



Adaptive Itô–Taylor algorithm can optimally approximate the Itô integrals of singular functions[☆]

Paweł Przybyłowicz^{*}

Department of Applied Mathematics, AGH University of Science and Technology, Al. Mickiewicza 30, 30-059 Cracow, Poland

ARTICLE INFO

Article history:

Received 21 July 2009

Received in revised form 10 March 2010

MSC:

68Q25

65C30

Keywords:

Stochastic Itô integrals

Singular problems

Optimal algorithm

Standard information

r -fold integrated Brownian motion

ABSTRACT

We deal with numerical approximation of stochastic Itô integrals of singular functions. We first consider the regular case of integrands belonging to the Hölder class with parameters r and ϱ . We show that in this case the classical Itô–Taylor algorithm has the optimal error $\mathcal{O}(n^{-(r+\varrho)})$. In the singular case, we consider a class of piecewise regular functions that have continuous derivatives, except for a finite number of unknown singular points. We show that any nonadaptive algorithm cannot efficiently handle such a problem, even in the case of a single singularity. The error of such algorithm is no less than $n^{-\min\{1/2, r+\varrho\}}$. Therefore, we must turn to adaptive algorithms. We construct the adaptive Itô–Taylor algorithm that, in the case of at most one singularity, has the optimal error $\mathcal{O}(n^{-(r+\varrho)})$. The best speed of convergence, known for regular functions, is thus preserved. For multiple singularities, we show that any adaptive algorithm has the error $\Omega(n^{-\min\{1/2, r+\varrho\}})$, and this bound is sharp.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

The goal of this paper is to derive optimal bounds on the minimal error for approximation of stochastic Itô integrals of the form

$$I(f, B) = \int_0^T f(t) dB_t, \quad (1)$$

where $T > 0$. We assume that B is a standard one-dimensional Brownian motion, and a deterministic function $f : [0, T] \rightarrow \mathbb{R}$ is possibly singular. By a singular function we mean a function f which may be discontinuous and/or have r discontinuous derivatives at a finite number of unknown singular points.

In order to compute an approximation of $I(f, B)$ we assume that standard information, given by values of integrands f and r -fold integrated Brownian motion $B^{(r)}$ at a finite number of points, is available. Such information about process $B^{(r)}$ is often used in the general Itô–Taylor schemes in order to approximate solutions of stochastic differential equations (see [1,2]). We will consider three classes of algorithms, denoted by Ψ_n , Ψ_n^* and Ψ_n^{**} . The index n denotes the total number of evaluations of f , $B^{(r)}$ and its derivatives used by algorithms from these classes. First class Ψ_n contains nonadaptive algorithms. Algorithms that are adaptive with respect to function f , but use the same sampling points for every trajectory of the process $B^{(r)}$, define the class Ψ_n^* . Finally, fully adaptive algorithms that choose sampling points in adaptive way with respect to function f and process $B^{(r)}$, define the class Ψ_n^{**} . The error of an algorithm is defined in the worst case sense by taking the square root of the maximum mean squared error over the class of integrands f that are considered in this paper.

[☆] The author was partially supported by the Polish Ministry of Science and Higher Education under Grant No. N N201 547738 and by AGH Grant No. 10.420.03.

^{*} Tel.: +48 794 428 645.

E-mail address: przybyl83@gmail.com.

A number of papers on singular problems have recently appeared in the literature, just to mention such articles as [3–6] or [7] which deal with approximation of scalar Lebesgue integrals, singular functions, and solutions of singular initial-value problems.

In this paper we deal with singular problems of the form (1). We start with the regular case when integrands f belong to the Hölder class $\mathcal{F}_{\text{reg}}^{r,\varrho}$ with regularity parameters $r \geq 0$ and $\varrho \in (0, 1]$. For the classical Itô–Taylor algorithm \mathcal{A}_n^{IT} we show that the error is $O(n^{-(r+\varrho)})$ (Proposition 1). Lower bounds are $\Omega(n^{-(r+1)})$ for algorithms from the class Ψ_n^{**} , and $\Omega(n^{-(r+\varrho)})$ for algorithms that belong to Ψ_n^* (see Corollary 1 and Theorem 2, respectively). This establishes the optimality of the nonadaptive Itô–Taylor scheme \mathcal{A}_n^{IT} (Corollary 2). This also means that we do not need to use adaptive algorithms in the regular case. In the singular case things go different. We consider the class $\mathcal{F}_{\text{sng},p}^{r,\varrho}$ of piecewise smooth functions f which have at most p unknown singular points. It turns out that nonadaptive algorithms cannot handle such problems, even in the case of a single singular point, and the error of such algorithms is no less than $n^{-\min\{1/2, r+\varrho\}}$ (Proposition 2). Therefore, we turn to adaptive algorithms. For $p = 1$ we construct the adaptive Itô–Taylor algorithm \mathcal{A}_n^{IT*} that detects a singular point and next suitably modifies starting discretization of the interval $[0, T]$. We show in Theorem 3 that the algorithm \mathcal{A}_n^{IT*} preserves the optimal error $O(n^{-(r+\varrho)})$, known from the regular case. For $p \geq 2$ we show that even adaptive algorithms from the class Ψ_n^* have the error $\Omega(n^{-\min\{1/2, r+\varrho\}})$ and the Itô–Taylor algorithm \mathcal{A}_n^{IT} is optimal in this case (Proposition 3 and Corollary 4). The question about lower bounds for the algorithms from the class Ψ_n^{**} for $p \geq 2$ is left open.

The paper is organized as follows. Problem formulation and basic definitions are given in Section 2. Regular case, along with upper and lower bounds on the n th minimal error, is investigated in Section 3. In Section 4 we define the optimal adaptive algorithm and present optimal bounds in the class $\mathcal{F}_{\text{sng},1}^{r,\varrho}$ of functions with at most one unknown singularity. The case of multiple singular points is discussed in Section 5. In Section 6 we summarize main results and give final remarks. Appendix contains the proof of some auxiliary result.

2. Problem formulation

For a given real number $T > 0$ we study optimal approximation of the following Itô integral,

$$I(f, B) = \int_0^T f(t) dB_t, \quad (2)$$

where $f : [0, T] \rightarrow \mathbb{R}$ and B is the standard one-dimensional Brownian motion.

We shall consider two cases: of regular or piecewise regular functions f . More precisely, we deal with the following classes of functions. For a given integer $r \geq 0$ and nonnegative real numbers L and $\varrho \in (0, 1]$ we define the class of Hölder regular functions

$$\mathcal{F}_{\text{reg}}^{r,\varrho} = \{f : [0, T] \rightarrow \mathbb{R} \mid f \in C^{(r)}([0, T]), |f^{(r)}(t) - f^{(r)}(s)| \leq L|s - t|^\varrho, s, t \in [0, T]\}.$$

A special case of (2), for $f \in \mathcal{F}_{\text{reg}}^{r,\varrho}$ with $r = 0$ and $\varrho = 1$, was investigated in [8], while the case with $\varrho = r = 1$ was considered in [9].

We also consider the class $\mathcal{F}_{\text{sng},p}^{r,\varrho}$ of scalar singular functions f . For given integers $r \geq 0$, $p \geq 1$ and nonnegative real numbers $\varrho \in (0, 1]$, M_0, M_1, \dots, M_r and L , function $f \in \mathcal{F}_{\text{sng},p}^{r,\varrho}$ satisfies the following conditions:

- Function f has r continuous derivatives in \mathbb{R} , except for at most p (unknown) singular points $0 < u_f^1 < u_f^2 < \dots < u_f^p < T$, i.e.,

$$f \in C^{(r)}([0, u_f^1)) \cap C^{(r)}([u_f^1, u_f^2)) \cap \dots \cap C^{(r)}([u_f^{p-1}, u_f^p)) \cap C^{(r)}([u_f^p, T]). \quad (3)$$

We set $f^{(j)}(u_f^q) = f^{(j)}(u_f^{q+})$ and we assume that there exists finite limits $f^{(j)}(u_f^{q-}) = \lim_{t \rightarrow u_f^q-} f^{(j)}(t)$ for $j = 0, 1, \dots, r$ and $q = 1, 2, \dots, p$.

- The jumps $\Delta_f^{j,q}$ at points u_f^q , defined by

$$\Delta_f^{j,q} = f^{(j)}(u_f^{q+}) - f^{(j)}(u_f^{q-}), \quad (4)$$

are bounded,

$$|\Delta_f^{j,q}| \leq M_j, \quad (5)$$

for $j = 0, 1, \dots, r$ and $q = 1, 2, \dots, p$.

- The r th derivative of f is a piecewise Hölder function:

$$|f^{(r)}(t) - f^{(r)}(s)| \leq L|t - s|^\varrho,$$

for all $t, s \in [0, u_f^1)$, $t, s \in [u_f^1, u_f^2)$, ..., $t, s \in [u_f^p, T]$.

The numbers $r, \varrho, p, M_0, M_1, \dots, M_r$ and L will be called parameters of the classes $\mathcal{F}_{\text{reg}}^{r,\varrho}$ or $\mathcal{F}_{\text{sng},p}^{r,\varrho}$. In general M_0, M_1, \dots, M_r and L are unknown and described algorithms will not use them as input parameters. Obviously, we have that $\mathcal{F}_{\text{reg}}^{r,\varrho} \subset \mathcal{F}_{\text{sng},p}^{r,\varrho}$ for all $p \geq 1$.

In order to describe information used, we recall definition of the s -fold integrated Brownian motion, denoted by $B^{(s)}$, $s \in \mathbb{N}_+$. It is defined by induction,

$$B_t^{(0)} = B_t, \quad B_t^{(s)} = \int_0^t B_u^{(s-1)} du, \quad (6)$$

for all $t \geq 0$ and $s \in \mathbb{N}_+$.

