amcs

ON THE CONVERGENCE OF THE WAVELET–GALERKIN METHOD
FOR NONLINEAR FILTERING

ŁUKASZ D. NOWAK *, MONIKA PASŁAWSKA-POŁUDNIAK **, KRYSTYNA TWARDOWSKA ***

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

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

*** Faculty of Applied Informatics and Mathematics
Warsaw University of Life Sciences—SGGW, ul. Nowoursynowska 159, 02–776 Warsaw, Poland
e-mail: krystyna_twardowska@sggw.pl

The aim of the paper is to examine the wavelet-Galerkin method for the solution of filtering equations. We use a wavelet biorthogonal basis with compact support for approximations of the solution. Then we compute the Zakai equation for our filtering problem and consider the implicit Euler scheme in time and the Galerkin scheme in space for the solution of the Zakai equation. We give theorems on convergence and its rate. The method is numerically much more efficient than the classical Galerkin method.

Keywords: Zakai equation, Galerkin method, wavelet basis, Euler scheme.

1. Introduction

The Zakai equation describes the evolution of unnormalized conditional density which solves the problem of nonlinear filtering of diffusion processes. It is a linear stochastic parabolic partial differential equation. The approximation of solutions to the Zakai equation is a very important problem for practical applications. The Zakai equation in the form which we use is given, e.g., in the papers (Pardoux, 1991; Kurtz and Ocone, 1988; Yau and Yau, 2000; 2008) and in the book (Liptser and Shiryaev, 1977). The conditions for the existence and uniqueness of the solution to the Zakai equation are given in (Pardoux, 1991). We refer also to the paper (Rozovskiĭ, 1991) for existence and uniqueness theorems of the Zakai equation.

There are a lot of methods of constructing approximations of the Zakai equation: splitting up method (Elliott and Glowinski, 1989), decomposition into Wiener integrals (Crisan *et al.*, 1998), discrete time approximations (Bennaton, 1985), some generalizations in Hilbert spaces (Germani and Picconi, 1984). There are also some modifications of the Galerkin method and various bases are

used, e.g., the Gaussian series basis, in the paper (Ahmed and Radaideh, 1997). The approximation of the Zakai equation in a nonlinear filtering problem with delay was considered in the paper (Twardowska *et al.*, 2003).

The Euler and Milshtein schemes for time discretization and the Galerkin scheme with the Hermite basis were given in (Itô, 1996). Some characterizations of the stochastic Euler and Milshtein schemes can be found in the book (Kloeden and Platen, 1992), while time and space approximations of parabolic partial differential equations in (Thomée, 1997).

In this paper we give a modification of the classical Galerkin method using a biorthogonal wavelet basis with compact support. Wavelets are used to transform the problem into an equivalent one which is well-posed in the space l_2. The construction of such a basis can be found in the books (Bramble *et al.*, 2003; Cohen, 2003; Daubechies, 1992; Wojtaszczyk, 1997) and in the papers (Dahmen, 1997; Cohen *et al.*, 1992; Wang, 2002). The main advantage of wavelet methods is better efficiency of numerical calculations (see Section 5). The

Lévy-Ciesielski representation (Ciesielski, 1961; McKean, 1969) of the Wiener process is used in the paper.

It is known that the wavelet basis improves the condition number of a problem but the convergence rate is not changed.

Indeed, the Galerkin discretization in finite element spaces is ill conditioned in the sense that its condition number grows like $1/h^{2s}$, where h is the mesh size and s is the order of the corresponding Sobolev space, i.e., $2s$ is the order of the operator (see (Cohen, 2003)). But the stiffness matrix relative to the wavelet hierarchical basis is even diagonal and a simple diagonal scaling would yield uniformly bounded condition numbers independent of the mesh size h (see (Bramble *et al.*, 2003)).

In our paper we show how to use the wavelet basis in the nonlinear filtering problem and we compute the convergence rate. Therefore, we can conclude that the condition number of our problem is better than in a classical basis, but the convergence rate is not improved.

There are many papers in which the condition number is computed when a wavelet basis is used. The problem is well conditioned if the condition number is small so it does not depend on the mesh size. This is done by preconditioning in wavelet methods.

For example, in the paper (Beuchler *et al.*, 2004), the mass and stiffness matrices in a wavelet basis are built. Then the matrices that can be used as preconditioners are introduced. Estimates for the condition numbers of such matrices are built in some degenerate elliptic problems and parabolic problems from finance. Some examples of preconditioning matrices are considered.

Further, in the paper (Hilbert *et al.*, 2004), the authors introduce the pricing problem in terms of parabolic partial differential equations. They show how one can construct optimal diagonal preconditioners based on wavelet norm equivalences.

Finally, (Yserentant, 1990) shows that the spectral condition numbers for a discretized second order elliptic boundary problem should remain bounded independently of the gridsize. There are two preconditioners that meet these requirements especially well. Both preconditioners utilize a multi-level structure. The condition number estimates are computed.

In Section 2 we derive the Zakai equation for our filtering problem, and in Section 3 we give the variational formulation for it. We describe our wavelet basis in Section 4. Section 5 is devoted to the wavelet formulation of the problem and the discussion of the condition number. Namely, we give an equivalent formulation in the wavelet coordinates for our original variational problem arising from the filtering equations and next from the Zakai equation. The transformed problem can be arranged to be well-posed due to the norm equivalences introduced by the wavelet basis and the compression of the operator matrix in the variational formulation, see (Dahmen and Schnei-

der, 1999; von Petersdorff and Schwab, 1996; 2003). This allows us to obtain convergent iterative schemes for the infinite-dimensional problem over the space l_2. Finally, in Section 6 we discuss convergence and its rate.

2. Nonlinear filtering problem and the Zakai equation

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, P)$ be a probability space with a usual filtration $\mathcal{F}_t$. Denote by

$$b : \mathbb{R}^d \to \mathbb{R}^d, \quad g : \mathbb{R}^d \to \mathbb{R}^{d \times d}, \quad f : \mathbb{R}^d \to \mathbb{R}^p$$

some bounded and Lipschitz functions. Let $\sigma = (\sigma_{\alpha\beta})_{\alpha,\beta=1,\dots,p}$ be a nonsingular matrix of constants. We consider the following system of Itô equations:

$$\mathrm{d}X(t) = b(X(t), Y(t))\, \mathrm{d}t + g(X(t))\, \mathrm{d}W(t),$$
$$X(0) = X_0, \quad t \geq 0, \tag{1}$$
$$\mathrm{d}\overline{Y}(t) = f(X(t))\, \mathrm{d}t + \sigma\, \mathrm{d}V(t),$$
$$Y(0) = 0, \quad t \geq 0, \tag{2}$$

where the state $X(t) \in \mathbb{R}^d$ is the unobserved process to be estimated and $Y(t) \in \mathbb{R}^p$ is the observation process. Moreover, (W, V) are $\mathcal{F}_t$-adapted, $\mathbb{R}^d \times \mathbb{R}^p$-valued independent Wiener processes with covariances I. Furthermore, X_0 is an $\mathcal{F}_0$-measurable vector, independent of the Wiener processes W and V.

Let us write

$$\mathrm{d}Y(t) = \frac{\mathrm{d}\overline{Y}(t)}{\sigma} = \sigma^{-1} f(X(t))\, \mathrm{d}t + \mathrm{d}V(t). \tag{3}$$

Further, let $C_b = C_b(\mathbb{R}^d, \mathbb{R})$ be the space of continuous bounded functions on $\mathbb{R}^d$ with values in $\mathbb{R}$ and $C_b^1 = C_b^1(\mathbb{R}^d, \mathbb{R})$, $C_b^2 = C_b^2(\mathbb{R}^d, \mathbb{R})$ be the spaces of continuous bounded functions with continuous first and second derivatives, respectively.

Let $L^2(\mathbb{R}^d)$ be the Hilbert space, with the inner product

$$(u, v) = \int_{\mathbb{R}^d} u(x) v(x)\, \mathrm{d}x,$$

and let $H^1(\mathbb{R}^d)$ denote the Sobolev space

$$\left\{ u \in L^2(\mathbb{R}^d) : \frac{\partial u}{\partial x_i} \in L^2(\mathbb{R}^d), \quad i = 1, \dots, d \right\}$$

with the norm

$$\|u\|_{H^1(\mathbb{R}^d)} = \left(\|u\|_{L^2(\mathbb{R}^d)}^2 + \sum_{i=1}^d \left\| \frac{\partial u}{\partial x_i} \right\|_{L^2(\mathbb{R}^d)}^2 \right)^{1/2}.$$

Set $H = L^2(\mathbb{R}^d)$, $V = H^1(\mathbb{R}^d)$ and $V^* = H^{-1}(\mathbb{R}^d)$ as the dual space of V. Moreover, H^* is identified with H. Thus

$$V \subset H = H^* \subset V^*.$$

We denote the dual product of $V \times V^*$ by $\langle \cdot, \cdot \rangle$.

Let ϕ be any C_b^2 function. Our problem is to estimate $\phi(X(t))$ when the history of Y up to time t is given. Define

$$E\{\phi(X(t)) \,|\, \mathcal{F}_t^Y\} = \frac{\int_{\mathbb{R}^d} \phi(x)u(t,x)\,\mathrm{d}x}{\int_{\mathbb{R}^d} u(t,x)\,\mathrm{d}x},$$

where $\mathcal{F}_t^Y$ is the σ-field generated by $Y(s)$ for $s \leq t$, complete with respect to the measure P and $u(t) = u(t,x)$ (the unnnormalized conditional probability).

Proposition 1. (McKean, 1969; Rozovskiĭ, 1991) . *Under the assumptions that the coefficients in the system* (1), (3) *are bounded and measurable, the function* $u(t)$ *is the solution to the following so-called Zakai equation associated with* (1), (3):

$$\mathrm{d}u(t) + \widetilde{L}(Y)u(t)\,\mathrm{d}t = \widetilde{B}u(t)\,\mathrm{d}Y(t),$$
$$u(0) = u_0 \in L^2(\mathbb{R}^d) \quad (4)$$

for all $t \in [0,T]$*, where on the space* $L^2(\mathbb{R}^d)$ *the linear operators* $\widetilde{L}(Y)$ *and* $\widetilde{B}$ *are defined as follows for* $x = (x_1, \ldots, x_d)$:

$$\widetilde{L}(Y)\phi(x) = -\sum_{i,l=1}^d a_{il}(x)\frac{\partial^2}{\partial x_i \partial x_l}\phi(x)$$
$$-\sum_{i=1}^d b_i(x)\frac{\partial}{\partial x_i}\phi(x), \quad (5)$$

$$\widetilde{B}\phi(x) = \sigma^{-1}f(x)\phi(x), \quad (6)$$

with $a_{il} = (\frac{1}{2}gg^*)_{i,l=1,\ldots,d}$*, where* * *denotes the transpose matrix. Therefore, we do not need any boundary conditions.*

For the existence and uniqueness of solutions of Eqn. (4) and their regularity properties, we need the following lemma.

Lemma 1. *(Ahmed and Radaideh, 1997; Elliott and Glowinski, 1989; Pardoux, 1991) Suppose that*

(i) $b_i \in C_b$, $i = 1, \ldots, d$, $(gg^*)_{il} \in C_b^1$, $i, l = 1, \ldots, d$, *and there exists* $\alpha > 0$ *such that*

$$((gg^*)\xi, \xi) \geq \alpha \|\xi\|^2 \quad (7)$$

for all $\xi \in \mathbb{R}^d$.

Then $(\widetilde{L}(Y))^* \in L(V, V^*)$*, and there exist some constant* $\rho \geq 0$ *such that the coercivity condition*

$$\alpha \|\phi\|_V^2 + \langle(\widetilde{L}(Y))^*\phi, \phi\rangle \leq \rho\|\phi\|_H^2 \quad (8)$$

is satisfied for all $\phi \in V$ *and uniformly in* Y*. Moreover, there exist some positive constants* β *and* γ *such that*

$$\langle\widetilde{L}(Y)\phi, \phi\rangle + \frac{1}{2}\|\widetilde{B}\phi\|_H^2 + \beta\|\phi\|_H^2 \geq \frac{1}{2}\gamma\|\phi\|_V^2 \quad (9)$$

for all $\phi \in V$ *and uniformly in* Y.

