A regularity theorem for quasilinear parabolic systems under random perturbations

Lisa Beck*   Franco Flandoli†

September 25, 2013

Abstract

Several aspects of regularity theory for parabolic systems are investigated under the effect of random perturbations. The deterministic theory, when strict parabolicity is assumed, presents both classes of systems where all weak solutions are in fact more regular, and examples of systems with weak solutions which develop singularities in finite time. Our main result is the extension of a regularity result due to Kalita to the stochastic case, which concerns local Hölder continuity of weak solutions in the vectorial case. For the proof we apply stochastic versions of methods which are classical in the deterministic case (such as difference quotient techniques, higher integrability by embedding theorems and a version of Moser's iteration technique). This might be of interest on their own.

MSC (2010): 60H15 (primary); 60H30, 35B65, 35R60 (secondary)

1 Introduction

Nonlinear parabolic systems of the form

$$\partial_t u = \operatorname{div} A(x, t, u, Du), \qquad u|_{t=0} = u_0 \tag{1.1}$$

on a cylindrical domain $D \times (0, T)$, with $D \subset \mathbb{R}^n$ a bounded, regular domain, $u \colon D \times [0, T] \to \mathbb{R}^N$ a vector-valued function, $A \colon D \times [0, T] \times \mathbb{R}^N \times \mathbb{R}^{nN} \to \mathbb{R}^{nN}$ a vector field, have been investigated intensively in the past decades. Various techniques were developed which then were applicable in different settings or under different structure assumptions and which eventually led to a better understanding of the issues of existence and regularity for weak solutions to such systems. A key feature in the vectorial case $N > 1$ is that, under the strict parabolicity assumption

$$\sum_{i,j=1}^{n} \sum_{\alpha,\beta=1}^{N} \frac{\partial A_i^\alpha}{\partial z_j^\beta}(x, t, u, z)\, \xi_i^\alpha \xi_j^\beta \geq \lambda_0 |\xi|^2 \qquad \text{for all } \xi \in \mathbb{R}^{nN}$$

and some differentiability assumption on A with respect to the (x, u)-variable, there are classes of vector fields $A(x, t, u, z)$ such that all weak solutions to (1.1) are in fact more regular, and examples of systems such that there exist weak solutions with singularities; this dichotomy does not happen for single equations, the case $N = 1$, where regularity of weak solutions is always true, due to the (elliptic and parabolic) works based on the fundamental results of De Giorgi, Nash and Moser [8, 30, 29].

The aim of the present paper is to expand some of these by now classical techniques to the framework of parabolic systems under random perturbations and to discuss as an example of its applications the extension of a classical regularity result for parabolic systems to the stochastic case. For this reason, we start by giving first some background on the deterministic theory (without stochastic noise), by sketching some basic regularity results obtained in the elliptic and parabolic theory. Since the interesting phenomena already occur for very simple systems, we here state the results only for some particular cases in order to allow a clearer exposition, and we refer to the literature for a more detailed discussion of the general results.

*Hausdorff Center for Mathematics, Universität Bonn, Germany. Email address: `lisa.beck@hcm.uni-bonn.de`.

†Dipartimento di Matematica "U. Dini", Università di Pisa, Italy. Email address: `flandoli@dma.unipi.it`.

Elliptic systems. For reasons of comparison we begin the discussion with the case of elliptic systems and stationary weak solutions (i. e. system and solution do not depend on the time variable). The notion of weak solution here refers always to a function u which belongs to a suitable Sobolev space and which solves the system in an integral sense, i. e. $\int_D A(x, u, Du) \cdot D\varphi \, dx = 0$ for all $\varphi \in C_0^\infty(D, \mathbb{R}^N)$. As already mentioned above, in the scalar case $N = 1$ weak solutions are locally continuous under quite general assumptions, that is under a suitable ellipticity condition and a corresponding growth condition on the vector field $A(x, u, z)$, see [8, 30, 29]. Roughly speaking, weak solutions are always regular, and the degree of regularity (Hölder regularity and differentiability) is determined by the regularity of the system. In the vectorial setting $N > 1$, which will be studied in the present paper, it turns out that the different component functions of a weak solution might interact in a nontrivial way, in the sense that full regularity cannot be expected. In fact, counterexamples by De Giorgi, Giusti, Miranda [9, 19] and others revealed that discontinuous weak solutions (having a smooth trace on ∂D) exist, even for analytic vector fields which are linear in the gradient variable. Instead, again under suitable additional assumptions on growth and regularity of the vector field $A(x, u, z)$, there are partial regularity results available, yielding Hölder regularity of the solution u (or of its gradient Du) outside of a negligible set, the singular set of u (or of Du). Hence, for general systems, the best regularity to hope for is partial regularity of Du, with an estimate for the Hausdorff dimension of the singular set strictly below the space dimension n. This general partial regularity theory including estimates on the Hausdorff dimension of the singular set was developed in a series of papers initiated from [17, 21], and for an overview on the state of the art we refer to the survey paper [28]. In order to obtain regularity of u on larger sets, one needs stronger assumptions, such as some a priori information on the regularity of the solution (for example that the integrability exponent of the solution is coupled to the space dimension) or special structure assumptions (such as vector fields which are linear in the gradient variable), see e. g. [5, 18]. Full regularity result instead are only possible if even more restrictive structural assumptions are imposed. In the nonlinear case, full regularity of Du can for instance still be obtained if the nonlinear part of the vector field depends only on the modulus of z, such as for the p-Laplace system see [38]. An easier (and very classical) example is a linear elliptic system with constant coefficients. Moreover, it is sometimes also sufficient to be sufficiently close to one of these systems, for example in terms of the eigenvalues of $D_z A$.

For simplicity we now focus on quasilinear systems with a vector field of the form $a(x)z$, i. e. to weak solutions of

$$\operatorname{div}\big(a(x)Du\big) = 0 \,, \tag{1.2}$$

with coefficients $a \in L^\infty(D, \mathbb{R}^{Nn \times Nn})$ with ellipticity constant λ_0 and the upper bound λ_1

$$\lambda_0 |\xi|^2 \leq \langle\, a(x)\xi, \xi \,\rangle, \qquad |a(x)\xi| \leq \lambda_1 |\xi| \,. \tag{1.3}$$

One restriction leading to full regularity imposes the ratio λ_0/λ_1 to be sufficiently close to 1, meaning that the different component functions cannot interact too much. This is sometimes referred to as "Cordes-type condition" since Cordes [6] studied quasilinear elliptic equations (that is $N = 1$) and succeeded in showing the availability of a priori Morrey-type estimates under such an assumption. In turn, Schauder fixed point theorems were applied and allowed to deduce the existence of regular solutions. Similar results were obtained later by Koshelev [23] in the vectorial case (here, uniform Morrey-type estimates in a regularization procedure of the original system play the crucial role), which for the toy case of the linear situation allows to identify two different regimes:

Theorem 1.1. *There exists a constant $c_e(n) < 1$ such that the following statement is true:*

(i) *whenever $a \in L^\infty(D, \mathbb{R}^{Nn \times Nn})$ is symmetric and satisfies (1.5) for $0 < \lambda_0 \leq \lambda_1$ with $\lambda_0/\lambda_1 > c_e(n)$, then every weak solution to system (1.2) is of class $C_{\mathrm{loc}}^{0,\alpha}(D \times [0, T], \mathbb{R}^N)$ for some $\alpha > 0$.*

(ii) *given two numbers $0 < \lambda_0 \leq \lambda_1$ with $\lambda_0/\lambda_1 \leq c_e(n)$, there exist symmetric coefficients $a \in L^\infty(D, \mathbb{R}^{Nn \times Nn})$ satisfying (1.5) such that the system (1.2) admits a discontinuous weak solution with regular boundary values on ∂D.*

We highlight that no continuity of the coefficients is assumed in (i) and that (ii) shows the sharpness of the constant c_e (the corresponding counterexamples are obtained by modifications of De Giorgi's famous

counterexample in [9], see [25, Section 2.5]). Furthermore, we note that Koshelev gave the condition $\lambda_0/\lambda_1 > c_e(n)$ explicitly as

$$\frac{\lambda_1 - \lambda_0}{\lambda_1 + \lambda_0} \sqrt{1 + \frac{(n-2)^2}{n-1}} < 1\,.$$

Parabolic systems. The regularity theory for parabolic systems is – to a certain extent – very similar to the elliptic one described above. In the scalar-valued case we again have full regularity under quite general assumptions; for an extension of Moser's iterative scheme to quasilinear parabolic systems we refer to the works of Aronson, Serrin, Ivanov and Kurihara [1, 20, 27]. In the general vectorial case only partial regularity results are available, provided that suitable assumptions on growth and regularity of the vector field $A(x, t, u, z)$ are satisfied. Moreover, an estimate for the Hausdorff dimension of the singular set strictly below the dimension of $\mathbb{R}^n \times [0, T]$ can be found, see [12]. Again, in order to obtain regularity of u on larger sets, restrictions on the space dimension or on the structure of the vector-field $A(x, t, u, z)$ need to be imposed, see e. g. [16, 4, 31] for results in low dimensions or for vector fields which are linear in the gradient variable. Concerning full regularity, we have the extension of the p-Laplacian system to the non-stationary case, which again allows to prove full regularity of the spatial Du, see [11]. Moreover, smoothness of solutions is obtained in the classical case of parabolic systems with constant coefficients. Furthermore, if the system is still sufficiently close to the Laplacian system, then we still get full regularity of u, see [24, 22].

Since the latter regularity result will be of great importance for our paper, we now go into more details and state an extension of Theorem 1.1 to the parabolic case. We will first point out some of the structural prerequisites of the positive (full) regularity theory, and we will then confront it with the existing examples of systems admitting a singular weak solutions. Similarly as in the elliptic case, we again focus on quasilinear problems with a vector field of the form $a(x, t)z$, i. e. to weak solutions of

$$\partial_t u = \operatorname{div}\big(a(x, t)Du\big)\,, \qquad u|_{t=0} = u_0 \tag{1.4}$$

in $D \times (0, T)$, with coefficients $a \in L^\infty(D \times (0, T), \mathbb{R}^{Nn \times Nn})$. Provided that the coupling of the single equations is sufficiently weak, which is ensured by a Cordes-type condition, Koshelev and Kalita [24, 22] observed that discontinuities of the weak solution can globally be excluded.

Theorem 1.2 ([22])**.** *Let $u_0 \in W^{1,q}(D, \mathbb{R}^N)$ for some $q > n$ and consider coefficients $a(x, t)$ which are of class C^1 in x, measurable in t and which satisfy*

$$\lambda_0|\xi|^2 \le \langle\, a(x, t)\xi, \xi\,\rangle, \quad |a(x, t)\xi| \le \lambda_1|\xi|\,, \quad and \quad |D_x a(x, t)| \le L \tag{1.5}$$

for all $\xi \in \mathbb{R}^{nN}$, $(x, t, z) \in D \times [0, T] \times \mathbb{R}^{nN}$ and some positive constants λ_0, λ_1, L. If $\lambda_0/\lambda_1 > 1 - 2/n$ holds, then every weak solution $u\colon D \times [0, T] \to \mathbb{R}^N$ to the initial boundary value problem (1.4) is of class $C^{0,\alpha}_{\mathrm{loc}}(D \times [0, T], \mathbb{R}^N)$ for some $\alpha > 0$.

We start with some comments on the statement of the result. It is important to mention that the original results –which is also true for the statement (i) of Theorem 1.1 – apply to more general systems, which are possibly nonlinear in the gradient variable, provided that the vector field $A(x, t, u, z)$ is sufficiently close to a quasilinear situation with small dispersion ratio. Secondly, we observe that in the parabolic case – in contrast to the elliptic case – an additional regularity with respect to the x-variable is required. It is not clear whether or not this prerequisite is actually needed, or whether it is due to the method of proof. However, before commenting on the strategies of proof, we note that by some perturbation arguments it is possible to obtain local continuity of all solutions, solely under the condition $\lambda_0/\lambda_1 > c_p(n)$, for some $c_p(n) < 1$ (but $c_p(n) > 1 - 2/n$ is still possible). The idea here is to consider $a(x, t)$ as a perturbation of the Laplace system and to carry the a priori regularity results (which are available for such systems) over to the original one, with arguments similar as in Campanato's paper [3]. Since $c_p(n)$ in this case is hardly determined explicitly, we omit the precise statement and prefer to give now a short exposition on the proofs of Theorem 1.2 under the additional regularity assumption on $D_x a$. First, Koshelev proved the existence of a regular solution (which in the situation above is already the unique one, but for nonlinear systems there might be more than one) by studying approximations of the system such that its solutions are regular and converge in a suitable (Morrey-type) norm to a solution of

the original system. This is also the line of arguments for the elliptic result of Theorem 1.1 (i). Kalita achieved later the regularity result for all solutions with a direct argument (and not as a consequence of a suitable approximating sequence). His approach is based on an equation satisfied by second space derivatives of u. Therefore, one first uses finite difference quotients in place of derivatives, a classical method in the study of regularity of weak solutions to elliptic or parabolic equations. Secondly, Moser's iterative method [29] is applied, which means proving higher integrability of a solution (here: the gradient of the solution) by testing with powers of the solution itself. In this way, regularity of weak solutions follows by standard embedding theorems (first slicewise for fixed times, and then by a classical argument the regularity follows locally in $D \times (0, T)$ via the system equation).

Under weaker assumptions than in the previous theorem, such a local regularity result can no longer be expected. In fact, in a very similar setting the following example of a system was proposed by Stará and John [37] (actually, the example was constructed on the full space and the solution can be traced back in time $t \to -\infty$), which admits a solution that starts from a regular – in particular Hölder continuous – initial data and develops a singularity in finite time in the interior of the parabolic cylinder.

Theorem 1.3 ([37]). *Let $n = N \geq 3$. There exist initial data $u_0 \in W^{1,2n}(B_1(0), \mathbb{R}^n)$ and symmetric coefficients $a \in L^\infty(B_1(0) \times [0, 1), \mathbb{R}^{n^2 \times n^2})$, which are elliptic and bounded in the sense of $(1.5)_{1,2}$ for all $(x, t) \in B_1(0) \times [0, 1)$, such that at least one of the solutions to the initial problem (1.4) develops a discontinuity in the origin $x = 0$ as $t \nearrow 1$.*

The coefficients constructed in [37] have a dispersion ratio $\lambda_0/\lambda_1 < 1 - 2/n$ strictly below the critical one investigated in [22], which was an essential ingredient in order to obtain globally Hölder continuous weak solutions. Moreover, the matrix constructed by Stará and John [37] also fails to satisfy the regularity with respect to x, i.e. the matrix A is not differentiable in x. For this reason it is not clear whether the counterexample could by constructed due to the small ellipticity ratio or the low regularity in x or a combination of both. As far as we know, in the literature neither this question is answered nor sharp conditions as in the elliptic case were found, so the occurrence of irregularities for parabolic systems is not yet understood completely. However, due to the positive results obtained by perturbation methods, we believe that the dispersion ratio λ_0/λ_1 plays an important role.

Parabolic systems with random noise.   The aim of this paper is to investigate parts of the parabolic theory presented above under the effect of random perturbations. The final aim of our research project, in analogy with recent results proved for other equations, is to show that the regularity theory of parabolic systems is, under random perturbations, in some sense not worse than the deterministic one (of course only up to a certain degree of regularity), and possibly better. As in the deterministic case there is more than one approach to the analysis of these problems, so we restrict here our attention only to a few directions. More precisely, we discuss in the present paper the extension of Kalita's result (which was displayed above) to the stochastic case. For this purpose we study systems with Itô noise of the form

$$du = \operatorname{div}\big(A(x, t)Du\big) \, dt + H(Du) \, dB_t \,, \qquad u|_{t=0} = u_0 \tag{1.6}$$

(with H Lipschitz), where $(B_t)_{t \geq 0}$ is a Brownian motion of suitable dimension. The passage from deterministic to stochastic of Kalita's approach contains at least one non trivial detail which is rather new in the stochastic setting: the weak solution u we start with is not, a priori, the limit of a sequence of smooth solutions of approximating equations (for instance, due to the nonlinearity, classical mollifiers are difficult to implement; in another direction, in some cases solutions exist as limits of Galerkin or other types of approximations, but we here start with a weak solution which a priori has not been constructed in that way) and thus it is not clear how to perform differential calculus on u. The methods described before in Kalita's approach (existence of higher order derivatives via difference quotient techniques or Moser's iteration techniques) are classical in the deterministic setting, but not common in the stochastic case. This leads to a number of technical novelties, mainly due to the fact that neither do we use any representation formulas for the solution nor can we work pathwise. Instead, we extract from the system information on the mean value of certain (space-time) integral expressions of the solution u or its spatial gradient Du as a process. This is accomplished by difference quotient techniques carried over to the stochastic setting, see Section 3.2 for the general setup. Moreover, the Moser iteration technique needs to be implemented in the stochastic case (via the application of Itô's formula in Banach spaces); we here

note that a stochastic version of Moser's iterative scheme and of a maximum principle was obtained in [10], but in contrast to our paper only the case of equations ($N = 1$) was studied, less general coefficients were investigated, and all the arguments were given on level of the solution, and not its derivatives. In our paper instead, it is ensured that the process u belongs P-almost surely to a suitable Sobolev space, and a pathwise regularity result then follows by embedding theorem for fixed times. Finally, with arguments similar as in the deterministic case, space-time regularity is obtained by exploiting the system equation, see Section 3.3. These technical aspects for stochastic partial differential equations might be of interest on their own even though they are not surprising. At the end we reach a full extension of Kalita result to a quite general stochastic case, which includes in particular perturbations in form of additive or of multiplicative noise. In the quasi-linear model case we obtain – as a particular case of the general Theorem 5.1 – the following result (the precise definition of weak solution is given in Definition 2.2 below).

