Consistency of the Stochastic Mesh Method

Yuri Kashtanov

Faculty of Mathematics and Mechanics
Saint Petersburg State University
Russia
Email: y.kashtanov@spbu.ru

***Abstract*—A Monte Carlo method for pricing high-dimensional American options is considered. The consistency of the stochastic mesh method is studied. Some ”natural” estimators of this method have infinite variance. A modification which gives consistent estimators for a diffusion model is proposed. It is shown that the variance of estimators is inverse proportional to the number of points in each layer of the mesh.**

***Keywords*–Optimal stopping; American option; Stochastic mesh.**

I. INTRODUCTION

Pricing of American options may be formulated as a problem of the theory of optimal stopping: if S_n is the price process of the underlying asset and f_n are the option payoffs then the option price C has the expession [1] $\mathcal{C} = \sup_{\tau<T} \mathbf{E} f_\tau(S_\tau)$, T being the expiry date, τ - an arbitrary Markov moment. In the case when S_n is a high-dimensional process, numerical methods with regular mesh are very time-consuming. Mark Broadie and Paul Glasserman [2] suggested a stochastic mesh method which does not depend on the dimension. In parallel, Athanassios N. Avramidis and Heinrich Matzinger [3] formulated general conditions under which estimators of the method are statistically consistent. Using Malliavin calculus approach, Vlad Bally and others [4] suggested analogous estimators for a generalized Black-Scholes model; these estimators may be transformed [5] to the formula (20) of the present paper. This slight modification of formula (5) allows to convert the estimator with infinite variance to the one with a finite variance. The issue of the method consistency applying to the general diffusion model was not investigated till now and is the goal of the paper. Section 2 describes the stochastic mesh method; in Section 3, theoretical assertions, concerning consistency, are formulated; in Section 4, numerical examples, which illustrate these results, are given; proofs are gathered in the Appendix.

II. THE STOCHASTIC MESH METHOD

Let prices of d securities $S_n = (S_{n1}, \dots, S_{nd})$, $S_n \in R^d = X$, be given at discrete moments $t_n = n\Delta t$, where $\Delta t = T/N$. Assume also, that discounted prices are martingales that means the following equalities are fulfilled

$$\mathbf{E}(S_{n+1,i} \mid S_1, \dots, S_n) = e^{r\Delta t} S_{ni}, \tag{1}$$

r being an interest rate, which is assumed to be constant. Onwards, it will be more convenient to deal with sequences $\xi_{ni} = \ln S_{ni} - rt_n$. Assume that $\xi_n = (\xi_{n1}, \dots, \xi_{nd})$ is a Markov chain with values in X and $p_n(x, dy)$ are transitional probabilities.

Let an American option be given with an expiration T and discounted payoffs of the form $f_n = f_n(\xi_n)$. Define successively functions $Y_n(x)$:

$$Y_N(x) = f_N(x), \quad Y_n(x) = \max\left(f_n(x), \mathbf{E}_{n,x} Y_{n+1}(\xi_{n+1})\right). \tag{2}$$

It is known [1] that the option price $\mathcal{C}$ satisfies $\mathcal{C} = Y_0(\xi_0)$.

For every time step n, a set of random points $\bar{x}_n = \left\{x_n^i\right\}_{i=1}^M$ (”mesh”) is constructed as a Markov chain with transitional probabilities

$$\bar{q}_n(\bar{x}, d\bar{y}) = q_{n,1}(\bar{x}, dy_1)...q_{n,M}(\bar{x}, dy_M). \tag{3}$$

Make an assumption on probabilities $q_{n,j}(\bar{x}, dy)$ that densities

$$\rho_{n,j}(\bar{x}, x, y) = p_n(x, dy)/q_{n,j}(\bar{x}, dy) \tag{4}$$

exist. For a shortage of further notations, introduce random variables defined on the mesh: $Y_n(j) = Y_n(x_n^j)$, $\rho_n(x, j) = \rho_{n,j}(\bar{x}_{n-1}, x, x_n^j)$, $\rho_n(i, j) = \rho_n(x_{n-1}^i, j)$.

Recursively construct random sequence $\check{Y}_n(x)$: first, set $\check{Y}_N(x) = f_N(x)$, then define

$$\check{Y}_n(x) = \max\left(f_n(x), \frac{1}{M}\sum_{j=1}^M \rho_{n+1}(x, j)\check{Y}_{n+1}(j)\right), \tag{5}$$

where using analogous notations $\check{Y}_{n+1}(j) = \check{Y}_{n+1}(x_{n+1}^j)$.

