Available online at www.sciencedirect.com

SCIENCE DIRECT®

Journal of Computational Physics 197 (2004) 275–298

JOURNAL OF COMPUTATIONAL PHYSICS

www.elsevier.com/locate/jcp

Evaluation of conditional Wiener integrals by numerical integration of stochastic differential equations

G.N. Milstein [a,b], M.V. Tretyakov [c,d,*]

[a] *Weierstrass Institute for Applied Analysis and Stochastics, Mohrenstr. 39, D-10117 Berlin, Germany*
[b] *Department of Mathematics, Ural State University, Lenin Str. 51, 620083 Ekaterinburg, Russia*
[c] *Department of Mathematics, University of Leicester, Leicester LE1 7RH, UK*
[d] *Institute of Mathematics and Mechanics, Russian Academy of Sciences, S. Kovalevskaya Str. 16, 620219 Ekaterinburg, Russia*

Received 7 July 2003; received in revised form 27 November 2003; accepted 2 December 2003
Available online 23 January 2004

Abstract

Numerical integration of stochastic differential equations together with the Monte Carlo technique is used to evaluate conditional Wiener integrals of exponential-type functionals. An explicit Runge–Kutta method of order four and implicit Runge–Kutta methods of order two are constructed. The corresponding convergence theorems are proved. To reduce the Monte Carlo error, a variance reduction technique is considered. Results of numerical experiments are presented.

Keywords: Conditional Wiener integrals; Feynman path integrals; Numerical integration of stochastic differential equations; Monte Carlo simulation

1. Introduction

We consider Wiener integrals

$$\mathscr{J} = \int_{C^d_{0,a;T,b}} F(x(\cdot))\, \mathrm{d}\mu^{T,b}_{0,a}(x) \tag{1.1}$$

of the exponential-type functionals

$$F(x(\cdot)) = \exp\left[\int_0^T f(t,x(t))\,\mathrm{d}t\right]. \tag{1.2}$$

* Corresponding author.
E-mail addresses: milstein@wias-berlin.de (G.N. Milstein), M.Tretiakov@le.ac.uk (M.V. Tretyakov).

0021-9991/$ - see front matter
doi:10.1016/j.jcp.2003.12.001

Here $\mu_{0,a}^{T,b}(x)$ is a conditional Wiener measure which corresponds to the Brownian paths $X_{0,a}^{T,b}(t)$ with fixed initial and final points, i.e., it corresponds to the d-dimensional Brownian bridge from a at the time $t = 0$ into b at the time $t = T$. The integral (1.1) is understood in the sense of Lebesgue integral with respect to the measure $\mu_{0,a}^{T,b}(x)$ and is taken over the set $C_{0,a;T,b}^{d}$ of all d-dimensional continuous vector-functions $x(t)$ satisfying the conditions $x(0) = a$, $x(T) = b$ (see, e.g. [4]).

A relation of such integrals with quantum physics and some equations of mathematical physics can be found, e.g., in [2–4,9,13]. In particular, the Feynman path integral of the form

$$\mathscr{I} = \int \exp\left(\int_0^T \left[\frac{m\dot{x}^2(t)}{2} - V(x(t))\right] \mathrm{d}t\right) \mathscr{D}x(t)$$

is equivalent to the integral (1.1), (1.2) with $f = -V$.

Numerical evaluation of Wiener integrals is an important and difficult task. Many approaches are proposed for solving this problem (see, e.g. [1,2,15] and references therein). As a rule, the known numerical methods reduce a path integral to a high dimensional integral which is then approximated using either classical or Monte Carlo methods. The high dimensionality of these integrals makes calculation of the Wiener integrals extremely difficult.

In [5,10,14], the approach using numerical integration (in the weak sense) of stochastic differential equations with application of the Monte Carlo technique is proposed for computation of Wiener integrals of the form

$$\mathscr{I} = \int_{C_{0,0}^{d}} F(x(\cdot))\, \mathrm{d}\mu_{0,0}(x), \tag{1.3}$$

where $\mu_{0,0}(x)$ is a Wiener measure corresponding to Brownian paths with the fixed initial point $(0, 0)$, $F(x(\cdot)) = \varphi(x(T), \int_0^T f(t, x(t))\, \mathrm{d}t)$, and $\varphi(x, z)$ is a function of $d + 1$ arguments from a sufficiently wide class. The approach is based on the following probabilistic representation of the integral (1.3):

$$\mathscr{I} = E\varphi(X_{0,0}(T), Z_{0,0,0}(T)), \tag{1.4}$$

where $X_{0,0}(t),\ Z_{0,0,0}(t),\ 0 \leqslant t \leqslant T$, is the solution of the $(d + 1)$-dimensional system of stochastic differential equations (SDEs)

$$\begin{aligned} \mathrm{d}X &= \mathrm{d}w(t), \quad X_{0,0}(0) = 0, \\ \mathrm{d}Z &= f(t, X^1, \ldots, X^d)\, \mathrm{d}t, \quad Z_{0,0,0}(0) = 0, \end{aligned} \tag{1.5}$$

and $w(t) = (w^1(t), \ldots, w^d(t))^{\mathsf{T}}$ is a d-dimensional standard Wiener process.

An efficiency of this approach is due to the fact that the system (1.5) has the fixed dimension $d + 1$ and the corresponding accuracy is reached by means of a choice of a method and a step of numerical integration and a number M of Monte Carlo simulations. Thus, the problem of calculating the infinite-dimensional Wiener integral $\mathscr{I}$ is reduced to the Cauchy problem (1.5). This problem can naturally be regarded as one-dimensional since it contains only one independent variable. We underline that in other methods the path integral is reduced to a high dimensional Riemann integral and the accuracy is reached by increasing its dimension. The approach based on the probabilistic representation (1.4), (1.5) is especially effective for evaluating Wiener integrals (1.3) in the case of exponential-type functionals because in this case there are fourth-order Runge–Kutta methods [5,10].

Here the approach of [5,10,14] is developed for evaluating the conditional Wiener integral (1.1), (1.2). The corresponding probabilistic representation contains a more complicated system than (1.5). The solution of this system gives a Markov representation of the Brownian bridge. The system is singular and this circumstance stipulates a certain complexity of theoretical proofs. Nevertheless the constructed fourth-order Runge–Kutta algorithms are equally simple and effective as in the case of the Wiener integral (1.3). The effectiveness of these algorithms allows us to evaluate integrals (1.1), (1.2) for a large dimension d.

In this paper, we restrict ourselves to conditional Wiener integrals of exponential-type functionals although the approach can also be applied to conditional Wiener integrals of functionals of an integral type (cf. [14]).

In Section 2, a fourth-order explicit Runge–Kutta method is constructed. Its one-step error is analyzed in Section 3. Implicit methods of order two are derived in Section 4. These methods have an implicitness with respect to the linear part only which is easily analytically resolved. In our approach, there are two types of errors: the error of numerical integration and the Monte Carlo error. To reduce the Monte Carlo error, the method of control variates is considered in Section 5. Some numerical tests are presented in Section 6. A summary of the obtained results is given in Section 7. Proofs of convergence theorems can be found in the appendix.

2. Explicit Runge–Kutta method of order four

As it is known [7,8], the d-dimensional Brownian bridge $X(t) = X_{0,a}(t) = X_{0,a}^{T,b}(t)$, $0 \leqslant t \leqslant T$, from a to b can be characterized as the pathwise unique solution of the system of SDEs

$$\mathrm{d}X = \frac{b - X}{T - t}\mathrm{d}t + \mathrm{d}w(t), \quad 0 \leqslant t < T, \; X(0) = a, \tag{2.1}$$

with

$$X(T) = b, \tag{2.2}$$

where $w(t) = (w^1(t), \ldots, w^d(t))^\mathsf{T}$ is a d-dimensional standard $\mathscr{F}_t$-Wiener process. The system is considered on a probability space $(\Omega, \mathscr{F}, P)$, and $\mathscr{F}_t$, $0 \leqslant t \leqslant T$, is a non-decreasing family of σ-algebras of $\mathscr{F}$.

Let us also introduce the scalar equation

$$\mathrm{d}Y = f(t, X(t))Y\,\mathrm{d}t, \quad 0 \leqslant t \leqslant T, \; Y(0) = 1, \tag{2.3}$$

where $X(t)$ is defined by (2.1), (2.2) and $f(t,x)$ is the same as in (1.2). Then the Wiener integral (1.1), (1.2) is equal to

$$\mathscr{J} = EY(T). \tag{2.4}$$

Thus, evaluation of the Wiener integral (1.1), (1.2) is reduced to the problem of numerical integration of the system (2.1)–(2.3).

Introduce a discretization of the time interval $[0, T]$, for definiteness the equidistant one with a time step $h > 0$:

$$t_k = kh, \quad k = 0, \ldots, N, \; t_N = T,$$

and let $t_{k+1/2} := t_k + h/2$.

To get a higher order method for (2.1)–(2.3), we need to simulate the solution of (2.1) exactly. The solution of (2.1) is

$$X(t) = a\frac{T - t}{T} + b\frac{t}{T} + (T - t)\int_0^t \frac{\mathrm{d}w(s)}{T - s}. \tag{2.5}$$

Hence

$$X(t+h) = X(t) + h\frac{b - X(t)}{T-t} + (T-t-h)\int_t^{t+h} \frac{\mathrm{d}w(s)}{T-s}. \tag{2.6}$$

We have

$$\begin{aligned} &E\left[(T-t-h)\int_t^{t+h} \frac{\mathrm{d}w(s)}{T-s}\middle|X(t)\right] = 0,\\ &E\left[(T-t-h)\int_t^{t+h} \frac{\mathrm{d}w(s)}{T-s}\middle|X(t)\right]^2 = \left(1-\frac{h}{T-t}\right)h. \end{aligned} \tag{2.7}$$

We can exactly simulate the solution of (2.1) by a simple recurrent procedure based on the formula

$$X(t+h) = X(t) + h\frac{b-X(t)}{T-t} + h^{1/2}\sqrt{\frac{T-t-h}{T-t}}\xi, \quad t < T, \tag{2.8}$$

where ξ is a random vector which components are Gaussian random variables with zero mean and unit variance and they are independent of $X(t)$.

Now let us formally apply a standard deterministic explicit fourth-order Runge–Kutta method to Eq. (2.3) assuming that $X(t)$ is a known function. Then, taking into account (2.8), we obtain the following algorithm for integrating the system (2.1)–(2.3):

$$\begin{aligned} &X(0) = a,\\ &X(t_{k+1/2}) = X(t_k) + \frac{h}{2}\frac{b - X(t_k)}{T-t_k} + \frac{h^{1/2}}{\sqrt{2}}\sqrt{\frac{T-t_{k+1/2}}{T-t_k}}\xi_{k+1/2}, \quad k = 0,\ldots,N-1,\\ &X(t_{k+1}) = X(t_{k+1/2}) + \frac{h}{2}\frac{b-X(t_{k+1/2})}{T-t_{k+1/2}} + \frac{h^{1/2}}{\sqrt{2}}\sqrt{\frac{T-t_{k+1}}{T-t_{k+1/2}}}\xi_{k+1}, \quad k = 0,\ldots,N-2,\\ &X(t_N) = b, \end{aligned} \tag{2.9}$$

$$\begin{aligned} &Y_0 = 1,\\ &k_1 = f(t_k, X(t_k))Y_k, \quad k_2 = f(t_{k+1/2}, X(t_{k+1/2}))[Y_k + hk_1/2],\\ &k_3 = f(t_{k+1/2}, X(t_{k+1/2}))[Y_k + hk_2/2], \quad k_4 = f(t_{k+1}, X(t_{k+1}))[Y_k + hk_3],\\ &Y_{k+1} = Y_k + \frac{h}{6}(k_1 + 2k_2 + 2k_3 + k_4), \quad k = 0,\ldots,N-1, \end{aligned} \tag{2.10}$$

where $\xi_{k+1/2}$, ξ_{k+1} are d-dimensional random vectors which components are mutually independent random variables with standard normal distribution $\mathcal{N}(0,1)$.

