A Small Time Approximation for the Solution to the Zakai Equation

Published Version:
A Small Time Approximation for the Solution to the Zakai Equation / Lanconelli A.; Scorolli R.. - In: POTENTIAL ANALYSIS. - ISSN 0926-2601. - ELETTRONICO. - 58:March(2023), pp. 561-571. [10.1007/s11118-021-09950-z]

Availability:
This version is available at: https://hdl.handle.net/11585/847098 since: 2022-01-24

Published:
DOI: http://doi.org/10.1007/s11118-021-09950-z

A small time approximation for the solution to the Zakai Equation

Alberto Lanconelli[*]   Ramiro Scorolli[†]

September 4, 2021

Abstract

We propose a novel small time approximation for the solution to the Zakai equation from nonlinear filtering theory. We prove that the unnormalized filtering density is well described over short time intervals by the solution of a deterministic partial differential equation of Kolmogorov type; the observation process appears in a pathwise manner through the degenerate component of the Kolmogorov's type operator. The rate of convergence of the approximation is of order one in the lenght of the interval. Our approach combines ideas from Wong-Zakai-type results and Wiener chaos approximations for the solution to the Zakai equation. The proof of our main theorem relies on the well-known Feynman-Kac representation for the unnormalized filtering density and careful estimates which lead to completely explicit bounds.

Key words and phrases: nonlinear filtering, Zakai equation, Feynmn-Kac formula, Wick product.

AMS 2000 classification: 60G35, 60H15, 60H07.

1 Introduction and statement of the main result

In this short note we derive a new small time approximation for the solution to the Zakai equation

$$u(t,x) = u_0(x) + \int_0^t \mathcal{L}_x^\star u(s,x)ds + \int_0^t h(x)u(s,x)dY_s, \quad t \in [0,1], x \in \mathbb{R}^d. \tag{1.1}$$

Here:

- $\mathcal{L}_x^\star$ is the formal adjoint of $\mathcal{L}_x$, generator of the d-dimensional *signal* process $\{X_t\}_{t\in[0,1]}$ which is assumed to solve the stochastic differential equation

$$X_t = X_0 + \int_0^t b(X_s)ds + \int_0^t \sigma(X_s)dB_s, \quad t \in [0,1]; \tag{1.2}$$

  the process $\{B_t\}_{t\in[0,1]}$ is a standard d-dimensional Brownian motion defined on the complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$;

[*]Dipartimento di Scienze Statistiche Paolo Fortunati, Università di Bologna, Bologna, Italy. **e-mail**: alberto.lanconelli2@unibo.it

[†]Dipartimento di Scienze Statistiche Paolo Fortunati, Università di Bologna, Bologna, Italy. **e-mail**: ramiro.scorolli2@unibo.it

- $\{Y_t\}_{t\in[0,1]}$ is the one dimensional *observation* process described by

$$Y_t = y_0 + \int_0^t h(X_s)ds + W_t, \quad t \in [0,1], \tag{1.3}$$

with $\{W_t\}_{t\in[0,1]}$ being a standard one-dimensional Brownian motion defined on $(\Omega, \mathcal{F}, \mathbb{P})$ and independent of $\{B_t\}_{t\geq 0}$.

The solution $\{u(t,x)\}_{t\in[0,1],x\in\mathbb{R}^d}$ to the Zakai equation (1.1), usually called *unnormalized filtering density*, plays a crucial role in the nonlinear filtering problem since it identifies uniquely the conditional distribution of X_t given $\mathcal{F}_t^Y := \sigma(Y_s, 0 \leq s \leq t)$. The reader is referred to the original paper [25] and the references quoted there; for an exhaustive treatment of the subject we suggest the excellent review [11], as well as the books [15] and [19].

Existence, uniqueness and regularity properties for the solution to (1.1) can be found for instance, under different sets of assumptions and solution concepts, in the classic works [5],[17],[18],[24] and the more recent paper [3]. We also mention a useful Feynman-Kac representation for the solution $\{u(t,x)\}_{t\in[0,1],x\in\mathbb{R}^d}$ obtained in [18] and, in a slightly different form, in [3]. This representation will play a crucial role in our investigation.

From the applications point of view, closed form expressions for the solution to the Zakai equation are certainly desirable; however, as pointed in [1] only few particular cases of (1.1) allow for explicit computations. The important issue of deriving simple approximation schemes for the solution to (1.1) have been considered in [2] and [9] which employ splitting up methods and time discretization, respectively; Wong-Zakai-type results were investigated in [6] and [13] while [5] and [20] proposed a Wiener chaos approach. We also mention the so called *pathwise filtering* that steams from the problem of having a robust, with respect to the observation process, filter; this has been discussed in [7] and [8].

