Signature SDEs as affine and polynomial processes

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

University of Vienna

Winter Seminar on Mathematical Finance

Online, January 24, 2024

Data driven models in finance

- Highly parametric and overparametrized models gain in importance: instead of a few parameters, learn the model's characteristics as a whole from data.
- Relying on different universal approximation theorems yields different classes of models. We consider here ...

Data driven models in finance

- Highly parametric and overparametrized models gain in importance: instead of a few parameters, learn the model's characteristics as a whole from data.
- Relying on different universal approximation theorems yields different classes of models. We consider here ...

⇒ Signature based models: the model itself or its characteristics are parameterized as linear functions of the signature of a primary underlying process, e.g.
  - a classical driving signal, e.g. Brownian motions or Lévy processes;
  - more general tractable processes describing well observable quantities.

- Compare e.g. with
  - I. Perez Arribas, C. Salvi, L. Szpruch "Sig-SDEs for quantitative finance"
  - T. Lyons, S. Nejad and I. Perez Arribas "Nonparametric pricing and hedging of exotic derivatives"

Data driven models in finance

- Highly parametric and overparametrized models gain in importance: instead of a few parameters, learn the model's characteristics as a whole from data.
- Relying on different universal approximation theorems yields different classes of models. We consider here ...

⇒ Signature based models: the model itself or its characteristics are parameterized as linear functions of the signature of a primary underlying process, e.g.
  - a classical driving signal, e.g. Brownian motions or Lévy processes;
  - more general tractable processes describing well observable quantities.

- Compare e.g. with
  - I. Perez Arribas, C. Salvi, L. Szpruch "Sig-SDEs for quantitative finance"
  - T. Lyons, S. Nejad and I. Perez Arribas "Nonparametric pricing and hedging of exotic derivatives"
- Other recent applications of signature methods in finance: Bayer et al. ('21) Bühler et al. ('20), Ni et al. ('20)), etc.

Part I

Signature based models for finance - theory and calibration

mainly based on

- "Signature based models for finance - theory and calibration" (joint work in progress with Guido Gazzani and Sara Svaluto-Ferro)

Key facts about signature - Universal approximation

Signature (see K. Chen ('57)) plays a prominent role in rough path theory (T. Lyons ('98), P. Friz & N. Victoir ('10), P. Friz & M. Hairer ('14)). It serves as linear regression basis for continuous path functionals.

Key facts about signature - Universal approximation

Signature (see K. Chen ('57)) plays a prominent role in rough path theory (T. Lyons ('98), P. Friz & N. Victoir ('10), P. Friz & M. Hairer ('14)). It serves as linear regression basis for continuous path functionals.

- Signature is point-separating:
  - The signature of a d-dimensional geometric rough path, including continuous semimartingales, uniquely determines the path up to tree-like equivalences (see H. Boedihardjo et al.('16)).
  - These tree-like equivalences can be avoided by adding time.

Key facts about signature - Universal approximation

Signature (see K. Chen ('57)) plays a prominent role in rough path theory (T. Lyons ('98), P. Friz & N. Victoir ('10), P. Friz & M. Hairer ('14)). It serves as linear regression basis for continuous path functionals.

- Signature is point-separating:
  - The signature of a d-dimensional geometric rough path, including continuous semimartingales, uniquely determines the path up to tree-like equivalences (see H. Boedihardjo et al.('16)).
  - These tree-like equivalences can be avoided by adding time.
- Linear functions on the signature form an algebra that contains 1:
  - Every polynomial on signature may be realized as a linear function via the so-called shuffle product $\sqcup\!\sqcup$.

Key facts about signature - Universal approximation

Signature (see K. Chen ('57)) plays a prominent role in rough path theory (T. Lyons ('98), P. Friz & N. Victoir ('10), P. Friz & M. Hairer ('14)). It serves as linear regression basis for continuous path functionals.

- Signature is point-separating:
  - The signature of a d-dimensional geometric rough path, including continuous semimartingales, uniquely determines the path up to tree-like equivalences (see H. Boedihardjo et al.('16)).
  - These tree-like equivalences can be avoided by adding time.
- Linear functions on the signature form an algebra that contains 1:
  - Every polynomial on signature may be realized as a linear function via the so-called shuffle product $\sqcup\!\sqcup$.

$\Rightarrow$ By the Stone-Weierstrass theorem continuous (with respect of to a variation distance of the lifted path) path functionals on compact sets can be uniformly approximated by a linear function of the time extended signature.

$\Rightarrow$ Universal approximation theorem (UAT).

Signature in a nutshell

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

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

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

Signature in a nutshell

- The signature takes values in the extended tensor algebra $T((\mathbb{R}^d))$ given by
$$T((\mathbb{R}^d)) := \{(a_0, a_1, \ldots, a_n, \ldots) \mid \text{ for all } n \geq 0,\ a_n \in (\mathbb{R}^d)^{\otimes n}\}.$$
Elements of $T((\mathbb{R}^d))$ are denoted in bold face, e.g. $\mathbf{a} = (a_0, a_1, \ldots, a_n, \ldots)$.
- Let $I = (i_1, \ldots, i_n)$ be a multi-index with entries in $\{1, \ldots, d\}$ and denote by $e_I = e_{i_1} \otimes \cdots \otimes e_{i_n}$ the basis elements of $(\mathbb{R}^d)^{\otimes n}$.
- We write $\langle e_I, \boldsymbol{a} \rangle$ to extract the I^{th} component from a_n. More generally we often write $\boldsymbol{u}(\boldsymbol{x}) = \langle \boldsymbol{u}, \boldsymbol{x} \rangle$ if $\sum_I |u_I x_I| < \infty$ and call this linear maps in $\boldsymbol{x}$.

