Stoch PDE: Anal Comp (2015) 3:52–83
DOI 10.1007/s40072-014-0042-6

On the solvability of degenerate stochastic partial differential equations in Sobolev spaces

Máté Gerencsér · István Gyöngy · Nicolai Krylov

Received: 22 April 2014 / Published online: 13 December 2014

Abstract Systems of parabolic, possibly degenerate parabolic SPDEs are considered. Existence and uniqueness are established in Sobolev spaces. Similar results are obtained for a class of equations generalizing the deterministic first order symmetric hyperbolic systems.

Keywords Cauchy problem · Degenerate stochastic parabolic PDEs · First order symmetric hyperbolic system

Mathematics Subject Classification 60H15 · 35K65 · 35K45 · 35F40

1 Introduction

In this paper we are interested in the solvability in L_p spaces of linear stochastic parabolic, possibly degenerate, PDEs and of systems of linear stochastic parabolic PDEs. The equations we consider are important in applications. They arise in nonlinear filtering of partially observable stochastic processes, in modelling of hydromagnetic dynamo evolving in fluids with random velocities, and in many other areas of physics and engineering.

M. Gerencsér · I. Gyöngy (✉)
School of Mathematics and Maxwell Institute, The University of Edinburgh, The King's Buildings, Peter Guthrie Tait Road, Edinburgh EH9 3FD, Scotland, UK
e-mail: I.Gyongy@ed.ac.uk; gyongy@maths.ed.ac.uk

M. Gerencsér
e-mail: m.gerencser@sms.ed.ac.uk

N. Krylov
University of Minnesota, 127 Vincent Hall, Minneapolis, MN 55455, USA
e-mail: krylov@math.umn.edu

Among several important results, an L_2-theory of degenerate linear elliptic and parabolic PDEs is presented in [25–27] and [28]. The solvability in L_2 spaces of linear degenerate stochastic PDEs of parabolic type was first studied in [20] (see also [29]).

Solving equations in W^m_p spaces for sufficiently high exponent p allows one to prove by Sobolev embedding better smoothness properties of the solutions than in the case of solving them in W^m_2 spaces. As it is mentioned above, the class of stochastic PDEs considered in this paper includes the equations of nonlinear filtering of partially observed diffusion processes. By our results one obtains the existence of the conditional density of the unobserved process, and its regularity properties, under minimal smoothness conditions on the coefficients.

The first existence and uniqueness theorem on solvability of these equations in W^m_p spaces, when they may also degenerate, is presented in [22]. This result is improved in [8].

In the present paper we fill in a gap in the proof of the existence and uniqueness theorems in [22] and [8]. Moreover, we essentially improve these theorems. In [22] the existence and uniqueness theorem for W^m_p-valued solutions is not separated from an existence and uniqueness theorem for W^m_2-valued solutions. In particular, it contains also conditions ensuring the existence and uniqueness of a W^m_2 solution. In [8] these conditions were removed, and for any $q \in (0, p]$ an estimate for $E\sup_{t\le T}|u|^q_{W^m_p}$ for the solution u is obtained. In the present paper we remove the extra conditions of the existence and uniqueness theorem in [22], remove the restriction $q \le p$ on the exponent q in the corresponding theorem in [8], and prove the uniqueness of the solution under weaker assumptions than those in [22] and [8] (see Theorem 2.1 below). Note that to have q-th moment estimates for any high q is useful, for example, in proving almost sure rate of convergence of numerical approximations of stochastic PDEs, see, e.g., [5]. Moreover, we not only improve the existence an uniqueness theorems in [22] and [8], but our main result, Theorem 3.1, extends them to degenerate stochastic parabolic systems. We present also an existence and uniqueness theorem, Theorem 3.2, on solvability in W^m_2 spaces for a larger class of stochastic parabolic systems, which, in particular, contains the first order symmetric hyperbolic systems. This result was indicated in [9].

We would like to emphasise that the equations we consider in this paper may degenerate and become first order equations. For non degenerate stochastic PDEs L_p- and $L_q(L_p)$-theories are developed, see e.g. [13,14,17,18] and [15], which give essentially stronger results on smoothness of the solutions.

There are many publications on stochastic PDEs driven by martingale measures, pioneered by [30]. (See also [2] and the references therein.) In [3] two set-ups for stochastic PDEs, concerning the driving noise are compared: a set-up when the driving noise is a martingale measure, and an other set-up when the equations are driven by martingales with values in infinite dimensional spaces. It is shown, in particular, that stochastic integrals with respect to martingale measures can be rewritten as stochastic Itô integrals with respect to martingales taking values in Hilbert spaces. Earlier this was proved in [6] in order to treat SDEs and stochastic PDEs driven by martingale measures as stochastic equations driven by martingales. In [16] super-Brownian motions in any dimension are constructed as solutions of SPDEs driven by infinite dimensional martingales, more precisely, by an infinite sequence of independent Wiener processes. As it is well-known, in the one-dimensional case the stochastic equation for the super-Brownian motion can be written as a stochastic PDE driven by a martingale measure, more precisely, by a space-time white noise, but as it is noted in [16], most likely this is not possible in higher dimensions.

Solvability of stochastic PDEs of parabolic type are often investigated in the sense of the *mild solution* concept, i.e., when solutions to stochastic PDEs are defined as solutions to a stochastic integral equation obtained via *Duhamel's principle*, called also *variation of constant formula* in the context of ODEs (see, e.g., [2] and [3]). For the theory of stochastic PDEs built on this approach, often called *semigroup approach*, we refer the reader to the monograph [4]. In this framework there are many results on solvability in various Banach spaces $\mathbb{B}$, including W^m_p spaces, when the linear operator in the drift term of the equation is an infinitesimal generator of a continuous semigroup of bounded linear operators acting on $\mathbb{B}$. The equations investigated in most papers, including [2] and [3], do not have a differential operator in their diffusion part, unlike the equations studied in this paper. In the case when the differential operator in the drift term is a time dependent random operator, serious problems arise in adaptation the semigroup approach. Thus the semigroup approach is not used to investigate the filtering equations of general signal and observation models, which are included in the class of equations considered in the present paper.

Finally we would like to mention that for some special degenerate stochastic PDEs, for example for the stochastic Euler equations, there are many results on solvability in the literature. See, for example, [1] and the references therein. Concerning the equation in [1] we note that its main term is non random, and its solution can be given in a sense explicitly.

In conclusion we introduce some notation used throughout the paper. All random elements will be given on a fixed probability space $(\Omega, \mathcal{F}, P)$, equipped with a filtration $(\mathcal{F}_t)_{t\geq 0}$ of σ-fields $\mathcal{F}_t \subset \mathcal{F}$. We suppose that this probability space carries a sequence of independent Wiener processes $(w^r)_{r=1}^{\infty}$, adapted to the filtration $(\mathcal{F}_t)_{t\geq 0}$, such that $w^r_t - w^r_s$ is independent of $\mathcal{F}_s$ for each r and any $0 \leq s \leq t$. It is assumed that $\mathcal{F}_0$ contains all P-null subsets of Ω, so that $(\Omega, \mathcal{F}, P)$ is a complete probability space and the σ-fields $\mathcal{F}_t$ are complete. By $\mathcal{P}$ we denote the predictable σ-field of subsets of $\Omega \times (0, \infty)$ generated by $(\mathcal{F}_t)_{t\geq 0}$. For basic notions in stochastic analysis, like continuous local martingales and their quadratic variation process, we refer to [12].

For $p \in [1, \infty)$, the space of measurable mappings f from $\mathbb{R}^d$ into a separable Hilbert space $\mathcal{H}$, such that

$$\|f\|_{L_p} = \left(\int_{\mathbb{R}^d} |f(x)|^p_{\mathcal{H}}\, dx\right)^{1/p} < \infty,$$

is denoted by $L_p(\mathbb{R}^d, \mathcal{H})$.

Remark 1.1 We did not include the symbol $\mathcal{H}$ in the notation of the norm in $L_p(\mathbb{R}^d, \mathcal{H})$. Which $\mathcal{H}$ is involved will be absolutely clear from the context. We do the same in other similar situations.

Often $\mathcal{H}$ will be l_2, or the space of infinite matrices $\{g^{ij} \in \mathbb{R} : i = 1, \ldots, M,\ j = 1, 2, \ldots\}$, or finite $M \times M$ matrices with the Hilbert–Schmidt norm. The space of functions from $L_p(\mathbb{R}^d, \mathcal{H})$, whose generalized derivatives up to order m are also in $L_p(\mathbb{R}^d, \mathcal{H})$, is denoted by $W^m_p(\mathbb{R}^d, \mathcal{H})$. By definition $W^0_p(\mathbb{R}^d, \mathcal{H}) = L_p(\mathbb{R}^d, \mathcal{H})$. The norm $|u|_{W^m_p}$ of u in $W^m_p(\mathbb{R}^d, \mathcal{H})$ is defined by

$$|u|^p_{W^m_p} = \sum_{|\alpha| \le m} |D^\alpha u|^p_{L_p}, \tag{1.1}$$

where $D^\alpha := D_1^{\alpha_1}, \ldots, D_d^{\alpha_d}$ for *multi-indices* $\alpha := (\alpha_1, \ldots, \alpha_d) \in \{0, 1, \ldots\}^d$ of *length* $|\alpha| := \alpha_1 + \alpha_2 + \cdots + \alpha_d$, and $D_i u$ is the generalized derivative of u with respect to x^i for $i = 1, 2 \ldots, d$. We also use the notation $D_{ij} = D_i D_j$ and $Du = (D_1 u, \ldots, D_d u)$. When we talk about "derivatives up to order m" of a function for some nonnegative integer m, then we always include the zeroth-order derivative, i.e. the function itself. Unless otherwise indicated, the summation convention with respect to repeated integer valued indices is used throughout the paper.

2 Formulation

In this section $\mathcal{H} = \mathbb{R}$ and we use a shorter notation

$$L_p = L_p(\mathbb{R}^d, \mathbb{R}), \quad W^m_p = W^m_p(\mathbb{R}^d, \mathbb{R}), \quad W^{m+1}_p(l_2) = W^{m+1}_p(\mathbb{R}^d, l_2).$$

Fix a $T \in (0, \infty)$ and consider the problem

$$du_t(x) = (L_t u_t(x) + f_t(x))\, dt + \left(M^r_t u_t(x) + g^r_t(x)\right)\, dw^r_t, \tag{2.1}$$

$(t, x) \in H_T := [0, T] \times \mathbb{R}^d$, with initial condition

$$u_0(x) = \psi(x), \quad x \in \mathbb{R}^d, \tag{2.2}$$

where

$$L_t = a^{ij}_t(x) D_{ij} + b^i_t(x) D_i + c_t(x), \quad M^r_t = \sigma^{ir}_t(x) D_i + \nu^r_t(x),$$

and all functions, given on $\Omega \times H_T$, are assumed to be real valued and satisfy the following assumptions in which $m \ge 0$ is an integer and K is a constant.

Assumption 2.1 The derivatives in $x \in \mathbb{R}^d$ of a^{ij} up to order $\max(m, 2)$ and of b^i and c up to order m are $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}^d)$-measurable functions, bounded by K for all $i, j \in \{1, 2, \ldots, d\}$. The functions $\sigma^i = (\sigma^{ir})^\infty_{r=1}$ and $\nu = (\nu^r)^\infty_{r=1}$ are l_2-valued and their derivatives in x up to order $m+1$ are $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}^d)$-measurable l_2-valued functions, bounded by K.

Assumption 2.2 The free data, f_t and $g_t = (g^r)_{r=1}^{\infty}$ are predictable processes with values in W_p^m and $W_p^{m+1}(l_2)$, respectively, such that almost surely

$$\mathcal{K}_{m,p}^p(T) = \int_0^T \left(|f_t|_{W_p^m}^p + |g_t|_{W_p^{m+1}}^p \right) dt < \infty. \tag{2.3}$$

The initial value, ψ is an $\mathcal{F}_0$-measurable random variable with values in W_p^m.

Assumption 2.3 For $P \otimes dt \otimes dx$-almost all $(\omega, t, x) \in \Omega \times [0, T] \times \mathbb{R}^d$

$$\alpha_t^{ij}(x) z^i z^j \geq 0$$

for all $z \in \mathbb{R}^d$, where

$$\alpha^{ij} = 2a^{ij} - \sigma^{ir}\sigma^{jr}.$$

This condition is a standard assumption in the theory of stochastic PDEs. If it is not satisfied then Eq. (2.1) may be solvable only for very special initial conditions and free terms. Notice that this assumption allows $\alpha = 0$, which can happen, for example, when $\sigma^{ik} = (\sqrt{2a})^{ik}$ for $i, k = 1, \ldots, d$ and $\sigma^{ik} = 0$ for $k > d$.

Let τ be a stopping time bounded by T.

Definition 2.1 A W_p^1-valued function u, defined on the stochastic interval $(\!(0, \tau]\!]$, is called a solution of (2.1)–(2.2) on $[0, \tau]$ if u is predictable on $(\!(0, \tau]\!]$,

$$\int_0^\tau |u_t|_{W_p^1}^p \, dt < \infty \, (a.s.),$$

and for each $\varphi \in C_0^\infty(\mathbb{R}^d)$ for almost all $\omega \in \Omega$

$$\begin{aligned} (u_t, \varphi) &= (\psi, \varphi) + \int_0^t \left\{ -(a_s^{ij} D_i u_s, D_j \varphi) + (\bar{b}_s^i D_i u_s + c_s u_s + f_s, \varphi) \right\} ds \\ &\quad + \int_0^t (\sigma_s^{ir} D_i u_s + \nu_s^r u_s + g_s^r, \varphi) \, dw_s^r \end{aligned}$$

for all $t \in [0, \tau(\omega)]$, where $\bar{b}^i = b^i - D_j a^{ij}$, and $(\cdot, \cdot)$ denotes the inner product in the Hilbert space of square integrable real-valued functions on $\mathbb{R}^d$.