Proposition 2. *see (Ahmed and Radaideh, 1997; Pardoux, 1991) Suppose that the assumptions of Lemma* 1 *hold and*

(ii) *there exists a positive constant* K *such that*

$$\|\widetilde{B}u\|_{L^2(\mathbb{R}^d, \mathbb{R}^p)}^2 \leq K\|u\|_H^2. \quad (10)$$

Then there exists a unique solution u *to the Zakai equation* (4) *and* $u \in L^\infty(0, T; H) \cap L^2(0, T; H^1)$*,* $u \in C(0, T; H)$*, P-almost surely.*

3. Variational formulation

The spatial domain for the solution of the Zakai equation (4) is the whole space $\mathbb{R}^d$. Computationally, it is impossible to consider such an unbounded domain. Since the initial density u_0 is approximately supported on a bounded subset of $\mathbb{R}^d$, the solution to the Zakai equation can also be considered possibly on another bounded (with the Lipschitz regular boundary) sufficiently large subset, e.g., $\mathcal{U} \subset \mathbb{R}^d$, where a significant part of the probability mass is concentrated. Therefore, we shall restrict the computations to the subset $\mathcal{U}$ (Ahmed and Radaideh, 1997; Elliott and Glowinski, 1989).

To create the variational form of Eqn. (4), we take the inner product of (4) by a function $v \in H_0^1(\mathcal{U})$, $\mathcal{U} \subset \mathbb{R}^d$. We also write

$$[u(t)](x) = u(t,x), \quad 0 \leq t \leq 1, \quad x \in \mathcal{U}.$$

Thus we get

$$\left(\frac{\partial}{\partial t}u(t), v\right) + \widetilde{\widetilde{B}}[u, v, t] = \left(\widetilde{B}u(t)\frac{\mathrm{d}Y(t)}{\mathrm{d}t}, v\right) \quad (11)$$

or

$$\left(\frac{\partial}{\partial t}u(t), v\right) + \widetilde{\widetilde{B}}[u, v, t] = \left(\sigma^{-1}f(X(t))u(t)\frac{\mathrm{d}Y(t)}{\mathrm{d}t}, v\right)$$

with the appropriate initial condition, where

$$\widetilde{\widetilde{B}}[u, v, t] = (\widetilde{L}u(x), v(x)). \quad (12)$$

We also put $\widetilde{l} : V \times V \to \mathbb{R}$ for

$$\widetilde{l}(u, v) = \langle\widetilde{L}u, v\rangle_{V^* \times V}, \quad u, v \in V. \quad (13)$$

Moreover, we shall need on V an equivalent energy norm defined by

$$\|u\|_{\widetilde{l}} = (\widetilde{l}(u, u))^{1/2} \sim \|u\|_V. \quad (14)$$

On $\partial\mathcal{U}$ we impose the zero boundary condition. Integrating by parts and taking into account the functions vanish-

ing on $\partial \mathcal{U}$, using (12) we get

$$\widetilde{\widetilde{B}}\,[u, v, t] = \int_{\mathcal{U}} \widetilde{L}(Y)u(x)v(x)\,\mathrm{d}x$$

$$= \int_{\mathcal{U}} \Big[-\sum_{i,l=1}^{d} a_{il} \frac{\partial^2}{\partial x_i \partial x_l} u(x)$$

$$- \sum_{i=1}^{d} b_i \frac{\partial}{\partial x_i} u(x) \Big] v(x)\,\mathrm{d}x \qquad (15)$$

$$= \int_{\mathcal{U}} \Big[\sum_{i,l=1}^{d} a_{il} \frac{\partial}{\partial x_i} u(x) \frac{\partial}{\partial x_l} v(x)$$

$$- \sum_{i=1}^{d} b_i \frac{\partial}{\partial x_i} u(x) v(x) \Big]\,\mathrm{d}x.$$

More precisely, from (11) we obtain

$$\int_{\mathcal{U}} \frac{\partial u(t)}{\partial t}(x)v(x)\,\mathrm{d}x + \widetilde{\widetilde{B}}\,[u, v, t]$$

$$= \int_{\mathcal{U}} \sigma^{-1} f(x)u(x) \frac{\mathrm{d}Y(t)}{\mathrm{d}t} v(x)\,\mathrm{d}x.$$

Now, let $\{e_s(x)\}$ denote an orthogonal basis of $H_0^1(\mathcal{U})$. For fixed n we define

$$u_n(t, x) = \sum_{s=1}^{n} \alpha_s^n(t)e_s(x), \qquad (16)$$

so $(u_n', e_s) = (\partial u_n(t)/\partial t, e_s) = (\alpha_s^n(t))'$. Further, we take $v(x) = e_s(x)$ in (11) and get

$$(u_n', e_s) - \widetilde{\widetilde{B}}\,[u_n, e_s, t] - \Big(\widetilde{B}u_n(t) \frac{\mathrm{d}Y(t)}{\mathrm{d}t}, e_s \Big) = 0. \quad (17)$$

Notice that

$$\widetilde{\widetilde{B}}\,[u_n, e_s, t] = \widetilde{\widetilde{B}}\,\Big[\sum_{h=1}^{n} \alpha_h^n(t)e_h(x), e_s, t \Big]$$

$$= \sum_{h=1}^{n} \alpha_h^n(t) \widetilde{\widetilde{B}}\,[e_h, e_s, t]. \quad (18)$$

So we obtain a system of ordinary differential equations

$$(\alpha_s^n(t))' - \sum_{h=1}^{n} \alpha_s^n(t) \widetilde{\widetilde{B}}\,[e_h, e_s, t]$$

$$- \Big(\widetilde{B}u_n(t) \frac{\mathrm{d}Y(t)}{\mathrm{d}t}, e_s \Big) = 0, \quad s = 1, \ldots, n \quad (19)$$

for unknown functions $\alpha_s^n(t)$, and with the initial condition

$$\sum_{s=1}^{n} \alpha_s^n(0)e_s(x) = u_n(0, x). \qquad (20)$$

It is obvious that there exists a unique absolutely continuous solution $\alpha^n(t) = (\alpha_1^n(t), \ldots, \alpha_n^n(t))$ to the system (19), (20). However, the performance of the iterative scheme depends on the condition number of the system (Bramble *et al.*, 2003; Dahmen, 1997) which is the quotient of the maximal and minimal eigenvalues of the system matrix. In practice, one would not apply an iterative scheme to the solution of the above system—it is too complicated.

In the next section we propose a set of basis wavelet functions to get the solution of the Zakai equation by the wavelet-Galerkin method.

4. Wavelet basis

The definitions are taken from the books (Bramble *et al.*, 2003; Cohen, 2003; Daubechies, 1992; Wang, 2002) and from the paper (Dahmen, 1997). Our goal is to create a biorthogonal wavelet basis.

In the paper, we further assume that $d = 1$ and we take as the scaling function the so-called "tent" function (Dahmen, 1997):

$$\varphi(x) = \begin{cases} 1 + x, & -1 \leq x \leq 0, \\ 1 - x, & 0 \leq x \leq 1, \\ 0, & \text{otherwise.} \end{cases} \qquad (21)$$

The simplest choice of the so-called trial spaces V_j are the spaces of the scaled tent functions

$$\varphi_{j,k}(x) = 2^{j/2}\varphi(2^j x - k), \quad k = 0, 1, \ldots, 2^j. \quad (22)$$

We choose $\varphi_{j,k}(x)$ as basis functions for the trial spaces V_j. They form a Riesz basis relative to the standard inner product in H.

Further (Dahmen, 1997, pp. 62–63), since

$$\varphi(x) = \frac{1}{2}\varphi(2x + 1) + \varphi(2x) + \frac{1}{2}\varphi(2x - 1), \quad (23)$$

that is,

$$\varphi_{j,k} = \frac{1}{2\sqrt{2}}\varphi_{j+1,2k-1} + \frac{1}{\sqrt{2}}\varphi_{j+1,2k} + \frac{1}{2\sqrt{2}}\varphi_{j+1,2k+1}, \quad (24)$$

V_j are nested and their union is dense in $L^2([0, 1])$. Let (Cohen, 2003, pp. 11–12)

$$\psi(x) = \varphi(2x - 1), \quad \psi_{j,k}(x) = 2^{j/2}\psi(2^j x - k),$$

$$j = 1, 2, \ldots, \quad k = 0, 1, \ldots, 2^j - 1. \quad (25)$$

In order to successively update from coarser grids, we consider the hierarchical decomposition of V_j. Let

$$P_j f = \sum_{k=0}^{2^j} 2^{-j/2} f(2^{-j}k)\varphi_{j,k}. \qquad (26)$$

Then $P_j f$ is the projection of f onto V_j and the complements

$$W_j = (P_{j+1} - P_j)V_{j+1} \tag{27}$$

are spanned by the tent functions on new grid points on the next higher scale

$$\Psi_j = \{\psi_{j,k} = \varphi_{j+1,2k+1} : \quad k = 0, 1, \dots, 2^j - 1\}. \tag{28}$$

The resulting multilevel basis has become known as the wavelet basis.

Further, let

$$c_{jk} = (f, \varphi_{j,k}), \quad d_{jk} = (f, \psi_{j,k}) \tag{29}$$

for any function f from a normed space be the scaling function coefficients and the wavelet coefficients, respectively.

The wavelet basis $\psi_{j,k}$ gives the advantage that the elements of V_j, for each j, have two equivalent representations, namely, in the scaling as well as wavelet coefficients (29).

Now let us denote the dilation by j operator and the translation by k operator by D^j and T^k, respectively. Thus

$$(D^j f)(x) = 2^{j/2} f(2^j x), \quad (T^k f)(x) = f(x - k), \tag{30}$$

for any function f and $j \in \mathbb{Z}$, $k \in \mathbb{R}$. They are the unitary operators so we can write the adjoint operators corresponding to them on the other side of the inner product.

It is obvious that $\psi_{j,k} = D^j T^k \psi$. Moreover, we have (Wojtaszczyk, 1997, §2.1; Dai and Larson, 1998, §3)

$$\begin{aligned}(\psi_{j,k}, \psi_{j',k'}) &= (D^j T^k \psi, D^{j'} T^{k'} \psi) \\ &= ((D^{j'-j} T - 2^{j'-j} k + k')\psi, \psi),\end{aligned} \tag{31}$$

so

$$\begin{aligned}&(\psi_{j,k}(x), \psi_{j',k'}(x)) \\ &= (2^{(j'-j)/2} \psi(2^{j'-j} x - 2^{j'-j} k + k'), \psi(x))\end{aligned} \tag{32}$$

for $j, j' = 1, 2, \dots, j \le j'$ and $k = 0, 1, \dots, 2^j - 1$, $k' = 0, 1, \dots, 2^{j'} - 1$.

From (31) and (32) we have, for example,

$$\begin{aligned}&(\psi_{1,1}(x), \psi_{2,2}(x)) \\ &= (DT\psi(x), D^2 T^2 \psi(x)) = (2^{1/2} \psi(2x), \psi(x)).\end{aligned}$$

Remark 1. If we construct the wavelets using the tent function (21) and we build the matrix, denoted here by $\mathcal{M}$, we get what follows:

(i) For the same level of D, that is, for $j = j'$, the three wavelets overlap (we can see this from the property (23)), so $\mathcal{M}$ is three-diagonal.

(ii) For $j' = j - 1$, five wavelets overlap so $\mathcal{M}$ is five-diagonal.

(iii) For general j and j', we have $2^{j-j'} + 1$ diagonals.

Moreover, see (Bramble *et al.*, 2003, p. 38), we have

$$\frac{d}{dx}\psi_{j,k}(x) = \frac{d}{dx}\varphi_{j+1,2k+1}(x) = 2^{j+\frac{3}{2}}\psi_{j,k}^H(x), \tag{33}$$

denoting by $\psi_{j,k}^H$ the Haar wavelets for the sake of distinction here.

Let us notice that the Haar wavelets have $\varphi(x) = \chi_{[0,1]}$ (the indicator function) as the scaling function (see (Bramble *et al.*, 2003, pp. 32–33)) and they satisfy the condition (Bramble *et al.*, 2003, p. 59)

$$(\psi_{j,k}^H, \psi_{j',k'}^H) = \delta_{j,j'}\delta_{k,k'}. \tag{34}$$

Therefore, from (34) we get the following (Bramble *et al.*, 2003, p. 63)

$$\begin{aligned}&\left(\frac{d}{dx}\psi_{j,k}, \frac{d}{dx}\psi_{j',k'}\right) \\ &= 2^{j+j'+3}(\psi_{j,k}^H, \psi_{j',k'}^H) = 2^{j+j'+3}\delta_{j,j'}\delta_{k,k'}. \tag{35}\end{aligned}$$

It is known that the Haar functions form a complete orthonormal basis in $L^2[0,1]$.

