Int. J. Appl. Math. Comput. Sci., 2003, Vol. 13, No. 2, 151–160

APPROXIMATION OF THE ZAKAI EQUATION IN A NONLINEAR FILTERING PROBLEM WITH DELAY

KRYSTYNA TWARDOWSKA*, TOMASZ MARNIK**
MONIKA PASŁAWSKA-POŁUDNIAK**

* Faculty of Mathematics and Information Science
Warsaw University of Technology
Plac Politechniki 1, 00–661 Warsaw, Poland
e-mail: tward@alpha.mini.pw.edu.pl

** Department of Mathematics
Technical University of Rzeszów
ul. W. Pola 2, 35–959 Rzeszów, Poland
e-mail: ma_poludniak@poczta.onet.pl

A nonlinear filtering problem with delays in the state and observation equations is considered. The unnormalized conditional probability density of the filtered diffusion process satisfies the so-called Zakai equation and solves the nonlinear filtering problem. We examine the solution of the Zakai equation using an approximation result. Our theoretical deliberations are illustrated by a numerical example.

Keywords: nonlinear filtering, stochastic differential equations with delay, Zakai's equation

1. Introduction

We study a nonlinear filtering problem with delay using an approximation result of the Wong-Zakai type for the corresponding Zakai equation with delay. The nonlinear filtering problem was considered in the literature, e.g., by Bucy (1965), Kushner (1967), Zakai (1969), Liptser and Shiryayev (1977), Pardoux (1979; 1989), Kallianpur (1980), and others. Their studies concentrated mainly on finding an equation for the conditional probability density of an unobserved process given an observed path. It is known that the conditional expectation gives the best estimate in the mean square sense. The conditional density can be computed by two methods. The first method gives the so-called Kushner equation (Kushner, 1967), which is a nonlinear stochastic partial differential equation. The second method gives the so-called Zakai equation (Bucy, 1965; Zakai, 1969), which is a linear stochastic partial differential equation for the unnormalized density. Therefore, the problem of constructing solutions of the Zakai equation is more important for practical applications because of the linearity.

In recent years, the Zakai equation has been examined by many authors, e.g., by Bensoussan *et al.* (1990) using a splitting method, by Lototsky *et al.* (1997) using a spectral approach, by Crisan *et al.* (1998) using a branching particle method, by Cohen de Lara (1998) using invariance group techniques, by Elliot and Moore (1998) in Hilbert spaces, and by Atar *et al.* (1999) using the Feynman-Kac formula.

In our study we apply the approximation problem of the Wong-Zakai type for stochastic partial differential equations. It was considered by Gyöngy (1989), Gyöngy and Pröhle (1990), Brzeźniak and Flandoli (1995), and Twardowska (1995). They showed that if in the Zakai equation we replace the disturbance by its good approximations, then the approximations converge to a limit equation with the so-called Itô correction term. The above problems were considered without delays.

The well-known result for the existence and uniqueness of a filtering problem with delays but in the linear case belongs to Kolmanovsky (1973), see also (Kolmanovsky *et al.*, 2002). The approximation result is not considered.

In this paper, the Zakai equation is a linear stochastic parabolic partial differential equation with delay. It corresponds to our nonlinear filtering problem with delay. We prove the existence and uniqueness theorem for this equation. Also, we establish the approximation result using the correction term derived in (Twardowska, 1991;

1993; 1995) in the approximation theorems of the Wong-Zakai type.

An important part of the present paper contains a numerical example showing that a good stability result is achieved because in the approximation sequence of equations we have added the appropriate correction term for stochastic linear differential equations with delay. Using the Galerkin technique and some numerical schemes (Kloeden and Platen, 1992; Sobczyk, 1991) we transform the Zakai equation to a simpler finite-multidimensional form. We solve this equation without any correction term and with a correction term in the approximation sequence. It is evident that the correction term has a crucial role and improves our approximation results.

In the paper by Ahmed and Radaideh (1997), a numerical method for the approximation of a nonlinear filtering problem was developed. Using the Galerkin technique, the solution of Zakai's equation was approximated by a sequence of nonstandard basis functions given by a parameterized family of Gaussian densities. We take some ideas from that paper. Other numerical techniques for the Zakai equation can be found in the papers by Beneš (1981), Elliot and Głowiński (1989), Florchinger and Le Gland (1995), and Itô (1996).

2. Definitions and Notation

We consider the probability space $(\Omega, \mathcal{F}, \mathcal{F}_{t\in[0,\infty)}, P)$ such that it is the cannonical space of a process $\{(X(t), Y(t)), t \in [0,\infty)\} \in \mathbb{R}^M \times \mathbb{R}^N$, where

$$\begin{aligned} \Omega &= \Omega_1 \times \Omega_2, \\ \Omega_1 &= C(\mathbb{R}_+, \mathbb{R}^M), \quad \Omega_2 = C(\mathbb{R}_+, \mathbb{R}^N), \\ X(t,\omega) &= \omega_1(t), \quad Y(t,\omega) = \omega_2(t), \\ \mathcal{F}_t &= \sigma\{(X(s), Y(s)), \quad 0 \le s \le t\} \cup \mathbb{N}, \end{aligned}$$

$\mathcal{F}$ is a σ-algebra of Borel sets on $\Omega \cup \mathbb{N}$, where $\mathbb{N}$ is a class of subsets with the P-measure equal to zero, P is the probability law of the process (X, Y), $C(\mathbb{R}_+, \mathbb{R}^M)$ is the class of continuous functions, and $C_b(\mathbb{R}_+, \mathbb{R}^M)$ denotes the class of bounded continuous functions.

For the stochastic process $X(t,\omega)$ and for a fixed $t \in [0,\infty)$ we define

$$X_t(\theta,\omega) = X(t+\theta,\omega), \quad \theta \in I = [-r, 0].$$

Therefore $X_t(\cdot,\omega)$ denotes the segment of the trajectory $X(\cdot,\omega)$ on $[t-r, t]$.

Let $\{(X(t), Y(t)), t \in [0,\infty)\}$ be the solution to the following system of stochastic equations with delay:

$$\begin{aligned} X(t,\omega) = X_0(\omega) &+ \int_0^t b(s, Y_s(\cdot,\omega), X_s(\cdot,\omega))\,\mathrm{d}s \\ &+ \int_0^t f\big(s, Y_s(\cdot,\omega), X_s(\cdot,\omega)\big)\,\mathrm{d}V(s) \\ &+ \int_0^t g\big(s, Y_s(\cdot,\omega), X_s(\cdot,\omega)\big)\,\mathrm{d}W(s), \end{aligned} \tag{1}$$

$$\begin{aligned} Y(t,\omega) = Y_0(\omega) &+ \int_0^t h\big(s, Y_s(\cdot,\omega), X_s(\cdot,\omega)\big)\,\mathrm{d}s \\ &+ W(t), \end{aligned} \tag{2}$$

where $X_0(\omega)$ is an initial constant random variable independent of the standard Wiener processes $\{(V(t), W(t)), t \in [0,\infty)\}$ with values in $\mathbb{R}^M \times \mathbb{R}^N$, $Y_0(\omega) = 0$. Moreover, b, f, g and h are measurable mappings from $\mathbb{R}_+ \times\ C(I, \mathbb{R}^N) \times C(I, \mathbb{R}^M)$ with values in $\mathbb{R}^M$, $\mathbb{R}^M$, $\mathbb{R}^{M\times N}$ and $\mathbb{R}^N$, respectively. We assume that they satisfy Lipschitz and growth conditions (see §4 below). Then the system of equations (1)–(2) has exactly one solution. The uniqueness is understood in the sense of trajectories. We shall call $X(t)$ the state and $Y(t)$ the observation process.