We want to prove the following existence and uniqueness theorem about the Cauchy problem (2.1)–(2.2).

Theorem 2.1 *Let Assumptions 2.3 and 2.1–2.2 with $m \geq 0$ hold. Then there exists at most one solution on $[0, T]$. If together with Assumptions 2.3, 2.1–2.2 hold with $m \geq 1$, then there exists a unique solution $u = (u_t)_{t \in [0,T]}$ on $[0, T]$. Moreover, u is a W_p^m-valued weakly continuous process, it is a strongly continuous process with values in W_p^{m-1}, and for every $q > 0$ and $n \in \{0, 1, \ldots, m\}$*

$$E \sup_{t\in[0,T]} |u_t|^q_{W^n_p} \le N(E|\psi|^q_{W^n_p} + E\mathcal{K}^q_n(T)), \tag{2.4}$$

where N is a constant depending only on K, T, d, m, p and q.

This result is proved in [22] in the case $q = p \ge 2$ under the additional assumptions that $E\mathcal{K}^r_{m,r}(T) < \infty$ and $E|\psi|^r_{W^m_r} < \infty$ for $r = p$ and $r = 2$ (see Theorem 3.1 therein). These additional assumptions are not supposed and a somewhat weaker version of the above theorem is obtained in [8] when $q \in (0, p]$. The proof of it in [8] uses Theorem 3.1 from [22], whose proof is based on an estimate for the derivatives of the solution u, formulated as Lemma 2.1 in [22]. The proof of this lemma, however, contains a gap. Our aim is to fill in this gap and also to improve the existence and uniqueness theorems from [22] and [8]. Since $Du = (D_1u, \dots, D_du)$ satisfies a system of SPDEs, it is natural to present and prove our results in the context of systems of stochastic PDEs.

3 Systems of stochastic PDEs

Let $M \ge 1$ be an integer, and let $\langle\cdot\,,\cdot\rangle$ and $\langle\cdot\rangle$ denote the scalar product and the norm in $\mathbb{R}^M$, respectively. By $\mathbb{T}^M$ we denote the set of $M \times M$ matrices, which we consider as a Euclidean space $\mathbb{R}^{M^2}$. For an integer $m \ge 1$ we define $l_2(\mathbb{R}^m)$ as the space of sequences $\nu = (\nu^1, \nu^2, \dots)$ with $\nu^k \in \mathbb{R}^m$, $k \ge 1$, and finite norm

$$\|\nu\|_{l_2} = \left(\sum_{k=1}^{\infty} |\nu|^2_{\mathbb{R}^m}\right)^{1/2}$$

(cf. Remark 1.1).

We look for $\mathbb{R}^M$-valued functions $u_t(x) = (u^1_t(x), \dots, u^M_t(x))$, of $\omega \in \Omega$, $t \in [0, T]$ and $x \in \mathbb{R}^d$, which satisfy the system of equations

$$du_t = \left[a^{ij}_t D_{ij}u_t + b^i_t D_i u_t + cu_t + f_t\right] dt \\ + \left[\sigma^{ik}_t D_i u_t + \nu^k_t u_t + g^k_t\right] dw^k_t, \tag{3.1}$$

and the initial condition

$$u_0 = \psi, \tag{3.2}$$

where $a_t = (a^{ij}_t(x))$ takes values in the set of $d \times d$ symmetric matrices,

$$\sigma^i_t = \left(\sigma^{ik}_t(x), k \ge 1\right) \in l_2, \quad b^i_t(x) \in \mathbb{T}^M, \quad c_t(x) \in \mathbb{T}^M, \\ \nu_t(x) \in l_2(\mathbb{T}^M), \quad f_t(x) \in \mathbb{R}^M, \quad g_t(x) \in l_2(\mathbb{R}^M) \tag{3.3}$$

for $i = 1, \dots, d$, for all $\omega \in \Omega$, $t \ge 0$, $x \in \mathbb{R}^d$.

Note that with the exception of a^{ij} and σ^{ik}, all 'coefficients' in Eq. (3.1) mix the coordinates of the process u.

Let m be a nonnegative integer, $p \in [2, \infty)$ and make the following assumptions, which are straightforward adaptations of Assumptions 2.1 and 2.2.

Assumption 3.1 The derivatives in $x \in \mathbb{R}^d$ of a^{ij} up to order $\max(m, 2)$ and of b^i and c up to order m are $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}^d)$-measurable functions, in magnitude bounded by K for all $i, j \in \{1, 2, \ldots, d\}$. The derivatives in x of the l_2-valued functions $\sigma^i = (\sigma^{ik})_{k=1}^{\infty}$ and the $l_2(\mathbb{T}^M)$-valued function ν up to order $m + 1$ are $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}^d)$-measurable l_2-valued and $l_2(\mathbb{T}^M)$-valued functions, respectively, in magnitude bounded by K.

Assumption 3.2 The free data, $(f_t)_{t\in[0,T]}$ and $(g_t)_{t\in[0,T]}$ are predictable processes with values in

$$W_p^m\left(\mathbb{R}^d, \mathbb{R}^M\right) \quad \text{and} \quad W_p^{m+1}\left(\mathbb{R}^d, l_2(\mathbb{R}^M)\right),$$

respectively, such that almost surely

$$\mathcal{K}_{m,p}^p(T) = \int_0^T \left(|f_t|_{W_p^m}^p + |g_t|_{W_p^{m+1}}^p\right) dt < \infty. \tag{3.4}$$

The initial value, ψ is an $\mathcal{F}_0$-measurable random variable with values in $W_p^m(\mathbb{R}^d, \mathbb{R}^M)$.

Set

$$\beta^i = b^i - \sigma^{ir}\nu^r, \quad i = 1, 2, \ldots, d,$$

and recall that $\alpha^{ij} = 2a^{ij} - \sigma^{ik}\sigma^{jk}$ for $i, j = 1, \ldots, d$. Instead of Assumption 2.3 we impose now the following condition, where δ^{kl} stands for the 'Kronecker δ', i.e., $\delta^{kl} = 1$ if $k = l$ and it is zero otherwise.

Assumption 3.3 There exist a constant $K_0 > 0$ and a $\mathcal{P} \times \mathcal{B}(\mathbb{R}^d)$-measurable $\mathbb{R}^d$-valued bounded function $h = (h_t^i(x))$, whose first order derivatives in x are bounded functions, such that for all $\omega \in \Omega$, $t \geq 0$ and $x \in \mathbb{R}^d$

$$|h| + |Dh| \leq K, \tag{3.5}$$

and for all $(\lambda_1, \ldots, \lambda_d) \in \mathbb{R}^d$

$$\left|\sum_{i=1}^d (\beta^{ikl} - \delta^{kl}h^i)\lambda_i\right|^2 \leq K_0 \sum_{i,j=1}^d \alpha^{ij}\lambda_i\lambda_j \quad \text{for } k, l = 1, \ldots, M. \tag{3.6}$$

Remark 3.1 Let Assumption 3.1 hold with $m = 0$ and the first order derivatives of b^i in x are bounded by K for each $i = 1, 2, \ldots, d$. Then notice that condition (3.6) is a natural extension of Assumption 2.3 to systems of stochastic PDEs. Indeed, when $M = 1$ then taking $h^i = \beta^i$ for $i = 1, \ldots, d$, we can see that Assumption 3.3 is equivalent to $\alpha \geq 0$. Let us analyse now Assumption 3.3 for arbitrary $M \geq 1$. Notice that it holds when α

is uniformly elliptic, i.e., $\alpha \geq \kappa I_d$ with a constant $\kappa > 0$ for all ω, $t \geq 0$ and $x \in \mathbb{R}^d$. Indeed, due to Assumption 3.1 there is a constant $N = N(K, d)$ such that

$$\left| \sum_{i=1}^{d} (\beta^{ikl} - \delta^{kl} h^i)\lambda_i \right|^2 \leq N \sum_{i=1}^{d} |\lambda_i|^2 \quad \text{for every } k, l = 1, 2, \ldots, M,$$

which together with the uniform ellipticity of α clearly implies (3.6). Notice also that (3.6) holds in many situations when instead of the strong ellipticity of α we only have $\alpha \geq 0$. Such examples arise, for example, when $a^{ij} = (\sigma^{ir}\sigma^{jr})/2$ for all $i, j = 1, \ldots, d$, and b and ν are such that β^i is a diagonal matrix for each $i = 1, \ldots, d$, and the diagonal elements together with their first order derivatives in x are bounded by a constant K. As a simple example, consider the system of equations

$$du_t(x) = \left\{ \tfrac{1}{2} D^2 u_t(x) + D v_t(x) \right\} dt + \{D u_t(x) + v_t(x)\}\, dw_t$$
$$dv_t(x) = \left\{ \tfrac{1}{2} D^2 v_t(x) - D u_t(x) \right\} dt + \{D v_t(x) - u_t(x)\}\, dw_t$$

for $t \in [0, T]$, $x \in \mathbb{R}$, for a 2-dimensional process $(u_t(x), v_t(x))$, where w is a one-dimensional Wiener process. In this example $\alpha = 0$ and $\beta = 0$. Thus clearly, condition (3.6) is satisfied.

In Sect. 5 it will be convenient to use condition (3.6) in an equivalent form, which we discuss in the next remark.

Remark 3.2 Notice that condition (3.6) in Assumption 3.3 can be reformulated as follows: There exists a constant K_0 such that for all values of the arguments and all continuously differentiable $\mathbb{R}^M$-valued functions $u = u(x)$ on $\mathbb{R}^d$ we have

$$\langle u, b^i D_i u\rangle - \sigma^{ik}\langle u, \nu^k D_i u\rangle \leq K_0 \left| \sum_{i,j=1}^{d} \alpha^{ij} \langle D_i u, D_j u\rangle \right|^{1/2} \langle u\rangle + h^i \langle D_i u, u\rangle. \quad (3.7)$$

Indeed, set $\hat{\beta}^i = \beta^i - h^i I_M$, where I_M is the $M \times M$ unit matrix, and observe that (3.7) means

$$\langle u, \hat{\beta}^i D_i u\rangle \leq K_0 \left| \sum_{i,j=1}^{d} \alpha^{ij} \langle D_i u, D_j u\rangle \right|^{1/2} \langle u\rangle.$$

By considering this relation at a fixed point x and noting that then one can choose u and Du independently, we conclude that

$$\left\langle \sum_i \hat{\beta}^i D_i u \right\rangle^2 \leq K_0^2 \alpha^{ij} \langle D_i u, D_j u\rangle \quad (3.8)$$

and (3.6) follows (with a different K_0) if we take $D_i u^k = \lambda_i \delta^{kl}$.

On the other hand, (3.6) means that for any l without summation on l

$$\Big|\sum_i \hat{\beta}^{ikl} D_i u^l\Big|^2 \leq K_0 \alpha^{ij}(D_i u^l) D_j u^l.$$

But then by Cauchy's inequality similar estimate holds after summation on l is done and carried inside the square on the left-hand side. This yields (3.8) (with a different constant K_0) and then leads to (3.7).

The notion of solution to (3.1)–(3.2) is a straightforward adaptation of Definition 2.1 to systems of equations. Namely, $u = (u^1, \dots, u^M)$ is a solution on $[0, \tau]$, for a stopping time $\tau \leq T$, if it is a $W^1_p(\mathbb{R}^d, \mathbb{R}^M)$-valued predictable function on $(\!(0, \tau]\!]$,

$$\int_0^\tau |u_t|^p_{W^1_p}\, dt < \infty \quad \text{(a.s.)},$$

and for each $\mathbb{R}^M$-valued $\varphi = (\varphi^1, \dots, \varphi^M)$ from $C_0(\mathbb{R}^d)$ with probability one

$$\begin{aligned}(u_t, \varphi) = (\psi, \varphi) + \int_0^t \Big\{ &- (a^{ij}_s D_i u_s, D_j \varphi) \\ &+ (\bar{b}^i_s D_i u_s + c_s u_s + f_s, \varphi)\Big\}\, ds \qquad (3.9)\\ &+ \int_0^t \Big(\sigma^{ir}_s D_i u_s + \nu^r_s u_s + g^r(s), \varphi\Big)\, dw^r_s \qquad (3.10)\end{aligned}$$

for all $t \in [0, \tau]$, where $\bar{b}^i = b^i - D_j a^{ij} I_M$. Here, and later on (Ψ, Φ) denotes the inner product in the L_2-space of $\mathbb{R}^M$-valued functions Ψ and Φ defined on $\mathbb{R}^d$.

The main result of the paper reads now just like Theorem 2.1 above.