Signature in a nutshell

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

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

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

- Let $I = (i_1, \dots, i_n)$ be a multi-index with entries in $\{1, \dots, d\}$ and denote by $e_I = e_{i_1} \otimes \dots \otimes e_{i_n}$ the basis elements of $(\mathbb{R}^d)^{\otimes n}$.
- We write $\langle e_I, \boldsymbol{a} \rangle$ to extract the I^{th} component from a_n. More generally we often write $\boldsymbol{u}(\boldsymbol{x}) = \langle \boldsymbol{u}, \boldsymbol{x} \rangle$ if $\sum_I |u_I x_I| < \infty$ and call this linear maps in $\boldsymbol{x}$.

Definition

The coordinate signature indexed by $I = (i_1, \dots, i_n)$ of an $\mathbb{R}^d$-valued semimartingale $\widehat{X}$ is defined via iterated Stratonovic/Marcus integrals (denoted by $\circ$)

$$\langle e_I, \widehat{\mathbb{X}}_T \rangle := \int_{0 < t_1 < \dots < t_n < T} \circ d\widehat{X}^{i_1}_{t_1} \cdots \circ d\widehat{X}^{i_n}_{t_n},$$

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

Model ingredients

- Goal: provide a good model for a set of traded assets $S = (S^1, \dots, S^m)$

Model ingredients

- Goal: provide a data-driven, universal, tractable and easy to calibrate model for a set of traded assets $S = (S^1, \ldots, S^m)$

Model ingredients

- Goal: provide a data-driven, universal, tractable and easy to calibrate model for a set of traded assets $S = (S^1, \dots, S^m)$
- Main ingredient: market's primary (underlying) process $\widehat{X}_t = (X_t, t)$, where X is a (continuous, multivariate) Itô-semimartingale such that its signature $\widehat{\mathbb{X}}$ serves a linear regression basis for S.

Model ingredients

- Goal: provide a data-driven, universal, tractable and easy to calibrate model for a set of traded assets $S = (S^1, \dots, S^m)$
- Main ingredient: market's primary (underlying) process $\widehat{X}_t = (X_t, t)$, where X is a (continuous, multivariate) Itô-semimartingale such that its signature $\widehat{\mathbb{X}}$ serves a linear regression basis for S.
- Examples for X:
  - Market inferred Brownian motion (under $\mathbb{Q}$, i.e. $\mathbb{P}$-BM + market price of risk) or Lévy process
  - assets whose dynamics are well understood and which can be described by so-called Sig-SDEs of the form

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

    where B is a Brownian motion, $(\widehat{\mathbb{X}}_t)_{t\geq 0}$ denotes the signature of $t \mapsto (X_t, t)$ and $\boldsymbol{b}$ and $\boldsymbol{a}$ are linear maps.

$\Rightarrow$ Truly general class of diffusions whose coefficients can depend on the whole path.

Model framework

The model

The traded assets $(S^1, \ldots, S^m)$ *are modeled via*

$$S^j_t(\ell^j) = \ell^j(\widehat{\mathbb{X}}_t) = \ell^j_0 + \sum_{0<|I|\leq n} \ell^j_I \langle e_I, \widehat{\mathbb{X}}_t \rangle, \quad j = 1, \ldots, m, \qquad \text{(Sig-model)}$$

- $\widehat{\mathbb{X}}$ *is the signature of* $\widehat{X}$,
- $n \in \mathbb{N}$ *is degree of truncation*
- ℓ^j *denotes a linear map. Here,* $\ell^j_0, \ell^j_I \in \mathbb{R}$ *the corresponding coefficients with respect to the basis elements* 1 *and* e_I *of the tensor algebra.*

The parameters ℓ can also be taken time-dependent.

Properties of signature based models

Universality Any classical model driven by Brownian motion can be arbitrarly well approximated; extensions to Lévy/Poisson random measure driven models are possible.

This is because the solution map of an SDE is a continuous map of the signature of the driving signal.

Properties of signature based models

Universality Any classical model driven by Brownian motion can be arbitrarly well approximated; extensions to Lévy/Poisson random measure driven models are possible.

This is because the solution map of an SDE is a continuous map of the signature of the driving signal.

No arbitrage ... can be easily guaranteed.

The model can also be expressed in terms of stochastic integrals with respect to local martingales, from which conditions for no-arbitrage can be easily deduced.

Properties of signature based models

Universality Any classical model driven by Brownian motion can be arbitrarly well approximated; extensions to Lévy/Poisson random measure driven models are possible.

This is because the solution map of an SDE is a continuous map of the signature of the driving signal.

No arbitrage ... can be easily guaranteed.

The model can also be expressed in terms of stochastic integrals with respect to local martingales, from which conditions for no-arbitrage can be easily deduced.

Calibration by regression when calibrating to time series data.

Properties of signature based models

Universality Any classical model driven by Brownian motion can be arbitrarly well approximated; extensions to Lévy/Poisson random measure driven models are possible.

This is because the solution map of an SDE is a continuous map of the signature of the driving signal.

No arbitrage ... can be easily guaranteed.

The model can also be expressed in terms of stochastic integrals with respect to local martingales, from which conditions for no-arbitrage can be easily deduced.

Calibration by regression when calibrating to time series data.

Calibration to options via precomputed Monte-Carlo samples of $\widehat{\mathbb{X}}$ exploiting linearity of the model

Properties of signature based models

Universality Any classical model driven by Brownian motion can be arbitrarly well approximated; extensions to Lévy/Poisson random measure driven models are possible.

This is because the solution map of an SDE is a continuous map of the signature of the driving signal.

No arbitrage ... can be easily guaranteed.

The model can also be expressed in terms of stochastic integrals with respect to local martingales, from which conditions for no-arbitrage can be easily deduced.

