Universal approximation theorems for continuous functions of càdlàg paths and Lévy-type signature models

Christa Cuchiero
(based on joint work with Francesca Primavera and Sara Svaluto-Ferro)

University of Vienna

Third Symposium on Machine Learning and Dynamical Systems

Toronto, Fields Institute, September 26, 2022

Outline

Data driven models

- Highly parametric and overparametrized models gain in importance: instead of a few parameters, the goal is rather to 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

- Highly parametric and overparametrized models gain in importance: instead of a few parameters, the goal is rather to 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 an underlying process, usually Brownian motion.

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

Data driven models

- Highly parametric and overparametrized models gain in importance: instead of a few parameters, the goal is rather to 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 an underlying process, usually Brownian motion.

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

- The signature models for asset prices proposed so far have only dealt with continuous trajectories.
  - How can we define signature-based models including jumps?

Signature as (non-random) reservoir

- Signature of continuous paths (T. Lyons ('98), P. Friz & N. Victoir ('10), P. Friz & M. Hairer ('14)) owes its importance to the fact that it serves as linear regression basis for continuous functionals of continuous paths.
- In spirit of reservoir computing is can thus be viewed as a (non-random) reservoir:
  - the input signal, in this case a path is fed into a fixed dynamical system, the signature transform, mapping the paths to an element in the extended tensor algebra (or rather group-like elements);
  - then a simple linear readout mechanism is trained with the goal to approximate a continuous functional of a continuous path uniformly on compact sets of paths.
- Extensions to randomized versions of signature (C.C., Gonon, Grigoryeva, Ortega, Teichmann)

Universal approximation via linear maps of signature

- This works due to the Stone-Weierstrass theorem since
  - signature is point-separating by adding time as first component;
  - linear functions on the signature form an algebra that contains 1, as every polynomial on signature may be realized as a linear function via the shuffle product $\sqcup\!\sqcup$.

Universal approximation via linear maps of signature

- This works due to the Stone-Weierstrass theorem since
  - signature is point-separating by adding time as first component;
  - linear functions on the signature form an algebra that contains 1, as every polynomial on signature may be realized as a linear function via the shuffle product $\sqcup\!\sqcup$.
- Do we get a similar universal approximation result for an appropriate notion of signature of a càdlàg path?
- In view of universal model classes this is essential for considering signature models based on Lévy processes.

Càdlàg rough paths

Our considerations rely on the Marcus signature for càdlàg rough paths:

Definition (P. Friz and A. Shekhar ('17))

Let $p \in [2,3)$ and $\Delta_1 := \{(s,t) \in [0,1]^2 \mid s \leq t\}$. A pair $\mathbf{X} = (X, \mathbb{X}^{(2)})$ is called càdlàg p-rough path over $\mathbb{R}^d$, in symbols $\mathbf{X} \in \mathcal{W}^p([0,1], \mathbb{R}^d)$, if

$$X : [0,1] \to \mathbb{R}^d, \qquad \mathbb{X}^{(2)} : \Delta_1 \to (\mathbb{R}^d)^{\otimes 2}$$

satisfy:

1. The map $[0,1] \ni t \longmapsto (X_{0,t}, \mathbb{X}^{(2)}_{0,t}) \in \mathbb{R}^d \oplus (\mathbb{R}^d)^{\otimes 2}$ is càdlàg .
2. Chen's relation holds:
$$\mathbb{X}^{(2)}_{s,t} = \mathbb{X}^{(2)}_{s,u} + \mathbb{X}^{(2)}_{u,t} + X_{s,u} \otimes X_{u,t} \text{ for } 0 \leq s < u < t \leq 1.$$
3. $\mathbf{X} = (X, \mathbb{X}^{(2)})$ is of finite p-variation in the rough path sense:
$$\|\mathbf{X}\|_{p-var} := \|X\|_{p-var} + \|\mathbb{X}^{(2)}\|^{1/2}_{p/2-var} < \infty.$$

Weakly geometric and Marcus-like càdlàg rough paths

Definition (P. Friz and A. Shekhar ('17))

Let $p \in [2, 3)$ and $\mathbf{X} \in \mathcal{W}^p([0,1], \mathbb{R}^d)$.

- $\mathbf{X}$ is said to be a weakly geometric càdlàg p-rough path over $\mathbb{R}^d$, in symbols $\mathbf{X} \in \mathcal{W}^p_g([0,1], \mathbb{R}^d)$, if for all $0 \leq s < t \leq 1$

$$Sym(\mathbb{X}^{(2)}_{s,t}) = \frac{1}{2} X_{s,t} \otimes X_{s,t}.$$

- If moreover

$$\lim_{s \uparrow t} Anti(\mathbb{X}^{(2)}_{s,t}) = 0, \text{ for all } t \in [0,1],$$

then $\mathbf{X}$ is said to be Marcus-like, in symbols $\mathbf{X} \in \mathcal{W}^p_M([0,1], \mathbb{R}^d)$.