We assume that we have available standard information about function f and about process $B^{(r)}$. This means that we can evaluate f and its derivatives $f^{(j)}$, for $j = 1, 2, \dots, r$, at points $a_l \in [0, T]$. We assume that if the argument of $f^{(j)}$ coincides with a singular point u_f^p then the procedure computing $f^{(j)}(u_f^p)$ returns both left- and right-hand side values $f^{(j)}(u_f^{p+})$ and $f^{(j)}(u_f^{p-})$. Moreover, we can evaluate process $B^{(r)}$ and its derivatives $\frac{d^i}{dt} B^{(r)} = B^{(r-i)}$, for $i = 1, 2, \dots, r$, at points $t_l \in [0, T]$. Hence, a vector of information used, denoted by $N_n(f, B)$, consists of values

$$f^{(j_1)}(a_1), \dots, f^{(j_k)}(a_k), B_{t_1}^{(r-i_1)}, \dots, B_{t_m}^{(r-i_m)}, \quad (7)$$

where $k(m)$ denote the total number of evaluations of $f(B^{(r)})$ and its derivatives, and where the total cardinality of information used $n = k + m$. In (7) we assume that if $\alpha \neq \beta$ then $(j_\alpha, a_\alpha) \neq (j_\beta, a_\beta)$ and $(i_\alpha, t_\alpha) \neq (i_\beta, t_\beta)$.

An algorithm \mathcal{A}_n using information N_n , that gives approximation of $I(f, B)$, is of the form

$$\mathcal{A}_n(f, B) = \psi(N_n(f, B)), \quad (8)$$

where $\psi: \mathbb{R}^n \rightarrow \mathbb{R}$ is a measurable function, see [18]. We consider three classes of algorithms. By Ψ_n we denote the class of all algorithms of above form, for which information points are given in nonadaptive way (i.e., independently of f and $B^{(r)}$). By Ψ_n^* we denote the class of adaptive algorithms that use points a_i which may depend on evaluations of f or its derivatives at the previous points, but not on the values of $B^{(r)}$ or its derivatives. That is, a_1 is fixed and we choose a_i , for $i \geq 2$, as a function of computed values of f or its derivatives with the restriction that the total cardinality is n (for the process $B^{(r)}$ we use points chosen in nonadaptive way with respect to $B^{(r)}$ which do not depend on a particular trajectory of the process $B^{(r)}$). Finally, by Ψ_n^{**} we denote the class of all algorithms for which information points a_i and t_i may be chosen in adaptive way with respect to f and $B^{(r)}$ as it was described above. We have that $\Psi_n \subset \Psi_n^* \subset \Psi_n^{**}$.

The error of an algorithm $\mathcal{A}_n \in \bar{\Psi}_n$ ($\bar{\Psi}_n = \Psi_n, \bar{\Psi}_n = \Psi_n^*$ or $\bar{\Psi}_n = \Psi_n^{**}$) for function f from a class \mathcal{F} ($\mathcal{F} = \mathcal{F}_{\text{reg}}^{r, \varrho}$ or $\mathcal{F} = \mathcal{F}_{\text{sing}, p}^{r, \varrho}$) is defined by

$$e(\mathcal{A}_n, f) = (\mathbb{E}(I(f, B) - \mathcal{A}_n(f, B))^2)^{1/2}, \quad (9)$$

and the error of an algorithm \mathcal{A}_n in a class \mathcal{F} ,

$$e(\mathcal{A}_n, \mathcal{F}) = \sup_{f \in \mathcal{F}} e(\mathcal{A}_n, f). \quad (10)$$

In respective classes of algorithms the n th minimal errors, in a class \mathcal{F} , are given by

$$e_n(\mathcal{F}) = \inf_{\mathcal{A}_n \in \bar{\Psi}_n} e(\mathcal{A}_n, \mathcal{F}), \quad e_n^*(\mathcal{F}) = \inf_{\mathcal{A}_n \in \bar{\Psi}_n^*} e(\mathcal{A}_n, \mathcal{F}), \quad e_n^{**}(\mathcal{F}) = \inf_{\mathcal{A}_n \in \bar{\Psi}_n^{**}} e(\mathcal{A}_n, \mathcal{F}). \quad (11)$$

Of course, we have that $e_n^{**}(\mathcal{F}) \leq e_n^*(\mathcal{F}) \leq e_n(\mathcal{F})$. Our aim is to find possibly sharp bounds on above n th minimal errors as $n \rightarrow \infty$, i.e., lower and upper bounds which match up to a constants. We also want to know algorithm, from respective class, for which the infimum in (11) is asymptotically attained.

Unless otherwise stated, all constants appearing in this paper (including those in the “ Θ ”, “ O ” and “ Ω ” notation) will only depend on the parameters of the certain class, and possibly on T . Statement “for sufficiently large n ” means “for $n \geq n_0$ ”, where n_0 also depends on the parameters of a class and T . The same symbol may be used for different constants.

3. Regular case—the class $\mathcal{F}_{\text{reg}}^{r, \varrho}$

We start with upper bound on the n th minimal error in the class $\mathcal{F}_{\text{reg}}^{r, \varrho}$.

3.1. Upper bound

It is known that the Itô integral (2) is the solution at time T of the stochastic differential equation, with zero drift coefficient and additive noise, given by

$$dX_t = f(t)dB_t, \quad (12)$$

with the initial condition $X_0 = 0$. Hence, to deliver upper bound on the n th minimal error we consider the general Itô–Taylor algorithm [1], in this paper denoted by \mathcal{A}_n^{IT} . Since, for problem (2) the Itô–Taylor expansion becomes ordinary Taylor series, we do not need to use the hierarchical sets, as it was in [1]. For our purpose the algorithm \mathcal{A}_n^{IT} is obtained by expanding integrand f in the Taylor series on the proper subintervals of $[0, T]$. Namely, take an integer $n \geq 1$ and let the sample points be defined as

$$a_i = (i-1)T/n, \quad i = 1, 2, \dots, n+1. \quad (13)$$

The algorithm \mathcal{A}_n^{IT} is defined by

$$\mathcal{A}_n^{IT}(f, B) = \sum_{i=1}^n A_{i,n}^{IT}(f, B), \quad (14)$$

where

$$A_{i,n}^{IT}(f, B) = \sum_{j=0}^r \frac{f^{(j)}(a_i)}{j!} \int_{a_i}^{a_{i+1}} (t - a_i)^j dB_t. \quad (15)$$

Parameters used by \mathcal{A}_n^{IT} are: the smoothness r and a positive integer n .

Note that the algorithm uses only available information about $B^{(r)}$ described in Section 2. One can see this from the following identity which can be proved by induction and integration by parts. For all $0 \leq a < b$, $r \in \mathbb{Z}_+$, $q = 1, 2, \dots, r$ and functions $g \in C^{(r)}([a, b])$, it holds

$$\int_a^b g(t) B_t dt = \sum_{j=0}^{q-1} (-1)^j \left(g^{(j)}(t) B_t^{(j+1)} \right) \Big|_a^b + (-1)^q \int_a^b g^{(q)}(t) B_t^{(q)} dt. \quad (16)$$

From (16) for the function $g(t) = (t - a_i)^{j-1}$ we have

$$\int_{a_i}^{a_{i+1}} (t - a_i)^j dB_t = (T/n)^j B_{a_{i+1}} - j \left(\sum_{s=0}^{j-1} (-1)^s g^{(s)}(a_{i+1}) B_{a_{i+1}}^{(s+1)} + (-1)^j (j-1)! B_{a_i}^{(j)} \right), \quad (17)$$

for $j = 1, 2, \dots, r$, and $\int_{a_i}^{a_{i+1}} (t - a_i)^0 dB_t = B_{a_{i+1}} - B_{a_i}$ for $j = 0$. Therefore, we can express every integral $\int_{a_i}^{a_{i+1}} (t - a_i)^j dB_t$ in terms of values of $B^{(s)}$ ($s = 0, 1, \dots, j$) at the sample points. By (14) and (17) the total cardinality of the information used by the algorithm is bounded from above by $2(r+1)n$. Hence, $\mathcal{A}_n^{IT} \in \Psi_{2(r+1)n}$.

We now ask about the upper bounds on the error of the algorithm \mathcal{A}_n^{IT} , in considered Hölder class of integrands f . In [1] the rate of strong convergence of order n^{-r} was established for the general Itô–Taylor scheme. However, bound delivered in [1] contains unspecified constants and cannot be used in our (worst case) setting. In the following proposition we show an upper bound $O(n^{-(r+\varrho)})$ on the error of the algorithm \mathcal{A}_n^{IT} in a slightly different class of functions than considered in [1] and using simpler proof technique, adapted to our problem. We show explicitly how the error depends on the cardinality n of information used and on the regularity parameters of the Hölder class $\mathcal{F}_{\text{reg}}^{r,\varrho}$.

Proposition 1. For the algorithm \mathcal{A}_n^{IT} we have¹

$$e(\mathcal{A}_n^{IT}, \mathcal{F}_{\text{reg}}^{r,\varrho}) \leq \frac{LT^{r+\varrho+1/2}}{(r-1)!(2(r+\varrho)+1)^{1/2}} \beta(\varrho+1, r) n^{-(r+\varrho)}, \quad (18)$$

where $\beta(x, y)$ is the beta-function.

