(Probability) measure-valued polynomial diffusions

Christa Cuchiero
(based on joint work in progress with Martin Larsson and Sara Svaluto-Ferro)

University of Vienna

Thera Stochastics
A Mathematics Conference in Honor of Ioannis Karatzas
June 1st, 2017

Motivation and goal

- Goal: Tractable dynamic modeling of probability measures

Motivation and goal

- Goal: Tractable dynamic modeling of probability measures
- Usual tractable model classes:
  - Lévy processes
  - Affine processes
  - Polynomial processes

Motivation and goal

- Goal: Tractable dynamic modeling of probability measures
- Even for discrete measures taking only finitely many values, this is only possible with
  - ~~Lévy processes~~
  - ~~Affine processes~~
  - Polynomial processes

Motivation and goal

- Goal: Tractable dynamic modeling of probability measures
- Even for discrete measures taking only finitely many values, this is only possible with
  - ~~Lévy processes~~
  - ~~Affine processes~~
  - Polynomial processes
- $\Rightarrow$ Probability measure-valued polynomial processes first canonical class to achieve the above goal

Motivation and goal

- Goal: Tractable dynamic modeling of probability measures
- Even for discrete measures taking only finitely many values, this is only possible with
  - ~~Lévy processes~~
  - ~~Affine processes~~
  - Polynomial processes
- $\Rightarrow$ Probability measure-valued polynomial processes first canonical class to achieve the above goal
- More general: polynomial processes taking values in subsets of signed measures, including for instance affine processes.

Motivation and goal

- Goal: Tractable dynamic modeling of probability measures
- Even for discrete measures taking only finitely many values, this is only possible with
  - ~~Lévy processes~~
  - ~~Affine processes~~
  - Polynomial processes
- $\Rightarrow$ Probability measure-valued polynomial processes first canonical class to achieve the above goal
- More general: polynomial processes taking values in subsets of signed measures, including for instance affine processes.
- Literature on measure valued processes: Dawson, Ethier, Etheridge, Fleming, Hochberg, Kurtz, Perkins, Viot, Watanabe, etc.

What does tractability actually mean?

- Consider first the finite dimensional case with a general Markov process on some subset of $\mathbb{R}^d$:

- For a general $\mathbb{R}^d$-valued Markov processes the Kolmogorov backward equation is a PIDE on $\mathbb{R}^d \times [0, \infty)$.
- Tractability:
  - Affine processes: For initial values of the form $x \mapsto \exp\langle u, x\rangle$, the Kolmogorov PIDE reduces to generalized Riccati ODEs on $\mathbb{R}^d$.
  - Polynomial processes: When the initial values are polynomials of degree k, the Kolmogorov PIDE reduces to a linear ODE on $\mathbb{R}^N$ with N the dimension of polynomials of degree k.

What does tractability actually mean?

- Consider first the finite dimensional case with a general Markov process on some subset of $\mathbb{R}^d$:

- Let E be some Polish space and consider $M(E)$ the space of finite signed measures.

- If E consists of d points, then $M(E)$ can be identified with $\mathbb{R}^d$.

- For a general $M(E)$-valued Markov processes the Kolmogorov backward equation is a PIDE on $M(E) \times [0, \infty)$.

- Tractability:
  - Affine processes: For initial values of the form $x \mapsto \exp\langle u, x \rangle$, the Kolmogorov PIDE reduces to generalized Riccati PDEs on E.
  - Polynomial processes: When the initial values are polynomials of degree k, the Kolmogorov PIDE reduces to a linear PIDE on E^k (in the case of probability measures).

What does tractability actually mean?

- Consider first the finite dimensional case with a general Markov process on some subset of $\mathbb{R}^d$:
- Let E be some Polish space and consider $M(E)$ the space of finite signed measures.
- If E consists of d points, then $M(E)$ can be identified with $\mathbb{R}^d$.
- For a general $M(E)$-valued Markov processes the Kolmogorov backward equation is a PIDE on $M(E) \times [0, \infty)$.
- Tractability:
  - Affine processes: For initial values of the form $x \mapsto \exp\langle u, x\rangle$, the Kolmogorov PIDE reduces to generalized Riccati PDEs on E.
  - Polynomial processes: When the initial values are polynomials of degree k, the Kolmogorov PIDE reduces to a linear PIDE on E^k (in the case of probability measures).

What does tractability actually mean?