Introduce $\mathcal{F}_n$ – a σ-algebra generated by values $\bar{x}_1, \dots, \bar{x}_n$. Denote a conditional expectation with respect to $\mathcal{F}_n$ as $\mathbf{E}_{\mathcal{F}_n}$. Assume that for every n, random variables j_n taking values $1, \dots, M$ with equal probabilities, are defined and which are independent in total and with respect to $\mathcal{F}_N$, then the last equality can be rewritten in the form

$$\check{Y}_n(x) = \max\left(f_n(x), \mathbf{E}_{\mathcal{F}_N} \rho_{n+1}(x, j_{n+1})\check{Y}_{n+1}(j_{n+1})\right). \tag{6}$$

It follows from (2) and (6) that

$$\left|\check{Y}_0(x_0) - Y_0(x_0)\right| \leqslant \sum_{n=0}^{N-1} \mathbf{E}_{\mathcal{F}_N} \prod_{k=1}^n \rho_k(j_{k-1}, j_k)|\Delta_n(j_n)|, \tag{7}$$

where

$$\Delta_n(x) = \mathbf{E}_{\mathcal{F}_N} \rho_{n+1}(x, j_{n+1})Y_{n+1}(j_{n+1}) - \mathbf{E}_{n,x} Y_{n+1}(\xi_{n+1}).$$

In virtue of (3), values $x_{n+1}^1, \dots, x_{n+1}^M$ are independent with regard to $\mathcal{F}_n$, therefore

$$\mathbf{E}_{\mathcal{F}_n} \Delta_n^2(x) \leqslant \frac{1}{M}\mathbf{E}\rho_{n+1}^2(x, j_{n+1})Y^2(j_{n+1}). \tag{8}$$

 ISBN: 978-1-61208-371-1

It follows in turn from (7) and (8) that under the condition

$$\mathbf{E}[\rho_1(x_0, j_)...\rho_n(j_{n-1}, j_n)Y_n(j_n)]^2 < \infty, \quad n = 1, \ldots, N; \tag{9}$$

the inequality

$$\mathbf{E}(\check{Y}_0 - \mathcal{C})^2 \leqslant C/M \tag{10}$$

is fulfilled ant thereby $\check{Y}_0$ is a consistent estimator for $\mathcal{C}$.

In [2] some substantiation is given for using the average distribution, which is defined by the expression

$$q_{n,j}(\bar{x}, dy) = M^{-1} \sum_i p_n(x^i, dy). \tag{11}$$

In some models (e.g. Black-Scholes), transitional probabilities for several steps $p_{k,n}(x, dy)$ are known, in this case one can use [2]

$$q_{n,j}(\bar{x}, dy) = p_{0,n}(x_0, dy). \tag{12}$$

The objective of the paper is to study the statistical consistency of these and other estimators for the general diffusion model.

III. Consistent estimators

First, consider a case when x_n are normal vectors in R^d defined by the sequence

$$x_{n+1} = x_n + a_n \varepsilon_{n+1} + b_n, \tag{13}$$

a_n being $d \times m$ - matrices, ε_n – independent standard normal vectors in R^m, $A_n = a_n a_n^*$, $b_n(i) = -0.5 A_n(i, i)$. Denote $\psi(A_n, z) = (A_n z, z)$ and suppose that

$$\sup_{||z||=1} \psi(A_n, z) \geqslant \bar{A}_0 > 0, \tag{14}$$

then densities $p_{k,n+1}(x, y)$ exist for transition from point x at the moment k to point y at the moment $n + 1$:

$$p_{k,n+1}(x, y) = c_{kn} \exp(-0.5\psi(\Sigma_{k,n}^{-1},\ x + B_{k,n} - y)),$$

where $c_{kn} = [(2\pi)^d \det \Sigma_{kn}]^{-1/2}$, $\Sigma_{k,n} = \sum_{i=k}^n A_i$ and $B_{k,n} = \sum_{i=k}^n b_i$. Suppose payoffs imply that

$$Y_n(x) \leqslant \sum_i c_i(e^{k_i x_i} + 1). \tag{15}$$

Theorem 1. *The inequality (10) for the estimator with average densities (11) is valid under conditions (14), (15).*

Proofs are given in the Appendix.

