Probabilistic solvers for ODE's and Bayesian inference of parametrized models

Giacomo Garegnani, Assyr Abdulle

giacomo.garegnani@epfl.ch, assyr.abdulle@epfl.ch

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Introduction

In this project we focus on the probabilistic interpretation of numerical solutions of ODE's and Bayesian inference inverse problems involving differential equations. In particular, we consider three main research topics.

Bayesian inference. Which are the most efficient inferential techniques?

Probabilistic integrators of ODE's [1]. What are their properties?

Bayesian inverse problems. Can we apply the probabilistic integrator in the context of Bayesian inference inverse problems?

Main contribution. We derive bounds for Monte Carlo estimators in a probabilistic extension of Runge-Kutta methods.

Bayesian inference and MCMC

Let $\theta \in \mathbb{R}^{N_p}$ be a parameter and $\mathcal{Y}$ a set of observations. Then Bayes' rule reads

$$\pi(\theta \mid \mathcal{Y}) \propto \mathcal{Q}(\theta)\mathcal{L}(\mathcal{Y} \mid \theta).$$

Goal. Generate samples $\{\theta^{(i)}\}_{i=1}^N$ such that for $g\colon \mathbb{R}^{N_p} \to \mathbb{R}$

$$\mathbb{E}^{\pi}[g(\theta)] \approx \frac{1}{N}\sum_{i=1}^{N} g(\theta^{(i)}).$$

Idea. Use Metropolis-Hastings (MH) Markov chain Monte Carlo method (MCMC) [2] to generate samples from π. Given $\theta^{(i)}$, the next element ϑ is chosen from a proposal distribution $q(x,y)$ and accepted to be $\theta^{(i+1)} = \vartheta$ with probability

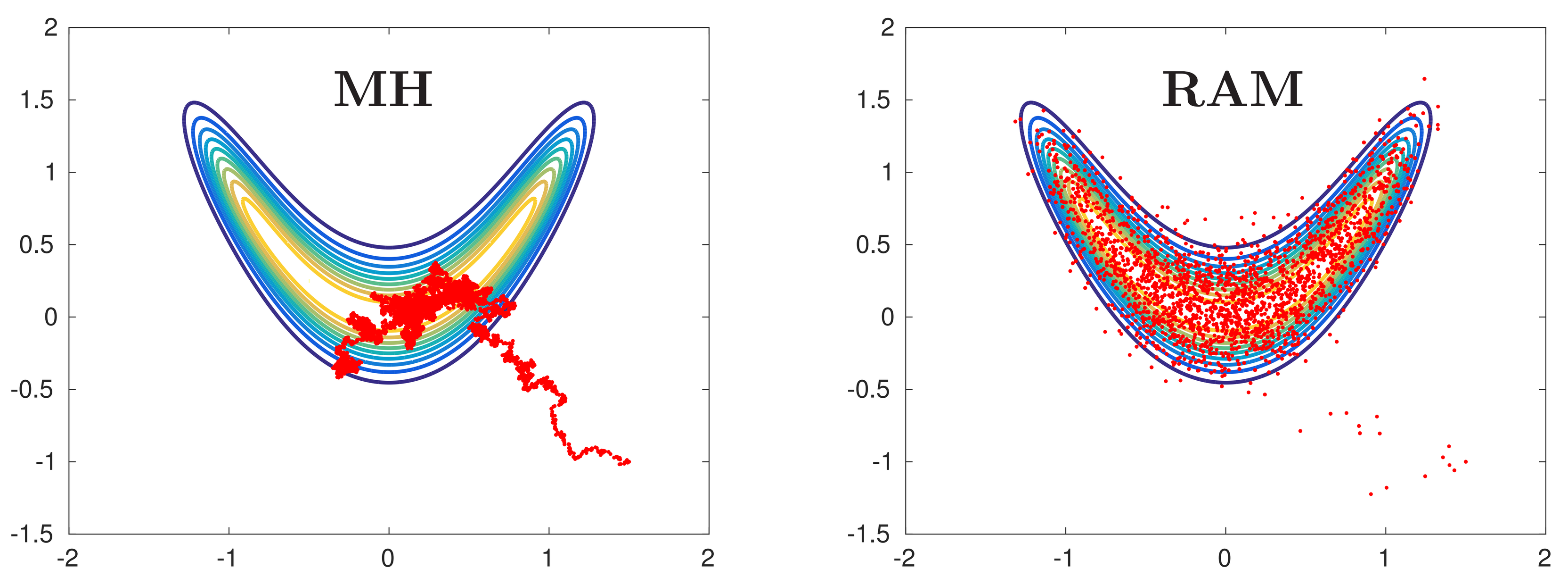
$$\alpha(\theta^{(i)}, \vartheta) = \min\left\{\frac{\pi(\vartheta)q(\vartheta,\theta^{(i)})}{\pi(\theta^{(i)})q(\theta^{(i)},\vartheta)}, 1\right\}.$$

