Bayesian Analysis (2016) **11**, Number 4, pp. 1279–1284

Comment on Article by Chkrebtii, Campbell, Calderhead, and Girolami*

Bani K. Mallick[†], Keren Yang[†,‡], Nilabja Guha[§], and Yalchin Efendiev[¶]

Abstract. This note is a discussion of the article "Bayesian Solution Uncertainty Quantification for Differential Equations" by Chkrebtii, Campbell, Calderhead, and Girolami. The authors propose stochastic models for differential equation discretizations. While appreciating the main concepts, we point out some possible extensions and modifications.

Keywords: Bayesian, uncertainty quantification, discretization errors, multiscale, variational formulation.

1 Introduction

In modeling a physical system governed by differential equation, quantifying uncertainty of the solution is an important issue and Bayesian methodology provides a natural framework in such problems Chakraborty et al. (2016); Stuart (2014); Konomi et al. (2014); Hoang et al. (2013); Holmes et al. (2012); Guha et al. (2015); Mondal et al. (2014); Efendiev et al. (2008b,a). This is a very nice contribution to develop structured models for uncertainty discretization within the forward problem and construction of efficient sequential algorithms. The authors used Gaussian process models for the unknown function as well as for the corresponding derivatives. The key component of modeling with Gaussian process is the specification of the covariance kernel. This specification could be very complex in higher dimensions and for complex multiscale problems. In the latter case, the stochastic modeling of discretization errors can be important as one can not represent all degrees of freedom. Typical subgrid basis functions representing the solutions over computational grids can not include all fine-grid information of the solution space. Some important subgrid information can be taken into account, while un-resolved scales and information can be modeled in a probabilistic fashion. The proposed framework can take this into account in a more systematic way if appropriate scales are resolved and un-resolved ones are modeled using appropriate probabilistic models. Though the applications presented in the paper are somewhat simplistic, the proposed concepts can be applied to more complex problems. In particular, with the help of appropriate subgrid models, one can obtain accurate representation of uncer-

*Main article DOI: 10.1214/16-BA1017.

[†]Department of Statistics, Texas A&M University, College Station, TX 77843, USA, bmallick@stat.tamu.edu

[‡]School of Mathematical Sciences, Fudan University, Shanghai 200433, China, kerenyangfudan@gmail.com

[§]Department of Statistics and Institute for Scientific Computation, Texas A&M University, College Station, TX 77843, USA, nguha@tamu.edu

[¶]Department of Mathematics, Texas A&M University, College Station, TX 77843, USA, efendiev@math.tamu.edu

 DOI: 10.1214/16-BA1038

tainties and the applications to partial differential equations (PDEs) can benefit from variational formulations.

2 Variational Approaches with Multiscale Basis Functions

An alternative way is to approximate the function u using basis functions and use a variational formulation of PDEs. As we mentioned, the basis functions can represent some subgrid effects or can be taken as splines or wavelets. As an illustrative example, we can consider

$$\frac{\partial u}{\partial t} = L(\kappa(x), u, \nabla u),$$

where, for example, $L(\kappa(x), u, \nabla u) = \text{div}(\kappa(x)\nabla u)$ with $\kappa(x)$ having a multiscale nature, or L is a nonlinear differential equation, which can generate subgrid effects (e.g., viscous fingering,...). In these applications, we can express the unknown solution function as

$$u(x,t) = \sum_i \beta_i(t^*)\ \phi_i(x,t),$$

where β_i's are the unknown regression coefficients defined in each computational time interval and ϕ_i are basis functions. We consider a simpler case, where ϕ_i is the spatial functions and $\beta_i(t)$ are time-varying functions. We note that it can be important to separate variables to avoid higher derivatives and introduce auxiliary variables. In the diffusion example, we can, for example, introduce $v = \nabla u$ and use additional subgrid basis functions for v, $v = \sum_i \gamma_i(t)\psi_i(x)$. Using an appropriate variational approach, one obtains

$$\frac{\partial[\beta;\gamma]}{\partial t} = A[\beta;\gamma],$$

where A can be a nonlinear operator, in general. An appropriate variational principle allows avoiding instability in spatial discretization, which could be a source of the error.

