbol{q})$ writes

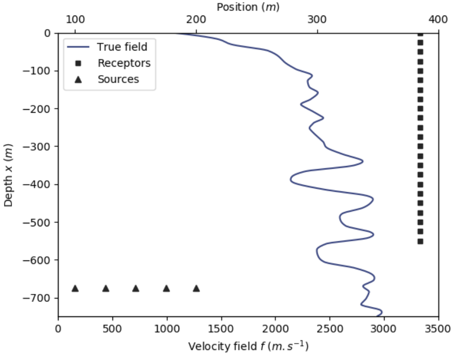
$$\forall 1 \leqslant i,j \leqslant r,\ \forall \boldsymbol{q} \in \mathbb{H}, \qquad \Sigma_{ij}(\boldsymbol{q}) := (\overline{\lambda}_i \overline{\lambda}_j)^{-1/2} \left\langle \langle k(\cdot,\cdot,\boldsymbol{q}), \overline{u}_j \rangle_\Omega, \overline{u}_i \right\rangle_\Omega .$$

Workflow



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

Application to seismic tomography

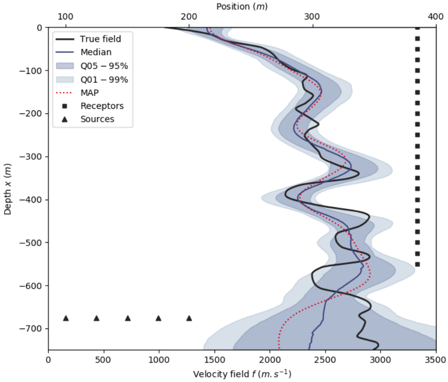


Application case: 1D section of Amoco model [O'Brien et al., 1994] and location of stations

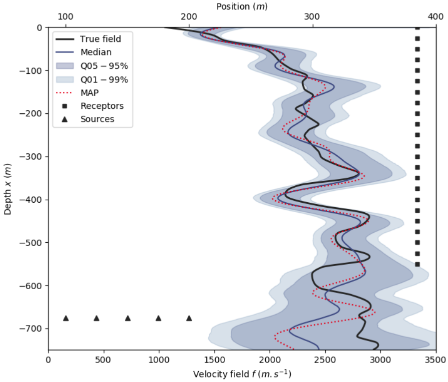
$\boldsymbol{d}^{\text{obs}}$: time of arrival, with noise level $\alpha = 0.002s$

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

Application to seismic tomography



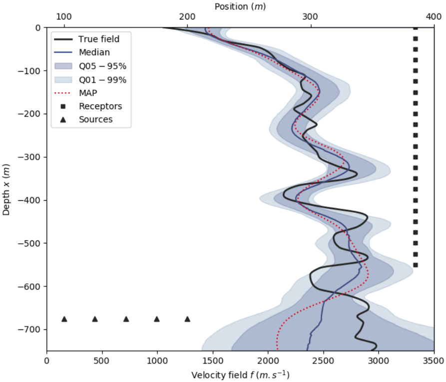
(a) Proposed method



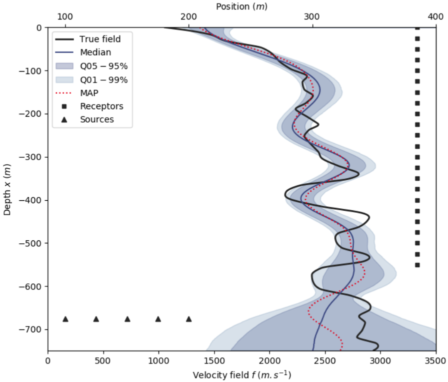
(b) $\ell = 20$

Comparison of inference results for different bases

Application to seismic tomography



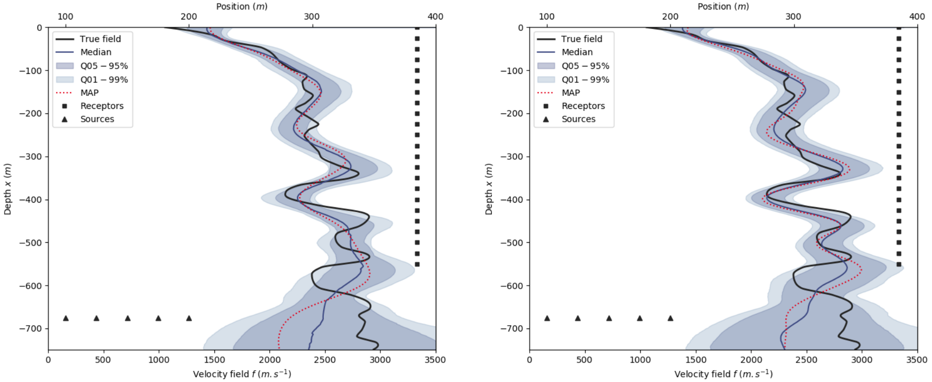
(a) Proposed method



(b) $\ell = 60$

Comparison of inference results for different bases

Application to seismic tomography



(a) Proposed method

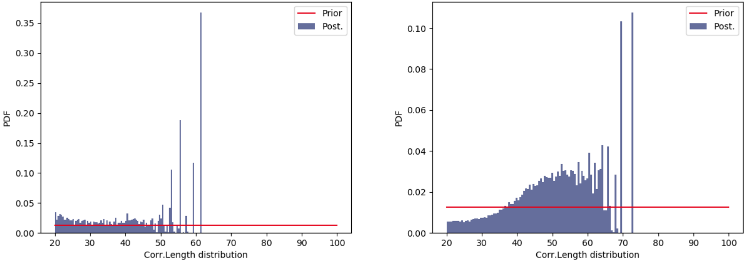
(b) $\ell_{\text{best}} = 36$

Comparison of inference results for different bases

Application to seismic tomography

Sampling joint law $(\boldsymbol{\xi}, \boldsymbol{q})$...

