Stationary Heston model: Calibration and Pricing of exotics using Product Recursive Quantization

Vincent Lemaire * Thibaut Montes *† Gilles Pagès *

November 30, 2021

Abstract

A major drawback of the Standard Heston model is that its implied volatility surface does not produce a steep enough smile when looking at short maturities. For that reason, we consider the Stationary Heston model where the deterministic initial condition of the volatility process is replaced by its invariant measure. We show, based on calibrated parameters, that this model produces a steeper smile for short maturities than the original Heston model. We also present numerical procedure based on Product Recursive Quantization for the pricing of exotic options (Bermudan and Barrier options).

Introduction

Originally introduced by Heston in [Hes93], the Heston model is a stochastic volatility model used in quantitative finance to model the joint dynamics of a stock, denoted $(S_t^{(x)})_{t\geqslant 0}$, and its volatility, denoted $(v_t^x)_{t\geqslant 0}$ where $v_0^x = x > 0$ stands for the initial volatility at time $t = 0$. Historically, the initial condition of the volatility x is considered as deterministic and is calibrated to market data like the other parameters of the model. This model received an important attention among practitioners for two reasons: first, it is a stochastic volatility model, hence it introduces smile in the implied volatility surface as observed in the market, which is not the case of models with constant volatility, and second, in its original form, we have access to a semi closed-form formula for the characteristic function which allows us to price European options (Call & Put) almost instantaneously using the Fast Fourier approach (Carr & Madan in [CM99]). Yet, a complaint often heard about the Heston model is that it fails to fit the implied volatility surface for short maturities because the model cannot produce a steep-enough smile for those maturities (see [Gat11]).

In [PP09] devoted to the computation by Langevin Monte Carlo simulation of stationary regime of ergodic diffusion, the authors introduce as an example what they called the Stationary Heston model: noting that the volatility process is ergodic with the Gamma distribution as unique invariant distribution $\nu = \gamma(\alpha, \beta)$, where α and β depend on the structure parameters of the volatility process, they assume the volatility evolves under this stationary regime rather than starting at time 0 from a deterministic value. The couple asset-volatility in this Stationary Heston model will be denoted $(S_t^{(\nu)}, v_t^{\nu})_{t\geqslant 0}$. Considering this avatar of Heston model is justified (on simulated data) by more realistic implied volatility surfaces produced for short maturities by the pricing of vanilla options. The resulting pricing method of path-dependent options turns

*Sorbonne Université, Laboratoire de Probabilités, Statistique et Modélisation, LPSM, Campus Pierre et Marie Curie, case 158, 4 place Jussieu, F-75252 Paris Cedex 5, France.

†The Independent Calculation Agent, The ICA, 112 Avenue Kleber, 75116 Paris, France.

out to be too slow for practical implementation and no attempt of calibration is performed. The aim of this paper is two-fold: first emphasize on market data that the calibration of the Stationary Heston model does fit implied volatility surface for short maturities in a satisfactory way and then propose efficient numerical schemes based on cubature formulas (vanilla options) and optimal quantization (American en path-dependent) to price derivatives in such a model.

The underlying idea is that replacing the initial condition of the volatility by the stationary measure does not modify the long-term behavior of the implied volatility surface but does inject more randomness into the volatility for short maturities. This tends to produce a steeper smile for short maturities than the original model, which is the kind of behavior we are looking for. Later, the short-time and long-time behavior of the implied volatility generated by such model has been studied by Jacquier & Shi in [JS17]. Other extensions of the Heston model have been suggested and extensively analyzed in order to reproduce the slope of the skew for short-term expiring options: the Rough Heston model where the volatility satisfies a Voltera equation driven by a "rough" Liouville process with H-Hölder paths, $H \simeq 0.1$ (see [JR16, GJRS18, GJR18, CGP21, GR19] for details on the model and numerical solutions).

In the beginning of the paper, we briefly recall the well-known methodology used for the pricing of European option in the regular Heston model. Based on that, we express the price I_0 of a European option on the asset $S_T^{(\nu)}$ as

$$I_0 = \mathbb{E}\left[\mathrm{e}^{-rT}\,\varphi(S_T^{(\nu)})\right] = \mathbb{E}\left[f(v_0^\nu)\right] \tag{0.1}$$

where $f(x)$ denotes the price of the European option in the regular Heston model with same parameters but starting from $v_0 = x \in (0, +\infty)$. The last expectation can be computed efficiently using quadrature formulas either based on optimal quantization of the Gamma distribution or on Laguerre polynomials.

Once we are able to price European options, we can think of calibrating our model to market data. Indeed the parameters of the model are calibrated using the implied volatility surface observed in the market. However, the calibration of the Heston model is highly depending on the initial guess we choose in the minimization problem. This is due to an over-parametrization of the model (see [GR09]). Hence, when we consider the Heston model in its stationary regime, there is one parameter less to calibrate as the initial value of the volatility is no longer deterministic.

In the second part of paper, we deal with the pricing of Exotic options such as Bermudan and Barrier options. We propose a method based on hybrid product recursive quantization. The "hybrid" term comes from the fact that we use two different types of schemes for the discretization of the volatility and the asset (Milstein and Euler-Maruyama). Recursive quantization was first introduced by Pagès & Sagna in [PS15]. It is a Markovian quantization (see [PPP04]) drastically improved by the introduction of fast deterministic optimization procedure of the quantization grids and the transition weights. This optimization allows them to drastically reduce the time complexity by an order of magnitude and build such trees in a few seconds. Originally devised for Euler-Maruyama scheme of one dimensional Brownian diffusion, it has been extended to one-dimensional higher-order schemes by [MRKP18] and to medium dimensions using product quantization (see [FSP18, RMKP17, CFG18, CFG17, PS20]). Then, once the quantization tree is built, we proceed by a backward induction using the Backward Dynamic Programming Principle for the price of Bermudan options and using the methodology detailed in [Sag10, Pag18] based on the conditional law of the Brownian Bridge for the price of Barrier options.

The paper is organized as follows. First, in Section 1, we recall the definition of the Heston model and the interesting features of the volatility diffusion which bring us to define the Stationary Heston model. In Section 2, we give a fast solution for the pricing of European options in the Stationary Heston model when there exists methods for the Heston model. Finally, once we are able to price European options, we can define the optimization problem of calibration on implied volatility surface. We perform the calibration of both models and compare their induced smile for short maturities options. In Section 3, we propose a numerical method based on hybrid product recursive quantization for the pricing of exotic financial products: Bermudan and Barrier options.

1 The Heston Model

The Heston model is a two-dimensional diffusion process $(S_t^{(x)}, v_t^x)_{t\geqslant 0}$ with $S_t^{(x)}$ is the price of the risky asset at time t, and v_t^x the volatility at time t, depending on $x > 0$ the inital value of the volatility. This two-dimensional process $(S_t^{(x)}, v_t^x)_{t\geqslant 0}$ is solution to the following Stochastic Differential Equation

$$\begin{cases} \dfrac{dS_t^{(x)}}{S_t^{(x)}} = (r-q)dt + \sqrt{v_t^x}dW_t, \\ \quad dv_t^x = \kappa(\theta - v_t^x)dt + \xi\sqrt{v_t^x}d\widetilde{W}_t \end{cases} \tag{1.1}$$

where

- $S_0^{(x)} = s_0 > 0$ is the initial value of the price of the risky asset and $v_0^x = x > 0$ is the deterministic initial condition of the volatility,
- $(W, \widetilde{W})$ is a two-dimensional correlated Brownian motion, with correlation $\rho \in [-1, 1]$ (correlation between the asset and the volatility),
- $r \in \mathbb{R}$ denotes the interest rate, and $q \geqslant 0$ is the dividend rate,
- $\theta > 0$ the long run average price variance,
- $\kappa > 0$ the rate at which v_t^x reverts to θ,
- $\xi > 0$ is the volatility of the volatility.

This model is widely used by practitioner for various reasons. One is that it leads to semi-closed forms for vanilla options. The other reason is that it represents well the observed mid and long-term market behavior of the implied volatility surface observed on the market. However, it fails producing or even fitting to the smile observed for short-term maturities.

Remark 1.1 (The volatility)**.** One can notice that the volatility process is autonomous thence we are facing a one dimensional problem. Moreover, the volatility process is following a Cox-Ingersoll-Ross (CIR) diffusion also known as the square root diffusion. Existence and uniqueness of a strong solution to this stochastic differential equation has been first shown in [IW81], if $x \geqslant 0$. Moreover, it has been shown, see e.g. [LL11], that, if the Feller condition

$$\xi^2 \leqslant 2\kappa\theta \tag{1.2}$$

is in force, then for every $x > 0$, then there exists a unique solution $(v_t^x)_{t\geqslant 0}$ to the volatility equation which satisfies

$$\forall t \geqslant 0, \quad \mathbb{P}(\tau_0^x = +\infty) = 1 \tag{1.3}$$

where τ_0^x is the first hitting time defined by

$$\tau_0^x = \inf\{t \geqslant 0 \mid v_t^x = 0\} \quad \text{where } \inf\varnothing = +\infty. \tag{1.4}$$

Moreover, the CIR diffusion admits, as a Markov process, a unique stationary regime, characterized by its invariant distribution, the Gamma distribution

$$\nu = \gamma(\alpha, \beta) \tag{1.5}$$

where

$$\alpha = \theta\beta \quad \text{and} \quad \beta = 2\kappa/\xi^2. \tag{1.6}$$

Based on the above remarks, the idea is to precisely consider the volatility process under its stationary regime, i.e., replacing the deterministic initial condition from the Heston model by a ν-distributed random variable independent of $(W, \widetilde{W})$. We will refer to this model as the Stationary Heston model. Doing so, we inject more randomness for short maturities (t small) into the volatility but also to reduce the number of free parameters to stabilize and robustify the calibration of the Heston model which is commonly known to be overparametrized (see e.g. [GR09]).

This model was first introduced by [PP09] (see also [IW81], p. 221). More recently, [JS17] studied its small-time and large-time behaviors of the implied volatility. The dynamic of the asset price $(S_t^{(\nu)})_{t\geqslant 0}$ and its stochastic volatility $(v_t^\nu)_{t\geqslant 0}$ in the Stationary Heston model are given by

$$\begin{cases} \dfrac{dS_t^{(\nu)}}{S_t^{(\nu)}} = (r-q)dt + \sqrt{v_t^\nu}dW_t \\ \quad dv_t^\nu = \kappa(\theta - v_t^\nu)dt + \xi\sqrt{v_t^\nu}d\widetilde{W}_t \end{cases} \tag{1.7}$$

where $v_0^\nu \sim \nu = \gamma(\alpha, \beta)$ with $\beta = 2\kappa/\xi^2$, $\alpha = \theta\beta$. Other parameters $S_0^{(\nu)}$, r and q are the same as those in (1.1) and the parameters ρ, θ, κ, θ and ξ can be described as in the Heston model.

2 Pricing of European Options and Calibration

In this section, we first calibrate both Stationary Heston and Heston models and then compare their short-term behaviors of their resulting implied volatility surfaces. For that purpose we relied on a dataset of options price on the Euro Stoxx 50 observed the 26th of September 2019 (see Figure 1). This is why, as a preliminary step we briefly recall the well-known methodology for the evaluation of European Call and Put in the Heston model. Based on that, we outline how to price these options in the Stationary Heston model. Then, we describe the methodology employed for the calibration of both models: the Stationary Heston model (1.7) and the Heston model (1.1), both under the Feller condition (1.2), and then we discuss the obtained parameters and compare their short-term behaviors.

2.1 European options pricing

The price of the European option with payoff φ on the asset $S_T^{(\nu)}$, under the Stationary Heston model, exercisable at time T is given by

$$I_0 = \mathbb{E}\left[\mathrm{e}^{-rT}\varphi(S_T^{(\nu)})\right]. \tag{2.1}$$

After preconditioning by v_0^ν, we have

$$I_0 = \mathbb{E}\Big[\mathbb{E}\left[\mathrm{e}^{-rT}\varphi(S_T^{(\nu)}) \mid \sigma(v_0^\nu)\right]\Big] = \mathbb{E}\left[f(v_0^\nu)\right] \tag{2.2}$$

where $f(v)$ is the price of the European option in the Heston model with deterministic initial conditions for the set of parameters $\lambda(v) = (s_0, r, q, \theta, \kappa, \xi, \rho, v)$.

Example 2.1 (Call)**.** If φ is the payoff of a Call option then f is simply the price given by Fourier transform in the Heston model of the European Call Option. The price at time 0, for a spot price s_0, of an European Call $C(\lambda(v), K, T)$ with expiry T and strike K under the Heston model with parameters $\lambda(v) = (s_0, r, q, \theta, \kappa, \xi, \rho, v)$ is

$$\begin{aligned} C(\lambda(v), K, T) &= \mathbb{E}\big[\, \mathrm{e}^{-rT}(S_T^{(v)} - K)_+\big] \\ &= \mathrm{e}^{-rT}\Big(\mathbb{E}\big[S_T^{(v)}\,\mathbb{1}_{S_T^{(v)}\geqslant K}\big] - K\,\mathbb{E}\big[\mathbb{1}_{S_T^{(v)}\geqslant K}\big]\Big) \\ &= s_0\,\mathrm{e}^{-qT}\, P_1\big(\lambda(v), K, T\big) - K\,\mathrm{e}^{-rT}\, P_2\big(\lambda(v), K, T\big) \end{aligned} \tag{2.3}$$

with $P_1\big(\lambda(v), K, T\big)$ and $P_2\big(\lambda(v), K, T\big)$ given by

$$\begin{aligned} P_1\big(\lambda(v), K, T\big) &= \frac{1}{2} + \frac{1}{\pi}\int_0^{+\infty} \mathrm{Re}\bigg(\frac{\mathrm{e}^{-\mathbf{i}u\log(K)}}{iu}\,\frac{\psi\big(\lambda(v), u - \mathbf{i}, T\big)}{s_0\,\mathrm{e}^{(r-q)T}}\bigg)du \\ P_2\big(\lambda(v), K, T\big) &= \frac{1}{2} + \frac{1}{\pi}\int_0^{+\infty} \mathrm{Re}\bigg(\frac{\mathrm{e}^{-\mathbf{i}u\log(K)}}{\mathbf{i}u}\psi\big(\lambda(v), u, T\big)\bigg)du \end{aligned} \tag{2.4}$$

where $\mathbf{i}$ is the imaginary unit s.t. $\mathbf{i}^2 = -1$, $\psi\big(\lambda(v), u, T\big)$ is the characteristic function of the logarithm of the stock price process at time T. Several representations of the characteristic function exist, we choose to use the one proposed by [SST04, Gat11, AMST07], which is numerically more stable. It reads

$$\begin{aligned} \psi\big(\lambda(v), u, T\big) &= \mathbb{E}\big[\,\mathrm{e}^{\mathbf{i}u\log(S_T^{(v)})} \mid S_0^{(v)}, x\big] \\ &= \mathrm{e}^{\mathbf{i}u(\log(s_0)+(r-q)T)}\,\mathrm{e}^{\theta\kappa\xi^{-2}\big((\kappa-\rho\xi u\mathbf{i}-d)T-2\log((1-g\,\mathrm{e}^{-dt})/(1-g))\big)} \\ &\qquad \times \mathrm{e}^{v^2\xi^{-2}(\kappa-\rho\xi u\mathbf{i}-d)(1-\mathrm{e}^{-dt})/(1-g\,\mathrm{e}^{-dt})} \end{aligned} \tag{2.5}$$

with

$$d = \sqrt{(\rho\xi u\mathbf{i} - \kappa)^2 - \xi^2(-u\mathbf{i} - u^2)} \quad \text{and} \quad g = (\kappa - \rho\xi u\mathbf{i} - d)/(\kappa - \rho\xi u\mathbf{i} + d). \tag{2.6}$$

Hence, in (2.2), $f(v)$ can be replaced by $C\big(\lambda(v), K, T\big)$, which yields

$$I_0 = \mathbb{E}\big[\,\mathrm{e}^{-rT}(S_T^{(\nu)} - K)_+\big] = \mathbb{E}\Big[C\big(\lambda(v_0^\nu), K, T\big)\Big]. \tag{2.7}$$

Now, we come to the pricing of European options in the Stationary Heston model, using the expression of the density of $v_0^\nu \sim \gamma\alpha, \beta)$, (2.2) reads

$$I_0 = \mathbb{E}\big[f(v_0^\nu)\big] = \int_0^{+\infty} f(v)\frac{\beta^\alpha}{\Gamma(\alpha)}v^{\alpha-1}\,\mathrm{e}^{-\beta v}\,dv. \tag{2.8}$$

Now, several approaches exists in order to approximate this integral on the positive real line.