Theorem 1.4. *Let $u_0 \in W^{1,q}(D, \mathbb{R}^N)$ for some $q > n$. Consider coefficients $a(x,t)$ which are of class C^1 in x, measurable in t and which satisfy (1.5) with $\lambda_0/\lambda_1 > 1 - 2/n$, and assume that H is Lipschitz continuous with Lipschitz constant $L_H < L_H^*$ for some sufficiently small L_H^* depending only on n, λ_0 and λ_1. Then there exists $\alpha > 0$ depending only on n, λ_0, λ_1 and q such that every weak solution $u \colon D \times [0,T] \times \Omega \to \mathbb{R}^N$ to the initial boundary value problem (1.6) is of class $C^{0,\alpha}_{\mathrm{loc}}(D \times [0,T], \mathbb{R}^N)$ with probability 1.*

We conclude this introductory part with some open questions and possible extensions. Concerning the existence of weak solutions, we could give a quite general result, but since it is related to the generalization of Koshelev's approach of [24, 25] to regularity, we did not pursue this research direction. Concerning other strategies for proving regularity we note that it might be interesting to generalize also the classical strategy via Campanato-type estimates to the stochastic case. However, we decided to concentrate here only on some selected aspects of regularity methods and postpone further development of classical techniques known from the deterministic setting to future works.

Finally, we mention an interesting open problem in a related research direction which might contribute to the understanding of the irregularity phenomenon for vectorial parabolic systems. So far, the present paper makes a contribution in only one direction, namely that regularity of weak solutions does not get worse under random perturbation. A natural question is the possibility that noise might actually prevent the emergence of singularities (for example by destroying coherent structures on which the construction of counterexamples is usually based). It has been recently proved that a Stratonovich bilinear multiplicative noise may have a regularizing effect on certain classes of PDEs, see [13] for a review, based on a number of works including [14, 15, 2]. In most cases, uniqueness by noise is the topic of these works. The problem of the interaction between noise and singularities is more difficult and less explored, but positive results from the linear transport equations and the point vortex motion associated to the 2D Euler equations were discovered. From these results it seems that each equation requires its own understanding, and no general method exists to investigate these kind of properties. In the context of parabolic systems one might hope to prove that, under assumptions on the vector field $A(x,t,u,z)$ such that weak solutions with singularities in the deterministic case might exist, there are no more singularities if a suitable noise is added. However, the answer to this question seems to be rather challenging, but we believe that it would be interesting also from the deterministic point of view.

2 Setting and assumptions

Consider $n, n' \in \mathbb{N}$ with $n \geq 2$, $T > 0$, and $D \subset \mathbb{R}^n$ a (regular) bounded domain. Let $(\Omega, \mathcal{F}, P)$ be a complete probability space with filtration $(\mathcal{F}_t)_{t \geq 0}$, and let $(B_t)_{t \geq 0}$ be a standard n'-dimensional Brownian motion. Let further $A \colon D \times [0,T] \times \mathbb{R}^N \times \mathbb{R}^{nN} \times \Omega \to \mathbb{R}^{nN}$ be a vector field satisfying the following properties:

- A is progressively measurable, i.e. for every $t \in [0,T]$ the restriction of A to $D \times [0,t] \times \mathbb{R}^N \times \mathbb{R}^{nN} \times \Omega \to \mathbb{R}^{nN}$ is $\mathcal{B}(D) \times \mathcal{B}([0,t]) \times \mathcal{B}(\mathbb{R}^N) \times \mathcal{B}(\mathbb{R}^{nN}) \times \mathcal{F}_t$ measurable;

- $A(x,t,u,z,\omega)$ (usually abbreviated by $A(x,t,u,z)$) is differentiable in x, u and z (with $\mathcal{F}_t$-adapted derivatives), and it satisfies for P-almost all $\omega \in \Omega$ the following assumptions concerning growth

and ellipticity:

$$\begin{cases} |A(x,t,u,z)| \le L\big(|z| + |u|^{\frac{n+2}{n}} + f^{\frac{a}{2}}(x,t)\big) \\ |\xi - \kappa D_z A(x,t,u,z)\xi|^2 \le (1-\nu^2)|\xi|^2 \\ |D_u A(x,t,u,z)| \le L\big(|z|^{\frac{2}{n+2}} + |u|^{\frac{2}{n}} + f(x,t)\big) \\ |D_x A(x,t,u,z)| \le L\big(|z| + |u|^{\frac{n+2}{n}} + f^2(x,t)\big) \end{cases} \tag{2.1}$$

for all $(x,t) \in D \times [0,T]$, $u \in \mathbb{R}^N$ and $z,\xi \in \mathbb{R}^{nN}$, some constants $\kappa,\nu,L > 0$, and an $\mathcal{F}_t$-adapted process f which with probability one belongs to $L^a(D \times [0,T])$ for a fixed number $a > n+2$.

Moreover, let $H : D \times [0,T] \times \mathbb{R}^{nN} \times \Omega \to \mathbb{R}^{n'N}$ be progressively measurable, of class C^1 in x, Lipschitz with respect to the gradient variable of at most linear growth, uniformly in (x,t), i.e.

$$\begin{cases} |H(x,t,z,\omega) - H(x,t,\widetilde{z},\omega)| \le L_H|z - \widetilde{z}|, \\ |H(x,t,z,\omega)| \le L\big(f_H(x,t,\omega) + |z|\big), \\ |D_x H(x,t,z,\omega)| \le L\big(f_H^{\frac{a}{a-2}}(x,t,\omega) + |z|\big) \end{cases} \tag{2.2}$$

for a constant L_H, all $(x,t) \in D \times [0,T]$, $z,\widetilde{z} \in \mathbb{R}^{nN}$, and almost every $\omega \in \Omega$. Here, f_H denotes another function in $L^a(D \times (0,T) \times \Omega)$.

Under these assumptions we consider a stochastic partial differential equation with noise of the form

$$du = \operatorname{div} A(x,t,u,Du)\, dt + H(x,t,Du)\, dB_t \qquad \text{in } D_T := D \times (0,T), \tag{2.3}$$

where $u \colon D_T \times \Omega \to \mathbb{R}^N$ is a random function. The stochastic integral is here understood in the Itô sense. According to the growth condition on the vector field A, we note that for P-almost every $\omega \in \Omega$ and all $t \in [0,T]$ we have $\operatorname{div} A(x,t,v,Dv) \in W^{-1,2}(D,\mathbb{R}^N)$ – the dual space to $W_0^{1,2}(D,\mathbb{R}^N)$ –, provided that $v \in W^{1,2}(D,\mathbb{R}^N)$.

Remark 2.1. *We have chosen this level of generality of the noise for several reasons: to keep a simple PDE structure instead of an abstract operator formulation, and to cover two interesting examples: additive noise (with $H(x,z)$ independent of z) and bilinear multiplicative noise with first derivatives of u (with $H(x,z)$ linear in z). A priori there is no conceptual obstacle to consider H depending also on u or to generalize to the case of a Brownian motion B in a Hilbert space U, with suitable assumptions on H, but for simplicity we restrict ourselves to the previous case.*

The function spaces that will be needed in the sequel are the Banach spaces

$$V^{m,p}(D_T,\mathbb{R}^N) := L^\infty\big(0,T; L^m(D,\mathbb{R}^N)\big) \cap L^p\big(0,T; W^{1,p}(D,\mathbb{R}^N)\big),$$
$$V_0^{m,p}(D_T,\mathbb{R}^N) := L^\infty\big(0,T; L^m(D,\mathbb{R}^N)\big) \cap L^p\big(0,T; W_0^{1,p}(D,\mathbb{R}^N)\big),$$

with $m,p \ge 1$, and they are equipped with the norm

$$\|u\|_{V^{m,p}(D_T,\mathbb{R}^N)} := \operatorname*{ess\,sup}_{t \in (0,T)} \|u(t)\|_{L^m(D,\mathbb{R}^N)} + \|Du\|_{L^p(D_T,\mathbb{R}^N)}.$$

When $m = p$ we shall use the abbreviations $V_{(0)}^p(D_T,\mathbb{R}^N) = V_{(0)}^{p,p}(D_T,\mathbb{R}^N)$. We remind that the spaces $V^{m,p}(D_T,\mathbb{R}^N)$ are embedded in the Lebesgue space $L^q(D_T,\mathbb{R}^N)$ with $q = p(n+m)/n > p$ (see [11, Propositions I.3.1, I.3.2]). We will need only the result concerning the cases $p \ge m = 2$ or $p = m \ge 2$. In the latter case, the embedding reads as follows (see [11, Propositions I.3.3, I.3.4]): let $v \in V_0^p(D_T,\mathbb{R}^N)$, $p < n$. Then there exists a constant c depending only on n and p such that

$$\|v\|_{L^q(D_T,\mathbb{R}^N)} \le c\|v\|_{V^p(D_T,\mathbb{R}^N)} \tag{2.4}$$

(and an analogous result holds without any restriction on the boundary values of v on $\partial D \times (0,T)$ if ∂D is assumed to be sufficiently regular).

We are now going to study the properties of weak (or variational) solutions to the system (2.3), which are to be understood in the following sense.

Definition 2.2. *An $\mathcal{F}_t$-progressively measurable process u on $[0,T] \times \Omega$ is called* weak solution *to the system* (2.3) *with initial values $u_0 \in L^2(D, \mathbb{R}^N)$ if P-a. e. path satisfies $u(\cdot, \omega) \in V^2(D_T, \mathbb{R}^N)$ and if for all $t \in [0,T]$, we have P-a. s. the identity*

$$\langle u(t) - u_0, \varphi \rangle_{L^2(D)} = \int_0^t \langle \operatorname{div} A(\cdot, s, u, Du), \varphi \rangle_{W^{-1,2}(D); W_0^{1,2}(D)} \, ds + \int_0^t \langle \varphi, H(\cdot, s, Du) \, dB_s \rangle_{L^2(D)}$$

for all $\varphi \in W_0^{1,2}(D, \mathbb{R}^N)$.

When a solution is progressively measurable with respect to the (completed) filtration associated to the Brownian motion, it is usually called a "strong" solution in the probabilistic sense, see [34, Section IX.1]. We do not require this condition, so our result will also apply to the so called "weak" solutions in the probabilistic sense (those for which there is a filtration $(\mathcal{F}_t)_{t \geq 0}$ such that u is $\mathcal{F}_t$-progressively measurable and B is an $\mathcal{F}_t$-Brownian motion). We further note that according to the definition above, a solution is defined as an equivalence class in the sense of versions (a process Y is a version or modification of a process X if for each time t we have P-a. s. $X_t = Y_t$). Hence, regularity of a weak solution is always to be understood as finding a regular representative in the corresponding equivalence class.

Moreover, we comment on the way in which the initial values are attained. Under mild assumptions on the growth of A and H with respect to the gradient variable one actually deduces from the equation itself that u belongs to $C^0(0, T; L^2(D', \mathbb{R}^N))$ P-a. s. for every $D' \Subset D$, compare the classical parabolic theory or formula (4.1) and the beginning of Step 3 on p. 16. Under further assumptions on the trace of u on $\partial D \times [0,T]$ this extends to continuity of the full L^2-norm, with $D' = D$. In this sense the term "initial value" in the definition of a weak solution as a function in the space V^2 is justified.

3 Preliminaries

In this section we recall some well-known facts and provide several technical tools. For convenience of the reader we state two suitable versions of Itô's formula. Furthermore, in analogy with the deterministic theory, we discuss a sufficient condition for the "existence of weak derivatives with probability one", and we further give a criterion which guarantees pathwise Hölder continuity of a process.

3.1 Itô formula

We first recall two versions of Itô's formula, the first one the standard version for N-dimensional processes and the second one for processes with values in Hilbert spaces. Consider (Ω, F, P) a complete probability space and let

$$dX(t) = a(t) \, dt + b(t) \, dB_t \tag{3.1}$$

be an N-dimensional Itô process which satisfies: a, b are $\mathcal{F}_t$-adapted (i. e., the maps $\omega \mapsto a(t, \omega), b(t, \omega)$ are $\mathcal{F}_t$ measurable), $(t, \omega) \mapsto b(t, \omega)$ is $\mathcal{B}([0,T]) \times \mathcal{F}$-measurable and

$$P\left(\int_0^T \left[|a(s, \omega)| + |b(s, \omega)|^2 \right] ds < \infty \right) = 1.$$

Then the following general Itô formula holds (see e. g. [32, Theorem 4.2.1]).

Theorem 3.1 (Itô's formula I). *Let $N' \in \mathbb{N}$. Let $g(t, z) = (g_1(t, z), \ldots, g_{N'}(t, z))$ be a map from $[0, T] \times \mathbb{R}^N$ to $\mathbb{R}^{N'}$ of class C^1 in t and of class C^2 in z. Then the process $Y(t, \omega) := g(t, X(t))$ with $X(t)$ defined in* (3.1) *is again an Itô process whose components are given by*

$$dY_k(t) = \frac{\partial g_k}{\partial t}(t, X) \, dt + \sum_{i=1}^N \frac{\partial g_k}{\partial y_i}(t, X) \, dX_i + \frac{1}{2} \sum_{i,j=1}^N \frac{\partial^2 g_k}{\partial y_i y_j}(t, X) \, d[X_i, X_j]_t$$

for all $k \in \{1, \ldots, N'\}$ and with $[X_i, X_j]_t$ the quadratic covariation of the processes X_i and X_j, with $d[B_i, B_j]_t = \delta_{ij} \, dt$ for all $i, j \in \{1, \ldots, N\}$.

In the sequel, we will also employ the following version of the Itô formula in Hilbert spaces that can be found in [26, Theorem 3.1] or [35, Chapter 1.4.2, Theorem 2].

Theorem 3.2 (Itô's formula II)**.** *Let $V \subset H \subset V'$ be a Gelfand triple, with H a separable Hilbert space. Assume that we have for $\mathcal{L}^1 \times P$ almost all $(t, \omega) \in [0, T] \times \Omega$*

$$\langle\, x(t), \varphi \,\rangle_H = \langle\, x(0), \varphi \,\rangle_H + \int_0^t \langle\, y(s), \varphi \,\rangle_{V', V}\, ds + \langle\, M_t, \varphi \,\rangle_H \tag{3.2}$$

for every $\varphi \in V$ where $x(t, \omega), y(t, \omega)$ are taking values in V and V', respectively, and are progressively measurable with

$$P\Big(\int_0^T \big[\|x(s, \omega)\|_V^2 + \|y(s, \omega)\|_{V'}^2 \big]\, ds < \infty \Big) = 1,$$

and where M_t is a continuous local martingale with values in H. Then there exists a set $\widetilde{\Omega} \subset \Omega$ with $P(\widetilde{\Omega}) = 1$ and a map $\widetilde{x}(t, \omega)$ with values in H such that:

(i) *$\widetilde{x}(t)$ is $\mathcal{F}_t$-adapted, continuous in $t \in [0, T]$ for every $\omega \in \widetilde{\Omega}$, and $x(t) = \widetilde{x}(t)$ P-almost surely;*

(ii) *for every $\omega \in \widetilde{\Omega}$, $t \in [0, T]$ and all $\varphi \in V$ there holds*

$$\langle\, \widetilde{x}(t), \varphi \,\rangle_H = \langle\, x(0), \varphi \,\rangle_H + \int_0^t \langle\, y(s), \varphi \,\rangle_{V', V}\, ds + \langle\, M_t, \varphi \,\rangle_H\,;$$

(iii) *for every $\omega \in \widetilde{\Omega}$, $t \in [0, T]$ there holds the equality*

$$\|\tilde{x}(t)\|_H^2 = \|x(0)\|_H^2 + 2 \int_0^t \langle\, y(s), x(s) \,\rangle_{V', V}\, ds + 2 \int_0^t \langle\, dM_s, \tilde{x}(s) \,\rangle_H + [M]_t$$

with $[M]_t = [M, M]_t$ denoting the quadratic variation of M.