We define

$$a(t,y,x) = f \circ f^*(t,y,x) + g \circ g^*(t,y,x) \tag{3}$$

for $t \in \mathbb{R}_+$, $y \in C(I, R^N)$ and $x \in C(I, \mathbb{R}^M)$, where f^* and g^* are the transpose matrices of f and g, respectively. Moreover,

$$\begin{aligned} Z(t) = \exp\Bigg(&\int_0^t \big(h(s, Y_s(\cdot,\omega), X_s(\cdot,\omega)\big)\,\mathrm{d}Y(s,\omega)\Bigg) \\ &- \frac{1}{2}\int_0^t \Big|h\big(s, Y_s(\cdot,\omega), X_s(\cdot,\omega)\big)\Big|^2\,\mathrm{d}s \end{aligned} \tag{4}$$

for $t \in [0, T]$.

We make the following assumptions:

(A1) For $t > 0$, $n \in \mathbb{N}$ and for a measurable function $\rho : \Omega_2 \to [0,1]$ such that

$$\rho(y) = 0 \quad \text{if} \quad \sup_{0\le t\le s} |\, y(s) \,| > n,$$

we have

$$E\Big[\rho(Y)\int_0^t \big|h\big(s, Y_s(\cdot,\omega), X_s(\cdot,\omega)\big)\big|^2\,\mathrm{d}s\Big] < \infty.$$

(A2) $E[Z(t)^{-1}] = 1$ for each $t \ge 0$.

(A3) The coefficients b, f, g and h are uniformly bounded by a constant c.

Having Assumption (A2), we define a new probability law P^0 on $(\Omega, \mathcal{F})$ by

$$\frac{\mathrm{d}P^0}{\mathrm{d}P}\Big|_{\mathcal{F}_t} = Z(t)^{-1}, \quad t \geq 0. \tag{5}$$

We know (Pardoux, 1989, p. 13) that for each $t \geq 0$, $\xi \in L^1(\Omega, \mathcal{F}_t, P)$ we then have $\xi Z(t) \in L^1(\Omega, \mathcal{F}_t, P^0)$ and

$$E(\xi \mid \mathcal{Y}_t) = \frac{E^0(\xi Z(t) \mid \mathcal{Y}_t)}{E^0(Z(t) \mid \mathcal{Y}_t)},$$

where $\mathcal{Y}_t = \sigma\{Y(s) : 0 \leq s \leq t\}$, E^0 being the conditional expectation operator under P^0.

Let $M_+(\mathbb{R}^M)$ denote the space of finite measures on $\mathbb{R}^M$. We define the processes $\{\zeta(t), t \geq 0\}$ and $\{\Pi(t), t \geq 0\}$ with values in $M_+(\mathbb{R}^M)$ by

$$\zeta(t)(\varphi) = E^0\big(\varphi(X(t))Z(t) \mid \mathcal{Y}_t\big) \tag{6}$$

and

$$\Pi(t)(\varphi) = E\big(\varphi(X(t)) \mid \mathcal{Y}_t\big) \tag{7}$$

for $t \geq 0$, and $\varphi \in C_b(\mathbb{R}_+, \mathbb{R}^M)$. The space $C_b(\mathbb{R}_+, \mathbb{R}^M)$ is endowed with the topology of the uniform convergence.

Let us remark that $\zeta(0) = \Pi(0) =$ law of $X(0)$. We introduce some families of partial differential operators indexed by $(t, y) \in \mathbb{R}_+ \times \Omega_2$ for $\varphi \in C_b^2(\mathbb{R}_+, \mathbb{R}^M)$, $y \in C(I, \mathbb{R}^N)$, $x \in C(I, \mathbb{R}^M)$:

$$\begin{aligned} L_{(t,y)}\varphi(x) = {} & \frac{1}{2}a^{ij}(t, y, x)\frac{\partial^2\varphi}{\partial x^i \partial x^j}(x) \\ & + b^i(t, y, x)\frac{\partial\varphi}{\partial x^i}(x), \end{aligned} \tag{8}$$

$$A^j_{(t,y)}\varphi(x) = f^{lj}(t, y, x)\frac{\partial\varphi}{\partial x^l}(x), \tag{9}$$

$$B^i_{(t,y)}\varphi(x) = g^{li}(t, y, x)\frac{\partial\varphi}{\partial x^l}(x) \tag{10}$$

and

$$L^i_{(t,y)}\varphi(x) = h^i(t, y, x)\varphi(x) + B^i_{(t,y)}\varphi(x) \tag{11}$$

for $i = 1, \dots, N$ and $j = 1, \dots, M$. We have used here the convention of repeated indices summation.

Now we are in a position to formulate the so-called Zakai equation in §3 (see Theorem 2.2.3 in (Pardoux, 1989; Chaleyat-Maurel, 1990) for the case without delay):

$$\begin{aligned} \zeta(t)(\varphi) = {} & \zeta(0)(\varphi) + \int_0^t \zeta(s)\big(L_{(s,Y)}\varphi\big)\,\mathrm{d}s \\ & + \int_0^t \zeta(s)\big(L^i_{(s,Y)}\varphi\big)\,\mathrm{d}Y^i(s) \end{aligned} \tag{12}$$

for every $\varphi \in C_b^2(\mathbb{R}_+, \mathbb{R}^M)$ if all coefficients of Eqns. (1)–(2) are bounded.

Note that this is a stochastic linear parabolic partial differential equation because of the form of the operator $L_{(t,y)}\varphi(x)$.

Let us introduce the normalized law by

$$\widetilde{\mu}(t)(\varphi) = E^0\big(\varphi(X(t))Z(1) \mid \mathcal{Y}_t\big). \tag{13}$$

The corresponding equation for the densities of the conditional probabilities Π cf. (7) can also be established. For the case without delay it is called the Kushner-Stratonovich equation (see, e.g., Pardoux, 1989).

