

Università degli Studi di Padova

Padua Research Archive - Institutional Repository

Quantization goes polynomial

Original Citation:

Availability:

This version is available at: 11577/3351878 since: 2021-09-03T13:05:40Z

Publisher:

Taylor and Francis

Published version:

DOI: 10.1080/14697688.2020.1828608

Terms of use:

Open Access

This article is made available under terms and conditions applicable to Open Access Guidelines, as described at <http://www.unipd.it/download/file/fid/55401> (Italian only)

(Article begins on next page)

Quantization goes Polynomial*

Giorgia Callegaro^{†‡} Lucio Fiorin[§] Andrea Pallavicini[¶]

September 18, 2020

Abstract

Quantization algorithms have been successfully adopted to option pricing in finance thanks to the high convergence rate of the numerical approximation. In particular, very recently, recursive marginal quantization has been proven to be a flexible and versatile tool when applied to stochastic volatility processes. In this paper we apply for the first time quantization techniques to the family of polynomial processes, by exploiting their peculiar nature. We focus our analysis on the stochastic volatility Jacobi process, by presenting two alternative quantization procedures: the first is a new discretization technique, whose foundation lies on the polynomial structure of the underlying process and which is suitable for vanilla option pricing, the second is based on recursive marginal quantization and it allows for pricing of (vanilla and) exotic derivatives. We prove theoretical results to assess the induced approximation errors, and we describe in numerical examples practical tools for fast vanilla and exotic option pricing.

JEL classification codes: C63, G13.

AMS classification codes: 65H35, 91G20, 91G60.

Keywords: Pricing, Quantization, Polynomial Models, Stochastic Volatility, Path-Dependent Options.

1 Introduction

Recently a new class of Markov processes, termed polynomial processes, has been introduced to model stock prices in view of financial applications. We refer to [Cuchiero et al. \(2012\)](#) and [Filipovic and Larsson \(2016\)](#) for an introduction and a review of the main properties of this family of processes, which includes, e.g., the Brownian motion, the geometric Brownian

***Acknowledgements:** this work was supported by the University of Padova research grant BIRD172407/17, “New perspectives in stochastic methods for finance and energy markets”. Authors are grateful to Damien Akerer, Damir Filipovic, Martino Grasselli, Martin Larsson and Daniele Marazzina for fruitful discussions.

[†]Department of Mathematics “Tullio Levi Civita”, University of Padova, via Trieste 63, 35121 Padova, Italy. Email: gcallega@math.unipd.it ORCID ID: 0000-0001-9026-5261

[‡]Corresponding author.

[§]Department of Mathematics “Tullio Levi Civita”, University of Padova, via Trieste 63, 35121 Padova, Italy. Email: luciofiorin@gmail.com ORCID ID: 0000-0002-0350-9473

[¶]Department of Mathematics, Imperial College, London SW7 2AZ, UK and Banca IMI, Largo Mattioli 3, 20121 Milano, Italy. Email: a.pallavicini@imperial.ac.uk.

motion, Ornstein-Uhlenbeck processes, Jacobi processes, Lévy processes and, more generally, affine processes. The main property of polynomial processes is that conditional expectations of polynomial functions of the process are again of polynomial type. Hereafter, we refer to this property as the “polynomial property”. In particular, expected values of any polynomial of the process is again a polynomial in the initial value of the process, so that moments of all orders can be easily computed in closed form, up to a matrix exponential, even if the characteristic function of the process may be not known.

In this paper we focus on a particular polynomial process, the Stochastic Volatility Jacobi process (hereafter SVJ) first introduced in [Ackerer et al. \(2018\)](#), but our results can be extended to any polynomial model. The SVJ model is a diffusion model for stock prices where the log-price squared volatility follows a Jacobi process with values in some compact interval. It includes as limiting cases the Black-Scholes model and the Heston model, so that it can be viewed as a possible alternative to these models in practical applications of option pricing.

The polynomial property allows to implement calibration algorithms for plain-vanilla options quoted by the market, which usually depend only on the marginal probability distribution of the underlying asset price at option expiry date. On the other hand, pricing exotic options also requires transition probabilities because of the path-dependent feature of such derivative products (think of time averages, continuous barriers and early redemptions). Joint probability distributions of the process at different times can be in principle derived by exploiting the polynomial property of the model, but the computational time rapidly explodes as the number of time observations increases. The work [Filipovic et al. \(2020\)](#) suggests one possible way to deal with this issue by introducing an approximating Markov chain via a moment matching condition. The space grid is made of the same set of points for all the discretization times. Since in general it is impossible to construct a non-trivial Markov chain satisfying an exact n -th moment matching, with $n \geq 4$, the authors overcome this severe restriction by relaxing the exact moment matching and by looking for an approximation of the moments.

Working with a discrete state space clearly reduces the dimensionality of the problem and this is also what happens here with our alternative methodology, based on quantization. It is crucial to understand that our approach is different. More precisely, while the final target is the same, namely discretizing a stochastic process, the way this is done is totally different: quantization allows to optimally (in an L^2 -sense) approximate the stochastic process at any time, so the output is an optimal grid which is not homogeneous with respect to time. We deem this is crucial when pricing path-dependent options over possibly long time horizons.

Quantization is a discretization technique for random variables, called vector quantization, and stochastic processes, known as functional quantization. The birth of quantization dates back to the 1950s, while its application in numerical probability and mathematical finance started in the 1990s (see e.g. [Gersho and Gray \(1991\)](#) and [Bally et al. \(2001\)](#)) and since then it has been applied to approximate conditional expectations, to solve optimal stopping problems, to nonlinear filtering, stochastic control, to numerically solve backward stochastic differential equations and, more recently, McKean-Vlasov equations. For a recent and rigorous introduction to quantization we also refer e.g. to ([Montes, 2020](#), Chapter 1) and references therein. When applied to random vectors, quantization provides the best, according to a distance that is commonly measured using the Euclidean norm, possible approximation to the original distribution via a discrete random vector taking a finite number of values. Many numerical procedures have been studied to obtain optimal quadratic quantizers of random

vectors even in high dimension and most of them are based on stochastic optimization algorithms, see Pagès (2015), which are typically very time consuming. Very recently vector quantization has been applied to recursively discretize stochastic processes. More precisely, the idea is approximating the random variables coming from the Euler scheme associated to a stochastic process, which is solution to a stochastic differential equation. Recursive marginal quantization, or fast quantization, introduced in Pagès and Sagna (2015), represents a new promising research field. Stationary quantizers of the stochastic process at fixed dates are obtained in a very fast and recursive way, to the point that recursive marginal quantization has been successfully applied to many models, starting from the classical ones and moving to local volatility models, such as in Callegaro et al. (2015), Bormetti et al. (2018), McWalter et al. (2018) and finally to stochastic volatility models, such as Stein and Stein, Stochastic Alpha Beta Rho, Heston and α -hypergeometric, as done in Callegaro et al. (2016), Fiorin et al. (2018).

In the present paper we adopt two different quantization techniques to price plain-vanilla options and exotic derivative contracts within the SVJ model. The first technique is designed for plain-vanilla or European options, and it is based on the direct quantization of the price probability distribution obtained by exploiting the polynomial property. In this way we provide an alternative approach to the algorithm proposed in Ackerer et al. (2018). The second one is designed for exotic products, hence providing an alternative to Filipovic et al. (2020), but it can be used also for European options, and it extends to multidimensional models the framework of Callegaro et al. (2016), based on recursive marginal quantization. This second procedure does not rely on the polynomial property of the model, so that we can use it for path-dependent products without encountering dimensionality problems. Our analysis on SVJ quantization provides both practical tools to develop fast exotic option pricing algorithms and theoretical results to assess the approximation errors. A final Conclusion follows.

The paper is organized as follows. In Section 2 the SVJ model is presented. An introduction to quantization of random variables is given in Section 3. Two different quantization approaches are then described: firstly, in Section 4 quantization techniques are adapted to polynomial models, leading to new pricing formulas for plain-vanilla options, whose approximation error is discussed. Then, in Section 5 recursive marginal quantization (which does not exploit the polynomial nature of our stochastic process) is introduced in a multidimensional setting and it is applied to price path-dependent exotic options. Numerical results for all the introduces algorithms, along with a discussion, are presented in Section 6.

2 The Stochastic Volatility Jacobi Model

We consider a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in [0, T]}, \mathbb{Q})$, where \mathbb{Q} is a risk neutral probability measure and where the filtration $(\mathcal{F}_t)_{t \in [0, T]}$ satisfies the usual hypotheses and models all the randomness in our model. We assume that the stock price process S follows a SVJ model as in Ackerer et al. (2018), namely we fix $0 \leq v_{min} < v_{max}$ and we define

$$S_t = e^{X_t} \tag{1}$$

where the dynamics of (V, X) follows the stochastic volatility model

$$\begin{cases} dV_t &= \kappa(\theta - V_t)dt + \sigma\sqrt{Q(V_t)}dW_t \\ dX_t &= (r - \delta - V_t/2)dt + \rho\sqrt{Q(V_t)}dW_t + \sqrt{V_t - \rho^2Q(V_t)}dW_t^\perp \end{cases} \tag{2}$$

with $X_0 = x_0 \in \mathbb{R}$, $V_0 = v_0 \in [v_{min}, v_{max}]$ and where the interest rate $r > 0$, $\rho \in [-1, 1]$, the mean reversion speed is $\kappa \geq 0$, the reversion level θ belongs to $[v_{min}, v_{max}]$,

$$Q(v) := \frac{(v - v_{min})(v_{max} - v)}{(\sqrt{v_{max}} - \sqrt{v_{min}})^2} \quad (3)$$

and where W and W^\perp are independent standard Brownian motions. Notice that, recalling Equation (3), we have $v - Q(v) \geq 0$ with equality if and only if $v = \sqrt{v_{min}v_{max}}$ and $Q(v) \geq 0$ for every $v \in [v_{min}, v_{max}]$, so that the square roots in Equation (2) do not create any issue. Clearly, we may assume that $\mathcal{F}_t = \mathcal{F}_t^W \vee \mathcal{F}_t^{W^\perp}$, $t \in [0, T]$, where \mathcal{F}_t^W and $\mathcal{F}_t^{W^\perp}$ are the natural filtrations associated with W and W^\perp completed with negligible sets. It is known that as special limiting cases of a SVJ model we obtain the Black-Scholes (take $v_0 = \theta = v_{max}$) and the Heston model (take $v_{min} = 0$ and $v_{max} \rightarrow \infty$).

Remark 2.1. (Existence and Uniqueness of SVJ SDE Solution)

The name SVJ is motivated by the model being clearly a stochastic volatility one, with the instantaneous squared volatility V having a dynamics of Jacobi type, taking values in the interval $[v_{min}, v_{max}]$. Indeed, the following result holds (see (Ackerer et al., 2018, Theorem 2.1)): for any deterministic initial state $(v_0, x_0) \in [v_{min}, v_{max}] \times \mathbb{R}$, there exists a unique solution (V, X) to the system (2), taking values in $[v_{min}, v_{max}] \times \mathbb{R}$. Furthermore, it is possible to show that if $(v_0, x_0) \in (v_{min}, v_{max}) \times \mathbb{R}$, then (V_t, X_t) takes values in $(v_{min}, v_{max}) \times \mathbb{R}$ if and only if $\frac{\sigma^2(v_{max} - v_{min})}{(\sqrt{v_{max}} - \sqrt{v_{min}})^2} \leq 2\kappa \min\{v_{max} - \theta, \theta - v_{min}\}$.

Moments in the SVJ model are known in closed form up to a matrix exponential. Indeed, if we write the generator \mathcal{G} of the SVJ process, namely

$$\mathcal{G}f(v, x) = b^\top(v) \nabla f(v, x) + \frac{1}{2} \text{Tr}(a(v) \nabla^2 f(v, x))$$

with drift vector $b(v)$ and the diffusion matrix $a(v)$ given by

$$b(v) = \begin{bmatrix} \kappa(\theta - v) \\ r - \delta - v/2 \end{bmatrix}, \quad a(v) = \begin{bmatrix} \sigma^2 Q(v) & \rho \sigma Q(v) \\ \rho \sigma Q(v) & v \end{bmatrix}$$

we have that \mathcal{G} maps any polynomial of degree n onto a polynomial of degree n or less as shown in Filipovic and Larsson (2016). As a consequence it is possible to evaluate the conditional moments of (V_T, X_T) as follows. Let Pol_n be the vector space of polynomials in (v, x) of degree less than or equal to n . For any positive integer N , we term $M = (N + 2)(N + 1)/2$ the dimension of Pol_N , we introduce a basis $h_1(v, x), \dots, h_M(v, x)$ of polynomials of Pol_N , and we denote by G the matrix representation of the linear map \mathcal{G} restricted to Pol_N with respect to this basis. Thus, from Theorem 3.1 in Filipovic and Larsson (2016) we get that for any polynomial $p \in \text{Pol}_N$ we have

$$\mathbb{E}[p(V_T, X_T) | \mathcal{F}_t] = [h_1(V_t, X_t) \quad \dots \quad h_M(V_t, X_t)] e^{(T-t)G} \vec{p}, \quad t \in [0, T],$$

where $\vec{p} \in \mathbb{R}^M$ is the column coordinate representation of the polynomial $p(v, x)$ with respect to the basis h . In this paper we term this relationship as *polynomial property*.