Further (see (Dahmen, 1997, p. 80)), every $\varphi_{j,k}$ has the following expansion for $\Delta_j = \{0, \dots, 2^j\}$:

$$\varphi_{j,k} = \sum_{l \in \Delta_{j+1}} m_{l,k}^j \varphi_{j+1,l},$$

where $m_k^j = \{m_{l,k}^j\}_{l \in \Delta_{j+1}} \in l_2(\Delta_{j+1})$ is the mask or filter sequence. We can write it in the matrix notation as

$$\Phi_j^T = \Phi_{j+1}^T M_{j,0}$$

and call it the refinement relation, where $M_{j,0}$ is the refinement matrix. For the tent function it is the $(2^{j+1} - 1) \times (2^j - 1)$ matrix $M_{j,0}$ below. Now we look for collections (Dahmen, 1997, p. 82)

$$\Psi_j = \{\psi_{j,k} : \quad k = 0, 1, \dots, 2^j - 1\} \subset W_j$$

for every $j = 1, 2, \dots$ which spans W_j, and Ψ being a union of such collections. This implies that there exists a matrix $M_{j,1}$ such that

$$\Psi_j^T = \Phi_{j+1}^T M_{j,1}.$$

For the tent function it is the $(2^{j+1} - 1) \times 2^j$ matrix $M_{j,1}$

below. The announced matrices are

$$M_{j,0} = \begin{pmatrix} \frac{1}{2\sqrt{2}} & 0 & 0 & 0 & \cdots \\ \frac{1}{\sqrt{2}} & 0 & 0 & 0 & \cdots \\ \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & 0 & 0 & \cdots \\ 0 & \frac{1}{\sqrt{2}} & 0 & \cdots \\ 0 & \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & 0 & \cdots \\ 0 & 0 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ & & & & \frac{1}{2\sqrt{2}} & 0 \\ & & & & \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} \\ 0 & \cdots & & & 0 & \frac{1}{\sqrt{2}} \\ 0 & \cdots & & & 0 & \frac{1}{2\sqrt{2}} \end{pmatrix},$$

$$M_{j,1} = \begin{pmatrix} 1\,0\,0 & & \\ 0\,0\,0 & & \\ 0\,1\,0 & & \\ \vdots\,\vdots\,\vdots & & \vdots \\ & & 0\,1\,0 \\ & & 0\,0\,0 \\ & & 0\,0\,1 \end{pmatrix}$$

Now we shall recall the Lévy-Ciesielski representation of the Wiener process (Ciesielski, 1961; McKean, 1969) and apply it to the description of the Wiener process $V(t)$ in (3).

Let us define the Haar functions for $k \geq 1$, odd $j < 2^k$ and $t \in [0,1]$:

$$h_0(t) = 1, \quad h_{jk}(t) = \begin{cases} \sqrt{2^{j-1}}, & t \in \left[\frac{k-1}{2^j}, \frac{k}{2^j}\right), \\ -\sqrt{2^{j-1}}, & t \in \left[\frac{k}{2^j}, \frac{k+1}{2^j}\right), \\ 0, & \text{elsewhere.} \end{cases}$$
(36)

Definition 1. The Schauder functions $\{H_0, H_{jk}\}$ are integrals of the Haar functions h_{jk}, that is,

$$H_{jk}(t) = \int_0^t h_{jk}(s)\, \mathrm{d}s.$$

Thus they are of the form

$$H_0(t) = t,$$
$$H_{jk}(t) = \begin{cases} \sqrt{2^{j-1}}\left[t - \frac{k-1}{2^j}\right], & t \in \left[\frac{k-1}{2^j}, \frac{k}{2^j}\right), \\ \sqrt{2^{j-1}}\left[\frac{k+1}{2^j} - t\right], & t \in \left[\frac{k}{2^j}, \frac{k+1}{2^j}\right), \\ 0, & \text{elsewhere.} \end{cases}$$
(37)

The Schauder functions are little hats or tents of height $\sqrt{2^{j+1}}$.

Theorem 1. (Ciesielski, 1961; McKean, 1969). *Let us define $\widetilde{W}^n(t)$ as*

$$\widetilde{W}^n(t) = X_0(\omega)H_0(t) + \sum_{j=1}^{n} \sum_{k=1}^{2^j-1} X_{jk}H_{jk}(t), \quad (38)$$

where $\{X_0, X_{jk}\}$, $j = 1, 2, \ldots, n$, $k = 1, 2, \ldots, 2^j - 1$ is a sequence of independent normal random variables with mean zero and variance one, defined on the probability space $(\Omega, \mathcal{F}, P)$, $t \in [0,1]$, $n = 1, 2, \ldots$. Then $\widetilde{W}^n(t)$ converges almost surely to some $\widetilde{W}(t)$ which is a standard Brownian motion, uniformly in t in the space $C([0,1])$.

5. Wavelet-Galerkin space discretization of the Zakai equation

We shall now approximate the exact solution to Eqn. (4) with respect to the space variable by the wavelet-Galerkin method using the biorthogonal wavelet basis.

In the general case, let $\mathcal{H}$ be a Hilbert space and $A(\cdot, \cdot) : \mathcal{H} \times \mathcal{H} \to \mathbb{R}$ a continuous bilinear form. Consider the following variational problem: Given $F \in \mathcal{H}^*$, find $U \in \mathcal{H}$ such that

$$A(V, U) = \langle V, F \rangle, \quad V \in \mathcal{H}. \tag{39}$$

Define the operator $\mathcal{L} : \mathcal{H} \to \mathcal{H}^*$ by

$$\langle V, \mathcal{L}U \rangle = A(V, U), \quad V \in \mathcal{H}, \tag{40}$$

so that Eqns. (39) and (40) are equivalent to

$$\mathcal{L}U = F. \tag{41}$$

Then the problem (39) is well-posed if the so-called mapping property is satisfied, that is, if there exist some positive constants $c_1, C_1 > 0$ such that

$$c_1\|V\|_{\mathcal{H}} \leq \|\mathcal{L}V\|_{\mathcal{H}^*} \leq C_1\|V\|_{\mathcal{H}} \quad \text{for } V \in \mathcal{H}. \tag{42}$$

In general, $\mathcal{H}$ will be a product space, i.e., $\mathcal{H} = H_{1,0} \times \cdots \times H_{m,0}$, where the component spaces $H_{i,0} \subseteq H_i$ will be closed subspaces of some Hilbert spaces H_i (e.g., some Sobolev spaces). Let the bilinear form $A(\cdot, \cdot)$ be given by $A(V, U) = (a_{i'l'}(v_{i'}, u_{l'}))_{i',l'=1,\ldots,m}$, so that the operator $\mathcal{L}$ is matrix-valued and $\mathcal{L} = (\mathcal{L}_{i'l'})_{i',l'=1,\ldots,m}$.

Define $\mathbf{D} = \mathrm{diag}(w_{j,k})$, $w_{j,k} = 2^j$ for the scale j and the spatial location k. Denote by $\mathbf{D}^{s'}$ a diagonal matrix with $w_{j,k} = 2^{s'j}$. Then $\mathbf{D}^{-s'}\Psi = \{2^{-s'j}\psi_{j,k}\}$, with Ψ given in Section 4. Now we shall describe how to solve the problem with the wavelet basis. Namely, we have the norm equivalence (Bramble *et al.*, 2003, p. 41) that we shall write generally for $\widetilde{\gamma}, \widetilde{\widetilde{\gamma}} > 0$ and $s \in (-\widetilde{\widetilde{\gamma}}, \widetilde{\gamma})$:

$$c_{s'}\|\mathbf{v}\|_{l_2} \leq \|\mathbf{v}\mathbf{D}^{-s'}\Psi\|_{H^{s'}} \leq C_{s'}\|\mathbf{v}\|_{l_2} \quad \text{for } \mathbf{v} \in l_2, \tag{43}$$

where $c_{s'}, C_{s'}$ are some positive bounded constants while $H^{s'}$ stands for a suitable subspace of a Sobolev space and $H^{-s'}$ for its dual space.

We will assume that for each $H_{i',0}$ one has suitable bases $\Psi_{i'}$ and scaling matrices $\mathbf{D}_{i'}$ such that (43) is satisfied. The scaled wavelet representation of the operators

$\mathcal{L}_{i'l'}$ is then given by

$$\mathbf{A}_{i'l'} = \mathbf{D}_{i'}^{-1} a_{i'l'}(\Psi_{i'}, \Psi_{l'}) \mathbf{D}_{l'}^{-1} \text{ for } i', l' = 1, \ldots, m, \quad (44)$$

$\mathbf{D} = \text{diag}(\mathbf{D}_1, \ldots, \mathbf{D}_m)$. Next, the scaled standard representation of $\mathcal{L}$ and the dual wavelet representation of the right-hand side data are given by

$$\mathbf{L} = (\mathbf{A}_{i'l'})_{i',l'=1}^m = \mathbf{D}^{-1}(\mathbf{\Psi}, \mathcal{L}\mathbf{\Psi}) \mathbf{D}^{-1}, \quad (45)$$

$$\mathbf{F} = \mathbf{D}^{-1}(\mathbf{\Psi}, F). \quad (46)$$

Important properties of the solution to our problem are given in the following two theorems.

Theorem 2. (Bramble *et al.*, 2003, pp. 66–67). *Let $\widehat{U} = \mathbf{U}^T \mathbf{D}^{-1} \mathbf{\Psi}$ be the scaled wavelet representation of the solution to (39). Then*

$$\mathcal{L}\widehat{U} = F \iff \mathbf{L}\mathbf{U} = \mathbf{F}. \quad (47)$$

The above result can be related to preconditioning. Namely, we have the next theorem.

Theorem 3. (Bramble *et al.*, 2003, pp. 67). *The mapping property (42) and the norm equivalence (43) imply a uniformly bounded condition number with respect to the mesh size during discretization:*

$$\text{cond}(\mathbf{L}) = \mathcal{O}(1) \quad \text{as } \sharp\Lambda \to \infty, \quad (48)$$

where Λ is a finite subset of $\mathcal{J} = \mathcal{J}_\varphi \cup \mathcal{J}_\psi$. Here $\mathcal{J}$ is an infinite index set, where $\sharp\mathcal{J}_\varphi < \infty$, representing the scaling functions. The index in $\mathcal{J}_\psi$ represents the wavelet spanning complements between refinement levels.

Any function $u(t, \cdot) \in U$ can be represented in the wavelet basis $\{\psi_{j,k}(x)\}$, $j = 0, 1, \ldots$, $k = 0, 1, \ldots, 2^j - 1$, as follows:

$$u(t, x) = \sum_{j=0}^{\infty} \sum_{k=0}^{2^j-1} \alpha_{jk}(t) \psi_{j,k}(x),$$

$$\alpha_{jk}(t) \in \mathbb{R}, \quad \alpha_{jk}(0) = \gamma_{jk}. \quad (49)$$

We define a projection $P_n : V \to V_n$ by truncating the wavelet expansion

$$u_n(t, x) = P_n u(t, x) = \sum_{j=0}^{n} \sum_{k=0}^{2^j-1} \alpha_{jk}(t) \psi_{j,k}(x) \quad (50)$$

for $u_n(t, \cdot) \in V_n$, where V_n is the finite-dimensional subspace of V that is spanned by the finite collection $\{\psi_{j,k}\}$, $j = 0, 1, \ldots, n$, and $k = 0, 1, \ldots, 2^j - 1$.

In our case the norm equivalence is of the form (von Petersdorff and Schwab, 2003, p. 168)

$$c_{s'} \|\mathbf{v}\|_{\mathcal{H}^{s'}(\mathcal{U})}^2 \leq \sum_{j=0}^{\infty} \sum_{k=0}^{2^j-1} |\alpha_{jk}|^2 2^{2js'} \leq C_{s'} \|\mathbf{v}\|_{\mathcal{H}^{s'}(\mathcal{U})}^2 \quad (51)$$

for $\mathbf{v} \in l_2$, $0 \leq s' \leq 1$.