The motivation for the definition of Marcus-like path comes from the Marcus integral. There jumps are replaced by straight lines, which do not create area.

Lie-group point of view on càdlàg rough path

- By extending weakly geometric and Marcus-like càdlàg rough paths by 1, i.e. $\mathbf{X} = (1, X, \mathbb{X}^{(2)})$, they can also be interpreted to take values in the free step-2 nilpotent Lie group $(G^2(\mathbb{R}^d), \otimes)$ defined as $\exp^{(2)}(\mathbb{R}^d \oplus [\mathbb{R}^d, \mathbb{R}^d])$ where $[\mathbf{g}, \mathbf{h}] = \mathbf{g} \otimes \mathbf{h} - \mathbf{h} \otimes \mathbf{g}$.
- For Lie group valued paths, the increments of $\mathbf{X}$ are defined via $\mathbf{X}_{s,t} := \mathbf{X}_s^{-1} \otimes \mathbf{X}_t$ and in turn the jumps as $\Delta \mathbf{X}_t := \lim_{s \uparrow t} \mathbf{X}_{s,t}$.
- For a weakly geometric path we thus have

$$\mathbf{X}_{s,t} = (1, X_{s,t}, \frac{1}{2} X_{s,t}^{\otimes 2} + Anti(\mathbb{X}_{s,t}^{(2)})) = \exp^{(2)}(X_{s,t}, Anti(\mathbb{X}_{s,t}^{(2)})) \in G^2(\mathbb{R}^d).$$

such that the Marcus condition implies

$$\log^{(2)}(\Delta \mathbf{X}_t) = \lim_{s \uparrow t} \log^{(2)}(\mathbf{X}_{s,t}) = \lim_{s \uparrow t}(0, X_{s,t}, Anti(\mathbb{X}_{s,t}^{(2)})) = (0, \Delta X_t, 0).$$

Towards signature - analogue of Lyons' extension theorem

- We recall the analogue of Lyons' extension theorem for càdlàg rough paths.
- Here, the free step-N-nilpotent Lie group $G^N(\mathbb{R}^d)$ is defined analogously as $G^2(\mathbb{R}^d)$, namely as exponential image of $\mathbb{R}^d \oplus [\mathbb{R}^d, \mathbb{R}^d] \oplus [\mathbb{R}^d, [\ldots, [\mathbb{R}^d, \mathbb{R}^d]]$.

Theorem (P. Friz and A. Shekhar ('17))

Let $p \in [2, 3)$ and $\mathbb{N} \ni N > 2$.
A weakly geometric càdlàg p-rough path $\mathbf{X} : [0, 1] \to G^2(\mathbb{R}^d)$ admits a unique extension to a càdlàg path $\mathbb{X}^N : [0, 1] \to G^N(\mathbb{R}^d)$, such that

- *$\mathbb{X}^N$ starts from $(1, 0, 0, \ldots) \in G^N(\mathbb{R}^d)$,*
- *it is of finite p-variation, (with respect to the Carnot-Caratheodory metric d_{CC} on $G^N(\mathbb{R}^d)$), and*
- *satisfies $\log^{(N)}(\Delta \mathbb{X}^N_t) = \log^{(2)}(\Delta \mathbf{X}_t)$ for all $t \in [0, 1]$.*

$\mathbb{X}^N$ is called minimal jump extension of $\mathbf{X}$ in $G^N(\mathbb{R}^d)$.

Remarks

- The proof of the above result relies on Marcus' idea of turning a càdlàg path into a continuous one by introducing an additional time interval at each jump time, and replacing the jumps by a straight line which connects the states before and after the jump.
- For computing the minimal jump extension the jumps of $\mathbf{X}$ are connected by the log-linear path-function $\varphi : G^2(\mathbb{R}^d) \times G^2(\mathbb{R}^d) \to C([0,1], G^2(\mathbb{R}^d))$. For $(\mathbf{X}_{u^-}, \mathbf{X}_u)$, φ explicitly reads as

$$\log^{(2)}(\varphi(\mathbf{X}_{u^-}, \mathbf{X}_u)) = s \mapsto (0, X_{u^-} + s\Delta X_u, Anti(\mathbb{X}^{(2)}_{0,u^-} + s\Delta \mathbb{X}^{(2)}_u)).$$

- The outcome of this is a continuous weakly geometric p-rough path $\mathbf{X}^\varphi$, which admits by Lyons' extension theorem a unique (continuous) extension $\mathbb{X}^{\varphi,N}$. The càdlàg extension is then obtained via a time-change, i.e. $\mathbb{X}^N := \mathbb{X}^{\varphi,N}_\tau$.

