Early Stage Assessment (ESA)

Department of Mathematics - Imperial College London
Yanni (John) Papandreou supervised by Dr Andrew Duncan
Project Title: *Bayesian formulations of PDE-based models, with an aim towards physics-informed machine learning*

1 Introduction

Mathematical models of physical systems are often expressed in terms of Partial Differential Equations (PDEs). In such models uncertainty is often introduced, either through a lack of knowledge of the parameters of the system or via inherent randomness in the system itself. Furthermore, these mathematical models and the associated computer simulations are often simplifications of the actual system leading to possible model misspecification [1]. The numerical algorithms used to simulate such models also induce uncertainty [2]. For example, it is often the case that a numerical method involves a finite-dimensional approximation of the unknown function. The computer simulations of these models can also often be very expensive computationally, and as such, these models are rarely used alone in the modelling procedure. Instead, observational data from the actual physical system via measurements is often incorporated as well. Data from measurements is now becoming increasingly available in almost every area of engineering and science, and failure to consider either the data or the model is clearly suboptimal. The issue of combining knowledge from both the model and data is thus of utmost importance and is often referred to as **data assimilation**, especially when the underlying mathematical model is a potentially stochastic dynamical system and the data may be time-ordered [3].

The Bayesian formulation of PDE-based models allows one to naturally incorporate all these sources of uncertainty (while still being able to identify the relevant source) and forces one to deal with modelling issues in a clear and precise manner. It allows measurement data to be considered in order to calibrate and tune mathematical models. Moreover, it allows a full characterization of all possible solutions to be found, together with their relative probabilities [4].

The statistical formulation of PDE-based models is necessary in many applications in order to handle in a precise manner the uncertainty present in the model. It allows this uncertainty to be propagated forward and gives clear answers for how much trust can be placed in the conclusions of the model. The finite element method (FEM) [5] is one of the most widely used methods for numerically approximating the solution of PDEs modelling natural and physical systems. Since FEM is an integral part of the study of many physical systems it is essential that we are able to fully quantify the uncertainty that using FEM introduces to our models and simulations.

In the recent paper "The Statistical Finite Element Method" Girolami et al. introduce a novel unifying approach which provides a fully statistical FEM in which both the finite element model and observational data are combined into a coherent inferential framework. This approach allows observational data to provide us with data adjusted FEM solutions. In particular, this paper considers a large class of linear PDEs and assumes that we have incomplete knowledge of the forcing term. The uncertainty resulting from this lack of knowledge is then formally accounted for by modelling the forcing as being random, having an appropriately defined Gaussian process distribution. The approximation of the linear PDE using the FE method then yields a multivariate Gaussian distribution on the resulting finite dimensional approximation. This probabilistic representation of the FE method is then used to condition the model on sensor data, providing a systematic methodology by which one can statistically update the Galerkin FEM solution.

Our work aims to provide more detailed error analysis explicitly quantifying the extent by which the distributions obtained using the "true" solution and the FEM solution differ. Once this is investigated we will then explore how this propagates through to any further inference. Our treatment aims to follow the guiding principle of "*avoiding discretization until the the last possible moment*" [4]. This principle is a very powerful one used throughout numerical analysis, and we will aim to highlight its importance in our work.

The remainder of the report is structured as follows. Section 2 provides a concise account of the most relevant background material for the project together with a brief survey of the literature behind the topic of this work. In Section 3 we introduce the general framework we will be considering. We then apply the general framework in the context of FEM for an elliptic boundary value problem in Section 4 and then to a parabolic problem in Section 5. Section 6 finally discusses in more general terms the aims of our research before outlining several ideas for future work.

2 Brief Overview of Background Material

Our project is focused on providing detailed analysis of the uncertainty introduced by utilising numerical approximations for solving potentially noisy PDEs and then investigating how this propagates through to further inference, when for instance observational data is incorporated. This project lies at the intersection of the fields of data assimilation, data-centric engineering, probabilistic numerics and Bayesian inference. We now provide a brief overview of the relevant background material for the project.