Now we consider Eqn. (11) putting $v(x) = \psi_{\tilde{l}, \tilde{s}}(x)$, $\tilde{l} = 0, 1, \ldots, \tilde{s} = 0, 1, \ldots, 2^{\tilde{l}} - 1$, and we use (38) for the Wiener process $V(t)$. We obtain

$$\left(\frac{\partial}{\partial t} u(t), \psi_{\tilde{l}, \tilde{s}}\right) + \widetilde{\widetilde{B}}[u, \psi_{\tilde{l}, \tilde{s}}, t]$$
$$- \left(\sigma^{-1} f(x) u(t) \frac{dY(t)}{dt}, \psi_{\tilde{l}, \tilde{s}}\right) = 0. \quad (52)$$

We substitute $u_n(t)$ for $u(t)$ and, using the fact that

$$\left(\frac{\partial}{\partial t} \sum_{j=0}^{n} \sum_{k=0}^{2^j-1} \alpha_{jk}(t) \psi_{j,k}(x), \psi_{\tilde{l}, \tilde{s}}(x)\right)$$
$$= \sum_{j=0}^{n} \sum_{k=0}^{2^j-1} \alpha'_{jk}(t)(\psi_{j,k}(x), \psi_{\tilde{l}, \tilde{s}}(x)), \quad (53)$$

from (17) we get

$$\sum_{j=0}^{n} \sum_{k=0}^{2^j-1} \alpha'_{jk}(t)(\psi_{j,k}(x), \psi_{\tilde{l}, \tilde{s}}(x))$$
$$+ \widetilde{\widetilde{B}}\left[\sum_{j=0}^{n} \sum_{k=0}^{2^j-1} \alpha_{jk}(t) \psi_{j,k}(x), \psi_{\tilde{l}, \tilde{s}}(x), t\right]$$
$$- \sum_{j=0}^{n} \sum_{k=0}^{2^j-1} \alpha_{jk}(t) \left(\sigma^{-1} f(x) \psi_{j,k}(x) \frac{dY(t)}{dt}, \psi_{\tilde{l}, \tilde{s}}(x)\right)$$
$$= 0. \quad (54)$$

Further, we shall need the derivative of $\psi_{j,k}(x)$:

$$\frac{\partial}{\partial x_i} \psi_{j,k}(x) = 2^{3j/2} \psi'_{x_i}(2^j x - k), \quad i = 1, \ldots, d. \quad (55)$$

From (12) and (50) and integration by parts we get

$$\widetilde{\widetilde{B}}[u_n, \psi_{\tilde{l}, \tilde{s}}, t]$$
$$= \widetilde{\widetilde{B}}\left[\sum_{j=0}^{n} \sum_{k=0}^{2^j-1} \alpha_{jk}(t) \psi_{j,k}(x), \psi_{\tilde{l}, \tilde{s}}(x), t\right]$$
$$= \sum_{j=0}^{n} \sum_{k=0}^{2^j-1} \alpha_{jk}(t) \int_{\mathcal{U}} \widetilde{L}(Y) \psi_{j,k}(x) \psi_{\tilde{l}, \tilde{s}}(x) \, dx$$
$$= \sum_{j=0}^{n} \sum_{k=0}^{2^j-1} \alpha_{jk}(t) \int_{\mathcal{U}} \left[\sum_{i,l=1}^{d} a_{il}(x) \frac{\partial}{\partial x_i} \psi_{j,k}(x) \frac{\partial}{\partial x_l} \psi_{\tilde{l}, \tilde{s}}(x)\right.$$
$$\left. + \sum_{i=1}^{d} b_i(x) \frac{\partial}{\partial x_i} \psi_{j,k}(x) \psi_{\tilde{l}, \tilde{s}}(x)\right] dx. \quad (56)$$

Due to the properties (33) and (35) we obtain

$$
\sum_{j=0}^{n} \sum_{k=0}^{2^j-1} \alpha'_{jk}(t)(\psi_{j,k}(x), \psi_{\tilde{l},\tilde{s}}(x))
$$

$$
+ \sum_{j=1}^{n} \sum_{k=0}^{2^j-1} \alpha_{jk}(t)
$$

$$
\times \int_{\mathcal{U}} \Bigg[\sum_{i,l=1}^{d} a_{il}(x) 2^{3j/2} \psi'_{x_i}(2^j x - k) 2^{3\tilde{l}/2} \psi'_{x_l}(2^{\tilde{l}} x - \tilde{s})
$$

$$
+ \sum_{i=1}^{d} b_i(x) 2^{3j/2} \psi'_{x_i}(2^j x - k) \Bigg] \mathrm{d}x
$$

$$
- \sum_{j=0}^{n} \sum_{k=0}^{2^j-1} \alpha_{jk}(t) \Big(\sigma^{-1} f(x) 2^{j/2} \psi(2^j x - k) \frac{\mathrm{d}Y(t)}{\mathrm{d}t},
$$

$$
2^{\tilde{l}/2} \psi(2^{\tilde{l}} x - \tilde{s}) \Big) = 0
\tag{57}
$$

with the initial condition

$$
\sum_{j=1}^{n} \sum_{k=0}^{2^j-1} \alpha_{jk}(0)(\psi_{j,k}(x), \psi_{\tilde{l},\tilde{s}}(x)) = (u_0, \psi_{\tilde{l},\tilde{s}}(x)) \tag{58}
$$

for all $\tilde{l} = 0, 1, \ldots, n$, $\tilde{s} = 0, 1, \ldots, 2^j - 1$. The last term has, by (3), the following form:

$$
\sum_{j=0}^{n} \sum_{k=0}^{2^j-1} \alpha_{jk}(t) \Big(\sigma^{-1} f(x) 2^{j/2} \psi(2^j x - k)
$$

$$
\times \frac{\mathrm{d}Y(t)}{\mathrm{d}t}, 2^{\tilde{l}/2} \psi(2^{\tilde{l}} x - \tilde{s}) \Big)
$$

$$
= \sum_{j=0}^{n} \sum_{k=0}^{2^j-1} \alpha_{jk}(t) \Big(\sigma^{-1} f(x) 2^{j/2} \psi(2^j x - k)
$$

$$
\times \Big[\sigma^{-1} f(x) + \frac{\mathrm{d}V(t)}{\mathrm{d}t} \Big], 2^{\tilde{l}/2} \psi(2^{\tilde{l}} x - \tilde{s}) \Big)
$$

$$
= \sum_{j=0}^{n} \sum_{k=0}^{2^j-1} \alpha_{jk}(t) \big((\sigma^{-1} f(x))^2 2^{j/2} \psi(2^j x - k),
$$

$$
2^{\tilde{l}/2} \psi(2^{\tilde{l}} x - \tilde{s}) \big)
$$

$$
+ \sum_{j=0}^{n} \sum_{k=0}^{2^j-1} \alpha_{jk}(t) \Big(\sigma^{-1} f(x) 2^{j/2} \psi(2^j x - k)
$$

$$
\times \frac{\mathrm{d}V(t)}{\mathrm{d}t}, 2^{\tilde{l}/2} \psi(2^{\tilde{l}} x - \tilde{s}) \Big).
\tag{59}
$$

Let $S_{j,k} = \operatorname{supp} \psi_{j,k}$. The diameter of $S_{j,k}$ is bounded by $C 2^{-k}$ and

$$
\delta_{k,\tilde{s}} = c \max\{2^{-n+C(2n-k-\tilde{s})}, 2^{-k}, 2^{-\tilde{s}}\}
$$

with some parameters $c, C > 0$. For the meaning of the compressed parameters c and C we refer to (von Peters-dorff and Schwab, 2003).

We define the compressed matrix $\widetilde{\widetilde{L}}(Y)$ and the corresponding bilinear form by replacing certain small matrix elements in $\widetilde{L}(Y)$ with zero:

$$
\widetilde{\widetilde{L}}_{(jk)(\tilde{l}\tilde{s})}(Y) = \begin{cases} \widetilde{L}_{(jk)(\tilde{l}\tilde{s})}(Y) \\ \quad \text{if } \operatorname{dist}(S_{j,k}; S_{\tilde{l},\tilde{s}}) \le \delta_{k,\tilde{s}} \\ \quad\quad \text{or } S_{j,k} \cap \partial\mathcal{U} \ne \emptyset, \\ 0 \quad\quad \text{otherwise.} \end{cases}
\tag{60}
$$

Finally, we get, in the matrix form, the following system, simpler than (19)–(20), to compute the unknown vector $\alpha(t)$:

$$
\mathcal{M}\alpha'(t) + \alpha(t)\mathbf{D}_i^{-1} \widetilde{\widetilde{L}}(Y)\mathbf{D}_l^{-1}
$$

$$
- \alpha(t)\mathbf{D}_i^{-1}\widetilde{\mathcal{B}}\mathbf{D}_l^{-1} = 0, \quad (61)
$$

for $t > 0$, where $\mathcal{M}$ is the mass matrix described in Remark 1, $\widetilde{\widetilde{L}}$ is the so-called stiffness matrix, $\widetilde{\mathcal{B}}$ is the matrix given in the last term of (57), and $\alpha(t)$ is the vector of unknowns. We notice that the stiffness and mass matrices are symmetric and positive-definite.

During time discretization of the increments of the Wiener process $V(t)$, appearing in $\mathrm{d}Y(t)$ in the last term of (57), we get the following expression using (3), (37) and (38) for $t_\kappa = \kappa\lambda$ and $\lambda = T/m$, $\kappa = 0, 1, \ldots, m$, $T > 0$, $n = 1, 2, \ldots$:

$$
V^n(t_\kappa) - V^n(t_{\kappa-1})
$$

$$
= X_0(\omega)H_0(t) + \sum_{j=1}^{n} \sum_{k=1}^{2^j-1} X_{jk}H_{jk}(t_\kappa)
$$

$$
- X_0(\omega)H_0(t) - \sum_{j=1}^{n} \sum_{k=1}^{2^j-1} X_{jk}H_{jk}(t_{\kappa-1})
$$

$$
= \sum_{j=1}^{n} \sum_{k=1}^{2^j-1} X_{jk}[H_{jk}(t_\kappa) - H_{jk}(t_{\kappa-1})]
$$

$$
= \sum_{j=1}^{n} \sum_{k=1}^{2^j-1} X_{jk}G_{jk},
\tag{62}
$$

where

$$
G_{jk} = \sqrt{2^{j-1}}\big[t_\kappa - 2^{-j}(k-1) - t_{\kappa-1} + 2^{-j}(k-1)\big]
$$

$$
= \sqrt{2^{j-1}}[t_\kappa - t_{\kappa-1}] = \sqrt{2^{j-1}}\lambda
$$

if $t_\kappa, t_{\kappa-1} \in [2^{-j}(k-1), 2^{-j}k)$,

$$
G_{jk} = \sqrt{2^{j-1}}\big[2^{-j}(k+1) - t_\kappa - 2^{-j}(k+1) + t_{\kappa-1}\big]
$$

$$
= -\sqrt{2^{j-1}}[t_\kappa - t_{\kappa-1}] = -\sqrt{2^{j-1}}\lambda
$$

if $t_\kappa, t_{\kappa-1} \in [2^{-j}k, 2^{-j}(k+1))$,

$$G_{jk} = \sqrt{2^{j-1}}\left[2^{-j}(k+1) - t_\kappa - 2^{-j}(k-1) + t_{\kappa-1}\right]$$
$$= \sqrt{2^{j-1}}\left[2^{1-j} - t_\kappa + t_{\kappa-1}\right], = \sqrt{2^{j-1}}\left[2^{1-j} - \lambda\right]$$

if $t_{\kappa-1} \in [2^{-j}(k-1), 2^{-j}k)$, $t_\kappa \in [2^{-j}k, 2^{-j}(k+1))$.

Observe that $\lambda = t_\kappa - t_{\kappa-1}$. Notice that in the above we can assume that λ is so small that $t_{\kappa-1}$ and t_κ are situated in the same dyadic subinterval $[2^{-j}(k-1), 2^{-j}k)$ or $[2^{-j}k, 2^{-j}(k+1))$ in (37) and (62). Thus we only have the first and the second situation, that is,

$$|V^n(t_\kappa) - V(t_{\kappa-1})| = 2^{(j-1)/2}(t_\kappa - t_{\kappa-1}) \tag{63}$$
$$= 2^{(j-1)/2}\lambda = \widetilde{c}(n)\lambda$$

and

$$E[V^n(t_\kappa) - V^n(t_{\kappa-1})]^2 = c(n)\lambda^2, \tag{64}$$

for $c(n) = (\widetilde{c}(n))^2$.

6. Convergence theorem

The outline of this section is as follows: First, we shall consider the convergence of the schemes defined in (65)–(69) for the implicit Euler approximation with respect to time and for the wavelet-Galerkin approximation with respect to space. Second, we shall also compute a constant of the rate of convergence.

Namely, in Proposition 3 we give the convergence result. Then we formulate Theorem 4 on the rate of the convergence of our two schemes. To prove Theorem 4 we shall need two auxiliary Lemmas 2–3, for the rate of the convergence for the Euler as well as the wavelet-Galerkin approximations, respectively. Third, we use the wavelet basis introduced in Section 4 and we apply the well known generalized minimal residual method with m_0 iterates to solve the linear systems that appear in our wavelet schemes. We denote this method by GMRES(m_0). To estimate this scheme (see Lemma 4), we shall proceed similarly to the technique of (von Petersdorff and Schwab, 1996, Lemma 6.1).

