

La méthode des éléments finis bayésienne

P. Kerfriden¹, I. Rocha², J. Rouse³

¹ Centre des Matériaux, Mines ParisTech, PSL university {pierre.kerfriden}@mines-paristech.fr

³ Faculty of Civil Engineering and Geosciences, TU Delft, The Netherlands

² Faculty of Engineering, University of Nottingham, UK

Résumé — Nous proposons la méthode des éléments finis bayésienne. L'algorithme délivre une densité de probabilité pour la solution du problème éléments finis, l'erreur numérique étant traitée comme une source d'incertitude épistémique. L'approche est hiérarchique, un prior Gaussien étant défini comme la solution d'une équation différentielle stochastique à une échelle fine, tandis qu'une échelle numérique plus grossière est utilisée comme grille d'observation dans un cadre d'assimilation de données probabiliste.

Mots clés — méthodes numériques probabilistes, estimation d'erreur, quantification d'incertitudes, assimilation de données.

Introduction

La méthode des éléments finis est aujourd'hui pratiquée dans la majeure partie des secteurs de l'industrie pour aider à la conception et à la certification des systèmes. Cette méthode est aujourd'hui relativement bien maîtrisée, notamment grâce à un certain nombre d'outils de vérification de la qualité des calculs, allant de la simple étude de convergence aux approches d'estimation d'erreur a posteriori et à l'adaptation locale des tailles de maille.

Pourtant, l'état de l'art en termes de vérification présente certaines limites. Les approches les plus avancées, basées sur l'estimation d'erreur a posteriori [1], ne permettent en général que d'évaluer l'erreur sur certaines quantités d'intérêt scalaires, définies comme des fonctionnelles du champ d'erreur. Ces approches sont le plus souvent restreintes aux problèmes linéaires, ou s'appuient très fortement sur une linéarisation des problèmes nonlinéaires, et sont donc fondamentalement basées sur une hypothèse de petites perturbations. Finalement, ces méthodes ne s'intègrent pas facilement dans des chaînes numériques probabilistes (e.g. assimilation de données, problèmes inverses Bayésiens), car elles délivrent des estimateurs ponctuels des erreurs en quantité d'intérêt, et non des distributions de probabilité comme il serait souhaitable pour qualifier, quantifier puis propager la méconnaissance épistémique due au solveur élément fini.

Nous proposons ici une approche alternative basée sur les méthodes numériques probabilistes bayésiennes (see e.g. [3,5]). Fondamentalement, la solution du problème élément fini est vue comme un problème d'assimilation de données, similaire à un problème de reconstruction de champ en géostatistique ou à un problème de régression par processus Gaussiens (e.g. Krigeage). Pour se faire, une densité de probabilité prior est définie pour le champ élément fini solution. Les équations à résoudre dans le cadre de la méthode des éléments finis sont alors interprétées comme des observations partielles de ce champ inconnu. Il est alors possible d'invoquer le théorème de Bayes pour obtenir la densité posterior, i.e. une densité de probabilité pour la solution du problème élément fini. Ainsi, la qualité de n'importe quelle quantité d'intérêt peut être naturellement évaluée en mesurant l'incertitude posterior associée à cette quantité. L'estimation d'erreur est, de fait, directement pratiquée sur le champ solution complet.

Techniquement, toutes les étapes de construction de la densité de probabilité posterior sont effectuées directement sur le système discret. Le prior est défini comme un processus Gaussien en utilisant les équations aux dérivées partielles stochastiques [10]. Ce choix permet d'interpréter le maximum a posteriori, car il est très fortement lié à la solution du problème élément fini déterministe. L'observation partielle de ce champ est obtenue n'écrivant la vérification de la condition de Galerkin qu'au sens de certains combinaisons des fonctions de forme élément fini. Il est naturel de retenir une formulation hié-

rarchique, ce qui est fait en cherchant des fonctions solutions dans un espace élément fini quadratique alors que l’orthogonalité de Galerkin ne sera assurée que par rapport à la restriction de cet espace aux fonctions linéaires sur chaque élément. Nous montrerons que ce choix permet de bien estimer la solution exacte sans avoir à réaliser d’étapes supplémentaires de calibration du prior par estimation d’erreur a posteriori. Finalement, un solveur à deux niveaux, qui ne sera pas décrit dans cet article court, se basant sur la construction hiérarchique de la méthode des élément fini bayésiens, est proposé afin de sampler la distribution posterior de manière efficace.

