Bayesian Analysis (2016) **11**, Number 4, pp. 1295–1299

Rejoinder*

Oksana A. Chkrebtii[†], David A. Campbell[‡],
Ben Calderhead[§], and Mark A. Girolami[¶]

1 Introduction

We thank Bruno Sansó and the BA editorial team for encouraging discussion on our paper, and the discussants for their interesting and valuable contributions. Our rejoinder is divided into six sections, which provide insight and clarification on the subjects that were raised by the discussants. The paper by Chkrebtii et al. (2016a) and the proposed formalism will hereafter be referred to as UQDE (uncertainty quantification for differential equations).

2 Uncertainty quantification for dynamical systems with the Markov property

Our probabilistic approach to modeling uncertainty in the unknown solution of a dynamical system is motivated by its Markov structure. For example, consider the initial value problem $du/dt = f(t, u)$ on $t \in [0, L]$ with initial condition $u(0) = u_0$. It can be shown that for $t_1 < t_2$, the solution $u(t_2)$ is a function of $u(t_1)$ that does not depend on $u(\tau), \tau \in [0, t_1)$ (e.g., Jazwinski, 1970). Thus, defining probability measures sequentially on a filtration of σ-algebras is a key feature of our proposal and an important distinction with the work of Skilling (1991). Such sequential probability models are also used in simulation of stochastic differential equation (SDE) sample paths, suggesting a relationship with the SDE literature, as described in the insightful discussion of Lysy (2016).

The Markov property is also relevant to the comment of Dass (2016). Equation (2) of UQDE expresses the probabilistic solution $[u, u_t \mid \theta, \Psi, N]$ as a continuous mixture of Gaussian processes obtained by marginalizing $[u, u_t, \mid \mathrm{f}_1, \ldots, \mathrm{f}_N, \theta, \Psi, N]$ over trajectories $\mathrm{f}_1, \ldots, \mathrm{f}_N$ with mixture weights $p(\mathrm{f}_1, \ldots, \mathrm{f}_N)$. Algorithm 1 samples from this mixture by effectively selecting a mixture component from $p(\mathrm{f}_1, \ldots, \mathrm{f}_N)$ and then drawing a sample from $[u, u_t \mid \mathrm{f}_1, \ldots, \mathrm{f}_N]$. However, the Markov structure of the solution u prevents conditioning the trajectory directly on samples from multiple sample paths simultaneously.

Cockayne (2016) suggests a different perspective on this problem. Instead of estimating $u : [0, L] \to \mathbb{R}^p$ given a known vector field function $f(t, \cdot) : \mathbb{R}^p \to \mathbb{R}^p$, the discussant considers estimation of the function $f(\cdot, u(\cdot)) : [0, L] \to \mathbb{R}^p$ directly. In both cases the

*Main article DOI: 10.1214/16-BA1017.
[†]Department of Statistics, The Ohio State University, Columbus, OH, USA, oksana@stat.osu.edu
[‡]Department of Statistics & Actuarial Science, Simon Fraser University, Canada, dac5@stat.sfu.ca
[§]Department of Mathematics, Imperial College London, London, UK, b.calderhead@imperial.ac.uk
[¶]Department of Statistics, The University of Warwick, Coventry, UK, M.Girolami@warwick.ac.uk

 DOI: 10.1214/16-BA1017REJ

Markov structure of u is inherent in the problem and must be enforced. UQDE does this by defining a probability model for u on a filtration. The solution is updated by conditioning on a set of sequentially updated model trajectories instead of unrelated samples of the solution at different time points and possibly belonging to different trajectories. In this way, we try to maintain as much of the structural information about u as possible. The inconsistency brought up in the discussion supports our point that neglecting to enforce the Markov property of u by, for example, sampling backwards in time, does not correctly account for the structure of the solution.

3 Uncertainty quantification for partial differential equations

The Markov property also holds for partial differential equations (PDEs) in the temporal input. For this reason, the well-known "forward in time, continuous in space" approach and its probabilistic analogue developed in UQDE are a reasonable choice for modeling both spatial and temporal uncertainty. On the other hand, the example described in Section 5 consists of applying the probabilistic solution technique to the spatially discretized Navier–Stokes equation projected into the Fourier domain, reducing the problem to a large number of coupled ordinary differential equations (ODEs). As suggested by Cockayne (2016), this of course does not take into account spatial uncertainty, which will be accounted for by a direct probabilistic method, such as the one introduced in Section 5.4 of UQDE.

