

A Kernel Density Estimation-Based Approach To Option Pricing

Soufiane Ouamaliche, Awatef Sayah

Abstract: When used in option pricing, a classical Monte Carlo method fails to deliver highly accurate results even when variance reduction techniques are introduced. This lack of accuracy is particularly striking when one is dealing with exotic options. In this paper we aim at improving the quality of the price estimates given by Monte Carlo simulations within the regular Black & Scholes framework, through the use of an approach in which suitable weights are applied to adjust the numerical evaluation of the expected value stated by the Feynman-Kac theorem. Computing the said weights requires the use of an empirical density estimation, namely we will be using a kernel density estimator coupled with various kernels some of which are based on known probability density functions and others based on orthogonal polynomials. The suggested technique was applied to pricing an arithmetic Asian option and the achieved results were compared to prices computed via a classical Monte Carlo procedure, the target price being a well-known numerical solution of the Roger & Shi PDE. Our method has proved its success in providing more accurate prices for all of the test cases which implies that the use of the technique would probably be adequate for a wide variety of high-dimensional pricing problems.

Index Terms: Asian Option, Feynman-Kac theorem, Kernel Density Estimation, Monte Carlo Simulations, Option Pricing, Roger & Shi PDE, Variance Reduction.

1 INTRODUCTION

Most of recent and less recent derivative securities pricing models are usually developed within the framework of the standard Black and Scholes model which often leads to partial differential equations; few of these equations if any have analytical solutions. For instance, one can consider American options [24], which give the investor the possibility to exercise the option at any point during the life of the contract, Bermudan options [11], which can only be exercised on predetermined dates that occur between the purchase date and the expiration date, barrier options [30], whose payoff depends on whether or not the underlying asset has reached or exceeded a pre-set barrier level, lookback options [7], whose payoff depends on the optimal value of the underlying asset and Asian options [27], whose payoff depends on the path taken by the underlying asset's price during a pre-set period of time. The lack of straightforward formulas, except the one for the simplest case which is the European option over an underlying that pays no dividends, has pushed theorists and practitioners toward considering simulations and numerical solutions to these problems. Aside from the binomial lattice method, which considers pricing vanilla options over one state variable, proposed by Cox, Ross and Rubinstein [8], the first serious numerical attempts addressing more exotic options were conducted during the nineties, first a binomial method for diffusion processes other than the geometric Brownian motion was proposed by Nelson and Ramaswamy [25] then trinomial methods were introduced by Hull and White [18]. Regarding simulations they were first introduced in the late seventies [2] and since then they have been applied in a successful way to pricing various options, for a general overview of the use of simulations in security pricing one can consult Boyle, Broadie and Glasserman [3].

In this paper we focus on Monte Carlo simulations combined with kernel density smoothing in order to generate realistic price estimates for various options. Adopting a Monte Carlo technique is mostly motivated by the well established fact that, since its introduction by Boyle [2], it has become a major numerical tool in the world of computational finance, mainly because it delivers reliable results in most cases where all other analytical and numerical methods fail to [12]. The use of Monte Carlo simulation is not only restricted to pricing options, which is the main concern of this paper, but it also extends to simulating hedging strategies, evaluating portfolio management rules and estimating Value at Risk, yet a potential drawback of this method is the computational burden, as a matter of fact a large number of replications is needed in order to refine the estimates and get satisfactory results. In the process of the suggested approach we will need to estimate the unknown density function of a large number of Monte Carlo draws, in order to carry out this estimation and for the sake of having the best possible approximation of the underlying statistical structure of the simulated data we have excluded the parametric methods that force us into making some premises that might hold some inaccuracies related to the dataset. Therefore, we have chosen to use a nonparametric density estimation method, namely, kernel density smoothing, this method is being mostly used for the purpose of graphical data exploration or in the context of nonparametric regression, therefore, and to the best of our knowledge using it in a Monte Carlo framework would be an original approach. The rest of this paper is organized as follows. In section 2 we describe the use of the classical Monte Carlo approach in the context of the Black and Scholes model and the Feynman-Kac theorem then we rapidly survey some variance reduction techniques that turn out to be useful when dealing with derivatives pricing. Next, in section 3 we introduce the modified Monte Carlo technique and we present some basics of kernel density estimation, after that we describe the kernels that we intend to use in order to address the problem at hand. Finally in section 4, we present some numerical experiments where we test the efficiency of the suggested method when pricing options.

- *Soufiane Ouamaliche: Mohammed V University in Rabat, Faculty of Sciences, Laboratory of Mathematics, Computing and Applications, BP1014RP, Rabat, Morocco.*
- *Awatef Sayah: Mohammed V University in Rabat, Faculty of Sciences, Laboratory of Mathematics, Computing and Applications, BP1014RP, Rabat, Morocco.*

2 CLASSICAL MC SIMULATIONS FOR OPTION VALUATION

Monte Carlo simulations seem to be a very convenient choice when one is trying to price options, especially when those options are exotic and there are some complicated features that we need to deal with. For instance, it offers the ability to deal effectively with high-dimensional problems either in the time parameter, when pricing path-dependent options, or in the state variable when pricing basket options, options on the maximum of three assets or more, etc. In addition to that, there is a relative ease of implementation in comparison to other approaches and it has a rate of convergence that is independent of the dimension of the problem treated and is equal to $O(n^{-1/2})$ where n is the number of simulations conducted.

