

SOLVING PARABOLIC STOCHASTIC PARTIAL DIFFERENTIAL EQUATIONS VIA AVERAGING OVER CHARACTERISTICS

G. N. MILSTEIN AND M. V. TRETYAKOV

ABSTRACT. The method of characteristics (the averaging over the characteristic formula) and the weak-sense numerical integration of ordinary stochastic differential equations together with the Monte Carlo technique are used to propose numerical methods for linear stochastic partial differential equations (SPDEs). Their orders of convergence in the mean-square sense and in the sense of almost sure convergence are obtained. A variance reduction technique for the Monte Carlo procedures is considered. Layer methods for linear and semilinear SPDEs are constructed and the corresponding convergence theorems are proved. The approach developed is supported by numerical experiments.

1. INTRODUCTION

A lot of attention has recently been paid to numerical methods for stochastic partial differential equations (SPDEs). Various numerical approaches for linear SPDEs are considered, e.g., in [31, 1, 19, 17, 34, 10, 11, 3, 7, 25, 5] (see also the references therein). The interest in linear SPDEs of parabolic type is mainly due to their well-known relation with the nonlinear filtering problem [18, 16, 28, 32, 31] although they have other applications as well (see, e.g., [32] and the references therein). At the same time, very little has been done in studying applications of the Monte Carlo technique to solving SPDEs while such a technique is the well-established tool for solving problems of mathematical physics associated with multi-dimensional (deterministic) partial differential equations (see, e.g., [23] and the references therein). The aim of this paper is to exploit the method of characteristics (the averaging over the characteristic formula, the generalization of the Feynman-Kac formula) and numerical integration of (ordinary) stochastic differential equations (SDEs) together with the Monte Carlo technique to propose numerical methods for linear SPDEs of parabolic type. The Monte Carlo methods are efficient for solving high-dimensional SPDEs when functionals or individual values of the solution have to be found. We note that the method of characteristics was exploited in [3] to propose a particle method for the Kushner-Stratonovich equation, and it was used in [25] for constructing Monte Carlo methods for a less general class of SPDEs than the one considered here.

Received by the editor May 30, 2007 and, in revised form, November 3, 2008.

2000 *Mathematics Subject Classification.* Primary 65C30, 60H15, 60H35, 60G35.

Key words and phrases. Probabilistic representations of solutions of stochastic partial differential equations, numerical integration of stochastic differential equations, Monte Carlo technique, mean-square and almost sure convergence, layer methods.

©2009 American Mathematical Society

The probabilistic approach based on the method of characteristics is also exploited to construct new layer methods for SPDEs (see this idea in the case of deterministic PDEs in [21, 23]). The layer methods are competitive with finite difference schemes (see [10, 34] and the references therein); they can be used when one needs to find the SPDE solution everywhere in the space-time domain $[t_0, T] \times \mathbf{R}^d$. It was shown in the deterministic case [21, 23] that layer methods have some advantages in comparison with conventional PDE solvers. We expect that they possess remarkable properties in the SPDE case as well. We construct layer methods both for linear and semilinear SPDEs. Semilinear SPDEs are used for modelling in physics, biology and chemistry (see [4, 13] and the references therein). For other numerical approaches to semilinear SPDEs, see, e.g., in [8, 12, 2] (see also the references therein).

In Section 2 we recall probabilistic representations (the method of characteristics) for SPDEs of parabolic type from [15, 16, 28, 32]. In Section 3 we propose a number of approximation methods for the SPDEs and study their mean-square and almost sure (a.s.) convergence. The methods are based on approximate solving of the characteristic SDEs, for which we exploit both the mean-square and weak-sense numerical integration. Section 4 deals with variance reduction methods that are important for any Monte Carlo procedures. We propose layer methods for linear SPDEs in Section 5 and for semilinear SPDEs in Section 6. Some results of numerical experiments are presented in Section 7. We solve numerically the Ornstein-Uhlenbeck SPDE and, in particular, demonstrate the effectiveness of the proposed variance reduction technique.

2. CONDITIONAL PROBABILISTIC REPRESENTATIONS OF SOLUTIONS TO LINEAR SPDES

Let $(\Omega, \mathcal{F}, P)$ be a probability space, $\mathcal{F}_t$, $T_0 \leq t \leq T$, be a nondecreasing family of σ -subalgebras of $\mathcal{F}$, $(w(t), \mathcal{F}_t) = ((w_1(t), \dots, w_q(t))^\top, \mathcal{F}_t)$ be a q -dimensional standard Wiener process. Consider the Cauchy problem for the backward SPDE

(2.1)

$$-dv = [\mathcal{L}v + f(t, x)] dt + \sum_{r=1}^q [\mathcal{M}_r v + \gamma_r(t, x)] * dw_r(t), \quad (t, x) \in [T_0, T] \times \mathbf{R}^d,$$

(2.2)

$$v(T, x) = \varphi(x), \quad x \in \mathbf{R}^d,$$

where

(2.3)

$$\begin{aligned} \mathcal{L}v(t, x) &:= \frac{1}{2} \sum_{i,j=1}^d a^{ij}(t, x) \frac{\partial^2}{\partial x^i \partial x^j} v(t, x) + \sum_{i=1}^d b^i(t, x) \frac{\partial}{\partial x^i} v(t, x) + c(t, x) v(t, x), \\ \mathcal{M}_r v(t, x) &:= \sum_{i=1}^d \alpha_r^i(t, x) \frac{\partial}{\partial x^i} v(t, x) + \beta_r(t, x) v(t, x), \quad r = 1, \dots, q. \end{aligned}$$

The notation “ $*dw_r$ ” in (2.1) means the backward Ito integral. We recall [32] that to define this integral one introduces the “backward” Wiener processes

$$(2.4) \quad \tilde{w}_r(t) := w_r(T) - w_r(T - (t - T_0)), \quad r = 1, \dots, q, \quad T_0 \leq t \leq T,$$

and a decreasing family of σ -subalgebras $\mathcal{F}_T^t$, $T_0 \leq t \leq T$, induced by the increments $w_r(T) - w_r(t')$, $r = 1, \dots, q$, $t' \geq t$. A σ -algebra induced by $\tilde{w}_r(t')$, $t' \leq t$, coincides

with $\mathcal{F}_T^{T-(t-T_0)}$. Then the backward Ito integral is defined as the Ito integral with respect to $\tilde{w}(s)$:

$$\int_t^{t'} \psi(t'') * dw_r(t'') := \int_{T-(t'-T_0)}^{T-(t-T_0)} \psi(T-(t''-T_0)) d\tilde{w}_r(t''), \quad T_0 \leq t \leq t' \leq T,$$

where $\psi(T-(t-T_0))$, $t \leq T$, is an $\mathcal{F}_T^{T-(t-T_0)}$ -adapted square-integrable function. The process $v(t, x)$ is $\mathcal{F}_T^t$ -adapted; it depends on $w(s) - w(t)$, $t \leq s \leq T$. We pay attention that the more precise notation for the solution of (2.1)-(2.2) is $v(t, x; \omega)$, $\omega \in \Omega$, but we use, as a rule, the shorter one, $v(t, x)$.

Assumption 2.1. We assume that the coefficients in (2.1) are sufficiently smooth and that their derivatives up to some order are bounded (in particular, it follows from here that the coefficients are globally Lipschitz). Furthermore, it is supposed that c , β_r , f , and γ_r are bounded themselves.

Assumption 2.2. We assume that the function $\varphi(x)$ is also sufficiently smooth and that $\varphi(x)$ and its derivatives up to some order belong to the class functions satisfying an inequality of the form

$$(2.5) \quad |\varphi(x)| \leq K(1 + |x|^\varkappa), \quad x \in \mathbf{R}^d,$$

where K and $\varkappa$ are positive constants.

Assumption 2.3. We assume that $a = \{a^{ij}\}$ is symmetric and that the matrix $a - \alpha\alpha^\top$ is nonnegative definite (the coercivity condition).

Assumptions 2.1-2.3 ensure the existence of a unique classical solution $v(t, x)$ of (2.1)-(2.2) which has derivatives in x^i , $i = 1, \dots, d$, up to a sufficiently high order satisfying an inequality of the form (2.5) a.s. with a positive random variable $K = K(\omega)$ for which the moments of a sufficiently high order are bounded (see [14, 32]). They are sufficient for all the statements in Sections 3 and 4 (some additional assumptions are needed in Section 5). At the same time, they are not necessary and the methods constructed can be used under broader conditions. We note that an attempt to weaken the conditions would inevitably lead to difficulties of a technical nature and, as a result, to a less clear exposition together with an unnecessary increase of the paper's length.

Let a $d \times p$ matrix $\sigma(t, x)$ be obtained from the equation

$$\sigma(t, x)\sigma^\top(t, x) = a(t, x) - \alpha(t, x)\alpha^\top(t, x).$$

The solution of the problem (2.1)-(2.2) has the following conditional probabilistic representation (the conditional Feynman-Kac formula) [15, 16, 28, 32]:

$$(2.6) \quad v(t, x) = E^w [\varphi(X_{t,x}(T))Y_{t,x,1}(T) + Z_{t,x,1,0}(T)], \quad T_0 \leq t \leq T,$$

where $X_{t,x}(s)$, $Y_{t,x,y}(s)$, $Z_{t,x,y,z}(s)$, $t \leq s \leq T$, is the solution of the SDEs

$$(2.7) \quad \begin{aligned} dX &= \left[b(s, X) - \sum_{r=1}^q \alpha_r(s, X)\beta_r(s, X) \right] ds \\ &+ \sum_{r=1}^p \sigma_r(s, X)dW_r(s) + \sum_{r=1}^q \alpha_r(s, X)dw_r(s), \\ X(t) &= x, \end{aligned}$$

$$dY = c(s, X)Yds + \sum_{r=1}^q \beta_r(s, X)Ydw_r(s), \quad Y(t) = y,$$

$$dZ = f(s, X)Yds + \sum_{r=1}^q \gamma_r(s, X)Ydw_r(s), \quad Z(t) = z,$$

and $W(s) = (W_1(s), \dots, W_p(s))^T$ is a p -dimensional standard Wiener process independent of $w(s)$, and the expectation E^w in (2.6) is taken over the realizations of $W(s)$, $t \leq s \leq T$, for a fixed $w(s)$, $t \leq s \leq T$; in other words, $E^w(\cdot)$ means the conditional expectation $E(\cdot | w(s) - w(t), t \leq s \leq T)$.

Remark 2.1. Introduce the new time variable $s := T - (t - T_0)$ and introduce the functions $u(s, x) := v(T + T_0 - s, x)$, $\tilde{a}^{ij}(s, x) := a^{ij}(T + T_0 - s, x)$, $\tilde{b}^i(s, x) := b^i(T + T_0 - s, x)$, and analogously $\tilde{c}(s, x)$, $\tilde{f}(s, x)$, $\tilde{\alpha}_r^i(s, x)$, $\tilde{\beta}_r(s, x)$, $\tilde{\gamma}_r(s, x)$. Then one can show that $u(s, x)$ is the solution of the Cauchy problem for the forward SPDE (see [32, p. 173]):

$$(2.8) \quad du = \left[\frac{1}{2} \sum_{i,j=1}^d \tilde{a}^{ij}(s, x) \frac{\partial^2}{\partial x^i \partial x^j} u + \sum_{i=1}^d \tilde{b}^i(s, x) \frac{\partial}{\partial x^i} u + \tilde{f}(s, x) \right] ds$$

$$+ \sum_{r=1}^q \left[\sum_{i=1}^d \tilde{\alpha}_r^i(s, x) \frac{\partial}{\partial x^i} u + \tilde{\beta}_r(s, x) u + \tilde{\gamma}_r(s, x) \right] d\tilde{w}_r(s), \quad (s, x) \in (T_0, T] \times \mathbf{R}^d,$$

$$(2.9) \quad u(T_0, x) = \varphi(x), \quad x \in \mathbf{R}^d,$$

where $\tilde{w}_r(s)$ are defined in (2.4). The process $u(s, x)$ is $\mathcal{F}_T^{T+T_0-t}$ -adapted; it depends on $\tilde{w}_r(s')$, $r = 1, \dots, q$, $T_0 \leq s' \leq s$. Analogously, for a given forward SPDE one can write down the corresponding backward SPDE. Thus, the methods for backward SPDEs considered in this paper can be used for solving forward SPDEs as well.

Remark 2.2. Consider an infinite-dimensional Wiener process $B(t)$ taking values in some Hilbert space H and with covariance operator Q (which is assumed to be a nuclear operator). Let e_r be unit eigenvectors of Q with nonzero eigenvalues $\lambda_r = (Qe_r, e_r)$. Then (see, e.g. [4, 14, 32]) $w_r(t) := (e_r, B(t))/\sqrt{\lambda_r}$ are independent standard Wiener processes, and the infinite-dimensional Wiener process is represented as

$$(2.10) \quad B(t) = \sum_{r=1}^{\infty} \sqrt{\lambda_r} w_r(t) e_r.$$

Furthermore, for any H -valued process $\psi(s)$ for which the integral $\int_0^t \psi(s) dB(s)$ is defined, one has [14, 32]

$$(2.11) \quad \int_0^t \psi(s) dB(s) = \sum_{r=1}^{\infty} \int_0^t \psi_r(s) dw_r(t)$$

with $\psi_r(s) = (\psi(s), e_r)\sqrt{\lambda_r}$. The Wiener process $B(t)$ and the integral $\int_0^t \psi(s) dB(s)$ can be approximated by truncating the expansions in (2.10) and (2.11) (see, e.g., [4]). Then it is possible to view the SPDEs (2.1) and (2.8) as approximations of SPDEs driven by the infinite-dimensional Wiener process. Consequently, the methods proposed in this paper for (2.1) can, in principle, be used for approximating

SPDEs with infinite-dimensional Wiener process, but we do not consider this aspect any further here.

3. NUMERICAL METHODS BASED ON THE CONDITIONAL FEYNMAN-KAC FORMULA (2.6)-(2.7)

Our purpose is to simulate approximately the random variable $v(t, x, \omega)$ under any fixed (t, x) using the probabilistic representation (2.6)-(2.7). In this section we construct mean-square approximations $\bar{v}(t, x, \omega)$ (see Section 3.1) and $\tilde{v}(t, x, \omega)$ (see Section 3.2) for $v(t, x, \omega)$; i.e., we construct $\bar{v}(t, x, \omega)$ (or $\tilde{v}(t, x, \omega)$) to be close to $v(t, x, \omega)$ in the mean-square sense. For a realization of the probabilistic representation (2.6), one can use numerical integration of SDEs (2.7) with respect to the Wiener process $W(s)$ both in the mean-square sense (Section 3.1; the approximation $\bar{v}(t, x)$) and in the weak sense (Section 3.2; the approximation $\tilde{v}(t, x)$). We prove a.s. convergence of the proposed methods.

3.1. Mean-square simulation of the representation (2.6)-(2.7). Consider the one-step approximation for (2.7):

$$(3.1) \quad \begin{aligned} \bar{X}_{t,x}(t+h) &= x + h \left[b(t, x) - \sum_{r=1}^q \alpha_r(t, x) \beta_r(t, x) \right] \\ &\quad + \sum_{r=1}^q \alpha_r(t, x) \Delta_t w_r + \sum_{r=1}^p \sigma_r(t, x) \Delta_t W_r, \\ \bar{Y}_{t,x,y}(t+h) &= y + hc(t, x)y + \sum_{r=1}^q \beta_r(t, x)y \Delta_t w_r, \\ \bar{Z}_{t,x,y,z}(t+h) &= z + hf(t, x)y + \sum_{r=1}^q \gamma_r(t, x)y \Delta_t w_r, \end{aligned}$$

where $\Delta_t w_r := w_r(t+h) - w_r(t)$, $\Delta_t W_r := W_r(t+h) - W_r(t)$.

The approximation (3.1) generates the strong Euler method in the usual way. We introduce a partition of the time interval $[t, T]$, for simplicity the equidistant one: $t = t_0 < \dots < t_N = T$, with step size $h = (T - t)/N$. An approximation of $(X_{t,x}(t_k), Y_{t,x,1}(t_k), Z_{t,x,1,0}(t_k))$ is denoted by $(\bar{X}_k, \bar{Y}_k, \bar{Z}_k)$.

