Using an SGFEM Surrogate to Accelerate Bayesian Inverse Uncertainty Quantification

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

SIAM UQ 2018, Los Angeles

james.rynn@manchester.ac.uk

16/04/18

National Physical Laboratory

Overview

- V. Hoang, C. Schwab, and A. Stuart, *Complexity Analysis of Accelerated MCMC Methods for Bayesian Inversion*, Inverse Problems, 29 (2013), p. 085010.
- Y. Marzouk and H. Najm, *Dimensionality Reduction and Polynomial Chaos Acceleration of Bayesian Inference in Inverse Problems*, Journal of Computational Physics, 228 (2009), pp. 1862–1902.
- Y. Marzouk, H. Najm, and L. Rahn, *Stochastic Spectral Methods for Efficient Bayesian Solution of Inverse Problems*, Journal of Computational Physics, 224 (2007), pp. 560–586.
- F. Nobile and R. Tempone, *Analysis and Implementation Issues for the Numerical Approximation of Parabolic Equations with Random Coefficients*, International Journal for Numerical Methods in Engineering, 80 (2009), pp. 979–1006.

Bayesian Inverse Problems

Find the unknown θ given n_z observations z, satisfying

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

where

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

Goal: Efficiently estimate the posterior density $\pi(\theta|\boldsymbol{z})$ for the unknowns θ given the data $\boldsymbol{z}$.

Bayes' Theorem

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

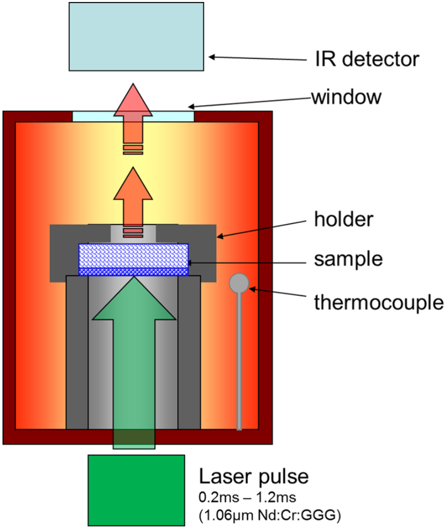
$$\begin{aligned}\pi(\theta|z) &\propto L(z|\theta)\ \pi_0(\theta)\\ &\propto \exp\left(-\frac{1}{2}\|z-\mathcal{G}(\theta)\|_\Sigma^2\right)\ \pi_0(\theta).\end{aligned}$$

Markov Chain Monte Carlo (MCMC) Methods

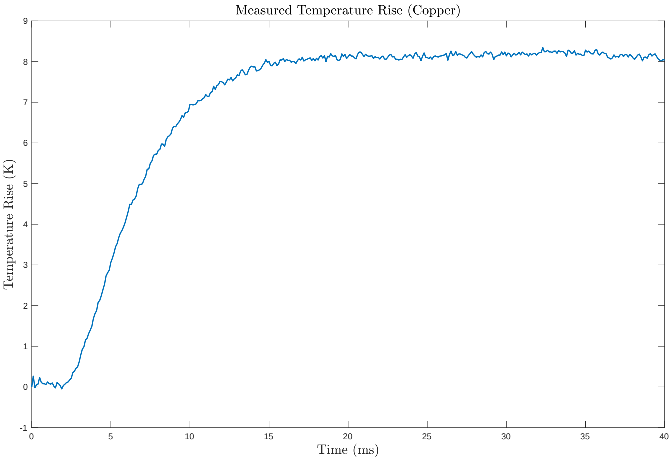
- We know $\pi(\theta|z)$ up to a constant of proportionality.
- Use MCMC algorithm to generates samples $\theta_1, \theta_2, \ldots, \theta_M$ from the posterior distribution.
- Use these samples to construct Monte Carlo estimates of quantities of interest (means, variances and/or probabilities),
-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

