Stochastics and Stochastics Reports, Vol. 35, pp. 233–256
Reprints available directly from the publisher
Photocopying permitted by license only

TIME-DISCRETIZATION OF THE ZAKAI EQUATION FOR DIFFUSION PROCESSES OBSERVED IN CORRELATED NOISE*

PATRICK FLORCHINGER†

Université de Metz, Département de Mathématiques, URA CNRS 399, Ile du Saulcy, F-57045 Metz Cédex, France

FRANÇOIS LE GLAND

INRIA Sophia-Antipolis, Route des Lucioles, F-06565 Valbonne Cédex, France

(*Received 6 April 1990; in final form 23 November 1990*)

A time discretization scheme is provided for the Zakai equation, a stochastic PDE which gives the conditional density of a diffusion process observed in white-noise. The case where the observation noise and the state noise are correlated, is considered. The numerical scheme is based on a Trotter-like product formula, which exhibits *prediction* and *correction* steps, and for which an error estimate of order δ is proved, where δ is the time discretization step. The *correction* step is associated with a degenerate second-order stochastic PDE, for which a representation result in terms of stochastic characteristics has been proved by Krylov–Rozovskii [13] and Kunita [15, 17]. A discretization scheme is then provided to approximate these stochastic characteristics. Under the additional assumption that the correlation coefficient is constant, an error estimate of order $\sqrt{\delta}$ is proved for the overall numerical scheme. This has been proved to be the best possible error estimate by Elliott–Glowinski [7].

KEY WORDS: Diffusion processes, correlated noises, nonlinear filtering, Zakai equation, stochastic PDE, stochastic characteristics, time discretization.

1. INTRODUCTION

The purpose of this paper is to present a *computable* time discretization scheme for the Zakai equation of nonlinear filtering with correlated noises, and to provide an estimate of the rate of convergence.

In the case of independent noises, the problem has been studied by Kushner [18], Newton [21], Korezlioglu–Mazziotto [11], Bennaton [1], DiMasi–Pratelli–Runggaldier [6], Picard [22], Bensoussan–Glowinski–Rascanu [2] and Le Gland [20]. Some of these authors have actually considered the associated Zakai equation. Time discretization schemes have been provided with a rate of convergence of order δ, where δ is the time discretization step.

In the case of correlated noises, the problem has been studied by Elliott–Glowinski [7]. The best approximation of the continuous filter based on the

*Research partially supported by USACCE under Contract DAJA45-90-C-0008.
†Also: INRIA Lorraine, CESCOM, Technopole de Metz 2000, 4 rue Marconi, F-57070 Metz, France.

values of the observation process at a regular partition (with mesh δ) has been considered, and it has been proved that the rate of convergence is of order $\sqrt{\delta}$. However, no algorithm is provided to actually *compute* this approximation.

The paper is organized as follows. In Section 2, the nonlinear filtering problem is presented. Some results on the Zakai equation, and on a related degenerate second-order stochastic PDE, are recalled in Section 3. A Trotter-like product formula is then considered, with an error estimate of order δ. However, this numerical scheme is not *computable.* In Section 4, a representation result in terms of *stochastic characteristics* is presented for the degenerate second-order stochastic PDE. This part follows mainly the work of Krylov–Rozovskii [13]—see also Kunita [15, 17]. A time discretization scheme is presented in Section 5, based on an approximation of the stochastic characteristics. Under the additional assumption that the correlation coefficient is constant, an error estimate of order $\sqrt{\delta}$ can be proved. In addition, this numerical scheme is actually *computable*, as far as time discretization is concerned, i.e. up to space discretization.

2. THE FILTERING PROBLEM

Consider the following stochastic differential system, defined on the probability space $(\Omega, \mathscr{F}, P)$

$$dX_t = b(X_t)\,dt + \sigma(X_t)\,dW_t + \rho(X_t)\,dV_t$$

$$dY_t = h(X_t)\,dt + dV_t$$

where the non observed component $\{X_t t \geqq 0\}$ takes values in $\mathbf{R}^m$, and the observation $\{Y_t, t \geqq 0\}$ takes values in $\mathbf{R}^d$. $\{W_t, t \geqq 0\}$ and $\{V_t, t \geqq 0\}$ are independent Wiener processes of appropriate dimension, with covariance matrix I (identity) and r respectively. For the clarity of exposition, it is assumed throughout the paper that $r = I$. In addition, the random variable X_0 is independent of the Wiener processes, with probability distribution $p_0(x)\,dx$.

Throughout the paper, it is assumed that the coefficients, b, σ, ρ and h are globally Lipschitz continuous functions defined on $\mathbf{R}^m$, so that the stochastic differential system has a unique strong solution. The following definitions are used: $a \triangleq \sigma\sigma^*$ and $c \triangleq \rho\rho^*$. In particular, it is not assumed that either a or c is uniformly elliptic.

With the diffusion process $\{X_t, t \geqq 0\}$ are associated the two partial differential operators

$$L \triangleq \frac{1}{2} \sum_{i,j=1}^{m} [a^{i,j} + c^{i,j}] \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{m} b^i \frac{\partial}{\partial x_i},$$

$$L_0 \triangleq \frac{1}{2} \sum_{i,j=1}^{m} a^{i,j} \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{m} b^i \frac{\partial}{\partial x_i}.$$

Another family of partial differential operators to be considered is

$$B_k \triangleq h_k + \sum_{i=1}^{m} \rho_k^i \frac{\partial}{\partial x_i}, \quad 1 \leqq k \leqq d.$$

Introducing

$$Z_t^s \triangleq \exp\left\{\int_s^t h^*(X_\tau)\, dY_\tau - \frac{1}{2}\int_s^t |h(X_\tau)|^2\, d\tau\right\}, \quad Z_t \triangleq Z_t^0,$$

it is standard that, for all $T > 0$ the original probability measure P is equivalent on $[0, T]$ to the *reference probability* measure $P\dagger$ with Radon–Nikodym derivative Z_T, so that under $P\dagger$

$$dX_t = b(X_t)\, dt + \sigma(X_t)\, dW_t + \rho(X_t)[dY_t - h(X_t)\, dt], \tag{2.1}$$

where $\{W_t, t \geqq 0\}$ and $\{Y_t, t \geqq 0\}$ are independent Wiener processes, with covariance matrix I (identity), and the random variable X_0 is independent of the Wiener processes, with probability distribution $p_0(x)\, dx$.

The Bayes formula gives

$$\mathbf{E}(f(X_t) \mid \mathscr{Y}_t) = \frac{\mathbf{E}\dagger(f(X_t)Z_t \mid \mathscr{Y}_t)}{\mathbf{E}\dagger(Z_t \mid \mathscr{Y}_t)},$$

and in addition

$$\mathbf{E}\dagger(f(X_t)Z_t \mid \mathscr{Y}_t) = \int f(x) p_t(x)\, dx,$$

where the unnormalized conditional density $\{p_t, t \geqq 0\}$ satisfies the Zakai equation [25]

$$dp_t = L^* p_t\, dt + \sum_{k=1}^{d} B_k^* p_t\, dY_t^k. \tag{2.2}$$

Consider then the following decomposition of the Zakai equation (2.2)

$$dp_t = L_0^* p_t\, dt + \Lambda^* p_t\, dt + \sum_{k=1}^{d} B_k^* p_t\, dY_t^k,$$

where

$$\Lambda \triangleq L - L_0 = \frac{1}{2}\sum_{i,j=1}^{m} c^{i,j} \frac{\partial^2}{\partial x_i\, \partial x_j}.$$

On one hand, the partial differential operator L_0 generates a strongly continuous semigroup $\{P_t, t \geqq 0\}$. On the other hand, it is possible to associate a stochastic semigroup $\{Q_t^s, 0 \leqq s \leqq t\}$ with the following degenerate second-order stochastic PDE

$$dp_t = \Lambda^* p_t\, dt + \sum_{k=1}^{d} B_k^* p_t\, dY_t^k, \tag{2.3}$$

which is studied below. Therefore, it is worth studying the following Trotter-like product formulas for approximating the original Zakai equation (2.2)

$$\begin{aligned} \bar{p}_{i+1} &= P_{\delta_i}^* Q_{t_{i+1}}^{t_i} \bar{p}_i, \\ \bar{p}_{i+1} &= Q_{t_{i+1}}^{t_i} P_{\delta_i}^* \bar{p}_i, \end{aligned} \tag{2.4}$$

where $\delta_i \triangleq t_{i+1} - t_i$, and $0 = t_0 < t_1 < \cdots < t_i < \cdots$
The main interest of such product formulas is that the original equation has been split into a second-order deterministic PDE (*prediction* step), and a degenerate second-order stochastic PDE (*correction* step). In the case of independent noises, this stochastic PDE reduces to a zero-order equation, for which there exists a straightforward explicit solution. In the case of correlated noises, a representation result is available by the method of stochastic characteristics (i.e. involving the stochastic flow of diffeomorphism associated with a SDE driven by the observation process), see Krylov–Rozovskii [13] and Section 4 below.

