István Gyöngy; Teresa Martínez

On stochastic differential equations with locally unbounded drift

Czechoslovak Mathematical Journal, Vol. 51 (2001), No. 4, 763–783

Persistent URL: `http://dml.cz/dmlcz/127685`

Czechoslovak Mathematical Journal, 51 (126) (2001), 763–783

ON STOCHASTIC DIFFERENTIAL EQUATIONS WITH LOCALLY UNBOUNDED DRIFT

István Gyöngy, Edinburgh, and Teresa Martínez, Madrid

(Received June 5, 2001)

Dedicated to Ivo Vrkoč on the occasion of his 70th birthday

Abstract. We study the regularizing effect of the noise on differential equations with irregular coefficients. We present existence and uniqueness theorems for stochastic differential equations with locally unbounded drift.

Keywords: stochastic differential equations, Krylov's estimate

MSC 2000: 60H10

1. Introduction

It is well-known that the ordinary differential equation

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = b\big(t, x(t)\big)$$

with initial condition $x(0) = x_0$ may have many solutions or may have no solution at all if b is not Lipschitz continuous. It is also known that for any bounded Borel function b one can regularize this equation by adding the white noise $\varepsilon\,\mathrm{d}W(t)/\mathrm{d}t$ to its right-hand side with any small constant $\varepsilon \neq 0$ and d-dimensional Wiener process W.

The work of the first author was partially supported by OTKA T 032932. The work of the second author was supported by the European TMR program "Harmonic Analysis", ERBFHRX-CT97-0159.

Namely, for any bounded Borel function $b\colon [0,\infty)\times\mathbb{R}^d\to\mathbb{R}^d$, vector $x_0\in\mathbb{R}^d$ and d-dimensional Wiener process W, there exists a unique solution of the equation

$$\frac{\mathrm{d}X(t)}{\mathrm{d}t}=b\big(t,X(t)\big)+\varepsilon\frac{\mathrm{d}W(t)}{\mathrm{d}t},\qquad X(0)=x_0 \tag{1.1}$$

for any $\varepsilon\neq 0$. This result is proved in [24] for $d=1$, and it is generalized in [22] to arbitrary dimension $d\geqslant 1$ for the stochastic differential equation

$$\mathrm{d}X(t)=b\big(t,X(t)\big)\,\mathrm{d}t+\sigma\big(t,X(t)\big)\mathrm{d}W(t),\ \ t\geqslant 0,\ \ X(0)=x_0\in\mathbb{R}^d, \tag{1.2}$$

with a bounded Borel function b and a bounded Lipschitz function σ, satisfying a non-degeneracy condition. In these theorems the existence and uniqueness of the solution is understood in the following sense. For any given probability space equipped with a Wiener process W there exists a unique non-anticipative transformation $X=F(W)$ of the given Wiener process W, which satisfies (1.2). In this case one says that equation (1.2) has a unique strong solution. We remark that by a recent result of [2] for equation (1.1), the uniqueness holds also in the class of pathwise solutions, obtained by solving (1.1) for (almost) every trajectory of W.

Equation (1.2) is often considered in a more general sense. We say that equation (1.2) has a solution in the wide sense if there exists a probability space $(\Omega,\mathscr{F},P)$ equipped with a Wiener martingale $\big(W(t),\mathscr{F}_t\big)$ and an $\mathscr{F}_t$-adapted process $X(t)$ such that equation (1.2) holds. This notion of solution is often called the weak solution in literature. The existence of a solution in this sense is known from Skorohod [19] if b and σ are continuous functions of the space variable and satisfy the linear growth condition. Moreover, the existence of such solutions and their uniqueness in law is known from Stroock and Varadhan if b is a bounded Borel function and σ is continuous in x, uniformly in t, and satisfies a non-degeneracy condition. By Krylov [9], [11] one knows that equation (1.2) always admits a weak solution when the coefficients b and σ are bounded measurable Borel functions and σ is nondegenerate.

In the present paper we are interested in solvability of equation (1.2) when b is locally unbounded. Such equations arise in stochastic mechanics, and have extensively been studied in literature (see e.g. [1], [14], [15], [16], [17], [21] and the references therein). Portenko [18] proves the existence and the uniqueness in law of the weak solution if $b\in L^p([0,T]\times\mathbb{R}^d)$ for some $p>d+2$ and σ is bounded, uniformly Hölder continuous and non-degenerate. This result is generalized in various directions in [3], [4], [21].

In this paper we generalize Portenko's result as follows. We prove the existence of a weak solution to equation (1.2) when σ is a bounded Borel function, $\sigma\sigma^*$ is strongly elliptic, and b is a Borel function such that

$$|b(t,x)|\leqslant K+F(t,x)\quad \mathrm{d}t\times\mathrm{d}x\text{-}a.e. \tag{1.3}$$

for some non-negative function $F \in L^{d+1}(\mathbb{R}_+ \times \mathbb{R}^d)$ and a constant $K \geqslant 0$. Moreover, if in addition to these assumptions σ is Lipschitz continuous and $b \in L^{2d+2}_{\mathrm{loc}}(\mathbb{R}_+ \times \mathbb{R}^d)$, then we get the existence and uniqueness of a strong solution. We present these results in Theorem 2.1 below. We derive this theorem from Theorem 2.2, in which we construct solutions via approximations. We note that this construction is the same as that used in Krylov [11] to obtain a solution in the wide sense to equation (1.2) with bounded measurable coefficients. The possibility to adapt this method to equations with locally unbounded drifts is indicated in [10], where an L^q-estimate on the distributions of continuous semimartingales is given, see Lemma 3.1 below, which plays a crucial role in the proof of Theorem 2.2. In order to show that for $b \in L^{2d+2}$ this construction gives a strong solution we first prove the pathwise uniqueness of the solution by adapting the approach of [22] and [24]. Hence a simple characterization of the convergence in probability in terms of tne convergence in distribution implies that the approximations converge in probability to a strong solution. This method, which is used for example in [7], [8], provides a constructive counterpart of the well-known result of Yamada and Watanabe, stating that the existence of a solution in the wide sense and the pathwise uniqueness imply the existence of a (unique) strong solution.

The paper is organized as follows. In Section 2 we formulate the main results of the paper. In Section 3 we present the main ingredients of the proofs. In Sections 4 and 5 the result on pathwise uniqueness and Theorem 2.2 are proved. In the last section we derive Theorem 2.1 from Theorem 2.2.

Finally, let us say a few words about the notation used in this paper. Except otherwise stated, C will denote a positive constant which may be different from one occurrence to another, $K \geqslant 0$ and $\delta \in (0,1)$ denote some fixed constants. We use the notation B_R for the open ball of radius R centered at the origin in $\mathbb{R}^d$, and $|\cdot|$ stands for the Hilbert-Schmidt norm either of a vector or a matrix. We denote by σ^* the transpose of a matrix σ. If F is a function defined on a domain $D \subset \mathbb{R}^d$ then $\|F\|_q$ denotes the $L^q(D)$-norm of F. For standard notation from the theory of partial differential equations we refer to [12]. Unless otherwise stated we use the summation convention with respect to repeated indices.

2. Main results

We consider the stochastic equation

$$X(t) = X_0 + \int_0^t b\big(s, X(s)\big)\,\mathrm{d}s + \int_0^t \sigma\big(s, X(s)\big)\,\mathrm{d}W(s) \tag{2.1}$$

on a complete probability space $(\Omega, \mathscr{F}, P)$ carrying a d_1-dimensional Wiener martingale $(W_t, \mathscr{F}_t)$, $t \geqslant 0$. Here X_0 is a vector in $\mathbb{R}^d$, b and σ are Borel measurable functions on $[0,\infty) \times \mathbb{R}^d$ with values in $\mathbb{R}^d$ and $\mathbb{R}^{d\times d_1}$, respectively. The stochastic integral is understood in Itô's sense.

Definition. An $\mathscr{F}_t$-adapted $\mathbb{R}^d$-valued continuous process $X = X(t)$ is called a solution of equation (2.1) if almost surely equation (2.1) holds for all $t \geqslant 0$. If for any probability space $(\Omega, \mathscr{F}, P)$ equipped with a Wiener martingale $(W_t, \mathscr{F}_t)$, $t \geqslant 0$ equation (2.1) has a solution then we say that (2.1) has a strong solution.

We say that equation (2.1) admits a weak solution if there exists a complete probability space $(\overline{\Omega}, \overline{\mathscr{F}}, \overline{P})$ carrying a d_1-dimensional Wiener martingale $(\overline{W}_t, \overline{\mathscr{F}}_t)$ and an $\overline{\mathscr{F}}_t$-adapted $\mathbb{R}^d$-valued continuous process $\overline{X} = \overline{X}(t)$, such that equation (2.1) holds $\overline{P}$-almost surely for all $t \geqslant 0$ with $\overline{X}$ and $\overline{W}$ in place of X and W, respectively.