Proof. We define the function $\hat{f} : [0, T] \rightarrow \mathbb{R}$ by

$$\hat{f}(t) = \sum_{i=1}^n \mathbf{1}_{[a_i, a_{i+1})}(t) \sum_{j=0}^r \frac{f^{(j)}(a_i)}{j!} (t - a_i)^j. \quad (19)$$

Then we can write the algorithm \mathcal{A}_n^{IT} in the integral form

$$\mathcal{A}_n^{IT}(f, B) = \int_0^T \hat{f}(t) dB_t.$$

If $r \geq 1$ then, by Taylor's theorem, we have for all $t \in [a_i, a_{i+1})$ that

$$f(t) - \hat{f}(t) = \int_0^1 (f^{(r)}(\theta t + (1-\theta)a_i) - f^{(r)}(a_i)) (t - a_i)^r \frac{(1-\theta)^{r-1}}{(r-1)!} d\theta. \quad (20)$$

Since $f^{(r)}$ satisfies Hölder condition, we have

$$|f(t) - \hat{f}(t)| \leq \frac{L\beta(\varrho+1, r)}{(r-1)!} (t - a_i)^{r+\varrho}, \quad (21)$$

¹ We use the convention that $\beta(x, 0) := 1$ and $(-1)! := 1$.

for all $r \geq 0$ and $t \in [a_i, a_{i+1})$. From the Itô isometry for Itô integrals we have that

$$(e(\mathcal{A}_n^T, f))^2 = \int_0^T (f(t) - \hat{f}(t))^2 dt = \sum_{i=1}^n e_i^2,$$

where

$$e_i^2 = \int_{a_i}^{a_{i+1}} (f(t) - \hat{f}(t))^2 dt.$$

From (21) we thus have that

$$e_i^2 \leq \frac{L^2 \beta^2(\varrho + 1, r)}{((r-1)!)^2 (2(r+\varrho) + 1)} (a_{i+1} - a_i)^{2(r+\varrho)+1}, \quad (22)$$

which finally gives

$$(e(\mathcal{A}_n^T, f))^2 \leq \frac{L^2 T^{2(r+\varrho)+1} \beta^2(\varrho + 1, r)}{((r-1)!)^2 (2(r+\varrho) + 1)} n^{-2(r+\varrho)},$$

for all $f \in \mathcal{F}_{\text{reg}}^{r, \varrho}$. This proves the thesis. \square

Remark 1. From the proof of Proposition 1, we see that for every discretization $0 = \tau_1 < \tau_2 < \dots < \tau_{n+1} = T$ (not necessary equidistant) we have

$$e(\mathcal{A}_n^T, \mathcal{F}_{\text{reg}}^{r, \varrho}) \leq C \left(\sum_{i=1}^n (\tau_{i+1} - \tau_i)^{2(r+\varrho)+1} \right)^{1/2},$$

where positive constant C depends only on the parameters of the class $\mathcal{F}_{\text{reg}}^{r, \varrho}$ and T .

In order to establish optimality of the Itô–Taylor algorithm, we need to investigate lower bounds on the n th minimal error in the considered class of integrands f .

3.2. Lower bounds

We prove a theorem that gives a lower bound on the error of any algorithm $\mathcal{A}_n \in \Psi_n^{**}$. This bound holds for any class \mathcal{F} of integrands containing functions of a special form. As we will see, such functions belong to the class $\mathcal{F}_{\text{reg}}^{r, \varrho}$.

Theorem 1. If a class \mathcal{F} contains function $h : [0, T] \rightarrow \mathbb{R}$ of the form $h(t) = \gamma_1 + \gamma_2 t^{r+1}$ for some $\gamma_1, \gamma_2 \in \mathbb{R}$ and $r \in \mathbb{Z}_+ \cup \{0\}$ with $\gamma_2 \neq 0$, then there exists positive constant C , depending only on r , such that for sufficiently large n and for any algorithm $\mathcal{A}_n \in \Psi_n^{**}$ we have

$$e(\mathcal{A}_n, \mathcal{F}) \geq C \gamma_2 T^{r+3/2} n^{-(r+1)}. \quad (23)$$

Proof. From the properties of the Itô integral and applying (16) to the function $g(t) = t^r$ we have

$$I(h, B) = (\gamma_1 + \gamma_2 T^{r+1}) B_T - \gamma_2 (r+1) \sum_{j=0}^{r-1} (-1)^j g^{(j)}(T) B_T^{(j+1)} + \gamma_2 (r+1)! (-1)^{r+1} \int_0^T B_t^{(r)} dt. \quad (24)$$

By changing variables we have

$$\int_0^T B_t^{(r)} dt = T^{r+3/2} \int_0^1 \bar{B}_t^{(r)} dt, \quad (25)$$

where $\bar{B}_t^{(r)} = T^{-(r+1/2)} B_{Tt}^{(r)}$ is also the r -fold integrated Brownian motion. Hence, for $\gamma_2 \neq 0$ approximating $I(h, B)$ is equivalent to the problem of approximating scalar Lebesgue integrals $\int_0^1 \bar{B}_t^{(r)} dt$ from the r -fold integrated Brownian motion $\bar{B}_t^{(r)}$. Such linear problems under Gaussian measures were studied in [10–13] and [17]. From the results of that papers we have that the n th minimal average error is no less than

$$C \gamma_2 T^{r+3/2} n^{-(r+1)}, \quad (26)$$

as $n \rightarrow \infty$, even for adaptive choice of evaluation points. In (26) positive constant C depends only on r . This proves (23). \square

We want to stress here that the lower bound from Theorem 1 holds even if we have complete information about the integrand.

Theorem 1 yields the following corollary.

Corollary 1. *There exists positive constant C , depending only on r , such that for sufficiently large n and for any algorithm $\mathcal{A}_n \in \Psi_n^{**}$ we have*

$$e(\mathcal{A}_n, \mathcal{F}_{\text{reg}}^{r,1}) \geq CLT^{r+3/2} n^{-(r+1)}. \quad (27)$$

Proof. For $\gamma_2 = L/(r+1)!$ and any γ_1 function $h(t) = \gamma_1 + \gamma_2 t^{r+1}$ belongs to the class $\mathcal{F}_{\text{reg}}^{r,1}$, therefore from **Theorem 1** we have the thesis. \square

Corollary 1 gives lower bounds for the largest class of algorithms Ψ_n^{**} , considered in this paper, in the Lipschitz class of integrands $\mathcal{F}_{\text{reg}}^{r,1}$. For the Hölder class $\mathcal{F}_{\text{reg}}^{r,\varrho}$, with any $\varrho \in (0, 1]$, we will give lower bound on the error for algorithms from a smaller class Ψ_n^* . Lower bound in the class $\mathcal{F}_{\text{reg}}^{r,\varrho}$, with $\varrho \in (0, 1)$, for algorithms that belong to Ψ_n^{**} is an open problem.

The following **Theorem 2** shows lower bound in the class Ψ_n^* . That is, we assume discrete adaptive information about f . Moreover, the lower bound holds even if we have complete information about trajectories of the process $B^{(r)}$.

Theorem 2. *There exist positive constants C and n_0 , depending only on the parameters of the class $\mathcal{F}_{\text{reg}}^{r,\varrho}$ and T , such that for all $n \geq n_0$ and any algorithm $\mathcal{A}_n \in \Psi_n^*$ we have*

$$e(\mathcal{A}_n, \mathcal{F}_{\text{reg}}^{r,\varrho}) \geq Cn^{-(r+\varrho)}. \quad (28)$$

Proof. Let a_1, a_2, \dots be the information points computed for the function $f \equiv 0$, where the total number of evaluations is n . We now construct the function $\tilde{f} \in \mathcal{F}_{\text{reg}}^{r,\varrho}$ which share with f the same information and for which the distance $\mathbb{E}|I(f, B) - I(\tilde{f}, B)|^2$ is as large as possible.

Let $\tilde{a}_0 = 0, \tilde{a}_{l+1} = T$ and let $\{\tilde{a}_i\}_{i=1}^l$ be those distinct (ordered in increasing order) information points, used by algorithm \mathcal{A}_n for f , which belong to the interval $(0, T)$. Of course, we have $0 \leq l \leq n$. We define the function \tilde{f} as follows.

For $i = 0, 1, \dots, l$ let $h_i \in \mathcal{F}_{\text{reg}}^{r,\varrho}$ be functions with the following properties:

- h_i has support $[\tilde{a}_i, \tilde{a}_{i+1}]$ and $h_i(t) \geq 0$,
- $\int_{\tilde{a}_i}^{\tilde{a}_{i+1}} h_i(t) dt = C(\tilde{a}_{i+1} - \tilde{a}_i)^{r+\varrho+1}$,

where C is known positive constant which depends only on the parameters of the class $\mathcal{F}_{\text{reg}}^{r,\varrho}$ and T (but not on i and n). Construction of such functions is well known, since these functions are often used in proving lower bounds, see [14,15].