- Consider first the finite dimensional case with a general Markov process on some subset of $\mathbb{R}^d$:

- Let E be some Polish space and consider $M(E)$ the space of finite signed measures.

- If E consists of d points, then $M(E)$ can be identified with $\mathbb{R}^d$.

- For a general $M(E)$-valued Markov processes the Kolmogorov backward equation is a PIDE on $M(E) \times [0, \infty)$.

- Tractability:

  - Affine processes: For initial values of the form $x \mapsto \exp\langle u, x\rangle$, the Kolmogorov PIDE reduces to generalized Riccati PDEs on E.

  - Polynomial processes: When the initial values are polynomials of degree k, the Kolmogorov PIDE reduces to a linear PIDE on E^k (in the case of probability measures). In certain cases it can be further reduced to an ODE.

Applications in finance

- Stochastic portfolio theory (SPT) (B. Fernholz, I.Karatzas, ...)
  - Large equity markets: joint stochastic modeling of a large finite (or even potentially infinite) number of stocks or (relative) market capitalizations constituting the major indices (e.g., 500 in the case of S&P 500)
  - Capital distribution curve modeling
- Term structure modeling of interest rates, variance swaps, commodities or electricity forward contracts involving potentially an uncountably infinite number of assets
- Polynomial Volterra processes in particular in view of rough volatility modeling
- Stochastic representations of (linear systems) of PIDEs

Large equity markets in SPT

- Consider a set of stocks with market capitalizations $S_t^1, \ldots, S_t^d$.
- In SPT the main quantity of interest are the market weights

$$\mu_t^i = \frac{S_t^i}{S_t^1 + \cdots + S_t^d}.$$

- $\mu_t = (\mu_t^1, \ldots, \mu_t^d)$ takes values in the unit simplex

$$\Delta^d = \left\{ z \in [0,1]^d \colon z_1 + \cdots + z_d = 1 \right\}.$$

- One is interested in the behavior of μ for large d!
- Possible approach: Linear factor models, i.e. view $(\mu^1, \ldots, \mu^d)$ as the projection of a single tractable infinite dimensional model.
  - Let X be a probability measure valued (polynomial) process.
  - For functions $g_i \geq 0$ such that $g_1 + \ldots + g_d \equiv 1$, set $\mu_t^i = \int g_i(x) X_t(dx)$.
  - Extensions to infinitely many assets are easily possible.

Capital distribution curves

- Probability measure valued processes can be used to describe the empirical measure of the capitalizations:

$$\frac{1}{d}\sum_{i=1}^{d}\delta_{S_t^i}(dx) \tag{1}$$

- There is a one to one correspondence between this empirical measure and the capital distribution curves which map the rank of the companies to their capitalizations . $\Rightarrow$ Analysis for specific models as $d \to \infty$. (e.g. by M. Shkolnikov, etc.)
- Empirically these curves proved to be of a specific shape and particularly stable over time with a certain fluctuating behavior.

Capital distribution curves

- Probability measure valued processes can be used to describe the empirical measure of the capitalizations:

$$\frac{1}{d}\sum_{i=1}^{d}\delta_{S_t^i}(dx) \tag{1}$$

- There is a one to one correspondence between this empirical measure and the capital distribution curves which map the rank of the companies to their capitalizations . $\Rightarrow$ Analysis for specific models as $d \to \infty$. (e.g. by M. Shkolnikov, etc.)
- Empirically these curves proved to be of a specific shape and particularly stable over time with a certain fluctuating behavior.

Question:

- For which models is (the limit of) (1) a probability measure valued polynomial process? Consistency with empirical features?

Term structure modeling

- Let us for instance consider modeling of bond prices $P(t, T)$ for $t \in [0, T^*]$ and $T \in [t, T^*]$ for some finite time horizon T^*.
- Let X be a probability measure valued (polynomial) process.
- Then, one possibility to define bond prices is

$$P(t, T) = \int_E g_t(T, x) X_t(dx),$$

where $g_t(\cdot, x) : [t, T^*] \to [0, 1]$ is a deterministic function with $g_t(t, \cdot) \equiv 1$, chosen to be decreasing if nonnegative short rates are to be enforced.

