From Lévy's stochastic area formula to universality of affine and polynomial processes

Christa Cuchiero

University of Vienna
Department of Statistics and Operations Research

International Seminar on SDEs and Related Topics

March 10, 2023

Outline

Part I An overview of affine and polynomial processes by means of Lévy's stochastic area formula

Outline

Part I An overview of affine and polynomial processes by means of Lévy's stochastic area formula

Part II Infinite dimensional Wishart processes

Outline

Part I An overview of affine and polynomial processes by means of Lévy's stochastic area formula

Part II Infinite dimensional Wishart processes

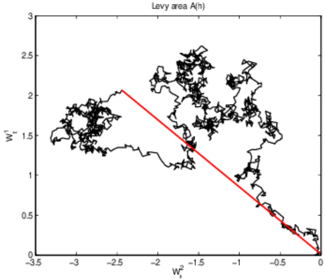
Part III Signature SDEs from an affine and polynomial perspective

Part I

An overview of affine and polynomial processes by means of Lévy's stochastic area formula

Lévy's stochastic area

- In the article "Le mouvement Brownien plan" (1940), P. Lévy began studying what he called the "stochastic area", i.e., the signed area enclosed by the trajectory of a 2-dimensional Brownian motion W and its chord.



Source: S. J. Malham, Anke Wiese

- In formulas, up to a factor of $\frac{1}{2}$ Lévy's stochastic area is thus given by

$$L_t := \int_0^t W_s^1 dW_s^2 - W_s^2 dW_s^1, \quad t \geq 0.$$

Lévy's stochastic area formula

In subsequent papers, P. Lévy then answered the following ...

Lévy's stochastic area formula

In subsequent papers, P. Lévy then answered the following ...

Questions:

- What is the characteristic function of L_t, i.e., $\mathbb{E}[e^{\mathrm{i}\lambda L_t}]$ for $\lambda \in \mathbb{R}$.
- What is the conditional characteristic function of L_t given W_t, i.e., $\mathbb{E}[e^{\mathrm{i}\lambda L_t}|W_t = y]$ for $y \in \mathbb{R}^2$ and $\lambda \in \mathbb{R}$?

Lévy's stochastic area formula

In subsequent papers, P. Lévy then answered the following ...

Questions:

- What is the characteristic function of L_t, i.e., $\mathbb{E}[e^{\mathrm{i}\lambda L_t}]$ for $\lambda \in \mathbb{R}$.
- What is the conditional characteristic function of L_t given W_t, i.e., $\mathbb{E}[e^{\mathrm{i}\lambda L_t} | W_t = y]$ for $y \in \mathbb{R}^2$ and $\lambda \in \mathbb{R}$?

- Lévy's formula was generalized and studied by many authors, see e.g., M. Yor (1980), Helmes and Schwane (1983) and the references therein.
- Alternatively to these proofs the above formulas can be derived via the theory of affine processes, in spirit of C.C., S. Svaluto-Ferro & J. Teichmann, "Signature SDEs from an affine and polynomial perspective" ('23).

Definition of affine diffusion processes

Simplest setting (for illustrative purposes): Itô diffusion with state space $S \subseteq \mathbb{R}^d$.

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

with $a : \mathbb{R}^d \to S^+(\mathbb{R}^d)$ and $b : \mathbb{R}^d \to \mathbb{R}^d$ continuous functions and W a Brownian motion on $\mathbb{R}^d$.

Definition of affine diffusion processes

Simplest setting (for illustrative purposes): Itô diffusion with state space $S \subseteq \mathbb{R}^d$.

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

with $a : \mathbb{R}^d \to S^+(\mathbb{R}^d)$ and $b : \mathbb{R}^d \to \mathbb{R}^d$ continuous functions and W a Brownian motion on $\mathbb{R}^d$.

Definition

A weak solution X of $(*)$ is called affine process if b and a are affine functions, i.e.,

$$b(x) = b + \sum_{i=1}^{d} \beta_i x_i, \quad a(x) = a + \sum_{i=1}^{d} \alpha_i x_i,$$

for characteristics b, $\beta_i \in \mathbb{R}^d$ and $a, \alpha_i \in \mathbb{R}^{d \times d}$.

Definition of affine diffusion processes

Simplest setting (for illustrative purposes): Itô diffusion with state space $S \subseteq \mathbb{R}^d$.

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

with $a : \mathbb{R}^d \to S^+(\mathbb{R}^d)$ and $b : \mathbb{R}^d \to \mathbb{R}^d$ continuous functions and W a Brownian motion on $\mathbb{R}^d$.

Definition

A weak solution X of $(*)$ is called affine process if b and a are affine functions, i.e.,

$$b(x) = b + \sum_{i=1}^{d} \beta_i x_i, \quad a(x) = a + \sum_{i=1}^{d} \alpha_i x_i,$$

for characteristics b, $\beta_i \in \mathbb{R}^d$ and $a, \alpha_i \in \mathbb{R}^{d \times d}$.

From this definition affine processes appear as a narrow class, whose universal character announced in the title of this talk is at this stage by no means visible.

Affine transform formula - Riccati ODEs

The remarkable implication is that exponential moments, i.e., $\mathbb{E}[\exp(\langle u, X_t \rangle)]$ for $u \in \mathbb{C}^d$, can be expressed as solutions of Riccati ordinary differential equations (ODEs).

Theorem (D. Duffie, D. Filipovic & W. Schachermayer ('03), C.C. & J. Teichmann ('13))

Let $(X_t)_{t \geq 0}$ be an affine process and let $u \in \mathbb{C}^d$ such that $\mathbb{E}[\exp(|\langle u, X_t \rangle|)] < \infty$. Then,

$$\mathbb{E}_x \left[\exp(\langle u, X_t \rangle)\right] = \Phi(t) \exp(\langle \psi(t), x \rangle),$$

where Φ and ψ solve the Riccati ODEs given by

$$\partial_t \Phi(t) = \Phi(t) F(\psi_t), \quad \Phi(0) = 1, \qquad \partial_t \psi(t) = R(\psi(t)), \quad \psi(0) = u,$$

where

$$F(u) = b^\top u + u^\top a u, \qquad R_i = \beta_i^\top u + u^\top \alpha_i u, \quad i = 1, \dots, d.$$

Back to Lévy's stochastic area formula

The affine transform formula is thus tailor-made to compute the characteristic function of the Lévy stochastic area L, if we can embed it within an affine process.

