The Hopf algebraic structure of stochastic expansions and efficient simulation

Kurusch Ebrahimi–Fard, Alexander Lundervold,
Simon J.A. Malham, Hans Munthe–Kaas and Anke Wiese

York: 15th October 2012

Introduction

SDE for $y_t \in \mathbb{R}^N$ in Stratonovich form

$$y_t = y_0 + \sum_{i=0}^{d} \int_0^t V_i(y_\tau)\, \mathrm{d}W_\tau^i.$$

- Driven by a d-dimensional Wiener process $(W^1, \ldots, W^d)$;
- Governed by drift V_0 and diffusion vector fields $V_1, \ldots, V_d$;
- Use convention $W_t^0 \equiv t$;
- Assume $t \in \mathbb{R}_+$ lies in the interval of existence.

Focus on solution series and strong simulation.

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*.

Flowmap

For the system of equations

$$\frac{\mathrm{d}y}{\mathrm{d}t} = V(y)$$

we can define the flowmap $\varphi_t \colon y_0 \mapsto y_t$. Chain rule $\implies$

$$\frac{\mathrm{d}}{\mathrm{d}t} f(y) = \underbrace{V(y) \cdot \nabla_y f(y)}_{V \circ f \circ y}$$

$$\begin{aligned}
\Leftrightarrow \quad & \frac{\mathrm{d}}{\mathrm{d}t}(f \circ \varphi_t \circ y_0) = V \circ f \circ \varphi_t \circ y_0 \\
\Leftrightarrow \quad & \frac{\mathrm{d}}{\mathrm{d}t}(f \circ \varphi_t) = V \circ f \circ \varphi_t \\
\Leftrightarrow \quad & f \circ \varphi_t = \exp(tV) \circ f
\end{aligned}$$

Flowmap for SDEs

Itô lemma for Stratonovich integrals $\Longrightarrow$

$$f \circ y_t = f \circ y_0 + \sum_{i=0}^{d} \int_0^t V_i \circ f \circ y_\tau \, \mathrm{d}W_\tau^i.$$

where

$$V_i \circ f = V_i \cdot \nabla_y f.$$

Hence for flowmap we have

$$f \circ \varphi_t = f \circ \mathrm{id} + \sum_{i=0}^{d} \int_0^t V_i \circ f \circ \varphi_\tau \, \mathrm{d}W_\tau^i.$$

Flowmap series

Non-autonomous linear functional differential equation for $f_t := f \circ \varphi_t$
$\implies$ set up the formal iterative procedure (with $f_t^{(0)} = f$):

$$f_t^{(n+1)} = f + \sum_{i=0}^{d} \int_0^t V_i \circ f_\tau^{(n)} \, \mathrm{d}W_\tau^i,$$

Iterating $\implies$

$$f \circ \varphi_t = f + \sum_{i=0}^{d} \underbrace{\left(\int_0^t \mathrm{d}W_t^i \right)}_{J_i(t)} V_i \circ f + \sum_{i,j=0}^{d} \underbrace{\left(\int_0^t \int_0^{\tau_1} \mathrm{d}W_{\tau_2}^i \, \mathrm{d}W_{\tau_1}^j \right)}_{J_{ij}(t)} \underbrace{V_i \circ V_j}_{V_{ij}} \circ f + \cdots .$$

Stochastic Taylor expansion

Stochastic Taylor expansion for the flowmap:

$$\varphi_t = \sum_{w \in \mathbb{A}^*} J_w(t)\, V_w.$$

- $w = a_1 \ldots a_n$ are words from alphabet $\mathbb{A} := \{0, 1, \ldots, d\}$;
- Scalar random variables

$$J_w(t) := \int_0^t \cdots \int_0^{\tau_{n-1}} \mathrm{d}W_{\tau_n}^{a_1} \cdots \mathrm{d}W_{\tau_1}^{a_n};$$

- Partial differential operators $V_w := V_{a_1} \circ \cdots \circ V_{a_n}$.