Calibration by regression when calibrating to time series data.

Calibration to options via precomputed Monte-Carlo samples of $\widehat{\mathbb{X}}$ exploiting linearity of the model

Tractable option pricing formulas relying on approximations via so-called signature payoffs of the form $\langle e_J, \widehat{\mathbb{S}}_T(\ell)\rangle$.

Generic primary processes $\widehat{X}$ of form (SigSDE) are projections of a extended tensor algebra valued affine and polynomial process.

$\Rightarrow \mathbb{E}_{\mathbb{Q}}[\widehat{\mathbb{S}}_T(\ell)]$ can thus be computed via polynomial technology, i.e. by solving a (usually infinite dimensional) linear ODE.

Calibration to time-series data

- Goal: match N market prices $(S^M_{t_1}, \dots, S^M_{t_N})$
- Assumption: Time series data $\widehat{X}_{t_1}, \dots, \widehat{X}_{t_N}$ is available.
- Procedure:
  - Compute from $\widehat{X}_{t_1}, \dots, \widehat{X}_{t_N}$ the path of the signature $\widehat{\mathbb{X}}$.
  - Use the path of $\widehat{\mathbb{X}}$ as linear regression basis to find $\boldsymbol{\ell}$ by matching the prices, i.e.

$$\operatorname*{argmin}_{\boldsymbol{\ell}} \sum_{i=1}^{N} \left(S_0 + \sum_{k=1}^{d} \sum_{0 \leq |I| \leq n} \ell_{I,k} \langle e_I, \widehat{\mathbb{X}}_{t_i} \rangle - S^M_{t_i} \right)^2$$

  - Important: since the dimension of $\boldsymbol{\ell}$ is typically high, introducing a regularization (Lasso, Ridge) is necessary.

Results for a Heston and a SABR market model

- Learn a Heston and a SABR market using the signature of estimated $\mathbb{Q}$-Brownian motions (i.e. $\mathbb{P}$-BM with market price of risk)
- Compare the learned (Sig-model) with new Heston/SABR trajectories.

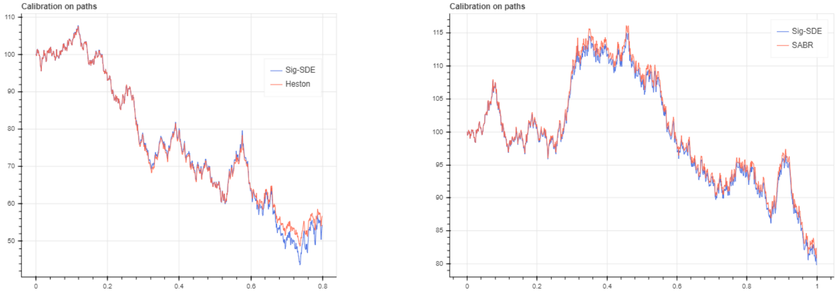
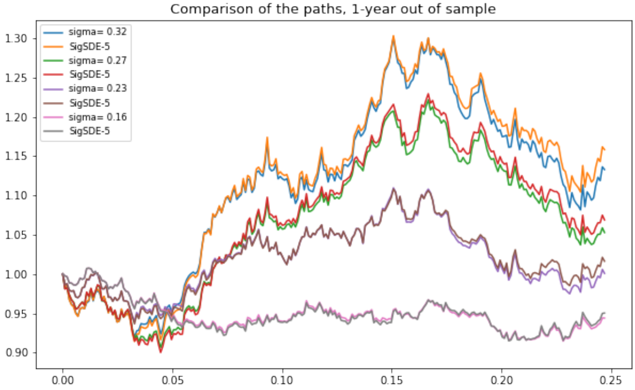


Figure: Out of sample comparison of the price trajectories

Multivariate Case - 4 dimensional correlated B&S model

- Learn a 4 dimensional Black and Scholes market using the signature of $\widehat{\mathbb{W}}$ up to order 5.
- Compare the learned (Sig-model) with new trajectories.



Calibration to options

- Goal: match N option prices $(\pi^1, \dots, \pi^N)$ with payoffs $F_i(S_{T_i}), i = 1, \dots, N$.
- Typically we calibrate to call and put options with different strikes and maturities, whose prices are expressed in terms of implied volatility.

Calibration to options

- Goal: match N option prices $(\pi^1, \ldots, \pi^N)$ with payoffs $F_i(S_{T_i}), i = 1, \ldots, N$.
- Typically we calibrate to call and put options with different strikes and maturities, whose prices are expressed in terms of implied volatility.
- Assumption: Monte-Carlo samples of $\widehat{\mathbb{X}}^j_{T_1}, \ldots, \widehat{\mathbb{X}}^j_{T_N}$ for $j = 1, \ldots M$ (under a pricing measure $\mathbb{Q}$) are available.
- Procedure: The calibration can then be formalized via

$$\operatorname*{argmin}_{\ell} \sum_{i=1}^{N} w^i \left(\frac{1}{M} \sum_{j=1}^{M} F_i(\ell(\widehat{\mathbb{X}}^j_{T_i})) - \pi_i \right)^2,$$

where w^i are weights, e.g. vega-weights to match implied volatility well.