The approach proposed in the current paper combines ideas from the Wong-Zakai approximation proposed in [13], where the signal process is smoothed through a polygonal approximation, and the Wiener chaos approach presented in [5] and [20], where one relates equation (1.1) to a system of nested deterministic partial differential equations solved by the kernels of the Cameron-Martin decomposition of the solution $\{u(t,x)\}_{t\in[0,1],x\in\mathbb{R}^d}$. We refer the reader to Remark 1.4 below for the heuristic idea supporting our analysis and its link to the aforementioned approaches.

The main novelty of our result is the connection between equation (1.1) and a deterministic partial differential equation of Kolmogorov type (see e.g. [23]), where the observation process enters as a degenerate component of the second order differential operator $\mathcal{L}_x^\star$. We prove that the solution $\{u(t,x)\}_{t\in[0,1],x\in\mathbb{R}^d}$ to the Zakai equation (1.1) can be approximated over small intervals of time by the solution of the aforementioned degenerate partial differential equation, with the observation process having a pathwise role. This approximation has the same rate of convergence of one obtained in [20] and is described by completely explicit constants.

To be more specific, we now introduce some notation and state our main result. In the sequel the following regularity conditions will be in force.

Assumption 1.1.

1. *For $1 \leq i,j \leq d$, the functions $b_i : \mathbb{R}^d \to \mathbb{R}^d$ and $a_{ij} : \mathbb{R}^d \to \mathbb{R}^d$, where*

$$a_{ij}(x) := \sum_{k=1}^d \sigma_{ik}(x)\sigma_{jk}(x), \quad x \in \mathbb{R}^d, \tag{1.4}$$

are bounded with bounded partial derivatives up to the third order. Moreover, the matrix $\{a_{ij}(x)\}_{1\leq i,j\leq d}$ is uniformly elliptic, i.e. there exists two positive constants $\mu_1 < \mu_2$ such that

$$\mu_1|z|^2 \leq \sum_{i,j=1}^d a_{ij}(x)z_iz_j \leq \mu_2|z|^2, \quad \textit{for all } z \in \mathbb{R}^d,$$

with $|z|^2 := z_1^2 + \cdots + z_d^2$.

2. *The initial data X_0 in* (1.2) *is random, independent of $\{B_t\}_{t\in[0,1]}$ and its distribution is absolutely continuous with respect to the d-dimensional Lebesgue measure; its density $u_0 : \mathbb{R}^d \to \mathbb{R}$ is bounded and acts as initial data in (1.1).*

3. *The function $h : \mathbb{R}^d \to \mathbb{R}$ is bounded and globally Lipschitz continuous.*

Remark 1.2. *We observe that, according to Assumption 1.1, there exists a positive constant L such that*

$$|h(x_1) - h(x_2)| \le L|x_1 - x_2|, \qquad \textit{for all } x_1, x_2 \in \mathbb{R}^d. \tag{1.5}$$

Moreover, there exists a positive constant M such that

$$\max\{|a(x)|^2, |b^\star(x)|\} \le M, \qquad \textit{for all } x \in \mathbb{R}^d, \tag{1.6}$$

where $b_i^\star(x) := \sum_{j=1}^d \partial_{x_j} a_{ij}(x) - b_i(x)$, $i = 1, ..., d$,. We will need these two constants in the statement of our main theorem.

According to the Girsanov theorem and thanks to the assumption of boundedness on h, the prescription

$$\mathbb{P}_1(A) := \int_A e^{-\int_0^1 h(X_s(\omega))dW_s(\omega) - \frac{1}{2}\int_0^1 h(X_s(\omega))^2 ds} d\mathbb{P}(\omega), \quad A \in \mathcal{F},$$

defines a probability measure on $(\Omega, \mathcal{F})$; moreover, the stochastic process $\{Y_t - y_0\}_{t\in[0,1]}$ in (1.3) becomes on the probability space $(\Omega, \mathcal{F}, \mathbb{P}_1)$ a one dimensional Brownian motion independent of $\{B_t\}_{t\ge 0}$. In the sequel we will write $\mathbb{E}_1$ to denote the expectation under the probability measure $\mathbb{P}_1$.

We are now ready to state our main result.

Theorem 1.3. *Let Assumption 1.1 be in force and, for $0 < T < 1$, let*

$$[0, T] \times \mathbb{R}^d \times \mathbb{R} \ni (t, x, y) \mapsto v(t, x, y)$$

be a classical solution of the Cauchy problem

$$\begin{cases} \partial_t v(t, x, y) = \mathcal{L}_x^\star v(t, x, y) - h(x)\partial_y v(t, x, y), & (t, x, y) \in]0, T] \times \mathbb{R}^d \times \mathbb{R}; \\ v(0, x, y) = u_0(x) e^{-\frac{y^2}{2T}}, & (x, y) \in \mathbb{R}^d \times \mathbb{R}. \end{cases} \tag{1.7}$$