Since the function $X(t)$ is non-smooth, the deterministic result on the accuracy order of the involved Runge–Kutta method is not applicable here and a separate convergence theorem is needed. The following theorem is proved under some assumptions on the function $f(t,x)$ (see them after (3.4)). Its proof is based on a thorough analysis of the one-step error which is made in the next section and in Appendix A. The proof of the convergence theorem itself is given in Appendix B.

Theorem 2.1. *The method* (2.9), (2.10) *applied to evaluation of the Wiener integral* (2.4) *is of fourth order of accuracy*, *i.e.*,

$$|\mathcal{J} - EY_N| = |EY(T) - EY_N| \leqslant Kh^4, \tag{2.11}$$

where the constant K *is independent of* h.

3. Theorem on one-step error

In this section, we consider a one-step error of the method (2.9), (2.10).

We say that a function $g(s,x)$, $s \in [0,T]$, $x \in \mathbf{R}^d$, belongs to the class $\mathbf{F}$ (with respect to the variable x), written as $g \in \mathbf{F}$, if there are constants $K > 0$ and $\kappa > 0$ such that for all $x \in \mathbf{R}^d$ the following inequality holds uniformly in $s \in [0,T]$:

$$|g(s,x)| \leqslant K(1+|x|^{\kappa}). \tag{3.1}$$

Introduce the operator

$$L = \frac{\partial}{\partial t} + \sum_{i=1}^{d} \frac{b^i - x^i}{T-t}\frac{\partial}{\partial x^i} + \frac{1}{2}\sum_{i=1}^{d}\frac{\partial^2}{(\partial x^i)^2}, \quad 0 \leqslant t < T. \tag{3.2}$$

We observe that this operator contains singularity since the denominator $T-t$ tends to zero as t goes to T.

Consider the function

$$u(t,x) = EY_{t,x,1}(T). \tag{3.3}$$

It satisfies the Cauchy problem

$$Lu + fu = 0, \quad 0 \leqslant t < T, \qquad x \in \mathbf{R}^d, \quad u(T,x) = 1. \tag{3.4}$$

We assume that the function $f(t,x)$ is sufficiently smooth, belongs to the class $\mathbf{F}$ together with its partial derivatives of a sufficiently high order and is such that the problem (3.4) has a unique solution which is sufficiently smooth and belongs to the class $\mathbf{F}$ together with its partial derivatives of a sufficiently high order. In addition, we suppose that $EY^2(t)$ exists and bounded on $[0,T]$ and that for all sufficiently small h the second moments EY_k^2 are uniformly bounded with respect to h. For instance, the latter conditions are satisfied when the function $f(t,x)$ is bounded. Therefore, theoretically, we can use Theorem 2.1, approximating $f(t,x)$ (if it is unbounded) by an appropriate bounded function.

Let g be a sufficiently smooth function belonging to the class $\mathbf{F}$ together with its partial derivatives up to a sufficiently high order. Then the expectations for a nonnegative integer m

$$EL^l g(\theta, X(\theta)), \quad l = 0,1,\ldots,m+1,$$

$$E\left(\frac{\partial}{\partial x^i}L^l g(\theta,X(\theta))\right)^2, \quad l=0,1,\ldots,m, \ i=1,\ldots,d,$$

exist and are continuous with respect to $0 \leqslant \theta < T$. And the following formulas are also valid for $t \leqslant s \leqslant t+h < T$:

$$\begin{aligned} E(g(t+h,X_{t,x}(t+h))|\mathscr{F}_s) &= g(s,X_{t,x}(s)) + (t+h-s)Lg(s,X_{t,x}(s)) + \cdots + \frac{(t+h-s)^m}{m!}L^m g(s,X_{t,x}(s)) \\ &\quad + \int_s^{t+h} \frac{(t+h-\theta)^m}{m!} E(L^{m+1}g(\theta,X_{t,x}(\theta))|\mathscr{F}_s)\,\mathrm{d}\theta, \end{aligned} \tag{3.5}$$

$$\begin{aligned} Eg(t+h,X_{t,x}(t+h)) &= g(t,x) + hLg(t,x) + \cdots + \frac{h^m}{m!}L^m g(t,x) \\ &\quad + \int_t^{t+h} \frac{(t+h-\theta)^m}{m!} EL^{m+1}g(\theta,X_{t,x}(\theta))\,\mathrm{d}\theta. \end{aligned} \tag{3.6}$$

The expansions (3.5) and (3.6) are analogous to an expansion of semigroups. Their proof is available in [10, p. 137].

It is convenient to introduce the additional notation for the approximation defined by (2.10): $\bar{Y}_{0,a,1}(t_k) = Y_k$ and also $\bar{Y}_{s,x,y}(t)$, $t \geqslant s$, by which we mean the approximation of (2.3) started from y at $t = s$ with $X(s) = x$.

It is not difficult to see that

$$Y_{t,x,y}(t+t') = yY_{t,x,1}(t+t'), \quad \bar{Y}_{t_k,x,y}(t_{k+k'}) = y\bar{Y}_{t_k,x,1}(t_{k+k'}), \quad EY_{t,x,y}(T) = yEY_{t,x,1}(T) = yu(t,x), \tag{3.7}$$

where $u(t,x)$ is the solution of the problem (3.4).

Recall that $t_0 = 0, X_0 = a,\ Y_0 = 1$. Using (3.3) and (3.7) and the fact that we simulate $X_k = X(t_k)$ exactly, we can represent the *global error* of the method (2.9), (2.10) (cf. (2.11)) in the form

$$\begin{aligned}
\left|EY_{0,a,1}(T) - E\bar{Y}_{0,a,1}(T)\right| &= \left|EY_{t_0,X_0,Y_0}(T) - EY_N\right| = |u(t_0,X_0)Y_0 - Eu(t_N,X_N)Y_N| \\
&= \left|\sum_{k=0}^{N-1}\left[Eu(t_k,X(t_k))Y_k - Eu(t_{k+1},X(t_{k+1}))\bar{Y}_{t_k,X_k,Y_k}(t_{k+1})\right]\right| \\
&= \left|\sum_{k=0}^{N-1} EY_k\left[u(t_k,X(t_k)) - u(t_{k+1},X(t_{k+1}))\bar{Y}_{t_k,X_k,1}(t_{k+1})\right]\right| \\
&\leqslant \sum_{k=0}^{N-1}\left|EY_k\left[u(t_k,X(t_k)) - u(t_{k+1},X_{t_k,X_k}(t_{k+1}))\bar{Y}_{t_k,X_k,1}(t_{k+1})\right]\right|.
\end{aligned} \tag{3.8}$$

We have

$$\begin{aligned}
R_k &:= \left|EY_k\left[u(t_k,X(t_k)) - u(t_{k+1},X_{t_k,X_k}(t_{k+1}))\bar{Y}_{t_k,X_k,1}(t_{k+1})\right]\right| \\
&= \left|EY_kE\left[u(t_k,X_k) - u(t_{k+1},X_{t_k,X_k}(t_{k+1}))\bar{Y}_{t_k,X_k,1}(t_{k+1})|\mathscr{F}_{t_k}\right]\right|.
\end{aligned} \tag{3.9}$$

First, we analyze R_k for $k = 0,\ldots,N-2$. To this end, we consider the *one-step error* for $0 \leqslant t < T-h$:

$$r(t,x) := Eu(t+h,X_{t,x}(t+h))\bar{Y}_{t,x,1}(t+h) - u(t,x). \tag{3.10}$$

We rewrite (2.10) on a single step in the form:

$$\begin{aligned}
\bar{Y}_{t,x,1}(t+h) = 1 &+ \frac{h}{6}(f_0 + 4f_{1/2} + f_1) + \frac{h^2}{6}\left(f_0f_{1/2} + f_{1/2}^2 + f_{1/2}f_1\right) \\
&+ \frac{h^3}{12}\left(f_0f_{1/2}^2 + f_{1/2}^2f_1\right) + \frac{h^4}{24}f_0f_{1/2}^2f_1,
\end{aligned} \tag{3.11}$$

where $f_0 := f(t,x)$, $f_{1/2} := f(t+h/2, X_{t,x}(t+h/2))$, and $f_1 := f(t+h, X_{t,x}(t+h))$.

Using (3.6), we get

$$\begin{aligned}
Eu(t+h,X_{t,x}(t+h)) &= u(t,x) + hLu(t,x) + \frac{h^2}{2}L^2u(t,x) + \frac{h^3}{6}L^3u(t,x) + \frac{h^4}{24}L^4u(t,x) \\
&+ \int_t^{t+h}\frac{(t+h-\theta)^4}{24}EL^5u(\theta,X_{t,x}(\theta))\,\mathrm{d}\theta,
\end{aligned} \tag{3.12}$$

$$
\begin{aligned}
Ef_0u(t+h,X_{t,x}(t+h)) &= f_0Eu(t+h,X_{t,x}(t+h)) \\
&= f_0\left[u(t,x)+hLu(t,x)+\frac{h^2}{2}L^2u(t,x)+\frac{h^3}{6}L^3u(t,x)\right. \\
&\quad \left.+\int_t^{t+h}\frac{(t+h-\theta)^3}{6}EL^4u(\theta,X_{t,x}(\theta))\,\mathrm{d}\theta\right],
\end{aligned} \tag{3.13}
$$

$$
\begin{aligned}
Ef_1u(t+h,X_{t,x}(t+h)) &= f_0u(t,x)+hL(fu)(t,x)+\frac{h^2}{2}L^2(fu)(t,x)+\frac{h^3}{6}L^3(fu)(t,x) \\
&\quad +\int_t^{t+h}\frac{(t+h-\theta)^3}{6}EL^4(fu)(\theta,X_{t,x}(\theta))\,\mathrm{d}\theta.
\end{aligned} \tag{3.14}
$$