Marcus-type RDEs

- This construction yields the concept of Marcus-type rough differential equation (RDEs).
- Roughly speaking, a solution to a Marcus-type RDE driven by $\mathbf{X}$ is defined as the time changed solution of the continuous RDE driven by $\mathbf{X}^{\varphi}$. The equation is denoted by

$$dY_t = V(Y_t) \diamond d\mathbf{X}_t,$$

  for suitable vector fields V.
- Since $\mathbb{X}^{\varphi,N}$ is the continuous Lyons' extension of $\mathbf{X}^{\varphi}$, it solves the linear signature RDE

$$d\mathbb{X}^{\varphi,N} = \mathbb{X}^{\varphi,N} \otimes d\mathbf{X},$$
$$\mathbb{X}_0^{\varphi,N} = (1, 0, 0, \dots) \in G^N(\mathbb{R}^d).$$

  Thus, $\mathbb{X}^N$ is by construction a solution of the corresponding Marcus-type linear RDE.

Marcus signature RDE

Corollary (P. Friz and A. Shekhar ('17))

Let $p \in [2,3)$, $\mathbb{N} \ni N > 2$, and $\mathbf{X} : [0,1] \to G^2(\mathbb{R}^d)$ be a weakly geometric càdlàg p-rough path. The minimal jump extension $\mathbb{X}^N$ with values in $G^N(\mathbb{R}^d)$ satisfies the Marcus-type RDE

$$d\mathbb{X}^N = \mathbb{X}^N \otimes \diamond d\mathbf{X},$$
$$\mathbb{X}_0^N = (1, 0, 0, \dots) \in G^N(\mathbb{R}^d),$$

whose explicit form is (by definition of the Marcus rough integration)

$$\mathbb{X}_t^N = 1 + \int_0^t \mathbb{X}_{s-}^N \otimes d\mathbf{X}_s$$
$$+ \sum_{0<s\leq t} \mathbb{X}_{s-}^N \otimes \left(-\frac{1}{2}(\Delta X_s)^{\otimes 2} + \sum_{k=2}^{N} \frac{1}{k!}(\Delta X_s, Anti(\Delta \mathbb{X}_s^{(2)}))^{\otimes k} \right).$$

The integral is understood as a rough integral and the summation term is well defined as an absolutely summable series.

Definition of signature

- Define the set of group-like elements as follows

$$G((\mathbb{R}^d)) := \{\boldsymbol{a} = (a_0, a_1, \dots, a_n, \dots) \in \prod_{n=0}^{\infty} (\mathbb{R}^d)^{\otimes n} \mid \pi_{\leq N}(\boldsymbol{a}) \in G^N(\mathbb{R}^d)\, \forall N\}.$$

  and denote for a multiindex $I = (i_1, \dots, i_n)$, $\epsilon_I = \epsilon_{i_1} \otimes \cdots \otimes \epsilon_{i_n}$ the basis elements of $(\mathbb{R}^d)^{\otimes n}$.

- Group-like elements satisfy the shuffle property: $\langle \epsilon_I, \boldsymbol{a} \rangle \langle \epsilon_J, \boldsymbol{a} \rangle = \langle \epsilon_I \sqcup\!\sqcup\, \epsilon_J, \boldsymbol{a} \rangle$, for $\boldsymbol{a} \in G((\mathbb{R}^d))$ and multiindices $I \in \{1, \dots, d\}^n$, $J \in \{1, \dots, d\}^m$.

- We can now introduce the signature of càdlàg paths without ambiguity.

Definition of signature

- Define the set of group-like elements as follows

$$G((\mathbb{R}^d)) := \{\boldsymbol{a} = (a_0, a_1, \dots, a_n, \dots) \in \prod_{n=0}^{\infty} (\mathbb{R}^d)^{\otimes n} \mid \pi_{\leq N}(\boldsymbol{a}) \in G^N(\mathbb{R}^d) \,\forall N\}.$$

  and denote for a multiindex $I = (i_1, \dots, i_n)$, $\epsilon_I = \epsilon_{i_1} \otimes \cdots \otimes \epsilon_{i_n}$ the basis elements of $(\mathbb{R}^d)^{\otimes n}$.