We recall here the technical results on the closed form pricing of European options, that we will need from now on. Let us define the weighted Lebesgue space

$$L_w^2 = \left\{ f \text{ measurable} : \|f\|_w^2 = \int_{\mathbb{R}} f^2(x) w(x) dx < \infty \right\},$$

equipped with the scalar product

$$\langle f, g \rangle_w = \int_{\mathbb{R}} f(x)g(x)w(x)dx,$$

where w is the Gaussian weight function, i.e., the Gaussian density with mean μ_w and variance σ_w^2 . The Hilbert space L_w^2 admits an orthonormal basis of generalized Hermite polynomials H_n , $n \geq 0$ (notice that the degree of H_n is equal to n), given by

$$H_n(x) = \frac{1}{\sqrt{n!}} \mathcal{H}_n \left(\frac{x - \mu_w}{\sigma_w} \right), \quad (4)$$

where \mathcal{H}_n are the probabilist Hermite polynomials defined as

$$\mathcal{H}_n(x) = (-1)^n e^{\frac{x^2}{2}} \frac{d^n}{dx^n} e^{-\frac{x^2}{2}}. \quad (5)$$

If we assume that g_T is the density of the log price X_T , then we can define $\ell(x) = \frac{g_T(x)}{w(x)}$.

Remark 2.2. (Assumptions on SVJ Parameters)

a) *The existence and regularity of the density g_T for the log price process X_T is studied in (Ackerer et al., 2018, Theorem 3.1), to which we refer. In particular, as soon as $v_{min} > 0$ and $\rho^2 < 1$ (see also b) below), the density g_T is k times continuously differentiable for any $k \in \mathbb{N}_{>0}$.*

b) *We need $\ell \in L_w^2$, so that from now on, applying Corollary 3.2 in Ackerer et al. (2018), we assume that*

$$v_{min} > 0, \quad \rho^2 < 1, \quad \sigma_w^2 > \frac{v_{max}T}{2}.$$

c) *The choice of generalized Hermite polynomials as orthonormal basis is quite natural, given the Gaussian nature of the log price X_T conditional on the knowledge of the volatility. Density series approximation is a broad (and not new) topic and in the case of polynomial processes we refer to (Ackerer et al., 2018, Introduction) and references therein to convince the reader about the robustness of the use of this Gram-Charlier A decomposition for the density g_T .*

Our aim being pricing a European option with payoff $f \in L_w^2$ (notice that both Call and Put options have payoffs belonging to L_w^2), we get

$$\mathbb{E}[f(X_T)] = \int_{\mathbb{R}} f(x)g_T(x)dx = \int_{\mathbb{R}} f(x)\ell(x)w(x)dx = \langle f, \ell \rangle_w.$$

Since L_w^2 is an Hilbert space with orthonormal basis H_n defined in Equation (4), we can rewrite the previous formula as

$$\mathbb{E}[f(X_T)] = \sum_{n \geq 0} f_n \ell_n, \quad (6)$$

for the *Fourier coefficients*

$$f_n = \langle f, H_n \rangle_w, \quad (7)$$

and the *Hermite moments*

$$\ell_n = \langle \ell, H_n \rangle_w = \int_{\mathbb{R}} H_n(x)\ell(x)w(x)dx = \int_{\mathbb{R}} H_n(x)g_T(x)dx. \quad (8)$$

Notice that the last equality shows that ℓ_n is a linear combination of moments of X_T , since H_n is a polynomial. It is then possible to compute ℓ_n in closed form, because of the polynomial nature of the process.

3 Essentials on Quantization

In this section we introduce optimal quadratic quantization of a random variable (also known as vector quantization), which will be a necessary tool toward a specific discretization of a stochastic process, known as recursive marginal quantization (henceforth RMQ). We refer to [Graf and Luschgy \(2000\)](#) and [Pagès \(2015\)](#) for vector quantization and to [Pagès and Sagna \(2015\)](#) for the first paper on RMQ¹.

Optimal quadratic quantization answers the following question: is it possible (and how) to optimally approximate, in an L^2 -sense, a continuous random variable X by a discrete one, \widehat{X} , taking a finite number of values?

The interest in such a discretization \widehat{X} is clear: expectations in the form $\mathbb{E}[h(X)]$ (for sufficiently regular functions h) would be approximated by finite sums. Let us be more precise. We consider a real valued random variable X defined on $(\Omega, \mathcal{F}, \mathbb{P})$, having probability distribution \mathbb{P}_X and admitting a finite second order moment. A quantization grid of level $N, N \geq 1$, is a subset of \mathbb{R} , $\Gamma = \{x_1, \dots, x_N\}$ (here N will be fixed, so that for simplicity we drop the dependence on N in the notation), of size at most N having pairwise distinct components. A quantization function, or *quantizer*, is a Γ -valued Borel function $q : \mathbb{R} \rightarrow \Gamma$ and quantizing X means projecting X on Γ following the closest neighbor rule

$$q(X) = \text{Proj}_\Gamma(X) := \sum_{i=1}^N x_i \mathbb{1}_{C_i(\Gamma)}(X) \quad (9)$$

where $(C_i(\Gamma))_{1 \leq i \leq N}$ is a Borel partition of $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, also known as Voronoi partition, satisfying $C_i(\Gamma) \subset \{\xi \in \mathbb{R} : |\xi - x_i| = \min_{1 \leq j \leq N} |\xi - x_j|\}$, $i = 1 \dots, N$. We will use the notation \widehat{X}^Γ or \widehat{X} (when no ambiguity is possible with respect to the grid) to denote the Voronoi Γ -quantization of X : $\widehat{X}^\Gamma = \widehat{X} = q(X)$.

The L^2 -error coming from such a discretization is given by

$$e_N(X, \Gamma) := \|X - q(X)\|_2 = \left\| \min_{1 \leq i \leq N} |X - x_i| \right\|_2$$

where $\|X\|_2 := \left[\mathbb{E}(|X|^2) \right]^{1/2}$ is the usual L^2 -norm and the aim of optimal quadratic quantization is finding a grid Γ , with size at most N , which minimizes the *distortion function* defined below (see ([Graf and Luschgy, 2000](#), Equation (3.4))).

Definition 3.1. *Let X be a real valued random variable belonging to $L^2(\mathbb{P})$. The L^2 -distortion function is a positive valued function defined on \mathbb{R}^N by*

$$D : (x_1, x_2, \dots, x_N) \mapsto \mathbb{E} \left[\min_{1 \leq i \leq N} |X - x_i|^2 \right] = e_N(X, \Gamma)^2. \quad (10)$$

Concerning the existence and uniqueness of an optimal grid, it is possible to show, see e.g. ([Pagès, 2015](#), Prop. 1.1), that if $X \in L^2(\mathbb{P})$, then the distortion function D attains (at least) one minimum Γ^* . The grid Γ^* and Proj_{Γ^*} are called *optimal quadratic quantizers*, respectively. In the case when $\text{card}(\text{supp}(\mathbb{P}_X)) \geq N$, then Γ^* has pairwise distinct components. Moreover

¹We also have to refer to the website: <http://www.quantize.maths-fi.com>, where grids of the d -dimensional Gaussian distributions $\mathcal{N}(0; I_d)$, for $N = 1$ up to 10^4 and for $d = 1, \dots, 10$ can be downloaded.

$\lim_{N \rightarrow +\infty} e_N(X) = 0$ and the convergence rate is given by the well-known Zador theorem (see [Graf and Luschgy \(2000\)](#))

$$\min_{\Gamma, |\Gamma|=N} \|X - q(X)\|_2 = \min_{\Gamma, |\Gamma|=N} e_N(X, \Gamma) = Q_2(\mathbb{P}_X)N^{-1} + o(N^{-1}) \quad (11)$$

where $Q_2(\mathbb{P}_X)$ is a nonnegative constant. Then, it is natural to approximate an expected value of the form $\mathbb{E}[h(X)]$ in the following way:

$$\mathbb{E}[h(X)] \approx \mathbb{E}[h(\widehat{X})] = \sum_{i=1}^N h(x_i) \mathbb{P}(\widehat{X} = x_i) = \sum_{i=1}^N h(x_i) \mathbb{P}(X \in C_i(\Gamma)). \quad (12)$$

Moreover, as mentioned in ([Graf and Luschgy, 2000](#), Theorem. 5.1) and originally proved in [Kieffer \(1982\)](#), as soon as \mathbb{P}_X is absolutely continuous with a log-concave density, then there exists exactly one optimal quantization grid of level N .

Remark 3.2. (Quantization vs. Monte Carlo Pricing Error)

When computing $\mathbb{E}[h(X)]$, where X is the value at maturity of an underlying asset and h is a Lipschitz (payoff) function with Lipschitz constant $[h]_{Lip}$, then

$$|\mathbb{E}[h(X)] - \mathbb{E}[h(\widehat{X}^\Gamma)]| \leq [h]_{Lip} e_N(X, \Gamma).$$

In particular, the error coming from pricing via quantization decays at the rate $\frac{1}{N}$, as opposed to the Monte Carlo error, ruled by the Central Limit Theorem, which is of order $\frac{1}{\sqrt{N}}$.

The last crucial point is how to obtain an optimal quantizer. It is known that the distortion function is differentiable at any N -tuple having pairwise distinct components $\Gamma = \{x_1, \dots, x_N\}$ (see ([Graf and Luschgy, 2000](#), Lemma 4.10) or ([Pagès, 2015](#), Prop. 1.1)), with differential

$$\nabla D(x_1, \dots, x_N) = 2 \left(\int_{C_i(\Gamma)} (x_i - \xi) d\mathbb{P}_X(\xi) \right)_{1 \leq i \leq N} = 2(\mathbb{E}[\mathbb{1}_{X \in C_i(\Gamma)}(x_i - X)])_{1 \leq i \leq N}, \quad (13)$$

so that many stochastic algorithms looking for zeros of the gradient of D have been developed. These include gradient descent and fixed point procedures and we refer to ([Pagès, 2015](#), Section 3) for a detailed overview. Critical points of the distortion function are called *stationary quantizers*. Optimal quantizers are stationary, but the converse is not true in general. Of course, stationary quantizers are in general not unique. From (13), the determination of stationary quantizers boils down to finding the solution of

$$\mathbb{E}[\mathbb{1}_{X \in C_i(\Gamma)} X] - x_i \mathbb{P}(X \in C_i(\Gamma)) = 0 \quad \forall i \in \{1, \dots, N\} \quad (14)$$

known as *Master Equation* (it is, indeed, a system of N equations in N unknowns x_1, \dots, x_N). Moreover, when the gradient itself is differentiable, it is possible to apply the classical Newton-Raphson procedure.

If the density of X is known, then it is possible to explicitly write the system in (14) in closed form and to find its solution. On the other hand, when X is the asset price at maturity, typically the density of the process is not explicitly known, except in trivial cases, and finding stationary quantizers becomes numerically interesting.

We conclude by sketching the ideas behind RMQ, which allows to use vector quantization to recursively discretize a stochastic process. More precisely, [Pagès and Sagna \(2015\)](#) introduced RMQ to discretize a stochastic process Y (in dimension one) by recursively working on the random variables $\{(Y_{t_k})\}_{k=0,\dots,M}$ associated with the Euler-Maruyama scheme. So, starting from a time discretization of a stochastic process on $\{t_k\}_{k=0,\dots,M}$, vector quantization provides a state space discretization of every random variable Y_{t_k} . The essence of RMQ lies on the knowledge of the conditional law of $(Y_{t_k}|Y_{t_{k-1}})$, $k = 1, \dots, M$, which allows to recursively quantize the marginals $\{(Y_{t_k})\}_{k=0,\dots,M}$ via a Newton-Raphson procedure (the gradient and the Hessian of the distortion function are explicit).

In the following Section 5 we provide more details on RMQ applied to our multidimensional setting.

4 Quantization of a Polynomial Process

We will now discuss how to deal with polynomial processes. In particular, we consider the SVJ model, even if our approach is general and flexible enough to be applied to any polynomial process. We only consider the SVJ model since it is a good representative for the class of diffusive polynomial processes. As a first approach, in this section, we exploit the polynomial property and we focus on the quantization of the log price process X at a fixed date. Then, in Section 5, in order to deal with path-dependent options, we forget about the polynomial nature of (V, X) and we extend the general framework in [Callegaro et al. \(2016\)](#) to discretize the bidimensional process (V, X) at a whole set of dates via RMQ.

4.1 Exploiting the Polynomial Property

In this section we consider the problem of finding a stationary quantizer of the log price process X at a given time T .

The main result of this section is the possibility of writing the Master equation (14) in closed form, thanks to the polynomial nature of our processes.

Theorem 4.1. (Polynomial Process Quantization)

Consider the Master equation

$$\mathbb{E}[(X_T - x_i) \mathbb{1}_{X_T \in C_i(\Gamma)}] = 0, \quad i = 1, \dots, N \quad (15)$$

and its i -th component

$$E_i(x_1, \dots, x_N) := \mathbb{E}[(X_T - x_i) \mathbb{1}_{X_T \in C_i(\Gamma)}] = 0 \quad (16)$$

where $C_i(\Gamma) = \left[\frac{x_{i-1} + x_i}{2}, \frac{x_i + x_{i+1}}{2} \right]$, $C_1(\Gamma) = [-\infty, \frac{x_1 + x_2}{2}]$ and $C_N(\Gamma) = \left[\frac{x_{N-1} + x_N}{2}, +\infty \right]$.

