

Universal optimal stochastic expansions

Simon J.A. Malham and Anke Wiese

Heriot-Watt University, Edinburgh, UK

Bergen: 6th June 2008

Outline

- 1 Stochastic differential equations
- 2 Stochastic Taylor expansion
- 3 Local and global errors
- 4 Exponential Lie series
- 5 Hopf algebra of words
- 6 Concatenation-shuffle operator algebra
- 7 Sinh-log remainder is smaller

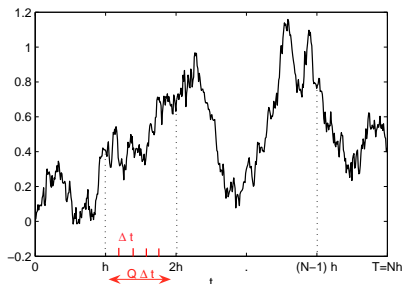
Stochastic differential equations

$$dy_t = V_0(y_t) dt + V_1(y_t) dW_t^1 + \cdots + V_d(y_t) dW_t^d$$

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

- ▶ Non-commuting vector fields: $V_i = \sum_{j=1}^n V_i^j(y) \partial_y^j$
- ▶ d -dimensional driving signal: (W^1, \dots, W^d)
- ▶ Convention: $W_t^0 \equiv t$
- ▶ Solution process: $y: \mathbb{R}_+ \rightarrow \mathbb{R}^N$,

Wiener process



- ▶ $W_t - W_s \sim N(0, \sqrt{t - s})$
- ▶ Independent increments
- ▶ Continuous, potentially nowhere differentiable

Stochastic chain rule

$$y_t = y_0 + \sum_i \int_0^t V_i \circ y_\tau dW_\tau^i$$

Itô lemma (stochastic chain rule) \Rightarrow

$$f \circ y_t = f \circ y_0 + \sum_j \int_0^t V_j \circ f \circ y_\tau dW_\tau^j$$

E.g. choose $f = V_i \Rightarrow$

$$V_i \circ y_t = V_i \circ y_0 + \sum_j \int_0^t V_j \circ V_i \circ y_\tau dW_\tau^j$$

$$y_t = y_0 + \sum_i \int_0^t \left(V_i \circ y_0 + \sum_j \int_0^{\tau_1} V_j \circ V_i \circ y_{\tau_2} dW_{\tau_2}^j \right) dW_{\tau_1}^i$$

Stochastic Taylor series

$$y_t = y_0 + \sum_i \int_0^t dW_{\tau_1}^i V_i \circ y_0 + \sum_{i,j} \int_0^t \int_0^{\tau_1} V_j \circ V_i \circ y_{\tau_2} dW_{\tau_2}^j dW_{\tau_1}^i$$

Now choose $f = V_j \circ V_i \Rightarrow$

$$y_t = y_0 + \sum_i \underbrace{\int_0^t dW_{\tau_1}^i}_{J_i(t)} V_i \circ y_0 + \sum_{i,j} \underbrace{\int_0^t \int_0^{\tau_1} dW_{\tau_2}^j dW_{\tau_1}^i}_{J_{ji}(t)} V_j \circ V_i \circ y_0 + \dots$$

- ▶ Euler-Maruyama, Milstein and Runge–Kutta methods (*Kloeden & Platen* 1999)
- ▶ Need approximations for iterated integrals for strong solution

Flow map

$$y_t = \varphi_t \circ y_0$$

Hence

$$\varphi_t = \text{id} + \sum_i J_i(t) V_i + \sum_{i,j} J_{ji}(t) V_j \circ V_i + \sum_{i,j,k} J_{kji}(t) V_k \circ V_j \circ V_i + \dots$$

i.e.

$$\varphi_t = \text{id} + \sum_{w \in \mathbb{A}^+} J_w(t) V_w$$

Here $\mathbb{A}^+ = \{\text{non-empty words over } \mathbb{A} = \{0, 1, \dots, d\}\}$