We set $\tilde{f}(t) = \gamma \sum_{i=0}^l h_i(t)$ with $\gamma > 0$. Then $\tilde{f} \in \mathcal{F}_{\text{reg}}^{r,\varrho}$ for suitably chosen γ that depends only on the parameters of the class $\mathcal{F}_{\text{reg}}^{r,\varrho}$ and T . From properties of h_i we see that f and \tilde{f} have the same information and therefore $\mathcal{A}_n(f_1, B) = \mathcal{A}_n(f_2, B)$. Hence

$$\begin{aligned} e(\mathcal{A}_n, \mathcal{F}_{\text{reg}}^{r,\varrho}) &\geq \frac{1}{2} \left(\mathbb{E} \left(I(f, B) - I(\tilde{f}, B) \right)^2 \right)^{1/2} = \frac{\gamma}{2} \left(\int_0^T \left(\sum_{i=0}^l h_i(t) \right)^2 dt \right)^{1/2} \\ &= \frac{\gamma}{2} \left(\sum_{i=0}^l \int_{\tilde{a}_i}^{\tilde{a}_{i+1}} (h_i(t))^2 dt \right)^{1/2} \geq \frac{\gamma}{2} \left(\sum_{i=0}^l (\tilde{a}_{i+1} - \tilde{a}_i)^{-1} \left(\int_{\tilde{a}_i}^{\tilde{a}_{i+1}} h_i(t) dt \right)^2 \right)^{1/2} \\ &\geq \frac{\gamma CT^{r+\varrho+1/2}}{2} (l+1)^{-(r+\varrho)}. \end{aligned}$$

Since $0 \leq l \leq n$, this proves the assertion of the theorem. \square

From the **Proposition 1** and **Theorem 2** we conclude the following result.

Corollary 2. *The n th minimal error in the class $\mathcal{F}_{\text{reg}}^{r,\varrho}$ satisfies*

$$e_n(\mathcal{F}_{\text{reg}}^{r,\varrho}) = \Theta(n^{-(r+\varrho)}). \quad (29)$$

Hence, optimal algorithm is the Itô–Taylor scheme $\mathcal{A}_n^{\text{IT}}$. Lower bound in (29) also holds for adaptive algorithms from the class Ψ_n^* .

Corollary 2 states that in the regular case the nonadaptive Itô–Taylor algorithm is optimal even among all adaptive algorithms. This means that we do not need to use adaptive algorithms when integrands f belong to the class of regular functions $\mathcal{F}_{\text{reg}}^{r,\varrho}$.

4. Singular case—the class $\mathcal{F}_{\text{sng},1}^{r,\varrho}$

In this section we investigate problem (2) in the class $\mathcal{F}_{\text{sng},1}^{r,\varrho}$. Since $\mathcal{F}_{\text{reg}}^{r,\varrho} \subset \mathcal{F}_{\text{sng},1}^{r,\varrho}$, we have

$$e_n(\mathcal{F}_{\text{sng},1}^{r,\varrho}) = \Omega(n^{-(r+\varrho)}). \quad (30)$$

We want to know if there exists an algorithm which preserves the error $O(n^{-(r+\varrho)})$ also in the singular class $\mathcal{F}_{\text{sng},1}^{r,\varrho}$. Such algorithm cannot be nonadaptive since we have the following proposition.

Proposition 2. For $r \geq 0, \varrho \in (0, 1]$ and any algorithm $\mathcal{A}_n \in \Psi_n$ we have

$$e(\mathcal{A}_n, \mathcal{F}_{\text{sng},1}^{r,\varrho}) = \Omega(n^{-1/2}). \quad (31)$$

For $r = 0$ we have that

$$e(\mathcal{A}_n, \mathcal{F}_{\text{sng},1}^{0,\varrho}) = \Omega(n^{-\min\{1/2, \varrho\}}). \quad (32)$$

Proof. Let $0 \leq a_1 \leq a_2 \leq \dots \leq a_n \leq T$ be the nonadaptive information points used by an algorithm \mathcal{A}_n for a function f . We define $k+1$ ($k \leq n+1$) points in the following way. Set $\bar{a}_0 := 0, \bar{a}_k := T$ and let $\{\bar{a}_i\}_{i=1}^{k-1}$ be those information points $\{a_i\}_{i=1}^n$ that belong to $(0, T)$, and for which $\bar{a}_i < \bar{a}_2 < \dots < \bar{a}_{k-1}$. Denote by $[\bar{a}_i, \bar{a}_{i+1}]$ the interval of maximal length, for which $\bar{a}_{i+1} - \bar{a}_i \geq T/k$ and let $[\tilde{a}_i, \tilde{a}_{i+1}]$ be a subinterval of $(\bar{a}_i, \bar{a}_{i+1})$, such that $\tilde{a}_{i+1} - \tilde{a}_i \geq T/(2k)$. Let us now consider two functions f_1 and f_2 from the class $\mathcal{F}_{\text{sng},1}^{r,\varrho}$, where $f_1 = M_0 \mathbf{1}_{[\bar{a}_i, T]}$ and $f_2 = M_0 \mathbf{1}_{[\tilde{a}_{i+1}, T]}$. Functions f_1 and f_2 differ only in $(\bar{a}_i, \tilde{a}_{i+1})$, and have singularities at \tilde{a}_i and \tilde{a}_{i+1} , respectively. Hence, they share the same information and the output of \mathcal{A}_n for f_1 and f_2 will be the same. Nevertheless $I(f_1, B) - I(f_2, B) = M_0(B_{\tilde{a}_{i+1}} - B_{\tilde{a}_i})$, and therefore

$$e(\mathcal{A}_n, \mathcal{F}_{\text{sng},1}^{r,\varrho}) \geq \frac{1}{2} (\mathbb{E} (I(f_1, B) - I(f_2, B))^2)^{1/2} \geq \left(\frac{M_0^2 T}{8k} \right)^{1/2}.$$

Since $k \leq n+1$, the proof of (31) is completed. To see (32) we use (31) and Theorem 2 with $r = 0$. \square

Proposition above states that if we use only nonadaptive information about f and $B^{(r)}$ to approximate (2) then the error is $\Omega(n^{-1/2})$, even for piecewise very regular functions with a single singularity. This means that we must turn to adaptive algorithms.

To define an adaptive algorithm, we first study the behavior of the algorithm \mathcal{A}_n^{IT} , when unknown singularities exist in the interval $(0, T)$. Next, in the case when $p = 1$, we use slightly modified bisection algorithm from [7], locating the singularity, and we define the adaptive algorithm (denoted by \mathcal{A}_n^{IT*}) that will have the desired error $O(n^{-(r+\varrho)})$.

4.1. Error of the Itô–Taylor algorithm in the presence of a singularity

In the singular case for the algorithm \mathcal{A}_n^{IT} , if the derivative $f^{(j)}(a_i)$ does not exist, by $f^{(j)}(a_i)$ we mean $f^{(j)}(a_i^+)$. Moreover, for simplicity, when $p = 1$ we will write u_f and Δ_f^j instead of u_f^1 and $\Delta_f^{j,1}$, respectively.

The following lemma shows an upper bound on the error of \mathcal{A}_n^{IT} in the presence of at most one singularity.

Lemma 1. There exist positive constants C and n_0 , depending only on the parameters of the class $\mathcal{F}_{\text{sng},1}^{r,\varrho}$ and T , such that for all $f \in \mathcal{F}_{\text{sng},1}^{r,\varrho}$ and $n \geq n_0$ we have

$$e(\mathcal{A}_n^{IT}, f) \leq C \left(n^{-(r+\varrho)} + n^{-1/2} \sum_{j=0}^r |\Delta_f^j| n^{-j} \right). \quad (33)$$

Before proving Lemma 1, we show the following result.

Lemma 2. There exist constants C, \bar{C} and n_0 , where C and n_0 depend only on r, ϱ, p, L and T , and \bar{C} depends only on $r, p, M_0, M_1, \dots, M_r$, such that for all $n \geq n_0$ we have

$$e(\mathcal{A}_n^{IT}, \mathcal{F}_{\text{sng},p}^{r,\varrho}) \leq C(n^{-(r+\varrho)} + \bar{C}n^{-1/2}). \quad (34)$$

Proof. Let $f \in \mathcal{F}_{\text{sng},p}^{r,\varrho}$ and let the function \hat{f} be defined by (19). Take $n \geq p$ and let discretization points a_i be defined by (13). Let us denote by $m(f, i)$ the number of singular points of f that are in interval (a_i, a_{i+1}) , $i = 1, 2, \dots, n$. By $\mathcal{N}(f)$ we mean the following set of indices:

$$\mathcal{N}(f) = \{i : 1 \leq i \leq n, m(f, i) > 0\}. \quad (35)$$

The cardinality of the set $\mathcal{N}(f)$, denoted by $\sharp \mathcal{N}(f)$, is the total number of intervals (a_i, a_{i+1}) that contain at least one singular point of f . If we have $\mathcal{N}(f) = \emptyset$, then either there is no singularity for f or every singularity coincides with some discretization point. In this case estimation of the error $\mathbb{E} \left(\int_0^T (f(t) - \hat{f}(t)) dB_t \right)^2$ goes exactly as in the regular case. Therefore, we can assume that $\mathcal{N}(f) \neq \emptyset$. Of course we have $\sharp \mathcal{N}(f) \leq p$ and $\sum_{i \in \mathcal{N}(f)} m(f, i) \leq p$. For $i \in \mathcal{N}(f)$ we denote by $u_f^{(\alpha, i)}$, $\alpha \in \{1, 2, \dots, m(f, i)\}$, those singular points $u_f^1, u_f^2, \dots, u_f^p$ which belong to (a_i, a_{i+1}) , ordered in increasing order, i.e., $u_f^{(1, i)} < u_f^{(2, i)} < \dots < u_f^{(m(f, i), i)}$. From Lemma 4 (in the Appendix), with $g := f$ and $p := m(f, i)$, we have for $t \in [a_i, a_{i+1})$ that

$$|f(t) - \hat{f}(t)| \leq C \left(n^{-(r+\varrho)} + \tilde{\delta}_{f,i} \right), \quad (36)$$