Encodes all stochastic and geometric information of system.

Basic Strategy

Is as follows:

1. Construct the series $\sigma_t = f(\varphi_t)$;
2. Truncate the series σ_t to $\hat{\sigma}_t$ according to a grading;
3. Compute $\hat{\varphi}_t = f^{-1}(\hat{\sigma}_t) \implies$ numerical scheme.

For example: Stochastic Taylor series implies

$$\varphi_t = \sum_{\mathrm{g}(w)\leq n} J_w V_w + \sum_{\mathrm{g}(w)\geqslant n+1} J_w V_w$$

$$\hat{\varphi}_t = \sum_{\mathrm{g}(w)\leq n} J_w V_w + \sum_{\mathrm{g}(w)=n+1} \bar{\mathrm{E}}(J_w) V_w$$

Example: Castell–Gaines method

Exponential Lie series $\psi_t = \log \varphi_t$ given by

$$\begin{aligned}\psi_t &= (\varphi_t - \mathrm{id}) - \tfrac{1}{2}(\varphi_t - \mathrm{id})^2 + \tfrac{1}{3}(\varphi_t - \mathrm{id})^3 + \cdots \\ &= \sum_{i=0}^{d} J_i V_i + \sum_{i>j} \tfrac{1}{2}(J_{ij} - J_{ji})[V_i, V_j] + \cdots .\end{aligned}$$

Truncate, then across $[t_n, t_{n+1}]$, we have

$$\hat{\psi}_{t_n,t_{n+1}} = \sum_{i=0}^{d} \Big(\Delta W^i(t_n, t_{n+1})\Big) V_i + \sum_{i>j} \hat{A}_{ij}(t_n, t_{n+1}) [V_i, V_j].$$

Then

$$\hat{y}_{t_{n+1}} \approx \exp(\hat{\psi}_{t_n,t_{n+1}}) \circ \hat{y}_{t_n}.$$

Example: Castell–Gaines method recovery

For each path simulated $\Delta W^i(t_n, t_{n+1})$ and $\hat{A}_{ij}(t_n, t_{n+1})$ are fixed constants $\implies$ $\hat{\psi}_{t_n,t_{n+1}}$ an autonomous vector field.

Thus, for $\tau \in [0, 1]$ and with $u(0) = \hat{y}_{t_n}$, solve

$$u'(\tau) = \hat{\psi}_{t_n,t_{n+1}} \circ u(\tau).$$

Using a suitable high order ODE integrator generates $u(1) \approx \hat{y}_{t_{n+1}}$.

Castell & Gaines: this method is *asymptotically efficient* in the sense of N. Newton.

Hopf algebra representation

Stochastic Taylor series for the flowmap:

$$\varphi = \sum_w w \otimes w,$$

lies in the product algebra (over $\mathbb{K} = \mathbb{R}$):

$$\mathbb{K}\langle\mathbb{A}\rangle_{\mathrm{sh}} \otimes \mathbb{K}\langle\mathbb{A}\rangle_{\mathrm{co}}.$$

See Reutenauer 1993.

$$(u \otimes x)(v \otimes y) = (u \sqcup\!\sqcup\, v) \otimes (xy),$$

Hopf algebra of words

$\mathbb{K}\langle\mathbb{A}\rangle_{\mathrm{co}}$:

- Product, concatenation: $u \otimes v \mapsto uv$;
- Coproduct, deshuffle: $w \mapsto \sum_{u \sqcup\!\sqcup v = w} u \otimes v$;
- Unit, counit and antipode S: $a_1 \dots a_n \mapsto (-1)^n a_n \dots a_1$.

$\mathbb{K}\langle\mathbb{A}\rangle_{\mathrm{sh}}$:

- Product, shuffle: $u \otimes v \mapsto u \sqcup\!\sqcup v$;
- Coproduct, deconcatenation: $w \mapsto \Delta(w) = \sum_{uv=w} u \otimes v$;
- Unit, counit and antipode same.