The Euler scheme takes the form

$$(3.2) \quad \begin{aligned} \bar{X}_0 &= x, \quad \bar{X}_{k+1} = \bar{X}_k + h \left[b(t_k, \bar{X}_k) - \sum_{r=1}^q \alpha_r(t_k, \bar{X}_k) \beta_r(t_k, \bar{X}_k) \right] \\ &\quad + \sum_{r=1}^q \alpha_r(t_k, \bar{X}_k) \Delta_k w_r + \sum_{r=1}^p \sigma_r(t_k, \bar{X}_k) \Delta_k W_r, \\ \bar{Y}_0 &= 1, \quad \bar{Y}_{k+1} = \bar{Y}_k + hc(t_k, \bar{X}_k) \bar{Y}_k + \sum_{r=1}^q \beta_r(t_k, \bar{X}_k) \bar{Y}_k \Delta_k w_r, \\ \bar{Z}_0 &= 0, \quad \bar{Z}_{k+1} = \bar{Z}_k + hf(t_k, \bar{X}_k) \bar{Y}_k + \sum_{r=1}^q \gamma_r(t_k, \bar{X}_k) \bar{Y}_k \Delta_k w_r, \quad k = 0, \dots, N-1, \end{aligned}$$

where $\Delta_k w_r := w_r(t_{k+1}) - w_r(t_k)$, $\Delta_k W_r := W_r(t_{k+1}) - W_r(t_k)$.

Let

$$(3.3) \quad \bar{v}(t, x) := E^w [\varphi(\bar{X}_N) \bar{Y}_N + \bar{Z}_N],$$

where $\bar{X}_N, \bar{Y}_N, \bar{Z}_N$ are from (3.2).

Below we use the same letter $C = C(x)$ or $C = C(x, \omega)$ for various expressions of the form $K(1 + |x|^\varkappa)$ (see (2.5)) with K being a positive constant or $K = K(\omega)$ being a positive random variable, respectively.

Theorem 3.1. *Let Assumptions 2.1-2.3 hold. The method (3.2)-(3.3) satisfies the inequality for $p \geq 1$:*

$$(3.4) \quad \left(E |\bar{v}(t, x) - v(t, x)|^{2p} \right)^{1/(2p)} \leq C(x) h^{1/2},$$

where C does not depend on the discretization step h , i.e., in particular, (3.2)-(3.3) is of mean-square order $1/2$.

For almost every trajectory $w(\cdot)$ and any $\varepsilon > 0$ there exists $C(x, \omega) > 0$ such that

$$(3.5) \quad |\bar{v}(t, x) - v(t, x)| \leq C(x, \omega) h^{1/2-\varepsilon},$$

where C does not depend on the discretization step h , i.e., the method (3.2)-(3.3) converges with order $1/2 - \varepsilon$ a.s.

Proof. Introduce $\mathbf{x} = (x, y, z)$, $\mathbf{X}(s) := (X(s), Y(s), Z(s))^\top$, $\bar{\mathbf{X}}_k := (\bar{X}_k, \bar{Y}_k, \bar{Z}_k)^\top$, and $\psi(\mathbf{x}) := \varphi(x)y + z$.

It is known that $\mathbf{X}(s)$ (see [6]) and $\bar{\mathbf{X}}_k$ (see [23]) have bounded moments of any order and also that (see [9]) for $p \geq 1$,

$$(3.6) \quad \left(E |\bar{\mathbf{X}}_k - \mathbf{X}(t_k)|^{2p} \right)^{1/(2p)} \leq C(x) h^{1/2},$$

where C does not depend on h .

By the conditional version of Jensen's inequality, we get

$$\begin{aligned} E |\bar{v}(t, x) - v(t, x)|^{2p} &= E |E^w \psi(\bar{\mathbf{X}}_N) - E^w \psi(\mathbf{X}(t_N))|^{2p} \\ &\leq E [E^w |\psi(\bar{\mathbf{X}}_N) - \psi(\mathbf{X}(t_N))|]^{2p} \\ &\leq E [E^w |\psi(\bar{\mathbf{X}}_N) - \psi(\mathbf{X}(t_N))|^{2p}] \\ &= E |\psi(\bar{\mathbf{X}}_N) - \psi(\mathbf{X}(t_N))|^{2p}. \end{aligned}$$

Using the smoothness of ψ and the assumption that derivatives of ψ satisfy an inequality of the form (2.5), we obtain

$$|\psi(\bar{\mathbf{X}}_N) - \psi(\mathbf{X}(t_N))| \leq K(1 + |\mathbf{X}(t_N)|^\varkappa + |\bar{\mathbf{X}}_N|^\varkappa) |\bar{\mathbf{X}}_N - \mathbf{X}(t_N)|,$$

where K is a positive constant. Then due to the boundedness of the moments, the Cauchy-Bunyakovskii inequality, and (3.6), we get

$$\begin{aligned} E |\bar{v}(t, x) - v(t, x)|^{2p} &\leq KE(1 + |\mathbf{X}(t_N)|^\varkappa + |\bar{\mathbf{X}}_N|^\varkappa)^{2p} |\bar{\mathbf{X}}_N - \mathbf{X}(t_N)|^{2p} \\ &\leq K \sqrt{E(1 + |\mathbf{X}(t_N)|^\varkappa + |\bar{\mathbf{X}}_N|^\varkappa)^{4p}} \sqrt{E |\bar{\mathbf{X}}_N - \mathbf{X}(t_N)|^{4p}} \leq C(x) h^p, \end{aligned}$$

whence (3.4) follows.

Now denote $R := |\bar{v}(t, x) - v(t, x)|$. The Markov inequality together with (3.4) implies

$$P(R > h^\gamma) \leq \frac{ER^{2p}}{h^{2p\gamma}} \leq C(x)h^{p(1-2\gamma)}.$$

Then for any $\gamma = 1/2 - \varepsilon$ there is a sufficiently large $p \geq 1$ such that (recall that $h = T/N$)

$$\sum_{N=1}^{\infty} P\left(R > \frac{T^\gamma}{N^\gamma}\right) \leq C(x)T^{p(1-2\gamma)} \sum_{N=1}^{\infty} \frac{1}{N^{p(1-2\gamma)}} < \infty.$$

Hence, due to the Borel-Cantelli lemma, the random variable $\varsigma := \sup_{h>0} h^{-\gamma} R$ is a.s. finite, which implies (3.5). $\square$

3.2. Weak simulation of the representation (2.6)-(2.7). Now consider another Euler-type scheme for (2.7):

(3.7)

$$\begin{aligned} \tilde{X}_0 &= x, \quad \tilde{X}_{k+1} = \tilde{X}_k + h \left[b(t_k, \tilde{X}_k) - \sum_{r=1}^q \alpha_r(t_k, \tilde{X}_k) \beta_r(t_k, \tilde{X}_k) \right] \\ &\quad + \sum_{r=1}^q \alpha_r(t_k, \tilde{X}_k) \Delta_k w_r + h^{1/2} \sum_{r=1}^p \sigma_r(t_k, \tilde{X}_k) \xi_{rk}, \\ \tilde{Y}_0 &= 1, \quad \tilde{Y}_{k+1} = \tilde{Y}_k + hc(t_k, \tilde{X}_k) \tilde{Y}_k + \sum_{r=1}^q \beta_r(t_k, \tilde{X}_k) \tilde{Y}_k \Delta_k w_r, \\ \tilde{Z}_0 &= 0, \quad \tilde{Z}_{k+1} = \tilde{Z}_k + hf(t_k, \tilde{X}_k) \tilde{Y}_k + \sum_{r=1}^q \gamma_r(t_k, \tilde{X}_k) \tilde{Y}_k \Delta_k w_r, \quad k = 0, \dots, N-1, \end{aligned}$$

where $\Delta_k w_r := w_r(t_{k+1}) - w_r(t_k)$ and ξ_{rk} are i.i.d. random variables with the moments

$$(3.8) \quad E\xi = 0, \quad E\xi^2 = 1, \quad E\xi^3 = 0, \quad E\xi^m < \infty,$$

for a sufficiently large integer $m \geq 4$. For instance, one can take ξ with the law

$$(3.9) \quad P(\xi = \pm 1) = 1/2.$$

We note that here a part of the SDE system (2.7) (which involves the Wiener process $W(t)$) is simulated weakly while the other part (which involves the Wiener process $w(t)$) is simulated in the mean-square sense.

Let

$$(3.10) \quad \tilde{v}(t, x) := E^w \left[\varphi(\tilde{X}_N) \tilde{Y}_N + \tilde{Z}_N \right],$$

where $\tilde{X}_N, \tilde{Y}_N, \tilde{Z}_N$ are from (3.7).

The same theorem as Theorem 3.1 is valid for $\tilde{v}(t, x)$ as well, however, with another proof.

Theorem 3.2. *Let Assumptions 2.1-2.3 hold. The method (3.7), (3.10) satisfies the inequality for $p \geq 1$:*

$$(3.11) \quad \left(E |\tilde{v}(t, x) - v(t, x)|^{2p} \right)^{1/(2p)} \leq C(x)h^{1/2},$$

where C does not depend on the discretization step h ; i.e., in particular, (3.7), (3.10) is of mean-square order $1/2$.

For almost every trajectory $w(\cdot)$ and any $\varepsilon > 0$ there exists $C(x, \omega) > 0$ such that

$$(3.12) \quad |\tilde{v}(t, x) - v(t, x)| \leq C(x, \omega) h^{1/2-\varepsilon},$$

where C does not depend on the discretization step h ; i.e., the method (3.7), (3.10) converges with order $1/2 - \varepsilon$ a.s.

Proof. We shall prove the inequality

$$(3.13) \quad \left(E |\tilde{v}(t, x) - \bar{v}(t, x)|^{2p} \right)^{1/(2p)} \leq C(x)h.$$

Due to Theorem 3.1 (see (3.4)), the inequality (3.11) follows from here.

The function $\bar{v}(t, x)$ (see (3.3)) is introduced for the time layer $t = t_0$. For any t_i let us introduce $\bar{v}(t_i, x)$. To this aim we use the scheme (3.2) starting at the time moment t_i from $\bar{X}_i = x$, $\bar{Y}_i = 1$, $\bar{Z}_i = 0$. Denote this solution of the scheme by $\bar{X}_{t_i, x}(t_k)$, $\bar{Y}_{t_i, x, 1}(t_k)$, $\bar{Z}_{t_i, x, 1, 0}(t_k)$, $k = i, \dots, N$. For instance, $\bar{X}_N$ in (3.3) in this notation is equal to $\bar{X}_N = \bar{X}_{t_0, x}(t_N) = \bar{X}_{t, x}(t_N)$. We set

$$(3.14) \quad \bar{v}(t_i, x) = E^w [\varphi(\bar{X}_{t_i, x}(t_N)) \bar{Y}_{t_i, x, 1}(t_N) + \bar{Z}_{t_i, x, 1, 0}(t_N)], \quad i = 0, \dots, N, \quad t_N = T.$$

One can prove that the function $\bar{v}(t_i, x)$ is sufficiently smooth in x and satisfies (together with its derivatives) the same inequality as the function $v(t, x)$ (see the inequality (2.5)).

Using the standard technique (see [23, p. 100]), we can write the difference $D_N := \tilde{v}(t, x) - \bar{v}(t, x)$ in the form

$$(3.15) \quad \begin{aligned} D_N &= E^w [\varphi(\tilde{X}_N) \tilde{Y}_N + \tilde{Z}_N] - E^w [\varphi(\bar{X}_{t_0, x}(t_N)) \bar{Y}_{t_0, x, 1}(t_N) + \bar{Z}_{t_0, x, 1, 0}(t_N)] \\ &= E^w \sum_{i=0}^{N-1} \rho_i(\tilde{X}_i, \tilde{Y}_i, \tilde{Z}_i), \end{aligned}$$

where

$$(3.16) \quad \begin{aligned} \rho_i(\tilde{X}_i, \tilde{Y}_i, \tilde{Z}_i) &= E^{w, \tilde{X}_i, \tilde{Y}_i, \tilde{Z}_i} [\bar{v}(t_{i+1}, \tilde{X}_{i+1}) \tilde{Y}_{i+1} + \tilde{Z}_{i+1} \\ &\quad - \bar{v}(t_{i+1}, \bar{X}_{t_i, \tilde{X}_i}(t_{i+1})) \bar{Y}_{t_i, \tilde{X}_i, \tilde{Y}_i}(t_{i+1}) - \bar{Z}_{t_i, \tilde{X}_i, \tilde{Y}_i, \tilde{Z}_i}(t_{i+1})], \end{aligned}$$

and $E^{w, \tilde{X}_i, \tilde{Y}_i, \tilde{Z}_i}[\cdot]$ means the conditional expectation $E[\cdot/w(s) - w(t), t \leq s \leq T; \tilde{X}_i, \tilde{Y}_i, \tilde{Z}_i]$. The presentation (3.15) follows from the equalities

$$(3.17) \quad \bar{v}(t_N, \tilde{X}_N) \tilde{Y}_N + \tilde{Z}_N = \varphi(\tilde{X}_N) \tilde{Y}_N + \tilde{Z}_N,$$

$$(3.18) \quad \begin{aligned} E^w [\bar{v}(t_1, \bar{X}_{t_0, \tilde{X}_0}(t_1)) \bar{Y}_{t_0, \tilde{X}_0, \tilde{Y}_0}(t_1) + \bar{Z}_{t_0, \tilde{X}_0, \tilde{Y}_0, \tilde{Z}_0}(t_1)] \\ = E^w [\varphi(\bar{X}_{t_0, x}(T)) \bar{Y}_{t_0, x, 1}(T) + \bar{Z}_{t_0, x, 1, 0}(T)], \end{aligned}$$

$$(3.19) \quad \begin{aligned} E^w [\bar{v}(t_k, \tilde{X}_k) \tilde{Y}_k + \tilde{Z}_k] &= E^w [\bar{v}(t_{k+1}, \bar{X}_{t_k, \tilde{X}_k}(t_{k+1})) \bar{Y}_{t_k, \tilde{X}_k, \tilde{Y}_k}(t_{k+1}) \\ &\quad + \bar{Z}_{t_k, \tilde{X}_k, \tilde{Y}_k, \tilde{Z}_k}(t_{k+1})], \\ k &= 1, \dots, N-1. \end{aligned}$$

The equality (3.17) is obvious. To prove (3.18) and (3.19), we use the following evident equalities:

$$(3.20) \quad \bar{Y}_{t_k, x, y}(\vartheta) = \bar{Y}_{t_k, x, 1}(\vartheta) \cdot y, \quad \bar{Z}_{t_k, x, y, z}(\vartheta) = z + \bar{Z}_{t_k, x, 1, 0}(\vartheta) \cdot y.$$

We have (see (3.14))

$$\begin{aligned}\bar{v}(t_k, \tilde{X}_k) &= E^{w, \tilde{X}_k}[\varphi(\bar{X}_{t_k, \tilde{X}_k}(T))\bar{Y}_{t_k, \tilde{X}_k, 1}(T) + \bar{Z}_{t_k, \tilde{X}_k, 1, 0}(T)] \\ &= E^{w, \tilde{X}_k}[\varphi(\bar{X}_{t_{k+1}, \tilde{X}_{t_k, \tilde{X}_k}(t_{k+1})}(T))\bar{Y}_{t_{k+1}, \tilde{X}_{t_k, \tilde{X}_k}(t_{k+1}), \bar{Y}_{t_k, \tilde{X}_k, 1}(t_{k+1})}(T) \\ &\quad + \bar{Z}_{t_{k+1}, \tilde{X}_{t_k, \tilde{X}_k}(t_{k+1}), \bar{Y}_{t_k, \tilde{X}_k, 1}(t_{k+1}), \bar{Z}_{t_k, \tilde{X}_k, 1, 0}(t_{k+1})}(T)].\end{aligned}$$

Therefore due to (3.20), we get

$$\begin{aligned}\bar{v}(t_k, \tilde{X}_k) &= E^{w, \tilde{X}_k}[\{\varphi(\bar{X}_{t_{k+1}, \tilde{X}_{t_k, \tilde{X}_k}(t_{k+1})}(T))\bar{Y}_{t_{k+1}, \tilde{X}_{t_k, \tilde{X}_k}(t_{k+1}), 1}(T) \\ &\quad + \bar{Z}_{t_{k+1}, \tilde{X}_{t_k, \tilde{X}_k}(t_{k+1}), 1, 0}(T)\} \cdot \bar{Y}_{t_k, \tilde{X}_k, 1}(t_{k+1}) + \bar{Z}_{t_k, \tilde{X}_k, 1, 0}(t_{k+1})] \\ &= E^{w, \tilde{X}_k}[\bar{v}(t_{k+1}, \bar{X}_{t_k, \tilde{X}_k}(t_{k+1})) \cdot \bar{Y}_{t_k, \tilde{X}_k, 1}(t_{k+1}) + \bar{Z}_{t_k, \tilde{X}_k, 1, 0}(t_{k+1})].\end{aligned}$$

Substituting this $\bar{v}(t_k, \tilde{X}_k)$ in $E^w[\bar{v}(t_k, \tilde{X}_k)\tilde{Y}_k + \tilde{Z}_k]$ and again using (3.20), we obtain (3.19). The equality (3.18) is proved analogously. So, the presentation (3.15) is proved.