Finally, we denote by $\widehat{u}_n^{\kappa,m}$ the inexact equation solution by GMRES iterations. We conclude that at each time step one can show, similarly as in (von Petersdorff and Schwab, 2003, pp. 178–179), that the solution $\widehat{u}_n^{\kappa,m}$ satisfies the same error bound as the solution $\widetilde{\widetilde{u}}_n^{\kappa,m}$ with the wavelet compressed operator and with all linear systems solved exactly.

At the beginning, let us mention that for the usual Hermite polynomial basis we have a similar time discretization in (Itô, 1996). We shall apply this result to the Wiener process representation given in (38).

We introduce the probability space $S = (\Omega \times [0, T], \mathcal{F} \times \mathcal{B}, \mathrm{d}P \times \mathrm{d}t)$. Throughout this section we set

$$u^\kappa = u(t_\kappa) \in V, \quad t_\kappa = \kappa\lambda \quad \text{for } \lambda = \frac{T}{m},$$

where $\kappa = 0, 1, \ldots, m$.

We consider the wavelet-Galerkin approximation scheme for Eqn. (4):

$$\mathrm{d}u_n(t) + \widetilde{L}_n(Y(t))u_n(t)\,\mathrm{d}t = \widetilde{B}_n(t)u_n(t)\,\mathrm{d}Y(t),$$
$$u_n(0) = u_{n,0}, \tag{65}$$

where $\widetilde{L}_n(Y(t))$ and $\widetilde{B}_n(t)$ are the projections of the operators given by (5) and (6), respectively. Moreover, $u_{n,0}$ is the orthogonal projection of u_0 onto V_n of H.

For $m \in N$ take the sequence $\{u^{\kappa,m}\}$ generated by the following Euler scheme for Eqn. (4):

$$\frac{u^{\kappa,m} - u^{\kappa-1,m}}{\lambda} + [\widetilde{L}(Y_{\kappa-1}) + \beta I]u^{\kappa,m}$$
$$= \beta u^{\kappa-1,m} + \widetilde{B}(t_{\kappa-1})u^{\kappa-1,m}\frac{Y^{\kappa,m} - Y^{\kappa-1,m}}{\lambda}, \tag{66}$$
$$u^{\kappa,m}(0) = u_0 \in L^2(\mathcal{U})$$

or

$$u^{\kappa,m} - u^{\kappa-1,m} + \lambda[\widetilde{L}(Y_{\kappa-1}) + \beta I]u^{\kappa,m}$$
$$= \lambda\beta u^{\kappa-1,m} + \widetilde{B}(t_{\kappa-1})u^{\kappa-1,m}(Y^{\kappa,m} - Y^{\kappa-1,m}),$$
$$u^{\kappa,m}(0) = u_0 \in L^2(\mathcal{U}), \tag{67}$$

where $Y_\kappa = Y(\kappa\lambda) = Y(t_\kappa)$. This method is defined by replacing the time derivative by a backward difference quotient.

Now, for $n, m \in \mathbb{N}$ take the sequence $\{u_n^{\kappa,m}\}$ approximating the solution to (4) by the wavelet-Galerkin as well as the Euler schemes:

$$u_n^{\kappa,m} - u_n^{\kappa-1,m} + \lambda[\widetilde{L}_n(Y_{\kappa-1}) + \beta I]u_n^{\kappa,m}$$
$$= \lambda\beta u^{\kappa-1,m} \tag{68}$$
$$+ \widetilde{B}_n(Y_{\kappa-1})u_n^{\kappa-1,m}(Y^{\kappa,m} - Y^{\kappa-1,m}).$$

Let

$$u_n^m(t) = u_n^{\kappa,m} \text{ on } [\kappa\lambda, (\kappa+1)\lambda). \tag{69}$$

Let us moreover write

$$u_n^{\kappa,m}(t) - u^\kappa(t) = (u_n^{\kappa,m}(t) - u^{\kappa,m}(t))$$
$$+ (u^{\kappa,m}(t) - u^\kappa(t)) = I + II, \tag{70}$$

where $u^{\kappa,m}(t)$ is the solution to Eqn. (67).

We have the following convergence result.

Proposition 3. *Suppose that the operators $\widetilde{L}_n(Y)$ and $\widetilde{B}_n$ satisfy the following approximation conditions*:

(i) $\|\widetilde{L}_n(Y)\phi\|_{V^*} \leq \eta\|\phi\|_V$ *and* $\|\widetilde{B}_n\phi\|_H \leq \eta\|\phi\|_V$, $\eta > 0$, *uniformly in n and $Y \in \mathbb{R}^p$*

(ii) $\langle\widetilde{L}_n(Y)\phi, \phi\rangle + \frac{1}{2}\|\widetilde{B}_n\phi\|_H^2 + \beta\|\phi\|_H^2 \geq \frac{1}{2}\gamma\|\phi\|_V^2$ *for all $\phi \in V, Y \in \mathbb{R}^d$*

(iii) $\|\widetilde{L}_n(Y)\phi - \widetilde{L}(Y)\phi\|_{V^*} + \|\widetilde{B}_n\phi - \widetilde{B}\phi\|_H \to 0$ *for all* $\phi \in V, Y \in \mathbb{R}^d$, *as* $n \to \infty$,

and $\|u_{n,0} - u_0\|_H \to 0$ *as* $n \to \infty$.

Then the sequence $\{u_n^m\}$ *of solutions to* (68)*, given in* (70) *and satisfying* (i)–(iii), *converges strongly in* $L^2(S, V)$ *and weak-star in* $L^\infty(0, T; L^2(\Omega, H, \mathrm{d}\widetilde{P}))$ *to the unique solution* $u(t)$ *to Eqn.* (4) *as* $m, n \to \infty$.

Proof. We notice that the operators $\widetilde{L}_n(Y)$ and $\widetilde{B}_n$ satisfy the above assumptions (i)–(iii). Indeed, it is obvious that condition (i) is satisfied. The condition (ii) follows from (9) since the left-hand side of (ii) is the restriction of (9) onto V_n.

The condition (iii) is satisfied because $\|P_V^n \phi - \phi\|_V \to 0$ as $n \to \infty$, where P_V^n is the orthogonal projection of V onto V_n.

Thus Theorem 2.1 of (Itô, 1996) can be applied to the Galerkin approximations. ■

Now we shall assume the consistency condition for $\widetilde{l}(\cdot, \cdot)$ given in (13), that is,

(iv) There exists a constant $\eta < 1$ independent of n such that

$$|\widetilde{l}(u_n, v_n) - \widetilde{\widetilde{l}}(u_n, v_n)| \leq \eta \|u_n\|_\alpha \|v_n\|_\alpha$$
$$\text{for every } u_n, v_n \in V_n. \quad (71)$$

The condition (71) shows that on $V_n \times V_n$ the form $\widetilde{l}(\cdot, \cdot)$ is equivalent to $\widetilde{\widetilde{l}}(\cdot, \cdot)$ in the following sense: for $\eta < 1$ in (71) we have for some constants $0 < \widetilde{\beta} < \widetilde{\alpha} < \infty$ independent of n:

$$\|\widetilde{\widetilde{l}}(u_n, v_n)\| \leq \eta \|u_n\|_\alpha \|v_n\|_\alpha$$
$$\text{for every } u_n, v_n \in V_n. \quad (72)$$

Moreover, we assume that

(v) We have

$$\int_0^T E\|\widetilde{L}(Y)u(t)\|_V^2 \,\mathrm{d}t \leq M, \; E\|\widetilde{B}u(t)\|_V^2 \leq M$$

for any $M > 0$, independently of $t \in [0, T]$.

Remark 2. Notice (Kloeden and Platen, 1992) that the order scheme for stochastic differential equations is sometimes less than for the corresponding deterministic differential equations. This is because the increments ΔW_n of the Wiener process are of root mean square order $\frac{1}{2}$ and not 1. One can prove (Kloeden and Platen, 1992) that for the classical definition of the Wiener process we have

$$E[W(t+\lambda) - W(t)]^{2n} = 1 \cdot 3 \cdots (2n-1)\lambda^n. \quad (73)$$

For example, the Euler scheme for stochastic differential equations has strong order equal to $1/2$, in contrast to the order 1 of the Euler scheme for deterministic ordinary differential equations. We stress that in the case of the representation (38) of the Wiener process, we get the property (64), different from (73).

The rate of the convergence of our two schemes is computed in the following theorem.

Theorem 4. *Assume that the operators* $\widetilde{L}_n(Y)$ *and* $\widetilde{B}_n$ *satisfy the assumptions* (i)–(iii) *of Proposition 3 together with the assumption* (v). *Then the rate of convergence of the numerical scheme* (68) *is of order* $\lambda^{1/2}$ *in the sense of the inequality*

$$E\|u_n^{\kappa,m} - u^\kappa\|_H^2 + \frac{\gamma}{2}E\sum_{\iota=1}^\kappa \lambda c^{\kappa-\iota}\|u_n^{\kappa,m} - u^\kappa\|_V^2$$

$$+ \frac{\gamma}{4}\int_0^t e^{2\beta(t-s)}E\|u_n^{\kappa,m}(s) - u^\kappa(s)\|_V^2 \,\mathrm{d}s$$

$$\leq \widetilde{C}E\|u_0\|_H^2 + C(n) + \widetilde{M}\lambda\frac{e^{2T\beta}-1}{2\beta}, \quad (74)$$

where $u_n^{\kappa,m}(t)$ *and* $u^\kappa(t) = u(t_\kappa)$ *are the solutions to Eqns.* (68) *and* (4)*, respectively. Moreover,* $\widetilde{M} = \widetilde{M}(\kappa) > 0$*,* $c = (1 + \lambda\beta)^2$*, and* $C(n)$ *is given by* (96) *in Lemma 3 below.*

First we shall prove two lemmas. We start from the estimation of II in (70).

Lemma 2. *Assume that the operators* $\widetilde{L}_n(Y)$ *and* $\widetilde{B}_n$ *satisfy the assumptions* (i)–(iii) *of Proposition 3 and, moreover, the assumption* (v). *Then*

$$E\|u^{\kappa,m} - u^\kappa\|_H^2 + \frac{\gamma}{2}E\sum_{\iota=1}^\kappa \lambda c^{\kappa-\iota}\|u_n^{\kappa,m} - u^\kappa\|_V^2$$

$$\leq \widetilde{M}\lambda\frac{e^{2T\beta}-1}{2\beta}, \quad (75)$$

where $u^{\kappa,m}(t)$ *and* $u^\kappa(t) = u(t_\kappa)$ *are the solutions to Eqns.* (67) *and* (4)*, respectively,* $\widetilde{M} = \widetilde{M}(n)$ *is given in the formula* (94) *below, and* $c = (1 + \lambda\beta)^2$.

Proof. To get an estimate of II in (70), we proceed similarly as in (Itô, 1996) in the following manner. For $\kappa \geq 0$ we define the approximation error

$$\varepsilon_\kappa = \varepsilon_\kappa^{(1)} - \varepsilon_\kappa^{(2)}, \quad (76)$$

with

$$\varepsilon_k^{(1)} = \int_{t_{\kappa-1}}^{t_\kappa} \widetilde{L}(u(t) - u(t_\kappa)) \,\mathrm{d}t$$
$$- \lambda\beta(u(t_\kappa) - u(t_{\kappa-1})), \quad (77)$$

$$\varepsilon_\kappa^{(2)} = \int_{t_{\kappa-1}}^{t_\kappa} \widetilde{B}(u(t) - u(t_\kappa)) \,\mathrm{d}Y(t), \quad (78)$$

where $u(t)$ is the unique solution to (4) and $t_\kappa = \kappa\lambda$, so $\lambda = t_\kappa - t_{\kappa-1}$. Further, we write for simplicity $\widetilde{L} = \widetilde{L}(Y(t_{\kappa-1}))$.

