Bayesian Inversion of Hierarchical PDE-Governed Inverse Problems Using INLA

Sonia Reilly

Georg Stadler

Courant Institute of Mathematical Sciences, New York University

Hierarchical PDE-Governed Bayesian Inverse Problems

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

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

- 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]

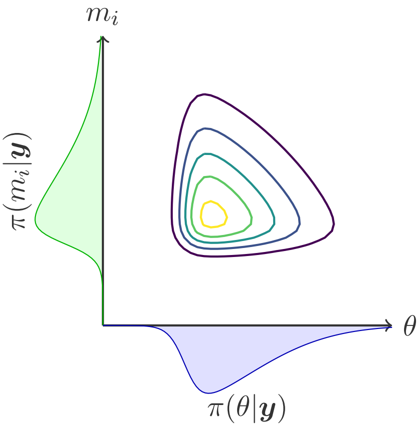
- Approximates $\pi(m_i|\boldsymbol{y})$ for LGMs

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 $\boldsymbol{\theta}_j$

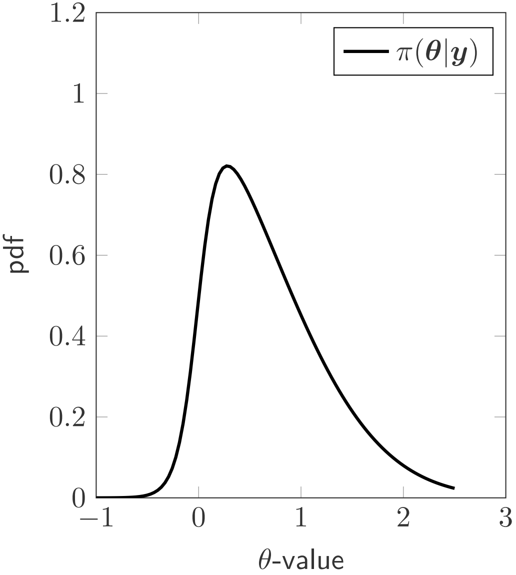
Step 3: compute $\pi(m_i|\boldsymbol{y})$ for all i using quadrature over $\boldsymbol{\theta}$:

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

Step 1: Finding the marginal of θ

High-dimensional integral:

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



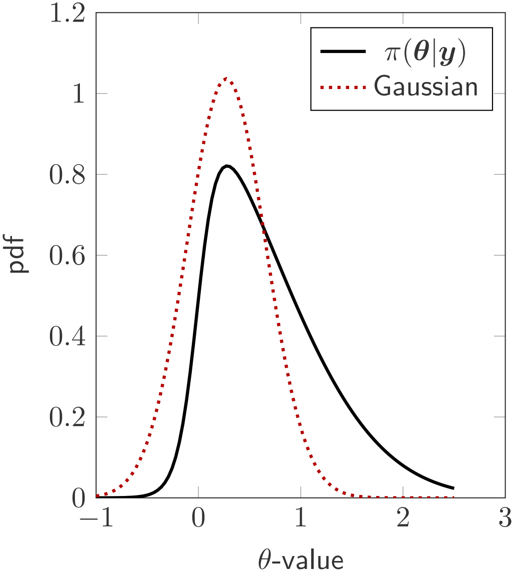
Step 1: Finding the marginal of θ

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:

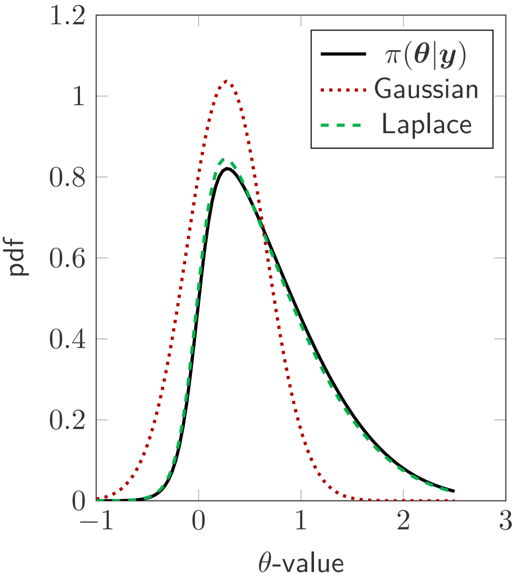
$$\pi(\boldsymbol{\theta}|\boldsymbol{y}) = \int \pi(\boldsymbol{m}, \boldsymbol{\theta}|\boldsymbol{y})\, d\boldsymbol{m}$$

- 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:

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

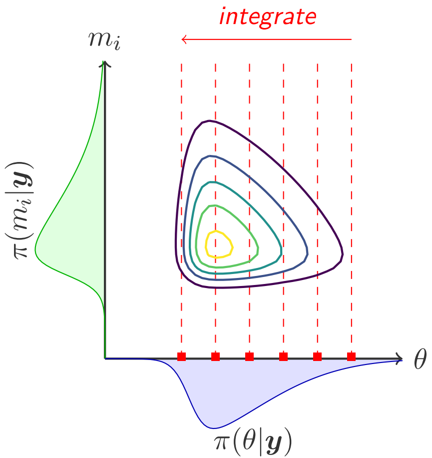


Steps 2 & 3: Finding the marginals of m_i

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
- computational cost dominated by inverting large matrices
- speed achieved via sparse linear algebra

PDE-Governed Problems

- PDE operator
- computational cost dominated by solving forward and adjoint PDE
- can't build matrices explicitly

Key insights:

- For PDE operators, precision update $\Gamma_{\text{post}}^{-1} - \Gamma_{\text{prior}}^{-1}$ is low-rank, so can precompute using randomized numerical linear algebra
- PDE problem allows a multilevel approach: use coarse mesh to quickly find quadrature points, then fine mesh for accuracy in $\pi(m_i|\boldsymbol{y})$.