Then, for any $q \ge 1$ and $K > 0$, we have

$$\sup_{|x|\le K} \mathbb{E}_1 \left[\left| u(T, x) - e^{\frac{(Y_T - y_0)^2}{2T}} v(T, x, Y_T - y_0) \right|^q \right]^{1/q} \le \mathcal{C}T, \tag{1.8}$$

with

$$\mathcal{C} := \frac{2}{\sqrt{3}} |u_0|_\infty e^{T\left(|c|_\infty + \frac{q_1 - 1}{2}|h|_\infty^2 + \sqrt{M} + M/2\right)} \left(\kappa(q_2) + \sqrt{T}|h|_\infty\right) L\sqrt{2(1 + K^2)(1 + T)}. \tag{1.9}$$

Here L and M are defined in (1.5) *and* (1.6)*, respectively; the constants $q_1, q_2 \ge 1$ verify the identity $\frac{1}{q_1} + \frac{1}{q_2} = \frac{1}{q}$; $\kappa(q_2)$ is given by $\sqrt{2}\left(\Gamma(\frac{q_2+1}{2})/\sqrt{\pi}\right)^{1/q_2}$; $|u_0|_\infty$ and $|h|_\infty$ denotes the $L^\infty(\mathbb{R}^d)$-norms of u_0 and h, respectively.*

Remark 1.4. *The heuristic idea that links equation* (1.1) *to equation* (1.7) *is as follows. Write* (1.1) *in the differential form*

$$\partial_t u(t, x) = \mathcal{L}_x^\star u(t, x) + h(x)u(t, x) \diamond \frac{dY_t}{dt}, \quad u(0, x) = u_0(x), \tag{1.10}$$

where $\diamond$ denotes the Wick product associated to the Brownian motion $\{Y_t - y_0\}_{t\in[0,1]}$ on the probability space $(\Omega, \mathcal{F}, \mathbb{P}_1)$. The use of the Wick product is dictated by the Itô's interpretation of (1.1) *(see [12] and [14] for a discussion on this issue and detailed analysis of the Wick product). If equation* (1.10) *is considered on a small time interval $[0,T]$, one may replace $\frac{dY_t}{dt}$ with $\frac{Y_T - y_0}{T}$ (this amounts at considering a Wong-Zakai approximation with the rudest possible partition of the interval $[0,T]$); this gives*

$$\partial_t u(t,x) = \mathcal{L}_x^{\star} u(t,x) + \frac{h(x)}{T} u(t,x) \diamond (Y_T - y_0), \quad u(0,x) = u_0(x). \tag{1.11}$$

In general, the Wick-multiplication between a random variable X and an element from the first order Wiener chaos, say $I(f)$, can be rewritten as

$$X \diamond I(f) = X \cdot I(f) - D_f X,$$

where $D_f X$ stands for the directional Malliavin derivative of X, in the direction $\int_0^{\cdot} f(s)ds$ (see [22]). Since, $Y_T - y_0 = \int_0^1 \mathbf{1}_{[0,T]}(s)dY_s$ is an element in the first Wiener chaos associated with the Brownian motion $\{Y_t - y_0\}_{t\in[0,1]}$ and probability space $(\Omega, \mathcal{F}, \mathbb{P}_1)$, we can transform equation (1.11) *into*

$$\partial_t u(t,x) = \mathcal{L}_x^{\star} u(t,x) + \frac{h(x)}{T} u(t,x)(Y_T - y_0) - \frac{h(x)}{T} D_{\mathbf{1}_{[0,T]}} u(t,x). \tag{1.12}$$

We now search for a solution $u(t,x)$ to equation (1.12) *of the form*

$$u(t,x,\omega) = \tilde{u}(t,x,Y_T(\omega) - y_0), \tag{1.13}$$

for some $\tilde{u} : [0,T] \times \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}$ to be determined. A substitution of (1.13) *in* (1.12) *yields, together with the chain rule for the Malliavin derivative,*

$$\begin{aligned}\partial_t \tilde{u}(t,x,Y_T - y_0) =& \mathcal{L}_x^{\star} \tilde{u}(t,x,Y_T - y_0) + \frac{h(x)}{T} \tilde{u}(t,x,Y_T - y_0)(Y_T - y_0) \\ & - h(x)\partial_y \tilde{u}(t,x,Y_T - y_0);\end{aligned}$$

note that here the term $Y_T - y_0$ can be tackled at a path-wise level. Equation (1.7) *is now obtained via the simple transformation*

$$v(t,x,y) := \tilde{u}(t,x,y) e^{-\frac{y^2}{2T}}, \quad t \in [0,T], x \in \mathbb{R}^d, y \in \mathbb{R}.$$

It is not difficult to see, using Theorem 4.12 in [14] and the Feynman-Kac representation for $\{u(t,x)\}_{t\in[0,1],x\in\mathbb{R}^d}$ in [3], that we also have

$$\mathbb{E}_1[u(T,x)|Y_T - y_0] = e^{\frac{(Y_T - y_0)^2}{2T}} v(T,x,Y_T - y_0);$$

this spots the analogy between our approach and the one in [20] where projections of $u(T,x)$ on suitable families of elements from the Wiener chaos were utilized to propose approximation schemes for the solution to (1.1).

