Sampling invariant distributions of SPDEs

Charles-Edouard Bréhier

CNRS & Université Lyon 1, Institut Camille Jordan

Advances in Computational Statistical Physics, September 2018

Setting

Consider Stochastic Partial Differential Equations

$$\begin{cases} \frac{\partial u}{\partial t} = \Delta u + f\big(u\big) + \xi, & \text{in } (0,\infty) \times \mathcal{D} \\ u(t,\cdot) = 0, & \text{in } \partial\mathcal{D} \\ u(0,\cdot) = u_0. \end{cases}$$

ξ is (Gaussian) space-time white noise.
Domain: $\mathcal{D} = (0,1)^d$.

Example: $f(\phi) = \phi - \phi^3$ (Φ^4 models).

Setting

Consider Stochastic Partial Differential Equations

$$\begin{cases} \frac{\partial u}{\partial t} = \Delta u + f\big(u\big) + \xi, \quad \text{in } (0,\infty)\times\mathcal{D} \\ u(t,\cdot) = 0, \quad \text{in } \partial\mathcal{D} \\ u(0,\cdot) = u_0. \end{cases}$$

ξ is (Gaussian) space-time white noise.
Domain: $\mathcal{D} = (0,1)^d$.

Example: $f(\phi) = \phi - \phi^3$ (Φ^4 models).

In this talk:

- $d = 1$: well-posedness in a “classical” sense
- f is globally Lipschitz continuous.

Abstract setting

Consider $u(t) = u(t, \cdot)$ as a process with values in $H = L^2(0, 1)$. It satisfies a Stochastic Evolution Equation

$$du(t) = Au(t)dt + F(u(t))dt + dW(t), \tag{1}$$

with solutions interpreted in the mild sense

$$u(t) = e^{tA}u_0 + \int_0^t e^{(t-s)A}F(u(s))ds + \int_0^t e^{(t-s)A}dW(s).$$

Abstract setting

Consider $u(t) = u(t, \cdot)$ as a process with values in $H = L^2(0,1)$. It satisfies a Stochastic Evolution Equation

$$du(t) = Au(t)dt + F(u(t))dt + dW(t), \tag{1}$$

with solutions interpreted in the mild sense

$$u(t) = e^{tA}u_0 + \int_0^t e^{(t-s)A}F(u(s))ds + \int_0^t e^{(t-s)A}dW(s).$$

If $\mathrm{Lip}(F)$ is sufficiently small, this Markov process admits a unique invariant distribution μ, and $u(t) \underset{t\to\infty}{\to} \mu$ in distribution.

Estimation of $\int \varphi d\mu$?

Cost in terms of temporal and spatial discretization errors?

Assumptions

- Linear operator: $Ae_n = -\lambda_n e_n$, for $n \in \mathbb{N}$, with $\lambda_n = (\pi n)^2$, $e_n(z) = \sqrt{2}\sin(n\pi z)$.
  It generates an analytic semi-group, $e^{tA} = \sum_{n\in\mathbb{N}} e^{-\lambda_n t}\langle e_n, \cdot\rangle e_n$.
- Nonlinear operator: F is Lipschitz continuous and bounded.
  For ergodicity: assume $\mathrm{Lip}(F) < \lambda_1$.
- Noise: $W(t) = \sum_{n\in\mathbb{N}} \beta_n(t) e_n$ is a cylindrical Wiener process.

Assumptions

- Linear operator: $Ae_n = -\lambda_n e_n$, for $n \in \mathbb{N}$, with $\lambda_n = (\pi n)^2$, $e_n(z) = \sqrt{2}\sin(n\pi z)$.
  It generates an analytic semi-group, $e^{tA} = \sum_{n\in\mathbb{N}} e^{-\lambda_n t}\langle e_n, \cdot\rangle e_n$.
- Nonlinear operator: F is Lipschitz continuous and bounded.
  For ergodicity: assume $\mathrm{Lip}(F) < \lambda_1$.
- Noise: $W(t) = \sum_{n\in\mathbb{N}} \beta_n(t) e_n$ is a cylindrical Wiener process.

Regularity of trajectories: $\frac{1}{4} - \epsilon$ Hölder in time, $\frac{1}{2} - \epsilon$ Hölder in space. In terms of Sobolev type spaces: $\mathbb{E}|(-A)^\alpha u(t)|^2 < \infty$ if and only if $\alpha < \frac{1}{4}$.

