Uncertainty quantification of numerical errors in geometric integration via random time steps

Assyr Abdulle, Giacomo Garegnani

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Tübingen, 6 March 2017

Outline

1. Geometric numerical integration
2. Stochastic methods for ODEs
   - Additive noise method
   - Random time steps
3. Geometric stochastic numerical integration
4. Bayesian inverse problems
5. Numerical experiments
   - Convergence
   - Geometric properties
   - Bayesian inverse problems

Outline

Notation

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

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

Flow of the equation $\varphi_t\colon \mathbb{R}^d \to \mathbb{R}^d$ such that

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

One-step method: numerical flow Ψ_h such that

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

Runge-Kutta methods: flow implicitly defined by

$$K_i = y_n + h\sum_{j=1}^{s} a_{ij} f(K_j),$$

$$\Psi_h(y_n) = y_n + h\sum_{i=1}^{s} b_i f(K_i).$$

Geometric numerical integration

Goal

Develop numerical methods that preserve geometric properties of certain dynamical systems.

First integral of motion $I\colon \mathbb{R}^d \to \mathbb{R}$

$$I(\varphi_t(y_0)) = I(y_0), \quad \forall t > 0.$$

Example: quadratic first integral, given $S \in \mathbb{R}^{d\times d}$, $v \in \mathbb{R}^d$

$$I(y) = y^\top S y + v^\top y,$$

conserved by all Gauss collocation methods (e.g., implicit midpoint, ...).

Theorem (Polynomial first integrals)

No Runge-Kutta method can conserve all polynomial first integrals of degree $\mathrm{Deg}(I) \geq 3$.

Geometric numerical integration

Goal

Develop numerical methods that preserve geometric properties of certain dynamical systems.

Hamiltonian systems: Given $Q\colon \mathbb{R}^{2d} \to \mathbb{R}$, define

$$y'(t) = J^{-1}\nabla Q(y), \qquad y(0) = y_0$$
$$J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}, \qquad I \text{ identity in } \mathbb{R}^{d\times d}$$

The flow φ_t is symplectic

$$\varphi_t'(y)^\top J \varphi_t'(y) = J \implies \text{Conservation of volumes}$$

Symplectic numerical methods

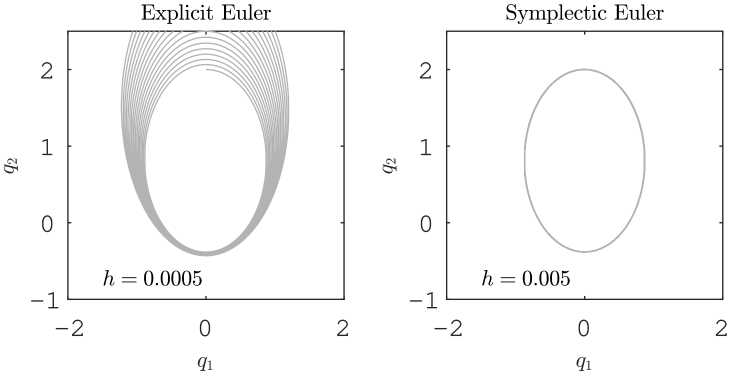
$$\Psi_h'(y)^\top J \Psi_h'(y) = J$$

Geometric numerical integration

Example

Two-body problem (planetary orbits), $y = (p, q)^\top \in \mathbb{R}^4$

$$Q(p, q) = \frac{1}{2}(p_1^2 + p_2^2) - \frac{1}{\sqrt{q_1^2 + q_2^2}}$$

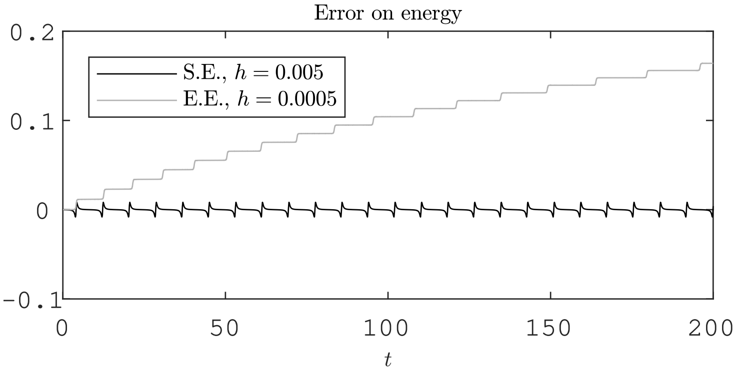


Geometric numerical integration

Example

Two-body problem (planetary orbits), $y = (p, q)^\top \in \mathbb{R}^4$

$$Q(p, q) = \frac{1}{2}(p_1^2 + p_2^2) - \frac{1}{\sqrt{q_1^2 + q_2^2}}$$



Geometric numerical integration

Given Hamiltonian $Q(p, q)$.

