

Nonlinear Filtering: Separation of Parameters and Observations Using Galerkin Approximation and Wiener Chaos Decomposition

C. P. Fung* S. Lototsky†

Published as IMA Preprint Series # 1458, February 1997

Abstract

The nonlinear filtering problem is considered for the time homogeneous diffusion model. An algorithm is proposed for computing a recursive approximation of the unnormalized filtering density, and the error of the approximation is estimated. The algorithm can have high resolution in time and, unlike most existing algorithms for nonlinear filtering, can be implemented in real time for large dimensions of the state process. The on-line part of the algorithm is simplified by performing the most time consuming operations off line.

1 Introduction

In many problems of stochastic analysis it is necessary to find the best mean square estimate of moments or other similar functionals of a partially observed diffusion process. Assume that $X = X(t)$, $t \geq 0$, is the unobservable component and the observable component $Y = Y(t)$, $t \geq 0$, is given by

$$Y(t) = \int_0^t h(X(s))ds + W(t),$$

where $W = W(t)$, $t \geq 0$, is a Wiener process independent of the process X . If $f = f(x)$ is a measurable function satisfying $\mathbf{E}|f(X(t))|^2 < \infty$, $t \geq 0$, then it is known [11, 13, 20] that under certain regularity assumptions the best mean square estimate \hat{f}_t of $f(X(t))$ given the trajectory $Y(s)$, $s \leq t$, can be written as

$$\hat{f}_t = \frac{\int f(x)u(t, x)dx}{\int u(t, x)dx},$$

*Center for Applied Mathematical Sciences, University of Southern California, Los Angeles, CA 90089-1113 (chingpan@cams-00.usc.edu).

†Institute for Mathematics and its Applications, University of Minnesota, 514 Vincent Hall, 206 Church Street S.E., Minneapolis, MN 55455 (lototsky@ima.umn.edu).

This work was partially supported by the ONR Grant #N00014-95-1-0229.

where $u = u(t, x)$ is a random field called the unnormalized filtering density (UFD). The problem of estimating $f(X(t))$ is thus reduced to the problem of computing the UFD u . It is also known that $u = u(t, x)$ is the solution of a certain stochastic partial differential equation, called the Zakai equation, driven by the observation process. The exact solution of the Zakai equation can be found only in a few special cases, and as a result the central part of the general nonlinear filtering problem is the numerical solution of the equation. In some applications, e.g. target tracking, the solution must be computed in real time, which puts additional restriction on the corresponding numerical scheme.

Most of the existing numerical schemes for the Zakai equation use various generalizations of the corresponding algorithms for the deterministic partial differential equations and therefore cannot be implemented in real time when the dimension of the state process is more than three because of the large amount of computations. Examples of the corresponding algorithms can be found in Bennaton [1], Florchinger and LeGland [5], Ito [10], etc.

When the parameters of the model are known in advance, an alternative approach is to separate the deterministic and stochastic components of the Zakai equation using the Wiener chaos decomposition and to do the computations related to the deterministic component in advance. Starting with the works of Kunita [12], Ocone [19], and Lo and Ng [14], this approach was further developed by Mikulevicius and Rozovskii [17, 18] and Budhiraja and Kallianpur [2]. The first algorithm to solve the Zakai equation using this approach so that no partial differential equations are solved on line was suggested in Lototsky et al. [15].

The objective of the current work is to develop another algorithm for solving the Zakai equation in such a way that no differential equations are solved on line and the on-line computations are performed recursively in time. The goal is achieved in two steps. First, the partial differential equation satisfied by the UFD $u(t, x)$ is approximated by a finite system of ordinary differential equations (Galerkin approximation). After that, the solution of the system is approximated using the Cameron-Martin version of the Wiener chaos decomposition. The error of each approximation and the overall error are estimated in terms of the asymptotic parameters of the algorithm. The approach was first suggested in [6] for a slightly different model, but no explicit error bound was given.

2 Filtering Problem

Let $\tilde{w} = \{\tilde{w}(t)\}_{t \geq 0}$ and $w = \{w(t)\}_{t \geq 0}$ be a d_1 - and an r -dimensional Wiener processes on a complete probability space $(\Omega, \mathcal{F}, \mathbf{P})$. Consider the d -dimensional state (or signal) process $X = X(t)$ of the form

$$\begin{aligned} dX_i(t) &= b_i(X(t))dt + \sum_{j=1}^{d_1} \sigma_{ij}(X(t))d\tilde{w}_j(t), \quad 0 < t \leq T; \\ X(0) &= X_0, \end{aligned} \tag{2.1}$$

where $b = (b_i(x))_{1 \leq i \leq d}$ is a d -dimensional vector function on \mathbb{R}^d and $\sigma = (\sigma_{ij}(x))_{1 \leq i \leq d, 1 \leq j \leq d_1}$ is a $d \times d_1$ dimensional matrix function on \mathbb{R}^d . The information about the state process is available only from the r -dimensional observation process $Y = Y(t)$ given by

$$Y(t) = \int_0^t h(X(s))ds + w(t), \quad 0 \leq t \leq T, \quad (2.2)$$

where $h = (h_l(x))_{1 \leq l \leq r}$ is an r -dimensional vector function on \mathbb{R}^d . The filtering problem studied in this paper is to compute the best mean square estimate \hat{f}_t of $f(X(t))$ given the observation trajectory $Y(s)$, $s \leq t$, where $f = f(x)$ is a function on \mathbb{R}^d such that $\mathbf{E}|f(X(t))|^2 < \infty$ for all $0 \leq t \leq T$.

The following is assumed about the model (2.1), (2.2):

- (A1) *the functions b , σ , and h are infinitely differentiable and bounded with all the derivatives;*
- (A2) *the processes w and \tilde{w} are independent;*
- (A3) *the random vector X_0 is independent of both w and \tilde{w} and has a density $p = p(x)$, $x \in \mathbb{R}^d$, so that the function p is infinitely differentiable and decays at infinity with all the derivatives faster than every power of $|x|$.*

Let \mathcal{F}_t^y be the σ -algebra generated by $Y(s)$, $s \leq t$. Define

$$\rho(t) = \exp \left\{ - \sum_{l=1}^r \int_0^t h_l(X(s))dw_l(s) - \frac{1}{2} \sum_{l=1}^r \int_0^t |h_l(X(s))|^2 ds \right\}.$$

It is well known [11, 13] that the measure $\tilde{\mathbf{P}}$ given by $d\tilde{\mathbf{P}} = \rho(T)d\mathbf{P}$ is a probability measure on (Ω, \mathcal{F}) with the properties:

- (i) On the reference probability space $(\Omega, \mathcal{F}, \tilde{\mathbf{P}})$, $Y(\cdot)$ is a Brownian motion independent of $X(\cdot)$;
- (ii) The optimal filter $\hat{f}_t = \mathbf{E}[f(X(t))|\mathcal{F}_t^y]$ is given by

$$\hat{f}_t = \frac{\tilde{\mathbf{E}}[f(X(t))\rho(t)^{-1}|\mathcal{F}_t^y]}{\tilde{\mathbf{E}}[\rho(t)^{-1}|\mathcal{F}_t^y]}, \quad (2.3)$$

where $\tilde{\mathbf{E}}$ is the expectation with respect to measure $\tilde{\mathbf{P}}$. If assumptions (A1 – A3) hold, then by Theorems 6.2.1 and 6.2.2 in [20] the unnormalized filtering measure $\Phi_t(dx) = \tilde{\mathbf{E}}[1_{\{X(t) \in dx\}}\rho(t)^{-1}|\mathcal{F}_t^y]$ admits the density $u(t, x) = \Phi_t(dx)/dx$, called **unnormalized filtering density** (UFD), which is a solution of the Zakai equation

$$du(t, x) = \mathcal{L}^*u(t, x)dt + \sum_{l=1}^r h_l(x)u(t, x)dY_l(t), \quad (2.4)$$

where

$$\mathcal{L}^*u := \frac{1}{2} \sum_{i,j=1}^d \frac{\partial^2}{\partial x_i \partial x_j} ((\sigma\sigma^*)_{ij}u) - \sum_{i=1}^d \frac{\partial}{\partial x_i} (b_i u).$$

By Theorem 4.3.2 in [20], there is a unique solution $u(t, x)$ of (2.4) and for every positive integer n and a multi-index $\gamma = (\gamma_1, \dots, \gamma_d)$,

$$\sup_{0 \leq t \leq T} \tilde{\mathbf{E}} \int_{\mathbb{R}^d} (1 + |x|^2)^n |D^\gamma u(t, x)|^2 dx < \infty, \quad (2.5)$$

where $D^{|\gamma|} = \frac{\partial^\gamma}{\partial x_1^{\gamma_1} \dots \partial x_d^{\gamma_d}}$ and $|\gamma| = \gamma_1 + \dots + \gamma_d$.

Throughout the paper, C will denote a constant depending on the parameters of the model, namely the length T of the time interval, dimensions d and r of the state and observation processes, and various bounds on the coefficients b , σ , h and the initial density p and their derivatives. The value of C can be different at different places. The norm and inner product in $L_2(\mathbb{R}^d)$ are denoted by $\|\cdot\|_0$ and $(\cdot, \cdot)_0$, respectively.

3 Galerkin approximation

In this section, a Galerkin approximation of equation (2.4) will be studied using the Hermite basis in $L_2(\mathbb{R}^d)$.