Problem : MCMC fails to draw prior hyperparameters distribution



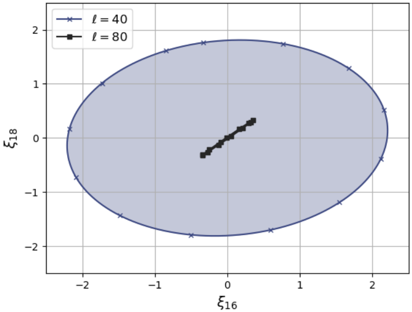
(a) Marginalized ℓ prior samples via MCMC

(b) Marginalized ℓ posterior samples via MCMC

Table of contents

Sampling problem

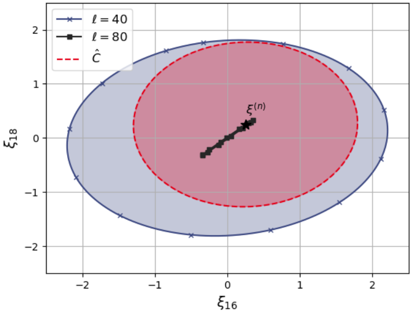
$\boldsymbol{\xi} \sim \mathcal{N}(0, \Sigma(\boldsymbol{q}))$: *The prior law of $\boldsymbol{\xi}$ is highly stretched along $\boldsymbol{q}$*



Prior law of $\boldsymbol{\xi}$ according to $\boldsymbol{q}$ (90%-confidence contour)

Sampling problem

$\boldsymbol{\xi} \sim \mathcal{N}(0, \Sigma(\boldsymbol{q}))$: *The prior law of $\boldsymbol{\xi}$ is highly stretched along $\boldsymbol{q}$*



Prior law of $\boldsymbol{\xi}$ according to $\boldsymbol{q}$ (90%-confidence contour) and covariance proposal

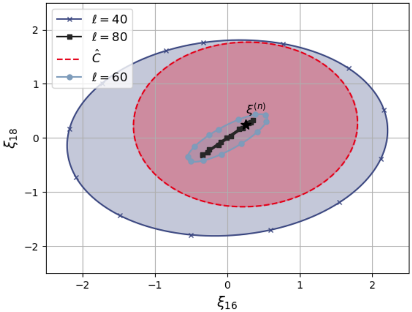
Initial sampling : Metropolis–Hastings random walk with adapted covariance proposal

$$Y^\star := (\boldsymbol{\xi}^\star, \boldsymbol{q}^\star) \sim \mathcal{N}(Y^{(n)}, \widehat{C}),$$

$$\text{with } \widehat{C} \propto \mathbb{C}\mathrm{ov}\left(Y^{(1)}, \ldots, Y^{(n)}\right)$$

Sampling problem

$\boldsymbol{\xi} \sim \mathcal{N}(0, \Sigma(\boldsymbol{q}))$: *The prior law of $\boldsymbol{\xi}$ is highly stretched along $\boldsymbol{q}$*



Prior law of $\boldsymbol{\xi}$ according to $\boldsymbol{q}$ (90%-confidence contour) and covariance proposal

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

$$Y^\star := (\boldsymbol{\xi}^\star, \boldsymbol{q}^\star) \sim \mathcal{N}(Y^{(n)}, \widehat{C}),$$

with $\widehat{C} \propto \mathbb{C}\text{ov}\left(Y^{(1)}, \dots, Y^{(n)}\right)$

$\Rightarrow$ The random walk with adapted covariance proposal $\widehat{C}$ needs to be adjusted along $\boldsymbol{q}$.

Presentation of adaptation

We want to consider a *$\boldsymbol{q}$-dependent sampling for ξ*.

$$\boldsymbol{q}^\star \sim \mathcal{N}(\boldsymbol{q}^{(n)}, \widehat{C}_{\boldsymbol{q}}), \text{ with } \widehat{C}_{\boldsymbol{q}} \propto \mathbb{C}\mathrm{ov}\left(\boldsymbol{q}^{(1)}, \dots, \boldsymbol{q}^{(n)}\right),$$

$$\xi^\star \sim \mathcal{N}(\xi^{(n)}, S(\boldsymbol{q}^\star)), \text{ where } S(\boldsymbol{q}) \text{ remains to define.}$$