We say that for equation (2.1) the pathwise uniqueness holds if for any probability space carrying a d_1-dimensional Wiener martingale $(W_t, \mathscr{F}_t)$, equation (2.1) cannot have more than one (strong) solution.

We assume the following conditions.

Assumption 2.1. *There exist a constant $K \geqslant 0$ and a non-negative function $F \in L^{d+1}(\mathbb{R}_+ \times \mathbb{R}^d)$ such that*

$$|b(t,x)| \leqslant K + F(t,x) \tag{2.2}$$

for $\mathrm{d}t \times \mathrm{d}x$-almost every $(t,x) \in \mathbb{R}_+ \times \mathbb{R}^d$.

Assumption 2.2. *There exists a constant $\delta \in (0,1)$ such that*

$$\delta I \leqslant \sigma\sigma^*(t,x) \leqslant \delta^{-1} I \tag{2.3}$$

for all $t \geqslant 0$, $x \in \mathbb{R}^d$, where I is the $d \times d$ identity matrix.

Assumption 2.3. *For every R there is a constant L_R such that*

$$|\sigma(t,x) - \sigma(t,y)| \leqslant L_R |x-y|$$

for all $t \geqslant 0$, $x \in \mathbb{R}^d$, $|x| \leqslant R$, $|y| \leqslant R$.

Remark 2.1. Notice that $|\sigma(t,x)| \leqslant \sqrt{d/\delta}$ for all t, x by Assumption 2.2.

Now we formulate the main result of the paper.

Theorem 2.1. *Suppose Assumptions* 2.1 *and* 2.2*. Then there exists a solution to equation* (2.1) *in the wide sense. If in addition to Assumptions* 2.1, 2.2, $b \in L^{2d+2}_{\mathrm{loc}}(\mathbb{R}_+ \times \mathbb{R}^d)$ *and Assumption* 2.3 *also holds, then equation* (2.1) *has a unique strong solution.*

The first statement of this theorem generalizes a result from [18]. The second statement generalizes an existence and uniqueness theorem from [22].

We will prove Theorem 2.1 by proving a limit theorem for the equation

$$X_n(t) = X_0 + \int_0^t b_n\big(s, X_n(s)\big)\,\mathrm{d}s + \int_0^t \sigma_n\big(s, X_n(s)\big)\,\mathrm{d}W(s) \tag{2.4}$$

as $n \to \infty$, where b_n and σ_n are Borel functions satisfying the following conditions.

Assumption 2.4. *There exist constants* $K \geqslant 0$, $\delta \in (0,1)$ *and a non-negative function* $F \in L^{d+1}(\mathbb{R}_+ \times \mathbb{R}^d)$ *such that for all* n

$$|b_n(t,x)| \leqslant K + F(t,x), \quad \delta I \leqslant \sigma_n\sigma_n^*(t,x) \leqslant \delta^{-1} I$$

for $\mathrm{d}t \times \mathrm{d}x$*-almost every* (t,x).

Assumption 2.5.

$$\lim_{n\to\infty} \sigma_n(t,x) = \sigma(t,x) \quad \text{and} \quad \lim_{n\to\infty} b_n(t,x) = b(t,x) \tag{2.5}$$

for $\mathrm{d}t \times \mathrm{d}x$*-almost every* (t,x).

Assumption 2.6. *For each* n *there is a constant* K_n *and for each* n *and* R *there exists a constant* $L_{n,R}$ *such that*

$$|b_n(t,x)| \leqslant K_n,$$
$$|b_n(t,z) - b_n(t,y)| + |\sigma_n(t,z) - \sigma_n(t,y)| \leqslant L_{n,R}|z-y|$$

for all $t \geqslant 0$ *and* $x, y, z \in \mathbb{R}^d$, $|z| \leqslant R$, $|y| \leqslant R$.

Notice that Assumptions 2.4 and 2.6 ensure the existence of a strong solution X_n to equation (2.4) by the classical existence and uniqueness theorem of Itô.

Theorem 2.2. *If Assumptions* 2.4, 2.5 *and* 2.6 *hold then* (X_n, W) *is tight in* $C([0,T]; \mathbb{R}^{d+d_1})$*, uniformly in* n*, for every finite* $T > 0$*. If* (X', W') *is the limit* *in distribution of a subsequence of* $(X_{n'}, W)$, *then there exists a probability space carrying a Wiener martingale* $\big(\tilde W(t), \tilde{\mathscr F}_t\big)$ *and an* $\tilde{\mathscr F}_t$*-adapted process* $\tilde X$ *such that* (X', W') *and* $(\tilde X, \tilde W)$ *have the same joint finite dimensional distributions, and*

$$\tilde X(t) = X_0 + \int_0^t b\big(s, \tilde X(s)\big)\,\mathrm{d}s + \int_0^t \sigma\big(s, \tilde X(s)\big)\,\mathrm{d}\tilde W(s) \tag{2.6}$$

almost surely for all $t \geqslant 0$. *If in addition to the above conditions,* $b \in L^{2d+2}_{\mathrm{loc}}(\mathbb{R}_+ \times \mathbb{R}^d)$ *and* σ *satisfies Assumption* 2.3, *then* X_n *converges in* $C([0,T])$ *in probability, for every* $T > 0$, *to a random process* X, *which is a strong solution of* (2.1).

The following result states the pathwise uniqueness for the solutions of equation (2.1), which we will use in the proof of Theorem 2.2.

Theorem 2.3. *Assume that* $b \in L^{2d+2}_{\mathrm{loc}}(\mathbb{R}_+ \times \mathbb{R}^d)$ *and that* σ *satisfies Assumptions* 2.2, 2.3. *Then for equation* (2.1) *the pathwise uniqueness holds.*

3. Preliminaries

First we invoke an important estimate, Lemma 5.1 from [10], on the distributions of continuous semimartingales.

Let $(\Omega, \mathscr F, P, \mathscr F_t)$ be a stochastic basis carrying a continuous $\mathbb{R}^d$-valued $\mathscr F_t$-local martingale $m = m_t$, a continuous increasing $\mathscr F_t$-adapted process $A = A(t)$ and a continuous $\mathbb{R}^d$-valued $\mathscr F_t$-adapted stochastic process $B = B(t)$ which has finite variation on finite intervals. Assume that $A(0) = 0$, $m(0) = B(0) = 0$ and $\mathrm{d}\langle m\rangle(t) \ll \mathrm{d}A(t)$. Let $r(t)$ and $c(t)$ be non negative progressively measurable stochastic processes such that

$$y(t) := \int_0^t r(s)\,\mathrm{d}A(s), \quad \varphi(t) := \int_0^t c(s)\,\mathrm{d}A(s)$$

are finite almost surely for all $t \geqslant 0$. Set $a^{ij}(t) := \mathrm{d}\langle m^i, m^j\rangle(t)/(2\,\mathrm{d}A(t))$, $X(t) := m(t) + B(t)$ and let τ^R be the first exit time of $X(t)$ from the ball B_R.

Lemma 3.1 [10]**.** *For every* $p \geqslant d$, *stopping time* γ *and nonnegative Borel function* $f\colon \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}_+$ *we have*

$$\begin{aligned} &E\int_0^{\gamma\wedge\tau^R} e^{-\varphi(t)}\big(c(t)^{p-d} r(t)\det a(t)\big)^{\frac{1}{p+1}} f\big(y(t), X(t)\big)\,\mathrm{d}A(t) \\ &\quad\leqslant N(d)(\mathbb{B}^2 + \mathbb{A})^{\frac{d}{2(p+1)}}\bigg(\int_0^\infty\int_{|x|\leqslant R} f^{p+1}(t,x)\,\mathrm{d}x\,\mathrm{d}t\bigg)^{\frac{1}{p+1}} \end{aligned}$$

where

$$\mathbb{A} := E\int_0^{\gamma\wedge\tau^R} \mathrm{e}^{-\varphi(t)}\,\mathrm{tr}\,a(t)\,\mathrm{d}A(t), \quad \mathbb{B} := E\int_0^{\gamma\wedge\tau^R} \mathrm{e}^{-\varphi(t)}|\mathrm{d}B(t)|,$$

and $N(d)$ is a constant depending only on the dimension d.

From this lemma we derive the following estimate for the occupation measure of the solutions of stochastic integral equations.

Corollary 3.2. *Let $X(t)$ be a solution of equation* (2.1) *under Assumptions* 2.1 *and* 2.2. *Then for any Borel function $f\colon \mathbb{R}^{d+1} \to \mathbb{R}_+$ and numbers $\lambda > 0$, $q \geqslant d+1$ we have*

$$E\int_0^\infty \mathrm{e}^{-\lambda t} f\big(t, X(t)\big)\,\mathrm{d}t \leqslant N\bigg(\int_0^\infty\int_{\mathbb{R}^d} f^q(t,x)\,\mathrm{d}t\,\mathrm{d}x\bigg)^{1/q}, \tag{3.1}$$

where N is a constant depending only on d, q, λ, δ, K and $\|F\|_{d+1}$, from Assumptions 2.1 *and* 2.2.