2.1 MC Simulation in the Black & Scholes Framework

The Black and Scholes framework that lays the foundation for more complex models focuses on a standard call option over a single state variable. This framework assumes that we are evolving in a perfectly frictionless market that is continuously open, has a constant riskless interest and in which no taxes or transaction costs are incurred, it is also based on the no arbitrage principle which states that two portfolios of financial assets that have the same payoffs at all times and in every possible scenarios must have the same price. Any other way, a sure profit can be made out of no investment at all. In the Black and Scholes model, the underlying asset is assumed to be given by the process $\{X(t), 0 \leq t \leq T\}$ which is described by the stochastic differential equation:

$$dX(t) = \mu X(t)dt + \sigma X(t)dW(t) \quad (1)$$

Where $W(t), 0 \leq t \leq T$ is a Wiener process and coefficients μ and σ are constants assumed to satisfy usual conditions ensuring the existence and uniqueness of a continuous adapted solution to equation (1). According to the Feynman-Kac theorem and under some very general conditions, the no arbitrage price of most financial derivatives (European, American, Asian ...) can be expressed as the mathematical expectation of the associated payoff "[13],[16],[17]", which is usually defined as a functional of the underlying asset process, discounted at the risk-free interest rate under a risk neutral probability measure where discounted security prices do not have a drift and therefore are martingales "[10],[21]". In order to formalize this, the price $p(t)$ of a financial derivative defined by the payoff function Φ and involving the times (t_1, t_2, \dots, t_n) is expressed as:

$$p(t) = E^Q\{\Phi(X(t_1), X(t_2), \dots, X(t_n)) | X(0) = x\} \quad (2)$$

Let us just note that $0 \leq t_1 \leq t_2 \leq \dots \leq t_n = T$, where T is the stopping time that corresponds to the expiration date of the contingent claim. Now, from equation (2) it is quite apparent that using a Monte-Carlo approach is a natural choice when trying to compute $p(t)$ and the most straightforward way of applying this approach consists of the following steps:

1. Simulate sample paths of the underlying state variable over the specified time horizon.
2. Evaluate the discounted payoffs of the derivative on each sample path as specified by the nature of the derivative in question.

3. Average the discounted payoffs over the number of generated sample paths.

As a matter of fact, this simple approach allows for the evaluation of multidimensional integrals, therefore it can very well be adapted and applied to a wide range of derivatives, for instance, Asian calls, American calls and basket calls whose prices are respectively expressed as $E[e^{-rT} ((1/T) \int_0^T S_t dt - K)^+]$, $\max_{\tau} E[e^{-r\tau} (S_{\tau} - K)^+]$ and $E[e^{-rT} (\sum \omega_i S_i - K)^+]$ where r is the risk-free interest rate, T is the time to expiration of the option, τ is the time to the specified exercise dates of the American option, S_t is the price of the underlying at date t and ω_i is the weight of the underlying i within the considered basket. However, as stated earlier, the Monte Carlo approach demonstrates an apparent need for a large number of simulations. This drawback can be fought through variance reduction techniques, some of which are briefly discussed below.

2.2 Variance reduction for MC simulations

Variance reduction is quite interesting when one is trying to conduct Monte Carlo simulations for it provides a substantial improvement in computational efficiency [15], in the following we present some specific variance reduction techniques while indicating how they can be used in a financial context.

2.2.1 Antithetic Variates

The antithetic variates technique is probably the simplest method to adopt while trying to reduce the variance; it is also rather easy to implement and to fuse with other techniques. The basic idea is that if we consider two negatively correlated random variables X_1, X_2 then the variance of their sum $Var(X_1 + X_2)$ would be smaller than if they were independent. Now, in order to illustrate the use of this technique in a financial context let us consider the formula:

$$X(t) = X(0) \exp[(\mu - \sigma^2 / 2)t + \sigma W(t)] \quad (3)$$

Which is the analytical solution to the sde in equation (1). This suggests that for the implementation of each scenario of the stock price we are going to need these two equations:

$$X(t + \Delta t) = X(t) \exp[(\mu - \sigma^2 / 2)\Delta t + \sigma \varepsilon \sqrt{\Delta t}] \quad (4)$$

$$X(t + \Delta t) = X(t) \exp[(\mu - \sigma^2 / 2)\Delta t - \sigma \varepsilon \sqrt{\Delta t}] \quad (5)$$

Where $\varepsilon \sim N(0, 1)$.

The outputs of equations (4) and (5) are considered as the antithetic variates since they are clearly negatively correlated. Therefore, when trying to price an option using these variates the payoff in each replication could be computed as the average of the payoffs obtained by each one of the two scenarios. This accelerates the convergence towards the true price of the option, for more on this method see [6].

2.2.2 Control Variates

This method is a widely used variance reduction technique based on the introduction of a control variate that is positively correlated with the quantity we desire to estimate. Let X and Y be two correlated random variables where Y is called a control variate of X . Considering that the expectation of Y is known let's construct the new variable \tilde{X} :

$$\tilde{X} = X + \beta(Y - E(Y)) \quad (6)$$

This implies $E[\tilde{X}] = E[X]$, thus we can estimate X using \tilde{X} . The procedure revolves around choosing β that minimizes the variance of \tilde{X} :

$$\text{Var}(\tilde{X}) = \text{Var}(X) + \beta^2 \text{Var}(Y) - 2\beta \text{Cov}(X, Y) \quad (7)$$

The optimal choice β^* that minimizes equation (7) is found to be:

$$\beta^* = \text{Cov}(X, Y) / \text{Var}(Y) \quad (8)$$

Then, equation (7) becomes:

$$\text{Var}(\tilde{X}) = \text{Var}(X) - \text{Cov}(X, Y)^2 / \text{Var}(Y) \quad (9)$$

And this clearly shows that $\text{Var}(\tilde{X}) \leq \text{Var}(X)$ which really is the point.

