Algebraic Structure of Stochastic Expansions and Efficient Simulation

Anke Wiese

Heriot–Watt University, Edinburgh, UK

SSSC2012: ICMAT, Madrid
November 5 - 9, 2012

Supported by EMS, IMA, LMS

Work in collaboration with:
Simon J.A. Malham (Proc. Roy. Soc. A, 2009),
Kurusch Ebrahimi–Fard, Alexander Lundervold, Simon J.A. Malham, Hans Munthe-Kaas (Proc. Roy. Soc. A, 2012)

Set-up

Suppose solution y of stochastic differential equation in $\mathbb{R}^N$ given by

$$y_t = y_0 + \int_0^t V_0(y_\tau)\,\mathrm{d}\tau + \sum_{i=1}^{d} \int_0^t V_i(y_\tau)\,\mathrm{d}W_i(\tau)$$

where

- $(W_1, \ldots W_d)$ standard Wiener process,
- smooth vector fields $V_i = \sum_{j=1}^N V_i^j \partial_{y_j}$ on $\mathbb{R}^N$, $i = 0, \ldots, d$,
- initial condition $y_0 \in \mathbb{R}^N$.

For notational convenience, $W_0(\tau) = \tau$.

Accuracy: Problem

Background: Solutions only in exceptional cases given explicitly. Typically, approximations are based on the stochastic Taylor expansion, truncated to include the necessary terms to achieve the desired order (e.g. Euler scheme, Milstein scheme).

Question: Are there other series such that any truncation generates an approximation that is always at least as accurate as the corresponding truncated stochastic Taylor series, independent of the vector fields and to all orders?

We will call such approximations *efficient integrators*.

Itô's Lemma for Stratonovich Integrals

Itô's Lemma: *Suppose f smooth. Then*

$$f(y_t) = f(y_0) + + \sum_{i=0}^{d} \int_0^t \underbrace{\sum_{j=1}^{N} V_i^j(y_\tau) \partial_{y_j} f(y_\tau)}_{= V_i \circ f \circ y_\tau} \, \mathrm{d}W_i(\tau)$$

$$= f(y_0) + \sum_{i=0}^{d} \int_0^t V_i \circ f \circ y_\tau \, \mathrm{d}W_i(\tau) \, ,$$

where

$$V_i \circ f = \sum_{j=1}^{N} V_i^j \partial_{y_j} f \, .$$

Stochastic Taylor Series (Stratonovich Form)

$$y_t = y_0 + \sum_{i=0}^{d} \int_0^t \mathrm{d}W_i(\tau) V_i \circ y_0 + \sum_{i,j=0}^{d} \underbrace{\int_0^t W_i(\tau)\,\mathrm{d}W_j(\tau)}_{=J_{ij}(t)} V_i \circ V_j \circ y_0$$

$$+ \cdots = \sum_w J_w(t)\, V_w \circ y_0,$$

where $w = (a_1, \ldots, a_n)$ word from alphabet $\{0, 1, \ldots, d\}$ (including the empty word),

$$V_w = V_{a_1} \circ V_{a_2} \circ \ldots \circ V_{a_n},$$

and

$$V_i \circ V_j = \sum_{k=1}^{N} V_i^k \partial_{y_k} V_j.$$

Stochastic Taylor flow (Lyons' signature)

$$\varphi_t = \sum_w J_w V_w.$$

Principal Idea

- Suppose f is invertible function with real series expansion

$$f(x) = \sum_{k=0}^{\infty} c_k\, x^k \qquad (c_k \text{ constant}).$$

Construct new series

$$\psi_t = f(\varphi_t) \equiv \sum_{k=0}^{\infty} c_k\, \varphi_t^k.$$

- Truncate to finite series $\hat{\psi}_t$.
- New approximation to flow map and to solution y_t

$$\hat{\varphi}_t = f^{-1}(\hat{\psi}_t), \quad \hat{y}_t = \hat{\varphi}_t \circ y_0.$$

- Mean-square error in this approximation

$$\|(\varphi_t - \hat{\varphi}_t) \circ y_0\|_{L^2}^2.$$

Aim: Find f, equivalently $(c_k)_k$, such that error is smaller than error for corresponding truncated Taylor series expansion.

