



# Stochastic PDE model identification from sparse noisy data

Fei Lu<sup>1</sup>, Nils Weitzel<sup>2</sup> and Adam Monahan<sup>3</sup>

<sup>1</sup>Department of Mathematics, Johns Hopkins University

<sup>2</sup>Institut für Geowissenschaften und Meteorologie, Universität Bonn

<sup>3</sup>School of Earth and Ocean Sciences, University of Victoria



## Introduction

**Ill-posedness of the inverse problem** and **sparse noisy data** are two major challenges in the modeling of high-dimensional spatiotemporal processes. We introduce a Bayesian inference method with a **strongly regularized posterior** to overcome these challenges, enabling **joint state-parameter estimation** with uncertainty quantification. We demonstrate the method on a physically motivated nonlinear stochastic partial differential equation (**SPDE**) arising from paleoclimate construction.

## The non-linear SEBM

We consider a stochastic energy balance model (SEBM) deduced from the process based a deterministic model of Fanning and Weaver (1996). It contains a diffusive transport term, an external forcing term, a linear term that models atmosphere-ocean fluxes, a fourth order term that corresponds to outgoing long-wave radiation, and a stochastic forcing:

$$\partial_t u(t, x) - v \Delta u(t, x) = (\theta_0 + \theta_1 u + \theta_2 u^4) + f(t, x).$$

The stochastic forcing  $f(t, x)$  is smooth in space and white in time, and it is represented by the Gaussian Markov random fields approximation of Matern noise [Lindgren et al. 2011].

The data are noisy observations at sparse nodes of the spatiotemporal process. We assume that the noises, representing measurement errors, are independent identically distributed Gaussian.

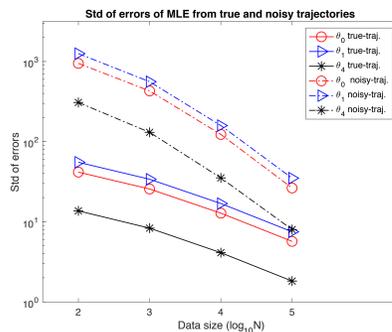
### Parameter settings (Prior)

We derive upper/lower bounds of parameters from physical principles, and use them to set a prior probability density: a Gaussian distribution with mean = (upper bound + lower bound)/2 and with std = (upper bound - lower bound)/2.

Bound	$\theta_0$	$\theta_1$	$\theta_2$
Lower	27.64	-25.46	-6.00
Upper	32.57	-22.70	-4.80

### Joint state and parameter estimation:

we estimate jointly the states and parameters.



### Inverse problem: ill-posed.

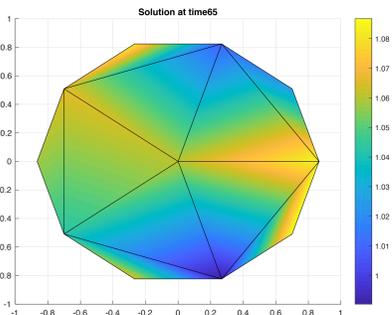
Due to the strong correlation between the terms in the parametric form, the inverse problem is ill-posed. Numerically, this means that the Fisher information matrix is ill-conditioned. As a result, large estimation errors appear for both perfect and noisy observations, leading to estimators out of physical ranges (see in the right figure) and failure of model identification.

## Numerical integration of the SEBM

The equation is discretized in space by finite element, and integrated in time by a semi-backward Euler scheme so as to be stable.

### Simplified protocol model settings:

- Consider here a tiny mesh with only 12 nodes and 20 elements on the unit sphere.
- A typical solution is shown in the right figure. The temperature oscillates around one (non-dimensionalized).
- Data: at each time, only 6 out of 12 nodes are observed (sparse), with additive independent Gaussian noises.



## Bayesian inference with strongly regularized posterior

$$\text{Posterior} = \text{Prior} \times \text{Normalized Likelihood}$$

$$p(\theta, u | \mathbf{y}) = p(\theta) \left( \frac{p(u | \theta) p(\mathbf{y} | u, \theta)}{p(\mathbf{y})} \right)^{1/N}$$

- The strongly regularized posterior normalizes the likelihood so that the prior can sufficiently constrain the parameters into the physical range. We also enforce a prior for the states according to the climatological distribution. The posterior **quantifies the uncertainty** in state and parameter estimation.
- The posterior is of extremely **high-dimensional**. For example, for a tiny mesh with 12 nodes and 100 time steps of observations, the discrete state variable is of dimension (1200 +3). This number grows to millions as the mesh refines and time steps increases. A major challenge is to generate samples for such a high dimensional distribution.

## Particle Markov chain Monte Carlo

A family of sampling methods that combines the advantages of both Sequential Monte Carlo (SMC) and Markov Chain Monte Carlo [Andrieu, Doucet & Holenstein 2010]:

- The chain: moves from one SMC ensemble to another
  - Each chain step: sample from the SMC ensemble
  - With the target distribution as invariant measure (improving SMC)
- Transition through **conditional SMC**
  - Previous chain as a reference path in the next SMC. The reference path is retained over the resampling steps, and interacts with other paths by contributing its weight.
- Particle Gibbs with Ancestor Sampling (PGAS)** [Lindstein, Jordan & Schon 2014]
  - Sample the ancestor of the reference path to increase mixing of the chain
  - Gibbs sampling for the parameter using the reference path