Cet article est organisé en trois parties. Nous introduisons tout d’abord la méthode des éléments fins classique, appliquée à un problème de Poisson. Dans un second temps, nous introduisons la méthode des éléments finis bayésienne. Finalement, nous présentons quelques résultats. Le reste de l’article est écrit en langue anglaise afin d’être compris des coauteurs non-anglophone de cette communication. Cependant, le premier auteur fournira des explications détaillées en français sur demande aux non anglophones.

Travaux apparentés

Le travail s’inscrit dans une tradition de méthodes numériques bayésiennes, et de méthodes numériques probabilistes plus généralement. La méthode proposée est notamment apparentée aux travaux suivants :

- Assimilation de données spatio-temporelles. La travail présenté trouve sa source dans l’assimilation de données par méthode de Kalman ou par d’autres méthodes statistiques permettant de fusionner des modèles spatio-temporels probabilistes et des données parcellaires [4,6].
- Solveurs bayésiens de type “meshfree”. Les solveurs d’équations aux dérivées partielles probabilistes à proprement parler peuvent être regroupés en deux catégories. D’une part, les solveurs de type collocation utilisant un processus gaussien dérivable au moins deux fois comme prior (e.g. versions bayésiennes des méthodes sans maillage) [2]. D’autre part, les solveurs de type Physics-Informed Neural Networks qui utilisent un prior de type réseau de neurones probabilistes, et toujours, sauf rares exceptions, des “observations” de type collocation de l’équation à résoudre [7,9]. A notre connaissance, le fruit de ce travail est la première extension bayésienne de la méthode des éléments finis (voir aussi [11] pour une alternative fréquentiste intéressante basée sur la perturbation aléatoire des maillages pour générer des ensembles de solutions).
- Reconstruction de champs d’erreur complets. Elle est notamment distincte de notre travail précédant en reconstruction de champs d’erreur complet par assimilation de données. En effet, dans l’approche proposée dans [8], la reconstruction statistique du champ d’erreur s’appuie sur des observations obtenues par estimation a posteriori d’erreur en quantité d’intérêt. Ces estimations sont réalisées par des méthodes classiques, dans une étape préliminaire à la reconstruction de champ. Cette étape n’est pas nécessaire dans le travail proposé ici.

1 Deterministic finite element formulation

We consider a finite element formulation of the standard Poisson problem in bounded domain Ω :

$$-\Delta u = f \quad (1)$$

Dirichlet and Neumann boundary conditions are applied over $\partial\Omega$. For the sake of simplicity of the derivations, we will assume that only Dirichlet conditions are applied. The weak form of the Poisson problem reads as

$$\int_{\Omega} \nabla u \cdot \nabla v d\Omega = \int_{\Omega} f v d\Omega \quad (2)$$

The corresponding finite element problem consists in looking for $\hat{u}(x) = \sum_{i=1}^N \psi_i(x) U_i$ where the $\{\psi_i\}$ ’s are finite element shape functions defined for $x \in \Omega$, and $\mathbf{U}$ is the solution of algebraic system

$$\mathbf{K}\mathbf{U} = \mathbf{F} \quad (3)$$

In the linear system above, the j^{th} line of matrix $\mathbf{K}$ and vector $\mathbf{F}$ are obtained by enforcing Galerkin’s condition

$$\int_{\Omega} \nabla \hat{u} \cdot \nabla \psi_j d\Omega = \int_{\Omega} f \psi_j d\Omega \quad (4)$$

Therefore, it follows that $\mathbf{K}_{ji} = \int_{\Omega} \nabla \psi_i(x) \cdot \nabla \psi_j(x) d\Omega$ and $\mathbf{F}_j = \int_{\Omega} f \psi_j(x) d\Omega$.