Remark 2.1 A similar prediction-correction numerical scheme was obtained by Kushner [18] in the case of independent noises.

Remark 2.2 Written in Stratonovich form, equation (2.3) is a first-order stochastic PDE. For such an equation, one can use the representation result of Kunita [15, 17], and translate the stochastic characteristics equations from Stratonovich form back to Itô form, to recover the representation result of [13].

As a consequence of the above discussion, there will be two steps in designing the approximation to the original Zakai equation (2.2)

- first use a Trotter-like product formula,
- then approximate the solution of the degenerate second-order stochastic PDE, by approximating the stochastic flow of diffeomorphisms involved in the stochastic characteristics method of [13].

It will be proved that the first step can be achieved with a rate of convergence of order δ, whereas the rate of convergence for the second step (and *a fortiori* for the global approximation procedure) is of order $\sqrt{\delta}$ only, where $\delta \triangleq \max_{i \geq 0} \delta_i$.

3. TROTTER-LIKE PRODUCT FORMULA

For all $n \geqq 0$, let H^n denote the space of real-valued Lebesgue-measurable functions on $\mathbf{R}^m$ whose generalized derivatives up to order n are square-integrable, with norm $\|\cdot\|_n$

$$\|u\|_n^2 \triangleq \sum_{0 \leq |\alpha| \leq n} \int |D^\alpha u(x)|^2 \, dx < \infty.$$

In addition, the following shorthand notations will be used throughout the paper: $|\cdot| \triangleq \|\cdot\|_0$ and $\|\cdot\| \triangleq \|\cdot\|_1$.

The beginning of this section is devoted to recall existence, uniqueness and regularity results for the Zakai equation

$$dp_t = L^* p_t \, dt + \sum_{k=1}^{d} B_k^* p_t \, dY_t^k, \tag{3.1}$$

and the degenerate second-order stochastic PDE

$$dp_t = \Lambda^* p_t \, dt + \sum_{k=1}^{d} B_k^* p_t \, dY_t^k, \tag{3.2}$$

with semigroup $\{Q_t^s, 0 \leq s \leq t\}$.

Although no coercivity hypothesis is satisfied, the following result is proved in Krylov–Rozovskii [13].

THEOREM 3.1 *Let $n \geq 1$ be fixed. Assume that*

- *a and c have bounded derivatives up to order* $\max(n, 2)$,
- *b, ρ and h have bounded derivatives up to order n,*
- *the initial condition satisfies $p_0 \in H^n$.*

Then both Eqs. (3.1) and (3.2) have a unique solution $p \in M^2(0, T; H^n)$. In addition

$$p \in L^2(\Omega; C_w([0, T]; H^n)),$$

and the following estimate holds

$$\mathbf{E}^\dagger \left[\sup_{0 \leq t \leq T} \|p_t\|_n^2 \right] \leq \|p_0\|_n^2 \, e^{CT}.$$

Similarly, for the Fokker–Planck equation

$$p_t' = L_0^* p_t, \tag{3.3}$$

and the following deterministic PDE associated with (3.2)

$$p_t' = \Lambda^* p_t, \tag{3.4}$$

with semigroup $\{P_t^*, t \geqq 0\}$ and $\{T_t^*, t \geqq 0\}$ respectively, it holds

THEOREM 3.2 *Let $n \geqq 1$ be fixed. Assume that*

- *a and c have bounded derivatives up to order* $\max(n, 2)$,
- *b has bounded derivatives up to order n,*
- *the initial condition satisfies $p_0 \in H^n$.*

Then both Eqs. (3.3) and (3.4) have a unique solution $p \in L^2(0, T; H^n)$. In addition

$$p \in C_w([0, T]; H^n),$$

and the following estimate holds

$$\sup_{0 \leqq t \leqq T} \|p_t\|_n^2 \leqq \|p_0\|_n^2 \, e^{CT}.$$

Remark 3.3 In the case where the coefficients a and c are uniformly elliptic, a slightly stronger theorem holds, see Krylov–Rozovskii [12] and Pardoux [23].

□ *Error Estimate*

The purpose here is to study one of the Trotter-like product formulas (2.4).

THEOREM 3.4 *Consider the following approximation scheme*

$$\bar{p}_{i+1} = P_{\delta_i}^* Q_{t_{i+1}}^{t_i} \bar{p}_i. \tag{3.5}$$

Assume that

- *a, c, b, ρ, and h have bounded derivatives up to order* 3,
- *the initial condition satisfies $p_0 \in H^3$.*

Then $\bar{p}_i$ approximates the solution p_{t_i} of the original Zakai equation (3.1) with a rate of convergence of order δ. Indeed

$$\{\mathbf{E}^\dagger |\bar{p}_i - p_{t_i}|^2\}^{1/2} \leqq C\delta \|p_0\|_3.$$

Proof The idea is to get an equation for $v_t \triangleq P_{t-s}^* Q_t^s \phi$ with ϕ smooth enough, that is similar to the original Zakai equation for p_t, except for some perturbation terms which have to be estimated. This gives an estimate of the one-step error, and the global estimate is obtained using the Gronwall lemma.

Differentiating with respect to t

$$dv_t = L_0^* v_t \, dt + P_{t-s}^* \left[\Lambda^* Q_t^s \phi \, dt + \sum_{k=1}^{d} B_k^* Q_t^s \phi \, dY_t^k \right]$$

$$= L_0^* v_t\, dt + \Lambda^* v_t\, dt + \sum_{k=1}^{d} B_k^* v_t\, dY_t^k$$

$$+ [P_{t-s}^* \Lambda^* - \Lambda^* P_{t-s}^*] Q_t^s \phi\, dt + \sum_{k=1}^{d} [P_{t-s}^* B_k^* - B_k^* P_{t-s}^*] Q_t^s \phi\, dY_t^k$$

$$= L^* v_t\, dt + \sum_{k=1}^{d} B_k^* v_t\, dY_t^k + f_t\, dt + \sum_{k=1}^{d} g_t^k\, dY_t^k,$$

where the perturbation terms are defined by

$$f_t \triangleq [P_{t-s}^* \Lambda^* - \Lambda^* P_{t-s}^*] Q_t^s \phi \quad \text{and} \quad g_t^k \triangleq [P_{t-s}^* B_k^* - B_k^* P_{t-s}^*] Q_t^s \phi,$$

respectively. The difference $\varepsilon_t \triangleq v_t - p_t$ satisfies

$$d\varepsilon_t = L^* \varepsilon_t\, dt + \sum_{k=1}^{d} B_k^* \varepsilon_t\, dY_t^k + f_t\, dt + \sum_{k=1}^{d} g_t^k\, dY_t^k.$$

Using estimates of [13]

$$\mathbf{E}^\dagger |\varepsilon_t|^2 \leqq \left[\mathbf{E}^\dagger |\varepsilon_s|^2 + C\mathbf{E}^\dagger \int_s^t |f_\tau|^2\, d\tau + C\mathbf{E}^\dagger \sum_{k=1}^{d} \int_s^t \|g_\tau^k\|^2\, d\tau \right] e^{C(t-s)}.$$

Assume that the following estimates hold

$$\mathbf{E}^\dagger |f_\tau|^2 \leqq C(\tau - s)^2 \mathbf{E}^\dagger \|\phi\|_3^2\, e^{C(\tau - s)}, \tag{3.6}$$

$$\mathbf{E}^\dagger \|g_\tau^k\|^2 \leqq C(\tau - s)^2 \mathbf{E}^\dagger \|\phi\|_3^2\, e^{C(\tau - s)}. \tag{3.7}$$

Then the Gronwall lemma would yield

$$\mathbf{E}^\dagger |\varepsilon_t|^2 \leqq [\mathbf{E}^\dagger |\varepsilon_s|^2 + C(t-s)^3 \mathbf{E}^\dagger \|\phi\|_3^2]\, e^{C(t-s)},$$

provided $\phi \in L^2(\Omega; H^3)$. Now, it follows from the assumptions and from Theorem 3.1, that $\bar{p}_i \in L^2(\Omega; H^3)$ for all i, so that setting $s = t_i$, $t = t_{i+1}$ and $\phi = \bar{p}_i$

$$\mathbf{E}^\dagger |\bar{p}_{i+1} - p_{t_{i+1}}|^2 \leqq [\mathbf{E}^\dagger |\bar{p}_i - p_{t_i}|^2 + C(t_{i+1} - t_i)^3 \mathbf{E}^\dagger \|\bar{p}_i\|_3^2]\, e^{C(t_{i+1} - t_i)},$$

and the result follows from the discrete Gronwall lemma. The end of the proof is devoted to proving estimates (3.6) and (3.7).