Now, consider the mesh generated according to the formula (12). Note that it is approximately 2 times less time-consuming for calculation than (11) but the variance of $\check{Y}_0$ may be infinite if N is big enough. Really, let $N \geqslant 3$ then

$$\begin{aligned}&\mathbf{E}\check{Y}_0^2 \geqslant \mathbf{E}\left[\rho_2(j_1, j_2)\rho_3(j_2, j_3)f_3(j_3)\right]^2 \\ &= \frac{1}{M^3} \int_{X^3} p(x_0, y_1) \frac{p^2(y_1, y_2)}{p_{0,2}(y_0, y_2)} \frac{p^2(y_2, z_3)}{p_{0,3}(y_0, y_3)} f_3^2(y_3) dy_1 dy_2 dy_3.\end{aligned}$$

Consider the case when $d = 1$ and a_n are constant, then after integrating by y_1, y_2 receive

$$\mathbf{E}\check{Y}_0^2 \geqslant c \int_R \exp\left(\frac{7y^2}{78a^2}\right) f_3^2(ya\sqrt{\Delta t} + 3b\Delta t) dy. \tag{16}$$

If $f_3(x) > \varepsilon$ for $x > K$ or $x < -K$ with some $\varepsilon, K > 0$ then the integral does not converge and thus the variance is infinite. Below, a modification of this estimator with finite variance will be constructed for a more general diffusion model.

Now, consider a discretization of a diffusion process according to the Euler scheme:

$$x_{n+1} = x_n + a_n(x_n)\varepsilon_{n+1} + b_n(x_n). \tag{17}$$

Suppose that for some $\bar{A}_0$, $\bar{A}_1$ the inequalities are fulfilled

$$\bar{A}_0|z|^2 \leqslant (A_n(x)z, z) \leqslant \bar{A}_1|z|^2; \tag{18}$$

then $|b_n(x)| \leqslant \bar{b}$ for some $\bar{b}$.

Let the mesh be produced by transition probabilities

$$q_{n,j}(\bar{x}, dy) = q_n(y)dy = c_n \exp(-|y - x_0|^2/(2s^2 n))dy. \tag{19}$$

Consider a following modification of the scheme (5) :

$$\check{Y}_n(i) = \max\left(f_n(i), \frac{\sum_j \rho_{n+1}(i, j)\ \check{Y}_{n+1}(j)}{\sum_j \rho_{n+1}(i, j)}\right). \tag{20}$$

Note that unlike the scheme (5), the estimator of the expectation (6) is biased here.

Theorem 2. *Suppose that $f_n(x) \leqslant F$ and $s^2 > 0.5\bar{A}_1$ then the inequality (10) is valid for the scheme (20).*

Note. Consider a more general model for which transitional densities may be estimated by a mixture

$$p_n(x, y) \leqslant C \sum_{k=1}^K p_n^{(k)}(x, y), \tag{21}$$

$p_n^{(k)}(x, y)$ being normal densities with correlation matrices $A_n^{(k)}(x)$. Suppose these matrices satisfy condition (18) with common constants A_0, A_1, then, as follows from the proof, the assertion of the theorem still holds true.

Now, suppose that instead of the condition $f_n(x) \leqslant F$ the the following condition is fulfilled: for some $g_n(x)$, G and F

$$f_n(x) \leqslant F g_n(x), \qquad \int_X dy\, p_n(x, y) g_n(y) \leqslant G g_{n-1}(x). \tag{22}$$

Denote $\widetilde{g}_n(x) = \int_X dy\, p_n(x, y) g_n(y)$ and consider the scheme

$$\check{Y}_n(i) = \max\left(f_n(i), \widetilde{g}_n(i) \frac{\sum_j \rho_{n+1}(i, j)\ \check{Y}_{n+1}(j)}{\sum_j \rho_{n+1}(i, j) g_{n+1}(j)}\right). \tag{23}$$

Theorem 3. *Suppose conditions (22) are fulfilled, densities $\widetilde{p}_n(x, y) = p_n(x, y) g_n(y)/\widetilde{g}_n(x)$ satisfy (21) and $s^2 > 0.5A_1$ then inequality (10) is valid for the scheme (23).*

Note 1. Though the variance decreases when M tends to infinity, it may be still quite big for a chosen M. To reduce it one can use a method of control variates [2], which implies the substitution of the term $\check{Y}_{n+1}(j)$ in formulas (5), (23) by the term $\check{Y}_{n+1}(j) + \nu_{n+1}(j)$ with appropriate functions $\nu_{n+1}(x)$, which satisfy the condition $\mathbf{E}_{n,x}\nu_{n+1}(\xi_{n+1}) = 0$.

