Bayesian Inversion of PDE-Governed Inverse Problems Using Integrated Nested Laplace Approximations

Sonia Reilly

Georg Stadler

Courant Institute of Mathematical Sciences, New York University

PDE-Governed Bayesian Inverse Problems

- Find parameter $m(x)$ from noisy measurements y of solution u of a PDE:

$$\underbrace{y_i}_{\text{data}} = \underbrace{u[m,\boldsymbol{\theta}](x_i)}_{\text{solution}} + \underbrace{\varepsilon_i}_{\text{noise}}, \quad \text{where e.g.,} \qquad -\nabla\cdot(m(x)\nabla u(x)) = f$$

- To infer discretization $\boldsymbol{m}$ of $m(x)$ with its uncertainty, compute posterior distribution using Bayes' rule

$$\underbrace{\pi(\boldsymbol{m},\boldsymbol{\theta}|\boldsymbol{y})}_{\text{posterior}} \propto \underbrace{\pi(\boldsymbol{\theta})\pi(\boldsymbol{m}|\boldsymbol{\theta})}_{\text{priors}}\underbrace{\pi(\boldsymbol{y}|\boldsymbol{m},\boldsymbol{\theta})}_{\text{likelihood}}$$

- Then marginalize over hyperparameters $\boldsymbol{\theta}$

$$\pi(\boldsymbol{m}|\boldsymbol{y}) = \int \pi(\boldsymbol{m},\boldsymbol{\theta}|\boldsymbol{y})\, d\boldsymbol{\theta}$$

High-Dimensional Integrals

Working with the posterior requires integrating over high-dimensional $\boldsymbol{m}$:

- normalizing: $\int \pi(\boldsymbol{m}|\boldsymbol{y})\, d\boldsymbol{m}$
- finding mean: $\int \boldsymbol{m}\, \pi(\boldsymbol{m}|\boldsymbol{y})\, d\boldsymbol{m}$
- finding covariance: $\int \boldsymbol{m}\boldsymbol{m}^T\, \pi(\boldsymbol{m}|\boldsymbol{y})\, d\boldsymbol{m}$

Quadrature suffers from the curse of dimensionality

Common approximation methods avoid quadrature, but incur tradeoffs:

- MCMC: converges prohibitively slowly in high dimensions
- Variational inference: limited to particular families of distributions

A Tractable Case

Assumptions

- Hyperparameter $\boldsymbol{\theta}$ is low-dimensional ($\lesssim$ 10)
- Prior for $\boldsymbol{m}$ is Gaussian (“latent Gaussian field")

Latent Gaussian Model

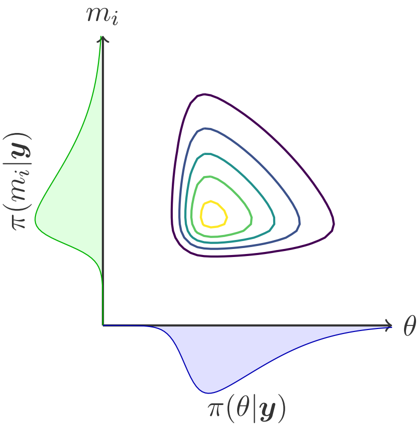
$$\begin{aligned} \boldsymbol{\theta} &\sim \pi(\boldsymbol{\theta}) && \text{hyperparameters} \\ \boldsymbol{m} \,|\, \boldsymbol{\theta} &\sim \mathcal{N}(\boldsymbol{\mu}(\boldsymbol{\theta}), Q^{-1}(\boldsymbol{\theta})) && \text{latent Gaussian field} \\ \boldsymbol{y} \,|\, \boldsymbol{m}, \boldsymbol{\theta} &\sim \prod_i \pi(y_i \,|\, u[\boldsymbol{m}, \boldsymbol{\theta}]_i) && \text{observations} \end{aligned}$$

for some forward operator $\boldsymbol{m}, \boldsymbol{\theta} \to \boldsymbol{u}[\boldsymbol{m}, \boldsymbol{\theta}]$

INLA: Fast Approximate Marginal Posteriors for LGMs

Integrated Nested Laplace Approximations [a]