□ *Estimate (3.6)*

The following perturbation result

$$[P^*_{\tau-s}\Lambda^* - \Lambda^* P^*_{\tau-s}]u = \int_s^\tau P^*_{\tau-\tau'}[L^*_0\Lambda^* - \Lambda^* L^*_0]P^*_{\tau'-s}u\,d\tau',$$

holds for u smooth enough. It follows from the assumptions, that the partial differential operator $D \triangleq [L^*_0\Lambda^* - \Lambda^* L^*_0]$ is bounded from H^3 to H^0. In addition, it follows from Theorem 3.2 that $\{P^*_t, t \geqq 0\}$ is a strongly continuous semigroup in both H^0 and H^3. Therefore

$$|f_\tau| \leqq \int_s^\tau |P^*_{\tau-\tau'} D P^*_{\tau'-s} Q^s_t \phi|\,d\tau' \leqq C(\tau-s)\|Q^s_t\phi\|_3\, e^{C(\tau-s)}.$$

Then

$$\mathbf{E}^\dagger|f_\tau|^2 \leqq C(\tau-s)^2\mathbf{E}^\dagger\|Q^s_t\phi\|^2_3\, e^{C(\tau-s)} \leqq C(\tau-s)^2\mathbf{E}^\dagger\|\phi\|^2_3\, e^{C(\tau-s)}.$$

□ *Estimate (3.7)*

Similarly, the following perturbation result

$$[P^*_{\tau-s}B^*_k - B^*_k P^*_{\tau-s}]u = \int_s^\tau P^*_{\tau-\tau'}[L^*_0 B^*_k - B^*_k L^*_0]P^*_{\tau'-s}u\,d\tau',$$

holds for u smooth enough. It follows from the assumptions, that the partial differential operator $D_k \triangleq [L^*_0 B^*_k - B^*_k L^*_0]$ is bounded from H^3 to H^1. In addition, it follows from Theorem 3.2 that $\{P^*_t, t \geqq 0\}$ is a strongly continuous semigroup in both H^1 and H^3. Therefore

$$\|g^k_\tau\| \leqq \int_s^\tau \|P^*_{\tau-\tau'} D_k P^*_{\tau'-s} Q^s_t\phi\|\,d\tau' \leqq C(\tau-s)\|Q^s_t\phi\|_3\, e^{C(\tau-s)}.$$

Then

$$\mathbf{E}^\dagger\|g^k_\tau\|^2 \leqq C(\tau-s)^2\mathbf{E}^\dagger\|Q^s_t\phi\|^2_3\, e^{C(\tau-s)} \leqq C(\tau-s)^2\mathbf{E}^\dagger\|\phi\|^2_3\, e^{C(\tau-s)}. \qquad \square$$

Remark 3.5 In the case where the coefficient a is uniformly elliptic, the same error estimate can be proved under weaker regularity assumptions on the coefficients and the initial condition, see Florchinger–LeGland [8].

Remark 3.6 It is possible to approximate the stochastic differential equation (2.1), in such a way that the approximation $\bar{p}_i$ given by (2.4), is actually the conditional density of the approximate process at time t_i, given the observations $\mathscr{Y}_{t_i}$. This problem will be addressed elsewhere.

The approximation scheme (3.5) is not yet *computable*. First, the Fokker–Planck equation (3.3) with semigroup $\{P^*_t, t \geqq 0\}$, has to be approximated: this is a rather standard problem, for which one can use e.g. the backward Euler scheme, or some other approximation scheme. On the other hand, some representation results are presented in the next section, which can be used for the approximation of the degenerate second-order stochastic PDE (3.2), with semigroup $\{Q_t^s, 0 \leqq s \leqq t\}$.

4. STOCHASTIC CHARACTERISTICS

Parallel to the decomposition of the stochastic PDE (2.2), there is a similar decomposition for the stochastic differential equation (2.1). With the first component

$$dX_t = b(X_t)\,dt + \sigma(X_t)\,dW_t,$$

is associated the partial differential operator L_0 and the Fokker–Planck equation (3.3). It is proved below that the second component

$$dX_t = \rho(X_t)[dY_t - h(X_t)\,dt], \tag{4.1}$$

is associated with the degenerate second-order stochastic PDE (3.2) and the corresponding deterministic PDE (3.4).

The beginning of this section is devoted to recall results concerning the stochastic flow of diffeomorphisms associated with the stochastic differential equation (4.1).

THEOREM 4.1 *Let $\xi_{s,t}(\cdot)$ be the stochastic flow associated with the forward stochastic differential equation*

$$d\xi_t = \rho(\xi_t)[dY_t - h(\xi_t)\,dt]. \tag{4.2}$$

Assume that the coefficients h and ρ have bounded derivatives up to order $(n+1)$. Then $\xi_{s,t}(\cdot)$ is a C^n-diffeomorphism in $\mathbf{R}^m$.

Under the assumption that the coefficient ρ has bounded derivatives up to order 2, the inverse map $\xi_{s,t}^{-1}(\cdot)$ is given explicitly as the (backward) stochastic flow $\eta_{t,s}(\cdot)$ associated with the backward stochastic differential equation

$$d\eta_t = \rho(\eta_t) \oplus [dY_t - h(\eta_t)\,dt] - \rho_0(\eta_t)\,dt, \tag{4.3}$$

with

$$\rho_0^i \triangleq \sum_{k=1}^{d} \sum_{j=1}^{m} \frac{\partial \rho_k^i}{\partial x_j} \rho_k^j, \quad 1 \leqq i \leqq m.$$

The regularity of $\xi_{s,t}(\cdot)$ was first proved by Blagoveschenskii–Freidlin [3], whereas the rest of the theorem is proved in Kunita [16].

PROPOSITION 4.2 *The Jacobian $J_{s,t}(\cdot)$ (i.e. the determinant of the Jacobian matrix) of the diffeomorphism $\xi_{s,t}(\cdot)$ satisfies*

$$J_{s,t}(x) \triangleq \exp\left\{\int_s^t \alpha^*(\xi_{s,\tau}(x))[dY_\tau - h(\xi_{s,\tau}(x))\,d\tau]\right.$$

$$\left. - \int_s^t h_0(\xi_{s,\tau}(x))\,d\tau - \int_s^t \bar{\alpha}(\xi_{s,\tau}(x))\,d\tau\right\} \tag{4.4}$$

with

$$\alpha_k \triangleq \sum_{i=1}^m \frac{\partial \rho_k^i}{\partial x_i} = \operatorname{div} \rho_k, \quad 1 \leqq k \leqq d$$

$$\bar{\alpha} \triangleq \frac{1}{2} \sum_{k=1}^d \sum_{i,j=1}^m \frac{\partial \rho_k^i}{\partial x_j} \frac{\partial \rho_k^j}{\partial x_i} \quad \text{and} \quad h_0 \triangleq \sum_{k=1}^d \sum_{i=1}^m \frac{\partial h_k}{\partial x_i} \rho_k^i.$$

Proof Transform first the stochastic differential equation (4.2) into Stratonovich form

$$d\xi_t = \rho(\xi_t) \circ [dY_t - h(\xi_t)\,dt] - \tfrac{1}{2}\rho_0(\xi_t)\,dt.$$

Similarly to the Liouville formula for ordinary differential equations, see Hartman [10], it holds

$$d \log J_{s,t}(x) = \alpha^*(\xi_{s,t}(x)) \circ [dY_t - h(\xi_{s,t}(x))\,dt] - h_0(\xi_{s,t}(x))\,dt - \tfrac{1}{2} \operatorname{div} \rho_0(\xi_{s,t}(x))\,dt.$$

Transforming back to Itô form

$$d \log J_{s,t}(x) = \alpha^*(\xi_{s,t}(x))[dY_t - h(\xi_{s,t}(x))\,dt] - h_0(\xi_{s,t}(x))\,dt$$

$$- \tfrac{1}{2} \operatorname{div} \rho_0(\xi_{s,t}(x))\,dt + \tfrac{1}{2}\alpha_0(\xi_{s,t}(x))\,dt.$$

Now it holds

$$\operatorname{div} \rho_0 = \sum_{k=1}^d \sum_{i,j=1}^m \frac{\partial}{\partial x_i}\left(\frac{\partial \rho_k^i}{\partial x_j} \rho_k^j\right),$$

$$\alpha_0 \triangleq \sum_{k=1}^d \sum_{i=1}^m \frac{\partial \alpha_k}{\partial x_i} \rho_k^i = \sum_{k=1}^d \sum_{i,j=1}^m \frac{\partial^2 \rho_k^j}{\partial x_i \partial x_j} \rho_k^i,$$

which finishes the proof. □

Remark 4.3 Note that $[J_{s,t}(\eta_{t,s}(\cdot))]^{-1}$ is actually the Jacobian of the inverse diffeomorphism $\eta_{t,s}(\cdot)$.