The Hermite basis in $L_2(\mathbb{R}^d)$ is constructed as follows. Let $\gamma = (\gamma_1, \dots, \gamma_d)$ be a multi-index with γ_i nonnegative integers. The set of all such multi-indices will be denoted by Γ . For every $\gamma \in \Gamma$ define

$$e_\gamma(x_1, \dots, x_d) = \prod_{i=1}^d e_{\gamma_i}(x_i), \quad (3.1)$$

where

$$e_{\gamma_i}(t) := \frac{1}{\sqrt{2^n \pi^{1/2} n!}} e^{-t^2/2} H_{\gamma_i}(t)$$

and

$$H_n(t) = (-1)^n e^{t^2} \frac{d^n}{dt^n} e^{-t^2}, \quad n \geq 0, \quad (3.2)$$

is the n -th Hermite polynomial. Then the set $\{e_\gamma\}_{\gamma \in \Gamma}$ is an orthonormal basis in $L_2(\mathbb{R}^d)$ [7, 8] and will be referred to as the Hermite basis.

For $\gamma \in \Gamma$ denote $|\gamma| := \sum_{i=1}^d \gamma_i$. Given a positive integer κ , define the set $\Gamma_\kappa := \{\gamma \in \Gamma : |\gamma| \leq \kappa\}$. Direct computations show that the number of elements $|\Gamma_\kappa|$ in the set Γ_κ is equal to $(\kappa + d)! / (\kappa! d!)$.

The Galerkin approximation u^κ of the solution of (2.4) is given by $u^\kappa(t, x) = \sum_{\gamma \in \Gamma_\kappa} u_\gamma^\kappa(t) e_\gamma(x)$, where the coefficients $\{u_\gamma^\kappa\}_{\gamma \in \Gamma_\kappa}$ satisfy the system of stochastic ordinary differential equations

$$\begin{aligned} du_\gamma^\kappa(t) &= \sum_{\tau \in \Gamma_\kappa} (\mathcal{L}^* e_\tau, e_\gamma)_0 u_\tau^\kappa(t) dt + \\ &\quad \sum_{l=1}^r \sum_{\tau \in \Gamma_\kappa} (h_l e_\tau, e_\gamma)_0 u_\tau^\kappa(t) dY_l(t), \quad 0 < t \leq T, \\ u_\gamma^\kappa(0) &= (p, e_\gamma)_0, \quad \gamma \in \Gamma. \end{aligned} \quad (3.3)$$

This is a linear system with constant coefficients, so the solution $\{u_\gamma^\kappa(t)\}_{\gamma \in \Gamma_\kappa}$ exists and is unique [13, Theorem 4.6]. Even though the Galerkin approximation has been used in various numerical schemes for the Zakai equation (e.g. Bennaton [1], Ito [10]), no estimates of the approximation error have been given.

In what follows, it is proved that under assumptions **(A1 – A3)** the approximation u^κ converges to u in $L_2(\Omega, \tilde{\mathbf{P}})$, i.e.

$$\lim_{\kappa \rightarrow \infty} \sup_{0 \leq t \leq T} \tilde{\mathbf{E}} \|u(t, \cdot) - u^\kappa(t, \cdot)\|_0^2 = 0,$$

and the rate of convergence is established.

Theorem 3.1 *If assumptions **(A1 – A3)** hold, then for every positive integer ν there is a constant $C(\nu)$ depending only on ν and the parameters of the model so that*

$$\sup_{0 \leq t \leq T} \tilde{\mathbf{E}} \|u(t, \cdot) - u^\kappa(t, \cdot)\|_0^2 \leq \frac{C(\nu)}{\kappa^\nu}. \quad (3.4)$$

Proof. If $u_\gamma(t) := (u(t, \cdot), e_\gamma)_0$, then

$$\tilde{\mathbf{E}} \|u(t, \cdot) - u^\kappa(t, \cdot)\|_0^2 = \sum_{\gamma \in \Gamma_\kappa} \tilde{\mathbf{E}} |u_\gamma(t) - u_\gamma^\kappa(t)|^2 + \sum_{\gamma \notin \Gamma_\kappa} \tilde{\mathbf{E}} |u_\gamma(t)|^2. \quad (3.5)$$

By Lemma A.1 in Appendix

$$\sup_{0 \leq t \leq T} \sum_{\gamma \notin \Gamma_\kappa} \tilde{\mathbf{E}} |u_\gamma(t)|^2 \leq \frac{C(\nu)}{\kappa^\nu}. \quad (3.6)$$

For $\gamma \in \Gamma_\kappa$ define $\delta_\gamma(t) := u_\gamma(t) - u_\gamma^\kappa(t)$, so that $\sum_{\gamma \in \Gamma_\kappa} \tilde{\mathbf{E}} |u_\gamma(t) - u_\gamma^\kappa(t)|^2 = \sum_{\gamma \in \Gamma_\kappa} \tilde{\mathbf{E}} |\delta_\gamma|^2$, and also

$$\delta_{1,\gamma}(t) := \sum_{\tau \notin \Gamma_\kappa} (\mathcal{L}^* e_\tau, e_\gamma)_0 u_\tau(t), \quad \delta_{2,\gamma}^l(t) := \sum_{\tau \notin \Gamma_\kappa} (h_l e_\tau, e_\gamma)_0 u_\tau(t).$$

Both $\delta_{1,\gamma}$ and $\delta_{2,\gamma}^l$ are well defined, because, according to Lemmas A.1 and A.3,

$$\sum_{\gamma \in \Gamma_\kappa} \int_0^T \tilde{\mathbf{E}} (\delta_{1,\gamma}(s))^2 ds + \sum_{l=1}^r \sum_{\gamma \in \Gamma_\kappa} \int_0^T \tilde{\mathbf{E}} (\delta_{2,\gamma}^l(s))^2 ds \leq \frac{C(\nu)}{\kappa^\nu}. \quad (3.7)$$

As a result,

$$\begin{aligned} d\delta_\gamma(t) &= \sum_{\tau \in \Gamma_\kappa} (\mathcal{L}^* e_\tau, e_\gamma)_0 \delta_\gamma(t) dt + \sum_{l=1}^r \sum_{\tau \in \Gamma_\kappa} (h_l e_\tau, e_\gamma)_0 \delta_\gamma(t) dY_l(t) + \\ &\quad \delta_{1,\gamma} ds + \sum_{l=1}^r \delta_{2,\gamma}^l dY_l(t), \quad 0 < t \leq T; \\ \delta_\gamma(0) &= 0, \quad \gamma \in \Gamma_\kappa, \end{aligned} \quad (3.8)$$

and by the Ito formula,

$$\begin{aligned}
\sum_{\gamma \in \Gamma_\kappa} \tilde{\mathbf{E}} |\delta_\gamma(t)|^2 &= 2 \int_0^t \sum_{\gamma, \tau \in \Gamma_\kappa} (\mathcal{L}^* e_\tau, e_\gamma)_0 \tilde{\mathbf{E}} \delta_\gamma(s) \delta_\tau(s) ds + \\
&\sum_{l=1}^r \sum_{\gamma \in \Gamma_\kappa} \int_0^t \tilde{\mathbf{E}} \left(\sum_{\tau \in \Gamma_\kappa} (h_l e_\tau, e_\gamma)_0 \delta_\tau(s) \right)^2 ds + 2 \sum_{\gamma \in \Gamma_\kappa} \int_0^t \tilde{\mathbf{E}} \delta_{1,\gamma}(s) \delta_\gamma(s) ds + \\
&2 \sum_{l=1}^r \sum_{\gamma, \tau \in \Gamma_\kappa} \int_0^t (h_l e_\tau, e_\gamma)_0 \tilde{\mathbf{E}} \delta_{2,\gamma}^l(s) \delta_\tau(s) ds + \sum_{l=1}^r \sum_{\gamma \in \Gamma_\kappa} \int_0^t \tilde{\mathbf{E}} (\delta_{2,\gamma}^l(s))^2 ds.
\end{aligned} \tag{3.9}$$

It follows from Lemma A.2 that

$$\begin{aligned}
&2 \int_0^t \sum_{\gamma, \tau \in \Gamma_\kappa} (\mathcal{L}^* e_\tau, e_\gamma)_0 \tilde{\mathbf{E}} \delta_\gamma(s) \delta_\tau(s) ds + \\
&\sum_{l=1}^r \sum_{\gamma \in \Gamma_\kappa} \int_0^t \tilde{\mathbf{E}} \left(\sum_{\tau \in \Gamma_\kappa} (h_l e_\tau, e_\gamma)_0 \delta_\tau(s) \right)^2 ds \leq C \sum_{\gamma \in \Gamma_\kappa} \int_0^t \tilde{\mathbf{E}} (\delta_\gamma(s))^2 ds.
\end{aligned} \tag{3.10}$$

Then (3.7), (3.9), (3.10), and the obvious inequality $2|ab| \leq a^2 + b^2$ imply

$$\sum_{\gamma \in \Gamma_\kappa} \tilde{\mathbf{E}} |\delta_\gamma(t)|^2 \leq \frac{C(\nu)}{\kappa^\nu} + C \sum_{\gamma \in \Gamma_\kappa} \int_0^t \tilde{\mathbf{E}} |\delta_\gamma(s)|^2 ds,$$

so that by the Gronwall inequality

$$\sup_{0 \leq t \leq T} \sum_{\gamma \in \Gamma_\kappa} \tilde{\mathbf{E}} |\delta_\gamma(t)|^2 \leq \frac{C(\nu)}{\kappa^\nu}.$$

Together with (3.5) and (3.6), the last inequality implies (3.4). \square

4 Approximate solution of the ODE

By Theorem 3.1, to find an approximate solution of the partial differential equation (2.4) it is sufficient to solve the finite system of ordinary differential equations (3.3). In this section, the system (3.3) is solved using the Cameron-Martin version of the Wiener chaos decomposition. An approximate solution is then constructed and the error of the approximation is estimated.