- Group-like elements satisfy the shuffle property: $\langle \epsilon_I, \boldsymbol{a} \rangle \langle \epsilon_J, \boldsymbol{a} \rangle = \langle \epsilon_I \sqcup\!\sqcup\, \epsilon_J, \boldsymbol{a} \rangle$, for $\boldsymbol{a} \in G((\mathbb{R}^d))$ and multiindices $I \in \{1, \dots, d\}^n$, $J \in \{1, \dots, d\}^m$.

- We can now introduce the signature of càdlàg paths without ambiguity.

Definition

Let $p \in [2, 3)$ and $\mathbf{X} : [0, 1] \to G^2(\mathbb{R}^d)$ be a weakly geometric càdlàg p-rough path. The signature of $\mathbf{X}$, denoted by $\mathbb{X}$, is the unique solution to the Marcus-type RDE in $G((\mathbb{R}^d))$

$$d\mathbb{X} = \mathbb{X} \otimes \diamond d\mathbf{X}, \qquad \mathbb{X}_0 = (1, 0, 0, \dots) \in G((\mathbb{R}^d)).$$

Càdlàg semimartingales as càdlàg rough paths

Càdlàg semimartingales fit well into the theory of càdlàg rough paths. Indeed, every semimartingale admits a canonical lift which is a.s. a Marcus-type càdlàg p-rough path for any $p > 2$.

Proposition (P. Friz and A. Shekhar ('17))

Let $p \in (2, 3)$, X be a $\mathbb{R}^d$-valued càdlàg semimartingale and $[X, X]^c$ its $(\mathbb{R}^d)^{\otimes 2}$-valued continuous quadratic variation. Then, $\mathbf{X}(\omega) = (X(\omega), \mathbb{X}^{(2)}(\omega)) \in \mathcal{W}_M^p([0, 1], \mathbb{R}^d)$ a.s., where, for $0 \leq s \leq t \leq 1$,

$$\mathbb{X}_{s,t}^{(2)} := \int_s^t X_{s,r^-} \otimes dX_r + \frac{1}{2}[X, X]_{s,t}^c + \sum_{s<u\leq t} \Delta X_u \otimes \Delta X_u,$$

and the integral is understood in Itô's sense. We call $\mathbf{X}$ Marcus lift of X.

Marcus integral for semimartingales

- Replacing the jumps by a straight line for a semimartingale X yields the so-called Marcus integral (see Kurtz, Pardoux & Protter (1995)), defined (here for $d = 1$) as follows:

$$\int_0^t f(X_{r-}) \diamond dX_r := \int_0^t f(X_{r-}) dX_r + \frac{1}{2} \int_0^t f'(X_{r-}) d[X, X]_r^c + \sum_{0 \leq r \leq t} \Delta X_r \int_0^1 f(X_{r-} + \theta \Delta X_r) - f(X_{r-}) d\theta,$$

where the first integral on the right is understood as Ito integral.

- This leads to a first order calculus, namely

$$f(X_t) - f(X_s) = \int_0^t f'(X_s) \diamond dX_s.$$

Marcus SDE

Proposition

Let X be an $\mathbb{R}^d$-valued semimartingale and **X** *its Marcus lift. It holds that the above Marcus-type RDE coincides a.s. with the Marcus SDE*

$$d\mathbb{X} = \mathbb{X} \otimes \diamond dX, \quad \mathbb{X}_0 = (1, 0, 0, \dots) \in G((\mathbb{R}^d)).$$

The explicit form of the Marcus SDE is given by

$$\begin{aligned} \mathbb{X}_t =& 1 + \int_0^t \mathbb{X}_{r-} \otimes dX_r + \frac{1}{2} \int_0^t \mathbb{X}_{r-} \otimes d[X,X]_r^c \\ &+ \sum_{0<r\leq t} \mathbb{X}_{r-} \otimes \{\exp(\Delta X_r) - \Delta X_r - 1\} \\ &= 1 + \int_0^t \mathbb{X}_{r-} \otimes dX_r + \frac{1}{2} \int_0^t \mathbb{X}_{r-} \otimes d[X,X]_r^c + \sum_{0<r\leq t} \mathbb{X}_{r-} \otimes \sum_{k=2}^{\infty} \frac{(\Delta X_r)^{\otimes k}}{k!}. \end{aligned}$$

For $d = 1$, this yields the Taylorpolynomials: $(1, X_t - X_0, \dots, , \frac{(X_t - X_0)^n}{n!}, \dots)$.

