ON THE EXISTENCE OF WEAK SOLUTIONS TO
STOCHASTIC DIFFERENTIAL EQUATIONS WITH
DEGENERATE DIFFUSION

Norbert CHRISTOPEIT
Institut für Ökonometrie und Operations Research, Universität Bonn, Adenauerallee 24–42, D5300
Bonn 1, West Germany

Received 2 July 1982

For a certain class of stochastic differential equations with nonlinear drift and degenerate
diffusion term existence of a weak solution is shown.

AMS 1980 Subject Classification: Primary 60H10
stochastic differential equations Ito processes
weak solutions weak convergence

1. Introduction

In [4] it is shown that the minimization of predicted miss (cf. [1, 5]) for a partially
observable linear system with control actions in the unit cube leads to the following
stochastic differential equation for the best estimate \( \hat{\xi}_t \) of the state \( \xi_t \):

\[
d\hat{\xi}_t = L(t)\hat{\xi}_t dt + B(t)u(t, \hat{\xi}_t) dt + K(t)G(t) d\omega_t, \tag{1.1}
\]

where \((\omega_t)\) is an \(r\)-dimensional standard Brownian motion and \(L(t), B(t), K(t)\) and
\(G(t)\) are continuous matrix valued functions of size \(d \times d, d \times m, d \times r\) and
\(r \times r\), respectively. \(G(t)\) is the covariance of the observation noise, assumed nonsingular,
while \(K(t)\) is the gain matrix. \(u(t, x)\) has the components

\[
u_i(t, x) = -\text{sign}(b_i(t)) \text{sign}(\xi_i, x), \quad i = 1, \ldots, m,
\]

where \(b(t) = B'(t)\xi_t\) and \(\xi_t\) is a nonzero vector evolving in a continuously differentiable
way in time. If the nonlinear drift term were not present, equation (1.1) would
have a unique strong (i.e. adapted to \((\omega_t)\)) solution for any initial value. In the
presence of the nonlinear term we could find a weak solution by means of the
Girsanov measure transformation method (cf. [10, 6]) provided \(K(t)G(t)\) has rank
\(d\) for all \(t\) (or, somewhat weaker, \(B(t) = K(t)G(t)\theta(t)\) for some matrix function
\(\theta(t)\)). But this would mean that the dimension of the observed vector, which is \(r,\)
should be at least as large as the dimension of the state vector, which is \(d,\) and,
moreover, that the observation matrix $F$ (in $d\eta_t = F(t)\xi_t \, dt + G(t) \, dv_t$, the equation for the observable output) should also have rank $d$.

There is, however, a way of avoiding this unpleasant assumption. It is based on the observation that actually the nonlinear term is only a function of the 1-dimensional statistic $\langle \xi_n, \dot{\xi}_n \rangle$. We shall do the analysis in a somewhat more general framework.

Consider the stochastic differential equation

$$d\xi_t = f(t, \xi_t) \, dt + g(t, \xi_t) \, dt + \sigma(t, \xi_t) \, dw_t, \quad 0 \leq t \leq T, \quad \xi_0 = x. \tag{1.2}$$

Let us make the following assumptions. Denote by $C^{\infty}$ the space of $\mathbb{R}^n$-valued continuous functions on $[0, T]$, endowed with the sup-norm topology and with the natural filtration $(\mathcal{G}_t)$.

(A1) $f:[0, T] \times C^d \rightarrow \mathbb{R}^d$ is measurable, adapted to $(\mathcal{G}_t^{d})$ and satisfies the Ito conditions

$$|f(t, x) - f(t, y)| \leq K\|x - y\|, \quad (1.3)$$

$$|f(t, x)| \leq K(1 + \|x\|), \quad (1.4)$$

for some constant $K$ (with $\|x\| = \sup_{t \leq T, x \in \mathbb{R}^n} |x|$, $\| \cdot \|$ denoting Euclidean matrix norm).

(A2) $\sigma:[0, T] \times C^d \rightarrow \mathbb{R}^d \times \mathbb{R}^q$ is measurable, adapted to $(\mathcal{G}_t^{d})$ and satisfies the Ito conditions $(1.3), (1.4)$.