Problem. Bad choices of the proposal distribution $q(x,y)$ lead to inefficient algorithms $\rightsquigarrow$ apply adaptive techniques as the robust adaptive Metropolis (RAM) [3]. Adapt a Gaussian proposal distribution $q(x,y)$ to obtain a desired acceptance rate α^*, defined as the ratio of accepted new parameter guesses ϑ.

Experiment. Consider the two-dimensional distribution π with density

$$\pi(X) \propto \exp(-10(X_1^2 - X_2)^2 - (X_1 - 0.25)^4),$$

and generate 5000 samples with standard MH or RAM. Choose Gaussian $q(x,y)$ with covariance structure $\Sigma = 0.01I$ for MH and for starting proposal for RAM ($\alpha^* = 0.4$). The posterior is not well described by MH, while for RAM we obtain a good approximation.



Probabilistic solvers for ODE's

Optimization-based method. Find $(u_1, u_2) \in H^1(\omega_1) \times H^1_\Gamma(\omega_2)$:

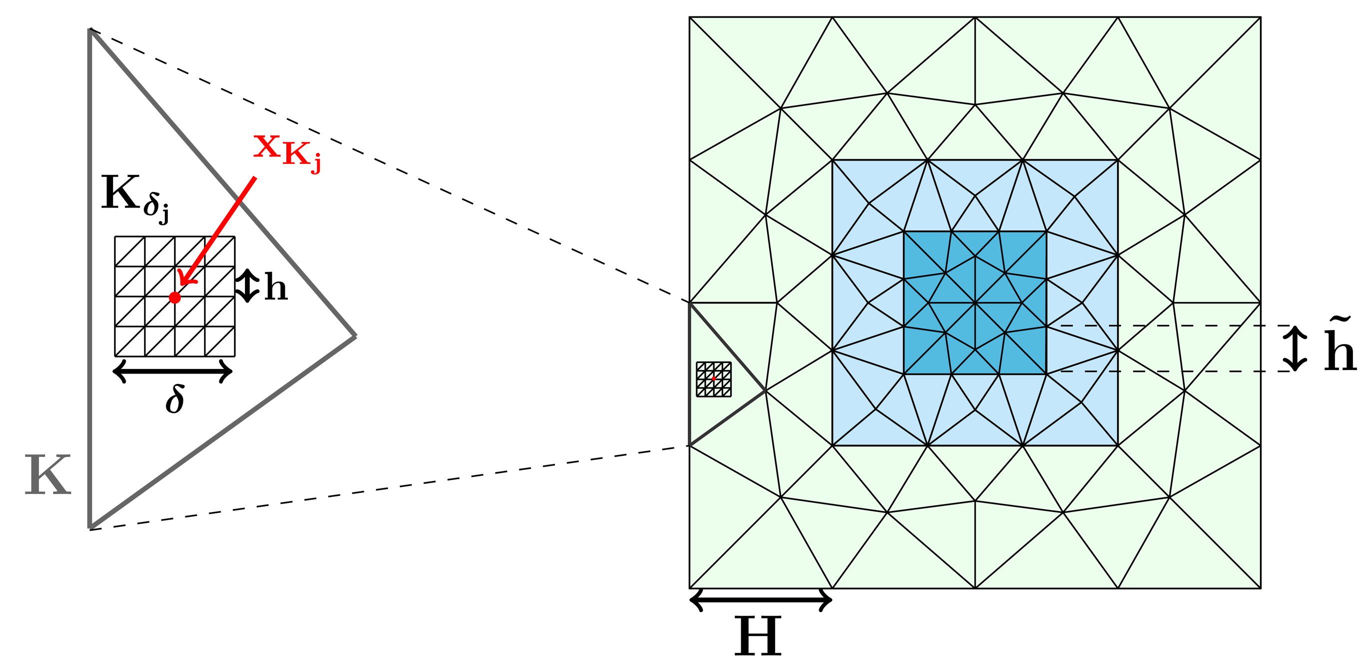
$$\min \frac{1}{2}\|u_1 - u_2\|^2_{L^2(\omega_0)} \text{ such that } \begin{cases} \boxed{-\mathrm{div}\,(a\nabla u_1) = f \quad \text{in } \omega_1,\ u_1 = \theta_1 \text{ on } \Gamma_1}, \\ \boxed{-\mathrm{div}\,(a^0\nabla u_2) = f^0 \text{ in } \omega_2,\ u_2 = \theta_2 \text{ on } \Gamma_2}. \end{cases}$$

Introducing Lagrange multipliers for each constraint, (u_1, u_2) is given as a critical point of a saddle point problem.