Let's also note that this technique can be applied using multiple control variates of X , this gives:

$$\tilde{X} = X + \beta^T (\bar{Y} - E_Y) \quad (10)$$

Where $\beta = (\beta_1, \beta_2, \dots, \beta_n)^T$ is the coefficients vector, $\bar{Y} = (Y_1, Y_2, \dots, Y_n)^T$ the control variates vector and $E_Y = (E[Y_1], E[Y_2], \dots, E[Y_n])^T$ is the Y_i 's expectations vector, with $i = 1, 2, \dots, n$.

An application in a financial context can be depicted through the example of pricing Asian options. Let V_{art} be the price of an arithmetic Asian option and V_{geo} the price of a geometric Asian option. We assume that these two options have the same specifications aside from the way in which the payoff is computed. In practice most Asian options payoffs are computed using arithmetic averaging. However, prices of arithmetic Asian options cannot be evaluated using closed forms, conversely, geometric Asian options can. This contrast can be used through the control variates method to compute V_{art} . Hence, we can write:

$$\hat{V}_{art} = V_{art} - \beta (\hat{V}_{geo} - V_{geo}) \quad (11)$$

Where \hat{V}_{art} is the simulated price of the considered arithmetic Asian option, \hat{V}_{geo} is the simulated price of the considered geometric Asian option and V_{geo} is the price computed using the analytic solution. And since the correlation between \hat{V}_{art} and its control variable is strong we consider β to be equal to 1. This leads to the following form:

$$\hat{V}_{art} = V_{geo} + (\hat{V}_{art} - \hat{V}_{geo}) \quad (12)$$

Now, equation (12) enables us to compute a price for the arithmetic Asian option while reducing the variance of the simulations. A more thorough analysis is provided by Broadie and Glasserman [5] and Glasserman [13].

2.2.3 Importance sampling

Importance sampling relies on a familiar idea in mathematical finance which is the use of a 'likelihood ratio' also known as 'Radon-Nikodym derivative'. This idea allows us to change the underlying probability measure in order to represent an expectation under one measure as an expectation under another measure that will emphasize the relevant events. In particular, one may use this technique when trying to simulate rare or unusual events so it will help avoid running an extremely large number of simulations.

Let us consider a random variable X with a density function f , we desire to estimate:

$$\theta = E_f[h(X)] = \int h(x)f(x)dx \quad (13)$$

Suppose we have another analytical density function then equation (13) can be rewritten as:

$$\theta = \int (h(x)/g(x))f(x)g(x)dx = E_g[h(x)f(x)/g(x)] \quad (14)$$

The 'likelihood ratio' mentioned above is the quantity $f(x)/g(x)$ that helps correct the change in probability measure.

Now, let us explain how changing the probability measure might be helpful in reducing the variance. Above, it has been stated that one can estimate θ using either of the two following expectations $E_f[h(X)]$, $E_g[h^*(X)]$ where:

$$h^*(x) = h(x)f(x)/g(x) \quad (15)$$

This means that the two estimators have the same expectation, let's check their variance:

$$\text{Var}_f[h(x)] = \int h^2(x)f(x)dx - \theta^2 \quad (16)$$

$$\text{Var}_g[h^*(x)] = \int h^{*2}(x)g(x)dx - \theta^2 \quad (17)$$

The choice of g has to be made in order to reduce the variance, from equation (17) it is apparent that if g is chosen to be:

$$g(x) = h(x)f(x)/\theta \quad (18)$$

Then $\text{Var}_g[h^*(x)] = 0$, however, using equation (18) requires the knowledge of θ which is unknown to us. All the same, one can always try to approximate the ideal density g so an optimal variance reduction can be attained, for more on the subject see [4]. In financial markets this approach can be adopted when trying to price out of the money options. Suppose we use a straightforward Monte Carlo simulation, then, if the option is out of the money, most of the simulated trajectories wouldn't be efficient in pricing the option since they would lead to a zero value. Thus it seems more practical to only choose useful paths, meaning those leading to a price higher than the strike. Given that we know f the density of the underlying at the exercise date T and we can analytically identify the probability q that the price of the underlying asset is higher than the strike K at time T , then instead of sampling from f we could simulate trajectories from $g = f/q$ which takes into consideration the condition that the price of the underlying is higher than the strike. For a comprehensive example see [14].

3 A MODIFIED MC METHOD BASED ON DENSITY ESTIMATION

In this section we present a modification of the ordinary Monte Carlo technique with the aim of improving the precision and increasing the convergence speed. The suggested technique adopts a weighting approach; the weights are computed through a density estimate and allow for an adjustment of the calculations.

3.1 MC Simulation in the Black & Scholes Framework

The outline presented in the following is a general approach which could be adapted to price a wide range of options, as it doesn't consider any specific payoff function but instead tries to evaluate the expected value of an unspecified functional

depending on a known process, meaning that it can be applied to most payoff functions. Thus, to simplify the presentation we focus on options on a single underlying asset, however, the method can easily be generalized to options on multiple underlying assets. Let $X_i (1 \leq i \leq n)$ be some independent draws made according to an unknown probability distribution law over a positive support $[a, b]$. Now, we are interested in using these draws to compute over $[a, b]$ the expectation $E[\Phi(Y)]$ where $\Phi(\cdot)$ is a continuous function on $[a, b]$ and Y is a random variable with a known probability distribution. In theory, a first approximation could be written as:

$$E[\Phi(Y)] \leftarrow (1/n) \sum_{i=1}^n \Phi(X_i) g(X_i) / f(X_i) \text{ where } n \rightarrow \infty \quad (19)$$

Where g and f are respectively the probability density functions of Y and X . As stated above f is unknown, this fact, renders equation (19) useless and raises a problem of density estimation that can be tackled through a parametric or a nonparametric approach, we have chosen to use the latter, mainly because it needs very few assumptions hence avoiding any restrictive hypotheses regarding the actual distribution of the random variable in question.