Remark 1.5. *The existence of a classical solution for the Cauchy problem* (1.7) *is actually not needed for the validity of Theorem 1.3 (the statement is presented this way for easiness of exposition). In fact, in the proof of our main result we deal with the Feynman-Kac representation for the solution to* (1.7) *(see formula* (2.2) *below) without using its differentiability properties with respect to t and x. The right hand side of* (2.2) *is well defined under mild conditions on the coefficients of equation* (1.7) *(largely covered by Assumption 1.1) and this makes our proof consistent. It is worth mentioning that the right hand side of* (2.2) *becomes a classical solution if suitable regularity assumptions on the coefficients of* (2.1) *are in force. For more details on this issue we refer the reader to [4] and [10] page 122.*

2 Proof of Theorem 1.3

We start with some notation. The generator $\mathcal{L}_x$ of the signal process $\{X_t\}_{t\in[0,1]}$ in (1.2) is

$$\mathcal{L}_x f(x) = \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}(x)\partial^2_{x_i x_j} f(x) + \sum_{i=1}^{d} b_i(x)\partial_{x_i} f(x),$$

where the $a_{ij}(x)$'s are defined in (1.4). The adjoint operator $\mathcal{L}^\star_x$ is given by

$$\mathcal{L}^\star_x f(x) = \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}(x)\partial^2_{x_i x_j} f(x) + \sum_{i=1}^{d} b^\star_i(x)\partial_{x_i} f(x) + c(x)f(x),$$

with

$$b^\star_i(x) := \sum_{j=1}^{d} \partial_{x_j} a_{ij}(x) - b_i(x), \quad i = 1, ..., d,$$

(see Remark 1.2) and

$$c(x) := \sum_{i=1}^{d}\left(\frac{1}{2}\sum_{j,k=1}^{d} \partial^2_{x_j x_k} a_{ij}(x) - \partial_{x_k} b_i(x)\right).$$

It is convenient to split the operator $\mathcal{L}^\star_x$ as

$$\mathcal{L}^\star_x f(x) = \mathtt{L}^\star_x f(x) + c(x)f(x)$$

where we set

$$\mathtt{L}^\star_x f(x) := \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}(x)\partial^2_{x_i x_j} f(x) + \sum_{i=1}^{d} b^\star_i(x)\partial_{x_i} f(x).$$

With this notation at hand, the Cauchy problem (1.7) takes the form

$$\begin{cases} \partial_t v(t,x,y) = \mathtt{L}^\star_x v(t,x,y) + c(x)v(t,x,y) - h(x)\partial_y v(t,x,y) \\ \quad (t,x,y) \in]0,T] \times \mathbb{R}^d \times \mathbb{R}; \\ v(0,x,y) = u_0(x)e^{-\frac{y^2}{2T}}, \quad (x,y) \in \mathbb{R}^d \times \mathbb{R}. \end{cases} \tag{2.1}$$

Now, assume

$$[0,T] \times \mathbb{R}^d \times \mathbb{R} \ni (t,x,y) \mapsto v(t,x,y)$$

to be a classical solution of (2.1). According to the Feynman-Kac formula (see, for instance, Theorem 1.1, page 120 in [10]) we can write

$$\begin{aligned} v(T,x,y) &= \hat{\mathbb{E}}\left[u_0(\hat{\xi}^x_T)e^{-\frac{\left(y-\int_0^T h(\hat{\xi}^x_s)ds\right)^2}{2T}} e^{\int_0^T c(\hat{\xi}^x_s)ds}\right] \\ &= e^{-\frac{y^2}{2T}}\hat{\mathbb{E}}\left[u_0(\hat{\xi}^x_T)e^{\int_0^T c(\hat{\xi}^x_s)ds} e^{\frac{y\int_0^T h(\hat{\xi}^x_s)ds}{T} - \frac{\left(\int_0^T h(\hat{\xi}^x_s)ds\right)^2}{2T}}\right], \end{aligned} \tag{2.2}$$

where $\{\hat{\xi}^x_s\}_{s\in[0,1]}$ solves the SDE

$$d\hat{\xi}^x_s = b^\star(\hat{\xi}^x_s) + \sigma(\hat{\xi}^x_s)d\hat{B}_s, \quad \hat{\xi}^x_0 = x, \tag{2.3}$$

on the auxiliary probability space $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{P}})$ with d-dimensional Brownian motion $\{\hat{B}_s\}_{s\in[0,1]}$. This gives

$$e^{\frac{(Y_T-y_0)^2}{2T}}v(T,x,Y_T-y_0)=\hat{\mathbb{E}}\left[u_0(\hat{\xi}^x_T)e^{\int_0^T c(\hat{\xi}^x_s)ds}e^{\frac{(Y_T-y_0)\int_0^T h(\hat{\xi}^x_s)ds}{T}-\frac{\left(\int_0^T h(\hat{\xi}^x_s)ds\right)^2}{2T}}\right]. \tag{2.4}$$

