Probabilistic geometric integration of ordinary differential equations

Assyr Abdulle, Giacomo Garegnani

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

MATHICSE Retreat, Sainte-Croix, June 2018

Outline

1. Motivation
2. Probabilistic methods for ODEs
3. Bayesian inverse problems

Motivation – Chaotic equations

Consider Lorenz equation (atmospheric convection)

$$\begin{aligned} x' &= \sigma(y - x), & x(0) &= -10, \\ y' &= x(\rho - z) - y, & y(0) &= -1, \\ z' &= xy - \beta z, & z(0) &= 40. \end{aligned}$$

For $\rho = 28$, $\sigma = 10$, $\beta = 8/3$ chaotic behaviour.

$\implies$ Numerical integration gives unreliable solutions.

Motivation – Chaotic equations

Consider Lorenz equation (atmospheric convection)

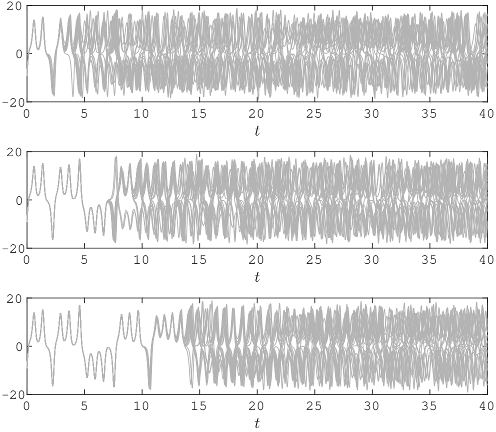
$$\begin{aligned} x' &= \sigma(y - x), & x(0) &= -10, \\ y' &= x(\rho - z) - y, & y(0) &= -1, \\ z' &= xy - \beta z, & z(0) &= 40. \end{aligned}$$

For $\rho = 28$, $\sigma = 10$, $\beta = 8/3$ chaotic behaviour.

$\implies$ Numerical integration gives unreliable solutions.

Goal: Understand reliability of numerical solutions.
Idea: Perturb the initial data e.g. with Gaussian noise on $x(0)$

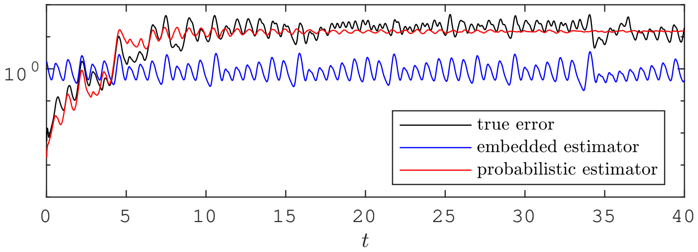
Motivation – Chaotic equations



Solutions of the Lorenz system (x component) – different perturbations

Which one has the correct magnitude wrt numerical error?

Motivation – Error estimators



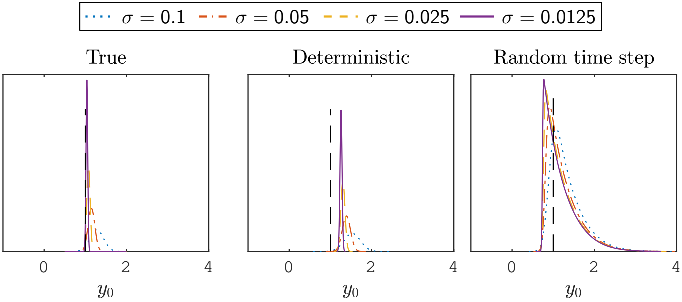
Error estimators for Lorenz given by

Probabilistic solution: use $\|\mathrm{Var}\, Y_n\|^{1/2}$ where Y_n is a probabilistic family

Classical embedded couple: Local errors don't show the true behaviour!

Goal: A posteriori error estimator $\mathrm{err}_n \approx \|\mathrm{Var}\, Y_n\|^{1/2} \rightsquigarrow$ Work in progress!

Motivation – Bayesian inverse problems



Posterior distributions (analytic) on y_0 for $y' = -y$, $y(0) = y_0$. One observation corrupted by noise $\mathcal{N}(0, \sigma^2)$ with truth y_0^*.
For $\sigma \to 0$:

- True solution: Posterior converging to $\delta_{y_0^*} \to$ good
- Runge-Kutta: Posterior converging to Dirac delta on wrong value $\to$ bad
- Probabilistic method: Posterior variance $\approx$ numerical error

Outline

Probabilistic methods for ODEs

Filtering methods for ODEs: fix a prior on $y(t)$ (Gaussian process), update with evaluations of $f(y)$.