Now we write (4) equivalently as

$$u(t_\kappa) - u(t_{\kappa-1}) + \int_{t_{\kappa-1}}^{t_\kappa} \widetilde{L}(Y(t))u(t)\,\mathrm{d}t$$
$$- \int_{t_{\kappa-1}}^{t_\kappa} \widetilde{B}(t)u(t)\,\mathrm{d}Y(t) = 0. \quad (79)$$

Define the error function $\delta(u_\kappa) = u^{\kappa,m} - u(t_\kappa)$. It satisfies the equation

$$\delta(u_\kappa) - \delta(u_{\kappa-1}) + \lambda[\widetilde{L} + \beta I]\delta(u_\kappa)$$
$$= \lambda\beta\delta(u_{\kappa-1}) + \widetilde{B}\delta(u_{\kappa-1})\Delta Y_\kappa - \varepsilon_\kappa, \quad (80)$$

where $\Delta Y_\kappa = Y(t_\kappa) - Y(t_{\kappa-1})$ is an $\mathbb{R}^p$-valued random variable with mean zero and covariance λ. Indeed, from (77)–(79) and (67) we get

$$\delta(u_\kappa) - \delta(u_{\kappa-1}) + \lambda[\widetilde{L} + \beta I]\delta(u_\kappa)$$
$$= \lambda\beta u_{\kappa-1} + \widetilde{B}u_{\kappa-1}\Delta Y_\kappa + \int_{t_{\kappa-1}}^{t_\kappa} \widetilde{L}(Y(t))u(t)\,\mathrm{d}t$$
$$- \int_{t_{\kappa-1}}^{t_\kappa} \widetilde{B}(t)u(t)\,\mathrm{d}Y(t) - \lambda\widetilde{L}u(t_\kappa) - \lambda\beta u(t_\kappa)$$

and also

$$\lambda\beta\delta(u_{\kappa-1}) + \widetilde{B}\delta(u_{\kappa-1})\Delta Y_\kappa - \varepsilon_\kappa = \lambda\beta u_{\kappa-1}$$
$$- \lambda\beta u(t_{\kappa-1}) + \widetilde{B}u_{\kappa-1}\Delta Y_\kappa - \widetilde{B}u(t_{\kappa-1})\Delta Y_\kappa$$
$$+ \int_{t_{\kappa-1}}^{t_\kappa} [\widetilde{L}(Y(t))u(t) - \widetilde{L}(Y(t))u(t_\kappa)]\,\mathrm{d}t$$
$$- \lambda\beta u(t_\kappa) + \lambda\beta u(t_{\kappa-1})$$
$$- \int_{t_{\kappa-1}}^{t_\kappa} [\widetilde{B}(t)u(t) - \widetilde{B}(t)u(t_{\kappa-1})]\,\mathrm{d}Y(t)$$
$$= \lambda\beta u_{\kappa-1} + \widetilde{B}u_{\kappa-1}\Delta Y_\kappa + \int_{t_{\kappa-1}}^{t_\kappa} \widetilde{L}(Y(t))u(t)\,\mathrm{d}t$$
$$- \int_{t_{\kappa-1}}^{t_\kappa} \widetilde{B}(t)u(t)\,\mathrm{d}Y(t) - \lambda\widetilde{L}u(t_\kappa) - \lambda\beta u(t_\kappa).$$

Multiplying (80) by $\delta(u_\kappa)$ we obtain

$$(\delta(u_\kappa) - \delta(u_{\kappa-1}), \delta(u_\kappa)) + (\lambda[\widetilde{L} + \beta I]\delta(u_k), \delta(u_k))$$
$$- (\lambda\beta\delta(u_{\kappa-1}), \delta(u_\kappa)) - (\widetilde{B}\delta(u_{\kappa-1})\Delta Y_\kappa, \delta(u_\kappa))$$
$$- (\varepsilon_\kappa, \delta(u_\kappa)) = 0. \quad (81)$$

Define a linear operator $\widetilde{L}_\beta = \widetilde{L} + \beta I$. Then by (8) we get $J_\lambda = [I + \lambda\widetilde{L}_\beta]^{-1} \in L(V^*, V)$.

For a V-valued random variable x, the equation

$$\widehat{x} - x + \lambda\widetilde{L}_\beta\widehat{x} = \lambda\beta x + \widetilde{B}x\Delta Y \quad (82)$$

has a unique solution $\widehat{x} = J_\lambda([1 + \lambda\beta]x + \widetilde{B}x\Delta Y_\kappa))$. We multiply (82) by $\widehat{x}$ and get

$$\left(\widehat{x} - (1 + \lambda\beta)x, \widehat{x}\right) + \lambda\langle\widetilde{L}_\beta\widehat{x}, \widehat{x}\rangle = \left(\widetilde{B}x\Delta Y_\kappa, \widehat{x}\right). \quad (83)$$

From this, using the definition of the norm $\|\cdot\|_H$, we obtain

$$\frac{1}{2}\left(\|\widehat{x}\|_H^2 - (1 + \lambda\beta)^2\|x\|_H^2 + \|\widehat{x} - (1 + \lambda\beta)x\|_H^2\right)$$
$$+ \lambda\langle\widetilde{L}_\beta\widehat{x}, \widehat{x}\rangle = (\widetilde{B}x\Delta Y_\kappa, \widehat{x}). \quad (84)$$

Now we set $z = (1 + \lambda\beta)x + \widetilde{B}x\Delta Y_\kappa$. Then from (82) we have $\widehat{x} = J_\lambda z$ and

$$\|J_\lambda z - z\|_H^2 = \|\widehat{x} - (1 + \lambda\beta)x - \widetilde{B}x\Delta Y_\kappa\|_H^2$$
$$= \|\widehat{x} - (1 + \lambda\beta)x\|_H^2$$
$$- 2\left(\widetilde{B}x\Delta Y_\kappa, \widehat{x} - (1 + \lambda\beta)x\right)$$
$$+ \|\widetilde{B}x\Delta Y_\kappa\|_H^2. \quad (85)$$

But $E(\widetilde{B}x\Delta Y_\kappa, x) = 0$ because x and ΔY_κ are independent random variables. Thus from (84) and (85) we get

$$\frac{1}{2}\left(E\|\widehat{x}\|_H^2 - (1 + \lambda\beta)^2 E\|x\|_H^2\right)$$
$$+ \lambda E\left(\langle\widetilde{L}_\beta\widehat{x}, \widehat{x}\rangle - \frac{1}{2}\|\widetilde{B}x\|_H^2\right) \leq 0. \quad (86)$$

Coming back to (81), we have, writing $\widehat{x} = \delta(u_\kappa)$, $x = \delta(u_{\kappa-1})$ and $c = (1 + \lambda\beta)^2$,

$$E\|\delta(u_\kappa)\|_H^2 - cE\|\delta(u_{\kappa-1})\|_H^2 + 2\lambda E\langle\widetilde{L}_\beta\delta(u_\kappa), \delta(u_\kappa)\rangle$$
$$\leq E\|\widetilde{B}\delta(u_{\kappa-1})\Delta Y_\kappa - \varepsilon_\kappa^{(2)}\|_H^2, \quad (87)$$

because $E(\delta(u_{k-1}), \varepsilon_k^{(2)}) = 0$.

From the assumption (ii) applied to the operators $\widetilde{L}$ and $\widetilde{B}$ it follows that there exists an $\varepsilon > 0$ such that

$$2\langle\widetilde{L}_\beta\phi, \phi\rangle - \|\widetilde{B}\phi\|_H^2 + \beta\|\phi\|_H^2 \geq \frac{3}{4}\|\phi\|_V^2 + \frac{1}{4}\beta\|\phi\|_V^2,$$

so

$$2\langle\widetilde{L}_\beta\phi, \phi\rangle - (1 + \varepsilon)\|\widetilde{B}\phi\|_H^2 \geq \frac{3}{4}\beta\|\phi\|_V^2 \quad (88)$$

for all $\phi \in V$. Moreover, from the Young inequality $ab \leq \frac{\varepsilon}{2}a^2 + \frac{1}{2\varepsilon}b^2$ we have

$$E\|\widetilde{B}\delta(u_{\kappa-1})\Delta Y_\kappa - \varepsilon_\kappa^{(2)}\|^2$$
$$= E[(\widetilde{B}\delta(u_{\kappa-1})\Delta Y_\kappa - \varepsilon_\kappa^{(2)}, \widetilde{B}\delta(u_{\kappa-1})\Delta Y_\kappa - \varepsilon_\kappa^{(2)})]$$
$$\leq (1 + \varepsilon)E\|\widetilde{B}\delta(u_{\kappa-1})\|_H^2 + \left(1 + \frac{1}{\varepsilon}\right)E\|\varepsilon_\kappa^{(2)}\|_H^2.$$

But, once more from the Young inequality, we see that

$$2(\delta(u_\kappa), \varepsilon_\kappa^{(1)}) \leq \frac{\lambda\gamma}{4} \|\delta(u_\kappa)\|_V^2 + \frac{4}{\lambda\gamma} \|\varepsilon_\kappa^{(1)}\|_V^2,$$

and from this we obtain

$$E(\|\delta(u_\kappa)\|_H^2 + \lambda\|\widetilde{B}(\delta u_\kappa)\|_H^2) + \frac{\lambda\gamma}{2} E\|\delta(u_\kappa)\|_V^2$$
$$\leq cE(\|\delta(u_{\kappa-1})\|_H^2 + \lambda\|\widetilde{B}\delta(u_{\kappa-1})\|_H^2)$$
$$+ \frac{4\lambda}{\gamma}\left\|\frac{\varepsilon_\kappa^{(1)}}{\lambda}\right\|_V^2 + \left(1 + \frac{1}{\varepsilon}\right)E\|\varepsilon_\kappa^{(2)}\|_H^2. \quad (89)$$

Now we multiply (89) by c^{-k} and take the sum over k. We obtain

$$c^{-k}E(\|\delta(u_\kappa)\|_H^2) + \frac{\gamma}{2}E\sum_{\iota=1}^{\kappa}\lambda c^{-\iota}\|\delta(u_\iota)\|_V^2$$
$$\leq E\sum_{\iota=1}^{\kappa}\lambda c^{-\iota}\left[\frac{4}{\gamma}\left\|\frac{\varepsilon_\iota^{(1)}}{\lambda}\right\|_{V^*}^2 + \frac{1+\frac{1}{\varepsilon}}{\lambda}E\|\varepsilon_\iota^{(2)}\|_H^2\right] \quad (90)$$

for $1 \leq \kappa \leq m$. Note that we can write (4) as

$$u(t) - u(s) + \int_s^t \widetilde{L}(Y(\tau))u(\tau)\,d\tau$$
$$- \int_s^t \widetilde{B}(\tau)u(\tau)\,dY(\tau) = 0, \quad u(0) = u_0 \quad (91)$$

for $s \leq t$. Moreover,

$$u(t_\kappa) - u(s) + \int_s^{t_\kappa} \widetilde{L}(Y(\tau))u(\tau)\,d\tau$$
$$- \int_s^{t_\kappa} \widetilde{B}(\tau)u(\tau)\,dY(\tau) = 0. \quad (92)$$

Subtracting the above equalities and taking the expectation we deduce from the assumption (v) that

$$E\|u(t_\kappa) - u(t)\|_V^2$$
$$\leq 2|t - t_\kappa|\int_t^{t_\kappa} E\|\widetilde{L}(Y(\tau))u(\tau)\|_V^2\,d\tau$$
$$+ 2\int_t^{t_\kappa} E\|\widetilde{B}u(\tau)\|_V^2\,d\tau \quad (93)$$
$$\leq 2M(|t - t_\kappa| + 1).$$

From (77) and (78) it follows that there exists a constant $\widetilde{M}$ such that we can estimate

$$\frac{4}{\gamma}E\left\|\frac{\varepsilon_\kappa^{(1)}}{\lambda}\right\|_{V^*}^2 + \frac{1+\frac{1}{\varepsilon}}{\lambda}E\|\varepsilon_\kappa^{(2)}\|_H^2$$
$$\leq \frac{4}{\gamma}M''' + \left(1 + \frac{1}{\varepsilon}\right)(M'\lambda + M''c(n)) \quad (94)$$
$$= \widetilde{M}(n)\lambda = \widetilde{M}\lambda$$

for $1 \leq \kappa \leq m$. Indeed, from the assumption (i) and (64) we have

$$\frac{4}{\gamma}E\left\|\frac{\varepsilon_\kappa^{(1)}}{\lambda}\right\|_{V^*}^2 \leq \frac{4}{\gamma}\frac{1}{\lambda^2}(\lambda^2 M' + \lambda^2 M'') = \frac{4}{\gamma}M''',$$

$$\frac{1+\frac{1}{\varepsilon}}{\lambda}E\|\varepsilon_\kappa^{(2)}\|_H^2$$
$$= \frac{1+\frac{1}{\varepsilon}}{\lambda}E\left\|\int_{t_{\kappa-1}}^{t_{\kappa-1}+\lambda}\widetilde{B}(u(t) - u(t_\kappa))\,dY(t)\right\|_H^2$$
$$= \frac{1+\frac{1}{\varepsilon}}{\lambda}E\left\|\int_{t_{\kappa-1}}^{t_{\kappa-1}+\lambda}\widetilde{B}(u(t) - u(t_\kappa))\sigma^{-1}f(x)\,dt\right\|_H^2$$
$$+ \frac{1+\frac{1}{\varepsilon}}{\lambda}E\left\|\int_{t_{\kappa-1}}^{t_{\kappa-1}+\lambda}\widetilde{B}(u(t) - u(t_\kappa))\,dV(t)\right\|_H^2$$
$$\leq \frac{1+\frac{1}{\varepsilon}}{\lambda}M'\lambda^2 + \frac{1+\frac{1}{\varepsilon}}{\lambda}M''c(n)\lambda \leq M'''(n)\lambda.$$

