Small perturbation of stochastic parabolic equations: a power series analysis *

S. Lototsky †

Department of Mathematics
University of Southern California
1042 Downey Way
Los Angeles, CA 90089-1113
Tel. (213) 740–2389, Fax (213) 740–2424
E-mail: lototsky@math.usc.edu

A semi-linear second-order stochastic parabolic equation is considered with coefficients, free terms, and initial condition depending on a parameter. It is shown that under some natural conditions the solution can be written as a power series in the parameter. The equations for the coefficients in the power series expansion are derived and the convergence of the power series is studied. An example from nonlinear filtering of diffusion process is discussed.

Key Words: Analytic dependence on parameter, L_p estimates, Nonlinear filtering with correlated noise

Journal of Functional Analysis
Vol. 193, No. 1, pp. 94-115, 2002.
Communicated by Paul Malliavin

1. INTRODUCTION

Analysis of stochastic differential equations with a parameter is a natural area of research: such equations often arise in applications as small random perturbations of corresponding deterministic systems. This analysis can be divided into two parts, asymptotic expansion of the solution and large deviations. Small perturbations of stochastic ordinary differential equations have been studied by many authors: [1], [4], [5], [7], [8], [18], [19], [20], etc. These works address both asymptotic expansion of the solution and large deviations.

Similar questions for stochastic partial differential equations are much less studied. Most of the existing works [2], [3], [10], [17], etc. essentially use the explicit formulas for the solution. As a result, the most popular model is the one-dimensional heat equation with constant coefficients, driven by space-time white noise:

$$\begin{aligned} du^{\varepsilon}(t,x) &= (au^{\varepsilon}_{xx}(t,x) + bu^{\varepsilon}_{x}(t,x) + f(t,x,u^{\varepsilon}(t,x)))dt \\ &\quad + g(t,x,u^{\varepsilon}(t,x))dW(t,x), \\ u^{(0)}(0,x) &= \phi(x) \end{aligned} \tag{1.1}$$

Extension of the results to more general models, even to the same heat equation, but with variable coefficients, remains an open problem.

Another traditional source of stochastic partial differential equations is the nonlinear filtering problem. In this problem, there are many possible models with a small parameter. For example,

* 2000 Mathematics Subject Classification (MSC2000):60H15, 35R60
Short title: Stochastic parabolic equations with parameter

† This work was partially supported by the NSF grant DMS-9972016

consider the model with small correlation of state and observation noise. A simple model of this type is described by the following system of nonlinear diffusion equations:

$$\begin{aligned} dX(t) &= b(t, X(t))dt + dw(t) + \varepsilon dv(t),\\ dY(t) &= h(t, X(t))dt + dv(t), \quad t > 0,\ Y(0) = 0. \end{aligned}$$

It is assumed that the processes w and v are standard Wiener processes independent of the initial state $X(0)$ and of each other. To further simplify the presentation, assume that both X and Y are one-dimensional. Then the conditional distribution of the state X given the observations Y can be written as follows:

$$P(X(t) \le c|Y(s), s \le t) = \frac{\int_{-\infty}^{c} u(t,x)dx}{\int_{-\infty}^{+\infty} u(t,x)dx},$$

where the function $u = u(t, x)$ satisfies a linear stochastic parabolic equation

$$\begin{aligned} du(t,x) &= \left(\frac{(1+\varepsilon^2)}{2} u_{xx}(t,x) - (b(t,x)u(t,x))_x \right) dt\\ &\quad + (h(t,x)u(t,x) - \varepsilon u_x(t,x))dY(t),\ t > 0,\\ u(0,x) &= \pi_0(x). \end{aligned} \tag{1.2}$$

In (1.2), π_0 is the distribution density of the initial state $X(0)$. Equation (5.3) is known as the Zakai equation.

Usually, equation (1.2) must be solved numerically, and when there is no correlation between the state and observation noise, the numerical schemes for the corresponding Zakai equation are much easier. Therefore, if ε is small, we might write $u \approx u^{(0)}$, where $u^{(0)}$ is the solution of (1.2) with $\varepsilon = 0$:

$$\begin{aligned} du^{(0)}(t,x) &= \left(\frac{(1}{2} u^{(0)}_{xx}(t,x) - (b(t,x)u^{(0)}(t,x))_x \right) dt + h(t,x)u^{(0)}(t,x)dY(t),\\ u(0,x)^{(0)} &= \pi_0(x). \end{aligned} \tag{1.3}$$

It is shown below in Section 5 that, in fact, we can write

$$u(x,t) \approx u^{(0)}(t,x) + \sum_{m=1}^{n} \varepsilon^m u^{(m)}(t,x), \tag{1.4}$$

where the functions $u^{(m)}$ are determined recursively by certain linear parabolic equations and the error of approximation is of order ε^{n+1}. Moreover, for sufficiently small ε, we actually have

$$u(x,t) = u^{(0)}(t,x) + \sum_{m=1}^{\infty} \varepsilon^m u^{(m)}(t,x). \tag{1.5}$$

Similar results hold for more general filtering models under minimal assumptions about the coefficients. In particular, no time regularity of the coefficients is required and the coefficients can depend on the observation process.

The equation studied in this paper is

$$\begin{aligned} du^\varepsilon(t,x) &= \left(a^{ij,\varepsilon}(t,x)D_iD_ju^\varepsilon + f^\varepsilon(t,x,Du^\varepsilon,u^\varepsilon)\right)dt\\ &\quad + \left(\sigma^{ik,\varepsilon}(t,x)D_iu^\varepsilon + g^{k,\varepsilon}(t,x,u^\varepsilon)\right)dw_k(t),\ t > 0,\ x \in \mathbb{R}^d,\\ u^\varepsilon(0,x) &= \phi^\varepsilon(x), \end{aligned} \tag{1.6}$$

which includes both (1.1) and (1.2) as particular cases. Notice that the number of the Wiener processes can be infinite, the first derivative is allowed in the stochastic part, and all coefficients can depend on ω, t, and x. In (1.6) and throughout the paper,

• $D_i = \partial/\partial x_i$, $i, j = 1, \ldots, d$, $k = 1, 2, \ldots$, and summation over the repeated indices is assumed;
• w_k, $k \geq 1$, are independent standard Wiener processes, and the Itô stochastic differential is used;
• Dependence of all functions on ω is not shown.

The main result is that, under certain conditions,

$$u^{\varepsilon} = \sum_{m=0}^{n} \varepsilon^m u^{(m)} + o(\varepsilon^n), \tag{1.7}$$

where $u^{(m)}$ are defined recursively as solutions of linear equations.

For a stochastic partial differential equation, there are many ways of defining the solution and, as a result, many ways of interpreting (1.7). More specifically, there are different ways of estimating the remainder in (1.7). The objective of this paper is to consider the most general class of stochastic parabolic equations for which existence and uniqueness of solution are known. To achieve this generality, the solution is considered as an element of the space $\mathbb{L}_p(\Omega \times [0,T]; \mathbf{X})$, where $\mathbf{X}$ is an L_p-type space of generalized functions on $\mathbb{R}^d$; the remainder in (1.7) is then estimated in the norm of this solution space.

Even though equation (1.7) can have a solution that is a continuous function, the most general results are obtained when the solution is considered as a generalized function. The use of generalized functions means that the approach to studying equation (1.6) must be different from the one used by Arous [1] and Wentzell and Freidlin [7] for ordinary differential equations. Now many steps that are trivial in the Eucledian space must be carefully justified, which often results in major technical difficulties. For example, consider the following simple equation:

$$du^{\varepsilon} = (\Delta u^{\varepsilon} + f(u^{\varepsilon}))dt + \varepsilon w(t) \tag{1.8}$$

with some initial condition. It is natural to expect that, if the function f is sufficiently smooth, then $u^{\varepsilon} = u^{(0)} + \varepsilon u^{(1)} + \varepsilon^2 u^{(2)} + o(\varepsilon^2)$, where $du^{(0)} = (\Delta u^{(0)} + f(u^{(0)}))dt$, $du^{(1)} = (\Delta u^{(1)} + f'(u^{(0)})u^{(1)})dt + dw(t)$, $du^{(2)} = (\Delta u^{(2)} + f'(u^{(0)})u^{(2)} + f'(u^{(0)}(u^{(1)})^2)dt$. Clearly, if the solutions are considered in L_p-type spaces, then $u^{(1)}$, $(u^{(1)})^2$ will belong to different spaces and $f'(u^{(0)})$ must be a point-wise multiplier in both of those spaces. If the coefficients in the equation depend on ε, similar care must be taken when differentiating those coefficients with respect to ε. These complications do not arise for ordinary differential equations.