- Kersting and Hennig (2016)
- Chkrebtii et al. (2016)
- Schober et al. (2014)
- ...

Randomised methods for ODEs: random perturbation of deterministic numerical solutions → sampling

- Conrad et al. (2016)
- Lie et al. (2017)
- Abdulle and Garegnani (2018)
- ...

Probabilistic methods for ODEs

Filtering methods for ODEs: fix a prior on $y(t)$ (Gaussian process), update with evaluations of $f(y)$.

- Kersting and Hennig (2016)
- Chkrebtii et al. (2016)
- Schober et al. (2014)
- ...

Randomised methods for ODEs: random perturbation of deterministic numerical solutions $\rightarrow$ sampling

- Conrad et al. (2016)
- Lie et al. (2017)
- Abdulle and Garegnani (2018)
- ...

Notation

Autonomous dynamical system, function $f : \mathbb{R}^d \to \mathbb{R}^d$ and the ODE

$$y' = f(y), \quad y(0) = y_0.$$

Flow of the equation φ_t

$$y(t) = \varphi_t(y_0).$$

One-step method (e.g. Runge Kutta): numerical flow Ψ_h

$$y_{n+1} = \Psi_h(y_n).$$

Additive noise method

Stochastic process $\{Y_n\}_{n=1,2,\ldots}$ with recurrence

$$Y_{n+1} = \underbrace{\Psi_h(Y_n)}_{\text{deterministic}} + \underbrace{\xi_n(h)}_{\text{random}}.$$

Main assumption: For $p > 1$ and $Q \in \mathbb{R}^{d \times d}$

$$\xi_n(h) \overset{\text{iid}}{\sim} \mathcal{N}(0, Qh^{2p+1}).$$

Additive noise method

Stochastic process $\{Y_n\}_{n=1,2,\dots}$ with recurrence

$$Y_{n+1} = \underbrace{\Psi_h(Y_n)}_{\text{deterministic}} + \underbrace{\xi_n(h)}_{\text{random}}.$$

Main assumption: For $p > 1$ and $Q \in \mathbb{R}^{d \times d}$

$$\xi_n(h) \overset{\text{iid}}{\sim} \mathcal{N}(0, Qh^{2p+1}).$$

Properties

If Ψ_h is of order q and for $\Phi \colon \mathbb{R}^d \to \mathbb{R}$ smooth

- Strong convergence: $\mathbb{E}\|y(hn) - Y_n\| \leq Ch^{\min\{p,q\}}$,
- Weak convergence: $|\Phi\big(y(hn)\big) - \mathbb{E}\,\Phi(Y_n)| \leq Ch^{\min\{2p,q\}}$,
- Good qualitative behavior in Bayesian inverse problems.

Additive noise method

Stochastic process $\{Y_n\}_{n=1,2,\dots}$ with recurrence

$$Y_{n+1} = \underbrace{\Psi_h(Y_n)}_{\text{deterministic}} + \underbrace{\xi_n(h)}_{\text{random}}.$$

Main assumption: For $p > 1$ and $Q \in \mathbb{R}^{d \times d}$

$$\xi_n(h) \overset{\text{iid}}{\sim} \mathcal{N}(0, Qh^{2p+1}).$$

Issues

- Robustness: $\Psi_h(Y_{n-1}) > 0 \;\not\Longrightarrow\; \mathbb{P}(Y_n < 0) = 0$,
- Geometric properties are not conserved from Ψ_h.

Random time steps

Intrinsic noise: Random time-stepping Runge-Kutta (RTS-RK)

$$Y_{n+1} = \Psi_{H_n}(Y_n),$$

Main assumption: $\{H_n\}_{n=0,1,\dots}$ iid such that for $h, C > 0$ and $p > 1$

$$H_n > 0 \text{ a.s.}, \quad \mathbb{E}\, H_n = h, \quad \operatorname{Var} H_n = Ch^{2p+1}.$$

Example: $H_n \overset{\text{iid}}{\sim} \mathcal{U}(h - h^{p+1/2}, h + h^{p+1/2})$.

