Premia 18

POLYNOMIAL PROCESSES AND THEIR APPLICATIONS TO MATHEMATICAL FINANCE – PREMIA DOCUMENTATION FOR AN IMPLEMENTATION OF THE BATES MODEL

CHRISTA CUCHIERO, MARTIN KELLER-RESSEL, JOSEF TEICHMANN

Abstract. We introduce a class of Markov processes called m-polynomial, for which the calculation of (mixed) moments up to order m only requires the computation of matrix exponentials. This class contains affine processes, Feller processes with quadratic squared diffusion coefficient, as well as Lévy-driven SDEs with affine vector fields. Thus, many popular models such as the classical Black-Scholes, exponential Lévy or affine models are covered by this setting. The applications range from statistical GMM estimation to option pricing. For instance, the efficient and easy computation of moments can successfully be used for variance reduction techniques in Monte Carlo simulations.

1. Introduction

Pricing and hedging of contingent claims is the crucial computation done within every model in mathematical finance. For European type claims this amounts to the computation of the expected value of a functional of the (discounted) price process under some martingale measure. Hedging portfolios are then constructed via appropriate derivatives of those expected values with respect to model parameters or to the prices, so called "Greeks". Let us denote the (discounted) price process at time T, a vector in $\mathbb{R}^n$, by X_T. We can roughly distinguish three cases of complexity for the mentioned computations:

(1) The probability distribution of X_T is known analytically.
(2) The characteristic function of X_T is known analytically.
(3) The local characteristics of X_T are known analytically.

In the first case, a numerical quadrature algorithm is sufficient for the efficient computation of the contingent claim's price $\mathbb{E}[\phi(X_T)]$, where ϕ denotes some payoff function.

In the second case, variants of Plancherel's theorem are applied in order to evaluate the price $\mathbb{E}[\phi(X_T)]$, for instance,

$$\mathbb{E}[\phi(X_T)] = \int_{\mathbb{R}^n} \widehat{\phi}(u)\mathbb{E}\left[\exp(i\langle u, X_T\rangle)\right] du,$$

where $\widehat{\phi}$ denotes the Fourier transform of the function ϕ. Remark that often modifications of the original payoff function are used to make the Fourier methodology applicable. This is numerically efficient, even though its implementation, in particular the complex integration, can take some time (see, e.g., [2]). Also there are

Key words and phrases. affine process, analytic tractability, Feller processes, pricing, hedging.
MSC 2000. 60J25, 91B70.
All authors gratefully acknowledge the support from the FWF-grant Y 328 (START prize from the Austrian Science Fund).

different levels of what it means to "know analytically" the characteristic function of X_T. If one, e.g. in affine models, has to solve a high-dimensional Riccati equation for each $u \in \mathbb{R}^n$ to calculate the characteristic function $u \mapsto \mathbb{E}[\exp(i\langle u, X_T\rangle)]$, the "analytic knowledge" means at least some precalculations which have to be performed efficiently, too. In other words, methods applying the characteristic function are efficient if the amount of precalculations is limited.

The third case is characterized by the use of Monte Carlo simulation methods: one samples from the (unknown) distribution of X_T by generating, for instance through Euler schemes, approximate distributions for X_T. From those approximate distributions, which should be easy to simulate, one draws a sufficient amount of samples. This two-step procedure is very robust, but takes a considerable amount of time.

In this article we would like to add a fourth case, which in the previous order would correspond to case $2\frac{1}{2}$. We can describe a class of processes, called "polynomial processes", where it is easy and efficient to compute moments of all orders of the random variable X_T, even though neither its probability distribution nor its characteristic function need to be known. We shall analyze this class and show that exponential Lévy processes, affine processes or Jacobi-type processes belong to it. The method is best explained by an example: consider a stochastic volatility model of SVJJ-type [10], i.e. both the logarithmic (discounted) price process and the stochastic volatility can jump. Such models can be described by stochastic differential equations of the type

$$\begin{aligned} dX_t &= (aX_t + bV_t + c)dt + \sqrt{V_t}dB_{t,1} + dZ_{t,1}, \\ dV_t &= (\alpha V_t + \beta)dt + \sqrt{V_t}dB_{t,2} + dZ_{t,2}, \end{aligned}$$

where (B_1, B_2) are possibly correlated Brownian motions and (Z_1, Z_2) are Lévy processes, independent of (B_1, B_2), where the second component Z_2 is a subordinator. For such models there is no easy-to-implement (explicit) formula for the characteristic function, even though they are affine models. Assuming now appropriate moment conditions for the jump measures, the Markov process (X, V) has the remarkable property that the expected value of *any polynomial* of the process (X, V) is *again a polynomial in* X_0 *and* V_0. The coefficients of this polynomial can be calculated efficiently by exponentiating a matrix, which can be easily deduced from the generator. In other words, there is a dense subset of claims in the set of "all" claims, where the prices and hedging ratios are explicitly known (up to matrix exponentials). This explicit knowledge allows to compute prices of general claims by variance reduction techniques, which is more efficient and easier to implement than Fourier methods with "unknown" characteristic function.

The article is devoted to a study of polynomial processes in some depths together with their most obvious applications to mathematical finance. As the most striking application we see the possibility to reduce the variance of Monte Carlo evaluations within a polynomial model by approximating the claim through elements of the dense subset of already priced claims. The remainder of our article is organized as follows: in Section 2 we formally introduce the class of m-polynomial processes and draw several basic conclusions. In Section 3 we analyze the conditions on the characteristics of a Feller semimartingale to be m-polynomial. Section 4 deals with examples from the class of m-polynomial processes and Section 5 with applications to pricing and hedging in mathematical finance. In Section 6 we then describe the

implementation of the Bates model in Premia. The algorithm is based on a Monte Carlo simulation with polynomial control variates.