3.2 Density estimation and Kernel smoothing

Probability density estimation is a major issue in statistics that can be addressed using two different ways, either a parametric or a nonparametric one. The parametric approach presumes that the unspecified density belongs to a known parametric family $\{\varphi(x, \theta) : \theta \in \Theta\}$ where $\varphi(\cdot, \cdot)$ is a given function and Θ is a subset of R^k with a fixed dimension k , then estimating the density goes back to estimating θ . It is clear that this technique binds the density function to a certain shape taken by $\varphi(\cdot, \cdot)$ whereas when dealing with a nonparametric approach such constraints are relaxed thus the only hypothesis regarding the density function f is the assumption that it belongs to a massive class of densities such as all the continuous probability densities on R . The nonparametric method we have chosen to work with is the kernel density estimation.

3.2.1 General overview

Kernel density estimation makes it possible to estimate the density f from a random i.i.d sample $\{X_1, X_2, \dots, X_n\}$ The basic idea of a kernel estimator is to retrieve the cardinal of the observations located around the point where we wish to estimate the density, this can be written as:

$$\hat{f}_n(x) = (1/nh) \sum_{i=1}^n K_0((X_i - x)/h) \quad (20)$$

where

$$K_0(u) = (1/2)I(-1 \leq u \leq 1) \quad (21)$$

The function K_0 is called the kernel function. Still, it is only the most naive example of a wide range of functions that could be considered as kernels. Actually, all function $K : R \rightarrow R$ satisfying the condition $\int K(u) du = 1$ are qualified to act as kernels, nonetheless, according to need, the kernel moments μ_i in equation (22) can be required to satisfy some conditions

related to symmetry and non-negativity.

$$\mu_i = \int u^i K(u) du \quad (22)$$

In the present paper, we will be using some unconventional kernels that have been introduced and tested in a previous work, these kernels have shown to be more efficient than the Gaussian, a widely used and popular kernel, when applied to financial data. The first one, the lognormal kernel, which is based upon the lognormal probability density function, takes the form:

$$K(u) = (1/(u\sqrt{2\pi})) \exp(-(\log u)^2 / 2) \quad (23)$$

In terms of parameter estimation, we have considered $\mu = 0$ and $\sigma = 1$ which are, respectively, the mean and the standard deviation of the associated normal distribution. Nevertheless, the parameters can be estimated through classical parametric fittings using the available data. As to the second and third kernels they are based upon Laguerre and Hermite orthogonal polynomials, these kernels can be of any desired order. But first, let us note that a kernel of order $l \in N^*$ is a function $K : R \rightarrow R$ where the functions $u \rightarrow u^i K(u)$, $i = 0, 1, 2, \dots, l$ are integrable and satisfy:

$$\int K(u) du = 1 ; \int u^i K(u) du = 0 ; i = 1, 2, \dots, l$$

Now, a kernel based on orthogonal polynomials can be given by:

$$K(u) = \sum_{i=1}^l \Pi_i(0) \Pi_i(u) \omega(u) \quad (24)$$

Where $\Pi_i(\cdot)_{i=1}^n$ is a normalized infinite orthogonal basis and $\omega(\cdot)$ is the weight function since some polynomials are only orthogonal with weights. Next, from equation (20), one can see that correctly using kernel density estimation requires estimating the parameter h called the bandwidth or smoothing parameter for it has a direct control over the smoothing applied to the data [28]. The value of this parameter has to be correctly estimated as to avoid under-smoothing and over-smoothing. In fact, under-smoothing might facilitate the appearance of some misleading features caused by the sampling process while over-smoothing can skip meaningful attributes of the underlying statistical structure. This brings us to the measurement of the error, considering that the main ideas behind the most used data-driven bandwidth selection methods are closely related to a common way of measuring the error which is the Mean Integrated Squared Error (MISE):

$$MISE(\hat{f}_n) = E[(\hat{f}_n - f)^2] \quad (25)$$

The MISE being simply the integrated squared error between the estimated and the actual density averaged over the set of realizations. This measure, despite its simplicity, has been proven to provide a profound analysis of the estimator's performance "[22],[23],[26]". In practice and under standard technical assumption [29], the MISE is usually approximated by using the first two terms of its asymptotic expansion (as $n \rightarrow \infty$), hence the Asymptotic Mean Integrated Squared Error (AMISE):

$$AMISE(\hat{f}_n) = (1/(nh))R(K) + (1/4)h^4 R(f'') \mu_2(K)^2 \quad (26)$$

Where the functional notations $\mu_2(K) = \int u^2 K(u) du$,

$R(K) = \int K(u)^2 du$ are used here and in the following.