An example: initial condition inference in advection-diffusion

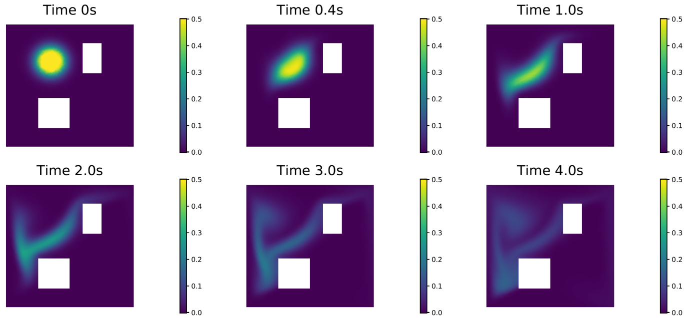
Forward problem:

- Given initial condition $u_0(x)$
- Find $u(x,t)$

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

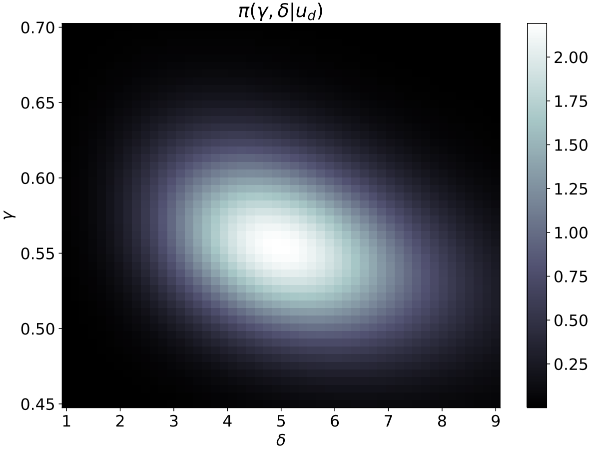
Inverse problem:

- Given data $u_d^i = u(x_i,t_i) + \varepsilon_i$ with Gaussian noise $\boldsymbol{\varepsilon}$, Gaussian prior on $\boldsymbol{u_0}$
- Find $\pi(\boldsymbol{u_0}|\boldsymbol{u_d})$



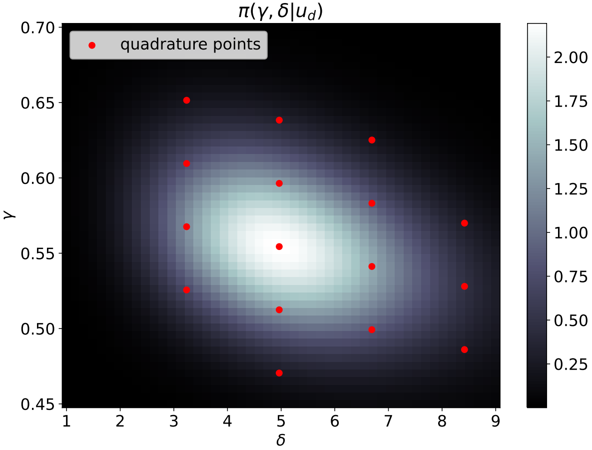
Hyperparameter Distribution

- Hyperparameters γ, δ in prior covariance $(\delta I - \gamma\Delta)^{-2}$
- γ/δ determines the length scale of the prior
- Other possible hyperparameters are magnitude of observation noise, or model parameters like diffusivity κ

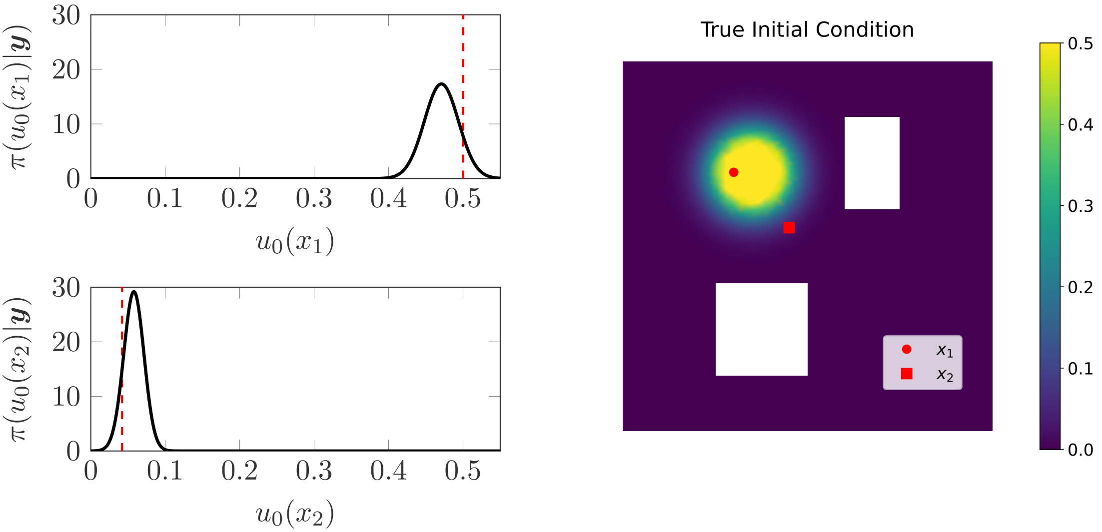


Hyperparameter Distribution

- Hyperparameters γ, δ in prior covariance $(\delta I - \gamma\Delta)^{-2}$
- γ/δ determines the length scale of the prior
- Other possible hyperparameters are magnitude of observation noise, or model parameters like diffusivity κ



Initial Condition Distributions



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

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.