It is now well established that the language of probabilistic inference can be applied to numerical problems in order to provide a more detailed notion of the uncertainty resulting from numerically approximating an intractable problem [6, 7, 8, 9]. Numerical algorithms can be viewed as estimation rules for a latent, often intractable quantity given the results of tractable computations. Such algorithms can be considered to perform inference and are thus open to being analysed using the formal framework of probability theory. The field of Probabilistic Numerics (PN) [10] involves the study of so called "probabilistic numerical methods"; these are numerical algorithms which take in a probability distribution over its inputs and gives out a probability distribution over its output. Several existing numerical methods have even recently been shown to arise from specific probabilistic models. It is worth pointing out that so far we have only been referring to problems of a deterministic nature and probability theory is used as a means of providing a notion of the uncertainty inherent in using a numerical approximation to the solution of an intractable deterministic problem. In our work we will not restrict attention to purely deterministic problems but instead will consider potentially noisy PDEs. We will then seek to analyse the problem from the viewpoint of PN.

Much work has already been undertaken in the field of PN into applications to differential equations, especially for ODEs. Classic numerical algorithms for solving initial value problems (IVPs) provide an approximate solution often defined on a grid of time points. This numerical solution is often computed iteratively by collecting information from evaluations of the vector field associated to the system of differential equations. Probabilistic numerical methods instead provide probability measures, as opposed to point estimates, over the space of possible solutions to the IVP. In the PN literature there are two main approaches to solving ODEs which we now briefly outline.

The first approach (including methods from [11, 2, 12, 13, 14, 15]) introduces probability measures to ODE solvers by representing the distribution of all numerically possible trajectories with a set of sample paths. The computation of these sample paths varies across these works. [11] draws them from a (Bayesian) Gaussian process regression while [2, 12, 13, 15] perturb classical estimates after an integration step with suitably scaled Gaussian noise and [14] instead perturbs the classical estimate via choosing a stochastic step size.

The second approach [16, 17, 18, 19, 20, 21] recasts IVPs as **stochastic filtering problems**. This method involves assuming *a priori* that the solution of the IVP and a prespecified number of its derivatives follow a Gauss-Markov process that solves a particular stochastic differential equation (SDE). The evaluations of the vector field of the IVP at numerical estimates of the true solution are then regarded as imperfect evaluations of the time derivative of the solution and are thus used as a Bayesian update for the Gauss-Markov process. This approach gives an algorithm very similar in structure to that of the Kalman filter.

Probabilistic numerical methods for PDEs are much more uncommon. However, some methods do exist, which we now briefly outline. In the book "An Introduction to Computational Stochastic PDEs" Lord et al. analyse several different methods for dealing with elliptic PDEs with random data. In particular in chapter 9 of [22] the following (random) elliptic boundary-value problem (BVP) on a domain $D \subset \mathbb{R}^2$ is considered:

$$\begin{aligned} -\nabla \cdot (a(x)\nabla u(x)) &= f(x), \quad \forall x \in D \\ u(x) &= g(x), \quad \forall x \in \partial D \end{aligned}$$

where $\{a(x)|x \in D\}$ and $\{f(x)|x \in D\}$ are second-order random fields. Lord et al. consider several methods for dealing with such a BVP. To start they first consider a variational formulation on D and show that under suitable assumptions on the diffusion coefficients there is a unique solution to the variational formulation almost surely. A Galerkin FE approximation is then established for this formulation. The FEM is then combined with the Monte Carlo method to yield what the authors call the "Monte Carlo Finite Element Method" (MCFEM) which can be used to estimate the expectation and variance of $u(x)$. This method essentially involves drawing *iid* samples from the random fields in the BVP and then applying the FEM element to the resulting elliptic BVPs. Following this a variational formulation on $D \times \Omega$ is instead considered, where Ω is the underlying sample space for the probability space where the random fields live. The associated weak form is not a convenient starting point for Galerkin approximation as it involves taking expectations with respect to the abstract set Ω and the associated probability measure. This leads the authors to instead consider that the noise arising from the fields comes from a finite number of random-variables (i.e. the random fields are so called *finite-dimensional noise*). Doing so yields an equivalent weak form on $D \times \Gamma$ where Γ is the range of the finite-dimensional noise. Having done this a Stochastic Galerkin FEM is developed to approximate the solution to this new weak form. Both a semi-discrete and fully-discrete version are considered (discretization can now occur in two spaces). After analysing this method the authors finally consider a stochastic collocation FEM which combines collocation on the range of the finite-dimensional noise and FEM approximations on D. In Section 4 we will consider a particular example of such a random elliptic BVP and introduce an alternative probabilistic numerical method to tackle this.

