Estimating Parametric Derivatives of First Exit Times of Diffusions by Approximation of Wiener Processes

Sergey A. Gusev*,
Institute of Computational Mathematics and Mathematical Geophysics SB RAS,
prospect Akademika Lavrentjeva, 6, Novosibirsk, 630090, Russia,
Email: sag@osmf.sscc.ru

***Abstract*—The problem of obtaining estimates of derivatives with respect to parameters of mathematical expectations of functionals of diffusion processes with absorbing boundary is considered in the paper. The problem demands to obtain the parametric derivatives of first exit times for the random processes. These derivatives can be obtained from the differentiation of the equation which is the result of applying the Ito's formula to some function that vanishes on the boundary. The problem of differentiating the Ito integral, that arises here, is solved by approximating the Wiener process by a Gaussian one with exponential correlation function, consistent with the step length of the Euler method.**

***Index Terms*—diffusion process, first exit time, parametric derivatives, exponential correlation function, Euler method.**

I. Introduction

Usually a study of a mathematical model is associated with the investigation of the nature of dependence of the model equations on the parameters. If the problem containing a diffusion process with absorbing boundary are in the consideration, then estimating the mean values of certain characteristics of the model is always connected with the first exit time τ of the process from the domain. This is due to the fact that theoretically the process can exit from a bounded domain at any time. Therefore, the first exit time of the diffusion process is a random value which depends on the parameters, and it is necessary to study this dependence. In particular, it is of interest if it is possible evaluating derivatives of τ with respect to the parameters. Evaluation of these derivatives is also important in solution of problems of stochastic optimization by gradient methods.

In [1] it was proposed an idea of finding estimates of the parametric derivatives of τ from the equation that is result of applying the Ito formula to a function g which vanishes on the boundary. The main problem here was differentiation with respect to upper limit of the stochastic integral over the Wiener process. To overcome this difficulty there was suggested in this paper to approximate the Wiener process in the stochastic integral by a Gaussian process with an exponential correlation function and use the Euler method for the numerical solution of the stochastic differential equation (SDE). Wherein this correlation function matched with the length of the step in the Euler method. In this paper, the method [1] has been further promoted and more fully justified.

In [2], [3] a variant of the method is considered when the function g vanishes on the boundary together with its first derivatives, but this option requires more stringent constraints on the smoothness of the input data.

* This work was supported by Russian Leading Science Schools Programme grant 5111.2014.1, Russian Foundation for Basic Research grant 14-01-00340-a,

II. Statement of the problem

Let $G \subset \mathbf{R}^d$ be a bounded domain with regular boundary ∂G. We suppose that we are given a complete filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$, $t \geq 0$. Let $W.$ be given d-dimensional Wiener process $W.$ which is measurable at any $t \geq 0$ with respect to $\mathcal{F}_t$ and $W_s - W_t$ for $s > t$ is independent on σ-algebra $\mathcal{F}_t$. Let us denote $E_{t,x}$ a mathematical expectation with respect to the probability measure that corresponds to the random process initiating from the point x at the time item t. Let us introduce for a given time segment $[0, T]$ a space of random functions $h(t)$ which are measurable at each $t \in [0, T]$ with respect to $\mathcal{F}_t$ and which have with probability one the finite integral $\int\limits_0^T g^2(t)dt$.

We also enter into consideration $Q_T = (0, T) \times G$ a cylinder in $\mathbf{R}^+ \times \mathbf{R}^d$ and an open set $U \subset \mathbf{R}^m$. For $x \in G$ and $t \in [0, T)$ we define a d-dimensional diffusion process $X.$ which depends on vector parameter $\theta \in U$, and this process is described by the following SDE

$$X_s(\theta) = x + \int\limits_t^s a(v, X_v(\theta), \theta)dv + \int\limits_t^s \sigma(v, X_v(\theta), \theta)dW_v \quad (1)$$

where $a : [0, \infty) \times \mathbf{R}^d \times U \to \mathbf{R}^d$ and $\sigma : [0, \infty) \times \mathbf{R}^d \times U \to \mathbf{R}^{d \times d}$ are measurable functions.

We assume that that the following condition holds for the coefficients of the equation (1).

(**A**) the functions a, σ are bounded and there exists a constant $\mathcal{K}$ such that for all $\theta \in U$, $v \geq 0$, $x, y \in \mathbf{R}^d$, $i, j \in \{1, \ldots, d\}$ the following inequality is valid

$$|a_i(v,x,\theta) - a_i(v,y,\theta)| + |\sigma_{ij}(v,x,\theta) - \sigma_{ij}(v,y,\theta)| \leq \mathcal{K}|x - y|.$$

It is supposed in the paper that the absorbing boundary condition holds, i.e. each path of $X.$ ends when the given time interval [0, T] exhausted, or the boundary ∂G reached. For such diffusions there are many applications when it is necessary to evaluate expectations of the type

$$u(t,x,\theta) = E_{t,x}\big[\varphi(X_T(\theta),\theta)\chi_{\tau>T} + \psi(\tau, X_\tau,\theta)\chi_{\tau<T} + \int_t^{T\wedge\tau} f(v, X_v(\theta),\theta)dv\big], \quad (2)$$

where $\tau = \inf\{v| \ v > t, \ X_v \notin G\}$ is the first exit time of the process $X.$ from G; χ_A is the set indicator function of a set A.

It is known that for any fixed $\theta \in U$ under some the value (2) at the point $(t,x) \in Q_T$ coincides with a solution of the parabolic boundary value problem

$$Lu + f(t,x,\theta) = 0, \quad t \in (0,T), \quad (t,x) \in Q_T, \quad (3)$$

$$u(T,x,\theta) = \varphi(x,\theta), \quad x \in G, \quad (4)$$

$$u(t,x,\theta) = \psi(t,x), \quad (t,x) \in S_T, \quad (5)$$

where $S_T = (0,T) \times \partial G$.

The operator L in the equation (3) is defined as

$$L \equiv \partial_t + \frac{1}{2}\sum_{i,j=1}^d b_{ij}(t,x,\theta)\partial^2_{x_i,x_j} + \sum_{i=1}^d a_i(t,x,\theta)\partial_{x_i}, \quad (6)$$

where b_{ij} the elements of the matrix $B \equiv \sigma\sigma^*$.

Sufficient conditions for existence and uniqueness of the solution of the problem can be found in [4].

We introduce the following assumptions providing existence and uniqueness of the solution of the problem (3) — (5) and the possibility of differentiation of the mathematical expectation (2):

(**B**) the matrix function $B(t,x,\theta)$ in (6) satisfies for all $(t,x) \in Q_T$, $\theta \in U$ to the condition:

$$B(t,x,\theta) > \alpha_0 I \quad (7)$$

for some $\alpha_0 > 0$.

(**C**) the derivatives $\frac{\partial a}{\partial x}, \frac{\partial^2 a}{\partial x^2}, \frac{\partial a}{\partial \theta_i}, \frac{\partial^2 a}{\partial x \partial \theta_i}, \frac{\partial \sigma}{\partial x}, \frac{\partial^2 \sigma}{\partial x^2}, \frac{\partial \sigma}{\partial \theta_i}, \frac{\partial^2 \sigma}{\partial x \partial \theta_i}, \frac{\partial a}{\partial t}, \frac{\partial \sigma}{\partial t}$ are bounded and continuous in $[0,\infty) \times R^d \times U$;

(**D**) there exist continuous derivatives $\frac{\partial \varphi}{\partial x}$ $\frac{\partial \varphi}{\partial \theta}$ for all $(x,\theta) \in Q \times U$;

(**E**) the function f is continuous on $[0,T] \times \overline{Q}_T$ for all $\theta \in U$ and there exist continuous derivatives $\frac{\partial f}{\partial x}$ $\frac{\partial f}{\partial \theta}$ for all $(t,x,\theta) \in Q_T \times U$.