Let $C^{1,2}$ denote the space of all functions $\varphi:[0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^q$ which are once continuously differentiable with respect to $t$ and twice with respect to $x$. We shall use the following notation:

$$\varphi'_t(t, x) = \frac{\partial \varphi}{\partial t}(t, x);$$

$$\nabla \varphi(t, x) = \left(\frac{\partial \varphi_i}{\partial x_j}(t, x)\right)_{i,j}, \quad i = 1, \ldots, q, \; j = 1, \ldots, d;$$

$$\Delta \varphi_l(t, x) = \left(\frac{\partial^2 \varphi_l}{\partial x_i \partial x_j}(t, x)\right)_{i,j}, \quad l = 1, \ldots, q; \; i, j = 1, \ldots, d.$$

Note in particular that, for scalar $\varphi$, $\nabla \varphi$ is a row vector.

(A3) There exist measurable functions $\gamma:[0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$, $\gamma^N:[0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$, $N = 1, 2, \ldots$, and $\varphi \in C^{1,2}$ such that the following holds:

(i) $g(t, x) = \gamma(t, \varphi(t, x))$ for all $(t, x) \in [0, T] \times \mathbb{R}^d$.

(ii) The $\gamma^N(t, z)$ are bounded by a constant uniformly in $(t, z)$ and $N$, and they satisfy the Ito condition $(1.3)$ with a constant $\tilde{K} = \tilde{K}_N$ depending on $N$.

(iii) For a.e. $0 \leq t \leq T$ there exists a set $I_t$ in $\mathbb{R}^q$, Lebesgue measure 0 such that, for $z \notin I_t$,

$$\gamma^N(t, z') \rightarrow \gamma(t, z) \quad \text{as } N \rightarrow \infty, \; |z' - z| \rightarrow 0.$$
(iv) For all \((t, x)\),
\[ |\varphi'_1(t, x)| + |\nabla \varphi(t, x)| + |\nabla \varphi_1(t, x)| \leq K(1 + |x|^p), \quad l = 1, \ldots, q, \]
for some nonnegative integer \(p\).

\(\text{(A4)}\) Denote \(a = \sigma \sigma'\). Then there exists a (measurable, \((\mathcal{F}^d_t)\)-adapted) diagonalization \(C(t, x)A(t, x)C'(t, x)\) of \(\nabla \varphi(t, x)A(t, x)\) of \(\nabla \varphi(t, x)A(t, x)\) (i.e. \(C\) orthogonal, \(A\) diagonal) such that \(\psi(t, x) = C(t, x)A^{1/2}(t, x)\) satisfies the Ito conditions and admits an inverse \(\psi^{-1}(t, x)\) satisfying the polynomial growth condition (iv). Moreover, the (unique) strong solution \((Y_t)\) of
\[
Y_t = Y_0 + \int_0^t \psi(s, Y_t) \, d\tilde{w}_s \quad (1.5)
\]
(where \((\tilde{w}_s)\) is a \(q\)-dimensional Brownian motion) has, for \(t > 0\), a distribution of \(Y_t\), which is absolutely continuous with respect to \(q\)-dimensional Lebesgue measure.

Apart from the technicalities, the decisive assumption is that the nonlinear drift term \(g\) can be expressed as a function \(\gamma\) of the statistic \(\varphi(t, x)\), and that \(\gamma\) itself can be approximated locally uniformly by smooth functions a.e. on \((t, z)\)-space. For the example considered initially, \(\varphi(t, x) = \langle \zeta, x \rangle\) and (A1)-(A3) are trivially satisfied if we take
\[ \gamma(t, z) = B(t) \text{sign}(b(t)) \cdot (-\text{sign}(z)) \]
(where \(\text{sign}(b(t))\) is the column vector with components \(\text{sign}(b_i(t))\)) and approximate \(-\text{sign}(z)\) by the functions
\[ \gamma^N(z) = 1_{[z < -1/N]} - Nz1_{[-1/N, z < 1/N]} - 1_{[z > 1/N]} \]
(\(\zeta_A = \text{indicator of } A\)). (A4) reduces then to the requirement that
\[ \psi^2(t) = \xi^t K(t)G(t)G'(t)K'(t)\xi_t > 0 \quad \text{for all } 0 < t < T. \]
For this implies that \(Y_t\) has a density for all \(t > 0\) (cf. [6]).

Let us now formulate the result we are going to prove.