Using the matrix-vector notation, system (3.3) can be written as

$$\begin{aligned}
du^\kappa(t) &= A^\kappa u^\kappa(t) dt + \sum_{l=1}^r B_l^\kappa u^\kappa(t) dY_l(t), \quad 0 < t \leq T; \\
u^\kappa(0) &= \bar{p}^\kappa,
\end{aligned} \tag{4.1}$$

where $u^\kappa(t)$ is the column vector $\{u_\gamma^\kappa(t)\}_{\gamma \in \Gamma_\kappa}$, \bar{p}^κ is the column vector $\{(p, e_\gamma)_0\}_{\gamma \in \Gamma_\kappa}$, and for every vector $\zeta \in \mathbb{R}^{|\Gamma_\kappa|}$,

$$\begin{aligned}
(A^\kappa \zeta)_\gamma &:= \sum_{\tau \in \Gamma_\kappa} (\mathcal{L}^* e_\tau, e_\gamma)_0 \zeta_\tau, \\
(B_l^\kappa \zeta)_\gamma &:= \sum_{l=1}^r \sum_{\tau \in \Gamma_\kappa} (h_l e_\tau, e_\gamma)_0 \zeta_\tau.
\end{aligned} \tag{4.2}$$

It follows from the definition that the matrices A^κ and B_l^κ do not, in general, commute with each other and therefore there is no *computable* explicit solution of (4.1).

Below, an approximate solution is constructed using the Cameron-Martin version of the Wiener chaos decomposition, i.e. an expansion of $u^\kappa(t)$ with respect to a certain orthonormal basis in $L_2(\Omega, \mathcal{F}_t^y, \tilde{\mathbf{P}})$. The basis is constructed as follows. Let α be an r -dimensional multi-index, i.e. a collection $\alpha = (\alpha_k^l)_{1 \leq l \leq r, k \geq 1}$ of nonnegative integers such that only finitely many of α_k^l are different from zero. The set of all such multi-indices will be denoted by J . For $\alpha \in J$ define:

$$|\alpha| := \sum_{l,k} \alpha_k^l \quad (\text{length of } \alpha);$$

$$d(\alpha) := \max\{k \geq 1 : \alpha_k^l > 0 \text{ for some } 1 \leq l \leq r\} \quad (\text{order of } \alpha);$$

$$\alpha! := \prod_{k,l} (\alpha_k^l!).$$

Fix $0 < t \leq T$, choose a complete orthonormal system $\{m_k\} = \{m_k(s)\}_{k \geq 1}$ in $L_2([0, t])$, and define

$$\xi_{k,l} = \int_0^t m_k(s) dY_l(s).$$

Due to property (i) of the measure $\tilde{\mathbf{P}}$, $\xi_{k,l}$ are independent Gaussian random variables with zero mean and unit variance.

Theorem 4.1 (Cameron and Martin [3])

If

$$\bar{H}(t) := 2^{-n/2} H_n(t/\sqrt{2}), \quad (4.3)$$

where H_n is the n -th Hermite polynomial (3.2), then the collection

$$\left\{ \xi_\alpha := \prod_{k,l} \left(\frac{\bar{H}_{\alpha_k^l}(\xi_{k,l})}{\sqrt{\alpha_k^l!}} \right), \quad \alpha \in J \right\}$$

is a complete orthonormal system in $L_2(\Omega, \mathcal{F}_t^y, \tilde{\mathbf{P}})$.

Theorem 4.2 For every $0 \leq s \leq t$ the solution of (4.1) can be written as

$$u^\kappa(s) = \sum_{\alpha \in J} \frac{1}{\sqrt{\alpha!}} \varphi_\alpha^\kappa(s, \bar{p}^\kappa) \xi_\alpha \quad (\mathbf{P}\text{- a.s.}) \quad (4.4)$$

The coefficients of the expansion satisfy the recursive system of deterministic ordinary differential equations

$$\begin{aligned} \frac{d\varphi_\alpha^\kappa(s, \bar{p}^\kappa)}{ds} &= A^\kappa \varphi_\alpha^\kappa(s, \bar{p}^\kappa) + \sum_{k,l} \alpha_k^l m_k(s) B_l^\kappa \varphi_{\alpha(k,l)}^\kappa(s, \bar{p}^\kappa), \quad 0 < s \leq t; \\ \varphi_\alpha^\kappa(0, \bar{p}^\kappa) &= \bar{p}^\kappa \mathbf{1}_{\{|\alpha|=0\}}, \end{aligned} \quad (4.5)$$

where $\alpha = (\alpha_k^l)_{1 \leq l \leq r, k \geq 1} \in J$ and $\alpha(i, j)$ stands for the multi-index $\tilde{\alpha} = (\tilde{\alpha}_k^l)_{1 \leq l \leq r, k \geq 1}$ with

$$\tilde{\alpha}_k^l = \begin{cases} \alpha_k^l & \text{if } k \neq i \text{ or } l \neq j \text{ or both} \\ \max(0, \alpha_k^j - 1) & \text{if } k = i \text{ and } l = j. \end{cases} \quad (4.6)$$

The series in (4.4) converges in $L_2(\Omega, \tilde{\mathbf{P}})$ and $L_1(\Omega, \mathbf{P})$ and the following Parseval's equality holds:

$$\tilde{\mathbf{E}}|u^\kappa(s)|^2 = \sum_{\alpha \in J} \frac{1}{\alpha!} |\varphi_\alpha^\kappa(s, \bar{p}^\kappa)|^2. \quad (4.7)$$

Proof. Since the initial condition $u^\kappa(0)$ and the matrices A^κ and B_l^κ are deterministic, the vector $u^\kappa(s)$ belongs to $L_2(\Omega, \mathcal{F}_t^y, \tilde{\mathbf{P}})$ for every $0 \leq s \leq t$ by Theorem 4.6 in [13], and therefore by Theorem 4.1

$$u^\kappa(s) = \sum_{\alpha \in J} \tilde{\mathbf{E}}[u^\kappa(s)\xi_\alpha]\xi_\alpha, \quad \tilde{\mathbf{E}}|u^\kappa(s)|^2 = \sum_{\alpha \in J} |\tilde{\mathbf{E}}[u^\kappa(s)\xi_\alpha]|^2 < \infty.$$

Denote

$$\varphi_\alpha^\kappa(s) := \sqrt{\alpha!} \tilde{\mathbf{E}}[u^\kappa(s)\xi_\alpha]$$

(to simplify the notations, \bar{p}^κ in the argument of φ_α^κ will be omitted). Direct computations show that

$$\xi_\alpha = \frac{1}{\sqrt{\alpha!}} \frac{\partial^\alpha}{\partial z^\alpha} P_t(z) \Big|_{z=0},$$

where

$$P_t(z) = \exp \left\{ \int_0^t \sum_{l=1}^r m_z^l(s) dY_l(s) - \frac{1}{2} \int_0^t \sum_{l=1}^r |m_z^l(s)|^2 ds \right\};$$

$m_z^l = \sum_{k \geq 1} m_k(s) z_k^l$; $\{z_k^l\}$, $l = 1, \dots, r$; $k = 1, 2, \dots$, is a sequence of real numbers such that $\sum_{k,l} |z_k^l|^2 < \infty$, and

$$\frac{\partial^\alpha}{\partial z^\alpha} := \prod_{k,l} \frac{\partial^{\alpha_k^l}}{(\partial z_k^l)^{\alpha_k^l}}.$$

Consequently,

$$\varphi_\alpha^\kappa(s) = \frac{\partial^\alpha}{\partial z^\alpha} \tilde{\mathbf{E}}[u^\kappa(s)P_t(z)] \Big|_{z=0} = \frac{\partial^\alpha}{\partial z^\alpha} \tilde{\mathbf{E}}[u^\kappa(s)P_s(z)] \Big|_{z=0},$$

where the second equality follows from the martingale property of $P_s(z)$ on $(\Omega, \tilde{\mathbf{P}})$. It also follows from the definition of $P_s(z)$ that

$$dP_s(z) = \sum_{l=1}^r m_z^l(s) P_s(z) dY_l(s), \quad s \leq t; \quad P_0(z) = 1.$$

Then (4.1) and the Ito formula imply

$$\begin{aligned} u^\kappa(t)P_t(z) &= \bar{p}^\kappa + \int_0^t \left(A^\kappa u^\kappa(s) + \sum_{l=1}^r B_l^\kappa u^\kappa(s) \right) m_z^l(s) P_s(z) ds + \\ &\quad \int_0^t \sum_{l=1}^r \left(B_l^\kappa u^\kappa(s) + u^\kappa(s) m_z^l(s) \right) P_s(z) dY_l(s). \end{aligned}$$

Taking expectation $\tilde{\mathbf{E}}$ on both sides of the last equality and setting $\varphi^\kappa(s, z) := \tilde{\mathbf{E}}u^\kappa(s)P_s(z)$ results in

$$\begin{aligned}\frac{\partial \varphi^\kappa(s, z)}{\partial s} &= A^\kappa \varphi_\alpha^\kappa(s) + \sum_{l=1}^r m_z^l(s) B_l^\kappa \varphi^\kappa(s, z), \quad 0 < s \leq t \\ \varphi^\kappa(0, z) &= \bar{p}^\kappa.\end{aligned}$$

Applying the operator $\frac{1}{\sqrt{\alpha!}} \frac{\partial^\alpha}{\partial z^\alpha}$ and setting $z = 0$ yields (4.5).