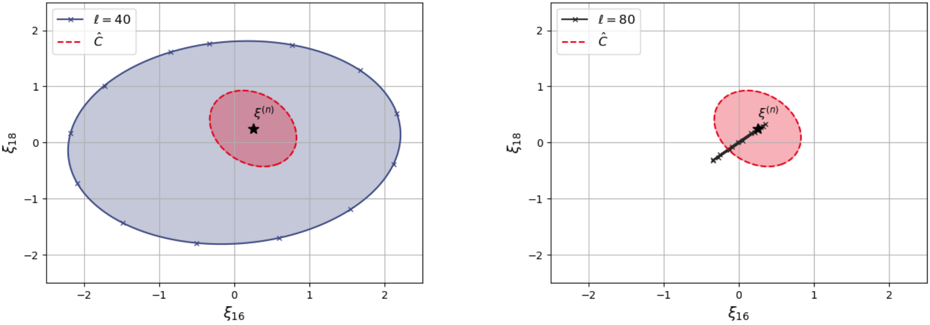
Three ideas are explained here:

- a simple *scaling*
- the use of the *conditional covariance*
- the use of a change of coordinates *preserving measure*

(1) Rescaling the covariance proposal

Objective: *restrict the movement along unfeasible directions.*
At each iteration,

- consider current scaled empirical covariance proposal $\widehat{C}$



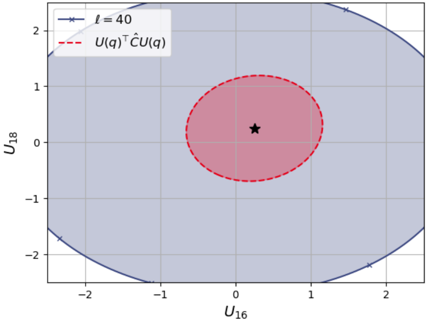
(a) $\ell = 40$ (b) $\ell = 80$

Illustration of the scaling process for two values of ℓ

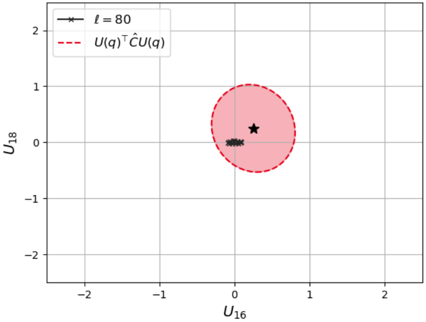
(1) Rescaling the covariance proposal

Objective: *restrict the movement along unfeasible directions.*
At each iteration,

- change of basis according to the prior $\Sigma(\boldsymbol{q})$ eigenelements



(a) $\ell = 40$



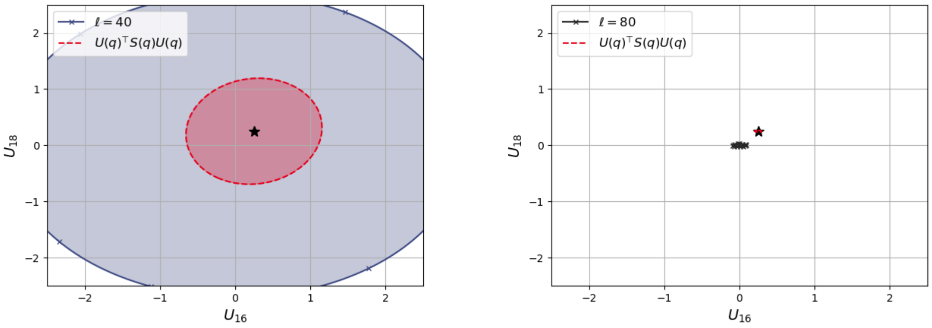
(b) $\ell = 80$

Illustration of the scaling process for two values of ℓ

(1) Rescaling the covariance proposal

Objective: *restrict the movement along unfeasible directions.*
At each iteration,

- scaling along eigendirections



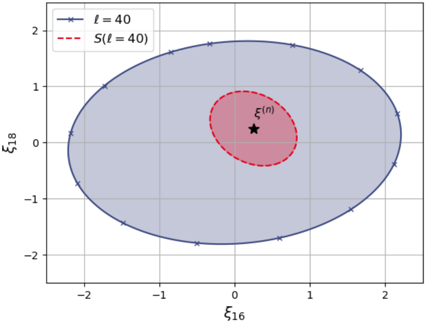
(a) $\ell = 40$ (b) $\ell = 80$

Illustration of the scaling process for two values of ℓ

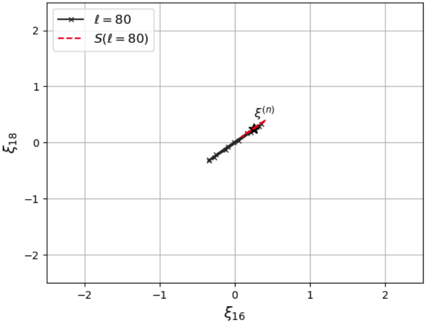
(1) Rescaling the covariance proposal

Objective: *restrict the movement along unfeasible directions.*
At each iteration,

- return to physical space: propose according $S(\boldsymbol{q})$



(a) $\ell = 40$



(b) $\ell = 80$

Illustration of the scaling process for two values of ℓ

(1) Rescaling the covariance proposal

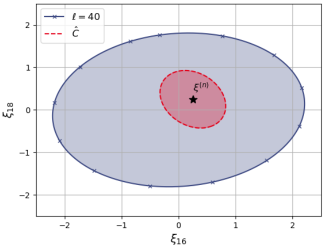
Objective: *restrict the movement along unfeasible directions*,

$$\forall u \in \mathbb{R}^r, \qquad \mathbb{V}_{S(\boldsymbol{q})}(u) \lesssim \mathbb{V}_{\Sigma(\boldsymbol{q})}(u).$$

