An Efficient Bayesian Computational Method Using Scalable Solvers For Stochastic PDEs

Sudhi Sharma [1] Ajay Kumar Verma [1] Pierre Jolivet [2] Victorita Dolean Maini [3] Abhijit Sarkar [1]

[1]Carleton University [2]Sorbonne Université, CNRS, Paris [3]Technical University of Eindhoven, The Netherlands

Abstract

The state or combined state and parameter estimation of extreme scale nonlinear stochastic computational models (generally arising from the discretization of stochastic PDEs), sampling based non-Gaussian filters (e.g. particle or ensemble Kalman filters [1]) can become computationally impractical. To alleviate the computational challenges at the forecast (prediction) step of nonlinear filtering for such models, a two- grid scalable solver using (sampling-free) intrusive stochastic Galerkin method is developed that leverages two-level Schwarz method for geometric decomposition and the algebraic multigrid solver (AMG) for the corresponding coarse problem [2]. This two-grid solver demonstrates scalable performance for high resolution spatial and temporal discretizations, a large number of input random variables and higher order stochastic (i.e. polynomial chaos) expansion terms for the output of nonlinear stochastic PDEs [3] Within the framework of the polynomial chaos Kalman filter, the distributed implementation of the update (analysis) step will be developed that exploits this two-grid scalable solver to condition the system state using observational data. We also investigate a specific case where only spatially averaged aggregated data in sub-regions are available. Such cases arise in the context of mechanistic geospatial modeling of infectious disease spread whereby the testing data for infected population averaged over public health units (defining a sub-region) are reported by public health agencies.

Bayesian Estimation using Nonlinear Filtering

Bayesian non-linear filtering involves propagating the random system state, parameters and model error in the forecast (prediction step). This is followed by assimilation (update step) that involves conditioning the system state, parameters and model error (using sparse and noisy observational data).

Model

$$\mathbf{x}_{k+1} = g_k(\mathbf{x_k}, \boldsymbol{\phi}, q_k); \qquad k = 0, 1, \cdots n$$

where $\mathbf{x}_{k+1}, \boldsymbol{\phi}, q_k$ are the random state, parameters and model error vectors respectively.

Measurement

$$\mathbf{d}_j = h_j(\mathbf{x}_{d(j)}, \boldsymbol{\phi}, \boldsymbol{\epsilon}_j); \qquad j = 0, 1, \cdots m$$

where $\mathbf{d}_j$ and $\boldsymbol{\epsilon}_j$ are the observational data and measurement noise respectively.

Estimation Problem

Jointly estimate the state $[\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n]$ and parameters $\boldsymbol{\phi}$.

For an effective calibration of parameters of the model, an efficient forecast and update step is essential [4]. Hence, we consider first the state estimation problem for high resolution PDE models tackled using a scalable sampling-free method with locally averaged measurements (relevant for infectious disease modelling).

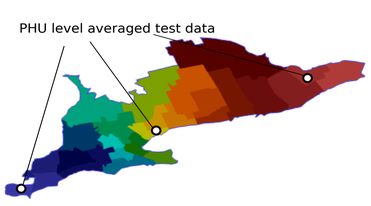


Figure 1. Infection data from Public Health Units (PHUs) in Southern Ontario

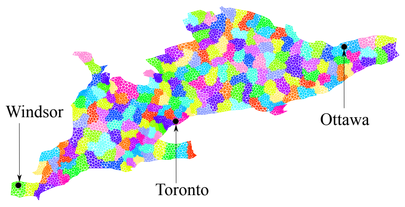


Figure 2. Computational domain of Southern Ontario partitioned into 200 subdomains

Forecast : Stochastic Galerkin Method

The sampling-free (intrusive) approach for forward uncertainty quantification (forecast) involves representing the random parameters of the model using a truncated polynomial chaos expansion (PCE) and the associated PC (deterministic) coefficients. For simplicity, a nonlinear Poisson problem with random diffusion coefficient $c(\mathbf{x}, \theta)$ is used to demonstrate the stochastic Galerkin method below.

$$-\nabla\cdot\Big(c(\mathbf{x},\theta)(1+\alpha u(\mathbf{x},\theta))\nabla u(\mathbf{x},\theta)\Big) = f(\mathbf{x}) \;\text{ on }\; \Omega\times\Phi$$

Random input and output terms are expanded as,

$$c(\mathbf{x},\theta) = \sum_{i=0}^{\infty}\bar{c}_i(\mathbf{x})\Psi_i(\boldsymbol{\xi}) \approx \sum_{i=0}^{L}\bar{c}_i(\mathbf{x})\Psi_i(\boldsymbol{\xi})$$
$$u(\mathbf{x},\theta) = \sum_{j=0}^{N}\bar{u}_j(\mathbf{x})\Psi_j(\boldsymbol{\xi})$$

where $\bar{c}_i(\mathbf{x}), \bar{u}_j(\mathbf{x})$ are the deterministic chaos coefficients and $\Psi_i(\boldsymbol{\xi})$ are the orthogonal functions of random variables $\boldsymbol{\xi}$. Galerkin projection leads to the following equations.