In our setting Equation (16) reads

$$\sum_{n \geq 0} f_n^i \ell_n = 0, \quad (17)$$

where ℓ_n are the Hermite moments defined in (8) and where the (Fourier) coefficients f_n^i are given by

$$f_n^i = h_n \left(\frac{x_{i-1} + x_i}{2} \right) - h_n \left(\frac{x_i + x_{i+1}}{2} \right) - x_i \left(l_n \left(\frac{x_{i-1} + x_i}{2} \right) - l_n \left(\frac{x_i + x_{i+1}}{2} \right) \right)$$

with

$$\begin{aligned}
h_0(K) &= \sigma_w \phi\left(\frac{K - \mu_w}{\sigma_w}\right) + \mu_w \Phi\left(\frac{\mu_w - K}{\sigma_w}\right) \\
h_1(K) &= \sigma_w \left[\frac{K - \mu_w}{\sigma_w} \phi\left(\frac{K - \mu_w}{\sigma_w}\right) + \Phi\left(\frac{K - \mu_w}{\sigma_w}\right) \right] + \mu_w \phi\left(\frac{K - \mu_w}{\sigma_w}\right) \\
h_n(K) &= \frac{1}{\sqrt{n!}} \phi\left(\frac{K - \mu_w}{\sigma_w}\right) \left[\sigma_w \mathcal{H}_n\left(\frac{K - \mu_w}{\sigma_w}\right) + n \sigma_w \mathcal{H}_{n-2}\left(\frac{K - \mu_w}{\sigma_w}\right) \right. \\
&\quad \left. + \mu_w \mathcal{H}_{n-1}\left(\frac{K - \mu_w}{\sigma_w}\right) \right], n \geq 2,
\end{aligned} \tag{18}$$

and

$$\begin{aligned}
l_0(K) &= \Phi\left(\frac{K - \mu_w}{\sigma_w}\right) \\
l_n(K) &= \frac{1}{\sqrt{n!}} \mathcal{H}_{n-1}\left(\frac{K - \mu_w}{\sigma_w}\right) \phi\left(\frac{K - \mu_w}{\sigma_w}\right), n \geq 1,
\end{aligned} \tag{19}$$

and where ϕ and Φ are, respectively, the density and the cumulative distribution functions of a standard univariate Gaussian random variable.

Proof. See Appendix A. \square

Remark 4.2. The results of Theorem 4.1 are general, in that they can be applied to other polynomial models, by suitably modifying the coefficients ℓ_n .

4.2 Calculation of a Stationary Quantizer

Even if Equation (16) can be written in closed form for every $i = 1, \dots, N$, it is impossible to find an analytical expression for the solution to the nonlinear system, which corresponds to the stationary quantizer. Hence, we need to solve this system numerically. As already noted in Section 3, the literature suggests the Newton-Raphson method as the best first choice to tackle the system of equations. The proposition below provides the Jacobian matrix to be used in the Newton-Raphson procedure.

Proposition 4.3. Consider the system of equations (15)

$$\begin{cases} E_1(x_1, \dots, x_N) := \mathbb{E}[(X_T - x_1) \mathbb{1}_{X_T \in C_1(\Gamma)}] = 0 \\ E_2(x_1, \dots, x_N) := \mathbb{E}[(X_T - x_2) \mathbb{1}_{X_T \in C_2(\Gamma)}] = 0 \\ \vdots \\ E_N(x_1, \dots, x_N) := \mathbb{E}[(X_T - x_N) \mathbb{1}_{X_T \in C_N(\Gamma)}] = 0. \end{cases} \tag{20}$$

When X is a polynomial process with Hermite moments ℓ_n , the Jacobian matrix J of the vector function $E = (E_1, \dots, E_N)$ is tridiagonal and symmetric, and its components have the following form:

$$\begin{aligned}
J_{i,i-1} &= \frac{1}{2} \left(\frac{x_i - x_{i-1}}{2} \right) g_T \left(\frac{x_{i-1} + x_i}{2} \right) \quad i = 2, \dots, N \\
J_{i,i} &= J_{i,i-1} + J_{i,i+1} - \mathbb{P}(X_T \in C_i(\Gamma)) \quad i = 1, \dots, N,
\end{aligned}$$

with $J_{1,0} = J_{N,N+1} = 0$ and where

$$g_T(x) = \sum_{n \geq 0} \ell_n H_n(x) w(x), \quad (21)$$

$$\mathbb{P}(X_T \in C_i(\Gamma)) = \sum_{n \geq 0} \ell_n \left(l_n \left(\frac{x_{i-1} + x_i}{2} \right) - l_n \left(\frac{x_i + x_{i+1}}{2} \right) \right), \quad (22)$$

and the coefficients l_n are computed in (19).

Proof. See Appendix A. □

The Newton-Raphson algorithm has then the following structure: starting from an initial guess $\Gamma^{(0)}$, the k -th iteration is

$$\Gamma^{(k)} = \Gamma^{(k-1)} - J^{-1} \left(\Gamma^{(k-1)} \right) E \left(\Gamma^{(k-1)} \right) \quad k = 1, 2, \dots \quad (23)$$

where $E = (E_1, \dots, E_N)$ is defined in Proposition 4.3 and it is computed thanks to Theorem 4.1. Given a stopping criterion, the final iteration gives the stationary quantization grid Γ^* .

Let us assume now that we have found the solution $\Gamma^* = \{x_1^*, \dots, x_N^*\}$, which is the stationary quantization grid associated to X_T . In order to compute an expected value as in (12), we need to know the weights associated to every Voronoi cell $C_i(\Gamma^*)$, for $i = 1, \dots, N$. The weights are straightforwardly given by (22).

4.3 Analysis of the Approximation Error

We focus on pricing at time 0 of a European option with payoff f . We consider, without loss of generality, a Call option written on S having expiry $T > 0$ and strike price K , i.e.,

$$f = f(S_T) = (S_T - K)^+.$$

Of course, as mentioned before, the results in this section are also valid for Put options. In what follows we will need the following three versions of the price:

- π_f is the exact price at time 0, i.e.,

$$\pi_f := \mathbb{E}^{\mathbb{Q}} [e^{-rT} (S_T - K)^+] = e^{-rT} \int_{\mathbb{R}} (e^x - K)^+ g_T(x) dx,$$

where g_T is the density of the log price X at time T , given by (21). This formula contains an infinite sum, so the function g_T function is not computable in closed form.

- $\pi_f^{(M)}$ is the price computed using the polynomial approximation at level M , i.e., approximating the density $g_T(x)$ with

$$g_T^{(M)}(x) = \sum_{n=0}^M \ell_n H_n(x) w(x), \quad (24)$$

namely

$$\pi_f^{(M)} = e^{-rT} \int_{\mathbb{R}} (e^x - K)^+ g_T^{(M)}(x) dx = \sum_{n=0}^M \ell_n f_n,$$

where here the f_n 's are the Fourier coefficients associated with the function $f(x) = (e^x - K)^+$.

- $\widehat{\pi}_f^{(M,N)}$ is the price computed by approximating the log-spot price at maturity by means of quantization on a grid with N points:

$$\begin{aligned}\widehat{\pi}_f^{(M,N)} &= e^{-rT} \sum_{i=1}^N \int_{C_i(\Gamma^X)} (e^{x_i} - K)^+ g_T^{(M)}(x) dx \\ &= e^{-rT} \sum_{i=1}^N (e^{x_i} - K)^+ \mathbb{P} \left(X_T^{(M)} \in C_i(\Gamma^X) \right)\end{aligned}\quad (25)$$

where $\Gamma^X = \{x_1, \dots, x_N\}$ is the optimal quantizer relative to $X_T^{(M)}$, the log price with (approximate) density $g_T^{(M)}$. We denote by $\widehat{X}_T^{(M,N)}$ the quantization of $X_T^{(M)}$. Notice that we also have

$$\widehat{\pi}_f^{(M,N)} = e^{-rT} \sum_{i=1}^N (e^{x_i} - K)^+ \mathbb{P} \left(\widehat{X}_T^{(M,N)} = x_i \right).$$

The accuracy of our methodology is studied by analyzing the (asymptotic) behavior of the price approximation, namely

$$\text{err}_{M,N} := \left| \pi_f - \widehat{\pi}_f^{(M,N)} \right|. \quad (26)$$

We split the error in two parts that we study separately:

$$\text{err}_{M,N} \leq \left| \pi_f - \pi_f^{(M)} \right| + \left| \pi_f^{(M)} - \widehat{\pi}_f^{(M,N)} \right| = \text{err}_M^{(1)} + \text{err}_{M,N}^{(2)}, \quad (27)$$

where $\text{err}_M^{(1)} := \left| \pi_f - \pi_f^{(M)} \right|$ is the truncation error $|\varepsilon^{(M)}|$ (depending only on M), studied in detail in (Ackerer et al., 2018, Section 3.2) and satisfying inequality (3.20) therein:

$$|\varepsilon^{(M)}| \leq \left(\|f\|_w^2 - \sum_{n=0}^M f_n^2 \right)^{\frac{1}{2}} \left(\|\ell\|_w^2 - \sum_{n=0}^M \ell_n^2 \right)^{\frac{1}{2}},$$

whose terms, as the authors state, can be estimated by Monte Carlo simulation, while $\text{err}_{M,N}^{(2)} := \left| \pi_f^{(M)} - \widehat{\pi}_f^{(M,N)} \right|$ is the quantization error, on which we focus in the remaining part of this section.

We study now $\text{err}_{M,N}^{(2)}$ via an intermediate lemma and a theorem.

Lemma 4.4. *The quantization error satisfies*

$$\text{err}_{M,N}^{(2)} \leq \left\| S_T^{(M)} - \widehat{S}_T^{(M,N)} \right\|_2,$$

where $S_T^{(M)} := e^{X_T^{(M)}}$ and $\widehat{S}_T^{(M,N)}$ is the N -quantizer relative to $S_T^{(M)}$.

Proof. See Appendix A. □

Now we are ready to size the quantization error. As known from Zador Theorem, see Equation (11), the distance $\|S_T^{(M)} - \widehat{S}_T^{(M,N)}\|_2$ has an asymptotic linear decay when N goes to infinity. A precise bound for the limit of the error is provided in the following theorem, whose proof is inspired by the one of (Callegaro et al., 2019, Theorem 2.11). We stress here that the recent error estimates given in (Callegaro et al., 2019, Theorem 2.11) were obtained in a different setting and under *ad hoc* assumptions (the second order moment of $S_T^{(M)}$ was required to be finite and the density of $S_T^{(M)}$ at 0 and at $+\infty$ had polynomial behavior), that can be relaxed here, thanks to the explicit form of the density $h_T^{(M)}$ of $S_T^{(M)}$. A thorough study of the error is also present in Fiorin and Schoutens (2020), in the field of conic finance.

Theorem 4.5. (Quantization Error Estimate)

In our setting, for any given $M > 0$, we have:

$$\lim_{N \rightarrow +\infty} N \text{err}_{M,N}^{(2)} \leq \frac{\left\| h_T^{(M)} \right\|_{\frac{1}{3}}^{\frac{1}{2}}}{2\sqrt{3}} \quad (28)$$

where $h_T^{(M)}$ is the density of $S_T^{(M)} = e^{X_T^{(M)}}$, $h_T^{(M)}(s) = \frac{g_T^{(M)}(\ln(s))}{s}$ for $s \in (0, +\infty)$.

Proof. See Appendix A. □

5 Pricing exotics: multidimensional RMQ

The quantization procedure described in the previous section, and based on the polynomial nature of the (V, X) , may encounter dimensionality problems when pricing exotic options. In this section we propose an alternative quantization procedure, known as recursive marginal quantization, which in the literature has been discussed for a wide class of stochastic volatility models (see e.g. Callegaro et al. (2016) and Callegaro et al. (2018)), and we extend it to a multidimensional setting.

The innovation behind this approach is twofold. First, the algorithm here is designed for systems of SDEs of any dimension d , while the approach used in Callegaro et al. (2016) and Callegaro et al. (2018) is tailored for systems of dimension 2. **The presentation of this section can be seen as a compact and robust alternative to Fiorin et al. (2018). In fact, the algorithm developed here does not depend on pre-computed multidimensional quantization grids, that Fiorin et al. (2018) heavily exploit.** The second innovation is in the computational cost required by the execution of the algorithm, which is reduced by a factor 5 with respect to Callegaro et al. (2016), see Section 6. Indeed, until now the volatility process was quantized independently, while the price process was then discretized using the information from the approximation of the variance. In the formulation that we propose here, the vector volatility-price process is quantized simultaneously in all its components. This allows, in the end, to more compact formulas and to a more efficient and parsimonious numerical routine. Here, we consider the quantization of a system of SDEs. We will present a general framework, that we will then apply to the case of the SVJ model. Let us consider the following d -dimensional SDE:

$$d\mathbf{X}_t = \mu(t, \mathbf{X}_t) dt + \Sigma(t, \mathbf{X}_t) d\mathbf{W}_t, \quad \mathbf{X}_0 = \mathbf{x}_0. \quad (29)$$

where $\mu : \mathbb{R}_+ \times \mathbb{R}^d \rightarrow \mathbb{R}^d$, $\Sigma : \mathbb{R}_+ \times \mathbb{R}^d \rightarrow \mathbb{R}^d \times \mathbb{R}^q$ and \mathbf{W} is a q -dimensional Brownian motion. We suppose that μ and Σ are sufficiently regular so that to ensure existence and uniqueness of a solution to the SDE (29).