Convergence of (4.4) in $L_1(\Omega, \mathbf{P})$ follows from its convergence in $L_2(\Omega, \tilde{\mathbf{P}})$ and the inequality

$$\mathbf{E}|\eta| = \tilde{\mathbf{E}}|\eta\rho(T)| \leq \sqrt{\tilde{\mathbf{E}}|\eta|^2} \sqrt{\tilde{\mathbf{E}}\rho^2(T)} \leq C \sqrt{\tilde{\mathbf{E}}|\eta|^2},$$

valid for every $\eta \in L_2(\Omega, \mathcal{F}_t^y, \tilde{\mathbf{P}})$ due to the boundedness of h_l . □

System (4.5) is recursive in $|\alpha|$: once the functions φ_α^κ are known for all α of length $|\alpha| = k$, it is possible to compute all φ_α^κ for $|\alpha| = k + 1$. The computations can be performed off line because (4.5) does not involve the observation process Y . It seems natural to use these features of expansion (4.4) in a numerical algorithm for solving (4.1). Since in all such algorithms the infinite summation in (4.4) must be truncated and the error due to the truncation can be expected to grow with t , it is desirable to have a recursive version of Theorem 4.2.

Let $0 = t_0 < t_1 < \dots < t_M = T$ be a uniform partition of the interval $[0, T]$ with step Δ (so that $t_i = i\Delta$, $i = 0, \dots, M$) and let $\{m_k\} = \{m_k(t)\}_{k \geq 1}$ be an orthonormal basis in $L_2([0, \Delta])$. Define random variables

$$\left\{ \xi_\alpha^i := \prod_{k,l} \left(\frac{\bar{H}_{\alpha_k^l}(\xi_{k,l}^i)}{\sqrt{\alpha_k^l!}} \right), \quad \alpha \in J \right\},$$

where

$$\xi_{k,l}^i = \int_{t_{i-1}}^{t_i} m_k(s - t_{i-1}) dY_l(s) \quad (4.8)$$

and \bar{H}_n is defined in (4.3).

Theorem 4.3 *If $u^\kappa(t_0) := \bar{p}^\kappa$ and $t_{i-1} \leq s \leq t_i$, then*

$$u^\kappa(s) = \sum_{\alpha \in J} \frac{1}{\sqrt{\alpha!}} \varphi_\alpha^\kappa(s - t_i, u^\kappa(t_{i-1})) \xi_\alpha^i, \quad i = 1, \dots, M \quad (\mathbf{P}\text{- a.s.}) \quad (4.9)$$

and the coefficients φ_α^κ satisfy (4.5) with the corresponding initial condition.

This series converges in $L_2(\Omega, \tilde{\mathbf{P}})$ and $L_1(\Omega, \mathbf{P})$ and the following Parseval's equality holds:

$$\tilde{\mathbf{E}}|u^\kappa(s)|^2 = \sum_{\alpha \in J} \frac{1}{\alpha!} \tilde{\mathbf{E}}|\varphi_\alpha^\kappa(s - t_i, u^\kappa(t_{i-1}))|^2, \quad i = 1, \dots, M.$$

Proof. Since the coefficients of (4.1) do not depend on time, for every $t_{i-1} \leq s \leq t_i$ the vector $u^\kappa(s)$ satisfies

$$u^\kappa(s) = u^\kappa(t_{i-1}) + \int_0^{s-t_{i-1}} A^\kappa u^\kappa(t+t_{i-1})dt + \sum_{l=1}^r \int_0^{s-t_{i-1}} B_l^\kappa u^\kappa(t+t_{i-1})d(Y_l(t+t_{i-1}) - Y_l(t_{i-1})). \quad (4.10)$$

Then both statements of the theorem follow from the corresponding statements of Theorem 4.2, because, under the measure $\tilde{\mathbf{P}}$, $Y_l(t+t_{i-1}) - Y_l(t_{i-1})$ is independent of $\mathcal{F}_{t_{i-1}}^y$. \square

To study the truncation of (4.4) or (4.9), it is necessary to note that the summation $\sum_{\alpha \in J}$ is double infinite: $\sum_{\alpha \in J} = \sum_{k=0}^{\infty} \sum_{|\alpha|=k}$ and there are infinitely many multi-indices α with $|\alpha| = k > 0$. A possible way to truncate the summation is to restrict α to the finite set $J_N^n := \{\alpha \in J : |\alpha| \leq N, d(\alpha) \leq n\}$. The error due to this truncation is given in the following theorem.

Theorem 4.4 *Given the positive integers n and N and the uniform partition $\{t_i\}$, $i = 0, \dots, M$ of $[0, T]$ with step Δ , define the sequence of vectors $\{u^\kappa(i, n, N)\} \subset \mathbb{R}^{|\Gamma^\kappa|}$, $i = 0, \dots, M$, by*

$$u^\kappa(0, n, N) := \bar{p}^\kappa, \quad u^\kappa(i, n, N) := \sum_{\alpha \in J_N^n} \frac{1}{\sqrt{\alpha!}} \varphi_\alpha^\kappa(\Delta, u^\kappa(i-1, n, N)) \xi_\alpha^i. \quad (4.11)$$

If $\{m_k\}$ is the Fourier cosine basis

$$m_1(s) = \frac{1}{\sqrt{\Delta}}; \quad m_k(s) = \sqrt{\frac{2}{\Delta}} \cos\left(\frac{\pi(k-1)s}{\Delta}\right), \quad k > 1; \quad 0 \leq s \leq \Delta, \quad (4.12)$$

then

$$\max_{0 \leq i \leq M} \tilde{\mathbf{E}} |u^\kappa(t_i) - u^\kappa(i, n, N)|^2 \leq \frac{(C\Delta)^N}{(N+1)!} + C \frac{\Delta + \kappa^2 \Delta^2}{n}. \quad (4.13)$$

The proof of this theorem is rather long and is given in Section 6.

For fixed Δ the error (4.13) grows with κ due to the stiffness of matrix A^κ . For fixed κ , the error tends to zero with Δ . In fact, since $N \geq 1$ and $n \geq 1$, (4.13) shows that the approximation $u^\kappa(i, n, N)$ has order of strong convergence in time equal to 0.5. If $n = 1$, then the approximation uses only the increments $Y_l(t_{i+1}) - Y_l(t_i)$ of the observation process, and it is known [4] that 0.5 is the highest possible rate of convergence under the circumstances. On the other hand, if $r = 1$ or the matrices B_l^κ commute (which, in general, can happen only if the functions h_l are multiples of each other), then, as the proof of the theorem shows, the term $C\Delta/n$ on the right hand side of (4.13) disappears and the order of convergence may be equal to 1.0.

If $n > 1$ and the observations $Y_l(t)$ are available continuously in time, then the random variables ξ_α^i can be computed with prescribed accuracy by reducing the Wiener integrals (4.8) to Riemann integrals using integration by parts; the details can be found in [16, Section 4.5.1].

5 The Algorithm

The approximation $u^\kappa(i, n, N)$ from Theorem 4.4 still does not meet the criteria stated in the Introduction because, starting with $i = 1$, the initial condition of (4.5) is random and therefore the system of differential equations must be solved on line. The following construction makes it possible to circumvent the difficulty.

If $\{\zeta^\gamma\}$, $\gamma \in \Gamma_\kappa$, is the standard unit basis in $\mathbb{R}^{|\Gamma_\kappa|}$, meaning that $\zeta_\tau^\gamma = 1$ if $\gamma = \tau$ and $\zeta_\tau^\gamma = 0$ if $\gamma \neq \tau$, then $u^\kappa(i, n, N) = \sum_{\gamma \in \Gamma_\kappa} u_\gamma^\kappa(i, n, N) \zeta^\gamma$. Recursive definition of $u^\kappa(i, n, N)$ and the linearity of system (4.5) imply that

$$u_\gamma^\kappa(0, n, N) = \bar{p}_\gamma^\kappa, \quad u_\gamma^\kappa(i, n, N) = \sum_{\tau \in \Gamma_\kappa} \sum_{\alpha \in J_N^n} Q_{\alpha, \gamma \tau}^\kappa u_\tau^\kappa(i-1, n, N) \xi_\alpha^i, \quad (5.1)$$

where

$$Q_{\alpha, \gamma \tau}^\kappa = \varphi_{\alpha, \gamma}^\kappa(\Delta, \zeta^\tau). \quad (5.2)$$

Using the matrix-vector notations, (5.1) can be written as

$$u^\kappa(i, n, N) = \sum_{\alpha \in J_N^n} Q_\alpha^\kappa u^\kappa(i-1, n, N) \xi_\alpha^i. \quad (5.3)$$

Since the matrices Q_α^κ can be computed off line, the on-line part of the algorithm is reduced to computing random variables ξ_α^i (which, in principle, can be implemented on the hardware level) and evaluating the sum in (5.3).