Power series function

Suppose we apply function to the flow-map φ:

$$
\begin{aligned}
f(\varphi) &= \sum_{k\geqslant 0} c_k\,\varphi^k \\
&= \sum_{k\geqslant 0} c_k \left(\sum_{w\in\mathbb{A}^*} w\otimes w \right)^k \\
&= \sum_{k\geqslant 0} c_k \sum_{u_1,\ldots,u_k\in\mathbb{A}^*} (u_1 ш \ldots ш u_k)\otimes(u_1\ldots u_k) \\
&= \sum_{w\in\mathbb{A}^*} \left(\sum_{k\geqslant 0} c_k \sum_{w=u_1\ldots u_k} u_1 ш \ldots ш u_k \right)\otimes w \\
&= \sum_{w\in\mathbb{A}^*} F(w)\otimes w
\end{aligned}
$$

Power series function recoding

In other words

$$f(\varphi) = \sum_{w \in \mathbb{A}^*} (F \circ w) \otimes w,$$

where

$$F \circ w = \sum_{k \geqslant 0} c_k \sum_{\substack{u_1, \ldots, u_k \\ w = u_1 \ldots u_k}} u_1 \sqcup\!\sqcup \ldots \sqcup\!\sqcup u_k.$$

$\Longrightarrow$ encode the action of f by $F \in \mathrm{End}(\mathbb{K}\langle\mathbb{A}\rangle_{\mathrm{sh}})$.

Note the stochastic Taylor series is $F = \mathrm{id}$.

Shuffle convolution algebra

Embedding $\text{End}(\mathbb{K}\langle\mathbb{A}\rangle_{\text{sh}}) \to \mathbb{K}\langle\mathbb{A}\rangle_{\text{sh}} \otimes \mathbb{K}\langle\mathbb{A}\rangle_{\text{co}}$ given by

$$X \mapsto \sum_{w} X(w) \otimes w,$$

is an algebra homomorphism for the convolution product:

$$X \star Y = \sqcup\!\sqcup \circ (X \otimes Y) \circ \Delta.$$

We henceforth denote $\mathbb{H} := \text{End}(\mathbb{K}\langle\mathbb{A}\rangle_{\text{sh}})$.

Unit in $\mathbb{H}$ is ν and the antipode

$$S \star \text{id} = \text{id} \star S = \nu.$$

Functions of the identity

$$F^{\star}(\mathrm{id}) = \sum_{k \geqslant 0} c_k \mathrm{id}^{\star k}$$

$$\begin{aligned}
\log^{\star}(\mathrm{id}) &= J - \tfrac{1}{2}J^{\star 2} + \tfrac{1}{3}J^{\star 3} - \cdots + \frac{(-1)^{k+1}}{k}J^{\star k} + \cdots \\
\operatorname{sinhlog}^{\star}(\mathrm{id}) &= \tfrac{1}{2}(\mathrm{id} - S) \\
S &= \nu - J + J^{\star 2} - J^{\star 3} + \cdots \\
\operatorname{sinhlog}^{\star}(\mathrm{id}) &= J - \tfrac{1}{2}J^{\star 2} + \tfrac{1}{2}J^{\star 3} - \cdots + (-1)^{k+1}\tfrac{1}{2}J^{\star k} + \cdots,
\end{aligned}$$

where $J := \mathrm{id} - \nu$ is the *augmented ideal projector*, it is the identity on non-empty words and zero on the empty word:

$$J^{\star k} \circ w = \sum_{w = u_1 \ldots u_k} u_1 \sqcup\!\sqcup \ldots \sqcup\!\sqcup u_k.$$

where all the words in the decomposition must be non-empty.