2. Polynomial Processes

We consider a time-homogeneous Markov process $X := (X_t^x)_{t\geq 0,\, x\in S}$ with state space $S \subseteq \mathbb{R}^n$, a closed subset of $\mathbb{R}^n$, and semigroup $(P_t)_{t\geq 0}$ defined by

$$P_t f(x) := \mathbb{E}[f(X_t^x)] = \int_S f(\xi) p_t(x, d\xi), \quad x \in S,$$

and acting on all functions $f : S \to \mathbb{R}$ which are integrable with respect to the family of Markov kernels $p_t(x, \cdot)$.

Notation 2.1. *Let* $\mathrm{Pol}_{\leq m}(S)$ *denote the finite dimensional vector space of polynomials up to degree* $m \geq 0$ *on* S*, i.e the restriction of polynomials on* $\mathbb{R}^n$ *to* S*, defined by*

$$\mathrm{Pol}_{\leq m}(S) := \left\{ \sum_{|\mathbf{k}|=0}^{m} \alpha_{\mathbf{k}} x^{\mathbf{k}} \Big| x \in S,\, \alpha_{\mathbf{k}} \in \mathbb{R} \right\}, \tag{2.1}$$

where we use the multi-index notation $\mathbf{k} = (k_1, \ldots, k_n) \in \mathbb{N}_0^n$*,* $|\mathbf{k}| = k_1 + \cdots + k_n$ *and* $x^{\mathbf{k}} = x_1^{k_1} \cdots x_n^{k_n}$*.* $\mathrm{Pol}_{\leq m}(S)$ *is endowed with some norm* $\|\cdot\|_{\mathrm{Pol}_{\leq m}}$ *and its dimension is denoted by* $N < \infty$*. Moreover,* $\mathrm{Pol}_m(S)$ *corresponds to the vector space of polynomials which are precisely of degree* m*.*

Definition 2.2. *We call an* S*-valued time-homogeneous Markov process* m-polynomial *if*

$$P_t f(x) \in \mathrm{Pol}_{\leq m}(S)$$

holds true for all $f \in \mathrm{Pol}_{\leq m}(S)$ *and* $t \geq 0$*. If* X *is* m*-polynomial for all* $m \geq 0$*, then it is called* polynomial*.*

Theorem 2.3. *Let* X *be a time-homogeneous Markov process with state space* S *and semigroup* (P_t)*, pointwise continuous at* $t = 0$*, i.e.* $t \mapsto P_t f(x)$ *is continuous at* $t = 0$ *for all* $x \in S$ *and* $f : S \to \mathbb{R}$ *where* $P_t f$ *exists as finitely valued function on* S*. Then the following assertions are equivalent:*

(i) X *is* m*-polynomial for some* $m \geq 0$*.*

(ii) *There exists a linear map* A *on* $\mathrm{Pol}_{\leq m}(S)$*, such that* (P_t) *restricted to* $\mathrm{Pol}_{\leq m}(S)$ *can be written as*

$$P_t|_{\mathrm{Pol}_{\leq m}(S)} = e^{tA}$$

for all $t \geq 0$*.*

(iii) *The infinitesimal generator* $\mathcal{A}$ *is well defined on* $\mathrm{Pol}_{\leq m}(S)$ *and maps* $\mathrm{Pol}_{\leq m}(S)$ *to itself.*

(iv) *The Kolmogorov backward equation for an initial value* $f(\cdot, 0) \in \mathrm{Pol}_{\leq m}(S)$

$$\frac{\partial f(x,t)}{\partial t} = \mathcal{A} f(x,t)$$

has a real analytic solution for all times $t \in \mathbb{R}$*. In particular,* $f(\cdot, t) \in \mathrm{Pol}_{\leq m}(S)$*.*

3. Polynomial Feller semimartingales

For the class of Feller processes, we want to find sufficient conditions for m-polynomial processes in terms of the infinitesimal generator. We consider therefore a Feller semigroup (P_t) on S. If $C_c^\infty(S)$, the space of smooth test functions, is contained in the domain of the infinitesimal generator $\mathcal{A}$, then it is well known that there exist real-valued functions a_{kl}, b_k, c and a kernel $K(x, d\xi)$ on $S \times \mathcal{B}(\mathbb{R}^n \backslash \{0\})$ such that for $u \in C_c^2(S)$, $\mathcal{A}$ is given by

$$\begin{aligned}\mathcal{A}u(x) = &\frac{1}{2}\sum_{k,l=1}^n a_{kl}(x)\frac{\partial^2 u(x)}{\partial x_k \partial x_l} + \sum_{k=1}^n b_k(x)\frac{\partial u(x)}{\partial x_k} - c(x)u(x) \\ &+ \int_{\mathbb{R}^n\backslash\{0\}} \left(u(x+\xi) - u(x) - \sum_{k=1}^n \chi_k(\xi)\frac{\partial u(x)}{\partial x_k} \right) K(x, d\xi),\end{aligned} \tag{3.1}$$

where $a(x) = (a_{kl}(x))_{k,l=1,\dots,n}$ is a symmetric positive semidefinite matrix, $b(x) \in \mathbb{R}^n$, c is non-negative, $K(x, \cdot)$ is a Radon measure on $S \times \mathcal{B}(\mathbb{R}^n\backslash\{0\})$ and $\chi : \mathbb{R}^n \to \mathbb{R}^n$ some bounded continuous truncation function with $\chi(\xi) = \xi$ in a neighborhood of 0 (see for example [13] or [16]). Clearly, the above parameters have to satisfy certain admissibility conditions guaranteeing the existence of the process in S.
As the definition of m-polynomial processes requires the existence of moments up to order m and thus the finiteness of X_t a.s., we henceforth assume X to be conservative, i.e. $c = 0$.
If X is additionally a semimartingale, which is automatically the case if $S = \mathbb{R}^n$ (see Theorem 7.16 in [4]), then its characteristics (B, C, ν) associated with the truncation function $\chi(\xi)$ are given by

$$B_t = \int_0^t b(X_s)ds, \quad C_t = \int_0^t a(X_s)ds, \quad \nu(dt, d\xi) = K(X_t, d\xi)dt.$$