$$\int_\Omega \sum_{i=0}^{L}\sum_{k=0}^{N}\langle\Psi_i\Psi_k\Psi_l\rangle\,(\bar{c}_i(\mathbf{x})\nabla\bar{u}_k(\mathbf{x})\cdot\nabla v_l(\mathbf{x}))\,d\mathbf{x}$$
$$+\int_\Omega \sum_{i=0}^{L}\sum_{j=0}^{N}\sum_{k=0}^{N}\langle\Psi_i\Psi_j\Psi_k\Psi_l\rangle\alpha\,(\bar{c}_i(\mathbf{x})\bar{u}_j(\mathbf{x})\nabla\bar{u}_k(\mathbf{x})\cdot\nabla v_l(\mathbf{x}))\,d\mathbf{x}$$
$$=\int_\Omega f(\mathbf{x})v_l(\mathbf{x})\langle\Psi_l\rangle\,d\mathbf{x}, \qquad l=0,1,2,\ldots N$$

Forecast : Multi-level Scalable Solver

A multi-level overlapping domain decomposition-based solver is developed by combining a restricted additive Schwarz method (RAS) in the fine-grid and algebraic multigrid for the coarse grid.

Overlapping Domain Decomposition

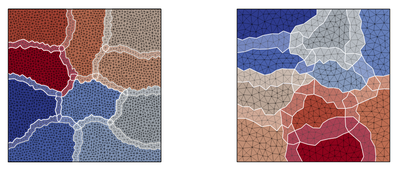


Figure 3. An overlapped decomposition : fine grid (left) and coarse grid (right).

Two-grid Preconditioner and Scalability

$$\mathbf{M}_2^{-1} = \mathbf{M}_1^{-1}\mathbf{P} + \mathbf{Q}\mathbf{P}\mathbf{Q}^{-1}\mathbf{M}_1^{-1} + \mathbf{Q} - \mathbf{M}_1^{-1}\mathbf{P}\mathbf{A}\mathbf{M}_1^{-1}$$
$$\mathbf{P} = (\mathbf{I} - \mathbf{A}\mathbf{Q}); \quad \mathbf{Q} = \mathbf{R}_0^T\mathbf{A}_c^{-1}\mathbf{R}_0; \quad \mathbf{M}_1^{-1} = \sum_{i=1}^{N}\mathbf{R}_i^T\mathbf{D}_i(\mathbf{R}_i\mathbf{A}\mathbf{R}_i^T)^{-1}\mathbf{R}_i.$$

where $\mathbf{M}_1^{-1}$ is the fine-grid preconditioner (RAS), $\mathbf{R}_i, \mathbf{R}_i^T$ are the restriction and prolongation operators, $\mathbf{A}_c$ is the coarse grid and $\mathbf{D}$ is a Boolean partition of unity matrix.

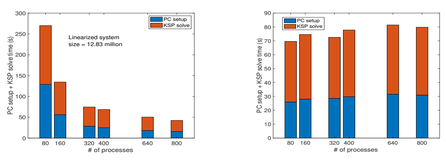


Figure 4. Strong and weak parallel scalability (max system size of 32 million) with respect to mesh size for stochastic nonlinear Poisson problem

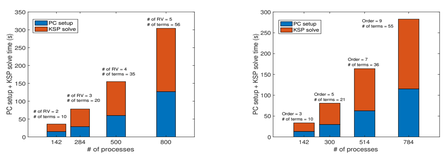


Figure 5. Weak scalability with respect to increasing number of random variables and order of expansion (max system size of 32 million).

Update : Polynomial Chaos Kalman Filter (PCKF)

Polynomial Chaos Kalman Filter [5]

The polynomial chaos expansion of the random state (after the forecast step) and measurement vector (with linear measurement operator $\mathbf{H}$) can be written as,

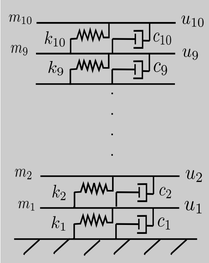
$$\mathbf{x}^f = \sum_{i=0}^{N}\mathbf{x}_i^f\Psi_i(\theta); \quad \mathbf{d} = \sum_{i=0}^{N}\mathbf{d}_i\Psi_i(\theta)$$