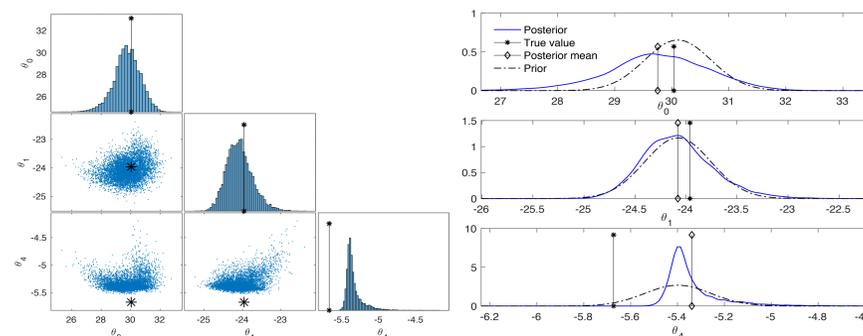
## Numerical experiments

### Settings:

- Observe 100 time steps, each time step observing 6 nodes. The noise to signal ratio is 1%.
- PGAS uses 5 particles for the SMC and a Markov chain of length 10000.

### Parameter estimation:

The posterior of the parameters is represented by the samples, shown below in a scatter plot and marginal density plots. Large uncertainty presents in the posterior, due to the ill-posedness and the need of regularization. The true values of the first two parameters are in the high probability region, but the true value of the last parameter is of low posterior possibility.

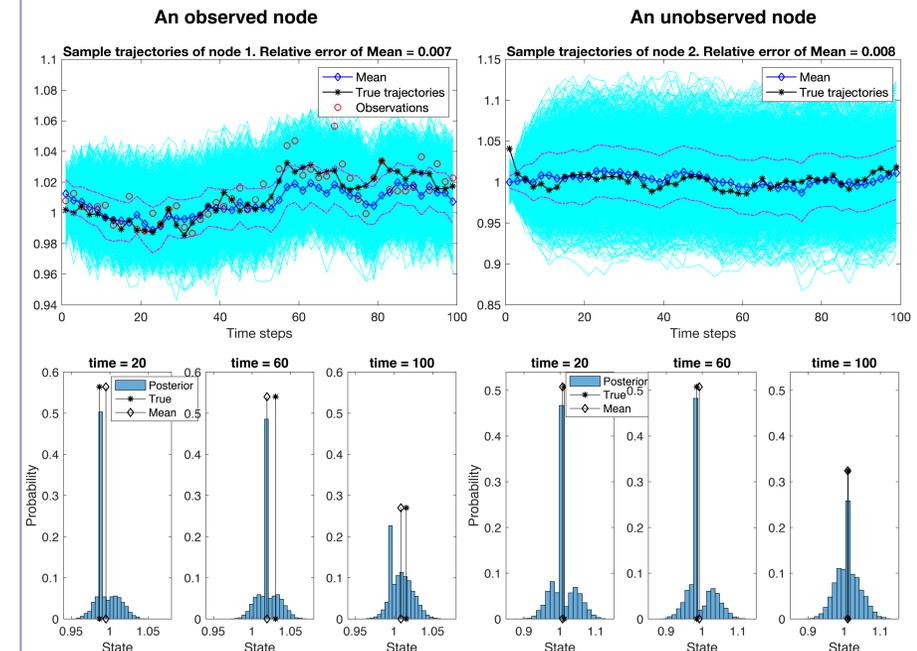


	$\theta_0$	$\theta_1$	$\theta_2$
Mean	-0.32	0.02	0.03
Std	0.61	0.42	0.21

The left table shows means and standard deviations of the errors of the maximum of posterior (MAP) in 100 independent simulations. Relatively large uncertainties in MAP's are seen, due to the ill-posedness. Therefore, it is important to quantify the uncertainty by the posterior.

### State estimation:

The state trajectory is well estimated by the sample ensemble. The figure below shows the ensemble of trajectories at an observed node (left) and an unobserved node (right), with marginal distributions at different time instants (bottom row). The ensemble trajectories (in cyan) spread around the true trajectory, and the ensemble mean (blue diamond with one standard deviation in magenta) is close to the true path for most of the time. The ensemble mean filtered out the noise in observations (red circles), reducing the relative error by 30%.



The mean and standard deviation of the relative errors in trajectory estimation in 100 independent experiments is 1.14% (0.41%) over all observed and unobserved nodes.

## Summary

We introduced a strongly regularized posterior that overcomes the ill-posedness in parameter estimation for an SPDE model from sparse noisy data, and investigate joint state-parameter estimation of a stochastic energy balance model arising in paleoclimate field reconstructions.

- Accurate state estimation from sparse and noisy data is achieved
- Large uncertainty presents in the parameter estimation, due to the ill-posedness

### Future work

- Re-parametrization or nonparametric inference to avoid ill-posedness
- Modify model and apply the method for cases with field data

## References

- Lu, F., Weitzel, N. & Monahan, A. (2019). Joint state-parameter estimation of a nonlinear stochastic energy balance model from sparse noisy data. <http://arxiv.org/abs/1904.05310>
- Fanning, A. F., & Weaver, A. J. (1996). An atmospheric energy-moisture balance model: Climatology, interpentadal climate change, and coupling to an ocean general circulation model. *Journal of Geophysical Research: Atmospheres*, 101(D10), 15111-15128.
- Lindgren, F., Rue, H., & Lindström, J. (2011). An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 73(4), 423-498.
- Andrieu, C., Doucet, A., & Holenstein, R. (2010). Particle markov chain monte carlo methods. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 72(3), 269-342.
- Lindsten, F., Jordan, M. I., & Schön, T. B. (2014). Particle Gibbs with ancestor sampling. *The Journal of Machine Learning Research*, 15(1), 2145-2184.