Universality of affine and polynomial processes

Christa Cuchiero
(based on ongoing joint works with Sara Svaluto-Ferro and Josef Teichmann)

University of Vienna

Cumulants in Stochastic Analysis

Berlin - Vienna - Zoom, February 25, 2021

Motivation

A plethora of stochastic models stem from the class of affine and polynomial processes, even though this is not always visible at first sight.

- Finite dimensional examples: Lévy processes, Ornstein-Uhlenbeck processes, Feller diffusion, Wishart processes, Black-Scholes model, Wright-Fisher diffusion (Jacobi process), ...
- Infinite dimensional examples:
  - measure valued processes: Dawson-Watanabe process, Fleming-Viot process, Markovian lifts of Volterra processes
  - Hilbert space valued processes: (forward) curve models, lifts of rough volatility models (rough Heston, rough Wishart or rough Bergomi)
  - sequence space valued processes: signature of Brownian motion

Motivation

A plethora of stochastic models stem from the class of affine and polynomial processes, even though this is not always visible at first sight.

- Finite dimensional examples: Lévy processes, Ornstein-Uhlenbeck processes, Feller diffusion, Wishart processes, Black-Scholes model, Wright-Fisher diffusion (Jacobi process), ...
- Infinite dimensional examples:
  - measure valued processes: Dawson-Watanabe process, Fleming-Viot process, Markovian lifts of Volterra processes
  - Hilbert space valued processes: (forward) curve models, lifts of rough volatility models (rough Heston, rough Wishart or rough Bergomi)
  - sequence space valued processes: signature of Brownian motion

$\Rightarrow$ Universal model classes?

$\Rightarrow$ Mathematically precise statements for this universality?

Objectives of this work

$\Rightarrow$ **One step in this direction by linearizing certain classes of SDEs via signature methods**

Objectives of this work

⇒ **One step in this direction by linearizing certain classes of SDEs via signature methods**

- Objectives
  - Develop an affine and polynomial theory for prolongations of generic stochastic processes. These prolongations can for instance consist of
    - ⋆ all monomials or
    - ⋆ all iterated integrals.
  - **Power series expansions** (in terms of the initial value) **for the cumulants as well as for the expected value of other analytic functions** for generic classes of diffusions via affine and polynomial technology

Recent related literature

- E. Alos, J. Gatheral & R. Radoicic ('20): "Exponentiation of conditional expectations under stochastic volatility"
- P. Friz, J. Gatheral & R. Radoicic ('20): "Forests, cumulants, martingales"
- P. Friz, P. Hager & N. Tapia ('21) "Unified Signature Cumulants and Generalized Magnus Expansions"
- H. Lacoin, R. Rhodes, and V. Vargas ('19): "A probabilistic approach of ultraviolet renormalisation in the boundary Sine-Gordon model"
- P. Bonnier & H. Oberhauser ('20): "Signature Cumulants, Ordered Partitions, and Independence of Stochastic Processes"
- K. Ebrahimi-Fard, F. Patras, N. Tapia & L. Zambotti ('18) "Hopf-algebraic Deformations of Products and Wick Polynomials"
- K. Ebrahimi-Fard & F. Patras ('15): "Cumulants, free cumulants and half-shuffles"
- I. Perez Arribas, C. Salvi, L. Szpruch ('20): "Sig-SDEs model for quantitative finance"

Implications

Modern models can be embedded in the affine and polynomial framework:

- Neural SDEs (see e.g. P. Gierjatowicz, M. Sabate-Vidales, D. Siska, L. Szpruch, Z. Zuric ('20) or C.C, W. Khosrawi, J. Teichmann ('20))

$$dX_t = b(X_t, \theta)dt + \sqrt{a(X_t, \theta)}dW_t,$$

  with a and b neural networks with analytic activation functions.

- Sig-SDE models (see e.g. I. Perez Arribas, C. Salvi, L. Szpruch ('20))

$$dX_t = \ell(\widehat{\mathbb{W}}_t)dW_t$$

  with ℓ linear functional of the time extended signature $\widehat{\mathbb{W}}$ of the driving Brownian motion W.

Implications

Modern models can be embedded in the affine and polynomial framework:

- Neural SDEs (see e.g. P. Gierjatowicz, M. Sabate-Vidales, D. Siska, L. Szpruch, Z. Zuric ('20) or C.C, W. Khosrawi, J. Teichmann ('20))

  $$dX_t = b(X_t, \theta)dt + \sqrt{a(X_t, \theta)}dW_t,$$

  with a and b neural networks with analytic activation functions.

