# From signature SDEs to affine and polynomial processes and back

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

University of Vienna

WU Wien - Institute for Statistics and Mathematics Vienna - Zoom,

January 22, 2021

# Data driven generative models in finance

- Generative models enter mathematical finance, in particular in view of market generation.
- Consider
  - as input some standard law $\mathbb{P}_I$, e.g. Wiener measure on path space, and
  - $\mathbb{P}_O$ denote some target output law, deduced from (time-series and option) data and not necessarily fully specified.
- A generative model can be viewed as a map $G^\theta$ depending on parameters $\theta$ which transports $\mathbb{P}_I$ to $\mathbb{P}_O$. We denote this push-forward by $G_*^\theta \mathbb{P}_I$.
- The goal is to find $\widehat{\theta}$ such that $G_*^{\widehat{\theta}} \mathbb{P}_I \approx \mathbb{P}_O$ which crucially depends on the parametrization of the transport map $G$.

# Learning the models' characteristics from data

- Highly parametric and overparametrized models gain in importance: instead of a few parameters, the goal is rather to learn the model characteristics as a whole.
- Relying on different universal approximation theorems yields then different classes of models. As an example of such a transport map $G$ we consider here..

# Learning the models' characteristics from data

- Highly parametric and overparametrized models gain in importance: instead of a few parameters, the goal is rather to learn the model characteristics as a whole.
- Relying on different universal approximation theorems yields then different classes of models. As an example of such a transport map $G$ we consider here..

⇒ Signature stochastic differential equations (Sig-SDEs): the model itself or its characteristics are parameterized as linear functions of the signature of a driving signal

- Compare with I. Perez Arribas, C. Salvi, L. Szpruch, 2020 "Sig-SDEs for quatitative finance"

# Part I

## Signature models in finance

based on

ongoing joint work with G. Gazzani and S. Svaluto-Ferro

# Towards signature SDEs in finance

- Consider a stochastic volatility model under a pricing measure $\mathbb{Q}$ of the following form

$$dS_t = \sigma(t, S_t, V_t)dW_t^1,$$
$$dV_t = \kappa^{\mathbb{Q}}(V_t)dt + \nu(V_t)dW_t^2,$$

with

  - $S$ the price process, $V$ the volatility or instantanous variance process,
  - $W^1$ and $W^2$ are correlated $\mathbb{Q}$-Brownian motions and
  - $\sigma, \kappa^{\mathbb{Q}}, \nu$ some functions.

# Towards signature SDEs in finance

- Consider a stochastic volatility model under a pricing measure $\mathbb{Q}$ of the following form

$$dS_t = \sigma(t, S_t, V_t) dW_t^1,$$
$$dV_t = \kappa^{\mathbb{Q}}(V_t) dt + \nu(V_t) dW_t^2,$$

  with
  - $S$ the price process, $V$ the volatility or instantanous variance process,
  - $W^1$ and $W^2$ are correlated $\mathbb{Q}$-Brownian motions and
  - $\sigma, \kappa^{\mathbb{Q}}, \nu$ some functions.
- As well known from the theory of rough paths (T. Lyons '98), solutions of such SDEs can be approximated arbitrarily well by linear maps of the so-called signature process of $t \mapsto (t, W_t^1, W_t^2)$, denoted by $\widehat{\mathbb{W}}$.
- Very briefly, the signature process is the (infinite dimensional tensor algebra valued) process of iterated integrals (in the Stratonovich sense).

# Signature SDEs (Sig-SDEs) in finance

- In other words signature serves as linear regression basis that allows to approximate continuous (with respect to a certain $p$-variation norm for $p \in (2,3)$) path functionals arbitrarily well.
- We can thus approximate the price process via

$$S_t = S_0 + \boldsymbol{\ell}(\widehat{\mathbb{W}}_t) = S_0 + \sum_{0<|I|\leq n} \ell_I \langle e_I, \widehat{\mathbb{W}}_t \rangle,$$

where $\boldsymbol{\ell}$ denotes a linear map (and $\ell_I \in \mathbb{R}$ the corresponding coefficients with respect to the basis elements $e_I$ of the tensor algebra).