Local error

Exact and approximate solutions:

$$y_t = \varphi_t \circ y_0 \quad \text{and} \quad \hat{y}_t = \hat{\varphi}_t \circ y_0$$

Flow remainder:

$$R = \varphi_t - \hat{\varphi}_t = \sum_{"w \in \mathbb{A}^+"} J_w(t) V_w$$

Local error measure:

$$\begin{aligned} \|R \circ y_0\|_{L^2}^2 &= \mathbb{E}((R \circ y_0)^T (R \circ y_0)) \\ &= \sum_{"u, v \in \mathbb{A}^+"} \mathbb{E}(J_u(t) J_v(t)) (V_u \circ y_0)^T (V_v \circ y_0) \end{aligned}$$

Stochastic integral properties

Expectations: $\mathbb{E}(J_i) = 0$, $\mathbb{E}(J_i^2) = h$

Stratonovich-to-Itô relations:

$$w = a_1 a_2: J_w = I_w + \frac{1}{2} \delta_{a_1 a_2} I_0$$

$$w = a_1 a_2 a_3: J_w = I_w + \frac{1}{2} (\delta_{a_1 a_2} I_{0 a_3} + \delta_{a_2 a_3} I_{a_1 0})$$

Shuffle relations: $J_u J_v = \sum_{w \in \text{sh}(u,v)} J_w$

$$J_{a_1} J_{a_2} = J_{a_1 a_2} + J_{a_2 a_1}$$

$$J_{a_1} J_{a_2 a_3} = J_{a_1 a_2 a_3} + J_{a_2 a_1 a_3} + J_{a_2 a_3 a_1}$$

Exponential Lie series

Set $\varphi_t = \exp \psi_t$ then

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

Local error: $R^{\text{ls}} = \exp \psi_t - \exp \hat{\psi}_t = \psi_t - \hat{\psi}_t + o(\psi_t - \hat{\psi}_t)$

- ▶ *Magnus* 1954, *Chen* 1957, *Kunita* 1980, *Ben Arous* 1989, *Castell* 1993. *Castell–Gaines* 1995, *Burrage* 1999, *Misawa* 2001, *P-C. Moan* 2004.

Castell–Gaines (ODE) method

Truncated exponential Lie series across $[t_n, t_{n+1}]$:

$$\hat{\psi}_{t_n, t_{n+1}} = \hat{J}_1 V_1 + \hat{J}_2 V_2 + \hat{J}_0 V_0 + \frac{1}{2}(\hat{J}_{12} - \hat{J}_{21})[V_1, V_2].$$

Approximate solution:

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

Castell–Gaines: solve ODE

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

across $\tau \in [0, 1]$ with $u(0) = y_{t_n}$ gives $u(1) \approx y_{t_{n+1}}$.

Error comparison

Set $R^{\text{st}} = R^{\text{ls}} + \bar{R}$

$$\begin{aligned}\|R^{\text{st}} \circ y_0\|_{L^2}^2 &= \mathbb{E} \left((R^{\text{ls}} + \bar{R}) \circ y_0 \right)^{\text{T}} \left((R^{\text{ls}} + \bar{R}) \circ y_0 \right) \\ &= \|R^{\text{ls}} \circ y_0\|_{L^2}^2 + \mathbb{E} \left(\bar{R} \circ y_0 \right)^{\text{T}} \left(R^{\text{ls}} \circ y_0 \right) \\ &\quad + \mathbb{E} \left(R^{\text{ls}} \circ y_0 \right)^{\text{T}} \left(\bar{R} \circ y_0 \right) + \mathbb{E} \left(\bar{R} \circ y_0 \right)^{\text{T}} \left(\bar{R} \circ y_0 \right) \\ &= \|R^{\text{ls}} \circ y_0\|_{L^2}^2 + E\end{aligned}$$

If $E \geq 0$ then $\|R^{\text{st}} \circ y_0\|_{L^2} \geq \|R^{\text{ls}} \circ y_0\|_{L^2}$