3. Zakai Equation

Theorem 1. *Let all coefficients in (1)–(2) be bounded. Then for every $\varphi \in C_b^2(\mathbb{R}_+, \mathbb{R}^M)$ the solution of (1)–(2) satisfies the Zakai equation (12).*

Proof. From (1) and (2) we have

$$\mathrm{d}W(t) = \mathrm{d}Y(t) - h\big(t, Y_t(\cdot), X_t(\cdot)\big)\mathrm{d}t.$$

From this we obtain the following relation:

$$\begin{aligned} & \int_0^t g\big(s, Y_s(\cdot,\omega), X_s(\cdot,\omega)\big)\,\mathrm{d}W(s) \\ & = \int_0^t g\big(s, Y_s(\cdot,\omega), X_s(\cdot,\omega)\big)\,\mathrm{d}Y(s) \\ & \quad - \int_0^t g\Big(\big(s, Y_s(\cdot,\omega), X_s(\cdot,\omega)\big) \\ & \qquad \times h\big(s, Y_s(\cdot,\omega), X_s(\cdot,\omega)\big)\Big)\,\mathrm{d}s. \end{aligned} \tag{14}$$

Using (14) we get

$$\begin{aligned} X(t) = {} & X_0 + \int_0^t \Big[b\big(s, Y_s(\cdot,\omega), X_s(\cdot,\omega)\big) \\ & - g\big(s, Y_s(\cdot,\omega), X_s(\cdot,\omega)\big) \\ & \times h\big(s, Y_s(\cdot,\omega), X_s(\cdot,\omega)\big)\Big]\,\mathrm{d}s \\ & + \int_0^t f\big(s, Y_s(\cdot,\omega), X_s(\cdot,\omega)\big)\,\mathrm{d}V(s) \\ & + \int_0^t g\big(s, Y_s(\cdot,\omega), X_s(\cdot,\omega)\big)\,\mathrm{d}Y(s). \end{aligned} \tag{15}$$

Using the Itô formula for the multidimensional case (see Liptser and Shiryayev, 1977), we obtain

$$\begin{aligned}
&\mathrm{d}\varphi(X(t))\\
&= \Big[\varphi'_x\Big(b\big(t, Y_t(\cdot), X_t(\cdot)\big) - g\big(t, Y_t(\cdot,\omega), X_t(\cdot,\omega)\big)\\
&\qquad \times h\big(t, Y_t(\cdot,\omega), X_t(\cdot,\omega)\big)\Big)\\
&\qquad + \frac{1}{2}\varphi''_{xx}\Big(f\circ f^*\big(t, Y_t(\cdot), X_t(\cdot)\big)\\
&\qquad + g\circ g^*\big(t, Y_t(\cdot), X_t(\cdot)\big)\Big) Big]\,\mathrm{d}t\\
&\qquad + \varphi'_x f\big(t, Y_t(\cdot), X_t(\cdot)\big)\,\mathrm{d}V(t)\\
&\qquad + \varphi'_x g\big(t, Y_t(\cdot), X_t(\cdot)\big)\,\mathrm{d}Y(t)\\
&= L_{(t,Y_t)}\varphi\big(X(t)\big)\,\mathrm{d}t - h^i\big(t, Y_t(\cdot,\omega), X_t(\cdot,\omega)\big)\\
&\qquad \times B^i_{(t,Y_t)}\varphi\big(X(t)\big)\,\mathrm{d}t + A^l_{(t,Y_t)}\varphi\big(X(t)\big)\,\mathrm{d}V^l(t)\\
&\qquad + B^i_{(t,Y_t)}\varphi\big(X(t)\big)\,\mathrm{d}Y^i(t). \qquad (16)
\end{aligned}$$

Writing the above equation in an integral form we have

$$\begin{aligned}
\varphi(X(t)) = \varphi(X_0) &+ \int_0^t L_{(s,Y_s)}\varphi\big(X(s)\big)\,\mathrm{d}s\\
&- \int_0^t h^i\big(s, Y_s(\cdot), X_s(\cdot)\big) B^i_{(s,Y_s)}\varphi\big(X(s)\big)\,\mathrm{d}t\\
&+ \int_0^t A^l_{(s,Y_s)}\varphi\big(X(s)\big)\,\mathrm{d}V^l(s)\\
&+ \int_0^t B^i_{(s,Y_s)}\varphi\big(X(s)\big)\,\mathrm{d}Y^i(s).
\end{aligned}$$

From the Girsanov theorem (see Liptser and Shiryayev, 1977), we have

$$Z(t) = 1 + \int_0^t Z(s)h^i\big(s, Y_s(\cdot), X_s(\cdot)\big)\,\mathrm{d}Y^i(s).$$

Using once more the Itô formula for the multidimensional case for $f(t, x_1, x_2) = x_1 \cdot x_2$, we get

$$\begin{aligned}
Z(t)\varphi(X(t)) = \varphi(X_0) &+ \int_0^t Z(s)L_{(s,Y_s)}\varphi\big(X(s)\big)\,\mathrm{d}s\\
&+ \int_0^t Z(s)A^l_{(s,Y_s)}\varphi\big(X(s)\big)\,\mathrm{d}V^l(s)\\
&+ \int_0^t Z(s)L^i_{(s,Y_s)}\varphi\big(X(s)\big)\,\mathrm{d}Y^i(s).
\end{aligned}$$

Taking the expected value $E^0(\cdot \mid \mathcal{Y})$ of both the sides and using Lemma 2.2.4 from (Pardoux, 1989), we have

$$E^0\Big(\int_0^t U(s)\,\mathrm{d}Y^i(s) \mid \mathcal{Y}\Big) = \int_0^t E^0\big(U(s) \mid \mathcal{Y}\big)\,\mathrm{d}Y^i(s)$$

and

$$E^0\Big(\int_0^t U(s)\,\mathrm{d}Y^j(s) \mid \mathcal{Y}\Big) = 0$$

for $t \ge 0,\ i = 1,\dots,N,\ j = 1,\dots,M$ and for a progressively measurable process $\{U(t), t \ge 0\}$. From the definition of $\zeta(t)(\varphi)$ we get

$$\begin{aligned}
\zeta(t)(\varphi) = \zeta(0)(\varphi) &+ \int_0^t \zeta(s)(L_{(s,Y)}\varphi)\,\mathrm{d}s\\
&+ \int_0^t \zeta(s)(L^i_{(s,Y)}\varphi)\,\mathrm{d}Y^i(s). \qquad \blacksquare
\end{aligned}$$

The existence and uniqueness of the solution of (12) follows, e.g., from the classical result of (Pardoux, 1979; Bensoussan *et al.*, 1990).

4. Approximation Results of the Wong-Zakai Type

We recall that for our numerical computations we shall need the approximation result of the Wong-Zakai type (Wong and Zakai, 1965) of our filtering problem when the noise in our Zakai equation is replaced by its polygonal approximations. In practice we obtain the "real observations" as a result of measurements of the process $Y(t)$. But then, instead of the observations $\{Y(t) : s \le t\}$, we obtain the paths $\{Y_n(t) : s \le t\}$, where the processes $Y_n(t)$ have bounded variations and they are approximations of $Y(t)$. Using real $Y_n(t)$ instead of $Y(t)$, we solve the approximate equations with the operator (11), i.e., we solve the equations

$$\begin{aligned}
\zeta_n(t)(\varphi) = \zeta_n(0)(\varphi) &+ \int_0^t \zeta_n(s)(L_{(s,Y_n)}\varphi)\,\mathrm{d}s\\
&+ \int_0^t \zeta_n(s)(L^i_{(s,Y_n)}\varphi)\,\mathrm{d}Y^i_n(s). \qquad (17)
\end{aligned}$$

So we obtain $\zeta_n(t)(\varphi)$ as the solutions and, consequently, we obtain the densities $p_n(t)(\varphi) = \mathrm{d}\zeta_n(t)(\varphi)/\mathrm{d}x$.

In our theorem we shall show that if $W_n(t) \to W(t)$ and so $Y_n(t) \to Y(t)$, in a certain sense, as $n \to \infty$, then also $\zeta_n(t)(\varphi) \to \zeta(t)(\varphi)$ in an appropriate sense.