We shall refer to (b, a, K) as *differential characteristics* of X (see [14]).
In order to specify the form of a, b and K such that $\mathcal{A}$ generates an m-polynomial process, we start by defining the following two conditions on the kernel $K(x, d\xi)$:

Condition A. *The kernel $K(x, d\xi)$ is of the form*

$$K(X_t, d\xi) := \mu_{00}(d\xi) + \sum_{i\in I} X_{t,i}\mu_{i0}(d\xi) + \sum_{(i,j)\in J} X_{t,i}X_{t,j}\mu_{ij}(d\xi),$$

where all μ_{ij} are Lévy measures on $\mathbb{R}^n$ with

$$\int_{\|\xi\|>1} \|\xi\|^m \mu_{ij}(d\xi) < \infty. \tag{3.2}$$

The index sets I and J are defined by

$$I = \{1 \le i \le n | S_i \subseteq \mathbb{R}_+\}$$

and

$$J = \{(i,j),\, i \le j | i = j \text{ or } S_i \times S_j \subseteq \mathbb{R}_+^2 \text{ or } S_i \times S_j \subseteq \mathbb{R}_-^2\},$$

where S_i stands for the projection on the i^{th} component.

Condition B. *Let $d \ge 1$ and let*

$$g : S \times \mathbb{R}^d \to \mathbb{R}^n, \quad (x, y) \mapsto g(x, y) = g^x(y) = H(y)x + h(y),$$

be an affine function in x. Here, $H : \mathbb{R}^d \to \mathbb{R}^{n\times n}$ and $h : \mathbb{R}^d \to \mathbb{R}^n$ are assumed to be measurable. The kernel $K(x, d\xi)$ satisfies

$$K(X_t, d\xi) := g_*^{X_t}\mu(d\xi),$$

where for each $x \in S$, $g_*^x\mu$ denotes the pushforward of the measure μ under the map g^x. Moreover, μ is a Lévy measure on $\mathbb{R}^d$ integrating

$$\int_{\mathbb{R}^d\setminus\{0\}} \left(\|H(y)\|^k + \|h(y)\|^k\right)\mu(dy) \quad \text{for all } 1 \le k \le m. \tag{3.3}$$

Remark 3.1. *The integrability condition (3.3) (in particular for $k = 1$) can be weakened for specific choices of H and h (see Example 4.3).*

Theorem 3.2. *Let $m \ge 2$ and X_t^x be a conservative Feller semimartingale on S whose infinitesimal generator on $C_c^\infty(S)$ is of form (3.1) with $c = 0$. Assume furthermore that $\mathbb{E}\left[\|(X_t^x)\|^m\right] < \infty$ for all $t \in [0, 1]$ (of course $[0, 1]$ could be replaced by $[0, \varepsilon]$ for any $\varepsilon > 0$). Then, X is m-polynomial if its differential characteristics (b, a, K) associated with the "truncation function" $\chi(\xi) = \xi$ are of the form*

$$b_t = b + \sum_{i=1}^n X_{t,i}\beta_i, \quad b,\, \beta_i \in \mathbb{R}^n,$$

$$a_t = a + \sum_{i=1}^n X_{t,i}\alpha_{i0} + \sum_{i\le j} X_{t,i}X_{t,j}\alpha_{ij}, \quad a,\, \alpha_{ij} \in \mathbb{R}^{n\times n},$$

with K satisfying either Condition A or B.

Remark 3.3. *It is important to note that the differential characteristics of X in Theorem 3.2 are specified with respect to the "truncation function" $\chi(\xi) = \xi$. While a and K do not depend on the choice of χ, the characteristic $b = b(\chi)$ does. So, if one chooses another truncation function $\widetilde{\chi}$ instead of ξ, then $b(\widetilde{\chi})$ transforms as follows*

$$b_t(\widetilde{\chi}(\xi)) = b_t(\xi) + \int_{\mathbb{R}^n\setminus\{0\}} (\widetilde{\chi}(\xi) - \xi)\, K(X_t, d\xi).$$

Thus, the requirement that a and K are as in Theorem 3.2 and

$$\left(b_t(\widetilde{\chi}(\xi)) + \int_{\mathbb{R}^n\setminus\{0\}} (\xi - \widetilde{\chi}(\xi))\, K(X_t, d\xi)\right) \in \mathrm{Pol}_{\le 1}(S)$$

is an equivalent condition guaranteeing that X is m-polynomial.