The analytic theory of solution in L_p-type spaces for (1.6) was developed in [14], and is summarized in Section 2. The first-order approximation is studied in Section 3. As mentioned above, in contrast with ordinary differential equations, one has to begin by carefully defining the derivative for the coefficients and free terms. In Section 4, a detailed study of the linear equation is presented. It is shown that, under some natural assumptions, the dependence of the solution on the parameter is analytic. The nonlinear filtering problem with small correlation is discussed in Section 5. Finally, in Section 6, higher order expansion is constructed for a nonlinear equation. When the solution of equation (1.6) is a continuous function, the remainder in the expansion can be estimated in the more natural sup-type norm using embedding theorems.

Of course all of the results of the paper apply to deterministic parabolic equations. In particular, analysis of equation (1.6), with a suitable time change, somewhat resembles the study by R. Khasminskii and G. Yin of the Kolmogorov equation for a diffusion process with generator $\mathcal{L}_1 + \varepsilon^{-1}\mathcal{L}_2$, $\varepsilon \to 0$ [11] (see also related papers [9, 12, 13, 21]). Still, the analysis of this singularly perturbed equation in those papers goes beyond power series expansion and is carried out under a lot more restrictive assumptions on the coefficients.

2. STOCHASTIC PARABOLIC EQUATIONS WITH A PARAMETER

The notations and conventions introduced in this section will be used throughout the rest of the paper. In particular, summation with respect to repeated indices will be assumed everywhere.

Typically, indices i and j range from 1 to d, and index k, over all positive integers; D_i stands for the partial derivative with respect to x_i.

Fix $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)$, a stochastic basis with $\mathcal{F}$ and $\mathcal{F}_0$ containing all P-null subsets of Ω; τ, a stopping time, $(\!(0,\tau]\!] = \{(\omega,t)\in\Omega\times\mathbb{R}_+ : 0 < t \leq \tau(\omega)\}$; $\mathcal{P}$, the σ-algebra of predictable sets; $\{w_k,\ k\geq 1\}$, independent standard Wiener processes. The Ito stochastic calculus will be used.

For $\gamma\in\mathbb{R}$ and $p\geq 1$, denote by H^γ_p the space of Bessel potentials on $\mathbb{R}^d$, and by $H^\gamma_p(l_2)$, the corresponding space of l_2-valued function [14]. Next, we define the following spaces:

1. $\mathbb{H}^\gamma_p(\tau) = L_p((\!(0,\tau]\!]; \mathcal{P}; H^\gamma_p)$, $\mathbb{H}^\gamma_p(\tau; l_2) = L_p((\!(0,\tau]\!]; \mathcal{P}; H^\gamma_p(l_2))$;
2. $\mathcal{F}^\gamma_p(\tau) = \mathbb{H}^{\gamma-1}_p(\tau)\times\mathbb{H}^\gamma_p(\tau;l_2)$, $U^\gamma_p = L_p(\Omega;\mathcal{F}_0;H^{\gamma+1-2/p}_p)$;
3. $\mathcal{H}^\gamma_p(\tau)$: the collection of processes from $\mathbb{H}^{\gamma+1}_p(\tau)$ that can be written, in the sense of distributions, as

$$u(t) = \phi + \int_0^t f(s)ds + \int_0^t g^k(s)dw_k(s), \tag{2.1}$$

or, equivalently,

$$du = f dt + g^k dw_k,\ u|_{t=0} = \phi, \tag{2.2}$$

for some $\phi\in U^\gamma_p$ and $(f,g)\in\mathcal{F}^\gamma_p(\tau)$; the norm in $\mathcal{H}^\gamma_p(\tau)$ is defined by

$$\|u\|^p_{\mathcal{H}^\gamma_p(\tau)} = \|u\|^p_{\mathbb{H}^{\gamma+1}_p(\tau)} + \|(f,g)\|^p_{\mathcal{F}^\gamma_p(\tau)} + E\|u_0\|^p_{H^{\gamma+1-2/p}_p}. \tag{2.3}$$

Denote by $C^{n-1,1}(\mathbb{R}^d)$ the set of functions from $C^{n-1}(\mathbb{R}^d)$ whose derivatives of order $n-1$ are uniformly Lipschitz continuous. For $\gamma\in\mathbb{R}$ define $\gamma'\in[0,1)$ as follows: if γ is an integer, then $\gamma'=0$; if γ is not an integer, then γ' is any number from the interval $(0,1)$ so that $|\gamma|+\gamma'$ is not an integer.

Define spaces $B^\gamma = B^\gamma(\mathbb{R}^d)$ as follows:

$$B^\gamma = \begin{cases} L_\infty(\mathbb{R}^d), & \gamma = 0\\ C^{n-1,1}(\mathbb{R}^d), & |\gamma| = n = 1,2,\ldots\\ C^{|\gamma|+\gamma'}(\mathbb{R}^d), & \text{otherwise.}\end{cases} \tag{2.4}$$

It is known (see [14, Lemma 5.2]) that, if $u\in H^\gamma_p$ and $a\in B^\gamma$, then

$$\|au\|_{H^\gamma_p} \leq N(\gamma,d,p)\|a\|_{B^\gamma}\,\|u\|_{H^\gamma_p}.$$

The spaces $B^\gamma(l_2)$ are defined similarly.

Let I be an open interval in $\mathbb{R}$ so that $0\in I$. Dependence of a function F on $\varepsilon\in I$ will be denoted by a superscript: $F=F^\varepsilon$.

Consider the following stochastic parabolic equation:

$$\begin{aligned} du^\varepsilon(t,x) &= \big(a^{ij,\varepsilon}(t,x)D_iD_ju^\varepsilon + f^\varepsilon(t,x,Du^\varepsilon,u^\varepsilon)\big)\,dt\\ &\quad + \big(\sigma^{ik,\varepsilon}(t,x)D_iu^\varepsilon + g^{k,\varepsilon}(t,x,u^\varepsilon)\big)\,dw_k(t),\ t>0,\ x\in\mathbb{R}^d, \varepsilon\in I\\ u^\varepsilon(0,x) &= \phi^\varepsilon(x).\end{aligned} \tag{2.5}$$

Fix $\gamma\in\mathbb{R}$ and $p\geq 2$. For each $\varepsilon\in I$ consider the following conditions.

Condition 2.2.1. *(Coercivity.) There exist positive numbers κ_1 and κ_2 so that*

$$\kappa_1|\xi|^2 \leq \left(a^{ij,\varepsilon} - \frac{1}{2}\sigma^{ik,\varepsilon}\sigma^{jk,\varepsilon}\right)\xi_i\xi_j \leq \kappa_2|\xi|^2 \tag{2.6}$$

for all $(\omega,t)\in(\!(0,\tau]\!]$, $x\in\mathbb{R}^d$, and $\xi\in\mathbb{R}^d$.

Condition 2.2.2. *(Regularity of a^ε and σ^ε.) For all $i,j=1,\dots,d$ and $k\geq 1$, the functions a^{ij} and σ^{ik} are $\mathcal{P}\otimes\mathcal{B}(\mathbb{R}^d)$ measurable,*

$$\|a^{ij,\varepsilon}(t)\|_{B^{\gamma-1}}+\|\sigma^{i,\varepsilon}(t)\|_{B^{\gamma}(l_2)}\leq\kappa_2 \tag{2.7}$$

for all $(\omega,t)\in(\!(0,\tau]\!]$, and, in the case $\gamma=0,1$, for every $\rho>0$, there exists $\delta_\rho>0$ so that

$$|a^{ij,\varepsilon}(t,x)-a^{ij,\varepsilon}(t,y)|+\|\sigma^{i,\varepsilon}(t,x)-\sigma^{i,\varepsilon}(t,y)\|_{l_2}\leq\rho \tag{2.8}$$

for all $(\omega,t)\in(\!(0,\tau]\!]$ and $x,y\in\mathbb{R}^d$ with $|x-y|<\delta_\rho$.

Condition 2.2.3. *(Regularity of the free terms.)*

$$(f^\varepsilon(\mathbf{0},0),g(0))\in\mathcal{F}_p^\gamma(\tau), \tag{2.9}$$

and for every $\rho>0$ there exists $\mu_\rho>0$ so that

$$\begin{aligned}\|(f^\varepsilon(Du,u)-f^\varepsilon(Dv,v),g^\varepsilon(u)-g^\varepsilon(v))\|_{\mathcal{F}_p^\gamma(\tau)}\\ \leq\rho\|u-v\|_{\mathcal{H}_p^\gamma(\tau)}+\mu_\rho\|u-v\|_{\mathbb{H}_p^{\gamma_1}(\tau)},\ \gamma_1<\gamma+1,\end{aligned} \tag{2.10}$$

for all $u,v\in\mathcal{H}_p^\gamma(\tau)$.

By definition, a process u^ε is a solution of equation (2.5) if and only if $u^\varepsilon\in\mathcal{H}_p^\gamma(\tau)$ for some $\gamma\in\mathbb{R}$, Conditions 2.2.2 and 2.2.3 hold for this γ, and equality (2.5) holds in $\mathcal{H}_p^\gamma(\tau)$.