Define

$$\Xi_{s,t}(x) \triangleq \exp\left\{\int_s^t h^*(\xi_{s,\tau}(x))\,dY_\tau - \frac{1}{2}\int_s^t |h(\xi_{s,\tau}(x))|^2\,d\tau\right\}, \tag{4.5}$$

and

$$\Theta_{s,t}(x) \triangleq \Xi_{s,t}(x)[J_{s,t}(x)]^{-1} = \exp\left\{\int_s^t h^*(\xi_{s,\tau}(x))\,dY_\tau - \frac{1}{2}\int_s^t |h(\xi_{s,\tau}(x))|^2\,d\tau\right.$$

$$\left. - \int_s^t \alpha^*(\xi_{s,\tau}(x))[dY_\tau - h(\xi_{s,\tau}(x))\,d\tau] + \int_s^t h_0(\xi_{s,\tau}(x))\,d\tau + \int_s^t \bar{\alpha}(\xi_{s,\tau}(x))\,d\tau\right\}.$$

Introduce the following *definition*

$$Q_t^s q(x) \triangleq q(\eta_{t,s}(x))\Theta_{s,t}(\eta_{t,s}(x)), \tag{4.6}$$

or equivalently

$$Q_t^s q(\xi_{s,t}(x)) = q(x)\Theta_{s,t}(x).$$

where the same notation has been used as in the previous section. This will be justified by the Theorem 4.8 to be proved below.

Remark 4.4 Under the additional assumption that the coefficient ρ has bounded derivatives up to order 2, the Lemma 6.2 of [16, Chapter 2] gives the following explicit expressions in terms of backward Itô stochastic integrals

$$\Xi_{s,t}(\eta_{t,s}(x)) = \exp\left\{\int_s^t h^*(\eta_{t,\tau}(x)) \oplus dY_\tau - \frac{1}{2}\int_s^t |h(\eta_{t,\tau}(x))|^2\,d\tau - \int_s^t h_0(\eta_{t,\tau}(x))\,d\tau\right\},$$

$$J_{s,t}(\eta_{t,s}(x)) = \exp\left\{\int_s^t \alpha^*(\eta_{t,\tau}(x)) \oplus [dY_\tau - h(\eta_{t,\tau}(x))\,d\tau]\right.$$

$$\left. - \int_s^t h_0(\eta_{t,\tau}(x))\,d\tau - \int_s^t \bar{\alpha}(\eta_{t,\tau}(x))\,d\tau - \int_s^t \alpha_0(\eta_{t,\tau}(x))\,d\tau\right\},$$

where the coefficients h_0 and α_0 have already been defined as

$$h_0 \triangleq \sum_{k=1}^d \sum_{i=1}^m \frac{\partial h_k}{\partial x_i}\rho_k^i \quad \text{and} \quad \alpha_0 \triangleq \sum_{k=1}^d \sum_{i=1}^m \frac{\partial \alpha_k}{\partial x_i}\rho_k^i = \sum_{k=1}^d \sum_{i,j=1}^m \frac{\partial^2 \rho_k^j}{\partial x_i \partial x_j}\rho_k^i.$$

Therefore

$$\Gamma_{t,s}(x) \triangleq \Theta_{s,t}(\eta_{t,s}(x)) = \exp\left\{\int_s^t h^*(\eta_{t,\tau}(x)) \oplus dY_\tau - \frac{1}{2}\int_s^t |h(\eta_{t,\tau}(x))|^2 \, d\tau\right.$$

$$\left. - \int_s^t \alpha^*(\eta_{t,\tau}(x)) \oplus [dY_\tau - h(\eta_{t,\tau}(x))\, d\tau] + \int_s^t \bar{\alpha}(\eta_{t,\tau}(x))\, d\tau + \int_s^t \alpha_0(\eta_{t,\tau}(x))\, d\tau\right\}. \quad (4.7)$$

Remark 4.5 If $\rho \equiv 0$, then $\xi_{s,t}(x) = x$ so that

$$Q_t^s q(x) = q(x) \exp\{h^*(x)(Y_t - Y_s) - \tfrac{1}{2}|h(x)|^2(t-s)\},$$

which is actually the explicit solution of the equation

$$dq_t = \sum_{k=1}^{d} h_k q_t \, dY_t^k,$$

with initial condition q at time s. In this case, (2.4) reduces to the discretization schemes considered in [2, 18, 20].

First, the following *stability* result holds

PROPOSITION 4.6 *Let $n \geqq 0$ be fixed. Assume that*

- *c, ρ and h have bounded derivatives up to order $(n+1)$,*
- *the initial condition satisfies $q \in H^n$.*

Then $Q_t^s q$ is a square integrable random variable with values in H^n. In addition, the following estimate holds

$$\{\mathbf{E}^\dagger \|Q_t^s q\|_n^2\}^{1/2} \leqq \|q\|_n e^{C(t-s)}.$$

Proof It is enough to prove the result for $n = 0$.

Using the change of variable $x = \eta_{t,s}(y)$ i.e. $y = \xi_{s,t}(x)$

$$\mathbf{E}^\dagger |Q_t^s q|^2 = \mathbf{E}^\dagger \int [|q(\eta_{t,s}(y))| \Theta_{s,t}(\eta_{t,s}(y))]^2 \, dy = \int |q(x)|^2 \mathbf{E}\{\Theta_{s,t}(x)\} \, dx,$$

and the result follows from the estimate

$$\sup_{x \in \mathbf{R}^m} \mathbf{E}\{\Theta_{s,t}(x)\} \leqq e^{C(t-s)}. \qquad \square$$

Another property of the two-parameter stochastic semigroup $\{Q_t^s, 0 \leqq s \leqq t\}$ is provided by the following

PROPOSITION 4.7 *Let $\{T_t, t \geqq 0\}$ be the semigroup generated by*

$$\Lambda = \frac{1}{2}\sum_{i,j=1}^{m} c^{i,j} \frac{\partial^2}{\partial x_i \partial x_j}.$$

Then

$$\mathbf{E}^{\dagger}Q_t^s = T_{t-s}^*.$$

Proof Using the same change of variable as in the proof of Proposition 4.6, it holds

$$(\mathbf{E}^{\dagger}(Q_t^s q), f) = \mathbf{E}^{\dagger} \int q(\eta_{t,s}(y))\Theta_{s,t}(\eta_{t,s}(y)) f(y)\, dy = \int q(x)\mathbf{E}[f(\xi_{s,t}(x))]\, dx,$$

for any test-function f. Now, under the original probability measure P

$$d\xi_t = \rho(\xi_t)\, dV_t,$$

where $\{V_t, t \geqq 0\}$ is a Wiener process with covariance matrix I. Therefore

$$(\mathbf{E}^{\dagger}(Q_t^s q), f) = (q, T_{t-s} f) = (T_{t-s}^* q, f). \qquad \square$$

The following representation result of the solution of Eq. (3.2) in terms of the *stochastic characteristics* $\eta_{t,s}(\cdot)$, is the stochastic counterpart of the usual method of characteristics for linear first-order PDE. It has been proved by Krylov–Rozovskii [13] and Kunita [15, 17].

THEOREM 4.8 *Let $\{Q_t^s, s \leqq t\}$ be defined by (4.6). Then, the unique solution of equation (3.2) satisfies*

$$q_t(x) = Q_t^s q_s(x). \tag{4.8}$$

Proof The proof given below is essentially that of [13]. Introduce

$$\zeta_t^s \triangleq \exp\left\{\int_s^t \phi_\tau^* \, dY_\tau - \frac{1}{2}\int_s^t |\phi_\tau|^2 \, d\tau\right\},$$

where $\{\phi_\tau, s \leqq \tau \leqq t\}$ is deterministic.