Full uncertainty quantification for systems governed by PDEs is expected to be very computationally expensive. To overcome this issue, Mallick et al. (2016) propose a hybrid method that uses a combination of numerical and probabilistic models on different scales. We discuss this strategy in the following section on prior specification.

4 Prior specification

The discussions of Yoo (2016) and Mallick et al. (2016) provide an alternative and flexible way of defining the prior process jointly on u and any partial derivatives via basis expansion with Gaussian priors on the coefficients. Although Gaussian processes can also be written in spectral form, the eigenfunctions of the covariance lack the interpretability of B-spline bases, and are difficult to adapt to different resolutions. Therefore, we believe that this suggestion is promising for modeling spatio-temporal uncertainty in the PDE forward problem.

When Gaussian processes are defined directly, as in UQDE, Briol et al. (2016) point out that differentiation is often simpler than integration as a technique of defining covariances over the state, u, and any derivatives. However, in this case, it is not straightforward to choose a covariance corresponding to an anisotropic prior that enforces $u(0) = u_0$ with probability one. Certainly, in the case of a single input, e.g. ODE boundary value problems, this is not strictly required as boundary conditions can be enforced by conditioning on the known boundary values. However, as pointed out

by Briol et al. (2016), this technique breaks down for PDE problems with functional boundaries.

5 Rate of convergence

An important advantage of the Bayesian formalism is that it encodes belief given a finite set of "data". However, consistency in the limit is an important feature of any Bayesian approach because it implies that with infinite information, we gain arbitrarily accurate knowledge about the unknown.

Our convergence result applies to the entire stochastic process $[u \mid \theta, \Psi, N] = \int [u, \mathbf{f} \mid \theta, \Psi, N]\, d\mathbf{f}$, which we term the "probabilistic solution". Therefore, as discussed in Dass (2016), the fact that $C_N \to 0$ alone is not enough to ensure convergence. In the Supplement (Chkrebtii et al., 2016b) we have shown that the posterior expectation $E(|u - u^*| \mid \theta, \Psi, N) = O(h)$ as $\alpha^{-1}, \lambda \to 0$ under the stated assumptions. Mallick et al. (2016) point out that this consistency result corresponds to the accuracy of a numerical solver. Additionally, convergence rates under different grid point distributions are certainly worth studying, as suggested by Yoo (2016).

6 Relationship to numerical solvers

There has been recent interest in characterizing uncertainty associated with a given numerical solver (Schober and Hennig, 2016) and in using the output of a numerical solver as information to update a prior distribution on the unknown solution (e.g., Dass, 2016). While the ability to take advantage of existing numerical infrastructure is appealing, our objective with UQDE is to show that the Bayesian formalism produces uncertainty quantification naturally without the need for numerical methods. The use of numerical output within a UQ algorithm (e.g., Dass, 2016) requires the conversion of numerical error bounds to probability statements, and may result in the loss of the functional structure of the error.

Mallick et al. (2016) suggest an interesting compromise between discrete approximations and probabilistic methods for multiscale models. They propose to recover large-scale "resolved" dynamics numerically, while the uncertainty in the smaller-scale "unresolved" dynamics, which is typically much greater, can be characterized through Gaussian process prior updating. This type of analysis would be applicable to many stiff systems with solutions that have structure at different scales.

Next, we discuss the points made by Briol et al. (2016) and Schober and Hennig (2016) and give an example of how a special case of UQDE can emulate a numerical method. We agree that analogies between numerical and probabilistic methods can provide insights into both techniques. Thus, the correspondence between methods must be carefully qualified, whether it is agreement in the mean or the rate of convergence (Schober and Hennig, 2016), or simply that the resulting sampling scheme and prior resemble a numerical solver, as with UQDE. Consider, for example, the following UQDE solver, which emulates an explicit first-order Euler scheme. This is constructed

by defining a GP prior on the state with integrated uniform covariance structure and length-scale equal to a single time step. This restricts possible trajectories to piecewise linear functions with derivative discontinuities over the sampling grid. This special case also illustrates the answer for a question of Briol et al. (2016) about how to best exploit sparsity of the covariance structures. In this example, the updating problem is reduced to a linear extrapolation at each solver step with a random component with predictive variance.

