Hyperparameters Bayesian inference for Inverse Problems

POLETTE Nadège*†, SOCHALA Pierre*, GESRET Alexandrine†, LE MAÎTRE Olivier+

*CEA DAM DIF, F-91297 Arpajon † Mines Paris, PSL University, Geosciences center, Fontainebleau
+ CMAP, CNRS, École Polytechnique, IPP, Palaiseau (France)

Context

The fight against nuclear proliferation consists in the first place in overseeing the respect of the international treaties. In that context, numerical tools are developed in order to monitor the environment and ***analyze seismic events***. In particular, when an earthquake occurs, the displacement of the ground is recorded by seismometers. The accurate identification of an event can be difficult, since it requires to solve an ***inverse problem*** where available observations and velocity field models used for forward simulations are ***uncertain***. We focus on the velocity field uncertainty quantification.

Set-up of the problem

- Goal: characterization of a physical field f thanks to indirect observations $\boldsymbol{d}^{\text{obs}}$ knowing a forward model $\mathcal{M}$.

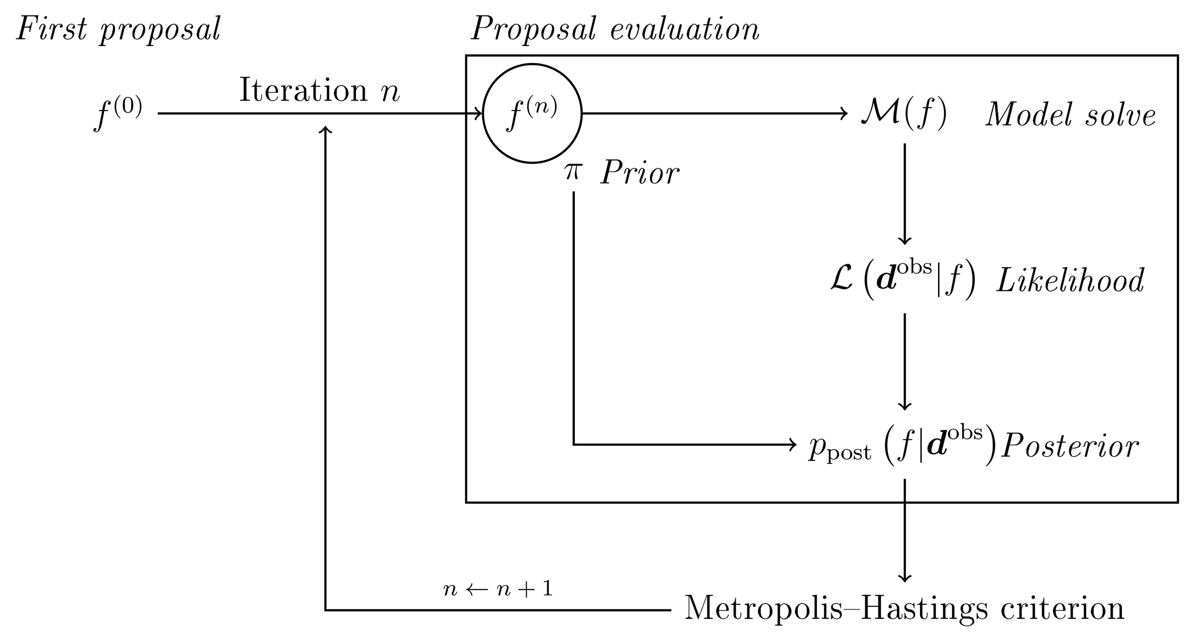
$$\boldsymbol{d}^{\text{obs}} = \mathcal{M}(f) + \varepsilon$$

- Probabilistic framework

Bayes rule

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

- Metropolis–Hastings algorithm



⇒ Difficulties: (i) infinite dimensional character of f
→ Karhunen–Loève decomposition [2, 3],
(ii) evaluation cost of $\mathcal{M}$ at each step
→ Polynomial chaos surrogate [1, 4].

- $f \in L^2(\Omega)$: particular realization of a Gaussian process $\mathcal{G} \sim \mathcal{N}(\mu, k)$

Karhunen–Loève decomposition

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

where $(\lambda_i, u_i)_{i\in\mathbb{N}^*}$ are the eigenelements of k and $\boldsymbol{\eta} \sim \mathcal{N}(0,1)$.