where

$$\begin{aligned} \tilde{\delta}_{f,i} = & \sum_{j_m(f,i)=0}^r |\tilde{\Delta}_f^{j_m(f,i), m(f,i)}| n^{-j_m(f,i)} + \sum_{j_m(f,i)=0}^r \sum_{j_m(f,i)-1=j_m(f,i)}^r |\tilde{\Delta}_f^{j_m(f,i)-1, m(f,i)-1}| n^{-j_m(f,i)-1} \\ & + \sum_{j_m(f,i)=0}^r \sum_{j_m(f,i)-1=j_m(f,i)}^r \sum_{j_m(f,i)-2=j_m(f,i)-1}^r |\tilde{\Delta}_f^{j_m(f,i)-2, m(f,i)-2}| n^{-j_m(f,i)-2} \\ & + \dots + \sum_{j_m(f,i)=0}^r \sum_{j_m(f,i)-1=j_m(f,i)}^r \sum_{j_m(f,i)-2=j_m(f,i)-1}^r \dots \sum_{j_2=j_3}^r \sum_{j_1=j_2}^r |\tilde{\Delta}_f^{j_1, 1}| n^{-j_1}, \end{aligned}$$

and where

$$\tilde{\Delta}_f^{j_m(f,i)-\beta, m(f,i)-\beta} := f^{(j_m(f,i)-\beta)} \left(u_f^{(m(f,i)-\beta, i)+} \right) - f^{(j_m(f,i)-\beta)} \left(u_f^{(m(f,i)-\beta, i)-} \right), \quad (37)$$

for $\beta \in \{0, 1, \dots, m(f, i) - 1\}$. Constant C in (36) depends only on r, L, ϱ, p and T . Since $f \in \mathcal{F}_{\text{sng},p}^{r,\varrho}$, there exists positive constant \tilde{C} , which depends only on $p, r, M_0, M_1, \dots, M_r$, such that we have

$$\tilde{\delta}_{f,i} \leq \tilde{C}, \quad (38)$$

for all $i \in \mathcal{N}(f)$. We have, as in proof of Theorem 1, that $\mathcal{A}_n^{\text{IT}}(f, B) = \int_0^T \hat{f}(t) dB_t$ and we can decompose

$$(e(\mathcal{A}_n^{\text{IT}}, f))^2 = \sum_{i \notin \mathcal{N}(f)} e_i^2 + \sum_{i \in \mathcal{N}(f)} e_i^2.$$

For $i \in \mathcal{N}(f)$ we have from (36)

$$e_i^2 = \int_{a_i}^{a_{i+1}} (f(t) - \hat{f}(t))^2 dt \leq C_1 (n^{-(2(r+\varrho)+1)} + n^{-1} (\tilde{\delta}_{f,i})^2),$$

where C_1 depends only on r, L, ϱ, p and T . If $i \notin \mathcal{N}(f)$ then we have (as in regular case) $e_i^2 \leq C_2 n^{-(2(r+\varrho)+1)}$ with constant C_2 that depends only on r, L, ϱ and T , see (22). Therefore

$$(e(\mathcal{A}_n^{\text{IT}}, f))^2 \leq C \left(n^{-2(r+\varrho)} + n^{-1} \sum_{i \in \mathcal{N}(f)} (\tilde{\delta}_{f,i})^2 \right) \quad (39)$$

$$\leq C (n^{-2(r+\varrho)} + \tilde{C} n^{-1}), \quad (40)$$

for $n \geq p$. In (40) constant $C > 0$ depends only on p, T, ϱ, L and r , while $\tilde{C} > 0$ depends on p, M_0, M_1, \dots, M_r and r . This ends the proof of (34). \square

The statement of Lemma 1 follows from (39) with $p = 1$.

Remark 2. Lemma 1 shows how the upper bound on the error of $\mathcal{A}_n^{\text{IT}}$ depends on the jumps Δ_f^j . For example, for discontinuous functions f ($\Delta_f^0 \neq 0$), the error of the $\mathcal{A}_n^{\text{IT}}$ is of order $|\Delta_f^0| n^{-\min\{1/2, r+\varrho\}}$. If f is continuous and has s continuous derivatives with $0 \leq s \leq r-1$ ($\Delta_f^\alpha = 0$ for $\alpha = 0, 1, \dots, s$) then the error is $O(n^{-(s+1+\min\{1/2, \varrho\})})$. For $s = r$ the error is $O(n^{-(r+\varrho)})$.

By [Lemma 1](#) the error of the algorithm $\mathcal{A}_n^{\text{IT}}$ in the class $\mathcal{F}_{\text{sng},1}^{r,\varrho}$ satisfies

$$e(\mathcal{A}_n^{\text{IT}}, \mathcal{F}_{\text{sng},1}^{r,\varrho}) = O(n^{-\min\{1/2, \varrho+r\}}). \quad (41)$$

Moreover, we can show explicitly the function $f \in \mathcal{F}_{\text{sng},1}^{r,1}$ for which the error (10) of the Itô–Taylor algorithm $\mathcal{A}_n^{\text{IT}}$ is $\Omega(n^{-1/2})$. Namely, for every $n \geq 2$ we construct the function f in the following way. Let $\{a_i\}_{i=1}^{n+1}$ be defined in (13). Set $i_0 = \lceil n/2 \rceil$ and $u_f = (a_{i_0} + a_{i_0+1})/2$. The function f is given by

$$f(t) = \begin{cases} 1, & t \in [0, u_f], \\ \gamma t^{r+1} + 1, & t \in [u_f, T], \end{cases} \quad (42)$$

where $\gamma > 0$ is suitably chosen constant that depends only on the parameters of the class $\mathcal{F}_{\text{sng},1}^{r,1}$ and T . Using the following identity

$$f(t) - \hat{f}(t) = \sum_{j=0}^r \frac{1}{j!} \Delta_f^j(t - u_f)^j + \sum_{j=0}^r (f^{(j)}(u_f^-) - \hat{f}^{(j)}(u_f))(t - u_f)^j + \gamma(t - u_f)^{r+1},$$

which follows from (70) in the proof of the [Lemma 4](#) (see the [Appendix](#)), it can be easily shown that for the function (42) we have

$$(e(\mathcal{A}_n^{\text{IT}}, f))^2 \geq \gamma^2 \left(\frac{T}{4}\right)^{2r+2} \frac{T}{2n}. \quad (43)$$

Identity (43) together with (41) finally gives

$$e(\mathcal{A}_n^{\text{IT}}, \mathcal{F}_{\text{sng},1}^{r,1}) = \Theta(n^{-1/2}). \quad (44)$$

Remark 3. It follows from (41) and [Theorem 2](#) that for $r = 0$ and $\varrho \in (0, 1/2]$ the bound (41) cannot be improved even if we use adaptive algorithms, since in this case we have that

$$e_n^*(\mathcal{F}_{\text{sng},1}^{0,\varrho}) = \Theta(n^{-\varrho}). \quad (45)$$

In the remaining cases, when $(r, \varrho) \in (\{0\} \times (1/2, 1]) \cup (\{1, 2, \dots\} \times (0, 1])$, we have from [Proposition 2](#) that the bound $O(n^{-\min\{1/2, \varrho+r\}})$ cannot be improved in the class of nonadaptive algorithms. Therefore, in the case when $r + \varrho > 1/2$ we must turn to adaptive algorithms.

In order to define an adaptive algorithm we introduce analogous classes of functions f as in [7], which we will use to distinguish between *mild* and *difficult* singularities for a given n . Let K be a positive constant which may only depends on the parameters of the class $\mathcal{F}_{\text{sng},1}^{r,\varrho}$ and T . For any $n \geq 1$, we define a subclass $M_K(n)$ of $\mathcal{F}_{\text{sng},1}^{r,\varrho}$ by

$$M_K(n) = \left\{ f \in \mathcal{F}_{\text{sng},1}^{r,\varrho} \mid \sum_{j=0}^r |\Delta_f^j| n^{-j} \leq K n^{-(r+\varrho)} \right\}.$$

We have that $M_K(\bar{n}) \subset M_K(n)$ for $n \leq \bar{n}$.

By [Lemma 1](#) for $f \in M_K(n)$ we obtain

$$e(\mathcal{A}_n^{\text{IT}}, f) \leq C(1 + K n^{-1/2}) n^{-(r+\varrho)}.$$

Therefore, up to a constant K , the error bound from the regular case is preserved. Singularities u_f for integrands $f \in M_K(n)$ will be called *mild* for a given n with constant K . Otherwise they will be referred to as *difficult* for n .

4.2. The adaptive Itô–Taylor algorithm

In order to define the adaptive Itô–Taylor algorithm we use a slightly modified bisection algorithm from [7], that was used to locate the unknown singularity for a singular initial-value problem. In [7] the detection mechanism was defined for a Lipschitz class but it is easy to see that, after some modifications, it also works for the Hölder class $\mathcal{F}_{\text{sng},1}^{r,\varrho}$.

For real numbers $c, d \in (0, T)$ ($c < d$) and function $f \in \mathcal{F}_{\text{sng},1}^{r,\varrho}$, we define as in [7] the following quantity

$$A_f(c, d) = \max_{0 \leq j \leq r} \frac{|f^{(j)}(d) - w_c^{(j)}(d)|}{(d - c)^{r-j+\varrho}}, \quad (46)$$

where $w_c(t) = \sum_{j=0}^r (1/j!) f^{(j)}(c)(t-c)^j$, $t \in [c, d]$. Then the following implication holds:

$$\text{if } u_f \notin [c, d] \text{ then } A_f(c, d) \leq L, \quad (47)$$

which is a direct consequence of Taylor's theorem.

We now recall the following lemma from [7] which gives upper bounds, in terms of quantity A_f , on unknown jumps Δ_f^j for singularities from an interval (c, d) .