Lemma

Let W be a 2-dimensional Brownian motion and consider the 4-dimensional process $(X_t)_{t\geq 0} = (x_1 + W_t^1, x_2 + W_t^2, x_3 + L_t, \|(x_1, x_2)^\top + W_t\|^2)_{t\geq 0}$.

Then X is an affine process with initial value $x = (x_1, x_2, x_3, \|(x_1, x_2)^\top\|^2) \in \mathbb{R}^4$ *and characteristics*

$$a = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad b = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 2 \end{pmatrix},$$

$$\alpha_1 = \begin{pmatrix} 0 & 0 & 0 & 2 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 2 & 0 & 0 & 0 \end{pmatrix}, \quad \alpha_2 = \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 2 \\ -1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \end{pmatrix}, \quad \alpha_4 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix},$$

and all others are 0.

Back to Lévy's stochastic area formula

The affine transform formula is thus tailor-made to compute the characteristic function of the Lévy stochastic area L, if we can embed it within an affine process.

Lemma

Let W be a 2-dimensional Brownian motion and consider the 4-dimensional process $(X_t)_{t\geq 0} = (x_1 + W_t^1, x_2 + W_t^2, x_3 + L_t, \|(x_1, x_2)^\top + W_t\|^2)_{t\geq 0}$.

Then X is an affine process with initial value $x = (x_1, x_2, x_3, \|(x_1, x_2)^\top\|^2) \in \mathbb{R}^4$ *and characteristics*

$$a = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad b = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 2 \end{pmatrix},$$

$$\alpha_1 = \begin{pmatrix} 0 & 0 & 0 & 2 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 2 & 0 & 0 & 0 \end{pmatrix}, \quad \alpha_2 = \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 2 \\ -1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \end{pmatrix}, \quad \alpha_4 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix},$$

and all others are 0.

Key idea: lift the process of interest to a higher-dimensional state to make it affine

Lévy's stochastic area formula via affine processes

To compute $\mathbb{E}[e^{\mathrm{i}\lambda L_t}]$, set $\mathbb{R}^4 \ni u = (0, 0, \mathrm{i}\lambda, 0)$. Then the Riccati ODEs reduce to $\psi_1 = \psi_2 = 0$, $\psi_3 = \mathrm{i}\lambda$ and

$$\partial_t \psi_4(t) = \frac{1}{2}(4(\psi_4(t))^2 - \lambda^2),\, \psi_4 = 0, \quad \partial_t \Phi(t) = 2\Phi(t)\psi_4(t),\, \Phi(t) = 1,$$

whose solutions are given by

$$\psi_4(t) = -\frac{\lambda \tanh(\lambda t)}{2}, \quad \Phi(t) = \frac{1}{\cosh(\lambda t)}.$$

Theorem

The characteristic function of $x_3 + L_t$ is given by

$$\mathbb{E}_x[e^{i\lambda(x_3+L_t)}] = \frac{1}{\cosh(\lambda t)} \exp(x_3 i\lambda + \|(x_1, x_2)^\top\|^2 \psi_4), \quad \lambda \in \mathbb{R}.$$

By setting $x = 0$, we then get the first Lévy stochastic area formula

$$\mathbb{E}[e^{i\lambda L_t}] = \frac{1}{\cosh(\lambda t)}, \quad \lambda \in \mathbb{R}.$$

Lévy's stochastic area formula via affine processes

- For the second one, we can compute the joint characteristic function $\mathbb{E}_0[e^{\mathrm{i}\lambda L_t + \mathrm{i}\langle v, W_t\rangle}]$ by solving additionally to ψ_4 from above the "Riccati" ODEs

$$\begin{pmatrix}\partial_t\psi_1(t)\\ \partial_t\psi_2(t)\end{pmatrix} = \begin{pmatrix}4\psi_4(t) & 2\mathrm{i}\lambda\\ -2\mathrm{i}\lambda & 4\psi_4(t)\end{pmatrix}\begin{pmatrix}\psi_1(t)\\ \psi_2(t)\end{pmatrix}, \qquad \begin{pmatrix}\psi_1(0)\\ \psi_2(0)\end{pmatrix} = \mathrm{i}v,$$
$$\partial_t\Phi(t) = \Phi(t)(\psi_1^2(t) + \psi_2^2(t) + 2\psi_4(t)), \qquad \Phi(0) = 1,$$

  which yields

$$\mathbb{E}_0[e^{\mathrm{i}\lambda L_t + \mathrm{i}\langle v, W_t\rangle}] = \frac{1}{\cosh(\lambda t)}\exp\Big(-\frac{\|v\|^2}{2\lambda\coth(\lambda t)}\Big)$$

- Since $\mathbb{E}_0[e^{\mathrm{i}\lambda L_t + \mathrm{i}\langle v, W_t\rangle}] = \int_{\mathbb{R}^2} e^{\mathrm{i}\langle v,y\rangle}\mathbb{E}[e^{\mathrm{i}\lambda L_t} | W_t = y]\frac{1}{2\pi t}e^{-\frac{1}{2t}\|y\|^2}dy$ holds, Fourier inversion yields...