Exponential Lie Series

Example: $f =$ logarithmic function:

$$\varphi_t = \exp(\psi_t) \ ,$$

where

$$\psi_t = \ln \varphi_t = \sum_{i=0}^{d} J_i(t) V_i + \sum_{i<j} \tfrac{1}{2}(J_{ij} - J_{ji})(t)[V_i, V_j] + \cdots \ .$$

exponential Lie series, Chen-Strichartz formula

(Magnus 1954, Chen 1957, Strichartz 1987 ...)

Accuracy: Exponential Lie Series

Theorem:
A (modified) exponential Lie series of order $\frac{1}{2}$, order 1 and order $1\frac{1}{2}$ is more accurate in the mean-square sense than the corresponding truncated Taylor series, **if** *the diffusion vector fields $V_1, \ldots, V_d$ commute. They do not need to commute with the drift vector field V_0.*

(Castell & Gaines 1995, Malham & W. 2008)

Remark: Counterexample for non-commuting vector fields (Lord, Malham & W. 2008).

Sinhlog Series

Theorem:
The solution series constructed by taking the hyperbolic sine of the logarithm of the stochastic Taylor flow-map,

$$\psi_t = \sinh\log(\varphi_t),$$

when truncated, generates an approximation,

$$\hat{\varphi}_t = \exp\sinh^{-1}(\hat{\psi}_t) \equiv \hat{\psi}_t + \sqrt{\mathrm{id} + \hat{\psi}_t^2}\,,$$

that is for all vector fields always at least as accurate as the corresponding truncated stochastic Taylor series expansion in the mean-square sense. (Malham & W. 2009)

Principal Idea

- Stochastic Taylor flow

$$\varphi_t = \sum_w J_w V_w \,.$$

$$f(\varphi_t) = \sum_{k=0}^{\infty} c_k \, \varphi_t^k = \sum_w F(w) V_w$$

(after rearrangement).

- Composition of vector fields is concatenation product,

$$V_{w_1} \circ V_{w_2} = V_{w_1 w_2}.$$

Itó's Product Formula: Multiplication of multiple Wiener integrals is a shuffle product. e.g.

$$J_{12} J_{34} = J_{1234} + J_{1324} + J_{3124} + J_{1342} + J_{3142} + J_{3412}.$$

Shuffle product $J_u J_v \mapsto u \sqcup\!\sqcup\, v$.

Hopf Algebra Structure - Recoding the Signature

- Representation of stochastic Taylor flow as

$$\varphi_t = \sum_w w \otimes w$$

in Hopf algebra given by tensor product of (algebras of) real series of words with shuffle product on left and concatenation product on right.

- Given $f(x) = \sum_{k=0}^{\infty} c_k x^k$,

$$f(\varphi_t) = \sum_{k \geq 0} c_k \left(\sum_w w \otimes w \right)^k$$
$$= \sum_w \underbrace{\left(\sum_{k=0}^{\infty} c_k \sum_{w=u_1 \dots u_k} u_1 \sqcup\!\sqcup \dots \sqcup\!\sqcup u_k \right)}_{=F(w)} \otimes w.$$

Convolution Shuffle Algebra

- Convolution algebra of linear endomorphisms of Hopf algebra (with shuffle product)

$$K_1 \star K_2(w) = \sum_{vv'=w} K_1(v) \sqcup\!\sqcup K_2(v').$$

- Representation

$$f(\varphi_t) = \sum_{w} F(w) \otimes w,$$

where F is the linear endomorphism

$$F = \sum_{k} c_k \mathrm{id}^{\star k} =: f^{\star}(\mathrm{id}).$$

Important: Action of f encoded by linear endomorphism $f^{\star}(\mathrm{id})$.

Generalization

- Construction through expansion about $\epsilon \in \mathbb{R}$

$$f^{\star}(X) = \sum_{k=0}^{\infty} (X - \epsilon\,\nu)^{\star k},$$

where ν is the unit in the convolution shuffle algebra (ν sends non-empty words to 0 and the empty word to itself).