Further,

$$
Ef_{1/2}u(t+h,X_{t,x}(t+h)) = E\big(f_{1/2}E\big[u(t+h,X_{t,x}(t+h))|\mathscr{F}_{t+h/2}\big]\big),
$$

and by (3.5) we obtain

$$
\begin{aligned}
E\big[u(t+h,X_{t,x}(t+h))|\mathscr{F}_{t+h/2}\big] &= u(t+h/2,X_{t,x}(t+h/2))+\frac{h}{2}Lu(t+h/2,X_{t,x}(t+h/2)) \\
&\quad +\frac{h^2}{8}L^2u(t+h/2,X_{t,x}(t+h/2))+\frac{h^3}{48}L^3u(t+h/2,X_{t,x}(t+h/2)) \\
&\quad +\int_{t+h/2}^{t+h}\frac{(t+h-\theta)^3}{6}E\big[L^4u(\theta,X_{t,x}(\theta))|\mathscr{F}_{t+h/2}\big]\mathrm{d}\theta,
\end{aligned}
$$

then

$$
\begin{aligned}
Ef_{1/2}u(t+h,X_{t,x}(t+h)) &= f_0u(t,x)+\frac{h}{2}L(fu)(t,x)+\frac{h^2}{8}L^2(fu)(t,x)+\frac{h^3}{48}L^3(fu)(t,x) \\
&\quad +\int_t^{t+h/2}\frac{(t+h/2-\theta)^3}{6}EL^4(fu)(\theta,X_{t,x}(\theta))\,\mathrm{d}\theta+\frac{h}{2}f_0Lu(t,x) \\
&\quad +\frac{h^2}{4}L(fLu)(t,x)+\frac{h^3}{16}L^2(fLu)(t,x) \\
&\quad +\frac{h}{2}\int_t^{t+h/2}\frac{(t+h/2-\theta)^2}{2}EL^3(fLu)(\theta,X_{t,x}(\theta))\,\mathrm{d}\theta+\frac{h^2}{8}f_0L^2u(t,x) \\
&\quad +\frac{h^3}{16}L\big(fL^2u\big)(t,x)+\frac{h^2}{8}\int_t^{t+h/2}(t+h/2-\theta)EL^2\big(fL^2u\big)(\theta,X_{t,x}(\theta))\,\mathrm{d}\theta \\
&\quad +\frac{h^3}{48}f_0L^3u(t,x)+\frac{h^3}{48}\int_t^{t+h/2}EL\big(fL^3u\big)(\theta,X_{t,x}(\theta))\,\mathrm{d}\theta \\
&\quad +Ef_{1/2}\int_{t+h/2}^{t+h}\frac{(t+h-\theta)^3}{6}L^4u(\theta,X_{t,x}(\theta))\,\mathrm{d}\theta.
\end{aligned} \tag{3.15}
$$

Analogously, we get

$$\begin{aligned} Ef_0f_{1/2}u(t+h,X_{t,x}(t+h)) &= f_0^2u(t,x)+\frac{h}{2}f_0L(fu)(t,x)+\frac{h^2}{8}f_0L^2(fu)(t,x) \\ &\quad + f_0\int_t^{t+h/2}\frac{(t+h/2-\theta)^2}{2}EL^3(fu)(\theta,X_{t,x}(\theta))\,\mathrm{d}\theta+\frac{h}{2}f_0^2Lu(t,x) \\ &\quad + \frac{h^2}{4}f_0L(fLu)(t,x)+\frac{h}{2}f_0\int_t^{t+h/2}(t+h/2-\theta)EL^2(fLu)(\theta,X_{t,x}(\theta))\,\mathrm{d}\theta \\ &\quad + \frac{h^2}{8}f_0^2L^2u(t,x)+\frac{h^2}{8}f_0\int_t^{t+h/2}EL\big(fL^2u\big)(\theta,X_{t,x}(\theta))\,\mathrm{d}\theta \\ &\quad + f_0Ef_{1/2}\int_{t+h/2}^{t+h}\frac{(t+h-\theta)^2}{2}L^3u(\theta,X_{t,x}(\theta))\,\mathrm{d}\theta, \end{aligned} \tag{3.16}$$

$$\begin{aligned} Ef_{1/2}^2u(t+h,X_{t,x}(t+h)) &= f_0^2u(t,x)+\frac{h}{2}L\big(f^2u\big)(t,x)+\frac{h^2}{8}L^2\big(f^2u\big)(t,x) \\ &\quad + \int_t^{t+h/2}\frac{(t+h/2-\theta)^2}{2}EL^3\big(f^2u\big)(\theta,X_{t,x}(\theta))\,\mathrm{d}\theta+\frac{h}{2}f_0^2Lu(t,x) \\ &\quad + \frac{h^2}{4}L\big(f^2Lu\big)(t,x)+\frac{h}{2}\int_t^{t+h/2}(t+h/2-\theta)EL^2\big(f^2Lu\big)(\theta,X_{t,x}(\theta))\,\mathrm{d}\theta \\ &\quad + \frac{h^2}{8}f_0^2L^2u(t,x)+\frac{h^2}{8}\int_t^{t+h/2}EL\big(f^2L^2u\big)(\theta,X_{t,x}(\theta))\,\mathrm{d}\theta \\ &\quad + Ef_{1/2}^2\int_{t+h/2}^{t+h}\frac{(t+h-\theta)^2}{2}L^3u(\theta,X_{t,x}(\theta))\,\mathrm{d}\theta, \end{aligned} \tag{3.17}$$

$$\begin{aligned} Ef_{1/2}f_1u(t+h,X_{t,x}(t+h)) &= f_0^2u(t,x)+\frac{h}{2}L\big(f^2u\big)(t,x)+\frac{h^2}{8}L^2\big(f^2u\big)(t,x) \\ &\quad + \int_t^{t+h/2}\frac{(t+h/2-\theta)^2}{2}EL^3\big(f^2u\big)(\theta,X_{t,x}(\theta))\,\mathrm{d}\theta+\frac{h}{2}f_0L(fu)(t,x) \\ &\quad + \frac{h^2}{4}L(fL(fu))(t,x)+\frac{h}{2}\int_t^{t+h/2}(t+h/2-\theta)EL^2(fL(fu))(\theta,X_{t,x}(\theta))\,\mathrm{d}\theta \\ &\quad + \frac{h^2}{8}f_0L^2(fu)(t,x)+\frac{h^2}{8}\int_t^{t+h/2}EL\big(fL^2(fu)\big)(\theta,X_{t,x}(\theta))\,\mathrm{d}\theta \\ &\quad + Ef_{1/2}\int_{t+h/2}^{t+h}\frac{(t+h-\theta)^2}{2}L^3(fu)(\theta,X_{t,x}(\theta))\,\mathrm{d}\theta, \end{aligned} \tag{3.18}$$

$$\begin{aligned} Ef_0f_{1/2}^2u(t+h,X_{t,x}(t+h)) &= f_0^3u(t,x)+\frac{h}{2}f_0L\big(f^2u\big)(t,x)+f_0\int_t^{t+h/2}(t+h/2-\theta)EL^2\big(f^2u\big)(\theta,X_{t,x}(\theta))\,\mathrm{d}\theta \\ &\quad + \frac{h}{2}f_0^3Lu(t,x)+\frac{h}{2}f_0\int_t^{t+h/2}EL\big(f^2Lu\big)(\theta,X_{t,x}(\theta))\,\mathrm{d}\theta \\ &\quad + f_0Ef_{1/2}^2\int_{t+h/2}^{t+h}(t+h-\theta)L^2u(\theta,X_{t,x}(\theta))\,\mathrm{d}\theta, \end{aligned} \tag{3.19}$$

$$
\begin{aligned}
Ef_{1/2}^2 f_1 u(t+h, X_{t,x}(t+h)) &= f_0^3 u(t,x) + \frac{h}{2} L\big(f^3 u\big)(t,x) + \int_t^{t+h/2} (t + h/2 - \theta) E L^2\big(f^3 u\big)(\theta, X_{t,x}(\theta))\,\mathrm{d}\theta \\
&\quad + \frac{h}{2} f_0^2 L(fu)(t,x) + \frac{h}{2} \int_t^{t+h/2} E L\big(f^2 L(fu)\big)(\theta, X_{t,x}(\theta))\,\mathrm{d}\theta \\
&\quad + Ef_{1/2}^2 \int_{t+h/2}^{t+h} (t+h-\theta) L^2(fu)(\theta, X_{t,x}(\theta))\,\mathrm{d}\theta,
\end{aligned} \tag{3.20}
$$

$$
\begin{aligned}
Ef_0 f_{1/2}^2 f_1 u(t+h, X_{t,x}(t+h)) &= f_0^4 u(t,x) + f_0 \int_t^{t+h/2} E L\big(f^3 u\big)(\theta, X_{t,x}(\theta))\,\mathrm{d}\theta \\
&\quad + f_0 E f_{1/2}^2 \int_{t+h/2}^{t+h} L(fu)(\theta, X_{t,x}(\theta))\,\mathrm{d}\theta.
\end{aligned} \tag{3.21}
$$

Substituting (3.11)–(3.21) in (3.10), we obtain

$$
\begin{aligned}
r &= h[Lu + fu] + \frac{h^2}{2}\left[L^2 u + L(fu) + fLu + f^2 u\right] + \frac{h^3}{6}\left[L^3 u + L^2(fu) + fL^2 u + fL(fu)\right. \\
&\quad \left. + L(fLu) + L\big(f^2 u\big) + f^2 Lu + f^3 u\right] + \frac{h^4}{24}[L^4 u + L^3(fu) + fL^3 u + fL^2(fu) + f^2 L^2 u + f^2 L(fu) \\
&\quad + f^3 Lu + f^4 u + L^2(fLu) + L^2(f^2 u) + fL(fLu) + fL(f^2 u) + L(fL^2 u) + L(fL(fu)) \\
&\quad + L(f^2 Lu) + L(f^3 u)] + \tilde{r},
\end{aligned} \tag{3.22}
$$

where all the operators and functions are evaluated at the point (t,x) and $\tilde{r}$ accumulates all the integrals present in (3.12)–(3.21) multiplied by h to the corresponding power. Taking into account that $u(t,x)$ satisfies the equation from (3.4), we get

$$
r(t,x) = \tilde{r}(t,x). \tag{3.23}
$$

If the terms in the one-step error $r(t,x)$ of the method (2.9), (2.10) (i.e., the terms in $\tilde{r}$) were bounded by $K(x)h^5$, $K(x) \in \mathbf{F}$ for all $t \leqslant T - h$, the relations (3.8)–(3.10) would imply that $\sum_{k=0}^{N-2} R_k \leqslant Ch^4$, where C is independent of h. But we see that the one-step error consists of integrals with integrands containing terms of the form $A(t,x) = L^n(q_1 L^l q_2)(t,x)$, where $q_1(t,x)$ and $q_2(t,x)$ are some functions from the class $\mathbf{F}$. The functions $A(t,x)$ belong to the class $\mathbf{F}$ for $t \in [0, T_*]$, where $T_* < T$ is a fixed (independent of h) time moment. Then $|r(t,x)| \leqslant K(x)h^5$, $K(x) \in \mathbf{F}$, $t \in [0, T_*]$, with $K(x)$ depending on T_*. However, the functions $A(t,x)$ do not belong to the class $\mathbf{F}$ for $t \in [0, T)$ due to the singularity in L (see (3.2)). Consequently, $r(t,x)$ can not be bounded by $K(x)h^5$, $K(x) \in \mathbf{F}$ for all $t < T$, and a more detailed analysis of the one-step error is required to prove the convergence theorem. In particular, we need to consider the structure of the functions $A(t,x)$ in detail. We always assume that L^0 is an identity operator. The following lemma is proved by induction.