Symplectic Euler method – order 1

$$
\begin{aligned}
p_{n+1} &= p_n - hQ_q(p_{n+1}, q_n), \\
q_{n+1} &= q_n + hQ_p(p_{n+1}, q_n).
\end{aligned}
$$

Störmer-Verlet scheme – order 2

$$
\begin{aligned}
p_{n+1/2} &= p_n - \frac{h}{2} Q_q(p_{n+1/2}, q_n), \\
q_{n+1} &= q_n + \frac{h}{2} \big(Q_p(p_{n+1/2}, q_n) + Q_p(p_{n+1/2}, q_{n+1}) \big), \\
p_{n+1} &= p_n - \frac{h}{2} Q_q(p_{n+1/2}, q_{n+1}).
\end{aligned}
$$

Geometric numerical integration

Given separable Hamiltonian $Q(p,q) = Q_1(p) + Q_2(q)$.

Symplectic Euler method – order 1, explicit

$$p_{n+1} = p_n - hQ_2'(q_n),$$
$$q_{n+1} = q_n + hQ_1'(p_{n+1}).$$

Störmer-Verlet scheme – order 2, explicit

$$p_{n+1/2} = p_n - \frac{h}{2}Q_2'(q_n),$$
$$q_{n+1} = q_n + hQ_1'(p_{n+1/2}),$$
$$p_{n+1} = p_n - \frac{h}{2}Q_2'(q_{n+1}).$$

Several examples of separable Hamiltonians (Two-body problem, ...)

Outline

Stochastic methods for ODEs

Probabilistic / Bayesian methods for ODEs: fix a prior on $y(t)$ (Gaussian process), update with evaluations of $f(y)$ [Kersting and Hennig, 2016]

Stochastic / Randomised methods for ODEs: random perturbation of deterministic numerical solutions $\to$ sampling [Conrad et al., 2016]

Stochastic methods for ODEs

Probabilistic / Bayesian methods for ODEs: fix a prior on $y(t)$ (Gaussian process), update with evaluations of $f(y)$ [Kersting and Hennig, 2016]

Stochastic / Randomised methods for ODEs: random perturbation of deterministic numerical solutions $\rightarrow$ sampling [Conrad et al., 2016]

Additive noise method [Conrad et al., 2016]

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: $\{\xi_n\}_{n=0,1,\dots}$ iid such that for $p > 1$ and $Q \in \mathbb{R}^{d\times d}$

$$\mathbb{E}\,\xi_n(h) = 0, \quad \mathbb{E}\,\xi_n(h)\xi_n(h)^T = Qh^{2p+1}.$$

Additive noise method [Conrad et al., 2016]

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: $\{\xi_n\}_{n=0,1,\ldots}$ iid such that for $p > 1$ and $Q \in \mathbb{R}^{d\times d}$

$$\mathbb{E}\,\xi_n(h) = 0, \quad \mathbb{E}\,\xi_n(h)\xi_n(h)^T = Qh^{2p+1}.$$

Properties

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

Additive noise method [Conrad et al., 2016]

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: $\{\xi_n\}_{n=0,1,\dots}$ iid such that for $p > 1$ and $Q \in \mathbb{R}^{d\times d}$

$$\mathbb{E}\,\xi_n(h) = 0, \quad \mathbb{E}\,\xi_n(h)\xi_n(h)^T = Qh^{2p+1}.$$

Properties

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

- Strong convergence: $\mathbb{E}\|y(hn) - Y_n\| \leq Ch^{\min\{p,q\}}$,

Additive noise method [Conrad et al., 2016]