Lévy's stochastic area formula via affine processes

- For the second one, we can compute the joint characteristic function $\mathbb{E}_0[e^{\mathrm{i}\lambda L_t + \mathrm{i}\langle v, W_t\rangle}]$ by solving additionally to ψ_4 from above the "Riccati" ODEs

$$\begin{pmatrix} \partial_t \psi_1(t) \\ \partial_t \psi_2(t) \end{pmatrix} = \begin{pmatrix} 4\psi_4(t) & 2\mathrm{i}\lambda \\ -2\mathrm{i}\lambda & 4\psi_4(t) \end{pmatrix} \begin{pmatrix} \psi_1(t) \\ \psi_2(t) \end{pmatrix}, \qquad \begin{pmatrix} \psi_1(0) \\ \psi_2(0) \end{pmatrix} = \mathrm{i}v,$$

$$\partial_t \Phi(t) = \Phi(t)(\psi_1^2(t) + \psi_2^2(t) + 2\psi_4(t)), \qquad \Phi(0) = 1,$$

  which yields

$$\mathbb{E}_0[e^{\mathrm{i}\lambda L_t + \mathrm{i}\langle v, W_t\rangle}] = \frac{1}{\cosh(\lambda t)} \exp(-\frac{\|v\|^2}{2\lambda \coth(\lambda t)})$$

- Since $\mathbb{E}_0[e^{\mathrm{i}\lambda L_t + \mathrm{i}\langle v, W_t\rangle}] = \int_{\mathbb{R}^2} e^{\mathrm{i}\langle v, y\rangle} \mathbb{E}[e^{\mathrm{i}\lambda L_t} | W_t = y] \frac{1}{2\pi t} e^{-\frac{1}{2t}\|y\|^2} dy$ holds, Fourier inversion yields...

Theorem

$$\mathbb{E}[e^{i\lambda L_t} | W_t = y] = \frac{\lambda t}{\sinh \lambda t} \exp\left(\|y\|^2 \frac{1 - \lambda t \coth(\lambda t)}{2t} \right), \quad \lambda \in \mathbb{R}.$$

Definition of polynomial diffusion processes

As any affine diffusion process is also a polynomial process, moments of the Lévy area can be computed by polynomial technology.

Definition

A weak solution X of $(*)$ is called polynomial process if b is affine and a quadratic.

- In this finite dimensional diffusion framework polynomial processes are always more general than affine ones. This does not necessarily hold true in the presence of jumps.
- As we shall see, in certain infinite dimensional setups the notions of affine and polynomial diffusion processes coincide.

Definition of polynomial diffusion processes

As any affine diffusion process is also a polynomial process, moments of the Lévy area can be computed by polynomial technology.

Definition

A weak solution X of $(*)$ is called polynomial process if b is affine and a quadratic.

- In this finite dimensional diffusion framework polynomial processes are always more general than affine ones. This does not necessarily hold true in the presence of jumps.
- As we shall see, in certain infinite dimensional setups the notions of affine and polynomial diffusion processes coincide.
- Denote by $\mathcal{P}_k$ polynomials on $S \subseteq \mathbb{R}^d$ up to degree $k \in \mathbb{N}$, i.e. $\mathcal{P}_k = \{x \mapsto \sum_{|\mathbf{i}|=0}^{k} u_{\mathbf{i}} x^{\mathbf{i}} \mid u_{\mathbf{i}} \in \mathbb{R}\}$, where we use multi-index notation $\mathbf{i} = (i_1, \ldots, i_n) \in \mathbb{N}^n$, $|\mathbf{i}| = i_1 + \cdots + i_n$ and $x^{\mathbf{i}} = x^{i_1} \cdots x^{i_n}$. The dimension of $\mathcal{P}_k$ is denoted by N.

Moment formula

- We write $u \in \mathbb{R}^N$ for the coefficients vector and define $p(x,u) := \sum_{|\mathbf{i}|=0}^{k} u_{\mathbf{i}} x^{\mathbf{i}}$.
- Note that for a polynomial process, the generator $\mathcal{A}$ maps $\mathcal{P}_k$ to $\mathcal{P}_k$, i.e. polynomials to polynomials of same or smaller degree.
- Hence there is a linear map L_N from $\mathbb{R}^N$ to $\mathbb{R}^N$ such that

$$\mathcal{A}(p(\cdot,u))(x) = p(x, L_N u).$$

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

Let $(X_t)_{t\in[0,T]}$ be a polynomial process. Denote by $c(t)$ the solution of the linear ODE given by

$$\partial_t c(t) = L_N c(t), \quad c(0) = u \in \mathbb{R}^N.$$

Then,

$$\mathbb{E}_x\Big[\sum_{|\mathbf{i}|=0}^{k} u_{\mathbf{i}} X_t^{\mathbf{i}}\Big] = \sum_{|\mathbf{i}|=0}^{k} c_{\mathbf{i}}(t) x^{\mathbf{i}}.$$

Affine and polynomial processes as universal model class?

- Despite the rather narrow definition of affine and polynomial processes, already the finite dimensional setting contains many well-known processes, e.g.,
  - Ornstein-Uhlenbeck, Feller-type and Wishart processes, the Black-Scholes and the Heston model, the Fisher-Snedecor process, the Wright-Fisher diffusion as well as all possible combinations thereof.

Affine and polynomial processes as universal model class?

- Despite the rather narrow definition of affine and polynomial processes, already the finite dimensional setting contains many well-known processes, e.g.,
  - Ornstein-Uhlenbeck, Feller-type and Wishart processes, the Black-Scholes and the Heston model, the Fisher-Snedecor process, the Wright-Fisher diffusion as well as all possible combinations thereof.
- As we shall see, the true universal character becomes visible in infinite dimensional setups.
- Infinite dimensional affine and polynomial processes appear either as infinite dimensional analogs of the finite dimensional ones, usually with a much more intricate structure, or as lifts, in spirit of the lift of the Lévy area.

Affine and polynomial processes as universal model class?

- Despite the rather narrow definition of affine and polynomial processes, already the finite dimensional setting contains many well-known processes, e.g.,
  - Ornstein-Uhlenbeck, Feller-type and Wishart processes, the Black-Scholes and the Heston model, the Fisher-Snedecor process, the Wright-Fisher diffusion as well as all possible combinations thereof.
- As we shall see, the true universal character becomes visible in infinite dimensional setups.
- Infinite dimensional affine and polynomial processes appear either as infinite dimensional analogs of the finite dimensional ones, usually with a much more intricate structure, or as lifts, in spirit of the lift of the Lévy area.
  - Infinite dimensional analogs can often be realized as measure valued or Hilbert space valued processes.
  - Markovian lifts appear for instance as lifts of stochastic Volterra processes or signature lifts.