UAT for continuous functionals of càdlàg rough paths

- Define the following set

$$\widehat{\mathcal{W}}_g^p([0,1],\mathbb{R}^{d+1}) := \{\hat{\mathbf{X}} = (\hat{X}, \hat{\mathbb{X}}^{(2)}) \in \mathcal{W}_g^p([0,1],\mathbb{R}^{d+1}) \mid \text{the first component of } \hat{X} \text{ is } t\}.$$

UAT for continuous functionals of càdlàg rough paths

- Define the following set

$$\widehat{\mathcal{W}_g^p}([0,1],\mathbb{R}^{d+1}) := \{\hat{\mathbf{X}} = (\hat{X}, \hat{\mathbb{X}}^{(2)}) \in \mathcal{W}_g^p([0,1],\mathbb{R}^{d+1}) \mid \text{the first component of } \hat{X} \text{ is } t\}.$$

- For a metric space (E, d) and let $D([0,1], E)$ be the space of càdlàg paths on it. Denote by Λ the set of all strictly increasing bijections of $[0,1]$ to itself. The Skorokhod J_1-metric is defined as follows: for $X, Y \in D([0,1], E)$

$$\sigma_\infty(X, Y) := \inf_{\lambda \in \Lambda} \{|\lambda| \vee \sup_{s \in [0,1]} d(X_{\lambda(s)}, Y_s)\},$$

  where, $|\lambda| := \sup_{s \in [0,1]} |\lambda(s) - s|$.

- In our case we consider $(E, d) = (G^N(\mathbb{R}^d), d_{CC})$ with d_{CC} denoting the Carnot-Caratheodory metric.

UAT for continuous functionals of weakly geometric càdlàg rough paths

Theorem (C. C., F. Primavera., S. Svaluto-Ferro (22'))

Let $K \subset \widehat{\mathcal{W}}_g^p([0,1], \mathbb{R}^{d+1})$ be a subset which is compact with respect to the topology induced by the J_1-metric and bounded with respect to the p-variation norm on the space of weakly geometric càdlàg rough paths.

Let $f : K \to \mathbb{R}$ be a continuous function with respect to the J_1-topology and, for each $\hat{\mathbf{X}} \in K$, denote by $\hat{\mathbb{X}}$ its signature. Then, for every $\epsilon > 0$ there exists a linear functional ℓ such that

$$\sup_{\hat{\mathbf{X}}_{[0,1]} \in K} |f(\hat{\mathbf{X}}_{[0,1]}) - \ell(\hat{\mathbb{X}}_1)| \leq \epsilon.$$

Sketch of the proof

- Apply the Stone-Weierstrass theorem to the set A given by

$$A := \text{span}\{\ell : K \to \mathbb{R}\ ;\ \hat{\mathbf{X}} \mapsto \langle \epsilon_I, \hat{\mathbb{X}}_1 \rangle : |I| \geq 0\}.$$

- Therefore, we have to prove that A

1. ... is a linear subspace of continuous functions from K to $\mathbb{R}$, which is a consequence of the fact that the solution map of the linear Marcus-type RDE

$$(K, \sigma_\infty) \to (D([0,1], G^N(\mathbb{R}^{d+1})), \sigma_\infty)$$
$$\hat{\mathbf{X}} \mapsto \hat{\mathbb{X}}^N$$

is continuous for every $N \geq 3$ (Chevyrev and Friz (2019)).

2. ... is a sub-algebra containing a non-zero constant function, which is true by the shuffle-property.

3. ... separates points, which follows from the fact that for a càglàd function $f : [0,1] \to \mathbb{R}$ with $f(0) = 0$ and $\int_0^1 f(s)s^n\, ds = 0$ for all $n \in \mathbb{N}$, it holds that $f \equiv 0$.

Towards Lévy type signature models

- Consider a Brownian motion W and an homogeneous Poisson random measure μ (independent from W) with intensity $\nu(dt, dx) = dt \times F(dx)$ such that

$$F(\{0\}) = 0, \quad \int_{\mathbb{R}} (|x|^2 \wedge 1) F(dx) < \infty, \quad \int_{|x|>1} |x|^k F(dx) < \infty \text{ for all } k \geq 1.$$

- Let $N \in \mathbb{N}, N \geq 2$ and define the $\mathbb{R}^{(N+2)}$-valued process $\hat{X}$ via

$$\hat{X}_t = (t, W_t, \int_0^t \int x(\mu - \nu)(ds, dx), \int_0^t \int x^2 \mu(ds, dx), \ldots, \int_0^t \int x^N \mu(ds, dx)).$$

- $\hat{X}$ encodes all the randomness of the model that we are going to define and call it thus primary underlying process. Moreover, it is a Lévy process, whose signature will be denoted by $\hat{\mathbb{X}}$.