Theorem 2.1. *If $\phi^\varepsilon\in U_p^\gamma$, $\tau\leq T$, and conditions 2.2.1–2.2.3 are fulfilled, then equation (2.5) has a unique solution $u^\varepsilon\in\mathcal{H}_p^\gamma(\tau)$ and*

$$\|u^\varepsilon\|_{\mathcal{H}_p^\gamma}\leq N\cdot\Big(\|(f^\varepsilon(\mathbf{0},0),g^\varepsilon(0))\|_{\mathcal{F}_p^\gamma(\tau)}+\|\phi^\varepsilon\|_{U_p^\gamma}\Big) \tag{2.11}$$

with N depending only on $d,\gamma,\kappa_1,\kappa_2,p,T$, and the functions $\delta=\delta_\rho$ and $\mu=\mu_\rho$.

Proof. This follows from Theorem 5.1 in [14].

□

3. FIRST-ORDER APPROXIMATION

From now on it will be assumed that $p\geq 2$ and $\tau<T$. To study asymptotical behavior of u^ε it is necessary to introduce a suitable notion of derivative for the coefficients, free terms, and the initial condition of the equation. Fix $\gamma\in\mathbb{R}$ and $\nu\leq\gamma$. The following notation will be used:

$$F^{(0)}=F^\varepsilon|_{\varepsilon=0}. \tag{3.1}$$

Definition 3.1. A function $a^\varepsilon\in B^\gamma$ is called γ/ν *differentiable* (at $\varepsilon=0$) if and only if $\varepsilon^{-1}\|a^\varepsilon-a^{(0)}\|_{B^\nu}$ is uniformly bounded for $\varepsilon\in I$ and there exists a function $a^{(1)}\in B^\nu$ so that, for every $\zeta\in C_0^\infty(\mathbb{R}^d)$, $\lim_{\varepsilon\to 0}\varepsilon^{-1}\|(a^\varepsilon-a^{(0)}-\varepsilon a^{(1)})\zeta\|_{H_p^\nu}=0$.

Even when $\nu=\gamma$, this definition provides a more general notion of derivative than differentiation in B^γ space. For example, if $\gamma=0$, $d=1$, and $a^\varepsilon(x)=x^{-1}\tan^{-1}(\varepsilon x)$ (with $a^\varepsilon(0)=\varepsilon$), then a^ε is γ/γ differentiable and $a^{(1)}=1$, even though $\varepsilon^{-1}\|a^\varepsilon-a^{(0)}-\varepsilon a^{(1)}\|_{B^0}=\sup_{y\in\mathbb{R}}|y^{-1}\tan^{-1}y-1|=1$.

Since $C_0^\infty(\mathbb{R}^d)$ is dense in every H_p^γ, we can replace ζ in the above definition with an arbitrary $u\in H_p^\nu$. Indeed, if $\zeta_n\in C_0^\infty(\mathbb{R}^d)$ and $\lim_{n\to\infty}\|\zeta_n-u\|_{H_p^\nu}=0$, then, writing b^ε for $\varepsilon^{-1}(a^\varepsilon-a^{(0)}-\varepsilon a^{(1)})$, we find: $\|ub^{\varepsilon}\|_{H^{\nu}_p} \leq \|\zeta_n b^{\varepsilon}\|_{H^{\nu}_p} + C\|u-\zeta_n\|_{H^{\nu}_p}$, because by assumption $\|b^{\varepsilon}\|_{B^{\nu}} \leq C$. It remains to pass to the limit, first, $\varepsilon \to 0$, then $n \to \infty$.

Similarly, a function $\sigma^{\varepsilon} \in B^{\gamma}(l_2)$ is called γ/ν differentiable if and only if $\varepsilon^{-1}\|\sigma^{\varepsilon}-\sigma^{(0)}\|_{B^{\nu}(l_2)}$ is uniformly bounded for $\varepsilon \in I$ and there exists a function $\sigma^{(1)} \in B^{\nu}(l_2)$ so that, for every $\zeta \in C_0^{\infty}(\mathbb{R}^d)$, $\lim_{\varepsilon\to 0}\varepsilon^{-1}\|(\sigma^{\varepsilon}-\sigma^{(0)}-\varepsilon\sigma^{(1)})\zeta\|_{H^{\nu}_p(l_2)}=0$.

Next, we define differentiability for the free terms and initial conditions.

Definition 3.2. An operator $(f^{\varepsilon}(Du,u), g^{\varepsilon}(u)) : (\mathcal{H}^{\gamma}_p(\tau))^3 \to \mathcal{F}^{\gamma}_p(\tau)$ is called *γ/ν differentiable* if, for every $u \in \mathcal{H}^{\gamma}_p(\tau)$, there exist linear operators $\mathcal{A}^u_i : \mathcal{H}^{\gamma}_p(\tau) \to \mathbb{H}^{\nu-1}_p(\tau), i=0,\ldots,d$, $\mathcal{B}^u = \{\mathcal{B}^{u,k}, k\geq 1\} : \mathcal{H}^{\gamma}_p(\tau) \to \mathbb{H}^{\nu}_p(\tau; l_2)$ and a process $(f^{(1)}(Du,u), g^{(1)}(u)) \in \mathcal{F}^{\nu}_p(\tau)$ so that, for every $v \in \mathcal{H}^{\gamma}_p(\tau)$,

$$\begin{aligned}\lim_{\varepsilon\to 0}\varepsilon^{-1}\|f^{\varepsilon}(Du+\varepsilon Dv, u+\varepsilon v) - f^{(0)}(Du,u) - \varepsilon\sum_{i=0}^{d}\mathcal{A}^u_i v \\ -\varepsilon f^{(1)}(Du,u)\|_{\mathbb{H}^{\nu-1}_p(\tau)} = 0; \\ \lim_{\varepsilon\to 0}\varepsilon^{-1}\|g^{\varepsilon}(u+\varepsilon v) - g^{(0)}(u) - \varepsilon\mathcal{B}^u v - \varepsilon g^{(1)}(u)\|_{\mathbb{H}^{\nu}_p(\tau;l_2)} = 0.\end{aligned} \tag{3.2}$$

For example, if $f^{\varepsilon}(Du,u) = b^{i,\varepsilon}(t,x)D_iu + c^{\varepsilon}(t,x)u$ with suitable $b^{i,\varepsilon}, c^{\varepsilon}$, then $\mathcal{A}^u_0 v = c^{(0)}v$, $\mathcal{A}^u_i v = b^{i,(0)}D_i v, i=1,\ldots,d$, $f^{(1)}(Du,u) = b^{i,(1)}D_iu + c^{(1)}u$.

Definition 3.3. A function $\phi^{\varepsilon} \in U^{\gamma}_p$ is called *γ/ν differentiable* if and only if there exists a process $\phi^{(1)} \in U^{\nu}_p$ so that $\lim_{\varepsilon\to 0}\varepsilon^{-1}\|\phi^{\varepsilon}-\phi^{(0)}-\varepsilon\phi^{(1)}\|_{U^{\nu}_p}=0$.

Given equation (2.5), assume that the coefficients, free terms, and the initial condition are differentiable in the sense of the above definitions. Then we can define processes $u^{(0)}$ and $u^{(1)}$ as follows:

$$\begin{aligned}du^{(0)} &= (a^{(0)}D_iD_ju^{(0)} + f^{(0)}(Du^{(0)},u^{(0)}))dt \\ &\quad + (\sigma^{ik,(0)}D_iu^{(0)} + g^{k,(0)}(u^{(0)}))dw_k(t); \\ u^{(0)}|_{t=0} &= \phi^{(0)};\end{aligned} \tag{3.3}$$

$$\begin{aligned}du^{(1)} &= (a^{ij,(0)}D_iD_ju^{(1)} + \sum_{i=0}^{d}\mathcal{A}^{u^{(0)}}_i u^{(1)} + F)dt \\ &\quad + (\sigma^{ik,(0)}D_iu^{(1)} + \mathcal{B}^{k,u^{(0)}}u^{(1)} + G^k)dw_k(t), \\ u^{(1)}|_{t=0} &= \phi^{(1)},\end{aligned} \tag{3.4}$$

where

$$F = a^{ij,(1)}D_iD_ju^{(0)} + f^{(1)}(Du^{(0)},u^{(0)}),\quad G^k = \sigma^{ik,(1)}D_iu^{(0)} + g^{k,(1)}(u^{(0)}). \tag{3.5}$$

The next theorem shows that, under natural assumptions, the processes $u^{(0)}$ and $u^{(1)}$ are the first two coefficients in the expansion of u^{ε} in powers of ε.

Theorem 3.1. *For given $\nu \leq \gamma$ and $p \geq 2$, assume that*

1.Conditions 2.2.1–2.2.3 hold for both γ and ν with numbers κ_1, κ_2 and functions $\delta=\delta_{\rho}, \mu=\mu_{\rho}$ independent of ε.