Infinite dimensional examples and their applications

- Measure-valued affine and polynomial processes:
  - Most prominent examples: Dawson-Watanabe and Flemming-Viot type processes
  - Measure-valued branching processes in the sense of Z. Li
  - Characterization of (probability) measure-valued affine and polynomial diffusions: C.C., L. Di Persio, F. Guida & S. Svaluto-Ferro ('21) and C.C., M. Larsson & S. Svaluto-Ferro ('19).
  - Applications: population genetics, chemistry, energy market modeling and stochastic portfolio theory

Infinite dimensional examples and their applications

- Measure-valued affine and polynomial processes:
  - Most prominent examples: Dawson-Watanabe and Flemming-Viot type processes
  - Measure-valued branching processes in the sense of Z. Li
  - Characterization of (probability) measure-valued affine and polynomial diffusions: C.C., L. Di Persio, F. Guida & S. Svaluto-Ferro ('21) and C.C., M. Larsson & S. Svaluto-Ferro ('19).
  - Applications: population genetics, chemistry, energy market modeling and stochastic portfolio theory
- Hilbert space valued processes:
  - e.g., T. Schmidt, S. Tappe and W. Yu ('20), S. Cox, S.Karbach & A. Khedher ('22) or C.C. & S. Svaluto-Ferro ('21)
  - S. Cox, C.C., A. Khedher ('23): Infinite dimensional Wishart processes
  - Applications: Infinite dim. covariance modeling, limits of random matrices (as e.g. in C. Bertucci, M. Debbah, J.-M. Lasry, and P.-L. Lions ('22))

Infinite dimensional examples and their applications

- Measure-valued affine and polynomial processes:
  - Most prominent examples: Dawson-Watanabe and Flemming-Viot type processes
  - Measure-valued branching processes in the sense of Z. Li
  - Characterization of (probability) measure-valued affine and polynomial diffusions: C.C., L. Di Persio, F. Guida & S. Svaluto-Ferro ('21) and C.C., M. Larsson & S. Svaluto-Ferro ('19).
  - Applications: population genetics, chemistry, energy market modeling and stochastic portfolio theory
- Hilbert space valued processes:
  - e.g., T. Schmidt, S. Tappe and W. Yu ('20), S. Cox, S.Karbach & A. Khedher ('22) or C.C. & S. Svaluto-Ferro ('21)
  - S. Cox, C.C., A. Khedher ('23): Infinite dimensional Wishart processes
  - Applications: Infinite dim. covariance modeling, limits of random matrices (as e.g. in C. Bertucci, M. Debbah, J.-M. Lasry, and P.-L. Lions ('22))

Infinite dimensional examples and their applications

- Markovian lifts of Volterra processes:
  - E. Abi Jaber & O. El Euch ('19): Markovian structure of the Volterra Heston model
  - C.C. & J. Teichmann ('20): Generalized Feller processes and Markovian lifts of stochastic Volterra processes: the affine case
  - Applications: Rough volatility modeling and forward curve modeling

Infinite dimensional examples and their applications

- Markovian lifts of Volterra processes:
  - E. Abi Jaber & O. El Euch ('19): Markovian structure of the Volterra Heston model
  - C.C. & J. Teichmann ('20): Generalized Feller processes and Markovian lifts of stochastic Volterra processes: the affine case
  - Applications: Rough volatility modeling and forward curve modeling
- Signature lifts:
  - I. Arribas, C. Salvi & L. Szpruch ('20): Sig-SDEs model for quantitative finance
  - C.C., S. Svaluto-Ferro & J. Teichmann ('23): Signature SDEs from an affine and polynomial perspective
  - Applications: Signature-based models, see e.g. C.C., G. Gazzani, J. Möller & S. Svaluto-Ferro ('23)

Infinite dimensional examples and their applications

- Markovian lifts of Volterra processes:
  - E. Abi Jaber & O. El Euch ('19): Markovian structure of the Volterra Heston model
  - C.C. & J. Teichmann ('20): Generalized Feller processes and Markovian lifts of stochastic Volterra processes: the affine case
  - Applications: Rough volatility modeling and forward curve modeling
- Signature lifts:
  - I. Arribas, C. Salvi & L. Szpruch ('20): Sig-SDEs model for quantitative finance
  - C.C., S. Svaluto-Ferro & J. Teichmann ('23): Signature SDEs from an affine and polynomial perspective
  - Applications: Signature-based models, see e.g. C.C., G. Gazzani, J. Möller & S. Svaluto-Ferro ('23)

Part II

Infinite dimensional Wishart processes

based on joint works with S. Cox and A. Khedher ('23)

Infinite dimensional Wishart processes

- Finite dimensional Wishart processes, introduced by F. Bru ('91), are affine processes taking values in the cone of positive semidefinite matrices $S^+(\mathbb{R}^n)$.
- The infinite dimensional analog of $S^+(\mathbb{R}^n)$ is $S_1^+(H)$, the cone of positive self-adjoint trace class operators on a separable real Hilbert space H.

Infinite dimensional Wishart processes

- Finite dimensional Wishart processes, introduced by F. Bru ('91), are affine processes taking values in the cone of positive semidefinite matrices $S^+(\mathbb{R}^n)$.
- The infinite dimensional analog of $S^+(\mathbb{R}^n)$ is $S_1^+(H)$, the cone of positive self-adjoint trace class operators on a separable real Hilbert space H.
- We define an infinite-dimensional Wishart process as an $S_1^+(H)$-valued solution (in an appropriate sense) to the following SDE:

$$dX_t = (\alpha Q + X_t A + A^* X_t)\, dt + \sqrt{X_t}\, dW_t \sqrt{Q} + \sqrt{Q}\, dW_t^* \sqrt{X_t}, \quad X_0 = x,$$

  - $\alpha \in \mathbb{R}$,
  - $A \colon D(A) \subset H \to H$ is the generator of a C_0-semigroup,
  - x and Q are a positive self-adjoint bounded operators,
  - $(W_t)_{t \geq 0}$ is an $L_2(H)$-cylindrical Brownian motion (where $L_2(H)$ is the space of Hilbert Schmidt operators on H).