Part I

Signed measure-valued polynomial diffusions

Polynomial diffusions on $\mathcal{S} \subseteq \mathbb{R}^d$

- $\mathrm{Pol}(\mathcal{S})$: vector space of all polynomial on $\mathcal{S}$

Definition

- A linear operator $L : \mathrm{Pol}(\mathbb{R}^d) \to \mathrm{Pol}(\mathcal{S})$ is called polynomial if $\deg(Lp) \leq \deg(p)$ for all $p \in \mathrm{Pol}(\mathbb{R}^d)$.
- Let L be a polynomial operator. Then a polynomial diffusion on $\mathcal{S}$ is a continuous $\mathcal{S}$-valued solution X to the martingale problem

$$p(X_t) - \int_0^t Lp(X_s)ds = (\text{martingale}), \qquad \forall p \in \mathrm{Pol}(\mathbb{R}^d).$$

- If the martingale problem is well posed it leads to a Markov process and thus to a polynomial process in the sense of (C., Keller-Ressel, Teichmann, '12).
- In this talk, the focus lies on $\mathcal{S} = \Delta^d$. In this case the martingale problem is always well-posed. Polynomial operators L generating diffusions on Δ^d have been completely characterized (Larsson, Filipović, '16).

Characterization and conditional moment formula

- Fix $k \in \mathbb{N}$ and let $H = (h_1, \ldots, h_N)$, $h_i \in \mathrm{Pol}(\mathcal{S})$, be a basis for $\mathrm{Pol}_k(\mathcal{S}) = \{p \in \mathrm{Pol}(\mathcal{S})\colon \deg(p) \leq k\}$.

Theorem (C., Keller-Ressel, Teichmann '12, Filipovic and Larsson '16)

Let L be a linear operator whose domain contains $Pol(\mathbb{R}^d)$ and assume that there is a continuous $\mathcal{S}$-valued solution X to the martingale problem for L. The following assertions are equivalent:

- *L is a polynomial.*
- *L is of the form* $Lp(x) = \nabla p(x)^\top \underbrace{b(x)}_{\text{affine in } x} + \frac{1}{2} Tr\Big(\underbrace{a(x)}_{\text{quadratic in } x} \nabla^2 p(x)\Big)$.
- *For every polynomial $p \in Pol_k(\mathcal{S})$ we have*

$$E[p(X_{t+s}) \mid \mathcal{F}_s] = H(X_s)^\top e^{tL} \vec{p},$$

where $\vec{p} \in \mathbb{R}^N$ is the vector representation of p, and we identify L with its $N \times N$ matrix representation on $Pol_k(\mathcal{S})$.

Goal of this talk

Develop a theory of measure valued polynomial processes:

- Questions:
  - How to define polynomials $p(\nu)$ with measures as argument?
  - What is a polynomial operator L in this setting?
  - How does this operator look like?
  - Specific state spaces: characterization or possible specification of L in the case of probability measures.
  - How does the moment formula look like?
  - How does the matrix exponential translate in this infinite dimensional setting?

Notation

E : compact Polish space.

$\widehat{C}(E^k)$: space of symmetric continuous functions $f : E^k \to \mathbb{R}$.

$M(E)$: space of finite signed measures on E with the topology of weak convergence.

$M_1(E)$: space of probability measures on E with the topology of weak convergence.

Polynomials of measure arguments

- A monomial of degree k on $M(E)$ is an expression of the form:

$$\nu \mapsto \int_{E^k} \underbrace{g(x_1, \ldots, x_k)}_{\text{coefficient of the monomial}} \nu(dx_1) \cdots \nu(dx_k) =: \langle g, \nu^k \rangle,$$

  for some $g \in \widehat{C}(E^k)$.

- A polynomial p of degree m on $M(E)$ is an expression of the form:

$$\nu \mapsto p(\nu) = \sum_{k=0}^{m} \langle g_k, \nu^k \rangle$$

  for some $g_k \in \widehat{C}(E^k)$.

- We denote the set of all polynomials on $\mathcal{S} \subseteq M(E)$ by $\mathrm{Pol}(\mathcal{S})$.

Derivatives of polynomials

- For a function $f : M(E) \to \mathbb{R}$ the directional derivative in direction δ_x at ν is given by

$$\partial_x f(\nu) := \lim_{\varepsilon \to 0} \frac{f(\nu + \varepsilon\delta_x) - f(\nu)}{\varepsilon}.$$

- The iterated derivative is then denoted by $\partial^2_{xy} f(\nu) = \partial_x \partial_y f(\nu)$.

Lemma

Consider the monomial $p(\nu) = \langle g, \nu^k \rangle$ for some $g \in \widehat{C}(E^k)$. Then

$$\partial_x p(\nu) = k\langle g(\cdot, x), \nu^{k-1} \rangle,$$

and the map $x \mapsto \partial_x p(\nu)$ lies in $C(E)$.