3.2.2 Bandwidth selection

As mentioned previously, choosing the bandwidth is an essential task that has to be performed correctly. In this section we are interested in computing a fixed bandwidth (i.e. constant at all locations) for a univariate kernel density estimation. Ordinarily, this can be done through numerous approaches that were grouped into 'first generation' and 'second generation' methods "[19], [20]". Nonetheless, an automated bandwidth selection method can only help provide a sensible starting point when one desires to perform a proper analysis and knows the needed amount of smoothing. Because, as it has been pointed out in [19], a manual interactive tuning of the smoothing parameter is the best way to handle data. Now, knowing that the MISE is a sound measure of error, it is obvious that the smoothing parameter has to be chosen so that it guarantees the smallest error. Generally, and due to the complicated form that the MISE may take, it's the AMISE that's minimized; moreover its minimizer can be shown to be:

$$h_{AMISE} = (R(K)/(nR(f^{(r)})\mu_2(K)^2))^{1/5} \quad (27)$$

As it might be seen, even though it is not possible to conduct a direct implementation of equation (27) since it requires the knowledge of $R(f^{(r)})$ which is computed using the second derivative of the unknown density, the h_{AMISE} formula offers some valuable insights as to the optimal bandwidth, for instance, the fact that a smaller h would probably be suitable for a larger n , mainly because when the amount of available information is large, the estimator has to be locally concentrated so important features wouldn't be overlooked. In what follows we give a quick outline of a method, that provides a way to compute the bandwidth. This method falls in between the first and the second generation considering it goes a little deeper than relying on an estimate of $f^{(r)}$ that is based on a parametric family, but at the same time it doesn't require the significant amount of time and computing power needed by the second generation methods (solve-the-equation plug in methods for example). The changes put forward consist in using the derivatives of a common kernel for the sake of estimating the derivatives of the unknown density [1]. Assuming we have a density estimator based on a ' r ' times differentiable kernel, the r^{th} derivative of the estimator may be written as:

$$\hat{f}_n^{(r)}(x) = (1/(nh^{r+1})) \sum^n K^{(r)}((x - X_i)/h) \quad (28)$$

The common kernel we will be using is the Gaussian which is based on the standard normal distribution:

$$K(u) = (1/\sqrt{2\pi}) \exp(-u^2/2) \quad (29)$$

It's ' r^{th} ' derivative is expressed as:

$$K^{(r)}(u) = (-1)^r H(r,u)K(u) \quad (30)$$

Where $H(r,u)$ is the Hermite polynomial of order ' r ' that takes the form:

$$H(r,u) = (-1)^r \exp(u^2) (\partial^r \exp(-u^2) / \partial u^r) \quad (31)$$

From the equations above the explicit form of the density estimator's ' r^{th} ' derivative can be written as:

$$\hat{f}_n^{(r)}(x) = ((-1)^r / (nh^{r+1} \sqrt{2\pi})) \sum_{i=1}^n H(r, ((x - X_i)/h)) \exp(-((x - X_i)/(h\sqrt{2}))^2) \quad (32)$$

And in order to keep the formula computationally tractable we

will use the normal scale smoother to estimate the corresponding bandwidth, this smoother is expressed as follows:

$$\hat{h}_{NSS} = [(8\sqrt{\pi}R(K))/(3\mu_2(K)^2n)]^{1/5} \hat{\sigma} \quad (33)$$

Where $\hat{\sigma}$ is mostly chosen as being the standardized interquartile range:

$$\hat{\sigma} = \text{sample's } sIQR / (\phi^{-1}(3/4) - \phi^{-1}(1/4)) \quad (34)$$

Next, we need to estimate properly the functional $R(f^{(r)})$, the integration by part formula gives:

$$R(f^{(r)}) = (-1)^r \int f^{(2r)}(x) f(x) dx \quad (35)$$

For an even ' r ' this quantity can be estimated by:

$$\hat{\rho}_r = \sum \hat{f}_n^{(r)}(x_i) / n \quad (36)$$

More clearly $\hat{\rho}_r$ is of the form:

$$\hat{\rho}_r = (1/(n^2 h^{r+1})) \sum \sum K^{(r)}((X_i - X_j)/h) \quad (37)$$

This enables us to write the formula for the suggested AMISE-optimal bandwidth:

$$\hat{h}_{AMISE} = [R(K)/(n\mu_2(K)^2 \hat{\rho}_4)]^{1/5} \quad (38)$$

Now that we have specified the kernels that we will be using and a potential method to estimate the corresponding bandwidth for each one of them, we can plug them into equation (19) where the density function will be replaced by its kernel estimate.

4 NUMERICAL EXPERIMENT

The main focus of the numerical experiments conducted below is to assess the efficiency of the suggested approach, when pricing options, and empirically establish whether or not there is any improvement over the classical Monte Carlo technique in terms of precision and speed of convergence. We consider here the pricing of Asian options. But first, before initiating the simulations, it should be recalled that in order to use correctly the generic kernel specified in equation (24) one should choose appropriate orthogonal polynomials, we have opted for the polynomials whose specifications are hereinafter described.

4.1 Orthogonal polynomials and related kernels

Let us just note that $H^N(\cdot)$, $Lag^N(\cdot)$ refer respectively to the normalized Hermite and Laguerre polynomials.

Hermite polynomials (limits of orthogonality: R):

Rodrigues Formula: $H(k,u) = (-1)^k \exp(u^2) D_u^k [\exp(-u^2)]$

Weight: $\omega(u) = \exp(-u^2)$

Norm: $\eta(k) = \sqrt{2^k k! \sqrt{\pi}}$

Related kernel: $K(u) = \sum_{k=0}^{\infty} H^N(k,0) H^N(k,u) \omega(u) \quad (39)$

Laguerre polynomials (limits of orthogonality: R):

Rodrigues Formula: $Lag(k,\alpha,u) = (\exp u) u^{-\alpha} / k!$

$D_u^k [\exp(-u) u^{k+\alpha}]; \alpha > -1$

Weight: $\omega(u,\alpha) = u^\alpha \exp(-u)$

Norm: $\eta(k,\alpha) = \sqrt{\Gamma(k+\alpha+1)/k}$

Related Kernel: $K(u) = \sum_{k=0}^{\infty} Lag^N(k,\alpha,0) Lag^N(k,\alpha,u) \omega(u,\alpha) \quad (40)$