Lemma 3.1. *Let $q_1(t,x)$ and $q_2(t,x)$ be sufficiently smooth functions belonging to the class* $\mathbf{F}$ *together with their partial derivatives of a sufficiently high order. Then for $0 \leqslant t < T$:*

$$
L^n(q_1 L^l q_2)(t,x) = g_0(t,x) + \sum_{j=1}^{m} \sum_{\alpha_j} g_{\alpha_j}(t,x) \psi^{\alpha_j}(t,x), \quad l, n = 0, 1, \ldots, \quad m = l + n, \tag{3.24}
$$

where α_j is a multi-index such that $\alpha_j = (i_1, \ldots, i_j)$ and each i_k is from $\{1, \ldots, d\}$, the summation in (3.24) *is over all possible values of α_j, g_0 and g_{α_j} are some functions from the class* $\mathbf{F}$, *and*

$$\psi^r = \frac{b^r - x^r}{T-t}, \quad r = 1,\ldots,d,$$
$$\psi^{\alpha_{j+1}} = \frac{b^{i_{j+1}} - x^{i_{j+1}}}{T-t}\psi^{\alpha_j} + \frac{\partial}{\partial x^r}\psi^{\alpha_j}, \quad \alpha_j = (i_1,\ldots,i_j), \ \alpha_{j+1} = (i_1,\ldots,i_j,i_{j+1}), \ j = 1,2,\ldots, \tag{3.25}$$

and for all α_j

$$L\psi^{\alpha_j} = 0. \tag{3.26}$$

Using specific properties of the functions ψ^{α_j}, the following theorem on one-step error is proved in Appendix A.

Theorem 3.2. *The one-step error of the method* (2.9), (2.10) *can be written in the form*

$$r(t,x) = \tilde{r}(t,x) = h^5 S(t,x) + E\rho(t,x;h), \tag{3.27}$$

where $S(t,x)$ *is a linear combination of the functions* $\psi^{\alpha_2}(t,x)$, $\psi^{\alpha_3}(t,x)$, $\psi^{\alpha_4}(t,x)$, $(T-t)\psi^{\alpha_4}(t,x)$, $h\psi^{\alpha_4}(t,x)$, $(T-t)\psi^{\alpha_5}(t,x)$, $h\psi^{\alpha_5}(t,x)$, $(T-t)^2\psi^{\alpha_6}(t,x)$, $(T-t)h\psi^{\alpha_6}(t,x)$, $h^2\psi^{\alpha_6}(t,x)$, *coefficients in this linear combination are independent of t, x, and h*; $\rho(t,x;h)$ *is such that*

$$\left(E[\rho(t,X_{0,a}(t);h)]^{2n}\right)^{1/2n} \leqslant \frac{Ch^5}{\sqrt{T-t-h}}, \quad t+h < T,$$

with a constant C independent of t and h.

We should emphasize that the most important part of this theorem consists in the equality $r(t,x) = \tilde{r}(t,x)$ which is due to Eqs. (3.10)–(3.22). Theorem 3.2 is a basis for the proof of Theorem 2.1 on the global error of the method (2.9), (2.10) (see Appendix B).

4. Implicit Runge–Kutta methods

From the point of view of possible applications, the most interesting case is when the function f is bounded from above, for example, when f is negative. In this case, the explicit Runge–Kutta method from Section 2 may cause some computational problems since, for instance, Y_{k+1} in (2.10) can become a large negative number while the exact $Y(t)$ is always positive. Apparently, this may occasionally lead to some instabilities and require a very small time step to achieve a reasonable accuracy. In such a situation an implicit method can behave better.

Let us formally apply the deterministic midpoint method to (2.3) provided $X(t)$ is a known function. As a result, we obtain

$$X(h/2) = a + \frac{h}{2}\frac{b-a}{T} + \sqrt{\frac{h}{2}}\sqrt{\frac{T-h/2}{T}}\xi_{1/2},$$
$$X(t_{k+1/2}) = X(t_{k-1/2}) + h\frac{b - X(t_{k-1/2})}{T - t_{k-1/2}} + \sqrt{h}\sqrt{\frac{T - t_{k+1/2}}{T - t_{k-1/2}}}\xi_{k+1/2}, \quad k = 1,\ldots,N-1, \tag{4.1}$$

$$Y_0 = 1,$$
$$Y_{k+1} = Y_k + hf(t_{k+1/2}, X(t_{k+1/2}))\frac{Y_k + Y_{k+1}}{2}, \quad k = 0,\ldots,N-1, \tag{4.2}$$

where $\xi_{k+1/2}$, $k=0,\dots,N-1$, are d-dimensional random vectors which components are mutually independent random variables with standard normal distribution $\mathcal{N}(0,1)$.

Resolving the implicitness in (4.2), we get

$$Y_{k+1} = Y_k \frac{1+\frac{h}{2}f(t_{k+1/2},X(t_{k+1/2}))}{1-\frac{h}{2}f(t_{k+1/2},X(t_{k+1/2}))}. \tag{4.3}$$

To ensure that the denominator in (4.3) does not vanish for all sufficiently small h, we should require that the function $f(t,x)$ is bounded from above, i.e., that $f(t,x)\leqslant c$ for all (t,x), c is a constant. In this case, for all sufficiently small h the denominator in (4.3) is positive. If $f(t,x)\leqslant 0$, then $-1\leqslant Y_k\leqslant 1$ for all k.

We prove the convergence theorem for the method (4.1), (4.2) under the same assumptions as in Section 3 (see p. 5). Note that in the case of $f(t,x)\leqslant 0$, the condition $EY_k^2\leqslant C$ is satisfied due to the uniform boundedness of the random variables Y_k.

Theorem 4.1. *The method* (4.1), (4.2) *applied to evaluation of the Wiener integral* (2.4) *is of second accuracy order, i.e.*,

$$|\mathcal{J}-EY_N| = |EY(T)-EY_N|\leqslant Kh^2, \tag{4.4}$$

where the constant K is independent of h.

The proof of this theorem is given in Appendix C.

If we formally apply the deterministic Gauss method of order four (see, e.g. [6, p. 71]) to (2.3), assuming that $X(t)$ is a known function, we obtain

$$X(\gamma h) = a+\gamma h\frac{b-a}{T}+\sqrt{\gamma h}\sqrt{\frac{T-\gamma h}{T}}\xi_\gamma, \tag{4.5}$$

$$X((1-\gamma)h) = X(\gamma h)+(1-2\gamma)h\frac{b-X(\gamma h)}{T-\gamma h}+\sqrt{(1-2\gamma)h}\sqrt{\frac{T-(1-\gamma)h}{T-\gamma h}}\xi_{1-\gamma},$$

$$X(t_k+\gamma h) = X(t_{k-1}+(1-\gamma)h)+2\gamma h\frac{b-X(t_{k-1}+(1-\gamma)h)}{T-t_k+\gamma h}+\sqrt{2\gamma h}\sqrt{\frac{T-t_k-\gamma h}{T-t_k+\gamma h}}\xi_{k+\gamma},$$

$$X(t_k+(1-\gamma)h) = X(t_k+\gamma h)+(1-2\gamma)h\frac{b-X(t_k+\gamma h)}{T-t_k-\gamma h}+\sqrt{(1-2\gamma)h}\sqrt{\frac{T-t_{k+1}+\gamma h}{T-t_k-\gamma h}}\xi_{k+1-\gamma},$$

$$k=1,\dots,N-1,$$

$$\begin{aligned}
&Y_0 = 1,\\
&k_1 = f(t_k+\gamma h,X(t_k+\gamma h))\left[Y_k+\frac{h}{4}k_1+\left(\frac{1}{4}-\frac{\sqrt{3}}{6}\right)hk_2\right],\\
&k_2 = f(t_k+(1-\gamma)h,X(t_k+(1-\gamma)h))\left[Y_k+\left(\frac{1}{4}+\frac{\sqrt{3}}{6}\right)hk_1+\frac{h}{4}k_2\right],\\
&Y_{k+1} = Y_k+\frac{h}{2}(k_1+k_2),\quad k=0,\dots,N-1,
\end{aligned} \tag{4.6}$$

where $\gamma = \frac{1}{2} - \frac{\sqrt{3}}{6}$ and $\xi_{k+\gamma}$, $\xi_{k+1-\gamma}$, $k = 0, \dots, N-1$, are d-dimensional random vectors which components are mutually independent random variables with standard normal distribution $\mathcal{N}(0,1)$.

Resolving (4.6) with respect to k_1 and k_2, we get

$$Y_{k+1} = Y_k \frac{1 + \frac{h}{4}(f_1 + f_2) + \frac{h^2}{12} f_1 f_2}{1 - \frac{h}{4}(f_1 + f_2) + \frac{h^2}{12} f_1 f_2}, \tag{4.7}$$

where $f_1 := f(t_k + \gamma h, X(t_k + \gamma h))$ and $f_2 := f(t_k + (1-\gamma)h, X(t_k + (1-\gamma)h))$.

The denominator in (4.7) does not vanish for all sufficiently small h for functions $f(t,x)$ being bounded from above. And if $f(t,x) \leqslant 0$, then $-1 \leqslant Y_k \leqslant 1$ for all k.

The intuition built on the previous analysis of the methods (2.9), (2.10) and (4.1), (4.2) tells us that the method (4.5), (4.6) should be of order four. But this assertion turned out to be wrong, the method is of order two only just as the method (4.1), (4.2). We have not found an implicit method for (2.4) that satisfies the condition $|Y_k| \leqslant 1$ for $f(t,x) \leqslant 0$ and has the fourth order of accuracy. In this search it was natural to restrict ourselves to standard fourth-order deterministic implicit methods for ordinary differential equations as a basis for potentially higher-order implicit methods for (2.4).

Analogously to Theorems 2.1 and 4.1, we prove the convergence theorem.

Theorem 4.2. *The method* (4.5), (4.6) *applied to evaluation of the Wiener integral* (2.4) *is of second order of accuracy, i.e.,*

$$|\mathcal{J} - EY_N| = |EY(T) - EY_N| \leqslant Kh^2, \tag{4.8}$$

where the constant K is independent of h.

Although the methods (4.1), (4.2) and (4.5), (4.6) are of the same order of convergence, in our numerical tests (see Section 6) the method (4.5), (4.6) gives more accurate results. Apparently, this is due to the fact that the constant K in (4.8) is, in general, less than its counterpart in (4.4). At the same time, the method (4.1), (4.2) requires one evaluation of f per step, while (4.5), (4.6) requires two evaluations of f per step.

5. Variance reduction

To evaluate $E\bar{Y}(T)$ in practice, we need to apply the Monte Carlo technique. As a result, in addition to the error of numerical integration considered in the previous sections, there is also the Monte Carlo error:

$$E\bar{Y}(T) = \frac{1}{M} \sum_{m=1}^{M} \bar{Y}^{(m)}(T) + R_{\mathrm{mc}}, \tag{5.1}$$

where M is the number of independent realizations $\bar{Y}^{(m)}(T)$ of $\bar{Y}(T)$. The Monte Carlo error R_{mc} has zero bias and its variance equals to

$$\mathrm{Var}(R_{\mathrm{mc}}) = \frac{\mathrm{Var}\,\bar{Y}(T)}{M} = \frac{\mathrm{Var}\,Y(T)}{M} + \mathrm{O}\left(\frac{h}{M}\right), \tag{5.2}$$

i.e., the simulated $\hat{Y} := (1/M)\sum_{m=1}^{M} \bar{Y}^{(m)}(T)$ belongs to the confidence interval:

$$\hat{Y} \in \left(E\bar{Y}(T) - c\sqrt{\mathrm{Var}(R_{\mathrm{mc}})}, E\bar{Y}(T) + c\sqrt{\mathrm{Var}(R_{\mathrm{mc}})}\right) \tag{5.3}$$

with the fiducial probability, for example, 0.997 for $c = 3$ and 0.95 for $c = 2$.