Note 2. Give an example when conditions of the theorem are fulfilled. For many multi-dimensional options (basket, geometrical average, maximum [2]), functions $g_n(x)$ may be chosen in the form $g_n(x) = \sum_{i=1}^d \exp(x_i)$. Since $\exp(\xi_{ni})$

 ISBN: 978-1-61208-371-1

are martingales then $\widetilde{g}_n(x) = g_n(x)$. It is easy to verify that the representation (21) takes place under $K = d$, correlation matrices $A_n^{(k)}(x) = A_n(x)$ and $b_n^{(k)}(x) = b_n(x) + A_n(x)e_k$.

IV. Numerical results

Consider a geometrical average option with payoff functions $f_n = ((S_{n1}..S_{nd})^{1/d} - K)^+$. In the case of Black-Scholes model, the calculation may be reduced to a one-dimensional American call with dividends and the solution may be obtained with high accuracy either by a regular mesh or by the analytical approximation [6]. Under parameters $d = 5$, $a = 0.3$, $r = 0.02$, $S = 100$, $K = 100$, $T = 1$, the price of American option in the continuous model is $\mathcal{C}_a = 4.63$, which is strictly greater than the price of the correspondent European option $\mathcal{C}_e = 4.46$. In discrete models the price of American option increases from $\mathcal{C}_e$ to $\mathcal{C}_a$ with increasing N. As a control variate, take the corresponding European call option.

TABLE I. Comparison of estimators

Estimator, density			(5), (11)		(23), (19)	
N	True	M	Av	Err	Av	Err
6	4.59	300	4.61	0.008	4.59	0.006
		1200	4.60	0.004	4.59	0.0026
12	4.61	300	4.7	0.017	4.61	0.0055
		1200	4.66	0.006	4.61	0.0023
24	4.62	300	4.86	0.04	4.63	0.005
		1200	4.81	0.015	4.62	0.0025

The table includes the true price, estimates and statistical errors, corresponding to 99% confidence level by 50 realizations. It illustrates the assertion that the variance is inverse proportional to M. It may be observed from the table (and is known from the theory) that the estimator (5), (11) is biased high and this bias increases with the number of time steps; to compensate it one should increase the mesh size. Note also that the first estimator is approximately 2 times more time-consuming than the second one.

Consider the same option in the following diffusion model $\Delta x_{ni} = \sqrt{\Delta t}\sigma(x_{n-1,i})[\alpha\varepsilon_{ni} + \rho\varepsilon_{n,d+1}] - 0.5\sigma^2(x_{n-1,i})\Delta t$, where $\alpha^2 + \rho^2 = 1$, $\rho = 0.5$, $\sigma(x)$ being a decreasing function, which varies from 0.6 to 0.1, $\sigma(x_0) = 0.3$.

TABLE II. Case of the general diffusion model

s	0.2		0.43		0.5	
M	Av	Err	Av	Err	Av	Err
300	7.13	0.012	7.19	0.009	7.21	0.012
600	7.13	0.008	7.18	0.007	7.19	0.009
1200	7.14	0.007	7.16	0.005	7.17	0.007

Rough 1-dimensional approximations give prices from 7.1 to 7.2. The simulation results for the estimator (19), (23) with $N = 12$ and different values of the parameter s are presented in the Table II. The value 0.43 is slightly greater than $A_1/\sqrt{2}$, which provides the variance finiteness.