If $u(t, x)$ is the exact solution of the Zakai equation (2.4) and $u_{n, N}^\kappa(i, x) := \sum_{\gamma \in \Gamma_\kappa} u_\gamma^\kappa(i, n, N) e_\gamma(x)$, then, by Theorems 3.1 and 4.4, the overall error of approximation is given by

$$\sup_{0 \leq i \leq M} \tilde{\mathbf{E}} \|u(i\Delta, \cdot) - u_{n, N}^\kappa(i, \cdot)\|_0^2 \leq \frac{C(\nu)}{\kappa^\nu} + \frac{(C\Delta)^N}{(N+1)!} + C \frac{\Delta + \kappa^2 \Delta^2}{n}. \quad (5.4)$$

The following is the complete description of the algorithm.

1. Off line (before the observations are available) compute

the matrices A^κ and B_1^κ according to (4.2);

the matrices Q_α^κ , $\alpha \in J_N^n$, according to (5.2).

the coefficients $u_\gamma^\kappa(0, n, N) = \int_{\mathbb{R}^d} p_0(x) e_\gamma(x) dx$, $\gamma \in \Gamma_\kappa$;

set $u_{n, N}^\kappa(0, x) = \sum_{\gamma \in \Gamma_\kappa} u_\gamma^\kappa(0, n, N) e_\gamma(x)$.

2. On line, i -th step (as the observations become available): compute

$$u_\gamma^\kappa(i, n, N) = \sum_{\alpha \in J_N^n} Q_\alpha^\kappa u_\gamma^\kappa(i-1, n, N) \xi_\alpha^i, \quad (5.5)$$

then, if necessary, compute

$$u_{n, N}^\kappa(i, x) = \sum_{\gamma \in \Gamma_\kappa} u_\gamma^\kappa(i, n, N) e_\gamma(x). \quad (5.6)$$

The main advantage of this algorithm as compared to the traditional schemes for solving the Zakai equation is that the time consuming computations of solving partial differential equations are performed off line, while the on-line part is relatively simple even when the dimension d of the state process is large. Here are some other features of the algorithm:

- (1) The overall amount of the off-line computations does not depend on the number of the on-line time steps;
- (2) Only the coefficients $u_\gamma^\kappa(i, n, N)$ must be computed at every time step while the approximate filter $\tilde{f}_{t_i, \kappa}$ and/or UFD $u_{n, N}^\kappa(i, x)$ can be computed as needed, e.g. at the final time moment.
- (3) The on-line part of the algorithm can be easily parallelized.

In the proposed algorithm, the Wiener chaos decomposition was used to solve the finite dimensional system obtained by projecting the original Zakai equation (2.4) on the subspace generated by $\{e_\gamma\}_{\gamma \in \Gamma_\kappa}$. The error estimate (5.4) shows that this approach should be used with sufficiently small time step Δ and moderate values of κ , e.g. when good resolution in time is required and the spatial resolution is not important. In [15], an alternative approach was suggested, in which the Wiener chaos decomposition was applied directly to the Zakai equation with subsequent projection of the result on the span of $\{e_\gamma\}_{\gamma \in \Gamma_\kappa}$. The error of the corresponding approximation is given by

$$\frac{C(\nu)}{\kappa^\nu \Delta} + \frac{C\Delta^2}{n} + \frac{(C\Delta)^N}{(N+1)!},$$

which means that this alternative approach should be efficient for moderate values of Δ and sufficiently large κ , e.g. when good resolution in space is required and the time resolution is not important. The choice between the two approaches will therefore depend on the specific features of the problem.

6 Proof of Theorem 4.4

Introduce the following notations:

$$\Phi_t^\kappa := e^{tA^\kappa};$$

$$s^k, \text{ the ordered set } (s_1, \dots, s_k); ds^k := ds_1 \dots ds_k;$$

$$l^k, \text{ the ordered set } (l_1, \dots, l_k);$$

$$F^\kappa(t; s^k; l^k) := \Phi_{t-s_k}^\kappa B_{l_k}^\kappa \Phi_{s_k-s_{k-1}}^\kappa \dots B_{l_1}^\kappa \Phi_{s_1}^\kappa \bar{p}^\kappa, \quad k \geq 1;$$

$$\int^{(k,t)} (\dots) ds^k := \int_0^t \int_0^{s_k} \dots \int_0^{s_2} (\dots) ds_1 \dots ds_k;$$

$$\sum_{l^k} := \sum_{l_1, \dots, l_k=1}^r .$$

To simplify the notations, the superscript κ will be omitted throughout the rest of the proof.

Every multi-index α with $|\alpha| = k$ can be identified with the set $K_\alpha = \{(i_1^\alpha, q_1^\alpha), \dots, (i_k^\alpha, q_k^\alpha)\}$ so that $i_1^\alpha \leq i_2^\alpha \leq \dots \leq i_k^\alpha$ and if $i_j^\alpha = i_{j+1}^\alpha$, then $q_j^\alpha \leq q_{j+1}^\alpha$. The first pair (i_1^α, q_1^α) in K_α is the position numbers (k, l) of the first nonzero element α_k^l of α . The second pair is the same as the first if the first nonzero element of α is greater than one; otherwise, the second pair is the position numbers of the second nonzero entry of α and so on. As a result, if $\alpha_j^q > 0$, then exactly α_j^q pairs in K_α are (j, q) . The set K_α will be referred to as *the characteristic set* of multi-index α . For example, if $r = 2$ and

$$\alpha = \begin{pmatrix} 0 & 1 & 0 & 2 & 3 & 0 & 0 & \dots \\ 1 & 2 & 0 & 0 & 0 & 1 & 0 & \dots \end{pmatrix},$$

then the nonzero elements are $\alpha_1^2 = \alpha_2^1 = \alpha_2^6 = 1$, $\alpha_2^2 = \alpha_4^1 = 2$, $\alpha_5^1 = 3$, and the characteristic set is

$K_\alpha = \{(1, 2), (2, 1), (2, 2), (2, 2), (4, 1), (4, 1), (5, 1), (5, 1), (5, 1), (6, 2)\}$. In the future, when there is no danger of confusion, the superscript α in i and q will be omitted (i.e. (i_j, q_j) will be used instead of (i_j^α, q_j^α)).

Let \mathcal{P}^k be the permutation group of the set $\{1, \dots, k\}$. For a given $\alpha \in J$ with $|\alpha| = k$ and the characteristic set $\{(i_1, q_1), \dots, (i_k, q_k)\}$ define

$$E_\alpha(s^k; l^k) := \sum_{\sigma \in \mathcal{P}^k} m_{i_1}(s_{\sigma(1)}) 1_{\{l_{\sigma(1)}=q_1\}} \cdots m_{i_k}(s_{\sigma(k)}) 1_{\{l_{\sigma(k)}=q_k\}}. \quad (6.1)$$

To prove (4.13) it suffices to show that

$$\tilde{\mathbf{E}}|u(t_i) - u(i, n, N)|^2 \leq \frac{(C\Delta)^{N+1}}{(N+1)!} + C \frac{\Delta^2 + \kappa^2 \Delta^3}{n}. \quad (6.2)$$

After that, (4.13) will follow directly from Gronwall's inequality. Moreover, (6.2) will follow from

$$\tilde{\mathbf{E}}|u(t_1) - u(1, n, N)|^2 \leq \left(\frac{(C\Delta)^{N+1}}{(N+1)!} + C \frac{\Delta^2 + \kappa^2 \Delta^3}{n} \right) |\bar{p}|^2, \quad (6.3)$$

because similar arguments can then be used to show that

$$\tilde{\mathbf{E}}|u(t_i) - u(i, n, N)|^2 \leq \left(\frac{(C\Delta)^{N+1}}{(N+1)!} + C \frac{\Delta^2 + \kappa^2 \Delta^3}{n} \right) \tilde{\mathbf{E}}|u(t_i)|^2,$$

and by Lemma A.4, $\max_{0 \leq i \leq M} \tilde{\mathbf{E}}|u(t_i)|^2 \leq C$. As a result, it suffices to prove (6.3). For the sake of simplicity, the argument \bar{p} in φ_α will be omitted.

Define

$$u(\Delta, N) := \sum_{|\alpha| \leq N} \frac{\varphi_\alpha(\Delta) \xi_\alpha^1}{\sqrt{\alpha!}}.$$

If it is known that

$$\tilde{\mathbf{E}}|u(\Delta) - u(\Delta, N)|^2 \leq \frac{(C\Delta)^{N+1}}{(N+1)!} |\bar{\rho}|^2 \quad (6.4)$$

and

$$\tilde{\mathbf{E}}|u(\Delta, N) - u(1, n, N)|^2 \leq C \left(\frac{\Delta^2 + \kappa^2 \Delta^3}{n} \right) |\bar{\rho}|^2, \quad (6.5)$$

then (6.3) will follow immediately from the inequality $(a+b)^2 \leq 2(a^2 + b^2)$.

Proof of (6.4). By Lemma A.5,

$$\sum_{|\alpha|=k} \frac{|\varphi_\alpha(\Delta)|^2}{\alpha!} = \sum_{l^k} \int^{(k, \Delta)} |F(\Delta; s^k; l^k)|^2 ds^k, \quad (6.6)$$

where φ_α is the solution of (4.5) with an arbitrary CONS $\{m_k\}$ in $L_2([0, \Delta])$.