3.2   Weak derivatives

For a vector-valued function $f \colon \mathbb{R}^n \supset D \to \mathbb{R}^N$, $k \in \{1, \dots, n\}$ and a real number $h \in \mathbb{R} \setminus \{0\}$ we denote by $\triangle_{k,h} f(x) := h^{-1}(f(x + he_k) - f(x))$ the finite different quotient in direction e_k and stepsize h (this makes sense as long as x, $x + he_k \in D$). Let $p > 1$, $f \in L^p(D)$, $k \in \{1, \dots, n\}$ and let $D_k f$ be the derivative of f in the direction k in the sense of distributions. Just for comparison let us recall the following lemma (not used below).

Lemma 3.3. *If there is $h_n \to 0$ and $g_k \in L^p(D)$ such that*

$$\lim_{n \to \infty} \int_D \big(\triangle_{k, h_n} f(x) - g_k(x) \big) \varphi(x)\, dx = 0$$

for every $\varphi \in C_0^\infty(D)$, then $D_k f$ is in $L^p(D)$ and is equal to g_k.

As an immediate consequence of this lemma and of the compactness of the L^p-spaces with $p > 1$ with respect to weak (or weak-$*$) convergence, we obtain a simple criterion for the existence of the weak derivative $D_k f$ in L^p, namely it is sufficient that $\|\triangle_{k,h} f\|_{L^p(D')}$ is bounded for every $D' \Subset D$ by some constant $C_{D'}$, uniformly for all h such that $|h| < \mathrm{dist}(D', \partial D)$.

Now this well-known principle shall be carried over to a probabilistic setting. Let (Ω, F, P) be a complete probability space and consider a function f in the Banach space $L^p(D \times \Omega)$. A function $g_k \in L^p(D \times \Omega)$ is said to be the weak derivative of f in the k-direction if

$$P\Big(\int_D f D_k \varphi\, dx = - \int_D g_k \varphi\, dx \Big) = 1$$

for every $\varphi \in C_0^\infty(D)$ (taking a countable sequence and using a density argument, the property "for every $\varphi \in C_0^\infty(D)$" can be written inside the probability). We then write $D_k f = g_k$. The previous lemma has a generalization to functions in $L^p(D \times \Omega)$.

Lemma 3.4. *If there is $h_n \to 0$ and $g_k \in L^p(D \times \Omega)$ such that*

$$\lim_{n \to \infty} \int \int_{D \times \Omega} \big(\triangle_{k,h_n} f(x,\omega) - g_k(x,\omega)\big) \varphi(x) X(\omega) \, dx \, dP(\omega) = 0$$

for every $\varphi \in C_0^\infty(D)$ and every bounded measurable $X : \Omega \to \mathbb{R}$, then $D_k f$ is in $L^p(D \times \Omega)$ and is equal to g_k.

Proof. Since X and φ are bounded, we may apply (first Fubini and then) Lebesgue's dominated convergence theorem, and we get

$$-E\Big[X \int_D \big(f D_k \varphi + g_k \varphi\big) dx\Big] = -\int \int_{D \times \Omega} X \big(f D_k \varphi + g_k \varphi\big) \, dx \, dP$$

$$= \lim_{n \to \infty} \int \int_{D \times \Omega} X \big(-f(x) \triangle_{k,-h_n} \varphi(x) - g_k(x) \varphi(x)\big) \, dx \, dP.$$

When $h_n < \mathrm{dist}(\mathrm{spt}\,\varphi, \partial D)$, this is equal to (we apply Fubini twice and a change of variables)

$$\lim_{n \to \infty} \int \int_{D \times \Omega} X \big(\triangle_{k,h_n} f(x,\omega) \varphi(x) - g_k(x,\omega) \varphi(x)\big) \, dx \, dP.$$

This limit is zero by assumption, hence

$$E\Big[X \int_D \big(f D_k \varphi + g_k \varphi\big) dx\Big] = 0.$$

The arbitrariness of X implies $\int_D \big(f D_k \varphi + g_k \varphi\big) dx = 0$, as a random variable on Ω. The proof is complete. $\qquad \square$

Corollary 3.5. *If there is a constant $C > 0$ such that*

$$E\Big[\int_{D'} |\triangle_{k,h} f(x)|^p \, dx\Big] \leq C$$

for all h and all $D' \Subset D$ such that $|h| < \mathrm{dist}(D', \partial D)$, then $D_k f$ is in $L^p(D \times \Omega)$.

Proof. The family $g_{k,h}(x,\omega) := \triangle_{k,h} f(x,\omega)$ is equibounded in $L^p(D' \times \Omega)$, hence there is a sequence $h_n \to 0$ such that g_{k,h_n} converges weakly in $L^p(D \times \Omega)$ to some function $g_k \in L^p(D \times \Omega)$. The product $\varphi(x) X(\omega)$ is in $L^{p'}(D \times \Omega)$ (with p' conjugate to p) for every $\varphi \in C_0^\infty(D)$ and every bounded measurable $X : \Omega \to \mathbb{R}$. Hence, we may apply the lemma and obtain the assertion. $\qquad \square$

First, for our later application, we replace D by $D \times [0,T]$ and we allow different integrability exponents with respect to the variables in $[0,T]$ and D, respectively. Let $f : \Omega \to L^q(0,T; L^p(D))$ be a measurable function with $p \in (1, \infty)$ and $q > 1$. We say that a function $g_k : \Omega \to L^q(0,T; L^p(D))$ is weak derivative of f in the k-direction with probability one if for a.e. $(t,\omega) \in [0,T] \times \Omega$ we have

$$\int_D f D_k \varphi \, dx = -\int_D g_k \varphi \, dx$$

for every $\varphi \in C_0^\infty(D)$, and we then write $D_k f = g_k$. Furthermore, let us generalize to a scheme where we relax the integrability in Ω.

Theorem 3.6. *Let $Y : [0,T] \times \Omega \to (0,1]$ be a positive random variable, with $P(\inf_{t \in [0,T]} Y > 0) = 1$. If there is a constant $C > 0$ such that*

$$E\Big[\big\|Y(t) \triangle_{k,h_n} f(x,t)\big\|_{L^q(0,T; L^p(D'))}^p\Big] \leq C$$

for all h and $D' \Subset D$ satisfying $|h| < \mathrm{dist}(D', \partial D)$, then $D_k f \in L^q(0,T; L^p(D))$ with probability one and there hold

$$Y \triangle_{k,h} f \to Y D_k f \quad \text{weakly in } L^p(\Omega; L^q(0,T; L^p(D))),$$
$$E\big[\big\|Y D_k f\big\|_{L^q(0,T; L^p(D))}^p\big] \leq C$$

with the same constant C.

Proof. The family $Z_{k,h}(x,t,\omega) := Y(t)\triangle_{k,h_n} f(x,t,\omega)$ is equibounded in $L^p(\Omega; L^q(0,T; L^p(D')))$, hence there is a sequence $h_n \to 0$ such that Z_{k,h_n} converges weakly in $L^p(\Omega; L^q(0,T; L^p(D)))$ (or weakly-$*$ if $q = \infty$) to some function $Z_k \in L^p(\Omega; L^q(0,T; L^p(D)))$. This implies (again with ψ, X bounded, measurable and φ smooth, compactly supported)

$$\lim_{n\to\infty} \int \int \int_{D\times[0,T]\times\Omega} \big(Y(t)\triangle_{k,h_n} f(x,t) - Z_k(x,t)\big)\varphi(x)\psi(t)X\, dx\, dt\, dP = 0\,.$$

Hence, by Fubini and change of variables as above, we find

$$\lim_{n\to\infty} \int \int \int_{D\times[0,T]\times\Omega} \big(Y(t)f(x,t)\triangle_{k,-h_n}\varphi(x) + Z_k(x,t)\varphi(x)\big)\psi(t)X\, dx\, dt\, dP = 0\,,$$

which in turn implies by Lebesgue's theorem

$$\int \int \int_{D\times[0,T]\times\Omega} \big(Y(t)f(x,t)D_k\varphi(x) + Z_k(x,t)\varphi(x)\big)\psi(t)X\, dx\, dt\, dP = 0\,.$$

Arbitrariness of X and ψ thus yields

$$\int_D \big(Y(t)f(x,t)D_k\varphi(x) + Z_k(x,t)\varphi(x)\big)\, dx = 0$$

for a. e. $(t,\omega) \in [0,T] \times \Omega$. Therefore, we have

$$\int_D \big(f(x,t)D_k\varphi(x) + g_k(x,t)\varphi(x)\big)\, dx = 0 \tag{3.3}$$

for a. e. $(t,\omega) \in [0,T] \times \Omega$, where $g_k = Y^{-1}Z_k$. Since Z_k belongs to $L^p(\Omega; L^q(0,T; L^p(D)))$, it is $L^q(0,T; L^p(D))$ for P-a. e. $\omega \in \Omega$. Hence, by assumption on Y, we also have $g_k \in L^q(0,T; L^p(D))$ for P-a. e. $\omega \in \Omega$. The only difference with the definition of g_k being the "weak derivative of f in the k-direction with probability one" is that the negligible set of $(t,\omega) \in [0,T] \times \Omega$ where (3.3) may fail depends on $\varphi \in C_0^\infty(D)$, until now. But $W^{1,p'}(D)$ (with p' conjugate to p) is separable and $C_0^\infty(D)$ is dense in it. Hence, there is a countable family $\{\varphi_n\} \subset C_0^\infty(D)$ which is dense in $W^{1,p'}(D)$. If we call N the countable union of all negligible sets of $(t,\omega) \in [0,T] \times \Omega$ where (3.3) may fail for $\{\varphi_n\}$, N is negligible, and on the complement we have (3.3) for every φ_n, hence by density for all $\varphi \in W^{1,p'}(D)$ and then for all $\varphi \in C_0^\infty(D)$. Having identified g_k as the weak derivative of f in the k-direction we take advantage of the lower semi-continuity of the norm with respect to weak (or weak-$*$) convergence and thus we find

$$E\big[\|YD_k f\|_{L^q(0,T;L^p(D))}^p\big] = E\big[\|Z_k\|_{L^q(0,T;L^p(D))}\big] \le C.$$

The proof is complete. $\square$

Remark 3.7. *This result will be applied later in the cases $p = q$ where the assumption then reads as*

$$E\Big[\int_0^T \int_D |Y(t)\triangle_{k,h} f_1(x,t)|^p\, dx\, dt\Big] \le C$$

and where we have $L^p(\Omega; L^q(0,T; L^p(D))) = L^p(D \times [0,T] \times \Omega)$, or in the case $q = \infty$ where we then require

$$E\Big[\sup_{t\in(0,T)} \int_D |Y(t)\triangle_{k,h} f_2(x,t)|^p\, dx\Big] \le C\,.$$

From the theorem we then conclude that $D_k f_1 \in L^p(D \times [0,T])$ and $D_k f_2 \in L^\infty(0,T; L^p(D))$ with probability one, respectively. In particular, if we take a function $f \in W^{1,p}(D)$ and if the previous assumptions are satisfied for $f_1 = Df$ and $f_2 = f$, then the conclusions are equivalent to $D_k f \in V^p(D_T)$.

3.3   A criterion for pathwise Hölder continuity

We next discuss a criterion which guarantees Hölder continuity of (a suitable representative of) a given functions $u\colon D \times [0,T] \to \mathbb{R}^N$. For example, Sobolev's embedding theorem provides a criterion which is easy to apply, namely that u belongs to a suitable Sobolev space $W^{1,q}(D \times [0,T], \mathbb{R}^N)$ – however, this is in general not satisfied for the solutions considered in our paper since derivatives in time need not exist. Instead, we now prove that it is sufficient to have only the spatial derivatives in a suitable Lebesgue space, provided that a weak form of continuity in time (here of the $L^2(D)$-norm) is available.

Lemma 3.8. *If a function $u\colon D \times [0,T] \to \mathbb{R}^N$ has the properties*

$$Du \in L^\infty(0,T; L^{n+\alpha}(D, \mathbb{R}^{nN})), \qquad u \in C^\beta(0,T; L^2(D, \mathbb{R}^N))$$

for some $\alpha, \beta > 0$, $D \subset \mathbb{R}^n$ a bounded, regular domain, then

$$u \in C^\gamma(D \times [0,T], \mathbb{R}^N)$$

for some $\gamma > 0$, depending only on α, β and n.

Proof. First, we deduce spatial Hölder continuity for every time slice. From the assumption $Du \in L^\infty(0,T; L^{n+\alpha}(D))$ we deduce $u \in L^\infty(0,T; C^\delta(D))$ for some $\delta > 0$, depending only on α and n, by Sobolev's embedding theorem; namely, there exists $C_1 > 0$ such that

$$|u(x,t) - u(y,t)| \le C_1 |x-y|^\delta \tag{3.4}$$

for all $t \in [0,T]$, $x, y \in D$.

Our next aim is Hölder continuity in time, at a fixed point. From the inequality

$$\|u(\cdot,t) - u(\cdot,s)\|_{L^2(D)} \le C_2 |t-s|^\beta$$

for $s, t \in [0,T]$, we infer for every set $B \subset D$

$$\inf_{x \in B} |u(x,t) - u(x,s)| \le \frac{1}{|B|} \int_B |u(x,t) - u(x,s)|\, dx \le \frac{1}{|B|^{1/2}} \|u(\cdot,t) - u(\cdot,s)\|_{L^2(D)} \le \frac{C_2 |t-s|^\beta}{|B|^{1/2}}.$$

Let $x_0 \in D$ be given. In order to prove Hölder continuity in time at x_0, we estimate

$$|u(x_0,t) - u(x_0,s)| \le |u(x_0,t) - u(x,t)| + |u(x,t) - u(x,s)| + |u(x,s) - u(x_0,s)|$$
$$\le 2C_1 |x-x_0|^\delta + |u(x,t) - u(x,s)|$$

for every $x \in D$. Hence, if we take x in a ball $B(x_0,\rho)$, we have

$$|u(x_0,t) - u(x_0,s)| \le 2C_1 \rho^\delta + \inf_{x \in B(x_0,\rho)} |u(x,t) - u(x,s)|$$
$$\le 2C_1 \rho^\delta + C_3 \frac{C_2 |t-s|^\beta}{\rho^{n/2}}$$

where C_3 is such that $|B(x_0,\rho)|\rho^n / C_3^2$. Let us now choose $\rho = |t-s|^\varepsilon$ for some $\varepsilon > 0$:

$$|u(x_0,t) - u(x_0,s)| \le 2C_1 |t-s|^{\varepsilon \delta} + C_3 C_2 |t-s|^{\beta - \varepsilon n/2}.$$

If we choose for instance $\varepsilon = \beta/n$, we get

$$|u(x_0,t) - u(x_0,s)| \le C_4 |t-s|^\eta \tag{3.5}$$

for some $\eta, C_4 > 0$, independently of $x_0 \in D$, $t, s \in [0,T]$. The exponent η depends only on β, δ and n.

From (3.4) and (3.5) it is now straightforward to deduce the claim of the lemma. $\qquad\square$

With the previous lemma at hand, we now give a continuity criterion which is adapted to weak solutions in the probabilistic setting, with (Ω, F, P) a complete probability space.

Proposition 3.9. *Let $u : D \times [0,T] \times \Omega \to \mathbb{R}^{nN}$ have the properties*

$$P\big(Du \in L^\infty(0,T; L^{n+\varepsilon}(D, \mathbb{R}^{nN}))\big) = 1, \qquad\qquad (3.6)$$

$$u(x,t) = u_0(x) + \int_0^t a(x,s)\, ds + \int_0^t b(x,s)\, dB_s,$$

for some $\varepsilon > 0$, $u_0 \in L^2(D)$, and with progressively measurable fields a, b such that

$$P\Big(\int_0^T \int_D |a(x,s)|^2\, dx\, ds + \int_0^T \Big(\int_D |b(x,s)|^2\, dx \Big)^{\frac{2+\varepsilon}{2}} ds < \infty \Big) = 1.$$

Then

$$P\big(u \in C^\gamma(D \times [0,T])\big) = 1$$

for some $\gamma > 0$ depending only on ε.