# Signature SDEs (Sig-SDEs) in finance

- In other words signature serves as linear regression basis that allows to approximate continuous (with respect to a certain $p$-variation norm for $p \in (2,3)$) path functionals arbitrarily well.
- We can thus approximate the price process via

$$S_t = S_0 + \boldsymbol{\ell}(\widehat{\mathbb{W}}_t) = S_0 + \sum_{0<|I|\leq n} \ell_I \langle e_I, \widehat{\mathbb{W}}_t \rangle,$$

$$= S_0 + \int_0^t \sum_{0<|I|\leq n} \ell_I \langle \widetilde{e}_I, \widehat{\mathbb{W}}_s \rangle dW_s^1, \qquad \text{(Price-Sig-SDE)}$$

  where $\boldsymbol{\ell}$ denotes a linear map (and $\ell_I \in \mathbb{R}$ the corresponding coefficients with respect to the basis elements $e_I$ and $\widetilde{e}_I$ respectively of the tensor algebra).
- Note that this covers also models with path dependent characteristics.
- Similar Sig-SDE models have been considered by I. Perez Arribas, C. Salvi, L. Szpruch ('20).

# Properties of Sig-SDEs

- Advantages:
  - Universality: any classical model (with path dependent charateristics) can be arbitrarly well approximated
  - Classical requirements from mathematical finance, like no arbitrage, can be easily solved, e.g. the price is a martingale due to (Price-Sig-SDE).
  - Appropriate to account for high dimensional option and time series data
  - Tractability: Sig-SDEs are projections of infinite dimensional affine and polynomial processes
- Disadvantages:
  - Parameters are no longer interpretable
  - Robustness of solutions? Ranges of exotic option prices?

# Signature of a path

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 (geometric) rough path uniquely determines the path up to tree-like equivalences (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)).
- Continuous path functional can be approximated by a linear function of the time extended signature arbitrarily well.
  $\Rightarrow$ Universal approximation theorem (UAT).

# Definition of the signature

The signature of a continuous path $X$ with values in $\mathbb{R}^d$ is defined via iterated integrals of the path as follows.

## 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

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

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 $(\mathbb{R}^d)^{\otimes n}$.
  - We write $e_I = e_{i_1} \otimes \cdots \otimes e_{i_n}$ for $I = (i_1, \dots, i_n)$.

# 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

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

Notationwise, we often write for linear functionals $\langle e_I, \mathbf{x} \rangle$ 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).

# Example

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

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

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

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

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

$$\dots,$$

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

# Shuffle product

- The crucial and remarkable property is that the pointwise product of two linear functionals (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)).

# Shuffle product

- The crucial and remarkable property is that the pointwise product of two linear functionals (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*

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

*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 $1 \sqcup\!\sqcup\, e_i := e_i$.*

# Towards data driven Sig-SDEs

- Suppose that we are given the trajectory of the price process $S$ under $\mathbb{P}$ in line with a stochastic volatility model of the form

$$dS_t = \mu(S_t, V_t)dt + S_t\sqrt{V_t}dW_t^{\mathbb{P},1}$$
$$dV_t = \kappa^{\mathbb{P}}(V_t)dt + \nu(V_t)dW_t^{\mathbb{P},2}.$$

  with correlated Brownian motions $(W_t^{\mathbb{P},1}, W_t^{\mathbb{P},2})$.

- Then we can infer from pathwise covariance estimation $t \mapsto V_t$ and in turn $t \mapsto \nu^2(V_t) = \langle V_t, V_t \rangle$.