Since ξ_α^1 are uncorrelated under $\tilde{\mathbf{P}}$,

$$\tilde{\mathbf{E}}|u(\Delta) - u(\Delta, N)|^2 = \sum_{|\alpha|=k} \frac{|\varphi_\alpha(\Delta)|^2}{\alpha!} = \sum_{k>N} \sum_{l^k} \int^{(k, \Delta)} |F(\Delta; s^k; l^k)|^2 ds^k.$$

It follows from the definition of F and Lemma A.5 that

$$\begin{aligned} |F(\Delta; s^k; l^k)| &\leq e^{C(\Delta-s_k)} |\Phi_{s_k-s_{k-1}} B_{l_k} \dots B_{l_1} \Phi_{s_1} \bar{\rho}| \leq \\ &\dots \leq C^k e^{C(\Delta-s_k+s_k-\dots+s_2-s_1+s_1)} |\bar{\rho}|^2 \leq C^k |\bar{\rho}|^2. \end{aligned}$$

Finally, $\int^{(k, \Delta)} ds^k = \Delta^k / k!$ implies

$$\tilde{\mathbf{E}}|u(\Delta) - u(\Delta, N)|^2 \leq \sum_{k>N} \frac{(Cr\Delta)^k}{k!} \leq \frac{(C\Delta)^{N+1}}{(N+1)!} e^{C\Delta},$$

and (6.4) follows. Note that it holds for every CONS $\{m_k\}$.

Proof of (6.5). If α is a multi-index with $|\alpha| = k$ and the characteristic set $\{(i_1^\alpha, q_1^\alpha), \dots, (i_k^\alpha, q_k^\alpha)\}$ then $i_k^\alpha = d(\alpha)$, the order of α , and so the set J_N^n can be described as $\{\alpha \in J : |\alpha| \leq N; i_{|\alpha|}^\alpha \leq n\}$. Thus

$$\tilde{\mathbf{E}}|u(1, n, N) - u(\Delta, N)|^2 = \sum_{b=n+1}^{\infty} \sum_{k=1}^N \sum_{|\alpha|=k; i_k^\alpha=b} \frac{|\varphi_\alpha(\Delta)|^2}{\alpha!}.$$

The problem is thus to estimate $\sum_{b=n+1}^{\infty} \sum_{k=1}^N \sum_{|\alpha|=k; i_k^\alpha=b} |\varphi_\alpha(\Delta)|^2 / \alpha!$.

By Lemma A.5 the corresponding solution φ_α of (4.5) can be written as

$$\varphi_\alpha(\Delta) = \sum_{l^k} \int^{(k, \Delta)} F(\Delta; s^k; l^k) E_\alpha(s^k, l^k) ds^k, \quad (6.7)$$

where E_α is given by (6.1). Since by definition (4.6) the characteristic set of $\alpha(i_k, q_k)$

is

$\{(i_1, q_1), \dots, (i_{k-1}, q_{k-1})\}$, it is possible to write

$$E_\alpha(s^k) = \sum_{j=1}^k m_{i_k}(s_j) 1_{\{l_j=q_k\}} E_{\alpha(i_k, q_k)}(s_j^k; l_j^k),$$

where s_j^k (resp. l_j^k) denotes the same set (s_1, \dots, s_k) (resp. (l_1, \dots, l_k)) with omitted s_j (resp. l_j); for example, $s_1^k = (s_2, \dots, s_k)$.

Then (6.7) can be rewritten as

$$\varphi_\alpha(\Delta) = \sum_{j=1}^k \sum_{l_j^k} \int^{(k-1, \Delta)} \left(\int_{s_{j-1}}^{s_{j+1}} F(\Delta; s^k; l^k) m_{i_k}(s_j) 1_{\{l_j=q_k\}} ds_j \right) E_{\alpha(i_k, q_k)}(s_j^k; l_j^k) ds_j^k, \quad (6.8)$$

where $s_0 := 0$; $s_{k+1} := \Delta$. (Change the order of integration in the multiple integral.)

Denote

$$M_k(s) := \frac{\sqrt{2\Delta}}{\pi(k-1)} \sin\left(\frac{\pi(k-1)}{\Delta}s\right); \quad k > 1, \quad 0 \leq s \leq \Delta,$$

and also $F_j := \partial F(\Delta; s^k; l^k) / \partial s_j$. Then, as long as $i_k = b > 1$, integration by parts in the inner integral on the right of (6.8) yields:

$$\int_{s_{j-1}}^{s_{j+1}} F(\Delta; s^k; l^k) m_b(s_j) ds_j = F(\Delta; s^k; l^k) M_b(s_j) \Big|_{s_j=s_{j-1}}^{s_j=s_{j+1}} - \int_{s_{j-1}}^{s_{j+1}} F_j(\Delta; s^k; l^k) M_b(s_j) ds_j.$$

For each j , the remaining variables s_j^k in (6.8) are renamed as follows: $t_i := s_i, i \leq j-1$; $t_i := s_{i+1}, i > j-1$, or, symbolically, $t^{k-1} := s_j^k$. Set $t_0 := 0, t_k := \Delta$ and denote by $t^{k-1, j}, j = 1, \dots, k-1$, the set t^{k-1} in which t_j is repeated twice (e.g. $t^{k-1, 1} = (t_1, t_1, \dots, t_{k-1})$, etc.); also $t^{k-1, 0} := (t_0, t_1, t_2, \dots, t_{k-1}), t^{k-1, k} := (t_1, \dots, t_{k-1}, t_k)$.

The similar changes will also be made with the set l^k : for fixed j , there are $k-1$ free indices $l_1, \dots, l_{j-1}, l_{j+1}, \dots, l_k$ and they are renamed just like s^k to form the set l^{k-1} (in this case, the same letters are used); symbol $l^{k-1, j}$ denotes the set $(l_1, \dots, l_{j-1}, q_j, l_j, \dots, l_{k-1})$. After these transformations, $E_{\alpha(i_k, q_k)}(s_j^k; l_j^k)$ becomes $E_{\alpha(i_k, q_k)}(t^{k-1}; l^{k-1})$ - independent of j , and

$$F(\Delta; s^k; l^k) 1_{\{l_j=q_k\}} M_b(s_j) \Big|_{s_j=s_{j-1}}^{s_j=s_{j+1}} = F(\Delta; t^{k-1, j}; l^{k-1, j}) M_b(t_j) - F(\Delta; t^{k-1, j-1}; l^{k-1, j}) M_b(t_{j-1}), \quad j = 1, \dots, k.$$

Therefore, if $d(\alpha) = b > 1$ and $|\alpha| = k > 0$, then

$$\varphi_\alpha(\Delta) = \sum_{l^{k-1}} \int^{(k-1, \Delta)} \left(f_b^{(1)}(\Delta; t^{k-1}; l^{k-1}) + f_b^{(2)}(\Delta; t^{k-1}; l^{k-1}) \right) E_{\alpha(i_k, q_k)}(t^{k-1}; l^{k-1}) dt^{k-1},$$

where

$$f_b^{(1)}(\Delta; t^{k-1}; l^{k-1}) = \sum_{j=1}^k \left(F(\Delta; t^{k-1, j}; l^{k-1, j}) M_b(t_j) - F(\Delta; t^{k-1, j-1}; l^{k-1, j}) M_b(t_{j-1}) \right), \quad \text{if } k > 1,$$

$f_b^{(1)} = 0$ if $k = 1$ – because $M_b(t_0) = M_b(t_k) = 0$ (this is the only place where the choice of $\{m_k\}$ really makes the difference), and

$$\begin{aligned} f_b^{(2)}(\Delta; t^{k-1}; l^{k-1}) = & - \int_0^{t_1} F_1(\Delta; s, t^{k-1}; q_k, l^{k-1}) M_b \Delta(s) ds - \\ & \sum_{j=2}^{k-1} \int_{t_{j-1}}^{t_j} F_j(\Delta; \dots, t_{j-1}, s, t_j, \dots; l^{k-1, j}) M_b(s) ds - \\ & \int_{t_{k-1}}^{t_k} F_k(\Delta; t^{k-1}, s; l^{k-1}, q_k) M_b(s) ds. \end{aligned}$$

Then, since $|\alpha(i_{|\alpha|}, q_{|\alpha|})| = |\alpha| - 1$ and $\alpha! \geq \alpha(i_{|\alpha|}, q_{|\alpha|})!$,

$$\begin{aligned} \sum_{|\alpha|=k; i_k^\alpha=b} \frac{|\varphi_\alpha(\Delta)|^2}{\alpha!} = & \sum_{q_k=1}^r \sum_{|\alpha|=k; i_k^\alpha=b} \left| \frac{1}{\sqrt{\alpha!}} \sum_{l^{k-1}}^{(k-1, \Delta)} \int (f_b^{(1)} + f_b^{(2)}) E_{\alpha(b, q_k)} dt^{k-1} \right|^2 \leq \\ & \sum_{q_k=1}^r \sum_{|\beta|=k-1} \left| \frac{1}{\sqrt{\beta!}} \sum_{l^{k-1}} \int^{(k-1, \Delta)} (f_b^{(1)} + f_b^{(2)}) E_\beta dt^{k-1} \right|^2, \end{aligned}$$

and arguments similar to those used in the proof of Lemma A.5 show that the last expression is equal to

$$\sum_{q_k=1}^r \sum_{l^{k-1}} \int^{(k-1, \Delta)} \left| f_b^{(1)}(\Delta; t^{k-1}; l^{k-1}) + f_b^{(2)}(\Delta; t^{k-1}; l^{k-1}) \right|^2 dt^{k-1}. \quad (6.9)$$