We shall further see that applying the Galerkin technique we shall obtain from (12) a finite multidimensional system of stochastic ordinary differential equations with delay (Ahmed and Radaideh, 1997).

So now we start from the investigation of a stochastic ordinary differential equation with delay (in a more general form, i.e., the stochastic functional differential equation when the delay is not constant with respect to time).

Let us restrict our deliberations to $t \in [0,T]$. For $J = (-\infty, 0]$ we introduce some metric spaces $C_- = C(J, \mathbb{R}^d)$, $C_1 = C((-\infty,T],\mathbb{R}^d)$ and $C_2^0 = C((-\infty,T],\mathbb{R}^m) = \widetilde{\Omega}$ of continuous functions. The space C_- is endowed with the metric

$$(f,g)_{C_-} = \sum_{n=1}^{\infty} 2^{-n} \frac{\|f-g\|_n}{1+\|f-g\|_n}$$

for $f,g \in C_-$, $\|h\|_n = \max_{-n\le t\le 0} h(t)$.

For further consideration we also set $I = [-r, 0]$, $0 < r < \infty$, and we introduce the norm spaces $C_- = C(I,\mathbb{R}^d)$, $C_1 = C([-r,T],\mathbb{R}^d)$ and $C_2^0 = C([-r,T],\mathbb{R}^m) = \widetilde{\Omega}$ of continuous functions with the usual norms of the uniform convergence.

Here d is the dimension of the state space and m is the dimension of the Wiener process; in the space C_2^0 all functions are equal to zero at zero.

Below we denote by $\mathcal{X}$ one of the above spaces. Let $\mathfrak{F}(\mathcal{X})$ denote the Borel σ-algebra of the space $\mathcal{X}$. It is obvious that C_2^0 is identical with the σ-algebra generated by the family of all Borel cylinder sets in $\mathcal{X}$ (see Ikeda and Watanabe, 1991). So we construct the Wiener space $(C_2^0, \mathfrak{B}(C_2^0), P^w)$, where P^w is a Wiener measure. The coordinate process $B(t,w) = w(t)$, $w \in C_2^0$, is an m-dimensional Wiener process.

The smallest Borel algebra that contains $\mathfrak{B}_1, \mathfrak{B}_2, \ldots$ is denoted by $\mathfrak{B}_1 \vee \mathfrak{B}_2 \vee \ldots$; $\mathfrak{B}_{u,v}(X)$ denotes the smallest Borel $\sigma-$algebra for which a given stochastic process $X(t)$ is measurable for every $t \in [u,v]$, and $\mathfrak{B}_{u,v}(\mathrm{d}B)$ denotes the smallest Borel algebra for which $B(s)-B(t)$ is measurable for every (t,s) with $u \le t \le s \le v$.

Let $B^n(t,w) = w_n(t)$ be the following piecewise linear approximation of $B(t,w) = w(t)$:

$$B^{n,p}(t,w) = w^p\left(\frac{k}{2^n}\right) + 2^n\left(t - \frac{k}{2^n}\right) \times \left[w^p\left(\frac{k+1}{2^n}\right) - w^p\left(\frac{k}{2^n}\right)\right] \quad (18)$$

for each $p = 1, \ldots, m$ and $kT/2^n \le t < (k+1)T/2^n$ for $k = 0, 1, \ldots, 2^n - 1$.

Now we consider $\widetilde{\Omega} = C_2^0$. Let X be a continuous stochastic process $X(t,w)$: $[-r,T] \times \Omega \to \mathbb{R}^d$, i.e., X: $\widetilde{\Omega} \to \mathcal{X} = C_1$. We take some fixed initial constant stochastic processes for $\theta \in J$ for $i = 1, \ldots, d$: $X^i(0+\theta, w) = X_0^i(w) = X_0^{n,i}(w) = Y_0^i(w)$.

We also consider operators $b: C_- \to \mathbb{R}^d$, $\sigma: C_- \to L(\mathbb{R}^m, \mathbb{R}^d)$ (where $L(\mathbb{R}^m,\mathbb{R}^d)$ is the Banach space of linear functions from $\mathbb{R}^m$ to $\mathbb{R}^d$ with the uniform operator norm $|\cdot|_L$).

In order to give a meaning to the stochastic integrals in (19) below, we introduce the following condition:

(A4) for every $t \in (-\infty, T]$ the algebra $\mathfrak{B}_{-\infty,t}(X) \vee \mathfrak{B}_{-\infty,t}(\mathrm{d}B)$ is independent of $\mathfrak{B}_{t,T}(\mathrm{d}B)$.

We consider the following stochastic functional differential equation:

$$X^i(t,w) = X_0^i + \int_0^t b^i\big(X_s(\cdot,w)\big)\,\mathrm{d}s + \sum_{p=1}^m \int_0^t \sigma^{ip}\big(X_s(\cdot,w)\big)\,\mathrm{d}w^p(s) \quad (19)$$

for $i = 1, \ldots, d$.

Replacing the Wiener process by B^n, we obtain the following approximations of (19):

$$X^{n,i}(t,w) = X_0^{n,i} + \int_0^t b^i\big(X_s^n(\cdot,w)\big)\,\mathrm{d}s + \sum_{p=1}^m \int_0^t \sigma^{ip}\big(X_s^n(\cdot,w)\big)\dot{B}^{n,p}(s,w)\,\mathrm{d}s. \quad (20)$$

We also introduce another stochastic differential equation:

$$\begin{aligned} Y^i(t,w) &= Y_0^i(w) + \int_0^t b^i\big(Y_s(\cdot,w)\big)\,\mathrm{d}s \\ &\quad + \sum_{p=1}^m \int_0^t \sigma^{ip}\big(Y_s(\cdot,w)\big)\,\mathrm{d}w^p(s) \\ &\quad + \frac{1}{2}\sum_{p=1}^m\sum_{j=1}^d \int_0^t \widetilde{D}_j\sigma^{ip}\big(Y_s(\cdot,w)\big)\sigma^{jp}\big(Y_s(\cdot,w)\big)\mathrm{d}s \end{aligned} \quad (21)$$

for every $i = 1, \ldots, d$, where the last term on the right-hand side of (21) is the so-called correction term that is described as follows (Twardowska, 1991; 1993):

Let $D\sigma^{ip}$ denote the Fréchet derivative from C_- to $L(C_-, \mathbb{R})$ (the necessary assumptions are given below). From the Riesz theorem it follows that there exists a family of measures $\mu = \mu_g^{ipj}$ of bounded variation such that

$$D\sigma^{ip}(g)(\Phi) = \sum_{j=1}^d \int_{-r}^0 \Phi_j(v)\mu_g^{ipj}(\mathrm{d}v)$$

is a directional derivative for any Φ, $g \in C_-$. The measure μ has the following decomposition:

$$\begin{aligned} \mu(A) &= \mu(A \cap (-\infty, 0)) + \mu(A \cap \{0\}) \\ &= \widetilde{\mu}(A) + \mu(\{0\})\delta_0(A), \end{aligned}$$

where δ_0 is the Dirac measure, $A \in \mathfrak{B}((-\infty, 0))$. We denote by $\widetilde{D}_j\sigma^{ip}(g)$ the value $\mu_g^{ipj}(\{0\})$, i.e.,

$$\widetilde{D}_j\sigma^{ip}\big(\xi_s(\cdot)\big) = \mu_g^{ipj}\big(\{0\}\big). \qquad (22)$$

The second integral in (21) is the Itô integral.