Theorem 3.1 *Let Assumption 3.3 hold. If Assumptions 3.1 and 3.2 also hold with* $m \geq 0$*, then there is at most one solution to* (3.1)–(3.2) *on* $[0, T]$*. If together with Assumption 3.3, Assumptions 3.1 and 3.2 hold with* $m \geq 1$*, then there is a unique solution* $u = (u^l)_{l=1}^M$ *to* (3.1)–(3.2) *on* $[0, T]$*. Moreover, u is a weakly continuous* $W^m_p(\mathbb{R}^d, \mathbb{R}^M)$*-valued process, it is strongly continuous as a* $W^{m-1}_p(\mathbb{R}^d, \mathbb{R}^M)$*-valued process, and for every* $q > 0$ *and* $n \in \{0, 1, \dots, m\}$

$$E \sup_{t\in[0,T]} |u_t|^q_{W^n_p} \leq N\Big(E|\psi|^q_{W^n_p} + E\mathcal{K}^q_{n,p}(T)\Big) \qquad (3.11)$$

with $N = N(m, p, q, d, M, K, T)$.

Example 3.1 In *hydromagnetic dynamo theory* the system of equations

$$\frac{\partial}{\partial t} B^k_t(x) = \lambda_t(x)\Delta B^k_t(x) + D_j v^k_t(x) B^j_t(x) - v^j_t(x) D_j B^k_t(x), \quad k = 1, 2, 3, \quad (3.12)$$

for $t \in [0, T]$ and $x \in \mathbb{R}^3$, called *induction equation*, describes the evolution of a magnetic field $B = (B^1, B^2, B^3)$ in a fluid flowing with velocity $v = (v^1, v^2, v^3)$, where $\lambda \geq 0$ is the *magnetic diffusivity* (see, for example, [23]). Notice that one can apply Theorem 3.1 to (3.12) to obtain its solvability in W^m_p spaces. To study effects in turbulent flows, induction equations with random velocity fields v have been investigated in the literature (see, for example, [24]). In [7] convergence of (3.12) to a system of stochastic PDEs is shown when the velocity fields are random and converge to a random field which is white noise in time. We note that Theorem 3.1 can be applied also to the system of stochastic PDEs obtained in this way.

In the case $p = 2$ we present also a modification of Assumption 3.3, in order to cover an important class of stochastic PDE systems, the *hyperbolic symmetric systems*.

Observe that if in (3.6) we replace β^{ikl} with β^{ilk}, nothing will change. By the convexity of t^2 condition (3.6) then holds if we replace β^{ilk} with $(1/2)[\beta^{ilk} + \beta^{ikl}]$. Since

$$|a - b|^2 \leq |a + b|^2 + 2a^2 + 2b^2$$

this implies that (3.6) also holds for

$$\bar{\beta}^{ikl} = (\beta^{ikl} - \beta^{ilk})/2$$

in place of β^{ikl}, which is the antisymmetric part of $\beta^i = b^i - \sigma^{ir}\nu^r$.

Hence the following condition is weaker than Assumption 3.3.

Assumption 3.4 There exist a constant $K_0 > 0$ and a $\mathcal{P} \times \mathcal{B}(\mathbb{R}^d)$-measurable $\mathbb{R}^M$-valued function $h = (h^i_t(x))$ such that (3.5) holds, and for all $\omega \in \Omega$, $t \geq 0$ and $x \in \mathbb{R}^d$ and for all $(\lambda_1, \ldots, \lambda_d) \in \mathbb{R}^d$

$$\left| \sum_{i=1}^{d} (\bar{\beta}^{ikl} - \delta^{kl} h^i)\lambda_i \right|^2 \leq K_0 \sum_{i,j=1}^{d} \alpha^{ij} \lambda_i \lambda_j \quad \text{for } k, l = 1, \ldots, M. \tag{3.13}$$

The following result in the special case of deterministic PDE systems is indicated and a proof is sketched in [9].

Theorem 3.2 *Take $p = 2$ and replace Assumption 3.3 with Assumption 3.4 in the conditions of Theorem 3.1. Then the conclusion of Theorem 3.1 holds with $p = 2$.*

Remark 3.3 Notice that Assumption 3.4 obviously holds with $h^i = 0$ if the matrices β^i are symmetric and $\alpha \geq 0$. When $a = 0$ and $\sigma = 0$ then the system is called a *first order symmetric hyperbolic system*.

Remark 3.4 If Assumption 3.4 does not hold then even simple first order deterministic systems with smooth coefficients may be ill-posed. Consider, for example, the system

$$\begin{aligned} du_t(x) &= Dv_t(x)\,dt \\ dv_t(x) &= -Du_t(x)\,dt \end{aligned} \tag{3.14}$$

for $(u_t(x), v_t(x))$, $t \in [0, T]$, $x \in \mathbb{R}$, with initial condition $u_0 = \psi$, $v_0 = \phi$, such that $\psi, \phi \in W_2^m \setminus W_2^{m+1}$ for an integer $m \geq 1$. Clearly, this system does not satisfy Assumption 3.4, and one can show that it does not have a solution with the initial condition $u_0 = \psi$, $v_0 = \phi$. We note, however, that it is not difficult to show that for any constant $\varepsilon \neq 0$ and Wiener process w the stochastic PDE system

$$\begin{aligned} du_t(x) &= Dv_t(x)\,dt + \varepsilon Dv_t(x)\,dw_t \\ dv_t(x) &= -Du_t(x)\,dt - \varepsilon Du_t(x)\,dw_t \end{aligned} \tag{3.15}$$

with initial condition $(u_0, v_0) = (\psi, \phi) \in W_2^m$ (for $m \geq 1$) has a unique solution $(u_t, v_t)_{t\in[0,T]}$, which is a W_2^m-valued continuous process. One can prove this statement and the statement about the nonexistence of a solution to (3.14) by using Fourier transform. We leave the details of the proof as exercises for those readers who find them interesting. Clearly, system (3.15) does not belong to the class of stochastic systems considered in this paper.

4 Preliminaries

First we discuss the solvability of (3.1)–(3.2) under the strong stochastic parabolicity condition.

Assumption 4.1 There is a constant $\kappa > 0$ such that

$$\alpha^{ij}\lambda_i\lambda_j \geq \kappa \sum_{i=1}^{d} |\lambda_i|^2$$

for all $\omega \in \Omega$, $t \geq 0$, $x \in \mathbb{R}^d$ and $(\lambda_1, \ldots, \lambda_d) \in \mathbb{R}^d$.

If the above non-degeneracy assumption holds then we need weaker regularity conditions on the coefficients and the data than in the degenerate case. Recall that $m \geq 0$ and make the following assumptions.

Assumption 4.2 The derivatives in $x \in \mathbb{R}^d$ of a^{ij} up to order $\max(m, 1)$ and of b^i and c up to order m are $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}^d)$-measurable functions, bounded by K for all $i, j \in \{1, 2, \ldots, d\}$. The derivatives in x of the l_2-valued functions σ^i and $l_2(\mathbb{T}^M)$-valued function ν up to order m are $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}^d)$-measurable l_2-valued and $l_2(\mathbb{T}^M)$-valued functions, respectively, in magnitude bounded by K.

Assumption 4.3 The free data, $(f_t)_{t\in[0,T]}$ and $(g_t)_{t\in[0,T]}$ are predictable processes with values in $W_2^{m-1}(\mathbb{R}^d, \mathbb{R}^M)$ and $W_2^m(\mathbb{R}^d, l_2(\mathbb{T}^M))$, respectively, such that almost surely

$$\mathcal{K}_{m-1,2}^2(T) = \int_0^T \left(|f_t|^2_{W_2^{m-1}} + |g_t|^2_{W_2^m}\right) dt < \infty.$$

The initial value, ψ is an $\mathcal{F}_0$-measurable random variable with values in $W_2^m(\mathbb{R}^d, \mathbb{R}^M)$.

The following is a standard result from the L_2-theory of stochastic PDEs. See, for example, [29]. Further results on solvability in W^1_2 spaces for non degenerate systems of stochastic PDEs in $\mathbb{R}^d$ and in domains of $\mathbb{R}^d$ can be found in [15].

Theorem 4.1 *Let Assumptions 4.1, 4.2 and 4.3 hold with $m \geq 0$. Then (3.1)–(3.2) has a unique solution u. Moreover, u is a continuous $W^m_2(\mathbb{R}^d, \mathbb{R}^M)$-valued process such that $u_t \in W^{m+1}(\mathbb{R}^d, \mathbb{R}^M)$ for $P \times dt$ everywhere, and*

$$E \sup_{t\in[0,T]} |u_t|^2_{W^m_2} + E\int_0^T |u_t|^2_{W^{m+1}_2}\,dt \leq N\left(E|\psi|^2_{W^m_2} + E\int_0^T \left(|f_t|^2_{W^{m-1}_2} + |g_t|^2_{W^m_2}\right)dt\right) \tag{4.1}$$

with $N = N(\kappa, m, d, M, K, T)$.

The crucial step in the proof of Theorem 2.1 is to obtain an apriori estimate, like estimate (2.4). In order to discuss the way how such estimate can be proved, take $q = p$, $M = 1$, and for simplicity assume that (a^{ij}) is nonnegative definite, it is bounded and has bounded derivatives up to a sufficiently high order, and that all the other coefficients and free terms in Eq. (2.1) are equal to zero. Thus we consider now the PDE

$$du(t,x) = a^{ij}(t,x)D_{ij}u(t,x)\,dt, \quad t \in [0,T], \quad x \in \mathbb{R}^d, \tag{4.2}$$

with initial condition (2.2), where we assume that ψ is a smooth function from W^1_p. We want to obtain the estimate

$$|u(t)|^p_{W^1_p} \leq N|\psi|^p_{W^1_p} \tag{4.3}$$

for smooth solutions u to (4.2)–(2.2).

After applying D_k to both sides of Eq. (4.2) and writing v_k in place of $D_k v$, by the chain rule we have

$$d\sum_k |u_k|^p = p|u_k|^{p-1}u_k\left(a^{ij}_k u_{ij} + a^{ij}u_{ijk}\right)dt.$$

Integrating over $\mathbb{R}^d$ we get

$$d\sum_k |u_k|^p_{L_p} = \int_{\mathbb{R}^d} Q(u)\,dx\,dt,$$

where

$$Q(u) = p|u_k|^{p-2}u_k\left(a^{ij}u_{ijk} + a^{ij}_k u_{ij}\right).$$

To obtain (4.3) we want to have the estimate

$$\int_{\mathbb{R}^d} Q(v)\,dx \le N||v||^p_{W^1_p} \tag{4.4}$$

for any smooth v with compact support. To prove this we write $\xi \sim \eta$ if ξ and η have identical integrals over $\mathbb{R}^d$ and we write $\xi \preceq \eta$ if $\xi \sim \eta + \zeta$ such that

$$\zeta \le N(|v|^p + |Dv|^p).$$

Then by integration by parts we have

$$\begin{aligned} |v_k|^{p-2}v_k a^{ij} v_{ijk} &\sim -(p-1)|v_k|^{p-2}a^{ij}v_{ki}v_{kj} - |v_k|^{p-2}v_k a^{ij}_i v_{jk} \\ &\sim -(p-1)|v_k|^{p-2}a^{ij}v_{ki}v_{kj} - p^{-1}D_j|v_k|^p a^{ij}_i \\ &\preceq -(p-1)|v_k|^{p-2}a^{ij}v_{ki}v_{kj}. \end{aligned}$$

By the simple inequality $\alpha\beta \le \varepsilon^{-1}\alpha^2 + \varepsilon\beta^2$ we have

$$|v_k|^{p-2}v_k a^{ij}_k v_{ij} \le \varepsilon^{-1}|v_k|^p + \varepsilon|v_k|^{p-2}|a^{ij}_k v_{ij}|^2$$

for any $\varepsilon > 0$. To estimate the term $|a^{ij}_k v_{ij}|^2$ we use the following lemma, which is well-known from [28].

Lemma 4.2 *Let $a = (a^{ij}(x))$ be a function defined on $\mathbb{R}^d$, with values in the set of non-negative $m \times m$ matrices, such that a and its derivatives in x up second order are bounded in magnitude by a constant K. Let V be a symmetric $m \times m$ matrix. Then*

$$|Da^{ij}V^{ij}|^2 \le Na^{ij}V^{ik}V^{jk}$$

for every $x \in \mathbb{R}^d$, where N is a constant depending only on K and d.

By this lemma $|a^{ij}_k v_{ij}|^2 \le Na^{ij}v_{il}v_{jl}$. Hence

$$|v_k|^{p-2}v_k a^{ij}_k v_{ij} \preceq N\varepsilon|v_k|^{p-2}a^{ij}v_{il}v_{jl}.$$

Thus for each fixed $k = 1, 2, \ldots, d$ we have

$$Q(v) \preceq -p(p-1)|v_k|^{p-2}a^{ij}v_{ki}v_{kj} + \varepsilon|v_k|^{p-2}a^{ij}v_{il}v_{jl} \tag{4.5}$$

for any $\varepsilon > 0$. Notice that for each fixed k there is a summation with respect to l over $\{1, 2, \ldots, d\}$ in the expression $\varepsilon|v_k|^{p-2}a^{ij}v_{il}v_{jl}$, and terms with $l \ne k$ cannot be killed by the expression

$$-p(p-1)|v_k|^{p-2}a^{ij}v_{ki}v_{kj}. \tag{4.6}$$

Hence we can get (4.4) when $d = 1$ or $p = 2$, but we does not get it for $p > 2$ and $d > 1$. To cancel every term in the sum $\varepsilon|v_k|^{p-2}a^{ij}v_{il}v_{jl}$ we need an expression like

$$-\nu|v_k|^{p-2}a^{ij}v_{li}v_{lj},$$

with a constant ν, in place of (4.6), for each $k \in \{1,\dots,d\}$ in the right-hand side of (4.5). This suggests to get (4.3) via an equation for $|\,|Du|^2|^{p/2}_{L_{p/2}}$ instead of that for $\sum_k |D_k u|^p_{L_p}$.