Example: Lie series more accurate

Order $\frac{1}{2}$:

$$\varphi = \text{id} + J_0 V_0 + \sum_{i=1}^d J_i V_i + \frac{1}{2} h \sum_{i=1}^d V_{ii} \quad \text{and} \quad \psi = J_0 V_0 + \sum_{i=1}^d J_i V_i$$

Remainders:

$$R^{\text{ls}} = \sum_{i>j} \frac{1}{2} (J_{ij} - J_{ji}) (V_{ij} - V_{ji})$$

$$\bar{R} = \sum_{i=1}^d (J_{ii} - \frac{1}{2} h) V_{ii} + \sum_{i>j} \frac{1}{2} (J_{ij} + J_{ji}) (V_{ij} + V_{ji})$$

R^{ls} and \bar{R} uncorrelated $\Rightarrow \|R^{\text{st}} \circ y_0\|_{L^2}^2 = \|R^{\text{ls}} \circ y_0\|_{L^2}^2 + \|\bar{R} \circ y_0\|_{L^2}^2$

Similarly for orders 1 and $\frac{3}{2}$ when assume $[V_i, V_j] = 0$ for $i, j \in \{1, \dots, d\}$

Counter-example

Order $\frac{3}{2}$: $[V_i, V_j] \neq 0$. Remainders:

$$R^{\text{st}} = \sum_i (J_{i0} V_{i0} + J_{0i} V_{0i}) + \sum_{i,j,k} J_{ijk} V_{ijk}$$

$$R^{\text{ls}} = \sum_i \frac{1}{2} (J_{i0} - J_{0i}) [V_i, V_0] + \sum_{i \neq j} (J_{jii} - \frac{1}{2} J_i J_{ji} + \frac{1}{12} J_i^2 J_j) [V_i, [V_i, V_j]] \\ + \sum_{i < j < k} (\dots) [V_i, [V_j, V_k]] + (\dots) [V_j, [V_i, V_k]]$$

Quadratic form: $U_{jii} = (V_{\text{perms}(ijj)}, V_{jjj}, V_{j0}, V_{0j})^T$ &
 $U_{ijk} = (U_{\text{perms}(ijk)})^T$

$$E = h^2 \sum_{i \neq j} U_{jii}^T B U_{jii} + h^2 \sum_{i \neq j \neq k} U_{ijk}^T C U_{ijk}$$

C is positive semi-definite, but B is not!

No drift

Now $\mathbb{A} = \{1, \dots, d\}$.

$$\varphi = \text{id} + \sum_{w \in \mathbb{A}^+} J_w V_w$$

Suppose: $\varphi = \text{id} + F(\psi) \Rightarrow$

$$\psi = F^{-1}(\varphi - \text{id}) = \sum_{k=1}^{\infty} C_k (\varphi - \text{id})^k = \sum_{w \in \mathbb{A}^+} K_w V_w$$

where

$$K_w = \sum_{k=1}^{|w|} C_k \sum_{\substack{u_1, \dots, u_k \in \mathbb{A}^+ \\ u_1 u_2 \dots u_k = w}} J_{u_1} J_{u_2} \dots J_{u_k}(t)$$

Goal: choose best C_k

Hopf algebra of words

\mathbb{A}^* = free monoid on \mathbb{A}

= endow set words over \mathbb{A} with *concat product*

$a_i \in \mathbb{A} \Rightarrow w = a_1 \dots a_n \in \mathbb{A}$

1 = empty word $\Rightarrow 1 \cdot w = w \cdot 1 = w$

\mathbb{K} = commutative ring with unit

$\mathbb{K}\langle\mathbb{A}\rangle$ = noncommutative polynomials/series on \mathbb{A} over \mathbb{K}

Polynomial: $P = \sum_{w \in \mathbb{A}} (P, w)w$

(Reutenauer)