P r o o f. First notice that by a simple argument from [10] it suffices to show (3.1) when $q = d+1$. Indeed, if $q > d+1$ then by using Hölder's inequality with conjugate exponents $q/(d+1)$ and $q/(q-d-1)$ we get

$$E\int_0^\infty \mathrm{e}^{-\lambda t} f\big(t, X(t)\big)\,\mathrm{d}t \leqslant C\bigg(E\int_0^\infty \mathrm{e}^{-\lambda t}\big|f\big(t,X(t)\big)\big|^{q/(d+1)}\,\mathrm{d}t\bigg)^{(d+1)/q}$$

with some constant C depending only on λ, q and d. Hence for $q > d+1$ estimate (3.1) follows from the same estimate for $q = d+1$. Notice also that because of the shift invariance of the Lebesgue measure we may assume that $X_0 = 0$. Applying Lemma 3.1 with $A(t) = t$, $\mathrm{d}B(t) = b\big(t, X(t)\big)\,\mathrm{d}t$, $\mathrm{d}m(t) = \sigma\big(t, X(t)\big)\,\mathrm{d}W(t)$, $r(t) = 1$ and $c(t) = \lambda$ we get

$$\begin{aligned} E\int_0^{\gamma_n\wedge\tau^R} \mathrm{e}^{-\lambda t} f\big(t, X(t)\big)\,\mathrm{d}t &\leqslant 2^{\frac{1}{d+1}}\delta^{\frac{-d}{d+1}} N(d)(\mathbb{B}_n^2 + \mathbb{A}_n)^{\frac{d}{2(d+1)}} \\ &\quad\times\bigg(\int_0^\infty\int_{|x|\leqslant R} |f(t,x)|^{d+1}\,\mathrm{d}x\,\mathrm{d}t\bigg)^{\frac{1}{(d+1)}} \end{aligned} \tag{3.2}$$

where

$$\mathbb{A}_n = E\int_0^{\gamma_n\wedge\tau^R} \mathrm{e}^{-\lambda t}\frac{1}{2}|\sigma|^2\,\mathrm{d}t \leqslant \frac{d}{2\delta}E\int_0^\infty \mathrm{e}^{-\lambda t}\,\mathrm{d}t \leqslant \frac{d}{2\lambda\delta},$$

$$\mathbb{B}_n = E\int_0^{\gamma_n\wedge\tau^R} \mathrm{e}^{-\varphi(t)}|\mathrm{d}B(t)| = E\int_0^{\gamma_n\wedge\tau^R} \mathrm{e}^{-\lambda t}\big|b\big(t,X(t)\big)\big|\,\mathrm{d}t,$$

and

$$\gamma_n := \inf\left\{t\geqslant 0\colon \int_0^t \big|b\big(s,X(s)\big)\big|\,\mathrm{d}s \geqslant n\right\}$$

for integers $n>0$. Clearly $\mathbb{B}_n$ is finite, and we need only to show that $\mathbb{B}_n$ is bounded by a constant N depending only on λ, d, δ, K and $\|F\|_{d+1}$. By Assumption 2.1 and Lemma 3.1

$$\begin{aligned}\mathbb{B}_n &\leqslant \frac{K}{\lambda} + E\int_0^{\gamma_n\wedge\tau^R} \mathrm{e}^{-\lambda t}F\big(t,X(t)\big)\,\mathrm{d}t \\ &\leqslant \frac{K}{\lambda} + N(\mathbb{B}_n^2+\mathbb{A}_n)^{\frac{d}{2(d+1)}}\|F\|_{d+1} \leqslant \frac{K}{\lambda} + N\mathbb{B}_n^{\frac{d}{d+1}}\|F\|_{d+1} + N\mathbb{A}_n^{\frac{d}{2d+2}}\|F\|_{d+1}\end{aligned}$$

with a constant N which is independent of R and n. Hence using the inequality $\mathbb{B}_n^{d/(d+1)} \leqslant \varepsilon^{(d+1)/d}\mathbb{B}_n + \varepsilon^{-d-1}$ with sufficiently small $\varepsilon>0$, we get $\mathbb{B}_n \leqslant \frac{1}{2}\mathbb{B}_n + N$ with a constant $N = N(\lambda,d,\delta,K,\|F\|_{d+1})$. Hence $\mathbb{B}_n \leqslant 2N$, and by (3.2)

$$E\int_0^{\gamma_n\wedge\tau^R} \mathrm{e}^{-\lambda t}f\big(t,X(t)\big)\,\mathrm{d}t \leqslant N\bigg(\int_0^\infty\int_{|x|\leqslant R} |f(t,x)|^{d+1}\,\mathrm{d}x\,\mathrm{d}t\bigg)^{1/(d+1)}$$

for all integers $n>0$, where N is a constant depending only on $\lambda,d,\delta,K,\|F\|_{d+1}$. Letting here $n\to\infty$ and $R\to\infty$ we get (3.1) with $q=d+1$ by Fatou's lemma. $\square$

The following lemma generalizes Krylov's extension of Itô's formula from [11]. To formulate it we recall that for a bounded domain $D\subset\mathbb{R}^d$ and a number $q\geqslant 1$ the Sobolev space $W_q^{1,2}\big((0,T)\times D\big)$ is defined as the Banach space of functions $u\colon (0,T)\times D\to\mathbb{R}$ whose generalized derivatives $D_t^\alpha D_x^\beta u$ are in $L^q\big((0,T)\times D\big)$ for all multi-indices (α,β) such that $2\alpha+|\beta| := 2\alpha+\sum_{i=1}^d \beta_i \leqslant 2$. We use the notation

$$\|u\|_{W_q^{1,2}} := \sum_{2\alpha+|\beta|\leqslant 2} \|D_t^\alpha D_x^\beta u\|_q$$

for the norm of u in $W_q^{1,2}\big((0,T)\times D\big)$, where $\|v\|_q$ denotes the L^q norm of v. (See, e.g., [12].)

Lemma 3.3. *Assume Assumptions* 2.1 *and* 2.2. *Let* $X(t)$ *be a solution to equation* (2.1), *and let* $R>0$ *be such that* $X_0\in B_R$. *Then for any* $u\colon [0,T]\times\mathbb{R}^d\longrightarrow\mathbb{R}$ *from the Sobolev space* $W_q^{1,2}\big((0,T)\times B_R\big)$, $q>d+2$ *we have*

$$
\begin{aligned}
u\big(t,X(t)\big) = u(0,X_0) &+ \int_0^t u_t\big(s,X(s)\big)\,\mathrm{d}s \\
&+ \int_0^t \big[b_i\big(s,X(s)\big)u_{x_i}\big(s,X(s)\big) \\
&+ \tfrac{1}{2}(\sigma\sigma^*)_{ij}\big(s,X(s)\big)u_{x_ix_j}\big(s,X(s)\big)\big]\,\mathrm{d}s \\
&+ \int_0^t u_{x_i}\big(s,X(s)\big)\sigma_{ij}\big(s,X(s)\big)\,\mathrm{d}W^j(s)
\end{aligned}
\tag{3.3}
$$

almost surely for $t\leqslant T\wedge\tau^R$, *where* τ^R *is the exit time of* X *from the ball* B_R.

Proof. By making use of Corollary 3.2 we can prove this lemma in the same way as Theorem 2.10.1 in [11] is proved. For the convenience of the reader we give the details of the proof. First we show that each integral in (3.3) is well-defined. To this end we note that by Sobolev's embedding there exists a constant N such that

$$\sup_{t\in[0,T]}\sup_{x\in B_R}\Big(|u(t,x)|+\sum_i|u_{x_i}(t,x)|\Big)\leqslant N\|u\|_{W_q^{1,2}}$$

for all $u\in W_q^{1,2}\big([0,T]\times B_R\big)$, $q>d+2$. (See, e.g., Lemma II.3.3 in [12].) Hence

$$E\int_0^{\tau^R\wedge T}\sigma_{ij}^2\big(s,X(s)\big)\big(u_{x_i}\big(s,X(s)\big)\big)^2\,\mathrm{d}s\leqslant K^2T\|u\|^2_{W_q^{1,2}} \tag{3.4}$$

by the boundedness of σ, and

$$
\begin{aligned}
E\int_0^{T\wedge\tau^R}&\big|b_i\big(s,X(s)\big)u_{x_i}(s,X_s)\big|\,\mathrm{d}s \\
&\leqslant C\|u\|_{W_q^{1,2}}\Big(E\int_0^{T\wedge\tau^R}F(s,X_s)\,\mathrm{d}s+KT\Big) \\
&\leqslant C\|u\|_{W_q^{1,2}}(N\mathrm{e}^T\|F\|_{L^q}+KT)
\end{aligned}
\tag{3.5}
$$

by Assumption 2.1 and Corollary 3.2. Moreover,

$$E\int_0^{T\wedge\tau^R}\big|u_t\big(s,X(s)\big)\big|\,\mathrm{d}s\leqslant N\mathrm{e}^T\|u_t\|_{L^q}\leqslant N\mathrm{e}^T\|u\|_{W_q^{1,2}}, \tag{3.6}$$

$$E\int_0^{T\wedge\tau^R}\big|(\sigma\sigma^*)_{ij}\big(s,X(s)\big)u_{x_ix_j}\big(s,X(s)\big)\big|\,\mathrm{d}s \tag{3.7}$$

$$\leqslant K^2E\int_0^{T\wedge\tau^R}\big|u_{x_ix_j}\big(s,X(s)\big)\big|\,\mathrm{d}s\leqslant K^2N\mathrm{e}^T\|u_{x_ix_j}\|_{L^q}\leqslant K^2N\mathrm{e}^T\|u\|_{W_q^{1,2}}$$

by Assumptions 2.1, 2.2 and Corollary 3.2. Consequently, the right-hand side of formula (3.3) is well-defined for $t \leqslant T \wedge \tau^R$.