Generating endomorphisms from endomorphisms

For any $X \in \mathbb{H}$:

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

$$\log^{\star}(X) = \sum_{k \geqslant 1} \frac{(-1)^{k+1}}{k} (X - \nu)^{\star k}$$

$$\exp^{\star}(X) = \sum_{k \geqslant 0} \frac{1}{k!} X^{\star k}$$

$$\operatorname{sinhlog}^{\star}(X) = \tfrac{1}{2}\Big(X - X^{\star(-1)}\Big)$$

$$\operatorname{coshlog}^{\star}(X) = \tfrac{1}{2}\Big(X + X^{\star(-1)}\Big)$$

Expectation endomorphism

Let $\mathbb{D}^* \subset \mathbb{A}^*$ be the free monoid of words on $\mathbb{D} = \{0, 11, 22, \ldots, dd\}$.

Definition (Expectation endomorphism)

Linear map $\mathrm{E} \in \mathbb{H}$ such that

$$\mathrm{E}\colon\ w \mapsto \begin{cases} \dfrac{t^{\mathrm{n}(w)}}{2^{\mathrm{d}(w)}\mathrm{n}(w)!}\,\mathbf{1}, & w \in \mathbb{D}^*, \\ 0\,\mathbf{1}, & w \in \mathbb{A}^* \backslash \mathbb{D}^*. \end{cases}$$

Here $\mathrm{d}(w)$ is the number of non-zero consecutive pairs in w from $\mathbb{D}$, $\mathrm{n}(w) = \mathrm{d}(w) + \mathrm{z}(w)$ where $\mathrm{z}(w)$ is the number of zeros in w.

Inner product

Definition (Inner product)

We define the *inner product* of $X, Y \in \mathbb{H}$ with respect to $\boldsymbol{V}$ to be

$$\langle X, Y\rangle_{\mathbb{H}} := \sum_{u,v\in\mathbb{A}^*} \bar{\mathrm{E}}\Big(X(u) \sqcup\!\sqcup\, Y(v)\Big)(u,v).$$

The norm of an endomorphism $X \in \mathbb{H}$ is $\|X\|_{\mathbb{H}} := \langle X, X\rangle^{1/2}$.

Motivation: if

$$x_t = \sum_{w\in\mathbb{A}^*} X(w)\, V_w(y_0) \qquad \text{and} \qquad y_t = \sum_{w\in\mathbb{A}^*} Y(w)\, V_w(y_0),$$

our definition is based on the L^2-inner product $\langle x_t, y_t\rangle_{L^2} = \bar{\mathrm{E}}\,(x_t, y_t)$.

Graded class

Definition (Grading map)

This is the linear map $\mathrm{g}\colon \mathbb{K}\langle\mathbb{A}\rangle_{\mathrm{sh}} \to \mathbb{Z}_+$ given by

$$\mathrm{g}\colon w \mapsto |w|.$$

Definition (Graded class)

For a given $n \in \mathbb{Z}_+$, we set $\mathbb{S}_n := \left\{w \in \mathbb{K}\langle\mathbb{A}\rangle_{\mathrm{sh}} : \mathrm{g}(w) = n\right\}$ and

$$\mathbb{S}_{\leqslant n} := \bigoplus_{k \leqslant n} \mathbb{S}_k \qquad \text{and} \qquad \mathbb{S}_{\geqslant n} := \bigoplus_{k \geqslant n} \mathbb{S}_k.$$

A subset $\mathbb{S} \subseteq \mathbb{A}^*$ is of *graded class* if, for a given $n \in \mathbb{Z}_+$, $\mathbb{S} = \mathbb{S}_n$, $\mathbb{S} = \mathbb{S}_{\leqslant n}$ or $\mathbb{S} = \mathbb{S}_{\geqslant n}$. We denote by $\pi_{\mathbb{S}}\colon \mathbb{K}\langle\mathbb{A}\rangle_{\mathrm{sh}} \to \mathbb{S}$ the canonical projection onto any graded class subspace $\mathbb{S}$,

Reversing Lemma

Lemma (M. & Wiese 2009)

For any pair $u, v \in \mathbb{A}^*$, we have $E(u \sqcup\!\sqcup\, v) \equiv E\big((|S| \circ u) \sqcup\!\sqcup\, (|S| \circ v)\big)$.

