Bayesian inference of multiscale diffusion processes

Assyr Abdulle [1] Giacomo Garegnani [1] Grigoris Pavliotis [2]

MATHICSE Retreat

Champéry – 11/13 June 2019

[1]Institute of Mathematics, EPFL

[2]Department of Mathematics, Imperial College London

Motivation

Bayesian inference

- Fit model to data
- Full UQ approach

Diffusion processes

- Intrinsically stochastic phenomena
- Analysis for BM noise

Multiscale

- Numerous real-world applications
- Theory of homogenization applies

Problem statement

Multiscale SDE – first order Langevin

$$\mathrm{d}x^{\varepsilon}(t) = -\underbrace{\alpha\nabla V_0\left(x^{\varepsilon}(t)\right)\mathrm{d}t}_{\text{large-scale potential}} - \underbrace{\frac{1}{\varepsilon}\nabla V_1\left(\frac{x^{\varepsilon}(t)}{\varepsilon}\right)\mathrm{d}t}_{\text{fluctuating potential}} + \underbrace{\sqrt{2\sigma}\,\mathrm{d}W(t)}_{\text{diffusion}}.$$

Problem statement

Multiscale SDE – first order Langevin

$$\mathrm{d}x^\varepsilon(t) = -\underbrace{\alpha \nabla V_0\big(x^\varepsilon(t)\big)\,\mathrm{d}t}_{\text{large-scale potential}} - \underbrace{\frac{1}{\varepsilon}\nabla V_1\Big(\frac{x^\varepsilon(t)}{\varepsilon}\Big)}_{\text{fluctuating potential}}\,\mathrm{d}t + \underbrace{\sqrt{2\sigma}\,\mathrm{d}W(t)}_{\text{diffusion}}.$$

Homogenized SDE

$$\mathrm{d}x^0(t) = -A\,\nabla V_0(x^0(t))\,\mathrm{d}t + \sqrt{2\Sigma}\,\mathrm{d}W(t), \quad A = K\alpha,\ \Sigma = K\sigma.$$

Homogenization result: $x^\varepsilon \Rightarrow x^0$ in $\mathcal{C}^0\big((0,T),\mathbb{R}^d\big)$ for $\varepsilon \to 0$.

Problem statement

Multiscale SDE – first order Langevin

$$\mathrm{d}x^\varepsilon(t) = -\underbrace{\alpha \nabla V_0\big(x^\varepsilon(t)\big)\,\mathrm{d}t}_{\text{large-scale potential}} - \underbrace{\frac{1}{\varepsilon}\nabla V_1\Big(\frac{x^\varepsilon(t)}{\varepsilon}\Big)}_{\text{fluctuating potential}}\,\mathrm{d}t + \underbrace{\sqrt{2\sigma}\,\mathrm{d}W(t)}_{\text{diffusion}}.$$

Homogenized SDE

$$\mathrm{d}x^0(t) = -A\,\nabla V_0(x^0(t))\,\mathrm{d}t + \sqrt{2\Sigma}\,\mathrm{d}W(t), \quad A = K\alpha,\ \Sigma = K\sigma.$$

Homogenization result: $x^\varepsilon \Rightarrow x^0$ in $\mathcal{C}^0\big((0,T),\mathbb{R}^d\big)$ for $\varepsilon \to 0$.

Inverse problem 1 – MS / MS

Find $\theta^\varepsilon = (\alpha, \sigma)$ given $\mathbf{y} = \mathbf{x}^\varepsilon(\theta^\varepsilon) + \boldsymbol{\eta}$, $\boldsymbol{\eta} \sim \rho_\eta$.

Notation: $\mathbb{R}^{Nd} \ni \mathbf{x}^\varepsilon = \big(x_1^\varepsilon, x_2^\varepsilon, \ldots, x_N^\varepsilon\big)$, $x_k^\varepsilon = x^\varepsilon(t_k)$.

Problem statement

Multiscale SDE – first order Langevin

$$\mathrm{d}x^\varepsilon(t) = -\underbrace{\alpha \nabla V_0\big(x^\varepsilon(t)\big)\,\mathrm{d}t}_{\text{large-scale potential}} - \underbrace{\frac{1}{\varepsilon}\nabla V_1\Big(\frac{x^\varepsilon(t)}{\varepsilon}\Big)}_{\text{fluctuating potential}}\,\mathrm{d}t + \underbrace{\sqrt{2\sigma}\,\mathrm{d}W(t)}_{\text{diffusion}}.$$

Homogenized SDE

$$\mathrm{d}x^0(t) = -A\,\nabla V_0(x^0(t))\,\mathrm{d}t + \sqrt{2\Sigma}\,\mathrm{d}W(t), \quad A = K\alpha,\ \Sigma = K\sigma.$$

Homogenization result: $x^\varepsilon \Rightarrow x^0$ in $\mathcal{C}^0\big((0,T),\mathbb{R}^d\big)$ for $\varepsilon \to 0$.