III. Applying the Euler method for the numerical approximation of the functional

We define on the interval $[t,T]$ a grid with knots $t_i = t + hi$ $(i = 0,\dots,N)$, $h = \frac{T-t}{N}$. The approximation of a random process, obtained by the Euler method, will be denoted by a bar above. Value of a function at a grid point will be denoted by the superscript that is equaled to the corresponding node number.

Using the notation adopted above, we write the Euler scheme of calculations

$$\overline{X}_{t_{n+1}} = \overline{X}_{t_n} + ha^n + \sqrt{h}\sum_{j=1}^d \sigma^n_j \xi^n_j, \quad (8)$$

where σ^n_j is j-th column of the matrix σ^n; ξ^n_j are independent standard normal random variables.

Let us denote by $\tau^N = \inf\{t_i : \overline{X}_{t_i} \notin G\}$ and by i^τ the corresponding node number, i.e. $t_{i^\tau} = \tau^N$.

For all $(t,x) \in (t,T) \times G$ we will define the approximated values $u(t,x)$ by the formula

$$u^N(t,x) = E_{t,x}\big[\varphi(\overline{X}_{t_N})\chi_{\tau^N>t_N} + \psi(\tau^N, \overline{X}_{t_N})\chi_{\tau^N>t_N} + \sum_{i=0}^{N-1}\chi_{\tau^N>t_i} f(t_i, \overline{X}_{t_i})(t_{i+1}-t_i)\big] \ . \quad (9)$$

IV. Determination of $\partial u/\partial\theta$

Further for simplicity we assume that θ is a scalar parameter and $U \subset \mathbf{R}$ is an interval. Sometimes for brevity formulas the argument θ is omitted.

Let us differentiate (2) with respect to θ

$$\frac{\partial u}{\partial \theta}(t,x) = E_{t,x}\Big[\Big(\frac{\partial \varphi}{\partial x}(T, X_T,\theta)Z_T + \frac{\partial \varphi}{\partial \theta}(T,X_T,\theta)\Big)\chi_{\tau>T} + \Big(\frac{\partial \psi}{\partial x}(\tau, X_\tau,\theta)Z_\tau + \frac{\partial \psi}{\partial \theta}(\tau, X_\tau,\theta)\Big)\chi_{\tau<T} + \int_t^{T\wedge\tau}\Big(\frac{\partial f}{\partial x}(v, X_v,\theta)Z_v + \frac{\partial f}{\partial \theta}(v,X_v,\theta)\Big)dv\Big] + \Phi(\theta) \ , \quad (10)$$

where

$$\Phi(\theta) := \lim_{\Delta\theta\to 0} E_{t,x}\left(\frac{\tau(\theta+\Delta\theta)-\tau(\theta)}{\Delta\theta}\big(f(\tau, X_\tau) + \frac{\partial \psi}{\partial t}(\tau, X_\tau)\big)\chi_{\tau<T}\right). \quad (11)$$

Our main problem in the paper is to prove that the limit (11) exists and to obtain a formula of $\partial u/\partial\theta$ which does not contain the derivative $\partial\tau/\partial\theta$.

The process

$$Z_s(\theta) = \int_t^s \left(\frac{\partial a}{\partial x}Z_v(\theta) + \frac{\partial a}{\partial \theta}\right)dv + \int_t^s \left(\frac{\partial \sigma}{\partial x}Z_v(\theta) + \frac{\partial \sigma}{\partial \theta}\right)dW(v) \quad (12)$$

is mean-square derivative $\frac{\partial X.}{\partial \theta}$. It is known that the assumptions made previously ensure existence of this derivative [5], and it can be obtained from the SDE system (1), (12).

In the book [6] it is considered a possibility of change of white noise in the Wiener integral by a stationary Gaussian process λ having an exponential correlation function

$$R_\lambda(t_1,t_2) = D \cdot exp(-\beta|t_2 - t_1|) \quad (13)$$

with sufficiently small correlation time.

In according to [7] a Gaussian process λ with zero expectation and exponential correlation function can be obtained from the SDE

$$d\lambda = -\beta\lambda dv + \vartheta dW(v) \quad , \tag{14}$$

where $\lambda(0) = \lambda_0$ is a Gaussuan random value with zero expectation and variance equaled to $\frac{\vartheta^2}{2\beta}$. At this λ_0 the process λ has constant disperion which is equald to $D = \frac{\vartheta^2}{2\beta}$.

The exact solution of the SDE (14) is the following

$$\lambda(t) = e^{-\beta t}\lambda_0 + e^{-\beta t}\int_0^t e^{\beta s}\vartheta dW(s). \tag{15}$$

To obtain estimates of expectation of the functional containing derivatives of τ with respect to θ, we use this type random processes for a change white noises in the Wiener integral.

Let $f \in H_2[0,T]$. We consider the integrals $I_W \equiv \int_0^T f(v)dW(v)$ and $I_\lambda \equiv \int_0^T f(v)\lambda_v dv$ as integrals defined in the mean-square sense.

We approximate the integrals I_W and I_λ by the corresponding integral sums defined on the uniform grid with the step $h = \frac{T}{N}$. So, $I_W^h(f) \equiv \sum_{n=0}^{N-1} f(t_n)\Delta W_{t_n}$ where $\Delta W_{t_n} = W(t_{n+1}) - W(t_n)$, and $I_\lambda^h(f) \equiv h\sum_{n=0}^{N-1} f(t_n)\lambda(t_{n+1})$. It is assumed that $f(t_0)$ is nonrandom or random value independent of λ_0.

It is shown in the theorem below that at specified parameters of the process λ the convergence in distribution of the integral sums $I_\lambda^h(f)$ to the integral sums $I_W^h(f)$ is possible.

Theorem 1: Let the process λ is a solution of the SDE

$$d\lambda = -\beta\lambda dv + \sqrt{\frac{2\beta}{h(1-e^{-2\beta h})}}dW(v). \tag{16}$$

Then for a given $h > 0$ the following representation of λ values at the grid points is valid

$$\lambda(t_n) = \varepsilon_n + \frac{1}{\sqrt{h}}\xi_n, \quad n = 1, \dots, N,$$

where ε_n are normal random variables having zero expectation and variance $\frac{\exp(-2\beta h)}{h\left(1-\exp(-2\beta h)\right)}$; ξ_n are mutually independent normal $N(0,1)$ random values. In this case, there is a convergence in distribution $I_\lambda^h(f) \to I_w^h(f)$ at $\beta \to \infty$.

Proof. We will define the parameter ϑ in (14) so that the approximate equality with sufficient accuracy is valid

$$\hbar\lambda((n+1)h) \approx h^{\frac{1}{2}}\xi_n, \tag{17}$$

where ξ_n are mutually independent $N(0,1)$ normal random values. At the same time we note that $\lambda((n+1)h)$ are correlated normally distributed random values.