FE Optimization Based Coupling

Let $V^p(\omega_1, \mathcal{T}_{\tilde h})$, $V^p_\Gamma(\omega_2, \mathcal{T}_H)$ and $S^q(K_{\delta_j}, \mathcal{T}_h)$ be FE spaces with

$$h, \tilde h < \varepsilon, \text{ and } H \gg h.$$



Fine scale solver Find $u_{1,\tilde h} = u_{1,0,\tilde h} + v_{1,\tilde h}$, with $u_{1,0,\tilde h} \in V^p_0(\omega_1, \mathcal{T}_{\tilde h})$ such that

$$B_1(u_{1,0,\tilde h}, w_{\tilde h}) := \int_{\omega_1} a\nabla u_{1,0,\tilde h}\cdot\nabla w_{\tilde h}\,\mathrm{d}x = \int_{\omega_1} f w_{\tilde h}\,\mathrm{d}x, \qquad \forall w_{\tilde h} \in V^p_0(\omega_1, \mathcal{T}_{\tilde h}).$$

FE-HMM solver Find $u_{2,H} = u_{2,0,H} + v_{2,H}$, with $u_{2,0,H} \in V^p_0(\omega_2, \mathcal{T}_H)$ such that, $\forall w_H \in V^p_0(\omega_2, \mathcal{T}_H)$

$$B_2(u_{2,0,H}, w_H) := \sum_{K\in\mathcal{T}_H}\sum_{j=1}^{J} \frac{w_{j,K}}{|K_{\delta_j}|}\int_{K_{\delta_j}} a\nabla u_j^h\cdot\nabla w_j^h\,\mathrm{d}x = \int_{\omega_2} f^0 w_H\,\mathrm{d}x,$$

where u_j^h, w_j^h are solutions of a micro problem involving the tensor a on K_{δ_j}. The solution $(v_{1,\tilde h}, v_{2,H})$ is the critical point of

$$\mathcal{L}(v_{1,\tilde h}, \lambda_{1,\tilde h}, v_{2,H}, \lambda_{2,H}) = \frac{1}{2}\|(u_{1,0,\tilde h} + v_{1,\tilde h}) - (u_{2,0,H} + v_{2,H})\|^2_{L^2(\omega_0)} - B_1(v_{1,\tilde h}, \lambda_{1,\tilde h}) - B_2(v_{2,H}, \lambda_{2,H}).$$

Algorithm

- Find $u^{\tilde h}_{1,0} \in V^p_0(\omega_1, \mathcal{T}_{\tilde h})$: $B^{\tilde h}_1(u^{\tilde h}_{1,0}, w^{\tilde h}) = \int_{\omega_1} f w^{\tilde h}\mathrm{d}x$, $\forall w^{\tilde h} \in V^p_0(\omega_1, \mathcal{T}_{\tilde h})$.
- Find $u^{H}_{2,0} \in V^p_0(\omega_2, \mathcal{T}_H)$: $B^{\tilde h}_2(u^{H}_{2,0}, w^{\tilde h}) = \int_{\omega_1} f w^{\tilde h}\mathrm{d}x$, $\forall w^{\tilde h} \in V^p_0(\omega_1, \mathcal{T}_{\tilde h})$.
- Find $u^{\tilde h}_{1,0} \in V^p_0(\omega_1, \mathcal{T}_{\tilde h})$: $B^{\tilde h}_1(u^{\tilde h}_{1,0}, w^{\tilde h}) = \int_{\omega_1} f w^{\tilde h}\mathrm{d}x$, $\forall w^{\tilde h} \in V^p_0(\omega_1, \mathcal{T}_{\tilde h})$.

A Priori Error Estimates

Fill here

References

[1] Conrad, P. R. et al. Statistical analysis of differential equations: introducing probability measures on numerical solutions. *Stat. Comput.*, 2016.

[2] J. Kaipio and E. Somersalo. Statistical and Computational Inverse Problems. *Applied Mathematical Sciences*, 2005.

[3] M. Vihola. Robust adaptive Metropolis algorithm with coerced acceptance rate. *Stat. Comput.*, 2012.