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

Industrial Example



Industrial Example



Industrial Example

Possible unknowns:

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

PDE Forward Problem

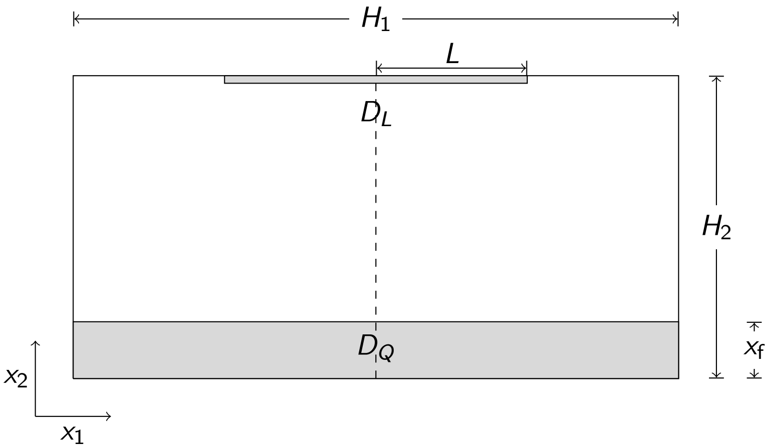
$$\rho c_\mathrm{p} \partial_t u(\boldsymbol{x}, t) - \nabla \cdot (\lambda \nabla u(\boldsymbol{x}, t)) = Q(\boldsymbol{x}, t), \quad (\boldsymbol{x}, t) \in D \times [0, T],$$

where $Q(\boldsymbol{x}, t) = I \cdot 1_{\{D_Q \times [0, t_\mathrm{f}]\}}(\boldsymbol{x}, t)$,

$$\begin{aligned} u(\boldsymbol{x}, 0) &= T_\mathrm{a}, & \boldsymbol{x} &\in \overline{D}, \\ \lambda \frac{\partial u}{\partial n}(\boldsymbol{x}, t) &= 0, & (\boldsymbol{x}, t) &\in \partial D_1 \times [0, T], \\ \lambda \frac{\partial u}{\partial n}(\boldsymbol{x}, t) &= \kappa \left(T_\mathrm{a} - u(\boldsymbol{x}, t) \right), & (\boldsymbol{x}, t) &\in \partial D_2 \times [0, T]. \end{aligned}$$

Goal: Find posterior density $\pi(\theta|\boldsymbol{z})$ for the unknowns $\theta := (\lambda, I)$, given observations $\boldsymbol{z}$ of the average (top) surface temperature $\bar{u}$ at the measurement times $t_1, t_2, \ldots, t_{n_z}$.

Physical Domain, D



Observation Operator

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

$$\mathcal{G}(\theta) = (\bar{u}(t_1; \theta), \bar{u}(t_2; \theta), \ldots, \bar{u}(t_{n_z}; \theta))^\top ,$$

and approximated by $\mathcal{G}_{h\tau}$ given by

$$\mathcal{G}_{h\tau}(\theta) = (\bar{u}_{h\tau}(t_1; \theta), \bar{u}_{h\tau}(t_2; \theta), \ldots, \bar{u}_{h\tau}(t_{n_z}; \theta))^\top ,$$

where

$$\bar{u}(t; \theta) := \frac{1}{\pi L^2} \int_{D_L} u(\cdot, t; \theta) \mathrm{d}A,$$

is the average temperature over surface D_L at time t.

Note: For each value of θ, to evaluate $\mathcal{G}_{h\tau}$ we are required to compute a FEM solve of a time-dependant PDE.

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** $\mathcal{G}_{h\tau}$ **(expensive!)**
  compute acceptance probability α
  accept proposal with probability α
end for
output chain $X = (\theta_0, \theta_1, \ldots, \theta_M)$

Here $M \gg 10^5$.

(Results)

Unfortunately we cannot compute these as producing the samples takes far too long!