- *Quantization based quadrature formulas.* One could use a quantization-based cubature formula with an optimal quantizer of v_0^ν with the methodology detailed in Appendix B. Given that optimal quantizer of size N, $\widehat{v}_0^N$, we approximate I_0 by $\widehat{I}_0^N$

$$\widehat{I}_0^N = \mathbb{E}\big[f(\widehat{v}_0^N)\big] = \sum_{i=1}^N f(v_{0,i}^N)\,\mathbb{P}\left(\widehat{v}_0^N = v_{0,i}^N\right). \tag{2.9}$$

Remarks 2.2. In one dimension, the minimization problem, that consists in building an optimal quantizer, commute with affine transformations. Hence applying a affine transformation T to an optimal quantizer at level N of a distribution μ makes it an optimal quantizer of $\mu \circ T^{-1}$. Thus , if we consider an optimal quantization $\widehat{X}^N$ of a standard normal distribution $\mathcal{N}(0,1)$ then $\mu + \sigma \widehat{X}^N$ is an optimal quantizer of $\mathcal{N}(\mu, \sigma^2)$ and the associated probabilities of each Voronoi centroid stay the same.

In our case, noticing that if we consider a Gamma random variable $X \sim \gamma(\alpha, 1)$ then the rescaling of X by $1/\beta$ yields $X/\beta \sim \gamma(\alpha, \beta)$. Hence, for building the optimal quantizer $\widehat{v}_0^N$ of v_0^ν, we can build an optimal quantizer of $X \sim \gamma(\alpha, 1)$ and then rescale it by $1/\beta$, yielding $\widehat{v}_0^N = \widehat{X}^N/\beta$. Our numerical tests showed that it is numerically more stable to use this approach.

In order to build the optimal quantizer, we use Lloyd's method detailed in Appendix B to $X \sim \gamma(\alpha, 1)$ with the cumulative distribution function $F_X(x) = \mathbb{P}(X \leqslant x)$ and the partial first moment $K_X(x) = \mathbb{E}[X \mathbb{1}_{X \leqslant x}]$ given by

$$\begin{aligned} F_X(x) &= \frac{\Gamma(\alpha, x)}{\Gamma(\alpha)}, \qquad & K_X(x) = \alpha F_X(x) - \frac{x^\alpha \,\mathrm{e}^{-x}}{\Gamma(\alpha)}, \quad x > 0 \\ F_X(x) &= K_X(x) = 0, & \text{otherwise.} \end{aligned} \tag{2.10}$$

where $\Gamma(\alpha, x) = \int_0^x t^{\alpha-1} \,\mathrm{e}^{-t} \, dt$ is the lower incomplete gamma function. The distribution (weights) of the optimal quantizer $\widehat{v}_0^N$ is given by (B.10)

$$\mathbb{P}\left(\widehat{v}_0^N = v_{0,i}^N\right) = \mathbb{P}\left(\widehat{X}^N = x_i^N\right) = F_X\left(x_{i+1/2}^N\right) - F_X\left(x_{i-1/2}^N\right) \tag{2.11}$$

where, for every $i \in \{2, \ldots, N\}$, $x_{i-1/2}^N = \frac{x_{i-1}^N + x_i^N}{2}$ and $x_{1/2}^N = 0$ and $x_{N+1/2}^N = +\infty$.

• *Quadrature formula from Laguerre polynomials.* One could also use an algorithm based on fixed point quadratures for the numerical integration. Indeed, noticing that the density we are integrating against is a gamma density which is exactly the Laguerre weighting function (up to a rescaling). Then, I_0 rewrites

$$I_0 = \int_0^{+\infty} f(v) \frac{\beta^\alpha}{\Gamma(\alpha)} v^{\alpha-1} \,\mathrm{e}^{-\beta v} \, dv = \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^{+\infty} f(v) \omega(v) dv \tag{2.12}$$

where $\omega(v) = v^{\alpha-1} \,\mathrm{e}^{-\beta v}$ is the Laguerre weighting function. Then, for a fixed integer $n \geqslant 1$ ([1]), I_0 is approximated by

$$\widetilde{I}_0^n = \frac{\beta^\alpha}{\Gamma(\alpha)} \sum_{i=1}^n \omega_i f(v_i) \tag{2.13}$$

where the v_i's are the associated Laguerre nodes and the ω_i's their weights ([2]).

2.2 Calibration

Now that we are able to compute the price of European options, we define the problem of minimization we wish to optimize in order to calibrate our models parameters. Let $\mathcal{P}_{SH}$ be the set of parameters of the Stationary Heston model that needs to be calibrated, defined by

$$\mathcal{P}_{SH} = \left\{ \phi = (\theta, \kappa, \xi, \rho) \in \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}_+ \times [-1, 1] \right\} \tag{2.14}$$

[1] In practice, we choose $n = 20$. This number of points allows us to reach a high precision while keeping the computation time under control.

[2] During our numerical tests, we used the numerical integration routine *gsl_integration_fixed_laguerre* developed in the C++ gsl library. See https://www.gnu.org/software/gsl/doc/html/integration.html for more details on the implementation.

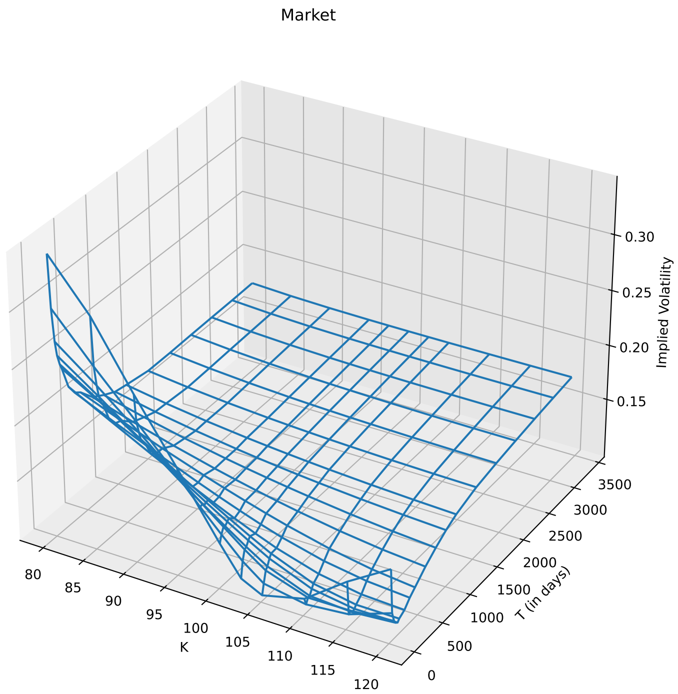

Figure 1: *Implied volatility surface of the* Euro Stoxx *50 as of the 26th of September 2019. (*$S_0 = 3541$*,* $r = -0.0032$ *and* $q = 0.00225$*) The expiries* T *are given in days and the strikes* K *in percentage of the spot.*

and let $\mathcal{P}_H$ be the set of parameters of the Heston model that needs to be calibrated, defined by

$$\mathcal{P}_H = \{\phi = (x, \theta, \kappa, \xi, \rho) \in \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}_+ \times [-1, 1]\}. \tag{2.15}$$

The others parameters are directly inferred from the market: we get $S_0 = 3541$, $r = -0.0032$ and $q = 0.00225$. Note that we did not directly included the Feller condition (1.2) in the set of parameters $\mathcal{P}_{SH}$ and $\mathcal{P}_H$ since this constraint will be introduced in the loss function of the calibration problem as a penalization (see further on). Indeed, preliminary calibration attempt tests performed without constraints yielded parameters really far from fulfilling the Feller condition. Which was inconsistent with the aim of pricing path-dependent or American style derivatives by any method.

In our case, we calibrate to option prices having all the same maturity. The problem can be formulated as follows: we search for the set of parameters $\phi^\star \in \mathcal{P}$ that minimizes the relative error between the implied volatility observed on the market and the implied volatility produced by the model for the given set of parameters, such that $\mathcal{P} = \mathcal{P}_{SH}$ for the Stationary Heston model and $\mathcal{P} = \mathcal{P}_H$ for the Heston model. There is no need to calibrate the parameters s_0, r and q since they are directly observable in the market.

Being interested in the short-term behaviors of the models, it is natural to calibrate both models based on options prices at a small expiry. Once the optimization procedures have been performed, we compare their performances for small expiries. For that, we calibrate using only the data on the volatility surface in Figure 1 with expiry 50 days ($T = 50/365$) and then we compare both models to the market implied volatility at expiry 22 days which is the smallest available in the data set.

Remark 2.3. The calibration is performed in C++ on a laptop with a 2,4 GHz 8-Core Intel Core i9 CPU using the randomized version of the simplex algorithm of [NM65] proposed in the

C++ library GSL. This algorithm is a derivative-free optimization method. It uses only the value of the function at each evaluation point. The computation time for calibrating the Heston model is around 20s and a bit more than a minute for the Stationary model. However, these computation times need to be considered carefully because the calibration time highly depends on the initial condition we choose for the minimizer and on the implementation of the Call pricer in the Heston model.

Let us be more precise. We want to find the set of parameter $\phi^\star$ that minimizes the relative error between the volatilities observed in the market and the ones generated by the model while enforcing the Feller condition. To take this constraint into account it appears that the most convenient compromise was to introduce it as a penalization in the loss function to be minimized, hence leading to the following problem

$$\min_{\phi\in\mathcal{P}} \sum_K \left(\frac{\sigma_{\mathrm{IV}}^{Market}(K,T) - \sigma_{\mathrm{IV}}^{Model}(\phi,K,T)}{\sigma_{\mathrm{IV}}^{Market}(K,T)} \right)^2 + \lambda \max(\xi^2 - 2\kappa\theta, 0) \tag{2.16}$$

where T is the expiry of the chosen options chosen a priori, K are their strikes and $\lambda > 0$ is the penalization factor to be adjusted during the procedure. $\sigma_{\mathrm{IV}}^{Market}(K,T)$ is the Mark-to-Market implied volatility taken from the observed implied volatility surface and the implied volatility $\sigma_{\mathrm{IV}}^{Model}(\phi,K,T)$ is the Black-Scholes volatility σ that matches the European Call price in this model to the price given by Heston or Stationary Heston model with the set of parameters ϕ.

The resulting parameters after calibration are summarized in Table 1. The Feller condition is still not fulfilled for both models but it is not far from being satisfied. We choose $\lambda = 0.01$ which seems to be right the compromise in order to avoid underfitting the model because of the constraint.

$\phi^\star$	ρ	v_0	θ	κ	ξ
Heston	-0.83	0.0045	0.17023	2.19	1.04
Stationary Heston	-0.99		0.02691	19.28	1.15

Table 1: *Parameters obtained for both models after calibration with penalization ($\lambda = 0.01$) for options with maturity 50 days ($S_0 = 3541$, $r = -0.0032$ and $q = 0.00225$).*

Figure 2 displays the resulting implied volatility curves at 50 days and 22 days for both calibrated models and observed in the market with calibration at 50 days.

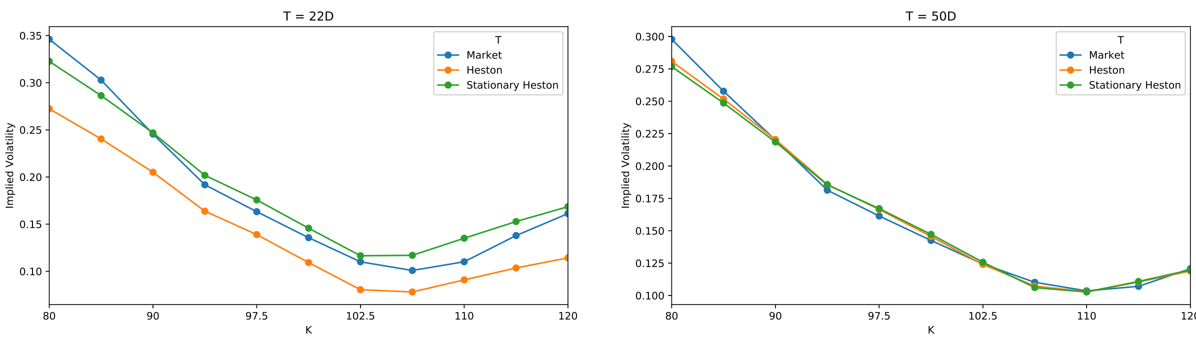

Figure 2: *Implied volatilities for 22 (left) and 50 (right) days expiry options after calibration at 50 days with penalization.*

Now, we extrapolate the implied volatility of both models for very short term maturities in Figure 3. The Stationary Heston model produces the desired smile, however the Heston model fails to produce prices sensibly different than 0 for strikes higher than 105 with this set of parameters, this is why there is no values in implied volatility curves.

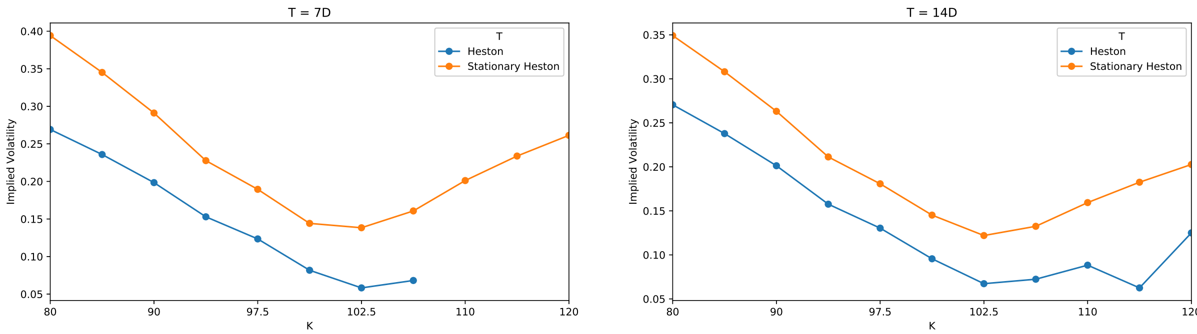

Figure 3: *Implied volatilities for* 7 *(left) and* 14 *(right) days expiry options after calibration at* 50 *days with penalization.*

From Figure 3 we can see that the Heston model fails at producing the desired smile for very small maturities whereas the Stationary model meets no difficulty to generate it. Figure 4 reproduces the term-structure of the implied volatility as a function of T in both models.

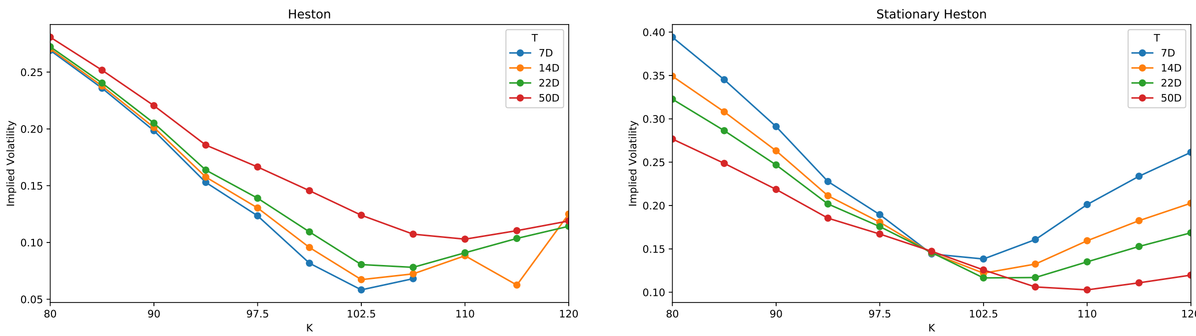

Figure 4: *Term-structure of the volatility as a function of T and K of both models (left: Heston and right: Stationary Heston) after calibration at* 50 *days with penalization.*

Figure 5 represents the relative error between the implied volatility given by the market and the one given by the models calibrated models at 50 days using a penalization. The Heston model completely fails to preserve the term-structure while being calibrated at 50 days. In comparison, the Stationary Heston behaves much better and the relative error does not explodes for long-term expiries, meaning that the long run average price variance is well caught.

