# Non-linear (PI)DEs and affine processes

Christa Cuchiero and Josef Teichmann
(joint work with Georg Grafendorfer)

ETH Zürich

Paris 2017

Often in mathematics it is fruitful to turn a successful point of view around:

- affine processes gained importance since their marginal distribution is known up to the solution of two non-linear ODEs, the *generalized Riccati equations*. Often the solutions of these ODEs are explicitly known.
- in turn one can apply affine processes to *represent stochastically (in a forward simulable manner)* the solution of non-linear ODEs, which means in particular that one obtains (Q)MC algorithms for the solution of non-linear ODEs of generalized Riccati type.

This point of view is classical in the theory of branching Markov processes or super-diffusions to represent stochastically (in a forward simulable manner) solutions of non-linear DEs. It is a linearization by a non-linear operation of generically non-linear equations of Lévy-Khintchine type (I shall come back to this point of view later).

# A simple Question

If one considers affine processes as stochastic representations of solutions of non-linear ODEs (PIDEs in case of infinite dimensional processes), then a simple question arises:

*Describe the class of equations, finite dimensional (ODEs) and infite dimensional (PIDEs) ones, which can be represented in this way, and estimate the complexity of simulating the representation.*

- Consider a set $\mathcal{D} \subset \mathbb{R}^{d+m}$ which will serve as state space for the affine stochastic process introduced as follows.
- We consider a diffusion process with jumps (in the sense of Jacod/Shirayev) $(N, Y)$ on the state space $\mathcal{D}$ whose differential semi-martingale characteristics $(b, c, F)$ (with respect to the truncation function 0) are given as linear functions in $N$, i.e.,

$$b_t = \beta N_t, \quad \beta \in \mathbb{R}^{d+m \times d},$$

$$c_t = \sum_{i=1}^{d} \alpha_i N_{i,t}, \quad \alpha_i \in \mathbb{S}^{d+m \times d+m},$$

$$F_t(dn, dy) = \langle N_t, \nu(dn, dy) \rangle,$$

where $\nu$ is a signed ($d$-dimensional) vector valued measure such that $\int_{\mathcal{D}} (\|(n, y)\| \wedge 1)(\nu_i^+(dn, dy) + \nu_i^-(dn, dy)) < \infty$ for all $i \in \{1, \ldots, d\}$.

- If the martingale problem corresponding to these characteristics is well-posed, the process $(N, Y)$ is an affine (actually linear) process in the classical sense. That is there exists a function $\psi : \mathcal{V} \to \mathbb{C}^{d+m}$ such that, for every initial value $(n, y) \in \mathcal{D}$ and for every $(t, (f, h)) \in \mathcal{V}$, it holds that

$$\mathbb{E}_{(n,y)}[e^{\langle f, N_t\rangle + \langle h, Y_t\rangle}] = e^{\langle \psi(t,f,h), n\rangle + hy},$$

where

$$\mathcal{V} := \{(t, \zeta) \in [0, \infty) \times \mathbb{C}^{d+m} : \zeta \in \mathfrak{U}_t\},$$

with

$$\mathfrak{U}_t := \left\{\zeta \in \mathbb{C}^{d+m} : \mathbb{E}\left[\left|e^{\langle \zeta, (N_s, Y_s)\rangle}\right|\right] < +\infty, \text{ for all } s \in [0, t]\right\}.$$

- The $\mathbb{C}^d$-valued function $\psi$ satisfies the following non-linear ODE:

$$\partial_t \psi(t) = \mathcal{R}(\psi(t)) = \beta^\top (\psi(t), h)^\top + \begin{pmatrix} (\psi(t), h)\alpha_1(\psi(t), h)^\top \\ \vdots \\ (\psi(t), h)\alpha_d(\psi(t), h)^\top \end{pmatrix} + \int_{\mathcal{D}} (e^{\langle \psi(t), n\rangle + \langle h, y\rangle} - 1)\nu(dn, dy).$$

- Its solution admits the following stochastic representation

$$\psi_i(t) = \log \mathbb{E}_{(e_i,0)}[e^{\langle f, N_t\rangle + \langle h, Y_t\rangle}], \quad i \in \{1, \dots, d\},$$

if $(f, h) \in \mathfrak{U}_t$.