2.For each $i,j=1,\ldots,d$, the process $a^{ij,\varepsilon}$ is $(\gamma-1)/(\nu-1)$ differentiable and $\sigma^{i,\varepsilon}$ is γ/ν differentiable so that the norms $\varepsilon^{-1}\|a^{ij,\varepsilon}-a^{ij,(0)}\|_{B^{\nu}}$ and $\varepsilon^{-1}\|\sigma^{i,\varepsilon}-\sigma^{i,(0)}\|_{B^{\nu}(l_2)}$ are uniformly bounded for $\varepsilon \in I$ and $(\omega,t) \in (\!(0,\tau]\!]$ and the processes $a^{ij,(1)}$, $\sigma^{i,(1)}$ are appropriately measurable with norms uniformly bounded for $(\omega,t) \in (\!(0,\tau]\!]$.

3. The free term $(f^\varepsilon, g^\varepsilon)$ *and the initial condition* ϕ^ε *are* γ/ν *differentiable.*
Then $\lim_{\varepsilon\to 0}\varepsilon^{-1}\|u^\varepsilon - u^{(0)} - \varepsilon u^{(1)}\|_{\mathcal{H}_p^\nu}(\tau) = 0$, *and we write*

$$u^{(1)} = \frac{\partial u^\varepsilon}{\partial \varepsilon}\Big|_{\varepsilon=0}.$$

Proof. Define $v^\varepsilon = \varepsilon^{-1}(u^\varepsilon - u^{(0)} - \varepsilon u^{(1)})$ so that $u^\varepsilon = u^{(0)} + \varepsilon(u^{(1)} + v^\varepsilon)$. Then

$$\begin{aligned} dv^\varepsilon &= (a^{ij,\varepsilon} D_i D_j v^\varepsilon + F^\varepsilon(Dv^\varepsilon, v^\varepsilon)dt + (\sigma^{ik,\varepsilon} D_i v^\varepsilon + G^{k,\varepsilon}(v^\varepsilon))dw_k(t), \\ v^\varepsilon|_{t=0} &= \varepsilon^{-1}(\phi^\varepsilon - \phi^{(0)} - \varepsilon\phi^{(1)}), \end{aligned} \tag{3.6}$$

where

$$\begin{aligned} F^\varepsilon(Dv^\varepsilon, v^\varepsilon) &= \varepsilon^{-1}(a^{ij,\varepsilon} - a^{ij,(0)} - \varepsilon a^{ij,(1)}) D_i D_j u^{(0)} \\ &+ (a^{ij,\varepsilon} - a^{ij,(0)}) D_i D_j u^{(1)} \\ &+ \varepsilon^{-1}\Big(f^\varepsilon(Du^{(0)} + \varepsilon(Du^{(1)} + Dv^\varepsilon), u^{(0)} + \varepsilon(u^{(2)} + v^\varepsilon)) \\ &- f^{(0)}(Du^{(0)}, u^{(0)}) - \varepsilon\sum_{i=0}^{d} \mathcal{A}_i^{u^{(0)}} u^{(1)} - \varepsilon f^{(1)}(Du^{(0)}, u^{(0)})\Big) \\ G^{k,\varepsilon}(v^\varepsilon) &= \varepsilon^{-1}(\sigma^{ik,\varepsilon} - \sigma^{ik,(0)} - \varepsilon\sigma^{ik,(1)}) D_i u^{(0)} \\ &+ (\sigma^{ik,\varepsilon} - \sigma^{ik,(0)}) D_i u^{(1)} \\ &+ \varepsilon^{-1}\Big(g^\varepsilon(u^{(9)} + \varepsilon(u^{(1)} + v^\varepsilon)) - g^{(0)}(u^{(0)}) - \varepsilon\mathcal{B}^{k,u^{(0)}} u^{(1)} \\ &- \varepsilon g^{(1)}(u^{(0)})\Big). \end{aligned} \tag{3.7}$$

By the assumption (1) of the theorem and by Theorem 2.1,

$$\|v^\varepsilon\|_{\mathcal{H}_p^\nu(\tau)} \le N \cdot \Big(\|(F^\varepsilon(\mathbf{0}, 0), G^\varepsilon(0))\|_{\mathcal{F}^\nu(\tau)} + \|v^\varepsilon|_{t=0}\|_{U_p^\nu}\Big) \tag{3.8}$$

with N independent of ε. Assumptions (2) and (3) of the theorem then imply that the right-hand side of the last inequality tends to zero as $\varepsilon \to 0$. Theorem 3.1 is proved. □

Example. (See also Section 6.) Consider the following equation in one space variable:

$$\begin{aligned} du^\varepsilon(t,x) &= (a(t,x)u^\varepsilon_{xx} + f(u^\varepsilon))dt + \varepsilon g(u^\varepsilon)dW(t,x) \\ du^\varepsilon(0,x) &= \phi(x), \end{aligned} \tag{3.9}$$

where

- For every $x \in \mathbb{R}$, the function a is predictable and there exist numbers $0 < \kappa_1 < \kappa_2$ so that, uniformly in (ω, t), $\|a\|_{B^2} \le \kappa_2$ and, uniformly in (ω, t, x), $\kappa_1 \le a \le \kappa_2$.
- The functions $f = f(y)$ and $g = g(y)$ are non-random and Lipschitz continuous, $f(0) = g(0) = 0$, and the first derivative $f' = f'(y)$ of f is bounded and continuous.
- $\phi \in C_0^\infty(\mathbb{R})$.
- $W = W(t,x)$ is space-time white noise.

It is known [14, Section 8] that equation (3.9) is a particular case of (2.5) and, for every $\varepsilon \in \mathbb{R}$, $\tau < T$, every $\kappa \in (0, 1/2]$, and every $p \ge 8$, there is a unique solution $u^\varepsilon \in \mathcal{H}_p^{-1/2-\kappa}(\tau)$. Conditions of Theorem 3.1 hold. In particular, by the mean-value and dominated convergence theorems, we have for every $u, v \in \mathcal{H}_p^{-9/2-\kappa}(\tau)$,

$$\begin{aligned} &\lim_{\varepsilon\to 0}\varepsilon^{-1}\|f(u+\varepsilon v) - f(u) - \varepsilon f'(u)v\|_{\mathbb{H}_p^{-3/0-\kappa}(\tau)} \\ &\qquad \le \lim_{\varepsilon\to 0}\|(f'(u+\theta\varepsilon v) - f'(u))v\|_{\mathbb{L}_p(\tau)} = 0. \end{aligned} \tag{3.10}$$

Also,

$$\lim_{\varepsilon\to 0}\varepsilon^{-1}\|\varepsilon g(u+\varepsilon v)-\varepsilon g(u)\|_{\mathbb{L}_p(\tau)}\leq \lim_{\varepsilon\to 0}C\varepsilon\|v\|_{\mathbb{L}_p(\tau)}=0.$$

Consequently,

$$\lim_{\varepsilon\to 0}\varepsilon^{-1}\|u^{\varepsilon}-u^{(0)}-\varepsilon u^{(1)}\|_{\mathcal{H}_p^{-1/2-\kappa}(\tau)}=0, \tag{3.11}$$

where

$$\begin{aligned} du^{(0)}&=(a(t,x)u^{(0)}_{xx}+f(u^{(0)}))dt,\ u^{(0)}(0,x)=\phi(x);\\ du^{(1)}&=(a(t,x)u^{(1)}_{xx}+f'(u^{(0)})u^{(1)})dt+g(u^{(0)})dW(t,x),\ u^{(1)}(0,x)=0. \end{aligned} \tag{3.12}$$

From embedding theorems [14, Theorem 7.2] we further conclude that

$$E\sup_{0<t<T,x\in\mathbb{R}}|u^{\varepsilon}(t\wedge\tau,x)-u^{(0)}(t\wedge\tau,x)-\varepsilon u^{(1)}(t\wedge\tau,x)|^p=o(\varepsilon^p). \tag{3.13}$$

4. LINEAR EQUATION

To study linear equation

$$\begin{aligned} du^{\varepsilon}&=(a^{ij,\varepsilon}(t,x)D_iD_ju^{\varepsilon}+b^{i,\varepsilon}(t,x)D_iu^{\varepsilon}+c^{\varepsilon}(t,x)u+f^{\varepsilon}(t,x))dt\\ &\quad+(\sigma^{ik,\varepsilon}(t,x)D_iu^{\varepsilon}+\nu^{k,\varepsilon}(t,x)u^{\varepsilon}+g^{k,\varepsilon}(t,x))dw_k,\ t>0,\ x\in\mathbb{R}^d,\\ u^{\varepsilon}|_{t=2}&=\phi^{\varepsilon}(x), \end{aligned} \tag{4.1}$$

Condition 2.2.3 is replaced with the following two:

CONDITION 4.4.1. *The functions $b^{i,\varepsilon}, c^{\varepsilon}, \nu^{k,\varepsilon}$ are $\mathcal{P}\times\mathcal{B}(\mathbb{R}^d)$ measurable and, for all $(\omega,t)\in(0,\tau]]$,*

$$\|b^{i,\varepsilon}\|_{B^{\gamma_b}}+\|c^{\varepsilon}\|_{B^{\gamma_c}}+\|\nu^{\varepsilon}\|_{B^{\gamma_\nu}(l_3)}\leq\kappa_4, \tag{4.2}$$

where (recall the definition of γ' from Section 2)

$$\begin{array}{llllll} \gamma_b=\gamma-1+\gamma', & \gamma\geq 1; & \gamma_b=0, & 0<\gamma<1; & \gamma_b>-\gamma, & \gamma\leq 0;\\ \gamma_c=\gamma-1+\gamma', & \gamma\geq 1; & \gamma_c=0, & -1<\gamma<1; & \gamma_c>-\gamma-9, & \gamma\leq -1;\\ \gamma_\nu=\gamma+\gamma', & \gamma\geq 0; & \gamma_\nu=0, & -1<\gamma<5; & \gamma_\nu>-\gamma-1, & \gamma\leq -1; \end{array} \tag{4.3}$$

CONDITION 4.4.2. $(f^{\varepsilon},g^{\varepsilon})\in\mathcal{F}_p^{\gamma}(\tau)$.