It is well known that the solution $u(t,x)$ to the Zakai equation (1.1) also possesses a Feynman-Kac representation: see formula (1.4) page 132 in [18]. Here, we use instead an equivalent formulation due to [3] (see formula (2.9) there), namely

$$u(T,x)=\hat{\mathbb{E}}\left[u_0(\hat{\xi}^x_T)e^{\int_0^T c(\hat{\xi}^x_s)ds}e^{\int_0^T h(\hat{\xi}^x_{T-s})dY_s-\frac{1}{2}\int_0^T h^2(\hat{\xi}^x_s)ds}\right], \tag{2.5}$$

where $\{\hat{\xi}^x_s\}_{s\in[0,1]}$ is defined in (2.3). A comparison between (2.4) and (2.5) gives

$$\begin{aligned}
&u(T,x)-e^{\frac{(Y_T-y_0)^2}{2T}}v(T,x,Y_T-y_0)\\
&\quad=\hat{\mathbb{E}}\left[u_0(\hat{\xi}^x_T)e^{\int_0^T c(\hat{\xi}^x_s)ds}e^{\int_0^T h(\hat{\xi}^x_{T-s})dY_s-\frac{1}{2}\int_0^T h^2(\hat{\xi}^x_s)ds}\right]\\
&\quad\quad-\hat{\mathbb{E}}\left[u_0(\hat{\xi}^x_T)e^{\int_0^T c(\hat{\xi}^x_s)ds}e^{\frac{(Y_T-y_0)\int_0^T h(\hat{\xi}^x_s)ds}{T}-\frac{\left(\int_0^T h(\hat{\xi}^x_s)ds\right)^2}{2T}}\right]\\
&\quad=\hat{\mathbb{E}}\left[u_0(\hat{\xi}^x_T)e^{\int_0^T c(\hat{\xi}^x_s)ds}\left(e^{\int_0^T h(\hat{\xi}^x_{T-s})dY_s-\frac{1}{2}\int_0^T h^2(\hat{\xi}^x_s)ds}-e^{\frac{(Y_T-y_0)\int_0^T h(\hat{\xi}^x_s)ds}{T}-\frac{\left(\int_0^T h(\hat{\xi}^x_s)ds\right)^2}{2T}}\right)\right],
\end{aligned}$$

and hence

$$\begin{aligned}
&\left|u(T,x)-e^{\frac{(Y_T-y_0)^2}{2T}}v(T,x,Y_T-y_0)\right|\\
&\quad\le\hat{\mathbb{E}}\left[|u_0(\hat{\xi}^x_T)|e^{\int_0^T c(\hat{\xi}^x_s)ds}\left|e^{\int_0^T h(\hat{\xi}^x_{T-s})dY_s-\frac{1}{2}\int_0^T h^2(\hat{\xi}^x_s)ds}-e^{\frac{(Y_T-y_0)\int_0^T h(\hat{\xi}^x_s)ds}{T}-\frac{\left(\int_0^T h(\hat{\xi}^x_s)ds\right)^2}{2T}}\right|\right]\\
&\quad\le|u_0|_\infty e^{T|c|_\infty}\hat{\mathbb{E}}\left[\left|e^{\int_0^T h(\hat{\xi}^x_{T-s})dY_s-\frac{1}{2}\int_0^T h^2(\hat{\xi}^x_s)ds}-e^{\frac{(Y_T-y_0)\int_0^T h(\hat{\xi}^x_s)ds}{T}-\frac{\left(\int_0^T h(\hat{\xi}^x_s)ds\right)^2}{2T}}\right|\right].
\end{aligned}$$

We now take $q\geq 1$ and compute the $L^q(\mathbb{P}_1)$-norm of the first and last members above; an application of Minkowsky's inequality gives

$$\begin{aligned}
&\left\|u(T,x)-e^{\frac{(Y_T-y_0)^2}{2T}}v(T,x,Y_T-y_0)\right\|_q\\
&\quad\le|u_0|_\infty e^{T|c|_\infty}\hat{\mathbb{E}}\left[\left\|e^{\int_0^T h(\hat{\xi}^x_{T-s})dY_s-\frac{1}{2}\int_0^T h^2(\hat{\xi}^x_s)ds}-e^{\frac{(Y_T-y_0)\int_0^T h(\hat{\xi}^x_s)ds}{T}-\frac{\left(\int_0^T h(\hat{\xi}^x_s)ds\right)^2}{2T}}\right\|_q\right].
\end{aligned} \tag{2.6}$$

We need the following result.