V. Appendix

Proof of theorem 1. First, note that $\rho_1(x_0, j) = 1$ for every j; further, $\mathbf{E}_{\mathcal{F}_N}\rho_2(j_1, j) = 1$ also for every j and by induction we receive that the right part in (7) is equal to $\sum_{n=0}^{N-1} |\Delta_n(j_n)|$. Therefore, to prove consistency it is sufficient to show that for every $n < N$ the following integrals are finite:

$$\mathbf{E}\rho_{n+1}^2(j_n, 1)Y_{n+1}^2(1) = \int_X dy\; Y_{n+1}^2(y)\mathbf{E}\frac{\sum_j p_{n+1}^2(x_n^j, y)}{\sum_j p_{n+1}(x_n^j, y)}. \tag{24}$$

To estimate the expectation under the sign of the integral, prove the following lemma:

Lemma. *Let η_i, $i = 1, \dots, M$, be positive, independent, similar distributed random variables, $\zeta_i = c\eta_i^{1+\varepsilon}$ $\varepsilon > 0$; denote $\bar\eta = M^{-1}\sum_j \eta_j$, $\bar\zeta = M^{-1}\sum_j \zeta_j$, then*

$$\mathbf{E}\bar\zeta/\bar\eta \leqslant 3\mathbf{E}\zeta/\mathbf{E}\eta.$$

Proof. From the equality

$$\frac{\bar\zeta}{\bar\eta} = \frac{\bar\zeta}{\mathbf{E}\eta} - \frac{1}{\mathbf{E}\eta}\frac{\bar\zeta}{\bar\eta}(\bar\eta - \mathbf{E}\eta) \tag{25}$$

the estimate follows:

$$\mathbf{E}\frac{\bar\zeta}{\bar\eta} \leqslant \frac{\mathbf{E}\zeta}{\mathbf{E}\eta} + \frac{1}{\mathbf{E}\eta}\mathbf{E}\frac{\bar\zeta}{\bar\eta}\left|\bar\eta - \mathbf{E}\eta\right|. \tag{26}$$

Further, from inequality $\sum_j \eta_j^{1+\varepsilon} \leqslant (\sum_j \eta_j)^{1+\varepsilon}$ it follows that

$$\bar\zeta/\bar\eta \leqslant \Big(\sum_j \eta_j^{1+\varepsilon}\Big)^{\frac{\varepsilon}{1+\varepsilon}}. \tag{27}$$

Using the last estimate and then Hölder's inequality receive

$$\begin{aligned} \mathbf{E}\frac{\bar\zeta}{\bar\eta}\left|\bar\eta - \mathbf{E}\eta\right| &\leqslant \mathbf{E}\Big(\sum_j \eta_j^{1+\varepsilon}\Big)^{\frac{\varepsilon}{1+\varepsilon}}\left|\bar\eta - \mathbf{E}\eta\right| \\ &\leqslant \mathbf{E}^{\frac{\varepsilon}{1+\varepsilon}}\Big(\sum_j \eta_j^{1+\varepsilon}\Big)\,\mathbf{E}^{\frac{1}{1+\varepsilon}}\left|\bar\eta - \mathbf{E}\eta\right|^{1+\varepsilon}. \end{aligned} \tag{28}$$

Now use the following theorem ([7], c.79):

Let $X_1, \dots, X_n$ be independent random variables with zero mean and finite absolute moments of order p ($1 \leqslant p \leqslant 2$); then

$$\mathbf{E}\left|\sum_{k=1}^n X_k\right|^p \leqslant \Big(2 - \frac{1}{n}\Big)\sum_{k=1}^n \mathbf{E}|X_k|^p.$$

Using this theorem one can estimate (28) by the value

$$M^{\frac{\varepsilon}{1+\varepsilon}}\mathbf{E}^{\frac{\varepsilon}{1+\varepsilon}}\eta^{1+\varepsilon}\;\frac{2}{M^{\frac{\varepsilon}{1+\varepsilon}}}\mathbf{E}^{\frac{1}{1+\varepsilon}}\eta^{1+\varepsilon} = 2\mathbf{E}\zeta.$$

Substituting the obtained estimates in (26), one receive the assertion of the lemma.

Now denote $\eta_{k,n+1}(j) = p_{k,n+1}(x_{k,j}, y)$; then applying lemma (with $\varepsilon = 1$) to expectation in (24), receive

$$\mathbf{E}\frac{\sum_j \eta_{n,n+1}^2(j)}{\sum_j \eta_{n,n+1}(j)} \leqslant 3\mathbf{E}\frac{\mathbf{E}_{\mathcal{F}_{n-1}}\eta_{n,n+1}^2(1)}{\mathbf{E}_{\mathcal{F}_{n-1}}\eta_{n,n+1}(1)}. \tag{29}$$