Our next step consists in estimating the smallness of the one-step error. It is easy to see that $\tilde{Z}_{i+1} = \bar{Z}_{t_i, \tilde{X}_i, \tilde{Y}_i, \tilde{Z}_i}(t_{i+1})$ and $\tilde{Y}_{i+1} = \bar{Y}_{t_i, \tilde{X}_i, \tilde{Y}_i}(t_{i+1})$. Then

$$\rho_i(\tilde{X}_i, \tilde{Y}_i, \tilde{Z}_i) = \rho_i(\tilde{X}_i, \tilde{Y}_i) = E^{w, \tilde{X}_i, \tilde{Y}_i} \tilde{Y}_{i+1} [\bar{v}(t_{i+1}, \tilde{X}_{i+1}) - \bar{v}(t_{i+1}, \bar{X}_{t_i, \tilde{X}_i}(t_{i+1}))].$$

Now we write the Taylor expansion of $\bar{v}(t_{i+1}, \tilde{X}_{i+1})$ with respect to powers of $\tilde{\Delta}_i := \tilde{X}_{i+1} - \tilde{X}_i$ in a neighborhood of $\tilde{X}_i$ and with the Lagrange remainder term containing terms of order four. We similarly expand $\bar{v}(t_{i+1}, \bar{X}_{t_i, \tilde{X}_i}(t_{i+1}))$ with respect to $\bar{\Delta}_i := \bar{X}_{t_i, \tilde{X}_i}(t_{i+1}) - \tilde{X}_i$. As a result, we get

$$\begin{aligned}\rho_i(\tilde{X}_i, \tilde{Y}_i) &= E^{w, \tilde{X}_i, \tilde{Y}_i} \tilde{Y}_{i+1} \left[\frac{1}{24} \sum_{j_1, j_2, j_3, j_4=1}^d \left\{ \frac{\partial^4 v}{\partial x^{j_1} \partial x^{j_2} \partial x^{j_3} \partial x^{j_4}}(\tilde{\theta}) \times \tilde{\Delta}_i^{j_1} \tilde{\Delta}_i^{j_2} \tilde{\Delta}_i^{j_3} \tilde{\Delta}_i^{j_4} \right. \right. \\ &\quad \left. \left. - \frac{\partial^4 v}{\partial x^{j_1} \partial x^{j_2} \partial x^{j_3} \partial x^{j_4}}(\bar{\theta}) \times \bar{\Delta}_i^{j_1} \bar{\Delta}_i^{j_2} \bar{\Delta}_i^{j_3} \bar{\Delta}_i^{j_4} \right\} \right],\end{aligned}$$

where $\tilde{\theta}$ is a point between $\tilde{X}_{i+1}$ and $\tilde{X}_i$, and $\bar{\theta}$ is a point between $\bar{X}_{t_i, \tilde{X}_i}(t_{i+1})$ and $\tilde{X}_i$. Then it is not difficult to obtain that for $p \geq 1$,

$$(3.21) \quad E|\rho_i(\tilde{X}_i, \tilde{Y}_i, \tilde{Z}_i)|^{2p} = O(h^{4p}).$$

Due to (3.15) and (3.21), we get

$$\begin{aligned}
 (3.22) \quad ED_N^{2p} &= E \left[\sum_{i=0}^{N-1} \rho_i(\tilde{X}_i, \tilde{Y}_i, \tilde{Z}_i) \right]^{2p} \\
 &\leq N^{2p-1} \sum_{i=0}^{N-1} E \left| \rho_i(\tilde{X}_i, \tilde{Y}_i, \tilde{Z}_i) \right|^{2p} \leq C(x) h^{2p},
 \end{aligned}$$

which implies (3.13) (and consequently (3.11)). The inequality (3.12) is proved exactly as its analogue in Theorem 3.1. $\square$

Remark 3.1. In some particular cases of the SPDE (2.1)-(2.2) the order of a.s. convergence of the Euler schemes (3.2), (3.3) and (3.7), (3.10) is higher. For instance, it is not difficult to modify the proofs of Theorems 3.1 and 3.2 to get that the Euler schemes have a.s. order of convergence $1 - \varepsilon$ if a and α_r are constant and $\beta_r = 0$ and $\gamma_r = 0$ (note that in this case (2.7) is a system with additive noise for which the standard Euler scheme for SDEs is of mean-square order 1 [23]). Furthermore, in [25] it is proved that the schemes considered there are also of a.s. order $1 - \varepsilon$ when $\alpha_r = 0$, $\gamma_r = 0$, $c = 0$, $f = 0$, $\beta_r \neq 0$.

Remark 3.2. Using the weak-sense numerical integration of SDEs in bounded domains (see [23, Chap. 6] and the references therein), the proposed approach can be carried over to the boundary value problems for SPDEs.

4. OTHER PROBABILISTIC REPRESENTATIONS, MONTE CARLO ERROR, AND VARIANCE REDUCTION

Using the Monte Carlo technique, we approximate the solution of the backward SPDE (2.1)-(2.2) as (see (2.6)):

$$\begin{aligned}
 (4.1) \quad v(t, x) &:= v(t, x; \omega) = E^w [\varphi(X_{t,x}(T)) Y_{t,x,1}(T) + Z_{t,x,1,0}(T)] \\
 &\approx \bar{v}(t, x) := E^w [\varphi(X_N) Y_N + Z_N] \\
 &\approx \hat{v}(t, x) := \frac{1}{M} \sum_{m=1}^M [\varphi({}_m X_N) {}_m Y_N + {}_m Z_N],
 \end{aligned}$$

where the first approximate equality involves an error due to replacing X , Y , Z by X_N , Y_N , Z_N (the error of numerical integration of (2.7) by (3.7) or (3.2)) and the error in the second approximate equality comes from the Monte Carlo technique; ${}_m X_N$, ${}_m Y_N$, ${}_m Z_N$, $m = 1, \dots, M$, are independent realizations of X_N , Y_N , Z_N .

The error of numerical integration was analyzed in the previous section. Now let us consider the Monte Carlo error. The error of the Monte Carlo method in (4.1) is evaluated by

$$\bar{\rho} = c \frac{[Var^w \{ \varphi(\bar{X}_{t,x}(T)) \bar{Y}_{t,x,1}(T) + \bar{Z}_{t,x,1,0}(T) \}]^{1/2}}{M^{1/2}},$$

where, e.g., the values $c = 1, 2, 3$ correspond to the fiducial probabilities 0.68, 0.95, 0.997, respectively. Since $Var^w \{ \varphi(\bar{X}_{t,x}(T)) \bar{Y}_{t,x,1}(T) + \bar{Z}_{t,x,1,0}(T) \}$ is close to the variance

$$V := Var^w [\varphi(X_{t,x}(T)) Y_{t,x,1}(T) + Z_{t,x,1,0}(T)],$$

we can assume that the error of the Monte Carlo method is estimated by

$$\rho = c \frac{V^{1/2}}{M^{1/2}}.$$

If V is large, then we have to simulate a very large number of trajectories to achieve a satisfactory accuracy. Fortunately, there exist other probabilistic representations for $v(t, x)$ which allow us to reduce the variance.

The solution of the problem (2.1)-(2.2) also has the following probabilistic representations:

$$(4.2) \quad v(t, x) = E^w [\varphi(X_{t,x}(T))Y_{t,x,1}(T) + Z_{t,x,1,0}(T)], \quad t_0 \leq t \leq T,$$

where $X_{t,x}(s)$, $Y_{t,x,y}(s)$, $Z_{t,x,y,z}(s)$, $t \leq s \leq T$, is the solution of the SDEs

$$(4.3) \quad \begin{aligned} dX &= \left[b(s, X) - \sum_{r=1}^q \alpha_r(s, X)\beta_r(s, X) - \sum_{r=1}^p \sigma_r(s, X)\mu_r(s, X) \right] ds \\ &\quad + \sum_{r=1}^p \sigma_r(s, X)dW_r(s) + \sum_{r=1}^q \alpha_r(s, X)dw_r(s), \quad X(t) = x, \\ dY &= c(s, X)Yds + \sum_{r=1}^p \mu_r(s, X)YdW_r(s) + \sum_{r=1}^q \beta_r(s, X)Ydw_r(s), \quad Y(t) = y, \\ dZ &= f(s, X)Yds + \sum_{r=1}^p \lambda_r(s, X)YdW_r(s) + \sum_{r=1}^q \gamma_r(s, X)Ydw_r(s), \quad Z(t) = z, \end{aligned}$$

and where $\mu = (\mu_1, \dots, \mu_p)^\top$ and $\lambda = (\lambda_1, \dots, \lambda_p)^\top$ are arbitrary p -dimensional vector functions satisfying some regularity assumptions. When $\mu = 0$ and $\lambda = 0$, we have the usual representation (2.6)-(2.7). The representation for $\mu = 0$, $\lambda \neq 0$ follows from the equality

$$E^w \int_t^T \lambda^\top(s, X(s))Y(s)dW(s) = 0.$$

For $\mu \neq 0$, it can be proved by arguments similar to those in [16, pp. 308-309], making use of Theorem 4.4.5 [16, p. 152] for the forward flow. We should note that X , Y , Z in (4.2)-(4.3) differ from X , Y , Z in (2.6)-(2.7); however, this does not lead to any ambiguity. While $v(t, x)$ does not depend on the choice of μ and λ , the variance

$$V = \text{Var}^w [\varphi(X_{t,x}(T))Y_{t,x,1}(T) + Z_{t,x,1,0}(T)]$$

does. Then one may hope to find μ and λ such that the variance V is relatively low and thus the Monte Carlo error can be reduced. Theorem 4.1 is helpful in this respect. Its use for variance reduction gives the method of importance sampling if $\lambda = 0$ and the method of control variates if $\mu = 0$. The form of Theorem 4.1 is very similar to the corresponding well-known result in the case of deterministic PDEs (see [20, 27, 22] and also [23, Section 2.4]). However, its proof is much more complicated in the case of SPDEs than for PDEs as can be seen below.

Theorem 4.1. *Let Assumptions 2.1-2.3 hold. Let μ_r and λ_r be such that for any $x \in \mathbf{R}^d$ there exists a solution to the system (4.3) on the interval $[t, T]$. Then*

$$(4.4) \quad \begin{aligned} &\text{Var}^w [\varphi(X_{t,x}(T))Y_{t,x,1}(T) + Z_{t,x,1,0}(T)] \\ &= E^w \int_t^T Y_{t,x,1}^2 \sum_{r=1}^p \left[\sum_{i=1}^d \sigma_r^i \frac{\partial v}{\partial x^i} + \mu_r v + \lambda_r \right]^2 d\theta \end{aligned}$$

provided that the expectation in (4.4) exists. In (4.4) all the functions in the integrand have $(\theta, X_{t,x}(\theta))$ as their argument.

Proof. It is convenient to introduce the notation

$$E_s^w(\cdot) := E(\cdot/w(s') - w(s), s \leq s' \leq T),$$

$$(4.5) \quad v(s, x) = v(s, x; \omega) := E_s^w[\varphi(X_{s,x}(T))Y_{s,x,1}(T) + Z_{s,x,1,0}(T)] \\ := E[\varphi(X_{s,x}(T))Y_{s,x,1}(T) + Z_{s,x,1,0}(T)/w(s') - w(s), s \leq s' \leq T].$$

Clearly, $E_t^w(\cdot) = E^w(\cdot)$. The notation $Var_s^w(\cdot)$ can be introduced analogously.

We get from (4.5):

$$(4.6) \quad v(s, X_{t,x}(s)) = E[\varphi(X_{s,X_{t,x}(s)}(T))Y_{s,X_{t,x}(s),1}(T) + Z_{s,X_{t,x}(s),1,0}(T)/ \\ W(s') - W(t), t \leq s' \leq s; w(s') - w(t), t \leq s' \leq T].$$

The presentation (4.6) follows from the assertion: let $\tilde{\mathcal{F}}$ and $\mathcal{F}'$ be independent σ -algebras and $\tilde{\mathcal{F}} \vee \mathcal{F}'$ be the minimal σ -algebra generated by $\tilde{\mathcal{F}}$ and $\mathcal{F}'$; if ξ is $\tilde{\mathcal{F}}$ -measurable, $f(x; \omega)$ does not depend on $\tilde{\mathcal{F}}$, and $E(f(x; \omega)/\mathcal{F}') = \psi(x; \omega)$, then $E(f(\xi; \omega)/\tilde{\mathcal{F}} \vee \mathcal{F}') = \psi(\xi; \omega)$ (for the trivial $\mathcal{F}'$ this assertion can be found, e.g., in [6, §10, Lemma 1], [4, Proposition 1.12], which proof can be straightforwardly generalized for a general $\mathcal{F}'$). In the case of (4.6) we have $\xi = X_{t,x}(s)$, $\tilde{\mathcal{F}}$ is the minimal σ -algebra generated by $\{W(s') - W(t), t \leq s' \leq s; w(s') - w(t), t \leq s' \leq s\}$, $\mathcal{F}'$ is the minimal σ -algebra generated by $\{w(s') - w(s), s \leq s' \leq T\}$, and $f(x; \omega) = v(s, x; \omega)$.

Furthermore, we obtain

$$E[v(s, X_{t,x}(s))Y_{t,x,1}(s) + Z_{t,x,1,0}(s)/w(s') - w(t), t \leq s' \leq T] \\ = E[E(\varphi(X_{s,X_{t,x}(s)}(T))Y_{s,X_{t,x}(s),1}(T) + Z_{s,X_{t,x}(s),1,0}(T)/ \\ W(s') - W(t), t \leq s' \leq s; w(s') - w(t), t \leq s' \leq T) \\ \times Y_{t,x,1}(s) + Z_{t,x,1,0}(s)/w(s') - w(t), t \leq s' \leq T] \\ = E[E(\varphi(X_{t,x}(T))Y_{t,x,1}(T) + Z_{t,x,1,0}(T)/W(s') - W(t), \\ t \leq s' \leq s; w(s') - w(t), t \leq s' \leq T)/w(s') - w(t), t \leq s' \leq T] \\ = E[\varphi(X_{t,x}(T))Y_{t,x,1}(T) + Z_{t,x,1,0}(T)/w(s') - w(t), t \leq s' \leq T] = v(t, x).$$

Thus, we prove that for any $s, t \leq s \leq T$, the following formula holds:

$$(4.7) \quad v(t, x) = E_t^w[v(s, X_{t,x}(s))Y_{t,x,1}(s) + Z_{t,x,1,0}(s)], \quad t \leq s.$$

Introduce the auxiliary function:

$$U(t, x, y, z) := v(t, x)y + z.$$

Partition the interval $[t, T]$ with the time step h and present the variance V in the following way (see a similar recipe in [29]):

$$V = Var_t^w[\varphi(X_{t,x}(T))Y_{t,x,1}(T) + Z_{t,x,1,0}(T)] \\ = E_t^w[U(T, X_{t,x}(T), Y_{t,x,1}(T), Z_{t,x,1,0}(T)) - U(t, x, 1, 0)]^2 \\ = E_t^w \left[\sum_{k=0}^{N-1} \{U(t_{k+1}, X_{t,x}(t_{k+1}), Y_{t,x,1}(t_{k+1}), Z_{t,x,1,0}(t_{k+1})) \right. \\ \left. - U(t_k, X_{t,x}(t_k), Y_{t,x,1}(t_k), Z_{t,x,1,0}(t_k))\} \right]^2.$$