Remark 3.4. *Note that an operator of form (3.1) with $\chi(\xi) = \xi$ can satisfy*

$$\mathcal{A}(\mathrm{Pol}_{\le m}(S)) \subseteq \mathrm{Pol}_{\le m}(S), \tag{3.4}$$

even though the conditions of Theorem 3.2 are not fulfilled. An example of such an operator is of the following form:

$$\begin{aligned}\mathcal{A}u(x) &= \frac{1}{2}\left(a + \alpha_{10}x + \alpha_{11}x^2\right)\frac{d^2u(x)}{dx^2} + (b + \beta x)\frac{du(x)}{dx} \\ &\quad + \frac{\lambda}{2x^2}\int_{\mathbb{R}}\left(u(x+\xi) - u(x) - \xi\frac{du(x)}{dx}\right)\delta_{-2x}(d\xi) \\ &= \frac{1}{2}\left(a + \alpha_{10}x + \alpha_{11}x^2\right)\frac{d^2u(x)}{dx^2} + \left(b + \beta x + \frac{\lambda}{x}\right)\frac{du(x)}{dx} + \frac{\lambda(u(-x) - u(x))}{2x^2}.\end{aligned}$$

The case $a = 2$, $\alpha_{10} = \alpha_{11} = b = \beta = 0$ corresponds to the generator of the so called Dunkl process [7, 9].

4. Examples

In order to apply Theorem 3.2 to the following examples, we assume throughout this section $m \geq 2$, i.e. the process X admits moments up to order at least 2.

Example 4.1 (Affine processes)**.** *Every conservative, regular affine process X on $S = \mathbb{R}^p_+ \times \mathbb{R}^{n-p}$ is m-polynomial, if $\mathbb{E}\left[\|(X^x_t)\|^m\right] < \infty$ for $t \in [0,1]$. For details on affine processes see* [5].

Example 4.2 (Lévy processes)**.** *Let L be a Lévy process on $\mathbb{R}^n$ with triplet (b, a, μ) satisfying $\int_{\|\xi\|>1} \|\xi\|^m \mu(d\xi) < \infty$. Then, the Markov process $X^x_t = x + L_t$ is m-polynomial.*

Example 4.3 (Exponential Lévy models)**.** *Exponential Lévy models are of the form*

$$X^x_t = xe^{L_t}, \tag{4.1}$$

where L_t is a Lévy process on $\mathbb{R}$ with triplet (b, a, μ). Under the integrability assumption

$$\int_{|y|>1} e^{my} \mu(dy) < \infty, \tag{4.2}$$

which guarantees the existence of $\mathbb{E}\left[|X^x_t|^m\right]$, exponential models are m-polynomial, since we have

$$\mathbb{E}\left[x^m e^{mL_t}\right] = x^m e^{t\psi(m)},$$

where ψ denotes the cumulant generating function of the Lévy process.

In financial applications X_t usually corresponds to the price process under a martingale measure. The absence of arbitrage then imposes conditions on b since the discounted price process $e^{-rt}X_t$ must be a martingale:

$$b = r - \frac{1}{2}a - \int_{\mathbb{R}} \left(e^y - 1 - \chi(y)\right) \mu(dy).$$

The infinitesimal generator of (4.1) is then given by

$$\mathcal{A}u(x) = \frac{ax^2}{2}\frac{d^2u(x)}{dx^2} + rx\frac{du(x)}{dx} + \int_{\mathbb{R}} \left(u\left(xe^y\right) - u(x) - x\left(e^y - 1\right)\frac{du(x)}{dx} \right) \mu(dy).$$

In terms of Theorem 3.2, we are in the situation of Condition B with $g(x, y) = H(y)x = (e^y - 1)x$.
Note also that the Black-Scholes model falls into the realm of this example.

Example 4.4 (Lévy driven SDEs)**.** *Let L_t denote a Lévy process on $\mathbb{R}^d$ with generating triplet (b, a, μ). Suppose furthermore that $V_1, \ldots, V_d$ are affine functions, i.e. we have*

$$V_i : S \to \mathbb{R}^n, \quad x \mapsto H_i x + h_i,$$

where $H_i \in \mathbb{R}^{n\times n}$ and $h_i \in \mathbb{R}^n$. A process X which solves the stochastic differential equation of type

$$dX_t = \sum_{i=1}^{d} V_i(X_{t-})dL_{t,i}, \quad X_0 = x \in S, \tag{4.3}$$

in S and which leaves S invariant is m-polynomial if $\int_{\|\xi\|>1} \|\xi\|^m \mu(d\xi) < \infty$. If $V_1, \dots, V_d$ are linear, then we even have $\mathcal{A}(\mathrm{Pol}_k(S)) \subseteq \mathrm{Pol}_k(S)$ for all $0 \le k \le m$ (in contrast to $\mathcal{A}(\mathrm{Pol}_{\le k}(S)) \subseteq \mathrm{Pol}_{\le k}(S)$).

Example 4.5 (Jacobi process)**.** *Another example of a polynomial process is the Jacobi process (see [12]) which is the solution of the stochastic differential equation*

$$dX_t = -\beta(X_t - \theta)dt + \sigma\sqrt{X_t(1-X_t)}dB_t, \quad X_0 = x \in [0,1],$$

on $S = [0,1]$, where $\theta \in [0,1]$ and $\beta, \sigma > 0$. Its generator is a self-adjoint operator satisfying the following eigenvalue equation

$$\mathcal{A}Q_n(x) = \lambda_n Q_n(x), \quad n \in \mathbb{N},$$

where the eigenfunctions Q_n are the Jacobi polynomials and the eigenvalues λ_n are given by $-\frac{\sigma^2}{2}n(n-1+\frac{2\beta}{\sigma^2})$.
This example can be extended by adding jumps, where the jump times correspond to those of a Poisson process N with intensity λ and the jump size is a function of the process level. Indeed, if a jump occurs, then the process is reflected at $\frac{1}{2}$ so that it remains in the interval $[0,1]$, i.e. we have

$$dX_t = -\beta(X_t - \theta)dt + \sigma\sqrt{X_t(1-X_t)}dB_t + (1-2X_t)dN_t, \quad X_0 = x \in [0,1],$$

whose generator is given by

$$\mathcal{A}u = \frac{1}{2}\sigma^2(x(1-x))\frac{d^2u(x)}{dx^2} - \beta(x-\theta)\frac{du(x)}{dx} + \lambda(u(1-x) - u(x)).$$

In terms of Condition B, we have here $g(x,y) = -2yx + y$ and $\mu(dy) = \lambda\delta_1(dy)$.