Calibration to options

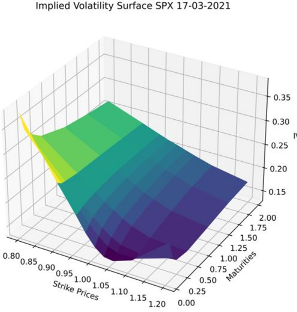
- Goal: match N option prices $(\pi^1, \dots, \pi^N)$ with payoffs $F_i(S_{T_i}), i = 1, \dots, N$.
- Typically we calibrate to call and put options with different strikes and maturities, whose prices are expressed in terms of implied volatility.
- Assumption: Monte-Carlo samples of $\widehat{\mathbb{X}}^j_{T_1}, \dots, \widehat{\mathbb{X}}^j_{T_N}$ for $j = 1, \dots M$ (under a pricing measure $\mathbb{Q}$) are available.
- Procedure: The calibration can then be formalized via

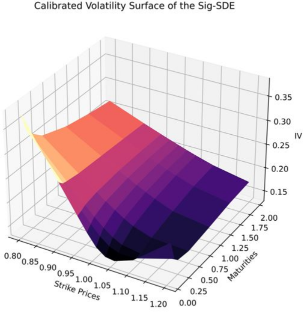
$$\operatorname*{argmin}_{\boldsymbol{\ell}} \sum_{i=1}^{N} w^i \left(\frac{1}{M} \sum_{j=1}^{M} F_i(\boldsymbol{\ell}(\widehat{\mathbb{X}}^j_{T_i})) - \pi_i \right)^2 ,$$

  where w^i are weights, e.g. vega-weights to match implied volatility well.
- Advantages:
  - All Monte-Carlo samples can be easily precomputed, no Monte-Carlo simulation in each optimization step!
  - For ℓ such that $\frac{1}{M} \sum_{j=1}^{M} F_i(\ell(\widehat{\mathbb{X}}^j_{T_i})) \geq \pi_i$ the optimization is convex for convex payoffs.

Calibration to real data

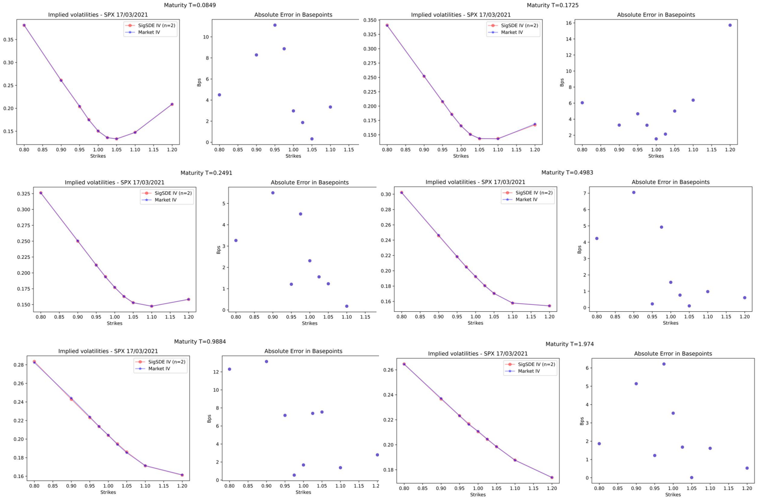
- Goal: reproduce S&P 500 volatility surface (here from 17-03-21) using as primary underlying process $\widehat{X}_t = (W_t^1, W_t^2, t)$.
- We consider here an extension of the model with time-dependent parameters and achieve a nearly perfect fit.





Calibration to real data - error analysis

- With 13 parameters per maturity the absolute error is in the range of 0 to 15 basis points.



Pricing of signature payoffs

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

The price of a signature payoff $\langle e_J, \widehat{\mathbb{S}}_{\mathbb{T}}(\boldsymbol{\ell})\rangle$ *can be expressed as*

$$\mathbb{E}_{\mathbb{Q}}[\langle e_J, \widehat{\mathbb{S}}_T(\boldsymbol{\ell})\rangle] = \langle e(J,\boldsymbol{\ell}), \mathbb{E}_{\mathbb{Q}}[\widehat{\mathbb{X}}_T]\rangle = \sum_I p_I(J,\boldsymbol{\ell})\langle e_I, \mathbb{E}_{\mathbb{Q}}[\widehat{\mathbb{X}}_T]\rangle,$$

where $p_I(J,\boldsymbol{\ell})$ *are polynomials in the coefficients of* $\boldsymbol{\ell}$*.*

- Path-dependent payoffs like Asian forwards correspond to signature payoffs.

Pricing of signature payoffs

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

The price of a signature payoff $\langle e_J, \widehat{\mathbb{S}}_{\mathbb{T}}(\boldsymbol{\ell})\rangle$ *can be expressed as*

$$\mathbb{E}_{\mathbb{Q}}[\langle e_J, \widehat{\mathbb{S}}_T(\boldsymbol{\ell})\rangle] = \langle e(J,\boldsymbol{\ell}), \mathbb{E}_{\mathbb{Q}}[\widehat{\mathbb{X}}_T]\rangle = \sum_I p_I(J,\boldsymbol{\ell})\langle e_I, \mathbb{E}_{\mathbb{Q}}[\widehat{\mathbb{X}}_T]\rangle,$$

where $p_I(J,\boldsymbol{\ell})$ *are polynomials in the coefficients of* $\boldsymbol{\ell}$*.*

- Path-dependent payoffs like Asian forwards correspond to signature payoffs.
- How can we compute $\mathbb{E}_{\mathbb{Q}}[\widehat{\mathbb{X}}_T]$?

Pricing of signature payoffs

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

The price of a signature payoff $\langle e_J, \widehat{\mathbb{S}}_{\mathbb{T}}(\boldsymbol{\ell})\rangle$ *can be expressed as*

$$\mathbb{E}_{\mathbb{Q}}[\langle e_J, \widehat{\mathbb{S}}_T(\boldsymbol{\ell})\rangle] = \langle e(J,\boldsymbol{\ell}), \mathbb{E}_{\mathbb{Q}}[\widehat{\mathbb{X}}_T]\rangle = \sum_I p_I(J,\boldsymbol{\ell})\langle e_I, \mathbb{E}_{\mathbb{Q}}[\widehat{\mathbb{X}}_T]\rangle,$$

where $p_I(J,\boldsymbol{\ell})$ *are polynomials in the coefficients of* $\boldsymbol{\ell}$.