- Note that if $\int_{\mathcal{D}} \|(n, y)\| \nu_i^{+/-}(dn, dy) = \infty$, then $\mathcal{R}$ is not Lipschitz.

- Consider the following polynomial ODE

$$\partial_t u(t) = \sum_{k=0}^{\infty} a_k u^k(t) - u(t), \quad u(0) = g,$$

with $\sum_{k=0}^{\infty} |a_k| < \infty$.

- Define $\psi(t) := \log(u(t))$. Then

$$\partial_t \psi(t) = \sum_{k=0}^{\infty} a_k u^{k-1}(t) - 1 = \sum_{k=0}^{\infty} a_k e^{\psi(t)(k-1)} - 1.$$

# Polynomial ODE (1d) in the $\psi$ picture

- This can be associated to an affine process $(N, Y_1, Y_2) \in \mathcal{D} = \mathbb{Z}_+ \times \mathbb{R}_+ \times \mathbb{Z}_+$. Indeed, taking $(h_1, h_2) = (1, \mathrm{i}\pi)$, the ODE can be rewritten as

$$\partial\psi(t) = \int_{\mathcal{D}} (e^{\psi(t)n + y_1 + \mathrm{i}\pi y_2} - 1)\nu(dn, dy_1, dy_2)$$

with

$$\nu(dn, dy_1, dy_2) = \sum_{k=0}^{\infty} p_k \delta_{\{k-1, \log(\frac{|a_k|}{p_k}), 1_{\{a_k<0\}}\}}(dn, dy_1, dy_2)$$

where $p_k = \frac{|a_k|}{\sum_{k=0}^{\infty} |a_k|}$.

## Stochastic representation of polynomial ODEs

- The stochastic representation is given by

$$\psi(t) = \log \mathbb{E}_{(1,0,0)}[e^{(\log g)N_t + Y_{1,t} + \mathrm{i}\pi Y_{2,t}}] .$$

or equivalently

$$u(t) = \mathbb{E}_{(1,0,0)}[g^{N_t} e^{Y_{1,t} + \mathrm{i}\pi Y_{2,t}}] .$$

if $(\log g, 1, \mathrm{i}\pi) \in \mathfrak{U}_t$. Note that the latter implies that

$$\partial_t u(t) = \sum_{k=0}^{\infty} |a_k| u^k(t) - u(t), \quad u(0) = g$$

admits also a stochastic representation given by $u(t) = \mathbb{E}_{(1,0)}[g^{N_t} e^{Y_{1,t}}]$.

- The process $(N, Y_1, Y_2)$ is a self-exciting jump process with intensity $N$. The component $N$ jumps with probability $p_k$ by $k-1$, $Y_1$ by $\log(\frac{|a_k|}{p_k})$ and $Y_2$ by 1 whenever $a_k$ is negative.

- Consider the following multivariate polynomial ODE

$$\partial_t u_i(t) = \sum_{|\mathbf{k}|=0}^{\infty} a^i_{\mathbf{k}} u^{\mathbf{k}}(t) - u_i(t), \quad u_i(0) = g_i, \quad i = 1, \ldots, d,$$

where $\mathbf{k} = (k_1, \ldots, k_d)$ denotes a multi-index and $u^{\mathbf{k}} = u_1^{k_1} \cdots u_d^{k_d}$ with $\sum_{|\mathbf{k}|=0}^{\infty} |a^i_{\mathbf{k}}| < \infty$.