Example 4.6 (Pearson diffusions)**.** *The above example 4.5 (without jumps) as well as Ornstein-Uhlenbeck and Cox-Ingersoll-Ross processes, all of them with mean-reverting drift, can be subsumed under the class of so called Pearson diffusions which are the solutions to SDEs of the form*

$$dX_t = -\beta(X_t - \theta)dt + \sqrt{(\alpha_{11}X_t^2 + \alpha_{10}X_t + a)}dB_t, \quad X_0 = x,$$

where $\beta > 0$ and α_{11}, α_{10} and a are specified such as the square root is well defined. Forman and Sørensen [8] give a complete classification of the different types of the Pearson diffusion in terms of their invariant distributions.

5. Applications

By Theorem 2.3 we know that there exists a linear map A such that moments of m-polynomial processes can simply be calculated by computing e^{tA}. Indeed, by choosing a basis $\langle e_1, \dots, e_N \rangle$ of $\mathrm{Pol}_{\le m}(S)$ the matrix corresponding to this linear map which we also denote by $A = (A_{ij})_{i,j=1,\dots,N}$ can be obtained through

$$\mathcal{A}e_i = \sum_{j=1}^{N} A_{ij}e_j.$$

Writing f as $\sum_{k=1}^{N} \alpha_k e_k$, we then have

$$P_t f = e^{tA}\left(\sum_{k=1}^{N} \alpha_k e_k\right) = (\alpha_1, \dots, \alpha_N)e^{tA}(e_1, \dots, e_N)', \tag{5.1}$$

which means that moments of polynomial processes can be evaluated simply by computing matrix exponentials.
By means of the one-dimensional Cox-Ingersoll-Ross process

$$dX_t = (b + \beta X_t)dt + \sigma\sqrt{X_t}dW_t, \quad b, \sigma \in \mathbb{R}_+, \quad \beta \in \mathbb{R},$$

we exemplify how moments of order m can be calculated. The generator is given by

$$\mathcal{A}u(x) = \frac{1}{2}\sigma^2 x \frac{d^2u(x)}{dx^2} + (b + \beta x)\frac{du(x)}{dx}.$$

Applying $\mathcal{A}$ to $(x^0, x^1, \dots, x^m)$ yields the following $(m+1) \times (m+1)$ matrix

$$A = \begin{pmatrix} 0 & \dots & & & & \\ b & \beta & 0 & \dots & & \\ 0 & 2b+\sigma^2 & 2\beta & 0 & \dots & \\ 0 & 0 & 3b+3\sigma^2 & 3\beta & 0 & \dots \\ & & & & \ddots & \\ 0 & \dots & & & mb + \frac{m(m-1)}{2}\sigma^2 & m\beta \end{pmatrix}.$$

Hence, $\mathbb{E}\big[(X_t^x)^k\big] = P_t x^k = (0, \dots, 1, \dots, 0)e^{tA}(x^0, \dots, x^k, \dots, x^m)'$.

Remark 5.1. *Note that A is a lower triangular matrix, whose eigenvalues are the diagonal elements. Since in this case they are all distinct, the matrix is diagonalizable. This holds true for all one-dimensional affine processes since the $(k+1)^{th}$ diagonal element is given by*

$$k\left(\beta + \int_{\mathbb{R}_+} (\xi - \chi(\xi))\mu_{10}(d\xi)\right),$$

where the integral part can only appear if the process is supported on $\mathbb{R}_+$. Recall the notation μ_{10} from Condition A. Of course there are many efficient algorithms to evaluate such matrix exponentials (see for example [11, 15]).

5.1. **Moment calculation - Generalized Method of Moments (GMM).** In view of this easy and fast technique of moment calculation for polynomial processes, the Generalized Method of Moments (GMM) qualifies for parameter estimation and thus, for model calibration. The implementation of a typical moment condition of the type

$$f(X_t, \theta) = \begin{pmatrix} X_t^{n_1} X_{t+s}^{m_1} - \mathbb{E}[X_t^{n_1} X_{t+s}^{m_1}] \\ \vdots \\ X_t^{n_q} X_{t+s}^{m_q} - \mathbb{E}[X_t^{n_q} X_{t+s}^{m_q}] \end{pmatrix}, \quad n_i, m_i \in \mathbb{N}, \quad 1 \le i \le q,$$

where θ is the set of parameters to be estimated, is simple since

$$\mathbb{E}\big[X_t^n X_{t+s}^m\big] = \mathbb{E}\big[X_t^n \mathbb{E}\big[X_{t+s}^m \big| X_t\big]\big]$$

can also be computed easily. In the case of one-dimensional jump-diffusions, Zhou [17] already uses this method for GMM estimation.