**Theorem.** Under assumptions (A1)-(A4) equation (1.2) has a weak solution.

2. Some prerequisites

Let us begin with some general remarks on Ito processes. Suppose that on some complete probability space \((\Omega, \mathcal{F}, P)\) there is given an \(r\)-dimensional Wiener process \((w_s, \mathcal{F}_t), 0 \leq t \leq T\) (with \(\mathcal{F}_0\) containing the \(\mathcal{F}\)-nullsets), and let \((\xi_t)\) be a \(q\)-dimensional Ito process given by
\[
\xi_t = \xi_0 + \int_0^t A_s(\omega) \, ds + \int_0^t B_s(\omega) \, dw_s, \quad 0 \leq t \leq T. \quad (2.1)
\]
The \((q \times 1)\)-respectively \((q \times r)\)-matrix valued processes \((A_i)\) and \((B_i)\) are supposed to be measurable and adapted to \((\mathcal{F}_t)\) and satisfy the following conditions:

\[ \int_0^T E|A_i(\omega)| \, dt < \infty; \quad (2.2) \]
\[ \int_0^T |B_i(\omega)|^2 \, dt < \infty \quad \text{a.e.} \quad (2.3) \]

It can then be shown (cf. [12], [10, Theorem 7.17]) that \(B_i(\omega)B_i'(\omega)\) ds is actually \(\mathcal{F}_t\)-measurable for all \(t\). Hence we may find \((\mathcal{F}_t^{\xi})\)-adapted processes \((C_i)\) and \((A_i)\), \(C_i\) orthogonal and \(C_i\) diagonal, that diagonalize \(B_iB_i'\):

\[ B_i(\omega)B_i'(\omega) = C_i(\omega)A_i(\omega)C_i'(\omega) \quad \text{a.e.} \]

for almost all \(t\) (cf. [10]). There exists then a measurable process \((\tilde{B}_i)\) (with values in \(\mathbb{R}^q \times \mathbb{R}^q\)) on \(C^q\), adapted to \((\mathcal{C}_t^{\xi, +})\), such that

\[ \tilde{B}_i(\xi(\omega)) = C_i(\omega)A_i^{1/2}(\omega) \]

for almost all \((t, \omega)\) (cf. [13]). For simplicity, we require right now (it is not yet necessary at this point, but we shall need it in the lemma anyway):

\[ \tilde{B}_i(x) \text{ is nonsingular for all } (t, x) \in [0, T] \times C^q. \quad (2.4) \]

It can then be shown that there exist a \(q\)-dimensional Wiener process \((\tilde{w}_s, \mathcal{F}_s^{\xi})\) and a measurable \((\mathcal{C}_t^{\xi, +})\)-adapted process \((\tilde{A}_i)\) on \(C^q\) such that \(\tilde{A}_i(\xi) = E\{A_i|\mathcal{F}_t^{\xi}\}\) a.e. for almost all \(t\) and the representation

\[ \tilde{\xi}_t = \xi_0 + \int_0^t \tilde{A}_i(\xi) \, ds + \int_0^t \tilde{B}_i(\xi) \, d\tilde{w}_s \]

holds (cf. [10, 12]).

Suppose now that

\[ \tilde{B}_i(x) \text{ satisfies the Ito conditions (1.3) and (1.4).} \quad (2.5) \]

Let \((\eta_i)\) be the (unique) strong solution of the equation

\[ \eta_t = \xi_0 + \int_0^t \tilde{B}_i(\eta) \, d\tilde{w}_s \quad (2.6) \]

and denote by \(\mu_\xi\) and \(\mu_\eta\) the measures induced on \(C^q\) by \((\xi_t)\) and \((\eta_t)\), respectively.

We need one more assumption:

\[ \int_0^T \tilde{A}_i'(x)[\tilde{B}_i'(x)\tilde{B}_i'(x)]^{-1}\tilde{A}_i(x) \, dt < \infty \quad \mu_\xi\text{-a.e.} \quad (2.7) \]

**Lemma.** Under assumptions (2.2)-(2.5) and (2.7) \(\mu_\xi\) is absolutely continuous with respect to \(\mu_\eta\): \(\mu_\xi \ll \mu_\eta\).