It follows from the Itô formula that $\bar{q}_\tau \triangleq \mathbf{E}^{\dagger}(\zeta_\tau^s \cdot q_\tau)$ satisfies

$$\bar{q}'_\tau = \Lambda^* \bar{q}_\tau + \sum_{k=1}^d B_k^* \bar{q}_\tau \phi_\tau^k, \tag{4.9}$$

with the initial condition $\bar{q}_s = \mathbf{E}^{\dagger}(q_s)$. On the other hand, define

$$\bar{w}_\tau(x) \triangleq \mathbf{E}^{\phi}\left[f(\xi_{\tau,t}(x)) \exp\left\{\int_\tau^t \phi_{\tau'}^* h(\xi_{\tau,\tau'}(x))\, d\tau'\right\}\right],$$

where under the probability P^{ϕ}

$$d\xi_\tau = \rho(\xi_\tau)[dV_\tau^\phi + \phi_\tau\, d\tau],$$

and $\{V_t^\phi, t \geqq 0\}$ is a Wiener process with covariance matrix I. By the Feynman–Kac formula, $\{\bar{w}_\tau, s \leqq \tau \leqq t\}$ satisfies a PDE which is dual to (4.9), so that $(\bar{q}_t, f) = (\bar{q}_s, \bar{w}_s)$. Consider now the right-hand side in the representation result (4.8). Then

$$\zeta_t^s \cdot Q_t^s q_s(x) = q_s(\eta_{t,s}(x))\Xi_{s,t}^\phi(\eta_{t,s}(x)) \exp\left\{\int_s^t \phi_\tau^* h(\eta_{t,\tau}(x))\, d\tau\right\}[J_{s,t}(\eta_{t,s}(x))]^{-1},$$

with

$$\Xi_{s,t}^\phi(x) \triangleq \exp\left\{\int_s^t [h(\xi_{s,\tau}(x)) + \phi_\tau]^*\, dY_\tau - \frac{1}{2}\int_s^t |h(\xi_{s,\tau}(x)) + \phi_\tau|^2\, d\tau\right\}.$$

Define next $\bar{v}_t \triangleq \mathbf{E}^\dagger(\zeta_t^s \cdot Q_t^s q_s)$. The Fubini theorem, the change of variable $x = \eta_{t,s}(y)$ and the Lemma 6.2 of [16, Chapter 2] give

$$\begin{aligned}
(\bar{v}_t, f) &= \mathbf{E}^\dagger \int f(y) q_s(\eta_{t,s}(y))\Xi_{s,t}^\phi(\eta_{t,s}(y)) \exp\left\{\int_s^t \phi_\tau^* h(\eta_{t,\tau}(y))\, d\tau\right\}[J_{s,t}(\eta_{t,s}(y))]^{-1}\, dy \\
&= \mathbf{E}^\dagger \int f(\xi_{s,t}(x)) q_s(x)\Xi_{s,t}^\phi(x) \exp\left\{\int_s^t \phi_\tau^* h(\xi_{s,\tau}(x))\, d\tau\right\} dx \\
&= \int \bar{q}_s(x)\mathbf{E}^\dagger\left[f(\xi_{s,t}(x))\Xi_{s,t}^\phi(x) \exp\left\{\int_s^t \phi_\tau^* h(\xi_{s,\tau}(x))\, d\tau\right\}\right] dx \\
&= \int \bar{q}_s(x)\mathbf{E}^\phi\left[f(\xi_{s,t}(x)) \exp\left\{\int_s^t \phi_\tau^* h(\xi_{s,\tau}(x))\, d\tau\right\}\right] dx = (\bar{q}_s, \bar{w}_s).
\end{aligned}$$

It follows that $(\bar{q}_t, f) = (\bar{v}_t, f)$ for arbitrary test-function f and arbitrary $\{\phi_\tau, s \leqq \tau \leqq t\}$, which finishes the proof. □

5. APPROXIMATION OF THE STOCHASTIC CHARACTERISTICS

It has been proved in Section 4 that the stochastic semigroup $\{Q_t^s,\ 0 \leqq s \leqq t\}$ associated with the degenerate second-order stochastic PDE (3.2) satisfies

$$Q_t^s \phi(x) = \phi(\eta_{t,s}(x))\Gamma_{t,s}(x), \tag{5.1}$$

where $\eta_{t,s}(\cdot)$ is the inverse of the stochastic flow of diffeomorphisms $\xi_{s,t}(\cdot)$ associated with the stochastic differential equation (4.2), and $\Gamma_{t,s}(x)$ has been defined in (4.7). The purpose of this section is to investigate approximations of (5.1).

Considering that $\eta_{t,s}(\cdot)$ is also the stochastic flow of diffeomorphisms associated with the backward stochastic differential equation (4.3), it is natural to consider the following approximation

$$\bar{Q}_t^s\phi(x) \triangleq \phi(\bar{\eta}_{t,s}(x))\bar{\Gamma}_{t,s}(x), \tag{5.2}$$

where

$$\bar{\eta}_{t,s}(x) \triangleq x - \rho(x)[Y_t - Y_s - h(x)(t-s)] + \rho_0(x)(t-s),$$

and

$$\bar{\Gamma}_{t,s}(x) \triangleq \exp\{h^*(x)(Y_t - Y_s) - \tfrac{1}{2}|h(x)|^2(t-s) - \alpha^*(x)[Y_t - Y_s - h(x)(t-s)]$$
$$+ \bar{\alpha}(x)(t-s) + \alpha_0(x)(t-s)\},$$

are *computable* approximations of $\eta_{t,s}(x)$ and $\Gamma_{t,s}(x)$ respectively, both depending only on the increments $(Y_t - Y_s)$.

Remark 5.1 One possible approach would be to approximate $\eta_{t,s}(\cdot)$ by the stochastic flow of diffeomorphisms associated with the ordinary differential equation obtained from (4.3) by replacing the observation sample-path $\{Y_t, 0 \leqq t \leqq T\}$ with some regular approximation, such as the Euler stepwise approximation or the polygonal interpolation. The numerical analysis of such an approximation should not be very difficult. However, the resulting approximation would not be *explicitly computable.*

The remainder of this section is devoted to studying the rate of convergence of this approximation. First, a *stability* result similar to Proposition 4.6 is needed.

CONDITION A *Let $n \geqq 0$ be fixed. Assume that the initial condition satisfies $q \in H^n$.*

Then $\bar{Q}_t^s q$ is a square integrable random variable with values in H^n. In addition, the following estimate holds

$$\{\mathbf{E}^\dagger \|\bar{Q}_t^s q\|_n^2\}^{1/2} \leqq \|q\|_n e^{C(t-s)}.$$

Remark 5.2 Because $\bar{\eta}_{t,s}(\cdot)$ is *not* a diffeomorphism, this stability result can not be proved in the same way as in the proof of Proposition 4.6. The following proposition, which is proved in the Appendix, shows that Condition (A) holds in the simple case where the correlation coefficient ρ is constant. Whether this remains true in the general case—or how to modify the approximation scheme in such a way that Condition (A) holds without any additional assumption on the correlation coefficient—is still an open problem (however, see Remark A.1 below).

PROPOSITION 5.3 *Let $n \geqq 0$ be fixed. Assume that*

- *ρ is constant,*
- *h has bounded derivatives up to order n.*

Then Condition (A) holds.

Remark 5.4 The approximations $\bar{\eta}_{t,s}(x)$ and $\bar{\Gamma}_{t,s}(x)$ are based on the explicit expressions for $\eta_{t,s}(x)$ and $\Gamma_{t,s}(x)$, given in (4.3) and (4.7) respectively. This explains why the regularity assumptions on the coefficient h are different in Proposition 4.6 and Proposition 5.3.

Remark 5.5 In the case where the correlation coefficient ρ is constant, the approximations of $\eta_{t,s}(x)$ and $\Gamma_{t,s}(x)$ take the simple form

$$\bar{\eta}_{t,s}(x) \triangleq x - \rho[Y_t - Y_s - h(x)(t-s)],$$

and

$$\bar{\Gamma}_{t,s}(x) \triangleq \exp\{h^*(x)(Y_t - Y_s) - \tfrac{1}{2}|h(x)|^2(t-s)\},$$

respectively.

Next, the following proposition provides an error estimate for commuting the operator $\bar{Q}_t^s$ and spatial derivatives.

PROPOSITION 5.6 *Let $n \geqq 0$ and α a multi-index, be fixed. Assume that*

- *ρ has bounded derivatives up to order $(n+|\alpha|+2)$,*
- *h has bounded derivatives up to order $(n+|\alpha|)$,*
- *the initial condition satisfies $q \in H^{n+|\alpha|}$.*

Then, under Condition (A)

$$\{\mathbf{E}^\dagger \|\bar{Q}_t^s D^\alpha q - D^\alpha \bar{Q}_t^s q\|_n^2\}^{1/2} \leqq C\sqrt{t-s}\,\|q\|_{n+|\alpha|}.$$

Here again, the proof of this proposition is given in the Appendix.

□ *Overall Error Estimate*

The main result of the paper is provided by the following

THEOREM 5.7 *Consider the following approximation scheme*

$$\bar{\bar{p}}_{i+1} = P_{\delta_i}^* \bar{Q}_{t_{i+1}}^{t_i} \bar{\bar{p}}_i.$$

Assume that

- *a and c have bounded derivatives up to order* 4,
- *b and ρ have bounded derivatives up to order* 3,
- *h has bounded derivatives up to order* 2,
- *the initial condition satisfies $p_0 \in H^2$.*

Then, under Condition (A), $\bar{\bar{p}}_i$ approximates the solution p_{t_i} of the original equation (3.1) with a rate of convergence of order $\sqrt{\delta}$. Indeed

$$\{\mathbf{E}^\dagger|\bar{\bar{p}}_i - p_{t_i}|^2\}^{1/2} \leqq C\sqrt{\delta}\|p_0\|_2.$$

Proof In view of Theorem 3.4, it is enough to prove that

$$\{\mathbf{E}^\dagger|\bar{\bar{p}}_i - \bar{p}_i|^2\}^{1/2} \leqq C\sqrt{\delta}\|p_0\|_2.$$

Similarly to the proof of Theorem 3.4, the idea is to get an equation for $\bar{Q}_t^s\phi$ with ϕ smooth enough, that is similar to the original Eq. (3.2) for $Q_t^s\psi$, except for the initial condition and for some perturbation terms which have to be estimated. This gives an estimate of the one-step error, and the global estimate is obtained using the Gronwall lemma. Throughout the proof, the summation convention over repeated indices i, j, is used.