- Path-dependent payoffs like Asian forwards correspond to signature payoffs.
- How can we compute $\mathbb{E}_{\mathbb{Q}}[\widehat{\mathbb{X}}_T]$?
  - Case of $\widehat{\mathbb{X}} = \widehat{\mathbb{W}}$: T. Fawcett ('04)
  - Semimartingale case: P. Friz, P. Hager & N. Tapia ('21)
  - An affine and polynomial process point view works for generic primary processes $\widehat{X}$ of form (SigSDE).
    $\Rightarrow \mathbb{E}_{\mathbb{Q}}[\widehat{\mathbb{X}}_T]$ can be computed via polynomial technology, i.e. by solving (usually infinite dimensional) linear ODEs.

Part II

An affine and polynomial perspective to signature based models

based on

- Universality of affine and polynomial processes (joint work in progress with S. Svaluto-Ferro and J. Teichmann)
- Signature based affine and polynomial jump diffusions (joint work in progress with F. Primavera and 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.

$\Rightarrow$ Universal model classes?

$\Rightarrow$ Mathematically precise statements for this universality?

$\Rightarrow$ Can we embed signature based models in this framework?

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.
- ... 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, i.e. the Feynman-Kac PDE reduces to a linear ODE.
  - 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 $\mathbb{E}[|\exp(uX_t)|] < \infty$, i.e. the Cole-Hopf transform of the Feynman-Kac PDE reduces to a Riccati ODE.

Key properties of affine and polynomial processes

From this definition, ...

- ... they appear as a narrow class.
- ... 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, i.e. the Feynman-Kac PDE reduces to a linear ODE.
  - 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 $\mathbb{E}[|\exp(uX_t)|] < \infty$, i.e. the Cole-Hopf transform of the Feynman-Kac PDE reduces to a Riccati ODE.

For much more generality and details about these processes classes we refer to

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

Linear processes

- We consider here an even simpler subset of affine (hence polynomial) processes where $b(x) = \beta x$ and $a(x) = \alpha x$ are just linear functions and call the corresponding stochastic processes linear processes.
- Their infinitesimal generator if given by $\mathcal{A}f(x) = f'(x)\beta x + \frac{1}{2}f''(x)\alpha x$ and acts on $f(x) = \exp(ux)$ and polynomials $f(x) = \sum_{i=0}^{k} c_i x^i$ as follows

Linear processes

- We consider here an even simpler subset of affine (hence polynomial) processes where $b(x) = \beta x$ and $a(x) = \alpha x$ are just linear functions and call the corresponding stochastic processes linear processes.
- Their infinitesimal generator if given by $\mathcal{A}f(x) = f'(x)\beta x + \frac{1}{2}f''(x)\alpha x$ and acts on $f(x) = \exp(ux)$ and polynomials $f(x) = \sum_{i=0}^{k} c_i x^i$ as follows
  - Affine property: $\mathcal{A}\exp(ux) = \exp(ux)R(u), \quad R(u) = \frac{1}{2}\alpha u^2 + \beta u$

Linear processes

- We consider here an even simpler subset of affine (hence polynomial) processes where $b(x) = \beta x$ and $a(x) = \alpha x$ are just linear functions and call the corresponding stochastic processes linear processes.
- Their infinitesimal generator if given by $\mathcal{A}f(x) = f'(x)\beta x + \frac{1}{2}f''(x)\alpha x$ and acts on $f(x) = \exp(ux)$ and polynomials $f(x) = \sum_{i=0}^k c_i x^i$ as follows
  - Affine property: $\mathcal{A}\exp(ux) = \exp(ux)R(u), \quad R(u) = \frac{1}{2}\alpha u^2 + \beta u$
  - Polynomial property: $\mathcal{A}\left(\sum_{i=0}^k c_i x^i\right) = \sum_{i=1}^k (ic_i\beta + \frac{1}{2}(i+1)ic_{i+1}\alpha 1_{\{i \leq k+1\}})x^i = \sum_{i=1}^k L(c)_i x^i$ with matrix L applied to the vector $c = (c_0, c_1, \ldots, c_k)^\top$

$$L = \begin{pmatrix} 0 & \cdots & & & & & 0 \\ 0 & \beta & \alpha & 0 & \cdots & & 0 \\ 0 & 0 & 2\beta & 3\alpha & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ & & & & i\beta & \frac{(i+1)i}{2}\alpha & \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

Affine transform formula and moment formula

Theorem (Duffie, Filipovic, Schachermayer ('03), C.C., Keller-Ressel, Teichmann ('12))

Let $T > 0$ be fixed and let X be a linear process.

- *Let $u \in \mathbb{C}$ such that $\mathbb{E}[|\exp(uX_T)|] < \infty$ and denote by $\psi(t)$ the solution of the following Riccati ODE*

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

  Then $\mathbb{E}\left[\exp(uX_T)\right] = \exp(\psi(t)X_0)$

Affine transform formula and moment formula

Theorem (Duffie, Filipovic, Schachermayer ('03), C.C., Keller-Ressel, Teichmann ('12))

Let $T > 0$ be fixed and let X be a linear process.