- The posterior quantity becomes

$$p_{\text{post}}\left(f = \mu + \sum_{i=1}^{r} \lambda_i^{1/2} u_i \eta_i \,\middle|\, \boldsymbol{d}^{\text{obs}}\right) \propto \widetilde{\mathcal{L}}(\boldsymbol{d}^{\text{obs}}|f)\pi(\boldsymbol{\eta}).$$

Change of measure method

In fact, k depends on hyperparameters $\boldsymbol{q}$ that are difficult to choose *a priori*:

$$p_{\text{post}}\left(f(\boldsymbol{q}) = \mu + \sum_{i=1}^{r} \lambda_i^{1/2}(\boldsymbol{q}) u_i(\boldsymbol{q}) \eta_i \,\middle|\, \boldsymbol{d}^{\text{obs}}\right) \propto \widetilde{\mathcal{L}}(\boldsymbol{d}^{\text{obs}}|f(\boldsymbol{q}))\pi(\boldsymbol{\eta}, \boldsymbol{q}).$$

We present a method to explore the hyperparameters space during the inference, while mitigating the computational cost.

- Introduction of a ***reference basis*** obtained from an average procedure over the hyperparameters space prior [6]:

$$f(x) = \mathcal{G}(x,\theta) \simeq \mu + \sum_{i=1}^{r} \overline{\lambda}_i^{1/2} \overline{u}_i(x)\xi_i(\theta), \text{ with } \xi_i(\theta) = \left\langle \mathcal{G}(x,\theta), \overline{\lambda}_i^{-1/2} \overline{u}_i \right\rangle_\Omega,$$

where $(\overline{\lambda}_i, \overline{u}_i)_{i\in\mathbb{N}^*}$ are the eigenelements of $\overline{k} = \mathbb{E}_{\mathbb{H}}(k(\boldsymbol{q})) = \int_{\mathbb{H}} k(\boldsymbol{q})\pi(\boldsymbol{q})d\boldsymbol{q}$.

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

$$p_{\text{post}}\left(f = \mu + \sum_{i=1}^{r} \overline{\lambda}_i^{1/2} \overline{u}_i \xi_i \,\middle|\, \boldsymbol{d}^{\text{obs}}\right) \propto \widetilde{\mathcal{L}}(\boldsymbol{d}^{\text{obs}}|f)\pi(\boldsymbol{\xi}(\boldsymbol{q}))\pi(\boldsymbol{q}).$$