Proof. **Step 1**. If we prove that, for some $\beta > 0$,

$$P\big(u \in C^\beta(0,T; L^2(D))\big) = 1,$$

then we get the claim of the proposition after the pathwise application of the previous Lemma 3.8 (using in particular the stated independence of the Hölder exponent). To this end we observe that the function u is the sum of two terms:

$$u_1(x,t) = u_0(x) + \int_0^t a(x,s)\, ds, \qquad u_2(x,t) = \int_0^t b(x,s)\, dB_s$$

The term u_1 is, with probability one, of class $W^{1,2}(0,T; L^2(D))$, hence it is of class $C^{1/2}(0,T; L^2(D))$:

$$\|u_1(t) - u_1(s)\|_{L^2(D)} = \Big\| \int_s^t a(\cdot, r)\, dr \Big\|_{L^2(D)} \leq |t-s|^{1/2} \Big(\int_0^T \int_D |a(x,r)|^2\, dx\, dr \Big)^{1/2}.$$

So it only remains to prove that, for some $\beta > 0$,

$$P\big(u_2 \in C^\beta(0,T; L^2(D))\big) = 1.$$

Step 2. For $R > 0$, let

$$\tau_R = \inf\Big\{ t \in (0,T) \colon \int_0^t \|b(\cdot,s)\|_{L^2(D)}^{2+\varepsilon}\, ds > R \Big\}$$

if the set is non empty, otherwise $\tau_R = T$. Let $\Omega_R \subset \Omega$ be the set where $\tau_R = T$. The family $\{\Omega_R\}_{R>0}$ is increasing, with

$$P\Big(\bigcup_{R>0} \Omega_R \Big) = 1$$

because by assumption we have $P\big(\int_0^T \|b(\cdot,s)\|_{L^2(D)}^{2+\varepsilon}\, ds < \infty \big) = 1$. We now set

$$b_R(x,s) = b(x,s) 1_{s \leq \tau_R} \qquad \text{and} \qquad u_{2,R}(t) = \int_0^t b_R(x,s)\, dB_s = \int_0^{t \wedge \tau_R} b(x,s)\, dB_s.$$

We then have

$$\int_0^T \|b_R(\cdot,s,\omega)\|_{L^2(D)}^{2+\varepsilon}\, ds \leq R$$

uniformly in ω. Hence, for every $p \geq 1$, we find

$$E\Big[\|u_{2,R}(t) - u_{2,R}(s)\|_{L^2(D)}^p \Big] = E\Big[\Big\| \int_s^t b_R(\cdot,r)\, dB_r \Big\|_{L^2(D)}^p \Big]$$

$$\leq C_p E\Big[\Big(\int_s^t \|b_R(\cdot,r)\|_{L^2(D)}^2\, dr \Big)^{\frac{p}{2}} \Big]$$

$$\leq C_p |t-s|^{\frac{p\varepsilon}{2(2+\varepsilon)}} E\Big[\Big(\int_s^t \|b_R(\cdot,r)\|_{L^2(D)}^{2+\varepsilon}\, dr \Big)^{\frac{p}{2+\varepsilon}} \Big] \leq C_p R^{\frac{p}{2+\varepsilon}} |t-s|^{\frac{p\varepsilon}{2(2+\varepsilon)}}.$$

This implies, for $p = p(\varepsilon)$ sufficiently large, by Kolmogorov's regularity theorem for processes taking values in $L^2(D)$ (see [7, Theorem 3.3] for a version in Banach spaces), that $u_{2,R}$ has a Hölder continuous version in $L^2(D)$

$$\|u_{2,R}(\cdot, t, \omega) - u_{2,R}(\cdot, s, \omega)\|_{L^2(D)} \le C_{\beta,R}(\omega)|t - s|^\beta$$

with β any Hölder exponent with $\beta < \frac{\varepsilon}{2(2+\varepsilon)}$. For $\omega \in \Omega_R$ we thus have (recalling the definition of $u_{2,R}$)

$$\|u_2(\cdot, t, \omega) - u_2(\cdot, s, \omega)\|_{L^2(D)} \le C_{\beta,R}(\omega)|t - s|^\beta.$$

Since $\bigcup_{R>0} \Omega_R$ is of full P-measure, we obtain $u_2 \in C^\beta(0, T; L^2(D))$ for P-a.e. $\omega \in \Omega$. Now the previous Lemma 3.8 can be applied, and the proof is complete. $\qquad\square$

3.4 A technical lemma

In Kalita's paper a crucial point is to show higher regularity (such as higher integrability and differentiability) not only for the solution, but also for powers of the solution (resp. its gradient). For this purpose the following technical lemma was essential.

Lemma 3.10 ([22])**.** *Let* $u\colon \mathbb{R}^n \to \mathbb{R}^N$ *be a function which is a.e. differentiable. Set* $v = u|u|^s$ *with* $s \in (-1, \infty)$*. Then, for* $\mu(s) := 1 - (\frac{s}{2+s})^2$*, we have a.e.*

$$Du \cdot Dv \ge \mu^{\frac{1}{2}}(s)|Du||Dv|\,.$$

We need the following modification of this result, which on the one hand allows to test the system with powers (truncated for large values) and which on the other hand satisfies an estimate corresponding to the one from Lemma 3.10.

Lemma 3.11. *For every* $K > 0$ *and every* $q \ge 1$ *there exists a* C^2*-function* $T_{q,K}\colon \mathbb{R}^+ \to \mathbb{R}^+$ *such that*

  (i) $T_{q,K}$ *is strictly increasing and convex on* $\mathbb{R}^+$*, and it satisfies* $T_{q,K}(t) = t^{2q}$ *for all* $t \le K$*;*

 (ii) *for all* $t \in \mathbb{R}^+$ *and a constant* $c(q)$ *the growth with respect to* t *is estimated by*

$$T_{q,K}(t) + T'_{q,K}(t)t + T''_{q,K}(t)t^2 \le c(q) \min\{K^{2q-2}t^2, t^{2q}\};$$

moreover, the inequalities $T''_{q,K}(t)t - T'_{q,K}(t) \le 2(q-1)T'_{q,K}(t)$ *as well as* $T''_{q,K}(t)t^2 \le c(q)T'_{q,K}(t)t \le c(q)T_{q,K}(t)$ *hold true on* $\mathbb{R}^+$*;*

(iii) *If* $u\colon \mathbb{R}^n \to \mathbb{R}^N$ *is a function which is a.e. differentiable and* $\mu(q) := 1 - (\frac{q-1}{q})^2$*, then for the function* $v = T'_{q,K}(|u|)|u|^{-1}u$ *the following inequality is satisfied a.e.:*

$$Du \cdot Dv \ge \sqrt{\mu(q)}|Du||Dv| \ge \sqrt{\mu(q)}T'_{q,K}(|u|)|u|^{-1}|Du|^2\,.$$

Proof. We first assume $K = 1$. We set

$$T_{q,1}(t) = \begin{cases} t^{2q} & \text{if } t \le 1 \\ at^2 + bt + c & \text{if } t > 1 \end{cases}$$

for some coefficients $a, b, c \in \mathbb{R}$ to be determined as follows. The C^2-regularity condition implies that the following linear system has to be satisfied:

$$\begin{pmatrix} 1 & 1 & 1 \\ 2 & 1 & 0 \\ 2 & 0 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 1 \\ 2q \\ 2q(2q-1) \end{pmatrix} \Rightarrow \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} q(2q-1) \\ -4q(q-1) \\ (2q-1)(q-1) \end{pmatrix}.$$

We now calculate some crucial quantities. We first observe that

$$T'_{q,1}(t) = \begin{cases} 2qt^{2q-1} & \text{if } t \le 1 \\ 2q(2q-1)t - 4q(q-1) & \text{if } t > 1 \end{cases}$$

is strictly increasing and positive on $\mathbb{R}^+$. Thus, we immediately obtain assertion (i) of the lemma. Furthermore, we have

$$T''_{q,1}(t)t - T'_{q,1}(t) = \begin{cases} 4q(q-1)t^{2q-1} & \text{if } t \le 1 \\ 4q(q-1) & \text{if } t > 1, \end{cases}$$

which is again positive on $\mathbb{R}^+$. Moreover, for all $t \in \mathbb{R}^+$ we get

$$T''_{q,1}(t)t - T'_{q,1}(t) \le 2(q-1)T'_{q,1}(t), \tag{3.7}$$

which in particular yields the inequality $T''_{q,1}(t)t^2 \le c(q)T'_{q,1}(t)t$ of assertion (ii). The last inequality $T'_{q,1}(t)t \le c(q)T_{q,1}(t)$ is also checked easily. For the function $v = T'_{q,K}(|u|)|u|^{-1}u$ given in (iii) we next compute

$$D_i v^\alpha = T'_{q,1}(|u|)\frac{D_i u^\alpha}{|u|} + \left(T''_{q,1}(|u|)|u| - T'_{q,1}(|u|)\right)\frac{D_i u \cdot u u^\alpha}{|u|^3},$$

$$Du \cdot Dv = T'_{q,1}(|u|)\frac{|Du|^2}{|u|} + \left(T''_{q,1}(|u|)|u| - T'_{q,1}(|u|)\right)\frac{|Du \cdot u|^2}{|u|^3}.$$

In particular, this shows $Du \cdot Dv \ge 0$, using again the positivity of $T''_{q,1}(t)t - T'_{q,1}(t)$ and of $T'_{q,1}(t)$ on $\mathbb{R}^+$. Furthermore, we obtain

$$|Dv|^2 = T'_{q,1}(|u|)^2\frac{|Du|^2}{|u|^2} + \left(T''_{q,1}(|u|)|u| - T'_{q,1}(|u|)\right)^2\frac{|Du \cdot u|^2}{|u|^4}$$
$$+ 2T'_{q,1}(|u|)\left(T''_{q,1}(|u|)|u| - T'_{q,1}(|u|)\right)\frac{|Du \cdot u|^2}{|u|^4} \ge T'_{q,1}(|u|)^2\frac{|Du|^2}{|u|^2},$$

which yields the second inequality in (iii). Now, keeping in mind the definition of $\mu(\cdot)$, we find via the previous estimate (3.7)

$$|Du \cdot Dv|^2 - \mu(q)|Du|^2|Dv|^2 = \left(\frac{q-1}{q}\right)^2 T'_{q,1}(|u|)^2\frac{|Du|^4}{|u|^2} + \left(T''_{q,1}(|u|)|u| - T'_{q,1}(|u|)\right)^2\frac{|Du \cdot u|^4}{|u|^6}$$
$$- \left(T''_{q,1}(|u|)|u| - T'_{q,1}(|u|)\right)^2\frac{2q-1}{q^2}\frac{|Du \cdot u|^2|Du|^2}{|u|^4}$$
$$+ 2\left(\frac{q-1}{q}\right)^2 T'_{q,1}(|u|)\left(T''_{q,1}(|u|)|u| - T'_{q,1}(|u|)\right)\frac{|Du \cdot u|^2|Du|^2}{|u|^4}$$
$$\ge \left(\frac{q-1}{q}\right)^2 T'_{q,1}(|u|)^2\frac{|Du|^4}{|u|^2} + \left(T''_{q,1}(|u|)|u| - T'_{q,1}(|u|)\right)^2\frac{|Du \cdot u|^4}{|u|^6}$$
$$- 2\frac{q-1}{q}T'_{q,1}(|u|)\left(T''_{q,1}(|u|)|u| - T'_{q,1}(|u|)\right)\frac{|Du \cdot u|^2|Du|^2}{|u|^4},$$

which is non-negative by Young's inequality. This finishes the proof of (iii) for the case $K = 1$. To complete the proof of the lemma it is sufficient to observe that for general $K > 0$ the coefficients a, b, c have to be replaced by $aK^{2q-2}, bK^{2q-1}, cK^{2q}$, and the conclusion then follows exactly as above. $\square$

4   Higher differentiability of weak solutions

In this section we start working on the solution u of the parabolic system (2.3) with noise. First, we prove an upper bound for the average of *weighted* norms of Du. This will be done in Section 4.1 and serves also to explain the general strategy to obtain such estimates. We will then extract higher regularity properties of the solution, still following the ideas given in Section 4.1. More precisely, as the final result of this section, we are interested in pathwise higher integrability of the gradient Du, which will be the core of the proof of the regularity result given later in Theorem 5.1.

4.1   An a priori estimate

From Definition 2.2 of a weak solution to the system (2.3), no a priori information is available on the expected value of the solution. In particular, we only know that every weak solution $u(\omega)$ belongs to the space $V^2(D_T, \mathbb{R}^N)$ for P-almost every $\omega \in \Omega$, but it is still possible that the average $E[\|u\|_{V^2(D_T, \mathbb{R}^N)}]$ is infinite. Even if this cannot be excluded in general, we can win an a priori information on the average of weighted norms of u.

The strategy for a priori estimates for deterministic elliptic or parabolic systems is simply to "test" the equation with the solution (or some modification of it), and estimates then follow by employing the regularity and growth properties of the system. For stochastic systems, such testing with an the appropriate modification of the solution is replaced by the application of an Itô's formula for Banach spaces. Then, a first pathwise estimate follows (Step 1 in the proof of the next lemma). Since we are interested in averages, the first estimate is rewritten (Step 2) by introducing weights depending on the solution itself. With these weights we can finally take the expectation (Step 3) and end up with the desired estimate, which we now state in its precise form (note that we usually suppress to indicate the precise path $\omega \in \Omega$ for the weak solution, in particular we simply write $u \in V^2(D_T, \mathbb{R}^N)$ and implicitly mean that this property holds with probability one, by definition of the term weak solution).

Lemma 4.1. *Let $u \in V^2(D_T, \mathbb{R}^N)$ be a weak solution to the initial boundary value problem to (2.3) under the assumptions (2.1)$_{1,2}$, (2.2)$_{1,2}$ and with $u(\,\cdot\,, 0) = u_0(\cdot)$ for deterministic initial values $u_0 \in L^2(D, \mathbb{R}^N)$. Suppose further that the smallness condition $L_H^2 < 2\kappa^{-1}(1 - (1 - \nu^2)^{1/2})$ is satisfied, and let $D_0 \subset D$ with $d_0 := \mathrm{dist}(D_0, \partial D) > 0$. Then there holds*

$$E\Big[\int_0^T e^{-\int_0^t c_0 G_0(u,f)\,ds} \|Du(t)\|_{L^2(D_0)}^2 \, dt \Big] \leq c_0 \big(\|u_0\|_{L^2(D)}^2 + 1 + E\big[\|f_H\|_{L^2(D_T)}^2\big] \big)$$

for a constant c_0 depending only on D, L, L_H, d_0, κ and ν, and a function $G_0(u, f)$ given by (4.3).