- Sig-SDE models (see e.g. I. Perez Arribas, C. Salvi, L. Szpruch ('20))

  $$dX_t = \ell(\widehat{\mathbb{W}}_t)dW_t$$

  with ℓ linear functional of the time extended signature $\widehat{\mathbb{W}}$ of the driving Brownian motion W.

Advantages:

⇒ Universality: all classical models can be arbitrarily well approximated

⇒ Tractability: computation of expected values via affine and polynomial technology; gradients can be provided analytically in optimization tasks

Definition of affine and polynomial processes (1d)

Simplest setting (for illustrative purposes): Itô diffusion in one dimension with state space S, some (bounded or unbounded) interval of $\mathbb{R}$:

$$dX_t = b(X_t)dt + \sqrt{a(X_t)}dB_t, \quad X_0 = x, \qquad (*)$$

with $a : \mathbb{R} \to \mathbb{R}_+$ and $b : \mathbb{R} \to \mathbb{R}$ continuous functions and B a Brownian motion.

Definition of affine and polynomial processes (1d)

Simplest setting (for illustrative purposes): Itô diffusion in one dimension with state space S, some (bounded or unbounded) interval of $\mathbb{R}$:

$$dX_t = b(X_t)dt + \sqrt{a(X_t)}dB_t, \quad X_0 = x, \qquad (*)$$

with $a : \mathbb{R} \to \mathbb{R}_+$ and $b : \mathbb{R} \to \mathbb{R}$ continuous functions and B a Brownian motion.

Definition

A weak solution X of $(*)$ is called polynomial process if

- b is an affine function, i.e. $b(x) = b + \beta x$ for some constants b and β and
- a is a quadratic function, i.e. $a(x) = a + \alpha x + Ax^2$ for some constants a, α and A.

If additionally $A = 0$, then the process is called affine.

- For affine processes we always set $b = a = 0$ and consider only linear ones.
- We denote by $\mathcal{A}$ the infinitesimal generator of a diffusion of form $(*)$, given by $\mathcal{A}f(x) = f'(x)b(x) + \frac{1}{2}f''(x)a(x)$.

Key properties of affine and polynomial processes

From this definition, ...

- ... they appear as narrow class, whose universal character is not visible.
- ... follow some remarkable implications.

Key properties of affine and polynomial processes

From this definition, ...

- ... they appear as narrow class, whose universal character is not visible.
- ... follow some remarkable implications.
  - All marginal moments of a polynomial process, can be computed by solving a system of linear ODEs, i.e.,

$$\mathbb{E}_x[\sum_{i=1}^{k} c_i X_T^i] = \sum_{i=1}^{k} c_i(T)x^i, \text{ with } \partial_t c(t) = L_k c(t), \; c(0) = \underbrace{(c_0, \ldots, c_k)'}_{:=c},$$

where the matrix L_k satisfies $\mathcal{A}(\sum_{i=1}^{k} c_i x^i) = \sum_{i=1}^{k} (L_k c)_i x^i$.

Key properties of affine and polynomial processes

From this definition, ...

- ... they appear as narrow class, whose universal character is not visible.
- ... follow some remarkable implications.
  - All marginal moments of a polynomial process, can be computed by solving a system of linear ODEs, i.e.,

$$\mathbb{E}_x[\sum_{i=1}^{k} c_i X_T^i] = \sum_{i=1}^{k} c_i(T)x^i, \text{ with } \partial_t c(t) = L_k c(t), \; c(0) = \underbrace{(c_0, \ldots, c_k)'}_{:=c},$$

where the matrix L_k satisfies $\mathcal{A}(\sum_{i=1}^{k} c_i x^i) = \sum_{i=1}^{k}(L_k c)_i x^i$.

  - Additionally, exponential moments of affine (linear) processes, can be expressed via solutions of Riccati ODEs, i.e., whenever $\mathbb{E}[|e^{uX_t}|] < \infty$,

$$\mathbb{E}_x[e^{uX_T})] = e^{\psi(T,u)x}, \text{ with } \partial_t \psi(t,u) = R(\psi(t,u)), \; \psi(0,u) = u,$$

where $\mathcal{A}\exp(ux) = R(u)x\exp(ux)$, i.e. $R(u) = \beta u + \frac{1}{2}\alpha u^2$.

Towards universality - signature of a path

Linear characteristics are the key feature of affine processes $\Rightarrow$ lift generic processes via signature methods to linearize their characteristics

Signature (see K. Chen ('57)) is prominent in rough path theory (T. Lyons ('98), P. Friz & N. Victoir ('10), P. Friz & M. Hairer ('14)) due to ...

Towards universality - signature of a path

Linear characteristics are the key feature of affine processes ⇒ lift generic processes via signature methods to linearize their characteristics

Signature (see K. Chen ('57)) is prominent in rough path theory (T. Lyons ('98), P. Friz & N. Victoir ('10), P. Friz & M. Hairer ('14)) due to ...