Let us take $t = (n+1)h$ in(15) and represent $\lambda((n+1)h)$ in the following sum

$$\lambda((n+1)h) = \varepsilon_n + \gamma_n, \tag{18}$$

where

$$\varepsilon_n = e^{-\beta(n+1)h}\lambda_0 + e^{-\beta(n+1)h}\int_0^{nh} e^{\beta s}\vartheta dW(s) \tag{19}$$

$$\gamma_n = e^{-\beta(n+1)h}\int_{nh}^{(n+1)h} e^{\beta s}\vartheta dW(s) \quad . \tag{20}$$

The random values ε_n are correlated , they are normally distributed, have zero expectation and the variance $D_{\varepsilon_n} = e^{-2\beta h}\frac{\vartheta^2}{2\beta}$. The random values γ_n are independent and normally distributed with zero expectation and the variance $D_{\gamma_n} = \frac{\vartheta^2}{2\beta}(1-e^{-2\beta h})$.

We define ϑ from the condition of the second moments equality

$$h^2\mathrm{E}(\gamma_n)^2 = \mathrm{E}\big(\Delta W(t_n)\big)^2 = h. \tag{21}$$

Using the property of the second central moment of an integral over the Wiener process we obtain the equality

$$h^2\mathrm{E}(\gamma_n)^2 = h^2 e^{-2\beta(n+1)h}\int_{nh}^{(n+1)h} e^{2\beta s}\vartheta^2 ds = \int_{nh}^{(n+1)h} ds. \tag{22}$$

Hence, we obtain $\vartheta = \left(\frac{2\beta}{h\left(1-e^{-2\beta h}\right)}\right)^{\frac{1}{2}}$. At this value ϑ the random values $h\gamma_n$ and $\Delta w(t_n)$ are identically distributed, and the variance ε_n is equaled to $\frac{e^{-2\beta h}}{h(1-e^{-2\beta h})}$. It is evident that by increasing β at the fixed value h it possible to decrease the contribution of the random value ε_n into $\lambda((n+1)h)$, and respectively to decrease error δ obtained as the result of replacing $\Delta w(t_n)$ on $h\lambda((n+1)h)$. This error is defined by the random values ε_n and it is equaled to

$$\delta = h\sum_{n=0}^{N-1} f(t_n)\varepsilon_n. \tag{23}$$

It follows from the definition ε_n in (19) that, as $\beta \to \infty$ the mathematical expectation δ goes to zero. Indeed, because the function f a.s. bounded on $[0,T]$ (as $f \in H_2[0,T]$) then with probability one $\max_n f^2(t_n) \le C_f$ for some $C_f > 0$. Using the Cauchy-Bunyakovskii inequality for mathematical expectations, we have

$$\mathrm{E}|\delta| \le h\sum_{n=0}^{N-1}\mathrm{E}|f(t_n)\varepsilon_n| \le h\sum_{n=0}^{N-1}\left(\mathrm{E}|f(t_n)|^2\right)^{\frac{1}{2}}\left(\mathrm{E}|\varepsilon_n|^2\right)^{\frac{1}{2}}$$
$$\le C_f^{\frac{1}{2}}\frac{Te^{-\beta h}}{\sqrt{h(1-e^{-2\beta h})}} \to 0 \qquad \beta \to \infty. \tag{24}$$

Let us consider the variance of δ

$$\mathrm{D}_\delta = h^2\mathrm{E}\Big[\sum_{n=0}^{N-1} f^2(t_n)\varepsilon_n^2 + \sum_{i\neq j} f(t_i)f(t_j)\varepsilon_i\varepsilon_j)\Big]. \tag{25}$$

The following inequality for $\mathrm{E}[f^2(t_n)\varepsilon_n^2]$, $n = 0, \dots N-1$ is valid

$$\mathrm{E}[f^2(t_n)\varepsilon_n^2] \le C_f\frac{e^{-2\beta h}}{h(1-e^{-2\beta h})}. \tag{26}$$

Then we define an evaluation of $\mathrm{E}[f(t_i)f(t_j)\varepsilon_i\varepsilon_j]$ $(i,j = 0,\dots,N; i \neq j)$. With the help of the Cauchy-Bunyakovskii inequality we have

$$\mathrm{E}[f(t_i)f(t_j)\varepsilon_i\varepsilon_j] \leq C_f\big(\mathrm{E}[\varepsilon_i^2\varepsilon_j^2]\big)^{\frac{1}{2}} \tag{27}$$

Let us write the expression for $\varepsilon_i^2\varepsilon_j^2$

$$\varepsilon_i^2\varepsilon_j^2 = e^{-2\beta(i+j+2)h}\Bigg[\lambda_0^4 + 2\lambda_0^3\Big[\int\limits_0^{ih} e^{\beta s}\vartheta dW_s + \int\limits_0^{jh} e^{\beta s}\vartheta dW_s\Big] + \lambda_0^2\Big[\Big(\int\limits_0^{ih} e^{\beta s}\vartheta dW_s\Big)^2 + \Big(\int\limits_0^{jh} e^{\beta s}\vartheta dW_s\Big)^2\Big]$$
$$+4\lambda_0^2\Big(\int\limits_0^{ih} e^{\beta s}\vartheta dW_s\Big)\Big(\int\limits_0^{jh} e^{\beta s}\vartheta dW_s\Big)$$
$$+2{\lambda_0}^2\Big[\Big(\int\limits_0^{ih} e^{\beta s}\vartheta dW_s\Big)\Big(\int\limits_0^{jh} e^{\beta s}\theta dW_s\Big)^2$$
$$+\Big(\int\limits_0^{ih} e^{\beta s}\vartheta dW_s\Big)^2\Big(\int\limits_0^{jh} e^{\beta s}\vartheta dW_s\Big)\Big]$$
$$+\Big(\int\limits_0^{ih} e^{\beta s}\vartheta dW_s\Big)^2\Big(\int\limits_0^{jh} e^{\beta s}\vartheta dW_s\Big)^2\Bigg]. \tag{28}$$

Hence we obtain the expectation of $\varepsilon_i^2\varepsilon_j^2$

$$\mathrm{E}[\varepsilon_i^2\varepsilon_j^2] = e^{-2\beta(i+j+2)h}\frac{\vartheta^4}{4\beta^2}\Big[e^{2\beta ih} + e^{2\beta jh} + 7e^{2\beta(i\wedge j)h}$$
$$+(e^{2\beta(i\wedge j)h} - 1)(e^{2\beta(i\vee j)h} - e^{2\beta(i\wedge j)h}) - 6\Big]$$
$$= e^{-4\beta h}\frac{\vartheta^4}{4\beta^2} + R_{ij}, \tag{29}$$

where R_{ij} are infinitesimal (at large values βh) compared to $e^{-4\beta h}$.

Thus, based on (27), (29), we obtain the inequality

$$\mathrm{E}[f(t_i)f(t_j)\varepsilon_i\varepsilon_j] \leq C_f\frac{e^{-2\beta h}}{h(1-e^{-2\beta h})}. \tag{30}$$

We obtain from (25), (26), (30) the estimate of the variance of δ

$$\mathrm{D}_\delta \leq \frac{C_f T^2 e^{-2\beta h}}{h(1-e^{-2\beta h})} \tag{31}$$

It follows from $\mathrm{E}|\delta| \to 0$ and $\mathrm{D}_\delta \to 0$ under $\beta \to \infty$ a.s. convergence $\delta \to 0$ when $\beta \to \infty$.

Because $I_\lambda^h(f) = \delta + h\sum_{i=0}^{N-1} f(t_k)\gamma_k$, the sum $h\sum_{i=0}^{N-1} f(t_k)\gamma_k$ is distributed as well as $I_w^h(f)$ and $\delta \to 0$ a.s. when $\beta \to \infty$ (hence convergence in probability). Then based on Remark 5 on the page 124 [8] $I_\lambda^h(f) \to I_w^h(f)$ when $\beta \to \infty$ in distribution. The proof is complete.

