Universality of affine and polynomial processes

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

University of Vienna

Bachelier Finance Society - One World Seminar

Vienna - Zoom, September 10, 2020

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?**

$\Rightarrow$ **This talk: one step in this direction by linearizing certain classes of SDEs via signature methods**

Objectives of this work - related literature

- Objectives
  - Develop an affine and polynomial theory for prolongations of generic stochastic processes.
  - These prolongations can for instance consist of the set of all monomials or iterated integrals.
  - **Power series expansions for the characteristic function/Laplace transform as well as for the moments** for generic classes of diffusions via affine and polynomial technology

Objectives of this work - related literature

- Objectives
  - Develop an affine and polynomial theory for prolongations of generic stochastic processes.
  - These prolongations can for instance consist of the set of all monomials or iterated integrals.
  - **Power series expansions for the characteristic function/Laplace transform as well as for the moments** for generic classes of diffusions via affine and polynomial technology
- Some 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. Bonnier & H. Oberhauser ('20): "Signature Cumulants, Ordered Partitions, and Independence of Stochastic Processes"

Some implications for mathematical finance

- 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))

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

    where a and b are neural networks depending on parameters θ with entire activation function.
  - **Sig-SDE models** (see I. Perez Arribas, C. Salvi, L. Szpruch ('20))

    $$dX_t = b(\mathbb{B}_t)dt + \sigma(\mathbb{B}_t)dB_t,$$

    where b and σ a linear functionals of the time extended signature $\mathbb{B}$ of Brownian motion.

Some implications for mathematical finance

- 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))

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

    where a and b are neural networks depending on parameters θ with entire activation function.
  - **Sig-SDE models** (see I. Perez Arribas, C. Salvi, L. Szpruch ('20))

    $$dX_t = b(\mathbb{B}_t)dt + \sigma(\mathbb{B}_t)dB_t,$$

    where b and σ a linear functionals of the time extended signature $\mathbb{B}$ of Brownian motion.
- Advantages:
  - universality: dynamics of classical models can be arbitrarily well approximated, while principles like no arbitrage are preserved
  - tractability: pricing, hedging, calibration (gradients can be provided analytically in optimization tasks)

Definition of affine and polynomial processes

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

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.[a]

[a] In this diffusion setting all affine processes are polynomial (in general this only holds true under moment conditions).

Key properties of affine and polynomial processes

From this definition, ...

- ... they appear as a narrow class, whose universal character is at this stage by no means visible.
- ... follow some remarkable implications.
  - All marginal **moments of a polynomial process,** i.e. $\mathbb{E}[X_t^n]$ can be computed by solving a system of **linear ODEs.**
  - Additionally, **exponential moments of affine processes,** i.e. $\mathbb{E}[\exp(uX_t)]$ for some $u \in \mathbb{C}$ can be expressed in terms of solutions of **Riccati ODEs** whenever $\mathbb{E}[|\exp(uX_t)|] < \infty$.

Key properties of affine and polynomial processes

From this definition, ...

- ... they appear as a narrow class, whose universal character is at this stage by no means visible.
- ... follow some remarkable implications.
  - All marginal **moments of a polynomial process,** i.e. $\mathbb{E}[X_t^n]$ can be computed by solving a system of **linear ODEs.**
  - Additionally, **exponential moments of affine processes,** i.e. $\mathbb{E}[\exp(uX_t)]$ for some $u \in \mathbb{C}$ can be expressed in terms of solutions of **Riccati ODEs** whenever $\mathbb{E}[|\exp(uX_t)|] < \infty$.

We here briefly present these implications from the point of view of **dual processes.** This is differs from the original papers

- D. Duffie, D. Filipović & W. Schachermayer ('03); D. Filipović & E. Mayerhofer ('09);
- C., M. Keller-Ressel & J. Teichmann ('12); D. Filipovic & M. Larsson ('16).