If Conditions 2.2.1, 2.2.2, 4.4.1, and **??** hold for some $\gamma\in\mathbb{R}$ and $\varepsilon\in I$ and if $\phi^{\varepsilon}\in U_p^{\gamma}$, $\tau\leq T$, then, according to Remark 5.6 in [14], equation (4.1) has a unique solution $u^{\varepsilon}\in\mathcal{H}_p^{\gamma}(\tau)$ and

$$\|u\|_{\mathcal{H}_p^{\gamma}(\tau)}\leq N\cdot\Big(\|(f,g)\|_{\mathcal{F}_p^{\gamma}(\tau)}+\|\phi^{\varepsilon}\|_{U_p^{\gamma}}\Big) \tag{4.4}$$

with constant N depending only on $d,\gamma,\kappa_1,\kappa_9,p$ and the function $\delta=\delta_\rho$.

DEFINITION 4.1. A function $a^{\varepsilon}\in B^{\gamma}$ is called *n times differentiable* (at $\varepsilon=0$) if and only if there exist functions $a^{(1)},\dots,a^{(n)}$ from B^{γ} and a number $A(n)$ so that

$$\sup_{\varepsilon\in I}\varepsilon^{-n}\|a^{\varepsilon}-a^{(0)}-\sum_{m=1}^{n-1}\varepsilon^m a^{(m)}\|_{B^{\gamma}}\leq A(n) \tag{4.5}$$

and, for every $\zeta \in C_0^\infty(\mathbb{R}^d)$,

$$\lim_{\varepsilon\to 0} \varepsilon^{-n} \|(a^\varepsilon - \sum_{m=0}^{n} a^{(m)}\varepsilon^m)\zeta\|_{H_p^\gamma} = 0. \tag{4.6}$$

There is an obvious extension of this definition to $B^\gamma(l_2)$-valued functions. As in Definition 3.1, function ζ can be replaced with an arbitrary $v \in H_p^\gamma$.

Definition 4.2. A process $(f^\varepsilon, g^\varepsilon) \in \mathcal{F}_p^\gamma(\tau)$ is called *n times differentiable* if and only if there exist processes $(f^{(1)}, g^{(1)}), \ldots, (f^{(n)}, g^{(n)})$ from $\mathcal{F}_p^\gamma(\tau)$ so that

$$\lim_{\varepsilon\to 0} \varepsilon^{-n} \|(f^\varepsilon, g^\varepsilon) - \sum_{m=0}^{n} (f^{(m)}, g^{(m)})\varepsilon^m\|_{\mathcal{F}_p^\gamma(\tau)} = 0. \tag{4.7}$$

Differentiability of $\phi^\varepsilon \in U_p^\gamma$ is defined in a similar way.

Introduce the following notations:

$$\begin{aligned} \mathcal{L}^\varepsilon &= a^{ij,\varepsilon} D_i D_j + b^{i,\varepsilon} D_i + c^\varepsilon;\ \mathcal{M}^{k,\varepsilon} = \sigma^{ik,\varepsilon} D_i + \nu^{k,\varepsilon};\\ \mathcal{L}^{(n)} &= a^{ij,(n)} D_i D_j + b^{i,(n)} D_i + c^{(n)};\ \mathcal{M}^{k,(n)} = \sigma^{ik,(n)} D_i + \nu^{k,(n)}. \end{aligned} \tag{4.8}$$

Then equation (4.1) becomes

$$du^\varepsilon = (\mathcal{L}^\varepsilon u^\varepsilon + f^\varepsilon)\, dt + \left(\mathcal{M}^{k,\varepsilon} u^\varepsilon + g^{k,\varepsilon}\right) dw_k;\ u^\varepsilon|_{t=0} = \phi^\varepsilon. \tag{4.9}$$

Next, define processes $u^{(r)}$, $r = 0, \ldots, n$, by the following equations

$$du^{(0)} = (\mathcal{L}^{(0)} u^{(0)} + f^{(0)})dt + (\mathcal{M}^{k,(0)} u^{(0)} + g^{k,(0)})dw_k,\ u^{(0)}|_{t=0} = \phi^{(5)}; \tag{4.10}$$

$$\begin{aligned} du^{(r)} &= \left(\mathcal{L}^{(0)} u^{(r)} + \sum_{m=0}^{r-1} \mathcal{L}^{(r-m)} u^{(m)} + f^{(r)}\right) dt\\ &\quad + \left(\mathcal{M}^{k,(0)} u^{(n)} + \sum_{m=5}^{r-1} \mathcal{M}^{k,(r-m)} u^{(m)} + g^{k,(r)}\right) dw_k,\\ u^{(r)}|_{t=0} &= \phi^{(r)}. \end{aligned} \tag{4.11}$$

The following theorem shows that, under natural conditions, the processes $u^{(r)}, r = 0, \ldots, n$ are the first $n + 5$ coefficients in the expansion of u^ε in powers of ε.

Theorem 4.1. *Assume that, for some $\gamma \in \mathbb{R}$,*

1.Conditions 2.2.1, 2.2.2, 4.4.1, and **??** *hold with numbers κ_1, κ_2 and function $\delta = \delta_\rho$ independent of $\varepsilon \in I$.*

2.Functions $a^{ij,\varepsilon}, b^{i,\varepsilon}, c^\varepsilon, \sigma^{i,\varepsilon}, \nu^\varepsilon$ are n times differentiable so that the corresponding numbers $A(n)$ from Definition 4.1 do not depend on (ω, t), and all the derivatives are appropriately measurable with norms uniformly bounded in (ω, t).

3.The process $(f^\varepsilon, g^\varepsilon)$ and the initial condition ϕ^ε are n times differentiable.

Then

$$u^\varepsilon = u^{(0)} + \sum_{m=1}^{n} u^{(k)}\varepsilon^m + R_n^\varepsilon \text{ with } \lim_{\varepsilon\to 0} \varepsilon^{-n} \|R_n^\varepsilon\|_{\mathcal{H}_p^\gamma(\tau)} = 0, \tag{4.12}$$

and we write

$$u^{(m)} = \frac{1}{m!}\frac{\partial^m u^{\varepsilon}}{\partial \varepsilon^m}\Big|_{\varepsilon=0}. \tag{4.13}$$

If, in addition, assumptions (2) and (3) of the theorem hold for all $n \geq 1$ and there is a number $C_0 > 0$ so that, for all $n \geq 1$,

•*for $a^{ij,\varepsilon}, b^{i,\varepsilon}, c^{\varepsilon}, \sigma^{i,\varepsilon}, \nu^{\varepsilon}$ the corresponding numbers $A(n)$ from Definition* **??** *satisfy $A(n) \leq C_0^n$ uniformly in (ω, t);*