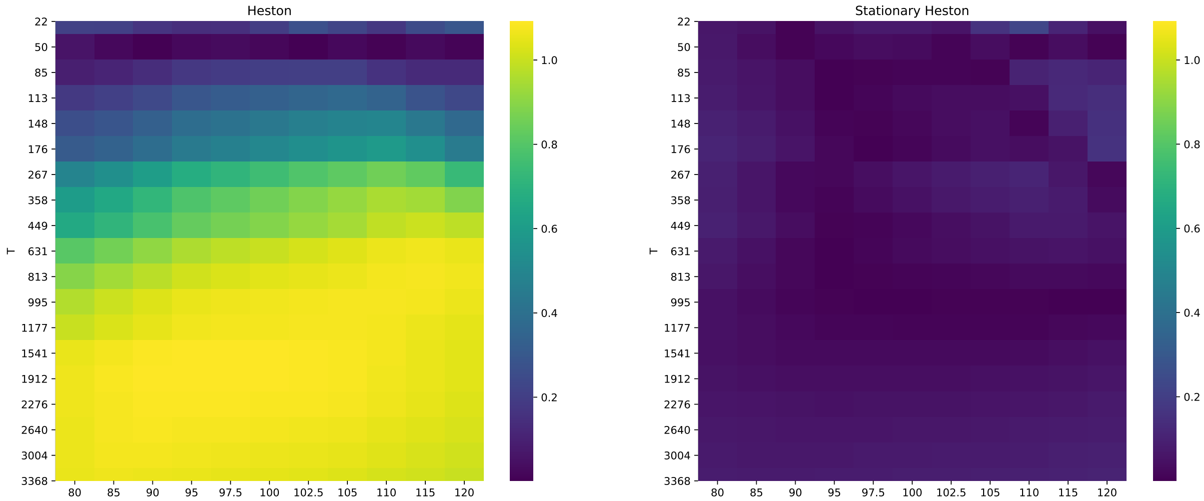

Figure 5: $(K,T) \longrightarrow \frac{|\sigma_{\mathrm{IV}}^{Market}(K,T)-\sigma_{\mathrm{IV}}^{Model}(\phi^\star,K,T)|}{\sigma_{\mathrm{IV}}^{Market}(K,T)}$ *for both models after calibration at* 50 *days with penalization. The expiries T are given in days and the strikes K are in percentage of the spot. (left: Heston and right: Stationary Heston).*

3 Toward the pricing of Exotic Options

In this Section, we evaluate first Bermudan options and then Barrier options under the Stationary Heston model. For both products, the pricing rely on a *Backward Dynamic Programming Principle*. The numerical solution we propose is based on a two-dimensional product recursive quantization scheme. We extend the methodology previously developed by [FSP18, CFG18, CFG17], where they considered an Euler-Maruyama scheme for both components. In this paper, we consider a hybrid scheme made up with an Euler-Maruyama scheme for the log-stock price dynamics and a Milstein scheme for the (boosted) volatility process. Finally, we apply the backward algorithm that corresponds to the financial product we are dealing with (the *Quantized Backward Dynamic Programming Principle* for Bermudan Options, see [BP03, BPP05, Pag18] and the algorithm by [Sag10, Pag18] for Barrier Options based on the conditional law of the Brownian motion).

3.1 Discretization scheme of a stochastic volatility model

We first present the time discretization schemes we use for the asset-volatility couple $(S_t^{(\nu)}, v_t^\nu)_{t\in[0,T]}$. For the volatility, we choose a Milstein on a *boosted* version of the process in order to preserve the positivity of the volatility and we select an Euler-Maruyama scheme for the log of the asset.

The boosted volatility. Based on the discussion in Appendix A, we will work with the following *boosted* volatility process: $Y_t = \mathrm{e}^{\kappa t}\, v_t^\nu, t \in [0,T]$ for some $\kappa > 0$, whose diffusion is given by

$$dY_t = \mathrm{e}^{\kappa t}\,\kappa\theta dt + \xi\,\mathrm{e}^{\kappa t/2}\sqrt{Y_t}d\widetilde{W}_t. \tag{3.1}$$

The Milstein discretization scheme of Y_t is given by

$$\bar{Y}_{t_{k+1}} = \mathcal{M}_{\tilde{b},\tilde{\sigma}}\big(t_k, \bar{Y}_{t_k}, Z_{k+1}\big) \tag{3.2}$$

with $(Z_k)_{k\geqslant 1}$ is an i.i.d. sequence of standard normal random variables, $t_k = \frac{Tk}{n}$ and $\widetilde{b}$ and $\widetilde{\sigma}$ are given by

$$\widetilde{b}(t,x) = \mathrm{e}^{\kappa t}\,\kappa\theta, \qquad \widetilde{\sigma}(t,x) = \xi\sqrt{x}\,\mathrm{e}^{\kappa t/2} \quad \text{and} \quad \widetilde{\sigma}'_x(t,x) = \frac{\xi\,\mathrm{e}^{\kappa t/2}}{2\sqrt{x}} \tag{3.3}$$

and $\mathcal{M}_{\widetilde{b},\widetilde{\sigma}}(t,x,z)$ defined by

$$\mathcal{M}_{\widetilde{b},\widetilde{\sigma}}(t,x,z) = x - \frac{\widetilde{\sigma}(t,x)}{2\widetilde{\sigma}'_x(t,x)} + h\Big(\widetilde{b}(t,x) - \frac{(\widetilde{\sigma}\widetilde{\sigma}'_x)(t,x)}{2}\Big) + \frac{(\widetilde{\sigma}\widetilde{\sigma}'_x)(t,x)h}{2}\Big(z + \frac{1}{\sqrt{h}\widetilde{\sigma}'_x(t,x)}\Big)^2. \tag{3.4}$$

We made this choice of scheme because, under the Feller condition, the positivity of $\mathcal{M}_{\widetilde{b},\widetilde{\sigma}}$ is ensured, since

$$\mathcal{M}_{\widetilde{b},\widetilde{\sigma}}(t,x,z) = h\,\mathrm{e}^{\kappa t}\left(\kappa\theta - \frac{\xi^2}{4}\right) + h\frac{\xi^2\,\mathrm{e}^{\kappa t}}{4}\left(z + \frac{2\sqrt{x}}{\sqrt{h}\xi\,\mathrm{e}^{\kappa t/2}}\right)^2 \tag{3.5}$$

and

$$\xi^2 \leqslant 2\kappa\theta \leqslant 4\kappa\theta.$$

Other schemes could have been used, see [Alf05] for an extensive review of the existing schemes for the discretization of the CIR model, but in our case we needed one allowing us to use the fast recursive quantization, i.e., where we can express explicitly and easily the cumulative distribution function and the first partial moment of the scheme, which is the case of the Milstein scheme (we give more details in Subsection 3.2).

Hence, as our time-discretized scheme is well defined because its positivity is ensured if the Feller condition is satisfied, we can start to think of the time-discretization of our process $(S_{t_k}^{(\nu)})_{k\in[\![0,n]\!]}$.

The log-**asset.** For the asset, the standard approach is to consider the process which is the logarithm of the asset $X_t = \log(S_t)$. Applying Itô's formula, the dynamics of X_t is given by

$$dX_t = \Big(r - q - \frac{v_t}{2}\Big)dt + \sqrt{v_t}dW_t. \tag{3.6}$$

Now, using a standard Euler-Maruyama scheme for the discretization of X_t, we have

$$\begin{cases} \bar{X}_{t_{k+1}} = \mathcal{E}_{b,\sigma}\big(t_k, \bar{X}_{t_k}, \bar{Y}_{t_k}, Z_{k+1}^1\big) \\ \bar{Y}_{t_{k+1}} = \mathcal{M}_{\widetilde{b},\widetilde{\sigma}}\big(t_k, \bar{Y}_{t_k}, Z_{k+1}^2\big) \end{cases} \tag{3.7}$$

where $(Z_k^1, Z_k^2)_{k\geqslant 1}$ is an i.i.d. sequence of bivariate normal random vectors with $Z_k^1 \sim \mathcal{N}(0,1)$, $Z_k^2 \sim \mathcal{N}(0,1)$, $\mathbb{C}\mathrm{orr}(Z_k^1, Z_k^2) = \rho$ and

$$\mathcal{E}_{b,\sigma}(t,x,y,z) = x + b(t,x,y)h + \sigma(t,x,y)\sqrt{h}\,z \tag{3.8}$$

with

$$b(t,x,y) = r - q - \frac{\mathrm{e}^{-\kappa t}\,y}{2} \qquad \text{and} \qquad \sigma(t,x,y) = \mathrm{e}^{-\kappa t/2}\,\sqrt{y}. \tag{3.9}$$

3.2 Hybrid Product Recursive Quantization

In this part, we describe the methodology used for the construction of the product recursive quantization tree of the couple log asset- boosted volatility in the Heston model.

In Figure 6, as an example, we synthesise the main idea behind the recursive quantization of a diffusion v_t which has been time-discretized with $F_0(t, x, z)$. We start at time $t_0 = 0$ with a quantizer $\widehat{v}_0$ taking values in the grid $\Gamma_{t_0} = \{v_1^0, \dots, v_{10}^0\}$ of size 10, where each point is represented by a black bullet ($\bullet$) with probability $p_i^0 = \mathbb{P}(\widehat{v}_0 = v_i^0)$ is represented by a bar. In the Stationary Heston model, $\widehat{v}_0$ is an optimal quantization of the Gamma distribution given by (1.5) and (1.6). Then, starting from this grid, we simulate the process from time t_0 to time $t_1 = 5$ days with our chosen time-discretization scheme $F_0(t, x, z)$, yielding $\widetilde{v}_1 = F_0(t_0, \widehat{v}_0, Z_1)$, where Z_1 is a standardized Gaussian random variable. Each trajectory starts from point v_i^0 with probability p_i^0. And finally we project the obtained distribution at time t_1 onto a grid $\Gamma_{t_1} = \{v_1^1, \dots, v_{10}^1\}$ of cardinality 10, represented by black triangles ($\blacktriangle$) such that $\widehat{v}_1$ is an optimal quantizer of the discretized and simulated process starting from quantizer $\widehat{v}_0$ at time $t_0 = 0$.

Remark 3.1. In practice, for low dimensions, we do not simulate trajectories. We use the information on the law of $\widetilde{v}_1$ conditionally of starting from $\widehat{v}_0$. The knowledge of the distribution allows us to use deterministic algorithms during the construction of the optimal quantizer of $\widetilde{v}_1$ that are a lot faster than algorithms based on simulation.

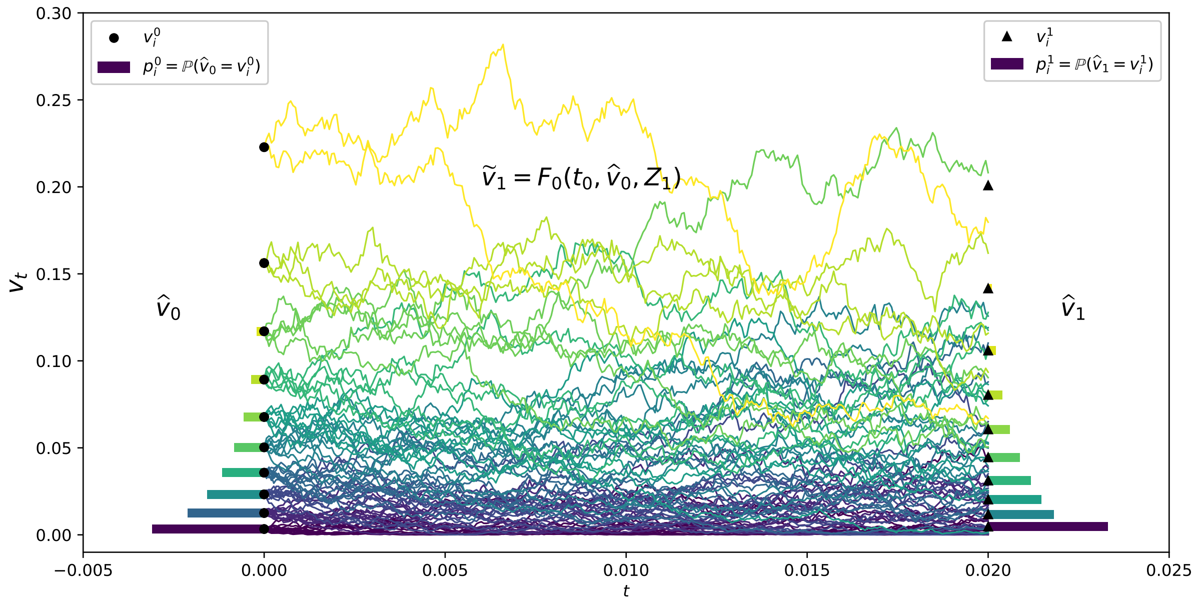

Figure 6: *Example of recursive quantization of the volatility process in the Heston model for one time-step.*

In our case, we consider the following stochastic volatility system

$$\begin{cases} dX_t = b(t, X_t, Y_t)dt + \sigma(t, X_t, Y_t)dW_t \\ dY_t = \widetilde{b}(t, Y_t)dt + \widetilde{\sigma}(t, Y_t)d\widetilde{W}_t \end{cases} \tag{3.10}$$

where W_t and $\widetilde{W}_t$ are two correlated Brownian motions with correlation $\rho \in [-1, 1]$, b and σ are defined in (3.9) and $\widetilde{b}$ and $\widetilde{\sigma}$ are defined in (3.3). Our aim is to build a quantization tree of the couple (X_t, Y_t) at given dates t_k, $k = 0, \dots, n$ based on a recursive product quantization scheme.

The product recursive quantization of such diffusion system has already been studied by [CFG17] and [RMKP17] in the case case where both processes are discretized using an Euler-Maruyama scheme.

One can notice that building the quantization tree $(\widehat{Y}_k)_{k\in[\![0,n]\!]}$ approximating $(Y_t)_{t\in[0,T]}$ is a one dimensional problem as the diffusion of Y_t is autonomous. Hence, based on our choice of discretization scheme, we will apply the fast recursive quantization (detailed above in Figure 6) that was introduced in [PS15] for one dimensional diffusion discretized by an Euler-Maruyama discretization scheme and then extended to higher order schemes, still in one dimension, by [MRKP18]. The minor difference with existing literature is that, in our problem, the initial condition y_0 is not deterministic.

Then, using the quantization tree of $(\widehat{Y}_k)_{k\in[\![0,n]\!]}$ we will be able to build the tree $(\widehat{X}_k)_{k\in[\![0,n]\!]}$ following ideas developed in [FSP18, RMKP17, CFG18, CFG17]. Indeed, once the quantization tree of the volatility is built, we are in a one-dimensional setting and we are able to use fast deterministic algorithms.

An estimate of the L^2-error induced by the above "hybrid" product recursive quantization approximation scheme has been established in [Mon20], Chapter 5. Classical results on quantization-based schemes on the L^2-error, showed in [FSP18, PS20], do not apply here. These estimates rely on Lipschitz continuity propagation of the semi-group induced by the scheme which is not satisfied here (du to presence of a square root in the diffusion term of the volatility in the CIR model (3.2)). The upper-bound obtained in [Mon20] should be considered with caution because it does not go to 0 when the number of time-steps goes to infinity. However, in practice, the number of time-steps is fixed and we study the behavior as a function of $N_{1,j}$ and $N_{2,j}$.

3.2.1 Quantizing the volatility (a one-dimensional case)

Let $(Y_t)_{t\in[0,T]}$ be a stochastic process in $\mathbb{R}$ and solution to the stochastic differential equation

$$dY_t = \widetilde{b}(t, Y_t)dt + \widetilde{\sigma}(t, Y_t)d\widetilde{W}_t \tag{3.11}$$

where Y_0 has the same law than the stationary measure ν: $\mathcal{L}(Y_0) = \nu$. In order to approximate our diffusion process, we choose a Milstein scheme for the time discretization, as defined in 3.4 and we build recursively the Markovian quantization tree $(\widehat{Y}_{t_k})_{k\in[\![0,n]\!]}$ where $\widehat{Y}_{t_{k+1}}$ is the Voronoi quantization of $\widetilde{Y}_{t_{k+1}}$ defined by

$$\widetilde{Y}_{t_{k+1}} = \mathcal{M}_{\widetilde{b},\widetilde{\sigma}}\big(t_k, \widehat{Y}_{t_k}, Z^2_{k+1}\big), \qquad \widehat{Y}_{t_{k+1}} = \mathrm{Proj}_{\Gamma^Y_{N_{2,k+1}}}\big(\widetilde{Y}_{t_{k+1}}\big) \tag{3.12}$$

and the projection operator $\mathrm{Proj}_{\Gamma^Y_{N_{2,k+1}}}(\cdot)$ is defined in (B.2), $\Gamma^Y_{N_{2,k+1}} = \big\{y^{k+1}_1, \dots, y^k_{N_{2,k+1}}\big\}$ is the grid of the optimal quantizer of $\widetilde{Y}_{t_{k+1}}$ and $Z^2_{k+1} \sim \mathcal{N}(0,1)$. In order alleviate the notations, we will denote $\widetilde{Y}_k$ and $\widehat{Y}_k$ in place of $\widetilde{Y}_{t_k}$ and $\widehat{Y}_{t_k}$.