Let us introduce the following conditions:

(A5) The initial stochastic process X_0 is $\mathfrak{F}_0$-measurable and $P(|X_0(w)| < \infty) = 1$, where $|X_0(w)| = \sum_{j=1}^{d} |X_0^i(w)|$, and $\mathfrak{B}_{-\infty,0}(X_0)$ is independent of $\mathfrak{B}_{0,T}(B)$;

(A6) For any $\varphi,\ \psi \in C_-$ the following Lipschitz condition is satisfied:

$$\begin{aligned}\big|b(\varphi) - b(\psi)\big|^2 + \big|\sigma(\varphi) - \sigma(\psi)\big|_L^2 &\\ \le L^1 \int_{-\infty}^{0} \big|\varphi(\theta) - \psi(\theta)\big|^2 \,\mathrm{d}K(\theta)&\\ + L^2\big|\varphi(0) - \psi(0)\big|^2,&\end{aligned}$$

where $K(\theta)$ is a certain bounded measure on J, and L^1, L^2 are some constants;

(A7) For every $\varphi,\ \psi \in C_-$ the following growth condition is satisfied:

$$\begin{aligned}\big|b(\varphi)\big|^2 + \big|\sigma(\varphi)\big|_L^2 \le L^1 \int_{-\infty}^{0} \big(1 + \varphi^2(\theta)\big) \,\mathrm{d}K(\theta)&\\ + L^2\big(1 + \varphi^2(0)\big),&\end{aligned}$$

where $\varphi^2(0) = \sum_{j=1}^{d} \varphi_i^2(0)$;

(A8) We have

$$P\Big(\int_0^T \big|b(X_s)\big| \,\mathrm{d}s < \infty\Big) \,=\, 1,$$

$$P\Big(\int_0^T \big|\sigma(X_s)\big|_L^2 \,\mathrm{d}s < \infty\Big) \,=\, 1;$$

(A9) Let $b^i,\ \sigma^{ip}$ be bounded functions and b^i, $\sigma^{ip} \in C^1$, for all $i = 1, \dots, d,\ p = 1, \dots, m$.

We say that a $d-$dimensional continuous stochastic process $X : (-\infty, T] \times C_2^0 \to \mathbb{R}^d$ is a *strong solution* of (19) for a given process $w(t)$ if Conditions (A4), (A5) and (A8) are satisfied and (19) is valid with probability 1 for all $t \in (-\infty, T]$. The uniqueness of strong solutions is understood in the sense of the trajectories:

An absolutely continuous stochastic process X^n : $(-\infty, T] \times C_2^0 \to \mathbb{R}^d$ is a solution of (20) if Conditions (A4) and (A5) are satisfied and (20) is valid with probability 1 for all $t \in (-\infty, T]$.

Notice that our conditions ensure the existence and uniqueness of the strong solution Y of (21) since $\widetilde{D}_j\sigma^{ip}(Y_t(\cdot, w))$ is a real number (it is a value of a measure). Moreover, for every $n \in \mathbb{N}$, there exists exactly one solution of the ordinary differential equation (20).

We have the following approximation theorem of the Wong-Zakai type for stochastic functional differential equations (Twardowska, 1991; 1993):

Theorem 2. *Let Conditions (A4)–(A7) be satisfied. Let $B^n(t, w)$ be an approximation of the type (18) of a Wiener process. We assume that X^n and Y are solutions of (20) and (21), respectively, with a constant initial stochastic process. Then Conditions (A4) and (A8) are satisfied and for every $\varepsilon > 0$ we have*

$$\lim_{n\to\infty} P\Big[\sup_{0\le t\le T} \big|X^n(t, \omega) - Y(t, \omega)\big|_H > \varepsilon\Big] = 0. \quad (23)$$

Remark 1. The proof in (Twardowska, 1991; 1993) is given for the interval $J = (-\infty, 0]$. Instead of $J = (-\infty, 0]$, we can consider $I = [-r, 0]$, $r > 0$. Then, instead of considering $X^i(t_i^n + s) - X^i(t_{i-1}^n + s)$ on the whole interval of the definition of time, we divide it into some parts (see Twardowska, 1993) and we estimate each part separately by expressions converging to zero.

For example, consider the initial equation

$$\begin{aligned}\mathrm{d}X(t) &= b(X_t)\,\mathrm{d}t + \sigma(X_t)\,\mathrm{d}w(t),\\ X_0(\theta, \omega) &= \eta(\omega) \quad \text{for} \quad \theta \in J,\end{aligned} \qquad (24)$$

where for some constants $b_0,\ b_1,\ \sigma_0,\ \sigma_1$ we define $b, \sigma : C_- \to \mathbb{R}$ as follows:

$$\begin{aligned}b(\varphi) &= b_0\varphi(0) + b_1\varphi(-r),\\ \sigma(\varphi) &= \sigma_0\varphi(0) + \sigma_1\varphi(-r).\end{aligned}$$

We note that $\varphi(0) = X_t(0) = X(t)$, $\varphi(-r) = X_t(-r) = X(t - r)$ and

$$\begin{aligned}\mathrm{d}X(t) = &\ \big(b_0X(t) + b_1X(t-r)\big)\,\mathrm{d}t\\ &+ \big(\sigma_0X(t) + \sigma_1X(t-r)\big)\,\mathrm{d}w(t), \quad (25)\\ X_0 = &\ \eta.\end{aligned}$$

Then the limit equation (21) takes on the form

$$\begin{aligned}\mathrm{d}Y(t) = &\ \big(b_0Y(t) + b_1Y(t-r)\big)\,\mathrm{d}t\\ &\big(\sigma_0Y(t) + \sigma_1Y(t-r)\big)\,\mathrm{d}w(t) \qquad (26)\\ &+ \frac{1}{2}\sigma_0\big(\sigma_0Y(t) + \sigma_1Y(t-r)\big)\,\mathrm{d}t,\\ Y_0 = &\ \eta\end{aligned}$$

because $\sigma_0X(t)$ is the only term for which the support of the measure contains zero. Therefore $\mu(\{0\}) = \sigma_0$.

Now we shall come back to our Zakai stochastic linear parabolic partial differential equation (11). We have the following approximation theorem of the Wong-Zakai type (Twardowska and Pasławska-Południak, 2003).