Because the normal distribution is uniquely determined by its expectation and variance, the random values $h\gamma_n$ with defined parameter ϑ from the condition (21) can be considered as increments of the Wiener process.

Let W^γ be designation of this process. Similarly, for an arbitrary set of points $\{t_1,\dots,t_n\} \subset [0,T]$ increments W^γ are defined as $\Delta W^\gamma_{t_k-t_{k-1}} \equiv \left(\frac{2\beta(t_k-t_{k-1})}{1-e^{-2\beta(t_k-t_{k-1})}}\right)^{\frac{1}{2}} e^{-\beta t_k}\int_{t_{k-1}}^{t_k} e^{\beta s}dW_s$.

Random values $\Delta W^\gamma_{t_k-t_{k-1}}$ are independent for disjoint intervals, they are normally distributed with zero expectation and variance $t_k - t_{k-1}$, i.e. they are distributed as increments of the Wiener process. Hence, finite-dimensional distributions W^γ coincide with finite-dimensional distributions of the Wiener process. Set of finite-dimensional distributions of the random process completely define its properties. By the theorem of Kolmogorov (see, for example, [9], page 110) there is a random process on $[0,T]$ that has the same finite-dimensional distributions as W^γ. We will for him to use the same notation W^γ. The process W^γ is the Wiener process.

Note, if increments of the random process are modeled using λ instead of the Wiener process on a non-uniform grid, then the variance of W^γ must change according to the length of the step of the grid. Therefore, it is necessary to define a new process λ with parameter ϑ corresponding to the new step length when the step of grid changes. With this, new process λ starts to move from the node in which the step changes.

It is shown in the next theorem that Ito's integral $\int_0^t f(v)dW^\gamma(v)$ can be approximated by the integral of the type $\int_0^t f(v)\lambda_v dv$.

Theorem 2: Let $f : [0,T]\times\Omega \to \mathbf{R}$ *be a random function from* $H_2[0,T]$*. Then for any* $\varepsilon > 0$ *the parameters of* λ *can be defined so that for each* $t \in [0,T]$ *the following inequality is valid*

$$E\Big|\int\limits_0^t f(v)dW^\gamma(v) - \int\limits_0^t f(v)\lambda(v)dv\Big| \leq \varepsilon.$$

Proof. Because $f \in H_2[0,T]$ there is a sequence of step functions f_m that it takes place the mean-square convergence

$$\Big|\int\limits_0^t f_m(v)dW^\gamma(v) - \int\limits_0^t f(v)dW^\gamma(v)\Big|^2 \to 0 \quad m \to \infty, \tag{32}$$

And there are many ways to construct such a sequence. Any function f_m from the sequence of step functions corresponds a grid $0 = t_0 < t_1 < \ldots < t_N = T$. We will consider in the theorm a sequence of step functions defined on uniform grids. Each function f_m we assign an integer N_m such that the corresponding step is defined from the equality $h_m = \frac{T}{N_m}$. And it is supposed that $N_m \to \infty$ as $m \to \infty$.

Each m specifies a step function f_m that is defined for any $t \in [0,T]$ as follows:

$$f_m(t,\omega) = \sum_{n=0}^{n_t-1} f(t_n,\omega)\chi_{t\in\Delta t_n} + f(t_{n_t},\omega)\chi_{t\geq t_{n_t}},$$

where $n_t = [\frac{t}{h_m}]$; $\Delta t_n = [t_n, t_{n+1})$. And the integral Ito over Wiener process is continuous function with respect to $t \in [0, T]$ and it is defined by the equality

$$I_{W^\gamma}(t) \equiv \int_0^t f_m(v)dW^\gamma(v) = \sum_{n<n_t} f(t_n,\omega)\Delta W^\gamma_{t_{n+1}-t_n} + f(t_{n_t},\omega)\Delta W^\gamma_{t-t_n}, \tag{33}$$

where$\Delta W^\gamma_{u-v} \equiv w^\gamma(u) - w^\gamma(v) = \sqrt{\frac{2\beta(u-v)}{1-e^{-2\beta(u-v)}}} e^{-\beta u} \cdot \int_v^u e^{\beta s} dW(s)$. For any $\beta > 0$ and $u > v$ the random value ΔW^γ_{u-v} is normally distributed and $\Delta W^\gamma_{u-v} \in N(0, (u-v)^{0.5})$. It follows from the Cauchy-Bunyakovskii inequality for mathematical expectations and mean-square convergency (32) that for a given $\varepsilon > 0$ there exists an integer number $m_1(\varepsilon)$ such that for $m \geq m_1(\varepsilon)$ the inequalities hold

$$E\Big|\int_0^t f(v)dW^\gamma(v) - \int_0^t f_m(v)dW^\gamma(v)\Big| \leq \Big(E\Big|\int_0^t f(v)dW^\gamma(v) - \int_0^t f_m(v)dW^\gamma(v)\Big|^2\Big)^{\frac{1}{2}} \leq \frac{\varepsilon}{9}. \tag{34}$$

Using the random values γ_n defined for an uniform grid by the equality (20), the integral of f_m over the process W^γ can be represented as follows:

$$\int_0^t f_m(v)dW^\gamma(v) = h_m \sum_{n=0}^{n_t-1} f(t_n)\gamma_n + (t-t_{n_t})f(t_{n_t})\tilde{\gamma}_{n_t}, \tag{35}$$

where $\tilde{\gamma}_{n_t} \in N\left(0, \frac{1}{(t-t_{n_t})^{1/2}}\right)$.

In (35) the length of the last step is less than h_m. Therefore, in accordance with the comments regarding the non-uniform grids made before the statement of this theorem it should define for $t_{n_t} < t < t_{n_t} + h_m$ the process λ with parameters corresponding to the length of the segment $[0, t - t_{n_t}]$. This process denoted as $\tilde{\lambda}$ is defined by the equation

$$\tilde{\lambda}_{t-t_{n_t}} = e^{-\beta(t-t_{n_t})}\tilde{\lambda}_{t_{n_t}} + \sqrt{\frac{2\beta}{(t-t_{n_t})(1-e^{-2\beta(t-t_{n_t})})}} \times e^{-\beta(t-t_{n_t})} \int_0^{t-t_{n_t}} e^{\beta s} dW(s+t_{n_t}), \tag{36}$$

where $\tilde{\lambda}_{t_{n_t}}$ is a normal distributed random value with zero expectation and the variance $\frac{1}{(t-t_{n_t})(1-e^{-2\beta(t-t_{n_t})})}$.

We obtain from (18), (35) the following

$$h_m \sum_{n=0}^{n_t-1} f(t_n)\lambda(t_{n+1}) + (t-t_{n_t})f(t_{n_t})\tilde{\lambda}(t-t_{n_t}) = h_m \sum_{n=0}^{n_t-1} f(t_n)\varepsilon_n + (t-t_{n_t})f(t_{n_t})e^{-\beta(t-t_{n_t})}\tilde{\lambda}_{t_{n_t}} + \int_0^T f_m(v)dW^\gamma(v). \tag{37}$$

As it follows from the proof of the theorem 1, the presence the factor $\exp(-\beta h_m)$ in the random values ε_n entails convergence to zero as $\beta \to \infty$ of the variance of the random value $h_m \sum\limits_{n=0}^{n_t-1} f(t_n)\varepsilon_n$.