- Define $\psi_i(t) := \log(u_i(t))$. Then

$$\begin{aligned}\partial_t \psi_i(t) &= \sum_{|\mathbf{k}|=0}^{\infty} a^i_{\mathbf{k}} u_1^{k_1} \cdots u_i^{k_i - 1} \cdots u_d^{k_d} - 1 \\ &= \sum_{|\mathbf{k}|=0}^{\infty} a^i_{\mathbf{k}} e^{\psi_1(t)k_1 + \cdots \psi_i(t)(k_i - 1) + \cdots \psi_d(t)k_d} - 1\,.\end{aligned}$$

- To recast it into the affine form, take a process $(N, Y_1, Y_2) \in \mathcal{D} = \mathbb{Z}_+^d \times \mathbb{R}_+^d \times \mathbb{Z}_+^d$, vectors $(h_1, h_2) = (\mathbf{1}, \mathrm{i}\pi\mathbf{1})$ and rewrite the ODE as

$$\partial\psi_i(t) = \int_{\mathcal{D}} \left( e^{\langle \psi(t), n\rangle + y_{1,i} + \mathrm{i}\pi y_{2,i}} - 1 \right) \nu_i(dn, dy_{1,i}, dy_{2,i})$$

where

$$\nu_i(dn, dy_{1,i}, dy_{2,i}) = \sum_{|\mathbf{k}|=0}^{\infty} p_{\mathbf{k}}^i \delta_{\{\mathbf{k}-e_i, \log(\frac{|a_{\mathbf{k}}^i|}{p_{\mathbf{k}}^i}), 1_{\{a_{\mathbf{k}}^i<0\}}\}}(dn, dy_{1,i}, dy_{2,i})$$

where $p_{\mathbf{k}}^i = \frac{|a_{\mathbf{k}}^i|}{\sum_{k=0}^{\infty} |a_{\mathbf{k}}^i|}$.

- The stochastic representation is then

$$\psi_i(t) = \log \mathbb{E}_{(e_i,0,0)}[e^{\langle \log g, N_t\rangle + \langle \mathbf{1}, Y_{1,t}\rangle + i\pi\langle \mathbf{1}, Y_{2,t}\rangle}] .$$

or equivalently

$$u_i(t) = \mathbb{E}_{(e_i,0,0)}[\prod_{j=1}^{d} g_j^{N_{j,t}} e^{\langle \mathbf{1}, Y_{1,t}\rangle + i\pi\langle \mathbf{1}, Y_{2,t}\rangle}] .$$

if $(\log(g), h_1, h_2) \in \mathfrak{U}_t$.

- The process $(N, Y_1, Y_2)$ is a self-exciting jump process with intensity $\langle N_t, \nu(\mathcal{D})\rangle$. The component $N$ jumps with probability $N_{(i,t)} p^i_{\mathbf{k}} / \langle N_t, \nu(\mathcal{D})\rangle$ by $\mathbf{k} - e_i$. In this case the $i^{th}$ component of $Y_1$ jumps by $\log(\frac{|a^i_{\mathbf{k}}|}{p^i_{\mathbf{k}}})$ and the $i^{th}$ component of $Y_2$ by 1 whenever $a^i_{\mathbf{k}}$ is negative.

Consider a state space $\mathcal{M} \subset \mathbb{R}^d$ and four vectors of Lévy measures $\nu_+^{\mathbf{r}}, \nu_-^{\mathbf{re}}, \nu_+^{\mathbf{im}}, \nu_-^{\mathbf{im}}$ corresponding to the characteristic vector fields $\mathcal{R}_\pm^{r/i}$. Only $\nu_+^{\mathbf{re}}$ is a generic Lévy measure of finite variation, all the others are assumed to be of finite activity. We assume the constant part $\mathcal{F}$ to vanish here since it is not important for the argument to come.

Assume furthermore that the sum over all measures

$$\nu = \nu_+^{\mathbf{re}} + \nu_-^{\mathbf{re}} + \nu_+^{\mathbf{im}} + \nu_-^{\mathbf{im}}$$

satisfies the admissibility conditions and describes a self-exciting pure jump affine (actually linear) process $N$ taking values in $\mathcal{M}$. Then one can construct a second affine process $\widetilde{N}$, actually a pure jump linear process, with state space $\mathcal{M} \times \mathbb{Z}^{2d}$ and corresponding Lévy measures $\widetilde{\nu}$ again being decomposable in four measures, too.