Theorem 3. *Let Conditions (A4)–(A7) be satisfied. Let* $B^n(t,w)$ *be an approximation of the type (18) of a Wiener process. We assume that* ζ *and* ζ_n *are solutions of (12) and (17), respectively, with a constant initial stochastic process, and also*

$$\begin{aligned}\widetilde{\zeta}(t)(\varphi) &= \widetilde{\zeta}(0)(\varphi) + \int_0^t \widetilde{\zeta}(s)(L_{(s,Y_n)}\varphi)\,\mathrm{d}s \\ &\quad + \int_0^t \widetilde{\zeta}(s)(L^i_{(s,Y_n)}\varphi)\,\mathrm{d}Y^i(s) \\ &\quad + \frac{1}{2}\int_0^t \widetilde{\zeta}(s)(\widetilde{D}L^i_{(s,Y_n)}\varphi)(L^i_{(s,Y_n)}\varphi)\,\mathrm{d}s, \end{aligned} \tag{27}$$

where the last term is the so-called correction term of the form (22). Then for every $t \geq 0$ *we have*

$$\lim_{n\to\infty} E\big|\zeta_n(t,\omega)(\varphi) - \widetilde{\zeta}(t,\omega)(\varphi)\big|^2 = 0. \tag{28}$$

Proof. For a proof of the Wong-Zakai type theorem for stochastic partial differential equations in Hilbert spaces, without delay, see (Twardowska, 1995). The convergence is of the type (28). The case of the nonlinear filtering equation (12) without delay is covered by the theorem which we can be found in the paper (Pardoux, 1975, pp. 130–131). Now the technique of proving the Wong-Zakai theorem with delay can be copied from (Twardowska, 1991; 1993). We get the convergence of the type $\lim_{n\to\infty} E(\sup_t |\zeta_n(t,\omega)(\varphi) - \widetilde{\zeta}(t,\omega)(\varphi)|^2) = 0$ but the convergence in (28) is weaker, so we prove (28) in our theorem. ■

5. Approximation Result for the Zakai Equation

From the numerical point of view, it is convenient to consider the Zakai equation (12) in the Stratonovich form (Dawidowicz and Twardowska, 1995), i.e., subtracting the correction term appearing in (27). Then, after the Wong-Zakai approximation, we will obtain a limit equation without a correction term.

First, to obtain a system of stochastic ordinary differential equations from our Zakai equation, we apply the Galerkin method. We follow the idea of Ahmed and Radaideh (1997, §3.3). Therefore, using the Galerkin method based on the Fourier coefficients $\{\psi_i^N\}$ and projecting the Zakai equation onto the space spanned by $\{w_i, 1 \leq i \leq N\}$ (see Ahmed and Radaideh, 1997, §3.2, Eqn. (8)) we can approximate the solution of (12) in the form

$$\zeta^N(t,x) = \sum_{i=1}^N \zeta^N(t) w_i(t)$$

and then we obtain a system of stochastic ordinary differential equations in a matrix form. In our case it is a system of linear stochastic ordinary differential equations with delay, so we can use the theory from §4.

6. Numerical Experiments

We start with the following filtering problem:

$$\begin{aligned}\mathrm{d}X(t) &= \big[b_0X^2(t) + b_1X^2(t-1)\big]\,\mathrm{d}t \\ &\quad + \big[\sigma_0X(t) + \sigma_1X^2(t-1)\big]\,\mathrm{d}W(t),\end{aligned} \tag{29a}$$

$$\mathrm{d}Y(t) = \big[a_0X(t) + a_1X^2(t-1)\big]\,\mathrm{d}t + \mathrm{d}W(t), \tag{29b}$$

where a_0, a_1 b_0, b_1, σ_0 and σ_1 are some constants, $X(t) \in \mathbb{R}$, $Y(t) \in \mathbb{R}$ and $W(t)$ is the one-dimensional Wiener process. We transform this problem to the following stochastic partial differential equation of the Zakai type (12):

$$\begin{aligned}\varphi_t = &\Big[\frac{1}{2}\sigma_0^2X^2(t) + \sigma_0\sigma_1X(t)X^2(t-1) \\ &+ \frac{1}{2}\sigma_1^2X^4(t-1)\Big]\varphi''_{xx} \\ &+ \big[b_0X^2(t) + b_1X^2(t-1)\big]\varphi'_x \\ &+ \big[a_0^2X^2(t) + 2a_0a_1X(t)X^2(t-1) \\ &+ a_1^2X^4(t-1)\big]\varphi \\ &+ \big[a_0X(t) + a_1X^2(t-1)\big]\,\mathrm{d}W(t)\end{aligned} \tag{30}$$

and the correction term is of the form (cf. (22))

$$\frac{1}{2}a_0\big[a_0X(t) + a_1X^2(t-1)\big]\,\mathrm{d}t.$$

After discretization (see §5 and Ahmed and Radaideh, 1997), we can restrict our analysis to the following stochastic ordinary differential equation with delay on the interval $[0,1]$:

$$\begin{aligned}\mathrm{d}X(t) &= \big(aX(t) + bX^2(t) + c\big)\,\mathrm{d}t \\ &\quad + \big(a_0X(t) + a_1\big)\,\mathrm{d}W(t),\end{aligned} \tag{31}$$

$$X_0(\theta) = X(0+\theta) = 1 \text{ for } \theta \in [-1,0],$$

$$X(t-1) = 1 \text{ on } [0,1] \text{ as } t-1 \in [-1,0],$$

where

$$a = \sigma_0\sigma_1 + 2a_0a_1,$$
$$b = \frac{1}{2}\sigma_0^2 + b_0 + a_0,$$
$$c = \frac{1}{2}\sigma_1^2 + b_1 + a_1^2.$$

We solve this equation with the following numerical methods: Euler, Milshtein and Runge-Kutta schemes (Kloeden and Platen, 1992; Sobczyk, 1991). But for our case of stochastic differential equations with delay, we modify the Milshtein scheme. It is well known that the Milshtein scheme can be obtained as the Euler scheme for the Stratonovich version of (31) using the relation for the transition between the Itô and Stratonovich integrals (Dawidowicz and Twardowska, 1995).

Below we present some numerical computations to confirm our theoretical result that the correction term plays a crucial role in numerical schemes, too.

Consider (31) with a_0, a_1, b_0, b_1, σ_0 and σ_1 given by

$$a_0 = 1, \quad a_1 = \frac{1}{2}, \quad b_0 = -2, \quad b_1 = \frac{1}{2},$$
$$\sigma_0 = \sqrt{2}, \quad \sigma_1 = -\frac{\sqrt{2}}{4}.$$

Then $b = \frac{1}{2}\sigma_0^2 + b_0 + a_0 = 0$. Equation (31) without the correction term has the following form:

$$\mathrm{d}X(t) = \left(\frac{1}{2}X(t) + \frac{13}{16}\right)\mathrm{d}t + (X(t) + \frac{1}{2})\,\mathrm{d}W(t). \tag{32}$$

Equation (31) with the correction term is

$$\mathrm{d}X(t) = \frac{9}{16}\,\mathrm{d}t + \big(X(t) + \frac{1}{2}\big)\,\mathrm{d}W(t). \tag{33}$$

First, we obtain an exact analytical formula for $t \in [0,1]$ in the so-called step method (see §4). We use the form of the solution derived for the linear equation (4.9), pp. 119-120 in the book by Kloeden and Platen (1992), i.e., for Eqn. (31) with $b = 0$. We have

$$X(t) = \Phi(t)\Big[X(0) + (c - a_0a_1)\int_0^t \Phi(s)^{-1}\,\mathrm{d}s + a_1\int_0^t \Phi(s)^{-1}\,\mathrm{d}W(s)\Big]$$

with the fundamental solution

$$\Phi(t) = \exp\left[(a - \frac{1}{2}a_0^2)t + a_0W(t)\right].$$

In our case

$$\Phi(t) = \exp\big(W(t)\big) \quad \text{and} \quad X(0) = 1.$$

So we obtain the following solution to (31) for $t \in [0,1]$:

$$X(t) = \exp\big(W(t)\big)\Big(1 + \frac{5}{16}\int_0^t \exp\big(-W(s)\big)\,\mathrm{d}s + \frac{1}{2}\int_0^t \exp\big(-W(s)\big)\,\mathrm{d}s\Big). \tag{34}$$