Let us show that for any $t \in [t_{n_t}, t_{n_t} + h_m]$ the variance of $\tilde{\varepsilon}_t \equiv (t-t_{n_t})f(t_{n_t})e^{-\beta(t-t_{n_t})}\tilde{\lambda}_{t_{n_t}}$ converges to zero as $\beta \to \infty$.

Indeed, the variance of $\tilde{\varepsilon}_t$ is

$$D\tilde{\varepsilon}_t = f^2(t_{n_t})\frac{(t-t_{n_t})e^{-2\beta(t-t_{n_t})}}{1-e^{-2\beta(t-t_{n_t})}}. \tag{38}$$

The limit (38) under $t \searrow t_{n_t}$ can be defined using L'Hopital's rule, and it is equaled $\frac{f^2(t_{n_t})}{2\beta}$.

The derivative $D\tilde{\varepsilon}_t$ with respect to t is defined from the equality

$$(D\tilde{\varepsilon}_t)'_t = f^2(t_{n_t})e^{-2\beta(t-t_{n_t})}\frac{1-2\beta(t-t_{n_t})-e^{-2\beta(t-t_{n_t})}}{(1-e^{-2\beta(t-t_{n_t})})^2}. \tag{39}$$

From (39) it is clear that the derivative of the function $D\tilde{\varepsilon}_t$ becomes negative on the half-interval $(t_{n_t}, t_{n_t} + h_m]$. Therefore, the variance of $\tilde{\varepsilon}_t$ decreases on $(t_{n_t}, t_{n_t} + h_m]$, and its values do not exceed $\frac{f^2(t_{n_t})}{2\beta}$. Because $f^2(t_{n_t})$ is a.s. bounded, it takes place the convergence $D\tilde{\varepsilon}_t \to 0$ as $\beta \to \infty$. From this fact and from the theorem 1 it follows that the variance of the random value

$$h \sum_{n=0}^{n_t-1} f(t_n)\varepsilon_n + (t-t_{n_t})f(t_{n_t})e^{-\beta(t-t_{n_t})}\tilde{\lambda}_{t_{n_t}}$$

converges to zero as $\beta \to \infty$. Consequently, there is a convergence

$$E\Big|h \sum_{n=0}^{n_t-1} f(t_n)\varepsilon_n + (t-t_{n_t})f(t_{n_t})e^{-\beta(t-t_{n_t})}\tilde{\lambda}_{t_{n_t}}\Big| \to 0 \quad \beta \to \infty \tag{40}$$

because the following inequality is valid

$$E\Big|h \sum_{n=0}^{n_t-1} f(t_n)\varepsilon_n + (t-t_{n_t})f(t_{n_t})e^{-\beta(t-t_{n_t})}\tilde{\lambda}_{t_{n_t}}\Big| \leq \Big(E\Big|h \sum_{n=0}^{n_t-1} f(t_n)\varepsilon_n + (t-t_{n_t})f(t_{n_t})e^{-\beta(t-t_{n_t})}\tilde{\lambda}_{t_{n_t}}\Big|^2\Big)^{\frac{1}{2}}. \tag{41}$$

Therefore, there exists $\beta(\varepsilon)$ that under $\beta > \beta(\varepsilon)$ it follows from (37) and (40)

$$E\Big|h_m \sum_{n=0}^{n_t-1} f(t_n)\lambda(t_{n+1}) + (t-t_{n_t})f(t_{n_t})\tilde{\lambda}(t-t_{n_t}) - \int_0^T f_m(v)dW^\gamma(v)\Big| = E\Big|h_m \sum_{n=0}^{n_t-1} f(t_n)\varepsilon_n + (t-t_{n_t})f(t_{n_t})e^{-\beta(t-t_{n_t})}\tilde{\lambda}_{t_{n_t}}\Big| \leq \frac{\varepsilon}{9} \ . \tag{42}$$

By definition, the integral $\int_0^T f(s)\lambda(s)ds$ is a mean-square limit of integrals of step functions. Therefore, there exits an integer $m_2(\varepsilon)$ that under $m > m_2((\varepsilon)$ the following inequality is valid

$$E\Big|\int\limits_0^T f(s)\lambda(s)ds - h_m\sum_{n=0}^{n_t-1} f(t_n)\lambda(t_{n+1})$$
$$+(t-t_{n_t})f(t_{n_t})\tilde{\lambda}(t-t_{n_t})\Big| \le \frac{\varepsilon}{9}. \qquad (43)$$

From the inequalities (34), (42), (43), in which $\beta \ge \beta(\varepsilon)$ and $m \ge \max\big(m_1(\varepsilon), m_2(\varepsilon)\big)$, it follows

$$E\Big|\int\limits_0^T f(v)dW^\gamma(v) - \int\limits_0^T f(v)\lambda(v)dv\Big| \le \varepsilon.$$

The proof is completed.

Further we use two lemmas from [3].

Lemma 1: . For any integer $p \ge 1$

$$E_{t,x}|\tau(\theta+\Delta\theta) - \tau(\theta)|^p \to 0 \text{ as } \Delta\theta \to 0 \ . \qquad (44)$$

Lemma 2: There exists a constant $K > 0$, such that, as $\Delta\theta \to 0$

$$E_{t,x}\Big|\frac{\tau(\theta+\Delta\theta) - \tau(\theta)}{\Delta\theta}\Big| < K. \qquad (45)$$

For a function $r(x,\theta)$ such that $r \in C^1(R^{d+1} \to R)$ we use the designation $\frac{d}{d\theta}r(X.(\theta),\theta) = \frac{\partial r}{\partial x}\frac{\partial X.(\theta)}{\partial\theta} + \frac{\partial r}{\partial\theta}$ where $\frac{\partial X.(\theta)}{\partial\theta}$ is a mean-square derivative.

We shall now proceed to the construction of a method for the determination of estimates of functional (2) using the change the Ito integral over Wiener process by the integral of the type $\int_0^T f(s)\lambda(s)ds$.

It is further assumed that the processes $X.$, $Z_s(\theta)$ are defined on the base of the Wiener process W^γ, i.e. $X.$, $Z_s(\theta)$ are solutions of the equations:

$$X_s(\theta) = x + \int\limits_t^s a(v, X_v(\theta), \theta)dv + \int\limits_t^s \sigma(v, X_v(\theta), \theta)dW^\gamma(v) \qquad (46)$$

$$Z_s(\theta) = \int\limits_t^s \left(\frac{\partial a}{\partial x}Z_v(\theta) + \frac{\partial a}{\partial\theta}\right)dv$$
$$+\int\limits_t^s \left(\frac{\partial\sigma}{\partial x}Z_v(\theta) + \frac{\partial\sigma}{\partial\theta}\right)dW^\gamma(v) \qquad (47)$$

We define a multidimensional stationary Gaussian process $\lambda = (\lambda_1, \dots, \lambda_d)$. Each component λ_j, $j = 1, \dots, d$ is a solution of the equation

$$d\lambda_j = -\beta\lambda_j dv + \vartheta dW_j(v) \quad , \qquad (48)$$

where $\lambda_j(0) = \lambda_{j0}$ is a Gaussian random value with zero expectation and the variance $\frac{\vartheta^2}{2\beta}$.

The formula for $\partial u/\partial\theta$ in the next theorem can be considered as a variant of the main result in [3], but it was obtained under less restrictive conditions than in [3] on the smoothness of the solved problem.