\(^1\) Apparently it suffices to require (2.4) \(\lambda \in \mu_\xi\text{-a.e.} \)
**Proof** (cf. [10, proof of Theorem 7.2], [7, Chapter III.2]). Denote

$$\beta_t(x) = \tilde{B}_t^{-1}(x) \tilde{A}_t(x)$$

(2.8)

(if (2.4) is required only $\lambda \otimes \mu_\xi$-a.e., put $\beta_t(x) = 0$ on the exceptional set). For $n = 1, 2, \ldots$, define

$$\tau_n = \begin{cases} \inf \{ t \leq T : \int_0^t |\beta_s(x)|^2 \, ds = n \} \\ T \quad \text{if } \int_0^T |\beta_s(x)|^2 \, ds < n, \end{cases}$$

and

$$A_t^{(n)}(x) = 1_{[t < \tau_n(x)]} \tilde{A}_t(x), \quad B_t^{(n)}(x) = 1_{[t < \tau_n(x)]} \tilde{B}_t(x), \quad \beta_t^{(n)}(x) = 1_{[t < \tau_n(x)]} \beta_t(x).$$

Then $\tau_n$ is a $(\mathcal{F}_t)$-stopping time, $(A_t^{(n)})$, $(B_t^{(n)})$ and $(\beta_t^{(n)})$ are measurable $(\mathcal{F}_{t^+})$-adapted, and

$$A_t^{(n)}(x) = B_t^{(n)}(x) \beta_t^{(n)}(x).$$

(2.9)

Let

$$\xi_t^{(n)} = \xi_0 + \int_0^t A_s^{(n)}(\xi) \, ds + \int_0^t B_s^{(n)}(\xi) \, d\tilde{w}_s, \quad \eta_t^{(n)} = \xi_0 + \int_0^t B_s^{(n)}(\eta) \, d\tilde{w}_s.$$

Then, since $\xi_t^{(n)} 1_{[t < \tau_n]} = \xi_t 1_{[t < \tau_n]}$ (and same for $\eta_t^{(n)}$),

$$\xi_t^{(n)} = \xi_0 + \int_0^t A_s^{(n)}(\xi_t^{(n)}) \, ds + \int_0^t B_s^{(n)}(\xi_t^{(n)}) \, d\tilde{w}_s,$$

(2.10)

$$\eta_t^{(n)} = \xi_0 + \int_0^t B_s^{(n)}(\eta_t^{(n)}) \, d\tilde{w}_s.$$  

(2.11)

Next we note that the solution to (2.11) is pathwise unique. This is true since, for $t \leq \tau_n$, any solution to (2.11) solves (2.6) and is constant for $t > \tau_n$. Since

$$P\left[ \int_0^T |\beta_t^{(n)}(\xi^{(n)})|^2 \, dt \leq n \right] = 1,$$

it follows that

$$E \exp \left[ \pm \int_0^T \beta_t^{(n)}(\xi^{(n)}) \, d\tilde{w}_t - \frac{1}{2} \int_0^T |\beta_t^{(n)}(\xi^{(n)})|^2 \, dt \right] = 1$$

(cf. Theorem 6.1 in [10]) and, making use of (2.9)–(2.11), we may argue as in the proof of Theorem 7.18 in [10] to show that $\mu^{\xi^{(n)}} \sim \mu^{\eta^{(n)}}$ with densities given by the appropriate exponential $\rho(n)$. 
Observing that $\xi^{(n)}(\omega) = \xi(\omega)$ on the set $[\omega, \tau_n(\omega) = T]$ and $\eta^{(n)}(\omega) = \eta(\omega)$ on $[\omega, \tau_n(\omega) = T]$, we find that, for any measurable $\Gamma \subset C^q$,

$$\mu_\xi(\Gamma) = \mu_\xi(\Gamma, \tau_n = T) + \mu_\xi(\Gamma, \tau_n < T) = \mu_\xi(\Gamma, \tau_n = T) + \mu_\xi(\Gamma, \tau_n < T)$$

$$= \int_{\Gamma \cap \{\tau_n = T\}} \rho^{(m)} d\mu_\eta^{(n)} + \mu_\xi(\Gamma, \tau_n < T)$$

$$= \int_{\Gamma \cap \{\tau_n = T\}} \rho^{(m)} d\mu_\eta + \mu_\xi(\Gamma, \tau_n < T).$$

Now, if $\mu_\eta(\Gamma) = 0$, the first term in the last line vanishes and

$$\mu_\xi(\Gamma) \leq \mu_\xi(\tau_n < T) \leq \mu_\xi\left(\int_0^T |\beta_t(x)|^2 \, dt \geq n\right) \to 0$$

as $n \to \infty$ by virtue of (2.7) and (2.8).