The invariant distribution

Ergodicity: synchronous coupling

$$\frac{1}{2}\frac{d|u(t,u_0^{(2)})-u(t,u_0^{(1)})|^2}{dt} \leq \big(-\lambda_1+\mathrm{Lip}(F)\big)|u(t,u_0^{(2)})-u(t,u_0^{(1)})|^2.$$

The invariant distribution

Ergodicity: synchronous coupling

$$\frac{1}{2}\frac{d|u(t,u_0^{(2)})-u(t,u_0^{(1)})|^2}{dt} \leq \big(-\lambda_1+\mathrm{Lip}(F)\big)|u(t,u_0^{(2)})-u(t,u_0^{(1)})|^2.$$

The gradient case: $F=-DV$. Then

$$\mu(du)=\frac{1}{Z}\exp(-2V(u))\nu(du)$$

where ν is a Gaussian distribution on H:

$$\nu=\mathcal{N}\big(0,\frac{1}{2}(-A)^{-1}\big).$$

ν is the distribution of the Brownian Bridge on $(0,1)$.
Sampling conditioned diffusions, cf works by Hairer, Stuart & Voss.

Numerical discretization

In time: linear implicit Euler scheme.

$$\begin{aligned} u_{n+1} &= u_n + \Delta t A u_{n+1} + \Delta t F(u_n) + \Delta W_n \\ &= S_{\Delta t} u_n + \Delta t S_{\Delta t} F(u_n) + S_{\Delta t} \Delta W_n, \end{aligned}$$

with $S_{\Delta t} = (I - \Delta t A)^{-1}$.

In space: spectral Galerkin method (M modes).

$$du^{(M)}(t) = Au^{(M)}(t)dt + \Delta t P_M F(u^{(M)}(t)) + P_M \Delta W_n.$$

For simulations: finite differences (mesh size Δx).
Fully-discrete: ergodic Markov chain $\left(u_n^{(M)}\right)_n$, the invariant distribution is denoted by $\mu^{\Delta t, M}$.

Orders of convergence

Theorem (B. 2014,B.-Kopec 2016)

For all $\alpha < \frac{1}{2}$, $\varphi : H \to \mathbb{R}$ *of class* $\mathcal{C}_b^2$,

$$\Big| \int \varphi d\mu^{\Delta t, M} - \int \varphi d\mu \Big| \leq C_\alpha(\|\varphi\|_2)\big(\Delta t^{\alpha} + \lambda_M^{-\alpha}\big).$$

Orders of convergence

Theorem (B. 2014, B.-Kopec 2016)

For all $\alpha < \frac{1}{2}$, $\varphi : H \to \mathbb{R}$ *of class* $\mathcal{C}_b^2$,

$$\left| \int \varphi d\mu^{\Delta t, M} - \int \varphi d\mu \right| \leq C_\alpha(\|\varphi\|_2)\big(\Delta t^\alpha + \lambda_M^{-\alpha}\big).$$

- Strategy: analysis of the weak error (Kolmogorov or Poisson equation approach), cf Talay, Mattingly-Stuart-Tretyakov, etc...
- Extension to SPDEs: specific regularity properties

Orders of convergence

Theorem (B. 2014,B.-Kopec 2016)

For all $\alpha < \frac{1}{2}$, $\varphi : H \to \mathbb{R}$ *of class* $\mathcal{C}_b^2$,

$$\Big|\int \varphi d\mu^{\Delta t,M} - \int \varphi d\mu\Big| \leq C_\alpha(\|\varphi\|_2)\big(\Delta t^\alpha + \lambda_M^{-\alpha}\big).$$

- Strategy: analysis of the weak error (Kolmogorov or Poisson equation approach), cf Talay, Mattingly-Stuart-Tretyakov, etc...
- Extension to SPDEs: specific regularity properties
- If φ is only $\mathcal{C}_b^1$ or $\mathcal{C}_b^0$: weak order is degraded,

$$\Big|\int \varphi d\mu^{\Delta t,M} - \int \varphi d\mu\Big| \leq C_\alpha(\|\varphi\|_1)\big(\Delta t^{\frac{\alpha}{2}} + \lambda_M^{-\frac{\alpha}{2}}\big),$$

$$\limsup_{\Delta t\to 0, M\to\infty} \sup_{\|\varphi\|_0\leq 1} \Big|\int \varphi d\mu^{\Delta t,M} - \int \varphi d\mu\Big| \geq 1.$$