- *Let $u \in \mathbb{C}$ such that $\mathbb{E}[|\exp(uX_T)|] < \infty$ and denote by $\psi(t)$ the solution of the following Riccati ODE*

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

  Then $\mathbb{E}\left[\exp(uX_T)\right] = \exp(\psi(t) X_0)$

- *Denote by $c(t) = (c_0(t), \ldots, c_k(t))^\top$ the solution of the following linear ODE*

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

  Then its moments are given by
  $\mathbb{E}\left[\sum_{i=0}^k c_i X_T^i\right] = \sum_{i=0}^k c_i(T) x^i = \sum_{i=0}^k (\exp(LT)c)_i X_0^i.$

One dimensional diffusions with analytic characteristics

- Consider a one-dimensional diffusion process X on S given by

$$dX_t = b(X_t)dt + \sqrt{a(X_t)}dB_t, \quad X_0 = x_0,$$

where the functions b and a are real analytic of the form

$$b(x) = \sum_{n=0}^{\infty} b_n x^n \qquad \text{and} \qquad a(x) = \sum_{n=0}^{\infty} a_n x^n, \qquad x \in S,$$

converging on an open neighborhood of S.

- Note that up to a slight reparametrization analytic functions are linear functions in the signature of X as $\mathbb{X}_t = (1, X_t - x_0, \frac{(X_t-x_0)^2}{2}, \frac{(X_t-x_0)^3}{3!}, \ldots)$.
- Let

$$\mathcal{U} := \{\mathbf{u} \in T((\mathbb{R})) \colon |\sum_{n=0}^{\infty} |u_n x^n| < \infty \text{ for all } x \in S + B_\varepsilon(0),\ \varepsilon > 0\}$$

and denote by $\star$ the discrete convolution, i.e.

$$(\mathbf{u} \star \mathbf{c})_i := \sum_{j_1+j_2=i} u_{j_1} c_{j_2}, \qquad \mathbf{u}, \mathbf{c} \in T((\mathbb{R})).$$

Affine case

Theorem (C.C, S. Svaluto-Ferro, J. Teichmann ('22))

The process $\mathbb{X} := (1, X, X^2, \dots, X^n, \dots)$ *is affine with respect to the operator* $R : \mathcal{U} \to T((\mathbb{R}))$ *given by*

$$R(\mathbf{u}) = \mathbf{b} \star \mathbf{u}^{(1)} + \frac{1}{2}\mathbf{a} \star \left(\mathbf{u}^{(2)} + \mathbf{u}^{(1)} \star \mathbf{u}^{(1)}\right), \quad u_k^{(\ell)} := u_{k+\ell}\frac{(k+\ell)!}{k!},$$

meaning that $\mathcal{A}\exp(\sum_{n=0}^\infty u_n x^n) = \exp(\sum_{n=0}^\infty u_n x^n)\sum_{n=0}^\infty R_n(\mathbf{u})x^n$.

Affine case

Theorem (C.C, S. Svaluto-Ferro, J. Teichmann ('22))

The process $\mathbb{X} := (1, X, X^2, \dots, X^n, \dots)$ *is affine with respect to the operator* $R : \mathcal{U} \to T((\mathbb{R}))$ *given by*

$$R(\mathbf{u}) = \mathbf{b} \star \mathbf{u}^{(1)} + \frac{1}{2}\mathbf{a} \star \left(\mathbf{u}^{(2)} + \mathbf{u}^{(1)} \star \mathbf{u}^{(1)}\right), \quad u_k^{(\ell)} := u_{k+\ell}\frac{(k+\ell)!}{k!},$$

meaning that $\mathcal{A}\exp(\sum_{n=0}^{\infty} u_n x^n) = \exp(\sum_{n=0}^{\infty} u_n x^n)\sum_{n=0}^{\infty} R_n(\mathbf{u})x^n$. *Suppose that the sequence valued Riccati equation*

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

admits an $\mathcal{U}$*-valued (weak) solution and that certain exponential moment conditions hold true. Then*

$$\mathbb{E}\left[\exp\left(\sum_{n=0}^{\infty} u_n X_T^n\right)\right] = \exp\left(\sum_{n=0}^{\infty} \psi_n(T) X_0^n\right).$$

Examples: Brownian motion & geometric Brownian motion

- Let $\mathcal{B}$ be the set of entire functions f such that $|f| \leq \exp(a(|\cdot|+1))$ on $\mathbb{R}$ for some $a \in \mathbb{R}_+$ and $\mathcal{D} := \{f \in \mathcal{B} \colon f', f'' \in \mathcal{B}\}$.
- Let $f \in \mathcal{D}$ such that $f(x) = \exp(\sum_{n=0}^{\infty} u_n x^n)$. Then

$$\mathbb{E}[\exp(\sum_{n=0}^{\infty} u_n(x+B_t)^n)] = \exp(\sum_{n=0}^{\infty} \psi_n(T)x^n).$$

  with $R(\mathbf{u}) = \frac{1}{2}(1,0,\ldots) \star \left(\mathbf{u}^{(2)} + \mathbf{u}^{(1)} \star \mathbf{u}^{(1)}\right)$.

- Choose $f(x) = \exp((i\lambda - \kappa)\exp(x))$ for $\lambda \in \mathbb{R}$ and $\kappa \in \mathbb{R}_+$, to obtain an expression for the Fourier-Laplace transform of geometric Brownian motion $S_t = \exp(x + B_t)$

$$\mathbb{E}[\exp((i\lambda - \kappa)S_t)] = \mathbb{E}[\exp((i\lambda - \kappa)\sum_{n=0}^{\infty} \frac{(x+B_t)^n}{n!}] = \exp(\sum_{n=0}^{\infty} \psi_n(T)\log(S_0)^n)$$

  with $\psi_n(0) = \frac{i\lambda - \kappa}{n!}$.

- We approximate the solution of the sequence-valued Riccati ODE by neural networks and deep learning methods for ODEs.