- The signature of a geometric p-rough path uniquely determines the path up to tree-like equivalences (see H. Boedihardjo et al.('16)).
- Linear functions on the range of signature form an algebra of real-valued functions via the shuffle product.

⇒ Continuous (with respect of to a p-variation distance of the lifted path) path functionals can be approximated by a linear function of the (time extended) signature arbitrarily well. ⇒ Universal approximation theorem (UAT).

- Under certain conditions, the expected signature of a stochastic process determines its law. (see I. Chevyrev & T. Lyons ('16), I.Chevyrev & H. Oberhauser ('18)).

Signature for continuous semimartingales

The signature of a continuous semimartingale X with values in $\mathbb{R}^d$ is defined via iterated Stratonovich integrals.

Definition

Let X be a continuous semimartingale with values in $\mathbb{R}^d$. Then its signature $\mathbb{X}$ at time $T > 0$ is given by

$$\mathbb{X}_T = (1, X_T^{(1)}, \ldots, X_T^{(n)}, \ldots),$$

where for each integer $n \geq 1$,

$$X_T^{(n)} := \int_{0<t_1<\cdots<t_n<T} dX_{t_1} \otimes \circ \cdots \otimes \circ dX_{t_n} \in (\mathbb{R}^d)^{\otimes n}, \quad n \geq 1,$$

it in the sense of the Stratonovich integral.

Tensor algebra and notation

- The signature takes values in the extended tensor algebra $T((\mathbb{R}^d))$ given by

$$T((\mathbb{R}^d)) := \{(a_0, a_1, \dots, a_n, \dots) \,|\, \text{ for all } n \geq 0,\ a_n \in (\mathbb{R}^d)^{\otimes n}\},$$

where by convention $(\mathbb{R}^d)^{\otimes 0} = \mathbb{R}$.

- Generic elements of $T((\mathbb{R}^d))$ are always denoted in bold face, e.g. $\mathbf{a} = (a_0, a_1, \dots, a_n, \dots)$.

- There is a natural pairing between $T((\mathbb{R}^d)^*) = T(\mathbb{R}^d)$ (space of all finite sequences of tensors) and $T((\mathbb{R}^d))$ given by

$$\langle \cdot, \cdot \rangle : T((\mathbb{R}^d)^*) \times T((\mathbb{R}^d)) \to \mathbb{R},$$

and used to denote linear functionals on $T((\mathbb{R}^d))$. We also write $\langle \mathbf{u}, \mathbf{x} \rangle$ when $\mathbf{u} \in T((\mathbb{R}^d))$ (formal dual space).

Coordinate signature

- $\mathcal{I}_d$: set of multi-indexes with entries in $\{1, \ldots, d\}$. The length of an index I is denoted by $|I|$.
- $(e_1, \ldots, e_d)$ is the canonicial basis of $\mathbb{R}^d$ and $(e_{i_1} \otimes \cdots \otimes e_{i_n})_{(i_1,\ldots,i_n)\in\{1,\ldots,d\}^n}$ forms a basis of $(\mathbb{R}^d)^{\otimes n}$ for $n \geq 1$.
- We write $e_I = e_{i_1} \otimes \cdots \otimes e_{i_n}$ for $I = (i_1, \ldots, i_n)$.

Coordinate signature

- $\mathcal{I}_d$: set of multi-indexes with entries in $\{1, \ldots, d\}$. The length of an index I is denoted by $|I|$.
- $(e_1, \ldots, e_d)$ is the canonicial basis of $\mathbb{R}^d$ and $(e_{i_1} \otimes \cdots \otimes e_{i_n})_{(i_1,\ldots,i_n)\in\{1,\ldots,d\}^n}$ forms a basis of $(\mathbb{R}^d)^{\otimes n}$ for $n \geq 1$.
- We write $e_I = e_{i_1} \otimes \cdots \otimes e_{i_n}$ for $I = (i_1, \ldots, i_n)$.

Definition

The coordinate signature indexed by $I = (i_1, \ldots, i_n)$ and denoted by $C_{I,T}(X)$ of a continuous semimartingale X with values in $\mathbb{R}^d$ is defined by

$$C_{I,T}(X) := \int_{0<t_1<\cdots<t_n<T} \circ dX^{i_1}_{t_1} \cdots \circ dX^{i_n}_{t_n},$$

where $\circ$ stands for the Stratonovic integral. Hence,

$$\mathbb{X}_T = 1 + \sum_{n=1}^{\infty} \sum_{|I|=n} C_{I,T}(X) e_I \in T((\mathbb{R}^d)).$$

Shuffle product

- One crucial property is that the product of two linear functions of the signature is a new linear function of the signature which can be made explicit via the shuffle product.
- In other words every polynomial on signature may be realized as a linear function of it, which is a consequence of the following theorem (Ree ('58)).