The first step consists in building $\widehat{Y}_0$, an optimal quantizer of size $N_{2,0}$ of Y_0. Noticing that $Y_0 = v^\nu_0$, we use the optimal quantizer we built for the pricing of European options. Then, we build recursively $(\widehat{Y}_k)_{k=1,\dots,n}$, where the $N_{2,k}$-tuple are defined by $y^k_{1:N_{2,k}} = \big(y^k_1, \dots, y^k_{N_{2,k}}\big)$, by solving iteratively the minimization problem defined in the Appendix B in (B.6), with the help of Lloyd's method I. Replacing X by $\widetilde{Y}_{k+1}$ in (B.6) yields

$$
\begin{aligned}
y_j^{k+1} &= \frac{\mathbb{E}\left[\widetilde{Y}_{k+1}\,\mathbb{1}_{Y_{k+1}\in C_j\left(\Gamma^Y_{N_{2,k+1}}\right)}\right]}{\mathbb{P}\left(\widetilde{Y}_{k+1}\in C_j\big(\Gamma^Y_{N_{2,k+1}}\big)\right)} \\
&= \frac{\mathbb{E}\left[\mathcal{M}_{\widetilde{b},\widetilde{\sigma}}\big(t_k,\widehat{Y}_k,Z^2_{k+1}\big)\,\mathbb{1}_{\mathcal{M}_{\widetilde{b},\widetilde{\sigma}}\left(t_k,\widehat{Y}_k,Z^2_{k+1}\right)\in C_j\left(\Gamma^Y_{N_{2,k+1}}\right)}\right]}{\mathbb{P}\left(\mathcal{M}_{\widetilde{b},\widetilde{\sigma}}\big(t_k,\widehat{Y}_k,Z^2_{k+1}\big)\in C_j\big(\Gamma^Y_{N_{2,k+1}}\big)\right)}.
\end{aligned}
\tag{3.13}
$$

Now, preconditioning by $\widehat{Y}_k$ in the numerator and the denominator and using $p_i^k = \mathbb{P}\left(\widehat{Y}_k = y_i^k\right)$, we have

$$
\begin{aligned}
y_j^{k+1} &= \frac{\mathbb{E}\left[\,\mathbb{E}\left[\mathcal{M}_{\widetilde{b},\widetilde{\sigma}}\big(t_k,\widehat{Y}_k,Z^2_{k+1}\big)\,\mathbb{1}_{\mathcal{M}_{\widetilde{b},\widetilde{\sigma}}\left(t_k,\widehat{Y}_k,Z^2_{k+1}\right)\in C_j\left(\Gamma^Y_{N_{2,k+1}}\right)}\mid \widehat{Y}_k\right]\right]}{\mathbb{E}\left[\,\mathbb{P}\left(\mathcal{M}_{\widetilde{b},\widetilde{\sigma}}\big(t_k,\widehat{Y}_k,Z^2_{k+1}\big)\in C_j\big(\Gamma^Y_{N_{2,k+1}}\big)\mid \widehat{Y}_k\right)\right]} \\
&= \frac{\sum_{i=1}^{N_{2,k}}\mathbb{E}\left[\mathcal{M}_{\widetilde{b},\widetilde{\sigma}}\big(t_k,y_i^k,Z^2_{k+1}\big)\,\mathbb{1}_{\mathcal{M}_{\widetilde{b},\widetilde{\sigma}}\left(t_k,y_i^k,Z^2_{k+1}\right)\in C_j\left(\Gamma^Y_{N_{2,k+1}}\right)}\right]p_i^k}{\sum_{i=1}^{N_{2,k}}\mathbb{P}\left(\mathcal{M}_{\widetilde{b},\widetilde{\sigma}}\big(t_k,y_i^k,Z^2_{k+1}\big)\in C_j\big(\Gamma^Y_{N_{2,k+1}}\big)\right)p_i^k} \\
&= \frac{\sum_{i=1}^{N_{2,k}}\left(K_i^k\big(y^{k+1}_{j+1/2}\big)-K_i^k\big(y^{k+1}_{j-1/2}\big)\right)p_i^k}{\sum_{i=1}^{N_{2,k}}\left(F_i^k\big(y^{k+1}_{j+1/2}\big)-F_i^k\big(y^{k+1}_{j-1/2}\big)\right)p_i^k}
\end{aligned}
\tag{3.14}
$$

where $C_j\big(\Gamma^Y_{N_{2,k+1}}\big) = \big(y^{k+1}_{j-1/2}, y^{k+1}_{j+1/2}\big]$ is defined in (B.1). F_i^k and K_i^k are the cumulative distribution function and the first partial moment function of $U_i^k \sim \mu_i^k + \kappa_i^k(Z^1_{k+1}+\lambda_i^k)^2$ respectively with

$$
\begin{aligned}
&\kappa_j^k = \frac{(\widetilde{\sigma}\widetilde{\sigma}'_x)(t_k,y_j^k)h}{2}, \qquad\qquad \lambda_j^k = \frac{1}{\sqrt{h}\,\widetilde{\sigma}'_x(t_k,y_j^k)}, \\
\text{and}\qquad &\mu_j^k = y_j^k - \frac{\sigma(t_k,y_j^k)}{2\widetilde{\sigma}'_x(t_k,y_j^k)} + h\left(\widetilde{b}(t_k,y_j^k) - \frac{(\widetilde{\sigma}\widetilde{\sigma}'_x)(t_k,y_j^k)}{2}\right).
\end{aligned}
\tag{3.15}
$$

The functions F_i^k and K_i^k can explicitly be determined in terms of the density and the cumulative distribution function of the normal distribution.

Lemma 3.2. *Let $U = \mu + \kappa(Z+\lambda)^2$, with $\mu,\kappa,\lambda \in \mathbb{R}$, $\lambda \geqslant 0$, $\kappa > 0$ and $Z \sim \mathcal{N}(0,1)$ then the cumulative distribution function F_X and the first partial moment K_U of U are given by*

$$
\begin{aligned}
F_U(x) &= \big(F_Z(x_+) - F_Z(x_-)\big)\,\mathbb{1}_{x>\mu} \\
K_U(x) &= \left(F_U(x)\big(\mu+\kappa(\lambda^2+1)\big) + \frac{\kappa}{\sqrt{2\pi}}\Big(x_-\,\mathrm{e}^{-\frac{x_+^2}{2}} - x_+\,\mathrm{e}^{-\frac{x_-^2}{2}}\Big)\right)\mathbb{1}_{x>\mu}
\end{aligned}
\tag{3.16}
$$

where $x_+ = \sqrt{\frac{x-\mu}{\kappa}} - \lambda$, $x_- = -\sqrt{\frac{x-\mu}{\kappa}} - \lambda$ and F_Z is the cumulative distribution function of Z.

Finally, we can apply the Lloyd algorithm defined in Appendix B.9 with F_X and K_X defined by

$$
F_X(x) = \sum_{i=1}^{N_{2,k}} p_i^k\,F_i^k(x) \qquad \text{and} \qquad K_X(x) = \sum_{i=1}^{N_{2,k}} p_i^k\,K_i^k(x). \tag{3.17}
$$

In order to be able to build recursively the tree quantization $(\widehat{Y}_k)_{k=0,\dots,n}$, we need to have access to the weights $p_i^k = \mathbb{P}\left(\widehat{Y}_k = y_i^k\right)$, which can be themselves computed recursively, as well as the conditional probabilities $p_{ij}^k = \mathbb{P}\left(\widehat{Y}_{k+1} = y_j^{k+1} \mid \widehat{Y}_k = y_i^k\right)$.

Lemma 3.3. *The conditional probabilities p_{ij}^k are given by*

$$p_{ij}^k = F_i^k\big(y_{j+1/2}^{k+1}\big) - F_i^k\big(y_{j-1/2}^{k+1}\big). \tag{3.18}$$

And the probabilities p_j^{k+1} are given by

$$p_j^{k+1} = \sum_{i=1}^{N_{2,k}} p_i^k p_{ij}^k. \tag{3.19}$$

Proof. The

$$\begin{aligned}
p_{ij}^k &= \mathbb{P}\left(\widehat{Y}_{k+1} = y_j^{k+1} \mid \widehat{Y}_k = y_i^k\right)\\
&= \mathbb{P}\left(\mathcal{M}_{\widetilde{b},\widetilde{\sigma}}\big(t_k, \widehat{Y}_k, Z_{k+1}^2\big) \in C_j\big(\Gamma_{N_{2,k+1}}^Y\big) \mid \widehat{Y}_k = y_i^k\right)\\
&= \mathbb{P}\left(\mathcal{M}_{\widetilde{b},\widetilde{\sigma}}\big(t_k, y_i^k, Z_{k+1}^2\big) \in C_j\big(\Gamma_{N_{2,k+1}}^Y\big)\right)\\
&= F_i^k\big(y_{j+1/2}^{k+1}\big) - F_i^k\big(y_{j-1/2}^{k+1}\big)
\end{aligned}$$

and

$$p_j^{k+1} = \mathbb{P}\left(\widehat{Y}_{k+1} = y_j^{k+1}\right) = \sum_{i=1}^{N_{2,k}} \mathbb{P}\left(\widehat{Y}_{k+1} = y_j^{k+1} \mid \widehat{Y}_k = y_i^k\right) \mathbb{P}\left(\widehat{Y}_k = y_i^k\right) = \sum_{i=1}^{N_{2,k}} p_i^k\, p_{ij}^k.$$

□

As an illustration, we display in Figure 7 the rescaled grids obtained after recursive quantization of the boosted-volatility, where $\widehat{v}_k = \mathrm{e}^{-\kappa t_k}\, \widehat{Y}_k$ and $(\widehat{Y}_k)_{k=1,\dots,n}$ are the quantizers built using the fast recursive quantization approach.

3.2.2 Quantizing the asset (a one-dimensional case again)

Now, using the fact that $(Y_t)_t$ has already been quantized and the Euler-Maruyama scheme of $(X_t)_t$, as defined (3.8), we define the Markov quantized scheme

$$\widetilde{X}_{t_{k+1}} = \mathcal{E}_{b,\sigma}\big(t_k, \widehat{X}_{t_k}, \widehat{Y}_{t_k}, Z_{k+1}^1\big), \qquad \widehat{X}_{t_{k+1}} = \mathrm{Proj}_{\Gamma_{N_{1,k+1}}^X}\big(\widetilde{X}_{t_{k+1}}\big) \tag{3.20}$$

where the projection operator $\mathrm{Proj}_{\Gamma_{N_{1,k+1}}^X}(\cdot)$ is defined in (B.2), $\Gamma_{N_{1,k+1}}^X$ is the optimal $N_{1,k+1}$-quantizer of $\widetilde{X}_{t_{k+1}}$ and $Z_{k+1}^1 \sim \mathcal{N}(0,1)$. Again, in order to simplify the notations, $\widetilde{X}_{t_k}$ and $\widehat{X}_{t_k}$ are denoted in what follows by $\widetilde{X}_k$ and $\widehat{X}_k$.

Note that we are still in an one-dimensional case, hence we can apply the same methodology as developed in Appendix B and build recursively the quantization $\big(\widehat{X}_k\big)_{k=0,\dots,n}$ as detailed above,

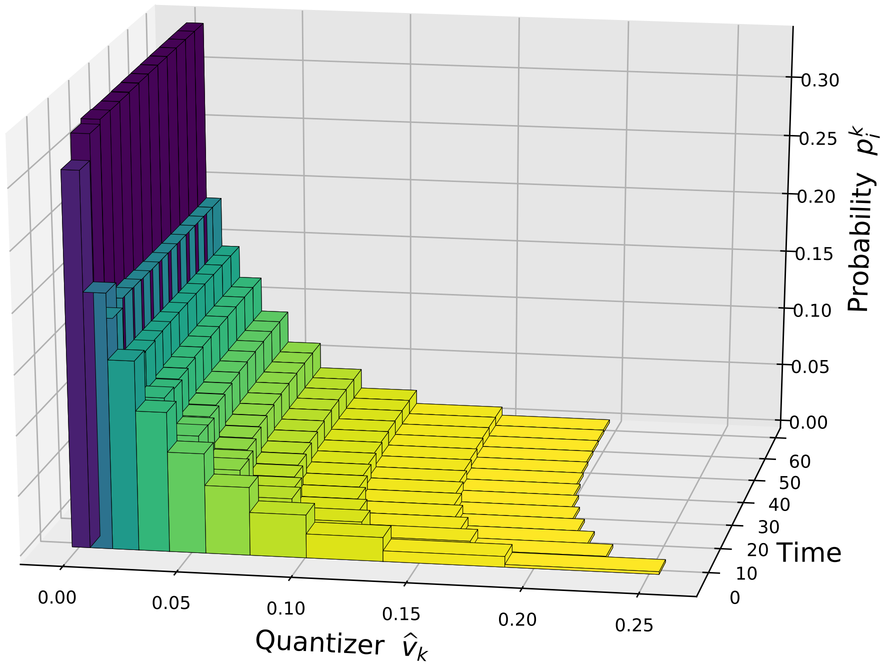

Figure 7: *Rescaled Recursive quantization of the boosted-volatility process with its associated weights from $t = 0$ to $t = 60$ days with a time step of 5 days with grids of size $N = 10$. The recursive quantization methodology is applied to $\widehat{Y}_k$ and then we display the rescaled volatility* $\widehat{v}_k = \mathrm{e}^{-\kappa t_k}\,\widehat{Y}_k$.

where the $N_{1,k}$-tuple are defined by $x^k_{1:N_{1,k}} = \big(x^k_1, \dots, x^k_{N_{1,k}}\big)$. Replacing X by $\widetilde{X}_k$ in (B.6) yield

$$\begin{aligned}
x^{k+1}_{j_1} &= \frac{\mathbb{E}\Big[\mathcal{E}_{b,\sigma}\big(t_k, \widehat{X}_{t_k}, \widehat{Y}_{t_k}, Z^1_{k+1}\big)\,\mathbb{1}_{\mathcal{E}_{b,\sigma}\left(t_k, \widehat{X}_{t_k}, \widehat{Y}_{t_k}, Z^1_{k+1}\right) \in C_{j_1}\left(\Gamma^X_{N_{1,k+1}}\right)}\Big]}{\mathbb{P}\Big(\mathcal{E}_{b,\sigma}\big(t_k, \widehat{X}_{t_k}, \widehat{Y}_{t_k}, Z^1_{k+1}\big) \in C_{j_1}\big(\Gamma^X_{N_{1,k+1}}\big)\Big)} \\
&= \frac{\displaystyle\sum_{i_1=1}^{N_{1,k}}\sum_{i_2=1}^{N_{2,k}} \mathbb{E}\Big[\mathcal{E}_{b,\sigma}\big(t_k, x^k_{i_1}, y^k_{i_2}, Z^1_{k+1}\big)\,\mathbb{1}_{\mathcal{E}_{b,\sigma}\left(t_k, x^k_{i_1}, y^k_{i_2}, Z^1_{k+1}\right) \in C_{j_1}\left(\Gamma^X_{N_{1,k+1}}\right)}\Big]\, p^k_{(i_1,i_2)}}{\displaystyle\sum_{i_1=1}^{N_{1,k}}\sum_{i_2=1}^{N_{2,k}} \mathbb{P}\Big(\mathcal{E}_{b,\sigma}\big(t_k, x^k_{i_1}, y^k_{i_2}, Z^1_{k+1}\big) \in C_{j_1}\big(\Gamma^X_{N_{1,k+1}}\big)\Big)\, p^k_{(i_1,i_2)}} \\
&= \frac{\displaystyle\sum_{i_1=1}^{N_{1,k}}\sum_{i_2=1}^{N_{2,k}} \Big(K^k_{(i_1,i_2)}\big(x^{k+1}_{j_1+1/2}\big) - K^k_{(i_1,i_2)}\big(x^{k+1}_{j_1-1/2}\big)\Big)\, p^k_{(i_1,i_2)}}{\displaystyle\sum_{i_1=1}^{N_{1,k}}\sum_{i_2=1}^{N_{2,k}} \Big(F^k_{(i_1,i_2)}\big(x^{k+1}_{j_1+1/2}\big) - F^k_{(i_1,i_2)}\big(x^{k+1}_{j_1-1/2}\big)\Big)\, p^k_{(i_1,i_2)}}
\end{aligned} \tag{3.21}$$

where $p^k_{(i_1,i_2)} = \mathbb{P}\big(\widehat{X}_k = x^k_{i_1}, \widehat{Y}_k = y^k_{i_2}\big)$ and $F^k_{(i_1,i_2)}$ and $K^k_{(i_1,i_2)}$ are the cumulative distribution function and the first partial moment function of the normal distribution $\mu^k_{(i_1,i_2)} + Z^1_{k+1}\sigma^k_{(i_1,i_2)}$

and they are defined by

$$
\begin{aligned}
F^k_{(i_1,i_2)}(x) &= F_Z\left(\frac{x-\mu^k_{(i_1,i_2)}}{\sigma^k_{(i_1,i_2)}}\right)\\
K^k_{(i_1,i_2)}(x) &= \mu^k_{(i_1,i_2)}F_Z\left(\frac{x-\mu^k_{(i_1,i_2)}}{\sigma^k_{(i_1,i_2)}}\right)+\sigma^k_{(i_1,i_2)}K_Z\left(\frac{x-\mu^k_{(i_1,i_2)}}{\sigma^k_{(i_1,i_2)}}\right)
\end{aligned}
\tag{3.22}
$$

with

$$
\mu^k_{(i_1,i_2)} = x^k_{i_1}+b(t_k,x^k_{i_1},y^k_{i_2})h \qquad \text{and} \qquad \sigma^k_{(i_1,i_2)} = \sigma(t_k,x^k_{i_1},y^k_{i_2})\sqrt{h} \tag{3.23}
$$

and F_Z and K_Z are the cumulative distribution function and the first partial moment of the standard normal distribution.