For $u \in W_q^{1,2}\big((0,T) \times B_R\big)$ there exists a sequence of functions $u^{(n)}$ from $C^{1,2}\big((0,T) \times \mathbb{R}^d\big)$ such that $u^{(n)} \to u$ in the norm of $W_q^{1,2}\big((0,T) \times B_R\big)$ and also for almost every $(t,x) \in (0,T) \times B_R$. By Itô's formula we have

$$
\begin{aligned}
(3.8)\qquad & u^{(n)}\big(t \wedge \gamma, X(t \wedge \gamma)\big) = u^{(n)}(0, X_0) \\
&+ \int_0^{t\wedge\gamma} u^{(n)}_{x_i}\big(s, X(s)\big)\sigma_{ij}\big(s, X(s)\big)\,\mathrm{d}W^j(s) + \int_0^{t\wedge\gamma} u^{(n)}_t\big(s, X(s)\big)\,\mathrm{d}s \\
&+ \int_0^{t\wedge\gamma}\Big[b_i\big(s, X(s)\big)u^{(n)}_{x_i}(s, X_s) + \frac{1}{2}(\sigma\sigma^*)_{ij}\big(s, X(s)\big)u^{(n)}_{x_i x_j}\big(s, X(s)\big)\Big]\,\mathrm{d}s
\end{aligned}
$$

for all $t \geqslant 0$, where $\gamma := T \wedge \tau^R$. Notice that inequalities (3.4), (3.5), (3.6) and (3.7) hold also with $u - u^{(n)}$ in place of u, with constants independent of n. Hence letting $n \to \infty$ in (3.8) we obtain (3.3). □

In the proof of the pathwise uniqueness we will adapt the method of transformation of the phase space (see [24]). To this end we consider the system of partial differential equations

$$
(3.9)\qquad u^k_t(t,x) + a^{ij}(t,x)u^k_{x^i x^j}(t,x) + b^i(t,x)u^k_{x^i}(t,x) = 0, \quad k = 1, 2, \ldots, d
$$

in the domain $\{(t,x) \in [0,T] \times \mathbb{R}^d\}$, with a boundary condition

$$
(3.10)\qquad u(T,x) = x, \quad x \in \mathbb{R}^d
$$

for a function $u = (u^k)\colon [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$, where $a^{ij} = \frac{1}{2}\sigma^{ik}\sigma^{jk}$, $b = (b^i)$ and $\sigma = (\sigma^{ij})$ are the drift and diffusion coefficients in equation (2.1).

Lemma 3.4. *Assume that $b \in L^{2d+2}(\mathbb{R}_+ \times \mathbb{R}^d)$ and that σ satisfies Assumptions 2.2 and 2.3. Then Problem (3.9)–(3.10) has a solution such that $u \in W^{1,2}_{2d+2}\big((0,T) \times B_R\big)$ for every $R > 0$. Moreover, if $T = T_0$ is a sufficiently small positive number, then there exists a constant N such that*

$$
|u(t,x) - u(t,y)| \geqslant N|x - y|
$$

for all $t \in [0, T_0]$ and $x, y \in \mathbb{R}^d$.

P r o o f. One can get the existence of the solution $u \in W^{1,2}_{2d+2}\big((0,T) \times B_R\big)$ for every $R > 0$ by Theorem IV.9.1 from [12], in which the condition on the continuity of $a(t,x) := \frac{1}{2}\sigma\sigma^*(t,x)$ in (t,x) can be relaxed by using the method of [23]. Hence in the same way as in [24] one can show that $u_x = (u^k_{x^i})$ is Hölder continuous in $(t,x) \in [0,T] \times \mathbb{R}^d$. In particular, there exist constants $\kappa > 0$, $\alpha \in (0,1)$ such that

$$|u_x(s,y) - u_x(t,y)| \leqslant \kappa |s-t|^\alpha$$

for all $s,t \in [0,T]$, $y \in \mathbb{R}^d$. Hence, noticing that $u_x(T,x) = I$, we have

$$\begin{aligned}
|u(t,x) - u(t,y)| &= \left| \int_0^1 u_x\big(t, \lambda x + (1-\lambda)y\big)(x-y)\,\mathrm{d}\lambda \right| \\
&\geqslant |x-y| \left(1 - \int_0^1 \big| u_x(t, \lambda x + (1-\lambda)y) - u_x\big(T, \lambda x + (1-\lambda)y\big) \big| \,\mathrm{d}\lambda \right) \\
&\geqslant |x-y| \left(1 - \int_0^1 \kappa |T-t|^\alpha \,\mathrm{d}\lambda \right) \\
&\geqslant |x-y|(1 - \kappa T^\alpha).
\end{aligned}$$

We can complete the proof by choosing $T := T_0 > 0$ such that $\kappa T_0^\alpha < 1$. □

We will make use of the following generalization of Gronwall's lemma. Its proof can be found, for example, in [6].

Lemma 3.5. *Let $Z(t)$ be an $\mathscr{F}_t$-adapted continuous process such that*

$$0 \leqslant Z(t) \leqslant \int_0^t Z(s)\,\mathrm{d}A(s) + M(t)$$

for all $t \geqslant 0$, where $A(t)$ is a continuous $\mathscr{F}_t$-adapted increasing process and $\big(M(t), \mathscr{F}_t\big)$ is a continuous local martingale such that $M(0) = A(0) = 0$. Then $Z(t) = 0$ (a.s.) for all $t \geqslant 0$.

4. Pathwise uniqueness

Let X^0 and X^1 be solutions of equation (2.1) on the same probability space with the same Wiener-martingale. By considering the processes $X^0(t) - X_0$, $X^1(t) - X_0$ and the coefficients $b(\cdot, X_0 + \cdot)$, $\sigma(\cdot, X_0 + \cdot)$ in place of $X^0(t)$, $X^1(t)$ and $b(\cdot,\cdot)$, $\sigma(\cdot,\cdot)$, respectively, we may assume that $X^0(0) = X^1(0) = 0$. Define for fixed $\varepsilon > 0$ and $R > 0$ the stopping times

$$\begin{aligned}
\tau^R &= \inf\{t \geqslant 0\colon\ |X^0(t)| + |X^1(t)| > R\}, \\
\tau_\varepsilon &= \inf\{t \geqslant 0\colon\ |X^0(t) - X^1(t)| > \varepsilon\},
\end{aligned}$$

and set $\tau^R_\varepsilon := \tau^R \wedge \tau_\varepsilon$ for $\varepsilon = \delta/(2L_R\sqrt{d/\delta})$, where δ and L_R are from Remark 2.1 and Assumption 2.3. Note that by Corollary 3.2

$$E\int_0^{T\wedge\tau^R} \mathrm{e}^{-\lambda t} f\bigl(t, X^i(t)\bigr)\,\mathrm{d}t \leqslant N\|f\|_{L^q((0,T)\times B_R)}, \quad i = 0, 1 \tag{4.1}$$

for $q \geqslant d+1$, for any $T > 0$ and any non-negative Borel function f, where N is a constant depending on d, q, λ, δ, K and $\|F\|_{d+1}$.

Set $X^\alpha(t) = \alpha X^1(t) + (1-\alpha)X^0(t)$ for $\alpha \in [0,1]$, $t \geqslant 0$. Then by the following lemma, estimate (4.1) holds also for X^α, with τ^R_ε in place of τ^R.

Lemma 4.1. *For any Borel function $f\colon \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}_+$, constants $q \geqslant d+1$, $\lambda > 0$ and $T > 0$ we have*

$$E\int_0^{\tau^R_\varepsilon\wedge T} \mathrm{e}^{-\lambda t} f\bigl(t, X^\alpha(t)\bigr)\,\mathrm{d}t \leqslant N\|f\|_{L^q((0,T)\times B_R)} \tag{4.2}$$

for all $\alpha \in [0,1]$, where N is a constant depending only on d, q, λ, δ, K and $\|F\|_{d+1}$.