Proof. **Step 1. A preliminary pathwise estimate.** We start by multiplying the equation (2.3) with a standard cut-off function $\eta \in C^\infty(D, [0,1])$ which satisfies $\eta \equiv 1$ on D_0 and $|D\eta| \leq c(d_0)$. Obviously, the map $\eta \triangle_{k,h} u$ has the same properties concerning integrability and measurability as u, and the Itô formula from Theorem 3.2 in Banach spaces may be applied with the Gelfand triple $W_0^{1,2}(D, \mathbb{R}^N) \subset L^2(D, \mathbb{R}^N) \subset W^{-1,2}(D, \mathbb{R}^N)$. This yields the existence of a subset $\Omega' \subset \Omega$ of full measure $P(\Omega') = 1$ and a function $u' : [0, T] \times \Omega \to W^{1,2}(D, \mathbb{R}^N)$ which satisfies: u' is $\mathcal{F}_t$-adapted on $[0, T] \times \Omega'$, continuous in t for every $\omega \in \Omega'$, and $u' = u\eta$ holds for $P \times \mathcal{L}^1$-almost all $(t, \omega) \in [0, T] \times \Omega$. Moreover, using the integration by parts formula, we have for every $\omega \in \Omega'$ and all $t \in [0, T]$

$$\|u'(t)\|_{L^2(D)}^2 + 2\int_0^t \langle D(u(s)\eta^2), A(\,\cdot\,, s, u(s), Du(s)) \rangle_{L^2(D)} \, ds$$

$$= \|u_0\eta\|_{L^2(D)}^2 + 2\int_0^t \langle u'(s)\eta, H(\,\cdot\,, s, Du(s))\, dB_s \rangle_{L^2(D)} + \int_0^t \|H(\,\cdot\,, s, Du(s))\eta\|_{L^2(D)}^2 \, ds \quad (4.1)$$

(with the convention $|M|^2 = \sum_{i,j=1}^n M_{ij}^2$ for every $n \times n$ matrix). Next we need to estimate the second integral on the left-hand side of the previous identity, employing the assumptions (2.1). For this purpose, we first observe with (2.1)$_{1,2}$ and Young's inequality that

$$\langle Du(s)\eta^2, A(\,\cdot\,, s, u(s), Du(s)) \rangle_{L^2(D)}$$

$$= \int_0^1 \langle Du(s)\eta^2, D_z A(\,\cdot\,, s, u(s), rDu(s))Du(s) \rangle_{L^2(D)} \, dr + \langle Du(s)\eta^2, A(\,\cdot\,, s, u(s), 0) \rangle_{L^2(D)}$$

$$= \frac{1}{\kappa} \langle Du(s)\eta, Du(s)\eta \rangle + \frac{1}{\kappa} \int_0^1 \langle Du(s)\eta^2, \kappa D_z A(\,\cdot\,, s, u(s), rDu(s))Du(s) - Du(s) \rangle_{L^2(D)} \, dr$$

$$\quad + \langle Du(s)\eta^2, A(\,\cdot\,, s, u(s), 0) \rangle_{L^2(D)}$$

$$\geq \frac{1}{\kappa} \big(1 - (1 - \nu^2)^{\frac{1}{2}} - \varepsilon\big) \|Du(s)\eta\|_{L^2(D)}^2 - c(\varepsilon^{-1}, L)\big(\|u(s)\|_{L^{\frac{2(n+2)}{n}}(D)}^{\frac{2(n+2)}{n}} + \|f(s)\|_{L^a(D)}^a \big)$$

for almost every $s \in (0, T)$ and all $\varepsilon > 0$. Moreover, again by assumption $(2.1)_1$ and Young's inequality, we find

$$
\begin{aligned}
&|\langle u(s) \otimes D\eta\eta, A(\cdot, s, u(s), Du(s)) \rangle_{L^2(D)}| \\
&\qquad \leq \frac{\varepsilon}{\kappa} \|Du(s)\eta\|^2_{L^2(D)} + c(L, \varepsilon^{-1}, \kappa, d_0)\big(\|u(s)\|^2_{L^2(D)} + \|u(s)\|^{\frac{2(n+2)}{n}}_{L^{\frac{2(n+2)}{n}}(D)} + \|f(s)\|^a_{L^a(D)}\big) \,.
\end{aligned}
$$

Next, the integrand of the last term on the right-hand side of (4.1) is bounded via $(2.2)_{1,2}$ by

$$
\begin{aligned}
&\|H(\cdot, s, Du(s))\eta\|^2_{L^2(D)} \\
&\qquad \leq \big(1 + \frac{\varepsilon}{\kappa L_H^2}\big)\|(H(\cdot, s, Du(s)) - H(\cdot, s, 0))\eta\|^2_{L^2(D)} + c(\varepsilon^{-1}, L_H, \kappa)\|H(\cdot, s, 0)\eta\|^2_{L^2(D)} \\
&\qquad \leq \big(L_H^2 + \frac{\varepsilon}{\kappa}\big)\|Du(s)\eta\|^2_{L^2(D)} + c(\varepsilon^{-1}, L_H, \kappa)\|f_H(s)\|^2_{L^2(D)} \,.
\end{aligned}
$$

Combining the last three inequalities (here enters the smallness assumption on L_H) with (4.1), choosing ε sufficiently small (in dependency of L_H and ν) and using Hölder's inequality, we thus end up with the announced preliminary pathwise estimate

$$
\begin{aligned}
&\|u'(t)\|^2_{L^2(D)} + c^{-1}(L_H, \kappa, \nu) \int_0^t \|Du(s)\eta\|^2_{L^2(D)} \, ds \\
&\qquad \leq \|u_0\eta\|^2_{L^2(D)} + 2\int_0^t \langle u'(s)\eta, H(\cdot, s, Du(s)) \, dB_s \rangle_{L^2(D)} \\
&\qquad\quad + c_0(D, L, L_H, d_0, \kappa, \nu) \int_0^t \big(1 + \|u(s)\|^{\frac{2(n+2)}{n}}_{L^{\frac{2(n+2)}{n}}(D)} + \|f(s)\|^a_{L^a(D)} + \|f_H(s)\|^2_{L^2(D)}\big) \, ds \,.
\end{aligned}
\tag{4.2}
$$

Step 2. An improved pathwise estimate. The next step consists in getting a pathwise estimate where the bound on the right-hand side contains a deterministic part almost *independent* of the weak solution and the function f, and a stochastic part which might still depend on the solution. We start by defining

$$
G_0(u, f)(s) = 1 + \|u(s)\|^{\frac{2(n+2)}{n}}_{L^{\frac{2(n+2)}{n}}(D)} + \|f(s)\|^a_{L^a(D)}
\tag{4.3}
$$

for $s \in (0, T)$. Obviously, G_0 belongs to $L^1(0, T)$ with probability one. Then we use a Gronwall-type argument, by applying the one-dimensional Itô-formula to $\exp(-\int_0^t c_0 G_0(u, f)(\tilde{s}) \, d\tilde{s})(1 + \|u'(t)\|^2_{L^2(D)})$, as e.g. in [36, Proof of Theorem 5.1]. Thus, we get

$$
\begin{aligned}
&e^{-\int_0^t c_0 G_0(u,f)(\tilde{s}) \, d\tilde{s}} \|u'(t)\|^2_{L^2(D)} + c^{-1}(L_H, \kappa, \nu) \int_0^t e^{-\int_0^s c_0 G_0(u,f)(\tilde{s}) \, d\tilde{s}} \|Du(s)\eta\|^2_{L^2(D)} \, ds \\
&\qquad \leq \|u_0\eta\|^2_{L^2(D)} + 1 + 2\int_0^t e^{-\int_0^s c_0 G_0(u,f)(\tilde{s}) \, d\tilde{s}} \langle u'(s)\eta, H(\cdot, s, Du(s)) \, dB_s \rangle_{L^2(D)} \\
&\qquad\quad + c_0 \int_0^t e^{-\int_0^s c_0 G_0(u,f)(\tilde{s}) \, d\tilde{s}} \|f_H(s)\|^2_{L^2(D)} \, ds \,.
\end{aligned}
\tag{4.4}
$$

Note that we here have omitted a negative term which appeared on the right-hand side and the positive term containing the 1 on the left-hand side. This is the desired improved pathwise estimate. We note that u and f still appear in the function G_0 in the deterministic integral on the right-hand side, but in a way that for greater values of u or f the integral gets smaller. At the same time obviously also the exponential factor on the left-hand side will get smaller, but this allows us now to proceed to Step 3.

Step 3. An estimate for the expected value with weights. Uniform estimates for the average of the weak solution (e.g. for expressions of the form $E[\|u\|]$ for some norm of u or Du) cannot be expected under such weak assumptions as we have supposed in the lemma. But the previous inequality (4.4) now allows us to get a weighted inequality, with no stochastic terms on the right-hand side. Since the expectation of the stochastic integral is not a priori known to vanish, we now apply a stopping time argument.

From identity (4.1) it follows that the process $\|u'(t)\|^2_{L^2(D)}$ has a continuous version in t, used in the following argument. For every $R > 0$ we introduce the random time

$$\tau_R := \inf\left\{t \in [0,T]: \int_0^t \|u'(s)\|^2_{L^2(D)}\|H(s, Du(s))\eta\|^2_{L^2(D)}\, ds > R\right\}$$

with $\tau_R = T$ when the set is empty. We note that $\|u'(s)\|^2_{L^2(D)}\|H(s, Du(s))\eta\|^2_{L^2(D)}$ is in $L^1(0,T)$ with probability one, because of the property $u \in V_0^2(D_T, \mathbb{R}^N)$ and the assumption $(2.2)_2$ on H. Hence, we have in particular $P(\lim_{R\to\infty} \tau_R = T) = 1$ and

$$P\left(\lim_{R\to\infty} \|u'(t \wedge \tau_R)\|^2_{L^2(D)} = \|u'(t)\|^2_{L^2(D)}\right) = 1$$

for every $t \in [0,T]$. Now we take inequality (4.4) at time $t \wedge \tau_R$ and get

$$e^{-\int_0^{t\wedge\tau_R} c_0 G_0(u,f)(\tilde{s})\, d\tilde{s}}\|u'(t \wedge \tau_R)\|^2_{L^2(D)} + c^{-1}\int_0^{t\wedge\tau_R} e^{-\int_0^s c_0 G_0(u,f)(\tilde{s})\, d\tilde{s}}\|Du(s)\eta\|^2_{L^2(D)}\, ds$$

$$\leq \|u_0\eta\|^2_{L^2(D)} + 1 + 2\int_0^{t\wedge\tau_R} e^{-\int_0^s c_0 G_0(u,f)(\tilde{s})\, d\tilde{s}}\langle\, u'(s)\eta, H(\cdot, s, Du(s))\, dB_s\,\rangle_{L^2(D)}$$

$$+ c_0\int_0^{t\wedge\tau_R} e^{-\int_0^s c_0 G_0(u,f)(\tilde{s})\, d\tilde{s}}\|f_H(s)\|^2_{L^2(D)}\, ds.$$

For the stochastic integral we now have

$$\int_0^{t\wedge\tau_R} e^{-\int_0^s c_0 G_0(u,f)(\tilde{s})\, d\tilde{s}}\langle\, u'(s)\eta, H(s, Du(s))\, dB_s\,\rangle_{L^2(D)}$$

$$= \int_0^t e^{-\int_0^s c_0 G_0(u,f)(\tilde{s})\, d\tilde{s}} 1_{s\leq\tau_R}\langle\, u'(s)\eta, H(s, Du(s))\, dB_s\,\rangle_{L^2(D)},$$

and we further have

$$\int_0^t e^{-2\int_0^s c_0 G_0(u,f)(\tilde{s})\, d\tilde{s}} 1_{s\leq\tau_R}\|u'(s)\|^2_{L^2(D)}\|H(s, Du(s))\eta\|^2_{L^2(D)}\, ds$$

$$\leq \int_0^{t\wedge\tau_R} \|u'(s)\|^2_{L^2(D)}\|H(s, Du(s))\eta\|^2_{L^2(D)}\, ds \leq R$$

by definition of τ_R. Thus the stopped stochastic integral above is a martingale, hence with expected value zero. This implies

$$E\left[e^{-\int_0^{t\wedge\tau_R} c_0 G_0(u,f)(\tilde{s})\, d\tilde{s}}\|u'(t \wedge \tau_R)\|^2_{L^2(D)}\right] + E\left[c^{-1}\int_0^{t\wedge\tau_R} e^{-\int_0^s c_0 G_0(u,f)(\tilde{s})\, d\tilde{s}}\|Du(s)\eta\|^2_{L^2(D)}\, ds\right]$$

$$\leq \|u_0\eta\|^2_{L^2(D)} + 1 + c_0 E\left[\int_0^T \|f_H(s)\|^2_{L^2(D)}\, ds\right].$$

On the left-hand side we now apply Fatou's lemma to the first term and the monotone convergence theorem to the second one, and we get

$$E\left[e^{-\int_0^t c_0 G_0(u,f)(\tilde{s})\, d\tilde{s}}\|u'(t)\|^2_{L^2(D)}\right] + E\left[c^{-1}\int_0^t e^{-\int_0^s c_0 G_0(u,f)(\tilde{s})\, d\tilde{s}}\|Du(s)\eta\|^2_{L^2(D)}\, ds\right]$$

$$\leq \|u_0\eta\|^2_{L^2(D)} + 1 + c_0 E\left[\int_0^T \|f_H(s)\|^2_{L^2(D)}\, ds\right]$$

for every $t \in [0,T]$. This proves the bound claimed in the lemma. $\qquad\qquad\square$

4.2   Existence of second order space derivatives

We next study the existence of second order space derivatives. For deterministic elliptic and parabolic partial differential equations it is a standard procedure to establish the existence of higher order derivatives by finite difference quotient methods. The basic idea in the deterministic case is the following. Once the norm of finite difference quotients of Du are kept under control *independently* of its step size, i.e. $\|\triangle_{k,h}Du\|_{L^p} \le C$ with C independent of h and with $p \in (1, \infty)$, then the weak derivative D_kDu exists and has finite norm in L^p (and as long as one is away from the boundary also the reverse it true), compare Section 3.2. So uniformly bounded difference quotients of Du can heuristically be considered as second derivatives D_kDu. This uniform bound in turn is usually achieved by "testing the system" with appropriate modifications of the solutions (formally one might think of $\triangle_{k,-h}\triangle_{k,h}u$) and relies on the one hand on the ellipticity of the vector field A and on the other hand on its regularity with respect to the x and u variables (we note that it seems mandatory to have at least Lipschitz-regularity in order to expect the existence of full second space derivatives).

For the stochastically perturbed system (2.3) the approach for proving the existence of higher order derivatives is still very similar, but we need some modifications of this method due to the stochastic terms. The above strategy (with testing replaced by the the use of the Itô formula in Banach spaces) applied to our stochastic system gives – after some standard, though very technical computations – a preliminary pathwise estimate for finite difference quotients of u and Du (this corresponds in some sense to Step 1 in the proof of the previous Lemma 4.1). But since this estimate still involves a stochastic integral, it is not yet possible to gain immediately any information on second order derivatives. In a second step, this pathwise estimate is rewritten (here again some Gronwall-type inequality is needed), which allows in the third step to take the expectation of a weighted version of $\|\triangle_{k,h}Du\|_{L^2}$ and to bound it independently of the stepsize h. This is still sufficient to deduce the existence of D_kDu with probability one (see Theorem 3.6 in Section 3.2).

Given a deterministic initial condition u_0 (sufficiently regular) we now give the precise statement on the boundedness of the expectation of finite difference quotients of Du.

Lemma 4.2. *Let $u \in V^2(D_T, \mathbb{R}^N)$ be a weak solution to the initial boundary value problem to (2.3) under the assumptions (2.1), (2.2) and with $u(\cdot, 0) = u_0(\cdot) \in W^{1,2}(D, \mathbb{R}^N)$. Suppose further that the smallness condition $L_H^2 < 2\kappa^{-1}(1 - (1 - \nu^2)^{1/2})$ is satisfied, and let $D' \subset D$ with $d' := \mathrm{dist}(D', \partial D) > 0$. Then there holds*

$$\sup_{|h|<d'} E\Big[\sup_{t\in(0,T)} e^{-\int_0^t c'G'(u,f)\,ds}\|\triangle_{k,h}u(t)\|_{L^2(D')}^2 + \int_0^T e^{-\int_0^t c'G'(u,f)\,ds}\|D\triangle_{k,h}u(t)\|_{L^2(D')}^2\,dt \Big]$$
$$\le c'\Big(\|D_ku_0\|_{L^2(D)}^2 + 1 + E\Big[\|f_H\|_{L^{\frac{2a}{a-2}}(D_T)}^{\frac{2a}{a-2}}\Big]\Big)$$

for every $k \in \{1, \ldots, n\}$, a constant c' depending only on $n, D, T, L, L_H, d', \kappa$, and ν, and a function $G'(u, f)$ given by (4.8) further below.

Proof. We proceed similarly to the proof of Lemma 4.1, with the main difference that instead of u we now need to estimate the difference quotients $\triangle_{k,h}u$.