Classes of polynomials

Restriction to specific sets of coefficients:

Definition

Let $D \subseteq C(E)$ be a dense linear subspace. Then

$$P^D = \{p \in \mathrm{Pol}(M(E)) : \text{ the coefficients of } p \text{ lie in } D^{\otimes k}\}.$$

Recall that $g \otimes \cdots \otimes g \in D^{\otimes k}$ denotes the map

$$(x_1, \ldots, x_k) \mapsto g(x_1) \cdots g(x_k).$$

Lemma

For any $p \in P^D$ and $\nu \in M(E)$: $\partial p(\nu) \in D$ and $\partial^2 p(\nu) \in D \otimes D$.

- The most relevant examples that we shall consider are $D = C^2(E)$ and $D = \mathrm{Pol}(E)$.

Polynomial diffusions on $\mathcal{S} \subseteq M(E)$

- Recall the finite dimensional definition:

Definition

- A linear operator $L : \mathrm{Pol}(\mathbb{R}^d) \to \mathrm{Pol}(\mathcal{S})$ is called polynomial if $\deg(Lp) \leq \deg(p)$ for all $p \in \mathrm{Pol}(\mathbb{R}^d)$.
- Let L be a polynomial operator. Then a polynomial diffusion on $\mathcal{S}$ is a continuous $\mathcal{S}$-valued solution X to the martingale problem

$$p(X_t) - \int_0^t Lp(X_s)ds = (\text{martingale}), \qquad \forall p \in \mathrm{Pol}(\mathbb{R}^d).$$

Polynomial diffusions on $\mathcal{S} \subseteq M(E)$

- Recall the finite dimensional definition:
- Completely analogously to the finite dimensional case we define:

Definition

- A linear operator $L : P^D \to \mathrm{Pol}(\mathcal{S})$ is called polynomial if $\deg(Lp) \leq \deg(p)$ for all $p \in P^D$.
- Let L be a polynomial operator. Then a polynomial diffusion on $\mathcal{S}$ is a continuous $\mathcal{S}$-valued solution X to the martingale problem

$$p(X_t) - \int_0^t Lp(X_s)ds = (\text{martingale}), \qquad \forall p \in P^D.$$

Polynomial operators generating diffusions

Theorem (C., Larsson, Svaluto-Ferro '17)

Let L be a linear operator whose domain contains P^D and assume that there is a continuous $\mathcal{S}$-valued solution of the martingale problem for L. Then the following assertions are equivalent.

- *L is a polynomial.*
- *L is of the form*

$$Lp(\nu) = \bar{B}(\partial p(\nu); \nu) + \frac{1}{2}\bar{Q}(\partial^2 p(\nu); \nu),$$

where $\bar{B} : D \times M(E) \to \mathbb{R}$ and $\bar{Q} : (D \otimes D) \times M(E) \to \mathbb{R}$ are given by

$$\bar{B}(g; \nu) = B_0(g) + \langle B_1(g), \nu \rangle$$
$$\bar{Q}(g \otimes g; \nu) = Q_0(g \otimes g) + \langle Q_1(g \otimes g), \nu \rangle + \langle Q_2(g \otimes g), \nu^2 \rangle$$

for some linear operators $B_0 : D \to \mathbb{R}$, $B_1 : D \to C(E)$, $Q_0 : D \otimes D \to \mathbb{R}$, $Q_1 : D \otimes D \to C(E)$, $Q_2 : D \otimes D \to \widehat{C}(E^2)$.

Polynomial operators generating diffusions on $M_1(E)$

Theorem (cont.)

In the case $\mathcal{S} = M_1(E)$, the form of L simplifies to

$$Lp(\nu) = \left\langle B\left(\partial p(\nu)\right), \nu \right\rangle + \frac{1}{2}\left\langle Q\left(\partial^2 p(\nu)\right), \nu^2 \right\rangle,$$

where B is a linear operator on D and Q is a linear operator on $D \otimes D$.

- The representation of B as linear and Q as quadratic monomials, comes from the fact that we work with probability measures, which allows to write every polynomial of degree k as a monomial of degree $n \geq k$.