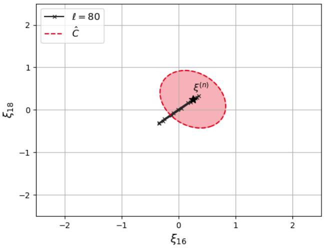
Approximation: *focus on the prior covariance eigendirections*.
$S(\boldsymbol{q})$ is the scaled empirical covariance of the $\boldsymbol{\xi}$ samples $\propto \mathbb{C}\mathrm{ov}(\boldsymbol{\xi}^{(1)}, \dots, \boldsymbol{\xi}^{(n)})$, scaled as follows

$$\text{We want that, } \forall 1 \leqslant i \leqslant r, \qquad \mathbb{V}_{S(\boldsymbol{q})}(U_i(\boldsymbol{q})) \lesssim \mathbb{V}_{\Sigma(\boldsymbol{q})}(U_i(\boldsymbol{q})),$$

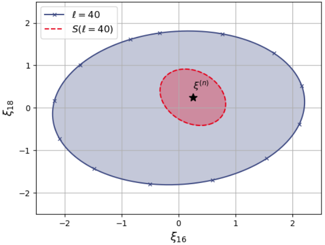
where $(U_i(\boldsymbol{q}), \Lambda_{ii}(\boldsymbol{q}))$ are the eigenelements of the prior covariance $\Sigma(\boldsymbol{q})$.



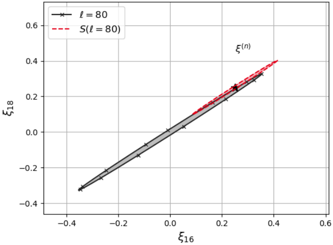
(a) $\ell = 40,\ \widehat{C}$



(b) $\ell = 80,\ \widehat{C}$

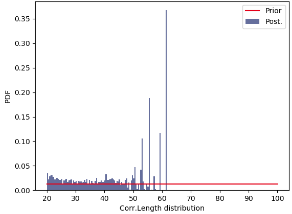


(c) $\ell = 40,\ S(\boldsymbol{q})$

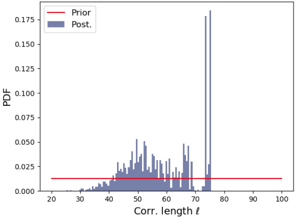


(d) $\ell = 80,\ S(\boldsymbol{q})$ (zoom)

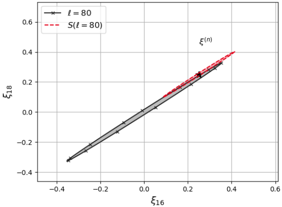
(1) Rescaling the covariance proposal



(a) Marginalized ℓ prior samples via MCMC (no scaling)



(b) Marginalized ℓ prior samples via MCMC (scaling)



Remaining problems :

Scaling along prior eigendirections only + how to move in the right direction/set automatically to zero ?

(2) Consider $\widehat{\mathbb{C}\mathrm{ov}}(\boldsymbol{\xi}|\boldsymbol{q})$

$$\boldsymbol{\xi}^\star \sim \mathcal{N}(\boldsymbol{\xi}^{(n)}, S(\boldsymbol{q})) \text{ with } S(\boldsymbol{q}) \propto \widehat{\mathbb{C}\mathrm{ov}}(\boldsymbol{\xi}|\boldsymbol{q}^\star)$$

Polynomial chaos surrogate

$$\begin{aligned}\widehat{\mathbb{C}\mathrm{ov}}(\boldsymbol{\xi}|\boldsymbol{q}^\star) &= \widehat{\mathbb{E}}(\boldsymbol{\xi}\boldsymbol{\xi}^\top|\boldsymbol{q}) - \widehat{\mathbb{E}}(\boldsymbol{\xi}|\boldsymbol{q})\widehat{\mathbb{E}}(\boldsymbol{\xi}|\boldsymbol{q})^\top \\ &\simeq \sum_{\alpha=1}^{K} \widehat{y}_\alpha \phi_\alpha(\boldsymbol{q}) - \sum_{\alpha,\beta=1}^{K} \widehat{x}_\alpha \widehat{x}_\beta \phi_\alpha(\boldsymbol{q})\phi_\beta(\boldsymbol{q}),\end{aligned}$$

α, β : set of multi-indexes; $(\phi_\alpha)_\alpha$: orthonormal polynomials;
$\widehat{y}_\alpha, \widehat{x}_\alpha$: PC coefficients obtained via *least squares regression*

$$\widehat{x}_\alpha = \underset{x_\alpha}{\mathrm{argmin}} \left\| \widehat{\mathbb{E}}(\boldsymbol{\xi}|\boldsymbol{q}) - \sum_{\alpha=1}^{K} x_\alpha \phi_\alpha(\boldsymbol{q}) \right\|^2$$

$$\text{and } \widehat{y_\alpha} = \underset{y_\alpha}{\mathrm{argmin}} \left\| \widehat{\mathbb{E}}(\boldsymbol{\xi}\boldsymbol{\xi}^\top|\boldsymbol{q}) - \sum_{\alpha=1}^{K} y_\alpha \phi_\alpha(\boldsymbol{q}) \right\|^2$$