Differentiating both sides of (5.2) with respect to t

$$\begin{aligned}
d\bar{Q}_t^s\phi(x) = {} & \frac{\partial\phi}{\partial x_i}(\bar{\eta}_{t,s}(x))\left[-\sum_{k=1}^d \rho_k^i(x)[dY_t^k - h_k(x)\,dt] + \rho_0^i(x)\,dt\right]\bar{\Gamma}_{t,s}(x) \\
& + \frac{1}{2}\frac{\partial^2\phi}{\partial x_i\,\partial x_j}(\bar{\eta}_{t,s}(x))\sum_{k=1}^d [\rho_k^i(x)\rho_k^j(x)\,dt]\bar{\Gamma}_{t,s}(x) \\
& + \phi(\bar{\eta}_{t,s}(x))\left[\sum_{k=1}^d h_k(x)\,dY_t^k - \frac{1}{2}|h(x)|^2\,dt - \sum_{k=1}^d \alpha_k(x)[dY_t^k - h_k(x)\,dt]\right. \\
& \left. + \bar{\alpha}(x)\,dt + \alpha_0(x)\,dt + \frac{1}{2}|h(x) - \alpha(x)|^2\,dt\right]\bar{\Gamma}_{t,s}(x) \\
& + \frac{\partial\phi}{\partial x_i}(\bar{\eta}_{t,s}(x))\left[-\sum_{k=1}^d \rho_k^i(x)[h_k(x) - \alpha_k(x)]\,dt\right]\bar{\Gamma}_{t,s}(x) \\
= {} & \frac{1}{2}c^{i,j}(x)\frac{\partial^2\phi}{\partial x_i\,\partial x_j}(\bar{\eta}_{t,s}(x))\bar{\Gamma}_{t,s}(x)\,dt \\
& + \left[\rho_0^i(x) + \sum_{k=1}^d \rho_k^i(x)\alpha_k(x)\right]\frac{\partial\phi}{\partial x_i}(\bar{\eta}_{t,s}(x))\bar{\Gamma}_{t,s}(x)\,dt \\
& + [\bar{\alpha}(x) + \alpha_0(x) + \tfrac{1}{2}|\alpha(x)|^2]\phi(\bar{\eta}_{t,s}(x))\bar{\Gamma}_{t,s}(x)\,dt \\
& + \sum_{k=1}^d [h_k(x) - \alpha_k(x)]\phi(\bar{\eta}_{t,s}(x))\bar{\Gamma}_{t,s}(x)\,dY_t^k
\end{aligned}$$

$$- \sum_{k=1}^{d} \rho_k^i(x) \frac{\partial \phi}{\partial x_i}(\bar{\eta}_{t,s}(x)) \bar{\Gamma}_{t,s}(x)\, dY_t^k.$$

Now, it can be checked that

$$\bar{\alpha} + \alpha_0 + \frac{1}{2}|\alpha|^2 = \frac{1}{2} \frac{\partial^2 c^{i,j}}{\partial x_i \partial x_j} \quad \text{and} \quad \rho_0^i + \sum_{k=1}^{d} \rho_k^i \alpha_k = \frac{\partial c^{i,j}}{\partial x_j}.$$

Therefore, it holds

$$\begin{aligned} d\bar{Q}_t^s\phi(x) = & \left[\frac{1}{2} c^{i,j}(x) \frac{\partial^2 \phi}{\partial x_i \partial x_j}(\bar{\eta}_{t,s}(x)) \right. \\ & \left. + \frac{\partial c^{i,j}}{\partial x_j}(x) \frac{\partial \phi}{\partial x_i}(\bar{\eta}_{t,s}(x)) + \frac{1}{2} \frac{\partial^2 c^{i,j}}{\partial x_i \partial x_j}(x) \phi(\bar{\eta}_{t,s}(x)) \right] \bar{\Gamma}_{t,s}(x)\, dt \\ & + \sum_{k=1}^{d} \left[h_k(x) \phi(\bar{\eta}_{t,s}(x)) - \rho_k^i(x) \frac{\partial \phi}{\partial x_i}(\bar{\eta}_{t,s}(x)) - \frac{\partial \rho_k^i}{\partial x_i}(x) \phi(\bar{\eta}_{t,s}(x)) \right] \bar{\Gamma}_{t,s}(x)\, dY_t^k \\ = & \left[\frac{1}{2} c^{i,j}(x) \bar{Q}_t^s \frac{\partial^2 \phi}{\partial x_i \partial x_j}(x) + \frac{\partial c^{i,j}}{\partial x_j}(x) \bar{Q}_t^s \frac{\partial \phi}{\partial x_i}(x) + \frac{1}{2} \frac{\partial^2 c^{i,j}}{\partial x_i \partial x_j}(x) \bar{Q}_t^s \phi(x) \right] dt \\ & + \sum_{k=1}^{d} \left[h_k(x) \bar{Q}_t^s \phi(x) - \rho_k^i(x) \bar{Q}_t^s \frac{\partial \phi}{\partial x_i}(x) - \frac{\partial \rho_k^i}{\partial x_i}(x) \bar{Q}_t^s \phi(x) \right] dY_t^k \end{aligned}$$

so that

$$\begin{aligned} d\bar{Q}_t^s\phi = & \Lambda^* \bar{Q}_t^s \phi\, dt + \sum_{k=1}^{d} B_k^* \bar{Q}_t^s \phi\, dY_t^k \\ & + \frac{1}{2} c^{i,j} \left[\bar{Q}_t^s \frac{\partial^2 \phi}{\partial x_i \partial x_j} - \frac{\partial^2}{\partial x_i \partial x_j} \bar{Q}_t^s \phi \right] dt + \frac{\partial c^{i,j}}{\partial x_j} \left[\bar{Q}_t^s \frac{\partial \phi}{\partial x_i} - \frac{\partial}{\partial x_i} \bar{Q}_t^s \phi \right] dt \\ & - \sum_{k=1}^{d} \rho_k^i \left[\bar{Q}_t^s \frac{\partial \phi}{\partial x_i} - \frac{\partial}{\partial x_i} \bar{Q}_t^s \phi \right] dY_t^k. \end{aligned}$$

The difference $\varepsilon_t \triangleq \bar{Q}_t^s \phi - Q_t^s \psi$ satisfies

$$d\varepsilon_t = \Lambda^* \varepsilon_t\, dt + \sum_{k=1}^{d} B_k^* \varepsilon_t\, dY_t^k + f_t\, dt + \sum_{k=1}^{d} g_t^k\, dY_t^k,$$

where the perturbation terms are defined by

$$f_t \triangleq \frac{1}{2} c^{i,j} \left[\bar{Q}_t^s \frac{\partial^2 \phi}{\partial x_i \partial x_j} - \frac{\partial^2}{\partial x_i \partial x_j} Q_t^s \phi \right] + \frac{\partial c^{i,j}}{\partial x_j} \left[\bar{Q}_t^s \frac{\partial \phi}{\partial x_i} - \frac{\partial}{\partial x_i} \bar{Q}_t^s \phi \right],$$

and

$$g_t^k \triangleq -\rho_k^i \left[\bar{Q}_t^s \frac{\partial \phi}{\partial x_i} - \frac{\partial}{\partial x_i} \bar{Q}_t^s \phi \right],$$

respectively. Using estimates of [13]

$$\mathbf{E}^\dagger |\varepsilon_t|^2 \leqq \left[\mathbf{E}^\dagger |\varepsilon_s|^2 + C\mathbf{E}^\dagger \int_s^t |f_\tau|^2 \, d\tau + C\mathbf{E}^\dagger \sum_{k=1}^d \int_s^t \|g_\tau^k\|^2 \, d\tau \right] e^{C(t-s)}.$$