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: $\{\xi_n\}_{n=0,1,\dots}$ iid such that for $p > 1$ and $Q \in \mathbb{R}^{d\times d}$

$$\mathbb{E}\,\xi_n(h) = 0, \quad \mathbb{E}\,\xi_n(h)\xi_n(h)^T = Qh^{2p+1}.$$

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: $|\Phi\big(y(hn)\big) - \mathbb{E}\,\Phi(Y_n)| \leq Ch^{\min\{2p,q\}}$,

Additive noise method [Conrad et al., 2016]

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: $\{\xi_n\}_{n=0,1,\dots}$ iid such that for $p > 1$ and $Q \in \mathbb{R}^{d\times d}$

$$\mathbb{E}\,\xi_n(h) = 0, \quad \mathbb{E}\,\xi_n(h)\xi_n(h)^T = Qh^{2p+1}.$$

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 [Conrad et al., 2016]

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: $\{\xi_n\}_{n=0,1,\ldots}$ iid such that for $p > 1$ and $Q \in \mathbb{R}^{d\times d}$

$$\mathbb{E}\,\xi_n(h) = 0, \quad \mathbb{E}\,\xi_n(h)\xi_n(h)^T = Qh^{2p+1}.$$

Issues

- Robustness: $\Psi_h(Y_{n-1}) > 0 \;\not\Longrightarrow\; \mathbb{P}(Y_n < 0) = 0$,

Additive noise method [Conrad et al., 2016]

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: $\{\xi_n\}_{n=0,1,\dots}$ iid such that for $p > 1$ and $Q \in \mathbb{R}^{d\times d}$

$$\mathbb{E}\,\xi_n(h) = 0, \quad \mathbb{E}\,\xi_n(h)\xi_n(h)^T = Qh^{2p+1}.$$

Issues

- Robustness: $\Psi_h(Y_{n-1}) > 0 \not\Rightarrow \mathbb{P}(Y_n < 0) = 0$,
- Geometric properties are not conserved from Ψ_h. For example if $I(y) = y^T S y$ and $I(\Psi_h(y_0)) = I(y_0)$

$$I(Y_1) = I(y_0) + 2\xi_0(h)^T S \Psi_h(y_0) + \xi_0(h)^T S \xi_0(h).$$

Random time steps [Abdulle and Garegnani, 2018]

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 $p > 1$

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

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

Random time steps [Abdulle and Garegnani, 2018]

Theorem (Weak convergence)

There exists $C > 0$ *independent of* h *such that for all smooth functions* $\Phi\colon \mathbb{R}^d \to \mathbb{R}$

$$|\mathbb{E}\,\Phi(Y_k) - \Phi(y(kh)))| \leq Ch^{\min\{2p-1,q\}},$$

for all $k = 1, 2, \ldots, N$*.*

Random time steps [Abdulle and Garegnani, 2018]

Theorem (Weak convergence)

There exists $C > 0$ independent of h such that for all smooth functions $\Phi\colon \mathbb{R}^d \to \mathbb{R}$

$$|\mathbb{E}\,\Phi(Y_k) - \Phi(y(kh)))| \leq Ch^{\min\{2p-1,q\}},$$

for all $k = 1, 2, \dots, N$.

Idea of the proof

- Taylor expansion of Φ and Ψ_h,
- Bound the one-step error considering the distance between the generators of φ_h and Ψ_{H_0},
- Consider the distance between φ_h, Ψ_h and Ψ_{H_0},
- Propagate in time (Markov property) to obtain a global estimate.

Random time steps [Abdulle and Garegnani, 2018]

Theorem (Mean square convergence)

There exists $C > 0$ independent of h such that

$$\left(\mathbb{E}\|Y_k - y(t_k)\|^2\right)^{1/2} \leq Ch^{\min\{p-1/2,q\}},$$

for all $k = 1, 2, \ldots, N$.

Random time steps [Abdulle and Garegnani, 2018]

Theorem (Mean square convergence)

There exists $C > 0$ independent of h such that

$$\left(\mathbb{E}\|Y_k - y(t_k)\|^2\right)^{1/2} \leq Ch^{\min\{p-1/2,q\}},$$

for all $k = 1, 2, \dots, N$.