Thus, if the variance $\mathrm{Var}\, Y(T)$ is big, a large number of trajectories M has to be simulated in order to reach a satisfactory accuracy. To reduce the Monte Carlo error, a variance reduction technique can be used. The basic idea of variance reduction techniques (see [5,10–12]) is to substitute $Y(T)$ by another random variable which has the same expectation as $Y(T)$ but a smaller variance. Two variance reduction methods are known: the method of importance sampling [5,10,11] and the method of control variates [11,12]. A combining method is given in [11]. The method of important sampling is based on Girsanov's transformation. In our case its application changes the linear system (2.1) for X to a system with, in general, a nonlinear drift. As a result, we lose the advantage of simulating $X(t)$ exactly and of approximating the conditional Wiener integral by higher-order numerical integrators from Sections 2 and 4. This shortcoming does not arise in the case of the method of control variates. That is why, we restrict ourselves here to this method only.

In connection with the evaluation of the Wiener integral (1.1), (1.2) consider the following system of Ito SDEs (cf. (2.1)–(2.3)):

$$\mathrm{d}X = \frac{b - X}{T - t}\,\mathrm{d}t + \mathrm{d}w(t), \quad X(s) = x, \tag{5.4}$$

$$\mathrm{d}Y = f(t, X(t))Y\,\mathrm{d}t, \quad Y(s) = y, \tag{5.5}$$

$$\mathrm{d}Z = G^\top(t, X)Y\,\mathrm{d}w(t), \quad Z(s) = z. \tag{5.6}$$

Here Z is a scalar and $G(t,x)$ is a column-vector of dimension d with good analytical properties, the other notation is the same as before.

It is clear that

$$u(s,x) = EY_{s,x,1}(T) = E[Y_{s,x,1}(T) + Z_{s,x,1,0}(T)].$$

As it is known [11]

$$\mathrm{Var}[Y_{s,x,1}(T) + Z_{s,x,1,0}(T)] = E\int_s^T Y_{s,x,1}^2(t)\sum_{i=1}^d \left(\frac{\partial u}{\partial x^i} + G^i\right)^2 \mathrm{d}t, \tag{5.7}$$

where $u(t,x)$ is the solution of (3.4). Then by choosing $G(t,x)$ as

$$G^i = -\frac{\partial u}{\partial x^i}, \quad j = 1, \ldots, d, \tag{5.8}$$

we obtain that the variance of $Y_{s,x,1}(T) + Z_{s,x,1,0}(T)$ is equal to zero.

Applying a numerical method to (5.4)–(5.6), we get the approximate $\bar{Y}_{s,x,1}(T)$ and $\bar{Z}_{s,x,1,0}(T)$. The variance $\mathrm{Var}[\bar{Y}_{s,x,1}(T) + \bar{Z}_{s,x,1,0}(T)]$ is close to $\mathrm{Var}[Y_{s,x,1}(T) + Z_{s,x,1,0}(T)]$, i.e., it is small in the case of G from (5.8), and, consequently, a smaller number of independent realizations M is needed to have a satisfactory accuracy.

Of course, in practice the solution $u(t,x)$ is not known. However, an approximate solution $\tilde{u}$ to the problem (3.4) can be known. In this case we can take $G(t,x)$ in the form of (5.8) with $\tilde{u}$ instead of u and we may expect a variance reduction. This is demonstrated in numerical examples (see the next section).

6. Numerical tests

1. We take $f(t,x)$ in the form

$$f(t,x) = (A(t)x, x) + (a_1(t), x) + a_0(t), \tag{6.1}$$

where $A(t)$ is a $d \times d$ symmetric matrix, $a_1(t)$ is a d-dimensional vector, and $a_0(t)$ is a scalar function.

Let $u(t,x)$ be the solution of (3.4) with f from (6.1). Introduce the function $P(t,x)$:

$$u(t,x) = \exp(P(t,x)). \tag{6.2}$$

This function satisfies the problem

$$LP + (A(t)x,x) + (a_1(t),x) + a_0(t) + \frac{1}{2}\sum_{i=1}^{d}\left(\frac{\partial P}{\partial x^i}\right)^2 = 0, \quad x \in \mathbf{R}^d, \ t < T, \ P(T,x) = 0. \tag{6.3}$$

We look for a solution of (6.3) in the form

$$P(t,x) = \frac{1}{2}(P(t)x,x) + (p(t),x) + q(t), \tag{6.4}$$

where $P(t)$ is a $d \times d$ symmetric matrix, $p(t)$ is a d-dimensional vector, and $q(t)$ is a scalar function.

Substituting (6.4) in (6.3) and collecting terms $(\cdot x,x)$, $(\cdot,x)$ and terms independent of x separately, we arrive at the system for $P(t)$, $p(t)$, and $q(t)$:

$$P'(t) - \frac{2}{T-t}P + 2A(t) + P^2(t) = 0, \quad P(T) = 0, \tag{6.5}$$

$$p'(t) - \frac{1}{T-t}p + \frac{1}{T-t}P(t)b + P(t)p + a_1(t) = 0, \quad p(T) = 0, \tag{6.6}$$

$$q'(t) + \frac{1}{T-t}(p(t),b) + \frac{1}{2}\operatorname{tr}P(t) + \frac{1}{2}(p(t),p(t)) + a_0(t) = 0, \quad q(T) = 0. \tag{6.7}$$

Note that if $a_1(t) \equiv 0$ and $b = 0$, then $p(t) \equiv 0$. And if in addition $a_0(t) \equiv 0$, then

$$q(t) = \frac{1}{2}\int_t^T \operatorname{tr}P(s)\,\mathrm{d}s.$$

The solution of (6.5) can be expanded in (positive) powers of $T-t$. If $A(t)$ is a constant matrix A, then this formal expansion starts with the terms

$$P(t) = \frac{2}{3}A \cdot (T-t) + \frac{4}{45}A^2 \cdot (T-t)^3 + \cdots.$$

For test purposes, it is convenient to have an exact solution of (6.5)–(6.7) in a closed analytical form. To this end, we choose a variable matrix $A(t)$ such that

$$A(t) = A - \frac{2}{9}A^2 \cdot (T-t)^2, \tag{6.8}$$

where A is a constant symmetric matrix. Then the exact solution of the system (6.5)–(6.7) with $b = 0$, $a_0(t) \equiv 0$, and $a_1(t) \equiv 0$ has the form

$$P(t) = \frac{2}{3}(T-t)A, \quad p(t) = 0, \quad q(t) = \frac{(T-t)^2}{6}\operatorname{tr}A. \tag{6.9}$$

Consequently, the solution of (6.4) is

$$P(t,x) = \frac{T-t}{3}(Ax,x) + \frac{(T-t)^2}{6}\operatorname{tr}A. \tag{6.10}$$

Then the conditional Wiener integral (1.1), (1.2) for f from (6.1) with $a_0 = 0$, $a_1 = 0$, $A(t)$ from (6.8) and for $a = b = 0$ is equal to

$$\mathscr{J} = u(0,0) = \exp\left(\frac{T^2}{6}\operatorname{tr}A\right).$$

In our experiments we take the dimension $d = 4$ and the following matrix A:

$$A = \begin{bmatrix} -1 & -0.5 & 0 & 0 \\ -0.5 & 2 & -0.5 & 0 \\ 0 & -0.5 & -2 & -0.5 \\ 0 & 0 & -0.5 & 1 \end{bmatrix}, \tag{6.11}$$

for which $\operatorname{tr}A = 0$.

In Table 1, we give results of simulation of the conditional Wiener integral (1.1), (1.2) for f from (6.1) with $a_0 = 0$, $a_1 = 0$, $A(t)$ from (6.8), (6.11) and for $a = b = 0$, $T = 1$ by the explicit Runge–Kutta method (2.9), (2.10) and the implicit Runge–Kutta methods (4.1), (4.3), and (4.5), (4.7). As it was mentioned in Section 5, we have two types of errors in numerical simulations here: the error of a method used and the Monte Carlo error. The results in the table are approximations of $E\bar{Y}(1)$ calculated as in (5.1)–(5.3) with $c = 2$. Note that the "$\pm$" reflects the Monte Carlo error only and it does not reflect the error of a method. The results obtained are in agreement with the proved convergence theorems (see also Table 2). Recall that the implicit methods (4.1), (4.2) and (4.5), (4.6) are both of order two. In our tests the method (4.5), (4.6) performs better. Apparently, this is due to the fact that the constant K in (4.8) is, in general, less than its counterpart in (4.4).

We also note that for the considered test problem we do not have any numerical instabilities and the explicit method is computationally effective. As has been discussed at the beginning of Section 4, implicit methods should be used in practice when explicit methods are affected by instabilities. A further investigation and tests are required in this direction.

2. To reduce the Monte Carlo error in simulation of the above test problem, we can use the variance reduction technique from Section 5. For f from (6.1) with $a_0 = 0$, $a_1 = 0$, $A(t)$ from (6.8), (6.11) and for $b = 0$, the solution $u(t,x)$ of (3.4) has the form (6.2), (6.10). Therefore, in this case the vector function G defined in (5.8) is equal to

$$G^i(t,x) = -\frac{2}{3}(T-t)\exp(P(t,x))\sum_{j=1}^{d} A^{ij}x^j, \quad i = 1,\ldots,d, \tag{6.12}$$

where $P(t,x)$ is from (6.10) and A is from (6.11).

Applying the Euler method to Eq. (5.6), we get

$$\begin{aligned} &Z_0 = 0, \\ &Z_{k+1} = Z_k + G^\top(t_k,X)Y_k\Delta w_k, \quad k = 1,\ldots,N-1. \end{aligned} \tag{6.13}$$

Table 1

The results of simulation of the conditional Wiener integral (1.1), (1.2) for f from (6.1) with $a_0 = 0$, $a_1 = 0$, $A(t)$ from (6.8), (6.11) and for $a = b = 0$, $T = 1$ by the explicit Runge–Kutta method (2.9), (2.10) and the implicit Runge–Kutta methods (4.1), (4.3) and (4.5), (4.7). The exact solution is 1

h	M	(2.9), (2.10)	(4.1), (4.3)	(4.5), (4.7)
0.2	10^6	0.9994 ± 0.0013	1.0176 ± 0.0044	1.0040 ± 0.0013
0.1	10^8	1.00002 ± 0.00013	1.00361 ± 0.00015	1.00093 ± 0.00013
0.05	10^8	0.99996 ± 0.00013	1.00089 ± 0.00013	1.00019 ± 0.00013

Table 2
The results of simulation of the conditional Wiener integral (1.1), (1.2) for f from (6.1) with $a_0 = 0$, $a_1 = 0$, $A(t)$ from (6.8), (6.11) and for $a = b = 0$, $T = 1$ by the explicit Runge–Kutta method (2.9), (2.10) and the implicit Runge–Kutta methods (4.1), (4.3) and (4.5), (4.7) using the variance reduction technique. The exact solution is 1

h	M	(2.9), (2.10)	(4.1), (4.3)	(4.5), (4.7)
0.1	10^7	0.99977 ± 0.00024	1.00396 ± 0.00050	1.00103 ± 0.00023
0.05	10^7	0.99992 ± 0.00017	1.00098 ± 0.00017	1.00023 ± 0.00016
0.05	10^8	0.99999 ± 0.00005	1.00088 ± 0.00005	1.00027 ± 0.00005
0.01	10^7	1.00003 ± 0.00007	1.00001 ± 0.00007	1.00003 ± 0.00007

If we approximate (5.4), (5.5) using the explicit fourth-order Runge–Kutta method (2.9), (2.10), then Y_k in (6.13) is from (2.10) and the Wiener increment is

$$\Delta w_k := w(t_{k+1}) - w(t_k) = \frac{h^{1/2}}{\sqrt{2}}\left(\xi_{k+1/2} + \xi_{k+1}\right),$$

where $\xi_{k+1/2}$ and ξ_{k+1} are the same as in (2.9), (2.10).