Part II

Probability measure-valued polynomial diffusions

$M_1(E)$-valued polynomial diffusions: characterization

Polynomial operators L generating polynomial diffusions on Δ^d are characterized (Filipovic and Larsson '16) as follows:

$$Lp(y) = \sum_{i=1}^{d} B_i^\top \nabla p(y) y_i + \frac{1}{2} \sum_{ij=1}^{d} \alpha_{ij} \Big(\partial_{ii}^2 p(y) + \partial_{jj}^2 p(y) - 2\partial_{ij}^2 p(y) \Big) y_i y_j$$

where B is a transition rate matrix, i.e. $B_{ij} \geq 0$ for $i \neq j$, $B_{ii} = -\sum_{j \neq i} B_{ij}$, and $\alpha_{ij} = \alpha_{ji} \geq 0$.

$M_1(E)$-valued polynomial diffusions: characterization

Polynomial operators L generating polynomial diffusions on Δ^d are characterized (Filipovic and Larsson '16) as follows:

$$Lp(y) = \sum_{i=1}^{d} B_i^\top \nabla p(y) y_i + \frac{1}{2} \sum_{ij=1}^{d} \alpha_{ij} \Big(\partial_{ii}^2 p(y) + \partial_{jj}^2 p(y) - 2\partial_{ij}^2 p(y) \Big) y_i y_j$$

where B is a transition rate matrix, i.e. $B_{ij} \geq 0$ for $i \neq j$, $B_{ii} = -\sum_{j \neq i} B_{ij}$, and $\alpha_{ij} = \alpha_{ji} \geq 0$.

Theorem (C., Larsson, Svaluto-Ferro '17)

Let $D = C(E)$, i.e. $P^D = Pol(M(E))$. A linear operator $L : P^D \to Pol(M_1(E))$ generates a polynomial diffusion on $M_1(E)$ if and only if

$$Lp(\nu) = \Big\langle B\big(\partial p(\nu)\big), \nu \Big\rangle + \frac{1}{2} \Big\langle \alpha(x,y) \Big(\partial_{xx}^2 p(\nu) + \partial_{yy}^2 p(\nu) - 2\partial_{xy}^2 p(\nu) \Big), \nu^2 \Big\rangle$$

where B is the generator of a jump-diffusion on E, $\alpha : E^2 \to \mathbb{R}$ is symmetric, nonnegative and continuous.

$M_1(E)$-valued polynomial diffusions: characterization

Polynomial operators L generating polynomial diffusions on Δ^d are characterized (Filipovic and Larsson '16) as follows:

$$Lp(x) = \sum_{i=1}^{d} B_i^\top \nabla p(x) x_i + \frac{1}{2} \sum_{ij=1}^{d} \alpha_{ij} \Big(\partial^2_{ii} p(x) + \partial^2_{jj} p(x) - 2\partial^2_{ij} p(x) \Big) x_i x_j$$

where B is a transition rate matrix, i.e. $B_{ij} \geq 0$ for $i \neq j$, $B_{ii} = -\sum_{j \neq i} B_{ij}$, and $\alpha_{ij} = \alpha_{ji} \geq 0$.

Theorem (C., Larsson, Svaluto-Ferro '17)

Let $D = C(E)$, i.e. $P^D = Pol(M(E))$. A linear operator $L : P^D \to Pol(M_1(E))$ generates a polynomial diffusion on $M_1(E)$ if and only if

$$Lp(\nu) = \Big\langle B\big(\partial p(\nu)\big), \nu \Big\rangle + \frac{1}{2} \Big\langle \alpha \Psi\big(\partial^2 p(\nu)\big)\Big), \nu^2 \Big\rangle$$

where B is the generator of a jump diffusion on E, $\alpha : E^2 \to \mathbb{R}$ is symmetric, nonnegative, continuous, and $\Psi g(x, y) = g(x, x) + g(y, y) - 2g(x, y)$.

$M_1(E)$-valued polynomial diffusions: characterization

Polynomial operators L generating polynomial diffusions on Δ^d are characterized (Filipovic and Larsson '16) as follows:

$$Lp(x) = \sum_{i=1}^{d} B_i^\top \nabla p(x) x_i + \frac{1}{2} \sum_{ij=1}^{d} \alpha_{ij} \Big(\partial_{ii}^2 p(x) + \partial_{jj}^2 p(x) - 2\partial_{ij}^2 p(x) \Big) x_i x_j$$

where B is a transition rate matrix, i.e. $B_{ij} \geq 0$ for $i \neq j$, $B_{ii} = -\sum_{j \neq i} B_{ij}$, and $\alpha_{ij} = \alpha_{ji} \geq 0$.