Methods to compute expected values

We can distinguish three different ways how to compute $\mathbb{E}_x[f(X_t)]$. Let $\mathcal{A}$ denote the infinitesimal generator of a diffusion of form $(*)$, i.e.
$\mathcal{A}f(x) = f'(x)b(x) + \frac{1}{2}f''(x)a(x)$.

1. Kolmogorov backward equation: $\mathbb{E}_x[f(X_t)] = g(t,x)$, where

$$\partial_t g(t,x) = \mathcal{A}g(t,x), \quad g(0,x) = f(x).$$

2. Duality method: Let $(U_t)_{t\geq 0}$ be an independent Markov process with state space U and infinitesimal generator $\mathcal{B}$. Assume that there is some $f : S \times U \to \mathbb{R}$ such that

$$\mathcal{A}f(\cdot,u)|_x = \mathcal{B}f(x,\cdot)|_u, \quad \text{for all } x \in S,\ u \in U,$$

then (modulo technicalities) $\mathbb{E}_x[f(X_t,u)] = \mathbb{E}_u[f(x,U_t)]$.

3. Kolmogorov backward equation for the dual: $\mathbb{E}_x[f(X_t,u)] = v(t,u,x)$, where

$$\partial_t v(t,u) = \mathcal{B}v(t,u,x), \quad v(0,u,x) = f(x,u).$$

Duality methods for affine and polynomial processes...

- ...work particularly well, because the dual process is deterministic (solution of ODEs) $\Rightarrow$ Tractability
- In the case of polynomial processes, the function family consists of polynomials, in the case of affine processes they are exponentials.
- Since these families are distribution determining (in the polynomial case under exponential moment conditions), existence of the dual process implies uniqueness of (the solution to the martingale problem for) the primal process X.

Polynomial case

- Denote by $\mathcal{P}_k$ polynomials up to degree $k \in \mathbb{N}$, i.e.,

$$\mathcal{P}_k = \left\{ x \mapsto \sum_{i=0}^{k} c_i x^i \mid c_i \in \mathbb{R} \right\}.$$

  - $c = (c_0, \ldots, c_k)^\top \in \mathbb{R}^{k+1}$ denotes the coefficients vector.
  - $\overline{x}$ stands for $(1, x, \ldots, x^k)^\top$ (without indicating the dependence on k).
  - We write $p(x, c) := \langle c, \overline{x} \rangle_k = \sum_{i=0}^k c_i x^i$.
- Note that for a polynomial process, $\mathcal{A}$ **maps $\mathcal{P}_k$ to $\mathcal{P}_k$, i.e. polynomials to polynomials of same or smaller degree**

$\Rightarrow$ Alternative defining property of polynomial processes.

- Dual polynomial operator $\mathcal{B}$: acting on $c \mapsto p(x, c)$ such that $\mathcal{A}p(\cdot, c)|_x = \mathcal{B}p(x, \cdot)|_c$. By the linearity in $\overline{x}$ we can identify $\mathcal{B}$ on $\mathcal{P}_k$ with a linear map L_k from $\mathbb{R}^{k+1}$ to $\mathbb{R}^{k+1}$ such that

$$\mathcal{A}p(\cdot, c)|_x = \langle L_k c, \overline{x} \rangle_k = p(x, L_k c), \quad \text{for all } x \in S.$$

Moment formula

Applying the duality method now to the deterministic dual process $c(t)$, given as a solution to a linear ODE, which describes the evolution of the coefficients vector yields ...