Polynomial case

Theorem (C.C, S. Svaluto-Ferro, J. Teichmann ('22))

The process $\mathbb{X} := (1, X, X^2, \dots, X^n, \dots)$ *is polynomial with respect to the operator* $L : \mathcal{U} \to T((\mathbb{R}))$ *given by*

$$L(\mathbf{u}) = \mathbf{b} \star \mathbf{u}^{(1)} + \frac{1}{2}\mathbf{a} \star \mathbf{u}^{(2)}, \quad u_k^{(\ell)} := u_{k+\ell}\frac{(k+\ell)!}{k!},$$

meaning that $\mathcal{A}(\sum_{n=0}^{\infty} u_n x^n) = \sum_{n=0}^{\infty} L_n(\mathbf{u}) x^n$.

Polynomial case

Theorem (C.C, S. Svaluto-Ferro, J. Teichmann ('22))

The process $\mathbb{X} := (1, X, X^2, \ldots, X^n, \ldots)$ *is polynomial with respect to the operator* $L : \mathcal{U} \to T((\mathbb{R}))$ *given by*

$$L(\mathbf{u}) = \mathbf{b} \star \mathbf{u}^{(1)} + \frac{1}{2}\mathbf{a} \star \mathbf{u}^{(2)}, \quad u_k^{(\ell)} := u_{k+\ell}\frac{(k+\ell)!}{k!},$$

meaning that $\mathcal{A}(\sum_{n=0}^\infty u_n x^n) = \sum_{n=0}^\infty L_n(\mathbf{u}) x^n$. *Suppose that the sequence valued linear ODE*

$$\partial_t \boldsymbol{c}(t) = L(\boldsymbol{c}(t)), \quad \boldsymbol{c}(0) = \mathbf{u}$$

admits an $\mathcal{U}$*-valued (weak) solution and that certain moment conditions hold true. Then*

$$\mathbb{E}\left[\sum_{n=0}^\infty u_n X_T^n\right] = \sum_{n=0}^\infty \boldsymbol{c}_n(T) X_0^n.$$

Examples

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

$$\mathbb{E}[\exp(uX_T)] = \sum_{n=0}^{\infty} c_n(T) X_0^n$$

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

Examples

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

$$\mathbb{E}[\exp(uX_T)] = \sum_{n=0}^{\infty} c_n(T) X_0^n$$

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

- Polynomial processes on compact state spaces
- 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 to the well known expression for the Laplace transform

$$\mathbb{E}[\exp(uX_T)] = \sum_{n=0}^{\infty} \underbrace{\frac{u^n}{(1 - \frac{a_1}{2} uT)^n n!}}_{c_n(T)} X_0^n = \exp\left(\frac{uX_0}{1 - \frac{a_1}{2} uT}\right).$$

Relation to signature based models?

Question: can the signature process of generic SDEs be treated as an infinite dimensional affine and/or polynomial process?

Relation to signature based models?

Question: can the signature process of generic SDEs be treated as an infinite dimensional affine and/or polynomial process?

Answer: very often this is the case!

Relation to signature based models?

Question: can the signature process of generic SDEs be treated as an infinite dimensional affine and/or polynomial process?

Answer: very often this is the case!

Indeed, if X is of the truly generic form

$$dX_t = \boldsymbol{b}(\widehat{\mathbb{X}}_t)dt + \sqrt{\mathbf{a}(\widehat{\mathbb{X}}_t)}dB_t, \quad X_0 \in S \subseteq \mathbb{R}^d \qquad \text{(SigSDE)}$$

where $(\widehat{\mathbb{X}}_t)_{t\geq 0}$ is the signature of $t \mapsto (X_t, t)$ and $\boldsymbol{b}$ and $\boldsymbol{a}$ are linear maps, then

- $\Rightarrow$ Ito's formula yields that the characteristics of $(\widehat{\mathbb{X}}_t)_{t\geq 0}$ are linear
- $\Rightarrow$ $\mathcal{A}\exp(\langle \boldsymbol{u}, \boldsymbol{x}\rangle) = \exp(\langle \boldsymbol{u}, \boldsymbol{x}\rangle)\langle R(\boldsymbol{u}), \boldsymbol{x}\rangle$ and $\mathcal{A}(\langle \boldsymbol{u}, \boldsymbol{x}\rangle) = \langle L(\boldsymbol{u}), \boldsymbol{x}\rangle$, $\boldsymbol{x} \in \mathcal{S}(S)$

$$R(\mathbf{u}) = \mathbf{b} \sqcup\mathchoice{\mkern-7mu}{\mkern-7mu}{\mkern-3.2mu}{\mkern-3.8mu}\sqcup \mathbf{u}^{(1)} + \frac{1}{2}\mathrm{tr}(\mathbf{a} \sqcup\mathchoice{\mkern-7mu}{\mkern-7mu}{\mkern-3.2mu}{\mkern-3.8mu}\sqcup (\mathbf{u}^{(2)} + \mathbf{u}^{(1)} \sqcup\mathchoice{\mkern-7mu}{\mkern-7mu}{\mkern-3.2mu}{\mkern-3.8mu}\sqcup \mathbf{u}^{(1)})),$$

$$L(\mathbf{u}) = \mathbf{b} \sqcup\mathchoice{\mkern-7mu}{\mkern-7mu}{\mkern-3.2mu}{\mkern-3.8mu}\sqcup \mathbf{u}^{(1)} + \frac{1}{2}\mathrm{tr}(\mathbf{a} \sqcup\mathchoice{\mkern-7mu}{\mkern-7mu}{\mkern-3.2mu}{\mkern-3.8mu}\sqcup \mathbf{u}^{(2)})$$

- $\Rightarrow$ $(\widehat{\mathbb{X}}_t)_{t\geq 0}$ is a $T((\mathbb{R}^d))$ valued linear, hence affine and polynomial process.

