Using Surrogate Models to Accelerate Bayesian Inverse Uncertainty Quantification

James Rynn
with Simon Cotter, Catherine Powell (University of Manchester)
& Louise Wright (National Physical Laboratory)

james.rynn@manchester.ac.uk

30/06/17

Overview

Bayesian Inverse Problems

Find the unknown θ given n_{obs} observations z, satisfying

$$z = \mathcal{G}(\theta) + \eta, \quad \eta \sim \mathcal{N}(0, \Sigma),$$

where

- $z \in \mathbb{R}^{n_{\text{obs}}}$ is a given vector of **observations**,
- $\mathcal{G} : \Theta \to \mathbb{R}^{n_{\text{obs}}}$ is the **observation operator**,
- $\theta \in \Theta$ is the **unknown**,
- $\eta \in \mathbb{R}^{n_{\text{obs}}}$ is a vector of **observational noise**.

We treat this as a probabilistic problem and search for a posterior distribution for θ.

Bayesian Inverse Problems

In the finite-dimensional case, from Bayes' Theorem we have

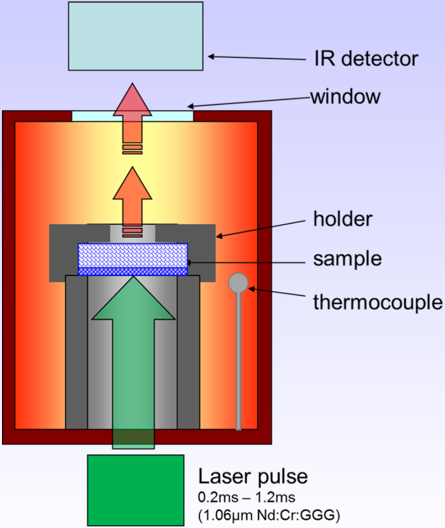
$$\begin{aligned} \pi(\theta|z) &\propto L(z|\theta)\ \pi_0(\theta) \\ &\propto \exp\left(-\frac{1}{2}\|z-\mathcal{G}(\theta)\|_\Sigma^2\right)\ \pi_0(\theta). \end{aligned}$$

Markov Chain Monte Carlo (MCMC) Methods

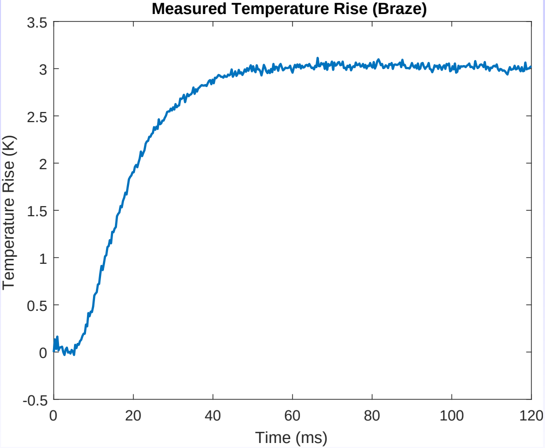
- We know $\pi(\theta|z)$ up to a constant of proportionality.
- Use MCMC algorithm to generates samples $\theta_1, \theta_2, \dots, \theta_M$ from the posterior distribution.
- Use these samples to construct Monte Carlo estimates of quantities of interest (means, variances and/or probabilities),
- e.g.

$$\mathbb{E}_\pi[\phi] = \int_\Theta \phi(\theta)\pi(\theta|z)\mathrm{d}\theta \approx \frac{1}{M}\sum_{i=1}^{M}\phi(\theta_i).$$

Motivation


IR detector
window
holder
sample
thermocouple
Laser pulse
0.2ms – 1.2ms
(1.06µm Nd:Cr:GGG)

Motivation


Measured Temperature Rise (Braze)
3.5
3
2.5
2
1.5
1
0.5
0
-0.5
Temperature Rise (K)
0
20
40
60
80
100
120
Time (ms)

Motivation

Possible unknowns:

- λ — thermal conductivity,
- I — laser intensity,
- k — boundary condition parameter,
- σ — standard deviation of measurement noise.