Theorem (C.C., M. Keller-Ressel, J. Teichmann ('12), D. Filipović, M. Larsson ('16))

Let $T > 0$ be fixed and let X be a polynomial process. Denote by $c(t) = (c_0(t), \ldots, c_k(t))^\top$ the solution of the following linear ODE

$$\partial_t c(t) = L_k c(t), \quad c(0) = c \in \mathbb{R}^{k+1}.$$

Then its moments are given by

$$\mathbb{E}_x \left[\sum_{i=0}^{k} c_i X_T^i \right] = \sum_{i=0}^{k} c_i(T) x^i = \langle \exp(L_k T) c, \overline{x} \rangle_k .$$

Affine case

- In the affine case, the family of functions to obtain duality are exponentials.
- For notational convenience we set $b = 0$ and $a = 0$ in the definiiton of the affine process so that we deal with purely linear processes.
- Dual affine operator $\mathcal{B}$: acting on $u \mapsto \exp(ux)$ such that

$$\mathcal{A} \exp(u\cdot)|_x = \mathcal{B} \exp(\cdot x)|_u, \quad x \in S.$$

- To explicitely compute the form of $\mathcal{B}$, define the function

$$R(u) := \frac{1}{2}\alpha u^2 + \beta u.$$

Then, by definition $\mathcal{B} \exp(ux) = \mathcal{A} \exp(ux) = (R(u)x) \exp(ux)$.

- Therefrom we can guess that $\mathcal{B}$ is the restriction of the following transport operator applied to function $g \in C^1(\mathbb{C}, \mathbb{C})$:

$$\mathcal{B}g(u) = R(u)g'(u).$$

Affine transform formula - transport PDE

Applying the third method, i.e. computing the Kolmogorov equation for the dual process, yields...

Theorem (D. Duffie, Filipović, Schachermayer ('03), C.C. and J. Teichmann ('18))

Let $T > 0$ be fixed and let X be an affine process. Let $u \in \mathbb{C}$ such that $\mathbb{E}[|\exp((u)X_T)|] < \infty$. Then,

$$\mathbb{E}_x \left[\exp(uX_T)\right] = v(T, u, x),$$

where $v(t, u, x)$ solves the following linear PDE of transport type

$$\partial_t v(t,u,x) = \mathcal{B}v(t,u,x) = R(u)\partial_u v(t,u,x), \quad v(0,u,x) = \exp(ux), \quad t \in [0,T].$$

Affine transform formula - Riccati ODE

Applying the duality method now to the deterministic dual process $\psi(t,u)$, given as a solution of a Riccati ODE, yields ...

Theorem (cont.)

The unique solution to this transport equation can be expressed by

$$v(t,u,x) = \exp(\psi(t,u)x),$$

where ψ (the dual process here) solves the following Riccati differential equation

$$\partial_t \psi(t,u) = R(\psi(t,u)), \qquad \psi(0,u) = u.$$

Hence, $\mathbb{E}_x\left[\exp(uX_T)\right] = \exp(\psi(T,u)x)$.

Affine transform formula - Riccati ODE

Applying the duality method now to the deterministic dual process $\psi(t, u)$, given as a solution of a Riccati ODE, yields ...

Theorem (cont.)

The unique solution to this transport equation can be expressed by

$$v(t, u, x) = \exp(\psi(t, u)x),$$

where ψ (the dual process here) solves the following Riccati differential equation

$$\partial_t \psi(t, u) = R(\psi(t, u)), \qquad \psi(0, u) = u.$$

Hence, $\mathbb{E}_x [\exp(uX_T)] = \exp(\psi(T, u)x)$.

We have here treated the one-dimensional diffusion setting, mainly to ease notation and technicalities. This is over now ...

Towards universality - signature of a path

- Key feature of affine processes: linear characteristics

$\Rightarrow$ Lift generic processes to linearize their characteristics

Towards universality - signature of a path

- Key feature of affine processes: linear characteristics

$\Rightarrow$ Lift generic processes to linearize their characteristics

- Signature, first studied by K. Chen ('57, '77), plays a prominent role in rough path theory (T. Lyons ('98), P. Friz & M. Hairer ('14)).
- It owes its relevance to the following three key facts:
  - The signature of a path (of bounded p-variation) uniquely determines the path up to tree-like equivalences (including reparametrizations; see H. Boedihardjo, X. Geng, T. Lyons, & D. Yang ('16)).
  - Under certain conditions, the expected signature of a stochastic process determines its law. (see I. Chevyrev & T. Lyons ('16), I.Chevyrev & H. Oberhauser ('18)).
  - **Every continuous path functional can be approximated by a linear function of the time extended signature** arbitrarily well $\Rightarrow$ Universal approximation theorem (UAT).