(2) Consider $\widehat{\mathbb{C}\mathrm{ov}}(\boldsymbol{\xi}|\boldsymbol{q})$

$$\boldsymbol{\xi}^\star \sim \mathcal{N}(\boldsymbol{\xi}^{(n)}, S(\boldsymbol{q})) \text{ with } S(\boldsymbol{q}) \propto \widehat{\mathbb{C}\mathrm{ov}}(\boldsymbol{\xi}|\boldsymbol{q}^\star)$$

Polynomial chaos surrogate

$$\begin{aligned}\widehat{\mathbb{C}\mathrm{ov}}(\boldsymbol{\xi}|\boldsymbol{q}^\star) &= \widehat{\mathbb{E}}(\boldsymbol{\xi}\boldsymbol{\xi}^\top|\boldsymbol{q}) - \widehat{\mathbb{E}}(\boldsymbol{\xi}|\boldsymbol{q})\widehat{\mathbb{E}}(\boldsymbol{\xi}|\boldsymbol{q})^\top \\ &\simeq \sum_{\alpha=1}^{K} \widehat{y}_\alpha \phi_\alpha(\boldsymbol{q}) - \sum_{\alpha,\beta=1}^{K} \widehat{x}_\alpha \widehat{x}_\beta \phi_\alpha(\boldsymbol{q})\phi_\beta(\boldsymbol{q}),\end{aligned}$$

α, β : set of multi-indexes; $(\phi_\alpha)_\alpha$: orthonormal polynomials;
$\widehat{y}_\alpha, \widehat{x}_\alpha$: PC coefficients obtained via *least squares regression*

$$\widehat{x}_\alpha = \underset{x_\alpha}{\mathrm{argmin}} \left\| \widehat{\mathbb{E}}(\boldsymbol{\xi}|\boldsymbol{q}) - \sum_{\alpha=1}^{K} x_\alpha \phi_\alpha(\boldsymbol{q}) \right\|^2$$

$$\text{and } \widehat{y_\alpha} = \underset{y_\alpha}{\mathrm{argmin}} \left\| \widehat{\mathbb{E}}(\boldsymbol{\xi}\boldsymbol{\xi}^\top|\boldsymbol{q}) - \sum_{\alpha=1}^{K} y_\alpha \phi_\alpha(\boldsymbol{q}) \right\|^2$$

(2) Consider $\widehat{\mathbb{C}\mathrm{ov}}(\boldsymbol{\xi}|\boldsymbol{q})$

$$\boldsymbol{\xi}^\star \sim \mathcal{N}(\boldsymbol{\xi}^{(n)}, S(\boldsymbol{q})) \text{ with } S(\boldsymbol{q}) \propto \widehat{\mathbb{C}\mathrm{ov}}(\boldsymbol{\xi}|\boldsymbol{q}^\star)$$

Polynomial chaos surrogate

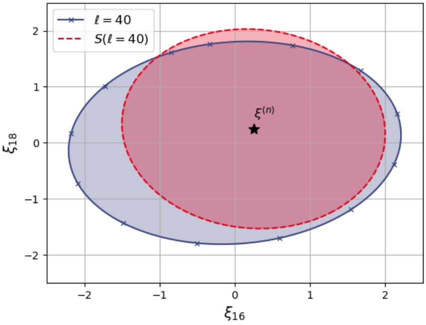
$$\widehat{\mathbb{C}\mathrm{ov}}(\boldsymbol{\xi}|\boldsymbol{q}^\star) = \widehat{\mathbb{E}}(\boldsymbol{\xi}\boldsymbol{\xi}^\top|\boldsymbol{q}) - \widehat{\mathbb{E}}(\boldsymbol{\xi}|\boldsymbol{q})\widehat{\mathbb{E}}(\boldsymbol{\xi}|\boldsymbol{q})^\top$$
$$\simeq \sum_{\alpha=1}^{K} \widehat{y}_\alpha \phi_\alpha(\boldsymbol{q}) - \sum_{\alpha,\beta=1}^{K} \widehat{x}_\alpha \widehat{x}_\beta \phi_\alpha(\boldsymbol{q})\phi_\beta(\boldsymbol{q}),$$

α, β : set of multi-indexes; $(\phi_\alpha)_\alpha$: orthonormal polynomials;
$\widehat{y}_\alpha, \widehat{x}_\alpha$: PC coefficients obtained via *least squares regression*