5.2. **Pricing - Variance reduction.** The fact that moments of polynomial processes are analytically known also gives rise to new and efficient techniques for pricing and hedging issues.
Let X be an m-polynomial process and $G : S \to \mathbb{R}^n$ a deterministic bi-measurable map such that the (discounted) price processes are given through

$$S_t = G(X_t)$$

under a martingale measure. Typically $G = \exp$, if X are log-prices. We denote by $F = \phi(S_T)$ a bounded measurable European claim for some maturity $T > 0$, whose price at $t \geq 0$ is given by the risk neutral evaluation formula

$$p_t^F = \mathbb{E}\big[\phi(S_T)\big|\mathcal{F}_t\big] = \mathbb{E}\big[(\phi \circ G)(X_T)\big|\mathcal{F}_t\big].$$

Obviously, claims of the form

$$F = f \circ G^{-1}(S_T) \tag{5.2}$$

for $f \in \mathrm{Pol}_{\leq m}(S)$ are analytically tractable, as we have

$$p_t^F = \mathbb{E}\big[(f \circ G^{-1})(S_T)\big|\mathcal{F}_t\big] = P_{T-t}f(G^{-1}(S_t)) = e^{(T-t)A}f(G^{-1}(S_t))$$

for $0 \leq t \leq T$, where A is the previously defined linear operator on $\mathrm{Pol}_{\leq m}(S)$. The sensitivities of the price process with respect to the factors of X can then be calculated by

$$\nabla p_t^F = \nabla P_{T-t}f(G^{-1}(S_t))\nabla G^{-1}(S_t). \tag{5.3}$$

Although claims are in practice not of form (5.2), the explicit knowledge of the price of polynomial claims can be used for variance reduction techniques based on control variates. Instead of using the estimator

$$\pi_0^F = \frac{1}{L}\sum_{i=1}^{N}(\phi \circ G)(X_T^i)$$

in a Monte-Carlo simulation, where $X_N^1, \ldots, X_T^N$ are L samples of X_T, we can use

$$\hat{\pi}_0^F = \frac{1}{N}\sum_{i=1}^{N}\left((\phi \circ G)(X_T^i) - \left(f(X_T^i) - \mathbb{E}[f(X_T)]\right)\right),$$

where $f \in \mathrm{Pol}_{\leq m}(S)$ is an approximation of $\phi \circ G$ and serves as control variate. Both estimators are unbiased and the second clearly outperforms the first since $\mathrm{Var}\left(\hat{\pi}_0^F\right) < \mathrm{Var}\left(\pi_0^F\right)$, where the ratio of the variances depends on the accuracy of the polynomial approximation.

It is worth mentioning that the previous pricing algorithm has also important consequences for hedging, since the greeks for "polynomial claims" $F = f(X_T)$ can be explicitly and efficiently calculated, again by matrix exponentials: the matrix exponential calculates the coefficients of the polynomial $x \mapsto \mathbb{E}[f(X_T)]$, calculating a derivative of this polynomial is then an algebraic question. These considerations are then applied in equation (5.3). Now assuming a complete market situation for the real-world claim $\phi(S_T) = \phi \circ G(X_T)$, i.e. there is a trading strategy η such that

$$\phi(S_T) = \mathbb{E}[\phi(S_T)] + \int_0^T \eta_t \bullet dS_t,$$

we can conclude that

$$\phi(S_T) - f(X_T) = \mathbb{E}[\phi(S_T)] - \mathbb{E}[f(X_T)] + \int_0^T (\eta - \nabla p_t^F) \bullet dS_t.$$

Therefore, if we assume that $\phi(S_T) - f(X_T)$ has a small variance, then also the stochastic integral representing the difference of the cumulative gains and losses of the two hedging portfolios, namely the one built by the unknown strategy η and the one built by the known strategy ∇p_t^F, is small.

Of course, this method of variance reduction can be applied to all polynomial processes, including all kinds of examples mentioned in Section 4. However, affine models for which the generalized Riccati ODEs (see [5]) cannot be explicitly solved are of particular interest. This is illustrated in the example below.

Example 5.2. *The following affine stochastic volatility model comprising two volatility factors is a modification of a model initially proposed by Bates* [1]*. The price process is specified as* $S_t = S_0 e^{X_t}$ *with dynamics*

$$\begin{aligned} d\begin{pmatrix} X_t \\ U_t \\ V_t \end{pmatrix} &= \begin{pmatrix} r - \frac{U_t}{2} - \frac{V_t}{2} - \lambda_1 U_t \int \left(e^{\xi_1} - 1\right) F_1(d\xi) - \lambda_2 V_t \int \left(e^{\xi_1} - 1\right) F_2(d\xi)) \\ b_1 - \beta_{11} U_t + \beta_{12} V_t \\ b_2 - \beta_{22} V_t + \beta_{21} U_t \end{pmatrix} dt \\ &+ \begin{pmatrix} \sqrt{U_t} & 0 & \sqrt{V_t} & 0 \\ \sigma_1 \rho_1 \sqrt{U_t} & \sigma_1 \sqrt{1-\rho_1^2}\sqrt{U_t} & 0 & 0 \\ 0 & 0 & \sigma_2 \rho_2 \sqrt{V_t} & \sigma_2 \sqrt{1-\rho_2^2}\sqrt{V_t} \end{pmatrix} \begin{pmatrix} dB_{t,1} \\ dB_{t,2} \\ dB_{t,3} \\ dB_{t,4} \end{pmatrix} \\ &+ dZ_{t,1} + dZ_{t,2}, \end{aligned}$$

where for $k = 1, 2$*,* Z_k *are pure jump processes in* $\mathbb{R} \times \mathbb{R}_+^2$ *with linear jump arrival intensity* $\lambda_1 u$*,* $\lambda_2 v$ *and trivariate jump size distribution* F_k *such that* $\int e^{\xi_1} F_k(d\xi) < \infty$*. The usual way to obtain the price of European options is to solve the following Riccati equations*