Concatenation and shuffle algebras

Concatenation product on $\mathbb{K}\langle\mathbb{A}\rangle$:

$$PQ = \sum_{u,v \in \mathbb{A}^*} (P, u)(Q, v)uv$$

Shuffle product: $u \sqcup v$

Extended to $\mathbb{K}\langle\mathbb{A}\rangle$ by:

$$P \sqcup Q = \sum_{u,v \in \mathbb{A}^*} (P, u)(Q, v)u \sqcup v$$

Associative, distributive wrt addition, shuffle algebra,

$$1 \sqcup w = w \sqcup 1 = w$$

Two bi-algebra structures: $(\mathbb{K}\langle\mathbb{A}\rangle, \eta, \varepsilon, c, \delta)$ and $(\mathbb{K}\langle\mathbb{A}\rangle, \eta, \varepsilon, s, \delta')$

Hopf algebra tensor product

Bialgebra with antipode $\alpha \in \text{End}(\mathbb{K}\langle A \rangle)$

Antipode here is linear signed reversal mapping:

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

Hence two Hopf algebra structures, consequently set:

$$\mathcal{H} = \mathbb{K}\langle A \rangle \otimes \mathbb{K}\langle A \rangle$$

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

Concatenation-shuffle operator algebra

$$u_{n_1} \sqcup u_{n_2} \sqcup \dots \sqcup u_{n_k} = \zeta \left(\underbrace{c^{n_1} s c^{n_2} s c^{n_3} \dots s c^{n_k}}_b \otimes \underbrace{u_{n_1} u_{n_2} \dots u_{n_k}}_w \right)$$

$$\mathbb{B} = \{c, s\} \longrightarrow \mathbb{B}^* \longrightarrow \mathbb{K}\langle \mathbb{B} \rangle$$

Shuffle gluing product: $g: b_1 \otimes b_2 \mapsto b_1 s b_2$

Associative tensor algebra: $\mathcal{K} = \bigoplus_{n \geq 0} \mathbb{K}\langle \mathbb{B} \rangle_n \otimes \mathbb{K}\langle \mathbb{A} \rangle_{n+1}$

$$(b_1 \otimes u_1)(b_2 \otimes u_2) = (b_1 s b_2) \otimes (u_1 u_2),$$

Homomorphism: $\zeta: \mathcal{K} \rightarrow \mathbb{K}\langle \mathbb{A} \rangle$

Homomorphic maps

Word-to-integral linear map: $\mu: \omega \mapsto J_\omega$ (shuffle homomorphism)

Word-to-vector field linear map: $\kappa: \omega \mapsto V_\omega$ (concat. homomorphism.)

Flow map:

$$\varphi = 1 \otimes \text{id}_V + \sum_{w \in \mathbb{A}^+} J_w \otimes V_w$$

Pullback to \mathcal{H} :

$$(\mu \otimes \kappa)^* \varphi = 1 \otimes 1 + \sum_{w \in \mathbb{A}^+} w \otimes w$$

Pullback to Hopf algebra

$$(\mu \otimes \kappa)^* \varphi = 1 \otimes 1 + \sum_{w \in \mathbb{A}^+} w \otimes w$$

$$\begin{aligned}(\mu \otimes \kappa)^* \psi &= \sum_{k \geq 1} C_k ((\mu \otimes \kappa)^* \varphi - 1 \otimes 1)^k \\ &= \sum_{k \geq 1} C_k \left(\sum_{w \in \mathbb{A}^+} w \otimes w \right)^k \\ &= \sum_{k \geq 1} C_k \sum_{u_1, \dots, u_k \in \mathbb{A}^+} (u_1 \sqcup \dots \sqcup u_k) \otimes (u_1 \dots u_k) \\ &= \sum_{w \in \mathbb{A}^*} \left(\sum_{k=1}^{|w|} C_k \sum_{\substack{u_1, \dots, u_k \in \mathbb{A}^+ \\ w = u_1 \dots u_k}} u_1 \sqcup \dots \sqcup u_k \right) \otimes w\end{aligned}$$