Lemma 2.1. *Let $f,g:[0,T]\to\mathbb{R}$ be bounded measurable deterministic functions. Then, for any $q\geq 1$ we have*

$$\begin{aligned}
&\mathbb{E}_1\left[\left|e^{\int_0^T f(s)dY_s-\frac{1}{2}|f|_2^2}-e^{\int_0^T g(s)dY_s-\frac{1}{2}|g|_2^2}\right|^q\right]^{\frac{1}{q}}\\
&\quad\le\left(e^{\frac{q_1-1}{2}T|f|_\infty^2}+e^{\frac{q_1-1}{2}T|g|_\infty^2}\right)\left(\kappa(q_2)+\frac{\sqrt{T}}{2}(|f|_\infty+|g|_\infty)\right)|f-g|_2,
\end{aligned}$$

where $q_1,q_2\geq 1$ satisfy $1/q_1+1/q_2=1/q$ while $\kappa(q_2):=\sqrt{2}\left(\Gamma(\frac{q_2+1}{2})/\sqrt{\pi}\right)^{1/q_2}$. Moreover, $|l|_2$ and $|l|_\infty$ stand for the norms in $L^2([0,T])$ and $L^\infty([0,T])$ of l, respectively.

Proof. By means of the elementary inequality $|e^a - e^b| \le (e^a + e^b)|a-b|$ we can write

$$
\begin{aligned}
&\left| e^{\int_0^T f(s)dY_s - \frac{1}{2}|f|_2^2} - e^{\int_0^T g(s)dY_s - \frac{1}{2}|g|_2^2} \right| \\
&\quad \le \left(e^{\int_0^T f(s)dY_s - \frac{1}{2}|f|_2^2} + e^{\int_0^T g(s)dY_s - \frac{1}{2}|g|_2^2} \right) \\
&\qquad \times \left| \int_0^T [f(s) - g(s)]dY_s - \frac{1}{2}\left(|f|_2^2 - |g|_2^2\right) \right| \\
&\quad \le \left(e^{\int_0^T f(s)dY_s - \frac{1}{2}|f|_2^2} + e^{\int_0^T g(s)dY_s - \frac{1}{2}|g|_2^2} \right) \\
&\qquad \times \left(\left| \int_0^T [f(s) - g(s)]dY_s \right| + \frac{1}{2}\left||f|_2^2 - |g|_2^2\right| \right).
\end{aligned}
$$

Now, for $q \ge 1$ we take the $L^q(\mathbb{P}_1)$-norm of the first and last members above and apply Hölder's inequality with exponents $q_1, q_2 \ge 1$ satisfying $1/q_1 + 1/q_2 = 1/q$. This gives

$$
\begin{aligned}
&\left\| e^{\int_0^T f(s)dY_s - \frac{1}{2}|f|_2^2} - e^{\int_0^T g(s)dY_s - \frac{1}{2}|g|_2^2} \right\|_q \\
&\quad \le \left\| e^{\int_0^T f(s)dY_s - \frac{1}{2}|f|_2^2} + e^{\int_0^T g(s)dY_s - \frac{1}{2}|g|_2^2} \right\|_{q_1} \\
&\qquad \times \left(\left\| \int_0^T [f(s) - g(s)]dY_s \right\|_{q_2} + \frac{1}{2}\left||f|_2^2 - |g|_2^2\right| \right).
\end{aligned} \tag{2.7}
$$

Under the measure $\mathbb{P}_1$, the random variables $\int_0^T f(s)dY_s$ and $\int_0^T g(s)dY_s$ are Gaussian with mean zero and variances $|f|_2^2$ and $|g|_2^2$, respectively. Hence,

$$
\begin{aligned}
&\left\| e^{\int_0^T f(s)dY_s - \frac{1}{2}|f|_2^2} + e^{\int_0^T g(s)dY_s - \frac{1}{2}|g|_2^2} \right\|_{q_1} \\
&\quad \le \left\| e^{\int_0^T f(s)dY_s - \frac{1}{2}|f|_2^2} \right\|_{q_1} + \left\| e^{\int_0^T g(s)dY_s - \frac{1}{2}|g|_2^2} \right\|_{q_1} \\
&\quad = e^{\frac{q_1-1}{2}|f|_2^2} + e^{\frac{q_1-1}{2}|g|_2^2} \\
&\quad \le e^{\frac{q_1-1}{2}T|f|_\infty^2} + e^{\frac{q_1-1}{2}T|g|_\infty^2}.
\end{aligned} \tag{2.8}
$$

Moreover, using once more the normality, under the measure $\mathbb{P}_1$, of the random variable $\int_0^T [f(s) - g(s)]dY_s$ we get

$$
\left\| \int_0^T [f(s) - g(s)]dY_s \right\|_{q_2} = \kappa(q_2)|f-g|_2, \tag{2.9}
$$

where $\kappa(q_2) := \sqrt{2}\left(\Gamma(\frac{q_2+1}{2})/\sqrt{\pi}\right)^{1/q_2}$ (see, for instance, Formula (1.1) in [14]). Furthermore,

$$
\begin{aligned}
\left||f|_2^2 - |g|_2^2\right| &= (|f|_2 + |g|_2)\,||f|_2 - |g|_2| \\
&\le (|f|_2 + |g|_2)\,|f-g|_2 \\
&\le \sqrt{T}(|f|_\infty + |g|_\infty)|f-g|_2.
\end{aligned} \tag{2.10}
$$