### 3. Proof of theorem

(For notational simplicity for scalar $\varphi$.) Denote $g^N(t, x) = \gamma^N(t, \varphi(t, x))$ and let $(\xi_i^N)$ be the (unique) strong solution of the equation

$$\xi_i^N = \xi_0 + \int_0^t f(s, \xi_i^N) \, ds + \int_0^t g^N(s, \xi_i^N) \, ds + \int_0^t \sigma(s, \xi_i^N) \, dw_s,$$

$$= \xi_0 + F^N(t) + G^N(t) + B^N(t), \quad (3.1)$$

where the random functions $F^N$, $G^N$ and $B^N$ are defined in an obvious manner. By Ito's formula, the process $X_i^N = \varphi(t, \xi_i^N)$ satisfies the equation

$$X_i^N = \varphi(0, \xi_0) + \int_0^t \left[ \varphi_t(s, \xi_i^N) + \nabla \varphi(s, \xi_i^N) f(s, \xi_i^N) + \frac{1}{2} \text{tr} A(s, \xi_i^N) a(s, \xi_i^N) \right] ds$$

$$+ \int_0^t \nabla \varphi(s, \xi_i^N) g^N(s, \xi_i^N) \, ds + \int_0^t \nabla \varphi(s, \xi_i^N) \sigma(s, \xi_i^N) \, dw_s,$$

$$= \varphi(0, \xi_0) + \tilde{F}^N(t) + \tilde{G}^N(t) + \tilde{B}^N(t). \quad (3.2)$$

with $\tilde{F}^N$, etc. again defined in an obvious manner. Consider now the random functions

$$\Phi^N = (\xi^N, X^N, F^N, G^N, B^N, \tilde{F}^N, \tilde{G}^N, \tilde{B}^N)$$

with values in $C^{d(d+1)}$. We shall show that the sequence $\Phi^N$, $N = 1, 2, \ldots$, is tight. To this end, note first that under assumptions (A1), (A2) and (A3) (ii), (iv), for every nonnegative integer $m$

$$E\|\xi^N\|_t^m \leq C_1(m) \quad (3.3)$$

(cf. [2]), where the constant $C_1(m)$ depends only on $K$, $d$, $r$, $T$ and the initial value
Let now $H^N(t) = \int_0^t h^N(s, \xi^N) \, ds$ stand for any of the random functions $F^N$, $G^N$, $\tilde{F}^N$ or $\tilde{G}^N$. Doing a little exercise in Hölder estimates we obtain that, uniformly in $0 \leq s \leq t \leq T$ and $N = 1, 2, \ldots$,

$$E|H^N(t) - H^N(s)|^4 \leq (t - s)^3 E \int_s^t |h^N(r, \xi^N)|^4 \, dr$$

$$\leq C_3(m)(t - s)^2 \int_s^t E(1 + \|\xi^N\|^{4m}) \, dr,$$

where the integer $m$ may be different for $F^N$, etc. Hence, from (3.3), and taking the maximum of the four $C_3(m)$ values,

$$E|H^N(t) - H^N(s)|^4 \leq C_4(t - s)^4.$$

(3.5)

Similarly, if $M^N(t) = \int_0^t m^N(s, \xi^N) \, dw$, stands for either $B^N(t)$ or $\tilde{B}^N(t)$,

$$E|M^N(t) - M^N(s)|^4 \leq 6(t - s)^3 \int_s^t E|m^N(r, \xi^N)|^4 \, dr$$

$$\leq (t - s)C_5(m) \int_s^t E(1 + \|\xi^N\|^{4m}) \, dr \leq C_6(t - s)^2.$$

(3.6)