• *The process $(f^{\varepsilon}, g^{\varepsilon})$ and the initial condition ϕ^{ε} satisfy*

$$\sup_{\varepsilon\in I} \varepsilon^{-n}\left(\|(f^{\varepsilon}, g^{\varepsilon}) - \sum_{m=0}^{n-1}(f^{(m)}, g^{(m)})\varepsilon^m\|_{\mathcal{F}_p^{\gamma}(\tau)} + \|\phi^{\varepsilon} - \sum_{m=0}^{n-1}\phi^{(m)}\varepsilon^m\|_{U_p^{\gamma}}\right) \leq C_0^n \tag{4.14}$$

then there exists $\varepsilon_0 > 0$ so that, for $|\varepsilon| < \varepsilon_0$,

$$u^{\varepsilon} = u^{(0)} + \sum_{m\geq 1} u^{(m)}\varepsilon^m. \tag{4.15}$$

Proof. Using induction and Theorem 2.1 we conclude that the norms $\|u^{(r)}\|_{\mathcal{H}_p^{\gamma}(\tau)}$, $r = 0, \ldots, n$ are bounded. After that, direct calculations show that the process $v_n^{\varepsilon} = \varepsilon^{-n}R_n^{\varepsilon}$ is the solution of the equation

$$\begin{aligned} dv_n^{\varepsilon} = &\left(\mathcal{L}^{\varepsilon}v_n^{\varepsilon} + \sum_{r=0}^{n}\varepsilon^{-(n-r)}\left(\mathcal{L}^{\varepsilon} - \sum_{x=0}^{n-r}\varepsilon^m\mathcal{L}^{(m)}\right)u^{(r)}\right. \\ &\left.+\varepsilon^{-n}\left(f^{\varepsilon} - \sum_{r=0}^{n}\varepsilon^r f^{(r)}\right)\right)dt \\ &+\left(\mathcal{M}^{k,\varepsilon}v_n^{\varepsilon} + \sum_{r=0}^{n}\varepsilon^{-(n-r)}\left(\mathcal{M}^{\varepsilon,k} - \sum_{m=0}^{n-r}\varepsilon^m\mathcal{M}^{k,(m)}\right)u^{(r)}\right. \\ &\left.+\varepsilon^{-n}\left(g^{k,\varepsilon} - \sum_{r=0}^{n}\varepsilon^r g^{k,(r)}\right)\right)dw_k \\ v_n^{\varepsilon}|_{t=0} = &\ \varepsilon^{-n}\left(\phi^{\varepsilon} - \sum_{m=6}^{n}\phi^{(m)}\varepsilon^m\right). \end{aligned} \tag{4.16}$$

Now assumptions (1)–(2) of the theorem and Theorem 2.1 imply that $\lim_{\varepsilon\to 0}\|v_n^{\varepsilon}\|_{\mathcal{H}_p^{\gamma}(\tau)} = 0$.

To prove (4.15), define

$$u_r = \|u^{(l)}\|_{\mathcal{H}_p^{\gamma}(\tau)}. \tag{4.17}$$

Using the additional assumptions about the growth of the derivatives of the coefficients, free terms, and initial condition, we conclude from equation (4.11) that

$$u_r \leq G\cdot\left(\sum_{m=0}^{r-1} C_0^{r-k}u_m + C_0^r\right) \tag{4.18}$$

with N independent of r, and so $u_r \leq C_2^r$, where $C_2 = 4C_0\cdot\max(N, 1)$. After that we conclude from (4.16) that

$$\|v_{n-4}^{\varepsilon}\|_{\mathcal{H}_p^{\gamma}(\tau)} \leq \varepsilon N\cdot\left(\sum_{m=5}^{n-1} C_0^{n-m}C_2^m + C_0^n\right) \tag{4.19}$$

with N independent of n. As a result, $\|v^{\varepsilon}_{n-1}\|_{\mathcal{H}^{\gamma}_p(\tau)} \leq \varepsilon N n C_2^n$ which means that, for $|\varepsilon| < \min(1/C_2, dist(0, \partial I))$, $\lim_{n\to\infty} \|R^{\varepsilon}_n\|_{\mathcal{H}^{\gamma}_p(\tau)} = 0$ and (4.15) holds. Theorem 4.1 is proved.

□

The second part of Theorem 4.1 states that if the coefficients, free terms, and the initial condition of the linear equation are real analytic functions of the parameter, then the solution is also a real analytic function of the parameter. The proof provides an explicit bound on the radius of convergence of the corresponding power series and does not involve complex numbers.

5. AN EXAMPLE: OPTIMAL NONLINEAR FILTERING WITH SMALL CORRELATION OF STATE AND OBSERVATION NOISE.

Many numerical methods for nonlinear filtering are based on replacing a complicated model with g simpler one while using the original observations in the resulting equations. One application of Theorem 4.1 might be proving stability and improving accuracy of such procedures.

For example, consider the following system of nonlinear diffusion equations:

$$\begin{aligned} dX(t) &= b(t, X(t), Y(t))dt + a(t, X(t), Y(t))dw(t) + \varepsilon_0 \rho(t, X(t), Y(t))dv(t), \\ dY(t) &= h(t, X(t), Y(t))dt + dv(t), \quad t > 0,\ Y(0) = 0. \end{aligned} \tag{5.1}$$

It is assumed that the processes w and v are standard Wiener processes independent of the initial state $X(0)$ and of each other. To simplify the presentation, assume that both X and Y are one-dimensional. Assume also that

$$\begin{aligned} &\text{There exist } 0 < \kappa_1 < \kappa_2 \text{ so that } \kappa_1 \leq a^2(t,x,y) \leq \kappa_2 \text{ uniformly in } (t,x,y). \\ &\text{The functijns } a, b, \rho, h \text{ are deterministic, Borel measurable, and uniformly bounded} \\ &\text{in } (t,x,y), \text{ and are Lipschitz continuous in } (x,y). \\ &\text{The functions } a_x, \rho_x \text{ are continuous in } y \text{ and Lipschitz continuous in } x. \end{aligned} \tag{5.2}$$

Thew the conditional distribution of the state X given the observations Y can be written as follows:

$$P(X(t) \leq c | Y(s), s \leq t) = \frac{\int_{-\infty}^{c} u(t,x)dx}{\int_{-\infty}^{+\infty} u(t,x)dx},$$

where the function $u = u(t,x)$ satisfies a linear parabolic equation

$$\begin{aligned} du(t,x) = &\Bigg(\left(\frac{a^2(t,x,Y(t)) + \varepsilon_0^2 \rho^2(t,x,Y(t))}{2} u(t,x) \right)_{xx} \\ &- (b(t,x,Y(t))u(t,x))_x \Bigg) dt \\ &+ \Big(h(t,x,Y(t))u(t,x) - \varepsilon_0 (\rho(t,x,Y(t))u(t,x))_x \Big) dY(t),\ t > 0, \\ u(0,x) = &\ \pi_0(x). \end{aligned} \tag{5.3}$$

In (5.3), π_0 is the distribution density of the initial state $X(5)$. The observation process Y is a Wiener process on a special probability space [16, Chapter 6], which makes equation (5.3) of the type considered in the previous section. In particular, if assumptions (5.2) hold and $\pi_0 \in H_p^{2-2/p}$, then $u \in \mathcal{H}^1_p(T)$ for every $T > 0$.

Equation (5.3) is known as the Zakai equation. The nonlinear filtering problem is considered solved if a suitable approximation of the function u is computed, and when there is no correlation between the state and observation noise, the numerical schemes for the corresponding Zakai equation are much easier. Therefore, a natural approximation for u is $u^{(0)}$ given by

$$\begin{aligned} du^{(0)}(t,x) &= \Big(\frac{1}{2}\Big(a^2(t,x,Y(t))u^{(0)}(t,x)\Big)_{xx} - (b(t,x,Y(t))u^{(0)}(t,x))_x\Big)dt \\ &\quad + h(t,x,Y(t))u^{(0)}(t,x)dY(t) \\ u^{(0)}(0,x) &= \pi_0(x). \end{aligned} \tag{5.4}$$

Note that (5.4) is the Zakai equation for the model (5.1) wihh ε_0, with the exception that the original observation process Y is used.

We can now use the results of the previous section to derive a better approximation of u. To this end, for $\varepsilon \in \mathbb{R}$, define process u^ε by rhe following equation:

$$\begin{aligned} du^\varepsilon(t,x) &= \Bigg(\Big(\frac{a^2(t,x,Y(t)) + \varepsilon^2\rho^2(t,x,Y(t))}{2}\, u^\varepsilon\Big)_{xx} \\ &\quad - (b(t,x,Y(t))u^\varepsilon)_x\Bigg)dt \\ &\quad + \Big(h(t,x,Y(t))u^\varepsilon - \varepsilon\Big(\rho(t,x,Y(t))u^\varepsilon\Big)_x\Big)dY(t),\ t>0, \\ u^\varepsilon(2,x) &= \pi_0(x). \end{aligned} \tag{5.5}$$

so that $u = u^{\varepsilon_0}$.