Idea of the proof

- Bound the one-step error (triangular inequality),
- Analyse the impact of discretisation and randomisation separately.
- Propagate in time to obtain a global estimate.

Random time steps [Abdulle and Garegnani, 2018]

Theorem (Mean square convergence)

There exists $C > 0$ independent of h such that

$$\left(\mathbb{E}\| Y_k - y(t_k)\|^2\right)^{1/2} \leq Ch^{\min\{p-1/2,q\}},$$

for all $k = 1, 2, \ldots, N$.

Idea of the proof

- Bound the one-step error (triangular inequality),
- Analyse the impact of discretisation and randomisation separately.
- Propagate in time to obtain a global estimate.

Consequences

- Reasonable choice $p = q + 1/2$
- $\mathbb{E}\| Y_k - y(t_k)\| \leq Ch^{\min\{p-1/2,q\}}$ (strong order)

Random time steps [Abdulle and Garegnani, 2018]

Theorem (Monte Carlo estimators)

For $\Phi\colon \mathbb{R}^d \to \mathbb{R}$ smooth, Monte Carlo estimators $\hat{Z} = M^{-1}\sum_{i=1}^{M}\Phi(Y_N^{(i)})$ of $Z = \Phi(Y_N)$ satisfy

$$\mathrm{MSE}(\hat{Z}) \leq C\Big(h^{2\min\{2p-1,q\}} + \frac{h^{2\min\{p-1/2,q\}}}{M}\Big),$$

where C is a positive constant independent of h and M and

$$\mathrm{MSE}(\hat{Z}) = \mathbb{E}\left(\hat{Z} - \Phi(y(t_N))\right)^2.$$

Random time steps [Abdulle and Garegnani, 2018]

Theorem (Monte Carlo estimators)

For $\Phi\colon \mathbb{R}^d \to \mathbb{R}$ *smooth, Monte Carlo estimators* $\hat{Z} = M^{-1} \sum_{i=1}^{M} \Phi(Y_N^{(i)})$ *of* $Z = \Phi(Y_N)$ *satisfy*

$$\mathrm{MSE}(\hat{Z}) \leq C\left(h^{2\min\{2p-1,q\}} + \frac{h^{2\min\{p-1/2,q\}}}{M}\right),$$

where C *is a positive constant independent of* h *and* M *and*

$$\mathrm{MSE}(\hat{Z}) = \mathbb{E}\left(\hat{Z} - \Phi(y(t_N))\right)^2.$$

Idea of the proof

Use the "bias-variance" decomposition of the MSE and apply weak and mean-square convergence results.

Random time steps [Abdulle and Garegnani, 2018]

Theorem (Monte Carlo estimators)

For $\Phi\colon \mathbb{R}^d \to \mathbb{R}$ *smooth, Monte Carlo estimators* $\hat{Z} = M^{-1}\sum_{i=1}^{M} \Phi(Y_N^{(i)})$ *of* $Z = \Phi(Y_N)$ *satisfy*

$$\mathrm{MSE}(\hat{Z}) \leq C\left(h^{2\min\{2p-1,q\}} + \frac{h^{2\min\{p-1/2,q\}}}{M}\right),$$

where C *is a positive constant independent of* h *and* M *and*

$$\mathrm{MSE}(\hat{Z}) = \mathbb{E}\left(\hat{Z} - \Phi(y(t_N))\right)^2.$$

Consequence

For reasonable choice $p = q + 1/2$, $\mathrm{MSE}(\hat{Z})$ converges independently of M with h (quality of the estimation independent of the number of paths)

Outline

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 – Random time steps

Theorem (Conservation of polynomial 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 – Random time steps

Theorem

If the flow Ψ_h of the deterministic integrator is symplectic, then the flow of the RTS-RK method is symplectic.

Idea of the proof

Adaptive time steps ruin symplectic properties if not carefully selected. Nonetheless, if the time steps are chosen independently of the solution, the flow is symplectic.

Symplecticity – Random time steps

Theorem

If the flow Ψ_h of the deterministic integrator is symplectic, then the flow of the RTS-RK method is symplectic.

Idea of the proof

Adaptive time steps ruin symplectic properties if not carefully selected. Nonetheless, if the time steps are chosen independently of the solution, the flow is symplectic.