Lemma: properties

Lemma

For $\mathbb{A} = \{0, 1, \ldots, d\}$ *and graded class subspace* $\mathbb{S} = \mathbb{S}_n$*:*

1. $\langle X, Y \rangle = \langle |S| \circ X, |S| \circ Y \rangle$;
2. $\langle |S|, |S| \rangle = \langle S, S \rangle = \langle id, id \rangle$;
3. $\langle sinhlog^{\star}(id), coshlog^{\star}(id) \rangle = 0$;
4. $\|id\|^2 = \|sinhlog^{\star}(id)\|^2 + \|coshlog^{\star}(id)\|^2$;
5. $\langle X, E \circ Y \rangle = \langle E \circ X, E \circ Y \rangle$;
6. $\langle E \circ id, E \circ id \rangle = \langle E \circ |S|, E \circ |S| \rangle = \langle E \circ S, E \circ S \rangle$;
7. $\langle E \circ sinhlog^{\star}(id), E \circ coshlog^{\star}(id) \rangle = 0$;
8. $\langle |S|, J^{\star n} \rangle = \langle id, J^{\star n} \rangle$,

Accuracy measurement

We measure the accuracy by $||r_t \circ y_0||_{L^2}$ where

$$r_t := \varphi_t - \hat{\varphi}_t.$$

A general stochastic integrator has the form

$$\widehat{\mathrm{id}} := f^{\star(-1)} \circ \pi_{\mathbb{S}_{\leqslant n}} \circ f^{\star}(\mathrm{id}).$$

The error is $||R||$ in the $\mathbb{H}$-norm where

$$R := \mathrm{id} - \widehat{\mathrm{id}}.$$

Definition (Pre-remainder)

$$Q := f^{\star}(\mathrm{id}) - \pi_{\mathbb{S}_{\leqslant n}} \circ f^{\star}(\mathrm{id}).$$

Efficient integrator

Definition (Efficient integrator)

We will say that a numerical approximation to the solution of an SDE is an *efficient integrator* if it generates a strong scheme that is more accurate in the root mean square sense than the corresponding stochastic Taylor scheme of the same strong order, independent of the governing vector fields and to all orders: i.e.

$$\|(\mathrm{id} - E) \circ R\|^2 \leqslant \|(\mathrm{id} - E) \circ \mathrm{id}\|^2.$$

Sinhlog integrator of strong order $n/2$:

$$\begin{aligned} P &:= \pi_{\mathbb{S}_{\leqslant n}} \circ \mathrm{sinhlog}^{\star}(\mathrm{id}) \\ &= \left(J - \tfrac{1}{2}J^{\star 2} + \cdots + \tfrac{1}{2}(-1)^{n+1}J^{\star n}\right) \circ \pi_{\mathbb{S}_{\leqslant n}}. \end{aligned}$$

$$Q = \mathrm{sinhlog}^{\star}(\mathrm{id}) \circ \pi_{\mathbb{S}_{\geqslant n+1}}.$$

Sinhlog pre-remainder and remainder

$$\begin{aligned} h^\star(X,+\nu) &= \left(X^{\star 2}+\nu\right)^{\star(1/2)} \\ &= \nu + \tfrac{1}{2}X^{\star 2} - \tfrac{1}{8}X^{\star 4} + \cdots \end{aligned}$$