Random time steps

Intrinsic noise: Random time-stepping Runge-Kutta (RTS-RK)

$$Y_{n+1} = \Psi_{H_n}(Y_n),$$

Main assumption: $\{H_n\}_{n=0,1,\dots}$ iid such that for $h, C > 0$ and $p > 1$

$$H_n > 0 \text{ a.s.}, \quad \mathbb{E}\, H_n = h, \quad \operatorname{Var} H_n = Ch^{2p+1}.$$

Example: $H_n \overset{\text{iid}}{\sim} \mathcal{U}(h - h^{p+1/2}, h + h^{p+1/2})$.

Properties

If Ψ_h is of order q and for $\Phi\colon \mathbb{R}^d \to \mathbb{R}$ smooth

- Strong convergence: $\mathbb{E}\|y(hn) - Y_n\| \leq Ch^{\min\{p,q\}}$,
- Weak convergence: $|\Phi\big(y(hn)\big) - \mathbb{E}\,\Phi(Y_n)| \leq Ch^{\min\{2p,q\}}$,
- Good qualitative behavior in Bayesian inverse problems.

Random time steps

Intrinsic noise: Random time-stepping Runge-Kutta (RTS-RK)

$$Y_{n+1} = \Psi_{H_n}(Y_n),$$

Main assumption: $\{H_n\}_{n=0,1,\dots}$ iid such that for $h, C > 0$ and $p > 1$

$$H_n > 0 \text{ a.s.}, \quad \mathbb{E}\, H_n = h, \quad \operatorname{Var} H_n = Ch^{2p+1}.$$

Example: $H_n \overset{\text{iid}}{\sim} \mathcal{U}(h - h^{p+1/2}, h + h^{p+1/2})$.

Properties – Geometric

- Conservation of (polynomial) first integrals is inherited by Ψ_h,
- Flow map is symplectic if Ψ_h is symplectic,
- Long-time conservation of energy in Hamiltonian systems.

Conservation of first integrals – Additive noise

Recall: $Y_{n+1} = \Psi_h(Y_n) + \xi_n(h)$, with $\mathbb{E}\,\xi_n(h)\xi_n(h)^\top = h^{2p+1}Q$

Linear first integrals: $I(y) = v^\top y$ such that $I(\Psi_h(Y_1)) = I(y_0)$. Then

$$I(Y_1) = v^\top(y_0 + \xi_0(h)) \implies \mathbb{E}\,I(Y_1) = I(y_0) \text{ iff } \mathbb{E}\,\xi_0(h) = 0.$$

Quadratic first integrals: $I(y) = y^\top Sy$ such that $I(\Psi_h(Y_1)) = I(y_0)$. Then

$$\begin{aligned} I(Y_1) &= I(y_0) + 2\xi_0(h)^T S\Psi_h(y_0) + \xi_0(h)^T S\xi_0(h), \\ \implies \mathbb{E}\,I(Y_1) &= I(y_0) + Q : Sh^{2p+1}, \quad (\text{with } \mathbb{E}\,\xi_0(h) = 0) \end{aligned}$$

Quadratic first integrals are not conserved on average!

Conservation of first integrals

Theorem (Conservation of invariants)

If the Runge-Kutta scheme defined by Ψ_h conserves an invariant $I(y)$ for an ODE, then the RTS-RK method conserves $I(y)$ for the same ODE.

Proof

If $I(\Psi_h(y)) = I(y)$ for any h, then $I(\Psi_{H_0}(y)) = I(y)$ for any value that H_0 can assume.

Symplecticity

Energy $Q\colon \mathbb{R}^{2d} \to \mathbb{R}$ and Hamiltonian system

$$y' = J^{-1}\nabla Q(y), \quad J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}.$$

Symplectic integrator Ψ_h of order q.

Theorem (Strong approximation of the Hamiltonian)

There exist positive constants κ, C_1, C_2, C_3, independent of h such that

$$\mathbb{E}|Q(Y_n) - Q(y_0)| \leq C_1 h^q + C_2 t^{1/2} h^{p+q}.$$

Symplecticity

Energy $Q\colon \mathbb{R}^{2d} \to \mathbb{R}$ and Hamiltonian system

$$y' = J^{-1}\nabla Q(y), \quad J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}.$$

Symplectic integrator Ψ_h of order q.

Theorem (Strong approximation of the Hamiltonian)

There exist positive constants κ, C_1, C_2, C_3, independent of h such that

$$\mathbb{E}|Q(Y_n) - Q(y_0)| \leq C_1 h^q + C_2 t^{1/2} h^{p+q}.$$

Consequence

If $q = p$, up to times $n = \mathcal{O}(h^{-2q})$ (balance between h^q and h^{2q} terms) same conservation as deterministic symplectic method.