Moreover, it follows from Proposition 5.6 that

$$\mathbf{E}^\dagger |f_\tau|^2 \leqq C(\tau - s) \mathbf{E}^\dagger \|\phi\|_2^2 \, e^{C(\tau - s)},$$

$$\mathbf{E}^\dagger \|g_\tau^k\|^2 \leqq C(\tau - s) \mathbf{E}^\dagger \|\phi\|_2^2 \, e^{C(\tau - s)},$$

and therefore the Gronwall lemma yields

$$\mathbf{E}^\dagger |\varepsilon_t|^2 \leqq [\mathbf{E}^\dagger |\varepsilon_s|^2 + C(t-s)^2 \mathbf{E}^\dagger \|\phi\|_2^2] \, e^{C(t-s)},$$

provided $\phi \in H^2$. Now, it follows from the assumptions and in particular Condition (A), that $\bar{p}_i \in L^2(\Omega; H^2)$ for all i, so that setting $s = t_i$, $t = t_{i+1}$, $\phi = \bar{\bar{p}}_i$ and $\psi = \bar{p}_i$

$$\mathbf{E}^\dagger |\bar{Q}_{t_{i+1}}^{t_i} \bar{\bar{p}}_i - Q_{t_{i+1}}^{t_i} \bar{p}_i|^2 \leqq [\mathbf{E}^\dagger |\bar{\bar{p}}_i - \bar{p}_i|^2 + C(t_{i+1} - t_i)^2 \mathbf{E}^\dagger \|\bar{\bar{p}}_i\|_2^2] \, e^{C(t_{i+1} - t_i)}.$$

Next

$$\begin{aligned} \mathbf{E}^\dagger |\bar{\bar{p}}_{i+1} - \bar{p}_{i+1}|^2 &= \mathbf{E}^\dagger |P_{\delta_i}^* [\bar{Q}_{t_{i+1}}^{t_i} \bar{\bar{p}}_i - Q_{t_{i+1}}^{t_i} \bar{p}_i]|^2 \\ &\leqq [\mathbf{E}^\dagger |\bar{\bar{p}}_i - \bar{p}_i|^2 + C(t_{i+1} - t_i)^2 \mathbf{E}^\dagger \|\bar{\bar{p}}_i\|_2^2] \, e^{C(t_{i+1} - t_i)}, \end{aligned}$$

and the result follows from the discrete Gronwall lemma. □

A further step in the time-discretization would consist in approximating the Fokker–Planck semigroup $\{P_t^*, t \geqq 0\}$, using some classical approximation scheme. For instance, using the backward Euler scheme would result in the following global approximation scheme

$$(I - \delta_i L_0^*) \bar{\bar{p}}_{i+1} = \bar{Q}_{t_{i+1}}^{t_i} \bar{\bar{p}}_i,$$

with the same error estimate.

□ *Particle Approximation*

Another possible approach to approximate the degenerate second-order stochastic PDE (3.2)—based also on the representation (5.1) in terms of stochastic characteristics—would be to use *particle methods*, adapting the results presented in Raviart [24] for deterministic first-order PDE. The basic idea is to solve exactly Eq. (3.2) for an approximation of the initial condition, rather than approximate the stochastic characteristics as was done before.

Suppose that, at time t_i an approximation of the conditional probability distribution $q(x)\,dx$ is available, in terms of a convex linear combination of Dirac masses sitting at some particle locations $\{x_i^k, k \in K\}$ with corresponding weights $\{a_i^k, k \in K\}$ i.e.

$$q(x)\,dx \sim \sum_{k \in K} a_i^k \delta(x - x_i^k). \tag{5.3}$$

Solving exactly Eq. (3.2) in weak sense, with the approximation (5.3) as initial condition, gives the following approximation

$$Q_{t_{i+1}}^{t_i} q(x)\,dx \sim \sum_{k \in K} a_{i+1}^k \delta(x - x_{i+1}^k)$$

for the solution at time t_{i+1}. The new particle locations $\{x_{i+1}^k, k \in K\}$ and the corresponding weights $\{a_{i+1}^k, k \in K\}$ are computed according to

$$x_{i+1}^k = \xi_{t_i, t_{i+1}}(x_i^k) \quad \text{and} \quad a_{i+1}^k = a_i^k \Xi_{t_i, t_{i+1}}(x_i^k),$$

where $\xi_{s,t}(\cdot)$ is the diffeomorphism associated with Eq. (4.2), and $\Xi_{s,t}(\cdot)$ has been defined in (4.5).

The error estimate associated with this particle approximation will be studied elsewhere.

6. CONCLUSION

A time–discretization scheme of the Zakai equation for diffusion processes observed in correlated noise has been proposed, based on the stochastic characteristics introduced in [13, 15, 17]. Under the additional assumption that the correlation coefficient is constant, it has been shown that the rate of convergence of this approximation is of order $\sqrt{\delta}$, where δ is the time discretization step.

The same rate of convergence has been obtained in Elliott–Glowinski [7] for a different approximation

- on one hand, the approximation considered in [7] has a probabilistic interpretation, which is not the case so far for the time discretization scheme presented here (however, see Remark 3.6 above),

- on the other hand, the latter is *actually computable*, whereas no numerical algorithm is provide to *compute* the approximation considered in [7].

Another point of interest would be to study some particle approximation for the degenerate second-order stochastic PDE, adapting the results presented in Raviart [24] for deterministic first-order PDE.

As was pointed out to the authors by Harold Kushner and the anonymous referee, one would have to discretize the space variable and to bound the state space, in order to get a completely computable numerical scheme. This is a different problem, for which several approaches have already been used: finite difference approximation, by Kushner [18] and DiMasi–Runggaldier [5], finite element method, by Bennaton [1] and Germani–Piccioni [9], with error estimate. The reference [9] also provides error estimate for bounding the state space, using weighted Sobolev spaces introduced by Krylov–Rozovskii [14]. Therefore, the time discretization scheme presented in the paper should be combined with such space discretization techniques, in order to be completely computable. To some extent, the choice of the space discretization scheme is dependent on the application: for instance, the method of characteristics (also called particle approximation in [24]) is well-adapted to first-order PDE arising in the filtering of noise-free processes, and has been recently used in target tracking applications, see Campillo–Le Gland [4] and Lasdas–Davis [19].

References

[1] J. F. Bennaton, Discrete time Galerkin approximations to the nonlinear filtering solution, *J. Math. Anal. Appl.* **110**, No. 2 (1985), 364–383.

[2] A. Bensoussan, R. Glowinski and A. Rascanu, Approximation of Zakai equation by the splitting-up method, In: *Stochastic Systems and Optimization* (*Warsaw–1988*) (ed. J. Zabczyk) 257–265, Springer-Verlag (LNCIS-136) (1989).

[3] Yu. N. Blagoveschenskii and M. I. Freidlin, Certain properties of diffusion processes depending on a parameter, *Soviet Math.* **2**, No. 3 (1961), 633–636.

[4] F. Campillo and F. Le Gland, Application du filtrage non linéaire en trajectographie passive, In: *12eme Colloque GRETSI* (*Juan les Pins–1989*) (1989), 197–200.

[5] G. B. Di Masi and W. J. Runggaldier, Continuous-time approximations for the nonlinear filtering problem, *Appl. Math. Optim.* **7**, No. 3 (1981), 233–245.

[6] G. B. Di Masi, M. Pratelli and W. J. Runggaldier, An approximation for the nonlinear filtering problem with error bound, *Stochastics* **14**, No. 4 (1985), 247–271.

[7] R. J. Elliott and R. Glowinski, Approximations to solutions of the Zakai filtering equation, *Stoch. Anal. Appl.* **7**, No. 2 (1988), 145–168.

[8] P. Florchinger and F. Le Gland, Time-discretization of the Zakai equation for diffusion processes observed in correlated noise, In: *Analysis and Optimization of Systems* (*Juan les Pins–1990*) (eds. A. Bensoussan and J. L. Lions) 228–237, Springer-Verlag (LNCIS–144) (1990) (also: *INRA Report 1222* (May, 1990)).

[9] A. Germani and M. Piccioni, Semi-discretization of stochastic partial differential equations on $\mathbf{R}^d$ by a finite-element technique, *Stochastics* **23**, No. 2 (1988), 131–148.

[10] Ph. Hartman, *Ordinary Differential Equations*, Birkhäuser (1982).

[11] H. Korezlioglu and G. Mazziotto, Approximations of the nonlinear filter by periodic sampling and quantization, In: *Analysis and Optimization of Systems, Part 1* (*Nice–1984*) (eds. A. Bensoussan and J. L. Lions) 553–567, Springer-Verlag (LNCIS–62) (1984).

[12] N. V. Krylov and B. L. Rozovskii, On the Cauchy problem for linear stochastic partial differential equations, *Math. USSR Izvestija* **11**, No. 6 (1977), 1267–1284.