$$\begin{aligned} R &= \mathrm{id} - \mathrm{sinhlog}^{\star(-1)} \circ P \\ &= \mathrm{sinhlog}^{\star(-1)} \circ (P+Q) - \mathrm{sinhlog}^{\star(-1)} \circ P \\ &= Q + h^\star(P+Q,+\nu) - h^\star(P,+\nu) \\ &= Q + \tfrac{1}{2}\Big((P+Q)^{\star 2} - P^{\star 2}\Big) + \cdots \\ &= Q + \tfrac{1}{2}\Big(P \star Q + Q \star P\Big) + \mathcal{O}\Big(Q^{\star 2}\Big) \\ &= Q \circ \pi_{\mathbb{S}_{n+1}}, \end{aligned}$$

since $\mathrm{sinhlog}^\star(\mathrm{id}) = J + \cdots$: $P \star Q = \Big(J \circ \pi_{\mathbb{S}_{\leqslant n}}\Big) \star \Big(J \circ \pi_{\mathbb{S}_{\geqslant n+1}}\Big)$.

Sinhlog efficient

Reversing Lemma with $\mathbb{S} = \mathbb{S}_{n+1} \implies$

$$\|\mathrm{id}\|^2 = \|Q\|^2 + \|\mathrm{coshlog}^{\star}(\mathrm{id})\|^2.$$

and

$$\left\|(\mathrm{id} - E) \circ \mathrm{id}\right\|^2 = \left\|(\mathrm{id} - E) \circ Q\right\|^2 + \left\|(\mathrm{id} - E) \circ \mathrm{coshlog}^{\star}(\mathrm{id})\right\|^2.$$

Indeed, for any ϵ:

$$\mathrm{sinhlog}_{\epsilon}^{\star}(\mathrm{id}) = J + \tfrac{1}{2}\sum_{k=2}^{\infty}(-1)^{k+1}J^{\star k} + \epsilon J^{\star(n+1)}$$

$$\|\mathrm{id}\|^2 = \|Q_{\epsilon}\|^2 + \|\mathrm{coshlog}^{\star}(\mathrm{id})\|^2 - \epsilon\left\langle \mathrm{id} - S, J^{\star(n+1)}\right\rangle - \epsilon^2\left\|J^{\star(n+1)}\right\|^2.$$

Optimality Theorem

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

$$f^{\star(-1)}(X;\epsilon) = X + h^{\star}(X, \epsilon\,\nu),$$

Theorem

For every $\epsilon > 0$ the class of integrators $f^{\star}(id;\epsilon)$ is efficent. When $\epsilon = 1$, the error of the integrator $f^{\star}(id;1)$ realizes its smallest possible value compared to the error of the corresponding stochastic Taylor integrator. Thus a strong stochastic integrator based on the sinhlog endomorphism is optimally efficient within this class, to all orders.

$$f^{\star}(\mathrm{id};\epsilon) = \tfrac{1}{2}(1-\epsilon)\nu + \tfrac{1}{2}(1+\epsilon)J - \tfrac{1}{2}\epsilon\Big(J^{\star 2} - J^{\star 3} + \cdots\Big).$$

Optimality Theorem proof: step 1

$$h\colon (x,y) \mapsto (x^2+y)^{1/2}.$$

$$h(x+q,y) - h(x,y) = \frac{x}{(x^2+y)^{1/2}} \cdot q + \cdots ,$$

$$P := \pi_{\mathbb{S}_{\leqslant n}} \circ f^{\star}(\mathrm{id};\epsilon)$$

$$\begin{aligned} R &= \mathrm{id} - \widehat{\mathrm{id}} \\ &= f^{\star(-1)} \circ (P+Q) - f^{\star(-1)} \circ P \\ &= Q + h^{\star}(P+Q,\epsilon\nu) - h^{\star}(P,\epsilon\nu). \end{aligned}$$