Definition of signature

The signature of a continuous path X with values in $\mathbb{R}^d$ is defined as a formal power series in d non-commutative indeterminates whose coefficients are iterated integrals of the path.

Definition

Let X be a path of finite p-variation such that the following integration makes sense. Then the signature $\mathbb{X}_T$ of X over the time interval $[0, T]$ is given by

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

where for each integer $n \geq 1$,

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

When X is a path of a continuous semimartingale we shall always define it in the sense of the Stratonovich integral (which is a first order calculus).

Tensor algebra

- The signature is an element of the tensor algebra space $T((\mathbb{R}^d))$ 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. $\mathbf{a} = (a_0, a_1, \dots, a_n, \dots)$.
- Notation:
  - $\mathcal{I}_d$: set of multi-indexes with entries in $\{1, \dots, d\}$. The length of an index I is denoted by $|I|$.
  - $(e_1, \dots, e_d)$ is the canoncial basis of $\mathbb{R}^d$.
  - For any positive integer n, $(e_{i_1} \otimes \cdots \otimes e_{i_n})_{(i_1,\dots,i_n)\in\{1,\dots,d\}^n}$ form a basis of $(\mathbb{R}^d)^{\otimes n}$.

Coordinate signature

Definition

The coordinate signature of X indexed by $I = (i_1, \ldots, i_n)$ denoted by $C_{I,T}(X)$ is defined to be

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

where $\circ$ stands here for a first order calculus, in particular to indicate the Stratonovic integral in the case of a semimartingale. Thus it follows that

$$\mathbb{X}_T = 1 + \sum_{n=1}^{\infty} \sum_{|I|=n} C_{I,T}(X) e_{i_1} \otimes \cdots \otimes e_{i_n} \in T((\mathbb{R}^d)).$$

Example

- Let X be a one-dimensional path of finite variation. Then, for every $n \geq 1$, the iterated integrals are given by

$$C_{\underbrace{(1, \ldots, 1)}_{n \text{ times}}, T}(X) = \frac{(X_T - X_0)^n}{n!}$$

and thus correspond to polynomials. This form translates one to one to semimartingales due to the Stratonovich integral.

- In higher dimension these expressions become more involved. Consider the two dimensional path $t \mapsto (t, B_t)$ for B a standard Brownian motion. Then

$$C_{(1),T} = T, \quad C_{(2),T} = B_T,$$

$$C_{(1,1),T} = \frac{T^2}{2}, C_{(1,2),T} = TB_T - \int_0^T B_t dt, C_{(2,1),T} = \int_0^T B_t dt, C_{(2,2),T} = \frac{B_T^2}{2}$$

$$\ldots,$$

so that we get expressions that depend on the whole path of the Brownian motion.

Linear functionals

- As signature should serve as linear regression basis, we need to introduce linear functionals.
- We denote by $T((\mathbb{R}^d))^*$ the space of linear functionals on $T((\mathbb{R}^d))$ induced by linear combinations of $(e_I)_{I\in\mathcal{I}_d}$, where $e_I = (e_{i_1} \otimes \cdots \otimes e_{i_n})$ (only finitely many are non-zero).
- For any multi-index $I \in \mathcal{I}_d$, we define $\ell_I \in T((\mathbb{R}^d))^*$ via

$$\ell_I(\mathbb{X}_T) = e_I(\mathbb{X}_T) = e_I(X_T^{(|I|)}) = C_{I,T}(X).$$

- Notationwise, we often write $\ell_I(\mathbf{x}) = \langle e_I, \mathbf{x}\rangle$ as well as $\langle \mathbf{u}, \mathbf{x}\rangle$ for $\mathbf{u}$ of the form $\mathbf{u} = \sum_{k\geq 0} u_k e_{I_k}$ (also infinite sums), where $u_k \in \mathbb{R}$ and I_k denotes some multi-index (formal dual space).