Under assumptions (5.2) we have $u^\varepsilon \in \mathcal{H}^1_p(T)$ for every $T > 0$. By Theorem 4.1 we can define process $u^{(1)}$ by

$$\begin{aligned} du^{(2)}(t,x) &= \Big(\frac{1}{2}(a^2(t,x,Y(t))u^{(1)}(t,x))_{xx} - (b(t,x,Y(t))u^{(1)}(t,x))_x\Big)dt \\ &\quad + (h(t,x,Y(t))u^{(1)}(t,x) - (\rho(t,x,Y(t))u^{(0)}(t,x))_x dY(t); \\ u^{(1)}(0,x) &= 0. \end{aligned} \tag{5.6}$$

and more generally, for $n \geq 2$, process $u^{(n)}$

$$\begin{aligned} du^{(n)}(t,x) &= \Big(\frac{1}{2}\Big(a^2(t,x,Y(t))u^{(n)}(t,x)\Big)_{xx} - (b(t,x,Y(t))u^{(n)}(t,x))_x \\ &\quad + \frac{1}{2}\Big(\rho^2(t,x,Y(t))u^{(n-2)}(t,x)\Big)_{xx}\Big)dt \\ &\quad + \Big(h(t,x,Y(t))u^{(n)}(t,x) - \Big(\rho(t,x,Y(t))u^{(n-1)}(t,x)\Big)_x\Big)dY(t), \\ u^{(n)}(0,x) &= 0;\quad n \geq 2. \end{aligned} \tag{5.7}$$

According to Theorem 4.1 we have

$$u \approx u^{(0)} + \sum_{m=1}^{n} u^{(m)}\varepsilon_0^m, \tag{5.8}$$

and the error of approximation is of order ε_0^{n+1}, that is

$$\|u - u^{(0)} - \sum_{m=1}^{n} u^{(m)}\varepsilon_0^m\|_{\mathcal{H}^1_p(T)} \leq C\varepsilon_0^{n+1}. \tag{5.9}$$

The proof of Theorem 4.1 provides an explicit bound on the constant C. Furthermore, by embedding theorems,

$$E\sup_{o<t<T, x\in\mathbb{R}} |u(t,x) - u^{(0)}(t,x) - \sum_{m=1}^{n} u^{(m)}(t,x)\varepsilon_0^m|^p \leq C\varepsilon_0^{(n+1)p}. \tag{5.10}$$

Notice that, by the second part of Theorem 4.1, for sufficiently small ε we actually have $u^{\varepsilon} = u^{(0)} + \sum_{m=1}^{\infty} u^{(m)}\varepsilon^m$, and there is an explicit bound on the radius of convergence.

6. NONLINEAR EQUATION: HIGHER ORDER APPROXIMATION.

Just as with the linear equation, we could apply Theorem 3.1 repeatedly to a general nonlinear equation and get higher order terms in the expansion of u^{ε}. The obvious difficulty is that the formulas become very complicated very quickly. As an example, consider the one-dimensional equation driven by space-time white noise:

$$\begin{aligned} du^{\varepsilon}(t,x) &= (a(t,x)u^{\varepsilon}_{xx} + b(t,x)u^{\varepsilon}_x + f(t,x,u^{\varepsilon}))dt \\ &\quad + \varepsilon g(t,x,u^{\varepsilon})dW(t,x),\ t>0,\ x\in\mathbb{R},\ \varepsilon\in I \\ u^{\varepsilon}(0,x) &= \phi(x) \end{aligned} \tag{6.1}$$

Recall [14, Section 8.3] that this equation is of the type (2.5) with $g^k = gh^k$, where $\{h^k, k\geq 1\}$ is an orthonormal basis in $L_2(\mathbb{R})$, and for this equation Conditions 2.2.1–2.2.3 are replaced with the following.

CONDITION 6.6.1. *(Regularity of the coefficients)*
For every $x\in\mathbb{R}$, the functions a, b are predictable and there exist numbers $0<\kappa_1<\kappa_2$ so that, uniformly in (ω,t), $\|a\|_{B^2}+\|b\|_{B^1}\leq\kappa_2$ and, uniformly in (ω,t,x), $\kappa_1\leq a\leq\kappa_2$.

CONDITION 6.6.2. *(Regularity of free terms)*
The functions $f=f(t,x,y)$ and $g=g(t,x,y)$ satisfy $|f(t,x,y_1)-f(t,x,y_2)|+|g(t,x,y_1)-g(t,x,y_2)|\leq\kappa_2|y_1-y_2|$ uniformly in (ω,t,x), $g(\cdot,\cdot,0)\in\mathbb{L}_p(\tau)=\mathbb{H}^0_p(\tau)$ for some $p\geq 2$, and there is $\kappa\in(0,1/2]$ so that $f(\cdot,\cdot,0)\in\mathbb{H}^{-3/2-\kappa}_p(\tau)$.

It is know [14, Theorem 8.5] that, if Conditions 6.6.1 and 6.6.2 hold and if $\phi\in U_p^{-1/2-\kappa}$, $\tau\leq T$, then, for every $\varepsilon\in\mathbb{R}$, equation (6.3) has a unique solution $u^{\varepsilon}\in\mathcal{H}_p^{-1/2-\kappa}(\tau)$ and

$$\|u^{\varepsilon}\|_{\mathcal{H}_p^{-1/2-\kappa}(\tau)} \leq N\cdot\Big(\|f(\cdot,\cdot,0)\|_{\mathbb{H}_p^{-3/2-\kappa}(\tau)} + \varepsilon\|g(\cdot,\cdot,0)\|_{\mathbb{L}_p(\tau)} + \|\phi\|_{U_p^{-1/2-\kappa}}\Big) \tag{6.2}$$

with constant N depending only on $\kappa,\kappa_1,\kappa_2,p,T$.

The starting point in the analysis is the zero-order approximation. Define $u^{(0)}$ by

$$\begin{aligned} du^{(0)} &= (au^{(0)}_{xx} + bu^{(0)}_x + f(u^{(0)}))dt,\ t>0,\ x\in\mathbb{R}, \\ u^{(0)}(0,x) &= \phi(x) \end{aligned} \tag{6.3}$$

It is obvious that $\lim_{\varepsilon\to 0}\|u^{\varepsilon}-u^{(0)}\|_{\mathcal{H}_p^{-1/2-\kappa}(\tau)}=0$.

Suppose next that, for every (ω,t,x,y), the partial derivative

$$f^{(1)}(t,x,y) = \partial f(t,x,y)/\partial y$$

exists, is uniformly bounded and, for every (ω,t,x), continuous in y. Then we can define $u^{(1)}$ by

$$\begin{aligned} du^{(1)} &= (au^{(1)}_{xx} + bu^{(1)}_x + f^{(1)}(u^{(0)})u^{(1)})dt + g(u^{(0)})dW(t,x) \\ u^{(1)}(0,x) &= 0. \end{aligned} \tag{6.4}$$

Then, similar to the example at the end of Section 3, we conclude from Theorem 3.1 that $\lim_{\varepsilon\to 0}\varepsilon^{-1}\|u^{\varepsilon}-u^{(0)}-\varepsilon u^{(1)}\|_{\mathcal{H}_p^{-1/2-\kappa}(\tau)}=0$.

To define higher order terms in the expansion, we need an additional construction. Denote by $A_m(n)$ the set of n-vectors $\vec{a}=(a_1,\dots,a_n)$ with non-negative integer components so that $\sum_{i=1}^n a_i =$

m and $\sum_{i=1}^{n} ia_i = n$. Note that if $m = 1$, then $a_n = 1$ and $a_i = 0$ for $i = 1, \dots, n-1$, whereas if $m \geq 2$, then $a_n = 0$. It is known [15, page 36] that, for every sufficiently smooth function F,

$$\frac{1}{n!} \frac{d^n F\Big(\sum_{i\geq 0} c_i \varepsilon^i\Big)}{d\varepsilon^n}\Bigg|_{\varepsilon=0} = c_n F^{(1)}(c_0) + \sum_{m=2}^{n} F^{(m)}(c_0) \sum_{\vec{a}\in A_m(n)} \frac{m!}{a_1! \cdots a_{n-1}!} c_1^{a_1} \cdots c_{n-1}^{a_{n-1}}, \tag{6.5}$$

where $F^{(m)} = \frac{1}{m!}\frac{d^m F(t)}{dt^m}$.

Assume now that all partial derivatives of $f(t,x,y)$ and $g(t,x,y)$ with respect to y exist and each one is uniformly bounded. Then we set

$$f^{(n)}(y) = \frac{1}{n!}\frac{\partial^n f(t,x,y)}{\partial y^n}, \quad g^{(n)}(y) = \frac{1}{n!}\frac{\partial^n g(t,x,y)}{\partial y^n}, \tag{6.6}$$

and define process $u^{(n)}$, $n \geq 2$, as the solution of the equation

$$\begin{aligned} du^{(n)} &= \Big(au^{(n)}_{xx} + bu^{(n)}_{x} + f^{(1)}(u^{(0)})u^{(n)} \\ &+ \sum_{m=2}^{n} f^{(m)}(u^{(0)}) \sum_{\vec{a}\in A_m(n)} \frac{m!}{a_1!\cdots a_{n-1}!}(u^{(1)})^{a_1}\cdots(u^{(n-1)})^{a_{n-1}}\Big)dt \\ &+\Big(\sum_{m=1}^{n-1} g^{(m)}(u^{(0)}) \sum_{\vec{a}\in A_m(n-1)} \frac{m!}{a_1!\cdots a_{n-1}!}(u^{(1)})^{a_1}\cdots(u^{(n-1)})^{a_{n-1}}\Big)dW(t,x) \\ u^{(n)}(0,x) &= 0, \end{aligned} \tag{6.7}$$

To have the process $u^{(n)}$ well-defined, the right-hand side of (6.7) must be in $\mathbb{L}_p(\tau)$. Since this right-hand side consists of products of $u^{(m)}$, $m < n$, we need higher order integrability for $u^{(m)}$. One way to ensure this integrability is to require Condition 6.6.2 to hold *for all* $p \geq 2$.