Using (4.7), we get

$$\begin{aligned}
& E(U(t_{n+1}, X_{t,x}(t_{n+1}), Y_{t,x,1}(t_{n+1}), Z_{t,x,1,0}(t_{n+1}))/w(s') - w(t), t \leq s' \leq T; \\
& \quad W(s'') - W(t), t \leq s'' \leq t_n) \\
& \quad = Y_{t,x,1}(t_n) \\
& \times E(U(t_{n+1}, X_{t_n, X_{t,x}(t_n)}(t_{n+1}), Y_{t_n, X_{t,x}(t_n), 1}(t_{n+1}), Z_{t_n, X_{t,x}(t_n), 1, 0}(t_{n+1})) \\
& \quad /w(s') - w(t), t \leq s' \leq T; W(s'') - W(t), t \leq s'' \leq t_n) + Z_{t,x,1,0}(t_n) \\
& \quad = Y_{t,x,1}(t_n)v(t_n, X_{t,x}(t_n)) + Z_{t,x,1,0}(t_n) \\
& \quad = U(t_n, X_{t,x}(t_n), Y_{t,x,1}(t_n), Z_{t,x,1,0}(t_n)),
\end{aligned}$$

whence for $k < n$,

$$\begin{aligned}
& E_t^w([U(t_{k+1}, X_{t,x}(t_{k+1}), Y_{t,x,1}(t_{k+1}), Z_{t,x,1,0}(t_{k+1})) \\
& \quad - U(t_k, X_{t,x}(t_k), Y_{t,x,1}(t_k), Z_{t,x,1,0}(t_k))] \\
& \quad \times [U(t_{n+1}, X_{t,x}(t_{n+1}), Y_{t,x,1}(t_{n+1}), Z_{t,x,1,0}(t_{n+1})) \\
& \quad - U(t_n, X_{t,x}(t_n), Y_{t,x,1}(t_n), Z_{t,x,1,0}(t_n))]) \\
& \quad = E_t^w([U(t_{k+1}, X_{t,x}(t_{k+1}), Y_{t,x,1}(t_{k+1}), Z_{t,x,1,0}(t_{k+1})) \\
& \quad - U(t_k, X_{t,x}(t_k), Y_{t,x,1}(t_k), Z_{t,x,1,0}(t_k))] \\
& \quad E(U(t_{n+1}, X_{t,x}(t_{n+1}), Y_{t,x,1}(t_{n+1}), Z_{t,x,1,0}(t_{n+1})) \\
& \quad - U(t_n, X_{t,x}(t_n), Y_{t,x,1}(t_n), Z_{t,x,1,0}(t_n)) \\
& \quad /w(s') - w(t), t \leq s' \leq T; W(s'') - W(t), t \leq s'' \leq t_n)) = 0.
\end{aligned}$$

Then

$$\begin{aligned}
(4.8) \quad V &= \sum_{k=0}^{N-1} E_t^w \{U(t_{k+1}, X_{t,x}(t_{k+1}), Y_{t,x,1}(t_{k+1}), Z_{t,x,1,0}(t_{k+1})) \\
& \quad - U(t_k, X_{t,x}(t_k), Y_{t,x,1}(t_k), Z_{t,x,1,0}(t_k))\}^2.
\end{aligned}$$

We rewrite the terms under the expectation in (4.8) as

$$\begin{aligned}
(4.9) \quad & U(t_{k+1}, X_{t,x}(t_{k+1}), Y_{t,x,1}(t_{k+1}), Z_{t,x,1,0}(t_{k+1})) \\
& - U(t_k, X_{t,x}(t_k), Y_{t,x,1}(t_k), Z_{t,x,1,0}(t_k)) \\
& = [U(t_{k+1}, X_{t,x}(t_{k+1}), Y_{t,x,1}(t_{k+1}), Z_{t,x,1,0}(t_{k+1})) \\
& \quad - U(t_{k+1}, X_{t,x}(t_k), Y_{t,x,1}(t_k), Z_{t,x,1,0}(t_k))] \\
& + [U(t_{k+1}, X_{t,x}(t_k), Y_{t,x,1}(t_k), Z_{t,x,1,0}(t_k)) \\
& \quad - U(t_k, X_{t,x}(t_k), Y_{t,x,1}(t_k), Z_{t,x,1,0}(t_k))] \\
& = [U(t_{k+1}, X_{t,x}(t_{k+1}), Y_{t,x,1}(t_{k+1}), Z_{t,x,1,0}(t_{k+1})) \\
& \quad - U(t_{k+1}, X_{t,x}(t_k), Y_{t,x,1}(t_k), Z_{t,x,1,0}(t_k))] \\
& + Y_{t,x,1}(t_k) [v(t_{k+1}, X_{t,x}(t_k)) - v(t_k, X_{t,x}(t_k))].
\end{aligned}$$

Let

$$\Lambda_r v(t, x) := \sum_{i=1}^d \sigma_r^i(t, x) \frac{\partial}{\partial x^i} v(t, x).$$

Applying the Ito formula to the first term on the right-hand side of (4.9), we get

$$\begin{aligned}
 (4.10) \quad & U(t_{k+1}, X_{t,x}(t_{k+1}), Y_{t,x,1}(t_{k+1}), Z_{t,x,1,0}(t_{k+1})) \\
 & - U(t_{k+1}, X_{t,x}(t_k), Y_{t,x,1}(t_k), Z_{t,x,1,0}(t_k)) \\
 & = \int_{t_k}^{t_{k+1}} Y_{t,x,1}(\theta) [\mathcal{L}v(t_{k+1}, X_{t,x}(\theta)) + f(\theta, X_{t,x}(\theta))] d\theta \\
 & + \sum_{r=1}^q \int_{t_k}^{t_{k+1}} Y_{t,x,1}(\theta) [\mathcal{M}_r v(t_{k+1}, X_{t,x}(\theta)) + \gamma_r(\theta, X_{t,x}(\theta))] dw_r(\theta) \\
 & + \sum_{r=1}^p \int_{t_k}^{t_{k+1}} Y_{t,x,1}(\theta) [\Lambda_r v(t_{k+1}, X_{t,x}(\theta)) + \mu_r(\theta, X_{t,x}(\theta))v(t_{k+1}, X_{t,x}(\theta)) \\
 & + \lambda_r(\theta, X_{t,x}(\theta))] dW_r(\theta).
 \end{aligned}$$

The coefficients of the operators $\mathcal{L}$ and $\mathcal{M}_r$ in (4.10) have $(\theta, X_{t,x}(\theta))$ as their argument. We note (see [29]) that the use of the Ito formula here is legitimate since $v(t_{k+1}, x)$ is $\mathcal{F}_T^{t_{k+1}}$ -adapted and independent of $\mathcal{F}_\theta$, $\theta \leq t_{k+1}$.

Further, since $v(t, x)$ satisfies the SPDE (2.1), we can present the second term on the right-hand side of (4.9) as

$$\begin{aligned}
 (4.11) \quad & Y_{t,x,1}(t_k) [v(t_{k+1}, X_{t,x}(t_k)) - v(t_k, X_{t,x}(t_k))] \\
 & = -Y_{t,x,1}(t_k) \left[\int_{t_k}^{t_{k+1}} [(\mathcal{L}v)(\theta, X_{t,x}(t_k)) + f(\theta, X_{t,x}(t_k))] d\theta \right. \\
 & \quad \left. + \sum_{r=1}^q \int_{t_k}^{t_{k+1}} [(\mathcal{M}_r v)(\theta, X_{t,x}(t_k)) + \gamma_r(\theta, X_{t,x}(t_k))] * dw_r(\theta) \right].
 \end{aligned}$$

We note that all the functions in the integrands in (4.11) have $(\theta, X_{t,x}(t_k))$ as their argument.

Using (4.9)-(4.11), properties of Ito integrals, and the independence of w and W , we obtain

$$\begin{aligned}
 (4.12) \quad & E_t^w \{ U(t_{k+1}, X_{t,x}(t_{k+1}), Y_{t,x,1}(t_{k+1}), Z_{t,x,1,0}(t_{k+1})) \\
 & - U(t_k, X_{t,x}(t_k), Y_{t,x,1}(t_k), Z_{t,x,1,0}(t_k)) \}^2 \\
 & = E_t^w \left[\int_{t_k}^{t_{k+1}} \{ Y_{t,x,1}(\theta) [\mathcal{L}v(t_{k+1}, X_{t,x}(\theta)) + f(\theta, X_{t,x}(\theta))] \right. \\
 & \quad - Y_{t,x,1}(t_k) [(\mathcal{L}v)(\theta, X_{t,x}(t_k)) + f(\theta, X_{t,x}(t_k))] \} d\theta \\
 & \quad + \sum_{r=1}^q \left\{ \int_{t_k}^{t_{k+1}} Y_{t,x,1}(\theta) [\mathcal{M}_r v(t_{k+1}, X_{t,x}(\theta)) + \gamma_r(\theta, X_{t,x}(\theta))] dw_r(\theta) \right. \\
 & \quad \left. \left. - Y_{t,x,1}(t_k) \int_{t_k}^{t_{k+1}} [(\mathcal{M}_r v)(\theta, X_{t,x}(t_k)) + \gamma_r(\theta, X_{t,x}(t_k))] * dw_r(\theta) \right\}^2 \right] \\
 & + 2E_t^w \sum_{r=1}^p \int_{t_k}^{t_{k+1}} Y_{t,x,1}(\theta) \\
 & \times [\Lambda_r v(t_{k+1}, X_{t,x}(\theta)) + \mu_r(\theta, X_{t,x}(\theta))v(t_{k+1}, X_{t,x}(\theta)) + \lambda_r(\theta, X_{t,x}(\theta))] dW_r(\theta)
 \end{aligned}$$

$$\begin{aligned}
& \times \left[\int_{t_k}^{t_{k+1}} Y_{t,x,1}(\theta) [\mathcal{L}v(t_{k+1}, X_{t,x}(\theta)) + f(\theta, X_{t,x}(\theta))] d\theta \right. \\
& \quad \left. + \sum_{r=1}^q \int_{t_k}^{t_{k+1}} Y_{t,x,1}(\theta) [\mathcal{M}_r v(t_{k+1}, X_{t,x}(\theta)) + \gamma_r(\theta, X_{t,x}(\theta))] dw_r(\theta) \right] \\
& \quad + \sum_{r=1}^p E_t^w \int_{t_k}^{t_{k+1}} Y_{t,x,1}^2(\theta) [\Lambda_r v(t_{k+1}, X_{t,x}(\theta)) + \mu_r(\theta, X_{t,x}(\theta))v(t_{k+1}, X_{t,x}(\theta)) \\
& \quad \quad \quad + \lambda_r(\theta, X_{t,x}(\theta))]^2 d\theta.
\end{aligned}$$

In this proof let us use the notation $O(h^q)$ for random variables which tend to zero at least as h^q a.s. when $h \rightarrow 0$. We have for any $0 < \varepsilon < 1/2$,

$$w_r(s+h) - w_r(s) = O(h^{1/2-\varepsilon}), \quad W_r(s+h) - W_r(s) = O(h^{1/2-\varepsilon}) \quad a.s.,$$

and also

$$v(s+h, x) - v(s, x) = O(h^{1/2-\varepsilon}) \quad a.s.$$

One can show that for $0 < \varepsilon < 1/2$,

$$\begin{aligned}
& \int_{t_k}^{t_{k+1}} Y_{t,x,1}(\theta) [\mathcal{L}v(t_{k+1}, X_{t,x}(\theta)) + f(\theta, X_{t,x}(\theta))] \\
& \quad - Y_{t,x,1}(t_k) [(\mathcal{L}v)(\theta, X_{t,x}(t_k)) + f(\theta, X_{t,x}(t_k))] d\theta = O(h), \\
& \int_{t_k}^{t_{k+1}} Y_{t,x,1}(\theta) [\mathcal{M}_r v(t_{k+1}, X_{t,x}(\theta)) + \gamma_r(\theta, X_{t,x}(\theta))] dw_r(\theta) \\
& \quad = Y_{t,x,1}(t_k) \int_{t_k}^{t_{k+1}} [(\mathcal{M}_r v)(t_{k+1}, X_{t,x}(t_k)) + \gamma_r(\theta, X_{t,x}(t_k))] dw_r(\theta) + O(h^{1-\varepsilon}) \\
& \quad = Y_{t,x,1}(t_k) \int_{t_k}^{t_{k+1}} [(\mathcal{M}_r v)(t_{k+1}, X_{t,x}(t_k)) + \gamma_r(\theta, X_{t,x}(t_k))] * dw_r(\theta) + O(h^{1-\varepsilon}) \\
& \quad = Y_{t,x,1}(t_k) \int_{t_k}^{t_{k+1}} [(\mathcal{M}_r v)(\theta, X_{t,x}(t_k)) + \gamma_r(\theta, X_{t,x}(t_k))] * dw_r(\theta) + O(h^{1-\varepsilon}), \\
& E_t^w \int_{t_k}^{t_{k+1}} Y_{t,x,1}(\theta) [\Lambda_r v(t_{k+1}, X_{t,x}(\theta)) + \mu_r(\theta, X_{t,x}(\theta))v(t_{k+1}, X_{t,x}(\theta)) \\
& \quad \quad \quad + \lambda_r(\theta, X_{t,x}(\theta))] dW_r(\theta) \\
& \quad \times \int_{t_k}^{t_{k+1}} Y_{t,x,1}(\theta) [\mathcal{L}v(t_{k+1}, X_{t,x}(\theta)) + f(\theta, X_{t,x}(\theta))] d\theta \\
& \quad = O(h^{3/2-\varepsilon}), \\
& E_t^w \int_{t_k}^{t_{k+1}} Y_{t,x,1}(\theta) [\Lambda_r v(t_{k+1}, X_{t,x}(\theta)) + \mu_r(\theta, X_{t,x}(\theta))v(t_{k+1}, X_{t,x}(\theta)) \\
& \quad + \lambda_r(\theta, X_{t,x}(\theta))] dW_r(\theta) \\
& \quad \times \int_{t_k}^{t_{k+1}} Y_{t,x,1}(\theta) [\mathcal{M}_r v(t_{k+1}, X_{t,x}(\theta)) + \gamma_r(\theta, X_{t,x}(\theta))] dw_r(\theta) \\
& \quad = E_t^w \int_{t_k}^{t_{k+1}} Y_{t,x,1}(\theta) [\Lambda_r v(t_{k+1}, X_{t,x}(\theta)) + \mu_r(\theta, X_{t,x}(\theta))v(t_{k+1}, X_{t,x}(\theta)) \\
& \quad + \lambda_r(\theta, X_{t,x}(\theta))] dW_r(\theta)
\end{aligned}$$

$$\begin{aligned}
& \times \int_{t_k}^{t_{k+1}} Y_{t,x,1}(t_k) [(\mathcal{M}_r v)(t_{k+1}, X_{t,x}(t_k)) + \gamma_r(\theta, X_{t,x}(t_k))] dw_r(\theta) \\
& + O(h^{3/2-\varepsilon}) = O(h^{3/2-\varepsilon}), \\
& E_t^w \int_{t_k}^{t_{k+1}} Y_{t,x,1}^2(\theta) [\Lambda_r v(t_{k+1}, X_{t,x}(\theta)) + \mu_r(\theta, X_{t,x}(\theta)) v(t_{k+1}, X_{t,x}(\theta)) \\
& \quad + \lambda_r(\theta, X_{t,x}(\theta))]^2 d\theta \\
& = E_t^w \int_{t_k}^{t_{k+1}} Y_{t,x,1}^2(\theta) [(\Lambda_r v)(\theta, X_{t,x}(\theta)) + \mu_r(\theta, X_{t,x}(\theta)) v(\theta, X_{t,x}(\theta)) \\
& \quad + \lambda_r(\theta, X_{t,x}(\theta))]^2 d\theta + O(h^{3/2-\varepsilon}).
\end{aligned}$$

Substituting the above relations in the right-hand side of (4.12), we obtain

$$\begin{aligned}
& E_t^w \{U(t_{k+1}, X_{t,x}(t_{k+1}), Y_{t,x,1}(t_{k+1}), Z_{t,x,1,0}(t_{k+1})) \\
& \quad - U(t_k, X_{t,x}(t_k), Y_{t,x,1}(t_k), Z_{t,x,1,0}(t_k))\}^2 \\
& = \sum_{r=1}^p E_t^w \int_{t_k}^{t_{k+1}} Y_{t,x,1}^2(\theta) [(\Lambda_r v)(\theta, X_{t,x}(\theta)) + \mu_r(\theta, X_{t,x}(\theta)) v(\theta, X_{t,x}(\theta)) \\
& \quad + \lambda_r(\theta, X_{t,x}(\theta))]^2 d\theta + O(h^{3/2-\varepsilon}),
\end{aligned}$$

which substitution in (4.8) gives

$$\begin{aligned}
V & = \sum_{r=1}^p E_t^w \int_t^T Y_{t,x,1}^2(\theta) [(\Lambda_r v)(\theta, X_{t,x}(\theta)) + \mu_r(\theta, X_{t,x}(\theta)) v(\theta, X_{t,x}(\theta)) \\
& \quad + \lambda_r(\theta, X_{t,x}(\theta))]^2 d\theta + \sum_{k=0}^{N-1} O(h^{3/2-\varepsilon}),
\end{aligned}$$

and tending h to 0 we arrive at (4.4). $\square$

Theorem 4.1 implies that if μ_r and λ_r are such that

$$(4.13) \quad \sum_{i=1}^d \sigma_r^i \frac{\partial v}{\partial x^i} + \mu_r v + \lambda_r = 0, \quad r = 1, \dots, p,$$

then the right-hand side of (4.4) is zero and, consequently, the variance is zero. We recall that $v(s, x)$ depends on $w(\theta) - w(s)$, $s \leq \theta \leq T$. However, we need to require that $\mu_r(s, x)$ does not depend on $w(\theta) - w(s)$, $s \leq \theta \leq T$; otherwise, the coefficients in (4.3) depend on $w(\theta) - w(s)$, $s \leq \theta \leq T$, and we are facing the difficulty with interpreting (4.3) except the case when α_r is independent of x and $\beta_r = \gamma_r = 0$. In the case mentioned, all the integrals appearing in the integral form of (4.3) can be given the usual meaning in the sense of Ito calculus. At the same time, dependence of the function $\lambda_r(s, x)$ on $w(\theta) - w(s)$, $s \leq \theta \leq T$, does not cause any trouble in interpreting the third equation in (4.3), and the identity (4.13) can, in principle, be reached. We note that Theorem 4.1 and its proof are valid for $\lambda_r(s, x)$ depending on $w(\theta) - w(s)$, $s \leq \theta \leq T$. Furthermore, when (4.3) is discretely simulated, one can simulate the process $w(s)$ in advance and fix it. Then it is possible to get approximate solutions of (4.3) for μ_r depending on a current position and on the fixed $w(\theta) - w(s)$, $s \leq \theta \leq T$. Thus we obtain a heuristic interpretation of (4.3) for

μ_r depending on the future of w and, apparently, the result of Theorem 4.1 can be used for such μ_r as well.