Shuffle product

- The crucial and remarkable property is now that the pointwise product of two linear functionals ℓ_I and ℓ_J (which is clearly a quadratic functional) is still a linear functional when restricted to the space of signatures.
- In other words every polynomial on signatures may be realized as a linear functional 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

$$\ell_I(\mathbb{X}_T)\ell_J(\mathbb{X}_T) = (\ell_I \sqcup\!\sqcup\, \ell_J)(\mathbb{X}_T),$$

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 tensor algebra space

- State space $\mathcal{S} \subseteq T((\mathbb{R}^d))$
- $\mathcal{S}^* = \{\mathbf{u} \in T((\mathbb{R}^d)) + \mathrm{i}\,T((\mathbb{R}^d)) \,|\, |\langle \mathbf{u}, \mathbf{x} \rangle| < \infty$ for all $\mathbf{x} \in \mathcal{S}\}$ ($\mathbf{u}$ is not necessarily in $T((\mathbb{R}^d))^* + \mathrm{i}\,T((\mathbb{R}^d))^*$)
- $\widehat{\mathcal{U}} := \{\mathbf{u} \in T((\mathbb{R}^d)) + \mathrm{i}\,T((\mathbb{R}^d)) \,|\, \mathbf{x} \mapsto |\exp(\langle \mathbf{u}, \mathbf{x} \rangle)|$ is bounded on $\mathcal{S}\}$
- $\widehat{\mathcal{U}}^m := \{\mathbf{u} \in T((\mathbb{R}^d)) + \mathrm{i}\,T((\mathbb{R}^d)) \,|\, |\exp(\langle \mathbf{u}, \mathbf{x} \rangle)| \leq m$ for all $\mathbf{x} \in \mathcal{S}\}$

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. $\mathbb{X}_0 = \mathbf{x}$ $\mathbb{P}$-a.s. for some initial value $\mathbf{x} \in \mathcal{S}$,
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

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

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 ('20))

Fix $T > 0$ and let $(\mathbb{X}_t)$ be a $\mathcal{S}$-valued affine process with $\mathbb{E}[\sup_{t\leq T} |\langle e_I, \mathbb{X}_t\rangle|] < \infty$ for all multi-indices I. For fixed $m > 0$, let $\mathcal{U}^m \subseteq \widehat{\mathcal{U}}^m \cap \mathcal{U}$ be such that for all $\mathbf{u} \in \mathcal{U}^m$, there is a function $g(\mathbf{u}, \cdot) : \mathcal{S} \to \mathbb{R}$ with $\mathbb{E}_{\mathbf{x}}[\sup_{t\leq T} g(\mathbf{u}, \mathbb{X}_t)] < \infty$ and $|\langle R^{(n)}(\mathbf{u}), \mathbf{x}^{(n)}\rangle| \leq g(\mathbf{u}, \mathbf{x})$ for all $n \in \mathbb{N}$ and $\mathbf{x} \in \mathcal{S}$. Then for all $\mathbf{u} \in \mathcal{U}^m$

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

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

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

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

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

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

Generic diffusions with path dependent coefficients

- 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

$$dX_{t,i} = \langle \boldsymbol{b}_i, \mathbb{X}_t \rangle dt + \sum_{j=1}^{d-1} \langle \boldsymbol{\sigma}_{ij}, \mathbb{X}_t \rangle dB_{t,j}, \quad i \in \{1, \ldots, d-1\}, \qquad \text{(SigSDE)}$$