Let us test this idea. From

$$du_k = \left(a^{ij}u_{ijk} + a^{ij}_k u_{ij}\right)\,dt$$

by the chain rule and Lemma 4.2 we have

$$\begin{aligned} d|Du|^2 &= 2u_k a^{ij}u_{ijk}\,dt + 2u_k a^{ij}_k u_{ij}\,dt \le a^{ij}\left[|Du|^2\right]_{ij}\,dt - 2a^{ij}u_{ik}u_{jk}\,dt \\ &\quad + N|Du|\left[a^{ij}u_{ik}u_{jk}\right]^{1/2}\,dt \le a^{ij}\left[|Du|^2\right]_{ij}\,dt + N|Du|^2\,dt \end{aligned}$$

with a constant N. Hence

$$d\left(|Du|^2\right)^{p/2} \le (p/2)|Du|^{p-2}a^{ij}\left[|Du|^2\right]_{ij}\,dt + N|Du|^p\,dt,$$

where

$$\begin{aligned} |Du|^{p-2}a^{ij}[|Du|^2]_{ij} &\sim -|Du|^{p-2}a^{ij}_j\left[|Du|^2\right]_i \\ &\quad - ((p-2)/2)|Du|^{p-4}a^{ij}\left[|Du|^2\right]_i\left[|Du|^2\right]_j \\ &\le -(2/p)a^{ij}_j\left[|Du|^p\right]_i \preceq N|Du|^p, \end{aligned} \tag{4.7}$$

which implies

$$|\,|Du|^2|^{p/2}_{L_{p/2}} \le N|\,|D\psi|^2|^{p/2}_{L_{p/2}},$$

by Gronwall's lemma. Consequently, estimate (4.3) follows, since it is not difficult to see that

$$|u(t)|^p_{L_p} \le N|\psi|^p_{L_p}$$

holds. The careful reader may notice that though the computations in (4.7) are justified only for $p \ge 4$, by approximating the function $|t|^{p-2}$, $t \in \mathbb{R}^d$ by smooth functions we can extend them to get the desired estimate for all $p \ge 2$.

The following lemma on Itô's formula in the special case $M = 1$ is Theorem 2.1 from [19]. The proof of this multidimensional variant goes the same way, and therefore will be omitted. Note that for $p \geq 2$ the second derivative, $D_{ij}\langle x\rangle^p$ of the function $(x_1, x_2, \ldots, x_M) \to \langle x\rangle^p$ for $p \geq 2$ is

$$p(p-2)\langle x\rangle^{p-4}x_i x_j + p\langle x\rangle^{p-2}\delta_{ij},$$

which makes the last term in (4.8) below natural. Here and later on we use the convention $0 \cdot 0^{-1} := 0$ whenever such terms occur.

Lemma 4.3 *Let $p \geq 2$ and let $\psi = (\psi^k)_{k=1}^M$ be an $L_p(\mathbb{R}^d, \mathbb{R}^M)$-valued $\mathcal{F}_0$-measurable random variable. For $i = 0, 1, 2, \ldots, d$ and $k = 1, \ldots, M$ let f^{ki} and $(g^{kr})_{r=1}^\infty$ be predictable functions on $\Omega \times (0, T]$, with values in L_p and in $L_p(l_2)$, respectively, such that*

$$\int_0^T \left(\sum_{i,k} |f_t^{ki}|_{L_p}^p + \sum_k |g_t^{k\cdot}|_{L_p}^p \right) dt < \infty \quad \text{(a.s.)}.$$

Suppose that for each $k = 1, \ldots, M$ we are given a W_p^1-valued predictable function u^k on $\Omega \times (0, T]$ such that

$$\int_0^T |u_t^k|_{W_p^1}^p \, dt < \infty \, (a.s.),$$

and for any $\phi \in C_0^\infty$ with probability 1 for all $t \in [0, T]$ we have

$$\big(u_t^k, \phi\big) = \big(\psi^k, \phi\big) + \int_0^t \big(g_s^{kr}, \phi\big)\, dw_s^r + \int_0^t \big((f_s^{k0}, \phi) - (f_s^{ki}, D_i\phi)\big)\, ds.$$

Then there exists a set $\Omega' \subset \Omega$ of full probability such that

$$u = \mathbf{1}_{\Omega'}\big(u^1, \ldots, u^k\big)_{t\in[0,T]}$$

is a continuous $L_p(\mathbb{R}^d, \mathbb{R}^M)$-valued process, and for all $t \in [0, T]$

$$\begin{aligned}
\int_{\mathbb{R}^d} \langle u_t\rangle^p \, dx = {} & \int_{\mathbb{R}^d} \langle \psi\rangle^p \, dx + \int_0^t \int_{\mathbb{R}^d} p\langle u_s\rangle^{p-2}\langle u_s, g_s^r\rangle \, dx \, dw_s^r \\
& + \int_0^t \int_{\mathbb{R}^d} \Big(p\langle u_s\rangle^{p-2}\langle u_s, f_s^0\rangle - p\langle u_s\rangle^{p-2}\langle D_i u_s, f_s^i\rangle \\
& - (1/2)p(p-2)\langle u_s\rangle^{p-4}\langle u_s, f_s^i\rangle D_i\langle u_s\rangle^2 \\
& + \sum_r \big[(1/2)p(p-2)\langle u_s\rangle^{p-4}\langle u_s, g_s^r\rangle^2 + (1/2)p\langle u_s\rangle^{p-2}\langle g_s^r\rangle^2\big]\Big)\, dx\, ds,
\end{aligned} \tag{4.8}$$

where $f^i := (f^{ki})_{k=1}^M$ and $g^r := (g^{kr})_{k=1}^M$ for all $i = 0, 1, \ldots, d$ and $r = 1, 2, \ldots$.

5 The main estimate

Here we consider the problem (3.1)-(3.2) with $a_t = (a_t^{ij}(x))$ taking values in the set of nonnegative symmetric $d \times d$ matrices and the other coefficients and the data are described in (3.3). The following lemma presents the crucial estimate to prove solvability in L_p spaces. It generalises the estimate for Du explained in section 4 for a solution u to a simple PDE.

Lemma 5.1 *Suppose that Assumptions 3.1, 3.2, and 3.3 hold with* $m \geq 0$*. Assume that* $u = (u_t)_{t\in[0,T]}$ *is a solution of* (3.1)–(3.2) *on* $[0, T]$ *(as defined before Theorem 3.1). Then (a.s.)* u *is a continuous* $L_p(\mathbb{R}^d, \mathbb{R}^M)$*-valued process, and there is a constant* $N = N(p, K, d, M, K_0)$ *such that*

$$
\begin{aligned}
&d \int_{\mathbb{R}^d} \langle u_t\rangle^p \, dx + (p/4) \int_{\mathbb{R}^d} \langle u_t\rangle^{p-2} \alpha_t^{ij} \langle D_i u_t, D_j u_t\rangle \, dx \, dt \\
&\leq p \int_{\mathbb{R}^d} \langle u_t\rangle^{p-2} \left\langle u_t, \sigma^{ik} D_i u_t + \nu_t^k u_t + g_t^k \right\rangle dx \, dw_t^k \\
&\quad + N \int_{\mathbb{R}^d} \left[\langle u_t\rangle^p + \langle f_t\rangle^p + \left(\sum_k \langle g_t^k\rangle^2\right)^{p/2} + \left(\sum_k \langle D g_t^k\rangle^2\right)^{p/2} \right] dx \, dt. \qquad (5.1)
\end{aligned}
$$

Proof By Lemma 4.3 (a.s.) u is a continuous $L_p(\mathbb{R}^d, \mathbb{R}^M)$-valued process and

$$
\begin{aligned}
d \int_{\mathbb{R}^d} \langle u_t\rangle^p \, dx = &\int_{\mathbb{R}^d} p \langle u_t\rangle^{p-2} \langle u_t, \sigma^{ik} D_i u_t + \nu_t^k u_t + g_t^k\rangle \, dx \, dw_t^k \\
&+ \int_{\mathbb{R}^d} \Big(p \langle u_t\rangle^{p-2} \langle u_t, b_t^i D_i u_t + c_t u_t + f_t - D_i a_t^{ij} D_j u_t\rangle \\
&- p \langle u_t\rangle^{p-2} \langle D_i u_t, a_t^{ij} D_j u_t\rangle \\
&- (1/2) p(p-2) \langle u_t\rangle^{p-4} D_i \langle u_t\rangle^2 \langle u_t, a_t^{ij} D_j u_t\rangle \\
&+ \sum_k \Big\{ (1/2) p(p-2) \langle u_t\rangle^{p-4} \langle u_t, \sigma_t^{ik} D_i u_t + \nu_t^k u_t + g_t^k\rangle^2 \\
&+ (1/2) p \langle u_t\rangle^{p-2} \langle \sigma_t^{ik} D_i u_t + \nu_t^k u_t + g_t^k\rangle^2 \Big\} \Big) \, dx \, dt. \qquad (5.2)
\end{aligned}
$$

Observe that

$$
\langle u_t\rangle^{p-2} \langle u_t, f_t\rangle \leq \langle u_t\rangle^p + \langle f_t\rangle^p, \quad \langle u_t\rangle^{p-2} \sum_k \langle g_t^k\rangle^2 \leq \langle u_t\rangle^p + \Big(\sum_k \langle g_t^k\rangle^2\Big)^{p/2},
$$

$$
\langle u_t\rangle^{p-2} \sum_k \langle \nu_t^k u_t, g_t^k\rangle \leq N \langle u_t\rangle^{p-1} \left(\sum_k \langle g_t^k\rangle^2\right)^{1/2} \leq N \langle u_t\rangle^p + N \left(\sum_k \langle g_t^k\rangle^2\right)^{p/2},
$$

$$
\langle u_t\rangle^{p-4} \sum_k \langle u_t, g_t^k\rangle^2 \leq \langle u_t\rangle^{p-2} \sum_k \langle g_t^k\rangle^2 \leq \langle u_t\rangle^p + \left(\sum_k \langle g_t^k\rangle^2\right)^{p/2},
$$

$$\langle u_t\rangle^{p-4}\sum_k\langle u_t, \nu_t^k u_t\rangle\langle u_t, g_t^k\rangle \le N\langle u_t\rangle^{p-1}\Big(\sum_k\langle g_t^k\rangle^2\Big)^{1/2} \le \langle u_t\rangle^p + \Big(\sum_k\langle g_t^k\rangle^2\Big)^{p/2},$$
$$\langle u_t\rangle^{p-2}\langle u_t, c_t u_t\rangle \le \langle u_t\rangle^{p-1}\langle c_t u_t\rangle \le |c_t|\langle u_t\rangle^p,$$

where $|c|$ denotes the (Hilbert–Schmidt) norm of c.

This shows how to estimate a few terms on the right in (5.2). We write $\xi \sim \eta$ if ξ and η have identical integrals over $\mathbb{R}^d$ and we write $\xi \preceq \eta$ if $\xi \sim \eta + \zeta$ and the integral of ζ over $\mathbb{R}^d$ can be estimated by the coefficient of dt in the right-hand side of (5.1). For instance, integrating by parts and using the smoothness of σ_t^{ik} and g_t^k we get

$$\begin{aligned} p\langle u_t\rangle^{p-2}\langle\sigma_t^{ik}D_i u_t, g_t^k\rangle &\preceq -p\sigma_t^{ik}(D_i\langle u_t\rangle^{p-2})\langle u_t, g_t^k\rangle \\ &= -p(p-2)\langle u_t\rangle^{p-4}\langle u_t, \sigma_t^{ik}D_i u_t\rangle\langle u_t, g_t^k\rangle, \end{aligned} \tag{5.3}$$

where the first expression comes from the last occurrence of g_t^k in (5.2), and the last one with an opposite sign appears in the evaluation of the first term behind the summation over k in (5.2). Notice, however, that these calculations are not justified when p is close to 2, since in this case $\langle u_t\rangle^{p-2}$ may not be absolutely continuous with respect to x^i and it is not clear either if 0/0 should be defined as 0 when it occurs in the second line. For $p = 2$ we clearly have $\langle\sigma_t^{ik}D_i u_t, g_t^k\rangle \preceq 0$. For $p > 2$ we modify the above calculations by approximating the function $\langle t\rangle^{p-2}$, $t \in \mathbb{R}^M$, by continuously differentiable functions $\phi_n(t) = \varphi_n(\langle t\rangle^2)$ such that

$$\lim_{n\to\infty}\varphi_n(r) = |r|^{(p-2)/2}, \quad \lim_{n\to\infty}\varphi_n'(r) = (p-2)\mathrm{sign}(r)|r|^{(p-4)/2}/2$$

for all $r \in \mathbb{R}$, and

$$|\varphi_n(r)| \le N|r|^{(p-2)/2}, \quad |\varphi_n'(r)| \le N|r|^{(p-4)/2}$$

for all $r \in \mathbb{R}$ and integers $n \ge 1$, where $\varphi_n' := d\varphi_n/dr$ and N is a constant independent of n. Thus instead of (5.3) we have

$$p\varphi_n(\langle u_t\rangle^2)\langle\sigma_t^{ik}D_i u_t, g_t^k\rangle \preceq -2p\varphi_n'(\langle u_t\rangle^2)\langle u_t, \sigma_t^{ik}D_i u_t\rangle\langle u_t, g_t^k\rangle, \tag{5.4}$$

where

$$|\varphi_n'(\langle u_t\rangle^2)\langle u_t, \sigma_t^{ik}D_i u_t\rangle\langle u_t, g_t^k\rangle| \le N\langle u_t\rangle^{p-2}\langle D_i u_t\rangle\langle g_t^k\rangle \tag{5.5}$$

with a constant N independent of n. Letting $n \to \infty$ in (5.4) we get

$$p\langle u_t\rangle^{p-2}\langle\sigma_t^{ik}D_i u_t, g_t^k\rangle \preceq -p(p-2)\langle u_t\rangle^{p-4}\langle u_t, \sigma_t^{ik}D_i u_t\rangle\langle u_t, g_t^k\rangle,$$

where, due to (5.5), 0/0 means 0 when it occurs .