The first three odd orders of the kernels in equations (39) and (40) are displayed below [figs. 1 and 2] each over a suitable

interval

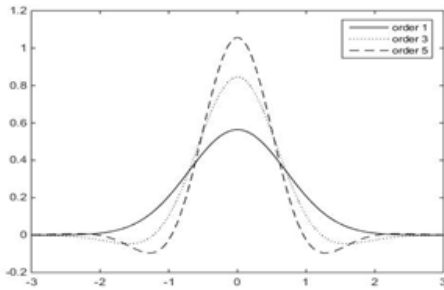


Figure 1: Hermite kernel of order 1, 3 and 5

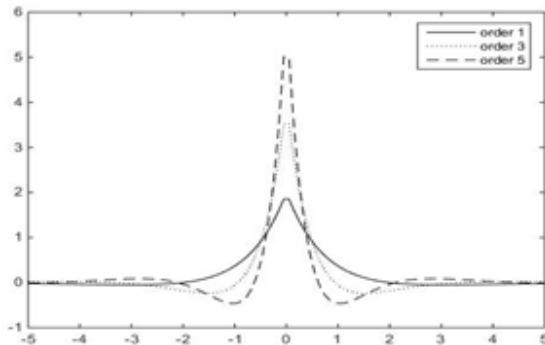


Figure 2: Laguerre kernel of orders 1, 3 and 5

4.2 Pricing an Asian option

Now, we desire to price three Asian options with the same volatility and duration but different strikes, the obtained prices will be compared to target prices which are the results in [9], where the option was valued using a pde approach. We therefore proceed as follows, first, the option is priced using the classical Monte Carlo technique and then, the approach involving kernel density estimation is tested using various kernels, amongst them the Gaussian which can be considered as a benchmark to judge how effective the new kernels are. When the estimated density is introduced, an approximation of $E[\phi(Y)]$ can be given by:

$$\sum_{i=1}^n ((\phi(X_i)g(X_i))/(\sum_{j=1}^n K_h(X_i - X_j))) \tag{41}$$

Where $K_h(X_i - X_j) = (1/h)K[(X_i - X_j)/h]$. Furthermore, in this instance, $\phi(\cdot)$ is the option's payoff function expressed as $((1/T)\int S_t dt - K)^+$ and $g(\cdot)$ is considered to be the lognormal pdf whose parameters are estimated through a lognormal fit, since, in the framework of Black and Scholes, the dynamic of stock prices is supposed to follow a lognormal distribution. Using equation (41) prices were computed for the considered Asian call with three different strikes and are presented in [figs. 3 to 8]. For each example, the results obtained using the kernels based on the probability density functions and the kernels based on orthogonal polynomials are displayed separately. When analyzing the figures, one can observe that, on the one hand, for the considered cases, the classical Monte Carlo technique is inclined to overestimate the target price, which is, the numerical solution to the pde of Roger and Shi, with no sign of the slightest improvement even when the number of simulations is increased in a considerable manner.

On the other hand, although the results given by the simulations using kernel density estimation start far from the desired price, the estimation shows a gradual convergence in the direction of the targeted value when the number of simulations is increased until the results get stabilized around 10000 replications. Of course, one can also notice that the behavior of the kernel based simulations is different from one kernel to another but the conclusion that can be drawn is that when approaching the target price, in contrast to the simulations using kernels based on orthogonal polynomials, those using kernels based on probability density functions register some relatively high discrepancies before returning to display some minor fluctuations.

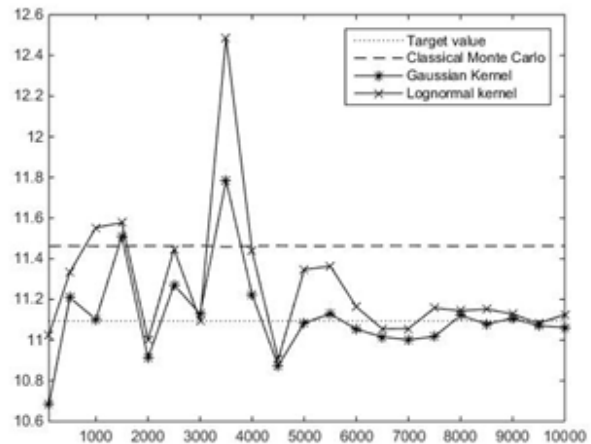


Figure 3: Values obtained for the price of an arithmetic Asian call using Classical MC and MC combined with the Gaussian and lognormal kernels with $r = 0.15$, $\sigma = 0.05$, $K = 95$, $S_0 = 100$.

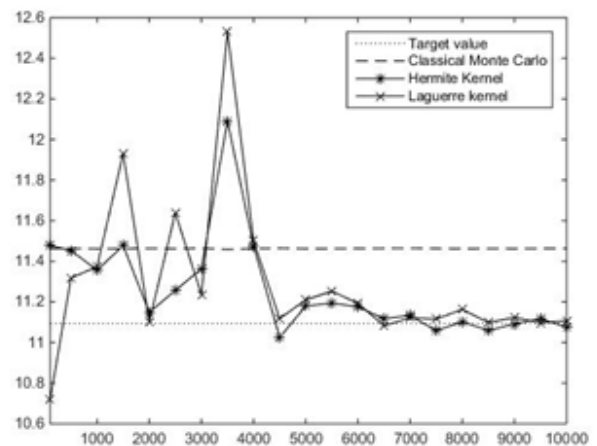


Figure 4: Values obtained for the price of an arithmetic Asian call using Classical MC and MC combined with Hermite and Laguerre kernels with $r = 0.15$, $\sigma = 0.05$, $K = 95$, $S_0 = 100$.