Lemma 3 ([7]). *There exist positive constants K_j , $j = 0, 1, \dots, r$, dependent only on j, r and L such that for all $c, d \in [0, T]$, $c < d$, the following implication holds:*

$$\text{if } u_f \in (c, d) \text{ then } |\Delta_f^j| \leq K_j(A_f(c, d) + 1)(d - c)^{r-j+q}, \quad j = 0, 1, \dots, r. \quad (48)$$

We will be locating difficult singularity for $f \in \mathcal{F}_{\text{sing},1}^{r,q}$ using algorithm from [7], suitably modified for our purpose. Let us now recall this algorithm with minor modifications necessary in our context. Properties of the recalled algorithm are discussed in [7].

Take arbitrary $D > 0$ and proceed as follows.

L1 Compute $f^{(j)}(a_i)$ for $i = 1, 2, \dots, n+1$, $j = 0, 1, \dots, r$, where points $\{a_i\}$ are given by (13). If $u_f = a_i$ for some i , then go to STOP.

L2 Compute \bar{A}_f given by

$$\bar{A}_f = \max_{1 \leq i \leq n} A_f(a_i, a_{i+1}). \quad (49)$$

If, for some distinct i and j , $\bar{A}_f = A_f(a_i, a_{i+1}) = A_f(a_j, a_{j+1})$ then go to STOP and use the algorithm $\mathcal{A}_n^{\text{IT}}(f, B)$, from the regular case, to approximate $I(f, B)$.

L3 Check if $\bar{A}_f \leq D$. If YES, go to STOP and use the algorithm $\mathcal{A}_n^{\text{IT}}(f, B)$ to approximate $I(f, B)$. If NO, choose interval $(c, d) = (a_i, a_{i+1})$ for which the maximum in (49) is achieved. Set $s = 0$ and $s_{\max} = \lceil (2(r+q) - 1) \log_2(n/T) \rceil$.

L4 Set $s := s + 1$. Compute $v = (c + d)/2$ and $f^{(j)}(v)$ for $j = 0, 1, \dots, r$. If $v = u_f$ then go to STOP.

L5 Compute the values $A_f(c, v)$ and $A_f(v, d)$, and check if $A_f(c, v) = A_f(v, d)$. If YES then go to STOP and use the algorithm $\mathcal{A}_n^{\text{IT}}(f, B)$ to approximate $I(f, B)$. If NO then choose the next (c, d) in the following way:

$$(c, d) = \begin{cases} (c, v), & \text{if } A_f(c, v) > A_f(v, d), \\ (v, d), & \text{if } A_f(c, v) < A_f(v, d). \end{cases} \quad (50)$$

If $s < s_{\max}$, then go to L4.

STOP.

In the algorithm above the quantity s_{\max} is strictly positive since, according to Remark 3, in the adaptive case we assume that $r + q > 1/2$.

Information cost of using the algorithm is bounded from above by $(n+1)(r+1) + \lceil (2(r+q) - 1) \log_2 n + \log_2(1/T)^{2(r+q)-1} \rceil (r+1)$ evaluations of f or its derivatives. The first term comes from step L1, the second one is consequence of a logarithmic number of repetitions of steps L4–L5.

Remark 4. Note that the Lipschitz constant L , which is in general unknown, is not used by the algorithm. If we knew L , then the best choice for D would be $D = L$.

While performing the bisection algorithm, the following cases may occur:

Case 1. For some distinct i and j , $\bar{A}_f = A_f(a_i, a_{i+1}) = A_f(a_j, a_{j+1})$. Then $\bar{A}_f \leq L$ and hence, we have

$$e(\mathcal{A}_n^{\text{IT}}, f) \leq Cn^{-(r+q)} \left(1 + n^{-1/2}(L+1) \max_{0 \leq j \leq r} K_j T^q \frac{T^{r+1} - 1}{T - 1} \right), \quad (51)$$

where $C > 0$ depends only on r, q, L and T .

Case 2. $\bar{A}_f \leq D$. Then there exists an interval $(\tilde{c}, \tilde{d}) \subset (0, T)$, $\tilde{d} - \tilde{c} = T/n$, such that $u_f \in (\tilde{c}, \tilde{d})$ and $A_f(\tilde{c}, \tilde{d}) \leq \max\{D, L\}$. Therefore from Lemmas 1 and 3 we obtain for $\mathcal{A}_n^{\text{IT}}(f, B)$ that

$$e(\mathcal{A}_n^{\text{IT}}, f) \leq Cn^{-(r+q)} \left(1 + n^{-1/2}(\max\{D, L\} + 1) \max_{0 \leq j \leq r} K_j T^q \frac{T^{r+1} - 1}{T - 1} \right), \quad (52)$$

where positive constant C depends only on r, q, L and T .

Case 3. $\bar{A}_f = A_f(a_i, a_{i+1}) > D$ and at some stage in L4–L5 we have $A_f(c, v) = A_f(v, d)$. Then $A_f(c, v) \leq L$ and (51) holds.

As we can see in all above cases, in order to preserve optimal error from the regular case, it suffices to use the nonadaptive Itô–Taylor algorithm \mathcal{A}_n^{IT} . If none of the these cases happened then, after at most s_{\max} steps, we obtain the interval (c, d) or the singular point u_f itself. The length of this interval is $d - c \leq (T/n)^{2(r+\varrho)}$. (If the algorithm gives us the singular point u_f itself then we can choose interval (c, d) to be arbitrarily small.) The main property of the algorithm is that we have

$$\text{if } u_f \notin (c, d) \text{ then } \exists (\bar{c}, \bar{d}) \subset (0, T) : u_f \in (\bar{c}, \bar{d}), \bar{d} - \bar{c} \leq T/n \text{ and } A_f(\bar{c}, \bar{d}) \leq L. \quad (53)$$

Hence, if $u_f \notin (c, d)$ then from (53) and Lemma 3 we obtain

$$\sum_{j=0}^r |\Delta_f^j| n^{-j} \leq (L+1) \max_{0 \leq j \leq r} K_j T^\varrho \frac{T^{r+1} - 1}{T - 1} n^{-(r+\varrho)}, \quad (54)$$

which means that such singularity is mild.

Having the information about the interval (c, d) , that contains difficult singularity for f , we can modify the starting discretization $\{a_i\}$ of the interval $[0, T]$ in the following way.

Let l be the index for which the maximum in (49) is achieved. Using the interval $(c, d) \subset [a_l, a_{l+1}]$, given by the bisection algorithm, we define the new mesh points $\{\bar{a}_i\}$ as follows: $\bar{a}_i = a_i$, for $i = 1, 2, \dots, l$, $\bar{a}_{l+1} = c$, $\bar{a}_{l+2} = d$ and $\bar{a}_{i+2} = a_i$ for $i = l+1, l+2, \dots, n+1$. Therefore, the new discretization $\{\bar{a}_i\}$ contains at most $n+3$ points. Note that $\bar{a}_i \neq u_f$ for $i = 1, 2, \dots, n+3$.

The adaptive Itô–Taylor scheme \mathcal{A}_n^{IT*} is defined in the following way. For $f \in \mathcal{F}_{\text{sg},1}^{r,\varrho}$ let

$$\mathcal{A}_n^{IT*}(f, B) = \sum_{i=1}^{n+2} A_{i,n}^{IT*}(f, B), \quad (55)$$

where

$$A_{i,n}^{IT*}(f, B) = \sum_{j=0}^r \frac{f^{(j)}(\bar{a}_i)}{j!} \int_{\bar{a}_i}^{\bar{a}_{i+1}} (t - \bar{a}_i)^j dB_t. \quad (56)$$

Parameters of the algorithm \mathcal{A}_n^{IT*} are: the smoothness parameters r and ϱ , a positive integer n and a threshold D .

For sufficiently large n the informational cost of the algorithm \mathcal{A}_n^{IT*} , which first detects the difficult singularity for f and next performs the classical Itô–Taylor algorithm on the modified mesh, is bounded from above by cn evaluations of function f and process $B^{(r)}$ along with their derivatives. Positive constant c depends only on r, ϱ and T . Hence, we have that $\mathcal{A}_n^{IT*} \in \Psi_{cn}^*$.

We now show that the algorithm \mathcal{A}_n^{IT*} has the optimal order of convergence in the class $\mathcal{F}_{\text{sg},1}^{r,\varrho}$ of singular functions, which is the same as for the class of regular functions $\mathcal{F}_{\text{reg}}^{r,\varrho}$.