Finally, we apply the Lloyd method defined in Appendix (B.9) with F_X and K_X defined by

$$
F_X(x) = \sum_{i_1=1}^{N_{1,k}}\sum_{i_2=1}^{N_{2,k}} p^k_{(i_1,i_2)}\,F^k_{(i_1,i_2)}(x) \qquad \text{and} \qquad K_X(x) = \sum_{i_1=1}^{N_{1,k}}\sum_{i_2=1}^{N_{2,k}} p^k_{(i_1,i_2)}\,K^k_{(i_1,i_2)}(x). \tag{3.24}
$$

The sensitive part concerns the computation of the joint probabilities $p^k_{(i_1,i_2)}$. Indeed, they are needed at each step in order to be able to design recursively the quantization tree.

Lemma 3.4. *The joint probabilities $p^k_{(i_1,i_2)}$ are given by the following forward induction*

$$
p^{k+1}_{(j_1,j_2)} = \sum_{i=1}^{N_{1,k}}\sum_{j=1}^{N_{2,k}} p^k_{(i_1,i_2)}\,\mathbb{P}\left(\widehat{X}_{k+1}=x^{k+1}_{j_1},\widehat{Y}_{k+1}=y^{k+1}_{j_2}\mid \widehat{X}_k=x^k_{i_1},\widehat{Y}_k=y^k_{i_2}\right) \tag{3.25}
$$

where the joint conditional probabilities $\mathbb{P}\left(\widehat{X}_{k+1}=x^{k+1}_{j_1},\widehat{Y}_{k+1}=y^{k+1}_{j_2}\mid \widehat{X}_k=x^k_{i_1},\widehat{Y}_k=y^k_{i_2}\right)$ are given by the formulas below, depending on the correlation

- *if* $\mathrm{Corr}(Z^1_{k+1},Z^2_{k+1})=\rho=0$

$$
\mathbb{P}\left(\widehat{X}_{k+1}=x^{k+1}_{j_1},\widehat{Y}_{k+1}=y^{k+1}_{j_2}\mid \widehat{X}_k=x^k_{i_1},\widehat{Y}_k=y^k_{i_2}\right) = p^k_{i_2j_2}\Big[\mathcal{N}\left(x^k_{i_1,i_2,j_1,+}\right)-\mathcal{N}\left(x^k_{i_1,i_2,j_1,-}\right)\Big], \tag{3.26}
$$

where $p^k_{i_2j_2}$ is defined in (3.18) *and*

$$
x^k_{i_1,i_2,j_1,-} = \frac{x^{k+1}_{j_1-1/2}-\mu^k_{(i_1,i_2)}}{\sigma^k_{(i_1,i_2)}}, \qquad x^k_{i_1,i_2,j_1,+} = \frac{x^{k+1}_{j_1+1/2}-\mu^k_{(i_1,i_2)}}{\sigma^k_{(i_1,i_2)}}, \tag{3.27}
$$

with $\mu^k_{(i_1,i_2)}$ and $\sigma^k_{(i_1,i_2)}$ defined in (3.23).

- *if* $\mathrm{Corr}(Z^1_{k+1},Z^2_{k+1})=\rho\neq 0$

$$
\begin{aligned}
&\mathbb{P}\big(\widehat{X}_{k+1}=x^{k+1}_{j_1},\widehat{Y}_{k+1}=y^{k+1}_{j_2}\mid \widehat{X}_k=x^k_{i_1},\widehat{Y}_k=y^k_{i_2}\big)\\
&\quad=\mathbb{P}\left(Z^1_{k+1}\in\big(x^k_{i_1,i_2,j_1,-},x^k_{i_1,i_2,j_1,+}\big],Z^2_{k+1}\in\left(\sqrt{y^k_{i_2,j_2,-}}-\lambda^k_{i_2},\sqrt{y^k_{i_2,j_2,+}}-\lambda^k_{i_2}\right]\right)\\
&\qquad+\mathbb{P}\left(Z^1_{k+1}\in\big(x^k_{i_1,i_2,j_1,-},x^k_{i_1,i_2,j_1,+}\big],Z^2_{k+1}\in\left[-\sqrt{y^k_{i_2,j_2,+}}-\lambda^k_{i_2},-\sqrt{y^k_{i_2,j_2,-}}-\lambda^k_{i_2}\right)\right)
\end{aligned}
\tag{3.28}
$$

where

$$
y^k_{i_2,j_2,-} = 0\vee\frac{y^{k+1}_{j_2-1/2}-\mu^k_{i_2}}{\kappa^k_{i_2}}, \qquad y^k_{i_2,j_2,+} = 0\vee\frac{y^{k+1}_{j_2+1/2}-\mu^k_{i_2}}{\kappa^k_{i_2}}, \tag{3.29}
$$

with $\mu^k_{i_2}$, $\kappa^k_{i_2}$ and $\lambda^k_{i_2}$ defined in (3.15).

Remark 3.5. The probability in the right hand side of (3.28) can be computed using the cumulative distribution function of a correlated bivariate normal distribution[3]. Indeed, let

$$F_\rho(x_1, x_2) = \mathbb{P}(X_1 \leqslant x_1, X_2 \leqslant x_2)$$

the cumulative distribution function of the correlated centered Gaussian vector (X_1, X_2) with unit variance and correlation ρ, we have

$$\mathbb{P}\left(X_1 \in [a,b], X_2 \in [c,d]\right) = F_\rho(b,d) - F_\rho(b,c) - F_\rho(a,d) + F_\rho(a,c) \tag{3.30}$$

with $a, c \geqslant -\infty$ and $b, d \leqslant +\infty$.

Proof.

$$\begin{aligned}
p^{k+1}_{(j_1,j_2)} &= \mathbb{P}\left(\widehat{X}_{k+1} = x^{k+1}_{j_1}, \widehat{Y}_{k+1} = y^{k+1}_{j_2}\right) \\
&= \sum_{i=1}^{N_{1,k}} \sum_{j=1}^{N_{2,k}} \mathbb{P}\left(\widehat{X}_{k+1} = x^{k+1}_{j_1}, \widehat{Y}_{k+1} = y^{k+1}_{j_2} \mid \widehat{X}_k = x^k_{i_1}, \widehat{Y}_k = y^k_{i_2}\right) \mathbb{P}\left(\widehat{X}_k = x^k_{i_1}, \widehat{Y}_k = y^k_{i_2}\right) \\
&= \sum_{i=1}^{N_{1,k}} \sum_{j=1}^{N_{2,k}} p^k_{(i_1,i_2)} \mathbb{P}\left(\widehat{X}_{k+1} = x^{k+1}_{j_1}, \widehat{Y}_{k+1} = y^{k+1}_{j_2} \mid \widehat{X}_k = x^k_{i_1}, \widehat{Y}_k = y^k_{i_2}\right).
\end{aligned}$$

- if $\mathbb{C}\mathrm{orr}(Z^1_{k+1}, Z^2_{k+1}) = \rho = 0$

$$\begin{aligned}
&\mathbb{P}\left(\widehat{X}_{k+1} = x^{k+1}_{j_1}, \widehat{Y}_{k+1} = y^{k+1}_{j_2} \mid \widehat{X}_k = x^k_{i_1}, \widehat{Y}_k = y^k_{i_2}\right) \\
&\qquad = p^k_{i_2 j_2} \mathbb{P}\left(\widehat{X}_{k+1} = x^{k+1}_{j_1} \mid \widehat{X}_k = x^k_{i_1}, \widehat{Y}_k = y^k_{i_2}\right) \\
&\qquad = p^k_{i_2 j_2} \mathbb{P}\left(\bar{X}_{k+1} \in \left(x^{k+1}_{j_1-1/2}, x^{k+1}_{j_1+1/2}\right] \mid \widehat{X}_k = x^k_{i_1}, \widehat{Y}_k = y^k_{i_2}\right) \\
&\qquad = p^k_{i_2 j_2} \mathbb{P}\left(\mathcal{E}_{b,\sigma}\big(t_k, x^k_{i_1}, y^k_{i_2}, Z^1_{k+1}\big) \in \left(x^{k+1}_{j_1-1/2}, x^{k+1}_{j_1+1/2}\right]\right) \\
&\qquad = p^k_{i_2 j_2} \Big[\mathcal{N}\left(x^k_{i_1,i_2,j_1,+}\right) - \mathcal{N}\left(x^k_{i_1,i_2,j_1,-}\right)\Big],
\end{aligned}$$

- if $\mathbb{C}\mathrm{orr}(Z^1_{k+1}, Z^2_{k+1}) = \rho \neq 0$

$$\begin{aligned}
&\mathbb{P}\big(\widehat{X}_{k+1} = x^{k+1}_{j_1}, \widehat{Y}_{k+1} = y^{k+1}_{j_2} \mid \widehat{X}_k = x^k_{i_1}, \widehat{Y}_k = y^k_{i_2}\big) \\
&\quad = \mathbb{P}\left(\mathcal{E}_{b,\sigma}\big(t_k, x^k_{i_1}, y^k_{i_2}, Z^1_{k+1}\big) \in \left(x^{k+1}_{j_1-1/2}, x^{k+1}_{j_1+1/2}\right], \mathcal{M}_{\tilde{b},\tilde{\sigma}}\big(t_k, y^k_{i_2}, Z^2_{k+1}\big) \in \left(y^{k+1}_{j_2-1/2}, y^{k+1}_{j_2+1/2}\right]\right) \\
&\quad = \mathbb{P}\left(\mu^k_{(i_1,i_2)} + \sigma^k_{(i_1,i_2)} Z^1_{k+1} \in \left(x^{k+1}_{j_1-1/2}, x^{k+1}_{j_1+1/2}\right], \mu^k_{i_2} + \kappa^k_{i_2}(Z^2_{k+1} + \lambda^k_{i_2})^2 \in \left(y^{k+1}_{j_2-1/2}, y^{k+1}_{j_2+1/2}\right]\right) \\
&\quad = \mathbb{P}\left(Z^1_{k+1} \in \left(x^k_{i_1,i_2,j_1,-}, x^k_{i_1,i_2,j_1,+}\right], (Z^2_{k+1} + \lambda^k_{i_2})^2 \in \left(y^k_{i_2,j_2,-}, y^k_{i_2,j_2,+}\right]\right) \\
&\quad = \mathbb{P}\left(Z^1_{k+1} \in \left(x^k_{i_1,i_2,j_1,-}, x^k_{i_1,i_2,j_1,+}\right], Z^2_{k+1} \in \left(\sqrt{y^k_{i_2,j_2,-}} - \lambda^k_{i_2}, \sqrt{y^k_{i_2,j_2,+}} - \lambda^k_{i_2}\right]\right) \\
&\qquad + \mathbb{P}\left(Z^1_{k+1} \in \left(x^k_{i_1,i_2,j_1,-}, x^k_{i_1,i_2,j_1,+}\right], Z^2_{k+1} \in \left[-\sqrt{y^k_{i_2,j_2,+}} - \lambda^k_{i_2}, -\sqrt{y^k_{i_2,j_2,-}} - \lambda^k_{i_2}\right)\right).
\end{aligned}$$

□

[3] C++ implementation of the upper right tail of a bivariate normal distribution can be found in John Burkardt's website `https://people.sc.fsu.edu/~jburkardt/cpp_src/toms462/toms462.html`.

Remark 3.6. Another possibility for the quantization of the Stationary Heston model could be to use optimal quantizers for the volatility at each date t_k in place of using recursive quantization. Indeed, the volatility $(v_t)_t$ being stationary and the fact that we required the volatility to start at time 0 from the invariant measure, we could use the grid of the optimal quantization $\widehat{v}_0$ of size N of the stationary measure with its associated weights for every dates, hence setting $\widehat{v}_k = \widehat{v}_0$. We need as well the transitions from time t_k to t_{k+1} defined by

$$\mathbb{P}\left(\widehat{v}_{k+1} = v_{j_2}^{k+1} \mid \widehat{v}_k = v_{i_2}^k\right). \tag{3.31}$$

These probabilities can be computed using the conditional law of the CIR process described in [CIJR05, And07], which is a non-central chi-square distribution. Then, we would build the recursive quantizer of the log-asset at date $\widehat{X}_{k+1}$ with the standard methodology of recursive quantization using the already built quantizers of the volatility $\widehat{v}_k$ and the log-asset $\widehat{X}_k$ at time t_k, i.e.

$$\widetilde{X}_{k+1} = \mathcal{E}_{b,\sigma}\big(t_k, \widehat{X}_k, \widehat{v}_k, Z_{k+1}^1\big) \quad \text{and} \quad \widehat{X}_{k+1} = \mathrm{Proj}_{\Gamma^X_{N_{1,k+1}}}\big(\widetilde{X}_{k+1}\big) \tag{3.32}$$

where, this time, the Euler scheme is not defined as a function of the boosted-volatility but directly as a function of the volatility and is given by

$$\mathcal{E}_{b,\sigma}\big(t, x, v, z\big) = x + h\Big(r - q - \frac{v}{2}\Big) + \sqrt{v}\sqrt{h}z. \tag{3.33}$$

However, the difficulties with this approach come from the computation of the couple transitions

$$\mathbb{P}\left(\widehat{X}_{k+1} = x_{j_1}^{k+1}, \widehat{v}_{k+1} = v_{j_2}^{k+1} \mid \widehat{X}_k = x_{i_1}^k, \widehat{v}_k = v_{i_2}^k\right). \tag{3.34}$$

Indeed, these probability weights would not be as straightforward to compute as the methodology we adopt in this paper, namely using time-discretization schemes for both components. Our approach allows us to express the conditional probability of the couple as the probability that a correlated bi-variate Gaussian vector lies in a rectangle domain and this can be easily be computed numerically.

3.3 Backward algorithm for Bermudan and Barrier options

Bermudan Options. A Bermudan option is a financial derivative product that gives the right to its owner to buy or sell (or to enter to, in the case of a swap) an underlying product with a given payoff $\psi_t(\cdot,\cdot)$ at predefined exercise dates $\{t_0, \cdots, t_n\}$. Its price, at time $t_0 = 0$, is given by

$$\sup_{\tau \in \{t_0, \cdots, t_n\}} \mathbb{E}\left[\mathrm{e}^{-r\tau}\, \psi_\tau(X_\tau, Y_\tau) \mid \mathcal{F}_{t_0}\right]$$

where X_t and Y_t are solutions to the system defined in (3.10).

In this part, we follow the numerical solution first introduced by [BPP05, BP03]. They proposed to solve discrete-time optimal stopping problems using a quantization tree of the risk factors X_t and Y_t.