- For an endomorphism X define the sinhlog endomorphism by

$$\text{sinhlog}^{\star}(X) = \frac{1}{2}\big(X - X^{\star(-1)}\big).$$

Thus

$$\text{sinhlog}^{\star}(\mathrm{id}) = \frac{1}{2}\big(\mathrm{id} - \mathrm{id}^{\star(-1)}\big) = \frac{1}{2}\big(\mathrm{id} - S\big),$$

where

$$S \circ (a_1 \dots a_n) = (-1)^n a_n \dots a_1$$

is the *antipode* (signed reverse).

Sinhlog Endomorphism

- Series representation

$$\text{sinhlog}^{\star}(\mathrm{id}) = J - \tfrac{1}{2}J^{\star 2} + \tfrac{1}{2}J^{\star 3} - \cdots + (-1)^{k+1}\tfrac{1}{2}J^{\star k} + \cdots ,$$

where $J := \nu - \mathrm{id}$ (*augmented ideal projector*).

- The compositional inverse of sinhlog* is given by

$$\text{sinhlog}^{-1}(X) = X + \left(X^{\star 2} + \nu\right)^{\star(1/2)}.$$

- Similar, the coshlog endomorphism is

$$\text{coshlog}^{\star}(X) = \frac{1}{2}\left(X + X^{\star(-1)}\right).$$

Sinhlog Series

On computational interval $[t_n, t_{n+1}]$:

$$\text{sinhlog}(\varphi_t) = \sum_{i=0}^{d} J_i(t_n, t_{n+1}) V_i + \sum_{i,j=0}^{d} \tfrac{1}{2}\big(J_{ij} - J_{ji}\big)(t_n, t_{n+1}) V_{ij}$$

$$+ \sum_{i,k,k=0}^{d} \tfrac{1}{2}\big(J_{ijk} + J_{kji}\big)(t_n, t_{n+1}) V_{ijk} + \ldots$$

Inner Product of Endomorphism

- Measure accuracy by mean-square error: for stochastic processes $X_t = \sum_u X(u) V_u(y_0)$ and $Y_t = \sum_v Y(v) V_v(y_0)$

$$\langle X_t, Y_t \rangle = E\left[X_t^\dagger Y_t\right] = \sum_{u,v} V_u^\dagger(y_0) E\left[X(u)Y(v)\right] V_v(y_0),$$

  where $\dagger$ denotes matrix transpose.
- Corresponding inner product for endomorphisms X and Y:
  - Write $X(w) = \sum_u \mathsf{X}_{w,u} u$ with real-valued coefficients $X_{w,u}$. Define matrix
$$\mathsf{X} = \left(X_{w,u}\right)_{w,u}.$$
  - Define matrix $\mathsf{W} = \left(W_{w,u}\right)_{w,u}$, where $W_{w,u} = E\left[w \sqcup\!\sqcup u\right]$.
  - Define matrix $\mathsf{V} = \left(V_{w,u}\right)_{w,u}$, where $V_{w,u} = V_w^\dagger V_u$ (inner product).

$$\langle X, Y \rangle := \mathrm{tr}\left(\mathsf{X}\,\mathsf{W}\,\mathsf{Y}^\dagger\,\mathsf{V}^\dagger\right).$$

- Note that the definition of the inner product depends on the V; all our subsequent results hold independently of V.

Orthogonality of Sinhlog and Coshlog

- $\mathrm{id} = \mathsf{sinhlog}^\star(\mathrm{id}) + \mathsf{coshlog}^\star(\mathrm{id})$
- For any n, consider the restriction of the following endomorphisms on the subspace $\mathbb{S}_n$ generated by words of length n. Then
  1. $\big\langle |S|, |S| \big\rangle = \langle S, S \rangle = \langle \mathrm{id}, \mathrm{id} \rangle$;
  2. $\langle \mathsf{sinhlog}^\star(\mathrm{id}), \mathsf{coshlog}^\star(\mathrm{id}) \rangle = 0$;
  3. $\|\mathrm{id}\|^2 = \|\mathsf{sinhlog}^\star(\mathrm{id})\|^2 + \|\mathsf{coshlog}^\star(\mathrm{id})\|^2$.
- Proof is based on the following identity (see Malham & W. 2009): For any pair u, v, we have $E[J_u J_v] = E(u \sqcup\!\sqcup\, v) \equiv E\big((|S| \circ u) \sqcup\!\sqcup\, (|S| \circ v)\big) = E[J_{|S| \circ u} J_{|S| \circ v}]$.