Cost of a Monte-Carlo simulation

Natural estimator of $\int \varphi d\mu$ is $\frac{1}{K}\sum_{k=1}^{K}\varphi(u_N^{(M),k})$.
To have a mean-square error of size ϵ^2, cost is of size

$$K\frac{T}{\Delta t}M \propto \epsilon^{-2}\frac{|\log(\epsilon)|}{\epsilon^{\frac{1}{\alpha}}}\epsilon^{-\frac{1}{2\alpha}} \propto \epsilon^{-5-}.$$

Cost of a Monte-Carlo simulation

Natural estimator of $\int \varphi d\mu$ is $\frac{1}{K}\sum_{k=1}^{K}\varphi(u_N^{(M),k})$.
To have a mean-square error of size ϵ^2, cost is of size

$$K\frac{T}{\Delta t}M \propto \epsilon^{-2}\frac{|\log(\epsilon)|}{\epsilon^{\frac{1}{\alpha}}}\epsilon^{-\frac{1}{2\alpha}} \propto \epsilon^{-5-}.$$

Objective: improve the order of convergence with respect to the time-step size.

Analysis with no spatial discretization.

Two ideas: **postprocessing** and **preconditioning**.

First technique: Post-processing

Postprocessed scheme [B.-Vilmart 2016]:

$$u_{n+1} = S_{\Delta t}\Big(u_n + \Delta t F\big(u_n + \frac{1}{2} S_{\Delta t}\Delta W_n\big) + \Delta W_n\Big).$$

Post-processing: $\overline{u}_N = u_N + \frac{1}{2} J_{\Delta t} \Delta W_N$,

with linear operators

$$S_{\Delta t} = (I - \Delta t A)^{-1}, \quad , \quad J_{\Delta t} = (I - \frac{\Delta t}{2} A)^{-1/2}.$$

New approximation for $\int \varphi d\mu_\infty$: $\mathbb{E}[\varphi(\overline{u}_n)] \underset{n\to\infty}{\to} \int \varphi d\overline{\mu}_\infty^{\Delta t}$.

First technique: Post-processing

Postprocessed scheme [B.-Vilmart 2016]:

$$u_{n+1} = S_{\Delta t}\Big(u_n + \Delta t F\big(u_n + \frac{1}{2} S_{\Delta t}\Delta W_n\big) + \Delta W_n\Big).$$

Post-processing: $\overline{u}_N = u_N + \frac{1}{2} J_{\Delta t}\Delta W_N$,

with linear operators

$$S_{\Delta t} = (I - \Delta t A)^{-1}, \quad , \quad J_{\Delta t} = (I - \frac{\Delta t}{2} A)^{-1/2}.$$

New approximation for $\int \varphi d\mu_\infty$: $\mathbb{E}[\varphi(\overline{u}_n)] \underset{n\to\infty}{\to} \int \varphi d\overline{\mu}_\infty^{\Delta t}$.

If $F = 0$, then $\overline{\mu}_\infty^{\Delta t} = \mu = \nu$.

Proof: $\varphi(u) = \langle u, e_p\rangle^2$,

$$\int \varphi d\mu_\infty - \int \varphi d\mu_\infty^{\Delta t} = \frac{1}{2\lambda_p}\big(1 - \frac{2}{2 + \lambda_p \Delta t}\big) = \mathbb{E}\langle \frac{1}{2} J_{\Delta t}\Delta W_N, e_p\rangle^2.$$

Numerical simulations. Conjecture: order $\frac{3}{2}$.