Shuffle product

- One crucial property is that the product of two linear functions of the signature is a new linear function of the signature which can be made explicit via the shuffle product.
- In other words every polynomial on signature may be realized as a linear function of it, which is a consequence of the following theorem (Ree ('58)).

Theorem

Fix two multi-indices $I = (i_1, \dots, i_n)$ and $J = (j_1, \dots, j_m)$. Then

$$\langle e_I, \mathbb{X}_T \rangle \langle e_J, \mathbb{X}_T \rangle = \langle e_I \sqcup\!\sqcup\ e_J, \mathbb{X}_T \rangle,$$

where the shuffle product $\sqcup\!\sqcup$ is recursively defined as

$$e_I \sqcup\!\sqcup\ e_J = e_{i_1} \otimes ((e_{i_2} \otimes \cdots \otimes e_{i_n}) \sqcup\!\sqcup\ e_J) + e_{j_1} \otimes (e_I \sqcup\!\sqcup\ (e_{j_2} \otimes \cdots \otimes e_{j_m})),$$

with $e_i \sqcup\!\sqcup\ 1 := e_i$ and $1 \sqcup\!\sqcup\ e_i := e_i$.

This will be the crucial property for the universality of affine processes!

Affine processes on the extended tensor algebra

To establish universality properties of affine processes we define them on subspaces of the extended tensor algebra $T((\mathbb{R}^d))$ given by

$$T((\mathbb{R}^d)) := \{(a_0, a_1, \ldots, a_n, \ldots) \mid \text{ for all } n \geq 0,\ a_n \in (\mathbb{R}^d)^{\otimes n}\},$$

- State space $\mathcal{S} \subseteq T((\mathbb{R}^d))$
- $\mathcal{S}^* = \{\mathbf{u} \in T((\mathbb{R}^d)) + \mathrm{i}\,T((\mathbb{R}^d)) \,|\, |\langle \mathbf{u}, \mathbf{x}\rangle| < \infty \text{ for all } \mathbf{x} \in \mathcal{S}\}$
- $\widehat{\mathcal{U}} := \{\mathbf{u} \in T((\mathbb{R}^d)) + \mathrm{i}\,T((\mathbb{R}^d)) \,|\, \mathbf{x} \mapsto |\exp(\langle \mathbf{u}, \mathbf{x}\rangle)| \text{ is bounded on } \mathcal{S}\}$

Definition

We call a linear operator $\mathcal{L}$ of affine type if there exists a distribution determining subset $\mathcal{U} \subseteq \widehat{\mathcal{U}}$ and a map $R : \mathcal{U} \to \mathcal{S}^*, \mathbf{u} \mapsto R(\mathbf{u})$ such that

$$\mathcal{L} \exp(\langle \mathbf{u}, \mathbf{x}\rangle) = \exp(\langle \mathbf{u}, \mathbf{x}\rangle)\langle R(\mathbf{u}), \mathbf{x}\rangle$$

on the family of functions $\{\mathbf{x} \mapsto \exp(\langle \mathbf{u}, \mathbf{x}\rangle) \,|\, \mathbf{u} \in \mathcal{U}\}$.

Affine processes on the tensor algebra space

An $\mathcal{S}$-valued process $(\mathbb{X}_t)_{t\geq 0}$ defined on some filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ is called a solution to the martingale problem for $\mathcal{L}$ if

1. $\mathbb{X}_0 = \mathbf{x}_0$ $\mathbb{P}$-a.s. for some initial value $\mathbf{x}_0 \in \mathcal{S}$,
2. for every $\mathbf{u} \in \mathcal{U}$ there exists a càdlàg version of $(\langle \mathbf{u}, \mathbb{X}_t \rangle)_{t\geq 0}$ and $(\langle R(\mathbf{u}), \mathbb{X}_t \rangle)_{t\geq 0}$ and
3. the process

$$M_t^{\mathbf{u}} := \exp(\langle \mathbf{u}, \mathbb{X}_t \rangle) - \exp(\langle \mathbf{u}, \mathbf{x}_0 \rangle) - \int_0^t \mathcal{L} \exp(\langle \mathbf{u}, \mathbb{X}_s \rangle) ds$$

defines a local martingale for every $\mathbf{u} \in \mathcal{U}$.