Example

Consider the one-dimensional steady state heat equation,

$$-\frac{\mathrm{d}}{\mathrm{d}x}\left(\lambda\frac{\mathrm{d}u}{\mathrm{d}x}(x)\right) = 1, \quad x \in [0, H],$$

with homogeneous Dirichlet boundary conditions,

$$u(0) = u(H) = 0,$$

where $\lambda = e^{\theta}$ is the unknown thermal conductivity.

We wish to find a posterior distribution for λ (equivalently θ), given observations of $u(x)$ at $x_1, x_2, \ldots, x_{n_{obs}} \in [0, H]$.

Example

Here, our observation operator $\mathcal{G}$ is of the form

$$\mathcal{G}(\theta) = (u(x_1;\theta), u(x_2;\theta), \ldots, u(x_{n_{obs}};\theta))^T,$$

and approximated by $\mathcal{G}_h$ given by

$$\mathcal{G}_h(\theta) = (u_h(x_1;\theta), u_h(x_2;\theta), \ldots, u_h(x_{n_{obs}};\theta))^T,$$

where u_h is the finite element solution to the ODE on a mesh of width h.

Note: For each value of θ, to evaluate $\mathcal{G}_h$ we are required to compute a FEM solve.

Random Walk Metropolis Hastings Algorithm (FEM)

Algorithm 1: RWMH Algorithm

set initial state $X^{(0)} = \theta_0$
for $m = 1, 2, \ldots, M$ **do**
  draw proposal
  evaluate likelihood by **computing** $\mathcal{G}_h$ **(expensive!)**
  compute acceptance probability α
  accept proposal with probability α
output chain $X = (\theta_0, \theta_1, \ldots, \theta_M)$

Here $M \gg 10^5$.

Results

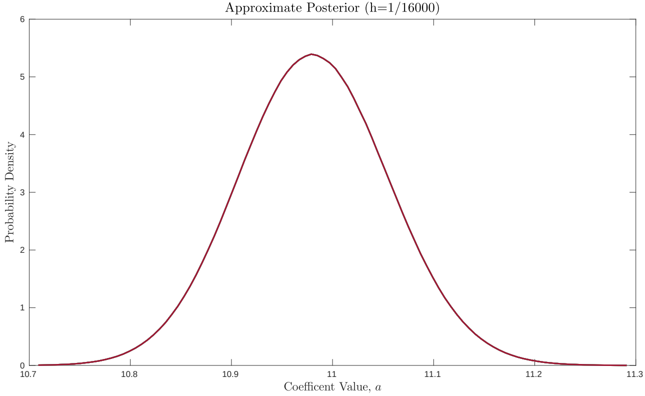


Figure: Approximate posterior density π_h with $h = 1/16000$ from 160 million samples.

Stochastic Galerkin Finite Element Method

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and consider the problem

$$-\frac{\mathrm{d}}{\mathrm{d}x}\left(e^{\theta(\omega)}\frac{\mathrm{d}u}{\mathrm{d}x}(x,\omega)\right) = 1, \quad x \in [0, H], \quad \omega \in \Omega,$$

with homogeneous Dirichlet boundary conditions,

$$u(0,\omega) = u(H,\omega) = 0, \quad \omega \in \Omega.$$

Assuming θ is of the form

$$\theta(\omega) = \theta(\xi(\omega)),$$

we can transform this into a parametric equation on $[0, H] \times \xi(\Omega)$.