These manipulations allow us to take care of the terms containing f and g and show that to prove the lemma we have to prove

$$\begin{aligned} & p(I_0 + I_1 + I_2) + (p/2)I_3 + [p(p-2)/2](I_4 + I_5) \\ & \quad \preceq -(p/4)\langle u_t\rangle^{p-2}\alpha_t^{ij}\langle D_i u_t, D_j u_t\rangle, \end{aligned} \tag{5.6}$$

where

$$\begin{aligned} & I_0 = -\langle u_t\rangle^{p-2} D_i a_t^{ij}\langle u_t, D_j u_t\rangle, \quad I_1 = -\langle u_t\rangle^{p-2} a_t^{ij}\langle D_i u_t, D_j u_t\rangle \\ & I_2 = \langle u_t\rangle^{p-2}\langle u_t, b_t^i D_i u_t\rangle, \quad I_3 = \langle u_t\rangle^{p-2}\sum_k \langle \sigma_t^{ik} D_i u_t + \nu_t^k u_t\rangle^2, \\ & I_4 = \langle u_t\rangle^{p-4}\sum_k \Big\langle u_t, \sigma_t^{ik} D_i u_t + \nu_t^k u_t\Big\rangle^2, \quad I_5 = -\langle u_t\rangle^{p-4} D_i\langle u_t\rangle^2\langle u_t, a_t^{ij} D_j u_t\rangle. \end{aligned}$$

Observe that

$$I_0 = -(1/2)\langle u_t\rangle^{p-2} D_i a_t^{ij} D_j\langle u_t\rangle^2 = -(1/p) D_j\langle u_t\rangle^p D_i a_t^{ij} \preceq 0,$$

by the smoothness of a. Also notice that

$$I_3 \preceq \langle u_t\rangle^{p-2}\sigma_t^{ik}\sigma_t^{jk}\langle D_i u_t, D_j u_t\rangle + I_6,$$

where

$$I_6 = 2\langle u_t\rangle^{p-2}\sigma_t^{ik}\langle D_i u_t, \nu^k u_t\rangle.$$

It follows that

$$pI_1 + (p/2)I_3 \preceq -(p/2)\langle u_t\rangle^{p-2}\alpha_t^{ij}\langle D_i u_t, D_j u_t\rangle + (p/2)I_6.$$

Next,

$$\begin{aligned} I_4 & \preceq \langle u_t\rangle^{p-4}\sigma_t^{ik}\sigma_t^{jk}\langle u_t, D_i u_t\rangle\langle u_t, D_j u_t\rangle + 2\langle u_t\rangle^{p-4}\sigma_t^{ik}\langle u_t, D_i u_t\rangle\langle u_t, \nu_t^k u_t\rangle \\ & = (1/4)\langle u_t\rangle^{p-4}\sigma_t^{ik}\sigma_t^{jk} D_i\langle u_t\rangle^2 D_j\langle u_t\rangle^2 + [2/(p-2)](D_i\langle u_t\rangle^{p-2})\sigma_t^{ik}\langle u_t, \nu_t^k u_t\rangle \\ & \preceq (1/4)\langle u_t\rangle^{p-4}\sigma_t^{ik}\sigma_t^{jk} D_i\langle u_t\rangle^2 D_j\langle u_t\rangle^2 - [1/(p-2)]I_6 - [2/(p-2)]I_7, \end{aligned}$$

where

$$I_7 = \langle u_t\rangle^{p-2}\sigma_t^{ik}\langle u_t, \nu_t^k D_i u_t\rangle.$$

Hence

$$\begin{aligned} pI_1 & + (p/2)I_3 + [p(p-2)/2](I_4 + I_5) \preceq -(p/2)\langle u_t\rangle^{p-2}\alpha_t^{ij}\langle D_i u_t, D_j u_t\rangle \\ & - [p(p-2)/8]\langle u_t\rangle^{p-4}\alpha_t^{ij} D_i\langle u_t\rangle^2 D_j\langle u_t\rangle^2 - pI_7, \end{aligned}$$

and

$$I_2 - I_7 = \langle u_t\rangle^{p-2}(\langle u_t, b^i_t D_i u_t\rangle - \sigma^{ik}_t\langle u_t, \nu^k_t D_i u_t\rangle) = \langle u_t\rangle^{p-2}\langle u_t, \beta^i_t D_i u_t\rangle,$$

with $\beta^i = b^i - \sigma^{ik}\nu^k$. It follows by Remark 3.2 that the left-hand side of (5.6) is estimated in the order defined by $\preceq$ by

$$\begin{aligned}
&-(p/2)\langle u_t\rangle^{p-2}\alpha^{ij}_t\langle D_i u_t, D_j u_t\rangle\\
&-[p(p-2)/8]\langle u_t\rangle^{p-4}\alpha^{ij}_t D_i\langle u_t\rangle^2 D_j\langle u_t\rangle^2\\
&+K_0 p\langle u_t\rangle^{p-2}\left|\sum_{i,j=1}^d \alpha^{ij}_t\langle D_i u_t, D_j u_t\rangle\right|^{1/2}\langle u_t\rangle + h^i D_i\langle u_t\rangle^p\\
&\preceq -(p/4)\langle u_t\rangle^{p-2}\alpha^{ij}_t\langle D_i u_t, D_j u_t\rangle\\
&-[p(p-2)/8]\langle u_t\rangle^{p-4}\alpha^{ij}_t D_i\langle u_t\rangle^2 D_j\langle u_t\rangle^2\rangle,
\end{aligned} \tag{5.7}$$

where the last relation follows from the elementary inequality $ab \le \varepsilon a^2 + \varepsilon^{-1}b^2$. The lemma is proved. □

Remark 5.1 In the case that $p = 2$ one can replace condition (3.6) with the following:

There are constant $K_0, N \ge 0$ such that for all continuously differentiable $\mathbb{R}^M$-valued functions $u = u(x)$ with compact support in $\mathbb{R}^d$ and all values of the arguments we have

$$\begin{aligned}
\int_{\mathbb{R}^d}\langle u, \beta^i D_i u\rangle\,dx \le{}& N\int_{\mathbb{R}^d}\langle u\rangle^2\,dx\\
&+K_0\int_{\mathbb{R}^d}\left(\left|\sum_{i,j=1}^d \alpha^{ij}\langle D_i u, D_j u\rangle\right|^{1/2}\langle u\rangle + h^i\langle D_i u, u\rangle\right)dx.
\end{aligned} \tag{5.8}$$

This condition is weaker than (3.6) as follows from Remark 3.2 and still by inspecting the above proof we get that u is a continuous $L_2(\mathbb{R}^d, \mathbb{R}^M)$-valued process, and there is a constant $N = N(K, d, M, K_0)$ such that (5.1) holds with $p = 2$.

Remark 5.2 In the case that $p = 2$ and the magnitudes of the first derivatives of b^i are bounded by K one can further replace condition (5.8) with a more tractable one, which is Assumption 3.4.

Indeed, for $\varepsilon > 0$

$$\begin{aligned}
R := \langle u, (\beta^i - h^i I_M)D_i u\rangle &= \tfrac{1}{2}\beta^{ikl}D_i(u^k u^l) + \langle u, (\bar\beta^i - h^i I_M)D_i u\rangle\\
&\le \tfrac{1}{2}\beta^{ikl}D_i(u^k u^l) + \varepsilon\langle(\bar\beta^i - h^i I_M)D_i u\rangle^2/2 + \varepsilon^{-1}\langle u\rangle^2/2.
\end{aligned}$$

Using Assumption 3.4 we get

$$R \le \tfrac{1}{2}\beta^{ikl}D_i(u^k u^l) + \varepsilon M K_0\alpha^{ij}\langle D_i u, D_j u\rangle/2 + \varepsilon^{-1}\langle u\rangle^2/2$$

for every $\varepsilon > 0$. Hence by integration by parts we have

$$\int_{\mathbb{R}^d} \langle u, \beta^i D_i u\rangle \, dx \le N \int_{\mathbb{R}^d} \langle u\rangle^2 \, dx + \int_{\mathbb{R}^d} \langle u, h^i I_M D_i u\rangle \, dx$$
$$+ M K_0 \int_{\mathbb{R}^d} (\varepsilon/2)\alpha^{ij} \langle D_i u_t, D_j u_t\rangle + (\varepsilon^{-1}/2)\langle u\rangle^2 \, dx.$$

Minimising here over $\varepsilon > 0$ we get (5.8). In that case again u is a continuous $L_2(\mathbb{R}^d, \mathbb{R}^M)$-valued process, and there is a constant $N = N(K, d, M, K_0)$ such that (5.1) holds with $p = 2$.

Remark 5.3 If $M = 1$, then condition (3.7) is obviously satisfied with $K_0 = 0$ and $h^i = b^i - \sigma^{ik}\nu^k$.

Also note that in the general case, if the coefficients are smoother, then by *formally* differentiating equation (3.1) with respect to x^i we obtain a new system of equations for the $M \times d$ matrix-valued function

$$v_t = (v_t^{nm}) = Du_t = (D_m u_t^n).$$

We treat the space of $M \times d$ matrices as a Euclidean Md-dimensional space, the coordinates in which are organized in a special way. The inner product in this space is then just $\langle\langle A, B\rangle\rangle = \mathrm{tr} AB^*$. Naturally, linear operators in this space will be given by matrices like $(T^{(nm)(pj)})$, which transforms an $M \times d$ matrix (A^{pj}) into an $M \times d$ matrix (B^{nm}) by the formula

$$B^{nm} = \sum_{p=1}^{m} \sum_{j=1}^{d} T^{(nm)(pj)} A^{pj}.$$

We claim that the coefficients, the initial value and free terms of the system for v_t satisfy Assumptions 3.1, 3.2, and 3.3 with $m \ge 0$ if Assumptions 3.1, 3.2, and 3.3 are satisfied with $m \ge 1$ for the coefficients, the initial value and free terms of the original system for u_t.

Indeed, as is easy to see, v_t satisfies (3.1) with the same σ and a and with $\tilde{b}^i$, $\tilde{c}$, $\tilde{f}$, $\tilde{\nu}^k$, $\tilde{g}^k$ in place of b^i, c, f, ν^k, g^k, respectively, where

$$\tilde{b}^{i(nm)(pj)} = D_m a^{ij}\delta^{pn} + b^{inp}\delta^{jm}, \quad \tilde{c}^{(nm)(pj)} = c^{np}\delta^{mj} + D_m b^{jnp}, \tag{5.9}$$
$$\tilde{f}^{nm} = D_m f^n + u^r D_m c^{nr}, \quad \tilde{\nu}^{k(nm)(pj)} = D_m \sigma^{jk}\delta^{np} + \nu^{knp}\delta^{mj},$$
$$\tilde{g}^{knm} = D_m g^{kn} + u^r D_m \nu^{knr}. \tag{5.10}$$

Then the left-hand side of the counterpart of (3.7) for v is

$$\sum_{m=1}^{d} K_m + \sum_{n=1}^{M} J_n,$$

where (no summation with respect to m)

$$K_m = v^{nm} b^{inr} D_i v^{rm} - \sigma^{ik} v^{nm} \nu^{knr} D_i v^{rm}$$

and (no summation with respect to n)

$$J_n = v^{nm} D_m a^{ij} D_i v^{nj} - \sigma^{ik} v^{nm} D_m \sigma^{jk} D_i v^{nj}.$$

Observe that $D_i v^{nj} = D_{ij} u^n$ implying that

$$\sigma^{ik} D_m \sigma^{jk} D_i v^{nj} = (1/2) D_m (\sigma^{ik} \sigma^{jk}) D_{ij} u^n,$$
$$J_n = (1/2) v^{nm} D_m \alpha^{ij} D_{ij} u^n.$$

By Lemma 4.2 for any $\varepsilon > 0$ and n (still no summation with respect to n)

$$J_n \leq N \varepsilon^{-1} \langle\langle v \rangle\rangle^2 + \varepsilon \alpha^{ij} D_{ik} u^n D_{jk} u^n,$$

which along with the fact that $D_{ik} u^n = D_i v^{nk}$ yields

$$\sum_{n=1}^{M} J_n \leq N \varepsilon^{-1} \langle\langle v \rangle\rangle^2 + \varepsilon \alpha^{ij} \langle\langle D_i v, D_j v \rangle\rangle.$$

Upon minimizing with respect to ε we find

$$\sum_{n=1}^{M} J_n \leq N \left(\sum_{i,j=1}^{d} \alpha^{ij} \langle\langle D_i v, D_j v \rangle\rangle \right)^{1/2} \langle\langle v \rangle\rangle.$$

Next, by assumption for any $\varepsilon > 0$ and m (still no summation with respect to m)

$$K_m \leq N \varepsilon^{-1} \langle\langle v \rangle\rangle^2 + \varepsilon \alpha^{ij} D_i v^{rm} D_j v^{rm} + (1/2) h^i D_i \sum_{r=1}^{M} (v^{rm})^2.$$

We conclude as above that

$$\sum_{m=1}^{d} K_m \leq N \left(\sum_{i,j=1}^{d} \alpha^{ij} \langle\langle D_i v, D_j v \rangle\rangle \right)^{1/2} \langle\langle v \rangle\rangle + h^i \langle\langle D_i v, v \rangle\rangle$$

and this proves our claim.