Modeling subgrid errors in each computational domain can be a challenging task. Multiscale techniques Efendiev and Hou (2009); Chung et al. (2016a,b); Calo et al. (2016) are designed to extract important features (typically non-local) in each computational grid and couple them via a global formulation. For example, if in each computational grid K (in 2D or 3D), we use basis functions ϕ_i^K, where $i = 1, \ldots, N_K$ is the number of basis functions. The "computable" part of the solution can be sought as

$$u_c(x,t) = \sum_{i,K} \beta_i^K(t)\ \phi_i^K(x),$$

where β_i^K can be found from the variational formulation. To model un-resolved scales, one can consider several approaches. One of them includes writing the solution as

$$u = u_c + u_r,$$

where u_r is the un-resolved part. The equation for u_r can be formulated and solved. This approach is similar of modeling the error Kennedy and O'Hagan (2001). Another approach (which is closer to the proposed method) is to model subgrid errors for each basis functions. These new subgrid basis functions can be adaptively defined in space and time. In this case, the solution is sought as

$$u(x,t) = \sum_{i,K} \beta_i^K(t)\ (\phi_i^K(x) + r_i^K(x,t)),$$

where $r_i^K(x,t)$ account for un-resolved scales. In this case, one can seek the solution as

$$u(x,t) = \sum_{i,K} \widetilde{\beta}_i^K(t)\ \phi_i^K(x),$$

with $\widetilde{\beta}_i^K(t)$ contain uncertainties due to un-resolved scales. This can be considered as one of the applications of the proposed method, which provides a framework.

In addition to subgrid errors, the proposed method designs efficient sequential algorithm. The solution representation can be used within a natural variational formulation for underlying PDEs, which results a model equation for β. In the iterative framework, let the discretized solution of a PDE at time $t+1$ be u^{t+1}, which is a vectorized value at grid points on the spatial domain. If we can write the solution as the following linearized form,

$$u^{t+1} = Au^t,$$

with the given initial condition u^0. Then, we can quantify the uncertainty of the solution, through specifying Gaussian distribution structure, or specifying the covariance function u_i^{t+1} and u_j^t, where j in the subscript denote the value at the j th grid point. The algorithm will be very fast due to inherent Markov structure. Furthermore, we can add the approximation error as $u^t = R\beta^t + \epsilon$ with a Gaussian distribution for ϵ. When we put a Gaussian prior on β_i, we also make u become a Gaussian distribution which could be used to quantify the uncertainty.

Next, we would like to comment on using multiscale approaches, which incorporate subgrid information into basis functions. We consider the example in Section 5.4. In this case, multiscale basis functions for 1D case can be constructed by solving local problems with boundary conditions $\phi_i(x_j) = \delta_{ij}$ (see Efendiev and Hou (2009)). One can regard these approaches as seeking solutions' values at the coarse-grid computational nodes. The information on the fine grid in this case is approximated via basis functions containing subgrid information. In a multiscale approach, we can use the Gaussian updates on coarse nodes and write the solution at $\{x, t_j\}$ (for a generalized equation for (5.6)) as

$$u(x,t_j) = \sum_{i=1}^{N_1} \alpha_i(t)\phi_i(x),$$

where $\phi_i(x)$ is the multiscale basis at coarse grid points. Therefore, by updating on α by the Gaussian kernel as given in Algorithm 8, we would be able to capture the local structure while attaining computational efficiency.

For high dimensional multiscale PDEs, a separability assumption can be used in solving the differential equation over the whole parameter domain (Yang et al. (2016)). We write the solution as

$$u(x) = \sum_{i=1}^{N} \prod_{j=1}^{M} \phi_{i,j}(x_j),$$

where $x = (x_1, x_2, \cdots, x_M)$, which can include the parameters and spatial/temporal variables. This can be used to estimate the mean function and uncertainty quantification, by assuming Gaussian processes on the separated functions $\phi_{i,j}$. We can also get probabilistic solutions for high-dimensional PDE by this separable scheme. Similar separability assumptions can be made on the Gaussian Process covariance kernel but it is difficult to evaluate its effect due to misspecification.

3 Consistency Result

The Gaussian process is characterized by the covariance kernel, which is parametrized by prior precision α^{-1} and correlation length λ. Here, α controls the variability of solution path at some grid points and λ controls the effect of other grid points through the covariance kernel. From the consistency theorem (Theorem 3.1 in Chakraborty et al. (2016)), with α^{-1} and λ going to zero, the consistency actually comes to the accuracy of the numerical solver, which is a function of grid size when we deal with smooth solutions. Otherwise, they can be related to local approximation with multiscale basis functions. Hence, as prior variance, correlation length, and grid size converge to zero, the solution converges to the unique solution of the differential equation. It can also be helpful to look into posterior consistency and convergence rate (Vollmer (2013), Barron et al. (1999)) of θ, the unknown initial condition. The posterior consistency would imply that with more observed data and dense grids the posterior distribution of θ would be concentrated in a small ball around unknown true θ^*, where the radius of the ball goes to zero at some appropriate rate given by the rate of convergence.

4 Concluding Remarks

The proposed method characterizes the uncertainty in the discretization of the solutions of differential equations. Using space–time basis functions and variational formulations, we can use the proposed concepts and construct the mean function by appropriately incorporating the subgrid effects. Furthermore, using Gaussian process structure, we can describe the uncertainty around the mean function and this allows characterizing the uncertainty of the solution. Comparing methods based on basis functions with subgrid models for un-resolved scales in the case of multiscale problems with the sequential interrogation can be an interesting exercise.