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

- Here, $\boldsymbol{b}_i, \boldsymbol{\sigma}_{ij} \in T((\mathbb{R}^d))$, such that $\langle \boldsymbol{b}_i, \mathbf{x} \rangle < \infty$ and $\langle \boldsymbol{\sigma}_{ij}, \mathbf{x} \rangle < \infty$ for all $\mathbf{x} \in \mathcal{S}$, (e.g. $\boldsymbol{b}_i, \boldsymbol{\sigma}_{ij} \in T((\mathbb{R}^d))^*$ or entire functions).
- For Sig-SDE models in finance see I. Perez Arribas, C. Salvi, L. Szpruch ('20)
- Choosing $\boldsymbol{b}$ and $\boldsymbol{\sigma}$ appropriately allows to approximate any continuous path functional arbitrarily well (a consequence of UAT). $\Rightarrow$ **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.

Generic diffusions are (formally) affine processes

Lemma

Consider the signature process $(\mathbb{X}_t)_{t\geq 0}$ of $t \mapsto (X_t, t)$ with X given by (SigSDE). Set $a = \sigma\sigma^\top$ and let $\mathbf{u} = \sum_{k\geq 1} u_k e_{I_k} \in \widehat{\mathcal{U}}$. Define $\mathcal{L}\exp(\langle \mathbf{u}, \mathbf{x}\rangle)$ by

$$
\begin{aligned}
&\mathcal{L}\exp(\langle \mathbf{u}, \mathbf{x}\rangle) \\
&= \Big(\sum_{k\geq 1} \langle \frac{1}{2}(e_{i_1} \otimes \cdots \otimes e_{i_{|I_k|-2}}) ⧢ a_{i_{|I_k|-1} i_{|I_k|}} + (e_{i_1} \otimes \cdots \otimes e_{i_{|I_k|-1}}) ⧢ b_{i_{|I_k|}}, \mathbf{x}\rangle u_k \\
&+ \frac{1}{2} \sum_{k,l\geq 1} \langle (e_{i_1} \otimes \cdots \otimes e_{i_{|I_k|-1}}) ⧢ (e_{j_1} \otimes \cdots \otimes e_{j_{|I_l|-1}}) ⧢ a_{i_{|I_k|}, j_{|I_l|}}, \mathbf{x}\rangle u_k u_l \Big) \exp(\langle \mathbf{u}, \mathbf{x}\rangle) \\
&=: \langle R(\mathbf{u}), \mathbf{x}\rangle \exp(\langle \mathbf{u}, \mathbf{x}\rangle)
\end{aligned}
$$

If $R(\mathbf{u}) \in \mathcal{S}^*$, then $\exp(\langle \mathbf{u}, \mathbb{X}_t\rangle) - \exp(\langle \mathbf{u}, \mathbb{X}_0\rangle) - \int_0^t \mathcal{L}\exp(\langle \mathbf{u}, \mathbb{X}_s\rangle) ds$ is a local martingal and $\mathcal{L}$ is of affine type.

Generic diffusions are (formally) 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 $(\mathbb{X}_t)_{t \geq 0}$ of $t \mapsto (X_t, t)$ is an affine process taking values in $T((\mathbb{R}^d))$. Hence X is the projection of an affine process.*

- The main difficulty is determining the set $\mathcal{U}$ and verifying the conditions on R, which are needed to guarantee that the affine transform formula holds.
- It is a generic methodology, to obtain power series expansions of the logarithm of the characteristic function/Laplace transform with coefficients solving an infinite dimensional Riccati equation. The corresponding convergence radii have to be determined (compare P. Friz, J. Gatheral, R. Radoicic ('20)).
- Back to one-dimensional diffusion processes to illustrate this...

One dimensional diffusions with entire characteristics ...

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

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

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

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

  where the above conditions translate (at least for $S = \mathbb{R}_+$) to b and a being entire functions, i.e.

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

  with infinite convergence radius. We further assume that they are either bounded on $S \subseteq \mathbb{R}_+$ or polynomials.