85 seconds per (time-dependant) PDE solve
$\implies$ 5m samples takes 4.25×10^8 seconds $=$ 13.5 years (single CPU)

Parametric Forward Problem

Assume now that both λ and l may be expressed in terms of uniform random variables of mean zero and unit variance. That is,

$$\lambda = \mu_\lambda + w_\lambda \xi_1, \qquad l = \mu_l + w_l \xi_2,$$

for some given $\mu_\lambda, \mu_l, w_\lambda, w_l \in \mathbb{R}^+$ with

$$\xi_1, \xi_2 \sim \mathcal{U}(-\sqrt{3}, \sqrt{3}), \qquad p(\xi) = \frac{1}{2\sqrt{3}},$$

$$\boldsymbol{y} := (\xi_1(\omega), \xi_2(\omega))^\top \in \Gamma := (-\sqrt{3}, \sqrt{3})^2.$$

Parametric PDE

$$\rho c_{\mathrm{p}} \partial_t u(\boldsymbol{x}, t, \boldsymbol{y}) - \nabla \cdot (\lambda(\boldsymbol{y}) \nabla u(\boldsymbol{x}, t, \boldsymbol{y})) = Q(\boldsymbol{x}, t, \boldsymbol{y}),$$

for $(\boldsymbol{x}, t, \boldsymbol{y}) \in D \times [0, T] \times \Gamma$, where

$$Q(\boldsymbol{x}, t, \boldsymbol{y}) = I(\boldsymbol{y}) \cdot 1_{\{D_Q \times [0, t_{\mathrm{f}}]\}}(\boldsymbol{x}, t),$$

with IC and BCs given by

$$\begin{aligned} u(\boldsymbol{x}, 0, \boldsymbol{y}) &= T_{\mathrm{a}}, && (\boldsymbol{x}, \boldsymbol{y}) \in \overline{D} \times \Gamma, \\ \lambda(\boldsymbol{y}) \frac{\partial u}{\partial n}(\boldsymbol{x}, t, \boldsymbol{y}) &= 0, && (\boldsymbol{x}, t, \boldsymbol{y}) \in \partial D_1 \times [0, T] \times \Gamma, \\ \lambda(\boldsymbol{y}) \frac{\partial u}{\partial n}(\boldsymbol{x}, t, \boldsymbol{y}) &= \kappa \left(T_{\mathrm{a}} - u(\boldsymbol{x}, t, \boldsymbol{y}) \right), && (\boldsymbol{x}, t, \boldsymbol{y}) \in \partial D_2 \times [0, T] \times \Gamma. \end{aligned}$$

SGFEM Solution

Compute a finite-dimensional approximation at time steps $\tau_1, \tau_2, \ldots, \tau_{n_t}$ such that for each $n = 1, 2, \ldots, n_t$,

$$u_{hk\tau}(\boldsymbol{x}, \tau_n, \boldsymbol{y}) = \sum_{i=1}^{n_h} \sum_{j=1}^{n_k} u_{ij}(\tau_n) \phi_i(\boldsymbol{x}) \Psi_j(\boldsymbol{y}) \in \mathcal{X}^h \otimes S^k,$$

where

$$\mathcal{X}^h := \text{span}\{\phi_1, \phi_2, \ldots, \phi_{n_h}\} \subseteq H^1(D), \qquad |\mathcal{X}^h| = n_h,$$
$$S^k := \text{span}\{\Psi_1, \Psi_2, \ldots, \Psi_{n_k}\} \subseteq L^2_p(\Gamma), \qquad |S^h| = n_k.$$

Approximate observation operator $\mathcal{G}_{hk\tau}$

$$\mathcal{G}_{hk\tau}(\boldsymbol{y}) = (\bar{u}_{hk\tau}(t_1; \boldsymbol{y}), \bar{u}_{hk\tau}(t_2; \boldsymbol{y}), \ldots, \bar{u}_{hk\tau}(t_{n_z}; \boldsymbol{y}))^\top.$$