- From this we can recover estimates of correlated $\mathbb{Q}$-Brownian motions

$$W_t^1 = \int_0^t \frac{\mu(S_s, V_s)}{S_s V_s} dt + W_t^{\mathbb{P},1}$$
$$W_t^2 = \int_0^t \frac{\kappa^{\mathbb{P}}(V_s)}{\nu(V_s)} dt + W_t^{\mathbb{P},2},$$

  which would lead to a stochastic volatility model under $\mathbb{Q}$ with $\mu = \kappa^{\mathbb{Q}} = 0$.

# Calibration to time-series data

- With these Brownian motions obtained from market data, we compute the signature process $\widehat{\mathbb{W}}$.
- The first goal is now to find coefficients $\ell_I$ such that the Sig-SDE model

$$\begin{aligned} S_t &= S_0 + \sum_{0<|I|\leq n} \ell_I \langle e_I, \widehat{\mathbb{W}}_t \rangle \\ &= S_0 + \int_0^t \sum_{0<|I|\leq n} \ell_I \langle \widetilde{e}_I, \widehat{\mathbb{W}}_s \rangle dW_s^1 \\ &= S_0 + \int_0^t \sum_{0<|I|\leq n} \ell_I \langle \widetilde{e}_I, \widehat{\mathbb{W}}_s \rangle \left( \frac{\mu(S_s, V_s)}{S_s V_s} ds + dW_s^{\mathbb{P},1} \right) \end{aligned}$$

matches say $N$ observed market prices $(S_{t_1}^M, \dots, S_{t_N}^M)$.

# Calibration to time-series data

- This means either matching directly the prices

$$\operatorname*{argmin}_{\ell} \sum_{i=1}^{N} \left( \sum_{I} \ell_I \langle e_I, \widehat{\mathbb{W}}_{t_i} \rangle - (S_{t_i}^M - S_0) \right)^2$$

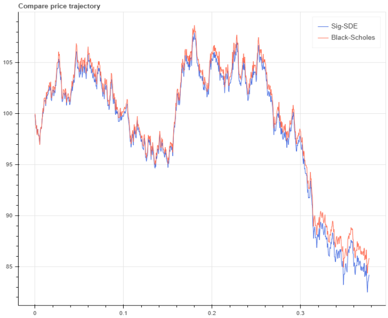
or the volatility

$$\operatorname*{argmin}_{\ell} \sum_{i=1}^{N} \left( \sum_{I} \ell_I \langle \widetilde{e}_I, \widehat{\mathbb{W}}_{t_i} \rangle - (S_{t_i}^M \sqrt{V_{t_i}^M}) \right)^2$$

- In both cases is just a linear regression on the components of the signature.

# First results

- Learn a Black-Scholes market (using the signature computed from the estimated Brownian motion)
- Compare the learned Sig-SDE model with a new Black Scholes trajectory.



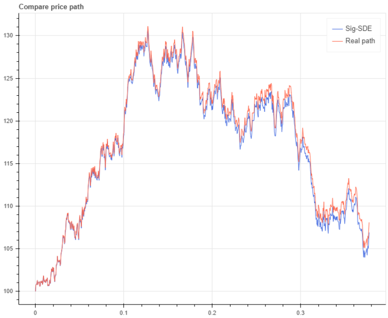




Figure: Out of sample comparison using regression on the price (left) and regression on the volatility (right)

# Towards calibration to option data

- We say that $F$ is a signature payoff if it is a linear function of the signature of $t \mapsto (t, S_t)$, i.e.

$$F(\widehat{\mathbb{S}}_T) = \sum_{|I| \leq m} f_I \langle e_I, \widehat{\mathbb{S}}_T \rangle.$$

- Since linear functions on the signature are dense in the space continuous path functionals, we can approximate any (exotic) payoff (here on $S$) by signature payoffs.
- Asian forwards are for instance signature payoffs.
- The price of a signature payoff is given by $\sum_{|I| \leq m} f_I \langle e_I, \mathbb{E}[\widehat{\mathbb{S}}_T] \rangle$ provided that $\mathbb{E}[\widehat{\mathbb{S}}_T] < \infty$ for all relevant components.