Definition

Suppose that $\mathcal{L}$ is of affine type and that the corresponding martingale problem admits a unique solution $(\mathbb{X}_t)_{t\geq 0}$. Then $(\mathbb{X}_t)_{t\geq 0}$ is called $\mathcal{S}$-valued affine process.

Affine transform formula

Theorem (C.C., S. Svaluto-Ferro, J. Teichmann ('21))

Let $(\mathbb{X}_t)_{t\geq 0}$ *be an* $\mathcal{S}$*-valued affine process with initial value* $\mathbf{x}_0$ *Set*

$$g(\mathbf{u},\mathbf{x}) := \sup_{n\in\mathbb{N}} |\langle R(\mathbf{u})^{(n)}, \mathbf{x}^{(n)}\rangle|, \qquad \mathbf{u}\in\mathcal{U}, \mathbf{x}\in\mathcal{S}$$

and suppose that for each $\mathbf{u}\in\mathcal{U}$ *and* $I\in\mathcal{I}_d$

$$\mathbb{E}[\sup_{t\leq T} g(\mathbf{u},\mathbb{X}_t)|\exp(\langle\mathbf{u},\mathbb{X}_t\rangle)|] < \infty, \text{ and } \mathbb{E}[\sup_{t\leq T}(1+|\langle e_I,\mathbb{X}_t\rangle|)|\exp(\langle\mathbf{u},\mathbb{X}_t\rangle)|] < \infty.$$

Then for all $\mathbf{u}\in\mathcal{U}$

$$\mathbb{E}_{\mathbf{x}_0}[\exp(\langle\mathbf{u},\mathbb{X}_T\rangle)] = v(T,\mathbf{u},\mathbf{x}_0),$$

where $v(t,\mathbf{u})$ *is a solution to the following transport equation*

$$\partial_t v(t,\mathbf{u},\mathbf{x}_0) = \langle R(\mathbf{u}), \nabla_{\mathbf{u}} v(t,\mathbf{u},\mathbf{x}_0)\rangle, \quad v(0,\mathbf{u},\mathbf{x}_0) = \exp(\langle\mathbf{u},\mathbf{x}_0\rangle).$$

Affine transform formula

Theorem (cont.)

Suppose that there exists a solution of the tensor algebra valued Riccati equation up to time T with values in $\mathcal{U}$ such that

$$\partial_t \langle \psi(t, \mathbf{u}), \mathbf{x} \rangle = \langle R(\psi(t, \mathbf{u})), \mathbf{x} \rangle, \quad \psi(0, \mathbf{u}) = \mathbf{u}.$$

Then, if $\mathbb{E}[\sup_{s,t \leq T} |\langle R(\psi(s, \mathbf{u})), \mathbb{X}_t \rangle \exp(\langle \psi(s, \mathbf{u}), \mathbb{X}_t \rangle)|] < \infty$ *, it holds that*

$$\mathbb{E}[\exp(\langle \mathbf{u}, \mathbb{X}_T \rangle)] = \exp(\langle \psi(T, \mathbf{u}), \mathbf{x}_0 \rangle).$$

Generic SDEs with path-dependent coefficients (Sig-SDEs)

- Consider a generic class of diffusion type models with state space $S \subseteq \mathbb{R}^{d-1}$ driven by some $d-1$ dimensional Brownian motion B, given by

$$dX_t = \boldsymbol{b}(\widehat{\mathbb{X}}_t)dt + \sqrt{\mathbf{a}(\widehat{\mathbb{X}}_t)}dB_t, \qquad \text{(SigSDE)}$$

  with $(\widehat{\mathbb{X}}_t)_{t \geq 0}$ the signature of $t \mapsto (X_t, t)$ on the state space $\mathcal{S}$.

- Here, $\boldsymbol{b} : \mathcal{S} \to \mathbb{R}^{d-1}$ with $b_i(\boldsymbol{x}) = \langle \boldsymbol{b}_i, \boldsymbol{x} \rangle$ and $\boldsymbol{a} : \mathcal{S} \to \mathbb{S}_+^{d-1}$ with $a_{ij}(\boldsymbol{x}) = \langle \boldsymbol{a}_{ij}, \boldsymbol{x} \rangle$, where $\boldsymbol{b}_i, \boldsymbol{a}_{ij} \in T((\mathbb{R}^d))$.

- Choosing $\boldsymbol{b}$ and $\boldsymbol{a}$ appropriately allows to approximate any continuous path functional arbitrarily well $\Rightarrow$ Truly general class of diffusions whose coefficients can depend on the whole path.

- We suppose that a solution to (SigSDE) exists uniquely on an appropriate state space S.