$$\widetilde{\nu} = \widetilde{\nu}_+^{\mathbf{re}} + \widetilde{\nu}_-^{\mathbf{re}} + \widetilde{\nu}_+^{\mathbf{im}} + \widetilde{\nu}_-^{\mathbf{im}}$$

Fix $i = 1, \ldots, d$: coordinate $i$ of the measure $\widetilde{\nu}_+^{\mathbf{re}}$ corresponds to the push forward along $\mathcal{M} \ni m \mapsto (m, 0, 0) \in \mathcal{M} \times \mathbb{Z}^{2d}$ of coordinate $i$ of $\nu_+^{\mathbf{re}}$; coordinate $i$ of the measure $\widetilde{\nu}_-^{\mathbf{re}}$ corresponds to the push forward along $\mathcal{M} \ni m \mapsto (m, e_i, 0) \in \mathcal{M} \times \mathbb{Z}^{2d}$ of coordinate $i$ of $\nu_-^{\mathbf{re}}$; coordinate $i$ of the measure $\widetilde{\nu}_+^{\mathbf{im}}$ corresponds to the push forward along $\mathcal{M} \ni m \mapsto (m, 0, e_i) \in \mathcal{M} \times \mathbb{Z}^{2d}$ of coordinate $i$ of $\nu_+^{\mathbf{im}}$; whereas coordinate $i$ of the measure $\widetilde{\nu}_-^{\mathbf{im}}$ corresponds to the push forward along $\mathcal{M} \ni m \mapsto (m, e_i, e_i) \in \mathcal{M} \times \mathbb{Z}^{2d}$ of coordinate $i$ of $\nu_-^{\mathbf{im}}$. All other jump measures necessary to fully specify the affine process $\widetilde{N}$ vanish.

The non-trivial components of the $\psi$ function of $\widetilde{N}$ started at $(f, \mathrm{i}\pi, \ldots, \mathrm{i}\pi, \mathrm{i}\pi/2, \ldots, \mathrm{i}\pi/2)$ actually solve

$$\partial\psi_t = \mathcal{R}_+^{\mathbf{re}}(\psi_t) - \mathcal{R}_i^{\mathbf{re}}(\psi_t) + \mathrm{i}\mathcal{R}_+^{\mathbf{im}}(\psi_t) - \mathrm{i}\mathcal{R}_-^{\mathbf{im}}(\psi_t) \tag{1}$$

$$= \int \big(\exp(\langle \psi_t, \xi \rangle) - 1)\eta(d\xi) = \mathcal{R}(\psi_t)\,, \tag{2}$$

where $\eta = \nu_+^{\mathbf{re}} - \nu_-^{\mathbf{re}} + \mathrm{i}\nu_+^{\mathbf{im}} - \mathrm{i}\nu_-^{\mathbf{im}}$ is a complex measure.

In other words loosely speaking we have stochastic representations for non-linear ODEs with vector fields being Fourier-Laplace transforms of finite complex-valued measures on a certain state space up to an explosion time depending on the initial value $f$.

Notice that we can also add an additive noise $W$ to the equation

$$d\psi_t = \mathcal{R}(\psi_t)\, dt + dW_t\,,$$

which finally leads to the stochastic representation

$$\exp\left(\psi^i(t,f)\right) =$$
$$= \mathbb{E}_{(e_i,0)}\Big[\exp(\langle \widetilde{N}_t, (f, \mathrm{i}\pi, \mathrm{i}\pi/2)\rangle)\exp(\int_0^t \langle N_s, d\,W_s\rangle)|\sigma(W)_t\Big]$$

up to an explosion time depending on $f$ and the trajectory of $W$.

# Some comments on simulation

In contrast to classical algorithms for ODEs the stochastic representation allows for parallelization. There are several competing techniques to simulate affine processes with different complexities depending on the situation:

- Euler scheme on determinstic grid.
- Euler schemes on stochastic grids.
- Random time change techniques.
- Branching Markov process techniques.

*Random time change techniques* and *Branching techniques* can be particularly interesting due to low complexity in many important situations, in our example we just applied the unbiased Euler scheme on a stochastic grid in dimension $10^6$ (sic!) with $i$th vector field coordinate depending on $\{i-1, i, i+1\}$ for $2 \leq i \leq 10^6 - 1$.