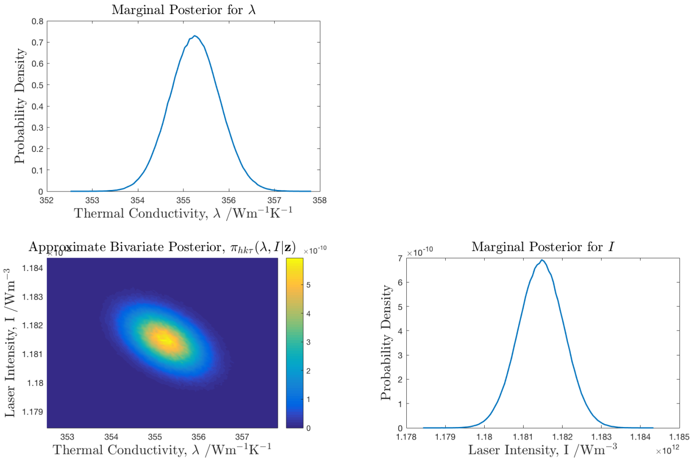
Random Walk Metropolis Hastings Algorithm (SGFEM)

Algorithm 2 RWMH Algorithm with SGFEM Surrogate

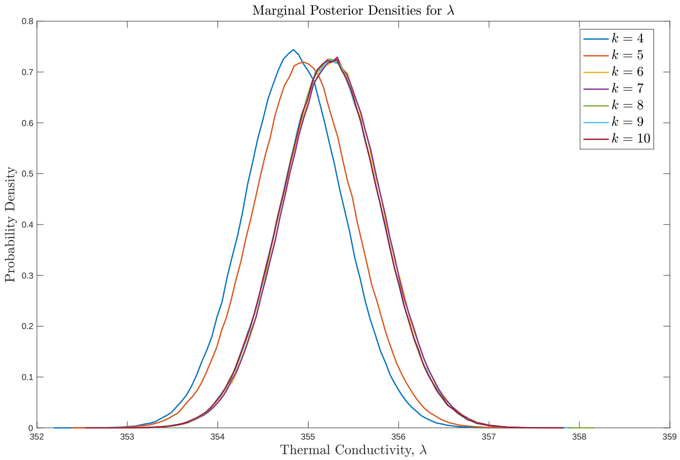
compute SGFEM solution $u_{hk\tau}$
set initial state $X^{(0)} = \theta_0$
for $m = 1, 2, \dots, M$ **do**
  draw proposal
  evaluate likelihood by **evaluating** $\mathcal{G}_{hk\tau}$ (**cheap!**)
  compute acceptance probability α
  accept proposal with probability α
end for
output chain $X = (\theta_0,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

Here $M \gg 10^5$.