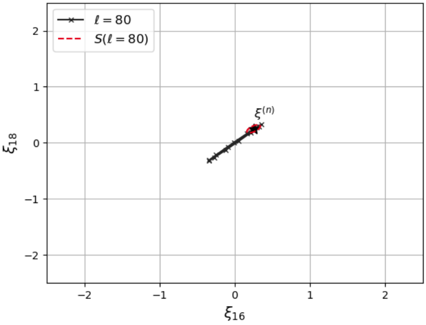
$$\widehat{x}_\alpha = \underset{x_\alpha}{\mathrm{argmin}} \sum_{i=1}^{N} \left\| \boldsymbol{\xi}^{(i)} - \sum_{\alpha=1}^{K} x_\alpha \phi_\alpha(\boldsymbol{q}^{(i)}) \right\|^2$$

$$\text{and } \widehat{y_\alpha} = \underset{y_\alpha}{\mathrm{argmin}} \sum_{i=1}^{N} \left\| \boldsymbol{\xi}^{(i)} \boldsymbol{\xi}^{(i)\top} - \sum_{\alpha=1}^{K} y_\alpha \phi_\alpha(\boldsymbol{q}^{(i)}) \right\|^2$$

(2) Consider $\widehat{\mathbb{C}\mathrm{ov}}(\boldsymbol{\xi}|\boldsymbol{q})$



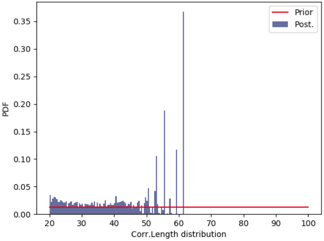
(a) $\ell = 40$



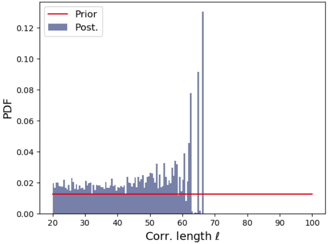
(b) $\ell = 80$

Resulting covariance proposal after burn-in phase

(2) Consider $\widehat{\mathbb{C}\mathrm{ov}}(\boldsymbol{\xi}|\boldsymbol{q})$



(a) With $\widehat{C}$



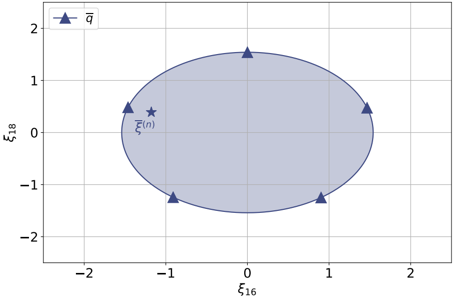
(b) With $\widehat{\mathbb{C}\mathrm{ov}}(\boldsymbol{\xi}|\boldsymbol{q})$

Marginalized ℓ prior samples via MCMC

- Surrogate building: expensive + order choice + decomposition at each step
- Surrogate can be non pertinent at non sampled space
- Still remains the problem of setting automatically to zero

(3) Sampling preserving measure

Idea : We want to sample from a unique distribution: sample $\overline{\boldsymbol{\xi}} \sim \mathcal{N}(0, \Sigma(\overline{\boldsymbol{q}}))$.

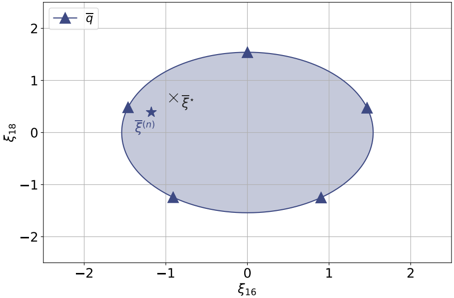


At each iteration:

- we know $(\overline{\boldsymbol{\xi}}^{(n)}, \boldsymbol{q}^{(n)})$

(3) Sampling preserving measure

Idea : We want to sample from a unique distribution: sample $\overline{\boldsymbol{\xi}} \sim \mathcal{N}(0, \Sigma(\overline{\boldsymbol{q}}))$.

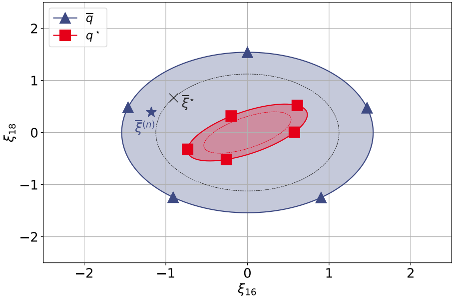


At each iteration:

- we know $(\overline{\boldsymbol{\xi}}^{(n)}, \boldsymbol{q}^{(n)})$
- Draw $\overline{\boldsymbol{\xi}}^{\star} \sim \mathcal{N}(\overline{\boldsymbol{\xi}}, K)$

(3) Sampling preserving measure

Idea : We want to sample from a unique distribution: sample $\overline{\boldsymbol{\xi}} \sim \mathcal{N}(0, \Sigma(\overline{\boldsymbol{q}}))$.

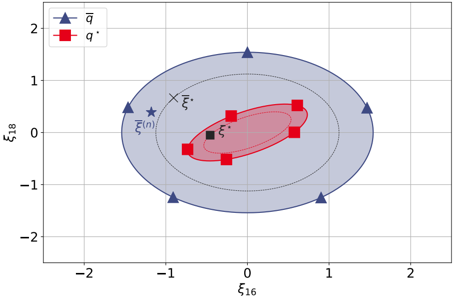


At each iteration:

- we know $(\overline{\boldsymbol{\xi}}^{(n)}, \boldsymbol{q}^{(n)})$
- Draw $\overline{\boldsymbol{\xi}}^\star \sim \mathcal{N}(\overline{\boldsymbol{\xi}}, K)$
- Draw $\boldsymbol{q}^\star \sim \mathcal{N}(\boldsymbol{q}, \widehat{C}_{\boldsymbol{q}})$

(3) Sampling preserving measure

Idea : We want to sample from a unique distribution: sample $\overline{\boldsymbol{\xi}} \sim \mathcal{N}(0, \Sigma(\overline{\boldsymbol{q}}))$.