Optimality Theorem proof: step 2

$$\begin{aligned}
R &= Q + h^{\star}(P+Q,\epsilon\,\nu) - h^{\star}(P,\epsilon\,\nu) \\
&= Q + \Big((P+Q)^{\star 2} + \epsilon\,\nu\Big)^{\star(1/2)} - \Big(P^{\star 2} + \epsilon\,\nu\Big)^{\star(1/2)} \\
&= Q + \tfrac{1}{2}(P^{\star 2} + \epsilon\,\nu)^{\star(-1/2)} \star \Big(P \star Q + Q \star P\Big) + \mathcal{O}\Big(Q^{\star 2}\Big).
\end{aligned}$$

$$\begin{aligned}
P &= \tfrac{1}{2}(1-\epsilon)\,\nu + \mathcal{O}(J), \\
(P^{\star 2} + \epsilon\,\nu)^{\star(-1/2)} &= \Big(\tfrac{1}{2}(1+\epsilon)\Big)^{-1}\,\nu + \mathcal{O}(J).
\end{aligned}$$

$$\begin{aligned}
R &= Q + \frac{1-\epsilon}{1+\epsilon} Q + \mathcal{O}(J \star Q) \\
&= \frac{2}{1+\epsilon} Q \circ \pi_{\mathbb{S}_{n+1}}.
\end{aligned}$$

Optimality Theorem proof: step 3

On $\mathbb{S}_{n+1}$:

$$\begin{aligned}
\|\mathrm{id}\|^2 &= \langle R + \mathrm{id} - R, R + \mathrm{id} - R\rangle \\
&= \|R\|^2 + 2\langle R, \mathrm{id} - R\rangle + \langle \mathrm{id} - R, \mathrm{id} - R\rangle \\
&= \|R\|^2 + 2\left\langle \mathrm{id} - \frac{1}{1+\epsilon}(\mathrm{id} - \epsilon S), \frac{1}{1+\epsilon}(\mathrm{id} - \epsilon S)\right\rangle \\
&\qquad + \left\langle \mathrm{id} - \frac{1}{1+\epsilon}(\mathrm{id} - \epsilon S), \mathrm{id} - \frac{1}{1+\epsilon}(\mathrm{id} - \epsilon S)\right\rangle \\
&= \|R\|^2 + 1 - \frac{1}{(1+\epsilon)^2}\left(1 - 2\epsilon\langle \mathrm{id}, S\rangle + \epsilon^2\right) \\
&= \|R\|^2 + \frac{2\epsilon}{(1+\epsilon)^2}\left(1 + \langle \mathrm{id}, S\rangle\right).
\end{aligned}$$

Optimality Theorem proof: step 4

$$R = Q + \tfrac{1}{2}(P^{\star 2} - \nu)^{\star(-1/2)} \star \Big(P \star Q + Q \star P\Big) + \mathcal{O}\Big(Q^{\star 2}\Big).$$

$$P = \nu + \tfrac{1}{2} J^{\star 2} + \mathcal{O}\Big(J^{\star 3}\Big).$$

$$(P^{\star 2} - \nu)^{\star(1/2)} = \sqrt{2}\,(P - \nu)^{\star(1/2)} \star \Big(\nu + \tfrac{1}{4}(P - \nu) + \mathcal{O}\Big((P - \nu)^{\star 2}\Big)\Big).$$

$$R = Q + \Big(J^{\star(-1)} \circ \pi_{\mathbb{S}_{\leqslant n}}\Big) \star Q + \mathcal{O}\Big(J \star Q\Big).$$

$\Longrightarrow$ order reduction.

References

- 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. 2008 Stochastic Lie group integrators, *SIAM J. Sci. Comput.* **30**(2), 597–617
- Malham, S.J.A., Wiese, A. 2009 Stochastic expansions and Hopf algebras. *Proc. R. Soc. A* **465**, 3729–3749.
- Malham, S.J.A., Wiese, A. 2011 Lévy area logistic expansion and simulation, arXiv:1107.0151v1
- Newton, N.J. 1991, *SIAM J. Appl. Math.* **51**, 542–567
- Reutenauer, C. 1993 *Free Lie algebras*, London Mathematical Society Monographs New Series 7, Oxford Science Publications