P r o o f. Clearly

$$X^\alpha(t) = X_0 + \int_0^t b_\alpha(s)\,\mathrm{d}s + \int_0^t \sigma_\alpha(s)\,\mathrm{d}W(s),$$

where

$$\begin{aligned} b_\alpha(t) &= \alpha b\bigl(t, X^1(t)\bigr) + (1-\alpha)b\bigl(t, X^0(t)\bigr),\\ \sigma_\alpha(t) &= \alpha\sigma\bigl(t, X^1(t)\bigr) + (1-\alpha)\sigma\bigl(t, X^0(t)\bigr). \end{aligned}$$

For every $z \in \mathbb{R}^d$ we have

$$\begin{aligned} z^*\sigma_\alpha\sigma_\alpha^*(t)z &= \alpha^2 z^*\sigma\sigma^*\bigl(t, X^1(t)\bigr)z + (1-\alpha)^2 z^*\sigma\sigma^*\bigl(t, X^0(t)\bigr)z\\ &\quad + 2\alpha(1-\alpha)z^*\bigl[\sigma\bigl(t, X^1(t)\bigr)\sigma^*\bigl(t, X^0(t)\bigr)\bigr]z\\ &\geqslant \delta|z|^2 + 2\alpha(1-\alpha)z^*\Bigl[\bigl[\sigma\bigl(t, X^1(t)\bigr) - \sigma\bigl(t, X^0(t)\bigr)\bigr]\sigma^*\bigl(t, X^0(t)\bigr)\Bigr]z \end{aligned}$$

for all t, ω. By Remark 2.1 and Assumption 2.3

$$\begin{aligned} &\bigl|z^*\bigl[\sigma\bigl(t, X^1(t)\bigr) - \sigma\bigl(t, X^0(t)\bigr)\bigr]\sigma^*\bigl(t, X^0(t)\bigr)z\bigr|\\ &\leqslant L_R\sqrt{d/\delta}|X^1(t) - X^0(t)|\,|z|^2 \leqslant \frac{\delta}{2}|z|^2 \end{aligned}$$

for $t \leqslant \tau_\varepsilon^R$ and for all z. Hence $a(t) := \frac{1}{2}\sigma_\alpha\sigma_\alpha^*(t) \geqslant \frac{\delta}{4}I$ for all $t \leqslant \tau_\varepsilon^R$. Applying Lemma 3.1 with $r(t) = 1$, $c(t) = \lambda$, $A(t) = t$, $\mathrm{d}B(t) = b_\alpha(t)\,\mathrm{d}t$, $\mathrm{d}M(t) = \sigma_\alpha(t)\,\mathrm{d}W(t)$ we have

$$E\int_0^{T\wedge\tau_\varepsilon^R} \mathrm{e}^{-\lambda t} f\big(t, X^\alpha(t)\big)\,\mathrm{d}t \leqslant \lambda^{(d+1-q)/q}\Big(\frac{4}{\delta}\Big)^{d/q} N(d)(\mathbb{B}^2 + \mathbb{A})^{d/2q}\|f\|_{L^q((0,T)\times B_R)},$$

where

$$\mathbb{A} := E\int_0^{T\wedge\tau_\varepsilon^R} \mathrm{e}^{-\lambda t}\,\mathrm{tr}\,a(t)\,\mathrm{d}t = E\int_0^{T\wedge\tau_\varepsilon^R} \mathrm{e}^{-\lambda t}\,\tfrac{1}{2}|\sigma_\alpha|^2\,\mathrm{d}t \leqslant \frac{d}{2\delta\lambda},$$

and

$$\begin{aligned}\mathbb{B} &:= E\int_0^{T\wedge\tau_\varepsilon^R} \mathrm{e}^{-\lambda t}|b_\alpha(s)|\,\mathrm{d}s\\ &\leqslant \alpha E\int_0^{T\wedge\tau^R} \mathrm{e}^{-\lambda t}\big|b\big(t, X^1(t)\big)\big|\,\mathrm{d}t + (1-\alpha)E\int_0^{T\wedge\tau^R} \mathrm{e}^{-\lambda t}\big|b\big(t, X^0(t)\big)\big|\,\mathrm{d}t\\ &\leqslant \frac{K}{\lambda} + N\|F\|_{L^{d+1}((0,T)\times B_R)}\end{aligned}$$

by $|\sigma_\alpha|^2 \leqslant d/\delta$ and by estimate (4.1). □

P r o o f of Theorem 2.3. For fixed R let b_R denote the function which is equal to b in the ball B_R and to 0 elsewhere. Take $T_0 > 0$ and the function u from Lemma 3.4 with b_R in place of b. Then by Lemma 3.4

$$|X^1(t) - X^0(t)| \leqslant N\big|u\big(t, X^1(t)\big) - u\big(t, X^0(t)\big)\big|,$$

and by Itô's formula (Lemma 3.3)

$$\mathrm{d}u^k\big(t, X^i(t)\big) = u^k_{x_j}\big(t, X^i(t)\big)\sigma_{jl}\big(t, X^i(t)\big)\mathrm{d}W^l(t), \quad i = 0, 1$$

for $t \in (0, T_0 \wedge \tau^R]$. Hence by Itô's formula for $\big|u\big(t, X^1(t)\big) - u\big(t, X^0(t)\big)\big|^2$ we get

$$\begin{aligned}|X^1(t\wedge\gamma) - X^0(t\wedge\gamma)|^2 \leqslant N^2 &\int_0^{t\wedge\gamma} \sum_k \big|\sigma^* u^k_x\big(s, X^1(s)\big) - \sigma^* u^k_x\big(s, X^0(s)\big)\big|^2\,\mathrm{d}s\\ &+ 2N^2\int_0^{t\wedge\gamma} \big[u^k\big(s, X^1(s)\big) - u^k\big(s, X^0(s)\big)\big]\\ &\times \big[\sigma_{ij}u^k_{x_i}\big(s, X^1(s)\big) - \sigma_{ij}u^k_{x_i}\big(s, X^0(s)\big)\big]\,\mathrm{d}W^j(s)\end{aligned}$$

for all $t \geqslant 0$, where $\gamma := \tau^R \wedge \tau_\varepsilon \wedge T_0$. Since $u \in W^{1,2}_{2d+2}(D)$ with $D = (0, T_0) \times B_R$,

$$\sup_{t\in[0,T_0]}\sup_{x\in B_R} |u_x(t,x)| < \infty$$

by Lemma II.3.3 from [12]. Hence by the boundedness and Lipschitz continuity of σ

$$
\begin{aligned}
&\big|\sigma^* u_x^k\big(s, X^1(s)\big) - \sigma^* u_x^k\big(s, X^0(s)\big)\big|^2 \\
&\leqslant C\big|\sigma^*\big(s, X^1(s)\big) - \sigma^*\big(s, X^0(s)\big)\big|^2 \big|u_x^k\big(s, X^1(s)\big)\big|^2 \\
&\quad + C\big|\sigma^*\big(s, X^0(s)\big)\big|^2 \big|u_x^k\big(s, X^1(s)\big) - u_x^k\big(s, X^0(s)\big)\big|^2 \\
&\leqslant C|X^1(s) - X^0(s)|^2 \\
&\quad + C\sum_i \bigg|\Big\langle X^1(s) - X^0(s), \int_0^1 \nabla_x u_x^k\big(s, \alpha X^1(s) + (1-\alpha)X^0(s)\big)\,\mathrm{d}\alpha\Big\rangle\bigg|^2.
\end{aligned}
$$

Consequently,

$$
|X^1(t\wedge\gamma) - X^0(t\wedge\gamma)|^2 \leqslant \int_0^t |X^1(s\wedge\gamma) - X^0(s\wedge\gamma)|^2\,\mathrm{d}A(s) + M(t),
$$

where $M(t)$ is a continuous martingale and

$$
A(t) := C\int_0^{t\wedge\tau} \bigg[\sum_{k,i,j}\int_0^1 \big|u_{x^i x^j}^k\big(s, Z^\alpha(s)\big)\big|^2\,\mathrm{d}\alpha + 1\bigg]\,\mathrm{d}s
$$

with

$$
Z^\alpha(s) := \alpha X^1(s) + (1-\alpha)X^0(s).
$$

By Lemma 4.1 we have

$$
\begin{aligned}
&E\int_0^{t\wedge\gamma}\int_0^1 \big|\nabla_x u_{x^i}^k\big(s, Z^\alpha(s)\big)\big|^2\,\mathrm{d}\alpha\,\mathrm{d}s \\
&\leqslant C\sum_{j=1}^d \int_0^1 \|\,|u_{x^i x^j}^k|^2\|_{L^{d+1}(D)}\,\mathrm{d}\alpha \leqslant C\|u\|_{W^{1,2}_{2d+2}(D)} < \infty,
\end{aligned}
$$

which implies that almost surely $A(t) < \infty$ for all $t > 0$. Therefore $Z(t) := |X^1(t\wedge\gamma) - X^0(t\wedge\gamma)|^2 = 0$ by Lemma 3.5, and we have $X^1(t) = X^0(t)$ for $t \leqslant \gamma$. Hence $X^1(t) = X^0(t)$ for $t \leqslant T_0 \wedge \tau^R$ by the definition of τ_ε. Since R can be as large as we want, $X^1(t) = X^0(t)$ for $t \leqslant T_0$. Hence we get $X^1(t) = X^0(t)$ for all $t \geqslant 0$ by the standard argument of shifting the time.