According to Chapman-Kolmogorov equations, the denominator is equal to $\sum_j \eta_{n-1,n+1}(j)$. Calculate $\mathbf{E}_{\mathcal{F}_{k-1}}\eta_{k,n+1}^m(1)$ with $m > 1$:

$$\mathbf{E}_{\mathcal{F}_{k-1}}\eta_{k,n+1}^m(1) = \widetilde{c}_{kn}\sum_j e^{-\psi(\widetilde\Sigma_{k-1,n}^{-1},\; x_{k-1}^j + B_{k-1,n} - y)/2},$$

 ISBN: 978-1-61208-371-1

where

$$\widetilde{\Sigma}_{k-1,n} = A_{k-1} + \frac{1}{m}\Sigma_{k,n} = \Sigma_{k-1,n} - \frac{m-1}{m}\Sigma_{k,n}. \quad (30)$$

It follows from (30) that $\widetilde{\Sigma}^{-1}_{k-1,n} = \Sigma^{-1}_{k-1,n}(I - D_{k,n})^{-1}$, where $D_{k,n} = (m-1)/m\ \Sigma_{k,n}\Sigma^{-1}_{k-1,n}$, I being a unit $d \times d$ matrix. Since $||D_{k,n}|| < 1$, one can estimate the inverse matrix $\widetilde{\Sigma}^{-1}_{k-1,n} \geqslant \Sigma^{-1}_{k-1,n}(I + D_{k,n})$. On the other hand, $D_{k,n} \geqslant \varepsilon I$ with some value $\varepsilon = \varepsilon_{k,n,m} > 0$; therefore, an estimation from below is $\widetilde{\Sigma}^{-1}_{k-1,n} \geqslant (1+\varepsilon)\Sigma^{-1}_{k-1,n}$. Particularly, it follows from this estimate that

$$\frac{\mathbf{E}_{\mathcal{F}_{n-1}}\eta^2_{n,n+1}(1)}{\mathbf{E}_{\mathcal{F}_{n-1}}\eta_{n,n+1}(1)} \leqslant C_1 \frac{\sum_j \eta^{1+\varepsilon}_{n-1,n+1}(j)}{\sum_j \eta_{n-1,n+1}(j)}$$

with $\varepsilon = \varepsilon_{n-1,n,2}$. The lemma may be applied to the obtained expression as well, and after n iterations receive

$$\mathbf{E}\frac{\sum_j p^2_{n+1}(x^j_n, y)}{\sum_j p_{n+1}(x^j_n, y)} \leqslant C_n p^{\varepsilon}_{0,n+1}(x_0, y).$$

Under condition (15) the integral in (24) is finite. Thus, the proof is complete.

Proof of theorem 2. Introduce an additional notation

$$\bar{\rho}_n(i,j) = \rho_n(i,j)/\sum_k \rho_n(i,k)$$

and in the same way as in (7) obtain

$$\left|\check{Y}_0 - Y_0\right| \leqslant \sum_{n=0}^{N-1}\sum_{i_1,\ldots,i_n} \bar{\rho}_1(0,i_1)\ldots\bar{\rho}_n(i_{n-1},i_n)|\Delta_n(i_n)|,$$

where $\Delta_n(i) = \sum_j \bar{\rho}_{n+1}(i,j)Y_{n+1}(j) - \mathbf{E}_{n,x_{ni}}Y_{n+1}(\xi_{n+1})$. Squaring both parts and taking into account that $\bar{\rho}_n(i,j)$ is a distribution by j, due to Hölder inequality obtain

$$\mathbf{E}\left(\check{Y}_0 - Y_0\right)^2 \leqslant N\sum_{n=0}^{N-1}\sum_{i_1,\ldots,i_n}\mathbf{E}\prod_{k=1}^{n}\bar{\rho}_k(i_{k-1},i_k)\Delta^2_n(i_n). \quad (31)$$

Define values $d_{kn}(i)$:

$$\begin{aligned} d_{nn}(i) &= \mathbf{E}_{\mathcal{F}_n}\Delta^2_n(i), \\ d_{k-1,n}(i) &= \mathbf{E}_{\mathcal{F}_{k-1}}\bar{\rho}_k(i,j_k)d_{k,n}(j_k). \quad k < n, \end{aligned}$$

and show that the following estimate takes place

$$d_{kn}(i) \leqslant \min\left(F^2, \frac{C}{M}\sum_{m=k}^{n}\phi_{m+1,m+1-k}(x_{ki})\right), \quad (32)$$

where $\phi_{rl}(x) = e^{\frac{(|x-x_0|+\beta_l)^2}{s^2 r - 0.5\bar{A}_1 l}}$, $\beta_l = \bar{b}\sum_{i=1}^{l} d^{(i-1)/2}$.