Stochastic Galerkin Finite Element Method

Parametric form:

$$-\frac{\mathrm{d}}{\mathrm{d}x}\left(e^{\theta(y)}\frac{\mathrm{d}u}{\mathrm{d}x}(x,y)\right) = f(x), \quad x \in [0,H], \quad y \in \Gamma := \xi(\Omega),$$

with homogeneous Dirichlet boundary conditions,

$$u(0,y) = u(H,y) = 0, \quad y \in \Gamma.$$

Construct a stochastic Galerkin FEM solution u_{hk} on a finite dimensional subspace of $L^2(\Gamma, H^1_g(D)) \cong L^2(\Gamma) \otimes H^1_0(D)$ of size $(k+1) \times N_h$.

$$\mathcal{G}_{hk}(y) = \left(u_{hk}(x_1,y), u_{hk}(x_2,y), \ldots, u_{hk}(x_{n_{\text{obs}}},y)\right)^T.$$

Random Walk Metropolis Hastings Algorithm (SGFEM)

Algorithm 2: RWMH Algorithm with SGFEM Surrogate

compute SGFEM solution u_{hk}
set initial state $X^{(0)} = \theta_0$
for $m = 1, 2, \dots, M$ **do**
  draw proposal
  evaluate likelihood by **evaluating** $\mathcal{G}_{hk}$ (**cheap!**)
  compute acceptance probability α
  accept proposal with probability α
output chain $X = (\theta_0, \theta_1, \dots, \theta_M)$

Here $M \gg 10^5$.

Posterior Convergence in k (Polynomial Degree)

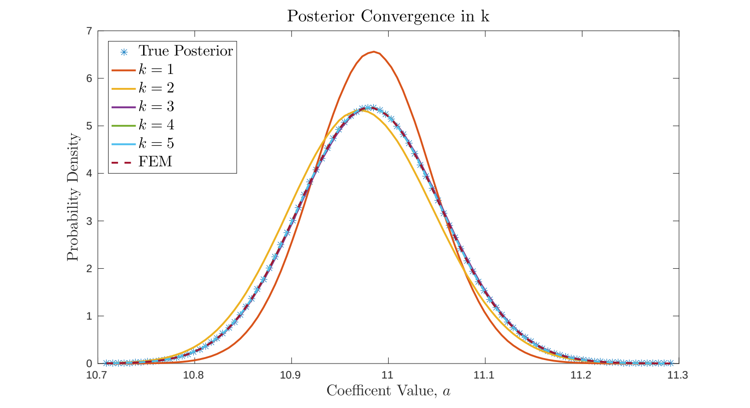


Figure: Approximate posterior densities π_{hk} with $h = 1/16000$ from 160 million samples with various values of k along with corresponding π_h produced using standard FEM approach.

Posterior Convergence in M (Number of Samples)

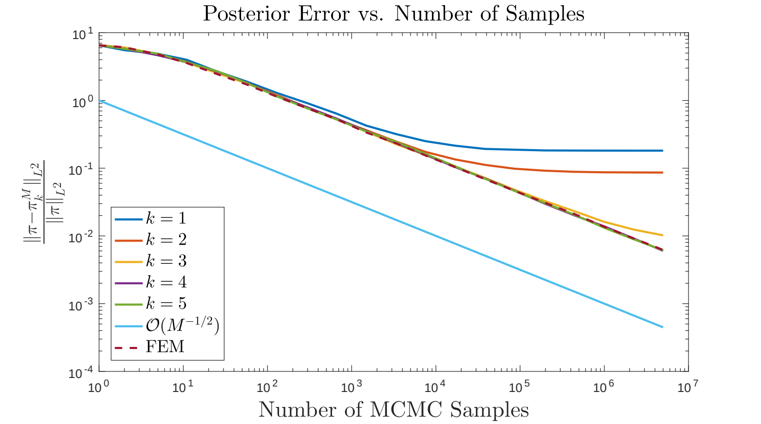


Figure: Relative L^2 errors in the approximate posteriors π_{hk} (for various k) and π_h with $h = 1/16000$.

Time Saving

Time to compute **160 million MCMC samples** using MH algorithm on a (fine) mesh of width $h = 1/16000$:

Standard FEM approach: $\approx$ **40 hours**,

SGFEM surrogate approach ($k = 5$): $\approx$ **10 minutes**.