Let $\mathcal{F}^{X,Y} = (\mathcal{F})_{0 \leqslant k \leqslant n}$ the natural filtration of X and Y. Hence, we can define recursively the sequence of random variable L^p-integrable $(V_k)_{0 \leqslant k \leqslant n}$

$$\begin{cases} V_n = \mathrm{e}^{-rt_n}\, \psi_n(X_n, Y_n), \\ V_k = \max\big(\mathrm{e}^{-rt_k}\, \psi_k(X_k, Y_k), \mathbb{E}[V_{k+1} \mid \mathcal{F}_k]\big), \qquad 0 \leqslant k \leqslant n-1 \end{cases} \tag{3.35}$$

called *Backward Dynamic Programming Principle.* Then

$$V_0 = \sup\big\{\, \mathbb{E}[\mathrm{e}^{-r\tau}\, \psi_\tau(X_\tau, Y_\tau) \mid \mathcal{F}_0], \tau \in \Theta_{0,n}\big\}$$

with $\Theta_{0,n}$ the set of all stopping times taking values in $\{t_0, \cdots, t_n\}$. The sequence $(V_k)_{0 \leqslant k \leqslant n}$ is also known as the Snell envelope of the obstacle process $\big(\mathrm{e}^{-rt_k}\, \psi_k(X_k, Y_k)\big)_{0 \leqslant k \leqslant n}$. In the end, $\mathbb{E}[V_0]$ is the quantity we are interested in. Indeed, $\mathbb{E}[V_0]$ is the price of the Bermudan option whose payoff is ψ_k and is exercisable at dates $\{t_1, \cdots, t_n\}$.

Following what was defined in (3.35), in order to compute $\mathbb{E}[V_0]$, we will need to use the previously defined quantizer of X_k and Y_k: $\widehat{X}_k$ and $\widehat{Y}_k$. Hence, for a given global budget $N = N_{1,0}N_{2,0} + \cdots + N_{1,n}N_{2,n}$, the total number of nodes of the tree by the couple $(\widehat{X}_k, \widehat{Y}_k)_{0 \leqslant k \leqslant n}$, we can approximate the *Backward Dynamic Programming Principle* (3.35) by the following sequence involving the couple $(\widehat{X}_k, \widehat{Y}_k)_{0 \leqslant k \leqslant n}$

$$\begin{cases} \widehat{V}_n = \mathrm{e}^{-rt_n}\, \psi_n(\widehat{X}_n, \widehat{Y}_n), \\ \widehat{V}_k = \max\big(\mathrm{e}^{-rt_k}\, \psi_k(\widehat{X}_k, \widehat{Y}_k), \mathbb{E}[\widehat{V}_{k+1} \mid (\widehat{X}_k, \widehat{Y}_k)]\big), \qquad k = 0, \dots, n-1. \end{cases} \tag{3.36}$$

Remark 3.7. A direct consequence of choosing recursive Markovian Quantization to spatially discretize the problem is that the sequence $(\widehat{X}_k, \widehat{Y}_k)_{0 \leqslant k \leqslant n}$ is Markovian. Hence $(\widehat{V}_k)_{0 \leqslant k \leqslant n}$ defined in (3.36) obeying a *Backward Dynamic Programming Principle* is the Snell envelope of $\big(\mathrm{e}^{-rt_k}\, \psi_k(\widehat{X}_k, \widehat{Y}_k)\big)_{0 \leqslant k \leqslant n}$. This is the main difference with the first approach of [BPP05, BP03], where in there case they only had a pseudo-Snell envelope of $\big(\mathrm{e}^{-rt_k}\, \psi_k(\widehat{X}_k, \widehat{Y}_k)\big)_{0 \leqslant k \leqslant n}$.

Using the discrete feature of the quantizers, (3.36) can be rewritten

$$\begin{cases} \widehat{v}_n(x_{i_1}^n, y_{i_2}^n) = \mathrm{e}^{-rt_n}\, \psi_n(x_{i_1}^n, y_{i_2}^n), & \begin{array}{l} i_1 = 1, \dots, N_{1,n} \\ i_2 = 1, \dots, N_{2,n} \end{array} \\ \widehat{v}_k(x_{i_1}^k, y_{i_2}^k) = \max\Big(\mathrm{e}^{-rt_k}\, \psi_k(x_{i_1}^k, y_{i_2}^k), \displaystyle\sum_{j_1=1}^{N_{1,k+1}} \sum_{j_2=1}^{N_{2,k+1}} \pi_{(i_1,i_2),(j_1,j_2)}^k\, \widehat{v}_{k+1}(x_{j_1}^{k+1}, y_{j_2}^{k+1})\Big), & \begin{array}{l} k = 0, \dots, n-1 \\ i_1 = 1, \dots, N_{1,k} \\ i_2 = 1, \dots, N_{2,k} \end{array} \end{cases} \tag{3.37}$$

where $\pi_{(i_1,i_2),(j_1,j_2)}^k = \mathbb{P}\big(\widehat{X}_{k+1} = x_{j_1}^{k+1}, \widehat{Y}_{k+1} = y_{j_2}^{k+1} \mid \widehat{X}_k = x_{i_1}^k, \widehat{Y}_k = y_{i_2}^k\big)$ is the conditional probability weight given in (3.28). Finally, the approximation of the price of the Bermudan option is given by

$$\mathbb{E}\big[\widehat{v}_0(x_0, \widehat{Y}_0)\big] = \sum_{i=1}^{N_{2,0}} p_i\, \widehat{v}_0(x_0, y_i^0) \tag{3.38}$$

with $p_i = \mathbb{P}\big(\widehat{Y}_0 = y_i^0\big)$ given by (2.11).

Barrier Options. A Barrier option is a path-dependent financial product whose payoff at maturity date T depends on the value of the process X_T at time T and its maximum or minimum over the period $[0, T]$. More precisely, we are interested by options with the following types of payoff h

$$h = f(X_T)\, \mathbb{1}_{\{\sup_{t \in [0,T]} X_t \in I\}} \qquad \textit{or} \qquad h = f(X_T)\, \mathbb{1}_{\{\inf_{t \in [0,T]} X_t \in I\}} \tag{3.39}$$

where I is an unbounded interval of $\mathbb{R}$, usually of the forme $(-\infty, L]$ or $[L, +\infty)$ (L is the barrier) and f can be any vanilla payoff function (Call, Put, Spread, Butterfly, ...).

In this part, we follow the methodology initiated in [Sag10] in the case of functional quantization. This work is based on the Brownian bridge method applied to the Euler-Maruyama scheme as described e.g. in [Pag18]. We generalize it to stochastic volatility models and product Markovian recursive quantization. X_t being discretized by an Euler-Maruyama scheme, yielding

$\bar{X}_k$ with $k = 0, \ldots, n$, we can determine the law of $\max_{t\in[0,T]} \bar{X}_t$ and $\min_{t\in[0,T]} \bar{X}_t$ given the values $\bar{X}_k = x_k, \bar{Y}_k = y_k, k = 0, \ldots, n$

$$\mathcal{L}\Big(\max_{t\in[0,T]} \bar{X}_t \mid \bar{X}_k = x_k, \bar{Y}_k = y_k, k = 0, \ldots, n\Big) = \mathcal{L}\Big(\max_{k=0,\ldots,n-1} \big(G^k_{(x_k,y_k),x_{k+1}}\big)^{-1}(U_k)\Big) \quad (3.40)$$

and

$$\mathcal{L}\Big(\min_{t\in[0,T]} \bar{X}_t \mid \bar{X}_k = x_k, \bar{Y}_k = y_k, k = 0, \ldots, n\Big) = \mathcal{L}\Big(\max_{k=0,\ldots,n-1} \big(F^k_{(x_k,y_k),x_{k+1}}\big)^{-1}(U_k)\Big) \quad (3.41)$$

where $(U_k)_{k=0,\ldots,n-1}$ are i.i.d uniformly distributed random variables over the unit interval and $(G^k_{(x,y),z})^{-1}$ and $(F^k_{(x,y),z})^{-1}$ are the inverse of the conditional distribution functions $G^k_{(x,y),z}$ and $F^k_{(x,y),z}$ defined by

$$G^k_{(x,y),z}(u) = \Big(1 - \mathrm{e}^{-2n\frac{(x-u)(z-u)}{T\sigma^2(t_k,x,y)}}\Big) \mathbb{1}_{\{u\geqslant \max(x,z)\}} \quad (3.42)$$

and

$$F^k_{(x,y),z}(u) = 1 - \Big(1 - \mathrm{e}^{-2n\frac{(x-u)(z-u)}{T\sigma^2(t_k,x,y)}}\Big) \mathbb{1}_{\{u\leqslant \min(x,z)\}}\,. \quad (3.43)$$

Now, using the resulting representation formula for $\mathbb{E}\, f(\bar{X}_T, \max_{t\in[0,T]} \bar{X}_t)$ (see e.g. [Sag10, Pag18]), we have a new representation formula for the price of up-and-out options $\bar{P}_{UO}$ and down-and-out options $\bar{P}_{DO}$

$$\bar{P}_{UO} = \mathrm{e}^{-rT}\, \mathbb{E}\big[f(\bar{X}_T)\, \mathbb{1}_{\sup_{t\in[0,T]} \bar{X}_t \leqslant L}\big] = \mathrm{e}^{-rT}\, \mathbb{E}\Big[f(\bar{X}_T) \prod_{k=0}^{n-1} G^k_{(\bar{X}_k,\bar{Y}_k),\bar{X}_{k+1}}(L)\Big] \quad (3.44)$$

and

$$\bar{P}_{DO} = \mathrm{e}^{-rT}\, \mathbb{E}\big[f(\bar{X}_T)\, \mathbb{1}_{\inf_{t\in[0,T]} \bar{X}_t \geqslant L}\big] = \mathrm{e}^{-rT}\, \mathbb{E}\Big[f(\bar{X}_T) \prod_{k=0}^{n-1} \Big(1 - F^k_{(\bar{X}_k,\bar{Y}_k),\bar{X}_{k+1}}(L)\Big)\Big] \quad (3.45)$$

where L is the barrier.

Finally, replace $\bar{X}_k$ and $\bar{Y}_k$ by $\widehat{X}_k$ and $\widehat{Y}_k$ and apply the recursive algorithm in order to approximate $\bar{P}_{UO}$ or $\bar{P}_{DO}$ by $\mathbb{E}[\widehat{V}_0]$ or equivalently $\mathbb{E}[\widehat{v}_0(x_0, \widehat{Y}_0)]$

$$\begin{cases} \widehat{V}_n = \mathrm{e}^{-rT} f(\widehat{X}_n), \\ \widehat{V}_k = \mathbb{E}\big[g_k(\widehat{X}_k, \widehat{Y}_k, \widehat{X}_{k+1})\widehat{V}_{k+1} \mid (\widehat{X}_k, \widehat{Y}_k)\big], \qquad 0 \leqslant k \leqslant n-1 \end{cases} \quad (3.46)$$

that can be rewritten

$$\begin{cases} \widehat{v}_n(x^n_{i_1}, y^n_{i_2}) = \mathrm{e}^{-rT} f(x^n_i), & \begin{array}{l} i = 1, \ldots, N_{1,n} \\ j = 1, \ldots, N_{2,n} \end{array} \\ \widehat{v}_k(x^k_{i_1}, y^n_{i_2}) = \displaystyle\sum_{j_1=1}^{N_{1,k+1}} \sum_{j_2=1}^{N_{2,k+1}} \pi^k_{(i_1,i_2),(j_1,j_2)} \widehat{v}_{k+1}(x^{k+1}_{j_1}, y^{k+1}_{j_2}) g_k(x^k_{i_1}, y^k_{i_2}, x^{k+1}_{j_1}), & \begin{array}{l} k = 0, \ldots, n-1 \\ i = 1, \ldots, N_{1,k} \\ j = 1, \ldots, N_{2,k} \end{array} \end{cases} \quad (3.47)$$

with $\pi^k_{(i_1,i_2),(j_1,j_2)} = \mathbb{P}\big(\widehat{X}_{k+1} = x^{k+1}_{j_1}, \widehat{Y}_{k+1} = y^{k+1}_{j_2} \mid \widehat{X}_k = x^k_{i_1}, \widehat{Y}_k = y^k_{i_2}\big)$ the conditional probabilities given in (3.28) and $g_k(x, y, z)$ is either equal to $G^k_{(x,y),z}(L)$ or $1 - F^k_{(x,y),z}(L)$ depending on the option type. Finally, the approximation of the price of the barrier option is given by

$$\mathbb{E}[\widehat{V}_0] = \mathbb{E}\big[\widehat{v}_0(x_0, \widehat{Y}_0)\big] = \sum_{i=1}^{N_{2,0}} p_i\, \widehat{v}_0(x_0, y^0_i) \quad (3.48)$$

with $p_i = \mathbb{P}\big(\widehat{Y}_0 = y^0_i\big)$ given by (2.11).

3.4 Numerical illustrations

In this part, we deal with numerical experiments[4] in the Stationary Heston model. Throughout this section we implement the recursive quantization algorithm with size n (the number of time steps) and quantization grids of constant size: for all $k = 0, \dots, n$, we take $N_{1,k} = N_1$ and $N_{2,k} = N_2$ where n is the number of time steps. First, we use the European options framework to numerically evaluate the errors of recursive quantization as a function of quantization grids sizes. Then, we evaluate Bermudan options and Barrier options.

Impact of N_1, N_2 and n. In this paragraph, we want to evaluate the impact that have the tree parameters N_1, N_2 and n on the accuracy of the prices computed by product recursive quantization. For that purpose, we consider pricing of European options in the Stationary Heston model with parameters given in Table 1 (obtained after the penalized calibration procedure). We set $S_0 = 100$ in order to get prices of an order we are used to. The benchmark prices of 10 European options with various strikes K: 5 calls and 5 puts, all in-the-money, with maturity $t_n = T = 0.5$ (6 months) are computed using the quadrature formula based on Laguerre polynomials as explained in Section 2. At this point let us emphasize that we are aware that neither quantization nor Monte Carlo should be used in practice to compute premia of Eropean options in a Heston model.

In place of using the backward algorithm (3.36) (without the max function) for computing the expectation at the expiry date, we use the weights $p^n_{(i_1,i_2)}$ defined in (3.25) and computed by a forward induction, in order to use the quantization based cubature formula

$$\mathbb{E}\left[\psi_n(\widehat{X}_n, \widehat{Y}_n)\right] = \sum_{i_1=1}^{N_1} \sum_{i_2=1}^{N_2} \psi_n(x^n_{i_1}, y^n_{i_2}) p^n_{(i_1,i_2)}. \tag{3.49}$$

The results are reported in Tables 2 and 3. The relative error induced by the quantization-based approximation is given between brackets. The computation time is split into the time taken to construct the tree and the pricing time (Equation (3.4)).

First, we compare in Table 2 the prices for a fixed number of time-steps: $n = 180$ and we make vary the parameters N_1 and N_2. Then in Table 3 we fix $(N_1, N_2) = (50, 10)$ and make n vary.

We notice from Table 3 that the error is mostly explained by the number of the time-steps n. By comparison, a Monte-Carlo pricer with $M = 300,000$ paths reaches the same order of precision than our quantization tree with $(N_1, N_2) = (100, 10)$ and $n = 180$ in $3.3s$. Once the quantification tree is built, the price calculation is extremely fast and if we compare it with the Monte Carlo estimator we have a ratio of $3.3/0.173 \simeq 19$ in calculation speed. Note also that it is much more efficient to store the quantization tree than to store all the paths of a Monte Carlo estimator.

[4]All the numerical tests have been carried out in `C++` on a laptop with a 2,4 GHz 8–Core Intel Core i9 CPU. The computations of the transition probabilities are parallelized on the CPU.