Sig-SDEs as universal models in finance

- Similar Sig-SDE models for quantitative finance have been considered by I. Perez Arribas, C. Salvi, L. Szpruch ('20).
- They take values in $\mathbb{R}$ and are of the form

$$dS_t = \ell(\widehat{\mathbb{W}}_t)dW_t,$$

where $(\widehat{\mathbb{W}}_t)_{t\geq 0}$ denotes the signature of $t \mapsto (W_t, t)$ with W a one-dimensional Brownian motion and ℓ a linear functional.

- This can be embedded in the above framework by setting $(B^1, B^2) = (W, B^2)$, $a_{11} = 1$, $a_{22}(\mathbf{x}) = (\ell \sqcup\!\sqcup \ell)(\mathbf{w})$ and $a_{12}(\boldsymbol{x}) = a_{21}(\boldsymbol{x}) = \ell(\boldsymbol{w})$ and $\mathbf{b} = 0$ so that $X = (W, S)$, i.e.

$$dX_t = \begin{pmatrix} 1 & 0 \\ \ell(\widehat{\mathbb{W}}_t) & 0 \end{pmatrix} \begin{pmatrix} dW_t \\ dB_t^2 \end{pmatrix}.$$

Sig-SDEs are affine processes

Lemma

Consider the signature process $\widehat{\mathbb{X}}_t$ of $t \mapsto (X_t, t)$ with X given by (SigSDE). *Suppose that for some $\mathcal{U} \subseteq \widehat{\mathcal{U}}$ the map $R : \mathcal{U} \to T((\mathbb{R}^d))$ given by*

$$R(\mathbf{u}) = \sum_{I \in \mathcal{I}_d} \left(\frac{1}{2} (e_{i_1} \otimes \cdots \otimes e_{i_{|I|-2}}) \sqcup\!\sqcup\, \boldsymbol{a}_{i_{|I|-1} i_{|I|}} + (e_{i_1} \otimes \cdots \otimes e_{i_{|I|-1}}) \sqcup\!\sqcup\, \boldsymbol{b}_{i_{|I|}} \right) \mathbf{u}_I + \frac{1}{2} \sum_{I,J \in \mathcal{I}_d} \left((e_{i_1} \otimes \cdots \otimes e_{i_{|I|-1}}) \sqcup\!\sqcup\, (e_{j_1} \otimes \cdots \otimes e_{j_{|J|-1}}) \sqcup\!\sqcup\, \boldsymbol{a}_{i_{|I|} j_{|J|}} \right) \mathbf{u}_I \mathbf{u}_J,$$

satisfies $R(\mathbf{u}) \in \mathcal{S}^*$ for each $\mathbf{u} \in \mathcal{U}$. Fix then $\mathbf{u} \in \mathcal{U}$ and set $\mathcal{L} \exp(\langle \mathbf{u}, \mathbf{x} \rangle) = \exp(\langle \mathbf{u}, \mathbf{x} \rangle) \langle R(\mathbf{u}), \mathbf{x} \rangle$ for each $\mathbf{x} \in \mathcal{S}$. Then

$$\exp(\langle \mathbf{u}, \widehat{\mathbb{X}}_t \rangle) - \exp(\langle \mathbf{u}, 1 \rangle) - \int_0^t \mathcal{L} \exp(\langle \mathbf{u}, \widehat{\mathbb{X}}_s \rangle) ds$$

is a local martingale and $\mathcal{L}$ is of affine type.

Sig-SDEs are affine processes

Corollary

Let X be given by (SigSDE) *and R as of the previous lemma. Suppose there exists a distribution determining set $\mathcal{U} \subseteq \widehat{\mathcal{U}}$ such that $R(\mathcal{U}) \subseteq \mathcal{S}^*$. Then*

- *the signature process $(\widehat{\mathbb{X}}_t)_{t \geq 0}$ of $t \mapsto (X_t, t)$ is an affine process taking values in $T((\mathbb{R}^d))$;*
- *X is the projection of an affine process.*

- Difficulty: determine the set $\mathcal{U}$ and verify the conditions on R, which are needed to guarantee that the affine transform formula holds.
- Generic methodology to obtain representations of the cumulants of X of the form

$$\log \mathbb{E}[\exp(\langle u, X_T \rangle)] = \langle \psi(T, \mathbf{u}), \underbrace{\widehat{\mathbb{X}}_0}_{(1,0,0,\dots)} \rangle, \quad \mathbf{u} = (ux_0, u, 0, \dots),$$

  where ψ solves the extended tensor algebra valued Riccati ODEs.
- Open question: existence of solutions to these ODEs

Sig-SDEs as polynomial processes and expected signature

Note that in this framework affine and polynomial processes coincide, and we can therefore also apply polynomial technology.