- Approximates $\pi(m_i|\boldsymbol{y})$ for LGMs



Pros

- Deterministic
- Not restricted to a family of distributions
- Fast and typically very accurate

Cons

- Does not give full joint $\pi(\boldsymbol{m}, \boldsymbol{\theta}|\boldsymbol{y})$
- **Forward model usually simple**

[a] Rue et al., *Approx. Bayesian Inference*, JRSS, 2009

INLA: A Bird's Eye View

Step 1: approximate hyperparameter distribution $\pi(\boldsymbol{\theta}|\boldsymbol{y})$

Step 2: approximate $\pi(m_i|\boldsymbol{\theta}, \boldsymbol{y})$ for all i at quadrature points $\boldsymbol{\theta}_j$

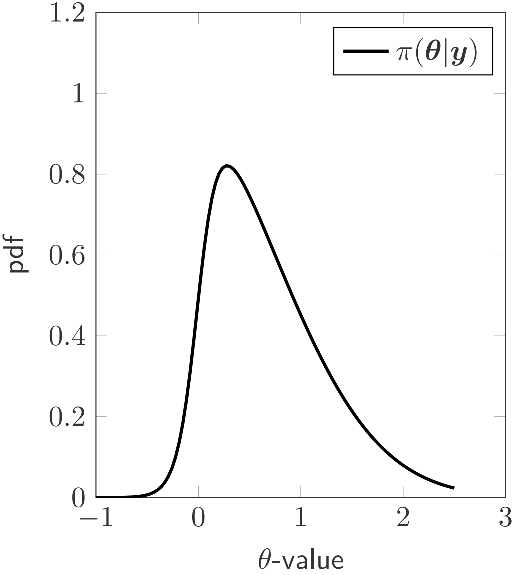
Step 3: compute $\pi(m_i|\boldsymbol{y})$ for all i using quadrature over $\boldsymbol{\theta}$:

$$\begin{aligned}\pi(m_i|\boldsymbol{y}) &= \int \pi(m_i, \boldsymbol{\theta}|\boldsymbol{y})\, d\boldsymbol{\theta} \\ &= \int \underbrace{\pi(m_i|\boldsymbol{\theta}, \boldsymbol{y})}_{\text{Step 2}}\, \underbrace{\pi(\boldsymbol{\theta}|\boldsymbol{y})}_{\text{Step 1}}\, d\boldsymbol{\theta}\end{aligned}$$

Step 1: Finding the marginal of θ

High-dimensional integral:

$$\pi(\boldsymbol{\theta}|\boldsymbol{y}) = \int \pi(\boldsymbol{m}, \boldsymbol{\theta}|\boldsymbol{y})\, d\boldsymbol{m}$$



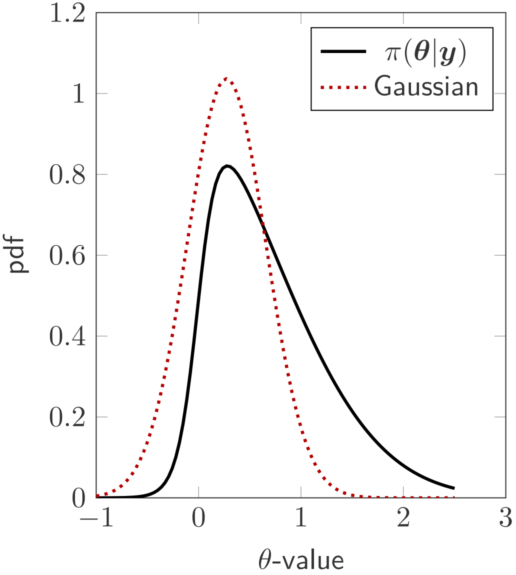
Step 1: Finding the marginal of θ

High-dimensional integral:

$$\pi(\boldsymbol{\theta}|\boldsymbol{y}) = \int \pi(\boldsymbol{m}, \boldsymbol{\theta}|\boldsymbol{y})\, d\boldsymbol{m}$$