Law of reference coordinates

$$\boldsymbol{\xi} \sim \mathcal{N}(0, \Sigma(\boldsymbol{q})),$$

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

- Σ is smooth along $\boldsymbol{q}$: possible surrogate $\widetilde{\Sigma}$

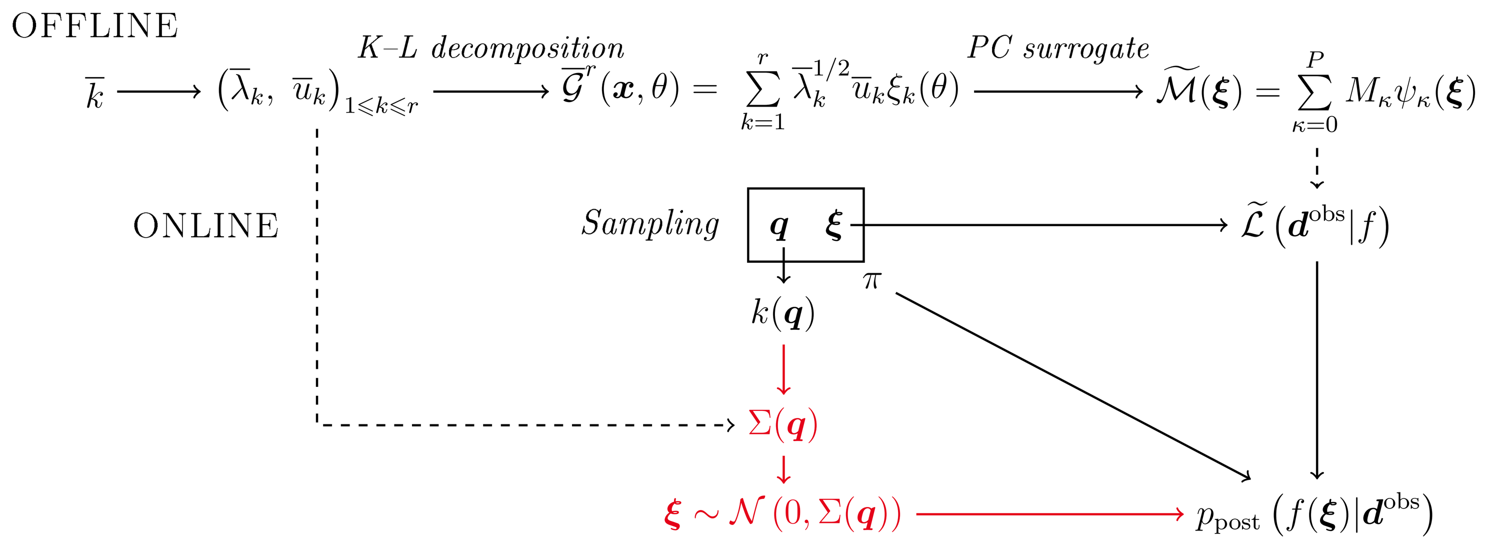


Fig. 1 – Worflow for the change of measure method

References

[1] R. G. Ghanem et al. "Stochastic Finite Element Method: Response Statistics". In: Stochastic Finite Elements: A Spectral Approach. Springer New York, 1991.
[2] K. Karhunen. "Zur Spektraltheorie Stochastischer Prozesse". In: Annales Academiae Scientiarum Fennicae (1946).
[3] M. Loève. Probability Theory I. Graduate Texts in Mathematics. Springer New York, 1977.
[4] Youssef Marzouk et al. "Dimensionality reduction and polynomial chaos acceleration of Bayesian inference in inverse problems". In: Journal of Computational Physics (2009).
[5] P. Sochala et al. "Polynomial surrogates for Bayesian traveltime tomography". In: GEM - International Journal on Geomathematics (2021).
[6] 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 (2016).

We thank M. Noble for the 2D eikonal solver.

Application to seismic tomography

The method is applied on a seismic tomography problem. We consider a 1D continuous velocity field inspired by the realistic model Amoco Tulsa Research Lab (O'Brien, 1994) and generalizing layered fields [5]. The posterior distribution of the velocity field is plotted in Fig.2, and is compared to the results of inferences with fixed hyperparameters (Fig.3(a)-(b)-(c)). We observe that, for a close computational cost, the exploration of the hyperparameters space yields a better estimation of uncertainties.