Therefore, combining (2.7) with (2.8), (2.9) and (2.10) we get

$$
\begin{aligned}
&\left\| e^{\int_0^T f(s)dY_s - \frac{1}{2}|f|_2^2} - e^{\int_0^T g(s)dY_s - \frac{1}{2}|g|_2^2} \right\|_q \\
&\le \left(e^{\frac{q_1-1}{2}T|f|_\infty^2} + e^{\frac{q_1-1}{2}T|g|_\infty^2} \right) \left(\kappa(q_2) + \frac{\sqrt{T}}{2}(|f|_\infty + |g|_\infty) \right) |f-g|_2.
\end{aligned}
$$

The proof is complete. $\square$

Thanks to the identities

$$\frac{(Y_T - y_0)\int_0^T h(\hat{\xi}_s^x)ds}{T} = \int_0^T \left(\frac{1}{T}\int_0^T h(\hat{\xi}_r^x)dr\right) dY_s,$$

and

$$\int_0^T \left(\frac{1}{T}\int_0^T h(\hat{\xi}_r^x)dr\right)^2 ds = \frac{\left(\int_0^T h(\hat{\xi}_r^x)dr\right)^2}{T},$$

we are in a position to apply Lemma 2.1 to the last term in (2.6) with

$$f(s) := h(\hat{\xi}_{T-s}^x) \quad \text{and} \quad g(s) := \frac{1}{T}\int_0^T h(\hat{\xi}_r^x)dr;$$

note that such choices imply $|f|_\infty \le |h|_\infty$ and $|g|_\infty \le |h|_\infty$ (here, the norms are on the corresponding domains). Therefore,

$$\begin{aligned}
&\left\| u(T,x) - e^{\frac{(Y_T-y_0)^2}{2T}} v(T,x,Y_T-y_0)\right\|_q \\
&\quad \le 2|u_0|_\infty e^{T\left(|c|_\infty + \frac{q_1-1}{2}|h|_\infty^2\right)} \left(\kappa(q_2) + \sqrt{T}|h|_\infty\right) \\
&\quad\quad \times \hat{\mathbb{E}}\left[\left(\int_0^T \left| h(\hat{\xi}_{T-s}^x) - \frac{1}{T}\int_0^T h(\hat{\xi}_r^x)dr\right|^2 ds\right)^{1/2}\right].
\end{aligned}$$

We now focus on the last expectation; using a combination of Jensen's inequalities and Tonelli's theorem we get

$$\begin{aligned}
&\hat{\mathbb{E}}\left[\left(\int_0^T \left| h(\hat{\xi}_{T-s}^x) - \frac{1}{T}\int_0^T h(\hat{\xi}_r^x)dr\right|^2 ds\right)^{1/2}\right] \\
&\quad \le \left(\hat{\mathbb{E}}\left[\int_0^T \left| h(\hat{\xi}_{T-s}^x) - \frac{1}{T}\int_0^T h(\hat{\xi}_r^x)dr\right|^2 ds\right]\right)^{1/2} \\
&\quad = \left(\int_0^T \hat{\mathbb{E}}\left[\left| h(\hat{\xi}_{T-s}^x) - \frac{1}{T}\int_0^T h(\hat{\xi}_r^x)dr\right|^2\right] ds\right)^{1/2} \\
&\quad = \left(\int_0^T \hat{\mathbb{E}}\left[\left| \frac{1}{T}\int_0^T (h(\hat{\xi}_{T-s}^x) - h(\hat{\xi}_r^x))dr\right|^2\right] ds\right)^{1/2} \\
&\quad \le \left(\int_0^T \hat{\mathbb{E}}\left[\frac{1}{T}\int_0^T |h(\hat{\xi}_{T-s}^x) - h(\hat{\xi}_r^x)|^2 dr\right] ds\right)^{1/2} \\
&\quad = \left(\int_0^T \left(\frac{1}{T}\int_0^T \hat{\mathbb{E}}\left[|h(\hat{\xi}_{T-s}^x) - h(\hat{\xi}_r^x)|^2\right] dr\right) ds\right)^{1/2}.
\end{aligned}$$

The Lipschitz continuity of h and Theorem 4.3, Chapter 2 in [21] yield

$$\hat{\mathbb{E}}\left[|h(\hat{\xi}_{T-s}^x) - h(\hat{\xi}_r^x)|^2\right] \le L^2 \hat{\mathbb{E}}\left[|\hat{\xi}_{T-s}^x - \hat{\xi}_r^x|^2\right]$$

$$\leq 2L^2(1+|x|^2)(1+T)e^{2(\sqrt{M}+M/2)T}|T-s-r|,$$

where L and M come from (1.5) and (1.6). Moreover,

$$\left(\int_0^T \left(\frac{1}{T}\int_0^T |T-s-r|dr\right) ds\right)^{1/2} = \frac{T}{\sqrt{3}}.$$

Combining all our estimates we obtain

$$\begin{aligned}
&\left\| u(T,x) - e^{\frac{(Y_T-y_0)^2}{2T}} v(T,x,Y_T-y_0)\right\|_q \\
&\leq \frac{2}{\sqrt{3}}|u_0|_\infty e^{T\left(|c|_\infty + \frac{q_1-1}{2}|h|_\infty^2 + \sqrt{M}+M/2\right)}\left(\kappa(q_2)+\sqrt{T}|h|_\infty\right) \\
&\quad \times L\sqrt{2(1+|x|^2)(1+T)}T,
\end{aligned}$$

as desired.