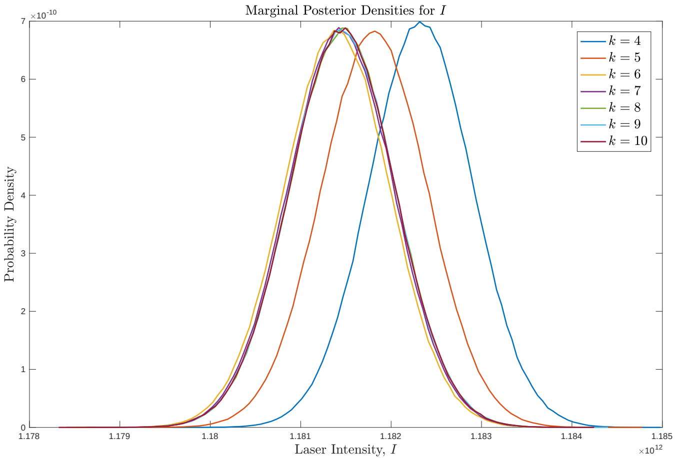
Posterior Density, $\pi(\theta|z)$



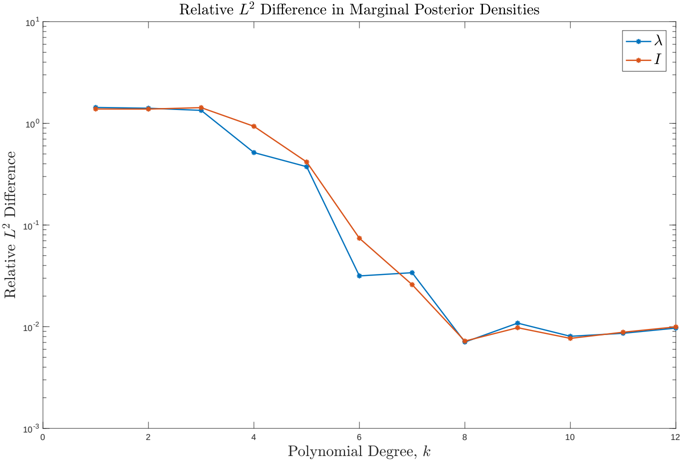
Posterior Convergence in k (Polynomial Degree)



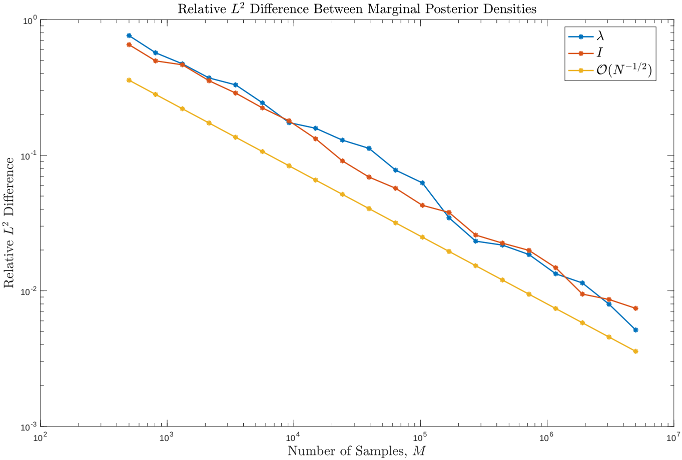
Posterior Convergence in k (Polynomial Degree)



Posterior Convergence in *k* (Polynomial Degree)



Posterior Convergence in M (Number of Samples)



Computational Time

Offline: Compute SGFEM solution with around 1.3 billion DOF ($n_h \times n_t \times n_k = 38397 \times 120 \times 28$): 1289 seconds

Online: Generate 5 million samples using SGFEM-RWMH: 452 seconds

Total: 1741 seconds (3.48×10^{-4} seconds per sample)

Future Work

- More realistic/complex forward problem:
  - spatial varying random variables
  - multi-layered material
  - express boundary heat loss parameter as random variable
- More sophisticated MCMC algorithm
- Error analysis
- Paper (in production): *"Surrogate accelerated Bayesian inversion for the determination of the thermal diffusivity of a material"*

- V. Hoang, C. Schwab, and A. Stuart, *Complexity Analysis of Accelerated MCMC Methods for Bayesian Inversion*, Inverse Problems, 29 (2013), p. 085010.
- Y. Marzouk and H. Najm, *Dimensionality Reduction and Polynomial Chaos Acceleration of Bayesian Inference in Inverse Problems*, Journal of Computational Physics, 228 (2009), pp. 1862–1902.
- Y. Marzouk, H. Najm, and L. Rahn, *Stochastic Spectral Methods for Efficient Bayesian Solution of Inverse Problems*, Journal of Computational Physics, 224 (2007), pp. 560–586.
- F. Nobile and R. Tempone, *Analysis and Implementation Issues for the Numerical Approximation of Parabolic Equations with Random Coefficients*, International Journal for Numerical Methods in Engineering, 80 (2009), pp. 979–1006.

Choice of Prior