Sig-SDEs as affine and polynomial processes

- This means that under appropriate conditions...
  - ... $\mathbb{E}[\widehat{\mathbb{X}}_T]$ can be computed via polynomial technology, i.e. by solving an infinite dimensional linear ODE.
  - ... $\log \mathbb{E}[\exp(\langle \boldsymbol{u}, \widehat{\mathbb{X}}_T \rangle)]$ can be computed via affine technology, i.e. by solving an infinite dimensional Riccati ODE.

Sig-SDEs as affine and polynomial processes

- This means that under appropriate conditions...
  - ... $\mathbb{E}[\widehat{\mathbb{X}}_T]$ can be computed via polynomial technology, i.e. by solving an infinite dimensional linear ODE.
  - ... $\log \mathbb{E}[\exp(\langle \boldsymbol{u}, \widehat{\mathbb{X}}_T \rangle)]$ can be computed via affine technology, i.e. by solving an infinite dimensional Riccati ODE.
- The SigSDE setting goes beyond Markovian settings due to possibly path-dependent coefficients. The signature itself remains Markovian with linear characteristics, which is the essential feature.

Sig-SDEs as affine and polynomial processes

- This means that under appropriate conditions...
  - ... $\mathbb{E}[\widehat{\mathbb{X}}_T]$ can be computed via polynomial technology, i.e. by solving an infinite dimensional linear ODE.
  - ... $\log \mathbb{E}[\exp(\langle \boldsymbol{u}, \widehat{\mathbb{X}}_T \rangle)]$ can be computed via affine technology, i.e. by solving an infinite dimensional Riccati ODE.
- The SigSDE setting goes beyond Markovian settings due to possibly path-dependent coefficients. The signature itself remains Markovian with linear characteristics, which is the essential feature.
- Special cases of (SigSDE) are Makovian SDEs with $\boldsymbol{b}$ and $\boldsymbol{a}$ analytic in X. The infinite dimensional linear PDE for the expected signature (valid for any Markovian diffusion) reduces to an infinite dimensional linear ODE.
- If $\boldsymbol{b}$ and $\boldsymbol{a}$ only depend on the signature up to order 1 and 2 respectively, then $(\widehat{\mathbb{X}}_t^{\leq N})_{t\geq 0}$ is a finite dimensional polynomial process. This holds true in particular for X being a classical polynomial processes.

Jump Sig-SDEs as affine and polynomial processes

- In the case of jumps we use the Marcus signature, given as solution of

$$d\widehat{\mathbb{X}} = \sum_{i=1}^{d} \widehat{\mathbb{X}} \otimes \diamond d\widehat{X}^i, \quad \widehat{\mathbb{X}}_0 = (1, 0, 0, \dots) \in T((\mathbb{R}^d)),$$

where $\diamond$ denotes the Marcus integral (giving rise to a first order calculus)
$\int_0^t f(Z_s) \diamond dZ_s :=$
$\int_0^t f(Z_{s-}) dZ_s + \frac{1}{2} \int_0^t f'(Z_{s-}) d[Z^c, Z^c]_s + \sum_{0<s\leq t} \Delta Z_s \left(\int_0^1 f(Z_{s-} + \theta \Delta Z_s) - f(Z_{s-}) d\theta \right).$

Jump Sig-SDEs as affine and polynomial processes

- In the case of jumps we use the Marcus signature, given as solution of

$$d\widehat{\mathbb{X}} = \sum_{i=1}^{d} \widehat{\mathbb{X}} \otimes \diamond d\widehat{X}^i, \quad \widehat{\mathbb{X}}_0 = (1, 0, 0, \dots) \in T((\mathbb{R}^d)),$$

where $\diamond$ denotes the Marcus integral (giving rise to a first order calculus)

$\int_0^t f(Z_s) \diamond dZ_s :=$

$\int_0^t f(Z_{s-}) dZ_s + \frac{1}{2} \int_0^t f'(Z_{s-}) d[Z^c, Z^c]_s + \sum_{0<s\leq t} \Delta Z_s \left(\int_0^1 f(Z_{s-} + \theta \Delta Z_s) - f(Z_{s-}) d\theta \right).$

- Then analogous statements hold true for Sig-SDEs with jumps of the form

$$dX_t = \boldsymbol{b}(\widehat{\mathbb{X}}_t)dt + \sqrt{\mathbf{a}(\widehat{\mathbb{X}}_t)}dB_t + \int \xi(\mu^X(d\xi, dt) - \boldsymbol{K}(\widehat{\mathbb{X}}_t, d\xi)dt),$$

where the compensator $\boldsymbol{K}$ is such that $\mathbf{x} \mapsto \boldsymbol{K}(\mathbf{x}, d\xi)$ is a linear map.

$\Rightarrow$ $(\widehat{\mathbb{X}}_t)_{t\geq 0}$ is a $T((\mathbb{R}^d))$ valued affine and polynomial process.

$\Rightarrow$ If X is a classical polynomial process, then $(\widehat{\mathbb{X}}_t^{\leq N})_{t\geq 0}$ is a finite dimensional polynomial process. The infinite dimensional linear PIDE for the expected signature becomes a finite dimensional ODE.

Conclusion

- Signature based models distinguish themselves in
  - universality, as the dynamics of all classical models can be approximated
  - criteria for no-arbitrage
  - efficient pricing, hedging and calibration (also extension to VIX options).
- Extension to Lévy type signature models is possible (joint work with F.Primavera and S.Svaluto-Ferro).
- Generic classes of SDEs can be proved to be affine and polynomial, in particular SDEs with analytic coefficients $\Rightarrow$ one step in the direction of universality of affine processes
- For (jump) SigSDEs
  - its expected signature can be computed via polynomial technology
  - the Fourier-Laplace transform of its signature can be computed via affine technology

Thank you for your attention!