Numerical experiments – Geometric properties

Consider the perturbed Kepler equation (model for two-body problem)

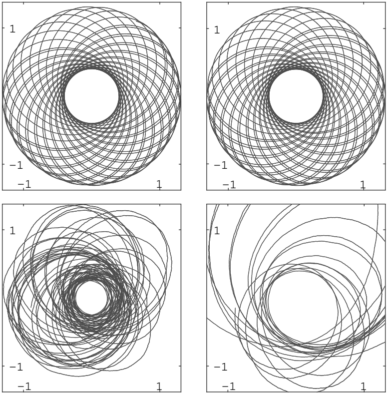
$$w_1' = v_1, \quad v_1' = -\frac{w_1}{\|w\|^3} - \frac{\delta w_1}{\|w\|^5},$$
$$w_2' = v_2, \quad v_2' = -\frac{w_2}{\|w\|^3} - \frac{\delta w_2}{\|w\|^5}.$$

The angular momentum is conserved (quadratic first integral)

$$I(v, w) = w_1 v_2 - w_2 v_1$$

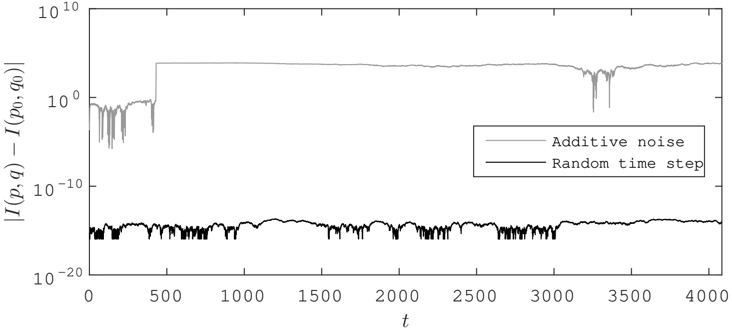
$\rightarrow$ employ a Gauss method (implicit midpoint rule).

Numerical experiments – Geometric properties



RTS-RK (first row), Additive noise (second row). Time $0 \leq t \leq 200$ and $200 \leq t \leq 400$ (left and right)

Numerical experiments – Geometric properties



Conservation of the angular momentum (quadratic first integral)

Numerical experiments – Geometric properties

Consider the pendulum system, Hamiltonian with energy

$$Q(v,w) = \frac{1}{2}v^2 - \cos(w).$$

Energy is separable: $Q(v,w) = Q_1(v) + Q_2(w)$.

Störmer-Verlet scheme – order 2

$$\begin{aligned} v_{n+1/2} &= v_n - \frac{h}{2}Q_w(v_{n+1/2}, w_n), \\ w_{n+1} &= w_n + \frac{h}{2}\big(Q_v(v_{n+1/2}, w_n) + Q_v(v_{n+1/2}, w_{n+1})\big), \\ v_{n+1} &= v_n - \frac{h}{2}Q_w(v_{n+1/2}, w_{n+1}). \end{aligned}$$

Numerical experiments – Geometric properties

Consider the pendulum system, Hamiltonian with energy

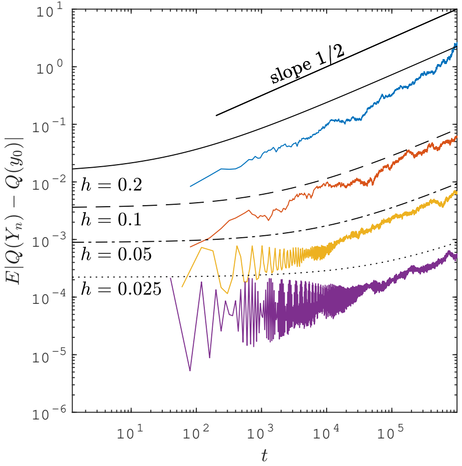
$$Q(v, w) = \frac{1}{2}v^2 - \cos(w).$$

Energy is separable: $Q(v, w) = Q_1(v) + Q_2(w)$.

Störmer-Verlet scheme – order 2, explicit

$$\begin{aligned} v_{n+1/2} &= v_n - \frac{h}{2} Q_2'(w_n), \\ w_{n+1} &= w_n + h Q_1'(v_{n+1/2}), \\ v_{n+1} &= v_n - \frac{h}{2} Q_2'(w_{n+1}). \end{aligned}$$

Numerical experiments – Geometric properties



Mean error on the Hamiltonian for different values of the time step h.