Step 1. We first set $D_0 := \{x \in D \colon \mathrm{dist}(x, \partial D) \ge d'/2\}$ and note $D' \subset D_0 \subset D$. We now observe that if u is a solution of (2.3), then, for all $t \in [0, T]$, by definition also the following identity holds true P-a.s.:

$$\langle \eta\triangle_{k,h}u(t) - \eta\triangle_{k,h}u_0, \varphi \rangle_{L^2(D)} = \int_0^t \langle \eta\,\mathrm{div}\,\triangle_{k,h}A(\cdot, s, u(s), Du(s)), \varphi \rangle_{W^{-1,2}(D);W_0^{1,2}(D)}\,ds$$
$$+ \int_0^t \langle \varphi, \eta\triangle_{k,h}H(\cdot, s, Du(s)) \rangle\,dB_s \rangle_{L^2(D)}$$

for all $\varphi \in W_0^{1,2}(D, \mathbb{R}^N)$. Here $k \in \{1, \ldots, n\}$ is arbitrary, $h \in \mathbb{R} \setminus \{0\}$ with $|h| < d'/2$, and η denotes a standard cut-off function in $C^\infty(D_0, [0, 1])$ satisfying $\eta \equiv 1$ on D' and $|D\eta| \le c(d')$. Therefore, the map $\eta\triangle_{k,h}u$ has the same properties concerning integrability and measurability as u, and the Itô formula from Theorem 3.2 in Banach spaces is again applied with the Gelfand triple $W_0^{1,2}(D, \mathbb{R}^N) \subset$

$L^2(D, \mathbb{R}^N) \subset W^{-1,2}(D, \mathbb{R}^N)$. We hence get a subset $\Omega' \subset \Omega$ of full measure $P(\Omega') = 1$ and a function $u'_k \colon [0, T] \times \Omega \to W^{1,2}(D, \mathbb{R}^N)$ with the following properties: u'_k is $\mathcal{F}_t$-adapted on $[0, T] \times \Omega'$, continuous in t for every $\omega \in \Omega'$, and satisfies $u'_k = \triangle_{k,h} u\eta$ for $P \times \mathcal{L}^1$-almost all $(t, \omega) \in [0, T] \times \Omega$. Moreover, for every $\omega \in \Omega'$ and all $t \in [0, T]$ we have

$$\|u'_k(t)\|^2_{L^2(D)} + 2 \int_0^t \langle\, D(\triangle_{k,h} u(s)\eta^2), \triangle_{k,h} A(\cdot, s, u(s), Du(s))\,\rangle_{L^2(D)} \, ds$$

$$= \|\triangle_{k,h} u_0 \eta\|^2_{L^2(D)} + 2 \int_0^t \langle\, u'_k(s)\eta, \triangle_{k,h} H(\cdot, s, Du(s)) \, dB_s\,\rangle_{L^2(D)}$$

$$+ \int_0^t \|\triangle_{k,h} H(\cdot, s, Du(s))\eta\|^2_{L^2(D)} \, ds \,. \tag{4.5}$$

Our first aim is to deduce a pathwise estimate for finite differences of u and Du, respectively. To this end we start by studying in detail the second term on the left-hand side. For almost every $t \in [0, T]$ we decompose the finite difference quotient applied on $A(x, t, u, Du)$ as follows

$$\begin{aligned}
\triangle_{k,h} &A(x, t, u(x), Du(x)) \\
&= h^{-1}\big[A(x + he_k, t, u(x + he_k), Du(x + he_k)) - A(x + he_k, t, u(x + he_k), Du(x))\big] \\
&\quad + h^{-1}\big[A(x + he_k, t, u(x + he_k), Du(x)) - A(x + e_k, t, u(x), Du(x))\big] \\
&\quad + h^{-1}\big[A(x + he_k, t, u(x), Du(x)) - A(x, t, u(x), Du(x))\big] \\
&= \int_0^1 D_z A(x + he_k, t, u(x + he_k), Du(x) + rh\triangle_{k,h} Du(x)) \, dr \, \triangle_{k,h} Du(x) \\
&\quad + \int_0^1 D_u A(x + he_k, t, u(x) + rh\triangle_{k,h} u(x), Du(x)) \, dr \, \triangle_{k,h} u(x) \\
&\quad + \int_0^1 D_x A(x + rhe_k, t, u(x), Du(x)) \, dr \\
&=: \mathcal{A}(h) + \mathcal{B}(h) + \mathcal{C}(h)
\end{aligned} \tag{4.6}$$

with the obvious abbreviations. Using the assumptions (2.1), Hölder's and Young's inequality, we now estimate the different terms arising from this decomposition in equation (4.5) on time slices $t \in (0, T)$ (on such slices we omit the notion of t). We first find for almost every $t \in (0, T)$ and every $\varepsilon > 0$

$$\begin{aligned}
\langle\, D&(\triangle_{k,h} u\eta^2), \mathcal{A}(h)\,\rangle_{L^2(D)} \\
&= \kappa^{-1}\langle\, D(\triangle_{k,h} u\eta^2), D\triangle_{k,h} u\,\rangle_{L^2(D)} - \kappa^{-1}\langle\, D(\triangle_{k,h} u\eta^2), D\triangle_{k,h} u - \kappa\mathcal{A}(h)\,\rangle_{L^2(D)} \\
&\geq \kappa^{-1}\|D\triangle_{k,h} u\eta\|^2_{L^2(D)} - 2\varepsilon\|D\triangle_{k,h} u\eta\|^2_{L^2(D)} - c(\kappa, \varepsilon)\|\triangle_{k,h} u D\eta\|^2_{L^2(D)} \\
&\quad - \kappa^{-1}(1 - \nu^2)^{\frac{1}{2}}\|D\triangle_{k,h} u\eta\|^2_{L^2(D)} \\
&\geq \big(\kappa^{-1}(1 - (1 - \nu^2)^{\frac{1}{2}}) - 2\varepsilon\big)\|D\triangle_{k,h} u\eta\|^2_{L^2(D)} - c(\kappa, \varepsilon, \|D\eta\|_{L^\infty(D)})\|D_k u\|^2_{L^2(D)}
\end{aligned}$$

where in the last line we have used the fact that the norm of the finite difference quotient of a compactly supported function is always bounded by the norm of the partial derivative (provided that the stepsize is sufficiently small). This lower bound will be crucial (and can be understood as some ellipticity of the vector field A up to lower order terms). We next observe with Hölder's inequality that given arbitrary exponents $p_i > 0$, $q_i \geq \max\{1, 1/p_i\}$ and arbitrary functions $f_i \in L^{q_i}(D)$, with $i \in \{1, \ldots, m\}$ for some $m \in \mathbb{N}$, we have

$$\big\| \prod_{i=1}^m |f_i|^{p_i} \big\|_{L^1(D)} \leq \prod_{i=1}^m \|f_i\|^{p_i}_{L^{q_i p_i}(D)}$$

provided that $\sum_{i=1}^m q_i^{-1} = 1$ holds. Combining the growth condition (2.1) for $D_u A$ and this inequality with $m = 4$ and exponents $2, 2/(1-\theta), 2n/((n-2)\theta), n/\theta$ as q_i's, we then infer from the Sobolev-Poincaré

embedding (applied on every time-slice)

$$\left|\langle D(\triangle_{k,h}u\eta^2), \mathcal{B}(h)\rangle_{L^2(D)}\right|$$

$$\leq L\langle |D\triangle_{k,h}u\eta| + 2|\triangle_{k,h}uD\eta|, \left(|Du|^{\frac{2}{n+2}} + \int_0^1 |u(x) + rh\triangle_{k,h}u|^{\frac{2}{n}}\,dr + |f|\right)|\triangle_{k,h}u|\rangle_{L^2(D)}$$

$$\leq L\left(\|D\triangle_{k,h}u\eta\|_{L^2(D)} + 2\|\triangle_{k,h}uD\eta\|_{L^2(D)}\right)\|\triangle_{k,h}u\eta\|^{\theta}_{L^{\frac{2n}{n-2}}(D)}$$

$$\times \|\triangle_{k,h}u\eta\|^{1-\theta}_{L^2(D)}\left(\|Du\|^{\frac{2}{n+2}}_{L^{\frac{2n}{(n+2)\theta}}(D)} + \|u\|^{\frac{2}{n}}_{L^{\frac{2}{\theta}}(D)} + \|f\|_{L^{\frac{n}{\theta}}(D)}\right)$$

$$\leq \left(\|D\triangle_{k,h}u\eta\|^{1+\theta}_{L^2(D)} + \|\triangle_{k,h}uD\eta\|^{1+\theta}_{L^2(D)}\right)$$

$$\times c(n,D,L)\|\triangle_{k,h}u\eta\|^{1-\theta}_{L^2(D)}\left(\|Du\|^{\frac{2}{n+2}}_{L^{\frac{2n}{(n+2)\theta}}(D)} + \|u\|^{\frac{2}{n}}_{L^{\frac{2}{\theta}}(D)} + \|f\|_{L^{\frac{n}{\theta}}(D)}\right) \tag{4.7}$$

for every $\theta \in (0,1)$; in the two-dimensional case $n = 2$, $2n/(n-2)$ shall be interpreted as any arbitrary number greater than 2. We note that we have omitted the step of passing from the shifted to the original domain in the first inequality and we have applied $cd^\theta + c^\theta d \leq c^{1+\theta} + d^{1+\theta}$ for all $c, d \geq 0$ to get from the first to the second inequality. To estimate further we choose $\theta = n/(n+2)$, according to the integrability assumptions of u (using the embedding given in (2.4)), Du and f. Thus, Young's inequality gives

$$\left|\langle D(\triangle_{k,h}u\eta^2), \mathcal{B}(h)\rangle_{L^2(D)}\right|$$

$$\leq \varepsilon\|D\triangle_{k,h}u\eta\|^2_{L^2(D)} + c(n,D,L,\|D\eta\|_{L^\infty(D)},\varepsilon)\left(\|\triangle_{k,h}u\eta\|^2_{L^2(D)} + 1\right)$$

$$\times \left(\|Du\|^2_{L^2(D)} + \|u\|^{\frac{2(n+2)}{n}}_{L^{\frac{2(n+2)}{n}}(D)} + \|f\|^{n+2}_{L^{n+2}(D)}\right).$$

Finally, using Young's inequality and standard properties of finite difference quotients, we estimate the last term involving $\mathcal{C}(h)$ by

$$\left|\langle D(\triangle_{k,h}u\eta^2), \mathcal{C}(h)\rangle_{L^2(D)}\right|$$

$$\leq L\left(\|D\triangle_{k,h}u\eta\|_{L^2(D)} + 2\|\triangle_{k,h}uD\eta\|_{L^2(D)}\right)\left(\|Du\|_{L^2(D)} + \|u\|^{\frac{n+2}{n}}_{L^{\frac{2(n+2)}{n}}(D)} + \|f\|^2_{L^4(D)}\right)$$

$$\leq \varepsilon\|D\triangle_{k,h}u\eta\|^2_{L^2(D)} + c(D,L,\|D\eta\|_{L^\infty(D)},\varepsilon)\left(1 + \|Du\|^2_{L^2(D)} + \|u\|^{\frac{2(n+2)}{n}}_{L^{\frac{2(n+2)}{n}}(D)} + \|f\|^{n+2}_{L^{n+2}(D)}\right).$$

Now we have estimated all terms coming from the integral involving $\triangle_{k,h}A(x,t,u,Du)$. Next we study the last integral in equation (4.5). Employing the properties (2.2), we find

$$\|\triangle_{k,h}H(s,Du(s))\eta\|_{L^2(D)} \leq \||f_H|^{\frac{a}{a-2}}(s) + |Du(s)|\eta\|_{L^2(D)} + L_H\|D\triangle_{k,h}u(s)\eta\|_{L^2(D)}.$$

Before summarizing the previous estimates for the single terms, we introduce, for ease of notation, the function

$$G'(u,f)(s) := 1 + \|Du(s)\|^2_{L^2(D)} + \|u(s)\|^{\frac{2(n+2)}{n}}_{L^{\frac{2(n+2)}{n}}(D)} + \|f(s)\|^a_{L^a(D)} \tag{4.8}$$

with $s \in (0,T)$, which belongs to $L^1(0,T)$ almost surely. Note that by definition we have $G'(u,f) \geq G_0(u,f)$ with $G_0(u,f)$ denoting the function introduced in Lemma 4.1. Combining the latter estimates with the decomposition given in (4.6), using standard properties for finite difference quotients and choosing $\varepsilon = \varepsilon(\kappa,\nu,L_H)$ sufficiently small, we hence infer from (4.5) that for every $\omega \in \Omega'$ there holds

$$\|u'_k(t)\|^2_{L^2(D)} + c^{-1}(L_H,\kappa,\nu)\int_0^t \|D\triangle_{k,h}u(s)\eta\|^2_{L^2(D)}\,ds$$

$$\leq \|\triangle_{k,h}u_0\eta\|^2_{L^2(D)} + c'\int_0^t \left(\|\triangle_{k,h}u(s)\eta\|^2_{L^2(D)} + 1\right)G'(u,f)(s)\,ds$$

$$+ 2\int_0^t \langle u'_k(s)\eta, \triangle_{k,h}H(\cdot,s,Du(s))\,dB_s\rangle_{L^2(D)} + 2\int_0^t \|f_H(s)\|^{\frac{2a}{a-2}}_{L^{\frac{2a}{a-2}}(D)}\,ds, \tag{4.9}$$

and the constant c' depends only on n, D, L, L_H, d', κ and ν. Here we assume $c' \geq c_0$ with c_0 denoting the constant given in Lemma 4.1. This is the preliminary pathwise estimate on the finite difference quotients (which however involves the stochastic integral) and concludes the Step 1.

Step 2. Before passing to the expectation value as described in the beginning we still need the announced Gronwall-type argument, similarly as in the proof of Lemma 4.1. However, we here observe that the second integral on the right-hand side is in general not known to be finite, but the first factor of the integrand "almost" happens to appear in the sum of its left-hand side (in the sense that u'_k differs from $\triangle_{k,h} u\eta$ only on a negligible set). Hence, in order to get rid of this possibly uncontrollable term we apply the one-dimensional Itô-formula (recalling $a > n + 2$), and we obtain

$$
e^{-\int_0^t c'G'(u,f)(\tilde{s})\, d\tilde{s}} \|u'_k(t)\|_{L^2(D)}^2 + c^{-1} \int_0^t e^{-\int_0^s c'G'(u,f)(\tilde{s})\, d\tilde{s}} \|D\triangle_{k,h} u(s)\eta\|_{L^2(D)}^2 \, ds
$$
$$
\leq c' \int_0^t e^{-\int_0^s c'G'(u,f)(\tilde{s})\, d\tilde{s}} G'(u,f) \big(\|\triangle_{k,h} u(s)\eta\|_{L^2(D)}^2 - \|u'_k(s)\|_{L^2(D)}^2 \big)\, ds
$$
$$
+ \|\triangle_{k,h} u_0\eta\|_{L^2(D)}^2 + 1 + 2 \int_0^t e^{-\int_0^s c'G'(u,f)(\tilde{s})\, d\tilde{s}} \langle u'_k(s)\eta, \triangle_{k,h} H(\cdot, s, Du(s))\, dB_s \rangle_{L^2(D)}
$$
$$
+ 2 \int_0^t e^{-\int_0^s c'G'(u,f)(\tilde{s})\, d\tilde{s}} \|f_H(s)\|_{L^{\frac{2a}{a-2}}(D)}^{\frac{2a}{a-2}}\, ds\,. \tag{4.10}
$$

Here, we again note that the average of the first integral on the right-hand side vanishes, due to the fact that u'_k and $\triangle_{k,h} u\eta$ coincide on $[0, T] \times \Omega$ except for a set of $\mathcal{L}^1 \times P$-measure zero.

Step 3. We now derive the desired estimate for the average with weights. In contrast to Lemma 4.1, we now derive in a first step an estimate for the (weighted) average of the $L^2(L^2)$-norm of $D\triangle_{k,h} u$ (which proceeds in exactly the same way as before). Then, we use this estimate to get also an upper bound for the (weighted) average of the $L^\infty(L^2)$-norm of $\triangle_{k,h} u$.