Let us now fix a time discretization grid $t_k = k\Delta, k = 0, \dots, L$, with $\Delta = \frac{T}{L}$, where T is a given maturity, Δ is the time step size and L is the number of discretization points of the time grid. A general discretization scheme can be written in the following iterative form:

$$\tilde{\mathbf{X}}_{t_{k+1}} = A(t_k, \Delta, \tilde{\mathbf{X}}_{t_k}) + B(t_k, \Delta, \tilde{\mathbf{X}}_{t_k}, \Delta \mathbf{W}_{t_k}), \quad \tilde{\mathbf{X}}_{t_0} = x_0, \quad (30)$$

where $A : \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ and $B : \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}^d \times \mathbb{R}^q \rightarrow \mathbb{R}^d$ depend on the discretization scheme and $\Delta \mathbf{W}_{t_k} = \frac{1}{\sqrt{\Delta}} (\mathbf{W}_{t_{k+1}} - \mathbf{W}_{t_k})$ is a q -dimensional Gaussian vector with mean $\mathbf{0}$ and variance-covariance matrix I_q .

Depending on the time discretization scheme in use, it is possible to know the law of $(\tilde{\mathbf{X}}_{t_{k+1}} | \tilde{\mathbf{X}}_{t_k})$, that clearly depends on A and B . In particular, in the case of the Euler-Maruyama (or simply Euler) scheme, that we will choose, $(\tilde{\mathbf{X}}_{t_{k+1}} | \{\tilde{\mathbf{X}}_{t_k} = \mathbf{x}\})$, $\mathbf{x} \in \mathbb{R}^d$, has a multivariate Gaussian distribution, while in the case of the Milstein scheme it has a generalized Chi-squared distribution. For higher order schemes, the conditional distribution has to be determined on a case by case basis.

5.1 Mathematical Foundation of the Algorithm

Henceforth we consider the Euler scheme, so that, conditioning on $\{\tilde{\mathbf{X}}_{t_k} = \mathbf{x}\}$, we have (recall Equation (30))

$$A(t_k, \Delta, \mathbf{x}) = \mathbf{x} + \mu(t_k, \mathbf{x})\Delta, \quad B(t_k, \Delta, \mathbf{x}, \Delta \mathbf{W}_{t_k}) = \sqrt{\Delta} \Sigma(t_k, \mathbf{x}) \Delta \mathbf{W}_{t_k}, \quad (31)$$

and the following lemma holds:

Lemma 5.1. *For every $0 \leq k \leq L$, conditionally on the event $\{\tilde{\mathbf{X}}_{t_k} = \mathbf{x}\}$, the random vector $\tilde{\mathbf{X}}_{t_{k+1}}$ is Gaussian:*

$$\mathcal{L}(\tilde{\mathbf{X}}_{t_{k+1}} | \{\tilde{\mathbf{X}}_{t_k} = \mathbf{x}\}) \sim \mathcal{N}(\mathbf{x} + \mu(t_k, \mathbf{x})\Delta, \Delta(\Sigma \Sigma^T)(t_k, \mathbf{x})). \quad (32)$$

In particular, if $\mathbf{X}_t = (X_t^1, \dots, X_t^d)$ and $\tilde{\mathbf{X}}_{t_k} = (\tilde{X}_{t_k}^1, \dots, \tilde{X}_{t_k}^d)$ for $k = 0, \dots, L$, and $\mathbf{x} = (x^1, \dots, x^d)$, we have that for every $i = 1, \dots, d$

$$\mathcal{L}(\tilde{X}_{t_{k+1}}^i | \{\tilde{\mathbf{X}}_{t_k} = \mathbf{x}\}) \sim \mathcal{N}(m_i(t_k, \mathbf{x}), \varsigma_i(t_k, \mathbf{x})), \quad (33)$$

where

$$m_i(t_k, \mathbf{x}) := x^i + \mu_i(t_k, \mathbf{x})\Delta$$

is the i -th component of the vector $\mathbf{x} + \mu(t_k, \mathbf{x})$ and $\varsigma_i(t_k, \mathbf{x})$ is the i -th diagonal element of the (symmetric) matrix $\Delta \Sigma \Sigma^T$.

It is then possible to write the distribution of $\tilde{X}_{t_{k+1}}^i$ in a closed form:

$$\mathbb{P}(\tilde{X}_{t_{k+1}}^i \in dx_{k+1}^i) = \int_{\mathbb{R}^d} \phi_{m_i(t_k, \mathbf{x}_k), \varsigma_i(t_k, \mathbf{x}_k)}(x_{k+1}^i) \mathbb{P}(\tilde{\mathbf{X}}_{t_k} \in d\mathbf{x}_k), \quad (34)$$

where $\phi_{m,\varsigma}$ is the probability density function of a one dimensional Gaussian variable with mean m and variance ς .

Let us fix a quantization grid $\mathbf{\Gamma}_{i,k+1} = \{\gamma_{i,k+1}^1, \dots, \gamma_{i,k+1}^N\}$ of size N relative to $\tilde{X}_{t_{k+1}}^i$. The distortion function associated with $\mathbf{\Gamma}_{i,k+1}$ reads

$$D_{i,k+1}(\mathbf{\Gamma}_{i,k+1}) = \sum_{j=1}^N \int_{C_j(\mathbf{\Gamma}_{i,k+1})} \left(x_{k+1}^i - \gamma_{i,k+1}^j\right)^2 \mathbb{P}\left(\tilde{X}_{t_{k+1}}^i \in dx_{k+1}^i\right) \quad (35)$$

where $(C_j(\mathbf{\Gamma}_{i,k+1}))_{j=1,\dots,N}$ is the Voronoi tessellation associated with the grid $\mathbf{\Gamma}_{i,k+1}$.

It is now possible to write the recursive quantization algorithm. Having quantized every i -th component of the vector $\tilde{\mathbf{X}}_{t_k}$, **via a grid of size N^i** , it is possible to approximate the distribution in (34) as

$$\mathbb{P}\left(\tilde{X}_{t_{k+1}}^i \in dx_{k+1}^i\right) \approx \sum_{j_1=1}^{N^1} \cdots \sum_{j_d=1}^{N^d} \phi_{m_i(t_k, x_{1,k}^{j_1}, \dots, x_{d,k}^{j_d}), \varsigma_i(t_k, x_{1,k}^{j_1}, \dots, x_{d,k}^{j_d})}(x_{k+1}^i) \mathbb{P}\left(\tilde{\mathbf{X}}_{t_k} = \left(x_{1,k}^{j_1}, \dots, x_{d,k}^{j_d}\right)\right), \quad (36)$$

where $x_{\ell,k}^{j_\ell}$ corresponds to the j_ℓ -th element of the optimal quantization grid of the ℓ -th component of the vector $\tilde{\mathbf{X}}_{t_k}$. It is immediate to see that it is possible to compute in closed form also the distribution of the vector $\tilde{\mathbf{X}}_{t_{k+1}}$: indeed, we have that

$$\mathbb{P}\left(\tilde{\mathbf{X}}_{t_{k+1}} \in dx_{k+1}\right) \approx \sum_{j_1=1}^{N^1} \cdots \sum_{j_d=1}^{N^d} \bar{\phi}_{m(t_k, x_{1,k}^{j_1}, \dots, x_{d,k}^{j_d}), \varsigma(t_k, x_{1,k}^{j_1}, \dots, x_{d,k}^{j_d})}(x_{k+1}) \mathbb{P}\left(\tilde{\mathbf{X}}_{t_k} = \left(x_{1,k}^{j_1}, \dots, x_{d,k}^{j_d}\right)\right), \quad (37)$$

where $\bar{\phi}$ is the density function of a d -dimensional Gaussian random variable with mean $m(t_k, x_{1,k}^{j_1}, \dots, x_{d,k}^{j_d}) = \left(x_{1,k}^{j_1}, \dots, x_{d,k}^{j_d}\right) + \Delta\mu\left(t_k, x_{1,k}^{j_1}, \dots, x_{d,k}^{j_d}\right)$ and variance-covariance matrix $\varsigma(t_k, x_{1,k}^{j_1}, \dots, x_{d,k}^{j_d}) = \Delta(\Sigma\Sigma^T)\left(t_k, x_{1,k}^{j_1}, \dots, x_{d,k}^{j_d}\right)$.

Having computed all these elements, it is possible to obtain the (approximate) distortion function (35), its gradient and its Hessian function and to implement the Newton-Raphson algorithm as in [Callegaro et al. \(2015\)](#) and [Callegaro et al. \(2016\)](#).

5.2 Recursive Quantization of the SVJ Model

We focus now on the application of the arguments in Section 5.1 to the specific model considered. We consider the Euler scheme of the price S , instead of the log price X , since quantizing S instead of X is crucial if we want to be in the setting of Section 4.3 devoted to the study of the numerical error of our procedure. Using the notation of the previous section, $\mathbf{X}_t = (V_t, S_t)$ and $\tilde{X}_{t_k} = \left(\tilde{V}_{t_k}, \tilde{S}_{t_k}\right)$, and the Euler scheme reads

$$\begin{pmatrix} \tilde{V}_{t_{k+1}} \\ \tilde{S}_{t_{k+1}} \end{pmatrix} = \begin{pmatrix} \tilde{V}_{t_k} \\ \tilde{S}_{t_k} \end{pmatrix} + \begin{pmatrix} \kappa(\theta - \tilde{V}_{t_k}) \Delta \\ (r - \delta) \Delta \end{pmatrix} + \sqrt{\Delta} \begin{pmatrix} \sigma \sqrt{Q(\tilde{V}_{t_k})} & 0 \\ \rho \tilde{S}_{t_k} \sqrt{Q(\tilde{V}_{t_k})} & \tilde{S}_{t_k} \sqrt{\tilde{V}_{t_k} - \rho^2 Q(\tilde{V}_{t_k})} \end{pmatrix} \begin{pmatrix} \Delta W_k^1 \\ \Delta W_k^2 \end{pmatrix}. \quad (38)$$

We have then that

$$\mathbb{P}\left(\tilde{V}_{t_{k+1}} \in dv_{k+1}\right) = \int_{\mathbb{R}} \int_{\mathbb{R}} \phi_{m_1(t_k, v_k, s_k), \varsigma_1(t_k, v_k, s_k)}(v_{k+1}) \mathbb{P}\left(\tilde{V}_{t_k} \in dv_k, \tilde{S}_{t_k} \in ds_k\right), \quad (39)$$

where $m_1(t_k, v_k, s_k) = v_k + \kappa(\theta - v_k)\Delta = m_1(t_k, v_k)$ and $\varsigma_1(t_k, v_k, s_k) = \sigma^2 Q(v_k) = \varsigma_1(t_k, v_k)$. In the case of the price process

$$\mathbb{P}\left(\tilde{S}_{t_{k+1}} \in ds_{k+1}\right) = \int_{\mathbb{R}} \int_{\mathbb{R}} \phi_{m_2(t_k, v_k, s_k), \varsigma_2(t_k, v_k, s_k)}(s_{k+1}) \mathbb{P}\left(\tilde{V}_{t_k} \in dv_k, \tilde{S}_{t_k} \in ds_k\right), \quad (40)$$

where $m_2(t_k, v_k, s_k) = s_k + (r - \delta)\Delta$ and $\varsigma_2(t_k, v_k, s_k) = (s_k)^2 v_k$. Moreover, we notice that, since m_1 and ς_1 do not depend on s_k , we can simplify (39):

$$\mathbb{P}\left(\tilde{V}_{t_{k+1}} \in dv_{k+1}\right) = \int_{\mathbb{R}} \phi_{m_1(t_k, v_k), \varsigma_1(t_k, v_k)}(v_{k+1}) \mathbb{P}\left(\tilde{V}_{t_k} \in dv_k\right). \quad (41)$$

This allows to use the technique developed in [Pagès and Sagna \(2015\)](#) and [Callegaro et al. \(2015\)](#) for the quantization of the variance process, which is one dimensional and it can be discretized independently of S . On the other hand, of course, the quantization grids for S will depend on the ones for V .

We now give an idea on how it is possible to recursively obtain the quantization grids $\mathbf{\Gamma}_{1,k} =: \mathbf{\Gamma}_{V,k}$ and $\mathbf{\Gamma}_{2,k} =: \mathbf{\Gamma}_{S,k}$, $k = 1, \dots, L$. We suppose that the cardinality of the grids is fixed: $|\mathbf{\Gamma}_{1,k}| = N^V$ and $|\mathbf{\Gamma}_{2,k}| = N^S$, for every k . Moreover, we recall that the quantization grids at time $t_0 = 0$, $\mathbf{\Gamma}_{V,0}$ and $\mathbf{\Gamma}_{S,0}$, are vectors whose components correspond, respectively, with v_0 and S_0 .