  - This implies property 1.
  - Property 2 follows using the bilinearity of the inner product:

$$\begin{aligned}\langle \mathsf{sinhlog}^\star(\mathrm{id}), \mathsf{coshlog}^\star(\mathrm{id}) \rangle &= \langle \frac{1}{2}(\mathrm{id} - S), \frac{1}{2}(\mathrm{id} + S) \rangle \\ &= \frac{1}{4}\big(\langle \mathrm{id}, \mathrm{id} \rangle - \langle S, S \rangle\big) = 0.\end{aligned}$$

  - Property 3 now immediate.

Construction of Stochastic Approximations

- Function f with formal power series given;
- corresponding endomorphism $f^{\star}(\mathrm{id})$;
- truncate to include only words up to length n, denote this

$$\Pi_{\mathbb{S}_{\leq n}} \circ f^{\star}(\mathrm{id});$$

- Apply compositional inverse

$$f^{-1} \circ \pi_{\mathbb{S}_{\leq n}} \circ f^{\star}(\mathrm{id});$$

- error in approximation

$$\mathrm{id} - f^{-1} \circ \pi_{\mathbb{S}_{\leq n}} \circ f^{\star}(\mathrm{id});$$

- measure error in norm defined on endomorphisms.
- compare with error in corresponding truncated Taylor approximation

$$\mathrm{id} - \mathrm{id} \circ \pi_{\mathbb{S}_{\leq n}} \circ \mathrm{id};$$

Efficiency

A numerical approximation to the solution of an SDE is an *efficient integrator* if it generates a strong numerical integration scheme that is more accurate in the mean square sense than the corresponding truncated stochastic Taylor integration scheme of the same order (according to word length), independent of the governing vector fields and to all orders.

Efficiency of Sinhlog Integrator

Theorem:

The integrator based on the sinhlog endomorphism is efficient.

Idea of Proof:

- Define

$$P := \pi_{\mathbb{S}_{\leq n}} \circ \mathsf{sinhlog}^{\star}(\mathrm{id})$$
$$Q := \pi_{\mathbb{S}_{\geq n+1}} \circ \mathsf{sinhlog}^{\star}(\mathrm{id}),$$

  where $\pi_{\mathbb{S}_{\geq n+1}}$ is projector on space generated by words of length $n+1$ or more.

- Error in approximation

$$\mathsf{sinhlog}^{-1} \circ (P+Q) - \mathsf{sinhlog}^{-1} \circ P = Q + \tfrac{1}{2}\big(P \star Q + Q \star P\big) + \mathcal{O}\big(Q^{\star 2}\big).$$

  Leading order term $Q \circ \pi_{\mathbb{S}_{n+1}}$.

- Leading order term in truncated Taylor series is $\mathrm{id} \circ \pi_{\mathbb{S}_{n+1}}$, and

$$\begin{aligned}\|\mathrm{id} \circ \pi_{\mathbb{S}_{n+1}}\| &= \|Q \circ \pi_{\mathbb{S}_{n+1}}\| + \|\mathsf{coshlog}^{\star}(\mathrm{id}) \circ \pi_{\mathbb{S}_{n+1}}\| \\ &> \|Q \circ \pi_{\mathbb{S}_{n+1}}\|.\end{aligned}$$

New Class of Integrators

Define the class of endomorphisms

$$f^{\star}(X;\epsilon)=\tfrac{1}{2}(X-\epsilon X^{\star(-1)}).$$

for any $\epsilon\in\mathbb{R}$.

- $f^{\star}(\mathrm{id},+1)$ is the sinhlog endomorphism;
- $f^{\star}(\mathrm{id},-1)$ is the coshlog endomorphism.