# Pricing of signature payoffs

## Proposition (C.C, G. Gazzani, S.Svaluto-Ferro ('21))

*In a Sig-SDE model of the form*

$$S_t = S_0 + \int_0^t \sum_{0<|I|\leq n} \ell_I \langle \widetilde{e}_I, \widehat{\mathbb{W}}_s \rangle dW_s^1$$

*the price of a signature payoff* $F(\widehat{\mathbb{S}}_T) = \sum_{|I|\leq m} f_I \langle e_I, \widehat{\mathbb{S}}_{\mathbb{T}} \rangle$ *can be expressed as*

$$\mathbb{E}_{\mathbb{Q}}[F(\widehat{\mathbb{S}}_T)] = \sum_J \widehat{f}_J p_J(\ell) \langle \widehat{e}_J, \mathbb{E}_{\mathbb{Q}}[\widehat{\mathbb{W}}_T] \rangle,$$

*where* $p_J(\ell)$ *are polynomials in the coefficients of* $\ell$ *and where* $\widehat{f}_J$ *only depends on* $F$*.*

# Calibration to option data

- The calibration to $N$ options with signature payoffs and market prices $(\pi^1, \ldots, \pi^N)$ can thus be formalized via

$$\operatorname*{argmin}_{\boldsymbol{\ell}} \sum_{i=1}^{N} w^i \left( \sum_J \widehat{f}_J^i p_J^i(\boldsymbol{\ell}) \langle \widehat{e}_J, \mathbb{E}[\widehat{\mathbb{W}}_T] \rangle - \pi^i \right)^2,$$

  where $w^i$ are certain weights.

- Advantages
  - The crucial point is here that $\mathbb{E}[\widehat{\mathbb{W}}_T]$ only needs to be computed once! (No Monte Carlo integration in each optimization step!)
  - This criterion can be easily combined with the time series criterion.
- Disadvantages
  - Approximation of general payoffs is comparable with one-dimensional approximation by polynomials.

# Calibration to option data

- The calibration to $N$ options with signature payoffs and market prices $(\pi^1, \ldots, \pi^N)$ can thus be formalized via

$$\operatorname*{argmin}_{\boldsymbol{\ell}} \sum_{i=1}^{N} w^i \left( \sum_J \widehat{f}^i_J p^i_J(\boldsymbol{\ell}) \langle \widehat{e}_J, \mathbb{E}[\widehat{\mathbb{W}}_T] \rangle - \pi^i \right)^2,$$

  where $w^i$ are certain weights.
- Advantages
  - The crucial point is here that $\mathbb{E}[\widehat{\mathbb{W}}_T]$ only needs to be computed once! (No Monte Carlo integration in each optimization step!)
  - This criterion can be easily combined with the time series criterion.
- Disadvantages
  - Approximation of general payoffs is comparable with one-dimensional approximation by polynomials.
- How can we compute $\mathbb{E}[\widehat{\mathbb{W}}_T]$?

# Calibration to option data

- The calibration to $N$ options with signature payoffs and market prices $(\pi^1, \dots, \pi^N)$ can thus be formalized via

$$\operatorname*{argmin}_{\boldsymbol{\ell}} \sum_{i=1}^{N} w^i \left( \sum_J \widehat{f}_J^i p_J^i(\boldsymbol{\ell}) \langle \widehat{e}_J, \mathbb{E}[\widehat{\mathbb{W}}_T] \rangle - \pi^i \right)^2,$$

  where $w^i$ are certain weights.