Infinite dimensional Wishart processes

- Finite dimensional Wishart processes, introduced by F. Bru ('91), are affine processes taking values in the cone of positive semidefinite matrices $S^+(\mathbb{R}^n)$.
- The infinite dimensional analog of $S^+(\mathbb{R}^n)$ is $S_1^+(H)$, the cone of positive self-adjoint trace class operators on a separable real Hilbert space H.
- We define an infinite-dimensional Wishart process as an $S_1^+(H)$-valued solution (in an appropriate sense) to the following SDE:

$$dX_t = (\alpha Q + X_t A + A^* X_t)\, dt + \sqrt{X_t}\, dW_t \sqrt{Q} + \sqrt{Q}\, dW_t^* \sqrt{X_t}, \quad X_0 = x,$$

  - $\alpha \in \mathbb{R}$,
  - $A\colon D(A) \subset H \to H$ is the generator of a C_0-semigroup,
  - x and Q are a positive self-adjoint bounded operators,
  - $(W_t)_{t \geq 0}$ is an $L_2(H)$-cylindrical Brownian motion (where $L_2(H)$ is the space of Hilbert Schmidt operators on H).
- Note that the affine structure is analogous to the finite dimensional case and visible from the characteristics.

Questions and challenges

Questions:

- Under which parameter conditions do infinite dimensional Wishart processes exist?
- Do infinite rank Wishart processes exist?
- What is the evolution of the eigenvalues?

Questions and challenges

Questions:

- Under which parameter conditions do infinite dimensional Wishart processes exist?
- Do infinite rank Wishart processes exist?
- What is the evolution of the eigenvalues?

- To explain the challenges consider the finite-dimensional setting, with $A \in \mathbb{R}^{n\times n}$, $Q \in S^+(\mathbb{R}^n)$, W is a standard $\mathbb{R}^{n\times n}$-valued Brownian motion.
- If Q is injective, then a finite-dimensional Wishart process exists if and only if either
$$\alpha \in [n-1, \infty) \text{ or } \alpha \in \{0, \dots, n-2\}$$
and $\mathrm{rank}(x) \leq \alpha$. In case of the latter one has $\mathrm{rank}(X_t) \leq \alpha$ a.s. for all $t \geq 0$; see P. Graczyk, J. Malecki, and E. Mayerhofer ('18).
- When translated to the infinite-dimensional setting this suggests that Wishart processes of infinite rank might not exist.

Main result

Theorem (S. Cox, C.C. & A. Khedher ('23))

If Q and A are such that $\int_0^t \|e^{sA}\sqrt{Q}\|_{L_2(H)}\, ds < \infty$ for all $t > 0$ and if additionally

- *Q is injective and*
- *there exists a $t > 0$ such that the semigroup e^{tA} is injective,*

then an analytically and probabilistically weak solution to the Wishart SDE exists if and only if

$$\alpha \in \mathbb{N} \text{ and } \operatorname{rank}(x) \leq \alpha.$$

Main result

Theorem (S. Cox, C.C. & A. Khedher ('23))

If Q and A are such that $\int_0^t \|e^{sA}\sqrt{Q}\|_{L_2(H)}\, ds < \infty$ for all $t > 0$ and if additionally

- *Q is injective and*
- *there exists a $t > 0$ such that the semigroup e^{tA} is injective,*

then an analytically and probabilistically weak solution to the Wishart SDE exists if and only if

$$\alpha \in \mathbb{N} \text{ and } \operatorname{rank}(x) \leq \alpha.$$

In this case, $\operatorname{rank}(X_t) = \alpha$ a.s. for almost all $t > 0$, the solution is an affine process, whose Laplace transform is given by

$$\mathbb{E}_x[\exp(-\operatorname{Tr}(uX_t)] = \det(I_H + 2\sqrt{u}Q_t\sqrt{u})^{-\frac{\alpha}{2}} e^{\left(-\operatorname{Tr}\left(e^{tA}\sqrt{u}(I_H+2\sqrt{u}Q_t\sqrt{u})^{-1}\sqrt{u}e^{tA^*}x\right)\right)}$$

with $Q_t = \int_0^t e^{sA^*} Q e^{sA}\, ds$.

Main result

Theorem (S. Cox, C.C. & A. Khedher ('23))

If Q and A are such that $\int_0^t \|e^{sA}\sqrt{Q}\|_{L_2(H)}\, ds < \infty$ for all $t > 0$ and if additionally

- *Q is injective and*
- *there exists a $t > 0$ such that the semigroup e^{tA} is injective,*

then an analytically and probabilistically weak solution to the Wishart SDE exists if and only if

$$\alpha \in \mathbb{N} \text{ and } \operatorname{rank}(x) \leq \alpha.$$

In this case, $\operatorname{rank}(X_t) = \alpha$ a.s. for almost all $t > 0$, the solution is an affine process, whose Laplace transform is given by

$$\mathbb{E}_x[\exp(-\operatorname{Tr}(uX_t)] = \det(I_H + 2\sqrt{u}Q_t\sqrt{u})^{-\frac{\alpha}{2}} e^{\left(-\operatorname{Tr}\left(e^{tA}\sqrt{u}(I_H+2\sqrt{u}Q_t\sqrt{u})^{-1}\sqrt{u}e^{tA^*}x\right)\right)}$$

with $Q_t = \int_0^t e^{sA^*} Q e^{sA}\, ds$.

Remark: If A is bounded or selfadjoint, the injectivity condition on e^{tA} is satisfied.

Eigenvalue equation

The following eigenvalue equation generalizes also the finite dimensional results.

Theorem (S. Cox, C.C. & A. Khedher ('23))

Under the above assumptions, each eigenvalue λ^i for $i = 1, \dots, \alpha$ satisfies

$$d\lambda_t^i = 2\sqrt{\lambda_t^i (V_t^* Q V_t)_{ii}}\, dB_t^i + \mathrm{Tr}(Q) dt + 2\lambda_t^i (V_t^* A V_t)_{ii} dt + \sum_{k \neq i} \frac{1}{\lambda_t^i - \lambda_t^k} (\lambda_t^k (V_t^* Q V_t)_{kk} + \lambda_t^i (V_t^* Q V_t)_{ii}) dt,$$

until the stopping time where they collide. Here, V denotes the orthonormal operator containing the eigenvectors of X and B is an α-dim. Brownian motion.