Of course, μ_r and λ_r satisfying (4.13) cannot be constructed without knowing $v(s, x)$, $t \leq s \leq T$, $x \in \mathbf{R}^d$. Nevertheless, Theorem 4.1 claims a general possibility of variance reduction by a proper choice of the functions μ_r and λ_r . Theorem 4.1 can be used, for example, if we know a function $\tilde{v}(s, x)$ being close to $v(s, x)$ (see a practical approach to constructing such approximate functions in the case of deterministic PDEs in [26]). Then we take any μ_r and λ_r satisfying

$$(4.14) \quad \sum_{i=1}^d \sigma_r^i \frac{\partial \tilde{v}}{\partial x^i} + \mu_r \tilde{v} + \lambda_r = 0,$$

and we expect that the variance $\text{Var}^w [\varphi(X_{t,x}(T))Y_{t,x,1}(T) + Z_{t,x,1,0}(T)]$ is small although not zero.

5. LAYER METHODS FOR LINEAR SPDES

In the previous sections we use the averaging over the characteristic formula to propose Monte Carlo methods for linear SPDEs. In this section we exploit probabilistic representations to construct some layer methods for linear SPDEs (see them in the case of deterministic PDEs in [21, 23]). Layer methods for semilinear SPDEs are considered in the next section. The layer methods are competitive with finite difference schemes [10, 34]. They can be used for relatively low-dimensional SPDEs when one needs to find the solution $v(t, x)$ everywhere in $[t_0, T] \times \mathbf{R}^d$. The bottleneck for using the layer methods in the case of higher-dimensional SPDEs is interpolation, which should be used to restore the function $\bar{v}(t_k, x)$, $x \in \mathbf{R}^d$, by its values $\bar{v}(t_k, x_j)$ at a set of nodes x_j . The linear interpolation, which is successfully applied for low-dimensional problems, is not effective for the higher-dimensional ones. However, layer methods can be turned into numerical algorithms by exploiting other approximations of functions. New developments in the theory of multidimensional approximation (see, e.g., [33] and the references therein) give approximations which can work well in relatively high dimensions.

The method (3.7), (3.10) can be turned into a layer method. Indeed, analogously to the probabilistic representation (2.6)-(2.7) one can write down the local probabilistic representation of the solution to (2.1)-(2.2) (see (4.7)):

$$(5.1) \quad v(t_k, x) = E^w [v(t_{k+1}, X_{t_k,x}(t_{k+1}))Y_{t_k,x,1}(t_{k+1}) + Z_{t_k,x,1,0}(t_{k+1})],$$

where $X_{t,x}(s)$, $Y_{t,x,y}(s)$, $Z_{t,x,y,z}(s)$, $t \leq s \leq T$, is the solution of the SDEs (2.7). Replacing $X_{t_k,x}(t_{k+1})$, $Y_{t_k,x,1}(t_{k+1})$, $Z_{t_k,x,1,0}(t_{k+1})$, by the one-step Euler approximation $\bar{X}_{t_k,x}(t_{k+1})$, $\bar{Y}_{t_k,x,1}(t_{k+1})$, $\bar{Z}_{t_k,x,1,0}(t_{k+1})$ as in (3.7), (3.9), we obtain the one-step approximation $\check{v}(t_k, x)$ for $v(t_k, x)$:

$$(5.2) \quad \check{v}(t_k, x) = E^w [v(t_{k+1}, \bar{X}_{t_k,x}(t_{k+1}))\bar{Y}_{t_k,x,1}(t_{k+1}) + \bar{Z}_{t_k,x,1,0}(t_{k+1})].$$

Now for simplicity in writing let us consider the case $d = 1$ and $q = 1$. Thanks to the use of the discrete random variables (3.9), the expectation in (5.2) can be evaluated exactly:

(5.3)

$$\begin{aligned} \tilde{v}(t_k, x) = & \left\{ \frac{1}{2} v \left(t_{k+1}, x + h [b(t_k, x) - \alpha(t_k, x) \beta(t_k, x)] + h^{1/2} \sigma(t_k, x) + \alpha(t_k, x) \Delta_k w \right) \right. \\ & \left. + \frac{1}{2} v \left(t_{k+1}, x + h [b(t_k, x) - \alpha(t_k, x) \beta(t_k, x)] - h^{1/2} \sigma(t_k, x) + \alpha(t_k, x) \Delta_k w \right) \right\} \\ & \times [1 + hc(t_k, x) + \beta(t_k, x) \Delta_k w] + hf(t_k, x) + \gamma(t_k, x) \Delta_k w. \end{aligned}$$

Based on the one-step approximation (5.3), we obtain the layer method for (2.1)-(2.2) with $d = 1$, $q = 1$:

(5.4)

$$\begin{aligned} \tilde{v}(t_N, x) &= \varphi(x), \\ \tilde{v}(t_k, x) &= \left\{ \frac{1}{2} \tilde{v} \left(t_{k+1}, x + h [b(t_k, x) - \alpha(t_k, x) \beta(t_k, x)] + h^{1/2} \sigma(t_k, x) + \alpha(t_k, x) \Delta_k w \right) \right. \\ &\quad \left. + \frac{1}{2} \tilde{v} \left(t_{k+1}, x + h [b(t_k, x) - \alpha(t_k, x) \beta(t_k, x)] - h^{1/2} \sigma(t_k, x) + \alpha(t_k, x) \Delta_k w \right) \right\} \\ &\quad \times [1 + hc(t_k, x) + \beta(t_k, x) \Delta_k w] + hf(t_k, x) + \gamma(t_k, x) \Delta_k w, \quad k = N-1, \dots, 0. \end{aligned}$$

It is not difficult to see that $\tilde{v}(t_k, x)$ from (5.4) coincides with the $\tilde{v}(t_k, x)$ due to (3.10). Then according to Theorem 3.2, the estimates (3.11) and (3.12) are valid for $\tilde{v}(t_k, x)$ from (5.4).

To realize (5.4) numerically, it is sufficient to calculate the functions $\tilde{v}(t_k, x)$ at some knots x_j applying some kind of interpolation at every layer. So, to become a numerical algorithm, the layer method (5.4) needs a discretization in the variable x . Consider the equidistant space discretization:

(5.5)

$$x_j = x_0 + jh_x, \quad j = 0, \pm 1, \pm 2, \dots,$$

where x_0 is a point on $\mathbf{R}$ and $h_x > 0$ is a step of the space discretization. Using, for instance, linear interpolation, we construct the following algorithm on the basis of the layer method (5.4):

(5.6)

$$\begin{aligned} \bar{v}(t_N, x) &= \varphi(x), \\ \bar{v}(t_k, x_j) &= \left\{ \frac{1}{2} \bar{v} \left(t_{k+1}, x_j + h [b(t_k, x_j) - \alpha(t_k, x_j) \beta(t_k, x_j)] \right. \right. \\ &\quad \left. \left. + h^{1/2} \sigma(t_k, x_j) + \alpha(t_k, x_j) \Delta_k w \right) \right. \\ &\quad \left. + \frac{1}{2} \bar{v} \left(t_{k+1}, x_j + h [b(t_k, x_j) - \alpha(t_k, x_j) \beta(t_k, x_j)] - h^{1/2} \sigma(t_k, x_j) + \alpha(t_k, x_j) \Delta_k w \right) \right\} \\ &\quad \times [1 + hc(t_k, x_j) + \beta(t_k, x_j) \Delta_k w] + hf(t_k, x_j) + \gamma(t_k, x_j) \Delta_k w, \quad j = 0, \pm 1, \pm 2, \dots, \\ (5.7) \quad \bar{v}(t_k, x) &= \frac{x_{j+1} - x}{h_x} \bar{v}(t_k, x_j) + \frac{x - x_j}{h_x} \bar{v}(t_k, x_{j+1}), \quad x_j \leq x \leq x_{j+1}, \\ &\quad k = N-1, \dots, 1, 0. \end{aligned}$$

We use the same notation $\bar{v}$ for the two different functions: one is defined by the layer algorithm (5.6)-(5.7) and the other is defined in Section 3.1 for (3.3), but this should not cause any confusion since we do not use the approximation (3.3) in the current section.

If we need the solution of (2.1)-(2.2) for all points (t_k, x_i) , we can use (5.6)-(5.7) to find $\bar{v}(t_k, x_j)$ layerwise. But if we need the solution at a particular point (t_k, x) , the formula (3.10) is more convenient.

Remark 5.1. Thanks to the probabilistic approach, we do not need any stability criteria for the layer methods in comparison with finite-difference schemes, on which the Lax-Richtmyer equivalence theorem imposes a requirement on the relation between the time step h and the space step h_x . E.g., for $b = 0$, $\alpha = 0$, $c = 0$, $\beta = 0$, and σ is a constant, the method (5.4) (we note that it does not need an interpolation in this case) coincides with the stable finite-difference scheme (see [10] and the references therein)

$$\frac{\bar{v}(t_k, x_j) - \bar{v}(t_{k+1}, x_j)}{h} = \frac{\sigma^2}{2} \frac{\bar{v}(t_{k+1}, x_{j+1}) - 2\bar{v}(t_{k+1}, x_j) + \bar{v}(t_{k+1}, x_{j-1}))}{h_x^2} + f(t_k, x_j) + \gamma(t_k, x_j) \frac{\Delta_k w}{h}$$

with $h_x = \sigma\sqrt{h}$. In layer methods the suitable choice of points at which $\bar{v}$ has to be evaluated is achieved automatically by taking into account the coefficient dependence on the space variables and a relationship between various terms (driving noise, diffusion, and advection) in an intrinsic manner.

To prove the next convergence theorem, we need to impose additional assumptions on the considered problem to Assumptions 2.1-2.3 from Section 2.

Assumption 5.1. We assume that the coefficients a , b , α_r are uniformly bounded and that the function φ has bounded derivatives up to a sufficiently high order (i.e., it is globally Lipschitz).

Theorem 5.1. Let Assumptions 2.1, 2.3, and 5.1 hold. The algorithm (5.6)-(5.7) with $h_x = \varkappa h^{3/4}$, $\varkappa > 0$, satisfies the inequality for $p \geq 1$:

$$(5.8) \quad \left(E |\bar{v}(t_k, x) - v(t_k, x)|^{2p} \right)^{1/(2p)} \leq Kh^{1/2},$$

where K does not depend on x , h , k , i.e., in particular, (5.6)-(5.7) is of mean-square order $1/2$.

For almost every trajectory $w(\cdot)$ and any $\varepsilon > 0$ the algorithm (5.6)-(5.7) with $h_x = \varkappa h^{3/4}$ converges with order $1/2 - \varepsilon$, i.e.,

$$(5.9) \quad |v(t_k, x) - \bar{v}(t_k, x)| \leq C(\omega)h^{1/2-\varepsilon} \quad a.s.,$$

where C does not depend on x , h , k .

Proof. We first prove that

$$(5.10) \quad \left(E |\bar{v}(t_k, x_j) - v(t_k, x_j)|^{2p} \right)^{1/(2p)} \leq Kh^{1/2}.$$

In connection with the algorithm (5.6)-(5.7), we introduce the random sequence $\check{X}_i, \check{Y}_i, \check{Z}_i$, $i = k, \dots, N$, in the following way. We fix k and put $\check{X}_k = x_j$, $\check{Y}_k = 1$, $\check{Z}_k = 0$ (to avoid confusion, we note that the index k of $\check{X}_k, \check{Y}_k, \check{Z}_k$ means that $\check{X}_k, \check{Y}_k, \check{Z}_k$ belong to the k^{th} time layer, while the index j of x_j corresponds to the space discretization (5.5)). Knowing $\check{X}_i$, we introduce the auxiliary values $\check{X}_{i+1}^\pm$ at the $(i+1)$ -th step, $i = k, \dots, N-1$:

$$(5.11) \quad \check{X}_{i+1}^\pm := \check{X}_i + h \left[b(t_i, \check{X}_i) - \alpha(t_i, \check{X}_i)\beta(t_i, \check{X}_i) \right] \pm h^{1/2}\sigma(t_i, \check{X}_i) + \alpha(t_i, \check{X}_i)\Delta_i w.$$

Then $\check{X}_{i+1}$ for $i = k, \dots, N-2$ is defined as the random variable distributed according to the law:

$$P(\check{X}_{i+1} = x_l) = \frac{1}{2} \frac{x_{l+1} - \check{X}_{i+1}^-}{h_x}, \quad P(\check{X}_{i+1} = x_{l+1}) = \frac{1}{2} \frac{\check{X}_{i+1}^- - x_l}{h_x},$$

$$P(\check{X}_{i+1} = x_m) = \frac{1}{2} \frac{x_{m+1} - \check{X}_{i+1}^+}{h_x}, \quad P(\check{X}_{i+1} = x_{m+1}) = \frac{1}{2} \frac{\check{X}_{i+1}^+ - x_m}{h_x},$$

where $x_l, x_{l+1}, x_m, x_{m+1}$ are such that $x_l \leq \check{X}_{i+1}^- < x_{l+1}$, $x_m < \check{X}_{i+1}^+ \leq x_{m+1}$; and $\check{X}_N$ is distributed as

$$P(\check{X}_N = \check{X}_N^-) = P(\check{X}_N = \check{X}_N^+) = \frac{1}{2};$$

$\check{Y}_{i+1}$ and $\check{Z}_{i+1}$ for $i = k, \dots, N-1$ are defined in the following way:

$$\check{Y}_{i+1} = \check{Y}_i \left[1 + hc(t_i, \check{X}_i) + \beta(t_i, \check{X}_i) \Delta_i w \right],$$

$$\check{Z}_{i+1} = \check{Z}_i + hf(t_i, \check{X}_i) \check{Y}_i + \gamma(t_i, \check{X}_i) \check{Y}_i \Delta_i w.$$

It can be directly verified that

$$(5.12) \quad \bar{v}(t_k, x_j) = E^w \left[\bar{v}(t_{k+1}, \check{X}_{k+1}) \check{Y}_{k+1} + \check{Z}_{k+1} \right].$$

Using equalities of the type (3.20) and the fact that $\check{X}_{k+1}$ takes the values on the set of knots of the space discretization, we get

$$\bar{v}(t_k, x_j) = E^w \left[\bar{v}(t_{k+2}, \check{X}_{k+2}) \check{Y}_{k+2} + \check{Z}_{k+2} \right].$$

Carrying on in this way, we obtain that $\bar{v}(t_k, x_j)$ admits the following probabilistic representation:

$$(5.13) \quad \bar{v}(t_k, x_j) = E^w \left[\bar{v}(t_N, \check{X}_N) \check{Y}_N + \check{Z}_N \right] = E^w \left[\varphi(\check{X}_N) \check{Y}_N + \check{Z}_N \right].$$

Now we shall act analogously to the proof of Theorem 3.2 and prove the inequality

$$(5.14) \quad \left(E |\bar{v}(t_k, x_j) - \tilde{v}(t_k, x_j)|^{2p} \right)^{1/(2p)} \leq Kh^{1/2}.$$

Then the inequality (5.10) follows from Theorem 3.2 (see (3.11)).