- Advantages
    - The crucial point is here that $\mathbb{E}[\widehat{\mathbb{W}}_T]$ only needs to be computed once! (No Monte Carlo integration in each optimization step!)
    - This criterion can be easily combined with the time series criterion.
- Disadvantages
    - Approximation of general payoffs is comparable with one-dimensional approximation by polynomials.
- How can we compute $\mathbb{E}[\widehat{\mathbb{W}}_T]$? see e.g. Friz & Hairer ('14)
- An affine and polynomial process point view helps generically...

# Part II

## An affine and polynomial perspective on signature SDEs

based on

- Universality of affine and polynomial processes (ongoing joint work with S. Svaluto-Ferro and J. Teichmann)
- Infinite dimensional polynomial processes (joint work with S. Svaluto-Ferro)

# 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?

$\Rightarrow$ Can we embed signature SDEs in this framework?

$\Rightarrow$ Method: linearize certain classes of SDEs via signature methods

# 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 $\mathbb{R}$:

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

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

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

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 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 $u \in \mathbb{C}$ 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 $u \in \mathbb{C}$ 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 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 $\mathbb{E}_x[f(X_t)]$.

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

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

# Moment formula for polynomial processes

- For a polynomial of degree $k$ with coefficients vector $c = (c_0, \ldots, c_k) \in \mathbb{R}^{k+1}$ we write $p(x, c) := \langle c, \overline{x} \rangle_k = \sum_{i=0}^{k} c_i x^i$.
- Dual polynomial operator $\mathcal{B}$: acting on $c \mapsto p(x, c)$ s.t. $\mathcal{A}p(\cdot, c)|_x = \mathcal{B}p(x, \cdot)|_c$. We can identify $\mathcal{B}$ with a linear map $L_k$ from $\mathbb{R}^{k+1}$ to $\mathbb{R}^{k+1}$ such that $\mathcal{A}p(\cdot, c)|_x = \langle L_k c, \overline{x} \rangle_k = p(x, L_k c)$ for all $x \in S$.

# Moment formula for polynomial processes

- For a polynomial of degree $k$ with coefficients vector $c = (c_0, \ldots, c_k) \in \mathbb{R}^{k+1}$ we write $p(x, c) := \langle c, \overline{x} \rangle_k = \sum_{i=0}^k c_i x^i$.
- Dual polynomial operator $\mathcal{B}$: acting on $c \mapsto p(x, c)$ s.t. $\mathcal{A}p(\cdot, c)|_x = \mathcal{B}p(x, \cdot)|_c$. We can identify $\mathcal{B}$ with a linear map $L_k$ from $\mathbb{R}^{k+1}$ to $\mathbb{R}^{k+1}$ such that $\mathcal{A}p(\cdot, c)|_x = \langle L_k c, \overline{x} \rangle_k = p(x, L_k c)$ for all $x \in S$.

## Theorem (C.C., M. Keller-Ressel, J. Teichmann ('12))

*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 function family of interest are exponentials.
- For notational convenience we set $b = 0$ and $a = 0$ in the definiton 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 $\mathbb{E}[|\exp(uX_T)|] < \infty$. 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 solution of the Riccati ODE $\partial_t \psi(t, u) = R(\psi(t, u))$, yields ...

**Theorem (cont.)**

*The unique solution to this transport equation can be expressed by*

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

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

# Affine transform formula - Riccati ODE

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

## Theorem (cont.)

*The unique solution to this transport equation can be expressed by*

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

*where $\psi$ (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 ...

# Affine processes on the extended tensor algebra

- 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*

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

# Back to Sig-SDE models

- Consider a generalization of the previous Sig SDE model with state space $S \subseteq \mathbb{R}^{d-1}$ given by

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

  where $B$ is a $d-1$ dimensional standard Brownian motion $B$ and $(\widehat{\mathbb{X}}_t)_{t\geq 0}$ denotes the signature of $t \mapsto (X_t, t)$.

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

$\Rightarrow$ 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 $S$.
- Note the PriceSigSDE model from before can be embedded in this framework by considering the process $X = (W^1, W^2, S)$.