Outline

1. Motivation

2. Probabilistic methods for ODEs

3. Bayesian inverse problems

Bayesian inverse problems

Goal

Given $\vartheta \in \mathbb{R}^n$, $f_\vartheta \colon \mathbb{R}^d \to \mathbb{R}^d$ and the ODE

$$y' = f_\vartheta(y), \quad y(0) = y_{0,\vartheta} \in \mathbb{R}^d,$$

retrieve the true value ϑ^* from observations of $y(t)$, $t > 0$.

Bayesian inverse problems

Goal

Given $\vartheta \in \mathbb{R}^n$, $f_\vartheta \colon \mathbb{R}^d \to \mathbb{R}^d$ and the ODE

$$y' = f_\vartheta(y), \quad y(0) = y_{0,\vartheta} \in \mathbb{R}^d,$$

retrieve the true value ϑ^* from observations of $y(t)$, $t > 0$.

Bayesian setting: fix prior $\pi_{\mathrm{prior}}(\vartheta)$, consider the forward operator $\mathcal{G}$ and model observations as

$$\underbrace{\mathcal{Y}}_{\text{observations}} = \underbrace{\mathcal{G}(\vartheta^*)}_{\text{forward}} + \underbrace{\varepsilon}_{\text{noise}}, \quad \varepsilon \sim \pi_{\mathrm{noise}},$$

then the posterior distribution (density) is

$$\pi(\vartheta \mid \mathcal{Y}) \propto \pi_{\mathrm{prior}}(\vartheta)\pi_{\mathrm{noise}}(\mathcal{Y} - \mathcal{G}(\vartheta)).$$

Bayesian inverse problems

The posterior $\pi(\vartheta \mid \mathcal{Y})$ is not computable, approximate with

$$\pi^{h}(\vartheta \mid \mathcal{Y}) \propto \pi_{\text{prior}}(\vartheta)\pi_{\text{noise}}(\mathcal{Y} - \mathcal{G}^{h}(\vartheta)).$$

Bayesian inverse problems

The posterior $\pi(\vartheta \mid \mathcal{Y})$ is not computable, approximate with

$$\pi^h(\vartheta \mid \mathcal{Y}) \propto \pi_{\mathrm{prior}}(\vartheta)\pi_{\mathrm{noise}}(\mathcal{Y} - \mathcal{G}^h(\vartheta)).$$

Properties

$\pi^h \to \pi$ for $h \to 0$ (in the Hellinger distance).

Issue

- π^h concentrated around values "far" from ϑ^* $\to$ non-predictive posterior

Bayesian inverse problems

The posterior $\pi(\vartheta \mid \mathcal{Y})$ is not computable, approximate with

$$\pi^{\boldsymbol{h},\mathrm{RTS}}(\vartheta \mid \mathcal{Y}) \propto \pi_{\mathrm{prior}}(\vartheta)\mathbb{E}^{\mathbf{H}}\pi_{\mathrm{noise}}(\mathcal{Y} - \mathcal{G}^{\mathbf{H}}(\vartheta)),$$

where $\mathbf{H} = (H_0, H_1, \dots)$ is the vector of all time steps chosen in one run.

Bayesian inverse problems

The posterior $\pi(\vartheta \mid \mathcal{Y})$ is not computable, approximate with

$$\pi^{h,\mathrm{RTS}}(\vartheta \mid \mathcal{Y}) \propto \pi_{\mathrm{prior}}(\vartheta)\mathbb{E}^{\mathbf{H}}\pi_{\mathrm{noise}}(\mathcal{Y} - \mathcal{G}^{\mathbf{H}}(\vartheta)),$$

where $\mathbf{H} = (H_0, H_1, \ldots)$ is the vector of all time steps chosen in one run.