Theorem 3. *There exist positive constants C and n_0 , depending only on the parameters of the class $\mathcal{F}_{\text{sg},1}^{r,\varrho}$ and T , such that for $n \geq n_0$ we have*

$$e(\mathcal{A}_n^{IT*}, \mathcal{F}_{\text{sg},1}^{r,\varrho}) \leq Cn^{-(r+\varrho)}. \quad (57)$$

Proof. We can write the algorithm $\mathcal{A}_n^{IT*}(f, B)$ in the integral form:

$$\mathcal{A}_n^{IT*}(f, B) = \int_0^T \bar{f}(t) dB_t, \quad (58)$$

where the function $\bar{f} : [0, T] \rightarrow \mathbb{R}$ is defined as follows

$$\bar{f}(t) = \sum_{i=1}^{n+2} \mathbf{1}_{[\bar{a}_i, \bar{a}_{i+1})}(t) \sum_{j=0}^r \frac{f^{(j)}(\bar{a}_i)}{j!} (t - \bar{a}_i)^j. \quad (59)$$

From the Itô isometry we have that

$$\mathbb{E} (I(f, B) - \mathcal{A}_n^{IT*}(f, B))^2 = \sum_{i=1}^{n+2} \bar{e}_i^2, \quad \text{where } \bar{e}_i^2 = \int_{\bar{a}_i}^{\bar{a}_{i+1}} (f(t) - \bar{f}(t))^2 dt.$$

If $u_f \notin (c, d)$ then we have from (53) and (54) that the singularity u_f is mild. Therefore, $\bar{e}_i^2 = O(n^{-(2(r+\varrho)+1)})$, for $i = 1, 2, \dots, n+2$, and this gives

$$\mathbb{E} (I(f, B) - \mathcal{A}_n^{IT*}(f, B))^2 = O(n^{-2(r+\varrho)}).$$

Let now $u_f \in (c, d)$. Since $|d_f^j| \leq M_j$, for $j = 0, 1, \dots, r$, and $d - c \leq (T/n)^{2(r+\varrho)}$, we have $\bar{e}_{l+1}^2 = O(n^{-2(r+\varrho)})$. Since function f has no singularities in the remaining intervals, we have as in the regular case that $\bar{e}_i^2 = O(n^{-(2(r+\varrho)+1)})$ for $i \neq l+1$. Hence

$$\mathbb{E} (I(f, B) - \mathcal{A}_n^{IT*}(f, B))^2 = \sum_{i=1, i \neq l+1}^{n+2} \bar{e}_i^2 + \bar{e}_{l+1}^2 \leq Cn^{-2(r+\varrho)},$$

for all $n \geq n_0$, where constants C and n_0 depend only on the parameters of the class $\mathcal{F}_{\text{sng},1}^{r,\varrho}$ and T . This gives (57). \square

Theorems 2 and 3, together with the fact that $\mathcal{F}_{\text{reg}}^{r,\varrho} \subset \mathcal{F}_{\text{sng},1}^{r,\varrho}$, imply matching upper and lower bounds on $e_n^*(\mathcal{F}_{\text{sng},1}^{r,\varrho})$.

Corollary 3. In the class $\mathcal{F}_{\text{sng},1}^{r,\varrho}$ we have the following bounds on the n th minimal error:

$$e_n^*(\mathcal{F}_{\text{sng},1}^{r,\varrho}) = \Theta(n^{-(r+\varrho)}).$$

The optimal algorithm is the adaptive Itô–Taylor scheme \mathcal{A}_n^{IT*} .

Remark 5. We want to stress that it is possible to approximate the Itô integral (2) in other classes of singular functions than described in this paper. For example, in [6] it is shown how optimally approximate scalar functions from the class F_r^1 of piecewise Sobolev functions, that have at most one singular point. Using algorithm from [6] we can define approximation of integral (2) as follows:

$$\bar{\mathcal{A}}_n(f, B) = \int_0^T (\mathcal{A}_n^{\text{ad}} f)(t) dB_t, \quad f \in F_r^1, \quad (60)$$

where $\mathcal{A}_n^{\text{ad}}$ is the algorithm from [6] which gives approximation of a singular function from F_r^1 . It is possible to show for the algorithm $\bar{\mathcal{A}}_n$ that

$$\lim_{n \rightarrow \infty} e(\bar{\mathcal{A}}_n, f) \cdot n^r = C(r, T) \cdot \|f^{(r)}\|_{L^2(0,T)}.$$

Theorem 3 shows that optimal approximation of the Itô integrals of singular functions from the Hölder class can be done by a slightly modified Itô–Taylor algorithm \mathcal{A}_n^{IT*} .

5. Singular case—the class $\mathcal{F}_{\text{sng},p}^{r,\varrho}$ with $p \geq 2$

In this section we deal with bounds on the n th minimal error in classes of functions with multiple singularities. We shall show that the optimal error is $\Theta(n^{-\min\{1/2, \varrho+r\}})$ in the class of algorithms Ψ_n^* .

Proposition 3. Let $p \geq 2$. For $r \geq 0$, $\varrho \in (0, 1]$ and any algorithm $\mathcal{A}_n \in \Psi_n^*$ we have

$$e(\mathcal{A}_n, \mathcal{F}_{\text{sng},p}^{r,\varrho}) \geq e(\mathcal{A}_n, \mathcal{F}_{\text{sng},2}^{r,\varrho}) = \Omega(n^{-1/2}). \quad (61)$$

For $r = 0$, we have that

$$e(\mathcal{A}_n, \mathcal{F}_{\text{sng},p}^{0,\varrho}) = \Omega(n^{-\min\{1/2, \varrho\}}). \quad (62)$$

Proof. Since $\mathcal{F}_{\text{sng},2}^{r,\varrho} \subset \mathcal{F}_{\text{sng},p}^{r,\varrho}$ for $p \geq 2$, it suffices to consider the case $p = 2$. Let a_1, a_2, \dots be the information points computed for the function $f \equiv 0$ (the total number of evaluations is n). The points $\{\tilde{a}_i\}_{i=0}^k$ and the intervals $[\tilde{a}_i, \tilde{a}_{i+1}]$, $[\tilde{a}_i, \tilde{a}_{i+1}]$ are defined in the same way as in the proof of Proposition 2. Let us now consider two functions f_1 and f_2 , where $f_1(t) = M_0 \mathbf{1}_{[\tilde{a}_i, \tilde{a}_{i+1})}(t)$ and $f_2(t) = -f_1(t)$. Each function f_1 and f_2 have two singular points, in \tilde{a}_i and \tilde{a}_{i+1} , and belongs to the class $\mathcal{F}_{\text{sng},2}^{r,\varrho}$. Moreover $I(f_1, B) - I(f_2, B) = 2M_0(B_{\tilde{a}_{i+1}} - B_{\tilde{a}_i})$. Since both functions are equal to 0 at the information points a_i , the algorithm uses those points to evaluate both functions. Therefore $\mathcal{A}_n(f_1, B) = \mathcal{A}_n(f_2, B)$, and hence

$$e(\mathcal{A}_n, \mathcal{F}_{\text{sng},2}^{r,\varrho}) \geq \frac{1}{2} (\mathbb{E} (I(f_1, B) - I(f_2, B))^2)^{1/2} = M_0 (\mathbb{E} |B_{\tilde{a}_{i+1}} - B_{\tilde{a}_i}|^2)^{1/2} \geq M_0 \left(\frac{T}{2k} \right)^{1/2}.$$

This and the fact that $k \leq n+1$ complete the proof of (61). To see (62) we use (61) and Theorem 2 with $r = 0$. \square

Lower bounds from Proposition 3 are sharp, since from Lemma 2 we have the following corollary.

Corollary 4. In the class $\mathcal{F}_{\text{sng},p}^{r,\varrho}$, $p \geq 2$, we have the following bounds on the n th minimal error:

$$e_n^*(\mathcal{F}_{\text{sng},p}^{r,\varrho}) = \Theta(n^{-\min\{1/2, r+\varrho\}}). \quad (63)$$

The optimal algorithm is the Itô–Taylor scheme \mathcal{A}_n^{IT*} .

6. Summary and final remarks

The main results of this paper can be summarized as follows:

Class of functions f	Class of algorithms \mathcal{A}_n	n th minimal error
$\mathcal{F}_{\text{reg}}^{r,\varrho}$	Ψ_n	$\Theta(n^{-(r+\varrho)})$
$\mathcal{F}_{\text{reg}}^{r,\varrho}$	Ψ_n^*	$\Theta(n^{-(r+\varrho)})$
$\mathcal{F}_{\text{reg}}^{r,1}$	Ψ_n^{**}	$\Theta(n^{-(r+1)})$
$\mathcal{F}_{\text{sng},1}^{r,\varrho}$	Ψ_n	$\Theta(n^{-\min\{1/2, r+\varrho\}})$
$\mathcal{F}_{\text{sng},1}^{r,\varrho}$	Ψ_n^*	$\Theta(n^{-(r+\varrho)})$
$\mathcal{F}_{\text{sng},1}^{r,1}$	Ψ_n^{**}	$\Theta(n^{-(r+1)})$
$\mathcal{F}_{\text{sng},p}^{r,\varrho}, 2 \leq p < \infty$	Ψ_n^*	$\Theta(n^{-\min\{1/2, r+\varrho\}})$

We see that adaption with respect to f helps, in the worst case setting, only when there is at most one singularity and $r + \varrho > 1/2$. If $r = 0$ and $\varrho \in (0, 1/2]$ then the error is $\Theta(n^{-\varrho})$ both in the regular and singular case, even if we use adaptive algorithms. In our future work we are planning to consider the asymptotic setting. We believe that the error $O(n^{-(r+\varrho)})$ can be preserved in this case, even when multiple singularities exist.

Results presented in this paper can easily be used for the approximation of the Itô integrals (2) where $f = [f_1, f_2, \dots, f_d]^T$, with $f_k : [0, T] \rightarrow \mathbb{R}$ for $k = 1, 2, \dots, d$, and $B = [B^1, B^2, \dots, B^d]$ is the d -dimensional Brownian motion, $d \geq 2$. It is known that Itô integrals with such f and B are defined as a sum of independent stochastic integrals, each driven by the one-dimensional Brownian motion:

$$\int_0^T f(t) dB_t = \sum_{k=1}^d I(f_k, B^k), \quad (64)$$

where $I(f_k, B^k) = \int_0^T f_k(t) dB_t^k$, see [16]. Provided that each component f_k belongs to $\mathcal{F}_{\text{reg}}^{r,\varrho}$ or $\mathcal{F}_{\text{sng},p}^{r,\varrho}$, we can define approximation of (64) by taking $\mathcal{A}_n^{IT}(f, B) = \sum_{k=1}^d \mathcal{A}_n^{IT}(f_k, B^k)$ or $\mathcal{A}_n^{IT*}(f, B) = \sum_{k=1}^d \mathcal{A}_n^{IT*}(f_k, B^k)$, respectively. The $\mathcal{A}_n^{IT}(f_k, B^k)$ is defined in (14), while $\mathcal{A}_n^{IT*}(f_k, B^k)$ in (55). Using the Itô isometry for integrals (64) one can show that the error of both algorithms $\mathcal{A}_n^{IT}(f, B)$ and $\mathcal{A}_n^{IT*}(f, B)$ is again $O(n^{-(r+\varrho)})$, where constant in the “O” notation depends on \sqrt{d} . Moreover, this bound is optimal which can be easily shown by considering for lower bounds functions $f = [f_1, 0, \dots, 0]^T$, with f_1 defined as in proofs of lower bounds in the scalar case. (The constant \sqrt{d} may also be obtained in the lower bounds by taking functions $f = f_1 \cdot [1, 1, \dots, 1]^T$).