From (3.5) and (3.6) it follows that

$$E|\Phi^N(t) - \Phi^N(s)|^4 \leq C_7(t - s)^2$$

for all $0 \leq s \leq t \leq T$ such that $t - s \leq 1$ and all $N = 1, 2, \ldots$, with a constant $C_7$ depending only on $d$, $r$, $K$, $T$ and $\xi_0$. This implies that the random functions $\Psi^N$, $N = 1, 2, \ldots$, are tight in $C^{(d + 1)}$ (cf. [9]). Hence we may extract a subsequence $(N') \subset (N)$ such that the $\Phi^N$ converge to some limit $\Phi$ in the sense of weak convergence of measures, and, by Skorokhod’s imbedding theorem (cf. [10]), there exist processes

$$\Phi^N = (\tilde{\xi}^N, \tilde{X}^N, \tilde{F}^N, \tilde{G}^N, \tilde{\tilde{F}}^N, \tilde{\tilde{G}}^N, \tilde{\tilde{B}}^N),$$

$$\Phi = (\tilde{\xi}, \tilde{X}, \tilde{F}, \tilde{G}, \tilde{\tilde{F}}, \tilde{\tilde{G}}, \tilde{\tilde{B}}),$$

all defined on the same probability space, such that $\Phi^N$ is stochastically equivalent to $\Phi^N$ and $\Phi^N \to \Phi$ a.e. in the norm of $C^{(d + 1)}$. Note that all relations between the components of $\Phi^N$ are preserved for the $\Phi$; in particular, (3.1) and (3.2) continue to hold, and $\phi(t, \tilde{\xi}^N) = \tilde{X}^N$. For more details cf. [2, 3, 8, 9]. Hence, reebelling, let us assume henceforth that $\Phi^N \to \Phi$ a.e. in the norm of $C^{(d + 1)}$ for some limit process $\Phi = (\xi, X, F, G, B, \tilde{F}, \tilde{G}, \tilde{B})$. It then follows that

$$\xi_t = \xi_0 + F(t) + G(t) + B(t),$$

(3.7)

$$X_t = \phi(0, \xi_0) + \tilde{F}(t) + \tilde{G}(t) + \tilde{B}(t).$$

(3.8)
Moreover, the estimates (3.3)--(3.6) carry over to the limit:

\[ E\|\xi\|_m^2 \leq C_1(m), \quad E\|X\|_m^2 \leq C_2(m), \]

\[ E|H(t) - H(s)|^4 \leq C_7(t - s)^4, \quad E|M(t) - M(s)|^4 \leq C_7(t - s)^2 \]

(\(H\) any of \(F, G, \dot{F}, \dot{G}\), and \(M\) any of \(B, \dot{B}\)). Finally, for all \(t\),

\[ X_t = \varphi(t, \xi_t) \text{ a.e.} \]  

(3.11)

It can be shown (cf. [2, 8]) that \(\bar{B}(t) = (B(t)', \dot{B}(t)')'\) is a continuous vector valued martingale with respect to the filtration \((\bar{\mathcal{F}}_t) = (\sigma\{\xi_s, X_s, B(s), \dot{B}(s), s \leq t\})\) whose quadratic variation is given by \(\langle \bar{B} \rangle_t = \int_0^t \Sigma(s, \xi) \, ds\) with

\[ \Sigma(s, \xi) = \begin{pmatrix} a(s, \xi) & \nabla \varphi(s, \xi) \cdot a(s, \xi) \\ \nabla \varphi(s, \xi) \cdot a(s, \xi) & \nabla \varphi(s, \xi) \cdot \nabla \varphi(s, \xi) \end{pmatrix}. \]

Let \(C(s, \xi)\Lambda(s, \eta)C'(s, \xi)\) be a (measurable and \((\bar{\mathcal{F}}_t)\)-adapted) diagonalization of \(\Sigma(s, \xi)\). Adjoining an independent \((d + q)\)-dimensional Brownian motion, we obtain the representation

\[ \bar{B}(t) = \int_0^t C(s, \xi) \Lambda^{1/2}(s, \xi) \, d\hat{w}_s \]

for some Wiener process \((\hat{w}_n, \bar{\mathcal{F}}_t)\) (with respect to which \(\xi, X, B\) and \(\dot{B}\) are nonanticipative) (cf. [9, 12]), and, by Lemma 10.4 in [10], adjoining one more \(r\)-dimensional Brownian motion, we arrive at the representation

\[ B(t) = \int_0^t \sigma(s, \xi) \, d\hat{w}_s, \quad \dot{B}(t) = \int_0^t \nabla \varphi(s, \xi) \cdot \sigma(s, \xi) \, d\hat{w}_s, \]  

(3.12)

for some \(r\)-dimensional Wiener process \((\hat{w}_n, \bar{\mathcal{F}}_t)\) (again \(\xi, X, B\) and \(\dot{B}\) are adapted to \((\bar{\mathcal{F}}_t)\)). So we find that the martingale terms in (3.7) and (3.8) are of the desired form.