Theorem 3: Let $g(t,x)$ be a function twice continuously differentiable with respect to x and continuously differentiable with respect to t. The function g vanishes on ∂G, and Lg is not equaled to zero on ∂G.

*Let the coefficients of the SDE (1) and the functions φ, f satisfy to the conditions **(A)** — **(E)**.*

Then estimates of $\frac{\partial u}{\partial\theta}$ can be obtained from the equation

$$\frac{\partial u}{\partial\theta}(t,x) = E_{t,x}\Big[\Big(\frac{\partial\varphi}{\partial x}(T, X_T, \theta)Z_T + \frac{\partial\varphi}{\partial\theta}(T, X_T, \theta)\Big)\chi_{\tau>T}$$
$$+\Big(\frac{\partial\psi}{\partial x}(\tau, X_\tau, \theta)Z_\tau + \frac{\partial\psi}{\partial\theta}(\tau, X_\tau, \theta)\Big)\chi_{\tau<T}$$
$$+\int\limits_t^{T\wedge\tau}\Big(\frac{\partial f}{\partial x}(v, X_v, \theta)Z_v + \frac{\partial f}{\partial\theta}(v, X_v, \theta)\Big)dv\Big] \qquad (49)$$

$$-\Big(\frac{f(\tau, X_\tau) + \frac{\partial\psi}{\partial t}(\tau, X_\tau)}{R(\tau, X_\tau)}\int\limits_t^\tau \frac{d}{d\theta}R(v, X(v))dv\Big)\chi(\tau < T)) \qquad (50)$$

where

$$R(v, X(v)) = \Big(Lg + \sum_{ij}\frac{\partial g}{\partial x_i}\sigma_{ij}\lambda_j\Big)\Big|_{(v,X(v))}, \qquad (51)$$

λ_j is a solution of the SDE

$$d\lambda_j = -\beta\lambda_j dv + \sqrt{\frac{2\beta}{h(1-e^{-2\beta h})}}dW_j(v), \qquad (52)$$

β is a constant such that $\beta \gg \frac{1}{h}$.

Proof. We introduce the designation $R_j^\mu(v,\theta) = \sum\limits_l \frac{\partial\mu(X_v(\theta))}{\partial x_l} \cdot\sigma_{lj}(v, X_v(\theta), \theta)$ for a function $\mu \in C^1\big(R \times R^d \to R\big)$.

Applying the Ito formula to $g(t,x)$ we obtain the equality:

$$0 = g(t,x) + \int\limits_t^\tau Lg(v, X_v)dv + \int\limits_t^\tau\sum_j R_j^g dW_j^\gamma(v) \qquad (53)$$

We add and subtract in (53) the integral $\int\limits_t^{\tau(\theta)}\sum\limits_j R_j^g(v,\theta)\cdot\lambda_j(v)dv$ and rewrite this equation as follows:

$$0 = g(t,x) + \int\limits_t^\tau R(v, X(v))dv + \int\limits_t^\tau\sum_j R_j^g(v,\theta)dW_j^\gamma(v)$$
$$-\int\limits_t^\tau\sum_j R_j^g(v,\theta)\lambda_j(v)dv. \qquad (54)$$

We designate $\tilde{\tau}(\theta,\Delta\theta)\equiv\tau(\theta)\wedge\tau(\theta+\Delta\theta)$. Using equations obtained from (54) for θ and $\theta+\Delta\theta$ we write the equality

$$0=\frac{1}{\Delta\theta}\Big[\int\limits_t^{\tilde{\tau}(\theta,\Delta\theta)}\left(R(v,X_v(\theta+\Delta\theta),\theta+\Delta\theta)-R(v,X_v(\theta),\theta)\right)dv$$
$$+\int\limits_{\tilde{\tau}(\theta,\Delta\theta)}^{\tau(\theta+\Delta\theta)}R(v,X_v(\theta+\Delta\theta),\theta+\Delta\theta)dv-\int\limits_{\tilde{\tau}(\theta,\Delta\theta)}^{\tau(\theta)}R(v,X_v(\theta),\theta)dv$$
$$+\int\limits_t^{\tilde{\tau}(\theta,\Delta\theta)}\sum_j R_j^g(v,\theta+\Delta\theta)dW_j^\gamma(v)$$
$$-\int\limits_t^{\tilde{\tau}(\theta,\Delta\theta)}\sum_j R_j^g(v,\theta+\Delta\theta)\lambda_j(v)dv$$
$$-\int\limits_t^{\tilde{\tau}(\theta,\Delta\theta)}\sum_j R_j^g(v,\theta)dW_j^\gamma(v)+\int\limits_t^{\tilde{\tau}(\theta,\Delta\theta)}\sum_j R_j^g(v,\theta)\lambda_j(v)dv$$
$$+\int\limits_{\tilde{\tau}(\theta,\Delta\theta)}^{\tau(\theta+\Delta\theta)}\sum_j R_j^g(v,\theta+\Delta\theta)dW_j^\gamma(v)$$
$$-\int\limits_{\tilde{\tau}(\theta,\Delta\theta)}^{\tau(\theta+\Delta\theta)}\sum_j R_j^g(v,\theta+\Delta\theta)\lambda_j(v)dv$$
$$-\int\limits_{\tilde{\tau}(\theta,\Delta\theta)}^{\tau(\theta)}\sum_j R_j^g(v,\theta)dW_j^\gamma(v)+\int\limits_{\tilde{\tau}(\theta,\Delta\theta)}^{\tau(\theta)}\sum_j R_j^g(v,\theta)\lambda_j(v)dv\Big]\ . \tag{55}$$

Let us multiply (55) by $\chi_{\tau<T}\frac{f(\tau,X_\tau)+\frac{\partial\psi}{\partial t}(\tau,X_\tau)}{R(\tau,X_\tau)}$ where τ is taken at the parameter value θ. Then we consider a limit as $\Delta\theta\to0$ of the expectation of the right-hand send of the obtained equality.

Let us show that

$$\lim_{\Delta\theta\to0}E_{t,x}\Big[\chi_{\tau(\theta)<T}\frac{f(\tau,X_\tau)+\frac{\partial\psi}{\partial t}(\tau,X_\tau)}{R(\tau,X_\tau)}\cdot$$
$$\int\limits_t^{\tilde{\tau}(\theta,\Delta\theta)}\frac{R(v,X_v(\theta+\Delta\theta),\theta+\Delta\theta)-R(v,X_v(\theta),\theta)}{\Delta\theta}dv\Big]$$
$$=E_{t,x}\Big[\chi_{\tau(\theta)<T}\frac{f(\tau,X_\tau)+\frac{\partial\psi}{\partial t}(\tau,X_\tau)}{R(\tau,X_\tau)}\cdot$$
$$\int\limits_t^{\tau(\theta)}\frac{d}{d\theta}R(v,X_v(\theta),\theta)dv\Big]\ . \tag{56}$$

For this object we write the inequality which is followed from the fact that the random value $\chi_{\tau(\theta)<T}\frac{f(\tau,X_\tau)+\frac{\partial\psi}{\partial t}(\tau,X_\tau)}{R(\tau,X_\tau)}$ is a.s. bounded.