5. Proof of Theorem 2.2

We adapt the method of the proof of Theorem 6.1 from [11]. First we need two lemmas due to Skorohod. For their proof we refer to [19].

Lemma 5.1. *Let $\{\psi_n\}_{n=1}^{\infty}$ be a sequence of d-dimensional processes defined on some probability space. Assume that for each $T \geqslant 0$,*

$$\lim_{c\to\infty} \sup_n \sup_{t\leqslant T} P(|\psi_n(t)| > c) = 0, \tag{5.1}$$

$$\lim_{\theta\downarrow 0} \sup_n \sup_{t,s} \bigl\{ P\bigl(|\psi_n(t) - \psi_n(s)| > \varepsilon\bigr) \colon\ t, s \leqslant T,\ \ |t-s| \leqslant \theta \bigr\} = 0 \tag{5.2}$$

for every $\varepsilon > 0$. Then there exist a sequence $\{n_k\}_{k=1}^{\infty}$, a probability space and random processes X, X_k ($k \geqslant 1$), such that the finite dimensional distributions of the processes $\psi_{n_k}(t)$ and $X_k(t)$ coincide, and $X_k(t)$ converges to $X(t)$ in probability for every $t \geqslant 0$.

Lemma 5.2. *Let $\{\eta_n\}_{n=1}^{\infty}$, $\eta(t)$ be uniformly bounded $\mathbb{R}^{d\times k}$-valued random processes and let W_n, W be Wiener processes such that the stochastic Itô integrals $I_n(t) = \int_0^t \eta_n(s)\,\mathrm{d}W_n(s)$ and $I(t) = \int_0^t \eta(s)\,\mathrm{d}W(s)$ are well-defined. Assume that $\eta_n(t) \to \eta(t)$ and $W_n(t) \to W(t)$ in probability for every $t \geqslant 0$. Then*

$$\lim_{n\to\infty} P\Bigl(\sup_{t\in[0,T]} |I_n(t) - I(t)| \geqslant \varepsilon\Bigr) = 0$$

for every $\varepsilon > 0$ and $T > 0$.

We show now that the processes (X_n, W) satisfy the assumptions of Lemma 5.1. Obviously we need to verify the assumptions only for X_n. To this end notice that estimate (3.1) of Corollary 3.2 holds for the process X_n with the same constant N for all n, due to Assumption 2.4. Hence by using the Davis inequality we obtain

$$\begin{aligned} E|X_n(t)| &\leqslant |X_0| + E\sup_{s\leqslant T}\left|\int_0^s \sigma_n\bigl(u, X_n(u)\bigr)\,\mathrm{d}W(u)\right| + E\int_0^T \bigl|b_n\bigl(s, X_n(s)\bigr)\bigr|\,\mathrm{d}s \\ &\leqslant |X_0| + 3E\left(\int_0^T \bigl|\sigma_n\bigl(s, X_n(s)\bigr)\bigr|^2\,\mathrm{d}s\right)^{1/2} + \mathrm{e}^T N\|F\|_{d+1} + KT \\ &\leqslant L, \end{aligned}$$

where $\|F\|_{d+1}$ is the $L^{d+1}([0,T]\times\mathbb{R}^d)$-norm of F and L is a constant which does not depend on n and t for $t \leqslant T$. By Chebyshev's inequality

$$\sup_{t\leqslant T} P(|X_n(t)| \geqslant C) \leqslant \frac{1}{C}\sup_{t\leqslant T} E(|X_n(t)|) \leqslant \frac{L}{C},$$

which implies (5.1). Similarly, for $0 \leqslant s \leqslant t \leqslant T$ we have

$$\begin{aligned}
E|X_n(t) - X_n(s)| &\leqslant E\left|\int_s^t \sigma_n\big(u, X_n(u)\big)\,\mathrm{d}W(u)\right| + E\int_s^t \big|b_n\big(u, X_n(u)\big)\big|\,\mathrm{d}u \\
&\leqslant 3E\bigg(\int_s^t \big|\sigma_n\big(u, X_n(u)\big)\big|^2\,\mathrm{d}u\bigg)^{\frac{1}{2}} \\
&\quad + E\int_s^t F\big(u, X_n(u)\big)\,\mathrm{d}u + K(t-s) \\
&\leqslant 3K|t-s|^{1/2} + N\mathrm{e}^T\bigg(\int_s^t\int_{\mathbb{R}^d} |F(u,x)|^{d+1}\,\mathrm{d}x\,\mathrm{d}u\bigg)^{\frac{1}{d+1}} \\
&\quad + K|t-s|.
\end{aligned}$$

Hence by Chebyshev's inequality we obtain (5.2), since

$$\lim_{\vartheta\to 0}\sup\bigg\{\int_s^t\int_{\mathbb{R}^d} |F(u,x)|^{d+1}\,\mathrm{d}x\,\mathrm{d}u\colon\ s \leqslant t \leqslant T,\ \ |t-s| \leqslant \vartheta\bigg\} = 0$$

due to $\|F\|_{d+1} < \infty$. By virtue of Lemma 5.1 we have a subsequence of (X_n, W), which we keep denoting by (X_n, W), a sequence of random processes $(\tilde{X}_n, \tilde{W}_n)$ and a process $(\tilde{X}, \tilde{W})$ such that the finite dimensional distributions of $(\tilde{X}_n, \tilde{W}_n)$ and (X_n, W) coincide, and $\lim\limits_{n\to\infty} \tilde{X}_n(t) = \tilde{X}(t)$, $\lim\limits_{n\to\infty} \tilde{W}_n(t) = \tilde{W}(t)$ in probability for every $t \geqslant 0$. Define $\tilde{\mathscr{F}}_t^n$ as the completion of the σ-algebra generated by the random variables $\{\tilde{X}_n(s), \tilde{W}_n(s);\ \ s \leqslant t\}$. Obviously, $\tilde{X}_n(t)$ is $\tilde{\mathscr{F}}_t^n$-adapted and since it is stochastically continuous, we may assume that it is a separable process. Clearly $\tilde{W}_n$ is a Wiener process and $\tilde{\mathscr{F}}_t^n$ is independent of the increments of $\tilde{W}_n$ after time t. Thus, $\big(\tilde{W}_n(t), \tilde{\mathscr{F}}_t^n\big)$ is a Wiener martingale. Consequently, the stochastic integral $\int_0^t \sigma_n(s, \tilde{X}_n(s))\,\mathrm{d}\tilde{W}_n(s)$ is well-defined. It is not difficult to show that

$$\tilde{X}_n(t) = X_0 + \int_0^t b_n\big(s, \tilde{X}_n(s)\big)\,\mathrm{d}s + \int_0^t \sigma_n\big(s, \tilde{X}_n(s)\big)\,\mathrm{d}\tilde{W}_n(s) \tag{5.3}$$

for every $t \geqslant 0$, since the finite dimensional distributions of (X_n, W) and $(\tilde{X}_n, \tilde{W}_n)$ are the same and X_n satisfies equation (2.4). Now we are going to take the limit $n \to \infty$ in equation (5.3). Let us fix $T > 0$ and consider $t \in [0, T]$. First we show how to take the limit in the drift term. Fix an index n_0 and set

$$\begin{aligned}
I_1(t) &= \int_0^t \big|b_n\big(s, \tilde{X}_n(s)\big) - b_{n_0}\big(s, \tilde{X}_n(s)\big)\big|\,\mathrm{d}s, \\
I_2(t) &= \int_0^t \big|b_{n_0}\big(s, \tilde{X}_n(s)\big) - b_{n_0}\big(s, \tilde{X}(s)\big)\big|\,\mathrm{d}s, \\
I_3(t) &= \int_0^t \big|b_{n_0}\big(s, \tilde{X}(s)\big) - b\big(s, \tilde{X}(s)\big)\big|\,\mathrm{d}s.
\end{aligned}$$