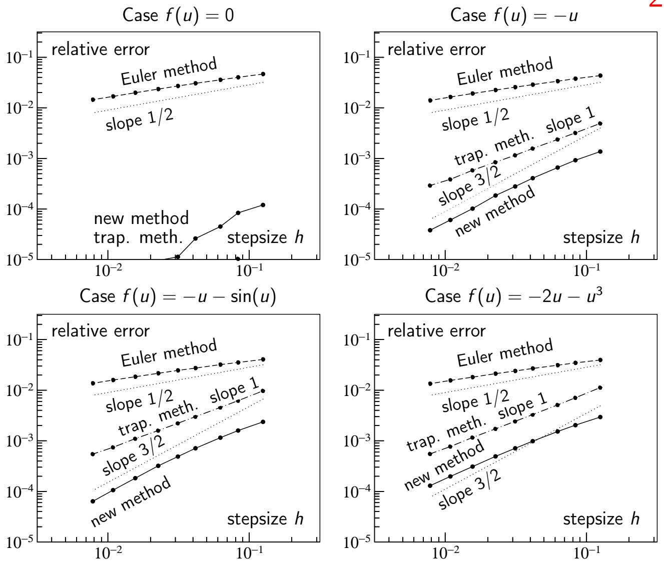


Figure: Orders of convergence, $\varphi(u) = \exp(-\|u\|^2)$.

Qualitative behavior

One realization of the standard and of the postprocessed schemes:

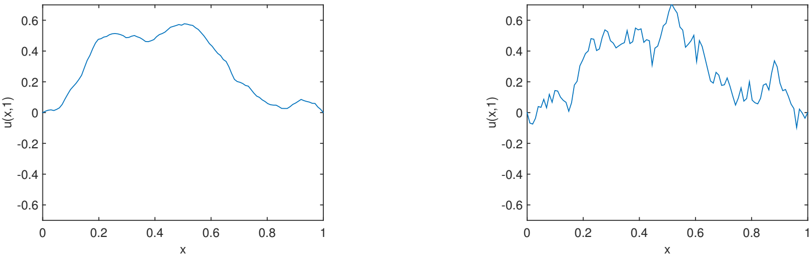


Figure: $f(u) = -u - sin(u)$, $T = 1$, $\Delta t = 0.01$, $\Delta x = 0.01$.

Qualitative behavior

One realization of the standard and of the postprocessed schemes:

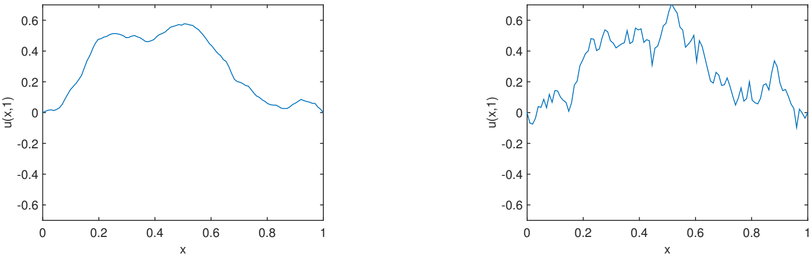


Figure: $f(u) = -u - sin(u)$, $T = 1$, $\Delta t = 0.01$, $\Delta x = 0.01$.

Noise $(I - \frac{\Delta t}{2} A)^{-\frac{1}{2}} \Delta W_n$ is rougher than $(I - \Delta A)^{-1} \Delta W_n$.

Qualitative behavior

One realization of the standard and of the postprocessed schemes:

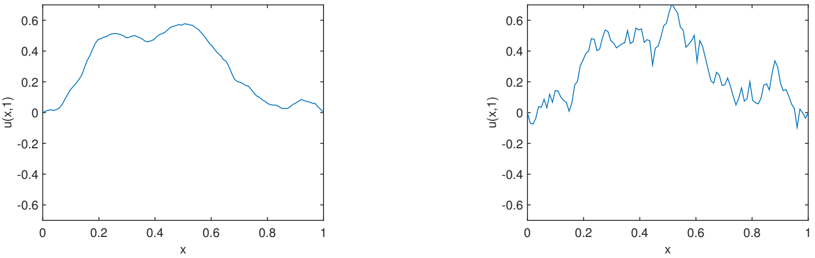


Figure: $f(u) = -u - sin(u)$, $T = 1$, $\Delta t = 0.01$, $\Delta x = 0.01$.

Noise $(I - \frac{\Delta t}{2} A)^{-\frac{1}{2}} \Delta W_n$ is rougher than $(I - \Delta A)^{-1} \Delta W_n$.
The postprocessing gives the correct regularity.

Analysis of Postprocessing for SDEs

Stochastic Differential Equation: $dX_t = f(X_t)dt + dB_t$.

