Non-linear (PI)DEs and affine processes

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

ETH Zürich

Autumn 2016

Why non-linear PIDEs?

Credit valuation problems, utility optimization problems, super-hedging problems, American option problems lead to non-linear PIDEs.

For instance the pricing of an American Option with payoff $g : D \to \mathbb{R}$ and time horizon $T > 0$ with respect to a Markovian model X on state space D leads to a value function $v : D \times [0, T] \to \mathbb{R}$ satisfying

$$\big(\partial_t + \mathcal{L}\big)v(x,t) = 1_{\{v(x,t)<g(x)\}}1_{\{(\partial_t+\mathcal{L})v(x,t)>0\}}\mathcal{L}g(x)$$

with boundary condition $v(x, T) = g(x)$ for $x \in D$. Here $\mathcal{L}$ denotes the generator of the Markov process X.

In other words: $(v(X_t, t))_{0\leq t\leq T}$ is the minimal super-martingale above $g(X_t)_{0\leq t\leq T}$.

Linear PIDEs of the form

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

with boundary condition $u(x,T) = g(x)$ allow for (Q)MC algorithms, i.e. for every $x \in D$ the value function has a stochastic representation of the form

$$u(x,t) = \mathbb{E}_{(x,t)}\big[g(X_T)\big]\,.$$

This expectation can be approximated by pseudo- or quasi-random samples of the Markov process Xthrough

$$\frac{1}{N}\sum_{i=1}^{N} g(X_T^{(i)})\,.$$

This is a robust, universal and almost dimension-free method to approximate the solution u if the increments of the Markov process X can be simulated with low complexity, even though slow, i.e. with rate of convergence $1/\sqrt{N}$ which bounds the error in a probabilistic sense.

Alternatives are finite difference or finite element methods with higher convergence rates bounding the error in a deterministic way, however, their complexity depends heavily on dimension.

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

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

leads to the short time asymptotics

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

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.

Forward stochastic representation for semi-linear PIDEs

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.

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 goes back to Henry-Labordere-Touzi-Wang, but roots in works of Dynkin, McKean, LeJan, Sznithman, etc.

At this point several questions arise:

- The distribution of individuals corresponds to an integer-measure-valued affine process: how is it possible that the branching measure can be replaced by a signed measure?
- Is the previous construction passing from measures to signed measures generic for affine processes on general state spaces?
- What is the relation of the auxiliary process and the functional solving the equation?

To demonstrate the essential structure let us assume for a moment $\mathcal{L} = 0$, then we are dealing with a branching process with just constant state.

We consider a Cole-Hopf transform together with time reversal $u_t = \exp(\psi_{T-t})$, then

$$\partial_t \psi_t \exp(\psi_t) = \sum_{k=0}^{M} p_k \exp(k\psi_t) - \exp(\psi_t)$$

with initial value $\psi_0 = \log g$. Let us denote by ν the law which takes the value $k-1$ (one ancestor dies!) with probability p_k, for $k = 1, \ldots, m$, then

$$\partial \psi_t = \mathcal{R}(\psi_t)$$

with

$$\mathcal{R}(f) := \int (\exp(f\xi) - 1)\, \nu(d\xi)\,.$$

This is a generalized Riccati equation for a one-dimensional self-exciting affine (actually linear) process, namely the number of individuals at a certain point in time in the branching process picture..

- stochastically continuous, time homogenous Markov processes N with state space $\mathcal{M}$ (usually a cone which is not necessarily pointed).
- the Fourier-Laplace transform of the marginal distribution of N is of exponential affine form

  $$\mathbb{E}_{(n,0)}\big[\exp(\langle N_t, f\rangle)\big] = \exp\big(\phi(t,f) + \langle \psi(t,f), n\rangle\big)$$

  where the functions ϕ and ψ satisfy so called generalized Riccati equations

  $$\partial_t \phi_t = \mathcal{F}(\psi_t) \text{ and } \partial_t \psi_t = \mathcal{R}(\psi_t)$$

  with $\phi(0,f) = 0$ and $\psi(0,f) = f$.

The vector fields $\mathcal{F}$ and $\mathcal{R}$ are of Lévy-Khintchine form. The classification of the specific form is fully understood on symmetric cones, $\mathbb{R}^m \times \mathbb{R}^n$, etc, leading to the so called admissibility conditions, *BUT* the Lévy measures are of course never signed. This can be found in the works of Cuchiero, Duffie, Filipovic, Keller-Ressel, Mayerhofer, Schachermayer, etc.

As often in mathematics it is fruitful to turn a point of view around:

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

The fact that in the Henry-Labordere-Touzi-Wang representation signed Lévy measures appear must have a meaning in the world of affine processes. The result of Henry-Labordere-Touzi-Wong suggests that there are more ODE types than generalized Riccati ones, which allow for representations.

A general result on non-linear ODEs

Consider a state space $\mathcal{M} \subset \mathbb{R}^d$ and four vectors of Lévy measures $\nu_+^{re}, \nu_-^{re}, \nu_+^{im}, \nu_-^{im}$ corresponding to the characteristic vector fields $\mathcal{R}_\pm^{r/i}$. Only ν_+^{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

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

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.

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

Fix $i = 1, \dots, d$: coordinate i of the measure $\widetilde{\nu}_+^{re}$ 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 ν_+^{re}; coordinate i of the measure $\widetilde{\nu}_-^{re}$ 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 ν_-^{re}; coordinate i of the measure $\widetilde{\nu}_+^{im}$ 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 ν_+^{im}; whereas coordinate i of the measure $\widetilde{\nu}_-^{im}$ 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 ν_-^{im}. All other jump measures necessary to fully specify the affine process $\widetilde{N}$ vanish.

The non-trivial components of the ψ 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

$$\begin{aligned} \partial\psi_t &= \mathcal{R}_+^{re}(\psi_t) - \mathcal{R}_i^{re}(\psi_t) + \mathrm{i}\mathcal{R}_+^{im}(\psi_t) - \mathrm{i}\mathcal{R}_-^{im}(\psi_t) \qquad (1) \\ &= \int \big(\exp(\langle \psi_t, \xi \rangle) - 1) \eta(d\xi) = \mathcal{R}(\psi_t)\,, \qquad (2) \end{aligned}$$

where $\eta = \nu_+^{re} - \nu_-^{re} + \mathrm{i}\nu_+^{im} - \mathrm{i}\nu_-^{im}$ 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.

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

$$\exp\left(\psi^i(t,f)\right) = \tag{3}$$

$$= \mathbb{E}_{(e_i,0)}\big[\exp(\langle \widetilde{N}_t, (f, \mathrm{i}\pi, \mathrm{i}\pi/2)\rangle) \exp(\int_0^t \langle N_s, d\,W_s\rangle)|\sigma(W)_t\big]\,. \tag{4}$$