We recall that in the step method we set $X(t-1) = 1$ for $t \in [0,1]$, so $(t-1) \in [-1,0]$. We have also used the following formula (Kloeden and Platen, 1992, p. 101):

$$\int_0^t \exp\big(-W(s)\big)\,\mathrm{d}W(s) = U\big(W(t)\big) - U\big(W(0)\big) - \frac{1}{2}\int_0^t h'(W(s))\,\mathrm{d}s,$$

where

$$h(x) = \exp(-x), \quad U'(x) = h(x).$$

This solution is used to test and compare numerical methods in this paper. We solve the stochastic differential equation numerically by the simulation of the approximation of discrete trajectories in time. To construct a solution for a given discretization $t_0 = 0 < t_1 < \cdots < t_N = T$ we used the Euler and Milshtein methods. We modified the recursive formulae for the Milshtein method taking into consideration the delayed argument.

The Euler approximation for (29) is generated recursively by

$$Y_{n+1} = Y_n + (b_0Y_n + b_1Y_{n-k})\Delta_n + (\sigma_0Y_n + \sigma_1Y_{n-k})\Delta W_n \tag{35}$$

for $n = k+1,\ k+2, \ldots, N-1$ with initial values $Y_0 = Y_1 = \cdots = Y_k = 1$ and $\Delta_n = T/N$ (equidistant step size), $k = 1/\Delta_n$ (an integer parameter related to the delay), $\Delta W_n = W_{t_{n+1}} - W_{t_n}$.

The random variables ΔW_n are independently $\mathcal{N}(0,1)$-normally distributed random variables. We have generated such random variables in simulations from independent and uniformly distributed random variables on [0,1] which are provided by a pseudorandom number generator on a computer. The generation of the sample paths of the process $W(t)$ may be realized by $W(0) = 0$, $W(t) = \sqrt{\Delta_n}(\xi_1 + \cdots + \xi_{\frac{t}{\Delta_n}})$, where ξ_i are independent and identically $\mathcal{N}(0,1)$-normally distributed random variables.

The Milhstein approximation scheme has the modified form (see Kloeden and Platen, 1992) for the correction term

$$
\begin{aligned}
Y_{n+1} = {} & Y_n + (b_0 Y_n + b_1 Y_{n-k})\Delta_n \\
& + (\sigma_0 Y_n + \sigma_1 Y_{n-k})\Delta W_n \\
& + \frac{1}{2}\sigma_0(\sigma_0 Y_n + \sigma_1 Y_{n-k})(\Delta W_n^2 - \Delta_n). \quad (36)
\end{aligned}
$$

The Runge-Kutta approximation scheme (Kloeden and Platen, 1992) is of the form

$$
\begin{aligned}
Y_{n+1} = {} & Y_n + (b_0 Y_n + b_1 Y_{n-k})\Delta_n \\
& + (\sigma_0 Y_n + \sigma_1 Y_{n-k})\Delta W_n \\
& + \frac{1}{2}\sigma_0(\sigma_0 \widetilde{\Gamma}_n - \sigma_0 Y_n + \sigma_1 Y_{n-k}) \\
& \times (\Delta W_n^2 - \Delta_n)\Delta_n^{-1/2}, \quad (37)
\end{aligned}
$$

where $\widetilde{\Gamma}_n = Y_n + b\Delta_n^{1/2}$.

We say that the approximating process Y converges in the strong sense to the process X with the order $\gamma \in (0,\infty]$ if there exist some finite constants K and $\delta_0 \geq 0$ such that

$$E\big(|X_T - Y_N|\big)K\delta^\gamma$$

for any time discretization with the maximum step size $\delta \in (0, \delta_0)$.

In (Kloeden and Platen, 1992) it is proved that the Euler scheme has the strong order $\gamma = 0.5$ and the Milshtein scheme converges with the strong order $\gamma = 1$ (under some regularity conditions).

Our computations were performed using the MATLAB package. Figure 1 summarizes graphically the numerical experiment with Eqn. (33). It compares simulated trajectories of the examined Euler, Milhstein and Runge-Kutta schemes with the exact solution (34) of (32) for the same sample path of the Wiener process. The solid line represents the exact solution, the dotted line the Euler method, the dashed line the Milshtein method and the dotted-dashed line the Runge-Kutta method. In Fig. 2 we solve Eqn. (32) without the corretion term and we compare it with the exact solution (34) of (32). We can observe that the simulated trajectories in Fig. 1 are close to the exact solution because in (33) the correction term occurs. The results with the correction term in Fig. 1 are better.

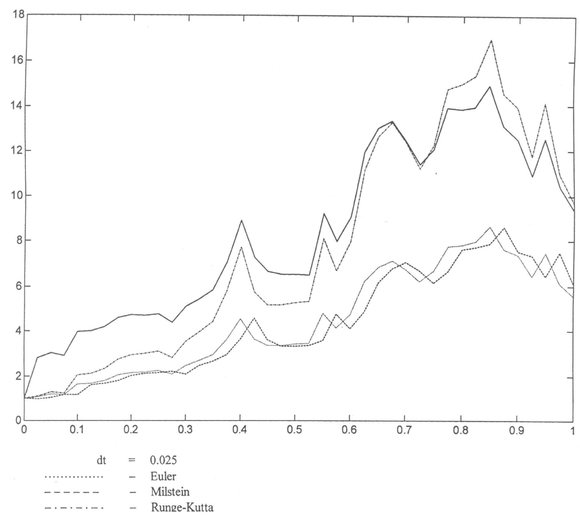

Fig. 1. Simulated trajectory of Euler, Milhstein and Runge-Kutta schemes.

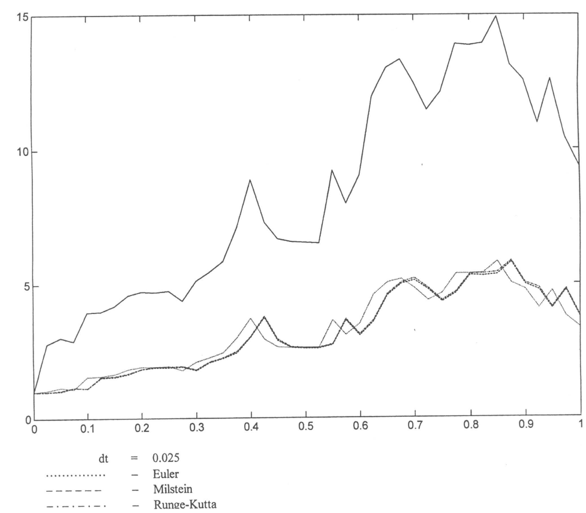

Fig. 2. Solution of Eqn. (32) without the correction term along the exact solution.