Infinitesimal generator: $\mathcal{L}\phi(x) = \langle f(x), D\phi(x)\rangle + \frac{1}{2}\Delta\phi(x)$.
Weak Taylor expansion:

$$\mathbb{E}_x\big[\varphi(X(h))\big] = \big(e^{h\mathcal{L}}\varphi\big)(x) = \varphi(x) + h\mathcal{L}\varphi(x) + \frac{h^2}{2}\mathcal{L}^2\varphi(x) + \mathrm{O}(h^3)$$

One-step integrator $x \mapsto X_1$:

$$\mathbb{E}_x\big[\varphi(X_1)\big] = \varphi(x) + h\mathcal{L}\varphi(x) + h^2 A_1\varphi(x) + \mathrm{O}(h^3)$$

Postprocessing: $x \mapsto \overline{X}_1$

$$\mathbb{E}_x\big[\varphi(\overline{X}_1)\big] = \varphi(x) + h\overline{A}_1\varphi(x) + \mathrm{O}(h^2)$$

Error analysis

- Integrator $X_k = \Phi^h(X_{k-1}) = (\Phi^h)^k(X_0)$; compute $X_1, \ldots, X_n$.
- Postprocessing: $\overline{X}_n = \overline{\Phi}^h(X_n)$.
- Convergence to invariant distributions: $X_n \underset{n\to\infty}{\to} \mu_\infty^h$ and $\overline{X}_n \underset{n\to\infty}{\to} \overline{\mu}_\infty^h$.

Error analysis

- Integrator $X_k = \Phi^h(X_{k-1}) = (\Phi^h)^k(X_0)$; compute $X_1, \ldots, X_n$.
- Postprocessing: $\overline{X}_n = \overline{\Phi}^h(X_n)$.
- Convergence to invariant distributions: $X_n \underset{n\to\infty}{\to} \mu_\infty^h$ and $\overline{X}_n \underset{n\to\infty}{\to} \overline{\mu}_\infty^h$.

Results

- Weak order is equal to 1: $\int \varphi d\mu_\infty^h - \int \varphi d\mu_\infty = \mathrm{O}(h)$
- Leading order term [Talay-Tubaro 1990]:

$$\int \varphi d\mu_\infty^h - \int \varphi d\mu_\infty = C(\varphi)h + \mathrm{O}(h^2)$$

- With a postprocessing [Vilmart 2015]:

$$\int \varphi d\overline{\mu}_\infty^h - \int \varphi d\mu_\infty = \overline{C(\varphi)}h + \mathrm{O}(h^2)$$

Error analysis

Results:

- Weak order 1: $\int \varphi d\mu_\infty^h - \int \varphi d\mu_\infty = \mathrm{O}(h)$
- Leading order: $\int \varphi d\mu_\infty^h - \int \varphi d\mu_\infty = C(\varphi)h + \mathrm{O}(h^2)$
- Postprocessing: $\int \varphi d\overline{\mu}_\infty^h - \int \varphi d\mu_\infty = \overline{C(\varphi)}h + \mathrm{O}(h^2)$

with the expressions

$$C(\varphi) = -\int A_1 \Psi d\mu_\infty \quad , \quad \overline{C(\varphi)} = -\int \big(A_1 + \big[\mathcal{L}, \overline{A}_1\big]\big)\Psi d\mu_\infty,$$

where Ψ is solution of the Poisson equation $\mathcal{L}\Psi = \varphi - \int \varphi d\mu_\infty$:

$$\Psi(x) = -\int_0^\infty \mathbb{E}_x\big[\varphi(X(t)) - \int \varphi d\mu_\infty\big] dt$$

Order conditions

$$C(\varphi) = -\int A_1 \Psi d\mu_\infty \quad , \quad \overline{C(\varphi)} = -\int \left(A_1 + \left[\mathcal{L}, \overline{A}_1\right]\right) \Psi d\mu_\infty,$$

Order conditions

- $\int \varphi d\mu_\infty^h - \int \varphi d\mu_\infty = \mathrm{O}(h^2)$ for all φ iff $A_1^\star \mu_\infty = 0$.
- $\int \varphi d\overline{\mu}_\infty^h - \int \varphi d\mu_\infty = \mathrm{O}(h^2)$ for all φ iff $\left(A_1 + \left[\mathcal{L}, \overline{A}_1\right]\right)^\star \mu_\infty = 0$.