Observe that $c = (1 + \lambda\beta)^2 > 1$. We also use the fact that $1 + u < e^u$ for $u > 0$, that is, $(1 + u)^{2\kappa} < e^{2\kappa u}$. Now from the formula for the sum of κ terms in the geometric series with quotient $1/c < 1$ we get

$$\sum_{\iota=1}^{\kappa}\lambda c^{\kappa-\iota} = \frac{[(1+\lambda\beta)^2]^\kappa - 1}{2\beta}$$
$$< \frac{e^{2\kappa\frac{T}{m}\beta} - 1}{2\beta}$$
$$\leq \frac{e^{2T\beta} - 1}{2\beta}$$

for $1 \leq \kappa \leq m$. From (90) we have

$$E\|\delta(u_\kappa)\|^2 + \frac{\gamma}{2}E\sum_{\iota=1}^{\kappa}\lambda c^{\kappa-\iota}\|\delta(u_\iota)\|_V^2$$
$$\leq E\sum_{\iota=1}^{\kappa}\lambda c^{\kappa-\iota}\left[\frac{4}{\gamma}\left\|\frac{\varepsilon_\iota^{(1)}}{\lambda}\right\|_V^2 + \frac{1+\frac{1}{\varepsilon}}{\lambda}\|\varepsilon_\iota^{(2)}\|_H^2\right]$$
$$\leq \widehat{M}\lambda\sum_{\iota=1}^{\kappa}\lambda c^{\kappa-\iota} \leq \widehat{M}\lambda\frac{e^{2T\beta} - 1}{2\beta},$$

which completes the proof. ∎

To estimate the term I, we notice that, instead of $I = u_n^{\kappa,m} - u_n^\kappa$, it is sufficient to estimate the Galerkin scheme only, that is, $I' = u_n(t) - u(t)$. Therefore, we have what follows.

Lemma 3. *Assume that the operators $\widetilde{L}_n(Y)$ and $\widetilde{B}_n$ satisfy the assumptions* (i)–(iii) *of Proposition* 3 *together with the assumption* (v). *Then*

$$E\|u_n(t) - u(t)\|_H^2$$
$$+ \frac{\gamma}{4}\int_0^t e^{2\beta(t-s)}E\|u_n(s) - u(s)\|_V^2\,ds$$
$$\leq \widetilde{C}E\|u_0\|_H^2 + C(n), \quad (95)$$

where $u_n(t)$ and $u(t)$ are the solutions to Eqns. (65) and (4), respectively, and $C(n)$ is given by the formula (96) below.

Proof. Let

$$u_n(t) = u_{n,0} - \int_0^t \widetilde{L}_n(Y(s))u_n(s)\,\mathrm{d}\tau$$
$$+ \int_0^t \widetilde{B}(s)u_n(s)\,\mathrm{d}Y(s), \quad u_n(0) = u_{n,0}.$$

From Itô's lemma (Krylov and Rozovskiĭ, 1981, Theorem 3.23; Liptser and Shiryaev, 1977) applied to the function $\varphi(u_n(t)) = e^{2\beta t}E\|u_n(t) - u(t)\|_H^2$ it follows that

$$E\|u_n(t) - u(t)\|_H^2$$
$$+ \int_0^t e^{2\beta(t-s)}E\Big(2\langle\widetilde{L}_{n,\beta}(Y(s))(u_n(s) - u(s))$$
$$+ \widetilde{L}_{n,\beta}(Y(s))u(s) - \widetilde{L}_\beta(Y(s))u(s), u_n(s) - u(s)\rangle$$
$$- \|\widetilde{B}_n(s)(u_n(s) - u(s))$$
$$+ \widetilde{B}_n(s)u(s) - \widetilde{B}(s)u(s)\|_H^2\Big)\,\mathrm{d}s$$
$$= e^{2\beta t}E\|u_{n,0} - u_0\|_H^2.$$

From the assumption (i) for the operator $\widetilde{L}_{n,\beta}(Y)$ and from the Young inequality $xy \leq \frac{\varepsilon}{2}x^2 + \frac{1}{2\varepsilon}y^2$, $\varepsilon > 0$, for $\varepsilon = \gamma/2$ we obtain

$$\langle\widetilde{L}_{n,\beta}(Y(s))u_n(s) - \widetilde{L}(Y(s))u(s), u_n(s) - u(s)\rangle$$
$$\leq \frac{\gamma}{4}\|u_n(s) - u(s)\|_V^2$$
$$+ \frac{1}{\gamma}\|\widetilde{L}_{n,\beta}(Y(s))u_n(s) - \widetilde{L}_\beta(Y(s))u(s)\|_V^2.$$

Similarly, from the assumption (i) for the operator $\widetilde{B}_n$ and from the same inequality of Young for $\varepsilon = \gamma/2$ we obtain

$$\big(\widetilde{B}_n(s)(u_n(s) - u(s)), \widetilde{B}_n(s)u(s) - \widetilde{B}(s)u(s)\big)$$
$$\leq \frac{\gamma}{4}\|u_n(s) - u(s)\|_V^2$$
$$+ \frac{1}{\gamma}\big(\eta^2\|\widetilde{B}_n(s)u_n(s) - \widetilde{B}(s)u(s)\|_H^2\big).$$

From the assumptions (ii) and (iii) we get

$$E\|u_n(t) - u(t)\|_H^2$$
$$+ \frac{\gamma}{4}\int_0^t e^{2\beta(t-s)}E\|u_n(s) - u(s)\|_V^2\,\mathrm{d}s$$
$$\leq e^{2\beta t}E\|u_{n,0} - u_0\|_H^2 + \int_0^t e^{2\beta(t-s)}$$
$$\times E\Big(\frac{2}{\gamma}\|\widetilde{L}_{n,\beta}(Y(s))u(s) - \widetilde{L}_\beta(Y(s))u(s)\|_{V^*}^2$$
$$+ \Big(\frac{\eta^2}{\gamma} + 1\Big)\|\widetilde{B}_n(s)u(s) - \widetilde{B}(s)u(s)\|_H^2\Big)\,\mathrm{d}s.$$

Therefore, denoting the last integral in the above inequality by $C(n)$, that is,

$$C(n) = \int_0^t e^{2\beta(t-s)}$$
$$\times E\Big(\frac{2}{\gamma}\|\widetilde{L}_{n,\beta}(Y(s))u(s) - \widetilde{L}_\beta(Y(s))u(s)\|_{V^*}^2$$
$$+ \Big(\frac{\eta^2}{\gamma} + 1\Big)\|\widetilde{B}_n(s)u(s) - \widetilde{B}(s)u(s)\|_H^2\Big)\,\mathrm{d}s, \quad (96)$$

we arrive at

$$E\|u_n(t) - u(t)\|_H^2$$
$$+ \frac{\gamma}{4}\int_0^t e^{2\beta(t-s)}E\|u_n(s) - u(s)\|_V^2\,\mathrm{d}s$$
$$\leq \widetilde{C}\|u_0\|_H^2 + C(n),$$

where $\widetilde{C} = e^{2\beta T}$, which completes the proof. ∎

Proof of Theorem 4. Together with inequalities (75) and (95) this completes the proof of the theorem. ∎

Now, we apply the wavelet basis. First we compute the mass matrix $\mathcal{M}$ and the stiffness matrix $\widetilde{\widetilde{L}}$, both in our wavelet basis. With the compressed matrix $\widetilde{\widetilde{L}}$, we associate the solution $\widetilde{\widetilde{u}}_n^{\kappa,m}$. We consider the difference

$$\widetilde{\widetilde{w}}_n^{\kappa,m} = \widetilde{\widetilde{u}}_n^{\kappa+1,m} - \widetilde{\widetilde{u}}_n^{\kappa,m} \in V_n$$

satisfying the equation (derived from (68))

$$\lambda^{-1}(\widetilde{\widetilde{w}}_n^{\kappa,m}, v_n) + ([\widetilde{\widetilde{L}}_n(Y_{\kappa-1}) + \beta I]\widetilde{\widetilde{u}}_n^{\kappa,m}, v_n)$$
$$= \beta(\widetilde{\widetilde{u}}_n^{\kappa,m}\widetilde{u}^{\kappa-1,m}, v_n)$$
$$+ \lambda^{-1}\big(\widetilde{B}_n\widetilde{\widetilde{u}}_n^{\kappa-1,m}(Y^{\kappa,m} - Y^{\kappa-1,m}), v_n\big)$$

for every $v_n \in V_n$. From this, adding and subtracting the term $\widetilde{\widetilde{L}}_n(Y_{\kappa-1})\widetilde{\widetilde{u}}_n^{\kappa-1,m}$, we get

$$\lambda^{-1}(\widetilde{\widetilde{w}}_n^{\kappa,m}, v_n) + \widetilde{\widetilde{B}}_n[\widetilde{\widetilde{w}}_n^{\kappa,m}, v_n, t] + \beta(\widetilde{\widetilde{w}}_n^{\kappa,m}, v_n)$$
$$= -\widetilde{\widetilde{B}}_n[\widetilde{\widetilde{u}}_n^{\kappa-1,m}, v_n, t]$$
$$+ \lambda^{-1}\big(\widetilde{B}_n\widetilde{\widetilde{u}}_n^{\kappa-1,m}(Y^{\kappa,m} - Y^{\kappa-1,m}), v_n\big). \quad (97)$$

Denote by

$$\mathcal{K} = \lambda^{-1}\mathcal{M} + [\widetilde{\widetilde{B}} + \beta I] \quad (98)$$

the left-hand side of (97). Then update $\widetilde{\widetilde{u}}_n^{\kappa+1,m} = \widetilde{\widetilde{w}}_n^{\kappa,m} + \widetilde{\widetilde{u}}_n^{\kappa,m}$.

Let $\widetilde{\widetilde{w}}^{\,\kappa,m}$ denote the coefficient vectors of $\widetilde{\widetilde{w}}^{\,\kappa,m}_n$ with respect to the wavelet basis. Therefore, we get a linear system

$$\mathcal{K}\widetilde{\widetilde{w}}^{\,\kappa,m}_n = \mathcal{B}^{\kappa,m}$$

with the known right-hand side of (98) denoted by $\mathcal{B}^{\kappa,m}$.

Further, we shall use the well-known generalized minimal residual method with m_0 iterates to solve the linear systems. We denote this method by GMRES(m_0). We shall proceed similarly to (von Petersdorff and Schwab, 1996, Lemma 6.1).

Lemma 4. *Consider the linear system $\mathcal{K}x = \mathcal{B}$ with the matrix $\mathcal{K}$ given by (98) and with known right-hand side. Let the assumptions* (i)–(iii) *of Proposition* 3 *be satisfied. Denote by x_j, for $j \in \mathbb{N}$, the iterates obtained by the* GMRES(m_0) *method with initial guess x_0. Then*

$$E\|x - x_j\|_H^2 + E\|x - x_j\|_V^2 \le \widetilde{C}q^j\|x - x_0\|, \quad (99)$$

where $\widetilde{C}$ and $q < 1$ are independent of λ, β, L.