References

Barron, A., Schervish, M. J., Wasserman, L., et al. (1999). "The consistency of posterior distributions in nonparametric problems." *The Annals of Statistics*, 27(2): 536–561. MR1714718. doi: http://dx.doi.org/10.1214/aos/1018031206. 1282

Calo, V. M., Efendiev, Y., Galvis, J., and Li, G. (2016). "Randomized oversampling for generalized multiscale finite element methods." *Multiscale Modeling & Simulation*, 14(1): 482–501. MR3477310. doi: http://dx.doi.org/10.1137/140988826. 1280

Chakraborty, A., Bingham, D., Dhavala, S. S., Kuranz, C. C., Drake, R. P., Grosskopf, M. J., Rutter, E. M., Torralva, B. R., Holloway, J. P., McClarren, R. G., et al. (2016). "Emulation of numerical models with over-specified basis functions." *Technometrics*, doi: http://dx.doi.org/10.1080/00401706.2016.1164078. 1279, 1282

Chung, E., Efendiev, Y., and Hou, T. Y. (2016a). "Adaptive multiscale model reduction with generalized multiscale finite element methods." *Journal of Computational Physics*, 320: 69–95. MR3506953. doi: http://dx.doi.org/10.1016/j.jcp.2016.04.054. 1280

Chung, E., Efendiev, Y., Leung, W. T., and Li, G. (2016b). "Sparse Generalized Multiscale Finite Element Methods and their applications." *International Journal for Multiscale Computational Engineering*, 14(1). 1280

Efendiev, Y., Datta-Gupta, A., Hwang, K., Ma, X., and Mallick, B. (2008a). "Bayesian partition models for subsurface characterization." *Large-Scale Inverse Problems and Quantification of Uncertainty*, 107–122. MR2856653. 1279

Efendiev, Y., Datta-Gupta, A., Ma, X., and Mallick, B. (2008b). "Modified Markov Chain Monte Carlo method for dynamic data integration using streamline approach." *Mathematical Geosciences*, 40(2): 213–232. 1279

Efendiev, Y. and Hou, T. Y. (2009). "Multiscale finite element methods: Theory and Applications." *Springer Science & Business Media*, 2009, MR2477579. 1280, 1281

Guha, N., Wu, X., Efendiev, Y., Jin, B., and Mallick, B. K. (2015). "A variational Bayesian approach for inverse problems with skew-t error distributions." *Journal of Computational Physics*, 301: 377–393. MR3402736. doi: http://dx.doi.org/10.1016/j.jcp.2015.07.062. 1279

Hoang, V. H., Schwab, C., and Stuart, A. M. (2013). "Complexity analysis of accelerated MCMC methods for Bayesian inversion." *Inverse Problems*, 29(8): 085010. MR3084684. doi: http://dx.doi.org/10.1088/0266-5611/29/8/085010. 1279

Holmes, C., Denison, D. T., Ray, S., and Mallick, B. (2012). "Bayesian prediction via partitioning." *Journal of Computational and Graphical Statistics*. 14(4): 811–830, doi: http://dx.doi.org/10.1198/106186005X78107. 1279

Kennedy, M. C. and O'Hagan, A. (2001). "Bayesian calibration of computer models." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(3): 425–464. MR1858398. doi: http://dx.doi.org/10.1111/1467-9868.00294. 1281

Konomi, B. A., Sang, H., and Mallick, B. K. (2014). "Adaptive Bayesian nonstationary modeling for large spatial datasets using covariance approximations." *Journal of Computational and Graphical Statistics*, 23(3): 802–829. MR3224657. doi: http://dx.doi.org/10.1080/10618600.2013.812872. 1279

Mondal, A., Mallick, B., Efendiev, Y., and Datta-Gupta, A. (2014). "Bayesian Uncertainty Quantification for Subsurface Inversion Using a Multiscale Hierarchical Model." *Technometrics*, 56(3): 381–392. MR3238075. doi: http://dx.doi.org/10.1080/00401706.2013.838190. 1279

Stuart, A. (2014). "Uncertainty quantification in Bayesian inversion." *ICM2014. Invited Lecture*. 1279

Vollmer, S. J. (2013). "Posterior consistency for Bayesian inverse problems through stability and regression results." *Inverse Problems*, 29(12): 125011. MR3141858. doi: http://dx.doi.org/10.1088/0266-5611/29/12/125011. 1282

Yang, K., Guha, N., Efendiev, Y., and Mallick, B. (2016). "Bayesian and Variational Bayesian approaches for flows in heterogenous random media." arXiv:1611.01213. 1282