Let us assume now that we have computed the optimal grids for the variance and the price process, namely $\mathbf{\Gamma}_{V,k} = \{v_k^1, \dots, v_k^{N^V}\}$ for the variance process and $\mathbf{\Gamma}_{S,k} = \{s_k^1, \dots, s_k^{N^S}\}$ for the price process, up to time t_k and that we want to obtain $\mathbf{\Gamma}_{V,k+1}$ and $\mathbf{\Gamma}_{S,k+1}$. To do this, we look for the zeros of the gradient of the distortion function (35) when the probability (37) takes the form

$$\mathbb{P}\left(\tilde{V}_{t_{k+1}} \in dv_{k+1}, \tilde{S}_{t_{k+1}} \in ds_{k+1}\right) \approx \sum_{i=1}^{N^V} \sum_{j=1}^{N^S} \bar{\phi}_{m(t_k, v_k^i, s_k^j), \varsigma(t_k, v_k^i, s_k^j)}(v_{k+1}, s_{k+1}) \mathbb{P}\left(\tilde{V}_{t_k} = v_k^i, \tilde{S}_{t_k} = s_k^j\right), \quad (42)$$

where $\bar{\phi}$ is the density of a bivariate Gaussian with mean

$$m(t_k, v_k^i, s_k^j) = \begin{pmatrix} v_k^i + \kappa(\theta - v_k^i)\Delta \\ s_k^j + (r - \delta)\Delta \end{pmatrix}$$

and variance

$$\varsigma(t_k, v_k^i, s_k^j) = \Delta \begin{pmatrix} \sigma^2 Q(v_k^i) & \rho \sigma s_k^j Q(v_k^i) \\ \rho \sigma s_k^j Q(v_k^i) & (s_k^j)^2 v_k^i \end{pmatrix}.$$

Remark 5.2. (Calculation of Transition Probabilities)

As a byproduct of recursive quantization, we instantaneously get for free also the transition probabilities. Indeed, from (42) we immediately have the transition densities

$$\mathbb{P}\left(\tilde{V}_{t_{k+1}} \in dv_{k+1}, \tilde{S}_{t_{k+1}} \in ds_{k+1} | \tilde{V}_{t_k} = v_k^i, \tilde{S}_{t_k} = s_k^j\right) \approx \bar{\phi}_{m(t_k, v_k^i, s_k^j), \varsigma(t_k, v_k^i, s_k^j)}(v_{k+1}, s_{k+1}),$$

for $i = 1, \dots, N^V$ and $j = 1, \dots, N^S$.

6 Numerical Results

In this section we present numerical results on pricing of European and Bermudan options. Polynomial quantization is only used to price vanilla options, while recursive marginal quantization, allowing for an immediate approximation of the transition probabilities, is exploited to price both European and Bermudan derivatives. It is noteworthy that the pricing of vanilla and exotic options still represents a challenge from a numerical point of view when the dimension of the process considered is strictly greater than one. Numerical results have been obtained in Matlab 9.2, with a CPU 2.7 GHz and 4 Gb memory computer.

Remark 6.1. (Barrier and Asian options)

If interested in different types of path dependent options, such as barrier or Asian, we refer the reader to the methodology developed in [Bormetti et al. \(2018\)](#), where, on top of a grid for the underlying process at every intermediate date (together with transition probabilities from one time step to another) a backward Monte Carlo procedure is applied to price barrier, Asian and auto-callable options.

We choose the following values for the mean and the standard deviation of the Gaussian weight function w (also recall Remark 2.2), which are the same as in [Ackerer et al. \(2018\)](#):

$$\sigma_w = \sqrt{\frac{v_{max}T}{2}} + 10^{-4}, \quad \mu_w = \mathbb{E}[X_T]$$

and in all the numerical examples we will consider the parameters in Table 1:

$\kappa = 1.7$	$\theta = 0.06$	$\sigma = 0.5$	$\rho = -0.5$
$V_0 = 0.1$	$v_{min} = 10^{-2}$	$v_{max} = 1$	$r = 0.04$
$\delta = 0$	$S_0 = 100$	$T = 1$	$M = 100$

Table 1: Parameters of the SVJ model.

6.1 The choice of M

Before showing our results, we discuss the choice of M . This is crucial, since it is known that the density $g_T^{(M)}$ introduced in Equation (24) it is not guaranteed to be non-negative (see e.g. Figure 3 in the v2 ArXiv version of [Ackerer et al. \(2018\)](#)). In the literature this is a well-known issue and it is due to the polynomial nature of the terms involved in the Gram Charlier expansion of the density. While conditions on the expansion coefficients which ensure positivity of the density have been found in the case when the series is truncated at the fourth order moment (see e.g. [Jondeau and Rockinger \(1999\)](#)), no results are known, to our knowledge, in the general case. Various tests and alternative approaches and expansions have been proposed and we cite, among them, [Rompolis and Tzavalis \(2007\)](#), [León et al. \(2009\)](#), [Níguez and Perote \(2012\)](#), [Chateau and Dufresne \(2017\)](#) and [Schlogl \(2013\)](#).

Investigating this problem in details is out of the scope of the paper, since changing the series expansion would lead to a brand new contribution, so in this subsection we motivate our choice of M , based on an analysis of the produced vanilla prices. We proceed by computing the prices of call options for different strikes via a Monte Carlo simulation with $2 \cdot 10^6$ paths

Strike	CI lower bound	MC price	CI upper bound
80	25.8761	25.8984	25.9207
85	22.1012	22.1224	22.1436
90	18.5964	18.6164	18.6364
95	15.3991	15.4178	15.4364
100	12.5398	12.5570	12.5742
105	10.0396	10.0552	10.0709
110	7.9047	7.9188	7.9329
115	6.1252	6.1379	6.1505
120	4.6775	4.6886	4.6998

Table 2: Monte Carlo prices and 95% confidence interval of Call options obtained with $2 \cdot 10^6$ simulations and antithetic variates.

and antithetic variates. In Table 2 we show numerical results along with a 95% confidence interval for Monte Carlo prices.

In order to choose M , we then evaluate Call options via the procedure introduced in [Ackerer et al. \(2018\)](#), for the same strike values in Table 2 and for different values of M and we check for which values of M the obtained price belongs to the 95% Monte Carlo confidence interval in Table 2 for every strike. Results are presented in Table 3 and they show that a safe choice is $M \geq 90$. In the following, we choose the more conservative value of $M = 100$. Notice that this choice is in line with [Ackerer et al. \(2018\)](#), where they also truncate the series and study the pricing performance with respect to it.

Strike	$M = 20$	$M = 30$	$M = 40$	$M = 50$	$M = 60$	$M = 70$	$M = 80$	$M = 90$	$M = 100$
80	25.7122	25.7899	25.8457	25.8759	25.8906	25.8970	25.8992	25.8995	25.8991
85	22.0032	22.0213	22.0586	22.0857	22.1023	22.1119	22.1171	22.1198	22.1211
90	18.6058	18.5559	18.5642	18.5794	18.5919	18.6006	18.6064	18.6101	18.6125
95	15.5373	15.4222	15.3976	15.3958	15.3996	15.4040	15.4078	15.4107	15.4129
100	12.8039	12.6355	12.5806	12.5612	12.5546	12.5528	12.5529	12.5536	12.5544
105	10.4018	10.1979	10.1207	10.0870	10.0712	10.0634	10.0595	10.0576	10.0566
110	8.3189	8.1007	8.0121	7.9703	7.9487	7.9369	7.9301	7.9261	7.9237
115	6.5370	6.3259	6.2381	6.1954	6.1726	6.1597	6.1520	6.1473	6.1442
120	5.0336	4.8491	4.7734	4.7370	4.7176	4.7066	4.7001	4.6961	4.6934

Table 3: Pricing of Call options via the methodology in [Ackerer et al. \(2018\)](#), for different strikes and different values of M .

The truncation error $\text{err}_M^{(1)} = \varepsilon^{(M)}$, introduced in (27), takes into account the impact of the truncation level M on the pricing of options. We study in Figure 1 the rate of convergence of this error, i.e., we perform a Log-Log regression assuming that $\varepsilon^{(M)} \sim M^{-\alpha_M}$. Indeed, Figure 1 shows the value of α_M via a Log-Log regression of the relative error for different truncation levels M , when pricing a book of European Call options. The dispersion in the regression study is not negligible, and this is due to the characteristics of the payoff, as the strike of the options impacts the convergence rate. Nevertheless, the figure shows that the truncation error has a linear behaviour as a function of the truncation level M .

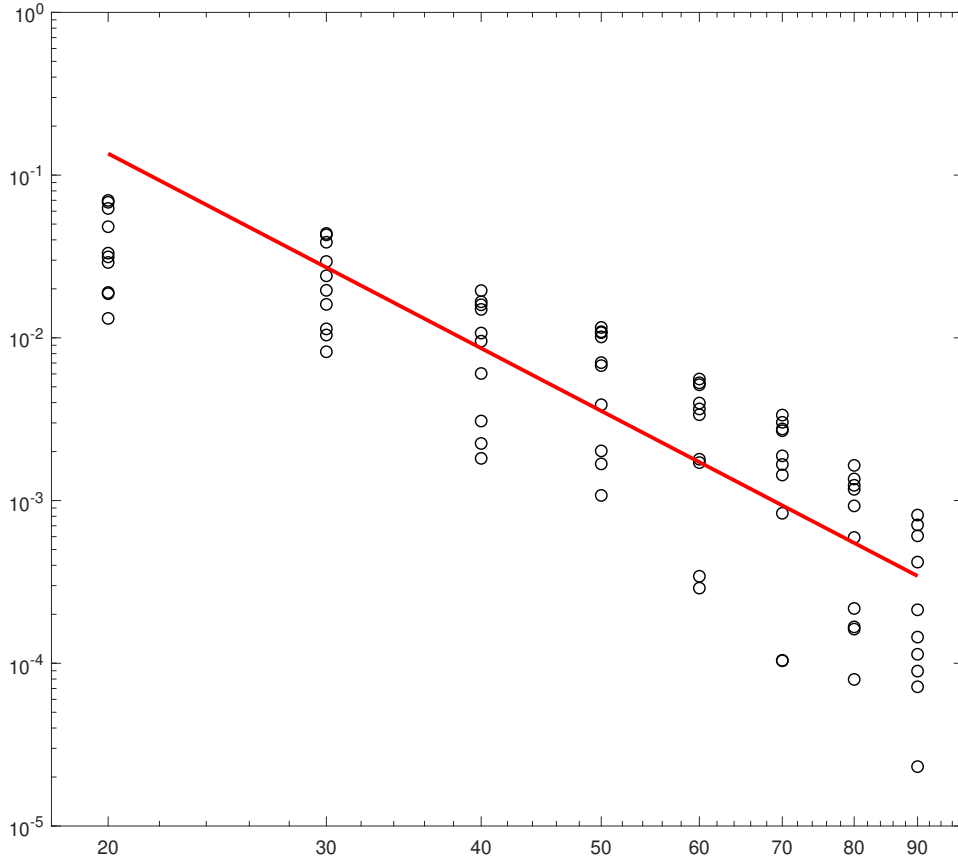


Figure 1: Log-Log regression of the relative error between the benchmark price and the price obtained via the methodology in [Ackerer et al. \(2018\)](#), as presented in [Table 3](#). The x-axis corresponds to the values of the truncation level M , for $M > 20$. The fit has been performed for $M > 20$. The slope α_M of the Log-Log regression line is 3.9740.

The quantization error $\text{err}_{M,N}^{(2)}$, introduced in [\(27\)](#), is dependent on both the quantization size N and the truncation level M . We assume that $\text{err}_{M,N}^{(2)} \sim C_M N^{-\alpha_N}$. According to [\(28\)](#), $\alpha_N = 1$ when $N \rightarrow \infty$. [Figure 2](#) shows the value of α_N via a Log-Log regression of the relative error when pricing an ATM European Call option for $M \in \{70, 80, 90\}$. The slope levels α_N are in line with the convergence rate of [Theorem 4.5](#).

Moreover, the regression shows that α_N is independent of the choice of the truncation level M , while only the intercept takes different values. In particular, the figure shows that, when implementing polynomial quantization, the choice of the quantization size N is almost independent of the choice of the truncation level M .

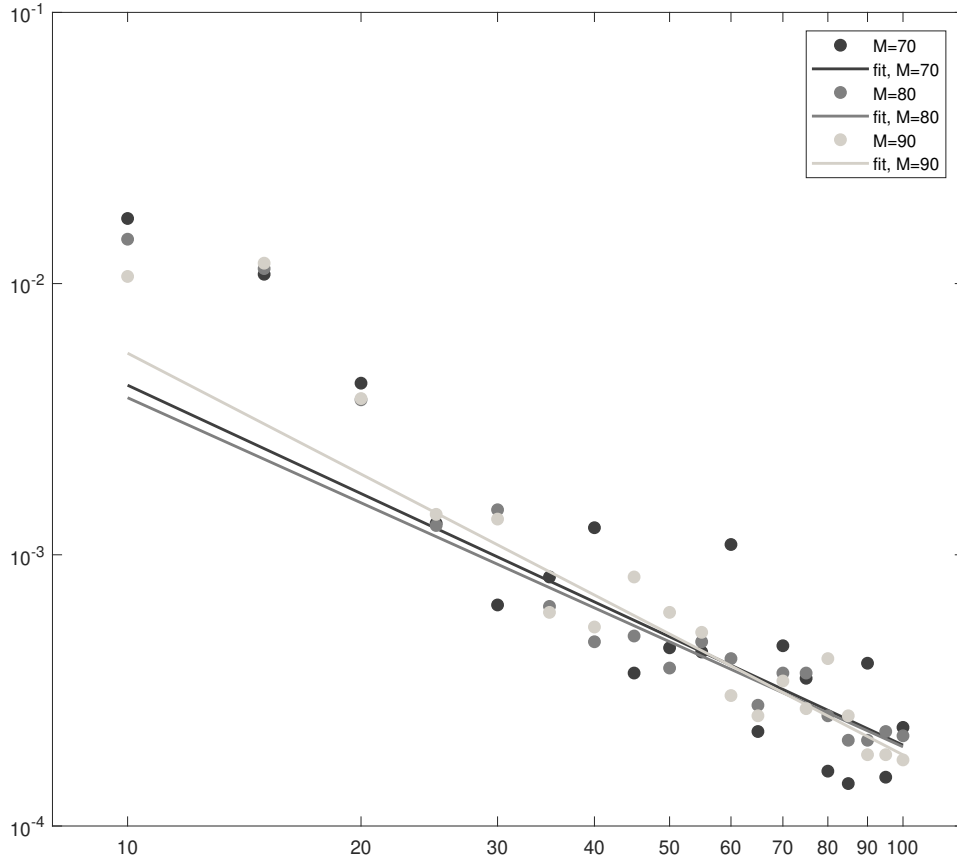


Figure 2: Log-Log regression of the relative error between the benchmark price and the price obtained via polynomial quantization of a European ATM Call option for the SVJ model with parameters as in Table 1, and for the truncation level $M \in \{70, 80, 90\}$. The x-axis corresponds to the values of the quantization size N . The fit has been performed for $N > 20$. The slopes α_N of the Log-Log regression lines are, respectively, $\{1.3270, 1.2870, 1.4811\}$.