We first note that identity (4.5) implies that the process $\|u'_k(t)\|_{L^2(D)}^2$ has a continuous version in t. Now, for every $R > 0$, we introduce the random time

$$
\tau_R := \inf \Big\{ t \in [0, T] \colon \int_0^t \|u'_k(s)\|_{L^2(D)}^2 \|\triangle_{k,h} H(s, Du(s))\eta\|_{L^2(D)}^2 \, ds > R \Big\}
$$

with $\tau_R = T$ when the set is empty. Notice that

$$
\int_0^T \|u'_k(s)\|_{L^2(D)}^2 \|\triangle_{k,h} H(s, Du(s))\eta\|_{L^2(D)}^2 \, ds < \infty
$$

with probability one, because of the property $u \in V_0^2(D_T, \mathbb{R}^N)$ combined with the assumption $(2.2)_2$ on H. Hence, when $R \to \infty$, τ_R is eventually equal to T, with probability one. In particular, $P(\lim_{R\to\infty} \tau_R = T) = 1$ and for every $t \in [0, T]$ we have

$$
P\Big(\lim_{R\to\infty} \|u'_k(t \wedge \tau_R)\|_{L^2(D)}^2 = \|u'_k(t)\|_{L^2(D)}^2 \Big) = 1\,.
$$

Step 3a. We compute inequality (4.10) at time $t \wedge \tau_R$ and get

$$
e^{-\int_0^{t \wedge \tau_R} c'G'(u,f)(\tilde{s})\, d\tilde{s}} \|u'_k(t \wedge \tau_R)\|_{L^2(D)}^2 + c^{-1} \int_0^{t \wedge \tau_R} e^{-\int_0^s c'G'(u,f)(\tilde{s})\, d\tilde{s}} \|D\triangle_{k,h} u(s)\eta\|_{L^2(D)}^2 \, ds
$$
$$
\leq c' \int_0^{t \wedge \tau_R} e^{-\int_0^s c'G'(u,f)(\tilde{s})\, d\tilde{s}} G'(u,f) \big(\|\triangle_{k,h} u(s)\eta\|_{L^2(D)}^2 - \|u'_k(s)\|_{L^2(D)}^2 \big)\, ds
$$
$$
+ \|\triangle_{k,h} u_0\eta\|_{L^2(D)}^2 + 1 + 2 \int_0^{t \wedge \tau_R} e^{-\int_0^s c'G'(u,f)(\tilde{s})\, d\tilde{s}} \langle u'_k(s)\eta, \triangle_{k,h} H(s, Du(s))\, dB_s \rangle_{L^2(D)}
$$
$$
+ 2 \int_0^{t \wedge \tau_R} e^{-\int_0^s c'G'(u,f)(\tilde{s})\, d\tilde{s}} \|f_H(s)\|_{L^{\frac{2a}{a-2}}(D)}^{\frac{2a}{a-2}}\, ds\,.
$$

Now we have

$$\int_0^{t\wedge\tau_R} e^{-\int_0^s c'G'(u,f)(\tilde{s})\,d\tilde{s}} \langle\, u_k'(s)\eta, \triangle_{k,h}H(s, Du(s))\, dB_s\,\rangle_{L^2(D)}$$

$$= \int_0^t e^{-\int_0^s c'G'(u,f)(\tilde{s})\,d\tilde{s}} 1_{s\leq\tau_R}\langle\, u_k'(s)\eta, \triangle_{k,h}H(s, Du(s))\, dB_s\,\rangle_{L^2(D)}$$

and

$$\int_0^t e^{-2\int_0^s c'G'(u,f)(\tilde{s})\,d\tilde{s}} 1_{s\leq\tau_R}\|u_k'(s)\|^2_{L^2(D)}\|\triangle_{k,h}H(s, Du(s))\eta\|^2_{L^2(D)}\,ds$$

$$\leq \int_0^{t\wedge\tau_R} \|u_k'(s)\|^2_{L^2(D)}\|\triangle_{k,h}H(s, Du(s))\eta\|^2_{L^2(D)}\,ds \leq R$$

by definition of τ_R. Thus, the stopped stochastic integral above is a martingale, hence with expected value zero. This implies (using that u_k' equals $\triangle_{k,h}u\eta$ outside a set of $\mathcal{L}^1\times P$-measure zero)

$$E\Big[e^{-\int_0^{t\wedge\tau_R} c'G'(u,f)(\tilde{s})\,d\tilde{s}}\|u_k'(t\wedge\tau_R)\|^2_{L^2(D)}\Big] + c^{-1}E\Big[\int_0^{t\wedge\tau_R} e^{-\int_0^s c'G'(u,f)(\tilde{s})\,d\tilde{s}}\|D\triangle_{k,h}u(s)\eta\|^2_{L^2(D)}\,ds\Big]$$

$$\leq \|\triangle_{k,h}u_0\eta\|^2_{L^2(D)} + 1 + 2E\Big[\int_0^T \|f_H(s)\|^{\frac{2a}{a-2}}_{L^{\frac{2a}{a-2}}(D)}\,ds\Big].$$

We apply Fatou's lemma to the first term and monotone convergence theorem to the second one on the left-hand side, and we get

$$E\Big[e^{-\int_0^t c'G'(u,f)(\tilde{s})\,d\tilde{s}}\|u_k'(t)\|^2_{L^2(D)}\Big] + c^{-1}E\Big[\int_0^t e^{-\int_0^s c'G'(u,f)(\tilde{s})\,d\tilde{s}}\|D\triangle_{k,h}u(s)\eta\|^2_{L^2(D)}\,ds\Big]$$

$$\leq \|\triangle_{k,h}u_0\eta\|^2_{L^2(D)} + 1 + 2E\Big[\int_0^T \|f_H(s)\|^{\frac{2a}{a-2}}_{L^{\frac{2a}{a-2}}(D)}\,ds\Big]$$

for every $t\in[0,T]$. This proves one of the two bounds claimed by the Lemma.

Step 3b. It is almost our final estimate except that we need the supremum in time inside the first expected value, and thus we have to repeat the previous computations by means of martingale inequalities. The previous estimate (as well as the a priori estimate from Lemma 4.1) will be used in the next one; we found it convenient to proceed in two steps. From the stopped inequality above we have

$$E\Big[\sup_{t\in[0,T]} e^{-\int_0^{t\wedge\tau_R} c'G'(u,f)(\tilde{s})\,d\tilde{s}}\|u_k'(t\wedge\tau_R)\|^2_{L^2(D)}\Big]$$

$$\leq \|\triangle_{k,h}u_0\eta\|^2_{L^2(D)} + 1 + 2E\Big[\int_0^T \|f_H(s)\|^{\frac{2a}{a-2}}_{L^{\frac{2a}{a-2}}(D)}\,ds\Big].$$

$$+ 2E\Big[\sup_{t\in[0,T]}\Big|\int_0^t e^{-\int_0^s c'G'(u,f)(\tilde{s})\,d\tilde{s}} 1_{s\leq\tau_R}\langle\, u_k'(s)\eta, \triangle_{k,h}H(s, Du(s))\, dB_s\,\rangle_{L^2(D)}\Big|\Big].$$

We apply again Fatou's lemma to the expected value on the left-hand side. The last term on the right-hand side is estimated, by means of the Burkholder-Davis-Gundy inequality, by

$$CE\Big[\Big(\int_0^T e^{-2\int_0^s c'G'(u,f)(\tilde{s})\,d\tilde{s}} 1_{s\leq\tau_R}\|u_k'(s)\|^2_{L^2(D)}\|\triangle_{k,h}H(s, Du(s))\eta\|^2_{L^2(D)}\,ds\Big)^{1/2}\Big]$$

$$= CE\Big[\Big(\int_0^T e^{-2\int_0^{s\wedge\tau_R} c'G'(u,f)(\tilde{s})\,d\tilde{s}} 1_{s\leq\tau_R}\|u_k'(s\wedge\tau_R)\|^2_{L^2(D)}\|\triangle_{k,h}H(s, Du(s))\eta\|^2_{L^2(D)}\,ds\Big)^{1/2}\Big]$$

$$\leq CE\Big[I_1^{1/2}I_2^{1/2}\Big] \leq \frac{1}{2}E[I_1] + \frac{C^2}{2}E[I_2]$$

where

$$I_1 = \sup_{t \in [0,T]} e^{-\int_0^{t \wedge \tau_R} c'G'(u,f)(\tilde{s})\,d\tilde{s}} \|u_k'(t \wedge \tau_R)\|_{L^2(D)}^2,$$

$$I_2 = \int_0^T e^{-\int_0^s c'G'(u,f)(\tilde{s})\,d\tilde{s}} \|\triangle_{k,h} H(s, Du(s))\eta\|_{L^2(D)}^2\,ds.$$

Hence, we have proved

$$\frac{1}{2} E[I_1] \leq \|\triangle_{k,h} u_0 \eta\|_{L^2(D)}^2 + 1 + 2E\Big[\int_0^T \|f_H(s)\|_{L^{\frac{2a}{a-2}}(D)}^{\frac{2a}{a-2}}\,ds\Big] + \frac{C^2}{2} E[I_2].$$

From the estimate above for $\|\triangle_{k,h} H(s, Du(s))\eta\|_{L^2(D)}^2$ and the estimate of Step 3a, we know that $E[I_2]$ is bounded from above via

$$E[I_2] \leq 4E\Big[\int_0^T e^{-\int_0^s c'G'(u,f)(\tilde{s})\,d\tilde{s}} \|Du(s)\eta\|_{L^2(D)}^2\,ds\Big]$$
$$+ c\Big(\|\triangle_{k,h} u_0 \eta\|_{L^2(D)}^2 + 1 + E\Big[\int_0^T \|f_H(s)\|_{L^{\frac{2a}{a-2}}(D)}^{\frac{2a}{a-2}}\,ds\Big]\Big).$$

With the initial choice of D_0 (in dependency of D and D') and keeping in mind $c'G'(u, f) \geq c_0 G_0(u, f)$ by construction, the first average on the right-hand side of the last inequality is bounded due to Lemma 4.1. Hence, keeping in mind that u_k' coincides with $\triangle_{k,h} u\eta$ as $W^{1,2}(D, \mathbb{R}^N)$-functions $P \times \mathcal{L}^1$-almost everywhere, we get from the bound for $E[I_1]$ the asserted inequality in the lemma. Since h was arbitrary and $\eta = 1$ on D', the proof is complete. $\qquad\square$

Remark 4.3. *For SPDEs having first space derivatives of the solution in the coefficient of the noise, the most general condition for existence of solutions in L^2, which becomes also a condition for an improvement of $W^{k,2}$-regularity, is more precise than just the control on the Lipschitz constant of H expressed by the statement of Lemma 4.2; see [33, 26]. However, when we go to $W^{k,p}$-regularity with $p > 2$, the computations are too involved and the algebraic simplicity of the condition of [33, 26] seems to be lost. For this reason we have simplified the estimate also for $p = 2$.*

Applying Theorem 3.6 with $(p, q) = (2, 2)$ and $(p, q) = (2, \infty)$ and summing over $k \in \{1, \ldots, n\}$ we then infer from the previous lemma that second order spatial derivatives of u exist almost surely. We should note that this result does not extend up to the boundary of D since the constant c' blows up for $\mathrm{dist}(D', D) \searrow 0$, but the result holds on any fixed subset $D' \Subset D$.

Corollary 4.4. *Let $u \in V^2(D_T, \mathbb{R}^N)$ be a weak solution under the assumptions of the Lemma 4.2. Then there holds $Du \in V^2(D_T', \mathbb{R}^N)$ with probability one, and*

$$E\Big[\sup_{t \in (0,T)} e^{-\int_0^t c'G'(u,f)\,ds} \|Du\|_{L^2(D')}^2 + \int_0^T e^{-\int_0^t c'G'(u,f)\,ds} \|D^2 u\|_{L^2(D')}^2\,dt\Big]$$
$$\leq c'\Big(\|Du_0\|_{L^2(D)}^2 + 1 + E\Big[\|f_H\|_{L^{\frac{2a}{a-2}}(D_T)}^{\frac{2a}{a-2}}\Big]\Big)$$

for the constant c' from Lemma 4.2. Moreover, we have for all $k \in \{1, \ldots, n\}$

$$e^{-\frac{1}{2}\int_0^t c'G'(u,f)\,ds} \triangle_{k,h} Du \to e^{-\frac{1}{2}\int_0^t c'G'(u,f)\,ds} D_k Du \quad \text{weakly in } L^2(D_T' \times \Omega).$$

4.3 Iteration

In the next step we iterate the procedure from the previous section, in a way such that we do not only know the spatial gradient Du to belong to the space V^2 with probability one, but that we get this result also up to a certain power of $|Du|$. For convenience we introduce the function

$$W_q \colon \mathbb{R}^k \to \mathbb{R} \quad \text{defined by} \quad W_q(\xi) := |\xi|^q$$

for every $q \geq 0$. We start by briefly describing the strategy how this regularity improvement is achieved. First we observe from the results in Section 4.2 that there exists a subset of Ω of full measure on which Du belongs to $V_{\mathrm{loc}}^2(D_T)$, hence we can now take advantage of higher integrability properties for u and Du. This shall be done with the following (formal) iteration scheme:

$$W_{q_j}(Du) \in V^2 \quad \longrightarrow \quad Du \in L^{2q_j \frac{n+2}{n}} \text{ and } u \in L^{2q_j(\frac{n+2}{n})^2} \cap L^\infty(L^{2q_j \frac{n}{n-2}})$$

$$\longrightarrow \quad D_u A(x,t,u,Du) \in L^{\min\{q_j \frac{(n+2)^2}{n}, a\}} \text{ and } D_x A(x,t,u,Du) \in L^{\min\{2q_j \frac{n+2}{n}, \frac{a}{2}\}}$$

$$\longrightarrow \quad W_{q_{j+1}}(Du) \in V^2$$

for a sequence $\{q_j\}_{j \in \mathbb{N}}$ of numbers $q_j \geq 1$ for all $j \in \mathbb{N}$. The first implication indeed follows from Sobolev's embedding for the space $V^{2,p}$, the second one from the growth conditions on the vector field A, the third one from the iteration (and a convergence result concerning finite difference quotients). After a finite number of steps we then arrive at a final (maximal) higher integrability exponent, which essentially reflects how close the vector field A is to the Laplace system. This should be understood in the following sense: the closer ν is to one (note that $\nu = 1$ corresponds to the case $A(x,t,u,z) = z$ plus potential lower order terms), the more integrability for Du can be gained in the iteration and the better will be the final regularity properties of u. Finally, we note that in every step of the iteration we will have to reduce the radius of the parabolic cylinder and we will also have to restrict ourselves to smaller subsets of Ω. Nevertheless, the higher integrability results will always be true on sets of probability one.

We now start with some preliminary remarks and consider again the equation (2.3)

$$du = \operatorname{div} A(x,t,u,Du)\,dt + H(x,t,Du)\,dB_t$$

in D_T. We observe that $\operatorname{div} A(x,t,u,Du)$ is well defined in view of the regularity assumptions (2.1) and the existence of second order spatial derivatives, see Corollary 4.4. More precisely, it is easy to check that for every weak solution $u \in V_0^2(D_T, \mathbb{R}^N)$ we have: $\operatorname{div} A(x,t,u,Du) \in L_{\mathrm{loc}}^2(D', \mathbb{R}^N)$ with probability one, and the equation above holds for $\mathcal{L}^n$-almost every $x \in D'$ for $\mathcal{L}^1 \times P$-almost all $(t,\omega) \in (0,T) \times \Omega$. Hence, we can now work immediately with this equation without passing to its weak formulation.

In the next lemma we provide the main step of the iteration argument:

Lemma 4.5. *Let $u \in V^2(D_T, \mathbb{R}^N)$ be a weak solution to the initial boundary value problem to (2.3) under the assumptions (2.1), (2.2) and with initial values $u(\,\cdot\,, 0) = u_0(\cdot) \in W^{1,2}(D, \mathbb{R}^N)$, and assume that*

$$E\left[\|Y_p^p W_p(Du)\|_{L^2 \frac{n+2}{n}(D_T')}^{\frac{2}{p}}\right] \leq C_p < \infty$$

for some $p \geq 1$, a set $D' \subset D$, and $Y_p \colon [0,T] \times \Omega \to (0,1]$ given by $Y_p(t,\omega) = \exp(-\int_0^t G_p(s,\omega)\,ds)$ for some function G_p which is in $L^1(0,T)$ with probability one. Let $D'' \subset D'$ with $d := \operatorname{dist}(D'', \partial D') > 0$. Then for every number $q \geq 1$ satisfying

$$q \leq \min\left\{p\frac{n+2}{n}, 1 + p\frac{n+2}{n}\frac{a-4}{a}, \frac{a-2}{2}\right\} \quad and \quad L_H^2 < \frac{1}{\kappa(q - \frac{1}{2})}\left(\left[1 - \left(\frac{q-1}{q}\right)^2\right]^{\frac{1}{2}} - \left[1 - \nu^2\right]^{\frac{1}{2}}\right), \quad (4.11)$$

all initial values $u_0 \in W^{1,2q}(D, \mathbb{R}^N)$, and every $k \in \{1, \ldots, n\}$ there holds

$$\sup_{|h| < d} E\left[\left(\sup_{t \in (0,T)} \|Y_q^q W_q(\triangle_{k,h} u)\|_{L^2(D'')}^2 + \int_0^T \|Y_q^q DW_q(\triangle_{k,h} u)\|_{L^2(D'')}^2\,dt\right)^{\frac{1}{q}}\right]$$

$$\leq c\left(\|W_q(D_k u_0)\|_{L^2(D)}^2 + 1 + E\left[\|f_H\|_{L^a(D_T)}^a\right]\right)^{\frac{1}{q}},$$

for $Y_q \colon [0,T] \times \Omega \to (0,1]$ given by $Y_q(t,\omega) = \exp(-\int_0^t G_q(s,\omega)\,ds)$ for some function G_q which is in $L^1(0,T)$ with probability one, and a constant c depending only on $n, p, D, T, L, L_H, d, \kappa, \nu,$ and C_p.