Lévy type signature models

Lévy-type signature models are jump-diffusion models whose characteristics are linear functions of the signature of the primary process $\hat{X}$.

Definition

We define a Lévy-type signature model (under a pricing measure $\mathbb{Q}$) as follows

$$S(\ell)_t = S_0 + \int_0^t \left(\sum_{|J| \leq n} \ell^{J_W} \langle \epsilon_J, \hat{\mathbb{X}}_{s-} \rangle \right) dW_s + \int_0^t \int_{\mathbb{R}} \left(\sum_{|J| \leq n} \ell^{J_\nu} \langle \epsilon_J, \hat{\mathbb{X}}_{s-} \rangle \right) y \, (\mu - \nu)(ds, dy),$$

where $\ell^{J_W}, \ell^{J_\nu} \in \mathbb{R}$ and J are multiindices of length at most n with entries in $\{-1, 0, 1, \ldots, d\}$ for d s.t. $d \leq N$. The index -1 corresponds to t and 0 to W.

The model is motivated by the above UAT, implying that the characteristics of $S(\ell)$ can be interpreted as approximations of continuous path functionals of X.

Sig-model representation

- The model itself can be represented as a linear function of $\hat{\mathbb{X}}$.

Sig-model representation

- The model itself can be represented as a linear function of $\hat{\mathbb{X}}$.

Proposition (C. C., F.Primavera, S. Svaluto-Ferro ('22))

If $N \geq nd + 1$, then

$$\int_0^t \langle \epsilon_J, \hat{\mathbb{X}}_{s-} \rangle dW_s = \langle (\epsilon_J; \epsilon_0)^\sim, \hat{\mathbb{X}}_t \rangle,$$

$$\int_0^t \int_{\mathbb{R}} \langle \epsilon_J, \hat{\mathbb{X}}_{s-} \rangle y \ (\mu - \nu)(ds, dy) = \langle (\epsilon_J; \epsilon_1)^\sim, \hat{\mathbb{X}}_t \rangle$$

and the Lévy signature model $S(\ell)$ admits the sig-model representation

$$S(\ell)_t = S_0 + \sum_{|J| \leq n} \ell^{J_W} \langle (\epsilon_J; \epsilon_0)^\sim, \hat{\mathbb{X}}_t \rangle + \ell^{J_\nu} \langle (\epsilon_J; \epsilon_1)^\sim, \hat{\mathbb{X}}_t \rangle.$$

Sig-model representation

- The model itself can be represented as a linear function of $\hat{\mathbb{X}}$.

Proposition (C. C., F.Primavera, S. Svaluto-Ferro ('22))

If $N \geq nd + 1$*, then*

$$\int_0^t \langle \epsilon_J, \hat{\mathbb{X}}_{s-} \rangle dW_s = \langle (\epsilon_J; \epsilon_0)^{\sim}, \hat{\mathbb{X}}_t \rangle,$$

$$\int_0^t \int_{\mathbb{R}} \langle \epsilon_J, \hat{\mathbb{X}}_{s-} \rangle y \, (\mu - \nu)(ds, dy) = \langle (\epsilon_J; \epsilon_1)^{\sim}, \hat{\mathbb{X}}_t \rangle$$

and the Lévy signature model $S(\ell)$ *admits the sig-model representation*

$$S(\ell)_t = S_0 + \sum_{|J| \leq n} \ell^{J_W} \langle (\epsilon_J; \epsilon_0)^{\sim}, \hat{\mathbb{X}}_t \rangle + \ell^{J_\nu} \langle (\epsilon_J; \epsilon_1)^{\sim}, \hat{\mathbb{X}}_t \rangle.$$

- The condition $N \geq nd + 1$ is the precise reason why we need to include moments of the jump measure.
- This result implies tractable pricing formulas and that time-series calibration reduces to a linear regression.

Towards pricing of signature payoffs

- As a result of the UAT, continuous (with respect to the J_1 topology) payoffs can be approximated by linear functions of the time-extended signature of the price process, so-called signature payoffs (see Lyons, S. Nejad, and I. Perez Arribas (2020)).
- Denote by $\hat{S}(\ell)_t = (t, S(\ell)_t)$ the time-extended price and by $\hat{\mathbb{S}}(\ell)$ its signature.

Towards pricing of signature payoffs

- As a result of the UAT, continuous (with respect to the J_1 topology) payoffs can be approximated by linear functions of the time-extended signature of the price process, so-called signature payoffs (see Lyons, S. Nejad, and I. Perez Arribas (2020)).
- Denote by $\hat{S}(\boldsymbol{\ell})_t = (t, S(\boldsymbol{\ell})_t)$ the time-extended price and by $\hat{\mathbb{S}}(\boldsymbol{\ell})$ its signature.