Inverse problem 2 – MS / HOM

Find $\theta^0 = (A, \Sigma)$ given $\mathbf{y} = \mathbf{x}^\varepsilon(\theta^\varepsilon) + \boldsymbol{\eta}$, $\boldsymbol{\eta} \sim \rho_\eta$.

Notation: $\mathbb{R}^{Nd} \ni \mathbf{x}^\varepsilon = \big(x_1^\varepsilon, x_2^\varepsilon, \dots, x_N^\varepsilon\big)$, $x_k^\varepsilon = x^\varepsilon(t_k)$.

Bayesian framework

Inverse problem 1 – MS / MS

$$\text{Find } \theta^\varepsilon = (\alpha, \sigma)^\top \text{ given } \mathbf{y} = \mathbf{x}^\varepsilon(\theta^\varepsilon) + \boldsymbol{\eta},\ \boldsymbol{\eta} \sim \rho_\eta.$$

Posterior distribution $\mu^\varepsilon(\theta^\varepsilon \mid \mathbf{y})$ with density

$$p^\varepsilon(\theta^\varepsilon \mid \mathbf{y}) = \frac{1}{Z^\varepsilon} \underbrace{p(\theta^\varepsilon)}_{\text{prior}} \underbrace{p^\varepsilon(\mathbf{y} \mid \theta^\varepsilon)}_{\text{likelihood}}, \quad Z^\varepsilon \text{ s.t. } \int p^\varepsilon(\theta \mid \mathbf{y}) \,\mathrm{d}\theta = 1.$$

Prior: Easy to evaluate (e.g. Gaussian), independent of ε

Likelihood: Needs more work

Bayesian framework

Inverse problem 1 – MS / MS

$$\text{Find } \theta^\varepsilon = (\alpha, \sigma)^\top \text{ given } \mathbf{y} = \mathbf{x}^\varepsilon(\theta^\varepsilon) + \boldsymbol{\eta},\ \boldsymbol{\eta} \sim \rho_\eta.$$

Likelihood: Needs more work $\Rightarrow$ marginalization

$$p^\varepsilon(\mathbf{y} \mid \theta^\varepsilon) = \int_{\mathbb{R}^{Nd}} p^\varepsilon(\mathbf{y} \mid \mathbf{x}, \theta^\varepsilon)\, p^\varepsilon(\mathbf{x} \mid \theta^\varepsilon)\, \mathrm{d}\mathbf{x}.$$

where (observation independence)

$$p^\varepsilon(\mathbf{y} \mid \mathbf{x}, \theta^\varepsilon) = \prod_{k=1}^{N} p^\varepsilon(y_k \mid x_k, \theta^\varepsilon).$$

Observation density: $p(y_k \mid x_k, \theta^\varepsilon) = \rho_\eta^{(k)}(y_k - x_k)$

Bayesian framework

Inverse problem 1 – MS / MS

Find $\theta^\varepsilon = (\alpha, \sigma)^\top$ given $\mathbf{y} = \mathbf{x}^\varepsilon(\theta^\varepsilon) + \boldsymbol{\eta}$, $\boldsymbol{\eta} \sim \rho_\eta$.

Likelihood: Needs more work $\Rightarrow$ marginalization

$$p^\varepsilon(\mathbf{y} \mid \theta^\varepsilon) = \int_{\mathbb{R}^{Nd}} p(\mathbf{y} \mid \mathbf{x}, \theta^\varepsilon)\, p^\varepsilon(\mathbf{x} \mid \theta^\varepsilon)\, \mathrm{d}\mathbf{x}.$$

where (observation independence)

$$p(\mathbf{y} \mid \mathbf{x}, \theta^\varepsilon) = \prod_{k=1}^{N} p(y_k \mid x_k, \theta^\varepsilon).$$

Observation density: $p(y_k \mid x_k, \theta^\varepsilon) = \rho_\eta^{(k)}(y_k - x_k) \Rightarrow$ independent of ε.