Fix the index i and represent $\Delta_n(i)$ in a form $\Delta_n(i) = \bar{\zeta}/\bar{\eta}$, where $\bar{\zeta} = \sum_j \zeta_j/M$, $\bar{\eta} = \sum_j \eta_j/M$,

$$\begin{aligned} &\zeta_j = \rho_{n+1}(i,j)\left[Y_{n+1}(j) - \mathbf{E}_{n,x_{ni}}\widetilde{Y}_{n+1}(S_{n+1})\right], \\ &\eta_j = \rho_{n+1}(i,j),\ \mathbf{E}_{\mathcal{F}_n}\zeta_j = 0,\ \mathbf{E}_{\mathcal{F}_n}\eta_j = 1. \end{aligned}$$

It follows from (26) that

$$\mathbf{E}_{\mathcal{F}_n}\left(\frac{\bar{\zeta}}{\bar{\eta}}\right)^2 \leqslant 2\mathbf{E}_{\mathcal{F}_n}\bar{\zeta}^2 + 2\mathbf{E}_{\mathcal{F}_n}\left(\frac{\bar{\zeta}}{\bar{\eta}}\right)^2(\bar{\eta}-1)^2.$$

First, note that $|\zeta_j| \leqslant F|\eta_j|$, and second that with respect to $\mathcal{F}_n$ values $\bar{\eta} - 1$ and $\bar{\zeta}$ are sample averages of independent random variables with zero mean, therefore

$$\mathbf{E}_{\mathcal{F}_n}\Delta^2_n(i) \leqslant \frac{C}{M^2}\sum_j \mathbf{E}_{\mathcal{F}_n}\eta^2_j.$$

Now calculate $\mathbf{E}_{\mathcal{F}_n}\eta^2_j$; denote $z^i_n = x^i_n + b_n(x^i_n) - x_0$ then

$$\begin{aligned} \mathbf{E}_{\mathcal{F}_n}\eta^2_j &= c_{ni}\exp(\psi([(n+1)s^2 I - 0.5A_n(x^i_n)]^{-1}, z^i_n)/2) \\ &\leqslant C\phi_{n+1,1}(x^i_n). \end{aligned} \quad (33)$$

Thus, (32) is fulfilled for $k = n$. Now suppose that the inequality is fulfilled for the index $k+1$; estimate $d_{kn}(i)$. Redenote $\zeta_j = \rho_{k+1}(i,j)d_{k+1,n}(j)$, $\eta_j = \rho_{k+1}(i,j)$, then $d_{kn}(i) = \mathbf{E}_{\mathcal{F}_k}\bar{\zeta}/\bar{\eta}$. Split the region of integration and use the induction assumption that $d_{k+1,n}(i) \leqslant F^2$:

$$\begin{aligned} \mathbf{E}_{\mathcal{F}_k}\bar{\zeta}/\bar{\eta} &= \mathbf{E}_{\mathcal{F}_k}\bar{\zeta}/\bar{\eta}\,1_{\{\bar{\eta}\geqslant 1/2\}} + \mathbf{E}_{\mathcal{F}_k}\bar{\zeta}/\bar{\eta}\,1_{\{\bar{\eta}<1/2\}} \\ &\leqslant 2\mathbf{E}_{\mathcal{F}_k}\bar{\zeta} + F^2\mathbf{E}_{\mathcal{F}_k}1_{\{|\bar{\eta}-1|>1/2\}}. \end{aligned} \quad (34)$$

For the first term use the induction assumption

$$\mathbf{E}_{\mathcal{F}_k}\bar{\zeta} \leqslant \frac{C}{M}\sum_{m=k+1}^{n}\int_X dy\ p_{k+1}(x^i_k, y)\phi_{m+1,m-k}(y),$$

and then the estimate $\phi_{r,l}(y) \leqslant \sum_{i=1}^{2d} e^{\frac{(x_0 - x + e_i\beta_l\sqrt{d})^2}{s^2 r - 0.5\bar{A}_1 l}}$, e_i being the basis vectors in R^d and opposite ones by sign. The integration leads to

$$\int_X dy\ p_{k+1}(x^i_k, y)\phi_{m+1,m-k}(y) \leqslant C\phi_{m+1,m-k+1}(x^i_k).$$