Proposition (C.C., Francesca Primavera, S. Svaluto-Ferro ('22))

Consider a multi-index $I \in \{-1, 1\}^{|I|}$ and let $N \geq |I|(dn+1)$. Then there exists a linear combination $U(I, \boldsymbol{\ell})$ of indices with entries in $\{-1, 0, 1, \dots, N\}$, such that the following equality holds

$$\langle \epsilon_I, \hat{\mathbb{S}}(\boldsymbol{\ell})_t \rangle = \langle \epsilon_{U(I,\boldsymbol{\ell})}, \hat{\mathbb{X}}_t \rangle.$$

Furthermore, $\langle \epsilon_{U(I,\boldsymbol{\ell})}, \mathbb{X}_t \rangle$ is a polynomial of degree $|I|$ in $(\ell^W_{1,\dots,M}, \ell^\nu_{1,\dots,M}) \in \mathbb{R}^{2M}$ with $M = \sum_{i=1}^n (d+2)^i$, and $U(I, \boldsymbol{\ell})$ can be computed recursively.

Pricing of signature payoffs

Corollary (C.C., F. Primavera, S. Svaluto-Ferro ('22))

Let $N \geq |I|(nd+1)$. Then the price of a signature payoff $\langle \epsilon_I, \hat{\mathbb{S}}(\boldsymbol{\ell})_T \rangle$ can be expressed as

$$\mathbb{E}_{\mathbb{Q}}[\langle \epsilon_I, \hat{\mathbb{S}}(\boldsymbol{\ell})_T \rangle] = \langle \epsilon_{U(I,\boldsymbol{\ell})}, \mathbb{E}_{\mathbb{Q}}[\hat{\mathbb{X}}_T] \rangle.$$

- The calibration to options on signature payoffs reduces to a polynomial optimization problem in the $\boldsymbol{\ell}$.
- $\hat{X}$ is a Lévy process, therefore its expected signature $\mathbb{E}_{\mathbb{Q}}[\hat{\mathbb{X}}_T]$ can be computed analytically by the methods in P. Friz and A. Shekhar ('17).

Pricing of signature payoffs

Corollary (C.C., F. Primavera, S. Svaluto-Ferro ('22))

Let $N \geq |I|(nd+1)$. *Then the price of a signature payoff* $\langle \epsilon_I, \hat{\mathbb{S}}(\boldsymbol{\ell})_T \rangle$ *can be expressed as*

$$\mathbb{E}_{\mathbb{Q}}[\langle \epsilon_I, \hat{\mathbb{S}}(\boldsymbol{\ell})_T \rangle] = \langle \epsilon_{U(I,\boldsymbol{\ell})}, \mathbb{E}_{\mathbb{Q}}[\hat{\mathbb{X}}_T] \rangle.$$

- The calibration to options on signature payoffs reduces to a polynomial optimization problem in the $\boldsymbol{\ell}$.
- $\hat{X}$ is a Lévy process, therefore its expected signature $\mathbb{E}_{\mathbb{Q}}[\hat{\mathbb{X}}_T]$ can be computed analytically by the methods in P. Friz and A. Shekhar ('17).
- $\hat{\mathbb{X}}^N$ is also a polynomial process. Therefore its expectation can be computed via polynomial technology, i.e. by solving a finite dimensional linear ODE.
- Roughly speaking, by a polynomial process we here mean a Markov process whose generator maps linear maps of at most "degree N", i.e. $\sum_{|I|\leq N} \alpha_I \langle e_I, \mathbf{x} \rangle$, to linear maps of at most "degree N".

Truncated signature as polynomial process

- This feature holds true for much more general processes than Lévy processes.

Truncated signature as polynomial process

- This feature holds true for much more general processes than Lévy processes.
- Indeed, if X is a Jump Sig-SDE of form

$$dX_t = \boldsymbol{b}(\mathbb{X}_{t-})dt + \sqrt{\mathbf{a}(\mathbb{X}_{t-})}dB_t + \int_{\mathbb{R}^d} \xi(\mu^X(d\xi, dt) - \boldsymbol{K}(\mathbb{X}_{t-}, d\xi)dt),$$

with $\boldsymbol{b}$, $\boldsymbol{a}$ and $\mathbf{x} \mapsto \int_{\mathbb{R}^d} \langle \epsilon_I, \xi^{\otimes I} \rangle \boldsymbol{K}(\mathbf{x}, d\xi)$ linear functions that depend on the signature up to order 1, 2 and $|I| \leq N$ respectively, then $(\mathbb{X}_t^N)_{t \geq 0}$ is a finite dimensional polynomial process on $G^N(\mathbb{R}^d)$.