Bayesian framework

Inverse problem 1 – MS / MS

Find $\theta^\varepsilon = (\alpha, \sigma)^\top$ given $\mathbf{y} = \mathbf{x}^\varepsilon(\theta^\varepsilon) + \boldsymbol{\eta}$, $\boldsymbol{\eta} \sim \rho_\eta$.

Likelihood: Needs more work $\Rightarrow$ marginalization

$$p^\varepsilon(\mathbf{y} \mid \theta^\varepsilon) = \int_{\mathbb{R}^{Nd}} p(\mathbf{y} \mid \mathbf{x}, \theta^\varepsilon)\, p^\varepsilon(\mathbf{x} \mid \theta^\varepsilon)\, \mathrm{d}\mathbf{x}.$$

where (Markov property)

$$p^\varepsilon(\mathbf{x} \mid \theta^\varepsilon) = p(x_0) \prod_{k=1}^{N} p^\varepsilon(x_k \mid x_{k-1}, \theta^\varepsilon).$$

Transition density: $p^\varepsilon(x_k \mid x_{k-1}, \theta^\varepsilon) \Rightarrow$ only "ingredient" depending on ε.

Bayesian framework

Inverse problem 1 – MS / MS

Find $\theta^\varepsilon = (\alpha, \sigma)^\top$ given $\mathbf{y} = \mathbf{x}^\varepsilon(\theta^\varepsilon) + \boldsymbol{\eta}$, $\boldsymbol{\eta} \sim \rho_\eta$.

Idea: Replace $p^0(\mathbf{x} \mid \theta^\varepsilon) \approx p^\varepsilon(\mathbf{x} \mid \theta^\varepsilon) \Rightarrow$ cheaper!

Result: Homogenized posterior $\mu^0(\theta \mid \mathbf{y})$ with density

$$p^0(\theta^\varepsilon \mid \mathbf{y}) = \frac{1}{Z^0} p(\theta^\varepsilon)\, p^0(\mathbf{y} \mid \theta^\varepsilon), \quad Z^0 \text{ s.t. } \int p^0(\theta \mid \mathbf{y})\, \mathrm{d}\theta = 1,$$

with

$$p^0(\mathbf{y} \mid \theta^\varepsilon) = \int_{\mathbb{R}^{Nd}} p(\mathbf{y} \mid \mathbf{x}, \theta^\varepsilon)\, p^0(\mathbf{x} \mid \theta^\varepsilon)\, \mathrm{d}\mathbf{x}.$$

High-dimensional integral $\Rightarrow$ Compute unbiased estimator $\hat{p}^0(\mathbf{y} \mid \theta^\varepsilon)$.

Bayesian framework

Inverse problem 1 – MS / MS

$$\text{Find } \theta^\varepsilon = (\alpha, \sigma)^\top \text{ given } \mathbf{y} = \mathbf{x}^\varepsilon(\theta^\varepsilon) + \boldsymbol{\eta},\ \boldsymbol{\eta} \sim \rho_\eta.$$

Idea: Replace $p^0(\mathbf{x} \mid \theta^\varepsilon) \approx p^\varepsilon(\mathbf{x} \mid \theta^\varepsilon) \Rightarrow$ cheaper!

Result: Homogenized posterior $\mu^0(\theta \mid \mathbf{y})$ with density

$$p^0(\theta^\varepsilon \mid \mathbf{y}) = \frac{1}{Z^0} p(\theta^\varepsilon)\, p^0(\mathbf{y} \mid \theta^\varepsilon), \quad Z^0 \text{ s.t. } \int p^0(\theta \mid \mathbf{y})\, \mathrm{d}\theta = 1,$$

Proposition

Hellinger distance $d_{\mathrm{Hell}}\big(\mu^\varepsilon(\cdot \mid \mathbf{y}), \mu^0(\cdot \mid \mathbf{y})\big) \to 0$ for $\varepsilon \to 0$.

Thank you for your attention!

References

Golightly, A. and Wilkinson, D. J. (2011). Bayesian parameter inference for stochastic biochemical network models using particle Markov chain Monte Carlo. *Interface focus*, 1(6):807–820.

Kalliadasis, S., Krumscheid, S., and Pavliotis, G. A. (2015). A new framework for extracting coarse-grained models from time series with multiscale structure. *J. Comput. Phys.*, 296:314–328.

Pavliotis, G. A. and Stuart, A. M. (2007). Parameter estimation for multiscale diffusions. *J. Stat. Phys.*, 127(4):741–781.

Pavliotis, G. A. and Stuart, A. M. (2008). *Multiscale methods: averaging and homogenization*, volume 53 of *Texts in Applied Mathematics*. Springer, New York.