These conditions are weaker than requiring order 2, *i.e.* $A_1 = \frac{\mathcal{L}^2}{2}$.

Examples: gradient SDEs

Gradient SDE: $dX_t = -\nabla V(X_t)dt + dB_t$:

$$X_{n+1} = X_n - h\nabla V(X_n + \frac{1}{2}\sqrt{h}\xi_n) + \sqrt{h}\xi_n \quad , \quad \overline{X}_n = X_n + \frac{1}{2}\sqrt{h}\xi_n.$$

Introduced in [Leimkuhler-Matthews 13], in the *non-Markovian* form $\overline{X}_{n+1} = \overline{X}_n - h\nabla V(\overline{X}_n) + \frac{\sqrt{h}}{2}(\xi_n + \xi_{n+1})$.

Examples: gradient SDEs

Gradient SDE: $dX_t = -\nabla V(X_t)dt + dB_t$:

$$X_{n+1} = X_n - h\nabla V(X_n + \frac{1}{2}\sqrt{h}\xi_n) + \sqrt{h}\xi_n \quad , \quad \overline{X}_n = X_n + \frac{1}{2}\sqrt{h}\xi_n.$$

Introduced in [Leimkuhler-Matthews 13], in the *non-Markovian* form $\overline{X}_{n+1} = \overline{X}_n - h\nabla V(\overline{X}_n) + \frac{\sqrt{h}}{2}(\xi_n + \xi_{n+1})$.

Construction of an IMEX scheme for gradient SPDEs:

$$u_{n+1} = S_{\Delta t}\Big(u_n - \Delta t\nabla V\big(u_n + \frac{1}{2}S_{\Delta t}\Delta W_n\big) + S_{\Delta t}\Delta W_n\Big),$$
$$\overline{u}_N = u_N + \frac{1}{2}J_{\Delta t}\Delta W_N,$$

Proved results:

- Order 2 for a SDE $dX_t = AX_t dt - \nabla V(X_t)dt + dB_t$.
- Well-posed, ergodic for SPDEs.

Second technique: Preconditioning

Gradient SPDE:

$$du(t) = Au(t)dt - DV(u(t))dt + dW(t).$$

Second technique: Preconditioning

Gradient SPDE:

$$du(t) = Au(t)dt - DV(u(t))dt + dW(t).$$

Preconditioned SPDE:

$$dv(t) = -v(t)dt + (-A)^{-1}DV(v(t))dt + (-A)^{-\frac{1}{2}}dW(t).$$

Key observation

The unique invariant distribution of $\big(v(t)\big)_{t\geq 0}$ is equal to μ_∞.

Improvements:

- A is replaced with the bounded operator $-I$,
- noise is trace-class, $\mathrm{Tr}(A) < \infty$.
- Trajectories of v are almost $\frac{1}{2}$-Hölder continuous in time.

Ã©

Construction of numerical schemes

Strategy

SDE integrators (and postprocessing) may be applied to the preconditioned SPDE to sample the invariant distribution μ_∞.

Construction of numerical schemes

Strategy

SDE integrators (and postprocessing) may be applied to the preconditioned SPDE to sample the invariant distribution μ_∞.

Example: **preconditioned Crank-Nicolson (pCN)**

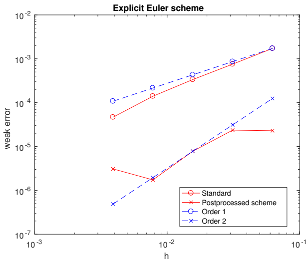
$$v_{n+1} = \frac{1-\frac{\Delta t}{2}}{1+\frac{\Delta t}{2}} v_n + \frac{\Delta t}{1+\frac{\Delta t}{2}} (-A)^{-1} F(v_n) + \frac{1}{1+\frac{\Delta t}{2}} (-A)^{-\frac{1}{2}} \Delta W_n.$$

With $F = 0$, this is the only known proposal kernel for MCMC simulation in infinite dimension, cf Stuart&al..