Truncated signature as polynomial process

- This feature holds true for much more general processes than Lévy processes.
- Indeed, if X is a Jump Sig-SDE of form

$$dX_t = \boldsymbol{b}(\mathbb{X}_{t-})dt + \sqrt{\mathbf{a}(\mathbb{X}_{t-})}dB_t + \int_{\mathbb{R}^d} \xi(\mu^X(d\xi, dt) - \boldsymbol{K}(\mathbb{X}_{t-}, d\xi)dt),$$

  with $\boldsymbol{b}$, $\boldsymbol{a}$ and $\mathbf{x} \mapsto \int_{\mathbb{R}^d} \langle \epsilon_I, \xi^{\otimes I} \rangle \boldsymbol{K}(\mathbf{x}, d\xi)$ linear functions that depend on the signature up to order 1, 2 and $|I| \leq N$ respectively, then $(\mathbb{X}_t^N)_{t \geq 0}$ is a finite dimensional polynomial process on $G^N(\mathbb{R}^d)$.
- This holds true in particular for X being a classical polynomial process.

Truncated signature as polynomial process

- This feature holds true for much more general processes than Lévy processes.
- Indeed, if X is a Jump Sig-SDE of form

$$dX_t = \boldsymbol{b}(\mathbb{X}_{t-})dt + \sqrt{\mathbf{a}(\mathbb{X}_{t-})}dB_t + \int_{\mathbb{R}^d} \xi(\mu^X(d\xi, dt) - \boldsymbol{K}(\mathbb{X}_{t-}, d\xi)dt),$$

  with $\boldsymbol{b}$, $\boldsymbol{a}$ and $\mathbf{x} \mapsto \int_{\mathbb{R}^d} \langle \epsilon_I, \xi^{\otimes I} \rangle \boldsymbol{K}(\mathbf{x}, d\xi)$ linear functions that depend on the signature up to order 1, 2 and $|I| \leq N$ respectively, then $(\mathbb{X}_t^N)_{t \geq 0}$ is a finite dimensional polynomial process on $G^N(\mathbb{R}^d)$.
- This holds true in particular for X being a classical polynomial process.
- The expected truncated signature can then be computed by solving a finite dimensional linear ODE, i.e. computing a matrix exponential.

Truncated signature as polynomial process

- This feature holds true for much more general processes than Lévy processes.
- Indeed, if X is a Jump Sig-SDE of form

$$dX_t = \boldsymbol{b}(\mathbb{X}_{t-})dt + \sqrt{\mathbf{a}(\mathbb{X}_{t-})}dB_t + \int_{\mathbb{R}^d} \xi(\mu^X(d\xi, dt) - \boldsymbol{K}(\mathbb{X}_{t-}, d\xi)dt),$$

  with $\boldsymbol{b}$, $\boldsymbol{a}$ and $\mathbf{x} \mapsto \int_{\mathbb{R}^d} \langle \epsilon_I, \xi^{\otimes I} \rangle \boldsymbol{K}(\mathbf{x}, d\xi)$ linear functions that depend on the signature up to order 1, 2 and $|I| \leq N$ respectively, then $(\mathbb{X}_t^N)_{t \geq 0}$ is a finite dimensional polynomial process on $G^N(\mathbb{R}^d)$.
- This holds true in particular for X being a classical polynomial process.
- The expected truncated signature can then be computed by solving a finite dimensional linear ODE, i.e. computing a matrix exponential.
- If $\boldsymbol{b}$, $\boldsymbol{a}$ and $\mathbf{x} \mapsto \int_{\mathbb{R}^d} \langle \epsilon_I, \xi^{\otimes I} \rangle \boldsymbol{K}(\mathbf{x}, d\xi)$ are functions corresponding to certain closures of linear maps (analogous to real analytic maps), then an infinite dimensional linear ODE (instead of a infinite dimensional PIDE) has to be solved.

Conclusion

- Universal approximation theorem for continuous functions of càdlàg paths
- Lévy type signature models distinguish themselves in
  - universality, as the characteristics of classical models can be approximated;
  - efficient pricing, hedging and calibration (through regression for time-series data and expected signature for options).
- Generic classes of jump SDEs, so-called jump SigSDEs, can be proved to be projections of (infinite dimensional) polynomial processes.
- For certain jump SigSDEs, in particular polynomial processes themselves, the truncated signature is again a finite dimensional polynomial process.

Thank you for your attention!