Remark

The symplecticity of the flow is not enough to guarantee good approximation of the Hamiltonian for long time spans.

Outline

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 $\mathcal{G} \colon \mathbb{R}^n \to \mathbb{R}^m$ and the observation model

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

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

Obtaining a sample $\{\vartheta^{(i)}\}_{i=0}^{N}$ from $\pi(\vartheta \mid \mathcal{Y})$.

Algorithm: Metropolis-Hastings.

Given $\vartheta^{(0)} \in \mathbb{R}^n$, proposal $q\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, $N \in \mathbb{N}$;

Compute $\pi(\vartheta^{(0)} \mid \mathcal{Y})$;

for $i = 0, \ldots, N$ **do**

Draw $\bar{\vartheta}$ from $q(\vartheta^{(i)}, \cdot)$;

Set $\vartheta^{(i+1)} = \bar{\vartheta}$ with probability

$$\alpha(\vartheta^{(i)}, \bar{\vartheta}) = \min\left\{1, \frac{\pi(\bar{\vartheta} \mid \mathcal{Y}) q(\vartheta^{(i)}, \bar{\vartheta})}{\pi(\vartheta^{(i)} \mid \mathcal{Y}) q(\bar{\vartheta}, \vartheta^{(i)})}\right\}$$

otherwise set $\vartheta^{(i+1)} = \vartheta^{(i)}$;

end

Bayesian inverse problems

Obtaining a sample $\{\vartheta^{(i)}\}_{i=0}^{N}$ from $\pi(\vartheta \mid \mathcal{Y})$.

Algorithm: Metropolis-Hastings.

Given $\vartheta^{(0)} \in \mathbb{R}^n$, proposal $q\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, $N \in \mathbb{N}$;

Compute $\pi(\vartheta^{(0)} \mid \mathcal{Y})$;

for $i = 0, \ldots, N$ **do**

Draw $\bar{\vartheta}$ from $q(\vartheta^{(i)}, \cdot)$;

Set $\vartheta^{(i+1)} = \bar{\vartheta}$ with probability

$$\alpha(\vartheta^{(i)}, \bar{\vartheta}) = \min\left\{1, \frac{\pi(\bar{\vartheta} \mid \mathcal{Y}) q(\vartheta^{(i)}, \bar{\vartheta})}{\pi(\vartheta^{(i)} \mid \mathcal{Y}) q(\bar{\vartheta}, \vartheta^{(i)})}\right\}$$

otherwise set $\vartheta^{(i+1)} = \vartheta^{(i)}$;

end

Bayesian inverse problems

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

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

Bayesian inverse problems

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

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

Properties

If Ψ_h is of order q

- $d_{\text{Hell}}(\pi^h, \pi) \to 0$ for $h \to 0$ with rate q [Stuart, 2010]
- fast MH iterations for explicit Ψ_h (and h coarse)
- explores complex posterior distributions

Bayesian inverse problems

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

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

Issue

- π^h concentrated around values "far" from $\vartheta^* \to$ 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 $\mathbf{H} = (H_0, H_1, \ldots)$.

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 $\mathbf{H} = (H_0, H_1, \ldots)$.

Properties

If $\Psi_h \to \varphi_h$ for $h \to 0$

- $d_{\mathrm{Hell}}(\pi^{h,\mathrm{RTS}}, \pi) \to 0$ for $h \to 0$ (it can be shown)
- "correct" the non-predictive behaviour of deterministic approximations
- explores complex posterior distributions

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 $\mathbf{H} = (H_0, H_1, \ldots)$.

Issues

- Approximation of $\mathbb{E}^{\mathbf{H}}\,\pi_{\mathrm{noise}}(\mathcal{Y} - \mathcal{G}^{\mathbf{H}}(\vartheta))$ is required
- Employ pseudo-marginal MH $\rightarrow$ slow mixing for small noise
- Employ noisy pseudo-marginal MH $\rightarrow$ inexact posterior distributions