- Could use Gaussian approximation:

$$\pi(\boldsymbol{\theta}|\boldsymbol{y}) \approx \int \pi_G(\boldsymbol{m}, \boldsymbol{\theta}|\boldsymbol{y})\, d\boldsymbol{m}$$



Step 1: Finding the marginal of θ

High-dimensional integral:

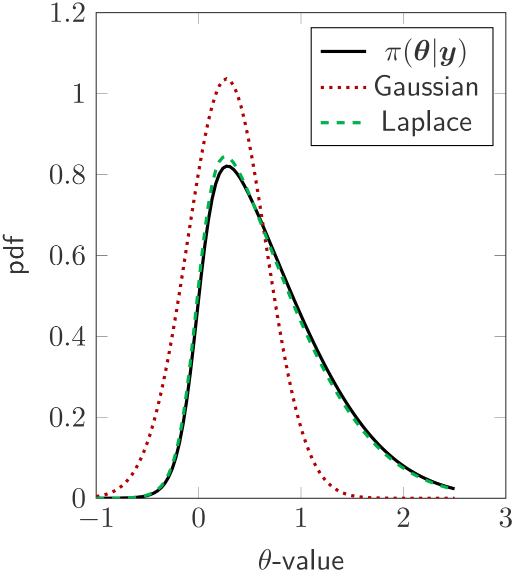
$$\pi(\boldsymbol{\theta}|\boldsymbol{y}) = \int \pi(\boldsymbol{m}, \boldsymbol{\theta}|\boldsymbol{y})\, d\boldsymbol{m}$$

- Could use Gaussian approximation:

$$\pi(\boldsymbol{\theta}|\boldsymbol{y}) \approx \int \pi_G(\boldsymbol{m}, \boldsymbol{\theta}|\boldsymbol{y})\, d\boldsymbol{m}$$

- Laplace approximation uses Gaussian approximations indirectly for much higher accuracy:

$$\pi(\boldsymbol{\theta}|\boldsymbol{y}) \propto \frac{\pi(\boldsymbol{m}, \boldsymbol{\theta}, \boldsymbol{y})}{\pi(\boldsymbol{m}|\boldsymbol{\theta}, \boldsymbol{y})} \approx \frac{\pi(\boldsymbol{m}, \boldsymbol{\theta}, \boldsymbol{y})}{\pi_G(\boldsymbol{m}|\boldsymbol{\theta}, \boldsymbol{y})}$$

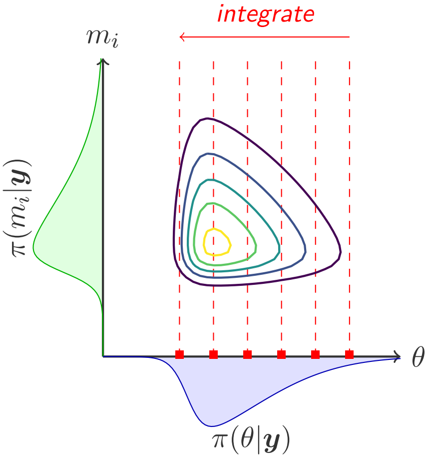


Steps 2 & 3: Finding the marginals of m_i

Low-dimensional integral:

$$\begin{aligned}\pi(m_i|\boldsymbol{y}) &= \int \pi(m_i, \boldsymbol{\theta}|\boldsymbol{y})\, d\boldsymbol{\theta} \\ &= \int \pi(m_i|\boldsymbol{\theta}, \boldsymbol{y})\, \underbrace{\pi(\boldsymbol{\theta}|\boldsymbol{y})}_{\text{Step 1}}\, d\boldsymbol{\theta}\end{aligned}$$

- Pick quadrature points in $\boldsymbol{\theta}$
- $\pi(m_i|\boldsymbol{\theta}, \boldsymbol{y}) \approx \pi_G(m_i|\boldsymbol{\theta}, \boldsymbol{y})$ or higher-order (2nd "nested" Laplace approx)
- Integrate with quadrature



Making INLA fast for PDE-governed problems

Classical INLA

- simple forward operator
- $\boldsymbol{\theta}$ usually prior hyperparameters
- computational cost dominated by inverting large matrices
- speed achieved via sparse linear algebra