Properties

- $\pi^{h,\mathrm{RTS}} \to \pi$ for $h \to 0$ (in the Hellinger distance). Lie et al. (2017)
- "correct" the non-predictive behaviour of deterministic approximations

Warning

- Approximation of $\mathbb{E}^{\mathbf{H}}\,\pi_{\mathrm{noise}}(\mathcal{Y} - \mathcal{G}^{\mathbf{H}}(\vartheta))$ is required

Numerical experiment – Bayesian inverse problems

Consider the Hénon-Heiles system (motion of a star around a galactic center), Hamiltonian with energy

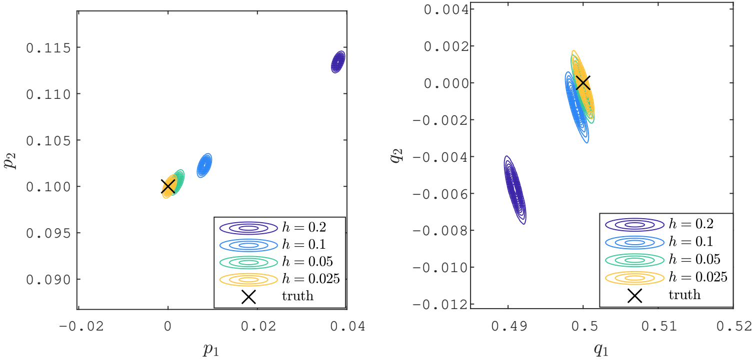
$$E(v, w) = \frac{1}{2}\|v\|^2 + \frac{1}{2}\|w\|^2 + w_1^2 w_2 - \frac{1}{3}w_2^3.$$

Chaotic problem for certain levels of energy.

Goal

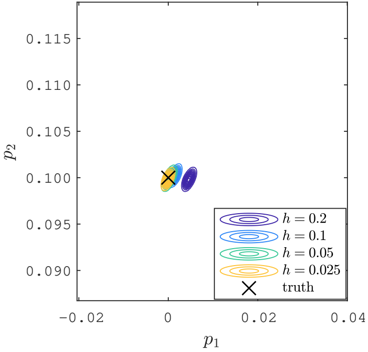
Find posterior $\pi((v_0, w_0) \mid \mathcal{Y})$ over the initial condition from a single observation of $(v(10), w(10))$

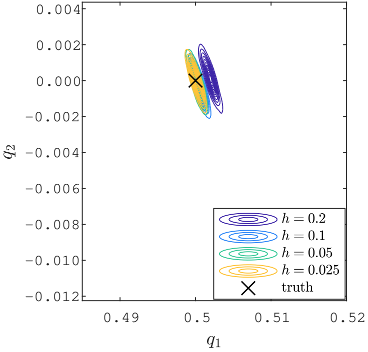
Numerical experiment – Bayesian inverse problems



Posterior distributions given by deterministic Heun method.

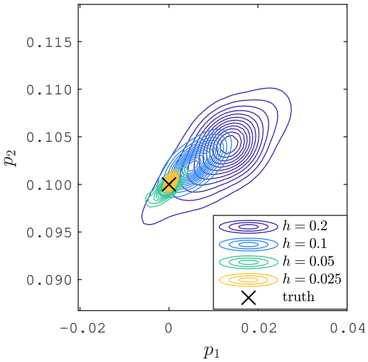
Numerical experiment – Bayesian inverse problems





Posterior distributions given by deterministic Störmer-Verlet method.

Numerical experiment – Bayesian inverse problems



Posterior distributions given by RTS-RK Störmer-Verlet method.

References

A. Abdulle, G. Garegnani (2018). Random time step probabilistic methods for uncertainty quantification in chaotic and geometric numerical integration. *preprint arXiv:1801.01340*

O. A. Chkrebtii, D. A. Campbell, B. Calderhead, M. A. Girolami (2016). Bayesian solution uncertainty quantification for differential equations. *Bayesian Anal.*

Conrad, P. R., Girolami, M., Särkkä, S., Stuart, A., and Zygalakis, K. (2016). Statistical analysis of differential equations: introducing probability measures on numerical solutions. *Stat. Comput.*

Kersting, H. and Hennig, P. (2016). Active uncertainty calibration in Bayesian ODE solvers. In *Proceedings of the 32nd Conference on Uncertainty in Artificial Intelligence (UAI 2016)*, pages 309–318. AUAI Press.

Lie, H. C., Sullivan, T. J., and Stuart, A. (2017). Strong convergence rates of probabilistic integrators for ordinary differential equations *preprint arXiv:1703.03680*

Lie, H. C., Sullivan, T. J., and Teckentrup, A. L. (2017). Random forward models and log-likelihoods in Bayesian inverse problems. *preprint arXiv:1712.05717*

Schober, M, Duvenaud, D. and Hennig, P. (2014). Probabilistic ODE Solvers with Runge-Kutta Means *preprint arXiv:1406.2582*