It's noteworthy that for large samples the AMISE-Optimal bandwidth tends to be very small, this causes equation (41) to severely underestimate the option's price. However, this

inconvenience can be overcome either by estimating the AMISE-Optimal bandwidth for a small sample then using this same bandwidth for larger samples or by manually tuning the bandwidth using a cross validation sample, especially given that a correct tuning is crucial for a stable convergence toward a satisfying price.

However, the suggested approach offers a considerable space for development since it may be adapted to deal with options with multiple underlyings having high volatilities or different volatility processes and described by a wide range of models.

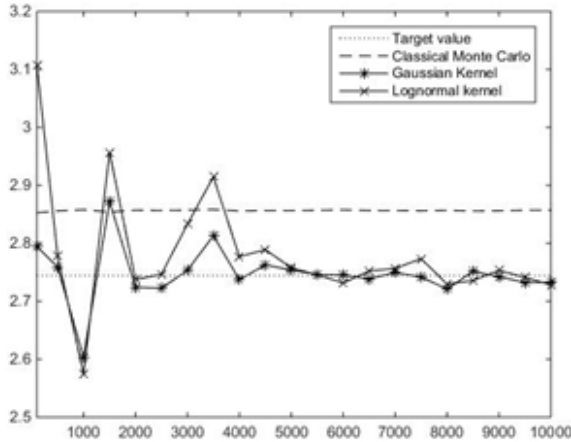


Figure 5: Values obtained for the price of an arithmetic Asian call using Classical MC and MC combined with the Gaussian and lognormal kernels with $r = 0.5$, $\sigma = 0.05$, $K = 105$, $S_0 = 100$.

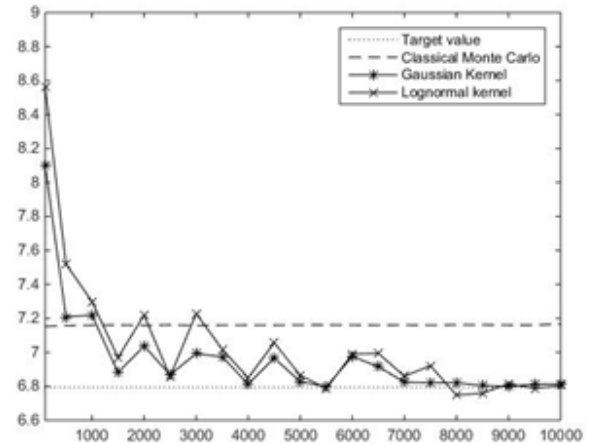


Figure 7: Values obtained for the price of an arithmetic Asian call using Classical MC and MC combined with the Gaussian and lognormal kernels with $r = 0.15$, $\sigma = 0.05$, $K = 100$, $S_0 = 100$.

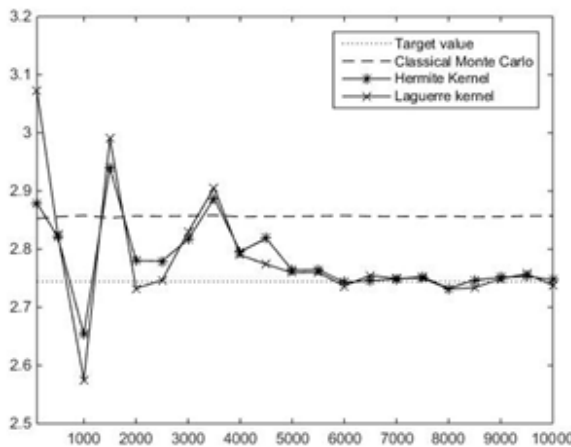


Figure 6: Values obtained for the price of an arithmetic Asian call using Classical MC and MC combined with Hermite and Laguerre kernels with $r = 0.5$, $\sigma = 0.05$, $K = 105$, $S_0 = 100$.

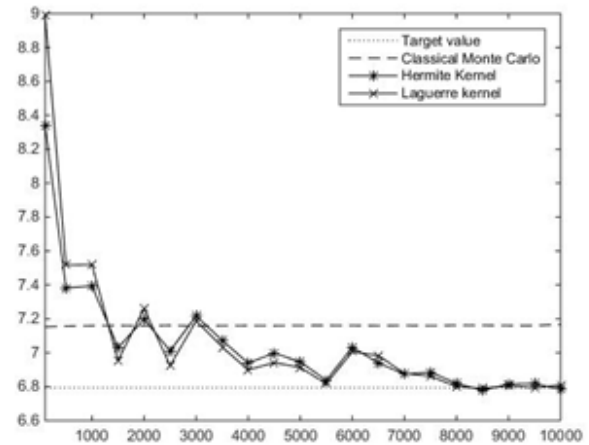


Figure 8: Values obtained for the price of an arithmetic Asian call using Classical MC and MC combined with Hermite and Laguerre kernels with $r = 0.15$, $\sigma = 0.05$, $K = 100$, $S_0 = 100$.

5 CONCLUSION

This paper presented a simple modification of the classical Monte Carlo technique for pricing options, based on introducing weights to adjust the crude averaging of payoffs computed using simulated asset paths with a reduced variance, these weights involve the use of an analytical density and an empirical density estimate. Even when simple probability density functions and low order kernels were used, the conducted tests have shown that the said adjustment lead to a significant improvement over the standard Monte Carlo in pricing an arithmetic Asian option in the case of low volatilities.