Three schemes of order 2 – ongoing work

Explicit Euler scheme and postprocessing

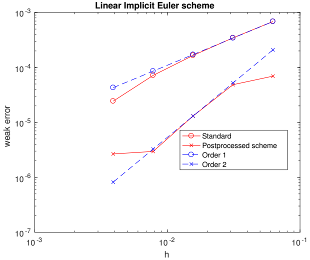
$$\begin{cases} v_{n+1} = v_n - \Delta t\big(v_n + \frac{1}{2}(-A)^{-\frac{1}{2}}\Delta W_n\big) + \Delta t(-A)^{-1}F\big(v_n + \frac{1}{2}(-A)^{-\frac{1}{2}}\Delta W_n\big) \\ \qquad\qquad + (-A)^{-\frac{1}{2}}\Delta W_n, \\ \overline{v}_N = v_N + \frac{1}{2}(-A)^{-\frac{1}{2}}\Delta W_N. \end{cases}$$



Three schemes of order 2 – ongoing work

Linear-Implicit Euler scheme and postprocessing

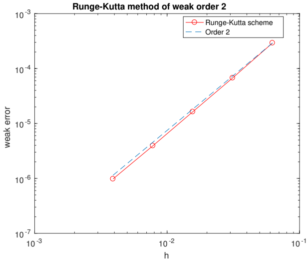
$$\begin{cases} v_{n+1} = \frac{1}{1+\Delta t} v_n + \frac{\Delta t}{1+\Delta t}(-A)^{-1} v\big(v_n + \frac{1}{2(1+\Delta t)}(-A)^{-\frac{1}{2}}\Delta W_n\big) \\ \qquad\qquad + \frac{1}{1+\Delta t}(-A)^{-\frac{1}{2}}\Delta W_n \\ \overline{v}_N = v_N + \frac{1}{2(1+\frac{\Delta t}{2})^{\frac{1}{2}}}(-A)^{-\frac{1}{2}}\Delta W_N. \end{cases}$$



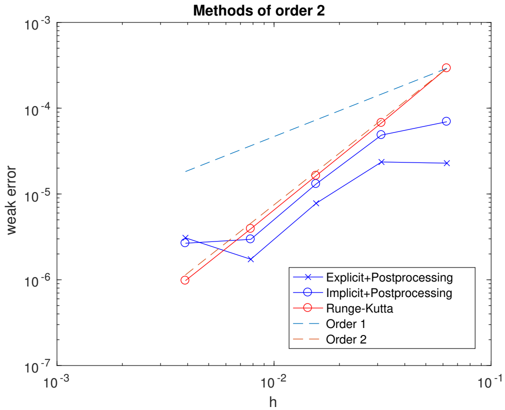
Three schemes of order 2 – ongoing work

Runge-Kutta method with weak order 2

$$\begin{cases} K_n = (1-\Delta t)v_n + \Delta t(-A)^{-1}F(v_n) + (-A)^{-\frac{1}{2}}\Delta W_n, \\ v_{n+1} = v_n + \frac{\Delta t}{2}\left(-v_n - K_n + (-A)^{-1}F(v_n) + (-A)^{-1}F(K_n)\right) + (-A)^{-\frac{1}{2}}\Delta W_n. \end{cases}$$



The three methods of order 2



Data of the experiment: $\Delta x = 0.02$, $F(u) = -u - 5\cos(u) - 2\sin(2u)$, $\varphi(u) = \int_0^1 e^{-5u(x)} dx$.

Multilevel Monte-Carlo is used to reduce the variance (10^7 realizations per level).

Conclusion

For **gradient** SPDEs, methods of order 2 for the invariant distribution can be built using preconditioning and postprocessing techniques.

Conclusion

For **gradient** SPDEs, methods of order 2 for the invariant distribution can be built using preconditioning and postprocessing techniques.

Improvement of the Monte-Carlo cost: with methods of order 2,

$$\epsilon^{-2}\frac{T}{\Delta t}M = \epsilon^{-3-\frac{1}{2}-}.$$

Conclusion

For **gradient** SPDEs, methods of order 2 for the invariant distribution can be built using preconditioning and postprocessing techniques.

Improvement of the Monte-Carlo cost: with methods of order 2,

$$\epsilon^{-2}\frac{T}{\Delta t}M = \epsilon^{-3-\frac{1}{2}-}.$$

Ongoing work and perspectives:

- theoretical analysis
- non-globally Lipschitz drift F?
- improve results for non-smooth test functions?

Thanks for your attention.