$$\begin{aligned} \frac{\partial \phi(t,x,u,v)}{\partial t} &= rx + b_1 \psi_1(t,x,u,v) + b_2 \psi_2(t,x,u,v), \\ \frac{\partial \psi_1(t,x,u,v)}{\partial t} &= \frac{1}{2}(x^2 - x) - \beta_{11}\psi_1(t,x,u,v) + \beta_{21}\psi_2(t,x,u,v) \\ &+ \frac{1}{2}\sigma_1^2 \psi_1^2(t,x,u,v) + \rho_1 \sigma_1 x \psi_1(t,x,u,v) \\ &+ \lambda_1 \left(\int_{\mathbb{R}\times\mathbb{R}_+^2} \left(e^{x\xi_1 + \psi_1(t,x,u,v)\xi_2 + \psi_2(t,x,u,v)\xi_3} - 1 \right) F_1(d\xi) \right. \\ &\left. - x \int_{\mathbb{R}\times\mathbb{R}_+^2} (e^{\xi_1} - 1) F_1(d\xi) \right). \\ \frac{\partial \psi_2(t,x,v)}{\partial t} &= \frac{1}{2}(x^2 - x) - \beta_{22}\psi_2(t,x,u,v) + \beta_{12}\psi_1(t,x,u,v) \\ &+ \frac{1}{2}\sigma_2^2 \psi_2^2(t,x,u,v) + \rho_2 \sigma_2 x \psi_2(t,x,u,v) \\ &+ \lambda_2 \left(\int_{\mathbb{R}\times\mathbb{R}_+^2} \left(e^{x\xi_1 + \psi_1(t,x,u,v)\xi_2 + \psi_2(t,x,u,v)\xi_3} - 1 \right) F_2(d\xi) \right. \\ &\left. - x \int_{\mathbb{R}\times\mathbb{R}_+^2} (e^{\xi_1} - 1) F_2(d\xi) \right). \end{aligned}$$

and to apply Fourier pricing methods as suggested in [6]. *As in this case explicit solutions of the Riccati equations are not available, our approach to use an approximating polynomial as control variate in the Monte-Carlo simulation is particularly expedient. For the calculation of the matrix exponential yielding the price of the polynomial claim, we only need to apply the infinitesimal generator* $\mathcal{A}$ *to polynomials of the form* $x^{k_1}u^{k_2}v^{k_3}$, $k_1+k_2+k_3 \leq m$ *in order to determine the matrix* A. *The price of the approximating polynomial claim is then simply calculated by means of* (5.1).

6. Implementation in Premia

6.1. **Model description.** In Premia, we implemented the above described algorithm to price an European call or put option under the original Bates model [1] (a simplified 2-dimensional version of Example 5.2). The function is called **MCCuchieroKellerResselTeichmann**.
In the Bates model the price process is given by $S_t = S_0 e^{X_t}$ with dynamics of the form

$$d\begin{pmatrix} X_t \\ V_t \end{pmatrix} = \begin{pmatrix} r - divid - \frac{V_t}{2} - \lambda(e^{\mu+\frac{\gamma_2}{2}} - 1) \\ \kappa(\theta - V_t) \end{pmatrix} dt + \begin{pmatrix} \sqrt{V_t} & 0 \\ \sigma_v \rho \sqrt{V_t} & \sigma_v \sqrt{1-\rho^2}\sqrt{V_t} \end{pmatrix} \begin{pmatrix} dB_{t,1} \\ dB_{t,2} \end{pmatrix} + \begin{pmatrix} dZ_t \\ 0 \end{pmatrix},$$

where Z is a pure jump process in $\mathbb{R}$ with jump intensity λ and normally distributed jump sizes with mean μ and variance γ_2. Additionally to the option and model parameters the function **MCCuchieroKellerResselTeichmann** requires the following inputs

- N: number of iterations in the Monte Carlo simulation,
- M: number of time steps in the Euler scheme,
- Nb_Degree_Pol: degree of approximating polynomial for variance reduction, between 0 and 8,
- generator: specification of the random generator,
- confidence: confidence level (for example 0.05, needed to determine the the upper and lower bounds for the option price).

and produces then the following outputs

- Price,
- Delta,
- Error Price,
- Error Delta,
- Inf Price,
- Sup Price,
- Inf Delta,
- Sup Delta.

Beside the procedures available in the PNL library, the function **MCCuchieroKellerResselTeichmann** calls

- **pol_approx** and
- **matrix_computation** (which in turn calls **moments_normal**).

These procedures are described below.

6.2. **Implementation.** As described above, we approximate the (undiscounted) European call/put payoff function by a polynomial whose degree can be chosen between 0 and 8. The function doing this job is called **pol_approx**. Therein we generate a vector x of equally spaced points with distance 0.01 in the interval $[\log(S_0/10), \log(10S_0)]$ and evaluate the payoff function at these points (vector y). The approximating polynomial is then obtained by applying the procedure **pnl_fit_least_squares**(coeff,x,y,&f,Nb_Degree_Pol+1), where f corresponds to the CANONICAL basis and Nb_Degree_Pol denotes the degree of the polynomial (between 0 and 8). The coefficients of the approximating polynomial are stored in coeff in increasing order starting with the coefficient corresponding to degree 0. The coefficients of the polynomial approximating the option's delta are simply calculated by taking the derivative of the polynomial approximating the payoff function.