For the second term in (34), apply the Chebyshev inequality and enequality (33):

$$\mathbf{E}_{\mathcal{F}_k}1_{\{|\bar{\eta}-1|>1/2\}} \leqslant 4\mathbf{E}_{\mathcal{F}_k}(\bar{\eta}-1)^2 \leqslant \frac{C}{M}\phi_{k+1,1}(x^i_k).$$

Thus, (32) is proven, which implies $d_{0n}(x_0) \leqslant C/M$. Finally, from (31) receive the assertion of the theorem:

$$\mathbf{E}\left(\check{Y}_0 - Y_0\right)^2 \leqslant N\sum_{n=0}^{N-1} d_{n0}(x_0) \leqslant C/M. \quad (35)$$

Proof of theorem 3. Introduce auxiliary payoff functions $\widetilde{f}_n(x) = f_n(x)/g_n(x)$ and consider the sequence

$$\widetilde{Y}_n(x) = \max\left(\widetilde{f}_n(x),\ \alpha_n(x)\int_X dy\ \widetilde{p}_{n+1}(x,y)\widetilde{Y}_{n+1}(y)\right), \quad (36)$$

where $\alpha_n(x) = \widetilde{g}_{n+1}(x)/g_n(x)$. One can construct the scheme analogous to (20) for the sequence (36):

$$Y'_n(i) = \max\left(\widetilde{f}_n(i), \alpha_n(i)\frac{\sum_j \widetilde{\rho}_{n+1}(i,j)\ Y'_{n+1}(j)}{\sum_j \widetilde{\rho}_{n+1}(i,j)}\right), \quad (37)$$

where $\widetilde{\rho}_n$ is defined by (4) with substitution $\widetilde{p}$ for p. Since $\widetilde{f}_n(x) \leqslant F$ and $\alpha_n(x) \leqslant G$, the theorem 2 may be applied to the scheme (37). Thus $\mathbf{E}(Y'_0(\xi_0) - \widetilde{Y}_0(\xi_0))^2 \leqslant CM^{-1}$. On the other hand, it is easy to verify that $\widetilde{Y}_n(x) = Y_n(x)/g_n(x)$ and $Y'_n(i) = \check{Y}_n(i)/g_n(i)$. The proof is complete.

 ISBN: 978-1-61208-371-1

VI. Conclusion and Future Work

It is shown in the paper that a simple modification from scheme (5) to schemes (20), (37) allows to prove the finiteness of the variance and, thus, consistency of the estimators in the general diffusion model. Besides, unpleasant bias vanishes, which allows to reduce the number of nodes in the mesh. The next step may be the extension of results to the jump-diffusion model, which is important for applications.

Acknowledgment

This work was supported by RFBR grant 14-01-00271.

References

[1] I. Karatzas, "On the pricing of American options. Applied Mathematics and Optimization," Applied Mathematics and Optimization, vol. 17, 1988, pp. 37–60.

[2] M. Broadie and P. Glasserman, "A Stochastic Mesh Method for Pricing High-Dimensional American Options," Journal of Computational Finance, vol. 7, 2004, pp. 35–72.

[3] A. Avramidis and H. Matzinger, "Convergence of the stochastic mesh estimator for pricing Bermudan options," Journal of Computational Finance, vol. 7, 2004, pp. 73–91.

[4] V. Bally, L. Caramellino, and A. Zanette, "Pricing and Hedging American Options by Monte Carlo Methods Using a Malliavin Calculus Approach," Journal of Monte Carlo Methods and Applications, vol. 11, 2005, pp. 97–134.

[5] Y. N. Kashtanov, Financial Mathematics Models and Statistical Simulation. VVM, St.-Petersburg, 2013, ISBN: 978-5-9651-0741-4.

[6] G. Barone-Adesi and R. E. Whaley, "Efficient Analitic Approximation of American Option Values," Journal of Finance, vol. 42, 1987, pp. 301–320.

[7] B. von Bahr and C.-G. Esseen, "Inequalities for the rth absolute moment of a some of random variables, $1 \leqslant r \leqslant 2$," Ann. Math. Statist, vol. 36, 1965, pp. 299–303.