6.2 Polynomial Quantization

We use the technique developed in Section 4.1, and we compute the quantization grid associated to the log price process at time T , i.e. we approximate X_T using an optimal grid $\Gamma^* = \{x_1^*, \dots, x_N^*\}$. The price of a Call option with maturity T and strike K is then approximated as

$$e^{-rT} \mathbb{E} \left[(e^{X_T} - K)^+ \right] \approx e^{-rT} \sum_{i=1}^N \left(e^{x_i^*} - K \right)^+ \mathbb{P}(X_T \in C_i(\Gamma^*)),$$

where $C_i(\Gamma^*) = \left[\frac{x_{i-1}^* + x_i^*}{2}, \frac{x_i^* + x_{i+1}^*}{2} \right]$ and the weights are given by (22). The results in Table 4 show that the quantization technique is accurate, when compared with the benchmark price, given by Ackerer et al. (2018). Moreover, the computational cost is comparable to

the execution time declared in [Ackerer et al. \(2018\)](#) ((see Figure 4)): computing Hermite moments requires the same time declared therein, getting the quantization grids costs 0.94 seconds and the price computation is immediate, so that the total computational times are equivalent.

Strike	Benchmark price	Quantization price	Relative error (%)
$K = 80$	25.8991	25.8646	0.1332
$K = 85$	22.1211	22.0980	0.1044
$K = 90$	18.6125	18.5866	0.1391
$K = 95$	15.4129	15.3648	0.3121
$K = 100$	12.5544	12.4710	0.6643
$K = 105$	10.0566	10.0165	0.3987
$K = 110$	7.9237	7.9066	0.2158
$K = 115$	6.1442	6.1093	0.5680
$K = 120$	4.6934	4.6626	0.6562

Table 4: Pricing comparison between the benchmark price and the price obtained via polynomial quantization of a European Call option for the SVJ model with parameters as in Table 1. The quantization grids have size $N = 20$.

6.3 Recursive Quantization

We use the methodology implemented in Section 5.2. Note that we do not exploit the fact that S is the exponential of a polynomial process, but we construct the optimal quantizers starting from the Euler scheme (38). We then compute, at every time step t_k , for $k = 1, \dots, L$, such that $t_L = T$, the quantization of the price process S at time t_k , that we call \widehat{S}_{t_k} . In order to price a European call option with strike K and maturity T we need only $\Gamma_{S,L}^* = \{s_1^L, \dots, s_N^L\}$, the optimal quantization grid associated to \widehat{S}_T , and we have the following approximation:

$$e^{-rT} \mathbb{E} \left[(e^{X_T} - K)^+ \right] \approx \sum_{j=1}^{N^S} \left(s_L^j - K \right)^+ \mathbb{P} \left(\widehat{S}_T = s_L^j \right),$$

where the weights are computed using (40). The results in Table 5 show that recursive quantization is efficient when compared to the benchmark methodology by [Ackerer et al. \(2018\)](#) and that it is a good alternative to polynomial quantization (see Table 4). Moreover, recursive quantization does not require the computation of Hermite moments, so the computational cost here is only relative to the computation of the quantization grids (recall, nevertheless, that here we have to both discretize the volatility and the price process), which corresponds to 2.24 seconds, with $N^S = 20$, $N^V = 10$ and $L = 12$ (so, a total of 360 points).

6.4 Bermudan Options

The advantage of the Recursive Marginal quantization algorithm developed in Section 5.1 is the possibility to price path dependent options, since we approximate the process at every time step of the Euler scheme, and the transition densities are given directly by the algorithm, as shown in (42). This motivates us to show an application of this methodology to the

Strike	Benchmark price	Quantization price	Relative error (%)
$K = 80$	25.8991	25.9082	0.0351
$K = 85$	22.1211	22.1462	0.1135
$K = 90$	18.6125	18.6430	0.1639
$K = 95$	15.4129	15.4395	0.1726
$K = 100$	12.5544	12.5677	0.1059
$K = 105$	10.0566	10.0789	0.2217
$K = 110$	7.9237	7.9508	0.3420
$K = 115$	6.1442	6.1692	0.4069
$K = 120$	4.6934	4.7106	0.3665

Table 5: Pricing comparison between the benchmark price and the price obtained via recursive quantization of a European Call option for the SVJ model with parameters as in Table 1. The quantization grids have size $N^S = 20$, $N^V = 10$ for every time step, and $L = 12$.

pricing of Bermudan options. Pricing such options can be done via a backward procedure on the multinomial tree obtained via quantization, as presented e.g. in (Bally et al., 2005, Proposition 2.1). As a first benchmark used for comparison we consider the Longstaff Schwarz algorithm, as done also in Filipovic et al. (2020) when dealing with American Put options in a Jacobi exchange rate model. The results in Table 6 show the accuracy of our methodology. The computational cost behind the pricing of Bermudan Options using recursive quantization is the one required by the computation of the quantization grids (which do not depend on the derivative’s strike), that is the same declared in Section 6.3 (2.24 seconds), and the Bermudan backward algorithm, that, due to the size of the grids in our example, is instantaneous.

Strike	Benchmark price	Quantization price	Relative error (%)
$K = 80$	3.0410	2.9984	1.3997
$K = 85$	4.1040	4.1077	0.0899
$K = 90$	5.4579	5.5012	0.7931
$K = 95$	7.1493	7.2222	1.0199
$K = 100$	9.2192	9.3151	1.0404
$K = 105$	11.6984	11.8285	1.1120
$K = 110$	14.6035	14.7564	1.0470
$K = 115$	17.9352	18.0969	0.9015
$K = 120$	21.6788	21.8295	0.6951

Table 6: Benchmark price (Longstaff Schwartz) and the price obtained via recursive quantization of a Bermudan Put option for the SVJ model with parameters as in Table 1. The quantization grids have size $N^S = 20$, $N^V = 10$ for every time step, and $L = 12$.

For completeness we also compare our Bermudan option prices with those in (Cui et al., 2018, Section 5.2.1), the benchmark there being still Longstaff Schwartz. In Table 7 we list the parameters used therein, which we also consider here for comparison. The computational time declared by our competitors for seven strikes is 4.7 seconds with Matlab 8.5 on a personal computer with an i7-6700 CPU @3.40 GHz. This is analogous to the time required by RMQ to obtain all the prices below.

$\kappa = 3$	$\theta = 0.13$	$\sigma = 0.4$	$\rho = -0.2$
$V_0 = 0.13$	$v_{min} = 10^{-2}$	$v_{max} = 0.25$	$r = 0.02$
$\delta = 0$	$S_0 = 10$	$T = 0.5$	

Table 7: Parameters of the SVJ model in Cui et al. (2018).

In Table 8 we show, for seven different strikes: the Longstaff-Schwartz benchmark price, our RMQ price, the price proposed by Cui et al. (2018) obtained via a continuous-time Markov chain approximation and the two absolute errors.

Strike	Benchmark (L.-S.)	RMQ price	Cui et al. price	Abs err RMQ	Abs err Cui et al.
$K = 8.5$	0.3527	0.3603	0.3522	0.0076	0.0005
$K = 9$	0.5125	0.5209	0.5139	0.0084	0.0014
$K = 9.5$	0.7157	0.7213	0.7163	0.0056	0.0006
$K = 10$	0.9591	0.9624	0.9599	0.0033	0.0008
$K = 10.5$	1.2407	1.2430	1.2435	0.0023	0.0028
$K = 11$	1.5625	1.5620	1.5643	0.0005	0.0018
$K = 11.5$	1.9163	1.9156	1.9189	0.0007	0.0026

Table 8: Benchmark price (Longstaff-Schwartz) provided in Cui et al. (2018), the price obtained via RMQ quantization, the price in Cui et al. (2018) and the absolute errors for a Bermudan Put option for the SVJ model with parameters as in Table 7. The quantization grids have size $N^S = 40$, $N^V = 10$ for every time step, and $L = 12$.

The results in Table 8 confirm (recall Table 6) the effectiveness of RMQ when pricing path-dependent options in the SJV model.

7 Conclusion

In this paper we presented how to efficiently apply quantization techniques to polynomial processes. In particular, we focused on the SVJ model, but our results can be extended to any polynomial model. Our analysis on SVJ quantization provided numerical tools to develop fast exotic option pricing algorithms. We presented two approaches. Firstly, we exploited the polynomial property, and we provided new theoretical results to study the approximation errors. As a result we obtained *ad-hoc* pricing tools for polynomial models and we provided numerical examples to assess their goodness with respect to the existing literature. Secondly, we applied RMQ to polynomial processes, by viewing them as a particular class of stochastic volatility processes. This allowed us to price exotic options and numerical examples for Bermudan options were given. Our conclusion is that quantization is a powerful discretization procedure, with respect to precision and speed, whose paradigm can be easily used in the field of polynomial processes. Moreover, when looking at the family of quantization procedures with pricing in view, recursive marginal quantization is undoubtedly the most powerful, given its flexibility and effectiveness.

A Proofs of the Main Results

We here provide the proofs of the main results obtained in the paper.

A.1 Proof of Theorem 4.1

Proof. First of all notice that, for every $i = 1, \dots, N$, the expectation in (15) can be rewritten as $\mathbb{E}[f^i(X_T)]$, with $f^i(y) := (y - x_i) \mathbb{1}_{y \in \left[\frac{x_{i-1}+x_i}{2}, \frac{x_i+x_{i+1}}{2}\right]}$. In order to exploit the polynomial nature of our setting and to use the result in Equation (6), we need f^i to be in L_w^2 . We have

$$\begin{aligned} \|f^i\|_w^2 &= \int_{\mathbb{R}} (f^i(y))^2 w(y) dy \leq \int_{\mathbb{R}} (y - x_i)^2 w(y) dy \\ &= \int_{\mathbb{R}} y^2 w(y) dy - 2x_i \int_{\mathbb{R}} y w(y) dy + x_i^2 \int_{\mathbb{R}} w(y) dy \\ &= \sigma_w^2 + \mu_w^2 - 2x_i \mu_w + x_i^2 \end{aligned}$$

which is finite for every $i = 1, \dots, N$.

We want to compute the expected value in (16). Using the polynomial property in (6), we now rewrite it in the form of (17), where (recalling Equation (7))

$$\begin{aligned} f_n^i &= \int_{\mathbb{R}} (y - x_i) \mathbb{1}_{y \in \left[\frac{x_{i-1}+x_i}{2}, \frac{x_i+x_{i+1}}{2}\right]} H_n(y) w(y) dy \\ &= \underbrace{\int_{\mathbb{R}} y \mathbb{1}_{\left[\frac{x_{i-1}+x_i}{2}, \frac{x_i+x_{i+1}}{2}\right]}(y) H_n(y) w(y) dy}_{a_n^i} - x_i \underbrace{\int_{\mathbb{R}} \mathbb{1}_{\left[\frac{x_{i-1}+x_i}{2}, \frac{x_i+x_{i+1}}{2}\right]}(y) H_n(y) w(y) dy}_{b_n^i} \end{aligned}$$

We focus first on the computation of a_n^i . Let us define

$$h_n(K) = \int_{\mathbb{R}} y \mathbb{1}_{[K, \infty)}(y) H_n(y) w(y) dy,$$

then $a_n^i = h_n\left(\frac{x_{i-1}+x_i}{2}\right) - h_n\left(\frac{x_i+x_{i+1}}{2}\right)$. When $n = 0$ we have, integrating by parts, that

$$\begin{aligned} h_0(K) &= \int_{\mathbb{R}} y \mathbb{1}_{[K, \infty)}(y) w(y) dy \\ &= \sigma_w \phi\left(\frac{K - \mu_w}{\sigma_w}\right) + \mu_w \Phi\left(\frac{\mu_w - K}{\sigma_w}\right). \end{aligned}$$

When $n \geq 1$ we have that

$$\begin{aligned} h_n(K) &= \int_K^\infty y H_n(y) w(y) dy \\ &= \frac{1}{\sqrt{n!}} \int_K^\infty y \mathcal{H}_n\left(\frac{y - \mu_w}{\sigma_w}\right) \frac{1}{\sigma_w} \phi\left(\frac{y - \mu_w}{\sigma_w}\right) dy \\ &= \frac{1}{\sqrt{n!}} \int_{\frac{K - \mu_w}{\sigma_w}}^\infty (\sigma_w z + \mu_w) \mathcal{H}_n(z) \phi(z) dz \\ &= \frac{\sigma_w}{\sqrt{n!}} \int_{\frac{K - \mu_w}{\sigma_w}}^\infty z \mathcal{H}_n(z) \phi(z) dz + \frac{\mu_w}{\sqrt{n!}} \int_{\frac{K - \mu_w}{\sigma_w}}^\infty \mathcal{H}_n(z) \phi(z) dz. \end{aligned}$$