Theorem (C., Larsson, Svaluto-Ferro '17)

Let $D = C(E)$, i.e. $P^D = Pol(M(E))$. A linear operator $L : P^D \to Pol(M_1(E))$ generates a polynomial diffusion on $M_1(E)$ if and only if

$$Lp(\nu) = \Big\langle B\big(\partial p(\nu)\big), \nu \Big\rangle + \frac{1}{2} \Big\langle \alpha \Psi\big(\partial^2 p(\nu)\big)\Big), \nu^2 \Big\rangle$$

where B is the generator of a jump diffusion on E, $\alpha : E^2 \to \mathbb{R}$ is symmetric, nonnegative, continuous, and $\Psi g(x,y) = g(x,x) + g(y,y) - 2g(x,y)$.
If the process generated by B is additionally Feller, then the polynomial diffusion generated by L is unique in law, i.e. the martingale problem is well posed.

Example: Fleming-Viot process ($\alpha = 1/2$)

- The most famous $\mathcal{M}_1(E)$-valued process is the Fleming-Viot process, which is actually polynomial with $\alpha = 1/2$.

Example: Fleming-Viot process ($\alpha = 1/2$)

- The most famous $\mathcal{M}_1(E)$-valued process is the Fleming-Viot process, which is actually polynomial with $\alpha = 1/2$.
- When E consists of d points, this process corresponds to a multivariate Jacobi-type process with infinitesimal generator

$$Lp(x) = \sum_{i=1}^{d} B_i^\top \nabla p(x) x_i + \frac{1}{2} \sum_{i,j \in E} \partial_{ij}^2 p(x) x_i(\delta_{ij} - x_j),$$

where B is the transition rate matrix of a continuous time Markov chain on E.

Example: Fleming-Viot process ($\alpha = 1/2$)

- The most famous $\mathcal{M}_1(E)$-valued process is the Fleming-Viot process, which is actually polynomial with $\alpha = 1/2$.
- When E consists of d points, this process corresponds to a multivariate Jacobi-type process with infinitesimal generator

$$Lp(x) = \sum_{i=1}^{d} B_i^\top \nabla p(x) x_i + \frac{1}{2} \sum_{i,j \in E} \partial_{ij}^2 p(x) x_i (\delta_{ij} - x_j),$$

  where B is the transition rate matrix of a continuous time Markov chain on E.
- In the general case, the corresponding operator is of the form

$$\begin{aligned} Lp(\nu) &= \int_E B(\partial p(\nu)) \nu(dx) + \frac{1}{2} \int_E \int_E \partial_{xy}^2 p(\nu) \nu(dx)(\delta_x(dy) - \nu(dy)) \\ &= \left\langle B\big(\partial p(\nu)\big), \nu \right\rangle + \frac{1}{4} \left\langle \Psi\big(\partial^2 p(\nu)\big), \nu^2 \right\rangle. \end{aligned}$$

  for $p \in P^D$ and D the domain of an operator B generating an E-valued Feller process.

Remarks

- We have a full characterization of $\mathcal{M}_1(E)$ valued diffusions for $D = C(E)$, in particular when E is finite dimensional we recover the characterization by Filipovic and Larsson (2016).
- Similarly, if D is general, but B does not contain a diffusion component, Q is necessarily of the above form.
- When $D \subseteq C^2(E)$, then other specifications are possible.

Specifications when $D \subseteq C^2(E)$

Proposition

Let $D \subseteq C^2(E)$. Consider the linear operator $L : P^D \to Pol(M_1(E))$ given by

$$Lp(\nu) = \left\langle B\left(\partial p(\nu)\right), \nu \right\rangle + \frac{1}{2}\left\langle Q\left(\partial^2 p(\nu)\right), \nu^2 \right\rangle$$

$$Bg(x) = B^0 g(x) + \frac{1}{2}\tau(x)^2 \frac{d^2}{dx^2} g(x)$$

$$Qg(x, y) = \alpha(x, y)\Psi g(x, y) + \tau(x)\tau(y)\frac{d^2}{dxdy} g(x, y).$$

for some B^0 generating a jump-diffusion on E, $\alpha \in \widehat{C}(E^2)$ nonnegative, and $\tau \in C(E)$ nonnegative and vanishing on ∂E.

Then L generates an $M_1(E)$-valued polynomial diffusion.

If the parameters satisfy some additional conditions and D is rich enough, then the diffusion generated by L is unique in law.