Coefficients in $\mathbb{K}\langle \mathbb{B} \rangle$

$w \in \mathbb{A}^+$ with $|w| = n + 1$:

$$\begin{aligned} K \circ w &= \sum_{k=1}^{|w|} C_k \sum_{\substack{u_1, \dots, u_k \in \mathbb{A}^+ \\ w = u_1 \dots u_k}} u_1 \sqcup \dots \sqcup u_k \\ &= \sum_{k=1}^{n+1} C_k \zeta((c^{n-(k-1)} \sqcup s^{k-1}) \otimes w) \\ &= \sum_{k=0}^n C_{k+1} \zeta((c^{n-k} \sqcup s^k) \otimes w) \\ &= \zeta\left(\left(\sum_{k=0}^n C_{k+1} (c^{n-k} \sqcup s^k)\right) \otimes w\right). \end{aligned}$$

Sinh-log coefficients

Lemma

With

$$C_k = \begin{cases} 1, & k = 1, \\ \frac{1}{2}(-1)^{k+1}, & k \geq 2, \\ \frac{1}{2}(-1)^{n+1} + \epsilon, & k = n, \end{cases}$$

then

$$K = \frac{1}{2}(c^n - \alpha^n) + \epsilon s^n$$

where for $w \in \mathbb{K}\langle \mathbb{A} \rangle_{n+1}$: $\alpha^n \circ w \equiv \alpha \circ w$

Proof: step 1

Partial integration formula:

$$a_1 \dots a_{n+1} = (a_1 \dots a_n) \sqcup a_{n+1} - (a_1 \dots a_{n-1}) \sqcup (a_{n+1} a_n) + \dots + (-1)^n a$$

$$c^n = -c^{n-1}s - c^{n-2}s\alpha - c^{n-3}s\alpha^2 - \dots - \alpha^n$$

$$\alpha^n = -c^n - \sum_{k=0}^{n-1} c^k s \alpha^{n-k-1}$$

Proof: step 2

Antipode polynomial:

$$\alpha^n \equiv -(c - s)^n.$$

True for $n = 1, 2$, assume true for $k = 1, 2, \dots, n - 1$. Direct expansion \Rightarrow

$$(c - s)^n = c^n - \sum_{k=0}^{n-1} c^k s (c - s)^{n-k-1}$$

Use induction.

Proof: step 3

Recall:

$$\begin{aligned}K &= \frac{1}{2}c^n + \frac{1}{2} \sum_{k=0}^n (-1)^k (c^{n-k} \sqcup s^k) + \epsilon s^n \\&= \frac{1}{2}c^n + \frac{1}{2}(c - s)^n + \epsilon s^n \\&= \frac{1}{2}c^n - \frac{1}{2}\alpha^n + \epsilon s^n\end{aligned}$$

Hence

$$K_w = \frac{1}{2}(J_w - J_{\alpha \circ w}) + \epsilon \prod_{i=1}^{n+1} J_{w_i}$$

Sinh-log remainder is smaller I

$$\|R^{\text{st}} \circ y_0\|_{L^2}^2 = \|R^{\text{sl}} \circ y_0\|_{L^2}^2 + E$$

where

$$E \equiv \mathbb{E} (\bar{R} \circ y_0)^{\text{T}} (R^{\text{sl}} \circ y_0) + \mathbb{E} (R^{\text{sl}} \circ y_0)^{\text{T}} (\bar{R} \circ y_0) + \mathbb{E} (\bar{R} \circ y_0)^{\text{T}} (\bar{R} \circ y_0)$$

$$R^{\text{sl}} = \sum_{\substack{w \in \mathbb{A}^+ \\ |w| \geq n+1}} K_w V_w + \dots$$

$$\bar{R} = \sum_{\substack{w \in \mathbb{A}^+ \\ |w|=n+1}} \bar{J}_w V_w$$

where $\bar{J}_w = J_w - K_w$

Sinh-log remainder is smaller II

$$E = \sum_{\substack{u, v \in \mathbb{A}^+ \\ |u|=|v|=n+1}} \mathbb{E} (\bar{J}_u K_v + K_u \bar{J}_v + \bar{J}_u \bar{J}_v) (V_u \circ y_0)^T (V_v \circ y_0)$$