			(N_1, N_2)			
	K	Benchmark	$(20, 5)$	$(50, 10)$	$(100, 10)$	$(150, 10)$
Call	80	20.17	19.71 (2.2%)	20.01 (0.8%)	20.07 (0.50%)	20.09 (0.42%)
	85	15.56	15.01 (3.5%)	15.38 (1.14%)	15.45 (0.70%)	15.46 (0.60%)
	90	11.24	10.64 (5.3%)	11.06 (1.58%)	11.13 (0.92%)	11.15 (0.77%)
	95	7.383	6.825 (7.5%)	7.233 (2.02%)	7.314 (0.93%)	7.331 (0.70%)
	100	4.196	3.769 (10.1%)	4.110 (2.06%)	4.196 (0.01%)	4.215 (0.44%)
Put	100	4.469	4.167 (6.75%)	4.402 (1.48%)	4.455 (0.31%)	4.464 (0.10%)
	105	7.171	7.031 (1.96%)	7.177 (0.08%)	7.234 (0.86%)	7.243 (0.99%)
	110	10.86	10.81 (0.39%)	10.90 (0.36%)	10.95 (0.87%)	10.96 (0.93%)
	115	15.38	15.39 (0.10%)	15.38 (0.005%)	15.41 (0.24%)	15.42 (0.26%)
	120	20.30	20.39 (0.43%)	20.29 (0.08%)	20.27 (0.17%)	20.27 (0.17%)
	Time		1.7s + 5ms	16.5s + 30ms	81s + 173ms	186s + 406ms

Table 2: *Comparison between European options prices, with maturity $T = 0.5$ (6 months), given by quantization and the benchmark, as a function of the strike K and (N_1, N_2) where we set $n = 180$. Relative error in percent between brackets.*

			n			
	K	Benchmark	30	60	90	180
Call	80	20.17	20.01 (0.82%)	20.04 (0.67%)	20.03 (0.68%)	20.01 (0.8%)
	85	15.56	15.33 (1.44%)	15.39 (1.07%)	15.40 (1.02%)	15.38 (1.14%)
	90	11.24	10.95 (2.58%)	11.04 (1.74%)	11.06 (1.54%)	11.06 (1.58%)
	95	7.383	7.047 (4.55%)	7.174 (2.82%)	7.211 (2.32%)	7.233 (2.02%)
	100	4.196	3.883 (7.46%)	4.021 (4.16%)	4.066 (3.09%)	4.110 (2.06%)
Put	100	4.469	4.161 (6.88%)	4.305 (3.65%)	4.356 (2.52%)	4.402 (1.48%)
	105	7.171	6.969 (2.81%)	7.080 (1.27%)	7.125 (0.64%)	7.177 (0.08%)
	110	10.86	10.81 (0.45%)	10.85 (0.09%)	10.86 (0.08%)	10.90 (0.36%)
	115	15.38	15.38 (0.04%)	15.38 (0.01%)	15.38 (0.04%)	15.38 (0.005%)
	120	20.30	20.28 (0.10%)	20.28 (0.12%)	20.29 (0.09%)	20.29 (0.08%)
	Time		3.6s + 30ms	6.5s + 30ms	9s + 30ms	16.5s + 30ms

Table 3: *Comparison between European options prices, with maturity $T = 0.5$ (6 months), given by quantization and the benchmark, as a function of the strike K and of the number n of time-steps where we set $(N_1, N_2) = (50, 10)$. Relative error in percent between brackets.*

Bermudan options. For Bermudan options, we do not use the parameters we calibrated in subsection 2.2 because the interest rate and the dividend rate are very close to 0. Hence the prices of Bermudan options with those parameters are equal to European ones. In place we use the parameter set No.5 defined in [FO11] which is a set of parameters commonly used in the financial literature for the pricing of Bermudan options that we recall in Table 4.

ρ	θ	κ	ξ	S_0	r	q
-0.64	0.0348	1.15	0.39	100	0.04	0

Table 4: *Test set No.5 from [FO11].*

Moreover, for the Bermudan Put we will use its European price as a control variate. This has to be compared to including the payoff function into the regression basis at each time-steps in regression-projection methods "à la Longstaff-Schwarz". The *Backward Dynamic Programming Principle* becomes

$$\begin{cases} \widetilde{V}_n = 0, \\ \widetilde{V}_k = \max\big(\mathrm{e}^{-rt_k}\,\psi_k(X_k, Y_k) - C^{EU}_{t_n - t_k}(\mathrm{e}^{X_k}), \mathbb{E}[\widetilde{V}_{k+1} \mid \mathcal{F}_k]\big), \qquad 0 \leqslant k \leqslant n-1, \end{cases} \tag{3.50}$$

where $C^{EU}_t(x)$ is the European price for maturity t with initial stock price x in the Stationary Heston model (see [BPP05] for more details). Finally, the price of the Bermudan option V_0 is equal to

$$V_0 = \widetilde{V}_0 + C^{EU}_{t_n}(S_0)$$

In Table 5, we display the prices obtain using product hybrid recursive quantization of Bermudan Call/Put with n equally spaced exercise dates, strike $K = 100$ and maturity $T = 0.25$. The prices of the European Call and Put are 3.6729 and 2.6779 respectively. QCall and QPut are the prices given by quantization without control variate and QPutCV are the prices obtained using the control variate as detailed in Equation (3.50).

n		20			40			60		
N_1	N_2	QCall	QPut	QPutCV	QCall	QPut	QPutCV	QCall	QPut	QPutCV
10	5	3.8048	2.9741	2.7335	3.6829	2.9529	2.7329	3.6750	3.0341	2.7266
20	10	3.7693	2.8613	2.7371	3.8152	2.9447	2.7383	3.7944	2.9649	2.7382
50	20	3.6997	2.7732	2.7377	3.7338	2.8156	2.7401	3.6811	2.8236	2.7402
100	20	3.6908	2.7635	2.7374	3.6447	2.7553	2.7384	3.6417	2.7767	2.7387
200	20	3.6782	2.7602	2.7375	3.6470	2.7488	2.7382	3.6484	2.7704	2.7384

Table 5: *Pricing of Bermudan options with maturity $T = 0.25$ and strike $K = 100$ in the Stationary Heston model with product hybrid recursive quantization.*

As $r > 0$, the Bermudan Call price is equal to the European one, namely 3.6729. Hence, we display the prices only for the sake of studying the convergence of the methodology. We notice in Table 5 that QCall converges to the European price but, as prescribed by theoretical error bounds, the more timesteps n we add in the quantization tree, the larger the grids need to be. For the Put, the methodology without control variate converges as well but control variate dramatically reduces the size of the grid needed in order to converge.

We are aware that the theoretical assumptions to obtain error bounds for quantization based space discretization schemes of the Milstein scheme (see [PS20]) are not satisfied by the volatility process due to the presence of a square root. However, considering larger grids allow overcome this problem in practice. Indeed we provide error bound for our model in a supplementary material available in arXiv which guided our choice for the ratio grid size of the volatility over grid size of the asset for a given time discretion time step (see also [Mon20]).

Heston vs Stationary Heston. In Figure 8, we compare the prices of Bermudan Call / Put with maturity $T = 0.25$ in the Heston model and the Stationary Heston model where both prices are given by product hybrid recursive quantization with $N_1 = 50$, $N_2 = 20$ and $n = 40$. Our aim is just to highlight the prices both models differ significantly, especially for short maturities. In order to compare the prices, we compute the relative difference between the 2 prices, namely $(P_{SH} - P_H)/P_H$ where P_H is the price in the Heston model and P_{SH} is the price in the Stationary Heston model. We used the same values given in Table 4 for the parameters in common and set $v_0 = \theta$ in the case of the Heston model. This means that we took as a starting value for the volatility of the Heston its long-term mean. This confirms that, as expected, both models yield different prices. To be more precise, prices in the Stationary Heston model are higher on the tails and lower close to moneyness.

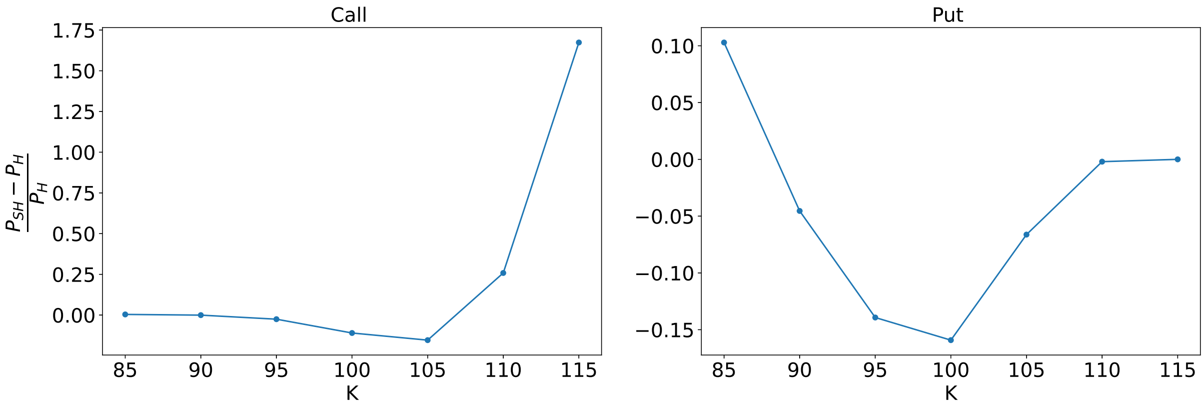

Figure 8: *Relative difference of Bermudan prices between the Heston model and the Stationary Heston model.*

Barrier options. Finally, we apply our recursive quantization algorithm for a path-dependent European option pricing. The parameters of the Stationary Heston model are given in Table 1 and $S_0 = 100$. The path-dependent option is an up-and-out Barrier option with strike $K = 100$, maturity $T = 0.5$ (6 months), barrier $L = 115$ and $N_2 = 10$. In Figure 9, we display the prices obtained for a fixed value $N_2 = 10$ and differents values of N_1 and n. Again, we can notice the impact of n on the approximated price.

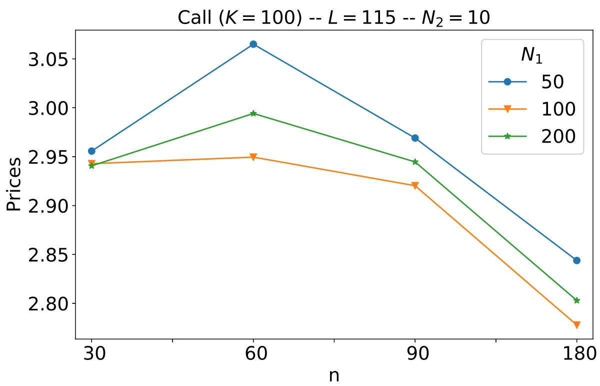

Figure 9: *Prices of a Barrier option in the Stationary Heston model given by product hybrid recursive quantization with fixed value* $N_2 = 10$.

Acknowledgment. The authors wish to thank Guillaume Aubert for fruitful discussion on the Heston model and Jean-Michel Fayolle for his advice on the calibration of the models. The PhD thesis of Thibaut Montes is funded by a CIFRE grand from The Independent Calculation Agent (The ICA) and French ANRT. The first an third author benefited for the support of the “Chaire Risques financiers” funded by the Société Générale, under the aegis of the Fondation du Risque.

References

[Alf05] Aurélien Alfonsi. On the discretization schemes for the CIR (and bessel squared) processes. *Monte Carlo Methods and Applications mcma*, 11(4):355–384, 2005.

[AMST07] Hansjörg Albrecher, Philipp Arnold Mayer, Wim Schoutens, and Jurgen Tistaert. The little Heston trap. *Wilmott*, (1):83–92, 2007.

[And65] Donald G Anderson. Iterative procedures for nonlinear integral equations. *Journal of the ACM*, 12(4):547–560, 1965.

[And07] Leif BG Andersen. Efficient simulation of the Heston stochastic volatility model. *SSRN Electronic Journal*, 2007.

[BP03] Vlad Bally and Gilles Pagès. A quantization algorithm for solving multidimensional discrete-time optimal stopping problems. *Bernoulli*, 9(6):1003–1049, 2003.

[BPP05] Vlad Bally, Gilles Pagès, and Jacques Printems. A quantization tree method for pricing and hedging multi-dimensional american options. *Mathematical Finance*, 15(1):119–168, 2005.

[CFG17] Giorgia Callegaro, Lucio Fiorin, and Martino Grasselli. Pricing via recursive quantization in stochastic volatility models. *Quantitative Finance*, 17(6):855–872, 2017.

[CFG18] Giorgia Callegaro, Lucio Fiorin, and Martino Grasselli. American quantized calibration in stochastic volatility. *Risk Magazine*, 2018.

[CGP21] Giorgia Callegaro, Martino Grasselli, and Gilles Pagès. Fast hybrid schemes for fractional Riccati equations (rough is not so tough). *Mathematics of Operations Research*, 46(1):221–254, 2021.

[CIJR05] John C Cox, Jonathan E Ingersoll Jr, and Stephen A Ross. A theory of the term structure of interest rates. In *Theory of Valuation*, pages 129–164. World Scientific, 2005.

[CM99] Peter Carr and Dilip Madan. Option valuation using the fast fourier transform. *Journal of computational finance*, 2(4):61–73, 1999.

[FO11] Fang Fang and Cornelis W Oosterlee. A fourier-based valuation method for Bermudan and barrier options under Heston's model. *SIAM Journal on Financial Mathematics*, 2(1):439–463, 2011.

[FSP18] Lucio Fiorin, Abass Sagna, and Gilles Pagès. Product markovian quantization of a diffusion process with applications to finance. *Methodology and Computing in Applied Probability*, pages 1–32, 2018.

[Gat11] Jim Gatheral. *The volatility surface: a practitioner's guide*, volume 357. John Wiley & Sons, 2011.

[GJR18] Jim Gatheral, Thibault Jaisson, and Mathieu Rosenbaum. Volatility is rough. *Quantitative Finance*, 18(6):933–949, 2018.

[GJRS18] Hamza Guennoun, Antoine Jacquier, Patrick Roome, and Fangwei Shi. Asymptotic behavior of the fractional Heston model. *SIAM Journal on Financial Mathematics*, 9(3):1017–1045, 2018.

[GL00] Siegfried Graf and Harald Luschgy. *Foundations of Quantization for Probability Distributions.* Springer-Verlag, Berlin, Heidelberg, 2000.

[GR09] Pierre Gauthier and Pierre-Yves Henri Rivaille. Fitting the smile, smart parameters for SABR and Heston. *SSRN Electronic Journal*, 2009.

[GR19] Jim Gatheral and Rados Radoicic. Rational approximation of the rough Heston solution. *International Journal of Theoretical and Applied Finance*, 22(3):1950010, 2019.

[Hes93] Steven L Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The review of financial studies*, 6(2):327–343, 1993.

[IW81] Nobuyuki Ikeda and Shinzo Watanabe. *Stochastic differential equations and diffusion processes*, volume 24. North Holland, 1981.

[JR16] Thibault Jaisson and Mathieu Rosenbaum. Rough fractional diffusions as scaling limits of nearly unstable heavy tailed hawkes processes. *The Annals of Applied Probability*, 26(5):2860–2882, 2016.

[JS17] Antoine Jacquier and Fangwei Shi. The randomized Heston model. *SIAM Journal on Financial Mathematics*, 10(1):89–129, 2017.

[Kie82] John C Kieffer. Exponential rate of convergence for lloyd's method i. *IEEE Transactions on Information Theory*, 28(2):205–210, 1982.

[LL11] Damien Lamberton and Bernard Lapeyre. *Introduction to stochastic calculus applied to finance*. Chapman and Hall/CRC, 2011.

[Llo82] Stuart Lloyd. Least squares quantization in pcm. *IEEE transactions on information theory*, 28(2):129–137, 1982.

[LMP20] Vincent Lemaire, Thibaut Montes, and Gilles Pagès. Stationary Heston model: Calibration and pricing of exotics using product recursive quantization, 2020.

[Mon20] Thibaut Montes. *Numerical methods by optimal quantization in finance*. Theses, Sorbonne Université, 2020.

[MRKP18] Thomas A McWalter, Ralph Rudd, Jörg Kienitz, and Eckhard Platen. Recursive marginal quantization of higher-order schemes. *Quantitative Finance*, 18(4):693–706, 2018.

[NM65] John A Nelder and Roger Mead. A simplex method for function minimization. *The computer journal*, 7(4):308–313, 1965.

[Pag18] Gilles Pagès. *Numerical Probability: An Introduction with Applications to Finance*. Springer, 2018.

[PP09] Gilles Pagès and Fabien Panloup. Approximation of the distribution of a stationary markov process with application to option pricing. *Bernoulli*, 15(1):146–177, 2009.

[PPP04] Gilles Pagès, Huyên Pham, and Jacques Printems. *Optimal Quantization Methods and Applications to Numerical Problems in Finance*, pages 253–297. Birkhäuser Boston, 2004.

[PS15] Gilles Pagès and Abass Sagna. Recursive marginal quantization of the euler scheme of a diffusion process. *Applied Mathematical Finance*, 22(5):463–498, 2015.

[PS20] Gilles Pagès and Abass Sagna. Weak and strong error analysis of recursive quantization: a general approach with an application to jump diffusions. *IMA Journal of Numerical Analysis*, 09 2020. draa033.