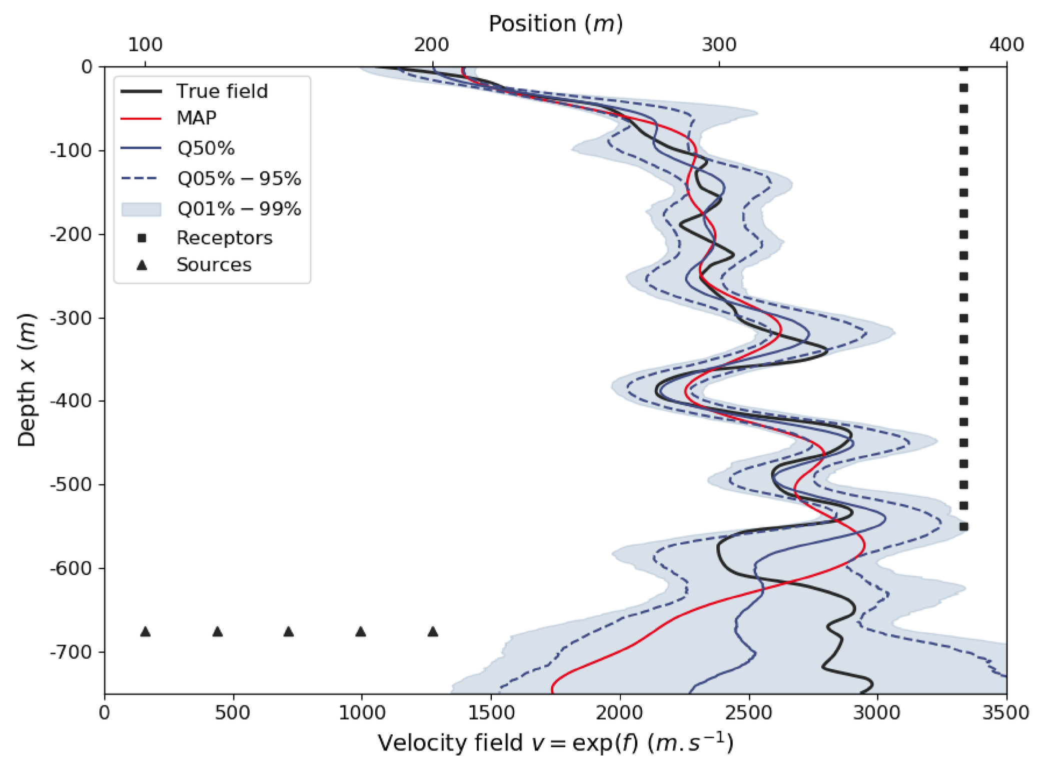


Fig. 2 – Result with change of measure

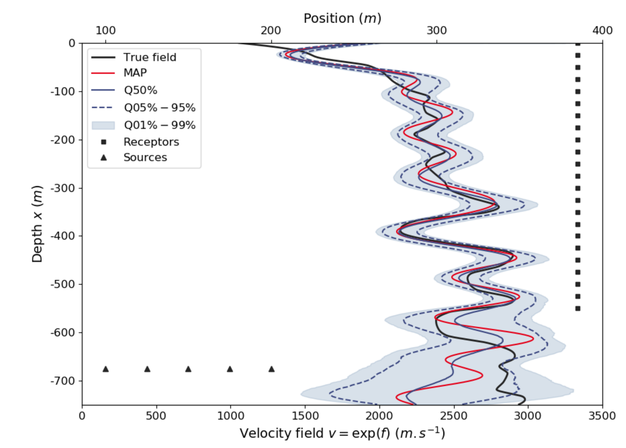


Fig. 3(a) – Inference with fixed small $\boldsymbol{q}$

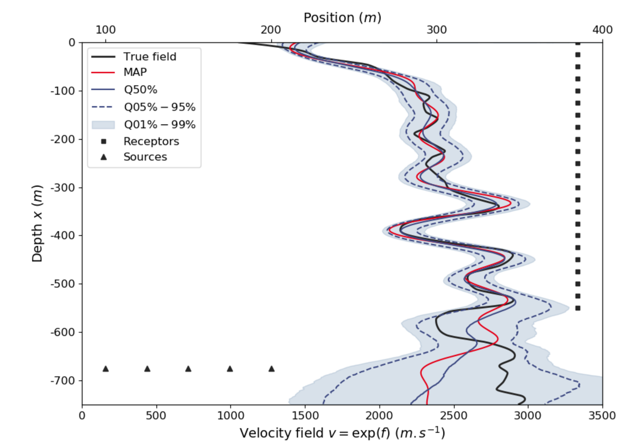


Fig. 3(b) – Inference with fixed medium $\boldsymbol{q}$

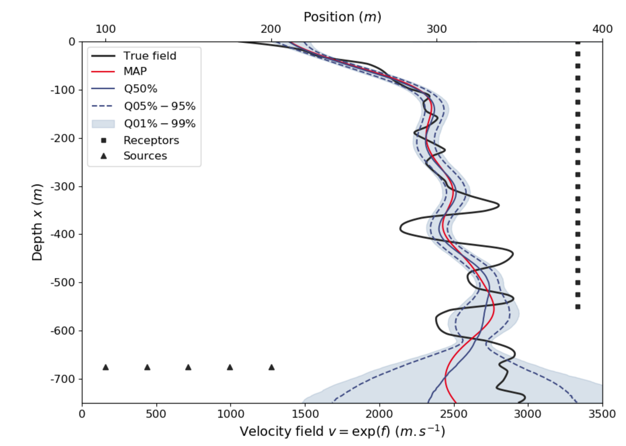


Fig. 3(c) – Inference with fixed large $\boldsymbol{q}$

Conclusion and perspectives

- Fast and accurate method to provide uncertainties on a physical field
- Further work:
  - Propagation of the uncertainties to source location
  - Use of surrogate maps for the forward model
  - Adaptive inference

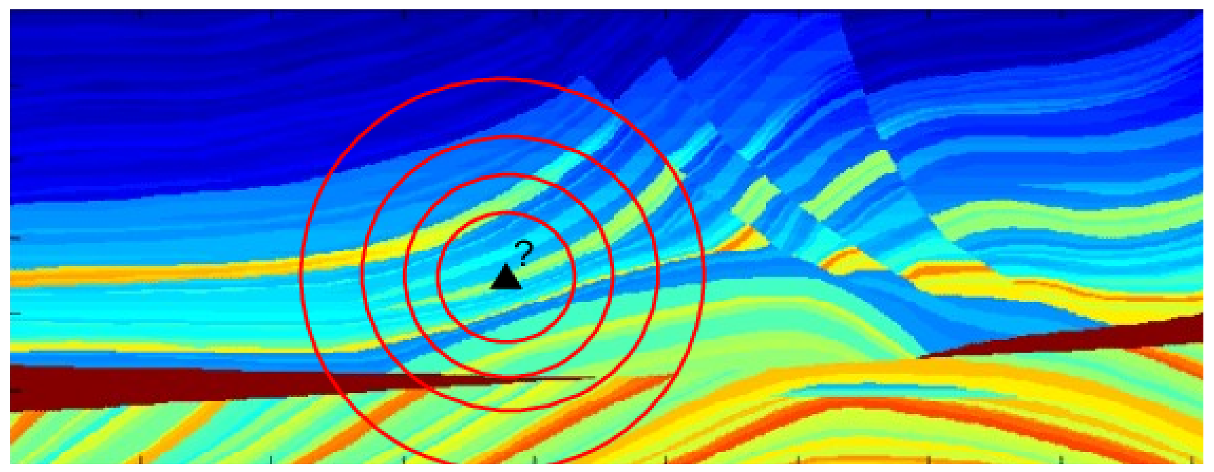


Fig. 4 – Illustration of the location source problem (Marmousi velocity model).