Eigenvalue equation

The following eigenvalue equation generalizes also the finite dimensional results.

Theorem (S. Cox, C.C. & A. Khedher ('23))

Under the above assumptions, each eigenvalue λ^i for $i = 1, \dots, \alpha$ satisfies

$$d\lambda_t^i = 2\sqrt{\lambda_t^i (V_t^* Q V_t)_{ii}} dB_t^i + \mathrm{Tr}(Q)dt + 2\lambda_t^i (V_t^* A V_t)_{ii} dt + \sum_{k \neq i} \frac{1}{\lambda_t^i - \lambda_t^k} (\lambda_t^k (V_t^* Q V_t)_{kk} + \lambda_t^i (V_t^* Q V_t)_{ii}) dt,$$

until the stopping time where they collide. Here, V denotes the orthonormal operator containing the eigenvectors of X and B is an α-dim. Brownian motion.

Remaining open questions:

- Do infinite rank Wishart processes exist when the injectivity conditions are not satisfied?
- How does the finite dimensional characterization of non-central Wishart distributions look like when Q is not of full rank?

Part III

Signature SDEs from an affine and polynomial perspective

based on joint work with S. Svaluto-Ferro and J. Teichmann ('23)

Signature SDEs from an affine and polynomial perspective

Research question:

- Prove universal approximation features of the affine and polynomial model class in the space of all stochastic processes driven by, say, Brownian motion.

Signature SDEs from an affine and polynomial perspective

Research question:

- Prove universal approximation features of the affine and polynomial model class in the space of all stochastic processes driven by, say, Brownian motion.

- One step in this direction by "linearizing" generic classes of stochastic processes called signature SDEs
- Develop an affine and polynomial theory for their prolongations, being the process' signature
- Introduce a novel and proper notion of entire and real-analytic functions on group-like elements, being the state space of the signature process
  - $\Rightarrow$ Essentially all real-analytic path-dependent characteristics become power (or rather linear) series in the signature components.
- Analysis of class that is universal within Itô processes with path-dependent characteristics and whose full law on path space can be characterized via the explicitly computable Fourier-Laplace transform

The signature of a continuous semimartingale

- Signature goes back to K. Chen ('57) and plays a prominent in rough path theory (T. Lyons ('98), P. Friz & N. Victoir ('10), P. Friz & M. Hairer ('14)).
- For an $\mathbb{R}^d$-valued continuous semimartingale X, its signature is given by iterated Stratonovich integrals, i.e.,

$$\mathbb{X}_t := \Big(1, \int_0^t \circ dX_s, \int_0^t \int_0^{s_2} \circ dX_{s_1} \otimes \circ dX_{s_2}, \dots, \\ \dots, \int_0^t \int_0^{s_n} \cdots \int_0^{s_2} \circ dX_{s_1} \otimes \cdots \otimes \circ dX_{s_n}, \dots \Big),$$

 which is an element in the extended tensor algebra $T((\mathbb{R}^d))$.
- Example (d=2)

$$\mathbb{X}_t = \left(1, \begin{pmatrix} X_t^1 - X_0^1 \\ X_t^2 - X_0^2 \end{pmatrix}, \begin{pmatrix} \frac{1}{2}(X_t^1 - X_0^1)^2 & \int_0^t \int_0^s \circ dX_u^1 \circ dX_s^2 \\ \int_0^t \int_0^s \circ dX_u^2 \circ dX_s^1 & \frac{1}{2}(X_t^2 - X_0^2)^2 \end{pmatrix}, \dots \right)$$

The signature of a continuous semimartingale

- Signature goes back to K. Chen ('57) and plays a prominent in rough path theory (T. Lyons ('98), P. Friz & N. Victoir ('10), P. Friz & M. Hairer ('14)).
- For an $\mathbb{R}^d$-valued continuous semimartingale X, its signature is given by iterated Stratonovich integrals, i.e.,

$$\mathbb{X}_t := \Big(1, \int_0^t \circ dX_s, \int_0^t \int_0^{s_2} \circ dX_{s_1} \otimes \circ dX_{s_2}, \dots,$$
$$\dots, \int_0^t \int_0^{s_n} \cdots \int_0^{s_2} \circ dX_{s_1} \otimes \cdots \otimes \circ dX_{s_n}, \dots \Big),$$

which is an element in the extended tensor algebra $T((\mathbb{R}^d))$.

- Example (d=2)

$$\mathbb{X}_t = \left(1, \begin{pmatrix} X_t^1 - X_0^1 \\ X_t^2 - X_0^2 \end{pmatrix}, \begin{pmatrix} \frac{1}{2}(X_t^1 - X_0^1)^2 & \int_0^t \int_0^s \circ dX_u^1 \circ dX_s^2 \\ \int_0^t \int_0^s \circ dX_u^2 \circ dX_s^1 & \frac{1}{2}(X_t^2 - X_0^2)^2 \end{pmatrix}, \dots \right)$$

- For a semimartingale with state space $S \subseteq \mathbb{R}^d$, we denote by $\mathcal{S}(S)$ the set of so-called group like elements of $T((\mathbb{R}^d))$ whose first level lies in S (i.e. the correct state space of the signature).

Linear functions on signature and universal approximation

- For a multi-index $I = \{i_1, ..., i_m\} \in \{1, \ldots, d\}^m$ we denote by $e_I := e_{i_1} \otimes ... \otimes e_{i_m}$ the basis elements of $(\mathbb{R}^d)^{\otimes m}$.
- We call

$$L(\mathbb{X}_t) = \sum_{0 \leq |I| \leq n} \alpha_I \langle e_I, \mathbb{X}_t \rangle \text{ for } n \in \mathbb{N}$$

with $\alpha_I \in \mathbb{R}$, linear functions of the signature.