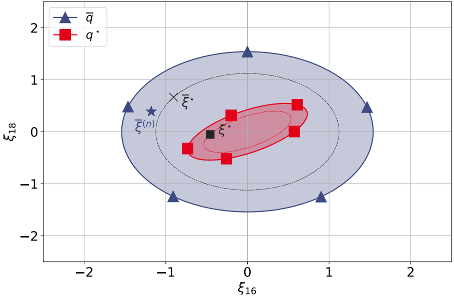


At each iteration:

- we know $(\overline{\boldsymbol{\xi}}^{(n)}, \boldsymbol{q}^{(n)})$
- Draw $\overline{\boldsymbol{\xi}}^\star \sim \mathcal{N}(\overline{\boldsymbol{\xi}}, K)$
- Draw $\boldsymbol{q}^\star \sim \mathcal{N}(\boldsymbol{q}, \widehat{C}_{\boldsymbol{q}})$
- $\boldsymbol{\xi}^\star = \mathrm{QCQO}(\overline{\boldsymbol{\xi}}^\star, \overline{\boldsymbol{q}} \to \boldsymbol{q}^\star)$

(3) Sampling preserving measure

Idea : We want to sample from a unique distribution: sample $\overline{\boldsymbol{\xi}} \sim \mathcal{N}(0, \Sigma(\overline{\boldsymbol{q}}))$.



At each iteration:

- we know $(\overline{\boldsymbol{\xi}}^{(n)}, \boldsymbol{q}^{(n)})$
- Draw $\overline{\boldsymbol{\xi}}^\star \sim \mathcal{N}(\overline{\boldsymbol{\xi}}, K)$
- Draw $\boldsymbol{q}^\star \sim \mathcal{N}(\boldsymbol{q}, \widehat{C}_{\boldsymbol{q}})$
- $\boldsymbol{\xi}^\star = \mathrm{QCQO}(\overline{\boldsymbol{\xi}}^\star, \overline{\boldsymbol{q}} \to \boldsymbol{q}^\star)$
- Compute $p_{\mathrm{post}}\left(f(\boldsymbol{\xi}^\star) | \boldsymbol{d}^{\mathrm{obs}}\right)$
- Metropolis–Hastings criterion to accept or not
- New iteration

(3) Sampling preserving measure

✗ Change of coordinates approach [Sraj et al., 2016]: $\boldsymbol{\xi}$ is seen as $\boldsymbol{\xi}(\boldsymbol{q}) = B(\boldsymbol{q})\eta$

- *Idea :* changing $\boldsymbol{q}$ should affect $\boldsymbol{\xi}$ in the sense that we want
  - to preserve the prior Mahalanobis distance
  - to induce the slightest change in terms of field variations
- $\mathrm{QCQO}(\overline{\boldsymbol{\xi}}^{\star}, \overline{\boldsymbol{q}} \rightarrow \boldsymbol{q}^{\star})$:

$$\boldsymbol{\xi}^{\star} = \underset{\eta}{\operatorname{argmin}} \left\| f(\overline{\boldsymbol{\xi}}^{\star}) - f(\eta) \right\|^2 = \underset{\eta}{\operatorname{argmin}}\, \overline{\boldsymbol{\xi}}^{\star\top} \overline{\boldsymbol{\lambda}} \eta \qquad \text{(field distance)}$$

$$\text{s.t. } \overline{\boldsymbol{\xi}}^{\star\top} \Sigma(\overline{\boldsymbol{q}})^{-1} \overline{\boldsymbol{\xi}}^{\star} = \eta^{\top} \Sigma(\boldsymbol{q}^{\star})^{-1} \eta \qquad \text{(prior distance)}$$

(3) Sampling preserving measure

✗ Change of coordinates approach [Sraj et al., 2016]: $\boldsymbol{\xi}$ is seen as $\boldsymbol{\xi}(\boldsymbol{q}) = B(\boldsymbol{q})\eta$

- *Idea :* changing $\boldsymbol{q}$ should affect $\boldsymbol{\xi}$ in the sense that we want
  - to preserve the prior Mahalanobis distance
  - to induce the slightest change in terms of field variations
- $\mathrm{QCQO}(\overline{\boldsymbol{\xi}}^{\star}, \overline{\boldsymbol{q}} \to \boldsymbol{q}^{\star})$:

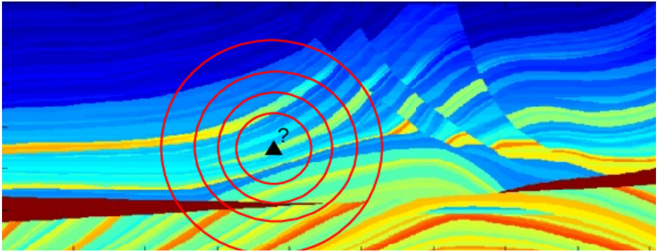
$$\boldsymbol{\xi}^{\star} = \underset{\eta}{\operatorname{argmin}} \left\| f(\overline{\boldsymbol{\xi}}^{\star}) - f(\eta) \right\|^2 = \underset{\eta}{\operatorname{argmin}}\, \overline{\boldsymbol{\xi}}^{\star\top} \overline{\boldsymbol{\lambda}} \eta \qquad \text{(field distance)}$$

$$\text{s.t. } \overline{\boldsymbol{\xi}}^{\star\top} \Sigma(\overline{\boldsymbol{q}})^{-1} \overline{\boldsymbol{\xi}}^{\star} = \eta^{\top} \Sigma(\boldsymbol{q}^{\star})^{-1} \eta \qquad \text{(prior distance)}$$

- Symmetric proposal
- Decomposition of the covariance $\rightsquigarrow$ QCQO problem solving

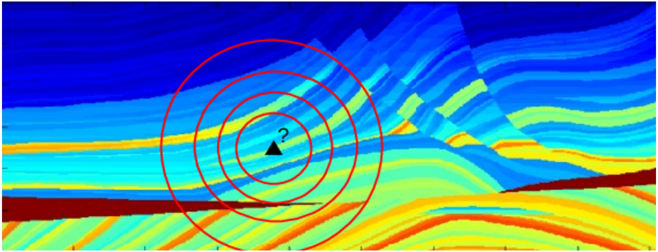
Conclusion & perspectives

- Other possible approaches : splitted sampling, use of derivatives (HMC)...
- Change of measure allows uncertainties estimation while remaining tractable
- Next implementation : QCQO, reinference
- Next application : extend to source location by using EOF



Conclusion & perspectives

- Other possible approaches : splitted sampling, use of derivatives (HMC)...
- Change of measure allows uncertainties estimation while remaining tractable
- Next implementation : QCQO, reinference
- Next application : extend to source location by using EOF



Thank you !

nadege.polette@minesparis.psl.eu

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.

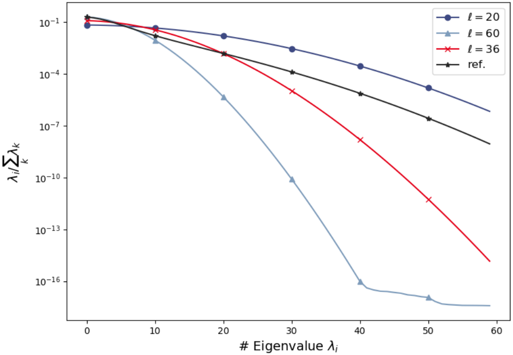
References II

M. Noble and A. Gesret. *FTeik2d_3.1*. 2011.

M. O'Brien and C. Regone. *Amoco*. 1994.

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.

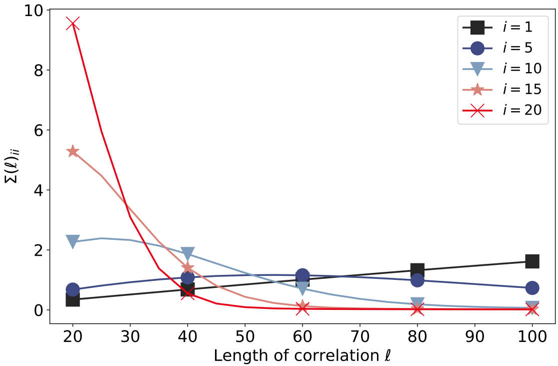
Eigenvalues according kernel



Decreasing of eigenvalues according to length of correlation considered

$\Rightarrow$ the higher the length of correlation, the smaller the number of modes needed to explain the field: last coordinates are likely to be close to zero.

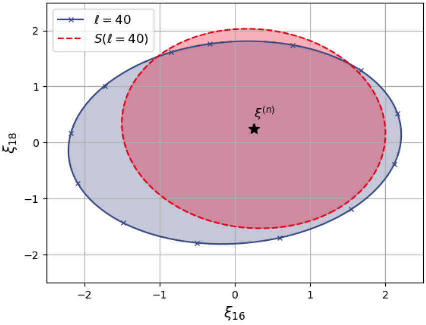
Variance according hyperparameter



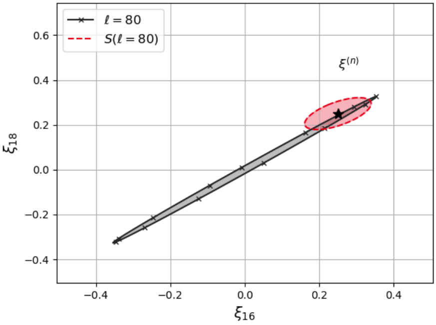
Variance of K–L coordinates ξ_i according to ℓ.

$\Rightarrow$ the higher the length of correlation, the smaller the variance of last coordinates: they are likely to be close to zero.

Results $\widehat{\mathbb{C}\mathrm{ov}}(\boldsymbol{\xi}|\boldsymbol{q})$ - zoom



(a) $\ell = 40$



(b) $\ell = 80$ (zoom)

Resulting covariance proposal after burn-in phase