Example : Empirical measures

- Let $X_t = \frac{1}{d}\sum_{i=1}^{d}\delta_{S_t^i}$, for

$$dS_t^i = b(S_t^i)dt + \sigma(S_t^i)dW_t^i + \tau(S_t^i)dW_t^0$$

where $(W^0, \ldots, W^d)$ is an $(d+1)$-dim Brownian Motion, b, σ and τ in $C(E)$.

- Then

$$p(X_t) := \langle g, X_t^k \rangle = \frac{1}{d^k}\sum_{i_1,\ldots,i_k=1}^{d} g(S_t^{i_1}, \ldots, S_t^{i_k})$$

- For $g \in C^2(E)$ (or equiv. $p \in P^D$ for $D \subseteq C^2(E)$) we can apply Itô's formula!

Example: Empirical measures

- This yields

$$
\begin{aligned}
&p(X_t) = \langle g, X_t^k \rangle = (martingale) \\
&+ \int_0^t \left\langle \left(b(x)\frac{d}{dx} + \frac{1}{2}(\tau(x)^2 + \sigma(x)^2)\frac{d^2}{dx^2} \right) (\partial_x p(X_s)), X_s \right\rangle ds \\
&+ \int_0^t \frac{1}{2}\left\langle \tau(x)\tau(y)\frac{d^2}{dxdy}(\partial^2_{xy} p(X_s)) + \frac{1}{d}\sigma^2(x)\frac{d^2}{dxdy}(\partial^2_{xy} p(X_s))1_{\{x=y\}}, X_s^2 \right\rangle ds \\
&= (martingale) + \int_0^t \left\langle B(\partial p(X_s)), X_s \right\rangle + \int_0^t \frac{1}{2}\left\langle Q(\partial^2 p(\nu)), X_s^2 \right\rangle,
\end{aligned}
$$

where

- $Bg(x) = b(x)\frac{d}{dx}g(x) + \frac{1}{2}(\sigma^2(x) + \tau^2(x))\frac{d^2}{dx^2}g(x)$ is the generator of S^i
- $Qg(x,y) = \tau(x)\tau(y)\frac{d^2}{dxdy}g(x,y) + \frac{1}{d}\sigma(x)^2\frac{d^2}{dxdy}g(x,x)1_{\{x=y\}}$

Example: Empirical measures

- This yields

$$p(X_t) = \langle g, X_t^k \rangle = (martingale)$$
$$+ \int_0^t \left\langle \left(b(x) \frac{d}{dx} + \frac{1}{2}(\tau(x)^2 + \sigma(x)^2) \frac{d^2}{dx^2} \right) (\partial_x p(X_s)), X_s \right\rangle ds$$
$$+ \int_0^t \frac{1}{2} \left\langle \tau(x)\tau(y) \frac{d^2}{dxdy} (\partial^2_{xy} p(X_s)) + \frac{1}{d} \sigma^2(x) \frac{d^2}{dxdy} (\partial^2_{xy} p(X_s)) 1_{\{x=y\}}, X_s^2 \right\rangle ds$$
$$= (martingale) + \int_0^t \left\langle B(\partial p(X_s)), X_s \right\rangle + \int_0^t \frac{1}{2} \left\langle Q(\partial^2 p(\nu)), X_s^2 \right\rangle,$$

where

  - $Bg(x) = b(x) \frac{d}{dx} g(x) + \frac{1}{2}(\sigma^2(x) + \tau^2(x)) \frac{d^2}{dx^2} g(x)$ is the generator of S^i
  - $Qg(x,y) = \tau(x)\tau(y) \frac{d^2}{dxdy} g(x,y) + \frac{1}{d} \sigma(x)^2 \frac{d^2}{dxdy} g(x,x) 1_{\{x=y\}}$

$\Rightarrow$ The empirical measure of

$$dS_t^i = b(S_t^i)dt + \sigma(S_t^i)dW_t^i + \tau(S_t^i)dW_t^0$$

is a polynomial process.

Towards the moment formula

- Let $p(\nu) = \langle g, \nu^k \rangle$ for some $g \in D^{\otimes k}$.
- Since Lp is a polynomial, we know that

$$Lp(\nu) = \left\langle h, \nu^k \right\rangle \qquad \forall\ \nu \in M_1(E)$$

  for some some $h \in \widehat{C}(E^k)$.
- We can thus define $L_k : D^{\otimes k} \to \widehat{C}(E^k)$ as the unique operator such that

$$Lp(\nu) = \left\langle L_k g, \nu^k \right\rangle.$$

- Fact: With the specifications given before, L_k is the generator of a jump-diffusion on E^k.