[PY16] Gilles Pagès and Jun Yu. Pointwise convergence of the lloyd i algorithm in higher dimension. *SIAM Journal on Control and Optimization*, 54(5):2354–2382, 2016.

[RMKP17] Ralph Rudd, Thomas McWalter, Jörg Kienitz, and Eckhard Platen. Fast quantization of stochastic volatility models. *SSRN Electronic Journal*, 2017.

[Sag10] Abass Sagna. Pricing of barrier options by marginal functional quantization. *Monte Carlo Methods and Applications*, 17(4):371–398, 2010.

[SST04] Wim Schoutens, Erwin Simons, and Jurgen Tistaert. A perfect calibration! now what? *Wilmott Magazine*, 2004.

[Tru82] Alexander V. Trushkin. Sufficient conditions for uniqueness of a locally optimal quantizer for a class of convex error weighting functions. *IEEE Trans. Inform. Theory*, 28(2):187–198, 1982.

[WN11] Homer F Walker and Peng Ni. Anderson acceleration for fixed-point iterations. *SIAM Journal on Numerical Analysis*, 49(4):1715–1735, 2011.

[Zad82] Paul Zador. Asymptotic quantization error of continuous signals and the quantization dimension. *IEEE Transactions on Information Theory*, 28(2):139–149, 1982.

Appendices

A Discretization scheme of the volatility preserving the positivity

We recall the dynamics of the volatility

$$dv_t = \kappa(\theta - v_t)dt + \xi\sqrt{v_t}d\widetilde{W}_t$$

with $\kappa > 0$, $\theta > 0$ and $\xi > 0$. In this section, we discuss the choice of the discretization scheme under the Feller condition, which ensures the positivity of the process.

Euler-Maruyama scheme. Let $(Z_k)_{k\geqslant 1}$ an i.i.d. sequence of standard normal distribution. Discretizing the volatility using an Euler-Maruyama scheme

$$\bar{v}_{t_{k+1}} = \bar{v}_{t_k} + \kappa(\theta - \bar{v}_{t_k})h + \xi\sqrt{\bar{v}_{t_k}}\sqrt{h}\, Z_{k+1}$$

with $t_k = kh$, $h = T/n$, may look natural. However, such a scheme clearly does not preserve positivity of the process even if the Feller condition is fulfilled since

$$\mathbb{P}\left(\bar{v}_{t_1} < 0\right) = \mathbb{P}\left(Z_1 < \frac{-v_0 - \kappa(\theta - v_0)h}{\xi\sqrt{v_0}\sqrt{h}}\right) > 0.$$

This suggests to introduce the Milstein scheme which is quite tractable in one dimension in absence of Lévy areas.

Milstein scheme and boosted Milstein scheme. The Milstein scheme of the stochastic volatility is given by

$$\bar{v}_{t_{k+1}} = \mathcal{M}_{b,\sigma}\big(t_k, \bar{v}_{t_{k+1}}, Z_{k+1}\big)$$

where (see (3.4))

$$\mathcal{M}_{b,\sigma}(t,x,z) = x - \frac{\sigma(x)}{2\sigma'_x(x)} + h\left(b(t,x) - \frac{(\sigma\sigma'_x)(x)}{2}\right) + \frac{(\sigma\sigma'_x)(x)h}{2}\left(z + \frac{1}{\sqrt{h}\sigma'_x(x)}\right)^2$$

with $b(x) = \kappa(\theta - x)$, $\sigma(x) = \xi\sqrt{x}$ and $\sigma'_x(x) = \frac{\xi}{2\sqrt{x}}$. Consequently, under the Feller condition, the positivity of $\mathcal{M}_{b,\sigma}(t,x,z)$ is ensured if

$$x \geqslant \frac{\sigma(x)}{2\sigma'_x(x)} \geqslant 0, \qquad b(t,x) \geqslant \frac{(\sigma\sigma'_x)(x)}{2} \geqslant 0.$$

In our case, if the first condition holds as an equally since $\frac{\sigma(x)}{2\sigma'_x(x)} = \frac{\xi\sqrt{x}}{2\frac{\xi}{2\sqrt{x}}} = x$ the second one fails. But

$$\frac{(\sigma\sigma'_x)(x)}{2} = \frac{\xi\sqrt{x}\frac{\xi}{2\sqrt{x}}}{2} = \frac{\xi^2}{4}$$

can be bigger than $b(t,x)$. To solve this problem, we consider the following *κ-boosted* volatility process

$$Y_t = \mathrm{e}^{\kappa t}\, v_t, \quad t \in [0,T], \tag{A.1}$$

which satisfies the stochastic differnetial equation

$$dY_t = \mathrm{e}^{\kappa t}\,\kappa\theta dt + \xi\,\mathrm{e}^{\kappa t/2}\sqrt{Y_t}d\widetilde{W}_t.$$

Remark A.1. Note that the the process $(Y_t)_{t\in[0,T]}$ will have a higher variance but, having in mind a quantized scheme, this has no real impact (by contrast with a Monte Carlo simulation).

Now, if we look at the Milstein discretization scheme of Y_t

$$\bar{Y}_{t_{k+1}} = \mathcal{M}_{\widetilde{b},\widetilde{\sigma}}\big(t_k, \bar{Y}_{t_k}, Z_{k+1}\big)$$

using the notation defined in (3.4) where drift and volatility term are now time-dependents and given by

$$\widetilde{b}(t,x) = \mathrm{e}^{\kappa t}\,\kappa\theta, \qquad \widetilde{\sigma}(t,x) = \xi\sqrt{x}\,\mathrm{e}^{\kappa t/2} \quad \text{ and } \quad \widetilde{\sigma}'_x(t,x) = \frac{\xi\,\mathrm{e}^{\kappa t/2}}{2\sqrt{x}}.$$

Under the Feller condition, the positivity of the above scheme is ensured, since

$$\frac{\widetilde{\sigma}(t,x)}{2\widetilde{\sigma}'_x(t,x)} = x \qquad \text{and} \qquad \frac{(\widetilde{\sigma}\widetilde{\sigma}'_x)(t,x)}{2} = \frac{\xi^2\,\mathrm{e}^{\kappa t}}{4} \leqslant \widetilde{b}(t,x) = \mathrm{e}^{\kappa t}\,\kappa\theta.$$

The last inequality is satisfied thanks to the condition $\frac{\xi^2}{2\kappa\theta} \leqslant 1$ ensuring the positivity of the scheme.

B Quadratic Optimal Quantization: Generic Approach

Let X be a $\mathbb{R}$-valued random variable with distribution $\mathbb{P}_X$ defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ such that $X \in L^2_{\mathbb{R}}(\Omega, \mathcal{A}, \mathbb{P})$.

Let $\Gamma_N = \{x_1^N, \dots, x_N^N\} \subset \mathbb{R}$ be a subset of size N indexed in an increasing order, called an N-quantizer. Any such N-quantizer is in a bijective correspondence with the N-tuple $x = (x_1^N, \dots, x_N^N)$lying in the simplex of $\mathbb{R}^N$. This leads us to sometimes replace Γ_N by x.

The *Voronoi partition* $(C_i(\Gamma_N))_{i=1,\dots,N}$ induced by Γ_N in such a one-dimensional setting is defined by

$$C_i(\Gamma_N) = \big(x^N_{i-1/2}, x^N_{i+1/2}\big], \qquad i \in \{1, \dots, N-1\}, \qquad C_N(\Gamma_N) = \big(x^N_{N-1/2}, x^N_{N+1/2}\big) \tag{B.1}$$

where $\forall i \in \{2, \dots, N\}, x^N_{i-1/2} = \frac{x^N_{i-1}+x^N_i}{2}$ and $x^N_{1/2} = -\infty$ and $x^N_{N+1/2} = +\infty$. Elements of the cell are closer to x_i than to any other x_j (up the endpoints).

The Voronoi quantization of X induced by Γ_N, denoted $\widehat{X}^{\Gamma_N}$ or $\widehat{X}^N$ when ambiguity, is defined as the nearest neighbour projection of X onto Γ_N based onto above Voronoi cells, namely

$$\widehat{X}^N = \mathrm{Proj}_{\Gamma_N}(X) = \sum_{i=1}^{N} x_i^N\,\mathbb{1}_{X\in C_i(\Gamma_N)}\,. \tag{B.2}$$

The probability distribution of $\widehat{X}^N$ (also know as weights) are given by

$$\mathbb{P}\left(\widehat{X}^N = x_i^N\right) = \mathbb{P}_X\left(C_i(\Gamma_N)\right) = \mathbb{P}\left(X \in \big(x^N_{i-1/2}, x^N_{i+1/2}\big]\right).$$

To measure the adequacy of a grid to the distribution of X, one introduces the *quadratic distortion* function at level N given by

$$\mathcal{Q}_{2,\mathrm{N}} : x = (x_1, \dots, x_N) \longmapsto \tfrac{1}{2}\,\mathbb{E}\Big[\min_{i\in\{1,\dots,N\}} |X - x_i^N|^2\Big] = \tfrac{1}{2}\,\mathbb{E}\big[\mathrm{dist}(X, \Gamma_N)^2\big] = \frac{1}{2}\|X - \widehat{X}^N\|_2^2.$$

Of course, the above result can be extended to the L^p case. We briefly recall some classical theoretical results, see [GL00, Pag18] for further details. The first one treats of existence of optimal quantizers.

Theorem B.1. *(Existence of optimal N-quantizers) Let $X \in L^2_{\mathbb{R}}(\mathbb{P})$ and $N \in \mathbb{N}^\star$.*

(a) The quadratic distortion function $\mathcal{Q}_{2,\mathrm{N}}$ at level N attains a minimum at a N-tuple $x^\star = (x_1^N, \dots, x_N^N)$ and $\Gamma_N^\star = \big\{x_i^N,\, i \in \{1, \dots, N\}\big\}$ is a quadratic optimal quantizer at level N.

(b) If the support of the distribution $\mathbb{P}_X$ of X has at least N elements, then $x^\star = (x_1^N, \ldots, x_N^N)$ has pairwise distinct components, $\mathbb{P}_X\big(C_i(\Gamma_N^\star)\big) > 0$, $i \in \{1, \ldots, N\}$. Furthermore, the sequence $N \mapsto \inf_{x\in(\mathbb{R})^N} \mathcal{Q}_{2,\mathrm{N}}(x)$ converges to 0 and is decreasing as long as it is positive.

Uniqueness of an optimal N-quantizer (defined as an N-tuple with increasing components), is due to Truskin [Tru82] was shown in dimension one under log-concavity assumptions on the density of the distribution $\mathbb{P}_X$ of X.

In what follows, we will drop the * subscript when speaking of optimal quantizers, $x^\star$ and $\Gamma_N^\star$ will be replaced by x and Γ_N.

The next result elucidates the asymptotic behavior of the distortion. We saw in Theorem B.1 that the infimum of the quadratic distortion converges to 0 as N goes to infinity. The next theorem, known as Zador's Theorem, establishes the sharp rate of convergence of the quantization error (stated here in the quadratic case).

Theorem B.2. *(Zador's Theorem) (a)* Sharp rate [Zad82, GL00]. *Let $X \in L^{p+\delta}_{\mathbb{R}}(\mathbb{P})$ for some $\delta > 0$. Let $\mathbb{P}_X(d\xi) = \varphi(\xi)\cdot\lambda(d\xi) + \nu(d\xi)$, where ν is singular with respect to the Lebesgue measure λ on $\mathbb{R}$.*

$$\lim_{N\to+\infty} N \min_{\Gamma_N\subset\mathbb{R}, |\Gamma_N|\leqslant N} \| X - \widehat{X}^N \|_2 = \frac{1}{\sqrt{12}} \left[\int_{\mathbb{R}} \varphi^{\frac{1}{3}} d\lambda \right]^{\frac{3}{4}}. \tag{B.3}$$

(b) Non asymptotic upper-bound [GL00, Pag18]. *Let $\delta > 0$. There exists a real constant $C_{1,p} \in (0, +\infty)$ such that, for every $\mathbb{R}$-valued random variable X,*

$$\forall N \geqslant 1, \qquad \min_{\Gamma_N\subset\mathbb{R}, |\Gamma_N|\leqslant N} \| X - \widehat{X}^N \|_2 \leqslant C_{1,p}\sigma_{2+\delta}(X) N^{-1} \tag{B.4}$$

where, for $r \in (0, +\infty), \sigma_r(X) = \min_{a\in\mathbb{R}} \|X - a\|_r < +\infty$ is the L^r-pseudo-standard deviation.

To compute optimal quantizers, we first differentiate $\mathcal{Q}_{2,\mathrm{N}}$, whose gradient is given by

$$\nabla\, \mathcal{Q}_{2,\mathrm{N}}(x) = \left(\mathbb{E}\left[(x_i^N - X)\, \mathbb{1}_{X\in\left(x_{i-1/2}^N, x_{i+1/2}^N\right]} \right] \right)_{i=1,\ldots,N}. \tag{B.5}$$

Then any critical point of $\mathcal{Q}_{2,\mathrm{N}}$, in particular its minima, satisfies

$$\begin{aligned} \nabla\, \mathcal{Q}_{2,\mathrm{N}}(x) = 0 \qquad &\iff \qquad x_i^N = \frac{\mathbb{E}\left[X\, \mathbb{1}_{X\in\left(x_{i-1/2}^N, x_{i+1/2}^N\right]} \right]}{\mathbb{P}\left(X \in \left(x_{i-1/2}^N, x_{i+1/2}^N\right] \right)}, \qquad i = 1, \ldots, N \\ &\iff \qquad x_i^N = \frac{K_X\left(x_{i+1/2}^N\right) - K_X\left(x_{i-1/2}^N\right)}{F_X\left(x_{i+1/2}^N\right) - F_X\left(x_{i-1/2}^N\right)}, \qquad i = 1, \ldots, N \end{aligned} \tag{B.6}$$

where $K_X(\cdot)$ and $F_X(\cdot)$ are the first partial moment and the cumulative distribution respectively, function of X, i.e.

$$K_X(x) = \mathbb{E}\left[X\, \mathbb{1}_{X\leqslant x} \right] \qquad \text{and} \qquad F_X(x) = \mathbb{P}\left(X \leqslant x \right). \tag{B.7}$$

Note equation (B.6) can be rewritten as the following so-called *stationarity equation*

$$\mathbb{E}\left[\widehat{X}^N \mid X \right] = \widehat{X}^N. \tag{B.8}$$

This equality in (B.6) is the starting point to the development of the first method devoted to the numerical computation of optimal quantizers: the Lloyd's method I. This fixed point method was first devised in 1957 by S.P. Lloyd and published later [Llo82]. Starting from a sorted N-tuple $x^{[0]}$ and with the knowledge of the first partial moment K_X and the cumulative distribution function F_X of X, the algorithm, which is essentially a deterministic fixed point method, is defined as follows

$$x_i^{N,[n+1]} = \frac{K_X\left(x_{i+1/2}^{N,[n]}\right) - K_X\left(x_{i-1/2}^{N,[n]}\right)}{F_X\left(x_{i+1/2}^{N,[n]}\right) - F_X\left(x_{i-1/2}^{N,[n]}\right)}, \qquad i = 1, \ldots, N. \tag{B.9}$$

In [Kie82], it has been shown that $\left(x^{[n]}\right)_{n\geqslant 1}$ converges exponentially fast toward x, the optimal quantizer, when the density φ of X is log-concave and not log-piecewise affine.

The convergence rate of this fixed point method can be improved e.g. by using Anderson acceleration (see [And65] for the original paper and [WN11] for details on the procedure).

Other algorithms exist, such as, stabbed gradient descent, the Newton-Raphson zero search procedure or its Levenberg-Marquardt variant which are deterministic procedures provided the density, the first partial moment and the cumulative distribution function of X are known. Additionally, we can cite stochastic procedures such as the CLVQ procedure (for Competitive Learning Vector Quantization) which is a zero search stochastic gradient and the randomized version of the Lloyd's method I. For more details, the reader can refer to [Pag18, PY16].

Once the algorithm (B.9) has been converging, we have at hand the quadratic optimal quantizer $\widehat{X}^N$ of X and its associated probabilities given by

$$\mathbb{P}\left(\widehat{X}^N = x_i^n\right) = F_{X}\left(x_{i+1/2}^N\right) - F_{X}\left(x_{i-1/2}^N\right), \qquad i = 1, \ldots, n. \tag{B.10}$$