References

Ahmed N.V. and Radaideh S.M. (1997): *A powerful numerical technique solving Zakai equation for nonlinear filtering*. — Dynam. Contr., Vol. 7, No. 3, pp. 293–308.

Atar R., Viens F. and Zeituni O. (1999): *Robustness of Zakai's equation via Feynman-Kac representation*, In: Stochastic Analysis, Control, Optimization and Applications (W.M. McEneaney, G.G. Yin and Q. Zhang, Eds.). — Boston: Birkhäuser, pp. 339–352.

Beneš V.E. (1981): *Exact finite-dimensional filters for certain diffusions with nonlinear drift*. — Stochastics, Vol. 5, No. 1–2, pp. 65–92.

Bensoussan A., Głowiński R. and Rascanu A. (1990): *Approximation of the Zakai equation by the splitting up method*. — SIAM J. Contr. Optim., Vol. 28, No. 6, pp. 1420–1431.

Brzeźniak Z. and Flandoli F. (1995): *Almost sure approximation of Wong-Zakai type for stochastic partial differential equations*. — Stoch. Proc. Appl., Vol. 55, No. 2, pp. 329–358.

Bucy R.S. (1965): *Nonlinear filtering theory*. — IEEE Trans. Automat. Contr., Vol. 10, No. 2, pp. 198–212.

Chaleyat-Maurel A., Michel D. and Pardoux E. (1990): *Un théorème d'unicité pour l'équation de Zakai*. — Stoch. Rep., Vol. 29, No. 1, pp. 1–12.

Cohen de Lara M. (1998): *Reduction of the Zakai equation by invariance group techniques*. — Stoch. Proc. Appl., Vol. 73, No. 1, pp. 119–130.

Crisan D., Gaines J. and Lyons T. (1998): *Convergence of a branching particle method to the solution of the Zakai equation*. — SIAM J. Appl. Math., Vol. 58, No. 5, pp. 1568–1590.

Dawidowicz A.L. and Twardowska K. (1995): *On the relation between the Stratonovich and Itô integrals with integrands of delayed argument*. — Demonstr. Math., Vol. 28, No. 2, pp. 456–478.

Elliot R.J. and Głowiński R. (1989): *Approximations to solutions of the Zakai filtering equation*. — Stoch. Anal. Appl., Vol. 7, No. 2, pp. 145–168.

Elliot R.J. and Moore J. (1998): *Zakai equations for Hilbert space valued processes*. — Stoch. Anal. Appl., Vol. 16, No. 4, pp. 597–605.

Elsgolc L.E. (1964): *Introduction to the Theory of Differential Equations with Delayed Argument*. — Moscow: Nauka (in Russian).

Florchinger P. and Le Gland F. (1991): Time-discretization of the Zakai equation for diffusion processes observed in correlated noise. — Stoch. Stoch. Rep., Vol. 35, No. 4, pp. 233–256.

Gyöngy I. (1989): *The stability of stochastic partial differential equations and applications. Theorems on supports*, In: Lecture Notes in Mathematics (G. Da Prato and L. Tubaro, Eds.). — Berlin: Springer, Vol. 1390, pp. 99–118.

Gyöngy I. and Pröhle T. (1990): *On the approximation of stochastic partial differential equations and Stroock-Varadhan's support theorem*. — Comput. Math. Appl., Vol. 19, No. 1, pp. 65–70.

Ikeda N. and Watanabe S. (1981): *Stochastic Differential Equations and Diffusion Processes*. — Amsterdam: North-Holland.

Itô K. (1996): *Approximation of the Zakai equation for nonlinear filtering theory*. — SIAM J. Contr. Optim., Vol. 34, No. 2, pp. 620–634.

Itô K. and Nisio M. (1964): *On stationary solutions of a stochastic differential equations*. — J. Math. Kyoto Univ., Vol. 4, No. 1, pp. 1–75.

Itô K. and Rozovskii B. (2000): *Approximation of the Kushner equation*. — SIAM J. Control Optim., v.38, No.3, pp.893-915.

Kallianpur G. (1980): *Stochastic Filtering Theory*. — Berlin: Springer.

Kallianpur G. (1996): *Some recent developments in nonlinear filtering theory*, In: *Itô stochastic calculus and probability theory* (N. Ikeda, Ed.). — Tokyo: Springer, pp. 157–170.

Kloeden P. and Platen E. (1992): *Numerical Solutions of Stochastic Differential Equations*. — Berlin: Springer.

Kolmanovsky V.B. (1974): *On filtration of certain stochastic processes with aftereffects*. — Avtomatika i Telemekhanika, Vol. 1, pp. 42–48.

Kolmanovsky V., Matasov A. and Borne P. (2002): *Mean-square filtering problem in hereditary systems with nonzero initial conditions*. — IMA J. Math. Contr. Inform., Vol. 19, No. 1–2, pp. 25–48.

Kushner H.J. (1967): *Nonlinear filtering: The exact dynamical equations satisfied by the conditional models*. — IEEE Trans. Automat. Contr., Vol. 12, No. 3, pp. 262–267.

Liptser R.S. and Shiryayev A.N. (1977): *Studies of Random Processes I and II*. — Berlin: Springer.

Lototsky S., Mikulevičius R. and Rozovskii B. (1997): *Nonlinear filtering revisited: A spectral approach*. — SIAM J. Contr. Optim., Vol. 35, No. 2, pp. 435–461.

Pardoux E. (1975): *Equations aux dérivées partielles stochastiques non linéaires monotones. Etude de solutions fortes de type Itô*. — Ph. D. thesis, Sci. Math., Univ. Paris Sud.

Pardoux E. (1989): *Filtrage non linéaire et équations aux dérivées partielles stochastiques associétes*. — Preprint, Ecole d'Été de Probabilités de Saint-Fleur, pp. 1–95.

Pardoux E. (1979): *Stochastic partial differential equations and filtering of diffusion processes*. — Stochastics, Vol. 3, pp. 127–167.

Sobczyk K. (1991): *Stochastic Differential Equations with Applications to Physics and Engineering*. — Dordrecht: Kluwer.

Twardowska K. (1993): *Approximation theorems of Wong-Zakai type for stochastic differential equations in infinite dimensions*. — Dissertationes Math., Vol. 325, pp. 1–54.

Twardowska K. (1995): *An approximation theorem of Wong-Zakai type for nonlinear stochastic partial differential equations*. — Stoch. Anal. Appl., v.13, No.5, pp.601-626.

Twardowska K. and Pasławska-Południak M. (2003): *Approximation theorems of Wong-Zakai type for stochastic partial differential equations with delay arising in filtering problems*. — to appear.

Twardowska K. (1991): *On the approximation theorem of Wong-Zakai type for the functional stochastic differential equations*. — Probab. Math. Statist., Vol. 12, No. 2, pp. 319–334.

Wong E. and Zakai M. (1965): *On the convergence of ordinary integrals to stochastic integrals*. — Ann. Math. Statist., Vol. 36, pp. 1560–1564.

Zakai M. (1969): *On the optimal filtering of diffusion processes*. — Z. Wahrsch. Verw. Geb., Vol. 11, pp. 230–243.

Received: 4 February 2002
Revised: 16 January 2003