$$\Big|E_{t,x}\Big[\chi_{\tau(\theta)<T}\frac{f(\tau,X_\tau)+\frac{\partial\psi}{\partial t}(\tau,X_\tau)}{R(\tau,X_\tau)}\cdot$$
$$\Big(\int\limits_t^{\tilde{\tau}(\theta,\Delta\theta)}\frac{R(v,X_v(\theta+\Delta\theta),\theta+\Delta\theta)-R(v,X_v(\theta),\theta)}{\Delta\theta}dv$$
$$-\int\limits_t^{\tau(\theta)}\frac{d}{d\theta}R(v,X_v(\theta),\theta)dv\Big)\Big]\Big|\le C\int\limits_t^T\Big|E_{t,x}\Big[\chi_{v<\tilde{\tau}(\theta,\Delta\theta)}\cdot$$
$$\Big(\frac{R(v,X_v(\theta+\Delta\theta),\theta+\Delta\theta)-R(v,X_v(\theta),\theta)}{\Delta\theta}$$
$$-\frac{d}{d\theta}R(v,X_v(\theta),\theta)\Big)\Big]\Big|dv$$
$$+C\Big|E_{t,x}\Big[\int\limits_{\tilde{\tau}(\theta,\Delta\theta)}^{\tau(\theta)}\frac{d}{d\theta}R(v,X_v(\theta),\theta)dv\Big]\Big|, \tag{57}$$

where $C=\sup\limits_{(v,x,\theta)\in[0,T]\times\partial G\times U}\left|\frac{f(v,x,\theta)+\frac{\partial\psi}{\partial t}(\tau,X_\tau)}{R(v,x,\theta)}\right|$.

The derivatives of a_i, σ_{ij}, g x, θ are continuous in $\overline{G}$ and hence they are bounded , then for $X.$ the conditions of mean-square differentiability with respect to θ are met. The processes λ_j have a bounded variance on $[0,T]$ and do not dependent on θ, therefore it is easy to show that the first term in the right-hand side of (57) converges to zero as $\Delta\theta\to0$.

Let us show that the second term in the right-hand side (57) converges to zero as $\Delta\theta\to0$. It follows by the definition of the function R that

$$\int\limits_{\tilde{\tau}(\theta,\Delta\theta)}^{\tau(\theta)}\frac{d}{d\theta}R(v,X_v(\theta),\theta)dv=\int\limits_{\tilde{\tau}(\theta,\Delta\theta)}^{\tau(\theta)}\frac{d}{d\theta}Lg(v,X_v(\theta),\theta)dv+$$
$$\int\limits_{\tilde{\tau}(\theta,\Delta\theta)}^{\tau(\theta)}\sum_{i,j}\frac{d}{d\theta}\Big(\frac{\partial g}{\partial x_i}\sigma_{ij}(v,X_v(\theta),\theta)\Big)\lambda_j(v)dv. \tag{58}$$

Because the derivatives of Lg are continuous in $\overline{G}$, the value $\frac{d}{d\theta}Lg(v,X_v(\theta),\theta)$ is bounded. Therefore, mathematical expectation of the first term in the right-hand side of (58) converges to zero as $\Delta\theta\to0$ by Lemma1

$$\Big|E_{t,x}\Big[\int\limits_{\tilde{\tau}(\theta,\Delta\theta)}^{\tau(\theta)}\frac{d}{d\theta}Lg(v,X_v(\theta),\theta)dv\Big]\Big|\le C_1E_{t,x}\Big|\tau(\theta)-\tilde{\tau}(\theta,\Delta\theta)\Big|$$
$$\le C_1E_{t,x}\big|\Delta\tau\big|\to0\quad\Delta\theta\to0\ ,$$

where $C_1=\sup\limits_{(v,x,\theta)\in[0,T]\times\partial G\times U}\Big|\frac{d}{d\theta}Lg(v,X_v(\theta),\theta)\Big|,\quad\Delta\tau=\tau(\theta+\Delta\theta)-\tau(\theta)$.

The convergence of the expectation of the second term in (58) as $\Delta\theta\to0$ follows from the inequalities

$$E_{t,x}\Big[\int\limits_{\tilde{\tau}(\theta,\Delta\theta)}^{\tau(\theta)}\sum_{i,j}\frac{d}{d\theta}\Big(\frac{\partial g}{\partial x_i}\sigma_{ij}\Big)\lambda_jdv\Big]$$
$$\le\sum_{i,j}E_{t,x}\Big[\Big(\int\limits_{\tilde{\tau}(\theta,\Delta\theta)}^{\tau(\theta)}\Big(\frac{d}{d\theta}\Big(\frac{\partial g}{\partial x_i}\sigma_{ij}\Big)\Big)^2dv\Big)^{\frac12}\Big(\int\limits_{\tilde{\tau}(\theta,\Delta\theta)}^{\tau(\theta)}\lambda_j^2dv\Big)^{\frac12}\Big]$$
$$\le\sum_{i,j}\Big[E_{t,x}\int\limits_{\tilde{\tau}(\theta,\Delta\theta)}^{\tau(\theta)}\Big(\frac{d}{d\theta}\Big(\frac{\partial g}{\partial x_i}\sigma_{ij}\Big)\Big)^2dv\Big]^{\frac12}\cdot\Big[E_{t,x}\int\limits_{\tilde{\tau}(\theta,\Delta\theta)}^{\tau(\theta)}\lambda_j^2dv\Big]^{\frac12}$$
$$\le C_2\Big(\sum_j\int\limits_t^TE_{t,x}\lambda_j^2dv\Big)E_{t,x}\big(\tau(\theta)-\tilde{\tau}(\theta,\Delta\theta)\big)^{\frac12}\to0 \tag{59}$$

as $\Delta\theta\to0,$ where $C_2=\max\limits_{\substack{i,j=1,\dots,d\\(v,x,\theta)\in[0,T]\times\bar{G}\times U}}\Big|\frac{d}{d\theta}\Big(\frac{\partial g}{\partial x_i}\sigma_{ij}\Big)\Big|.$

We use further the designation $R_\tau = R(\tau, X_\tau, \theta)$.

Let us prove that under $\Delta\theta \to 0$

$$E_{t,x}\left[\chi_{\tau(\theta)<T}\frac{f(\tau,X_\tau)+\frac{\partial\psi}{\partial t}(\tau,X_\tau)}{\Delta\theta R_\tau}\cdot\left(\int\limits_{\tilde\tau(\theta,\Delta\theta)}^{\tau(\theta+\Delta\theta)}R(v,X_v(\theta+\Delta\theta),\theta+\Delta\theta)dv-\int\limits_{\tilde\tau(\theta,\Delta\theta)}^{\tau(\theta)}R(v,X_v(\theta),\theta)dv\right)\right]$$
$$-E_{t,x}\left[\chi_{\tau(\theta)<T}\big(f(\tau,X_\tau)+\frac{\partial\psi}{\partial t}(\tau,X_\tau)\big)\frac{\Delta\tau}{\Delta\theta}\right]\to 0. \tag{60}$$

The following equality holds for the integrals in (60)

$$\int\limits_{\tilde\tau(\theta,\Delta\theta)}^{\tau(\theta+\Delta\theta)}R(v,X_v(\theta+\Delta\theta),\theta+\Delta\theta)dv-\int\limits_{\tilde\tau(\theta,\Delta\theta)}^{\tau(\theta)}R(v,X_v(\theta),\theta)dv$$
$$=\int\limits_{\tau(\theta)}^{\tau(\theta+\Delta\theta)}R(v,X_v(\theta+\eta\Delta\theta),\theta+\eta\Delta\theta)dv, \tag{61}$$

where $\eta=\begin{cases}1, & \text{if } \tau(\theta)<\tau(\theta+\Delta\theta),\\ 0, & \text{if } \tau(\theta+\Delta\theta)<\tau(\theta).\end{cases}$