We exploit the recursive relation between the Hermite polynomials for $n \geq 1$, : $z\mathcal{H}_n(z) = \mathcal{H}_{n+1}(z) + n\mathcal{H}_{n-1}(z)$, in order to get

$$h_n(K) = \frac{\sigma_w}{\sqrt{n!}} \int_{\frac{K-\mu_w}{\sigma_w}}^{\infty} \mathcal{H}_{n+1}(z) \phi(z) dz + \frac{n\sigma_w}{\sqrt{n!}} \int_{\frac{K-\mu_w}{\sigma_w}}^{\infty} \mathcal{H}_{n-1}(z) \phi(z) dz + \frac{\mu_w}{\sqrt{n!}} \int_{\frac{K-\mu_w}{\sigma_w}}^{\infty} \mathcal{H}_n(z) \phi(z) dz.$$

The case $n = 1$ can be obtained directly using integration by parts, while for $n \geq 1$ we exploit the relation $\int_x^{\infty} \mathcal{H}_n(z) \phi(z) dz = \mathcal{H}_{n-1}(x) \phi(x)$, so to have the result for $h_n(K)$. The b_n^i coefficients can be computed similarly. In fact, if we define

$$l_n(K) = \int_{\mathbb{R}} \mathbb{1}_{[K, \infty)}(y) H_n(y) w(y) dy,$$

then $b_n^i = l_n\left(\frac{x_{i-1}+x_i}{2}\right) - l_n\left(\frac{x_i+x_{i+1}}{2}\right)$. The case when $n = 0$ is trivial, instead when $n \geq 1$ we have that

$$\begin{aligned} l_n(K) &= \int_K^{\infty} H_n(y) w(y) dy \\ &= \frac{1}{\sqrt{n!}} \int_K^{\infty} \mathcal{H}_n\left(\frac{y-\mu_w}{\sigma_w}\right) \frac{1}{\sigma_w} \phi\left(\frac{y-\mu_w}{\sigma_w}\right) dy \\ &= \frac{1}{\sqrt{n!}} \int_{\frac{K-\mu_w}{\sigma_w}}^{\infty} \mathcal{H}_n(z) \phi(z) dz \\ &= \frac{1}{\sqrt{n!}} \mathcal{H}_{n-1}\left(\frac{K-\mu_w}{\sigma_w}\right) \phi\left(\frac{K-\mu_w}{\sigma_w}\right). \end{aligned}$$

□

A.2 Proof of Proposition 4.3

Proof. Remember that g_T is the density of X_T . We can then rewrite $E_i(x_1, \dots, x_N)$ as

$$E_i(x_1, \dots, x_N) = \int_{\frac{x_{i-1}+x_i}{2}}^{\frac{x_i+x_{i+1}}{2}} y g_T(y) dy - x_i \int_{\frac{x_{i-1}+x_i}{2}}^{\frac{x_i+x_{i+1}}{2}} g_T(y) dy.$$

This shows that E_i depends only on x_{i-1} , x_i and x_{i+1} , so that the Jacobian matrix J is tridiagonal. Moreover the lower diagonal has components:

$$\begin{aligned} J_{i,i-1} &= \frac{\partial E_i}{\partial x_{i-1}}(x_{i-1}, x_i, x_{i+1}) = -\frac{1}{2} \frac{x_{i-1}+x_i}{2} g_T\left(\frac{x_{i-1}+x_i}{2}\right) + \frac{1}{2} x_i g_T\left(\frac{x_{i-1}+x_i}{2}\right) \\ &= \frac{1}{2} \left(\frac{x_i - x_{i-1}}{2}\right) g_T\left(\frac{x_{i-1}+x_i}{2}\right) \end{aligned}$$

and the upper diagonal reads:

$$\begin{aligned} J_{i,i+1} &= \frac{\partial E_i}{\partial x_{i+1}}(x_{i-1}, x_i, x_{i+1}) = \frac{1}{2} \frac{x_i+x_{i+1}}{2} g_T\left(\frac{x_i+x_{i+1}}{2}\right) - \frac{1}{2} x_i g_T\left(\frac{x_i+x_{i+1}}{2}\right) \\ &= \frac{1}{2} \left(\frac{x_{i+1} - x_i}{2}\right) g_T\left(\frac{x_i+x_{i+1}}{2}\right). \end{aligned}$$

We can deduce immediately that $J_{i,i-1} = J_{i-1,i}$, so that J is also symmetric. Finally the diagonal has components:

$$\begin{aligned} J_{i,i} &= \frac{\partial E_i}{\partial x_i}(x_{i-1}, x_i, x_{i+1}) = \frac{1}{2} \frac{x_i + x_{i+1}}{2} g_T \left(\frac{x_i + x_{i+1}}{2} \right) - \frac{1}{2} \frac{x_i + x_{i+1}}{2} g_T \left(\frac{x_i + x_{i+1}}{2} \right) \\ &\quad - \frac{1}{2} x_i \left(g_T \left(\frac{x_i + x_{i+1}}{2} \right) - g_T \left(\frac{x_{i-1} + x_i}{2} \right) \right) - \int_{\frac{x_{i-1} + x_i}{2}}^{\frac{x_i + x_{i+1}}{2}} g_T(y) dy \\ &= J_{i,i-1} + J_{i,i+1} - \int_{\frac{x_{i-1} + x_i}{2}}^{\frac{x_i + x_{i+1}}{2}} g_T(y) dy, \end{aligned}$$

and the integral in the last equality is exactly the weight of the i -th Voronoi cell. The expression for the density in (21) comes from the following fact: the pricing of a derivative with payoff f is, recall Equation (6),

$$\begin{aligned} \mathbb{E}[f(X_T)] &= \sum_{n \geq 0} f_n \ell_n \\ &= \sum_{n \geq 0} \int_{\mathbb{R}} f(y) H_n(y) w(y) dy \ell_n \\ &= \int_{\mathbb{R}} f(y) \sum_{n \geq 0} H_n(y) \ell_n w(y) dy, \end{aligned}$$

where the fact that we can change the order of the infinite sum and the integral is proved in [Ackerer et al. \(2018\)](#). Since the price of the derivative can be seen also as

$$\int_{\mathbb{R}} f(y) g_T(y) dy,$$

(21) follows. Finally, the expression for $\mathbb{P}(X_T \in C_i(\Gamma))$ comes immediately from the proof of Theorem 4.1. \square

A.3 Proof of Lemma 4.4

Proof. First of all remember that

$$\text{err}_{M,N}^{(2)} = \left| e^{-rT} \int_{\mathbb{R}} (e^x - K)^+ g_T^{(M)}(x) dx - e^{-rT} \sum_{i=1}^N \int_{C_i(\Gamma^X)} (e^{x_i} - K)^+ g_T^{(M)}(x) dx \right|.$$

By introducing $s := e^x$ (notice that the payoff $(s - K)^+$ is Lipschitz with respect to s and this will be crucial), we have

$$\text{err}_{M,N}^{(2)} = e^{-rT} \left| \int_0^\infty (s - K)^+ \frac{g_T^{(M)}(\ln s)}{s} ds - \sum_{i=1}^N \int_{C_i(\Gamma^S)} (s_i - K)^+ \frac{g_T^{(M)}(\ln s)}{s} ds \right|$$

where $S_T^{(M)}$ is a random variable with density $h_T^{(M)}(s) := \frac{g_T^{(M)}(\ln s)}{s}$ for $s \in (0, +\infty)$ and where $\Gamma^S = \{s_1, \dots, s_N\}$ is an N -quantizer for $S_T^{(M)}$. We denote by $\widehat{S}_T^{(M,N)}$ the quantization of $S_T^{(M)}$ on Γ^S . We have

$$\mathbb{P}(\widehat{S}_T^{(M,N)} = s_i) = \int_{C_i(\Gamma^S)} h_T^{(M)}(s) ds.$$

Thus, working on the error $\text{err}_{M,N}^{(2)}$ corresponds to estimating the error coming from pricing a European Call option on $S_T^{(M)}$ via quantization. Now, for every Lipschitz function f with Lipschitz constant $[f]_{\text{Lip}}$, we have the following result:

$$\begin{aligned} \text{err}_{M,N}^{(2)} &= \left| \mathbb{E} \left[f \left(S_T^{(M)} \right) - f \left(\widehat{S}_T^{(M,N)} \right) \right] \right| \\ &\leq [f]_{\text{Lip}} \left\| S_T^{(M)} - \widehat{S}_T^{(M,N)} \right\|_1 \\ &\leq [f]_{\text{Lip}} \left\| S_T^{(M)} - \widehat{S}_T^{(M,N)} \right\|_2, \end{aligned}$$

where

$$\left\| S_T^{(M)} - \widehat{S}_T^{(M,N)} \right\|_r = \left(\sum_{i=1}^N \int_{C_i(\Gamma^S)} |s - s_i|^r h_T^{(M)}(s) ds \right)^{\frac{1}{r}}$$

is the L^r distance between the random variable with density $h_T^{(M)}$ and its quantization with N points. The Lipschitz constant for the payoff of a Call option is equal to one, and we have the result. \square

A.4 Proof of Theorem 4.5

Proof. It is worth noticing that in this polynomial setting, by definition (recall Equation (24)), the density $g_T^{(M)}(s)$, $s \in (0, +\infty)$, behaves like $s^M e^{-\frac{s^2}{2}}$, so that $h_T^{(M)}(s)$ behaves like $(\ln s)^M e^{-\frac{(\ln s)^2}{2}} \frac{1}{s} =: \widetilde{h}_T^{(M)}(s)$ at 0 and at infinity.

The proof of Callegaro et al. (2019, Theorem 2.11) consists of five steps, from zero to four. We now adapt it to our setting and in the case of quadratic quantization, namely in the case when p in the cited paper is equal to 2. The first three steps remain the same, so we briefly sketch them.

Step 0

We have to prove that $\|h_T^{(M)}\|_{\frac{1}{p+1}} = \|h_T^{(M)}\|_{\frac{1}{3}} < +\infty$. We hence study the convergence at 0 and at $+\infty$ of the integral of $\left(\widetilde{h}_T^{(M)}\right)^{\frac{1}{3}}$. In the rest of the proof, without loss of generality, we will assume that M is a multiple of 3, so that computations will be explicit. If we denote by $\overline{M} := \frac{M}{3}$, then, a primitive function is

$$\int \left(\widetilde{h}_T^{(M)}(s)\right)^{\frac{1}{3}} ds = \beta_{\overline{M}} \text{Erf} \left(\frac{-2 + \ln s}{\sqrt{6}} \right) + \sum_{n=1}^{\overline{M}} \alpha_n s^{\frac{2}{3}} e^{-\frac{1}{6} \ln^2 s} (\ln s)^{\overline{M}-n} =: \widetilde{H}_T^M(s),$$

where Erf is the error function, defined as $\text{Erf}(s) = \frac{2}{\pi} \int_0^s e^{-t^2} dt$, and the coefficients α_n , $n = 1, \dots, \overline{M}$ and $\beta_{\overline{M}}$ can be explicitly computed, e.g. using a symbolic programming language as *Mathematica*. Given that

$$\lim_{s \rightarrow +\infty} \widetilde{H}_T^M(s) = \beta_{\overline{M}}$$

and

$$\lim_{s \rightarrow 0} \widetilde{H}_T^M(s) = -\beta_{\overline{M}}$$

we obtain the finiteness of $\|h_T^{(M)}\|_{\frac{1}{3}}$.

Step 1

Here it can be shown the following estimation for the distortion function D , defined in (10), associated to $S_T^{(M)}$ and calculated in a generic grid $\Gamma = \{s_1, \dots, s_N\}$:

$$D(s_1, \dots, s_N) \leq \int_0^{s_1} (s_1 - y)^2 h_T^{(M)}(s) ds + \sum_{i=1}^{N-1} \frac{h_T^{(M)}(\xi_i) + h_T^{(M)}(\xi_{i+1})}{3} \left(\frac{s_{i+1} - s_i}{2} \right)^3 + \int_{s_N}^{+\infty} (y - s_N)^2 h_T^{(M)}(s) ds,$$

for some $\xi_1, \dots, \xi_N \in \mathbb{R}$.