7 Stochastic differential equations

In an insightful discussion, Lysy (2016) suggests possible strategies to adapt the UQDE method to the simulation of sample paths with uncertainty quantification for the stochastic differential equation (SDE) (i) initial value problem and (ii) boundary value problem, also called diffusion bridge sampling. Our suggestion is that diffusion bridge sampling could be used to interpolate subsequent model interrogations in method (i) to replace the likelihood, $p(\mathrm{f}_{i+1} \mid \mathrm{f}_i, X_t)$, which is misspecified for SDEs as suggested in the discussion. For (ii), we suggest incorporating boundary constraints into the prior measure. We agree that uncertainty quantification for SDE model simulation is a promising direction for research and we look forward to more work on this topic.

8 Calibration

One of the goals of UQDE is in addressing posterior undercoverage due to an optimistic view of numerical error (e.g., that it is negligible) in the forward problem. Dass (2016) and Mallick et al. (2016) pose an important question: is posterior coverage guaranteed in the limit? Intuitively, one would suppose that if the uncertainty quantification in the forward model does not underestimate discretization uncertainty, then this should be the case. However, theoretical studies are needed to confirm this fact. We encourage future investigation of this topic.

Regarding practical aspects of calibration, Dass (2016) has pointed out that the marginal posterior distribution of the length-scale hyperparameter λ will have support that decreases with the discretization step length. A similar phenomenon is observed in other nonparametric regression problems. This fact requires care in the sampling algorithm. Indeed, our implementation of Algorithm (2) adapts the proposal distribution variance during the burn-in phase to achieve reasonable acceptance rates (e.g. Roberts and Rosenthal, 2009). Additionally, we have noticed that prior hyperparameters in this setting often have complex posterior correlations with model parameters. For this reason, we suggest using and have provided in the Supplement a Parallel Tempering implementation (Geyer, 1991) for the calibration problem.

9 Conclusion

We are happy that this paper has generated so much discussion. It is encouraging to see a large number of other methods being put forward, many of which have been cited in

the discussions. We look forward to many new developments in probabilistic numerics in the future.

References

Briol, F.-X., Cockayne, J., Teymur, O., Yoo, W. W., Cockayne, J., Schober, M., and Hennig, P. (2016). "Contributed Discussion on Article by Chkrebtii, Campbell, Calderhead, and Girolami." *Bayesian Analysis*. doi: http://dx.doi.org/10.1214/16-BA1017A. 1295, 1296, 1297, 1298

Chkrebtii, O., Campbell, D., Calderhead, B., and Girolami, M. (2016a). "Bayesian Solution Uncertainty Quantification for Differential Equations." *Bayesian Analysis*. doi: http://dx.doi.org/10.1214/16-BA1017. 1295

Chkrebtii, O., Campbell, D., Calderhead, B., and Girolami, M. (2016b). "Supplementary Material for "Bayesian Solution Uncertainty Quantification for Differential Equations"." *Bayesian Analysis*. doi: http://dx.doi.org/10.1214/16-BA1017SUPP. 1297

Dass, S. (2016). "Comment on Article by Chkrebtii, Campbell, Calderhead, and Girolami." *Bayesian Analysis*. doi: http://dx.doi.org/10.1214/16-BA1037. 1295, 1297, 1298

Geyer, C. (1991). "Markov chain Monte Carlo maximum likelihood." In *Computing Science and Statistics, Proceedings of the 23rd Symposium on the Interface, 156*. American Statistical Association. 1298

Jazwinski, A. (1970). *Stochastic Processes and Filtering Theory*. Mathematics in Science and Engineering. Elsevier Science. 1295

Lysy, M. (2016). "Comment on Article by Chkrebtii, Campbell, Calderhead, and Girolami." *Bayesian Analysis*. doi: http://dx.doi.org/10.1214/16-BA1036. 1295, 1298

Mallick, B., Yang, K., Guha, N., and Efendiev, Y. (2016). "Comment on Article by Chkrebtii, Campbell, Calderhead, and Girolami." *Bayesian Analysis*. doi: http://dx.doi.org/10.1214/16-BA1038. 1296, 1297, 1298

Roberts, Gareth O., and Rosenthal, Jeffrey S. (2009). "Examples of adaptive MCMC." *Journal of Computational and Graphical Statistics*, 18(2): 349–367. MR2749836. doi: http://dx.doi.org/10.1198/jcgs.2009.06134. 1298

Skilling, J. (1991). *Bayesian Solution of Ordinary Differential Equations*, 23–37. Seattle: Kluwer Academic Publishers. 1295