$$\mathcal{A}^f = \left[\mathbf{x}_0^f, \mathbf{x}_1^f, \cdots, \mathbf{x}_N^f\right]; \quad \mathcal{A}^f \in \mathcal{R}^{n\times(N+1)}$$
$$\mathcal{D} = [\mathbf{d}_0, \mathbf{d}_1, \cdots, \mathbf{d}_N]; \quad \mathcal{D} \in \mathcal{R}^{m\times(N+1)}$$
$$\mathcal{A}^a = \mathcal{A}^f + \mathbf{P}^f\mathbf{H}^T\left(\mathbf{H}\mathbf{P}^f\mathbf{H}^T + \Gamma\right)^{-1}\left(\mathcal{D} - \mathbf{H}\mathcal{A}^f\right)$$

where $\mathbf{P}^f, \Gamma$ are the state and measurement noise covariance matrix respectively, computed as:

$$\mathbf{P}^f = \sum_{i=0}^{N}\mathbf{x}_i^f\mathbf{x}_i^{fT}\langle\Psi_i^2(\theta)\rangle; \quad \Gamma = \sum_{i=0}^{N}\mathbf{d}_i\mathbf{d}_i^T\langle\Psi_i^2(\theta)\rangle$$

Average Measurement for a 10 dof System



$$\mathbf{u}^f = u^0 + u^1\xi_1 + u^2\xi_2 + u^3(\xi_1^2 - 1) + u^4(\xi_1\xi_2) + u^5(\xi_2^2 - 1)$$

$$u^i = \begin{Bmatrix} u_1^i \\ u_2^i \\ \vdots \\ u_9^i \\ u_{10}^i \end{Bmatrix}; \quad i = 0, 1\cdots N$$

Measurement Cases

Case 1: $\mathbf{d}_0 = \{d_1^0\} = \{\sum_{k=1}^{10}u_k^0 + \sigma_1\epsilon_1\}$

Case 2: $\mathbf{d}_0 = \begin{Bmatrix} d_1^0 \\ d_2^0 \end{Bmatrix} = \begin{Bmatrix} \sum_{k=1}^{5}u_k^0 + \sigma_1\epsilon_1 \\ \sum_{k=6}^{10}u_k^0 + \sigma_2\epsilon_2 \end{Bmatrix}$

Case 3: $\mathbf{d}_0 = \begin{Bmatrix} d_1^0 \\ d_2^0 \\ d_3^0 \end{Bmatrix} = \begin{Bmatrix} \sum_{k=1}^{3}u_k^0 + \sigma_1\epsilon_1 \\ \sum_{k=4}^{6}u_k^0 + \sigma_2\epsilon_2 \\ \sum_{k=7}^{10}u_k^0 + \sigma_3\epsilon_3 \end{Bmatrix}$

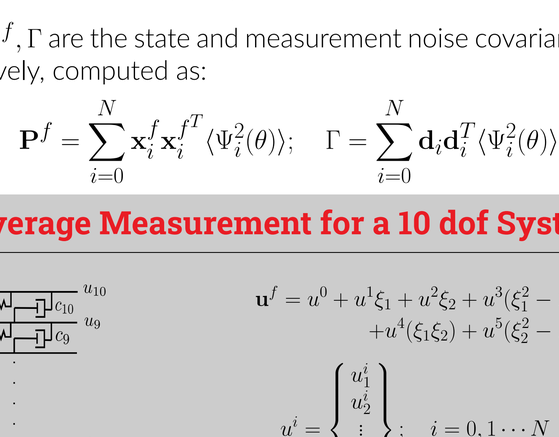


Figure 6. Prior and posterior for cases 1 (top row), 2 (middle row) and 3 (bottom row).

Stochastic Geospatial COVID Model

A five-compartment model (susceptible-exposed-infected-recovered-deceased; s-e-i-r-d) with random diffusion coefficients is considered as follows [2]:

$$\partial_t s(\mathbf{x},\theta) = \nabla\cdot(N\bar{\nu}_S(\theta)\nabla s) - \left(1-\frac{A}{N}\right)\beta_I\, si - \left(1-\frac{A}{N}\right)\beta_E\, se$$
$$\partial_t e(\mathbf{x},\theta) = \nabla\cdot(N\bar{\nu}_E(\theta)\nabla e) + \left(1-\frac{A}{N}\right)\beta_I\, si + \left(1-\frac{A}{N}\right)\beta_E\, se - \sigma e - \gamma_E e$$
$$\partial_t i(\mathbf{x},\theta) = \nabla\cdot(N\bar{\nu}_I(\theta)\nabla i) + \sigma e - \gamma_R i - \gamma_D i$$
$$\partial_t r(\mathbf{x},\theta) = \nabla\cdot(N\bar{\nu}_R(\theta)\nabla r) + \gamma_E e + \gamma_R i$$
$$\partial_t d(\mathbf{x},\theta) = \gamma_D i$$

where N is the total population density defined as,

$$N(\mathbf{x},\theta) = s(\mathbf{x},\theta) + e(\mathbf{x},\theta) + i(\mathbf{x},\theta) + r(\mathbf{x},\theta) + d(\mathbf{x},\theta).$$

Averaged Infection data for Southern Ontario

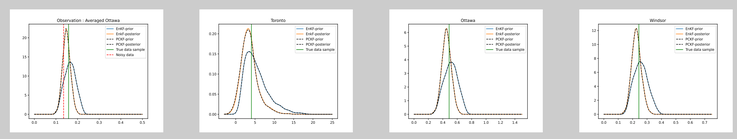


Figure 7. Prior and posterior with spatially averaged infection data from Ottawa PHU.