2 Probabilistic formulation : Bayesian state estimation setting

2.1 Linear observation operator

In the Bayesian finite element method, only selected linear combinations of the lines of the finite element system are to be enforced. To this end, we define

$$\mathbf{H}\mathbf{U} = \Phi^T \mathbf{K}\mathbf{U} \quad (5)$$

where $\mathbf{H}$ will be interpreted as an observation operator (i.e. in a data assimilation setting) and Φ^T is a rectangular coefficient matrix belonging to $\mathbb{R}^m \times \mathbb{R}^n$ with $m < n$.

We require that

$$\mathbf{d} := \Phi^T \mathbf{F} = \mathbf{H}\mathbf{U} + \epsilon \quad \epsilon \sim \mathcal{N}(0, \Sigma_\epsilon = \beta^2 \mathbf{I}) \quad (6)$$

where ϵ relaxes the enforcement of the Galerkin conditions. Typically, β^2 will be very small.

2.2 Prior : discrete stochastic differential equation

Due to partial enforcement of Galerkin's conditions, the finite element system is under-constrained. To regularise the problem, a prior probability density for $\mathbf{U}$ may be introduced. The chosen prior is defined by discrete stochastic partial differential equation

$$\mathbf{K}\mathbf{U} = \alpha \sqrt{\mathbf{M}} \mathbf{W} \quad \mathbf{W} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_d) \quad (7)$$

which leads to the following probability density function

$$\mathbf{U} \sim \mathcal{N}(\mathbf{0}, \Sigma) \quad \Sigma = \alpha^2 \mathbf{K}^{-1} \mathbf{M} \mathbf{K}^{-1} \quad (8)$$

In the previous expressions, the standard mass matrix $\mathbf{M}_{ij} = \int_{\Omega} \psi_i(x) \psi_j(x) d\Omega$ was introduced.

2.3 Analytical expression of the posterior distribution

With a Gaussian prior and a linear observation operator, the posterior distribution of the finite element solution is Gaussian,

$$\mathbf{U} \sim \mathcal{N}(\mathbf{U}^*, \Sigma^*) \quad (9)$$

where

$$\begin{aligned} \mathbf{U}^* &= \alpha^2 \mathbf{K}^{-1} \mathbf{M} \phi (\alpha^2 \phi^T \mathbf{M} \phi + \beta^2 \mathbf{I}_d)^{-1} \mathbf{d} \\ \Sigma^* &= \alpha^2 \mathbf{K}^{-1} \mathbf{M} \mathbf{K}^{-1} - \alpha^4 \mathbf{K}^{-1} \mathbf{M} \phi (\alpha^2 \phi^T \mathbf{M} \phi + \beta^2 \mathbf{I}_d)^{-1} \phi^T \mathbf{M} \mathbf{K}^{-1} \end{aligned} \quad (10)$$

Interestingly, if $\beta = 0$ the posterior mean may be written as $\mathbf{U}^* = \mathbf{K}^{-1} \mathbf{M} \phi (\phi^T \mathbf{M} \phi)^{-1} \mathbf{d} = \mathbf{K}^{-1} \mathbf{g}$ where $\mathbf{g}$ is the $\mathbf{M}$ -projection of the finite element right hand side onto the space spanned by the combination of shape functions implicitly defined by ϕ^T .

2.4 Choice of coarse space

We now specify observation operator $\phi^T \mathbf{K}$. Solution $\hat{u}(x) = \sum_{i=1}^N \psi_i(x) U_i$ is looked in the usual finite element space defined using quadratic Lagrange shape functions. We now introduce the usual finite element space defined using linear Lagrange shape functions. The corresponding coarse finite element shape functions $\{\phi_j\}$ can be shown to be linear combinations of the shape functions defining the quadratic Lagrange finite element space (hierarchical property of Lagrange finite element spaces) :

$$(\phi_1(x) \quad \phi_2(x) \quad \dots \quad \phi_M(x)) = (\psi_1(x) \quad \psi_2(x) \quad \dots \quad \psi_N(x)) (\Phi_1 \quad \Phi_2 \quad \dots \quad \Phi_M) \quad (11)$$