We write the difference $D_N := \bar{v}(t_k, x_j) - \tilde{v}(t_k, x_j)$ in the form (see (3.15)-(3.16)):

$$(5.15) \quad D_N = E^w \left[\varphi(\check{X}_{t_k, x_j}(t_N)) \check{Y}_{t_k, x_j, 1}(t_N) + \check{Z}_{t_k, x_j, 1, 0}(t_N) \right] - E^w \left[\varphi(\check{X}_N) \check{Y}_N + \check{Z}_N \right]$$

$$= E^w \sum_{i=0}^{N-1} \rho_i(\check{X}_i, \check{Y}_i, \check{Z}_i),$$

where

$$(5.16) \quad \rho_i(\check{X}_i, \check{Y}_i, \check{Z}_i) = E^{w, \check{X}_i, \check{Y}_i, \check{Z}_i} [\tilde{v}(t_{i+1}, \check{X}_{i+1}) \check{Y}_{i+1} + \check{Z}_{i+1}$$

$$- \tilde{v}(t_{i+1}, \check{X}_{t_i, \check{X}_i}(t_{i+1})) \check{Y}_{t_i, \check{X}_i, \check{Y}_i}(t_{i+1}) - \check{Z}_{t_i, \check{X}_i, \check{Y}_i, \check{Z}_i}(t_{i+1})].$$

One can prove that the function $\tilde{v}(t_i, x)$ from (3.10) is sufficiently smooth in x and satisfies (together with its derivatives) the same inequality as the function $v(t, x)$ (see the inequality (2.5)).

Our next step is to get an estimate for $\rho_i(\check{X}_i, \check{Y}_i, \check{Z}_i)$. Since $\check{Y}_{i+1} = \tilde{Y}_{t_i, \check{X}_i, \check{Y}_i}(t_{i+1})$, $\check{Z}_{i+1} = \tilde{Z}_{t_i, \check{X}_i, \check{Y}_i, \check{Z}_i}(t_{i+1})$, we have

$$\begin{aligned}\rho_i(\check{X}_i, \check{Y}_i, \check{Z}_i) &= \rho_i(\check{X}_i, \check{Y}_i) = E^{w, \check{X}_i, \check{Y}_i} \check{Y}_{i+1} [\tilde{v}(t_{i+1}, \check{X}_{i+1}) - \tilde{v}(t_{i+1}, \tilde{X}_{t_i, \check{X}_i}(t_{i+1}))] \\ &= \check{Y}_{i+1} E^{w, \check{X}_i, \check{Y}_i} [\tilde{v}(t_{i+1}, \check{X}_{i+1}) - \tilde{v}(t_{i+1}, \tilde{X}_{t_i, \check{X}_i}(t_{i+1}))].\end{aligned}$$

We expand $\tilde{v}(t_{i+1}, \check{X}_{i+1})$ with respect to powers of $\tilde{\Delta}_i := \check{X}_{i+1} - \check{X}_i$ in a neighborhood of $\check{X}_i$ and with the Lagrange remainder term containing terms of order four, and we similarly expand $\tilde{v}(t_{i+1}, \tilde{X}_{t_i, \check{X}_i}(t_{i+1}))$ with respect to the $\tilde{\Delta}_i := \tilde{X}_{t_i, \check{X}_i}(t_{i+1}) - \check{X}_i$. One can verify that for $p \geq 1$:

$$\begin{aligned}E^{w, \check{X}_i, \check{Y}_i} [\tilde{\Delta}_i - \tilde{\Delta}_i] &= E^{w, \check{X}_i, \check{Y}_i} [\check{X}_{i+1} - \tilde{X}_{t_i, \check{X}_i}(t_{i+1})] \\ &= E^{w, \check{X}_i} [\check{X}_{i+1} - \tilde{X}_{t_i, \check{X}_i}(t_{i+1})] = 0, \\ E \left(E^{w, \check{X}_i, \check{Y}_i} [\tilde{\Delta}_i^r - \tilde{\Delta}_i^r] \right)^{2p} &= O(h_x^{4p}), \quad r = 2, 3,\end{aligned}$$

and for $h_x \leq \varkappa\sqrt{h}$,

$$E \left(E^{w, \check{X}_i, \check{Y}_i} [\tilde{\Delta}_i^4] \right)^{2p} = O(h^{4p}), \quad E \left(E^{w, \check{X}_i, \check{Y}_i} [\tilde{\Delta}_i^4] \right)^{2p} = O(h^{4p}).$$

Then, taking into account $h_x = \varkappa h^{3/4}$, it is not difficult to prove that

$$E \left| \rho_i(\check{X}_i, \check{Y}_i, \check{Z}_i) \right|^{2p} \leq K h^{3p}.$$

From here, using the same arguments as in the proof of Theorem 3.2 (see (3.22)), we obtain (5.14) and, consequently, (5.10).

Due to the smoothness of $v(t, x)$ in x , we have

$$v(t_k, x) = \frac{x_{j+1} - x}{h_x} v(t_k, x_j) + \frac{x - x_j}{h_x} v(t_k, x_{j+1}) + O(h_x^2), \quad x_j \leq x \leq x_{j+1},$$

where $E|O(h_x^2)|^{2p} \leq K h_x^{4p}$. Then

$$\begin{aligned}\bar{v}(t_k, x) - v(t_k, x) &= \frac{x_{j+1} - x}{h_x} [\bar{v}(t_k, x_j) - v(t_k, x_j)] \\ &\quad + \frac{x - x_j}{h_x} [\bar{v}(t_k, x_{j+1}) - v(t_k, x_{j+1})] + O(h_x^2),\end{aligned}$$

and, using (5.10) and that $h_x = \varkappa h^{3/4}$, we obtain

$$\begin{aligned}E|\bar{v}(t_k, x) - v(t_k, x)|^{2p} &\leq K \cdot (E|\bar{v}(t_k, x_j) - v(t_k, x_j)|^{2p} \\ &\quad + E|\bar{v}(t_k, x_{j+1}) - v(t_k, x_{j+1})|^{2p}) + K h^{3p} \leq K h^p,\end{aligned}$$

whence (5.8) follows. The inequality (5.9) is proved as its analogue in Theorem 3.1. $\square$

We note that for some particular SPDEs (see Remark 3.1) the algorithm (5.6)-(5.7) with $h_x = \varkappa h$ converges with the weak order $1 - \varepsilon$.

Remark 5.2. The generalization to the multidimensional SPDE, $d > 1$, $q \geq 1$, is straightforward (see it in the case of deterministic PDEs in [21, 23]). We also note that other types of interpolation can be exploited here to construct layer algorithms [21, 23].

6. LAYER METHODS FOR SEMILINEAR SPDES

In this section, we generalize layer methods for the linear SPDEs considered in Section 5 to the case of semilinear SPDEs (see layer methods for semilinear deterministic PDEs in [21, 23] and for quasilinear deterministic PDEs in [24]). As in the previous section, we restrict ourselves to the one-dimensional ($d = 1$) and one noise ($q = 1$) case for simplicity in writing only. There is no difficulty in generalizing the methods of this section to the multi-dimensional multi-noise case (see how this is done in the case of deterministic PDEs in, e.g., [21, 23]).

We consider the Cauchy problem for the backward semilinear SPDE:

$$(6.1) \quad -dv = [Lv + f(t, x, v)] dt + \left[\alpha(t, x, v) \frac{\partial}{\partial x} v + \gamma(t, x, v) \right] * dw(t), \quad (t, x) \in [T_0, T] \times \mathbf{R},$$

$$(6.2) \quad v(T, x) = \varphi(x), \quad x \in \mathbf{R},$$

where

$$(6.3) \quad Lv(t, x) := \frac{1}{2} a(t, x, v(t, x)) \frac{\partial^2}{\partial x^2} v(t, x) + b(t, x, v(t, x)) \frac{\partial}{\partial x} v(t, x), \\ f(t, x, v) := f_0(t, x)v + f_1(t, x, v), \quad \gamma(t, x, v) := \gamma_0(t, x)v + \gamma_1(t, x, v).$$

Assumption 6.1. We assume that the coefficients $a(t, x, v)$, $b(t, x, v)$, $f_0(t, x)$, $f_1(t, x, v)$, $\alpha(t, x, v)$, $\gamma_0(t, x)$, and $\gamma_1(t, x, v)$ and the function $\varphi(x)$ are bounded and have bounded derivatives up to some order and that the coercivity condition is satisfied, i.e., $\sigma^2(t, x, v) := a(t, x, v) - \alpha^2(t, x, v) \geq 0$ for all t, x, v . We also suppose that the problem (6.1)-(6.2) has the classical solution $v(t, x)$, which has derivatives in x up to a sufficiently high order, and that the solution and its spatial derivatives have bounded moments up to some order.

We note that under Assumption 6.1 the existence of the classical solution is proved in the case $a(t, x, v) = a(t, x)$, $b(t, x, v) = b(t, x)$, $\alpha(t, x, v) = \alpha(t, x)$ in [14] and in the case $a(t, x, v) = a(t, x)$ in [28, 30].

Formally generalizing the one-step approximation (5.3), we get the one-step approximation for the semilinear problem (6.1)-(6.2):

$$(6.4) \quad \check{v}(t_k, x) = \frac{1}{2} v(t_{k+1}, x + hb(t_k, x, v(t_{k+1}, x))) \\ + h^{1/2} \sigma(t_k, x, v(t_{k+1}, x)) + \alpha(t_k, x, v(t_{k+1}, x)) \Delta_k w \\ + \frac{1}{2} v(t_{k+1}, x + hb(t_k, x, v(t_{k+1}, x))) \\ - h^{1/2} \sigma(t_k, x, v(t_{k+1}, x)) + \alpha(t_k, x, v(t_{k+1}, x)) \Delta_k w \\ + hf(t_k, x, v(t_{k+1}, x)) + \gamma(t_k, x, v(t_{k+1}, x)) \Delta_k w.$$

In the rest of this section we use the same letter K for various deterministic constants and $C = C(\omega)$ for various positive random variables.

Lemma 6.1. *Let Assumptions 6.1 hold. The error $\rho = \rho(t_k) = \check{v}(t_k, x) - v(t_k, x)$ of the one-step approximation (6.4) is estimated as*

$$(6.5) \quad |E(\rho | \mathcal{F}_T^{t_{k+1}})| \leq C(\omega) h^2, \quad |E(\Delta_k w \rho | \mathcal{F}_T^{t_{k+1}})| \leq C(\omega) h^2,$$

$$(6.6) \quad \left(E[\rho]^{2p}\right)^{1/(2p)} \leq Kh, \quad p \geq 1,$$

where $C(\omega)$ and K do not depend on h, k, x and $EC^2 < \infty$. In the case of additive noise, $\alpha = 0$ and $\gamma(t, x, v) = \gamma(t, x)$ (i.e., $\gamma_0(t, x) = 0$, $\gamma_1(t, x, v) = \gamma(t, x)$), the error ρ satisfies (6.5) and

$$(6.7) \quad \left(E[\rho]^{2p}\right)^{1/(2p)} \leq Kh^{3/2}, \quad p \geq 1.$$

Proof. We note that the increments $\Delta_k w$ and the values $v(t_{k+1}, x)$, $\partial^j v(t_{k+1}, x)/\partial x^j$ are independent. Using the assumptions, we first expand the right-hand side of (6.4) at (t_{k+1}, x) with the Lagrange remainder containing fourth-order spatial derivatives and then obtain

$$(6.8) \quad \begin{aligned} \check{v}(t_k, x) &= v(t_{k+1}, x) + [\alpha(t_k, x, v(t_{k+1}, x))\Delta_k w + hb(t_k, x, v(t_{k+1}, x))] \frac{\partial v}{\partial x}(t_{k+1}, x) \\ &\quad + \frac{1}{2} \left[h\sigma^2(t_k, x, v(t_{k+1}, x)) + \alpha^2(t_k, x, v(t_{k+1}, x)) (\Delta_k w)^2 \right] \frac{\partial^2 v}{\partial x^2}(t_{k+1}, x) \\ &\quad + hf(t_k, x, v(t_{k+1}, x)) + \gamma(t_k, x, v(t_{k+1}, x))\Delta_k w + r_1(t_{k+1}, x), \end{aligned}$$

where the remainder $r_1(t_{k+1}, x)$ satisfies the inequalities (in the general case)

$$(6.9) \quad \begin{aligned} |E(r_1(t_{k+1}, x)|\mathcal{F}_T^{t_{k+1}})| &\leq C(\omega)h^2, \quad |E(\Delta_k w r_1(t_{k+1}, x)|\mathcal{F}_T^{t_{k+1}})| \leq C(\omega)h^2, \\ \left(E[r_1(t_{k+1}, x)]^{2p}\right)^{1/(2p)} &\leq Kh^{3/2}. \end{aligned}$$

Introduce the operators:

$$\begin{aligned} \mathcal{L}^x \psi(t, z) &:= \frac{\partial}{\partial t} \psi(t, z) + [Lv(t, x) + f(t, x, v(t, x))] \frac{\partial}{\partial z} \psi(t, z) \\ &\quad + \frac{1}{2} \left[\alpha(t, x, v(t, x)) \frac{\partial}{\partial x} v + \gamma(t, x, v(t, x)) \right]^2 \frac{\partial^2}{\partial z^2} \psi(t, z), \\ \Lambda^x \psi(t, z) &:= \left[\alpha(t, x, v(t, x)) \frac{\partial}{\partial x} v + \gamma(t, x, v(t, x)) \right] \frac{\partial}{\partial z} \psi(t, z). \end{aligned}$$

Due to the backward Ito formula [32, 28], we have for a smooth $\psi(t, z)$ and $t \leq t_{k+1}$:

$$(6.10) \quad \begin{aligned} \psi(t, v(t, x)) &= \psi(t_{k+1}, v(t_{k+1}, x)) + \int_t^{t_{k+1}} \mathcal{L}^x \psi(s, v(s, x)) ds \\ &\quad + \int_t^{t_{k+1}} \Lambda^x \psi(s, v(s, x)) * dw(s). \end{aligned}$$

We write the solution $v(t, x)$, $T_0 \leq t \leq t_{k+1}$, of (6.1)-(6.2) as

$$(6.11) \quad \begin{aligned} v(t, x) &= v(t_{k+1}, x) + \int_t^{t_{k+1}} [Lv(s, x) + f(s, x, v(s, x))] ds \\ &\quad + \int_t^{t_{k+1}} \left[\alpha(s, x, v(s, x)) \frac{\partial v}{\partial x}(s, x) + \gamma(s, x, v(s, x)) \right] * dw(s) \end{aligned}$$

and, in particular,

$$(6.12) \quad \begin{aligned} v(t_k, x) &= v(t_{k+1}, x) + \int_{t_k}^{t_{k+1}} [Lv(s, x) + f(s, x, v(s, x))] ds \\ &\quad + \int_{t_k}^{t_{k+1}} \left[\alpha(s, x, v(s, x)) \frac{\partial v}{\partial x}(s, x) + \gamma(s, x, v(s, x)) \right] * dw(s). \end{aligned}$$

We substitute the expression for $v(s, x)$ from (6.11) in the parts of the integrand in (6.12):

(6.13)

$$\begin{aligned} Lv(s, x) &= L^{s,x}v(t_{k+1}, x) + \int_s^{t_{k+1}} L^{s,x} [Lv(s', x) + f(s', x, v(s', x))] ds' \\ &\quad + \int_s^{t_{k+1}} L^{s,x} \left[\alpha(s', x, v(s', x)) \frac{\partial v}{\partial x}(s', x) + \gamma(s', x, v(s', x)) \right] * dw(s'), \\ \frac{\partial v}{\partial x}(s, x) &= \frac{\partial v}{\partial x}(t_{k+1}, x) + \int_s^{t_{k+1}} \frac{\partial}{\partial x} [Lv(s', x) + f(s', x, v(s', x))] ds' \\ &\quad + \int_s^{t_{k+1}} \frac{\partial}{\partial x} \left[\alpha(s', x, v(s', x)) \frac{\partial v}{\partial x}(s', x) + \gamma(s', x, v(s', x)) \right] * dw(s'), \end{aligned}$$

where

$$L^{s,x}\psi(t, z) := \frac{1}{2}a(s, x, v(s, x)) \frac{\partial^2}{\partial z^2} \psi(t, z) + b(s, x, v(s, x)) \frac{\partial}{\partial z} \psi(t, z).$$