Let φ be a continuous function such that $\varphi(t,x)=1$ for $|(t,x)|\leqslant 1/2$, $\varphi(t,x)=0$ for $|(t,x)|\geqslant 1$ and $0\leqslant\varphi(t,x)\leqslant 1$ elsewhere. Then by Chebyshev's inequality and Corollary 3.2,

$$
\begin{aligned}
&P\Big(\sup_{t\leqslant T} I_1(t)\geqslant\frac{\varepsilon}{3}\Big)\\
&\leqslant\frac{3}{\varepsilon}E\int_0^T\varphi\Big(\frac{s}{R},\frac{\tilde X_n(s)}{R}\Big)\big|(b_n-b_{n_0})\big(s,\tilde X_n(s)\big)\big|\,\mathrm{d}s\\
&\quad+\frac{3}{\varepsilon}E\int_0^T\Big(1-\varphi\Big(\frac{s}{R},\frac{\tilde X_n(s)}{R}\Big)\Big)\big|(b_n-b_{n_0})\big(s,\tilde X_n(s)\big)\big|\,\mathrm{d}s\\
&\leqslant\frac{3}{\varepsilon}N\mathrm{e}^T\big\|\chi_{[0,T]\times B_R}|b_n-b_{n_0}|_{L^{d+1}}\big\|\\
&\quad+\frac{6K}{\varepsilon}E\int_0^T\Big(1-\varphi\Big(\frac{s}{R},\frac{\tilde X_n(s)}{R}\Big)\Big)\,\mathrm{d}s+\frac{6N\mathrm{e}^T}{\varepsilon}\|\chi_{B^c_{R/2}}F\|_{L^{d+1}},
\end{aligned}
$$

where $B^c_{R/2}=\big\{(t,x)\in\mathbb{R}^{d+1}\colon\ |(t,x)|>R/2\big\}$. Hence by Lebesgue's theorem on dominated convergence

$$
\begin{aligned}
\limsup_{n\to\infty}P\Big(\sup_{t\leqslant T}I_1(t)\geqslant\frac{\varepsilon}{3}\Big)&\leqslant\frac{3N\mathrm{e}^T}{\varepsilon}\big\|\chi_{(0,T)\times B_R}|b-b_{n_0}|\big\|_{L^{d+1}}\\
&+\frac{6K}{\varepsilon}E\int_0^T\Big(1-\varphi\Big(\frac{s}{R},\frac{\tilde X(s)}{R}\Big)\Big)\,\mathrm{d}s+\frac{6N\mathrm{e}^T}{\varepsilon}\|\chi_{B^c_{R/2}}F\|_{L^{d+1}}.
\end{aligned}
\tag{5.4}
$$

By Chebyshev's inequality and by Lebesgue's theorem on dominated convergence

$$
\lim_{n\to\infty}P\Big(\sup_{t\leqslant T}I_2(t)\geqslant\frac{\varepsilon}{3}\Big)\leqslant\lim_{n\to\infty}\frac{3}{\varepsilon}E\int_0^T\big|b_{n_0}\big(s,\tilde X_n(s)\big)-b_{n_0}\big(s,\tilde X(s)\big)\big|\,\mathrm{d}s=0
$$

since b_{n_0} is bounded and continuous. To treat the term $I_3(t)$ we first observe that the estimate in Corollary 3.2 holds also for the process $\tilde X(t)$. Therefore we obtain inequality (5.4) with I_3 in place of I_1 in the same way as we got it for I_1. Hence

$$
\begin{aligned}
&\limsup_{n\to\infty}P\Big(\sup_{t\leqslant T}\int_0^t\big|b_n\big(s,\tilde X_n(s)\big)-b\big(s,\tilde X(s)\big)\big|\,\mathrm{d}s\geqslant\varepsilon\Big)\\
&\leqslant\sum_{i=1}^3P\Big(\sup_{t\leqslant T}I_i(t)\geqslant\frac{\varepsilon}{3}\Big)\\
&\leqslant\frac{6N\mathrm{e}^T}{\varepsilon}\big\|\chi_{(0,T)\times B_R}|b-b_{n_0}|\big\|_{L^{d+1}}+\frac{12K}{\varepsilon}E\int_0^T\Big(1-\varphi\Big(\frac{s}{R},\frac{\tilde X(s)}{R}\Big)\Big)\,\mathrm{d}s\\
&\quad+\frac{12N\mathrm{e}^T}{\varepsilon}\|\chi_{B^c_{R/2}}F\|_{L^{d+1}}
\end{aligned}
$$

for any index n_0 and $R > 0$. Letting here first $n_0 \to \infty$ and then $R \to \infty$, we get that

$$\int_0^t b_n\big(s, \tilde{X}_n(s)\big)\,\mathrm{d}s \longrightarrow \int_0^t b\big(s, \tilde{X}(s)\big)\,\mathrm{d}s$$

in probability, uniformly in $t \in [0, T]$. We consider now the diffusion term in equation (5.3). Like in the case of the drift term, define

$$\begin{aligned}
I_1(t) &= \int_0^t \sigma_n\big(s, \tilde{X}_n(s)\big)\,\mathrm{d}\tilde{W}_n(s) - \int_0^t \sigma_{n_0}\big(s, \tilde{X}_n(s)\big)\,\mathrm{d}\tilde{W}_n(s), \\
I_2(t) &= \int_0^t \sigma_{n_0}\big(s, \tilde{X}_n(s)\big)\,\mathrm{d}\tilde{W}_n(s) - \int_0^t \sigma_{n_0}\big(s, \tilde{X}(s)\big)\,\mathrm{d}\tilde{W}(s), \\
I_3(t) &= \int_0^t \sigma_{n_0}\big(s, \tilde{X}(s)\big)\,\mathrm{d}\tilde{W}(s) - \int_0^t \sigma\big(s, \tilde{X}(s)\big)\,\mathrm{d}\tilde{W}(s)
\end{aligned}$$

for an index n_0. Similarly as before, by using Chebyshev's and Davis' inequalities we have

$$\begin{aligned}
\lim_{n\to\infty} & P\Big(\sup_{t\leqslant T} |I_1(t)| \geqslant \frac{\varepsilon}{3}\Big) + P\Big(\sup_{t\leqslant T} |I_3(t)| \geqslant \frac{\varepsilon}{3}\Big) \\
&\leqslant \frac{18}{\varepsilon}\Big(N\mathrm{e}^T \big\|\chi_{(0,T)\times B_R}|\sigma - \sigma_{n_0}|^2\big\|_{L^p} \\
&\qquad + 4K^2 E\int_0^T \Big(1 - \varphi\Big(\frac{s}{R}, \frac{\tilde{X}(s)}{R}\Big)\Big)\,\mathrm{d}s\Big)^{1/2}
\end{aligned}$$

for any n_0 and R. Letting here first $n_0 \to \infty$ and then $R \to \infty$, we see that $I_1(t)$ and $I_3(t)$ converge to zero in probability, uniformly in $t \in [0, T]$. By using Lemma 5.2 with

$$\eta_n(t) := \sigma_{n_0}\big(t, \tilde{X}_n(t)\big), \qquad \eta(t) := \sigma_{n_0}\big(t, \tilde{X}(t)\big)$$

we get

$$\int_0^t \sigma_{n_0}\big(s, \tilde{X}_n(s)\big)\,\mathrm{d}\tilde{W}_n(s) \to \int_0^t \sigma_{n_0}\big(s, \tilde{X}(s)\big)\,\mathrm{d}\tilde{W}(s)$$

in probability, uniformly in $t \in [0, T]$ for every fixed n_0. Hence

$$\int_0^t \sigma_n\big(s, \tilde{X}_n(s)\big)\,\mathrm{d}\tilde{W}_n(s) \to \int_0^t \sigma\big(s, \tilde{X}(s)\big)\,\mathrm{d}\tilde{W}(s)$$

in probability, uniformly in $t \in [0, T]$. Consequently, letting $t \to \infty$ in equation (5.3) we get

$$\tilde{X}(t) = X_0 + \int_0^t b\big(s, \tilde{X}(s)\big)\,\mathrm{d}s + \int_0^t \sigma\big(s, \tilde{X}(s)\big)\,\mathrm{d}\tilde{W}(s) \tag{5.5}$$

for every $t \geqslant 0$. We have proved that the right-hand side of equation (5.3) converges to the right-hand side of equation (5.5) in $C([0,T];\mathbb{R}^d)$, in probability, for every $T > 0$. Therefore $\tilde{X}_n$ converges also in this sense. Thus we have proved the first statement of Theorem 2.2. To prove the rest it remains to show that X_n converges in $C([0,T];\mathbb{R}^d)$ in probability, if $b \in L^{2d+2}_{\mathrm{loc}}(\mathbb{R}_+ \times \mathbb{R}^d)$. We show this by using the following simple observation (see [7]).