Compositional inverse of $f^{\star}(X;\epsilon)$ is

$$f^{-1}(X;\epsilon)=X+\left(X^{\star 2}+\epsilon\,\nu\right)^{\star(1/2)}.$$

Optimality of Sinhlog Integrator

Theorem:
For every $\epsilon > 0$ *the class of integrators* $f^\star(\mathrm{id};\epsilon)$ *is efficient. When* $\epsilon = 1$, *the error of the integrator* $f^\star(\mathrm{id};1)$ *realizes its smallest possible value compared to the error of the corresponding stochastic Taylor integrator.*

Coshlog Integrator

The coshlog endomorphism does *not* generate an efficient integrator.
The compositional inverse of the coshlog endomorphism is

$$\text{coshlog}^{-1}(X) = X + \left(X^{\star 2} - \nu\right)^{\star \frac{1}{2}}.$$

Define as before

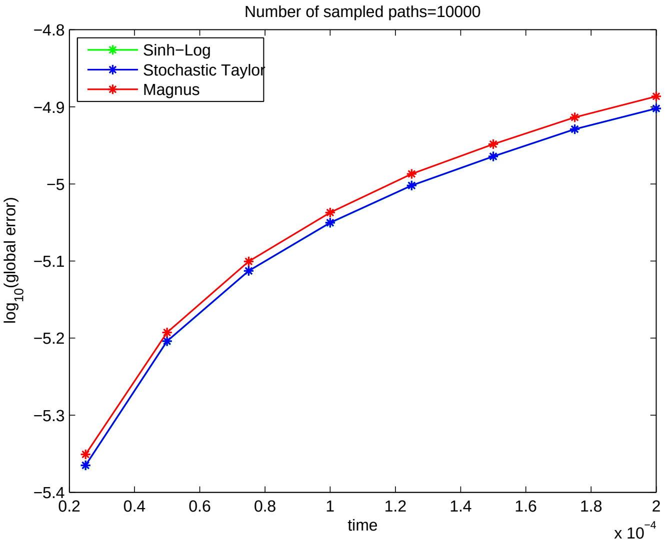
$$P := \pi_{\mathbb{S}_{\leq n}} \circ \text{coshlog}^{\star}(\mathrm{id})$$
$$Q := \pi_{\mathbb{S}_{\geq n+1}} \circ \text{coshlog}^{\star}(\mathrm{id})$$

Error in approximation

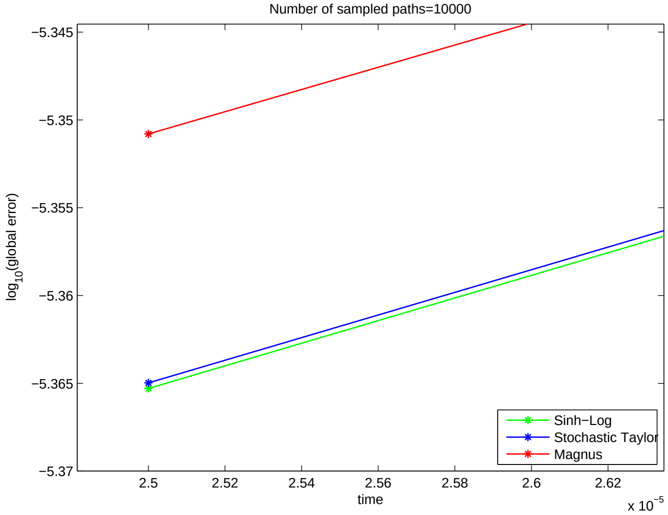
$$\text{coshlog}^{-1} \circ (P + Q) - \text{coshlog}^{-1} \circ P = Q + (J^{\star(-1)} \circ \pi_{\mathbb{S}_{\leq n}}) \star Q + \ldots.$$

$J^{\star(-1)}$ exists formally, but introduces terms of same order as remained in integrator: thus there is a from of order reduction.