Let us now turn to the drift terms. Let \(H\) and \(h^N\) be defined as above. Since, by (A1)--(A3) (ii), (iv) and the uniform convergence \(\xi^N \to \xi\), \(\sup_{N, t} |h^N(t, \xi^N(\omega))| < \infty\) for almost all \(\omega\), it is easily established that \(H(t)\) is absolutely continuous with probability one:

\[ H(t) = \int_0^t \tilde{h}(s, \omega) \, ds \]

(cf. [8]), with measurable \(\tilde{h}\) (=any of \(\bar{f}, \bar{g}, \bar{f}, \bar{g}\)) and

\[ |\tilde{h}(t, \omega)|^4 = \lim_{n \to 0} \Delta^{-4} |H(t + \Delta) - H(t)|^4 \]

for almost all \((t, \omega)\), from which, by Fatou's lemma and (3.10),

\[ E \int_0^T |\tilde{h}(t, \omega)|^4 \, dt \leq C_7 T. \]  

(3.13)
Further, it is easily established from the uniform convergence $\xi^N \to \xi$ that

$$ F(t) = \int_0^t f(s, \xi) \, ds, $$

$$ \hat{F}(t) = \int_0^t \left[ \varphi'_s(s, \xi_s) + \nabla \varphi(s, \xi_s)f(s, \xi) + \frac{1}{2} \text{tr} \Delta \varphi(s, \xi_s) a(s, \xi) \right] \, ds. $$

From (3.7), (3.8) and (3.12) it follows then that $G(t)$ and $\hat{G}(t)$ are adapted to $(\mathcal{F}_t)$. Hence, by Lemma 5.2 in [10], we may assume that $\tilde{g}(s, \omega)$ and $\tilde{g}'(s, \omega)$ in the representations

$$ G(t) = \int_0^t \tilde{g}(s, \omega) \, ds, \quad \hat{G}(t) = \int_0^t \tilde{g}'(s, \omega) \, ds $$

are adapted to $(\mathcal{F}_t)$. Collecting the results, we find that

$$ X_t = \varphi(0, \xi_0) + \int_0^t \left[ \varphi'_s(s, \xi_s) + \nabla \varphi(s, \xi_s)f(s, \xi) + \frac{1}{2} \text{tr} \Delta \varphi(s, \xi_s) a(s, \xi) \right] \, ds $$

$$ + \int_0^t \tilde{g}(s, \omega) \, ds + \int_0^t \nabla \varphi(s, \xi_s) \sigma(s, \xi) \, d\tilde{w}_s $$

is an Itô process of type (2.1). Conditions (2.2)–(2.5) and (2.7) are satisfied by (A1)–(A4) together with (3.9) and (3.13). Consequently, by Lemma 1, $\mu_X \ll \mu_Y$, where $Y$ is given by (1.5). By virtue of (A4), $Y$, and hence $X$, is (for $t > 0$) absolutely continuous with respect to Lebesgue measure. Hence, by (A3)(iii), for $s > 0$

$$ \gamma^N(s, X^N_s) \to \gamma(s, X_s) \quad \text{a.e.,} $$

whence, by virtue of (3.11),

$$ g^N(s, \xi^N_s) \to g(s, \xi_s) $$

for almost all $(s, \omega)$. This implies that for all $t$

$$ G^N(t) \to G(t) = \int_0^t g(s, \xi_s) \, ds $$

in probability. So, finally,

$$ \xi_t = \xi_0 + \int_0^t f(s, \xi) \, ds + \int_0^t g'(s, \xi_s) \, ds + \int_0^t \sigma(s, \xi) \, d\tilde{w}_s \quad \text{a.e.} \quad (3.14) $$

for all $0 \leq t \leq T$, and, by continuity, (3.14) holds for all $t$ with probability one.

References