Definition of $f_b^{(1)}$ and Lemmas A.2 and A.4 imply

$$|f_b^{(1)}|^2 = 0, \quad k = 1; \quad |f_b^{(1)}|^2 \leq \frac{kC\Delta}{(b-1)^2} |\bar{p}|^2, \quad k \geq 2. \quad (6.10)$$

Next, direct computations yield

$$\begin{aligned} F_j(\Delta; s^k; l^k) = & \Phi_{\Delta-s_k} B_{l_k} \dots \Phi_{s_{j+1}-s_j} B_{l_j} A \Phi_{s_j-s_{j-1}} \dots \Phi_{s_1} \bar{p} - \\ & \Phi_{\Delta-s_k} B_{l_k} \dots A \Phi_{s_{j+1}-s_j} B_{l_j} \Phi_{s_j-s_{j-1}} \dots \Phi_{s_1} \bar{p}, \end{aligned}$$

and then by Lemma A.4,

$$|F_j(\Delta; s^k; l^k)|^2 \leq \kappa^2 C^k |\bar{p}|^2.$$

After that the definition of $f_b^{(2)}$ and obvious inequalities

$$(a_1 + \dots + a_k)^2 \leq k(a_1^2 + \dots + a_k^2)$$

and

$$\left(\int_0^x f(y) dy \right)^2 \leq x \int_0^x (f(y))^2 dy$$

imply:

$$|f_b^{(2)}|^2 \leq \kappa^2 k C^k e^{C\Delta} |\bar{p}|^2 \Delta \int_0^\Delta (M_b(s))^2 ds \leq \kappa^2 \frac{k C^k \Delta^3}{(b-1)^2} |\bar{p}|^2;$$

so, since $\int^{(k-1, \Delta)} dt^{k-1} = \Delta^{k-1}/(k-1)!$, (6.9), (6.10), and the last inequality yield

$$\begin{aligned} \tilde{\mathbf{E}}|u(\Delta, N) - u(1, n, N)|^2 &= \sum_{b \geq n+1} \sum_{k=1}^N \sum_{|\alpha|=k; i_k^\alpha=b} \frac{|\varphi_\alpha(\Delta)|^2}{\alpha!} \leq \\ &C |\bar{p}|^2 \left[\Delta^2 \sum_{k \geq 0} \frac{k+2}{k+1} \frac{(C\Delta)^k}{k!} + \kappa^2 \Delta^3 \sum_{k \geq 0} \frac{(k+1)(C\Delta)^k}{k!} \right] \sum_{b \geq n} \frac{1}{b^2} \leq \\ &C \frac{\Delta^2 + \kappa^2 \Delta^3}{n} |\bar{p}|^2. \end{aligned}$$

This completes the proof of (6.5) and the theorem as a whole. \square

Acknowledgment. The authors would like to thank Professor B. L. Rozovskii who supervised the work. The second author is also thankful to the NSF for the financial support awarded through the Institute for Mathematics and its Applications.

References

- [1] J. F. Bennaton. Discrete Time Galerkin Approximation to the Nonlinear Filtering Solution. *Journal of Mathematical Analysis and Applications*, 110(2):364–383, 1985.
- [2] A. Budhiraja and G. Kallianpur. Approximations to the Solution of the Zakai Equations Using Multiple Wiener and Stratonovich Integral Expansions. *Stochastics and Stochastics Reports*, 56:271–315, 1996.
- [3] R. H. Cameron and W. T. Martin. The Orthogonal Development of Nonlinear Functionals in a Series of Fourier-Hermite Functions. *Annals of Mathematics*, 48(2):385–392, 1947.
- [4] J. M. C. Clark and R. J. Cameron. The Maximum Rate of Convergence of Discrete Approximation for Stochastic Differential Equations. *Springer Lecture Notes in Control and Inform. Sc.*, 25:162–171, 1980.
- [5] P. Florchinger and F. LeGland. Time Discretization of the Zakai Equation for Diffusion Processes Observed in Correlated Noise. *Stochastics and Stochastics Reports*, 35(4):233–256, 1991.
- [6] C. P. Fung. *New Numerical Algorithms for Nonlinear Filtering*. PhD thesis, University of Southern California, Los Angeles, CA, 90089, Nov. 1995.
- [7] D. Gottlieb and S. A. Orszag. *Numerical Analysis of Spectral Methods: Theory and Applications*. CBMS-NSF Regional Conference, Series in Applied Mathematics, Vol.26, 1977.

- [8] E. Hille and R. S. Phillips. *Functional Analysis and Semigroups*. Amer. Math. Soc. Colloq. Publ., Vol. XXXI, 1957.
- [9] K. Ito. Multiple Wiener integral. *Journal of the Mathematical Society of Japan*, 3:157–169, 1951.
- [10] K. Ito. Approximation of the Zakai Equation for Nonlinear Filtering. *SIAM Journal on Control and Optimization*, 34(2):620–634, 1996.
- [11] G. Kallianpur. *Stochastic Filtering Theory*. Springer, New York, 1980.
- [12] H. Kunita. Cauchy Problem for Stochastic Partial Differential Equations Arising in Nonlinear Filtering Theory. *System & Control Letters*, 1(1):37–41, 1981.
- [13] R. S. Liptser and A. N. Shiriyayev. *Statistics of Random Processes*. Springer, New York, 1992.
- [14] J. T.-H. Lo and S.-K. Ng. Optimal Orthogonal Expansion For Estimation I: Signal in White Gaussian Noise. In R. Bucy and J. Moura, editors, *Nonlinear Stochastic Problems*, pages 291–309, D. Reidel Publ. Company, 1983.
- [15] S. Lototsky, R. Mikulevicius, and B. L. Rozovskii. Nonlinear Filtering Revisited: A Spectral Approach. *SIAM Journal on Control and Optimization*, 35(2), 1997. To appear.
- [16] S. V. Lototsky. *Problems in Statistics of Stochastic Differential Equations*. PhD thesis, University of Southern California, Los Angeles, CA 90089, Aug. 1996.
- [17] R. Mikulevicius and B. L. Rozovskii. Separation of Observations and Parameters in Nonlinear Filtering. In *Proceedings of the 32nd IEEE Conference on Decision and Control, Part 2, San Antonio, 1993*, pages 1564–1569. IEEE Control Systems Society, 1993.
- [18] R. Mikulevicius and B. L. Rozovskii. Fourier-Hermite Expansion for Nonlinear Filtering. In A. A. Novikov, editor, *Festschrift in honor of A. N. Shiriyayev*, 1997.
- [19] D. Ocone. Multiple Integral Expansions for Nonlinear Filtering. *Stochastics*, 10:1–30, 1983.
- [20] B. L. Rozovskii. *Stochastic Evolution Systems*. Kluwer Academic Publishers, 1990.

Appendix

In what follows, $\{e_\gamma\}_{\gamma \in \Gamma}$ is the Hermite basis in $L_2(\mathbb{R}^d)$ and $\Gamma_\kappa = \{\gamma \in \Gamma : |\gamma| \leq \kappa\}$.

Lemma A.1 *If function f is infinitely differentiable and decays at infinity with all the derivatives faster than every power of $|x|$ then*

$$\sum_{\gamma \notin \Gamma_\kappa} |(f, e_\gamma)_0|^2 \leq \frac{C(\nu)}{\kappa^\nu}. \quad (\text{A.1})$$

In particular,

$$\sup_{0 \leq t \leq T} \sum_{\gamma \notin \Gamma_\kappa} \tilde{\mathbf{E}} |u_\gamma(t)|^2 \leq \frac{C(\nu)}{\kappa^\nu}. \quad (\text{A.2})$$

Proof. Introduce the operator $\Lambda := -\nabla^2 + 1 + |x|^2$, where ∇^2 is the Laplace operator. Direct computations show that $\Lambda e_\gamma = (2|\gamma| + d + 1)e_\gamma := \lambda_\gamma e_\gamma$. By assumption, $\Lambda^\nu f \in L_2(\mathbb{R}^d)$ for every positive integer ν and therefore

$$(f, e_\gamma)_0 = \frac{(f, \Lambda e_\gamma)_0}{\lambda_\gamma} = \frac{(\Lambda f, e_\gamma)_0}{\lambda_\gamma} = \dots = \frac{(\Lambda^\nu f, e_\gamma)_0}{\lambda_\gamma^\nu}.$$

This implies that

$$|(f, e_\gamma)_0|^2 \leq \frac{\|\Lambda^\nu f\|_0^2}{\lambda_\gamma^{2\nu}} \quad (\text{A.3})$$

and (A.1) follows because $|\Gamma_\kappa| \leq \kappa^d$. After that, (A.2) follows from (A.2) and (2.5). \square

Lemma A.2 *If assumption (A1) holds, then for every $\zeta \in \mathbb{R}^{|\Gamma_\kappa|}$*

$$\begin{aligned} \sum_{\gamma, \tau \in \Gamma_\kappa} (\mathcal{L}^* e_\tau, e_\gamma)_0 \zeta_\gamma \zeta_\tau &\leq C|\zeta|^2; \\ \sum_{\gamma \in \Gamma_\kappa} \left(\sum_{\tau \in \Gamma_\kappa} (h_l e_\tau, e_\gamma)_0 \zeta_\tau \right)^2 &\leq C|\zeta|^2. \end{aligned} \quad (\text{A.4})$$