Acknowledgements

I wish to thank Prof. Bolesław Kaciewicz for his valuable suggestions. I would also like to thank two anonymous referees for comments that helped to improve the quality of this paper.

Appendix

The following result gives upper bounds on the remainder term of Taylor expansion of scalar function with multiple singularities.

Lemma 4. Let $c, d \in [0, T]$ ($c < d$) and let for function $g \in \mathcal{F}_{\text{sng},p}^{r,\varrho}$, singularities $u_g^1, u_g^2, \dots, u_g^p \in (c, d)$. Then there exist constants C_k , depending only on the parameters of $\mathcal{F}_{\text{sng},p}^{r,\varrho}$ and k , such that for all $t \in [c, d]$ and $k = 0, 1, \dots, r$ we have

$$\begin{aligned} |g^{(k)}(t) - \hat{g}^{(k)}(t)| &\leq C_k(d-c)^{r-k+\varrho} + \sum_{j_p=k}^r |\Delta_g^{j_p,p}|(d-c)^{j_p-k} + \sum_{j_p=k}^r \sum_{j_{p-1}=j_p}^r |\Delta_g^{j_{p-1},p-1}|(d-c)^{j_{p-1}-k} \\ &\quad + \dots + \sum_{j_p=k}^r \sum_{j_{p-1}=j_p}^r \sum_{j_{p-2}=j_{p-1}}^r \dots \sum_{j_2=j_3}^r \sum_{j_1=j_2}^r |\Delta_g^{j_1,1}|(d-c)^{j_1-k}, \end{aligned} \quad (65)$$

where $\hat{g}(t) = \sum_{j=0}^r (g^{(j)}(c)/j!)(t-c)^j$, for $t \in [c, d]$.

Proof. Let us denote by $u_g^0 := c$ and $u_g^{p+1} := d$. It is easy to see that for $r = 0$ and $t \in [c, d]$ we have that

$$|g(t) - \hat{g}(t)| \leq C_0 \left((d - c)^\varrho + \sum_{j=1}^p |\Delta_g^{0,j}| \right). \quad (66)$$

Now, let $r \geq 1$. For $t \in [u_g^0, u_g^1]$ ($q = 0$) and $k = 0, 1, \dots, r - 1$, since $g^{(r)}$ is regular in this interval, we have

$$R_0^{(k)}(t) = g^{(k)}(t) - \hat{g}^{(k)}(t) = \int_0^1 (g^{(r)}(\theta t + (1 - \theta)u_g^0) - g^{(r)}(u_g^0)) (t - u_g^0)^{r-k} \frac{(1 - \theta)^{r-k-1}}{(r - k - 1)!} d\theta,$$

and $R_0^{(r)}(t) = g^{(r)}(t) - \hat{g}^{(r)}(t) = g^{(r)}(t) - g^{(r)}(c)$. Therefore

$$|R_0^{(k)}(t)| = |g^{(k)}(t) - \hat{g}^{(k)}(t)| \leq L(d - c)^{r-k+\varrho}, \quad (67)$$

for $k = 0, 1, \dots, r$ and $t \in [u_g^0, u_g^1]$.

Let $q \in \{1, 2, \dots, p\}$ and $t \in [u_g^q, u_g^{q+1}]$ (if $q = p$ then we take $t \in [u_g^p, u_g^{p+1}]$). Then we have by the Taylor's formula

$$g^{(k)}(t) = \sum_{j_q=k}^r \frac{g^{(j_q)}(u_g^{q+})}{(j_q - k)!} (t - u_g^q)^{j_q-k} + R_q^{(k)}(t),$$

where

$$R_q^{(k)}(t) = \int_0^1 (g^{(r)}(\theta t + (1 - \theta)u_g^{q+}) - g^{(r)}(u_g^{q+})) (t - u_g^q)^{r-k} \frac{(1 - \theta)^{r-k-1}}{(r - k - 1)!} d\theta, \quad (68)$$

for $k = 0, 1, \dots, r - 1$ and $R_q^{(r)}(t) = g^{(r)}(t) - g^{(r)}(u_g^{q+})$. Hence, we have that

$$|R_q^{(k)}(t)| \leq L(d - c)^{r-k+\varrho}, \quad (69)$$

for all $t \in [u_g^q, u_g^{q+1}]$ and $k = 0, 1, \dots, r$. Moreover, we can express function \hat{g} as

$$\hat{g}^{(k)}(t) = \sum_{j_q=k}^r \frac{\hat{g}^{(j_q)}(u_g^q)}{(j_q - k)!} (t - u_g^q)^{j_q-k},$$

for $k = 0, 1, \dots, r$. Thus we have for $q = 1, 2, \dots, p$, $t \in [u_g^q, u_g^{q+1}]$ and $k = 0, 1, \dots, r$ that

$$\begin{aligned} |g^{(k)}(t) - \hat{g}^{(k)}(t)| &= \left| \sum_{j_q=k}^r \frac{1}{(j_q - k)!} (g^{(j_q)}(u_g^{q+}) - \hat{g}^{(j_q)}(u_g^q)) (t - u_g^q)^{j_q-k} + R_q^{(k)}(t) \right| \\ &\leq \sum_{j_q=k}^r |\Delta_g^{j_q,q}| (d - c)^{j_q-k} + \sum_{j_q=k}^r |g^{(j_q)}(u_g^{q-}) - \hat{g}^{(j_q)}(u_g^q)| (d - c)^{j_q-k} + L(d - c)^{r-k+\varrho}. \end{aligned} \quad (70)$$

Recursive inequality (70) binds the error in the interval $[u_g^q, u_g^{q+1}]$ with the error on the previous interval $[u_g^{q-1}, u_g^q]$. From this inequality we get (65). \square

References

- [1] P.E. Kloeden, E. Platen, *Numerical Solution of Stochastic Differential Equations*, Springer-Verlag, Berlin, Heidelberg, 1992.
- [2] G.N. Milstein, M.V. Tretyakov, *Stochastic Numerics for Mathematical Physics*, Scientific Computation, Springer-Verlag, Berlin, Heidelberg, 2004.
- [3] L. Plaskota, G.W. Wasilkowski, Adaption allows efficient integration of functions with unknown singularities, *Numer. Math.* 102 (2005) 123–144.
- [4] L. Plaskota, G.W. Wasilkowski, Uniform approximation of piecewise r -smooth and globally continuous functions, *SIAM J. Numer. Anal.* 47 (2009) 762–785.
- [5] L. Plaskota, G.W. Wasilkowski, The power of adaptive algorithms for functions with singularities, *J. Fixed Point Theory Appl.* 6 (2009) 227–248.
- [6] L. Plaskota, G.W. Wasilkowski, Y. Zhao, The power of adaption for approximating functions with singularities, *Math. Comp.* 77 (2008) 2309–2338.
- [7] B. Kaciewicz, P. Przybyłowicz, Optimal adaptive solution of initial-value problems with unknown singularities, *J. Complexity* 24 (2008) 455–476.
- [8] G.W. Wasilkowski, H. Woźniakowski, On the complexity of stochastic integration, *Math. Comp.* 70 (2001) 685–698.
- [9] P. Przybyłowicz, Linear information for approximation of the Itô integrals, *Numer. Algorithms* 52 (2009) 677–699.
- [10] L. Plaskota, K. Ritter, G.W. Wasilkowski, Optimal designs for weighted approximation and integration of stochastic processes on $[0, \infty)$, *J. Complexity* 20 (2004) 108–131.
- [11] J. Sacks, D. Ylvisaker, Statistical designs and integral approximation, in: R. Pyke (Ed.), *Proc. 12th Bienn. Semin. Can. Math. Congr.*, Can. Math. Soc., Montreal, 1970.
- [12] J. Sacks, D. Ylvisaker, Design for regression problems with correlated errors III, *Ann. Math. Statist.* 41 (6) (1970) 2057–2074.
- [13] G.W. Wasilkowski, Information of varying cardinality, *J. Complexity* 2 (1986) 204–228.

- [14] B. Kacewicz, Improved bounds on the randomized and quantum complexity of initial-value problems, *J. Complexity* 21 (2005) 740–756.
- [15] E. Novak, Deterministic and Stochastic Error Bounds in Numerical Analysis, in: *Lecture Notes in Mathematics*, vol. 1349, Springer-Verlag, New York, 1988.
- [16] I. Karatzas, S.E. Shreve, *Brownian Motion and Stochastic Calculus*, 2nd ed., Springer-Verlag, New York, 1991.
- [17] K. Ritter, Average-Case Analysis of Numerical Problems, in: *Lecture Notes in Mathematics*, vol. 1733, Springer-Verlag, Berlin, 2000.
- [18] J.F. Traub, G.W. Wasilkowski, H. Woźniakowski, *Information-Based Complexity*, Academic Press, New York, 1988.