The coefficient matrix on the right hand side is ϕ , which is now fully specified. In other words, our Bayesian finite element procedure consists in cancelling the residual of the weak form when substituting the m linear shape functions for v , i.e.

$$\forall j, \quad \int_{\Omega} \nabla \hat{u} \cdot \nabla \phi_j d\Omega = \int_{\Omega} f \psi_j d\Omega - \epsilon_j \quad (12)$$

With this choice, the posterior mean becomes very close to the quadratic deterministic finite element solution, up to a $\mathbf{M}$ -projection of the right-hand side onto the linear finite element space.

3 Two-scale statistical sampler

The posterior distribution cannot be sampled with reasonable computational resources as for every realisation, the quadratic finite element problem would have to be solved twice. We therefore suggest a multiscale sampler whereby the action of $\mathbf{K}^{-1}$ is approximated using a couple of iterations of a projected Krylov solver, the projection space being the linear finite element space. Methodological details are not provided as this would lengthen this short contribution considerably.

4 Results

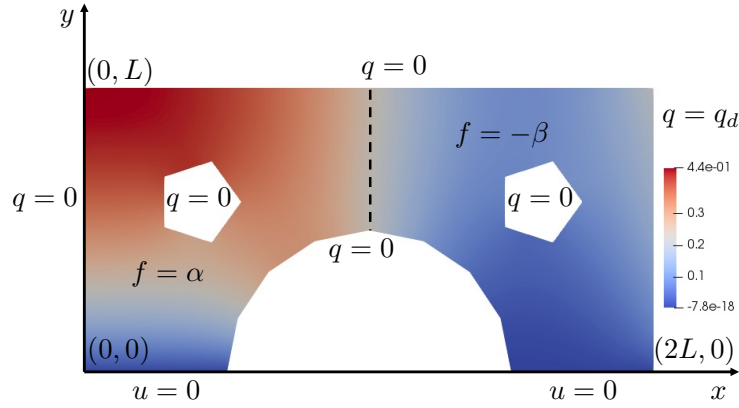


FIGURE 1 – Poisson problem with Dirichet and Neumann boundary conditions. The source term is positive in the left half of the computational domain, and negative in the right half. The colormap corresponds to the “truth” solution obtained by using a very refined finite element mesh

We showcase our methodology using the boundary value problem defined in figure .