Step 2

There exists a grid $\bar{\Gamma} = \{\bar{s}_1, \dots, \bar{s}_N\}$, and $\zeta_1, \dots, \zeta_{N-1}$, with $\zeta_i \in [\bar{s}_i, \bar{s}_{i+1}]$, such that

$$\int_0^{\bar{s}_i} \left(h_T^{(M)}(s) \right)^{\frac{1}{3}} ds = \int_{\bar{s}_N}^{+\infty} \left(h_T^{(M)}(s) \right)^{\frac{1}{3}} ds = \frac{1}{2N} \|h_T^{(M)}\|_{\frac{1}{3}},$$

and

$$(\bar{s}_{i+1} - \bar{s}_i)^2 = \frac{\|h_T^{(M)}\|_{\frac{1}{3}}^{\frac{2}{3}}}{\left(h_T^{(M)}(\zeta_i) \right)^{\frac{2}{3}} N^2}.$$

Step 3

We provide the following bound for the quantization error:

$$\begin{aligned} \left\| S_T^{(M)} - \widehat{S}_T^{(M,N)} \right\|_2^2 &\leq \frac{\|h_T^{(M)}\|_{\frac{1}{3}}^{\frac{2}{3}}}{24N^2} \sum_{i=1}^{N-1} \frac{h_T^{(M)}(\bar{\xi}_i) + h_T^{(M)}(\bar{\xi}_{i+1})}{\left(h_T^{(M)}(\zeta_i) \right)^{\frac{2}{3}}} (\bar{s}_{i+1} - \bar{s}_i) \\ &\quad + \int_0^{\bar{s}_1} (\bar{s}_1 - s)^2 h_T^{(M)}(s) ds + \int_{\bar{s}_N}^{+\infty} (s - \bar{s}_N)^2 h_T^{(M)}(s) ds \end{aligned}$$

Step 4

In Step 2 we have proved that

$$\frac{1}{N^2} = \frac{4}{\|h_T^{(M)}\|_{\frac{1}{3}}^{\frac{2}{3}}} \left(\int_0^{\bar{s}_i} \left(h_T^{(M)}(s) \right)^{\frac{1}{3}} ds \right)^2 = \frac{4}{\|h_T^{(M)}\|_{\frac{1}{3}}^{\frac{2}{3}}} \left(\int_{\bar{s}_N}^{+\infty} \left(h_T^{(M)}(s) \right)^{\frac{1}{3}} ds \right)^2,$$

So in order to prove that, when $N \rightarrow \infty$, $\int_0^{\bar{s}_1} (\bar{s}_1 - s)^2 h_T^{(M)}(s) ds$ and $\int_{\bar{s}_N}^{+\infty} (s - \bar{s}_N)^2 h_T^{(M)}(s) ds$ are $o\left(\frac{1}{N^2}\right)$, we just need to prove that

$$\lim_{y \rightarrow +\infty} \frac{\int_y^{+\infty} (s - y)^2 h_T^{(M)}(s) ds}{\left(\int_y^{+\infty} \left(h_T^{(M)}(s) \right)^{\frac{1}{3}} ds \right)^2} = 0$$

and that

$$\lim_{y \rightarrow 0} \frac{\int_0^y (s-y)^2 h_T^{(M)}(s) ds}{\left(\int_0^y \left(h_T^{(M)}(s) \right)^{\frac{1}{3}} ds \right)^2} = 0.$$

Since, both at 0 and at infinity, $h_T^{(M)} \sim \tilde{h}_T^{(M)}$, we can equivalently prove that

$$\lim_{y \rightarrow +\infty} \frac{\int_y^{+\infty} s^2 \tilde{h}_T^{(M)}(s) ds - 2y \int_y^{+\infty} s \tilde{h}_T^{(M)}(s) ds + y^2 \int_y^{+\infty} \tilde{h}_T^{(M)}(s) ds}{\left(\int_y^{+\infty} \left(\tilde{h}_T^{(M)}(s) \right)^{\frac{1}{3}} ds \right)^2} = 0$$

and that

$$\lim_{y \rightarrow 0} \frac{\int_0^y s^2 \tilde{h}_T^{(M)}(s) ds - 2y \int_0^y s \tilde{h}_T^{(M)}(s) dz + y^2 \int_0^y \tilde{h}_T^{(M)}(s) ds}{\left(\int_0^y \left(\tilde{h}_T^{(M)}(s) \right)^{\frac{1}{3}} ds \right)^2} = 0.$$

Up to a constant, we have that, for $\ell = 0, 1, 2$,

$$\int s^\ell \tilde{h}_T^{(M)}(s) ds = \beta_{\ell, M} \operatorname{Erf} \left(\frac{-\ell + \ln s}{\sqrt{2}} \right) + \sum_{n=0}^{M-1} \alpha_{\ell, n} e^{-\frac{1}{2}(\ln s)^2} s^\ell (\ln s)^n =: \tilde{H}_{\ell, T}^{(M)}(s),$$

where, as before, Erf is the error function and $\alpha_{\ell, n}$ and $\beta_{\ell, M}$ can be computed, for $\ell = 0, 1, 2$, with *Mathematica* in closed form. Please note that $\lim_{y \rightarrow +\infty} \tilde{H}_{\ell, T}^{(M)}(y) = \beta_{\ell, M}$ and $\lim_{y \rightarrow 0} \tilde{H}_{\ell, T}^{(M)}(y) = -\beta_{\ell, M}$, for $\ell = 0, 1, 2$. We have then that

$$\begin{aligned} \lim_{y \rightarrow +\infty} \frac{\int_y^{+\infty} s^2 \tilde{h}_T^{(M)}(s) dz - 2y \int_y^{+\infty} s \tilde{h}_T^{(M)}(s) ds + y^2 \int_y^{+\infty} \tilde{h}_T^{(M)}(s) ds}{\left(\int_y^{+\infty} \left(\tilde{h}_T^{(M)}(s) \right)^{\frac{1}{3}} ds \right)^2} = \\ \lim_{y \rightarrow +\infty} \frac{\beta_{2, M} - \tilde{H}_{2, T}^{(M)}(y) - 2y \left(\beta_{1, M} - \tilde{H}_{1, T}^{(M)}(y) \right) + y^2 \left(\beta_{0, M} - \tilde{H}_{0, T}^{(M)}(y) \right)}{\left(\beta_M - \tilde{H}_T^{(M)}(y) \right)^2} = 0, \end{aligned}$$

and, in a similar way,

$$\begin{aligned} \lim_{y \rightarrow 0} \frac{\int_0^y s^2 \tilde{h}_T^{(M)}(s) ds - 2y \int_0^y s \tilde{h}_T^{(M)}(s) ds + y^2 \int_0^y \tilde{h}_T^{(M)}(s) ds}{\left(\int_0^y \left(\tilde{h}_T^{(M)}(s) \right)^{\frac{1}{3}} ds \right)^2} = \\ \lim_{y \rightarrow 0} \frac{\tilde{H}_{2, T}^{(M)}(y) + \beta_{2, M} - 2y \left(\tilde{H}_{1, T}^{(M)}(y) + \beta_{1, M} \right) + y^2 \left(\tilde{H}_{0, T}^{(M)}(y) + \beta_{0, M} \right)}{\left(\tilde{H}_T^{(M)}(y) + \beta_M \right)^2} = 0. \end{aligned}$$

Since

$$\sum_{i=1}^{N-1} \frac{h_T^{(M)}(\bar{\xi}_i) + h_T^{(M)}(\bar{\xi}_{i+1})}{\left(h_T^{(M)}(\zeta_i)\right)^{\frac{2}{3}}} (\bar{s}_{i+1} - \bar{s}_i) \rightarrow 2 \|h_T^{(M)}\|_{\frac{1}{3}},$$

when $N \rightarrow +\infty$, we derive that

$$\lim_{N \rightarrow +\infty} N^2 \|S_T^{(M)} - \widehat{S}_T^{(M,N)}\|_2^2 \leq \frac{1}{12} \|h_T^{(M)}\|_{\frac{1}{3}},$$

and thanks to Lemma 4.4 we conclude. □

References

- Ackerer, D., Filipovic, D., and Pulido, S. (2018). The Jacobi stochastic volatility model. *Finance and Stochastics*, 22(3):667–700. DOI: [10.1007/s00780-018-0364-8](https://doi.org/10.1007/s00780-018-0364-8).
- Bally, V., Pagès, G., and Printems, J. (2001). A stochastic quantization method for nonlinear problems. *Monte Carlo Methods and Applications*, 1-2(7):21–33. DOI: [10.1515/mcma.2001.7.1-2.21](https://doi.org/10.1515/mcma.2001.7.1-2.21).
- Bally, V., Pagès, G., and Printems, J. (2005). A quantization tree method for pricing and hedging multidimensional American options. *Mathematical finance*, 15(1):119–168. DOI: [10.1111/j.0960-1627.2005.00213.x](https://doi.org/10.1111/j.0960-1627.2005.00213.x).
- Bormetti, G., Callegaro, G., Livieri, G., and Pallavicini, A. (2018). A backward Monte Carlo approach to exotic option pricing. *European Journal of Applied Mathematics*, 29(1):146–187. DOI: [10.1017/S0956792517000079](https://doi.org/10.1017/S0956792517000079).
- Callegaro, G., Fiorin, L., and Grasselli, M. (2015). Quantized calibration in local volatility. *Risk Magazine*, 28(4):62 – 67.
- Callegaro, G., Fiorin, L., and Grasselli, M. (2016). Pricing via recursive quantization in stochastic volatility models. *Quantitative Finance*, 17(6):855–872. DOI: [10.1080/14697688.2016.1255348](https://doi.org/10.1080/14697688.2016.1255348).
- Callegaro, G., Fiorin, L., and Grasselli, M. (2018). American quantized calibration in stochastic volatility. *Risk Magazine*, 31(2):84 – 88.
- Callegaro, G., Fiorin, L., and Grasselli, M. (2019). Quantization meets Fourier: a new technology for pricing options. *Annals of Operations Research*, 282(1):59–86. DOI: [10.1007/s10479-018-3048-z](https://doi.org/10.1007/s10479-018-3048-z).
- Chateau, J.-P. and Dufresne, D. (2017). Gram-Charlier processes and applications to option pricing. *Journal of Probability and Statistics*. DOI: [10.1155/2017/8690491](https://doi.org/10.1155/2017/8690491).
- Cuchiero, C., Keller-Ressel, M., and Teichmann, J. (2012). Polynomial processes and their applications to mathematical finance. *Finance and Stochastics*, 16(4):711–740. DOI: [10.1007/s00780-012-0188-x](https://doi.org/10.1007/s00780-012-0188-x).
- Cui, Z., Kirkby, J., and Nguyen, D. (2018). A general valuation framework for sabr and stochastic local volatility models. *SIAM J. Financial Mathematics*, 9(2):520–563. DOI: [10.1137/16M1106572](https://doi.org/10.1137/16M1106572).
- Filipovic, D. and Larsson, M. (2016). Polynomial diffusions and applications in finance. *Finance and Stochastics*, 20(4):931–972. DOI: [10.1007/s00780-016-0304-4](https://doi.org/10.1007/s00780-016-0304-4).
- Filipovic, D., Larsson, M., and Pulido, S. (2020). Markov cubature rules for polynomial processes. *Stochastic Processes and their Applications*, 130(4):1947 – 1971. DOI: [10.1016/j.spa.2019.06.010](https://doi.org/10.1016/j.spa.2019.06.010).
- Fiorin, L., Pagès, G., and Sagna, A. (2018). Product Markovian quantization of a diffusion process with applications to finance. *Methodology and Computing in Applied Probability*, 21(4):1087–1118. DOI: [10.1007/s11009-018-9652-1](https://doi.org/10.1007/s11009-018-9652-1).

- Fiorin, L. and Schoutens, W. (2020). Conic quantization: stochastic volatility and market implied liquidity. *Quantitative Finance*, 20(4):531–542. DOI: [10.1080/14697688.2019.1687928](https://doi.org/10.1080/14697688.2019.1687928).
- Gersho, A. and Gray, R. (1991). *Vector quantization and signal compression*. The Kluwer international series in engineering and computer science 159, Kluwer.
- Graf, S. and Luschgy, H. (2000). *Foundations of quantization for probability distributions*. Springer, New York. DOI: [10.1007/BFb0103945](https://doi.org/10.1007/BFb0103945).
- Jondeau, E. and Rockinger, M. (1999). Estimating Gram-Charlier expansions with positivity constraints. *Notes d'études et de recherche Banque de France*.
- Kieffer, J. (1982). Stochastic stability for feedback quantization schemes. *IEEE Transactions on Information Theory*, 28(2):248–254.
- León, A., Mencía, J., and Sentana, E. (2009). Parametric properties of semi-nonparametric distributions, with applications to option valuation. *Journal of Business & Economic Statistics*, 27(2):176–192. DOI: [10.1198/jbes.2009.0013](https://doi.org/10.1198/jbes.2009.0013).
- McWalter, T., Rudd, R., Kienitz, J., and Platen, E. (2018). Recursive Marginal Quantization of Higher-Order Schemes. *Quantitative Finance*, 18(4):693–706. DOI: [10.1080/14697688.2017.1402125](https://doi.org/10.1080/14697688.2017.1402125).
- Montes, T. (2020). *Numerical methods by optimal quantization in finance*. PhD thesis, Sorbonne Université - LPSM.
- Níguez, T. and Perote, J. (2012). Forecasting heavy-tailed densities with positive Edgeworth and Gram-Charlier expansions. *Oxford Bulletin of Economics and Statistics*, 74(4):600–627. DOI: [10.1111/j.1468-0084.2011.00663.x](https://doi.org/10.1111/j.1468-0084.2011.00663.x).
- Pagès, G. (2015). Introduction to vector quantization and its applications for numerics. *ESAIM: proceedings and surveys*, 48:29–79. DOI: [10.1051/proc/201448002](https://doi.org/10.1051/proc/201448002).
- Pagès, G. and Sagna, A. (2015). Recursive marginal quantization of the Euler scheme of a diffusion process. *Applied Mathematical Finance*, 22(5):463–498. DOI: [10.1080/1350486X.2015.1091741](https://doi.org/10.1080/1350486X.2015.1091741).
- Rompolis, L. and Tzavalis, E. (2007). Retrieving risk neutral densities based on risk neutral moments through a Gram-Charlier series expansion. *Mathematical and Computer Modelling*, 46(1-2):225–234. DOI: [10.1016/j.mcm.2006.12.021](https://doi.org/10.1016/j.mcm.2006.12.021).
- Schlogl, E. (2013). Option pricing where the underlying assets follow a Gram-Charlier density of arbitrary order. *Journal of Economic Dynamics and Control*, 37(3):611–632. DOI: [10.1016/j.jedc.2012.10.001](https://doi.org/10.1016/j.jedc.2012.10.001).