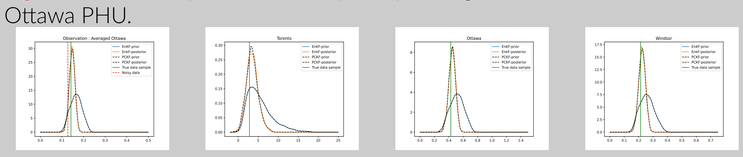


Figure 8. Prior and posterior with 3 spatially averaged infection data from Ottawa, Toronto and Windsor PHUs.

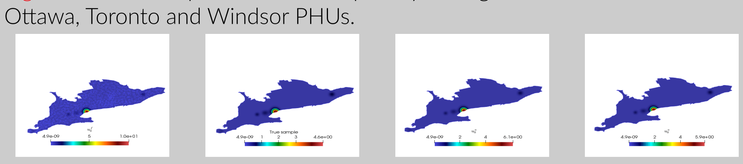


Figure 9. Prior mean, true sample and posterior mean for PCKF and EnKF (3 average measurements)

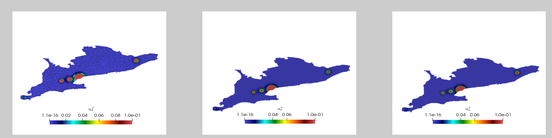


Figure 10. Prior and posterior standard deviations for PCKF and EnKF (3 average measurements)

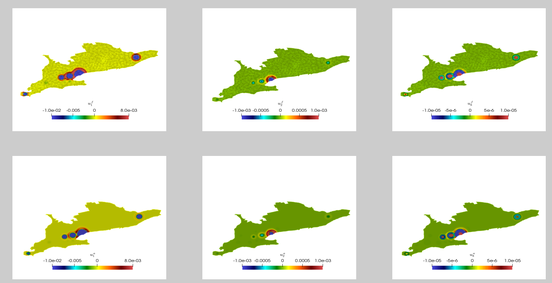


Figure 11. Prior (top row) and posterior (bottom row) polynomial chaos coefficients $(1, 2, 8)$ for 3 average measurements.

Concluding Remarks

A sampling-free approach to Bayesian nonlinear filtering for a dynamical system is proposed as a scalable alternative to sampling-based methods. Forecast step utilizes a multi-level scalable solver using stochastic Galerkin method, combining overlapping domain decomposition in the fine grid and algebraic multigrid for the coarse grid. This solver enables the use of large number of random variables and high resolution spatial/temporal discretizations. The update step is achieved using the polynomial chaos Kalman filter which can handle non-Gaussian states. Furthermore, numerical experiments demonstrate the efficiency of PCKF against ensemble Kalman filter. Numerical illustration considers the geospatial spread of COVID in Southern Ontario.

References

[1] Geir Evensen. Sequential data assimilation with a nonlinear quasi-geostrophic model using monte carlo methods to forecast error statistics. *Journal of Geophysical Research: Oceans*, 99(C5):10143–10162, 1994.

[2] Sudhi Sharma, Victorita Dolean, Pierre Jolivet, Brandon Robinson, Jodi D. Edwards, Tetyana Kendzerska, and Abhijit Sarkar. Scalable computational algorithms for geospatial covid-19 spread using high performance computing. *Mathematical Biosciences and Engineering*, 20(8):14634–14674, 2023. ISSN 1551-0018. doi: 10.3934/mbe.2023655. URL https://www.aimspress.com/article/doi/10.3934/mbe.2023655.

[3] Sudhi Sharma, Pierre Jolivet, Victorita Dolean, and Abhijit Sarkar. Multilevel scalable solvers for stochastic linear and nonlinear problems. *arXiv preprint arXiv:2310.14649*, 2023.

[4] Brandon Robinson, Philippe Bisaillon, Jodi D Edwards, Tetyana Kendzerska, Mohammad Khalil, Dominique Poirel, and Abhijit Sarkar. A bayesian model calibration framework for stochastic compartmental models with both time-varying and time-invariant parameters. *Infectious Disease Modelling*, 2024.

[5] George Saad and Roger Ghanem. Characterization of reservoir simulation models using a polynomial chaos-based ensemble kalman filter. *Water Resources Research*, 45(4), 2009.