Lemma 5.3. *Let Z_n be a sequence of random elements in a Polish space (S, ϱ) equipped with the Borel σ-algebra. Then Z_n converges in probability to an S-valued random element if and only if for every pair of subsequences Z_l and Z_m there exists a subsequence $v_k = (Z_{l(k)}, Z_{m(k)})$ converging in distribution to a random element v supported on the diagonal $\{(x,y) \in S \times S\colon\ x = y\}$.*

We have already proved that every pair of subsequences X_l and X_m is a tight sequence in $C([0,T];\mathbb{R}^{2d})$. Hence (X_l, X_m, W) is a tight sequence in $C([0,T];\mathbb{R}^{2d+d_1})$, and by the Skorohod representation theorem there exists a subsequence $(X_{l(k)}, X_{m(k)}, W)$ of it and a 'copy' of this subsequence $(\tilde{X}_{l(k)}, \tilde{X}_{m(k)}, \tilde{W}_k)$, on some probability space, such that their distributions coincide, and $\tilde{X}_{l(k)} \to \tilde{X}^1$, $\tilde{X}_{m(k)} \to \tilde{X}^2$, $\tilde{W}_k \to \tilde{W}$ in $C([0,T])$, in probability. Hence, as we have already seen,

$$\tilde{X}^i(t) = X_0 + \int_0^t b\big(s, \tilde{X}^i(s)\big)\,\mathrm{d}s + \int_0^t \sigma\big(s, \tilde{X}^i(s)\big)\,\mathrm{d}\tilde{W}(s)$$

follows for $i = 1, 2$. Consequently, $\tilde{X}^1(t) = \tilde{X}^2(t)$ almost surely for all $t \in [0,T]$ by Theorem 2.3 on the pathwise uniqueness. Hence by virtue of Lemma 5.3, X_n converges in $C([0,T];\mathbb{R}^d)$ in probability for every $T > 0$ to a process X which solves (2.1).

6. Proof of Theorem 2.1

Let $\varrho = \varrho(t,x)$ be a smooth non-negative mollifier supported in the unit ball. We set $a(t,x) := \sigma\sigma^*$ for $t \geqslant 0$, and $b(t,x) := 0$, $a(t,x) := \delta I$ for $t < 0$. We define

$$a_n(t,x) = (a * \varrho_n)(t,x), \quad \sigma_n = a_n^{1/2}, \quad b_n(t,x) = (b * \varrho_n)(t,x) \tag{6.1}$$

for every integer $n \geqslant 1$, where $\varrho_n(t,x) := n^{d+1}\varrho(ns, nx)$ and $*$ stands for the convolution. Then a_n and b_n are smooth bounded functions such that

$$\lim_{n\to\infty} \sigma_n(t,x) = (\sigma\sigma^*)^{1/2}(t,x) := \tilde{\sigma}(t,x), \quad \lim_{n\to\infty} b_n(t,x) = b(t,x)$$

almost everywhere with respect to the Lebesgue measure, and

$$\delta^{-1} I \geqslant \sigma_n \sigma_n^*(t,x) \geqslant \delta I$$

for all t, x. Moreover, we can choose a subsequence from b_n which satisfies Assumption 2.4. Namely, take n_k such that

$$\|F - F * \varrho_{n_k}\|_{d+1} \leqslant 2^{-k}.$$

Then $|b_{n_k}(t,x)| \leqslant K + G(t,x)$ with

$$G := \sum_{k=1}^{\infty} |F - F * \varrho_{n_k}| + F,$$

and clearly $\|G\|_{d+1} \leqslant 1 + \|F\|_{d+1}$. We get the first statement of Theorem 2.1 from Theorem 2.2 by considering the sequence $\{b_{n_k}\}_{k=1}^{\infty}$ and $\{\sigma_{n_k}\}_{k=1}^{\infty}$ in place of $\{b_n\}$ and $\{\sigma_n\}$. In fact a direct application of Theorem 2.2 shows the existence of a weak solution of equation (2.1) with $\tilde{\sigma}$ in place of σ. Notice, however, that $\sigma\sigma^* = \tilde{\sigma}\tilde{\sigma}^*$. Then it is well-known and easy to show (see, e.g., Lemma 10.4 from [13]) that equation (2.1) with the given diffusion coefficient σ has a weak solution. We obtain the second statement of Theorem 2.1 from Theorem 2.2 by taking the sequences $\{b_{n_k}\}_{k=1}^{\infty}$ and $\{\sigma_k = \sigma\}_{k=1}^{\infty}$ in place of $\{b_n\}$ and $\{\sigma_n\}$.

Acknowledgment. The authors thank Nicolai Krylov in Minnesota for useful discussions on the subject of the paper.

References

[1] *R. Aebi*: Diffusions with singular drift related to wave functions. Probab. Theory Related Fields *96* (1996), 107–121.

[2] *A. M. Davie*: Uniqueness of solutions of stochastic differential equations. 2001, preprint.

[3] *H. J. Engelbert and W. Schmidt*: On exponential local martingales connected with diffusion processes. Math. Nachr. *119* (1984), 97–115.

[4] *H. J. Engelbert and W. Schmidt*: On one-dimensional stochastic differential equations with generalized drift. Stochastic Differential Systems (Marseille-Luminy, 1984). Lecture Notes in Control and Information Sci. Vol. 69. Springer, Berlin-New York, 1985, pp. 143–155.

[5] *I. Gyöngy*: On stochastic differential equations with irregular coefficients. Technical Report Series of the Laboratory for Research in Statistics and Probability *91* (1986).

[6] *I. Gyöngy and N. V. Krylov*: Stochastic partial differential equations with unbounded coefficients and applications III. Stochastics Stochastics Rep. *40* (1992), 77–115.

[7] *I. Gyöngy and N. V. Krylov*: Existence of strong solutions for Itô's stochastic equations via approximations. Probab. Theory Related Fields *105* (1996), 143–158.

[8] *H. Kaneko and S. Nakao*: A note on approximation for stochastic differential equations. Séminaire de Probabilités XXII. Lecture Notes in Mathematics Vol. 1321. Springer, Berlin-New York, 1988, pp. 155–162.
[9] *N. V. Krylov*: On some estimates of the distribution of stochastic integrals. Izv. Akad. Nauk SSSR Ser. Math. *38* (1974), 228–248. (In Russian.)
[10] *N. V. Krylov*: Estimates of the maximum of the solution of a parabolic equation and estimates of the distribution of a semimartingale. Mat. Sb. (N.S.) *130 (172)* (1986), 207–221. (In Russian.)
[11] *N. V. Krylov*: Controlled Diffusion Processes. Nauka, Moscow, 1977, English transl.: Springer-Verlag, New York-Berlin, 1980.
[12] *O. A. Ladyzenskaya, V. A. Solonnikov and N. N. Ural'ceva*: Linear and Quasilinear Equations of Parabolic Type. Nauka, Moscow, 1967, English transl.: American Mathematical Society, Providence, R.I., 1968.
[13] *R. S. Liptser and A. N. Shiryaev*: Statistics of Random Processes I. Nauka, Moscow, 1974, English transl.: Springer-Verlag, New York-Heidelberg, 1977.
[14] *P. A. Meyer and W. A. Zheng*: Quelques résultats de "mécanique stochastique". Séminaire de Probabilités XVIII. Lecture Notes in Mathematics Vol. 1059. Springer, Berlin-New York, 1984, pp. 223–244.
[15] *P. A. Meyer and W. A. Zheng*: Sur la construction des certaines diffusions. Séminaire de Probabilités XX. Lecture Notes in Mathematics Vol. 1204. Springer, Berlin-New York, 1986, pp. 334–337.
[16] *M. Nagasawa*: Transformations of diffusions and Schrödinger processes. Probab. Theory Related Fields *82* (1989), 109-136.
[17] *M. Nagasawa, H. Tanaka*: Diffusion process in a singular mean-drift-field. Z. Wahrsch. Verw. Gebiete *68* (1985), 247–269.
[18] *N. I. Portenko*: Generalized Diffusion Processes. Naukova Dumka, Kiev, 1982, English transl.: American Mathematical Society, Providence, R.I., 1990.
[19] *A. V. Skorohod*: Studies in the Theory of Random Processes. Izdat. Kiev. Univ., Kiev, 1961, English transl.: Addison Wesley Publ. Co., Reading, Mass., 1965.
[20] *D. W. Stroock, S. R. S. Varadhan*: Multidimensional Diffusion Processes. Springer-Verlag, Berlin, 1979.
[21] *W. Stummer*: The Novikov and entropy conditions of multidimensional diffusion processes with singular drift. Probab. Theory Related Fields *97* (1993), 515–542.
[22] *A. Ju. Veretennikov*: On the strong solutions of stochastic differential equations. Theory Probab. Appl. *24* (1979), 354–366.
[23] *A. Yu. Veretennikov*: Parabolic equations and stochastic Itô differential equations with discontinuous coefficients in time. Mat. Zametki *31* (1982), 549–557. (In Russian.)
[24] *A. K. Zvonkin*: A transformation of the phase space of a diffusion process that will remove the drift. Mat. Sb. (N.S.) *93 (135)* (1974), 129–149. (In Russian.)

Authors' addresses: I. Gyöngy, Department of Mathematics and Statistics, University of Edinburgh, King's Buildings, Edinburgh, EH9 3JZ, United Kingdom, e-mail: `gyongy@maths.ed.ac.uk`; T. Martínez, Departamento de Matemáticas, Universidad Autonóma de Madrid, Campus de Cantoblanco 28049, Madrid, Spain, e-mail: `teresa.martinez@uam.es`.