Outline

Numerical experiments – Convergence

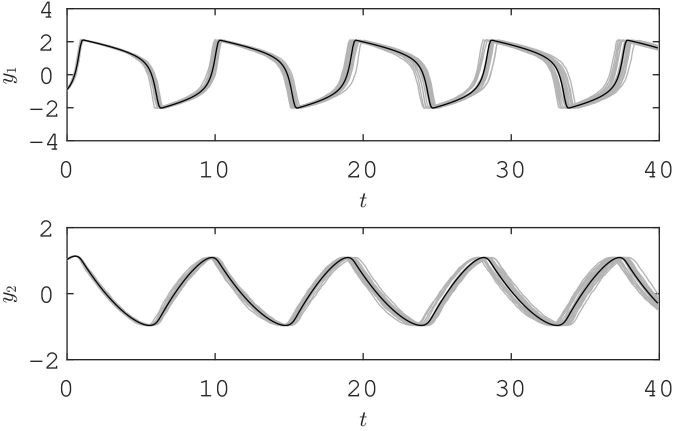
Consider FitzHug-Nagumo equation (model for neuron activity)

$$
\begin{aligned}
y_1' &= c\big(y_1 - \frac{y_1^3}{3} + y_2\big), && y_1(0) = -1, \\
y_2' &= -\frac{1}{c}(y_1 - a + b y_2), && y_2(0) = 1, \\
& a = 0.2, \quad b = 0.2 \quad c = 3.
\end{aligned}
$$

We verify numerically the results of convergence.

Numerical experiments – Convergence

Solution with Explicit Euler ($h = 0.1$, $T = 40$)



black – deterministic solution, grey – realizations of RTS-RK

Numerical experiments – Convergence

Method	Heun				RK4			
q	2				4			
p	1.5	2	2.5	3	3.5	4	4.5	5
$\min\{q, p - 1/2\}$	1	1.5	2	2	3	3.5	4	4
M.S. order	1.02	1.54	2.01	2.01	3.01	3.56	4.02	4.01

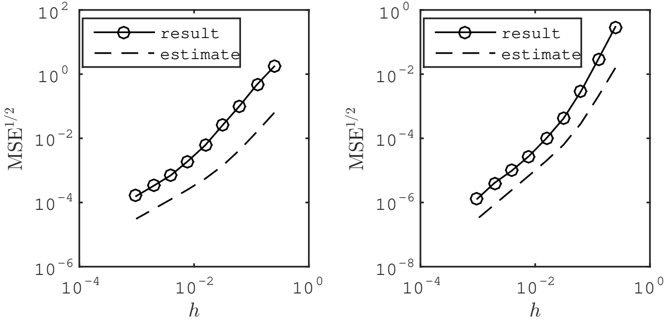
Numerical experiments – Convergence

Method	Heun				RK4			
q	2				4			
p	1.5	2	2.5	3	3.5	4	4.5	5
$\min\{q, p - 1/2\}$	1	1.5	2	2	3	3.5	4	4
M.S. order	1.02	1.54	2.01	2.01	3.01	3.56	4.02	4.01

Method	Heun			RK4			
q	2			4			
p	1	1.5	2	1.5	2	3	4
$\min\{q, 2p - 1\}$	1	2	2	2	3	4	4
Weak order ($\Phi = \|\cdot\|^{1/2}$)	0.98	2.06	2.12	1.96	3.01	3.97	4.08

Numerical experiments – Convergence

Convergence of Monte Carlo estimators, sub-optimal case.



Left: Heun method ($q = 2$) with $p = 1.5$, $M = 10^3$.
Right: RK4 method ($q = 4$) with $p = 2$, $M = 10^4$.
Recall: Convergence

$$\mathrm{MSE}(\hat{Z}) \leq C\left(h^{2\min\{2p-1,q\}} + \frac{h^{2\min\{p-1/2,q\}}}{M}\right)$$

Numerical experiments – Geometric properties

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

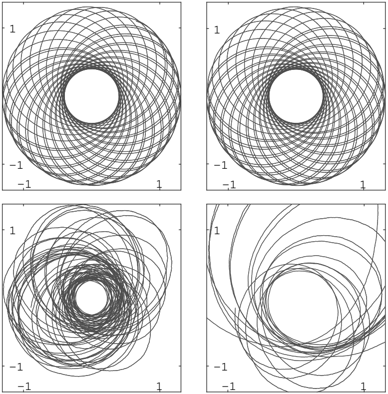
$$q_1' = p_1, \quad p_1' = -\frac{q_1}{\|q\|^3} - \frac{\delta q_1}{\|q\|^5},$$
$$q_2' = p_2, \quad p_2' = -\frac{q_2}{\|q\|^3} - \frac{\delta q_2}{\|q\|^5}.$$