THEOREM 6.1. *Fix* $\kappa \in (0, 1/2]$ *and assume that*

1.Conditions 6.6.1 and 6.6.2 hold for every $p \geq 2$.

2.For every $n \geq 1$, *n-th order partial derivatives of* f *and* g *with respect to the last argument exist and are uniformly bounded.*

Then, for every $n \geq 0$,

$$u^{\varepsilon} = u^{(0)} + \sum_{k=1}^{n} u^{(k)}\varepsilon^k + R^{\varepsilon}_n, \quad \|R^{\varepsilon}_n\|_{\mathcal{H}^{\gamma}_p(\tau)} \leq C|\varepsilon|^{n+1}. \tag{6.8}$$

Proof. By Theorem 3.1 we have $u^{(1)} = \frac{\partial u^{\varepsilon}}{\partial\varepsilon}\Big|_{\varepsilon=0}$. By changing ε to $\varepsilon + \varepsilon_0$, we conclude that, in fact, $u^{(1),\varepsilon} = \partial u^{\varepsilon}/\partial\varepsilon$ exists for every $\varepsilon \in \mathbb{R}$ and is the solution of

$$\begin{aligned} du^{(1),\varepsilon} &= (au^{(1)}_{xx} + bu^{(1)}_{x} + f^{(1)}(u^{\varepsilon})u^{(1),\varepsilon})dt + (\varepsilon g^{(1)}(u^{\varepsilon}) + g(u^{\varepsilon}))dW(t,x) \\ u^{(1),\varepsilon}(0,x) &= 0. \end{aligned} \tag{6.9}$$

After repeatedly applying Theorem 3.1, we see that $u^{(n),\varepsilon} = \frac{1}{n!}\frac{\partial^n u^\varepsilon}{\partial \varepsilon^n}$ is defined for every $n \geq 1$ and $\varepsilon \in \mathbb{R}$ and is the solution of

$$\begin{aligned} du^{(n),\varepsilon} &= \Big(a u^{(n),\varepsilon}_{xx} + b u^{(n),\varepsilon}_x + f^{(1)}(u^\varepsilon) u^{(n),\varepsilon} \\ &+ \sum_{m=2}^{n} f^{(m)}(u^\varepsilon) \sum_{\vec{a} \in A_m(n)} \frac{m!}{a_1! \cdots a_{n-1}!} (u^{(1),\varepsilon})^{a_1} \cdots (u^{(n-1),\varepsilon})^{a_{n-1}} \Big) dt \\ &+ \Big(\varepsilon g^{(1)} u^{(n),\varepsilon} + \varepsilon \sum_{m=2}^{n} g^{(m)}(u^\varepsilon) \sum_{\vec{a} \in A_m(n)} \frac{m!}{a_1! \cdots a_{n-1}!} \prod_{l=1}^{n-1} (u^{(l),\varepsilon})^{a_l} \\ &+ \sum_{m=1}^{n-1} g^{(m)}(u^\varepsilon) \sum_{\vec{a} \in A_m(n-1)} \frac{m!}{a_1! \cdots a_{n-1}!} \prod_{l=1}^{n-1} (u^{(l),\varepsilon})^{a_l} \Big) dW(t,x) \\ u^{(n),\varepsilon}|_{t=0} &= 0. \end{aligned} \tag{6.10}$$

Using (6.2) and the generalized Hölder inequality [6, p. 623] with $p_i = n/(ia_i)$, we conclude by induction that, for all $n \geq 1$,

$$\|u^{(n),\varepsilon}\|_{\mathcal{H}^\gamma_p(\tau)} \leq (1+\varepsilon) C \sum_{m=1}^{n-1} \sum_{\vec{a} \in A_m(n)} \prod_{i=1}^{n} \|u^{(i),\varepsilon}\|^{a_i}_{\mathbb{L}_{pk/i}(\tau)} \leq (1+\varepsilon) C, \tag{6.11}$$

where C depends on $\kappa, \kappa_1, \kappa_2, n, p, T$, and the corresponding norms of $f(t,x,0)$ and $g(t,x,0)$. Then (6.8) follows from the Taylor formula.

□

If conditions of Theorem 6.1 hold for every $\kappa \in (0, 1/2]$, then embedding theorems for the space $\mathcal{H}_p^{-1/2-\kappa}(\tau)$ [14, Theorem 7.2] imply that the remainder in (6.8) has the order $|\varepsilon|^{n+1}$ in the norm of the space $L_p(\Omega; C^{1/4-\delta, 1/2-\delta}_{t,x}(0,\tau))$ for every $\delta \in (0, 1/4)$; in particular,

$$E \sup_{0 \leq t \leq T, x \in \mathbb{R}} |u^\varepsilon(t \wedge \tau, x) - \sum_{m=0}^{n} u^{(m)}(t \wedge \tau, x) \varepsilon^m|^p \leq C |\varepsilon|^{(n+1)p}. \tag{6.12}$$

Even if the functions f and g are real analytic in the last variable, the $\mathcal{H}_p^{-1/2-\kappa}(\tau)$ norm of the remainder $R_n(\varepsilon)$ in (6.8) will not in general tend to zero as $n \to \infty$ for any $\varepsilon > 0$ due to the rapid growth of the moments of $W(t,x)$. This means that analyticity for nonlinear equation must be studied in a setting different from the one used for the linear equation. In [1], a similar analyticity question was studied for stochastic ordinary differential equations. It was proved that, *with probability one*, the solution of an ODE is an analytic function of the parameter as long as the coefficients are analytic. Almost sure convergence of $R_n(\varepsilon)$ in (6.8) as $n \to \infty$ remains an open problem.

REFERENCES

1. G. B. Arous. Flots et series de Taylor stochastiques. *Probability Theory and Related Fields*, 81:29–77, 1989.
2. C. Cardon-Weber. Large Deviations for a Burgers'-Type SPDE. *Stoch. Proc. Appl.*, 84:53–70, 1999.
3. F. Chenal and A. Millet. Uniform Large Deviations for Parabolic SPDEs and Applications. *Stoch. Proc. Appl.*, 72:161–186, 1997.
4. A. Dembo and O. Zeitouni. *Large Deviations Techniques and Applications.* Jones and Barlett Publishers, Boston, 1993.
5. J. D. Deuschel and D. W. Stroock. *Large Deviations.* Academic Press, 1989.
6. L. C. Evans. *Partial Differential Equations*, volume 19 of *Graduate Studies in Mathematics.* AMS, Providence, RI, 1998.
7. M. I. Freidlin and A. D. Wentzell. *Random Perturbation of Dynamical Systems.* Springer, 1984.

8. A. Friedman. *Stochastic Differential Equations and Applications, v. 2.* Academic Press, New York, 1976.

9. A. M. Il'in, R. Z. Khasminskii, and G. Yin. Singularly perturbed switching diffusions: rapid switchings and fast diffusions. *J. Optim. Theory Appl.*, 102(3):555–591, 1999.

10. G. Kallianpur and J. Xiong. Large Deviations for a Class of Stochastic Partial Differential Equations. *Ann. Probab.*, 24(1):320–345, 1996.

11. R. Z. Khasminskii and G. Yin. Asymptotic series for singularly perturbed Kolmogorov-Fokker-Planck equations. *SIAM J. Appl. Math.*, 56(6):1766–1793, 1996.

12. R. Z. Khasminskii and G. Yin. On transition densities of singularly perturbed diffusions with fast and slow components. *SIAM J. Appl. Math.*, 56(6):1794–1819, 1996.

13. R. Z. Khasminskii, G. Yin, and Q. Zhang. Asymptotic expansions of singularly perturbed systems involving rapidly fluctuating Markov chains. *SIAM J. Appl. Math.*, 56(1):277–293, 1996.

14. N. V. Krylov. An analytic approach to SPDEs. In B. L. Rozovskii and R. Carmona, editors, *Stochastic Partial Differential Equations. Six Perspectives, Mathematical Surveys and Monographs*, pages 185–242. AMS, 1999.

15. J. Riordan. *An introduction to combinatorial analysis.* Princeton University Press, Princeton, NJ, 1980.

16. B. L. Rozovskii. *Stochastic Evolution Systems.* Kluwer Academic Publishers, 1990.

17. R. B. Sowers. Large Deviations for a Reaction-Diffusion Equation with Non-Gaussian Perturbation. *Ann. Probab.*, 20(1):504–537, 1992.

18. D. W. Stroock. *An Introduction to the Theory of Large Deviations.* Springer, Berlin, 1984.

19. S. R. S. Varadhan. *Large Deviations and Applications.* SIAM, Philadelphia, 1984.

20. S. R. S. Varadhan. Asymptotic Probabilities and Differential Equations. *Comm. Pure Appl. Math.*, 19(3):261–286, 1966.

21. G. Yin and M. Kniazeva. Singularly perturbed multidimensional switching diffusions with fast and slow switchings. *J. Math. Anal. Appl.*, 229(2):605–630, 1999.