6 REFERENCES

- [1] P. Bhattacharya, "Estimation of a Probability Density Function and its Derivatives", Sankhya: The Indian Journal of Statistics, Series A, vol. 29, no. 4, pp. 373-382, 1967.
- [2] P. Boyle, "Options: A Monte Carlo Approach", Journal of Financial Economics, vol. 4, no. 3, pp. 323-338, 1977.
- [3] P. Boyle, M. Broadie and P. Glasserman, " Monte Carlo Methods for Security Pricing ", Journal of Economic Dynamics and Control, vol. 21, no. 8, pp. 1267-1321, 1997.

- [4] P. Brandimarte, *Numerical Methods in Finance and Economic*, 2nd ed. Hoboken, New Jersey: John Wiley & Sons, 2006.
- [5] M. Broadie, P. Glasserman, "Estimating Security Price Derivatives Using Simulation", *Management Science*, vol. 42, no. 2, pp. 269-285, 1996.
- [6] L. Clewlow and A. Carverhill, "On the Simulation of Contingent Claims", *The Journal of Derivatives*, vol. 2, no. 2, pp. 66-74, 1994.
- a. Conze, "Path Dependent Options: The Case of Lookback Options", *The Journal of Finance*, vol. 46, no. 5, pp. 1893-1907, 1991.
- [7] J.C. Cox, S.A. Ross and M. Rubinstein, "Option Pricing a Simplified Approach", *Journal of Financial Economics*, vol. 7, no. 3, pp. 229-263, 1979.
- [8] F. Dubois and T. Lelièvre, "Efficient Pricing of Asian Options by the PDE Approach", *Journal of computational finance*, vol. 8, no. 2, pp. 55-64, 2004.
- [9] D. Duffie, *Dynamic Asset pricing Theory*, 2nd ed. Princeton, New Jersey: University Press, 1996.
- [10] D. Egloff, M. Kohler and N. Todorovic, "A Dynamic Look-Ahead Monte Carlo Algorithm for Pricing Bermudan Options", *The Annals of Applied Probability*, vol. 17, no. 4, pp. 1138-1171, 2007.
- [11] G. Fusai and A. Roncoroni, *Implementing Models in quantitative Finance : Methods and Cases*. Berlin Heidelberg: Springer-Verlag, 2008.
- [12] P. Glasserman, *Monte Carlo Methods in Financial engineering*. New York: Springer-Verlag, 2003.
- [13] P. Glasserman, P. Heidelberger and N.P. Shahabuddin, "Pricing of Options Using Importance Sampling and Stratification/Quasi Monte Carlo", *US6, 381, 586B1*, 2002.
- [14] P.W. Glynn and W. Whitt, "The Asymptotic Efficiency of Simulation Estimators", *Operations Research*, vol. 40, no. 3, pp. 505-520, 1992.
- [15] J.M. Harrison and D.M. Kreps, "Martingales and Arbitrage in Multiperiod Securities Markets", *Journal of Economic Theory*, vol. 20, no. 3, pp. 381-408, 1979.
- [16] J.M. Harrison and S.R. Pliska, "Martingales and Stochastic Integrals in the Theory of Continuous Trading", *Stochastic Processes and Their Applications*, vol. 11, no. 3, pp. 215-260, 1981.
- [17] J. Hull and A. White, "One-Factor Interest Rate Models and the Valuation of Interest Rate Derivative Securities", *The Journal of Financial and Quantitative Analysis*, vol. 28, no. 2, pp. 235-254, 1993.
- [18] M.C. Jones, J.S. Marron and S.J. Sheather, "A Brief Survey of Bandwidth Selection for Density Estimation", *Journal of the American Statistical Association*, vol. 91, no. 433, pp. 401-407, 1996.
- [19] M.C. Jones, J.S. Marron and S. J. Sheather, "Progress in Data-Based Bandwidth Selection for Kernel Density Estimation", *Computational statistics*, vol. 11, no. 3, pp. 337-381, 1996.
- [20] Karatzas and S. Shreve, *Brownian Motion and Stochastic Calculus*, 2nd ed. New York: Springer-Verlag, 1991.
- [21] J.S. Marron and A. Tsybakov, "Visual Error Criteria for Qualitative Smoothing", *Journal of The American Statistical Association*, vol. 90, no. 430, pp. 499-507, 1995.
- [22] J.S. Marron and M.P. Wand, "Exact Mean Integrated Squared Error", *The Annals of Statistics*, vol. 20, no. 2, pp. 712-736, 1992.
- [23] R.C. Merton, M.J. Brennan and E.S. Schwartz, "The Valuation of American Put Options", *The Journal of Finance*, vol. 32, no. 2, pp. 449-462, 1977.
- [24] D.B. Nelson and K. Ramaswamy, "Simple Binomial Processes and Diffusion Approximations in Financial Models", *Review of Financial Studies*, vol. 3, no. 3, pp. 393-430, 1990.
- [25] B. Park and B. Turlach, "Practical Performance of Several Data-Driven Bandwidth Selectors", *Louvain: Université Catholique de Louvain, Center for Operations Research and Econometrics (CORE)*, 1992.
- [26] L.C.G. Roger and Z. Shi, "The Value of an Asian Option", *Journal of Applied Probability*, vol. 32, no. 4, pp. 1077-1088, 1995.
- [27] S.J. Sheather and M.C. Jones, "A Reliable Data-Based Bandwidth Selection Method for Kernel Density Estimation", *Journal of the Royal Statistical Society, Series B (Methodological)*, vol. 53, no. 3, pp. 683-690, 1991.
- [28] B. Silverman, *Density Estimation for Statistics and Data Analysis*. London ; New York: Chapman and Hall / CRC Press, 1986.
- [29] R. Zvan, K.R. Vetzal and P.A. Forsyth, "PDE Methods for Pricing Barrier Options", *Journal of Economics Dynamics and Control*, vol 24, no. 11, pp. 1563-1590, 2000.