The angular momentum is conserved (quadratic first integral)

$$I(p, q) = q_1 p_2 - q_2 p_1$$

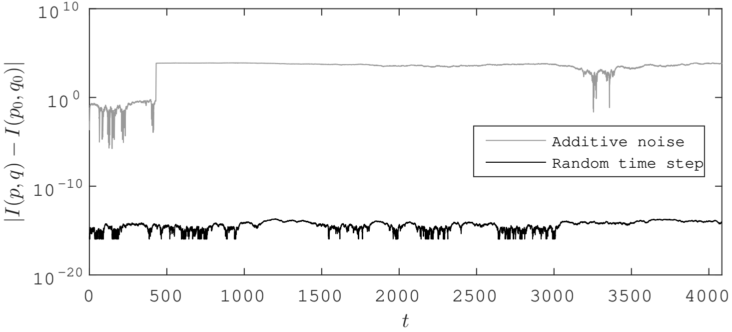
$\rightarrow$ employ a Gauss method (implicit midpoint rule).

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 harmonic oscillator, Hamiltonian system with energy

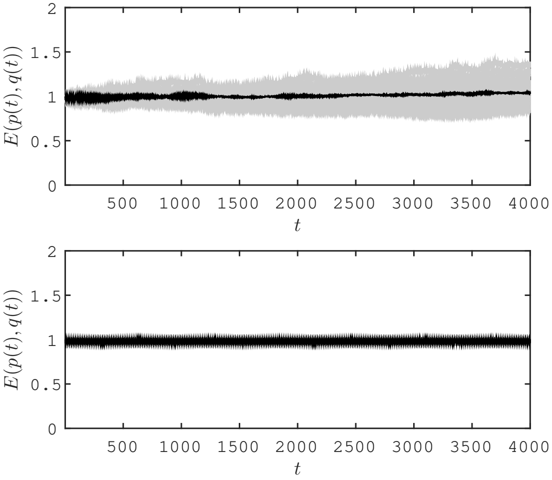
$$E(p, q) = \frac{1}{2}(p^2 + q^2).$$

Then the ODE reads for $y = (p, q)^\top$

$$y' = J^{-1}\nabla E(y), \quad y(0) = y_0,$$

Energy is separable $\rightarrow$ employ Störmer-Verlet (or symplectic Euler).

Numerical experiments – Geometric properties



Energy evolution in time for RTS-RK (top, mean energy in black) and for deterministic Störmer-Verlet (bottom).

Numerical experiments – Bayesian inverse problems

Consider again the FitzHug-Nagumo model

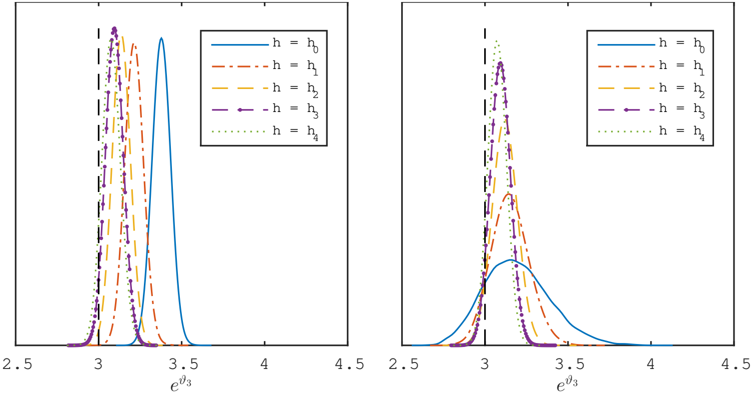
$$\begin{aligned} y_1' &= c\big(y_1 - \frac{y_1^3}{3} + y_2\big), \quad && y_1(0) = -1, \\ y_2' &= -\frac{1}{c}(y_1 - a + by_2), && y_2(0) = 1, \end{aligned}$$

with $\vartheta = (a, b, c)^\top$ unknown.