## Motivation from PHL (2012)

Non-linear PIDEs of the form

$$\big(\partial_t + \mathcal{L}\big)u(x,t) + F(u(t,x)) - u(t,x) = 0$$

with boundary condition $u(x,T) = g(x)$ allow for branching Markov process representations for certain types of non-linearities $F$.

Generically it holds true that

$$\begin{aligned} u(t,x) = \mathbb{E}_{(x,t)}\big[\exp(-(T-t))g(X_T)\big] + \\ + \int_t^T \mathbb{E}_{(x,t)}\big[\exp(-(s-t))F(u(s,X_s))\big]\,ds \end{aligned}$$

by the previous representation property. However, this is not a stochastic representation but rather a fixed point equation. Inserting the equation into itself leads towards a backwards algorithm or – under certain assumptions on $F$ – towards a branching tree representation.

Assume that $F$ is of the form

$$F(u) = \sum_{k=0}^{M} p_k u^k$$

with $p_k \geq 0$ and $\sum p_k = 1$, then the previous fixed point equation

$$u(t,x) = \mathbb{E}_{(x,t)}\big[\exp(-(T-t))g(X_T)\big] + \\ + \int_t^T \mathbb{E}_{(x,t)}\big[\exp(-(s-t))\sum_{k=0}^{M} p_k (u(s,X_s))^k\big]\, ds$$

leads to the short time asymptotics

$$
\begin{aligned}
u(t,x) = \exp(-(T-t))\mathbb{E}_{(x,t)}\big[g(X_T)\big] + \\
+ \sum_{k=0}^{M} p_k \int_t^T \exp(-(s-t))\mathbb{E}_{(x,t)}\big[\mathbb{E}_{(X_s,s)}[\prod_{j=1}^{k} g(X_T^{(j)})]\big]\, ds + \\
+ o\big((T-t)\big)\,,
\end{aligned}
$$

where $X^{(j)}$ denote independent copies of the Markov process $X$. We can now concatenate the short time asymptotics, since the expansion does not depend on $u$ anymore.

This leads to a *branching Markov process representation*, i.e. a Markov process whose state space at time $t$ is an integer number $k$ of individuals in state $(x^1, \ldots, x^k) \in D^k$. The particles move independently subject to the Markov process $X$ and they die at an exponential time with parameter 1 each after giving birth to a $l$ individuals with probability $p_l$ (which is called *branching*).

The number of particles in a measurable subset $A \subset D$ is an integer-measure-valued, self-exciting affine process. Let us denote the overall number of particles at time $T$ by $N_T$.

A similar consideration as before leads to the following stochastic representation formula

$$u(x,t) = \mathbb{E}_{(x,t)}\Big[\prod_{j=1}^{N_T} g(X_T^{(j)}) \times \prod_{k=0}^{M} \Big(\frac{a_k}{p_k}\Big)^{\#\{\text{branchings of type } k\}}\Big].$$

for equations with generic non-linearity

$$F(u) = \sum_{k=0}^{M} a_k u^k - u$$

and *auxiliary branching mechanism* $p_0, \ldots, p_M > 0$, $\sum p_k = 1$ governing the underlying branching process.

# Forward stochastic representation for semi-linear PIDEs

This is a stochastic representation by a Markov process, infinite dimensional though, which can be simulated forward such as in the linear case. The result has been brought to mathematical Finance by Pierre Henry-Labordere and extended by Henry-Labordere-Tan-Touzi, but roots in works of Dynkin, McKean, LeJan, Sznithman, etc.

- We can find forward stochastic representations for non-linear vector fields being of L-K-form on cones with respect to complex-valued measures even beyond local Lipschitz properties (which relates the result to Lions-DaPerna-Ambrosio theory).
- Similar representations also hold with additive noise.
- The results extend to semi-linear PIDEs when the underlying Markov process' state space (acting on types) is not finite anymore.
- The complexity of the method depends on the sparsity of the involved jump measures.