Theorem (C.C., S. Svaluto-Ferro, J. Teichmann ('21))

Consider the signature process $\widehat{\mathbb{X}}_t$ *of* $t \mapsto (X_t, t)$ *with* X *given by* (SigSDE). *For* $I = (i_1, \ldots, i_{|I|})$, *define a linear operator* L *by*

$$Le_I = \frac{1}{2}(e_{i_1} \otimes \cdots \otimes e_{i_{|I|-2}}) \sqcup\!\sqcup \, \boldsymbol{a}_{i_{|I|-1} i_{|I|}} + (e_{i_1} \otimes \cdots \otimes e_{i_{|I|-1}}) \sqcup\!\sqcup \, \boldsymbol{b}_{i_{|I|}}.$$

Suppose that the linear ODE $\partial_t \langle \mathbf{c}(t), \mathbf{x} \rangle = \langle L\mathbf{c}(t), \mathbf{x} \rangle$ *with* $\mathbf{c}(0) = e_I$ *admits a solution on* $[0, T]$. *If furthermore* $\mathbb{E}[\sup_{s,t \le T} |\langle \mathbf{c}(t), \widehat{\mathbb{X}}_s \rangle|] < \infty$ *and* $\mathbb{E}[\sup_{s,t \le T} |\langle L\mathbf{c}(t), \widehat{\mathbb{X}}_s \rangle|] < \infty$, *then*

$$\mathbb{E}[\langle e_I, \widehat{\mathbb{X}}_T \rangle] = \langle \mathbf{c}(T), \underbrace{\widehat{\mathbb{X}}_0}_{(1,0,0,\ldots)} \rangle.$$

One dimensional diffusions with analytic characteristics ...

- Consider a one-dimensional diffusion process X on $S \subseteq \mathbb{R}_+$ of the form

$$dX_t = \langle \mathbf{b}, \mathbb{X}_t \rangle dt + \sqrt{\langle \mathbf{a}, \mathbb{X}_t \rangle} dB_t, \quad X_0 = x_0,$$

where $(\mathbb{X}_t)_{t \geq 0}$ denotes its signature (without t part here) and $\mathbf{b}$, $\mathbf{a}$ are such that $\langle \mathbf{b}, \mathbf{x} \rangle < \infty$ and $\langle \mathbf{a}, \mathbf{x} \rangle < \infty$ for all $\mathbf{x} \in S$.

- Since $\mathbb{X}_t = (1, X_t - x_0, \frac{(X_t - x_0)^2}{2}, \ldots, \frac{(X_t - x_0)^n}{n!}, \ldots)$, we can reparametrize and write

$$dX_t = b(X_t)dt + \sqrt{a(X_t)} dB_t, \quad X_0 = x_0, \qquad \text{(SDE - 1d)}$$

where the above conditions translate to b and a being real analytic functions such that

$$b(x) = \sum_{n=0}^{\infty} b_n x^n, \quad a(x) = \sum_{n=0}^{\infty} a_n x^n,$$

converges on an open neighborhood of S.

... are projections of affine processes

Theorem (C.C., S. Svaluto-Ferro, J. Teichmann ('21))

- *Let* X *be specified by* (SDE - 1d) *and let* $\mathcal{A} : \mathcal{D}(\mathcal{A}) \subseteq C_0(S) \to C_0(S)$. *Set* $\mathcal{U} = \{\mathbf{u} = (u_n)_{n \in \mathbb{N}} \,|\, x \mapsto \exp(\sum_{n=0}^{\infty} u_n x^n) \in \mathcal{D}(\mathcal{A})\}$.
- *For fixed* T*, all* $n \in \mathbb{N}_0$ *and* $\mathbf{u} \in \mathcal{U}$, $\mathbb{E}[\sup_{t \leq T} |X_t|^n \exp(\sum_{n=0}^{\infty} u_n X_t^n)] < \infty$.

Then the process $(1, X_t, X_t^2, \ldots, X_t^n, \ldots)$ *is affine with*

$$R_n(\mathbf{u}) = \sum_{k=0}^{n} (n-k+1) b_k u_{n-k+1} + \frac{1}{2} \sum_{k=0}^{n} (n-k+2)(n-k+1) a_k u_{n-k+2}$$
$$+ \frac{1}{2} \sum_{j=0}^{n} \sum_{k=0}^{n-j} (n-j-k+1)(j+1) a_k u_{n-j-k+1} u_{j+1}.$$

Under the additional conditions of the affine transform formula from above

$$\mathbb{E}_{x_0}[\exp(\sum_{n=0}^{\infty} u_n X_t^n)] = \exp(\sum_{n=0}^{\infty} \psi_n(t, \mathbf{u}) x_0^n), \quad \textit{with } \partial_t \psi(t, \mathbf{u}) = R(\psi(t, \mathbf{u})).$$