[13] N. V. Krylov and B. L. Rozovskii, Characteristics of degenerating second-order parabolic Itô equations, *J. Soviet Math.* **32**, No. 4 (1982), 336–348.
[14] N. V. Krylov and B. L. Rozovskii, Stochastic partial differential equations and diffusion processes, *Russian Math. Surveys* **37**, No. 6 (1982), 81–105.
[15] H. Kunita, Stochastic partial differential equations connected with nonlinear filtering, In: *Nonlinear Filtering and Stochastic Control* (*Cortona–1981*) (eds. S. K. Mitter and A. Moro) 100–169, Springer-Verlag (LNM–972) (1982).
[16] H. Kunita, Stochastic differential equations and stochastic flows of diffeomorphisms, In: *Ecole d'Eté de Probabilités de St. Flour XII* (*1982*) (ed. P. L. Hennequin) 144–303, Springer-Verlag (LNM-1097) (1984).
[17] H. Kunita, First order partial differential equations, In: *Stochastic Analysis* (*Katata and Kyoto-1982*) (ed. K. Itô) 249–269, North-Holland (1984).
[18] H. J. Kushner, *Probability methods for approximations in stochastic control and for elliptic equations*, Academic Press (1977).
[19] V. Lasdas and M. H. A. Davis, A piecewise deterministic approach to target motion analysis, In: *28th IEEE CDC* (*Tampa–1989*) 1395–1396 (1989).
[20] F. Le Gland, Time discretization of nonlinear filtering equations, In: *28th IEEE CDC* (*Tampa–1989*) 2601-2606 (1989).
[21] N. J. Newton, Discrete approximations for Markov-chain filters, *Ph.D Thesis, Imperial College* (1983).
[22] J. Picard, Approximation of nonlinear filtering problems and order of convergence, In: *Filtering and Control of Random Processes* (*ENST/CNET–1983*) (eds. H. Korezlioglu, G. Mazziotto and J. Szpirglas) 219–236, Springer-Verlag (LNCIS–61) (1984).
[23] E. Pardoux, Stochastic partial differential equations and filtering of diffusion processes, *Stochastics* **3**, No. 2 (1979), 127–167.
[24] P. A. Raviart, An analysis of particle methods, In: *Numerical Methods in Fluid Dynamics* (*Como–1983*) (ed. F. Brezzi) 243–324, Springer-Verlag (LNM–1127) (1985).
[25] M. Zakai, On the optimal filtering of diffusion processes, *Z. Wahrschein. Verw. Geb.* **11**, No. 3 (1969), 230–243.

APPENDIX

Proof of Stability and Commutation Estimates

The purpose of this appendix is to prove the stability and commutation estimates for the approximation introduced in Section 5.

Proof of Proposition 5.3 It is enough to prove the result for $n=0$.

Since $\bar{\eta}_{t,s}(\cdot)$ is not a diffeomorphism, one can not use a change of variable as in the proof of Proposition 4.6. Instead, one uses the fact that $\bar{\eta}_{t,s}(x)$ and $\bar{\Gamma}_{t,s}(x)$ are very simple functions of the Gaussian random variable $(Y_t - Y_s)$. First

$$\mathbf{E}\dagger|\bar{Q}_t^s q|^2 = \mathbf{E}\dagger \int [|q(\bar{\eta}_{t,s}(x))|\bar{\Gamma}_{t,s}(x)]^2\,dx$$

$$= \frac{1}{[2\pi(t-s)]^{d/2}} \int\int |q(x-\rho(x)[w-h(x)(t-s)]+\rho_0(x)(t-s))|^2$$

$$\times \exp\{2h^*(x)w - |h(x)|^2(t-s) - 2\alpha^*(x)[w-h(x)(t-s)]$$

$$+2\bar{\alpha}(x)(t-s)+2\alpha_0(x)(t-s)\}\exp\left\{-\frac{|w|^2}{2(t-s)}\right\}dw\,dx.$$

Next

$$2[h(x)-\alpha(x)]^*w-\frac{|w|^2}{2(t-s)}=2|h(x)-\alpha(x)|^2(t-s)-\frac{|w-2[h(x)-\alpha(x)](t-s)|^2}{2(t-s)},$$

so that, using the new variables (x,v) with $v=w-2[h(x)-\alpha(x)](t-s)$

$$\mathbf{E}\dagger|\bar{Q}_t^s q|^2\leqq e^{C(t-s)}\frac{1}{[2\pi(t-s)]^{d/2}}\int\int|q(x-\rho(x)v$$

$$+\gamma(x)(t-s))|^2\exp\left\{-\frac{|v|^2}{2(t-s)}\right\}dv\,dx,$$

where $\gamma(x)\triangleq\rho_0(x)-\rho(x)[h(x)-2\alpha(x)]$.

In the particular case where $\rho(x)\equiv\rho$, the application $F(x)\triangleq x-\rho v+\gamma(x)(t-s)$ is a diffeomorphism provided $0\leqq(t-s)<1/C$, and moreover the Jacobian is bounded below by $[1-C(t-s)]$. Therefore, using the new variables (y,z) with $y=F(x)$

$$\mathbf{E}\dagger|\bar{Q}_t^s q|^2\leqq\frac{e^{C(t-s)}}{1-C(t-s)}\frac{1}{[2\pi(t-s)]^{d/2}}\int\int|q(y)|^2\exp\left\{-\frac{|v|^2}{2(t-s)}\right\}dv\,dy$$

$$\leqq\frac{e^{C(t-s)}}{1-C(t-s)}\int|q(y)|^2\,dy,$$

provided $0\leqq(t-s)\leqq\delta\leqq1/C$, which finishes the proof. □

Remark A.1 According to the detail of the proof above, it is enough for the Condition (A) to hold, that

$$\frac{1}{[2\pi(t-s)]^{d/2}}\int\int|q(x-\rho(x)v+\gamma(x)(t-s))|^2\exp\left\{-\frac{|v|^2}{2(t-s)}\right\}dv\,dx\leqq e^{C(t-s)}\int|q(y)|^2\,dy,$$

for any bounded function γ.

Proof of Proposition 5.6 Here again, it is enough to prove the result for $n=0$ and $|\alpha|=1$. Throughout the proof, the summation convention over repeated indices j is used.

For q smooth enough, it holds

$$\frac{\partial}{\partial x_i}\bar{Q}_t^s q(x)=\frac{\partial q}{\partial x_j}(\bar{\eta}_{t,s}(x))\Bigg[\delta^{i,j}-\sum_{k=1}^{d}\frac{\partial\rho_k^j}{\partial x_i}(x)[Y_t^k-Y_s^k-(t-s)h_k(x)]$$

$$+(t-s)\left(\sum_{k=1}^{d}\rho_k^j(x)\frac{\partial h_k}{\partial x_i}(x)+\frac{\partial\rho_0^j}{\partial x_i}(x)\right)\Bigg]\bar{\Gamma}_{t,s}(x)$$

$$+q(\bar{\eta}_{t,s}(x))\Bigg[\sum_{k=1}^{d}\left(\frac{\partial h_k}{\partial x_i}(x)-\frac{\partial\alpha_k}{\partial x_i}(x)\right)[Y_t^k-Y_s^k-(t-s)h_k(x)]$$

$$+(t-s)\left(\sum_{k=1}^{d}\alpha_k(x)\frac{\partial h_k}{\partial x_i}(x)+\frac{\partial\bar{\alpha}}{\partial x_i}(x)+\frac{\partial\alpha_0}{\partial x_i}(x)\right)\Bigg]\bar{\Gamma}_{t,s}(x)$$

$$=\bar{Q}_t^s\frac{\partial q}{\partial x_i}(x)-\sum_{k=1}^{d}\frac{\partial\rho_k^j}{\partial x_i}(x)[Y_t^k-Y_s^k-(t-s)h_k(x)]\bar{Q}_t^s\frac{\partial q}{\partial x_j}(x)$$

$$+(t-s)\left(\sum_{k=1}^{d}\rho_k^j(x)\frac{\partial h_k}{\partial x_i}(x)+\frac{\partial\rho_0^j}{\partial x_i}(x)\right)\bar{Q}_t^s\frac{\partial q}{\partial x_j}(x)$$

$$+\sum_{k=1}^{d}\left(\frac{\partial h_k}{\partial x_i}(x)-\frac{\partial\alpha_k}{\partial x_i}(x)\right)[Y_t^k-Y_s^k-(t-s)h_k(x)]\bar{Q}_t^s q(x)$$

$$+(t-s)\left(\sum_{k=1}^{d}\alpha_k(x)\frac{\partial h_k}{\partial x_i}(x)+\frac{\partial\bar{\alpha}}{\partial x_i}(x)+\frac{\partial\alpha_0}{\partial x_i}(x)\right)\bar{Q}_t^s q(x).$$

Therefore, under Condition (A)

$$\mathbf{E}^\dagger\left|\frac{\partial}{\partial x_i}\bar{Q}_t^s q-\bar{Q}_t^s\frac{\partial q}{\partial x_i}\right|^2\leqq C(t-s)\left[\mathbf{E}^\dagger\sum_{j=1}^{m}\left|\bar{Q}_t^s\frac{\partial q}{\partial x_j}\right|^2+\mathbf{E}^\dagger|\bar{Q}_t^s q|^2\right]$$

$$\leqq C(t-s)\left[\sum_{j=1}^{m}\left|\frac{\partial q}{\partial x_j}\right|^2+|q|^2\right]$$

$$\leqq C(t-s)\|q\|^2. \qquad \square$$