It is clear that $EZ_{k+1} = 0$. This implies that the method (2.9), (2.10), (6.13) applying to (5.4)–(5.6) to approximate the Wiener integral $\mathscr{J} = EY(T)$ is of order four, i.e., the above realization of the variance reduction technique does not affect the accuracy of the numerical method. The variance $\mathrm{Var}Y(T)$ is approximated with accuracy $O(h)$. Consequently, for a fixed number of realizations M the Monte Carlo error in simulations using the variance reduction technique is $\sim 1/\sqrt{h}$ times less than in simulations without variance reduction. In other words, in the case of variance reduction the Monte Carlo error is proportional to $\sqrt{h}/\sqrt{M}$. This is illustrated in Table 2. In particular, we see for $h = 0.05$ that to produce results of the same quality we need $M = 10^8$ independent trajectories without variance reduction and $M = 10^7$ independent realizations in the variance reduction case (compare Tables 1 and 2).

Remark 6.1. Recall that the implicit methods (4.1), (4.3) and (4.5), (4.7) do not contain simulation of $X(t_{k+1})$, and the random variables involved in these methods are not enough to evaluate the Wiener increments Δw_k on the intervals $[t_k, t_{k+1}]$. At the same time, these Wiener increments are needed to realize (6.13). Thus, to use the variance reduction technique in connection with the implicit methods (4.1), (4.3) and (4.5), (4.7), we introduce additional random variables and simulation of $X(t_{k+1})$ in the corresponding algorithms (see (C.1) in the case of the method (4.1), (4.3)).

3. Now we illustrate the assertion made at the end of Section 5. To this end we take the function $f(t,x)$ in the form (6.1) with the constant matrix $A(t) \equiv A$ from (6.11) and $a_0 = 0$, $a_1 = 0$. We also put $b = 0$. In this case, we do not know the exact solution $u(t,x)$ of (3.4). But for the variance reduction we can use an approximation $\tilde{u}(t,x)$ of the solution based on the formal expansion (6.9):

$$\tilde{u}(t,x) = \exp\left(\frac{1}{2}\left(\tilde{P}(t)x, x\right)\right), \tag{6.14}$$

where

$$\tilde{P}(t) = \frac{2}{3}A \cdot (T - t).$$

Deriving (6.14), we take into account that $\mathrm{tr}\,\tilde{P}(t) = 0$ because of the specific choice of the matrix A which is from (6.11).

Then we take the function G in (6.13) of the form

$$G^i(t,x) = -\frac{\partial \tilde{u}}{\partial x^i}, \quad j = 1,\ldots,d.$$

Putting $a = 0$ and $T = 1$, we evaluate the corresponding conditional Wiener integral (1.1), (1.2) by the fourth-order explicit Runge–Kutta method (2.9), (2.10) with time step $h = 0.01$ and we simulate $M = 10^5$ independent realizations. Without variance reduction, we get: $\mathscr{J} \doteq 1.1536 \pm 0.0093$, while applying the variance reduction technique (i.e., using the method (2.9), (2.10), (6.13) for (5.4)–(5.6) we obtain $\mathscr{J} \doteq 1.1482 \pm 0.0018$. We see that the Monte Carlo error is five times less when we use the variance reduction technique.

7. Summary

In this paper, we use numerical integration of SDEs together with the Monte Carlo technique to evaluate conditional Wiener integrals of exponential-type functionals. Other known methods reduce the infinite-dimensional integral to a high dimensional Riemann integral and the accuracy is reached by increasing its dimension. The high dimensionality of these Riemann integrals makes calculation of the Wiener integrals extremely difficult. An efficiency of the probabilistic approach considered here is due to the fact that the problem of calculating the infinite-dimensional Wiener integral is reduced to the Cauchy problem for a system of SDEs, which can naturally be regarded as one-dimensional. Moreover, due to the specific form of this system, we propose efficient fourth-order Runge–Kutta methods. In [5,10,14], the probabilistic approach was used for computation of Wiener integrals with respect to the usual Wiener measure. Here, in the case of conditional Wiener integrals, we deal with a more complicated system than in [5,10,14]. The solution of this system contains a Markov representation of the Brownian bridge. The system is singular and this circumstance stipulates a certain complexity of theoretical proofs although the constructed algorithms are simple and effective. The effectiveness of these algorithms allows us to evaluate conditional Wiener integrals for a large dimension of paths. There are two types of errors in our approach: the error of numerical integration and the Monte Carlo error. Both errors are analyzed in the paper: convergence theorems are proved for the methods proposed and such a variance reduction technique as the method of control variates is considered to reduce the Monte Carlo error. Finally, the algorithms are tested on a model problem.

Appendix A. Proof of the theorem on one-step error of the explicit method

The next corollary follows immediately from Lemma 3.1.

Corollary A.1. *We have for* $t \leqslant \theta < T$:

$$E\psi^{\alpha_j}(\theta, X_{t,x}(\theta)) = \psi^{\alpha_j}(t,x),$$

i.e., $\psi^{\alpha_j}(\theta, X_{0,a}(\theta))$ *is a martingale.*

Let us now consider some other properties of the functions $\psi^{\alpha_j}(t,x)$. We note that $\psi^{\alpha_j}(t,x)$ does not depend on the order of $i_1,\ldots,i_j$ in α_j (to see this it is enough to show that $\psi^{(i_1,\ldots,i_{j-2},l,r)}(t,x) = \psi^{(i_1,\ldots,i_{j-2},r,l)}(t,x)$). Introduce the function $\chi(r,\alpha_j)$, $\alpha_j = (i_1,\ldots,i_j)$, which is equal to the number of appearances of r in the set $\{i_1,\ldots,i_j\}$. In what follows we will sometimes denote by the same α_j different multi-indices having the length j, and therefore functions ψ^{α_j} may differ although they have the same notation. The next two lemmas are given without proofs.

Lemma A.2. *We have for $t < T$:*

$$\frac{\partial}{\partial x^r}\psi^{\alpha_j}(t,x) = -\frac{\chi(r,\alpha_j)}{T-t}\psi^{\alpha_{j-1}}(t,x), \quad j = 2,3,\ldots. \tag{A.1}$$

Corollary A.3. *We have for $t < T$:*

$$\psi^{\alpha_{j+1}}(t,x) = \frac{b^r - x^r}{T-t}\psi^{\alpha_j}(t,x) - \frac{\chi(r,\alpha_j)}{T-t}\psi^{\alpha_{j-1}}(t,x), \quad j > 1, \tag{A.2}$$

$$(b^r - x^r)\psi^{\alpha_j}(t,x) = (T-t)\psi^{\alpha_{j+1}}(t,x) + \chi(r,\alpha_j)\psi^{\alpha_{j-1}}(t,x), \quad j > 1, \tag{A.3}$$

$$\begin{aligned}\left(b^l - x^l\right)(b^r - x^r)\psi^{\alpha_j}(t,x) = {} & (T-t)^2\psi^{\alpha_{j+2}}(t,x) + \chi(l,\alpha_{j+1})(T-t)\psi^{\alpha_j}(t,x) + \chi(r,\alpha_j)(T-t)\psi^{\alpha_j}(t,x) \\ & + \chi(r,\alpha_j)\chi(l,\alpha_{j-1})\psi^{\alpha_{j-2}}(t,x), \quad j > 2.\end{aligned} \tag{A.4}$$

Note that ψ^{α_j} in (A.4) are, in general, different. We do not distinguish them because in the following analysis we will concern with the length of multi-indices only.

Lemma A.4. *We have for $\theta < T$:*

$$\begin{aligned}&\left(E\left[\left(b^{r_1} - X^{r_1}_{0,a}(\theta)\right) \times \cdots \times \left(b^{r_l} - X^{r_l}_{0,a}(\theta)\right) \times \psi^{\alpha_j}(\theta, X_{0,a}(\theta))\right]^{2n}\right)^{1/(2n)} \leqslant C \cdot (T-\theta)^{(l-j)/2}, \\ &\quad j = 1,2,\ldots, \quad l = 0,1,\ldots, \quad n = 1,2,\ldots,\end{aligned} \tag{A.5}$$

where the constant $C > 0$ is independent of θ (of course, it depends on n).

Now we are in position to make a detailed analysis of the remainder $r(t,x) = \tilde{r}(t,x)$ from (3.23). Let us recall that the one-step error consists of integrals with integrands containing terms of the form $A(t,x) = L^n(q_1 L^l q_2)(t,x)$, where $q_1(t,x)$ and $q_2(t,x)$ are some functions from the class **F**. Since $L^5 u = -L^4(fu)$ (cf. (3.4)), the number $m = l + n$ for all the terms $A(t,x)$ participating in the remainder is less than or equal to 4. Using Lemma 3.1, we can represent the term $A = L^n(q_1 L^l q_2)(\theta, X_{t,x}(\theta))$ as

$$A(\theta, X_{t,x}(\theta)) = g_0(\theta, X_{t,x}(\theta)) + \sum_{j=1}^{m}\sum_{\alpha_j} g_{\alpha_j}(\theta, X_{t,x}(\theta))\psi^{\alpha_j}(\theta, X_{t,x}(\theta)). \tag{A.6}$$

By Lemma A.4 (see (A.5) with $l = 0$), we get that

$$\left(E[A(\theta, X_{0,a}(\theta))]^{2n}\right)^{1/(2n)} \leqslant \frac{C}{(T-\theta)^{m/2}},$$

where the constant C is independent of θ.

Consequently (recall that $m \leqslant 4$), we obtain the following estimate:

$$\left(E[r(t, X_{0,a}(t); h)]^{2n}\right)^{1/2n} \leqslant \frac{Ch^5}{(T-t-h)^2}. \tag{A.7}$$

Using this rough estimate, we can show that the method (2.9), (2.10) is at least of order three. To prove the fourth-order of its convergence, a more sophisticated analysis based on extraction of singularity is needed.