Now we apply the Ito formula (6.10) to f , α and γ in (6.12) and to the coefficients a and b appearing in $L^{s,x}v(t_{k+1}, x)$ in (6.13). Then taking into account the assumptions and properties of Ito integrals, we obtain

$$\begin{aligned} (6.14) \quad v(t_k, x) &= v(t_{k+1}, x) + h \left[\frac{1}{2}a(t_{k+1}, x, v(t_{k+1}, x)) \frac{\partial^2}{\partial x^2} \right. \\ &\quad \left. + b(t_{k+1}, x, v(t_{k+1}, x)) \frac{\partial}{\partial x} \right] v(t_{k+1}, x) \\ &\quad + hf(t_{k+1}, x, v(t_{k+1}, x)) \\ &\quad + \left[\alpha(t_{k+1}, x, v(t_{k+1}, x)) \frac{\partial v}{\partial x}(t_{k+1}, x) \right. \\ &\quad \left. + \gamma(t_{k+1}, x, v(t_{k+1}, x)) \right] \Delta_k w + r_2(t_{k+1}, x), \end{aligned}$$

where the remainder $r_2(t_{k+1}, x)$ satisfies the inequalities of the form (6.5)-(6.6) in the general case and of the form (6.9) in the case $\alpha = 0$, $\gamma(t, x, v) = \gamma(t, x)$. Furthermore, in (6.14) we expand the coefficients at $(t_k, x, v(t_{k+1}, x))$ and the new remainder has the same properties as r_2 . The obtained result is subtracted from (6.8) and we get (taking into account that $a = \sigma^2 + \alpha^2$)

(6.15)

$$\rho = \check{v}(t_k, x) - v(t_k, x) = \frac{1}{2} \frac{\partial^2 v}{\partial x^2}(t_{k+1}, x) \alpha^2(t_k, x, v(t_{k+1}, x)) \left[(\Delta_k w)^2 - h \right] + r(t_{k+1}, x),$$

where the remainder $r(t_{k+1}, x)$ has the same properties as r_2 . Lemma 6.1 evidently follows from (6.15). $\square$

Based on the one-step approximation (6.4), we obtain the layer method:

$$\begin{aligned}
 (6.16) \quad \bar{v}(t_N, x) &= \varphi(x), \\
 \bar{v}(t_k, x) &= \frac{1}{2} \bar{v} \left(t_{k+1}, x + hb(t_k, x, \bar{v}(t_{k+1}, x)) \right. \\
 &\quad \left. + h^{1/2} \sigma(t_k, x, \bar{v}(t_{k+1}, x)) + \alpha(t_k, x, \bar{v}(t_{k+1}, x)) \Delta_k w \right) \\
 &\quad + \frac{1}{2} \bar{v} \left(t_{k+1}, x + hb(t_k, x, \bar{v}(t_{k+1}, x)) \right. \\
 &\quad \left. - h^{1/2} \sigma(t_k, x, \bar{v}(t_{k+1}, x)) + \alpha(t_k, x, \bar{v}(t_{k+1}, x)) \Delta_k w \right) \\
 &\quad + hf(t_k, x, \bar{v}(t_{k+1}, x)) + \gamma(t_k, x, \bar{v}(t_{k+1}, x)) \Delta_k w, \quad k = N-1, \dots, 1, 0.
 \end{aligned}$$

Let us prove the following technical lemma, which will be exploited in proving the convergence Theorem 6.1.

Lemma 6.2. *Let Assumptions 6.1 hold. Then*

$$\begin{aligned}
 (6.17) \quad R(t_k, x) &:= \bar{v}(t_k, x) - v(t_k, x) \\
 &= \frac{1}{2} R(t_{k+1}, x + \bar{b}h + \bar{\sigma}h^{1/2} + \bar{\alpha}\Delta_k w) + \frac{1}{2} R(t_{k+1}, x + \bar{b}h - \bar{\sigma}h^{1/2} + \bar{\alpha}\Delta_k w) \\
 &\quad + \Delta fh + \Delta \gamma \Delta_k w + \frac{\partial v}{\partial x}(t_{k+1}, x) (\Delta bh + \Delta \alpha \Delta_k w) + r + \rho,
 \end{aligned}$$

where ρ is the one-step error as in Lemma 6.1 and

$$|r| \leq C |R(t_{k+1}, x)| (h + \Delta_k^2 w).$$

Proof. From (6.16) and due to the equality $\bar{v}(t_{k+1}, \cdot) = v(t_{k+1}, \cdot) + R(t_{k+1}, \cdot)$, we get

$$\begin{aligned}
 (6.18) \quad \bar{v}(t_k, x) &= \frac{1}{2} v(t_{k+1}, x + \bar{b}h + \bar{\sigma}h^{1/2} + \bar{\alpha}\Delta_k w) + \frac{1}{2} v(t_{k+1}, x + \bar{b}h - \bar{\sigma}h^{1/2} + \bar{\alpha}\Delta_k w) \\
 &\quad + \frac{1}{2} R(t_{k+1}, x + \bar{b}h + \bar{\sigma}h^{1/2} + \bar{\alpha}\Delta_k w) + \frac{1}{2} R(t_{k+1}, x + \bar{b}h - \bar{\sigma}h^{1/2} + \bar{\alpha}\Delta_k w) \\
 &\quad + \bar{f}h + \bar{\gamma}\Delta_k w, \quad k = N-1, \dots, 1, 0,
 \end{aligned}$$

where $\bar{b}$, $\bar{\sigma}$, $\bar{\alpha}$, $\bar{f}$, $\bar{\gamma}$ are the coefficients $b(t, x, v), \dots, \gamma(t, x, v)$ calculated at $t = t_k$, $x = x$, $v = \bar{v}(t_{k+1}, x) = v(t_{k+1}, x) + R(t_{k+1}, x)$.

We have

$$(6.19) \quad \bar{b} = b(t_k, x, v(t_{k+1}, x) + R(t_{k+1}, x)) = b(t_k, x, v(t_{k+1}, x)) + \Delta b := b + \Delta b,$$

where Δb satisfies the inequality

$$(6.20) \quad |\Delta b| \leq K |R(t_{k+1}, x)|$$

with a deterministic constant K . Due to the uniform boundedness of b , we also get

$$(6.21) \quad |\Delta b|^2 \leq |\Delta b| \times K |R(t_{k+1}, x)| \leq K |R(t_{k+1}, x)|.$$

Analogously,

$$\begin{aligned}
 (6.22) \quad \bar{\sigma} &= \sigma + \Delta \sigma, \quad |\Delta \sigma| \leq K |R(t_{k+1}, x)|, \quad |\Delta \sigma|^2 \leq K |R(t_{k+1}, x)|, \\
 \bar{\alpha} &= \alpha + \Delta \alpha, \quad |\Delta \alpha| \leq K |R(t_{k+1}, x)|, \quad |\Delta \alpha|^2 \leq K |R(t_{k+1}, x)|.
 \end{aligned}$$

For Δf and $\Delta \gamma$, we have

$$(6.23) \quad |\Delta f| \leq K |R(t_{k+1}, x)|, \quad |\Delta \gamma| \leq K |R(t_{k+1}, x)|.$$

It is not difficult to see that (6.20)-(6.22) imply the equalities

$$(6.24) \quad v(t_{k+1}, x + \bar{b}h \pm \bar{\sigma}h^{1/2} + \bar{\alpha}\Delta_k w) = v(t_{k+1}, x + bh \pm \sigma h^{1/2} + \alpha\Delta_k w) \\ + \frac{\partial v}{\partial x}(t_{k+1}, x + bh \pm \sigma h^{1/2} + \alpha\Delta_k w)(\Delta bh \pm \Delta \sigma h^{1/2} + \Delta \alpha \Delta_k w) + r,$$

where

$$(6.25) \quad |r| \leq C |\Delta bh \pm \Delta \sigma h^{1/2} + \Delta \alpha \Delta_k w|^2 \leq C |R(t_{k+1}, x)| (h + \Delta_k^2 w),$$

and C is a random variable with uniformly bounded second moment EC^2 .

Furthermore,

$$(6.26) \quad \frac{\partial v}{\partial x}(t_{k+1}, x + bh \pm \sigma h^{1/2} + \alpha\Delta_k w) = \frac{\partial v}{\partial x}(t_{k+1}, x) + Ch^{1/2} + C\Delta_k w, \quad EC^2 < \infty.$$

Substituting (6.26) in (6.24) and then the new (6.24) in (6.18) and taking into account (6.20)-(6.22) with (6.25), we obtain

$$(6.27) \quad \bar{v}(t_k, x) = \frac{1}{2}v(t_{k+1}, x + bh + \sigma h^{1/2} + \alpha\Delta_k w) + \frac{1}{2}v(t_{k+1}, x + bh - \sigma h^{1/2} + \alpha\Delta_k w) \\ + fh + \gamma\Delta_k w + \Delta fh + \Delta \gamma \Delta_k w + \frac{\partial v}{\partial x}(t_{k+1}, x)(\Delta bh + \Delta \alpha \Delta_k w) \\ + \frac{1}{2}R(t_{k+1}, x + \bar{b}h + \bar{\sigma}h^{1/2} + \bar{\alpha}\Delta_k w) + \frac{1}{2}R(t_{k+1}, x + \bar{b}h - \bar{\sigma}h^{1/2} + \bar{\alpha}\Delta_k w) + r \\ = \check{v}(t_k, x) + \frac{1}{2}R(t_{k+1}, x + \bar{b}h + \bar{\sigma}h^{1/2} + \bar{\alpha}\Delta_k w) + \frac{1}{2}R(t_{k+1}, x + \bar{b}h - \bar{\sigma}h^{1/2} + \bar{\alpha}\Delta_k w) \\ + \Delta fh + \Delta \gamma \Delta_k w + \frac{\partial v}{\partial x}(t_{k+1}, x)(\Delta bh + \Delta \alpha \Delta_k w) + r,$$

where the new r also satisfies

$$|r| \leq C |R(t_{k+1}, x)| (h + \Delta_k^2 w).$$

At last using Lemma 6.1, we obtain

$$(6.28) \quad R(t_k, x) = \bar{v}(t_k, x) - v(t_k, x) = \check{v}(t_k, x) - v(t_k, x) \\ + \frac{1}{2}R(t_{k+1}, x + \bar{b}h + \bar{\sigma}h^{1/2} + \bar{\alpha}\Delta_k w) + \frac{1}{2}R(t_{k+1}, x + \bar{b}h - \bar{\sigma}h^{1/2} + \bar{\alpha}\Delta_k w) \\ + \Delta fh + \Delta \gamma \Delta_k w + \frac{\partial v}{\partial x}(t_{k+1}, x)(\Delta bh + \Delta \alpha \Delta_k w) + r \\ = \frac{1}{2}R(t_{k+1}, x + \bar{b}h + \bar{\sigma}h^{1/2} + \bar{\alpha}\Delta_k w) + \frac{1}{2}R(t_{k+1}, x + \bar{b}h - \bar{\sigma}h^{1/2} + \bar{\alpha}\Delta_k w) \\ + \Delta fh + \Delta \gamma \Delta_k w + \frac{\partial v}{\partial x}(t_{k+1}, x)(\Delta bh + \Delta \alpha \Delta_k w) + r + \rho,$$

whence (6.17) follows. $\square$

Due to Lemmas 6.1 and 6.2, one can expect that the method (6.16) converges with the mean-square order $1/2$. But we have not succeeded in proving the corresponding convergence theorem in the general case. Here we restrict ourselves to proving the following theorem.

Theorem 6.1. *Let Assumptions 6.1 hold. Let $\alpha = 0$ and b and σ be independent of v . Then*

$$(6.29) \quad (E(\bar{v}(t_k, x) - v(t_k, x))^2)^{1/2} \leq Kh^{1/2},$$

where K does not depend on h, k, x . If, in addition, $\gamma(t, x, v) = \gamma(t, x)$ (i.e., in the additive noise case), then

$$(6.30) \quad (E(\bar{v}(t_k, x) - v(t_k, x))^2)^{1/2} \leq Kh.$$

Proof. In the case $\alpha = 0$ and b and σ independent of v , the remainder r in (6.28) is equal to zero (see (6.25)). Then we get

$$(6.31) \quad R(t_k, x) = \frac{1}{2}R(t_{k+1}, x + bh + \sigma h^{1/2}) + \frac{1}{2}R(t_{k+1}, x + bh - \sigma h^{1/2}) + \Delta f h + \Delta \gamma \Delta_k w + \rho,$$

where ρ satisfies (6.6) in the case of general γ and (6.7) in the case $\gamma(t, x, v) = \gamma(t, x)$. Let us square (6.31) and then take conditional expectation with respect to the σ -algebra $\mathcal{F}_T^{t_{k+1}}$. We get

$$(6.32) \quad \begin{aligned} E(R^2(t_k, x) \mid \mathcal{F}_T^{t_{k+1}}) &= \frac{1}{4}R^2(t_{k+1}, x + bh + \sigma h^{1/2}) + \frac{1}{4}R^2(t_{k+1}, x + bh - \sigma h^{1/2}) \\ &\quad + \frac{1}{2}R(t_{k+1}, x + bh + \sigma h^{1/2})R(t_{k+1}, x + bh - \sigma h^{1/2}) \\ &\quad + h^2(\Delta f)^2 + h(\Delta \gamma)^2 + E(\rho^2 \mid \mathcal{F}_T^{t_{k+1}}) \\ &\quad + (R(t_{k+1}, x + bh + \sigma h^{1/2}) + R(t_{k+1}, x + bh - \sigma h^{1/2}))(\Delta f h + E(\rho \mid \mathcal{F}_T^{t_{k+1}})) \\ &\quad + 2\Delta f h E(\rho \mid \mathcal{F}_T^{t_{k+1}}) + 2\Delta \gamma E(\Delta_k w \rho \mid \mathcal{F}_T^{t_{k+1}}). \end{aligned}$$

Using Lemma 6.1 and the inequalities (6.23), one can estimate the terms in (6.32). For instance,

$$|\Delta \gamma E(\Delta_k w \rho \mid \mathcal{F}_T^{t_{k+1}})| \leq K |R(t_{k+1}, x)| C(\omega) h^2 \leq \frac{1}{2} K^2 R^2(t_{k+1}, x) h + \frac{1}{2} C^2(\omega) h^3.$$

Introduce the notation

$$m(t_k) := \sup_x E R^2(t_k, x).$$

After taking expectations in (6.32), we easily obtain

$$E R^2(t_k, x) \leq (1 + Kh) m(t_{k+1}) + E \rho^2(t_{k+1}) + Kh^3,$$

where the constant K does not depend on h, k, x . Therefore

$$m(t_k) \leq (1 + Kh) m(t_{k+1}) + E \rho^2(t_{k+1}) + Kh^3,$$

whence (6.29) follows because of $E \rho^2 \leq Kh^2$ in the multiplicative noise case (i.e., when γ depends on v) and (6.30) follows because of $E \rho^2 \leq Kh^3$ in the additive noise case. $\square$

Introduce the equidistant space discretization as in (5.5). Using linear interpolation, we construct the following algorithm on the basis of the layer method (6.16):

$$\begin{aligned}
(6.33) \quad & \bar{v}(t_N, x) = \varphi(x), \\
& \bar{v}(t_k, x_j) = \frac{1}{2} \bar{v} \left(t_{k+1}, x_j + hb(t_k, x_j, \bar{v}(t_{k+1}, x_j)) \right. \\
& \quad \left. + h^{1/2} \sigma(t_k, x_j, \bar{v}(t_{k+1}, x_j)) + \alpha(t_k, x_j, \bar{v}(t_{k+1}, x_j)) \Delta_k w \right) \\
& \quad + \frac{1}{2} \bar{v} \left(t_{k+1}, x_j + hb(t_k, x_j, \bar{v}(t_{k+1}, x_j)) \right. \\
& \quad \left. - h^{1/2} \sigma(t_k, x_j, \bar{v}(t_{k+1}, x_j)) + \alpha(t_k, x_j, \bar{v}(t_{k+1}, x_j)) \Delta_k w \right) \\
& \quad + hf(t_k, x_j, \bar{v}(t_{k+1}, x_j)) + \gamma(t_k, x_j, \bar{v}(t_{k+1}, x_j)) \Delta_k w, \quad j = 0, \pm 1, \pm 2, \dots, \\
(6.34) \quad & \bar{v}(t_k, x) = \frac{x_{j+1} - x}{h_x} \bar{v}(t_k, x_j) + \frac{x - x_j}{h_x} \bar{v}(t_k, x_{j+1}), \quad x_j \leq x \leq x_{j+1}, \\
& \quad k = N - 1, \dots, 1, 0.
\end{aligned}$$

The following convergence theorem is proved for this algorithm using ideas from the above proof of Theorem 6.1 and also from the proof of an analogous theorem for an algorithm in the deterministic case [23, Chapter 7].

Theorem 6.2. *Let Assumptions 6.1 hold. Let $\alpha = 0$ and b and σ be independent of v . Then the error of the one-step approximation (6.33)-(6.34) with $h_x = \varkappa h^{3/4}$, $\varkappa > 0$, is estimated as*

$$(6.35) \quad (E(\bar{v}(t_k, x) - v(t_k, x))^2)^{1/2} \leq Kh^{1/2},$$

where K does not depend on h, k, x . If, in addition, $\gamma(t, x, v) = \gamma(t, x)$ (i.e., in the additive noise case) and $h_x = \varkappa h$, $\varkappa > 0$, then

$$(6.36) \quad (E(\bar{v}(t_k, x) - v(t_k, x))^2)^{1/2} \leq Kh.$$

We note that Remark 5.2 is valid here.