Proof. If Π^κ is the $L_2(\mathbb{R}^d)$ -orthogonal projection on the span of $\{e_\gamma(x)\}_{\gamma \in \Gamma_\kappa}$ and $\tilde{\zeta} := \sum_{\gamma \in \Gamma_\kappa} \zeta_\gamma e_\gamma(x)$, then

$$\begin{aligned} \sum_{\gamma, \tau \in \Gamma_\kappa} (\mathcal{L}^* e_\tau, e_\gamma)_0 \zeta_\gamma \zeta_\tau &= (\mathcal{L}^* \tilde{\zeta}, \tilde{\zeta})_0; \\ \sum_{\gamma \in \Gamma_\kappa} \left(\sum_{\tau \in \Gamma_\kappa} (h_l e_\tau, e_\gamma)_0 \zeta_\tau \right)^2 &= \|\Pi^\kappa(h_l \tilde{\zeta})\|_0^2. \end{aligned}$$

Direct computations show that under assumption (A1)

$$(\mathcal{L}^* f, f)_0 \leq C\|f\|_0^2, \quad \|h_l f\|_0^2 \leq C\|f\|_0^2$$

for every sufficiently regular function f . Then estimates (A.4) follow, because

$$\|\tilde{\zeta}\|_0^2 = \sum_{\gamma \in \Gamma_\kappa} \zeta_\gamma^2 = |\zeta|^2.$$

\square

Lemma A.3 *Assumption (A1) implies the following estimates:*

$$\begin{aligned} \|\mathcal{L}^* e_\gamma\|_0 &\leq C|\gamma|; \\ \|h_l e_\gamma\|_0 &\leq C. \end{aligned}$$

Proof. The second inequality is obvious because $\|e_\gamma\|_0 = 1$. To prove the first, consider the operator $\Lambda = -\nabla^2 + 1 + |x|^2$. Then direct computations show that

$$\|\mathcal{L}^* f\|_0^2 \leq C\|(1 - \Delta)f\|_0^2 \leq C(\|\Lambda f\|_0^2 + \|f\|_0^2)$$

for every sufficiently regular function f . It follows from the proof of Lemma A.1 that $\|\Lambda e_\gamma\|_0 \leq C|\gamma|$, which implies the result. \square

Lemma A.4 *If assumption (A1) holds, then for every $\zeta \in \mathbb{R}^{|\Gamma_\kappa|}$*

$$|e^{tA^\kappa} \zeta| \leq |\zeta|e^{Ct}, \quad \sum_{l=1}^r |B_l^\kappa \zeta| \leq C|\zeta|, \quad |A^\kappa \zeta| \leq C\kappa|\zeta|.$$

(constant C does not depend on κ).

Proof. Given $\zeta \in \mathbb{R}^{|\Gamma_\kappa|}$, the vector $v(t) := e^{tA^\kappa} \zeta$ is the solution of

$$dv(t) = A^\kappa v(t), \quad v(0) = \zeta.$$

Then by the definition of matrix A^κ ,

$$d|v(t)|^2 = 2 \sum_{\gamma, \tau \in \Gamma_\kappa} (\mathcal{L}^* e_\tau, e_\gamma)_0 v_\tau(t) v_\gamma(t) dt$$

and by Lemma A.2, $d|v(t)|^2 \leq C|v(t)|^2 dt$. After that, Gronwall's inequality implies $|v(t)|^2 \leq |\zeta|^2 e^{Ct}$ and the first inequality follows. The second inequality follows directly from Lemma A.2 and the third – from Lemma A.3. \square

Lemma A.5 *If $\alpha \in J$ is a multi-index with $|\alpha| = k$ and the characteristic set $\{(i_1, q_1), \dots, (i_k, q_k)\}$, then the corresponding solution $\varphi_\alpha^\kappa(t, \bar{p}^\kappa)$ of (4.5) is given by*

$$\begin{aligned} \varphi_\alpha^\kappa(t, \bar{p}^\kappa) &= \\ & \sum_{\sigma \in \mathcal{P}^k} \sum_{l^k} \int^{(k,t)} F^\kappa(t; s^k; l^k) m_{i_{\sigma(k)}}(s_k) 1_{\{l_k=q_{\sigma(k)}\}} \cdots m_{i_{\sigma(1)}}(s_1) 1_{\{l_1=q_{\sigma(1)}\}} ds^k, \quad k > 1; \\ \varphi_\alpha^\kappa(t, \bar{p}^\kappa) &= \int_0^t \Phi_{t-s_1}^\kappa B_{q_1}^\kappa \Phi_{s_1} \bar{p}^\kappa m_{i_1}(s_1) ds_1, \quad k = 1; \\ \varphi_\alpha^\kappa(t, \bar{p}^\kappa) &= \Phi_t^\kappa \bar{p}^\kappa, \quad k = 0, \end{aligned} \tag{A.5}$$

(see beginning of Section 6 for notations).

In addition,

$$\sum_{|\alpha|=k} \frac{|\varphi_\alpha^\kappa(t, \bar{p}^\kappa)|^2}{\alpha!} = \sum_{l^k} \int^{(k,t)} |F^\kappa(t; s^k; l^k)|^2 ds^k. \tag{A.6}$$

Proof. For the sake of simplicity, the argument \bar{p} in φ_α will be omitted. Representation (A.5) is obviously true for $|\alpha| = 0$. Then the general case $|\alpha| \geq 1$ follows by induction from the variation of parameters formula.

To prove (A.6), first of all note that

$$\begin{aligned} & \sum_{\sigma \in \mathcal{P}^k} m_{i_{\sigma(k)}}(s_k) \mathbf{1}_{\{l_k=q_{\sigma(k)}\}} \cdots m_{i_{\sigma(1)}}(s_1) \mathbf{1}_{\{l_1=q_{\sigma(1)}\}} = \\ & \sum_{\sigma \in \mathcal{P}^k} m_{i_k}(s_{\sigma(k)}) \mathbf{1}_{\{l_{\sigma(k)}=q_k\}} \cdots m_{i_1}(s_{\sigma(1)}) \mathbf{1}_{\{l_{\sigma(1)}=q_1\}}. \end{aligned}$$

Indeed, every term on the left corresponding to a given $\sigma_0 \in \mathcal{P}^k$ is equal to the term on the right corresponding to $\sigma_0^{-1} \in \mathcal{P}^k$.

Then (A.5) can be written as

$$\varphi_\alpha^\kappa(t) = \sum_{l^k} \int^{(k,t)} F^\kappa(t; s^k; l^k) E_\alpha(s^k; l^k) ds^k,$$

where E_α is given by (6.1). Using the notation

$$G^\kappa(s^k; l^k) := \sum_{\sigma \in \mathcal{P}^k} \Phi_{t-s_{\sigma(k)}}^\kappa B_{l_{\sigma(k)}}^\kappa \cdots \Phi_{s_{\sigma(2)}-s_{\sigma(1)}}^\kappa B_{l_{\sigma(1)}}^\kappa \Phi_{s_{\sigma(1)}}^\kappa \bar{p}^k \mathbf{1}_{s_{\sigma(1)} < \dots < s_{\sigma(k)}},$$

it can be rewritten as

$$\varphi_\alpha^\kappa(t) = \frac{1}{k!} \sum_{l^k} \int_{[0,t]^k} G^\kappa(s^k; l^k) E_\alpha(s^k; l^k) ds^k. \quad (\text{A.7})$$

Since every component of the vector G^κ is a symmetric function from $L_2([0, t] \times \{1, \dots, r\}^k)$ and the collection $\{E_\alpha/\sqrt{\alpha!k!}, |\alpha| = k\}$ is a CONS for the symmetric part of the space,

$$G^\kappa = \sum_{|\alpha|=k} \frac{c_\alpha^\kappa E_\alpha}{\sqrt{\alpha!k!}}$$

with some $c_\alpha^\kappa \in \mathbb{R}^{|\Gamma_\kappa|}$. Then from (A.7) $|\varphi_\alpha^\kappa|^2/\alpha! = |c_\alpha^\kappa|^2/k!$ and so

$$\begin{aligned} \sum_{|\alpha|=k} \frac{|\varphi_\alpha^\kappa(t)|^2}{\alpha!} &= \frac{1}{k!} \sum_{|\alpha|=k} |c_\alpha^\kappa|^2 = \frac{1}{k!} \int_{[0,t]^k} |G^\kappa(s^k; l^k)|^2 ds^k = \\ & \frac{1}{k!} \sum_{l^k} \int_{[0,t]^k} \left| \sum_{\sigma \in \mathcal{P}^k} \Phi_{t-s_{\sigma(k)}}^\kappa B_{l_{\sigma(k)}}^\kappa \cdots \Phi_{s_{\sigma(2)}-s_{\sigma(1)}}^\kappa B_{l_{\sigma(1)}}^\kappa \Phi_{s_{\sigma(1)}}^\kappa \bar{p}^k \mathbf{1}_{s_{\sigma(1)} < \dots < s_{\sigma(k)}} \right|^2 ds^k = \\ & \sum_{l^k} \int^{(k,t)} |F^\kappa(t; s^k; l^k)|^2 ds^k, \end{aligned}$$

which proves (A.6). □

Remark. If $|\alpha| = k$, then by Theorem 3.1 in [9]

$$\xi_\alpha = \frac{1}{\sqrt{\alpha!}} \sum_{l^k} \int_0^t \int_0^{s_k} \cdots \int_0^{s_2} E_\alpha(s^k; l^k) dy(s_1) \cdots dy(s_k).$$

This gives an alternative (but equivalent) form of (4.4) in terms of multiple Wiener integrals.