The above calculations show also that the coefficients, the initial value and the free terms of the system for v_t satisfy Assumptions 3.1, 3.2, and 3.4 with $m - 1 \geq 0$ if Assumptions 3.1, 3.2, and 3.4 are satisfied with $m \geq 1$ for the coefficients, the initial value and free terms of the original equation for u_t. (Note that due to Assumptions 3.1

with $m \geq 1$, $\tilde{b}$, given in (5.9), has first order derivatives in x, which in magnitude are bounded by a constant.)

Now higher order derivatives of u are obviously estimated through lower order ones on the basis of this remark without any additional computations. However, we still need to be sure that we can differentiate equation (3.1).

By the help of the above remarks one can easily estimate the moments of the W^n_p-norms of u using of the following version of Gronwall's lemma.

Lemma 5.2 *Let $y = (y_t)_{t\in[0,T]}$ and $F = (F_t)_{t\in[0,T]}$ be adapted nonnegative stochastic processes and let $m = (m_t)_{t\in[0,T]}$ be a continuous local martingale such that*

$$dy_t \leq (Ny_t + F_t)\,dt + dm_t \quad on\,[0,T] \tag{5.11}$$

$$d[m]_t \leq (Ny_t^2 + y_t^{2(1-\rho)}G_t^{2\rho})\,dt \quad on[0,T], \tag{5.12}$$

with some constants $N \geq 0$ and $\rho \in [0, 1/2]$, and a nonnegative adapted stochastic process $G = (G_t)_{t\in[0,T]}$, such that

$$\int_0^T G_t\,dt < \infty\,(a.s.),$$

where $[m]$ is the quadratic variation process for m. Then for any $q > 0$

$$E\sup_{t\leq T} y_t^q \leq CEy_0^q + CE\left\{\int_0^T (F_t + G_t)\,dt\right\}^q$$

with a constant $C = C(N, q, \rho, T)$.

Proof This lemma improves Lemma 3.7 from [10]. Its proof goes in the same way as that in [10], and can be found in [11]. □

Lemma 5.3 *Let $m \geq 0$. Suppose that Assumptions 3.1, 3.2, and 3.3 are satisfied and assume that $u = (u_t)_{t\in[0,T]}$ is a solution of (3.1)-(3.2) on $[0,T]$ such that (a.s.)*

$$\int_0^T |u_t|^p_{W^{m+1}_p}\,dt < \infty.$$

Then (a.s.) u is a continuous $W^m_p(\mathbb{R}^d, \mathbb{R}^M)$-valued process and for any $q > 0$

$$E\sup_{t\in[0,T]} |u_t|^q_{W^m_p} \leq N(E|\psi|^q_{W^m_p} + E\mathcal{K}^q_{m,p}(T)) \tag{5.13}$$

with a constant $N = N(m, p, q, d, M, K, K_0, T)$. If $p = 2$ and instead of Assumption 3.3 Assumption 3.4 holds and (in case $m = 0$) the magnitudes of the first derivatives of b^i are bounded by K, then u is a continuous $W^m_2(\mathbb{R}^d, \mathbb{R}^M)$-valued process, and for any $q > 0$ estimate (5.13) holds (with $p = 2$).

Proof We are going to prove the lemma by induction on m. First let $m = 0$ and denote $y_t := |u_t|^p_{L_p}$. Then by virtue of Remark 5.2 and Lemma 5.1, the process $y = (y_t)_{t\in[0,T]}$ is an adapted L_p-valued continuous process, and (5.11) holds with

$$F_t := \int_{\mathbb{R}^d} \left[\langle f_t\rangle^p + \left(\sum_k \langle g^k_t\rangle^2\right)^{p/2} + \left(\sum_k \langle Dg^k_t\rangle^2\right)^{p/2} \right] dx,$$

$$m_t := p\int_0^t \int_{\mathbb{R}^d} \langle u_s\rangle^{p-2} \left\langle u_s, \sigma^{ik}_s D_i u_s + \nu^k_s u_s + g^k_s \right\rangle dx\, dw^k_s.$$

Notice that

$$d[m_t] = p^2 \sum_{r=1}^\infty \left(\int_{\mathbb{R}^d} \langle u_t\rangle^{p-2} \langle u_t, \sigma^{ir}_t D_i u_t + \nu^r_t u_t + g^r_t\rangle\, dx \right)^2 dt.$$
$$\leq 3p^2(A_t + B_t + C_t)\, dt,$$

with

$$A_t = \sum_{r=1}^\infty \left(p\int_{\mathbb{R}^d} \langle u_t\rangle^{p-2} \sigma^{ir}_t \langle u_t, D_i u_t\rangle\, dx\right)^2 = \sum_{r=1}^\infty \left(\int_{\mathbb{R}^d} \sigma^{ir}_t D_i \langle u_t\rangle^p\, dx\right)^2,$$

$$B_t = \sum_{r=1}^\infty \left(\int_{\mathbb{R}^d} \langle u_t\rangle^{p-2} \langle u_t, \nu^r_t u_t\rangle\, dx\right)^2, \quad C_t = \sum_{r=1}^\infty \left(\int_{\mathbb{R}^d} \langle u_t\rangle^{p-2} \langle u_t, g^r_t\rangle\, dx\right)^2.$$

Integrating by parts and then using Minkowski's inequality, due to Assumption 2.1, we get $A_t \leq N y_t^2$ with a constant $N = N(K, M, d)$. Using Minkowski's inequality and taking into account that

$$\sum_{r=1}^\infty \langle u, \nu^r u\rangle^2 \leq \langle u\rangle^4 \sum_{r=1}^\infty |\nu^r|^2 \leq N\langle u\rangle^4, \quad \sum_{r=1}^\infty \langle u, g^r\rangle^2 \leq \langle u\rangle^2 |g|,$$

we obtain

$$B_t \leq N y_t^2, \quad C_t \leq \left(\int_{\mathbb{R}^d} \langle u_t\rangle^{p-1} |g_t|\, dx\right)^2 \leq |y_t|^{2(p-1)/p} |g_t|^2_{L_p}.$$

Consequently, condition (5.12) holds with $G_t = |g_t|^p_{L_p}$, $\rho = 1/p$, and we get (5.13) with $m = 0$ by applying Lemma 5.2.

Let $m \geq 1$ and assume that the assertions of the lemma are valid for $m - 1$, in place of m, for any $M \geq 1$, $p \geq 2$ and $q > 0$, for any u, ψ, f and g satisfying the assumptions with $m - 1$ in place of m. Recall the notation $v = (v^{nl}_t) = (D_l u^n_t)$ from Remark 5.3, and that v_t satisfies (3.1) with the same σ and a and with $\tilde{b}^i$, $\tilde{c}$, $\tilde{f}$, $\tilde{\nu}^k$, $\tilde{g}^k$ in place of b^i, c, f, ν^k, g^k, respectively. By virtue of Remarks 5.3 and 5.2 the system for $v = (v_t)_{t\in[0,T]}$ satisfies Assumption 3.3, and it is easy to see that it satisfies also

Assumptions 3.1 and 3.2 with $m-1$ in place of m. Hence by the induction hypothesis v is a continuous $W_p^{m-1}(\mathbb{R}^d, \mathbb{R}^M)$-valued adapted process, and we have

$$E \sup_{t\in[0,T]} |v_t|^q_{W^{m-1}_p} \leq N\left(E|\tilde{\psi}|^q_{W^{m-1}_p} + E\tilde{\mathcal{K}}^q_{m-1,p}(T)\right) \tag{5.14}$$

with a constant $N = N(T, K, K_0, M, d, p, q)$, where $\tilde{\psi}^{nl} = D_l\psi^n$,

$$\tilde{\mathcal{K}}^p_{m-1,p}(T) := \int_0^T \left(|\tilde{f}_t|^p_{W^{m-1}_p} + |\tilde{g}_t|^p_{W^m_p}\right) dt.$$

It follows that $(u_t)_{t\in[0,T]}$ is a $W^m_p(\mathbb{R}^d, \mathbb{R}^M)$-valued continuous adapted process, and by using the induction hypothesis it is easy to see that

$$E\tilde{\mathcal{K}}^q_{m-1,p}(T)) \leq N\left(E|\psi|^q_{W^m_p} + E\mathcal{K}^q_{m,p}(T)\right).$$

Thus (5.13) follows.

If $p = 2$ and Assumption 3.3 is replaced with Assumptions 3.4, then the proof of the conclusion of the lemma goes in the same way with obvious changes. The proof is complete.

6 Proof of Theorems 3.1 and 3.2

First we prove uniqueness. Let $u^{(1)}$ and $u^{(2)}$ be solutions to (3.1)-(3.2), and let Assumptions 3.1, 3.2 and 3.3 hold with $m = 0$. Then $u := u^{(1)} - u^{(2)}$ solves (3.1) with $u_0 = 0$, $g = 0$ and $f = 0$ and Lemma 5.1 and Remark 5.2 are applicable to u. Then using Itô's formula for transforming $|u_t|^p_{L_p} \exp(-\lambda t)$ with a sufficiently large constant λ, after simple calculations we get that almost surely

$$0 \leq e^{-\lambda t}|u_t|^p_{L_p} \leq m_t \quad \text{for all } t \in [0, T],$$

where $m := (m_t)_{t\in[0,T]}$ is a continuous local martingale starting from 0. Hence almost surely $m_t = 0$ for all t, and it follows that almost surely $u^{(1)}_t(x) = u^{(2)}_t(x)$ for all t and almost every $x \in \mathbb{R}^d$. If $p = 2$ and Assumptions 3.1, 3.2 and 3.4 hold and the magnitudes of the first derivatives of b^i are bounded by K and $u^{(1)}$ and $u^{(2)}$ are solutions, then we can repeat the above argument with $p = 2$ to get $u^{(1)} = u^{(2)}$.

To show the existence of solutions we approximate the data of system (3.1) with smooth ones, satisfying also the strong stochastic parabolicity, Assumption 4.1. To this end we will use the approximation described in the following lemma.

Lemma 6.1 *Let Assumptions 3.1 and 3.3 (3.4, respectively) hold with* $m \geq 1$. *Then for every* $\varepsilon \in (0, 1)$ *there exist* $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}^d)$*-measurable smooth (in x) functions* $a^{\varepsilon ij}$, $b^{(\varepsilon)i}$, $c^{(\varepsilon)}$, $\sigma^{(\varepsilon)i}$, $\nu^{(\varepsilon)}$, $D_k a^{\varepsilon ij}$ *and* $h^{(\varepsilon)i}$, *satisfying the following conditions for every* $i, j, k = 1, \ldots, d$.

(i) *There is a constant* $N = N(K)$ *such that*

$$|a^{\varepsilon ij} - a^{ij}| + |b^{(\varepsilon)i} - b^i| + |c^{(\varepsilon)} - c| + |D_k a^{\varepsilon ij} - D_k a^{ij}| \le N\varepsilon,$$
$$|\sigma^{(\varepsilon)i} - \sigma^i| + |\nu^{(\varepsilon)} - \nu| \le N\varepsilon$$

for all (ω, t, x) *and* $i, j, k = 1, \ldots, d$.

(ii) *For every integer* $n \ge 0$ *the partial derivatives in* x *of* $a^{\varepsilon ij}$, $b^{(\varepsilon)i}$, $c^{(\varepsilon)}$, $\sigma^{(\varepsilon)i}$ *and* $\nu^{(\varepsilon)}$ *up to order* n *are* $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}^d)$*-measurable functions, in magnitude bounded by a constant. For* $n = m$ *this constant is independent of* ε*, it depends only on* m*,* M*,* d *and* K*;*

(iii) *For the matrix* $\alpha^{\varepsilon ij} := 2a^{\varepsilon ij} - \sigma^{(\varepsilon)ik}\sigma^{(\varepsilon)jk}$ *we have*

$$\alpha^{\varepsilon ij}\lambda^i\lambda^j \ge \varepsilon \sum_{i=1}^{d} |\lambda^i|^2 \quad \text{for all } \lambda = (\lambda^1, \ldots, \lambda^d) \in \mathbb{R}^d;$$

(iv) *Assumption 3.3 (3.4, respectively) holds for the functions* $\alpha^{\varepsilon ij}$, $\beta^{\varepsilon i} := b^{(\varepsilon)i} - \sigma^{(\varepsilon)ik}\nu^{(\varepsilon)k}$ *and* $h^{(\varepsilon)i}$ *in place of* α^{ij}, β^i *and* h^i*, respectively, with the same constant* K_0.