Goal

Find $\pi(\vartheta \mid \mathcal{Y})$ from observations $\mathcal{Y}$ of y (zero-mean Gaussian noise η, variance $\Sigma_\eta = (0.05)^2 I$).

Numerical experiments – Bayesian inverse problems



Marginal posterior distribution over $\vartheta_3 = c$ (truth $c^* = 3$), left explicit Euler (deterministic) right RTS-RK with explicit Euler. For both figures, $h_i = 0.1 \cdot 2^{-i}$.

Numerical experiments – Bayesian inverse problems

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

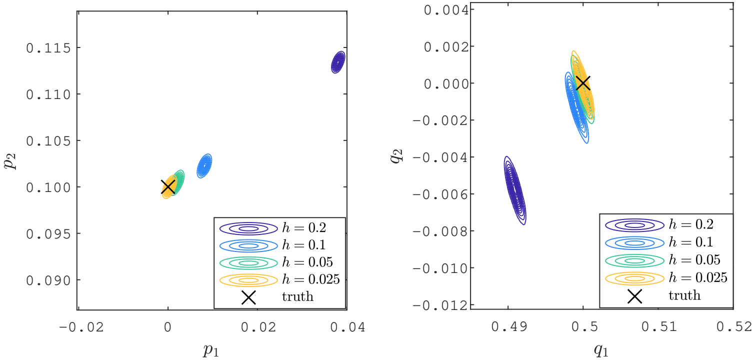
$$E(p, q) = \frac{1}{2}\|p\|^2 + \frac{1}{2}\|q\|^2 + q_1^2 q_2 - \frac{1}{3}q_2^3.$$

Chaotic problem for certain levels of energy.

Goal

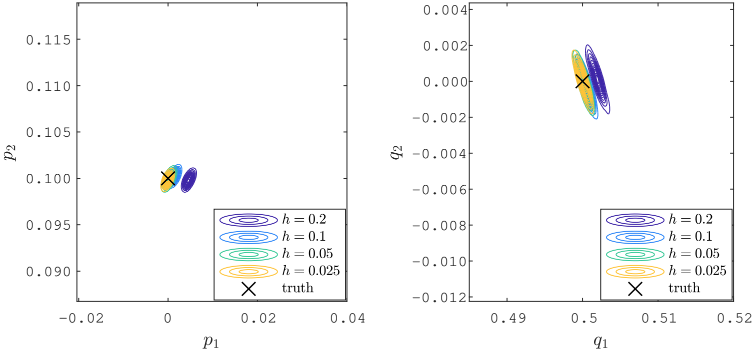
Find posterior $\pi((p_0, q_0) \mid \mathcal{Y})$ over the initial condition from a single observation of $(p(10), q(10))$

Numerical experiments – Bayesian inverse problems



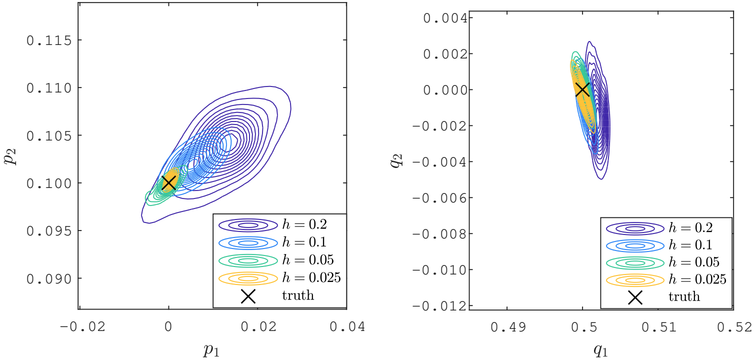
Posterior distributions given by deterministic Heun method.

Numerical experiments – Bayesian inverse problems



Posterior distributions given by deterministic Störmer-Verlet method.

Numerical experiments – Bayesian inverse problems



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

References I

[Abdulle and Garegnani, 2018] Abdulle, A. and Garegnani, G. (2018).
Random time step probabilistic methods for uncertainty quantification in chaotic and geometric numerical integration.

[Conrad et al., 2016] 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 and Hennig, 2016] 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.

References II

[Stuart, 2010] Stuart, A. M. (2010).
Inverse problems: A Bayesian perspective.
Acta Numer., 19:451–559.