Linear functions on signature and universal approximation

- For a multi-index $I = \{i_1, ..., i_m\} \in \{1, \dots, d\}^m$ we denote by $e_I := e_{i_1} \otimes ... \otimes e_{i_m}$ the basis elements of $(\mathbb{R}^d)^{\otimes m}$.
- We call

$$L(\mathbb{X}_t) = \sum_{0 \leq |I| \leq n} \alpha_I \langle e_I, \mathbb{X}_t \rangle \text{ for } n \in \mathbb{N}$$

with $\alpha_I \in \mathbb{R}$, linear functions of the signature.

Key properties to obtain a Universal Approximation Theorem (UAT) for linear functions of the signature

- Point-separation: for $\hat{X}_t := (t, X_t)$, its signature $\hat{\mathbb{X}}_T$ determines $(\hat{X}_t)_{t \in [0,T]}$ uniquely.
- Algebra: the product of linear functions of the signature is again a linear function of the signature, precisely $\langle e_I, \mathbb{X}_t \rangle \langle e_J, \mathbb{X}_t \rangle = \langle e_I ⧢ e_J, \mathbb{X}_t \rangle$.

$\Rightarrow$ Use Stone-Weierstrass Theorem to approximate continuous (with respect to a certain p-variation norm) path functionals $f(X_{[0,t]})$ via $L(\hat{\mathbb{X}}_t)$ uniformly on compact sets of paths.

Entire functions of the signature

- For $\mathbf{x} \in \mathcal{S}(S)$ and $\mathbf{u} \in T((\mathbb{R}^d)) + \mathrm{i}\, T((\mathbb{R}^d))$ set $|\mathbf{u}|_{\mathbf{x}} = \sum_{n=0}^{\infty} |\langle \pi_n(\mathbf{u}), \pi_n(\mathbf{x}) \rangle|$, where π_n denotes the projection on $(\mathbb{R}^d)^{\otimes n}$.
- Dual elements:
  $\mathcal{S}(S)^* := \{\mathbf{u} \in T((\mathbb{R}^d)) + \mathrm{i}\, T((\mathbb{R}^d)) \colon |\mathbf{u}|_{\mathbf{x}} < \infty \text{ for all } \mathbf{x} \in \mathcal{S}(S)\}$.
- For $\mathbf{u} \in \mathcal{S}(S)^*$ entire maps of group like elements are defined as

$$\mathcal{S}(S) \ni \mathbf{x} \mapsto \langle \mathbf{u}, \mathbf{x} \rangle := \lim_{N \to \infty} \sum_{n=0}^{N} \langle \pi_n(\mathbf{u}), \pi_n(\mathbf{x}) \rangle.$$

Products of entire functions are again entire functions which extends the algebra property of linear functions of the signature expressed via the shuffle product $\sqcup\!\sqcup$.

Proposition (Shuffle property)

Let $(X_t)_{t \in [0,T]}$ be a continuous $\mathbb{R}^d$-valued semimartingale and $\mathbf{u}, \mathbf{v} \in \mathcal{S}(S)^*$. Then $\mathbf{u} \sqcup\!\sqcup \mathbf{v} \in \mathcal{S}(S)^*$ and

$$\langle \mathbf{u}, \mathbb{X} \rangle \langle \mathbf{v}, \mathbb{X} \rangle = \langle \mathbf{u} \sqcup\!\sqcup \mathbf{v}, \mathbb{X} \rangle.$$

Signature SDEs

- We introduce signature SDEs with state space $S \subseteq \mathbb{R}^d$ driven by some d-dimensional Brownian motion W via

$$dX_t = b(\mathbb{X}_t)dt + \sqrt{a(\mathbb{X}_t)}dW_t, \quad X_0 = x. \tag{Sig-SDE}$$

  The coefficients $b : \mathcal{S}(S) \to \mathbb{R}^d$ and $a : \mathcal{S}(S) \to \mathbb{S}^d_+$ are componentwise entire maps of group-like elements, i.e.

$$b_i(\mathbf{x}) = \langle \mathbf{b}^i, \mathbf{x} \rangle \text{ and } a^{ij}(\mathbf{x}) = \langle \mathbf{a}^{ij}, \mathbf{x} \rangle,$$

  where $\mathbf{b}^i, \mathbf{a}^{ij} \in \mathcal{S}(S)^*$.

- In the one-dimensional case this corresponds to an SDE with real-analytic coefficients. Neural SDEs with real-analytic activation functions are also included.

Signature SDEs

- We introduce signature SDEs with state space $S \subseteq \mathbb{R}^d$ driven by some d-dimensional Brownian motion W via

$$dX_t = b(\mathbb{X}_t)dt + \sqrt{a(\mathbb{X}_t)}dW_t, \quad X_0 = x. \tag{Sig-SDE}$$

  The coefficients $b : \mathcal{S}(S) \to \mathbb{R}^d$ and $a : \mathcal{S}(S) \to \mathbb{S}^d_+$ are componentwise entire maps of group-like elements, i.e.

$$b_i(\mathbf{x}) = \langle \mathbf{b}^i, \mathbf{x} \rangle \text{ and } a^{ij}(\mathbf{x}) = \langle \mathbf{a}^{ij}, \mathbf{x} \rangle,$$

  where $\mathbf{b}^i, \mathbf{a}^{ij} \in \mathcal{S}(S)^*$.