PDE-Governed Problems

- PDE operator
- $\boldsymbol{\theta}$ usually also parameters of PDE
- computational cost dominated by solving forward and adjoint PDE
- can't build matrices explicitly

Key insight: For PDE operators, precision update $\Gamma_{\text{post}}^{-1} - \Gamma_{\text{prior}}^{-1}$ is low-rank

- Precompute low rank approx to $\Gamma_{\text{post}}^{-1} - \Gamma_{\text{prior}}^{-1}$ using randomized NLA
- Future work: speed up recomputing the low-rank approx for each $\boldsymbol{\theta}$

An example: initial condition inference in advection-diffusion

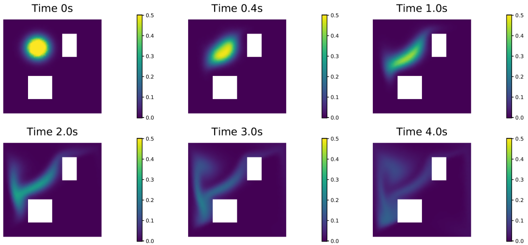
Forward problem:

- Given: diffusivity $\kappa > 0$ and IC $u_0(x)$
- Solve for $u(x,t)$:

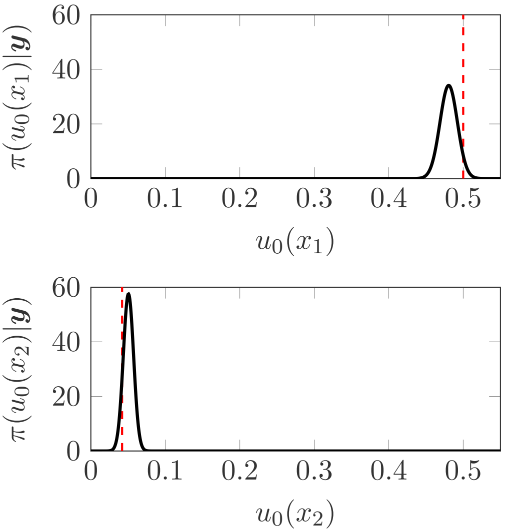
$$\begin{cases} u_t - \kappa \Delta u + \boldsymbol{v} \cdot \nabla u = 0 & \text{in} \quad \Omega \times [0,T] \\ u(0,x) = u_0(x) & \text{in} \quad \Omega \\ \kappa \nabla u \cdot \boldsymbol{n} = 0 & \text{on} \quad \partial\Omega. \end{cases}$$

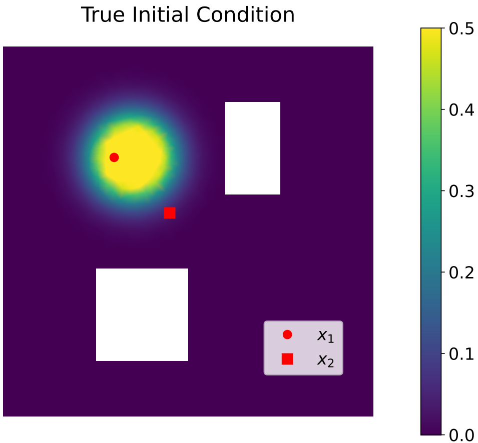
Inverse problem:

- Given: noisy data $u_d^i = u(x_i, t_i) + \varepsilon_i$, iid normally distributed noise $\boldsymbol{\varepsilon}$, normal prior on $\boldsymbol{u_0}$ and prior on κ
- Want to find: marginal posterior distribution $\pi(\boldsymbol{u_0} | \boldsymbol{u_d})$



Preliminary Results: hIPPYlib + INLA





Conclusion

- INLA is fast and accurate for computing pointwise uncertainty when hyperparameters are low-dimensional
- Uses Laplace approximations for high-dimensional integrals, quadrature for low-dimensional integrals
- Can be adapted to be fast for PDE-governed problems

NYU COURANT

This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research, Department of Energy, Computational Science Graduate Fellowship under Award Number DE-SC0022158.

DOE CSGF