... are projections of affine processes

Assumption

- *X given by* (SDE - 1d) *is a conservative Feller diffusion with generator* $(\mathcal{A}, \mathcal{D}(\mathcal{A}))$ *s.t.* $\mathbb{E}[\sup_{t \leq T} |X_t|^n] < \infty$ *for all* $n \in \mathbb{N}$ *and fixed* $T > 0$.
- $\mathcal{U} = \{\mathbf{u} = (u_n)_{n \in \mathbb{N}}$ *with* $u_n \leq 0 \,|\, x \mapsto \exp(\sum_{n=0}^{\infty} u_n x^n) \in \mathcal{D}(\mathcal{A}) \subseteq C_0\}$.

Theorem (C.C, S. Svaluto-Ferro, J. Teichmann ('20))

Under the above assumption, the process $(1, X_t, X_t^2, \ldots, X_t^n, \ldots)$ *is affine and*

$$\mathbb{E}_x[\exp(\sum_{n=0}^{\infty} u_n X_t^n)] = v(t, \mathbf{u}, x),$$

where $v(t, u, x)$ *solves the transport PDE* $\partial_t v(t, \mathbf{u}, x) = \sum_{n=0}^{\infty} R_n(\mathbf{u}) \partial_{u_n} v(t, \mathbf{u}, x)$ *with* $v(0, \mathbf{u}, x) = \exp(\sum_{n=0}^{\infty} u_n x^n)$. *If* $x \mapsto \mathbb{E}_x\left[\exp\left(\sum_{n=0}^{\infty} u_n X_t^n\right)\right]$ *is an* ***entire function,*** *then*

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

Classically non affine and non polynomial example

- Fleming Viot type process from population genetics with state space $[0, 1]$ (Spano & Gonazales ('16))

$$dX_t = \sum_{n=1}^{\infty} b_n(X_t^n - X_t^{n+1})dt + \sqrt{X_t(1 - X_t)}dB_t$$

  with b_n appropriate coefficients.

- In this case R is of the following form

$$\begin{aligned} R_n(\mathbf{u}) = &\sum_{k=1}^{n} b_k((n+1-k)u_{n+1-k} - (n-k)u_{n-k}) \\ &+ \frac{1}{2}n((n+1)u_{n+1} - (n-1)u_n) \\ &+ \frac{1}{2}\sum_{k,l \geq 1, k+l=n} k\, l\, u_k u_l + \frac{1}{2}\sum_{k,l \geq 1, k+l=n+1} k\, l\, u_k u_l. \end{aligned}$$

Some remarks and consequences

- Note that in this framework affine and polynomial processes coincide.
- In order to have convergence everywhere, the **semigroup** associated to the diffusion has to **map entire functions to entire functions.** ("Some Hilbert spaces of entire functions" by Louis de Branges ('60) could be useful.)
- In this case polynomial technology reappears:

$$\mathbb{E}[\sum_{n=0}^{\infty} c_n X_T^n] = \langle \exp(LT)\mathbf{c}, \mathbf{x} \rangle,$$

where L is the infinite matrix applied to the infinite coefficients vector such that

$$\mathcal{A}(\sum_{n=0}^{\infty} c_n x^n) = \mathcal{A}\langle \mathbf{c}, \mathbf{x} \rangle = \langle L\mathbf{c}, \mathbf{x} \rangle.$$

- Conjecture: The structure of L is in the case of (SDE - 1d) always of the form such that $\partial_t \mathbf{c}(t) = L\mathbf{c}(t)$ has a solution (Herzog ('98)), but convergence radii of $\sum_{n=0}^{\infty} c_n(t) x^n$ depend on X and $\mathbf{c}$.

Conclusion

- Generic classes of SDEs can be proved to be (formally) 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 for the Laplace transform/characteristic function and moments via affine and polynomial technology
- Develop a theory when the semigroup maps entire functions to entire functions (and also analytic ones)
- Large class of functions which qualify for duality methods
- Further tractability properties for neural SDEs and Sig-SDE models for applications in finance

Thank you for your attention!