Relation to polynomial technology

Theorem (C.C., S. Svaluto-Ferro, J.Teichmann ('21))

Let X be specified by (SDE - 1d) *and consider the following infinite matrix*

$$L = \begin{pmatrix} 0 & b_0 & a_0 & 0 & 0 & 0 & \cdots & \cdots \\ 0 & b_1 & a_1 + 2b_0 & 3a_0 & 0 & 0 & \cdots & \cdots \\ 0 & \vdots & & & & & & \\ 0 & b_n & a_n + 2b_{n-1} & 3a_{n-1} + 3b_{n-2} & \cdots & \cdots & \frac{(n+1)(n+2)}{2} a_0 & \cdots \\ 0 & \vdots & & & & & & \end{pmatrix}$$

Suppose that the linear ODE $\partial_t \langle \mathbf{c}(t), \mathbf{x} \rangle = \langle L\mathbf{c}(t), \mathbf{x} \rangle$ *with* $\mathbf{c}(0) = \mathbf{c}$ *admits a solution on* $[0, T]$*. Suppose furthermore that* $\mathbb{E}[\sup_{s,t \leq T} | \sum_{n=0}^{\infty} c_n(t) X_s^n |] < \infty$ *and* $\mathbb{E}[\sup_{s,t \leq T} | \sum_{n=0}^{\infty} (L\mathbf{c}(t))_n X_s^n |] < \infty$*. Then*

$$\mathbb{E}_{x_0}\left[\sum_{n=0}^{\infty} c_n X_T^n \right] = \sum_{n=0}^{\infty} c_n(T) x_0^n.$$

Examples

For the following examples we can e.g. compute the moment generating function

$$\mathbb{E}_{x_0}[\exp(uX_T)] = \sum_{n=0}^{\infty} c_n(T)x_0^n = \exp(\sum_{n=0}^{\infty} \psi_n(T, \mathbf{u})x_0^n)$$

for appropriate u by solving the above infinite dimensional linear ODE with inital value $\mathbf{c} = (1, u, \frac{u}{2}, \ldots, \frac{u^k}{k!}, \ldots)$ or the Riccati equation with inital value $\mathbf{u} = (0, u, 0 \ldots, 0)$.

- Classically non-affine and non-polynomial examples:
  - $dX_t = \sqrt{X_t}(1 - X_t)dB_t$ on $[0, 1]$
  - $dX_t = \kappa \sum_{i=1}^{\infty} \pi(X_t^i - X_t)dt + \sqrt{X_t(1 - X_t)}dB_t$ on $[0, 1]$

Examples

For the following examples we can e.g. compute the moment generating function

$$\mathbb{E}_{x_0}[\exp(uX_T)] = \sum_{n=0}^{\infty} c_n(T)x_0^n = \exp(\sum_{n=0}^{\infty} \psi_n(T, \mathbf{u})x_0^n)$$

for appropriate u by solving the above infinite dimensional linear ODE with inital value $\mathbf{c} = (1, u, \frac{u}{2}, \ldots, \frac{u^k}{k!}, \ldots)$ or the Riccati equation with inital value $\mathbf{u} = (0, u, 0 \ldots, 0)$.

- Classically non-affine and non-polynomial examples:
  - $dX_t = \sqrt{X_t}(1 - X_t)dB_t$ on $[0, 1]$
  - $dX_t = \kappa \sum_{i=1}^{\infty} \pi(X_t^i - X_t)dt + \sqrt{X_t(1 - X_t)}dB_t$ on $[0, 1]$
- Affine Feller diffusion: $dX_t = \sqrt{a_1 X_t}dB_t$ on $\mathbb{R}_+$. For $u < 0$, the solution of the linear ODE leads the well known expression for the Laplace transform

$$\mathbb{E}_{x_0}[\exp(uX_T)] = \sum_{n=0}^{\infty} \underbrace{\frac{u^n}{(1 - \frac{a_1}{2}uT)^n n!}}_{c_n(T)} x_0^n = \exp(\frac{ux_0}{1 - \frac{a_1}{2}uT}).$$

Conclusion

- Generic classes of SDEs can be proved to be affine by lifting them to the signature space where polynomials are linear functionals $\Rightarrow$ one step in the direction of universality of affine processes
- Power series expansions in terms of the initial value for cumulants and moments via affine and polynomial technology
- Develop a theory when the semigroup maps (real) analytic functions to (real) analytic functions together with an analysis of the convergence radii
- Extension to processes with jumps
- Tractability properties for neural SDEs and Sig-SDE models, in particular systematic polynomial way to compute expected signatures

Thank you for your attention!