3 General Framework

4 FEM for Elliptic Boundary Value Problem

5 FEM for Parabolic Problem

6 Conclusions and Research Plan

References

[1] Mark Girolami, Alastair Gregory, Ge Yin, and Fehmi Cirak. The statistical finite element method. *arXiv preprint arXiv:1905.06391*, 2019.

[2] Patrick R Conrad, Mark Girolami, Simo Särkkä, Andrew Stuart, and Konstantinos Zygalakis. Statistical analysis of differential equations: introducing probability measures on numerical solutions. *Statistics and Computing*, 27(4):1065–1082, 2017.

[3] Kody Law, Andrew Stuart, and Kostas Zygalakis. Data assimilation. *Cham, Switzerland: Springer*, 2015.

[4] Andrew M Stuart. Inverse problems: a bayesian perspective. *Acta numerica*, 19:451–559, 2010.

[5] Gilbert Strang and George J Fix. An analysis of the finite element method. *Journal of Applied Mathematics and Mechanics*, 1973.

[6] Persi Diaconis. Bayesian numerical analysis. *Statistical decision theory and related topics IV*, 1:163–175, 1988.

[7] A O'Hagan. Bayesian statistics 4. some bayesian numerical analysis, 1992.

[8] John Skilling. Bayesian solution of ordinary differential equations. In *Maximum entropy and Bayesian methods*, pages 23–37. Springer, 1992.

[9] Philipp Hennig, Michael A Osborne, and Mark Girolami. Probabilistic numerics and uncertainty in computations. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 471(2179): 20150142, 2015.

[10] URL `http://probabilistic-numerics.org/`.

[11] Oksana A Chkrebtii, David A Campbell, Ben Calderhead, Mark A Girolami, et al. Bayesian solution uncertainty quantification for differential equations. *Bayesian Analysis*, 11(4):1239–1267, 2016.

[12] Onur Teymur, Kostas Zygalakis, and Ben Calderhead. Probabilistic linear multistep methods. In *Advances in Neural Information Processing Systems*, pages 4321–4328, 2016.

[13] Han Cheng Lie, AM Stuart, and TJ Sullivan. Strong convergence rates of probabilistic integrators for ordinary differential equations. *Statistics and Computing*, 29(6):1265–1283, 2019.

[14] Assyr Abdulle and Giacomo Garegnani. Random time step probabilistic methods for uncertainty quantification in chaotic and geometric numerical integration. *Statistics and Computing*, pages 1–26, 2020.

[15] Onur Teymur, Han Cheng Lie, Tim Sullivan, and Ben Calderhead. Implicit probabilistic integrators for odes. In *Advances in Neural Information Processing Systems*, pages 7244–7253, 2018.

[16] Michael Schober, David K Duvenaud, and Philipp Hennig. Probabilistic ode solvers with runge-kutta means. In *Advances in neural information processing systems*, pages 739–747, 2014.

[17] Hans Kersting and Philipp Hennig. Active uncertainty calibration in bayesian ode solvers. *arXiv preprint arXiv:1605.03364*, 2016.

[18] Emilia Magnani, Hans Kersting, Michael Schober, and Philipp Hennig. Bayesian filtering for odes with bounded derivatives. *arXiv preprint arXiv:1709.08471*, 2017.

[19] Michael Schober, Simo Särkkä, and Philipp Hennig. A probabilistic model for the numerical solution of initial value problems. *Statistics and Computing*, 29(1):99–122, 2019.

[20] Filip Tronarp, Hans Kersting, Simo Särkkä, and Philipp Hennig. Probabilistic solutions to ordinary differential equations as nonlinear bayesian filtering: a new perspective. *Statistics and Computing*, 29(6): 1297–1315, 2019.

[21] Hans Kersting, TJ Sullivan, and Philipp Hennig. Convergence rates of gaussian ode filters. *arXiv preprint arXiv:1807.09737*, 2018.

[22] Gabriel J Lord, Catherine E Powell, and Tony Shardlow. *An introduction to computational stochastic PDEs*, volume 50. Cambridge University Press, 2014.