Proof. Let $\widetilde{c}_1, \ldots, \widetilde{c}_7$ be some positive constants independent of m and λ. Let D be the diagonal matrix, that is, $D = D_{(j,l)(j,l)} = \text{diag}\{2^{l\rho/2}\}$. Because of the norm equivalence (51) we have, for all $x, y \in \mathbb{R}^d$,

$$\widetilde{c}_1\|x\|_{l_2}^2 \le x^T\mathcal{M}x, \quad (100)$$

$$x^T\mathcal{M}y \le \widetilde{c}_2\|x\|_{l_2}^2\|y\|_{l_2}^2. \quad (101)$$

Further, using the consistency condition (71) and the assumption on continuity (72), we get

$$\widetilde{c}_3\|Dx\|_{l_2}^2 \le x^T\widetilde{\widetilde{L}}x, \quad (102)$$

$$x^T\widetilde{\widetilde{L}}y \le \widetilde{c}_4\|Dx\|_{l_2}^2\|Dy\|_{l_2}^2. \quad (103)$$

Thus the matrix $\mathcal{K}$ satisfies, because of (102) and (100), the inequality

$$\begin{aligned}
&\widetilde{c}_5 x^T(\lambda^{-1}I + \beta I + D^2)x \\
&= \widetilde{c}_5 x^T\lambda^{-1}x + \widetilde{c}_5 x^T\beta x + \widetilde{c}_5 x^T D^2 x \\
&\le \widetilde{c}_5 x^T(\lambda^{-1}\mathcal{M} + \beta I + \widetilde{\widetilde{L}})x = x^T\mathcal{K}x,
\end{aligned} \quad (104)$$

and using the Cauchy-Schwartz inequality we deduce from (101), (103) that

$$\begin{aligned}
x^T\mathcal{K}y &= x^T(\lambda^{-1}\mathcal{M} + [\widetilde{\widetilde{L}} + \beta I])y \\
&\le \lambda^{-1}\widetilde{c}_2\|x\|_{l_2}^2\|y\|_{l_2}^2 + \widetilde{c}_4\|Dx\|_{l_2}^2\|Dy\|_{l_2}^2 \\
&\quad + \beta\|x\|_{l_2}^2\|y\|_{l_2}^2 \\
&\le \widetilde{c}_6\big((\lambda^{-1} + \beta)\|x\|_{l_2}\|y\|_{l_2} + \|Dx\|_{l_2}\|Dy\|_{l_2}\big) \\
&= \widetilde{c}_6\big[\|x\|_{l_2}(\lambda^{-1}I + \beta I + D^2)\|y\|_{l_2}\big] \\
&\le \widetilde{c}_6(x^T(\lambda^{-1}I + \beta I + D^2)x)^{1/2} \\
&\quad \times (y^T(\lambda^{-1}I + \beta I + D^2)y)^{1/2},
\end{aligned} \quad (105)$$

where $\widetilde{c}_5 = \min(\widetilde{c}_1, \widetilde{c}_3)$, $\widetilde{c}_5 = \min(\widetilde{c}_2, \widetilde{c}_4)$. Moreover, scaling with the diagonal matrix $\widehat{D} = (\lambda^{-1}I + \beta I + D^2)^{1/2}$ and with $\widehat{\mathcal{K}} = (\widehat{D})^{-1}\mathcal{K}(\widehat{D})^{-1}$, $\widehat{x} = \widehat{D}x$, $\widehat{y} = \widehat{D}y$ we obtain

$$\widetilde{c}_5\|x\|_{l_2}^2 \le x^T\mathcal{K}x, \quad (106)$$

$$x^T\mathcal{K}y \le \widetilde{c}_6\|x\|_{l_2}^2\|y\|_{l_2}^2 \quad (107)$$

for all $x, y \in \mathbb{R}^d$. From this we get

$$\widetilde{c}_5 \le \lambda_{\min}\left(\frac{(\mathcal{K} + \widehat{\mathcal{K}}^T)}{2}\right) \quad (108)$$

and

$$\|\widehat{\mathcal{K}}\| \le \widetilde{c}_6.$$

Now, by (Eisenstat *et al.*, 1983, Theorem 4.3) the GMRES iterates and their residuals $r_{j+\nu} = \mathcal{B} - \widehat{\mathcal{K}}x_{j+\nu}$ satisfy for $\nu = 1, \ldots, m_0$ the inequality

$$\|r_{j+\nu}\| \le \left(1 - \frac{\widetilde{c}_5}{\widetilde{c}_6}\right)^{\nu/2}\|r_j\|. \quad (109)$$

Further,

$$\begin{aligned}
\widetilde{c}_5\|x_j - x\|_{l_2}^2 &\le (x_j - x)^T\widehat{\mathcal{K}}(x_j - x) \\
&\le \widetilde{c}_6\|x_j - x\|\|r_j\|
\end{aligned} \quad (110)$$

so

$$\|x - x_j\| \le \widetilde{C}q^j\|x - x_0\|$$

for $q < 1$, which completes the proof. ∎

Finally, we define the Euler scheme with the GMRES(m_0) approximation as follows. Set $m_0 = 1$ for the restart number and a value $\mathcal{N}$ for the number of GMRES iterations. Let $\widehat{u}_n^{\kappa,m}$ denote the inexact solution, where the linear system for each time step is solved with $\mathcal{N}$ steps in the GMRES(m_0) method, using zero as initial guess. At each time step, under the consistency condition (71), one can show, similarly as in (von Petersdorff and Schwab, 2003, pp. 178–179), that the solution $\widehat{u}_n^{\kappa,m}$ satisfies the same error bound as the solution $\widetilde{\widetilde{u}}^{\,\kappa,m}_n$ with the wavelet compressed operator and all linear systems solved exactly.

References

Ahmed, N. U. and Radaideh, S. M. (1997). A powerful numerical technique solving Zakai equation for nonlinear filtering, *Dynamics and Control* **7**(3): 293–308.

Bennaton, J. F. (1985). Discrete time Galerkin approximations to the nonlinear filtering solution, *Journal of Mathematical Analysis and Applications* **110**: 364–383.

Beuchler, S., Schneider, R. and Schwab, C. (2004). Multiresolution weighted norm equivalences and applications, *Numerische Mathematik* **98**(2): 67–97.

Bramble, J. H., Cohen, A. and Dahmen, W. (2003). *Multiscale Problems and Methods in Numerical Simulations. Lectures given at the C.I.M.E. Summer School, held in Martina Franca, Italy, September 9–15, 2001*, Lecture Notes in Mathematics, Vol. 1825, Springer, Berlin.

Ciesielski, Z. (1961). Hölder condition for realizations of Gaussian processes, *Transactions of American Mathematical Society* **99**: 403–413.

Cohen, A. (2003). *Numerical Analysis of Wavelet Methods*, North-Holland, Amsterdam.

Cohen, A., Daubechies, I. and Feauveau, J.-C. (1992). Biorthogonal bases of compactly supported wavelets, *Communications on Pure and Applied Mathematics* **45**(5): 485–560.

Crisan, D., Gaines, J. and Lyons, T. (1998). Convergence of a branching particle method to the solution of the Zakai equation, *SIAM Journal on Applied Mathematics* **58**(5): 1568–1590.

Dahmen, W. (1997). Wavelet and multiscale methods for operator equations, *Acta Numerica* **6**: 55–228.

Dahmen, W. and Schneider, R. (1999). Composite wavelet bases for operator equations, *Mathematics of Computation* **68**(228): 1533–1567.

Dai, X. and Larson, D. R. (1998). Wandering vectors for unitary systems and orthogonal wavelets, *Memoirs of the American Mathematical Society* **134**(640).

Daubechies, I. (1992). *Ten Lectures on Wavelets*, CBMS-NSF Regional Conference Series in Applied Mathematics, Vol. 61, SIAM, Philadelphia, PA.

Eisenstat, S. C., Elman, H. C. and Schultz, M. H. (1983). Variational iterative methods for nonsymmetric systems of linear equations, *SIAM Journal on Numerical Analysis* **20**: 345–357.

Elliott, R. J. and Glowinski, R. (1989). Approximations to solutions of the Zakai filtering equation, *Stochastic Analysis and Applications* **7**(2): 145–168.

Germani, A. and Picconi, M. (1984). A Galerkin approximation for the Zakai equation, *in* P. Thoft-Christensen (Ed.), *System Modelling and Optimization (Copenhagen, 1983)*, Lecture Notes in Control and Information Sciences, Vol. 59, Springer-Verlag, Berlin, pp. 415–423.

Hilbert, N., Matache, A.-M. and Schwab, C. (2004). Sparse wavelet methods for option pricing under stochastic volatility, *Technical Report 2004-07*, Seminar für angewandte Mathematik, Eidgenössische Technische Hochschule, Zürich.

Itô, K. (1996). Approximation of the Zakai equation for nonlinear filtering, *SIAM Journal on Control and Optimization* **34**(2): 620–634.

Kloeden, P. E. and Platen, E. (1992). *Numerical Solution of Stochastic Differential Equations*, Springer-Verlag, Berlin.

Krylov, N. V. and Rozovskiǐ, B. L. (1981). Stochastic evolution equations, *Journal of Soviet Mathematics* **14**: 1233–1277.

Kurtz, T. G. and Ocone, D. L. (1988). Unique characterization of conditional distributions in nonlinear filtering, *The Annals of Probability* **16**(1): 80–107.

Liptser, R. S. and Shiryaev, A. N. (1977). *Statistics of Random Processes. I. General Theory*, Springer-Verlag, New York, NY.

McKean, H. P. (1969). *Stochastic Integrals*, Academic Press, New York, NY.

Pardoux, E. (1991). Filtrage non linéaire et équations aux dérivées partielles stochastiques associées, *École d'Été de Probabilités de Saint-Flour XIX, 1989*, Lecture Notes in Mathematics, Vol. 1464, Springer-Verlag, Berlin, pp. 67–163.

Rozovskiǐ, B. L. (1991). A simple proof of uniqueness for Kushner and Zakai equations, *in* E. Mayer-Wolf, E. Merzbach and A. Shwartz (Eds), *Stochastic Analysis*, Academic Press, Boston, MA, pp. 449–458.

Thomée, V. (1997). *Galerkin Finite Element Methods for Parabolic Problems*, Springer-Verlag, Berlin.

Twardowska, K., Marnik, T. and Pasławska-Południak, M. (2003). Approximation of the Zakai equation in a nonlinear problem with delay, *International Journal of Applied Mathematics and Computer Science* **13**(2): 151–160.

von Petersdorff, T. and Schwab, C. (1996). Wavelet approximations for first kind boundary integral equations on polygons, *Numerische Mathematik* **74**(4): 479–519.

von Petersdorff, T. and Schwab, C. (2003). Wavelet discretizations of parabolic integrodifferential equations, *SIAM Journal on Numerical Analysis* **41**(1): 159–180.

Wang, J. (2002). Spline wavelets in numerical resolution of partial differential equations, *in* D. Deng, D. Huang, R.-Q. Jia, W. Lin and J. Wand (Eds), *Wavelet Analysis and Applications. Proceedings of an International Conference, Guangzhou, China, November 15–20, 1999*, AMS/IP Studies in Advanced Mathematics, Vol. 25, American Mathematical Society, Providence, RI, pp. 257–277.

Wojtaszczyk, P. (1997). *A Mathematical Introduction to Wavelets*, London Mathematical Society Student Texts, Vol. 37, Cambridge University Press, Cambridge.

Yau, S.-T. and Yau, S. S.-T. (2000). Real time solution of nonlinear filtering problem without memory I, *Mathematical Research Letters* **7**(5-6): 671–693.

Yau, S.-T. and Yau, S. S.-T. (2008). Real time solution of nonlinear filtering problem without memory II, *SIAM Journal on Control and Optimization* **47**: 163–195.

Yserentant, H. (1990). Two preconditioners based on the multilevel splitting of finite element spaces, *Numerische Mathematik* **58**(2): 163–184.

Łukasz D. Nowak is a Ph.D. student at the Warsaw University of Technology, Faculty of Mathematics and Information Science, and also at the Warsaw School of Economics. In 2004 he received the M.Sc. degree in mathematics from Warsaw University, Faculty of Mathematics, Mechanics and Computer Studies. His scientific interests include applications of filtering theory in finance and financial mathematics, particularly related to risk topics.

Monika Pasławska-Południak is a Ph.D. student at the Rzeszów University of Technology, Department of Mathematics. Her research interests include a nonlinear filtering problem with delay, stochastic differential equations and measures of noncompactness. Currently she works at the Department of Mathematics, Rzeszów University of Technology.

Krystyna Twardowska received the M.Sc. and Ph.D. degrees in mathematics from Jagiellonian University in 1970 and 1976, respectively, where she worked until 1990. She was a lecturer at the Warsaw University of Technology between 1990 and 2007. In 2000 she received a habilitation qualification in mathematics from Silesian University. Between 2003 and 2008 she was a professor at the College of Business and Social Sciences in Otwock, of which she was a rector. Since 2008 she has been a professor of mathematics at the Warsaw University of Life Sciences (SGGW). Her main research interests include differential equations, probability theory, numerical methods and stochastic differential equations with applications. She has published 50 scientific papers, four books and about 50 conferences papers. She is a member of the American Mathematical Society and a reviewer of *Mathematical Reviews*.