- In the one-dimensional case this corresponds to an SDE with real-analytic coefficients. Neural SDEs with real-analytic activation functions are also included.
- Universality within Itô-proceses as all continuous path-functionals can be approximated by linear and thus entire functions of the signature.
- By the shuffle property the characteristics of $\mathbb{X}$ are again entire functions. $\Rightarrow$ $\mathbb{X}$ is a $\mathcal{S}(S)$ valued affine and polynomial process

Main result

Theorem (C.C., S. Svaluto-Ferro, J.Teichmann ('23))

Let X be given by (Sig-SDE) *and fix* $\mathcal{U} \subseteq \mathcal{S}(S)^*$. *Consider the maps* $R : \mathcal{U} \to T((\mathbb{R}^d))$ *and* $L : \mathcal{U} \to T((\mathbb{R}^d))$ *given by*

$$R(\mathbf{u}) = \mathbf{b}^\top \sqcup\!\sqcup\, \mathbf{u}^{(1)} + \frac{1}{2} \operatorname{Tr}\left(\mathbf{a} \sqcup\!\sqcup \left(\mathbf{u}^{(2)} + \mathbf{u}^{(1)} \sqcup\!\sqcup\, (\mathbf{u}^{(1)})^\top\right)\right),$$

$$L(\mathbf{u}) = \mathbf{b}^\top \sqcup\!\sqcup\, \mathbf{u}^{(1)} + \frac{1}{2} \operatorname{Tr}\left(\mathbf{a} \sqcup\!\sqcup\, \mathbf{u}^{(2)}\right),$$

where $\mathbf{u}^{(1)}, \mathbf{u}^{(2)}$ *denotes certain shifts of* $\mathbf{u}$. *Under some technical conditions, $\mathbb{X}$ is an $\mathcal{S}(S)$-valued affine and polynomial process satisfying*

$$\mathbb{E}[\exp(\langle \mathbf{u}, \mathbb{X}_t \rangle] = \exp(\langle \psi(t), \mathbb{X}_0 \rangle), \quad \mathbb{E}[\langle \mathbf{u}, \mathbb{X}_t \rangle] = \langle \mathbf{c}(t), \mathbb{X}_0 \rangle,$$

where ψ and $\mathbf{c}$ are $\mathcal{U}$-valued solutions of the extended tensor algebra valued Riccati and linear ODEs, i.e.

$$\psi(t) = \mathbf{u} + \int_0^t R(\psi(s)ds, \quad \mathbf{c}(t) = \mathbf{u} + \int_0^t L(\mathbf{c}(s))ds.$$

Numerical illustration

- Computation of the Laplace transform of a geometric Brownian motion, i.e., $\mathbb{E}[\exp(-\lambda \exp(X_t))]$ where X is a one-dimensional Brownian motion and $\lambda \in \mathbb{R}$.
- In the one dimensional setup $\mathbb{X}_t := (1, X_t, \frac{X_t^2}{2!}, \ldots)$, and the function R is sequence-valued and here of the form

$$R(\mathbf{u})_k = \frac{1}{2}\Big(\mathbf{u}_{k+2} + \sum_{i+j=k} \binom{k}{i} \mathbf{u}_{i+1}\mathbf{u}_{j+1}\Big), \quad k \in \mathbb{N},$$

  such that $\mathbb{E}[\exp(-\lambda \exp(X_t))] = \mathbb{E}[\exp(\langle \mathbf{u}, \mathbb{X}_t \rangle)] = \exp(\langle \boldsymbol{\psi}(t), \mathbb{X}_0 \rangle)$, where $\mathbf{u} = -\lambda(1, 1, 1, \ldots)$ and $\partial_t \psi(t) = R(\psi(t))$.

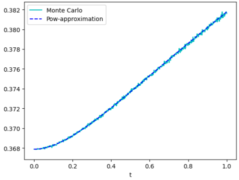
Numerical illustration

- Computation of the Laplace transform of a geometric Brownian motion, i.e., $\mathbb{E}[\exp(-\lambda \exp(X_t))]$ where X is a one-dimensional Brownian motion and $\lambda \in \mathbb{R}$.
- In the one dimensional setup $\mathbb{X}_t := (1, X_t, \frac{X_t^2}{2!}, \ldots)$, and the function R is sequence-valued and here of the form

$$R(\mathbf{u})_k = \frac{1}{2}\Big(\mathbf{u}_{k+2} + \sum_{i+j=k} \binom{k}{i} \mathbf{u}_{i+1}\mathbf{u}_{j+1}\Big), \quad k \in \mathbb{N},$$

such that $\mathbb{E}[\exp(-\lambda \exp(X_t))] = \mathbb{E}[\exp(\langle \mathbf{u}, \mathbb{X}_t \rangle)] = \exp(\langle \boldsymbol{\psi}(t), \mathbb{X}_0 \rangle)$, where $\mathbf{u} = -\lambda(1, 1, 1, \ldots)$ and $\partial_t \boldsymbol{\psi}(t) = R(\boldsymbol{\psi}(t))$.

Numerical implementation for $\lambda = 1$ via a standard ODE solver for the truncated Riccati ODEs with truncation level 20.



$t \mapsto \mathbb{E}[\exp(-\exp(W_t))]$

Conclusion

- Lévy's stochastic area formula as an early example from the literature showing the powerfulness of the unifying affine framework. It can actually be also embedded in the signature SDE framework (going just up to level 2).

Conclusion

- Lévy's stochastic area formula as an early example from the literature showing the powerfulness of the unifying affine framework. It can actually be also embedded in the signature SDE framework (going just up to level 2).
- Infinite dimensional Wishart processes as infinite dimensional affine process with intricate parameter restrictions and state space constraints

Conclusion

- Lévy's stochastic area formula as an early example from the literature showing the powerfulness of the unifying affine framework. It can actually be also embedded in the signature SDE framework (going just up to level 2).
- Infinite dimensional Wishart processes as infinite dimensional affine process with intricate parameter restrictions and state space constraints
- Signature SDEs as generic class of Itô-processes that is affine and polynomial when lifted to the state space of group-like elements
  - $\Rightarrow$ One step in the direction of universality of affine processes
  - $\Rightarrow$ Converging power series for the Fourier-Laplace transform and the expected value of entire functions of the signature process
  - $\Rightarrow$ Tractability properties for neural SDEs and Sig-SDE models, in particular systematic polynomial way to compute expected values which is important in many machine learning applications

Outlook and ongoing work

- Weighted Stone Weierstrass theorems to prove universality features without compactness criteria (joint work with P. Schmocker and J.Teichmann)
- Existence and uniqueness theory for signature SDEs
- Theory of entire and real-analytic processes, including jumps, extending the theory of polynomial processes to semigroups mapping real-analytic functions to real-analytic functions (joint work with F. Primavera and S. Svaluto-Ferro)
- Analysis of non-seminartingale setups
- Reading universality property backwards to obtain stochastic representations of generic non-linear partial differential equations via vector measure-valued affine process

Thank you for your attention!