$$\begin{aligned} & \bar{J}_u K_v + K_u \bar{J}_v + \bar{J}_u \bar{J}_v \\ &= \epsilon^0 \left(\frac{1}{2} (J_u J_v - J_{\alpha \circ u} J_{\alpha \circ v}) + \frac{1}{4} (J_u + J_{\alpha \circ u}) (J_v + J_{\alpha \circ v}) \right) \\ & \quad - \epsilon^1 \left(\frac{1}{2} (J_u - J_{\alpha \circ u}) \prod_{i=1}^{n+1} J_{v_i} + \frac{1}{2} (J_v - J_{\alpha \circ v}) \prod_{i=1}^{n+1} J_{u_i} \right) \\ & \quad - \epsilon^2 \left(\prod_{i,j=1}^{n+1} J_{u_i} J_{v_j} \right) \end{aligned}$$

Concluding remarks

- ▶ exponential series when diffusion vector fields commute
- ▶ optimal series when drift included

Global error: standard accumulation

$$\begin{aligned}\|y_T - \hat{y}_T\|_{L^2} &= \left\| \left(\prod_{n=N-1}^0 \varphi_{t_n, t_{n+1}} - \prod_{n=N-1}^0 \hat{\varphi}_{t_n, t_{n+1}} \right) \circ y_0 \right\|_{L^2} \\ &= \left\| \left(\prod_{n=N-1}^0 (\hat{\varphi}_{t_n, t_{n+1}} + R_n) - \prod_{n=N-1}^0 \hat{\varphi}_{t_n, t_{n+1}} \right) \circ y_0 \right\|_{L^2} \\ &= \left\| \left(\sum_{n=0}^{N-1} \hat{\varphi}_{t_{n+1}, t_N} R_n \hat{\varphi}_{t_0, t_n} \right) \circ y_0 \right\|_{L^2} + \text{h.o.t.}\end{aligned}$$

Global error: quadratic form

$$\begin{aligned} & \|y_T - \hat{y}_T\|_{L^2}^2 \\ &= \left\| \left(\sum_{n=0}^{N-1} \hat{\varphi}_{t_{n+1}, t_N} R_n \hat{\varphi}_{t_0, t_n} \right) \circ y_0 \right\|_{L^2}^2 \\ &= \mathbb{E} \left(\left(\sum_{n=0}^{N-1} \hat{\varphi}_{t_{n+1}, t_N} R_n \hat{\varphi}_{t_0, t_n} \right) \circ y_0 \right)^T \left(\left(\sum_{m=0}^{N-1} \hat{\varphi}_{t_{m+1}, t_N} R_m \hat{\varphi}_{t_0, t_m} \right) \circ y_0 \right) \\ &= \sum_{u,v} \left(\sum_n \mathbb{E}(J_u(t_n) J_v(t_n)) \mathbb{E} \left((\hat{\varphi}_{t_{n+1}, t_N} V_u \hat{\varphi}_{t_0, t_n} \circ y_0)^T (\hat{\varphi}_{t_{n+1}, t_N} V_v \hat{\varphi}_{t_0, t_n} \circ y_0) \right) \right. \\ &\quad \left. + \sum_{n \neq m} \mathbb{E}(J_u(t_n)) \mathbb{E}(J_v(t_m)) \mathbb{E} \left((\hat{\varphi}_{t_{n+1}, t_N} V_u \hat{\varphi}_{t_0, t_n} \circ y_0)^T (\hat{\varphi}_{t_{m+1}, t_N} V_v \hat{\varphi}_{t_0, t_m} \circ y_0) \right) \right) \end{aligned}$$

If $\mathbb{E}(J_w) \neq 0$ then set $J_w = \mathbb{E}(J_w) + J_w - \mathbb{E}(J_w)$