Proof The proofs of the two statements containing Assumptions 3.3 and 3.4, respectively, go in essentially the same way, therefore we only detail the former. Let ζ be a nonnegative smooth function on $\mathbb{R}^d$ with unit integral and support in the unit ball, and let $\zeta_\varepsilon(x) = \varepsilon^{-d}\zeta(x/\varepsilon)$. Define

$$b^{(\varepsilon)i} = b^i * \zeta_\varepsilon,\ c^{(\varepsilon)} = c * \zeta_\varepsilon,\ \sigma^{(\varepsilon)i} = \sigma^i * \zeta_\varepsilon,\ \nu^{(\varepsilon)} = \nu * \zeta_\varepsilon,\ h^{(\varepsilon)i} = h^i * \zeta_\varepsilon,$$

and $a^{\varepsilon ij} = a^{ij} * \zeta_\varepsilon + k\varepsilon\delta_{ij}$ with a constant $k > 0$ determined later, where δ_{ij} is the Kronecker symbol and '$*$' means the convolution in the variable $x \in \mathbb{R}^d$. Since we have mollified functions which are bounded and Lipschitz continuous, the mollified functions, together with $a^{\varepsilon ij}$ and $D_k a^{\varepsilon ij}$, satisfy conditions (i) and (ii). Furthermore,

$$|\sigma^{(\varepsilon)ir}\nu^{(\varepsilon)r} - \sigma^{ir}\nu^r| \le |\sigma^{(\varepsilon)i} - \sigma^i||\nu^{(\varepsilon)}| + |\sigma^i||\nu^{(\varepsilon)} - \nu| \le 2K^2\varepsilon,$$

for every $i = 1, \ldots, d$. Similarly,

$$|\sigma^{(\varepsilon)ir}\sigma^{(\varepsilon)jr} - \sigma^{ir}\sigma^{jr}| \le 2K^2\varepsilon, \quad |b^{(\varepsilon)i} - b^i| \le K\varepsilon, \quad |h^{(\varepsilon)i} - h^i| \le N\varepsilon$$

for all $i, j = 1, 2, \ldots, d$. Hence setting

$$B^{\varepsilon i} = b^{(\varepsilon)i} - \sigma^{(\varepsilon)ik}\nu^{(\varepsilon)k} - h^{(\varepsilon)i} I_M,$$

and using the notation B^i for the same expression without the superscript 'ε', we have

$$|B^{\varepsilon i} - B^i| \le |b^{(\varepsilon)i} - b^i| + |\sigma^{(\varepsilon)ir}\nu^{(\varepsilon)r} - \sigma^{ir}\nu^r| + \sqrt{M}|h^{(\varepsilon)i} - h^i| \le R\varepsilon,$$
$$|B^{(\varepsilon)i} + B^i| \le R$$

with a constant $R = R(M, K)$. Thus for any $z_1,\ldots,z_d$ vectors from $\mathbb{R}^M$

$$|\langle B^{\varepsilon i} z_i\rangle^2 - \langle B^i z_i\rangle^2| = |\langle (B^{\varepsilon i} - B^i) z_i, (B^{\varepsilon j} + B^j) z_j\rangle| \\ \leq |B^{\varepsilon i} - B^i||B^{\varepsilon j} + B^j|\langle z_i\rangle\langle z_j\rangle \leq dR^2\varepsilon\sum_{i=1}^{d}\langle z_i\rangle^2.$$

Therefore

$$\langle B^{\varepsilon i} z_i\rangle^2 \leq \langle B^i z_i\rangle^2 + C_1\varepsilon\sum_{i=1}^{d}\langle z_i\rangle^2$$

with a constant $C_1 = C_1(M, K, d)$. Similarly,

$$\sum_{i,j}\left(2a^{\varepsilon ij} - \sigma^{(\varepsilon)ik}\sigma^{(\varepsilon)jk}\right)\langle z_i, z_j\rangle \\ \geq \sum_{i,j}(2a^{ij} - \sigma^{ik}\sigma^{jk})\langle z_i, z_j\rangle + (k - C_2)\varepsilon\sum_i\langle z_i\rangle^2$$

with a constant $C_2 = C_2(K, m, d)$. Consequently,

$$\langle (\beta^{\varepsilon i} - h^{(\varepsilon)i} I_M) z_i\rangle^2 \leq \langle B^i z_i\rangle^2 + C_1\varepsilon\sum_{i=1}^{d}\langle z_i\rangle^2 \\ \leq K_0\sum_{i,j=1}^{d}\alpha^{ij}\langle z_i, z_j\rangle + C_1\varepsilon\sum_{i=1}^{d}\langle z_i\rangle^2 \\ \leq K_0\sum_{i,j=1}^{d}\alpha^{\varepsilon ij}\langle z_i, z_j\rangle + (K_0(C_2 - k) + C_1)\varepsilon\sum_{i=1}^{d}\langle z_i\rangle^2.$$

Choosing k such that $K_0(C_2 - k) + C_1 = -K_0$ we get

$$\langle (\beta^{\varepsilon i} - h^{(\varepsilon)i} I_M) z_i\rangle^2 + K_0\varepsilon\sum_{i=1}^{d}\langle z_i\rangle^2 \leq K_0\sum_{i,j=1}^{d}\alpha^{\varepsilon ij}\langle z_i, z_j\rangle.$$

Hence statements (iii) and (iv) follow immediately. □

Now we start with the proof of the existence of solutions which are $W^m_p(\mathbb{R}^d, \mathbb{R}^M)$-valued if the Assumptions 3.1, 3.2 and 3.3 hold with $m \geq 1$. First we make the additional assumptions that ψ, f and g vanish for $|x| \geq R$ for some $R > 0$, and that $q \in [2, \infty)$ and

$$E|\psi|^q_{W^m_p} + E\mathcal{K}^q_{m,q}(T) < \infty. \tag{6.1}$$

For each $\varepsilon > 0$ we consider the system

$$du_t^\varepsilon = \left[\sigma_t^{(\varepsilon)ir} D_i u_t^\varepsilon + \nu_t^{(\varepsilon)r} u_t^\varepsilon + g_t^{(\varepsilon)r}\right] dw_t^r \\ + \left[a_t^{\varepsilon ij} D_{ij} u_t^\varepsilon + b_t^{(\varepsilon)i} D_i u_t^\varepsilon + f_t^{(\varepsilon)}\right] dt \tag{6.2}$$

with initial condition

$$u_0^{(\varepsilon)} = \psi^{(\varepsilon)}, \tag{6.3}$$

where the coefficients are taken from Lemma 6.1, and $\psi^{(\epsilon)}$, $f^{(\epsilon)}$ and $g^{(\epsilon)}$ are defined as the convolution of ψ, f and g, respectively, with $\zeta_\varepsilon(\cdot) = \varepsilon^{-d}\zeta(\cdot/\varepsilon)$ for $\zeta \in C_0^\infty(\mathbb{R}^d)$ taken from the proof of Lemma 6.1. By Theorem 4.1 the above equation has a unique solution u^ε, which is a $W_2^n(\mathbb{R}^d, \mathbb{R}^M)$-valued continuous process for all n. Hence, by Sobolev embeddings, u^ε is a $W_p^{m+1}(\mathbb{R}^d, \mathbb{R}^M)$-valued continuous process, and therefore we can use Lemma 5.3 to get

$$E \sup_{t\in[0,T]} |u_t^\varepsilon|^q_{W^n_{p'}} \le N\left(E|\psi^{(\varepsilon)}|^q_{W^n_{p'}} + E(\mathcal{K}^\varepsilon_{n,p'})^q(T)\right) \tag{6.4}$$

for $p' \in \{p, 2\}$ and $n = 0, 1, 2, \ldots, m$, where $\mathcal{K}^\varepsilon_{n,p'}$ is defined by (3.4) with $f^{(\varepsilon)}$ and $g^{(\varepsilon)}$ in place of f and g, respectively. Keeping in mind that $T^{1/r} \le \max\{1, T\}$, and using basic properties of convolution, we can conclude that

$$E\left(\int_0^T |u_t^\varepsilon|^r_{W^n_{p'}}\,dt\right)^{q/r} \le N(E|\psi|^q_{W^n_{p'}} + E\mathcal{K}^q_{n,p'}(T)) \tag{6.5}$$

for any $r > 1$ and with $N = N(m, p, q, d, M, K, T)$ not depending on r.

For integers $n \ge 0$, and any $r, q \in (1, \infty)$ let $\mathbb{H}^n_{p,r,q}$ be the space of $\mathbb{R}^M$-valued functions $v = v_t(x) = (v_t^i(x))_{i=1}^M$ on $\Omega \times [0, T] \times \mathbb{R}^d$ such that $v = (v_t(\cdot))_{t\in[0,T]}$ are $W^n_p(\mathbb{R}^d, \mathbb{R}^M)$-valued predictable processes and

$$|v|^q_{\mathbb{H}^n_{p,r,q}} = E\left(\int_0^T |v_t|^r_{W^n_p}\,dt\right)^{q/r} < \infty.$$

Then $\mathbb{H}^n_{p,r,q}$ with the norm defined above is a reflexive Banach space for each $n \ge 0$ and $p, r, q \in (1, \infty)$. We use the notation $\mathbb{H}^n_{p,q}$ for $\mathbb{H}^n_{p,q,q}$.

By Assumption 3.2 the right-hand side of (6.5) is finite for $p' = p$ and also for $p = 2$ since ψ, f and g vanish for $|x| \ge R$. Thus there exists a sequence $(\varepsilon_k)_{k\in\mathbb{N}}$ such that $\varepsilon_k \to 0$ and for $p' = p, 2$ and integers $r > 1$ and $n \in [0, m]$ the sequence $v^k := u^{\varepsilon_k}$ converges weakly in $\mathbb{H}^n_{p'r,q}$ to some $v \in H^m_{p',r,q}$, which therefore also satisfies

$$E\left(\int_0^T |v_t|^r_{W^n_{p'}}\,dt\right)^{q/r} \le N(E|\psi|^q_{W^n_{p'}} + E\mathcal{K}^q_{n,q}(T))$$

for $p' = p, 2$ and integers $r > 1$. Using this with $p' = p$ and letting $r \to \infty$ by Fatou's lemma we obtain

$$E \operatorname*{ess\,sup}_{t\in[0,T]} |v_t|^q_{W^n_p} \le N(E|\psi|^q_{W^n_p} + E\mathcal{K}^q_{n,p}(T)) \quad \text{for } n = 0, 1, \ldots, m. \tag{6.6}$$

Now we are going to show that a suitable stochastic modification of v is a solution of (3.1)-(3.2). To this end we fix an $\mathbb{R}^M$-valued function φ in $C_0^\infty(\mathbb{R}^d)$ and a predictable real-valued process $(\eta_t)_{t\in[0,T]}$, which is bounded by some constant C, and define the functionals Φ, Φ_k, Ψ and Ψ_k over $\mathbb{H}^1_{p,q}$ by

$$\begin{aligned}
\Phi_k(u) &= E\int_0^T \eta_t \int_0^t \left\{-(a_s^{\varepsilon_k ij} D_i u_s, D_j\varphi) + (\bar{b}_s^{\varepsilon_k i} D_i u_s + c_s^{(\varepsilon_k)} u_s, \varphi)\right\} ds\, dt,\\
\Phi(u) &= E\int_0^T \eta_t \int_0^t \left\{-(a_s^{ij} D_i u_s, D_j\varphi) + (\bar{b}_s^{i} D_i u_s + c_s u_s, \varphi)\right\} ds\, dt,\\
\Psi(u) &= E\int_0^T \eta_t \int_0^t (\sigma_t^{ir} D_i u_t + \nu_t^r u_t, \varphi)\, dw_t^r\, dt\\
\Psi_k(u) &= E\int_0^T \eta_t \int_0^t \left(\sigma_t^{(\varepsilon_k)ir} D_i u_t + \nu_t^{(\varepsilon_k)r} u_t, \varphi\right) dw_t^r\, dt
\end{aligned}$$

for $u \in \mathbb{H}^1_{p,q}$ for each $k \ge 1$, where $\bar{b}^{\varepsilon i} = b^{(\varepsilon)i} - D_j a^{\varepsilon ij} I_M$. By the Bunyakovsky-Cauchy-Schwarz and the Burkholder-Davis-Gundy inequalities for all $u \in \mathbb{H}^1_{p,q}$ we have

$$\begin{aligned}
\Phi(u) &\le CNT^{2-1/q}|u|_{\mathbb{H}^1_{p,q}}|\varphi|_{W^1_{\bar{p}}},\\
\Psi(u) &\le CTE\sup_{t\le T}\left|\int_0^t (\sigma_t^{ir} D_i u_t + \nu_t^r u_t, \varphi)\, dw_t^r\right|\\
&\le 3CTE\left\{\int_0^T \sum_{r=1}^\infty (\sigma_t^{ir} D_i u_t + \nu_t^r u_t, \varphi)^2\, dt\right\}^{1/2}\\
&\le 3CTE\left\{\int_0^T \left(\int_{\mathbb{R}^d} |\langle \sigma_t^{ir} D_i u_t + \nu_t^r u_t, \varphi\rangle|_{l_2}\, dx\right)^2 dt\right\}^{1/2}\\
&\le CTNE\left\{\int_0^T |u_t|^2_{W^1_p}|\varphi|^2_{W^1_{\bar{p}}}\, dt\right\}^{1/2} \le CNT^{q/2}|u|_{\mathbb{H}^1_{p,q}}|\varphi|_{W^1_{\bar{p}}}
\end{aligned}$$

with a constant $N = N(K, d, M)$, where $\bar{p} = p/(p-1)$. (In the last inequality we make use of the assumption $q \ge 2$.) Consequently, Φ and Ψ are continuous linear functionals over $\mathbb{H}^1_{p,q}$, and therefore

$$\lim_{k\to\infty}\Phi(v^k) = \Phi(v), \quad \lim_{k\to\infty}\Psi(v^k) = \Psi(v). \tag{6.7}$$