The moment formula

Assume that L_k is the generator of a Feller process on E^k (which easily translates to conditions on B, τ, etc.) and let $\{Y_t^k\}$ be the corresponding Feller semigroup. In particular

$$L_k(Y_t^k g) = \frac{d}{dt}(Y_t^k g) \qquad \text{for all } g \in D^{\otimes k}.$$

Theorem

Let X be polynomial diffusion with generator L such that L_k is the generator of a Feller process on E^k. For any $k \in \mathbb{N}_0$ and any $g \in \widehat{C}(E^k)$ one has the representation

$$\mathbb{E}\left[\langle g, X_{t+s}^k\rangle | \mathcal{F}_s\right] = \langle Y_t^k g, X_s^k\rangle$$

of the conditional moments of X.

The moment formula - Remarks

- Moments up to order k can be computed by solving a linear PIDE in k variables. In the case of E consisting of d points this boils down to the usual linear ODE.
- For general measure valued processes computing moments would mean solving the Kolmogorov backward equation with measures as arguments.
- Even in the present case, when $D = \mathrm{Pol}(E)$ and L_k a polynomial operator on $D^{\otimes k}$, Y_t^k corresponds to a matrix exponentials.
- One can also view the moment formula as stochastic representation of PIDEs of the above type.

Example: pure drift process ($\alpha = \tau = 0$)

Let B be a generator of a Feller process Z and set

$$Lp(\nu) = \Big\langle B\big(\partial p(\nu)\big), \nu \Big\rangle.$$

Let X be the (unique) polynomial diffusion with generator L and initial value δ_{x_0}, for some $x_0 \in E$. Then

- $X_t = \mathbb{P}_{x_0}\big(Z_t \in \cdot\big)$.
- In particular, it is deterministic.
- $Y_t^1 g(x) = \mathbb{E}_x[g(Z_t)]$, or more generally

$$Y_t^k g^{\otimes k}(x_1, \ldots, x_k) = \mathbb{E}_{x_1}[g(Z_t)] \cdot \ldots \cdot \mathbb{E}_{x_k}[g(Z_t)].$$

Tractability and Flexibility

- Tractability
  - Comparison with polynomial diffusion in Δ^d for computing moments at T of order k (fixed):

    - $E = \{1, \dots, d\}$: linear ODE in $\mathbb{R}^N \times [0, T]$, $N = \dim \mathrm{Pol}_k(\Delta^d) = \binom{k+d-1}{k} \approx d^k$

    - $E = [0, 1]$: linear P(I)DE in $[0, 1]^k \times [0, T]$ Discretization of E: $\{\frac{i}{n} : \; i = 0, \dots, n\} \approx n^k$

  - Key additional structure: regularity in $x \in E$.
- Flexibility
  - Linear factor models being projections of an infinite dimensional process are a much richer class than polynomial models on the simplex.

Conclusions

- We defined polynomial processes as solution of a MP, whose operator L is polynomial, i.e. maps P^D to $\mathrm{Pol}(\mathcal{S})$.
- When $D = C(E)$ we characterize polynomial operators L, whose MP is well posed:

$$Lp(\nu) = \left\langle B\left(\partial p(\nu)\right), \nu \right\rangle + \frac{1}{2}\left\langle \alpha \Psi\left(\partial^2 p(\nu)\right), \nu^2 \right\rangle.$$

- We provide a moment formula, establishing a link between $M_1(E)$-valued polynomial diffusions X and linear PIDEs in $E^k \times [0, T]$:

$$E\left[\langle g, X^k_{t+s}\rangle \mid \mathcal{F}_s\right] = \langle Y^k_t g, X^k_s \rangle$$

- Polynomial measure-valued processes allow to exploit spatial regularity, which is not present in the finite dimensional setting.

Outlook

- Theoretical part
  - Full characterization for $D = C^2(E)$
  - Extension to locally compact E
  - Different state spaces - in particular nonnegative measures.
  - Work out numerical advantages, possibly also with respect to large finite dimensional simplexes
- Applications in stochastic portfolio theory building on linear factor models
  - Existence of arbitrages?
  - Existence of supermartingale deflators?
  - Functionally generated portfolios, in particular infinite dimension?
  - Itô type formulas and stochastic integration in the sense of Föllmer for measure valued processes?
  - Implications for capital distribution curve modeling?

Happy Birthday, Ioannis!