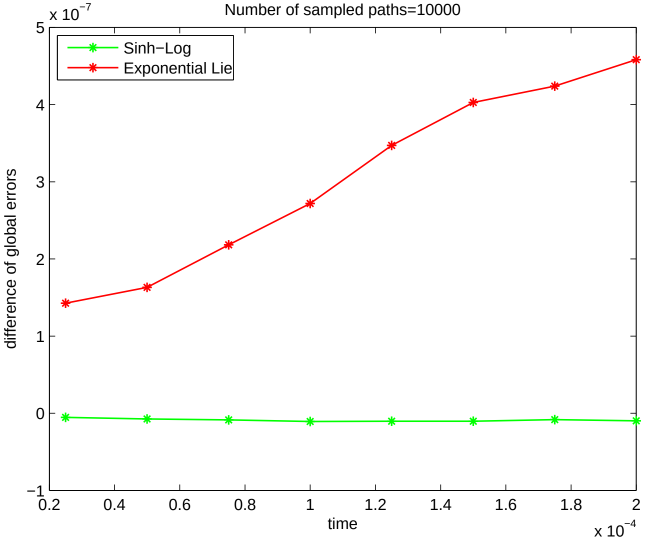
Example 1: Linear 2×2 System



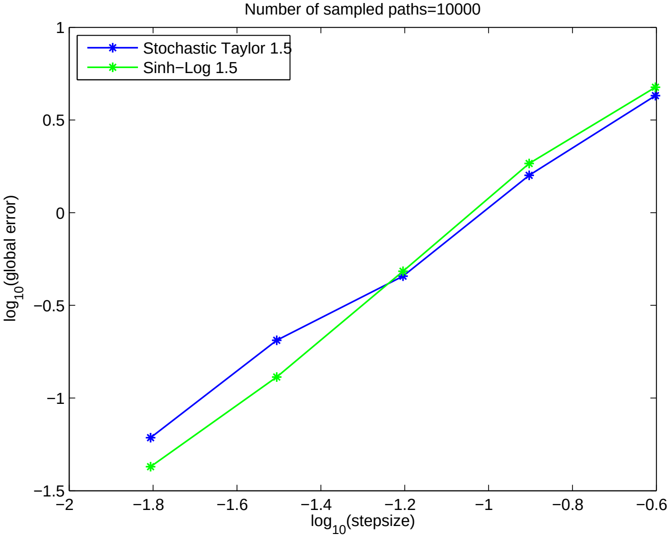
Example 1: Error



Example 1: Difference between Error and Stochastic Taylor Error



Example 2: Global Error



Conclusion

- Natural connection between stochastic approximations and convolution algebra of endomorphisms on Hopf shuffle algebra of words.
- Algebraic abstraction enables new class of efficient integrators.
- Within this class, the sinhlog integrator is optimal, that is the error of the integrator realizes its smallest value when compared with the error of the corresponding truncated stochastic Taylor integrator in the mean-square sense.
- Generalization to SDEs driven by Lévy processes and semimartingale algebra.

Future Directions

- Algebraic structure of BSDEs?
- Algebraic structure of SPDEs?

References

- Castell, F., Gaines, J. 1995 An efficient approximation method for stochastic differential equations by means of the exponential Lie series. *Math. Comp. Simulation* **38**, 13–19.
- Chen, K. T. 1957 Integration of paths, geometric invariants and a generalized Baker–Hausdorff formula. *Annals of Mathematics* **65**(1), 163–178.
- Curry, C., Ebrahimi-Fard, K., Malham, S.J.A. and Wiese, A. 2012 The algebra of semimartingales and efficient simulation of stochastic systems driven by Lévy processes.
- Ebrahimi-Fard, K., Lundervold, A., Malham, S.J.A., Munthe–Kaas, H. and Wiese, A. 2012 Algebraic structure of stochastic expansions and efficient simulation, *Proc. R. Soc. A*.
- Lord, G., Malham, S.J.A. & Wiese, A. 2008 Efficient strong integrators for linear stochastic systems, *SIAM J. Numer. Anal.* **46**(6), 2892–2919.
- Malham, S.J.A., Wiese, A. 2009 Stochastic expansions and Hopf algebras. *Proc. R. Soc. A* **465**, 3729–3749.

References

Magnus, W. 1954 On the exponential solution of differential equations for a linear operator. *Comm. Pure Appl. Math.* **7**, 649–673.

Reutenauer, C. 1993 *Free Lie algebras*, London Mathematical Society Monographs New Series 7, Oxford Science Publications