Variance-Informed Subspace: a Gradient-free Dimension Reduction for Adaptive Bayesian Inference

N. Polette[†,1,2], P. Sochala[§,1], O. Le Maître[§,3], A. Gesret[§,2]

[†] PhD student (presenting author). [§] PhD supervisor

PhD expected duration: Nov. 2022 – Oct. 2025

[1] CEA, DAM, DIF, F-91297 Arpajon, France
{pierre.sochala}@cea.fr

[2] Mines Paris PSL, Geosciences center, Fontainebleau, France
{nadege.polette,alexandrine.gesret}@minesparis.psl.eu

[3] CMAP, CNRS, Inria, École Polytechnique, IPP, Palaiseau, France
{olivier.le-maitre}@polytechnique.edu

Abstract

Inverse problems are encountered in many applications whenever one search for information about a physical system based on measurements [7]. In this work, we are interested in estimating a physical field thanks to a set of indirect observations $\boldsymbol{d}$. The Bayesian inference is an attractive approach for adressing such problems, as it provides a full estimation of the unknown parameters distributions. In that framework, the aim is to estimate the posterior probability of the field parameters $\boldsymbol{x}$ based on the observations

$$\pi_{\text{post}}(\boldsymbol{x}|\boldsymbol{d}) \propto \mathcal{L}(\boldsymbol{d}|\boldsymbol{x})\pi_{\text{prior}}(\boldsymbol{x}), \tag{1}$$

where $\mathcal{L}$ is the likelihood of the observations given a field and π_{prior} the prior probability of the field. The posterior distribution is then sampled with Markov Chain Monte–Carlo (MCMC) [4]. In order to accelerate the MCMC sampling, the forward model predictions are replaced with surrogate models based on polynomial chaos (PC) expansions [8, 5]. In order to reduce the input dimension of the surrogate model, a parsimonious representation of the field is introduced by means of the Karhunen-Loève (KL) decomposition, on the assumption that the field of interest is a particular realization of a Gaussian random field. Despite this parametrization, several hundred inputs could be required to represent accurately a two-dimensional field. This is expensive with regard to both the forward model surrogate training and the MCMC convergence.

Linear dimension reduction techniques have been developed to decrease the number of parameters to infer. These techniques assume that most of the information provided by the likelihood can be captured by a low-dimensional linear subspace. The input parameter space is decomposed into two subsets

$$\boldsymbol{x} = Ax_a + A_\perp x_i, \tag{2}$$

where x_a is informed by the likelihood, while x_i is constrained by the prior. The posterior distribution (1) rewrites

$$\pi_{\text{post}}(\boldsymbol{x}|\boldsymbol{d}) \propto \mathcal{L}(\boldsymbol{d}|x_a)\pi_{\text{prior}}(x_a)\pi_{\text{prior}}(x_i|x_a), \tag{3}$$

such that only x_a is sampled during the MCMC procedure. Several methods to define the linear transformation operator A have been developed. Cui et al. [3] build a Likelihood-Informed Subspace (LIS) which relies on the Hessian of the log-likelihood. The optimality of such construction

has been proven in [6] for the linear case. Constantine et al. [2] adapt the Active Subspace (AS) approach [1] to the Bayesian framework by using the misfit gradient. In both methods, the curvature of the log-posterior density is more constrained by the log-likelihood than by the prior along the subspace directions.

This study presents a new construction for the linear transformation operator A. The general idea is inspired from the work of [6] which states that, in the linear case, approximating the posterior covariance is equivalent to approximating its inverse. Instead of relying on the Hessian of the log-likelihood, the approximation of the inverse posterior covariance involves the ratio of the posterior and the prior variances.

For nonlinear inverse problems, we propose to generalize this variance ratio. The low-dimensional subspace is defined as the directions in which the posterior variance is drastically reduced in comparison to the prior variance. This method is gradient-free. We show on state-of-the-art examples that it is sufficient for unimodal posteriors, while some adjustments are required in the case of multimodal results. An application on a two-dimensional field inference case illustrates the interest of the method for high-dimensional problems.

Short biography (PhD student)

Nadège Polette studied applied mathematics at ENPC as well as at Sorbonne Université. Her PhD is funded by CEA. The supervision is ensured by Dr. A. Gesret (Geosciences center, École des Mines de Paris), Dr. P. Sochala (CEA), and Dr. O. Le Maître (CNRS, CMAP). Her PhD falls within the aim of the CTBTO to detect and analyse seismic events. The goal is to develop numerical methods for solving inverse problems applied to geophysical events analysis.

References

[1] P. G. Constantine. *Active Subspaces: Emerging Ideas for Dimension Reduction in Parameter Studies*. SIAM Spotlights. SIAM, 2015.

[2] P. G. Constantine, C. Kent, and T. Bui-Thanh. Accelerating MCMC with Active Subspaces. *SIAM Journal on Scientific Computing*, 38(5):A2779–A2805, January 2016.

[3] T. Cui, J. Martin, Y. M. Marzouk, A. Solonen, and A. Spantini. Likelihood-informed Dimension Reduction for Nonlinear Inverse Problems. *Inverse Problems*, 30(11):114015, November 2014.

[4] A. Doucet, A. Smith, N. de Freitas, and N. Gordon. *Sequential Monte Carlo Methods in Practice*. Information Science and Statistics. Springer New York, 2013.

[5] R. G. Ghanem and P. D. Spanos. Stochastic Finite Element Method: Response Statistics. In *Stochastic Finite Elements: A Spectral Approach*, pages 101–119. Springer New York, New York, NY, 1991.

[6] A. Spantini, A. Solonen, T. Cui, J. Martin, L. Tenorio, and Y. M. Marzouk. Optimal Low-Rank Approximations of Bayesian Linear Inverse Problems, July 2015.

[7] A. Tarantola. *Inverse Problem Theory and Methods for Model Parameter Estimation*, volume xii. SIAM, January 2005.

[8] N. Wiener. The Homogeneous Chaos. *American Journal of Mathematics*, 60(4):897–936, 1938.