Example 2D: Forward Problem

Consider the steady state heat equation with mixed boundary conditions and discontinuous unknown coefficient λ:

$$\begin{aligned} -\nabla \cdot (\lambda(x)\nabla u(x)) &= 1, \quad & x \in D := (0,1) \times (0,1) \subset \mathbb{R}^2, \\ u(x) &= 0 & x \in \{0\} \times (0,1), \\ u(x) &= 1 & x \in \{1\} \times (0,1), \\ \nabla u(x) \cdot n(x) &= 0 & x \in \{0,1\} \times (0,1). \end{aligned}$$

Here, $\lambda \colon (0,1) \times (0,1) \to \mathbb{R}$ is given by

$$\lambda(x) = \begin{cases} \theta, & x \in D \backslash (0.25, 0.75) \times (0.25, 0.75), \\ \lambda_0, & x \in (0.25, 0.75) \times (0.25, 0.75), \end{cases}$$

where λ_0 is known.

Example 2D: Forward Problem

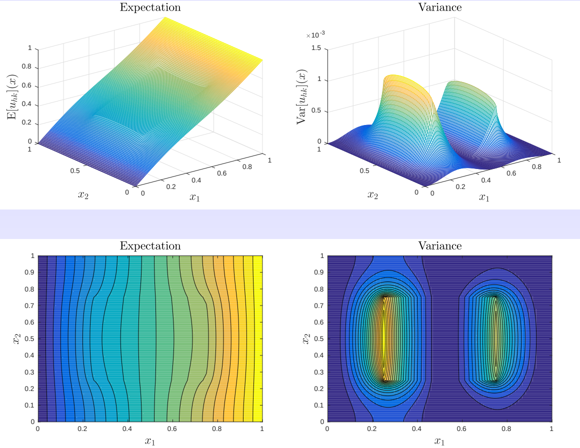


Figure: $\mathbb{E}[u_{hk}](x)$ and $\mathrm{Var}[u_{hk}](x)$: $h = 2^{-7}$ and $k = 4$.

Example 2D: Inverse Problem

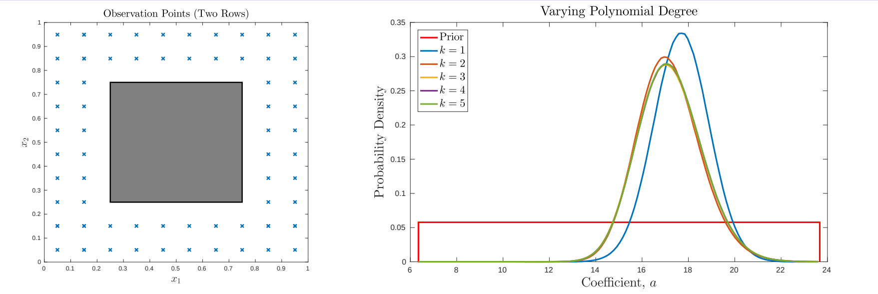


Figure: (L) Observation Points. (R) Approximate posterior densities with $h = 2^{-7}$ from 160 million samples for various values of k.

Time taken to produce **160 million samples** is $\approx$ **9 minutes**.

Future Work

- More realistic forward problem:
  - time-dependent PDE
  - multiple random variables
- More sophisticated MCMC algorithm
- Error analysis

References

I. Babuška, R. Tempone, and G. Zouraris, *Galerkin Finite Element Approximations of Stochastic Elliptic Partial Differential Equations*, SIAM Journal on Numerical Analysis, 42 (2004), pp. 800–825.

S. Brooks, A. Gelman, G. Jones, and X.-L. Meng, *Handbook of Markov Chain Monte Carlo*, Chapman and Hall/CRC, 2011.

A. Stuart, *Inverse Problems: A Bayesian Perspective*, Acta Numer., 19 (2010), pp. 451–559.

L. Wright, L. Chapman, and D. Partridge, *Laser Flash Experiment on Layered Materials: Parameter Estimation and Uncertainty Evaluation*.
National Physical Laboratory (NPL), Internal Communication.