In order to compute the expectation of the approximating polynomial, the next step is the implementation of the matrix A corresponding to the generator of the Bates model applied to the polynomials $x^i v^j$, $(i+j) \leq$ Nb_Degree_Pol, where x corresponds to the logprice and v to the variance. The procedure computing A is called **matrix_computation**. In our case the infinitesimal generator is given by

$$\begin{aligned} Au(x,v) = &\left(r - divid - \frac{v}{2} - \lambda(e^{\mu+\frac{\gamma_2}{2}} - 1)\right)\frac{\partial u}{\partial x} + \kappa\,(\theta - v)\,\frac{\partial u}{\partial v} \\ &+ \frac{v}{2}\frac{\partial^2 u}{\partial x^2} + \frac{v\sigma_v^2}{2}\frac{\partial^2 u}{\partial v^2} + \sigma_v \rho v \frac{\partial^2 u}{\partial v \partial x} \\ &+ \lambda \int (u(x+\xi, v) - u(x,v))\,\frac{1}{\gamma_2\sqrt{2\pi}} e^{-\left(\frac{(\xi-\mu)^2}{2\gamma_2^2}\right)} d\xi. \end{aligned} \tag{6.1}$$

When u is a polynomial, the evaluation of the integral in (6.1) is equivalent to computing the moments of the normal distribution. This computation is carried out in the procedure **moments_normal**. Finally, the expectation of the polynomial approximating the option's payoff (the option's delta) is denoted by polyprice (polydelta respectively) and calculated as described at the beginning of Section 5 by exponentiating the matrix A.

Having implemented the price of the polynomial payoff, we can start the Monte Carlo simulation. Using an Euler scheme we sample the log stock price at option maturity T and evaluate the approximating polynomial for the option (for the delta) at each of the replications. This value is then denoted by poly_sample (delta_poly_sample respectively). As described in Subsection 5.2, the estimator for the (call) option price is finally obtained by

$$\hat{\Pi}_0^{e^{-rT}(e^{X_T}-K)^+} = \frac{1}{N}\sum_{i=1}^{N} e^{-rT}\left((e^{X_T^i} - K)^+ - (\text{poly\_sample} - \text{polyprice})\right),$$

where K denotes the strike price. The estimator for the option's delta is computed analogously. This then gives the output values for "Price" and "Delta". The remaining output values simply correspond to the empirical standard deviations and the bounds for the confidence interval of "Price" and "Delta".

References

[1] D. S. Bates. Post-'87 crash fears in the S&P 500 futures option market. *Journal of Econometrics*, 94(1-2):181–238, 2000. 10, 11

[2] P. Carr and D. Madan. Option valuation using the fast Fourier transform. *Journal of Computational Finance*, 2(4):61–73, 1999. 1
[3] L. Chen, D. Filipović, and H. Vincent. Quadratic term structure models for risk-free and defaultable rates. *Math. Finance*, 14(4):515–536, 2004.
[4] E. Çinlar, J. Jacod, P. Protter, and M. J. Sharpe. Semimartingales and Markov processes. *Z. Wahrsch. Verw. Gebiete*, 54(2):161–219, 1980. 4
[5] D. Duffie, D. Filipović, and W. Schachermayer. Affine processes and applications in finance. *Ann. Appl. Probab.*, 13(3):984–1053, 2003. 6, 10
[6] D. Duffie, J. Pan, and K. Singleton. Transform analysis and asset pricing for affine jump-diffusions. *Econometrica*, 68(6):1343–1376, 2000. 11
[7] C. F. Dunkl. Hankel transforms associated to finite reflection groups. In *Hypergeometric functions on domains of positivity, Jack polynomials, and applications (Tampa, FL, 1991)*, volume 138 of *Contemp. Math.*, pages 123–138. Amer. Math. Soc., Providence, RI, 1992. 6
[8] J. L. Forman and M. Sørensen. The Pearson diffusions: A class of statistically tractable diffusion processes. *Economics Working Paper Series, University Of Aarhus*, 2007. 7
[9] L. Gallardo and M. Yor. Some remarkable properties of the Dunkl martingales. In *In memoriam Paul-André Meyer: Séminaire de Probabilités XXXIX*, volume 1874 of *Lecture Notes in Math.*, pages 337–356. Springer, Berlin, 2006. 6
[10] J. Gatheral. *The Volatility Surface: A Practitioner's Guide*. Wiley Finance, 2006. 2
[11] G. H. Golub and C. F. Van Loan. *Matrix computations*. Johns Hopkins Studies in the Mathematical Sciences. Johns Hopkins University Press, Baltimore, MD, third edition, 1996. 8
[12] C. Gourieroux and J. Jasiak. Multivariate Jacobi process with application to smooth transitions. *J. Econometrics*, 131(1-2):475–505, 2006. 7
[13] N. Jacob. *Pseudo differential operators and Markov processes. Vol. I*. Imperial College Press, London, 2001. Fourier analysis and semigroups. 4
[14] J. Kallsen. A didactic note on affine stochastic volatility models. In *From stochastic calculus to mathematical finance*, pages 343–368. Springer, Berlin, 2006. 4
[15] C. Moler and C. Van Loan. Nineteen dubious ways to compute the exponential of a matrix, twenty-five years later. *SIAM Rev.*, 45(1):3–49 (electronic), 2003. 8
[16] D. Revuz and M. Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1991. 4
[17] H. Zhou. Itô conditional moment generator and the estimation of short rate processes. *Journal of Financial Econometrics*, pages 250–271, 2003. 8

Vienna University of Technology, Wiedner Hauptstrasse 8–10, A-1040 Wien, Austria
E-mail address: christa.cuchiero@math.ethz.ch, martin.keller-ressel@math.ethz.ch, josef.teichmann@math.ethz.ch