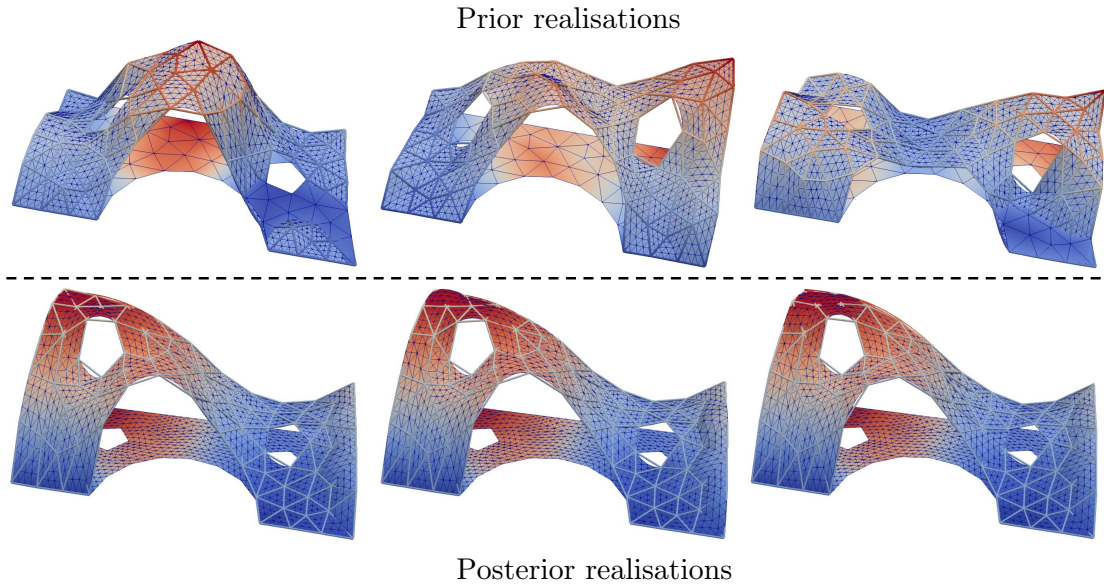


FIGURE 2 – Three realisations of the prior distribution (top row) and three realisations of the posterior distribution (bottom row). The ensemble of posterior solutions, weighted by the posterior probability density, constitutes the solution of the Bayesian finite element solveur.

In Figure 2, several realisations of the prior and posterior distributions are displayed. The amplitude of the prior, which is the only parameter of the method, has been calibrated using a maximum likelihood approach. The posterior realisations seem to waver around the solution of the deterministic finite element method, which is further investigated below

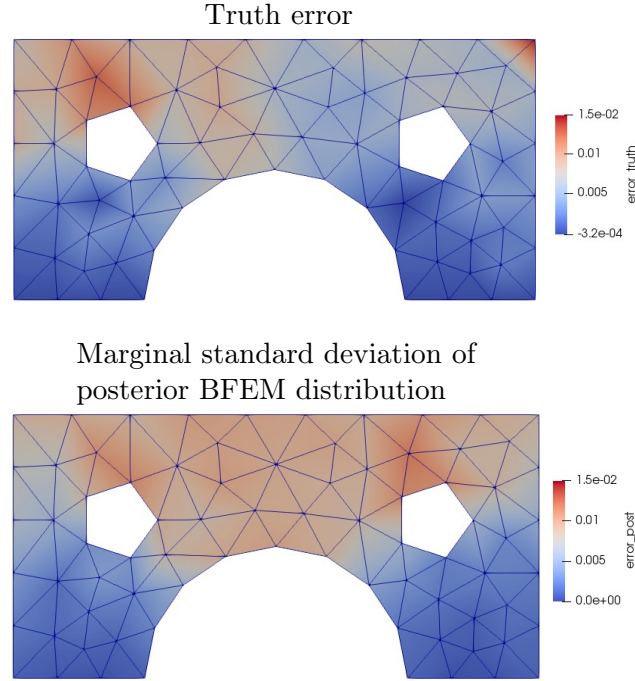


FIGURE 3 – Error distributions

in Figure 3, we show the difference between the deterministic finite element solution using quadratic shape functions and the deterministic finite element solution using linear shape functions. This field may be seen as a “truth” error, the quadratic finite element solution converging one order faster than the linear one with mesh refinement. We also report the field corresponding to the standard deviation of the distribution of the solution, marginally at every point of the computational domain. Clearly, the Bayesian FEM correctly predicts the truth error, without explicit error estimation operation.

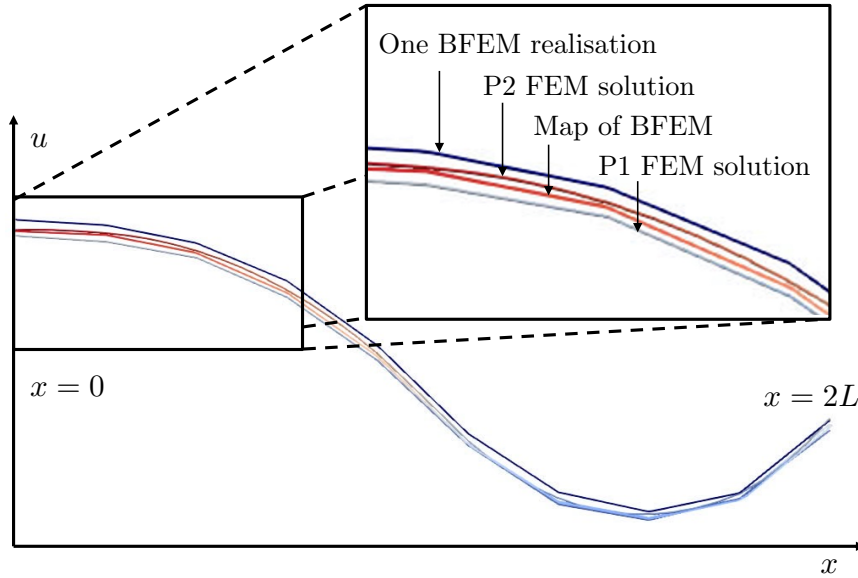


FIGURE 4 – Marginal distance between quadratic and linear finite element solutions (top) and error estimates provided by the Bayesian FEM, i.e. the standard deviation of the distribution of the solution, locally (bottom).