To clarify the matter, we consider, for example, the following integral from the remainder $r(t,x)$:

$$\int_t^{t+h} \frac{(t+h-\theta)^4}{24} EL^5 u(\theta, X_{t,x}(\theta))\,\mathrm{d}\theta = -\int_t^{t+h} \frac{(t+h-\theta)^4}{24} EL^4(fu)(\theta, X_{t,x}(\theta))\,\mathrm{d}\theta$$
$$= -\int_t^{t+h} \frac{(t+h-\theta)^4}{24} \Bigg[Eg_0(\theta, X_{t,x}(\theta)) + \sum_{j=1}^{4} \sum_{\alpha_j} Eg_{\alpha_j}(\theta, X_{t,x}(\theta)) \psi^{\alpha_j}(\theta, X_{t,x}(\theta)) \Bigg] \mathrm{d}\theta. \tag{A.8}$$

We will demonstrate the extraction of singularity analyzing the term with the highest singularity $g_{\alpha_4}(\theta, X_{t,x}(\theta))\psi^{\alpha_4}(\theta, X_{t,x}(\theta))$. The singularity of $\psi^{\alpha_4}(\theta, X_{t,x}(\theta))$ is of order two, i.e.,

$$\left(E[\psi^{\alpha_4}(\theta, X_{0,a}(\theta))]^2\right)^{1/2} \leqslant \frac{C}{(T-\theta)^2}.$$

At the same time, $\psi^{\alpha_4}(\theta, X_{0,a}(\theta))$ is a martingale (see Corollary A.1) and $E\psi^{\alpha_4}(\theta, X_{0,a}(\theta)) = \psi^{\alpha_4}(0,a)$, which is a constant independent of θ. To exploit this property of ψ^{α_4} in further analysis, we expand $g(\theta, X_{t,x}(\theta)) := g_{\alpha_4}(\theta, X_{t,x}(\theta))$ at (T,b):

$$\begin{aligned} g(\theta, X_{t,x}(\theta)) = {} & g(T,b) + \frac{\partial g}{\partial t}(T,b)(\theta - T) + \sum_{r=1}^{d} \frac{\partial g}{\partial x^r}(T,b)\left(X^r_{t,x}(\theta) - b^r\right) \\ & + \frac{1}{2} \sum_{r_1,r_2=1}^{d} \frac{\partial^2 g}{\partial x^{r_1} \partial x^{r_2}}(T,b)\left(X^{r_1}_{t,x}(\theta) - b^{r_1}\right)\left(X^{r_2}_{t,x}(\theta) - b^{r_2}\right) \\ & + \frac{1}{2} \frac{\partial^2 g}{\partial t^2}(\vartheta, X_{t,x}(\theta))(\theta - T)^2 + \sum_{r=1}^{d} \frac{\partial^2 g}{\partial t \partial x^r}(T, \eta_1)(\theta - T)\left(X^r_{t,x}(\theta) - b^r\right) \\ & + \frac{1}{6} \sum_{r_1,r_2,r_3=1}^{d} \frac{\partial^3 g}{\partial x^{r_1} \partial x^{r_2} \partial x^{r_3}}(T, \eta_2)\left(X^{r_1}_{t,x}(\theta) - b^{r_1}\right)\left(X^{r_2}_{t,x}(\theta) - b^{r_2}\right)\left(X^{r_3}_{t,x}(\theta) - b^{r_3}\right), \end{aligned} \tag{A.9}$$

where ϑ is a time between θ and T and η_1 and η_2 are points between $X_{t,x}(\theta)$ and b.

Then, using Corollaries A.1 and A.3, we obtain

$$\begin{aligned} & \int_t^{t+h} \frac{(t+h-\theta)^4}{24} Eg_{\alpha_4}(\theta, X_{t,x}(\theta)) \psi^{\alpha_4}(\theta, X_{t,x}(\theta))\,\mathrm{d}\theta \\ & = \int_t^{t+h} \Bigg[\frac{(t+h-\theta)^4}{24} g(T,b)\psi^{\alpha_4}(t,x) + \frac{\partial g}{\partial t}(T,b)(\theta - T)\psi^{\alpha_4}(t,x) \\ & - \sum_{r=1}^{d} \frac{\partial g}{\partial x^r}(T,b)\{(T-\theta)\psi^{\alpha_5}(t,x) + \chi(r,\alpha_4)\psi^{\alpha_3}(t,x)\} \\ & + \frac{1}{2} \sum_{r_1,r_2=1}^{d} \frac{\partial^2 g}{\partial x^{r_1}\partial x^{r_2}}(T,b)\Big\{(T-\theta)^2\psi^{\alpha_6}(t,x) + \chi(r_1,\alpha_5)(T-\theta)\psi^{\alpha_4}(t,x) \\ & + \chi(r_2,\alpha_4)(T-\theta)\psi^{\alpha_4}(t,x) + \chi(r_2,\alpha_4)\chi(r_1,\alpha_3)\psi^{\alpha_2}(t,x)\Big\}\Bigg]\mathrm{d}\theta \\ & + E\rho'(t,x;h) = h^5 S'(t,x) + E\rho'(t,x;h), \end{aligned} \tag{A.10}$$

where $S'(t,x)$ is a linear combination of the functions $\psi^{\alpha_2}(t,x)$, $\psi^{\alpha_3}(t,x)$, $\psi^{\alpha_4}(t,x)$, $(T-t)\psi^{\alpha_4}(t,x)$, $h\psi^{\alpha_4}(t,x)$, $(T-t)\psi^{\alpha_5}(t,x)$, $h\psi^{\alpha_5}(t,x)$, $(T-t)^2\psi^{\alpha_6}(t,x)$, $(T-t)h\psi^{\alpha_6}(t,x)$, $h^2\psi^{\alpha_6}(t,x)$, coefficients in this linear combination are independent of t, x, and h

$$E\rho'(t,x;h) = E\int_t^{t+h} \frac{(t+h-\theta)^4}{24}\psi^{\alpha_4}(\theta,X_{t,x}(\theta))\Bigg[\frac{1}{2}\frac{\partial^2 g}{\partial t^2}(\vartheta,X_{t,x}(\theta))(\theta-T)^2 + \sum_{r=1}^{d}\frac{\partial^2 g}{\partial t\partial x^r}(T,\eta_1)(\theta-T)\left(X_{t,x}^r(\theta)-b^r\right) + \frac{1}{6}\sum_{r_1,r_2,r_3=1}^{d}\frac{\partial^3 g}{\partial x^{r_1}\partial x^{r_2}\partial x^{r_3}}(T,\eta_2) \times (X_{t,x}^{r_1}(\theta)-b^{r_1})(X_{t,x}^{r_2}(\theta)-b^{r_2})(X_{t,x}^{r_3}(\theta)-b^{r_3})\Bigg]\mathrm{d}\theta.$$

Using the Cauchy–Bunyakovskii inequality and Lemma A.4, we obtain that

$$\left(E\left[\rho'(t,X_{0,a}(t);h)\right]^{2n}\right)^{1/2n} \leqslant \frac{Ch^5}{\sqrt{T-t-h}}. \tag{A.11}$$

Thus, we extract the singularity by presenting the integral (A.10) as the sum of the singular part $S'(t,x)$ and the remainder. The singular part contains singularities of order from one to four, while the remainder has non-singular terms and terms with singularity of order 1/2. By further expansion of $g(\theta,X_{t,x}(\theta))$ (cf. (A.9)), we could also include the singularity of order 1/2 in the singular part making the remainder singular-free. But for our purposes (i.e., for proving Theorem 2.1) the obtained expression (A.10) is sufficient.

We similarly analyze the other terms in the integral (A.8). Note that ψ^{α_1} has singularity of order 1/2 and we include it in the remainder. So, S' does not contain any ψ^{α_1}. Analogously, we consider all the integrals of the remainder $r(t,x)$. In the case of an integral from $t+h/2$ to $t+h$, we first take the conditional expectation of the term like A with respect to $\mathscr{F}_{t+h/2}$ in a similar way as above and then we repeat the procedure once again taking the expectation of the product of the conditional expectation and $f_{1/2}$ (or $f_{1/2}^2$). As a result, we obtain an expression like the right-hand side of (A.10). Thus, Theorem 3.2 on the one-step error is proved.

Appendix B. Proof of the convergence theorem for the explicit method

In this appendix we give a proof of the convergence Theorem 2.1 for the method (2.9), (2.10).

According to (3.9) and (3.10), we have

$$R_k = |EY_k r(t_k,X_k)|$$

with $r(t,x)$ from (3.27).

Using the rough estimate (A.7) and the Cauchy–Bunyakovskii inequality, we get straightforward that (recall that we assume uniform boundedness of the moments EY_k^2):

$$R_k \leqslant \frac{Kh^5}{(T-t_{k+1})^2}, \quad k=0,1,\ldots,N-2. \tag{B.1}$$

But to prove (2.11), a more accurate estimate of R_k is needed. We obtain such an estimate using Theorem 3.2.

Lemma B.1. *We have*

$$R_k \leqslant \frac{Kh^5}{\sqrt{T-t_{k+1}}}, \quad k=0,\ldots,N-2, \tag{B.2}$$

where K is independent of k and h.

Proof. Note that we use the same letters C and K for different constants which are independent of h and k.

By Theorem 3.2 and the Cauchy–Bunyakovskii inequality, we obtain

$$R_k = |EY_k r(t_k,X_k)| = \left|EY_k\left[h^5 S_k + \rho(t_k,X_k;h)\right]\right| \leqslant h^5|EY_kS_k| + \frac{Ch^5}{\sqrt{T-t_{k+1}}}, \tag{B.3}$$

where $S_k := S(t_k,X_k)$, $S(t,x)$ and $\rho(t,x;h)$ are from (3.27). Recall that S_k has singularity of order two, more precisely

$$\left(E[S_k]^2\right)^{1/2} \leqslant \frac{C}{(T-t_k)^2}.$$

Let $F_i := F(t_{i-1},t_{i-1/2},t_i,X_{i-1},X_{i-1/2},X_i)$ be the function defined by the method (2.10) (see also (3.11)), i.e., the last line of (2.10) for $k=i-1$ can be written as

$$Y_i = Y_{i-1} + hF_i.$$

Introduce $S_{k,i} := E(S_k|\mathscr{F}_{t_i})$, $i<k$. Due to Theorem 3.2 and Corollary A.1, $S_{k,i}$ is a linear combination of $\psi^{\alpha_2}(t_i,X_i)$, $\psi^{\alpha_3}(t_i,X_i)$, $\psi^{\alpha_4}(t_i,X_i)$, $(T-t_k)\psi^{\alpha_4}(t_i,X_i)$, $h\psi^{\alpha_4}(t_i,X_i)$, $(T-t_k)\psi^{\alpha_5}(t_i,X_i)$, $h\psi^{\alpha_5}(t_i,X_i)$, $(T-t_k)^2\psi^{\alpha_6}(t_i,X_i)$, $(T-t_k)h\psi^{\alpha_6}(t_i,X_i)$, $h^2\psi^{\alpha_6}(t_i,X_i)$, coefficients in this linear combination are independent of t_k, t_i, x, and h. Consequently (cf. (A.5) with $l=0$)

$$\left(E[S_{k,i}]^2\right)^{1/2} \leqslant \frac{C}{(T-t_i)^2}.$$

We see that though $S_{k,i}$ has the same order of singularity as S_k, the singularity is shifted. Roughly speaking, $S_{k,i}$ is less singular than $S_{k,i+1}$. Also note that $E\left(S_{k,i}|\mathscr{F}_{t_{i-1}}\right) = S_{k,i-1}$ since ψ^{α_j} are martingales (see Corollary A.1).