Data availability statement: Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

References

[1] V. E. Benes, Exact finite-dimensional filters for certain diffusions with nonlinear drift, *Stochastics* **5** 1-2 (1981) 65-92

[2] A. Bensoussan, R. Glowinski, and R. Rascanu, Approximations of the Zakai equation by splitting up method, *SIAM J. Control Optim.* **28** (1990) 1420-1431.

[3] F.E. Benth, T. Deck, J. Potthoff and G. Vage, Explicit strong solutions of SPDE's with applications to non-linear filtering, *Acta Appl. Math.* **51** (1998) 215–242.

[4] Y. N. Blagovescenskii and M. I. Freidlin, Certain properties of diffusion processes depending on a parameter, *Doklady Akad. Nauk SSSR* **138** (1962) 508-511. Translation in *Soviet Math.* **2** 633-636.

[5] A. Budhiraja and G. Kallianpur, Approximations to the Solution of the Zakai Equations Using Multiple Wiener and Stratonovich Integral Expansions, *Stochastics and Stochastics Rep.* **56** (1996) 271–315.

[6] M. Chaleyat-Maurel and D. Michel, A Stroock Varadhan support theorem in non-linear filtering theory, *Probab. Theory Related Fields* **84** (1990) 119–139.

[7] J. M. C. Clark, The design of robust approximations to the stochastic differential equations of nonlinear filtering, in *Communication Systems and Random Process Theory, J. K. Skwirzynski, Ed.*, Sijthoff-Noordhoff, 1978.

[8] M. H. A. Davis, A pathwise solution of the equations of nonlinear filtering, *Teor. Veroyatnost. Primenen.* **27** 1 (1982) 160–167.

[9] G. B. Di Masi , M. Pratelli and W. J. Runggaldier, An approximation for the nonlinear filtering problem, with error bound, *Stochastics* **14** 4 (1985) 247-271.

[10] M. I. Freidlin, *Functional Integration and Partial Differential Equations*, Princeton University Press, Princeton, 1985.

[11] A. J. Heunis, On the stochastic differential equations of filtering theory, *Appl. Math. Comput.* **39** 3 (1990) 3s–36s.

[12] H. Holden, B. Øksendal, J. Ubøe and T.-S. Zhang, *Stochastic Partial Differential Equations - II Edition*, Springer, New York, 2010.

[13] Y. Hu, G. Kallianpur and J. Xiong, An approximation for Zakai equation, *Applied Math. Optimiz.* **45** (2002) 23–44.

[14] S. Janson, *Gaussian Hilbert spaces*, Cambridge Tracts in Mathematics, 129. Cambridge University Press, Cambridge, 1997.

[15] G. Kallianpur, *Stochastic Filtering Theory*, Springer-Verlag, New York, 1980.

[16] I. Karatzas and S. E. Shreve, *Brownian motion and stochastic calculus*, Springer-Verlag, New York, 1991.

[17] N. V. Krylov and B. L. Rozovskii, On conditional distributions of diffusion processes, *Math. USSR-Izv.* **42** (1978) 336-356.

[18] H. Kunita, Stochastic partial differential equations connected with nonlinear filtering, in *Nonlinear Filtering and Stochastic Control*, Lecture Notes in Mathematics **972** (1982) Springer, Berlin, 100-169.

[19] R. S. Lipster and A. N. Shiryayev, *Statistics of Random Processes*, Springer, New York, 1992.

[20] S. Lototsky, R. Mikulevicius and B. Rozovskii, Nonlinear filtering revisited: a spectral approach, *SIAM J. Control Optim.* **35** 2 (1997) 435–461.

[21] X. Mao, *Stochastic Differential Equations and Applications*, Second edition, Horwood, Chichester, UK, 2008.

[22] D. Nualart, *Malliavin calculus and Related Topics - II Edition*, Springer, New York, 2006.

[23] O. A. Oleinik and E. V. Radkevich, *Second Order Equation with Nonnegative Characteristic Form*, Plennum Press, New York, USA, 1973.

[24] E. Pardoux, Stochastic partial differential equations and filtering of diffusion processes, *Stochastics* **3** (1979) 127-167.

[25] M. Zakai, On the optimal filtering of diffusion processes, *Z. Wahrscheinlichkeitstheorie und verw. Gebiete* **4** (1969) 230-233.