# Sig-SDEs are (formally) 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*

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

*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 (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 $(\widehat{\mathbb{X}}_t)_{t \geq 0}$ 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 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.

# 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). *Fix some multi-index* $I = (i_1, \ldots, i_{|I|})$ *and define an operator* $L$ *(corresponding to the dual polynomial operator) by*

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

*Consider* $\exp(TL)e_I = \sum_{k=0}^{\infty} \frac{T^k}{k!} L^k e_I$ *and suppose that* $\langle \exp(TL)e_I, \underbrace{\widehat{\mathbb{X}}_0}_{(1,0,0,\ldots)} \rangle < \infty$.

*If* $\mathbb{E}[\sup_{s,t \leq T} |\langle \exp(tL)e_I, \widehat{\mathbb{X}_s} \rangle|] < \infty$ *and* $\mathbb{E}[\sup_{s,t \leq T} |\langle L \exp(tL)e_I, \widehat{\mathbb{X}_s} \rangle|] < \infty$, *then*

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

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

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

  where the above conditions translate to $b$ and $a$ being analytic functions, i.e.

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

  converging on an open neighborhood of $S$.

# ... are projections of affine processes

## Assumption

- *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 $\mathbf{u} \in \mathcal{U}$,* $\mathbb{E}[\sup_{t\leq T} |X_t|^n \exp(\sum_{n=0}^{\infty} u_n X_t^n)] < \infty$.

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

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

# ... are projections of affine processes

## Assumption

- *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 $\mathbf{u} \in \mathcal{U}$,* $\mathbb{E}[\sup_{t\leq T} |X_t|^n \exp(\sum_{n=0}^{\infty} u_n X_t^n)] < \infty$.

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

*Under the above assumption, the process $(1, X_t, X_t^2, \ldots, X_t^n, \ldots)$ is affine. By further strenghtening the conditions, the affine transform formula holds*

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

*where $\psi$ solves an sequence valued Riccati equation.*

# ... are projections of affine processes

## Assumption

- *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 $\mathbf{u} \in \mathcal{U}$,* $\mathbb{E}[\sup_{t\leq T} |X_t|^n \exp(\sum_{n=0}^{\infty} u_n X_t^n)] < \infty$.

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

*Under the above assumption, the process* $(1, X_t, X_t^2, \ldots, X_t^n, \ldots)$ *is affine. By further strenghtening the conditions, the affine transform formula holds*

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

*where $\psi$ solves an sequence valued Riccati equation.*

Some recent related literature on expansions of moment generating functions:

- E. Alos, J.Gatheral & R. Radoicic ('20); P. Friz, J. Gatheral & R. Radoicic ('20): "Forests, cumulants, martingales"

# 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* $[0, T]$ *such that* $\langle \mathbf{c}(t), \mathbf{x} \rangle < \infty$ *for every* $x \in S$ *and* $t \in [0, T]$*. Suppose furthermore that* $\mathbb{E}[\sup_{s,t \le T} | \sum_{n=0}^{\infty} c_n(t) X_s^n|] < \infty$ *and* $\mathbb{E}[\sup_{s,t \le T} | \sum_{n=0}^{\infty} (L\mathbf{c}(t))_n X_s^n|] < \infty$*. 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 for instance compute the moment generating function

$$\mathbb{E}_{x_0}[\exp(uX_T)] = \sum c_n(T)x_0^n$$

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

- Classically 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 $u < 0$ we get 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 (formally) affine by lifting them to the signature space where polynomials are linear functionals.
- Power series expansions for the Laplace transform/characteristic function and moments via affine and polynomial technology
- Signature SDEs can be embedded in an affine and polynomial framework which in particular allows to compute expected signature via polynomial technology.

$\Rightarrow$ Sig-SDEs models thus distinguish themselves in

- universality, since the dynamics of all classical models can be arbitrarily well approximated
- efficient pricing, hedging and calibration (through expected signature).

**Thank you for your attention!**