Finally, in Figure 4, we provide deeper understanding in our results by showing several fields as a function of first spatial coordinate x , for $y = 1$. We see that the maximum a posterior (MAP) of the posterior distribution (which is also the mean for gaussian posteriors), is very close to the quadratic solution, which validates the previous observation about the analytical form of the posterior distribution. Realisations of the Bayesian FEM waver around the MAP with a standard deviation that is of the order of the distance between the quadratic and linear deterministic finite element solutions.

5 Conclusion

We have described a novel Bayesian finite element methodology that delivers distribution of finite element solutions. The prior is defined via a carefully chosen discrete stochastic partial differential equation, while the observation operator is obtained using nested Lagrange finite element spaces. These choices ensure (i) that the MAP of the posterior distribution of the Bayesian FEM is closely related to the deterministic FEM solution, (ii) that the prior needs only minimal adjusting to the data, typically using standard maximum likelihood (e.g. standard empirical Bayes approach). We show that the Bayesian FEM predicts posterior solution densities that are sensible, thereby circumventing the need for error estimation operations. Posterior sampling is intractable, and we suggest the use of an efficient multiscale algorithm that takes advantage of the nested finite element spaces that are at the root of the Bayesian FEM. Alternatives using Langevin-MCMC sampling are currently being studied.

Références

- [1] Chamoin, Ludovic ; Legoll, Frederic , *A pedagogical review on a posteriori error estimation in Finite Element computations*, arXiv :2110.02160, 2021
- [2] J. Cockayne, C. Oates, T. J. Sullivan, and M. Girolami, *Probabilistic meshless methods for partial differential equations and Bayesian inverse problems*, 2016.
- [3] Patrick R Conrad, Mark Girolami, Simo Särkkä, Andrew Stuart, Konstantinos Zygalakis, *Statistical analysis of differential equations : introducing probability measures on numerical solutions*, Stat Comput, 2017
- [4] Mark Girolami, Eky Febrianto, Ge Yin, Fehmi Cirak, *The statistical finite element method (statFEM) for coherent synthesis of observation data and model predictions*,
- [5] P. Hennig, *Probabilistic interpretation of linear solvers*, SIAM Journal on Optimization, 2015. Computer Methods in Applied Mechanics and Engineering, 2021
- [6] Nguyen, N.C., and J. Peraire, *Gaussian Functional Regression for Linear Partial Differential Equations.*, Computer Methods in Applied Mechanics and Engineering, 2015
- [7] Maziar Raissi, Paris Perdikaris, George Em Karniadakis, *Inferring solutions of differential equations using noisy multi-fidelity data*, Journal of Computational Physics, 2017.
- [8] J.P. Rouse, P. Kerfriden, and M. Hamadi, *A probabilistic hierarchical sub-modelling approach through a posteriori Bayesian state estimation of finite element error fields*, hal.archives-ouvertes.fr, 2021.
- [9] Liu Yang, Xuhui Meng, George Em Karniadakis, *B-PINNs : Bayesian physics-informed neural networks for forward and inverse PDE problems with noisy data*, Journal of Computational Physics, 2021.
- [10] L. Roininen, J. Huttunen and S. Lasanen, *Whittle-Matérn priors for Bayesian statistical inversion with applications in electrical impedance tomography*, Inverse Problems and Imaging, 2014.
- [11] Assyr Abdulle, Giacomo Garegnani, *A probabilistic finite element method based on random meshes : A posteriori error estimators and Bayesian inverse problems*, 2021.