7. NUMERICAL EXPERIMENTS

7.1. Model problem: Ornstein-Uhlenbeck equation. We consider the problem

$$(7.1) \quad -dv = \left[\frac{a^2}{2} \frac{\partial^2 v}{\partial x^2} + bx \frac{\partial v}{\partial x} \right] dt + \alpha \frac{\partial v}{\partial x} * dw(t), \quad (t, x) \in [T_0, T) \times \mathbf{R},$$

$$(7.2) \quad v(T, x) = \varphi(x), \quad x \in \mathbf{R},$$

where $w(t)$ is a standard scalar Wiener process, a, b, α are constants and

$$(7.3) \quad \sigma^2 = a^2 - \alpha^2 \geq 0.$$

The solution of the problem is given by (see (2.6)-(2.7)):

$$(7.4) \quad v(t, x) = E^w \varphi(X_{t,x}(T)),$$

$$(7.5) \quad dX = bXds + \sigma dW(s) + \alpha dw(s).$$

We have

$$(7.6) \quad X_{t,x}(T) = e^{b(T-t)}x + \int_t^T e^{b(T-\vartheta)} [\sigma dW(\vartheta) + \alpha dw(\vartheta)].$$

Clearly, the conditional distribution of $X_{t,x}(T)$ under $w(s)$, $t \leq s \leq T$, is Gaussian with the parameters

$$\begin{aligned}
 (7.7) \quad m(t, x) &:= E(X_{t,x}(T)/w(s) - w(t), t \leq s \leq T) \\
 &= e^{b(T-t)}x + \int_t^T e^{b(T-\vartheta)} \alpha dw(\vartheta) \\
 &= e^{b(T-t)}x + \alpha(w(T) - w(t)) + \alpha b \int_t^T e^{b(T-\vartheta)} (w(\vartheta) - w(t)) d\vartheta, \\
 (7.8) \quad \delta^2(t) &:= \text{Var}(X_{t,x}(T)/w(s) - w(t), t \leq s \leq T) \\
 &= \frac{\sigma^2}{2b} (e^{2b(T-t)} - 1).
 \end{aligned}$$

From here, we get the following explicit solution of the Ornstein-Uhlenbeck equation:

$$(7.9) \quad v(t, x) = E^w \varphi(X_{t,x}(T)) = \frac{1}{\sqrt{2\pi}\delta(t)} \int_{-\infty}^{\infty} \varphi(\xi) \exp\left(-\frac{(\xi - m(t, x))^2}{2\delta^2(t)}\right) d\xi.$$

We have

$$(7.10) \quad \frac{\partial v}{\partial x}(t, x) = \frac{e^{b(T-t)}}{\sqrt{2\pi}\delta^3(t)} \int_{-\infty}^{\infty} \varphi(\xi) [\xi - m(t, x)] \exp\left(-\frac{(\xi - m(t, x))^2}{2\delta^2(t)}\right) d\xi.$$

Now consider the problem which is a perturbation of (7.1)-(7.2):

$$(7.11) \quad -du = \left[\frac{a^2}{2} \frac{\partial^2 u}{\partial x^2} + (bx + \varepsilon b_1(x)) \frac{\partial u}{\partial x} \right] dt + \alpha \frac{\partial u}{\partial x} * dw(t), \quad (t, x) \in [T_0, T) \times \mathbf{R},$$

$$(7.12) \quad u(T, x) = \varphi(x), \quad x \in \mathbf{R}.$$

Aiming to reduce variance in the Monte Carlo procedure, let us use the representation (4.2)-(4.3) with $\mu = 0$ (hence $Y \equiv 1$) and with

$$(7.13) \quad \lambda(s, x) = -\sigma \frac{\partial v}{\partial x}(s, x),$$

where $\partial v / \partial x$ is from (7.10). We have

$$(7.14) \quad u(t, x) = E^w [\varphi(X_{t,x}(T)) + Z_{t,x,1,0}(T)],$$

where X, Z satisfy the system

$$\begin{aligned}
 (7.15) \quad dX &= [bX + \varepsilon b_1(X)] ds + \sigma dW(s) + \alpha dw(s), \\
 dZ &= -\sigma \frac{\partial v}{\partial x}(s, X) dW(s).
 \end{aligned}$$

To realize (7.14), we use both the mean-square and the weak Euler procedures for solving (7.15) with respect to the Wiener process $W(t)$. Due to Theorem 4.1, the variance is expected to be small if ε is relatively small. Here we exploit the control variates method (we put $\mu = 0$) for variance reduction. Since in the SPDEs (7.1)-(7.2) and (7.11)-(7.12) the coefficient α is constant and $\beta = \gamma = 0$, the method of important sampling can also be used without any theoretical difficulties (see the corresponding comment at the end of Section 4).

7.2. Numerical results. For definiteness, we simulate (7.1)-(7.2) with

$$(7.16) \quad \varphi(x) = x^2.$$

In the tests we fix a trajectory $w(t)$, $0 \leq t \leq T$, which is obtained with a small time step equal to 0.0001. To evaluate the exact solution given by (7.9), we simulate the integral in (7.7) by the trapezoidal rule with the step 0.0001. In Table 1 we present the results of simulating the solution of (7.1)-(7.2), (7.16) by the weak Euler-type scheme (3.7), (3.9). One can observe convergence with order one that is in good agreement with our theoretical results (see Remark 3.1 and note that in this example α and σ are constant). In the table the “ $\pm$ ” reflects the Monte Carlo error only, it gives the confidence interval for the corresponding value with probability 0.95. Similar results are obtained by the mean-square Euler scheme (3.2). For instance, for $h = 0.1$ (the other parameters are the same as in Table 1) we get $\hat{v}(0, 0) = 0.7550 \pm 0.0019$.

TABLE 1. *Ornstein-Uhlenbeck equation.* Evaluation of $v(t, x)$ from (7.1)-(7.2), (7.16) with various time steps h . Here $a = 1$, $b = -1$, $\alpha = 0.5$, and $T = 10$. The expectations are computed by the Monte Carlo technique simulating $M = 10^6$ independent realizations. The “ $\pm$ ” reflects the Monte Carlo error only, it does not reflect the error of the method. All simulations are done along the same sample path $w(t)$. The corresponding reference value is 0.73726, which is found due to (7.9).

h	$\hat{v}(0, 0)$
0.2	0.7735 ± 0.0019
0.1	0.7557 ± 0.0018
0.05	0.7456 ± 0.0018
0.02	0.7412 ± 0.0018
0.01	0.7391 ± 0.0018

To demonstrate the variance reduction technique from Section 4, we first simulate (7.1)-(7.2), (7.16) using a probabilistic representation of the form (7.14), (7.15) with $\varepsilon = 0$. In particular, we obtain that for $M = 10^4$, $h = 0.1$ (the other parameters are the same as in Table 1) $\hat{v}(0, 0) = 0.7562 \pm 0.0012$ and for $M = 100$, $h = 0.01$ the approximate value $\hat{v}(0, 0) = 0.7383 \pm 0.0010$. Recalling that the results in Table 1 are obtained with $M = 10^6$ Monte Carlo runs, one can see that we reach the same level of the Monte Carlo error with significantly fewer Monte Carlo runs. We note that although we use the optimal $\lambda(s, x)$ here, the variance (and, consequently, the Monte Carlo error) is not zero, which is due to the error of numerical integration of the equation for Z in (7.15).

Now we evaluate the solution of the perturbed Ornstein-Uhlenbeck equation (7.11)-(7.12). We take $\varphi(x)$ from (7.16) and we choose a small $\varepsilon > 0$ and

$$(7.17) \quad b_1(x) = -x^3.$$

We simulate (7.14), (7.15), (7.16), (7.17) both without employing the variance reduction technique (i.e., we put $Z(T) = 0$ in (7.14)) and with variance reduction (i.e., using $\lambda(s, x)$ from (7.13), (7.10)) by the Euler-type scheme (3.7), (3.9). The results of the experiments are presented in Table 2. When the variance reduction

TABLE 2. *Perturbed Ornstein-Uhlenbeck equation.* Evaluation of $u(t, x)$ from (7.11)-(7.12), (7.16), (7.17) at $(0, 0)$ with various time steps h . Here $a = 1$, $b = -1$, $\alpha = 0.5$, $\varepsilon = 0.1$, and $T = 10$. The expectations are computed by the Monte Carlo technique simulating $M = 10^4$ independent realizations. The corresponding reference value is 0.6006 ± 0.0004 , which is found by simulation with $h = 0.001$ and $M = 10^6$.

h	without variance reduction	with variance reduction
0.2	0.614 ± 0.014	0.6014 ± 0.0042
0.1	0.611 ± 0.014	0.6037 ± 0.0040
0.02	0.604 ± 0.014	0.6001 ± 0.0041

technique is used, the Monte Carlo error is 3.5 times less than in the standard simulation without variance reduction. In other words, to reach the same accuracy, we can run 12 times fewer trajectories in the case with variance reduction than without one, which is a significant gain of computational efficiency.

ACKNOWLEDGMENTS

This work was partially supported by the UK EPSRC Research Grant EP/D049792/1. MVT was also supported by a Leverhulme Research Fellowship. Part of this work was done while MVT was on study leave granted by the University of Leicester.

REFERENCES

- [1] E. J. Allen, S. J. Novosel, Z. Zhang. Finite element and difference approximation of some linear stochastic partial differential equations. *Stoch. Stoch. Rep.* 64 (1998), 117–142. MR1637047 (99d:60067)
- [2] A. de Bouard, A. Debussche. Weak and strong order of convergence of a semidiscrete scheme for the stochastic nonlinear Schrödinger equation. *J. Appl. Math. Optim.* 54 (2006), 369–399. MR2268663 (2008g:60208)
- [3] D. Crisan. Particle approximations for a class of stochastic partial differential equations. *J. Appl. Math. Optim.* 54 (2006), 293–314. MR2268660 (2007m:60174)
- [4] G. Da Prato, J. Zabczyk. *Stochastic Equations in Infinite Dimensions*. Cambridge University Press, 1992. MR1207136 (95g:60073)
- [5] A. Debussche, J. Printems. Weak order for the discretization of the stochastic heat equation. Preprint, 2007, arxiv.org/abs/0710.5450.
- [6] I.I. Gichman, A.V. Skorochod. *Stochastic Differential Equations*. Naukova Dumka, Kiev, 1968. MR0263172 (41:7777)
- [7] E. Gobet, G. Pages, H. Pham, J. Printems. Discretization and simulation of Zakai equation. *SIAM J. Numer. Anal.* 44 (2006), 2505–2538. MR2272604 (2008c:60063)
- [8] W. Grecksch, P. E. Kloeden. Time-discretised Galerkin approximations of parabolic stochastic PDEs. *Bull. Austral. Math. Soc.* 54 (1996), 79–85. MR1402994 (97g:60080)
- [9] I. Gyöngy. A note on Euler's approximations. *Potential Analysis* 8 (1998), 205–216. MR1625576 (99d:60060)
- [10] I. Gyöngy. Approximations of stochastic partial differential equations. In: *Stochastic partial differential equations and applications (Trento, 2002)*, *Lecture Notes in Pure and Appl. Math.*, Vol. 227, Dekker, New York, 2002, 287–307. MR1919514 (2003b:60096)
- [11] I. Gyöngy, N. Krylov. On the splitting-up method and stochastic partial differential equations. *Ann. Appl. Prob.* 31 (2003), 564–591. MR1964941 (2004c:60182)
- [12] I. Gyöngy, A. Millet. On discretization schemes for stochastic evolution equations. *Potential Analysis* 23 (2005), 99–134. MR2139212 (2006a:60115)

- [13] P. Kotelenetz. *Stochastic Ordinary and Stochastic Partial Differential Equations: Transition from microscopic to macroscopic equations*. Springer, 2008. MR2370567
- [14] N. V. Krylov. *An Analytic Approach to SPDEs*. In: Stochastic Partial Differential Equations, Six Perspectives, R.A. Carmona and B. Rozovskii, eds., Mathematical Surveys and Monographs, Vol. 64, AMS, 1999, 185–242. MR1661766 (99j:60093)
- [15] N. V. Krylov, B.L. Rozovskii. On the characteristics of degenerate second order parabolic Ito equations. *J. Soviet Math.* 32 (1986), 336–348.
- [16] H. Kunita. *Stochastic Flows and Stochastic Differential Equations*. Cambridge University Press, 1990. MR1070361 (91m:60107)
- [17] F. Le Gland. Splitting-up approximation for SPDEs and SDEs with application to nonlinear filtering. *Lecture Notes in Control and Inform. Sc.*, Vol. 176, Springer, 1992, 177–187. MR1176783
- [18] R. S. Liptser, A.N. Shirayev. *Statistics of Random Processes*. Springer, I. General theory, 1977, II. Applications, 1978. MR0474486 (57:14125); MR0488267 (38:7827)
- [19] R. Mikulevicius, B.L. Rozovskii. Linear parabolic stochastic PDEs and Wiener chaos. *SIAM J. Math. Anal.* 29 (1998), 452–480. MR1616515 (2000c:60095)
- [20] G. N. Milstein. *Numerical Integration of Stochastic Differential Equations*. Ural State University, Sverdlovsk, 1988. English transl.: Kluwer Academic Publishers, 1995. MR1335454 (96e:65003)
- [21] G. N. Milstein. The probability approach to numerical solution of nonlinear parabolic equations. *Numer. Meth. PDEs* 18 (2002), 490–522. MR1906636 (2003g:65129)
- [22] G. N. Milstein, J.G.M. Schoenmakers, V. Spokoyny. Transition density estimation for stochastic differential equations via forward-reverse representations. *Bernoulli* 10 (2004), 281–312. MR2046775 (2005f:62130)
- [23] G. N. Milstein, M.V. Tretyakov. *Stochastic Numerics for Mathematical Physics*. Springer, 2004. MR2069903 (2005f:60004)
- [24] G. N. Milstein, M.V. Tretyakov. Discretization of forward-backward stochastic differential equations and related quasilinear parabolic equations. *IMA J. Numer. Anal.* 27 (2007), 24–44. MR2289270 (2008a:65013)
- [25] G. N. Milstein, M.V. Tretyakov. *Monte Carlo algorithms for backward equations in nonlinear filtering*. Advances in Applied Probability 41 (2009), no. 1.
- [26] G. N. Milstein, M.V. Tretyakov. Practical variance reduction via regression for simulating diffusions. *SIAM J. Numer. Anal.* (accepted).
- [27] N. J. Newton. Variance reduction for simulated diffusions. *SIAM J. Appl. Math.* 54 (1994), 1780–1805. MR1301282 (96e:65004)
- [28] E. Pardoux. Stochastic partial differential equations and filtering of diffusion processes. *Stochastics* 3 (1979), 127–167. MR553909 (81b:60059)
- [29] E. Pardoux. Nonlinear filtering, prediction and smoothing. In: M. Hazewinkel, J.C. Willems (Eds.), *Stochastic Systems: The Mathematics of Filtering and Identification and Applications*, NATO Advanced Study Institute series, D. Reidel, 1981, 529–557. MR674341 (84h:93070)
- [30] E. Pardoux, S. Peng. Backward doubly stochastic differential equations and systems of quasilinear SPDEs. *Probab. Theory Relat. Fields* 98 (1994), 209–227. MR1258986 (94m:60120)
- [31] J. Picard. Approximation of nonlinear filtering problems and order of convergence. *Lecture Notes in Contr. Inform. Sc.*, Vol. 61, Springer, 1984, 219–236. MR874832 (87m:93128)
- [32] B. L. Rozovskii. *Stochastic Evolution Systems, Linear Theory and Application to Nonlinear Filtering*. Kluwer Academic Publishers, Dordrecht, 1991. MR1135324 (92k:60136)
- [33] H. Wendland. *Scattered Data Approximation*. Cambridge University Press, Cambridge, 2005. MR2131724 (2006i:41002)
- [34] H. Yoo. Semi-discretization of stochastic partial differential equations on R^1 by a finite-difference method. *Math. Comp.* 69 (2000), 653–666. MR1654030 (2000i:65129)

URAL STATE UNIVERSITY, LENIN STR. 51, 620083 EKATERINBURG, RUSSIA
E-mail address: Grigori.Milstein@usu.ru

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF LEICESTER, LEICESTER LE1 7RH, UNITED KINGDOM
E-mail address: M.Tretyakov@le.ac.uk