Using statement (i) of Lemma 6.1, we get

$$|\Phi_k(u) - \Phi(u)| + |\Psi_k(u) - \Psi(u)| \leq N\varepsilon_k |u|_{\mathbb{H}^1_{p,q}} |\varphi|_{W^1_{\bar{p}}} \tag{6.8}$$

for all $u \in \mathbb{H}^1_{p,q}$ with a constant $N = N(k, d, M)$. Since u^ε is the solution of (6.2)-(6.3), we have

$$E \int_0^T \eta_t(v^k_t, \varphi)\, dt = E \int_0^T \eta_t(\psi^k, \varphi)\, dt + \Phi(v^k) + \Psi(v^k)$$
$$+ F(f^{(\varepsilon_k)}) + G(g^{(\varepsilon_k)}) \tag{6.9}$$

for each k, where

$$F(f^{(\varepsilon_k)}) = E \int_0^T \eta_t \int_0^t \left(f^{(\varepsilon_k)}_s, \varphi \right) ds\, dt,$$
$$G(g^{(\varepsilon_k)}) = E \int_0^T \eta_t \int_0^t \left(g^{(\varepsilon_k)r}_s, \varphi \right) dw^r_s\, dt.$$

Taking into account that $|v^k|_{\mathbb{H}^1_{p,q}}$ is a bounded sequence, from (6.7) and (6.8) we obtain

$$\lim_{k\to\infty} \Phi_n(v^k) = \Phi(v), \quad \lim_{k\to\infty} \Psi_k(v^k) = \Psi(v). \tag{6.10}$$

One can see similarly (in fact easier), that

$$\lim_{k\to\infty} E \int_0^T \eta_t(v^k_t, \varphi)\, dt = E \int_0^T \eta_t(v_t, \varphi)\, dt, \tag{6.11}$$

$$\lim_{k\to\infty} E \int_0^T \eta_t(\psi^{(\varepsilon_k)}_t, \varphi)\, dt = E \int_0^T \eta_t(\psi, \varphi)\, dt, \tag{6.12}$$

$$\lim_{k\to\infty} F(f^{(\varepsilon_k)}) = F(f), \quad \lim_{k\to\infty} G(g^{(\varepsilon_k)}) = G(g). \tag{6.13}$$

Letting $k \to \infty$ in (6.9), and using (6.10) through (6.13) we obtain

$$E \int_0^T \eta_t(v_t, \varphi)\, dt$$
$$= E \int_0^T \eta_t \Big\{ (\psi, \varphi) + \int_0^t \big[-(a^{ij}_s D_i u_s, D_j \varphi) + (\bar{b}^i_s D_i u_s + c_s u_s + f_s, \varphi) \big] ds$$
$$+ \int_0^t (\sigma^{ir} D_i v_s + \nu^r v_s, \varphi)\, dw^r_s \Big\}\, dt$$

for every bounded predictable process $(\eta_t)_{t\in[0,T]}$ and φ from C_0^∞. Hence for each $\varphi\in C_0^\infty$

$$(v_t,\varphi)=(\psi,\varphi)+\int_0^t\big[-(a_s^{ij}D_iv_s,D_j\varphi)+(\bar b_s^iD_iv_s+c_sv_s+f_s,\varphi)\big]\,ds$$
$$+\int_0^t(\sigma^{ir}D_iv_s+\nu^rv_s+g_s^r,\varphi)\,dw_s^r$$

holds for $P\times dt$ almost every $(\omega,t)\in\Omega\times[0,T]$. Substituting here $(-1)^{|\alpha|}D^\alpha\varphi$ in place of φ for a multi-index $\alpha=(\alpha_1,\dots,\alpha_d)$ of length $|\alpha|\le m-1$ and integrating by parts, we see that

$$(D^\alpha v_t,\varphi)=(D^\alpha\psi,\varphi)+\int_0^t\big[-(F_s^j,D_j\varphi)+(F_s^0,\varphi)\big]\,ds+\int_0^t(G_s^r,\varphi)\,dw_s^r \tag{6.14}$$

for $P\times dt$ almost every $(\omega,t)\in\Omega\times[0,T]$, where, owing to the fact that (6.6) also holds with 2 in place of p, F^i and $(G^r)_{r=1}^\infty$ are predictable processes with values in L_2-spaces for $i=0,1,\dots,d$, such that

$$\int_0^T\left(\sum_{i=0}^d|F_s^i|_{L_2}^2+|G_s|_{L_2}^2\right)ds<\infty\quad\text{(a.s.)}.$$

Hence the theorem on Itô's formula from [21] implies that in the equivalence class of v in $\mathbb{H}_{2,q}^m$ there is a $W_2^{m-1}(\mathbb{R}^d,\mathbb{R}^M)$-valued continuous process, $u=(u_t)_{t\in[0,T]}$, and (6.14) with u in place of v holds for any $\varphi\in C_0^\infty(\mathbb{R}^d)$ almost surely for all $t\in[0,T]$. After that an application of Lemma 4.3 to $D^\alpha u$ for $|\alpha|\le m-1$ yields that $D^\alpha u$ is an $L_p(\mathbb{R}^d,\mathbb{R}^M)$-valued, strongly continuous process for every $|\alpha|\le m-1$, i.e., u is a $W_p^{m-1}(\mathbb{R}^d,\mathbb{R}^M)$-valued strongly continuous process. This, (6.6), and the denseness of C_0^∞ in $W_p^m(\mathbb{R}^d,\mathbb{R}^M)$ implies that (a.s.) u is a $W_p^m(\mathbb{R}^d,\mathbb{R}^M)$-valued weakly continuous process and (3.11) holds.

To prove the theorem without the assumption that ψ, f and g have compact support, we take a $\zeta\in C_0^\infty(\mathbb{R}^d)$ such that $\zeta(x)=1$ for $|x|\le1$ and $\zeta(x)=0$ for $|x|\ge2$, and define $\zeta_n(\cdot)=\zeta(\cdot/n)$ for $n>0$. Let $u(n)=(u_t(n))_{t\in[0,T]}$ denote the solution of (3.1)-(3.2) with $\zeta_n\psi$, ζ_nf and ζ_ng in place of ψ, f and g, respectively. By virtue of what we have proved above, $u(n)$ is a weakly continuous $W_p^m(\mathbb{R}^d,\mathbb{R}^M)$-valued process, and

$$E\sup_{t\in[0,T]}|u_t(n)-u_t(l)|_{W_p^m}^q\le NE|(\zeta_n-\zeta_l)\psi|_{W_p^m}^q$$
$$+NE\left(\int_0^T\left\{|(\zeta_n-\zeta_l)f_s|_{W_p^m}^p+|(\zeta_n-\zeta_l)g_s|_{W_p^{m+1}}^p\right\}ds\right)^{q/p}.$$

Letting here $n,l\to\infty$ and applying Lebesgue's theorem on dominated convergence in the left-hand side, we see that the right-hand side of the inequality tends to zero. Thus for a subsequence $n_k \to \infty$ we have that $u_t(n_k)$ converges strongly in $W^m_p(\mathbb{R}^d, \mathbb{R}^M)$, uniformly in $t \in [0, T]$, to a process u. Hence u is a weakly continuous $W^m_p(\mathbb{R}^d, \mathbb{R}^M)$-valued process. It is easy to show that it solves (3.1)–(3.2) and satisfies (3.11).

By using a standard stopping time argument we can dispense with condition (6.1). Finally we can prove estimate (3.11) for $q \in (0, 2)$ by applying Lemma 3.2 from [8] in the same way as it is used there to prove the corresponding estimate in the case $M = 1$. The proof of the Theorem 3.1 is complete. We have already showed the uniqueness statement of Theorem 3.2, the proof of the other assertions goes in the above way with obvious changes.

Acknowledgments The results of this paper were presented at the 9th International Meeting on "Stochastic Partial Differential Equations and Applications" in Levico Terme in Italy, in January, 2014, and at the meeting on "Stochastic Processes and Differential Equations in Infinite Dimensional Spaces" in King's College London, in March, 2014. The authors would like to thank the organisers for these possibilities. The authors are grateful to the referee whose comments and suggestions helped to improve the presentation of the paper. The work of the third author was partially supported by NSF Grant DMS-1160569.

References

1. Brzeźniak, Z., Flandoli, F., Maurelli, M.: Existence and uniqueness for stochastic 2D Euler flows with bounded vorticity. arXiv:1401.5938
2. Dalang, R.C.: Extending martingale measure stochastic integral, with applications to spatially homogeneous S.P.D.E's. Electron. J. Probab. **4**, 1–21 (1999)
3. Dalang, R.C., Quer-Sardanyons, L.: Stochastic integrals for spde's: a comparison. Expositiones Mathematicae **29**, 67–109 (2011)
4. Da Prato, G., Zabczyk, J.: Stochastic Equations in Infinite Dimensions. Cambridge University Press, Cambridge (1992)
5. Gerencsér, M., Gyöngy, I.: On stochastic finite difference schemes. Appl. Math. Opt. (2014). doi:10.1007/s00245-014-9272-2
6. Gyöngy, I., Krylov, N.V.: On stochastic equations with respect to semimartingales I. Stochastics **4**(1), 1–21 (1980)
7. Gyöngy, I., Krylov, N.V.: Stochastic partial differential equations with unbounded coefficients and applications III. Stoch. Stoch. Rep. **40**, 77–115 (1992)
8. Gyöngy, I., Krylov, N.V.: On the rate of convergence of splitting-up approximations for SPDEs. In Progress in Probability, vol. 56, pp. 301–321. Birkhäuser, Berlin (2003)
9. Gyöngy, I., Krylov, N.V.: Expansion of solutions of parametrized equations and acceleration of numerical methods. Ill. J. Math. **50**, 473–514 (2006). Special Volume in Memory of Joseph Doob (1910–2004)
10. Gyöngy, I., Shmatkov, A.: Rate of convergence of Wong–Zakai approximations for stochastic partial differential equations. Appl. Math. Opt. **54**(3), 315–341 (2006)
11. Gyöngy, I.: An introduction to the theory of stochastic partial differential equations. In preparation
12. Ikeda, N., Watanabe, S.: Stochastic Differential Equations and Diffusion Processes. North-Holland, Amsterdam (1981)
13. Kim, K.: L_p estimates for SPDE with discontinuous coefficients in domains. Electron. J. Probab. **10**, 1–20 (2005)
14. Kim, K.: $L_q(L_p)$-theory of parabolic PDEs with variable coefficients. Bull. Korean Math. Soc. **45**(1), 169–190 (2008)
15. Kim, K., Lee, K.: A W^1_2-theory of stochastic partial differential systems of divergence type on C^1 domains. Electron. J. Probab. **16**, 1296–1317 (2011)
16. Krylov, N.V.: On SPDE's and superdiffusions. Ann. Probab. **25**, 1789–1809 (1997)
17. Krylov, N.V.: Stochastic Partial Differential Equations: Six Perspectives, Mathematical Surveys and Monographs. An analytic approach to SPDEs, pp. 185–242. American Mathematical Society, Providence, RI (1999)
18. Krylov, N.V.: SPDEs in $L_q((0, \tau], L_p)$ spaces. Electron. J. Prob. **5**(13), 1–29 (2000)

19. Krylov, N.V.: Itô's formula for the L_p-norm of stochastic W_p^1-valued processes. Probab. Theory Relat. Fields **147**, 583–605 (2010)
20. Krylov, N.V., Rozovskii, B.L.: On the Cauchy problem for linear stochastic partial differential equations. Math. USSR Izvestija **11**(6), 1267–1284 (1977)
21. Krylov, N.V., Rosovskii, B.L.: Stochastic evolution equations. J. Soviet Math. **16**, 1233–1277 (1981)
22. Krylov, N.V., Rozovskii, B.L.: Characteristics of degenerating second-order parabolic Itô equations. J. Soviet Maths. **32**, 336–348 (1986). Translated from Trudy Seminara imeni I.G. Petrovskogo, No. 8. pp. 153–168, (1982)
23. Moffatt, H.K.: Magnetic Field Generation in Electrically Conducting Fluids. Cambridge University Press, Cambridge (1978)
24. Molchanov, S.A., Ruzmaikin, A.A., Sokolov, D.D.: Short-correlated random flow as a fast dynamo. Dokl. Akad. Nauk SSSR **295**(3), 576–579 (1987)
25. Oleĭnik, O.A.: Alcuni risultati sulle equazioni lineari e quasi lineari ellittico-paraboliche a derivate parziali del secondo ordine. (Italian) Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur. **40**(8), 775–784 (1966)
26. Oleĭnik, O.A.: On the smoothness of solutions of degenerating elliptic and parabolic equations. Dokl. Akad. Nauk SSSR **163**, 577–580 (1965). in Russian; English translation in Soviet Mat. Dokl., Vol. 6, No. 3, 972–976 (1965)
27. Oleĭnik, O.A., Radkevič, E.V.: Second order equations with nonnegative characteristic form, Mathematical Analysis, 1969, pp. 7–252. (errata insert) Akad. Nauk SSSR, Vsesojuzn. Inst. Naučn. i Tehn. Informacii, Moscow, : in Russian, p. 1973. Plenum Press, New York-London, English translation (1971)
28. Olejnik, O.A., Radkevich, E.V.: Second Order Equations with Nonnegative Characteristic Form. AMS, Providence (1973)
29. Rozovskii, B.L.: Stochastic Evolution Systems. Linear Theory and Applications to Nonlinear Filtering. Kluwer, Dordrecht (1990)
30. Walsh, J.B.: An introduction to stochastic partial differential equations. d'Eté de Prob. de St-Flour XIV, 1984. Lecture Notes in Mathematics, vol. 1180. Springer, New York (1986)