Because R is continuous, then by the mean value theorem there exits γ, $\gamma\in[\tau(\theta)\wedge\tau(\theta+\Delta\theta),\tau(\theta)\vee\tau(\theta+\Delta\theta)]$ such, that the equality holds

$$\int\limits_{\tau(\theta)}^{\tau(\theta+\Delta\theta)}R(v,X_v(\theta+\eta\Delta\theta),\theta+\eta\Delta\theta)dv$$
$$=R(\gamma,X_\gamma(\theta+\eta\Delta\theta),\theta+\eta\Delta\theta)\Delta\tau. \tag{62}$$

Let us denote $C_3=\sup\limits_{(v,x,\theta)\in[0,T]\times\partial G\times U}\left|\frac{f(v,x,\theta)}{R(v,x,\theta)}\right|$. We rewrite the difference of the mathematical expectations in (60) taking into account (61), (62) and obtain the following estimate

$$E_{t,x}\left[\chi_{\tau(\theta)<T}\frac{f(\tau,X_\tau)+\frac{\partial\psi}{\partial t}(\tau,X_\tau)}{\Delta\theta R_\tau}\cdot\right.$$
$$\left.\left(R(\gamma,X_\gamma(\theta+\eta\Delta\theta),\theta+\eta\Delta\theta)-R_\tau\right)\Delta\tau\right]\le$$
$$C_3\left(E_{t,x}\Big[\frac{R(\gamma,X_\gamma(\theta+\eta\Delta\theta),\theta+\eta\Delta\theta)-R_\tau}{\Delta\theta}\Big]^2\Delta\theta\right)^{\frac12}\cdot$$
$$\left(E_{t,x}\frac{(\Delta\tau)^2}{\Delta\theta}\right)^{\frac12}. \tag{63}$$

The value of $\frac{R(\gamma,X_\gamma(\theta+\eta\Delta\theta),\theta+\eta\Delta\theta)-R_\tau}{\Delta\theta}$ is bounded when $\Delta\theta\to 0$ because of the continuously of derivatives R in all arguments and the existence of the mean-square derivative $\frac{\partial X}{\partial\theta}$.

The value $E_{t,x}\Big[\frac{(\Delta\tau)^2}{\Delta\theta}\Big]$ is bounded when $\Delta\theta\to 0$, because the inequality (15) from [3] for any integer p is valid $E_{t,x}|\tau(\theta+\Delta\theta)-\tau(\theta)|^p \le C(p)E_{t,x}|X_{\tilde\tau(\theta,\Delta\theta)}(\theta+\Delta\theta)-X_{\tilde\tau(\theta,\Delta\theta)}(\theta)|$, where $\tilde\tau(\theta,\Delta\theta)=\tau(\theta)\wedge\tau(\theta+\Delta\theta)$, $C(p)$ is a constant depending on p.

So, we have

$$E_{t,x}\Big[\frac{(\Delta\tau)^2}{\Delta\theta}\Big]\le C(2)E_{t,x}\Big|\frac{X_{\tilde\tau(\theta,\Delta\theta)}(\theta+\Delta\theta)-X_{\tilde\tau(\theta,\Delta\theta)}(\theta)}{\Delta\theta}\Big|\le$$
$$C(2)\Big(E_{t,x}\Big(\frac{X_{\tilde\tau(\theta,\Delta\theta)}(\theta+\Delta\theta)-X_{\tilde\tau(\theta,\Delta\theta)}(\theta)}{\Delta\theta}\Big)^2\Big)^{\frac12}$$
$$\to C(2)\Big(E_{t,x}\Big(\frac{\partial X_\tau}{\partial\Delta\theta}\Big)^2\Big)^{\frac12}. \tag{64}$$

Thus, it has been shown that under $\Delta\theta\to 0$ the right-hand side of (63) converges to zero, and, consequently the assertion (60) is valid.

Let us show that under $\Delta\theta\to 0$ the sum of terms containing integrals with R_j^g can be close to zero by selecting parameters of λ_j.

Really, we have under $\Delta\theta\to 0$

$$\frac{1}{\Delta\theta}\Big[\int\limits_t^{\tilde\tau(\theta,\Delta\theta)}\sum_j R_j^g(v,\theta+\Delta\theta)dW_j^\gamma(v)$$
$$-\int\limits_t^{\tilde\tau(\theta,\Delta\theta)}\sum_j R_j^g(v,\theta+\Delta\theta)\lambda_j(v)dv$$
$$-\int\limits_t^{\tilde\tau(\theta,\Delta\theta)}\sum_j R_j^g(v,\theta)dW_j^\gamma(v)+\int\limits_t^{\tilde\tau(\theta,\Delta\theta)}\sum_j R_j^g(v,\theta)\lambda_j(v)dv\Big]\to$$
$$\int\limits_t^{\tilde\tau(\theta,\Delta\theta)}\sum_j\frac{d}{d\theta}R_j^g(v,\theta)dW_j^\gamma(v)-\int\limits_t^{\tilde\tau(\theta,\Delta\theta)}\sum_j\frac{d}{d\theta}R_j^g(v,\theta)\lambda_j(v)dv. \tag{65}$$

By the Theorem1 the difference $\int\limits_t^{\tilde\tau(\theta,\Delta\theta)}\sum_j\frac{d}{d\theta}R_j^g(v,\theta)dW_j^\gamma(v)-\int\limits_t^{\tilde\tau(\theta,\Delta\theta)}\sum_j\frac{d}{d\theta}R_j^g(v,\theta)\lambda_j(v)dv$ can be approached to zero by selecting the parameters of λ_j.

Based on the above considerations, the convergence to zero of the sum of terms in (55) containing the rest integrals with the functions R_j^g in the integrands is evident. The proof of the theorem is complete.

V. Conclusion

The paper proposed and justified a numerical statistical method for estimating parametric derivatives of functionals of diffusion processes with absorbing boundary conditions.

References

[1] S. A. Gusev, "Monte Carlo estimates of the solution of a parabolic equation and its derivatives made by solving stochastic differential equations", Communications in Nonlinear Science and Numerical Simulation, Vol.9, Issue 2, pp. 177-185, April 2004.

[2] S. A. Gusev, "Estimation of derivatives with respect to parameters of a functional of a diffusion process moving in a domain with absorbing boundary", Numer. Anal. Appl., 1, pp. 314331, 2008 .

[3] S. A. Gusev and N. G. Dokuchaev, "On Differentiation of Functionals Containing the First Exit of a Diffusion Process from a Domain", Theory Probab. Appl., 59(1), 136-144, 2015.

[4] O. A. Ladyzhenskaya, V. A. Solonnikov and N.N. Uraltseva *Linear and quasilinear equations of the parabolic type*, Nauka, Moskow, 1967 (in Russian).

[5] I. I. Gikhman and A. V. Skorokhod, *Stochastic Differential Equations*, Naukova Dumka, Kiev, 1968 (in Russian).

[6] V. I. Tikhonov, V. A. Mihronov, *Markovian processes*, Sovetskoe radio, Moskow, 1977 (in Russian).

[7] T. A. Averina, S. S. Artemiev, "Simulation stationary random processes with given one-dimensional distribution and exponential correlation function," Preprint No 495, Computing Center, USSR Akad. Sci., Novosibirsk, 1984 (in Russian).

[8] A. A. Borovkov, *Probability theory*, Editorial URSS, Moskow,1999 (in Russian).
[9] I. I. Gikhman and A. V. Skorokhod, *Introduction to the Theory of Random Processes*, Nauka, Moscow, 1977 (in Russian).