We fix $k>0$ and consider $B_i := |EY_iS_{k,i}|$, $i=k,k-1,\ldots,1$:

$$B_i = |EY_iS_{k,i}| = |EY_{i-1}[1+hF_i]S_{k,i}| \leqslant |EY_{i-1}S_{k,i-1}| + h|EY_{i-1}F_iS_{k,i}|. \tag{B.4}$$

We expand the terms, which form F_i, at (T,b) up to terms of first order, i.e., we write F_i as a constant plus a remainder consisting of products of $f(t,x)$, some its derivatives and $X_j^r - b^r$ or $t_j - T$ with $j=i$, $i-1/2$, or $i-1$. Then, using the Cauchy–Bunyakovskii inequality and Lemma A.4, we get:

$$|EY_{i-1}F_iS_{k,i}| \leqslant K|EY_{i-1}S_{k,i}| + \frac{C}{(T-t_i)^{3/2}} = K|EY_{i-1}S_{k,i-1}| + \frac{C}{(T-t_i)^{3/2}}.$$

Hence, due to (B.4), we obtain

$$B_i \leqslant B_{i-1} + KhB_{i-1} + \frac{Ch}{(T-t_i)^{3/2}}, \quad i=k, \quad k-1,\ldots,1, \tag{B.5}$$

where B_0 is evidently a constant.

Therefore,

$$\begin{aligned} B_k &\leqslant (1+Kh)^k B_0 + (1+Kh)^{k-1}\frac{Ch}{(T-t_1)^{3/2}} + (1+Kh)^{k-2}\frac{Ch}{(T-t_2)^{3/2}} + \cdots \\ &+ \frac{Ch}{(T-t_k)^{3/2}} \leqslant B_0\,\mathrm{e}^{KT} + C\mathrm{e}^{KT}h\sum_{i=1}^{k}\frac{1}{(T-t_i)^{3/2}} \leqslant \frac{C}{\sqrt{T-t_{k+1}}}, \end{aligned} \tag{B.6}$$

which together with (B.3) implies (B.2). □

Remark B.2. It is possible to prove that $R_k \leqslant Kh^5,\ k = 0, \ldots, N-2$, with the constant K independent of h and k. But we restrict ourselves here to the estimate (B.2) since it is sufficient for proving Theorem 2.1 and is obtained by less efforts than it would be needed for a more accurate estimate.

Since the operator L is not defined at $t = T$, we need a separate analysis of the error on the last step R_{N-1}. It is true that $R_{N-1} \leqslant Kh^5$ but for our purposes it is enough that

$$R_{N-1} \leqslant Kh^4, \tag{B.7}$$

where K is independent of h. We omit the proof of (B.7) here.

Now we are in position to prove the convergence theorem.

Proof of Theorem 2.1. Lemma B.1, estimate (B.7) and the relations (3.8), (3.9) imply

$$\left| EY_{0,a,1}(T) - E\bar{Y}_{0,a,1}(T) \right| \leqslant \sum_{k=0}^{N-2} \frac{Kh^5}{\sqrt{T - t_{k+1}}} + Kh^4.$$

Since $\sum_{k=0}^{N-2} \frac{h}{\sqrt{T - t_{k+1}}} \leqslant C$, we get

$$\left| EY_{0,a,1}(T) - E\bar{Y}_{0,a,1}(T) \right| \leqslant Kh^4,$$

i.e., we have proved that the method (2.9), (2.10) is of order four. □

Appendix C. Proof of the convergence theorem for the implicit method

This appendix contains the proof of convergence of the implicit Runge–Kutta method (4.1), (4.2).

Proof of Theorem 4.1. In the method (4.1), (4.2), the approximation $\bar{Y}(t)$ is evaluated at $t = t_k,\ k = 1, \ldots, N$, while $X(t)$ is simulated at $t = t_{k+1/2}$ and $X(t_k)$ is not used in the algorithm. Due to this reason, we cannot directly make use of relations like (3.8), (3.9) to prove convergence of the method (4.1), (4.2). To overcome this difficulty, we consider the other algorithm:

$$\begin{aligned} X(t_{k+1/2}) &= X(t_k) + \frac{h}{2}\frac{b - X(t_k)}{T - t_k} + \sqrt{\frac{h}{2}}\sqrt{\frac{T - t_{k+1/2}}{T - t_k}}\xi_{k+1/2}, \qquad k = 0, \ldots, N-1, \\ X(t_{k+1}) &= X(t_{k+1/2}) + \frac{h}{2}\frac{b - X(t_{k+1/2})}{T - t_{k+1/2}} + \sqrt{\frac{h}{2}}\sqrt{\frac{T - t_{k+1}}{T - t_{k+1/2}}}\xi_{k+1},\ k = 0, \ldots, N-2, \quad X(t_N) = b, \end{aligned} \tag{C.1}$$

and Y_k, $k = 0, \ldots, N-1$, are simulated by the same formulas as in (4.2) (or, what is the same, (4.3)). In (C.1), $\xi_{k+1/2}$ and ξ_{k+1} are d-dimensional random vectors which components are mutually independent random variables with standard normal distribution $\mathcal{N}(0,1)$.

Since $X(t)$ is simulated exactly both in (4.1) and (C.1) and, in particular, $X(t_{k+1/2})$ from (4.1) have the same distributions as their counterparts in (C.1), it is clear that the estimate (4.4) for the algorithm (C.1), (4.2) implies this estimate for (4.1), (4.2). At the same time, due to the presence of $X(t_{k+1})$ in (C.1), we can make use of relations like (3.8), (3.9) to estimate the error of the algorithm (C.1), (4.2). In what follows, we prove (4.4) for (C.1), (4.2).

We write the global error of (C.1), (4.2) in the form (3.8), (3.9) and introduce the one-step error of (C.1), (4.2) as in (3.10):

$$r(t,x) := Eu(t+h, X_{t,x}(t+h))\bar{Y}_{t,x,1}(t+h) - u(t,x). \tag{C.2}$$

We rewrite (4.3) on a single step and expand it as

$$\bar{Y}_{t,x,1}(t+h) = 1 + hf_{1/2} + \frac{h^2}{2}f_{1/2}^2 + \rho, \tag{C.3}$$

where $f_{1/2} := f(t+h/2, X_{t,x}(t+h/2))$ and the random variable ρ is such that

$$\left(E\rho^2\right)^{1/2} \leqslant Ch^3. \tag{C.4}$$

We substitute (C.3) in (C.2) and then expand the terms in the obtained relation using (3.5), (3.6) as we did in the proof of Theorem 3.2 (see pp. 7–9). In fact, the expansions are simpler here since we are proving the second order of convergence only. Then, taking into account that $u(t,x)$ is a solution of (3.4), we arrive at

$$\begin{aligned} r(t,x) = &-\int_t^{t+h} \frac{(t+h-\theta)^2}{2} EL^2(fu)(\theta, X_{t,x}(\theta))\,\mathrm{d}\theta + h\int_t^{t+h/2}(t+h/2-\theta)EL^2(fu)(\theta, X_{t,x}(\theta))\,\mathrm{d}\theta \\ &+ hEf_{1/2}\int_{t+h/2}^{t+h}(t+h-\theta)L^2u(\theta, X_{t,x}(\theta))\,\mathrm{d}\theta + \frac{h^2}{2}Ef_{1/2}^2\int_{t+h/2}^{t+h}Lu(\theta, X_{t,x}(\theta))\,\mathrm{d}\theta \\ &+ Eu(t+h, X_{t,x}(t+h))\rho. \end{aligned} \tag{C.5}$$

Using Lemmas 3.1 and A.4 (cf. the proof of Theorem 3.2), we obtain that the one-step error of the method (4.1), (4.2) can be written in the form

$$r(t,x) = h^3 S(t,x) + E\tilde{\rho}(t,x;h), \tag{C.6}$$

where $S(t,x)$ is a linear combination of the functions $\psi^{\alpha_2}(t,x)$, coefficients in this linear combination are independent of t, x, and h; $\tilde{\rho}(t,x;h)$ is such that

$$\left(E\left[\tilde{\rho}(t, X_{0,a}(t); h)\right]^{2n}\right)^{1/2n} \leqslant \frac{Ch^3}{\sqrt{T-t-h}}$$

with a constant C independent of t and h. We see that $S(t,x)$ in (C.6) and, consequently, the one-step error $r(t,x)$, has singularity of order one.

Further, using arguments similar to those in the proofs of Lemma B.1 and Theorem 2.1 (in fact, due to the lower order of convergence and lower order of singularity, much simpler calculations are needed here), we obtain (4.4). □

Acknowledgements

We acknowledge partial support from the UK EPSRC Research Grant GR/R90017/01.

References

[1] S.D. Bond, B.B. Laird, B.J. Leimkuhler, On the approximation of Feynman–Kac path integrals, J. Comp. Phys. 185 (2003) 472–483.

[2] A.D. Egorov, P.I. Sobolevsky, L.A. Yanovich, Functional Integrals: Approximate Evaluation and Applications, Kluwer Academic Publishers, 1993.

[3] R.P. Feynman, A.R. Hibbs, Quantum Mechanics and Path Integrals, McGraw-Hill, New York, 1965.

[4] I.M. Gelfand, A.M. Yaglom, Integration in functional spaces and its application in quantum physics, J. Math. Phys. 1 (1960) 48–69 (English Trans. from Russian, Uspechi Mat. Nauk 11 (1956) 77–114).
[5] S.A. Gladyshev, G.N. Milstein, Runge–Kutta method for evaluation of Wiener integrals of exponential-type functionals, Zh. Vychisl. Mat. Mat. Fiz. 24 (1984) 1136–1149.
[6] E. Hairer, G. Wanner, Solving ordinary differential equations. II. Stiff and differential-algebraic problems, Springer, Berlin, 1996.
[7] N. Ikeda, S. Watanabe, Stochastic Differential Equations and Diffusion Processes, North-Holland, Amsterdam, 1981.
[8] I. Karatzas, S.E. Shreve, Brownian Motion and Stochastic Calculus, Springer, Berlin, 1988.
[9] H. Kleinert, Path Integrals in Quantum Mechanics, Statistics and Polymer Physics, World Scientific, New York, 1995.
[10] G.N. Milstein, Numerical Integration of Stochastic Differential Equations, Kluwer Academic Publishers, 1995 (Engl. Transl. from Russian 1988).
[11] G.N. Milstein, J.G.M. Schoenmakers, Numerical construction of a hedging strategy against the multi-asset European claim, Stoch. Stoch. Rep. 73 (2002) 125–157.
[12] N.J. Newton, Continuous-time Monte Carlo methods and variance reduction, in: L.C.G. Rogers, D. Talay (Eds.), Numerical Methods in Finance, Cambridge University Press, Cambridge, 1997, pp. 22–42.
[13] G. Roepstorff, Path Integral Approach to Quantum Physics, Springer, Berlin, 1994.
[14] A.D. Ventzel, S.A. Gladyshev, G.N. Milstein, Piecewise constant approximation for Monte Carlo calculation of Wiener integrals, Theor. Prob. Appl. 29 (1984) 744–752.
[15] W. Wagner, Unbiased multi-step estimators for the Monte Carlo evaluation of certain functional integrals, J. Comp. Phys. 79 (1988) 336–352.