Remarks 4.6. *In the case of additive noise (with $L_H = 0$) the second condition (4.11) for the restriction on the integrability exponent q reduces to the inequality $q < 1/(1-\nu)$. For multiplicative noise instead,*

the right-hand side in the second inequality (4.11) *is decreasing in q (note that for q = 1 it just reproduces the condition required in Lemma 4.2) and allows the following interpretation. Obviously, the previous restriction $q < 1/(1 - \nu)$ for additive noise remains valid, and in fact the more multiplicative noise is considered (in the sense that L_H should not be too small), the smaller will be the maximal integrability exponent, which however still needs to satisfy both inequalities in* (4.11). *For this reason multiplicative noise might destroy some regularity in form of a loss of integrability of the gradient Du.*

Moreover, we comment on the scaling of the hypothesis and the assertion with respect to u and u_0, respectively, in order to avoid confusion. In view of the definition of W_p it is easy to see that Lemma 4.5 is stated in a way such that an weighted average of a quadratic quantity in Du gives an information about the weighted average of a quadratic quantity in $\triangle_{k,h} u$. In this sense, the scaling is the natural one.

Proof. We now follow the line of arguments from the proof of Lemma 4.2 (and of Lemma 4.1), with the essential novelty that we estimate this time powers of the difference quotients $\triangle_{k,h} u$.

Step 1. We consider $k \in \{1, \ldots, n\}$ arbitrary, $h \in \mathbb{R}$ with $|h| < d$, and $\eta \in C^\infty(D', [0, 1])$ a standard cut-off function satisfying $\eta \equiv 1$ on $D'' \Subset D'$ and $|D\eta| \leq c(d)$. We first observe that, by the integrability assumptions on Du and on G_p (which in particular implies strict positivity of $\inf_{t \in [0,T]} Y_p$ for P-almost every ω), we have

$$u \in L_{\text{loc}}^{2p(\frac{n+2}{n})^2}(D_T', \mathbb{R}^N) \cap L_{\text{loc}}^\infty(0, T; L^{2p\frac{n}{n-2}}(D', \mathbb{R}^N))$$

with probability one. Furthermore, due to the restriction on q, it is guaranteed that $W_q(Du)$ belongs to L^2 locally on D_T' with probability one. For almost every (fixed) $x \in D$ we first consider finite differences in direction e_k and stepsize h of the differential equation (2.3), i.e.

$$d\, \eta^{\frac{1}{q}} \triangle_{k,h} u(x, t) = \eta^{\frac{1}{q}} \, \text{div} \, \triangle_{k,h} A(x, t, u, Du) \, dt + \eta^{\frac{1}{q}} \triangle_{k,h} H(x, t, Du) \, dB_t$$

in $(0, T)$ for $q \geq 1$. We next introduce (for technical reasons) for $K > 0$ the approximating function $T_{q,K}(\cdot)$ of class C^2 according to Lemma 3.11, and we recall that $T_{q,K}$ satisfies in particular the polynomial growth conditions $T_{q,K}(t) = t^{2q}$ for all $t \leq K$ and $T_{q,K}(t) \leq c(q)K^{2q-2}t^2$ for all $t \in \mathbb{R}$. Employing the one-dimensional Itô formula (note that $\text{div} \, A(x, t, u, Du)$ is as a composition of $\mathcal{F}_t$-adapted functions again $\mathcal{F}_t$-adapted) from Theorem 3.1, applied with $g(t, u(x, t)) = \eta^2 T_{q,K}(|\triangle_{k,h} u(x, t)|)$, we obtain the identity

$$
\begin{aligned}
d\big(&\eta^2 T_{q,K}(|\triangle_{k,h} u(x, t)|)\big) \\
&= \eta^2 T_K'(|\triangle_{k,h} u(x, t)|)|\triangle_{k,h} u(x, t)|^{-1} \langle \triangle_{k,h} u(x, t), \text{div} \, \triangle_{k,h} A(x, t, u, Du) \rangle_{\mathbb{R}^N} \, dt \\
&\quad + \frac{1}{2}\eta^2 \big[T_{q,K}''(|\triangle_{k,h} u(x, t)|)|\triangle_{k,h} u(x, t)| - T_{q,K}'(|\triangle_{k,h} u(x, t)|)\big] \\
&\qquad \times |\triangle_{k,h} u(x, t)|^{-3}|\langle \triangle_{k,h} u(x, t), \triangle_{k,h} H(x, t, Du) \rangle|^2 \, dt \\
&\quad + \frac{1}{2}\eta^2 T_{q,K}'(|\triangle_{k,h} u(x, t)|)|\triangle_{k,h} u(x, t)|^{-1}|\triangle_{k,h} H(x, t, Du)|^2 \, dt \\
&\quad + \eta^2 T_{q,K}'(|\triangle_{k,h} u(x, t)|)|\triangle_{k,h} u(x, t)|^{-1} \langle \triangle_{k,h} u(x, t), \triangle_{k,h} H(x, t, Du) dB_t \rangle_{\mathbb{R}^N}.
\end{aligned}
$$

In order to prove the assertion of the lemma, we start with a simple observation concerning the terms involving $\triangle_{k,h} H(x, t, Du)$. Taking into account the properties of the function $T_{q,K}$, see Lemma 3.11, we estimate

$$
\begin{aligned}
\big[T_{q,K}''(|\triangle_{k,h} u(x, t)|)&|\triangle_{k,h} u(x, t)| - T_{q,K}'(|\triangle_{k,h} u(x, t)|)\big]|\triangle_{k,h} u(x, t)|^{-3}|\langle \triangle_{k,h} u(x, t), \triangle_{k,h} H(x, t, Du) \rangle|^2 \\
&\leq 2(q - 1)T_{q,K}'(|\triangle_{k,h} u(x, t)|)|\triangle_{k,h} u(x, t)|^{-1}|\triangle_{k,h} H(x, t, Du)|^2 .
\end{aligned}
$$

We next introduce the abbreviation

$$V(\xi) := T_{q,K}'(|\xi|)|\xi|^{-1}\xi$$

for all $\xi \in \mathbb{R}^N$, and we note $|V(\xi)| = T_{q,K}'(|\xi|)$. Now we integrate over $x \in D$, and we then apply Fubini which is always allowed due to the truncation procedure, see Lemma 3.11 ii). Applying the integration

by parts formula, we hence obtain

$$\|(T_{q,K}|\triangle_{k,h}u(t)|)^{\frac{1}{2}}\eta\|^2_{L^2(D)} + \int_0^t \langle D\big(V(\triangle_{k,h}u(s))\eta^2\big), \triangle_{k,h}A(\cdot, s, u, Du)\rangle_{L^2(D)}\, ds$$

$$\leq \|(T_{q,K}|\triangle_{k,h}u_0|)^{\frac{1}{2}}\eta\|^2_{L^2(D)}$$

$$+ (q - 2^{-1})\int_0^t \big\|T'_{q,K}(|\triangle_{k,h}u(s)|)^{\frac{1}{2}}|\triangle_{k,h}u(s)|^{-\frac{1}{2}}\triangle_{k,h}H(\cdot, s, Du)\eta\big\|^2_{L^2(D)}\, ds$$

$$+ \int_0^t \langle V(\triangle_{k,h}u(x,s))\eta^2, \triangle_{k,h}H(\cdot, s, Du)\, dB_s\rangle_{L^2(D)}. \tag{4.12}$$

Now the second term on the left-hand side of this inequality shall be estimated. Using the decomposition introduced in (4.6) and applying Lemma 3.11, we first find for every $\varepsilon > 0$:

$$\langle D\big(V(\triangle_{k,h}u(s))\eta^2\big), \mathcal{A}(h)\rangle_{L^2(D)}$$

$$= \kappa^{-1}\langle D\big(V(\triangle_{k,h}u(s))\eta^2\big), D\triangle_{k,h}u\rangle_{L^2(D)} - \kappa^{-1}\langle D\big(V(\triangle_{k,h}u(x,s))\eta^2\big), D\triangle_{k,h}u - \kappa\mathcal{A}(h)\rangle_{L^2(D)}$$

$$\geq \kappa^{-1}\|D\big(V(\triangle_{k,h}u(s))\big) \cdot D\triangle_{k,h}u\eta^2\|_{L^1(D)} - 2\varepsilon\|T'_{q,K}(|\triangle_{k,h}u|)^{\frac{1}{2}}|\triangle_{k,h}u|^{-\frac{1}{2}}D\triangle_{k,h}u\eta\|^2_{L^2(D)}$$

$$- c(q, \kappa, \varepsilon)\||\triangle_{k,h}u|^q D\eta\|^2_{L^2(D)} - \kappa^{-1}(1 - \nu^2)^{\frac{1}{2}}\|D\big(V(\triangle_{k,h}u(s))\big)|D\triangle_{k,h}u|\eta^2\|_{L^1(D)}$$

$$\geq \big(\kappa^{-1}\mu^{\frac{1}{2}}(q) - \kappa^{-1}(1 - \nu^2)^{\frac{1}{2}} - 2\varepsilon\big)\|T'_{q,K}(|\triangle_{k,h}u|)^{\frac{1}{2}}|\triangle_{k,h}u|^{-\frac{1}{2}}D\triangle_{k,h}u\eta\|^2_{L^2(D)}$$

$$- c(q, \kappa, \varepsilon, \|D\eta\|_{L^\infty(D)})\|W_q(\triangle_{k,h}u)\|^2_{L^2(D')}.$$

We observe from the definition of $\mu(q)$ and the second bound in (4.11) on q that the factor $\mu^{\frac{1}{2}}(q) - (1 - \nu^2)^{\frac{1}{2}}$ appearing in the previous inequality is always strictly positive. Now, for the second term in the decomposition (4.6) we proceed analogously to (4.7); in fact, via Hölder's inequality, the inequalities $T''_{q,K}(t)t^2 \leq c(q)T'_{q,K}(t)t \leq c(q)T_{q,K}(t)$ on $\mathbb{R}^+$ and the Sobolev-Poincaré embedding (applied on every time-slice), we obtain

$$\big|\langle D\big(V(\triangle_{k,h}u(s))\eta^2\big), \mathcal{B}(h)\rangle_{L^2(D)}\big|$$

$$\leq c(L, q)\big(\|T'_{q,K}(|\triangle_{k,h}u|)^{\frac{1}{2}}|\triangle_{k,h}u|^{-\frac{1}{2}}D\triangle_{k,h}u\eta\|_{L^2(D)} + \|T'_{q,K}(|\triangle_{k,h}u|)^{\frac{1}{2}}|\triangle_{k,h}u|^{\frac{1}{2}}D\eta\|_{L^2(D)}\big)$$

$$\times \|T'_{q,K}(|\triangle_{k,h}u|)^{\frac{1}{2}}|\triangle_{k,h}u|^{-\frac{1}{2}}\mathcal{B}(h)\eta\|_{L^2(D)}$$

$$\leq c(L, q)\big(\|T'_{q,K}(|\triangle_{k,h}u|)^{\frac{1}{2}}|\triangle_{k,h}u|^{-\frac{1}{2}}D\triangle_{k,h}u\eta\|_{L^2(D)} + \|T_{q,K}(|\triangle_{k,h}u|)^{\frac{1}{2}}D\eta\|_{L^2(D)}\big)$$

$$\times \|T'_{q,K}(|\triangle_{k,h}u|)^{\frac{1}{2}}|\triangle_{k,h}u|^{-\frac{1}{2}}\triangle_{k,h}u\eta\|^\theta_{L^{\frac{2n}{n-2}}(D)}\|T'_{q,K}(|\triangle_{k,h}u|)^{\frac{1}{2}}|\triangle_{k,h}u|^{-\frac{1}{2}}\triangle_{k,h}u\eta\|^{1-\theta}_{L^2(D)}$$

$$\times \big(\|Du\|^{\frac{2}{n+2}}_{L^{\frac{2n}{(n+2)\theta}}(\text{spt}\,\eta)} + \|u\|^{\frac{2}{n}}_{L^{\frac{2}{\theta}}(\text{spt}\,\eta)} + \|f\|_{L^{\frac{n}{\theta}}(D)}\big)$$

$$\leq \big(\|T'_{q,K}(|\triangle_{k,h}u|)^{\frac{1}{2}}|\triangle_{k,h}u|^{-\frac{1}{2}}D\triangle_{k,h}u\eta\|^{1+\theta}_{L^2(D)} + \|T_{q,K}(|\triangle_{k,h}u|)^{\frac{1}{2}}D\eta\|^{1+\theta}_{L^2(D)}\big)$$

$$\times c(n, D, T, L, q)\|T_{q,K}(|\triangle_{k,h}u|)^{\frac{1}{2}}\eta\|^{1-\theta}_{L^2(D)}\big(\|Du\|^{\frac{2}{n+2}}_{L^{\frac{2n}{(n+2)\theta}}(\text{spt}\,\eta)} + \|u\|^{\frac{2}{n}}_{L^{\frac{2}{\theta}}(\text{spt}\,\eta)} + \|f\|_{L^{\frac{n}{\theta}}(D)}\big)$$

for every $\theta \in (0, 1)$. We now choose $\theta = \max\{p^{-1}n^2(n + 2)^{-2}, na^{-1}\}$, for which the last expression in brackets of the previous inequality is consequently bounded with probability one, according to the integrability assumptions on f, Du and the consequences on the integrability of u explained at the beginning of the proof. Young's inequality then implies

$$\big|\langle D\big(V(\triangle_{k,h}u(x,s))\eta^2\big), \mathcal{B}(h)\rangle_{L^2(D)}\big| \leq \varepsilon\|T'_{q,K}(|\triangle_{k,h}u|)^{\frac{1}{2}}|\triangle_{k,h}u|^{-\frac{1}{2}}D\triangle_{k,h}u\eta\|^2_{L^2(D)}$$

$$+ c(n, D, T, L, q, \|D\eta\|_{L^\infty(D)}, \varepsilon)\big(\|T_{q,K}(|\triangle_{k,h}u|)^{\frac{1}{2}}\eta\|^2_{L^2(D)} + 1\big)$$

$$\times \big(1 + \|Du\|^{2p\frac{n+2}{n}}_{L^{2p\frac{n+2}{n}}(\text{spt}\,\eta)} + \|u\|^{2p(\frac{n+2}{n})^2}_{L^{2p(\frac{n+2}{n})^2}(\text{spt}\,\eta)} + \|f\|^a_{L^a(D)}\big).$$

Finally, via the bounds for q in terms of n, p, a and ν, the last term in the decomposition involving $\mathcal{C}(h)$

is estimated with Young's inequality and the well-known estimates for finite difference quotients by

$$
\begin{aligned}
&\left|\left\langle D\big(V(\triangle_{k,h}u(s))\eta^2\big), \mathcal{C}(h)\right\rangle_{L^2(D)}\right| \\
&\qquad \leq c\big(\|T'_{q,K}(|\triangle_{k,h}u|)^{\frac{1}{2}}|\triangle_{k,h}u|^{-\frac{1}{2}}D\triangle_{k,h}u\eta\|_{L^2(D)} + \|T'_{q,K}(|\triangle_{k,h}u|)^{\frac{1}{2}}D\eta\|_{L^2(D)}\big) \\
&\qquad\qquad \times \|T'_{q,K}(|\triangle_{k,h}u|)^{\frac{1}{2}}|\triangle_{k,h}u|^{-\frac{1}{2}}\mathcal{C}(h)\eta\|_{L^2(D)} \\
&\qquad \leq \varepsilon\|T'_{q,K}(|\triangle_{k,h}u|)^{\frac{1}{2}}|\triangle_{k,h}u|^{-\frac{1}{2}}D\triangle_{k,h}u\eta\|_{L^2(D)}^2 \\
&\qquad\qquad + c\big(D,T,L,q,\|D\eta\|_{L^\infty(D)},\varepsilon\big)\big(1 + \|Du\|_{L^{2p\frac{n+2}{n}}(\mathrm{spt}\,\eta)}^{2p\frac{n+2}{n}} + \|u\|_{L^{2p(\frac{n+2}{n})^2}(\mathrm{spt}\,\eta)}^{2p(\frac{n+2}{n})^2} + \|f\|_{L^a(D)}^a\big),
\end{aligned}
$$

provided that $4q \leq a$. For the general case, one again has to argue more subtle, using the Sobolev embedding on time slices as for the term with $\mathcal{B}(h)$. With the analogous calculations as before this yields

$$
\begin{aligned}
&\left|\left\langle D\big(V(\triangle_{k,h}u(s))\eta^2\big), \mathcal{C}(h)\right\rangle_{L^2(D)}\right| \\
&\qquad \leq \varepsilon\|T'_{q,K}(|\triangle_{k,h}u|)^{\frac{1}{2}}|\triangle_{k,h}u|^{-\frac{1}{2}}D\triangle_{k,h}u\eta\|_{L^2(D)}^2 \\
&\qquad\qquad + c\big(n,D,T,L,q,\|D\eta\|_{L^\infty(D)},\varepsilon\big)\big(\|T_{q,K}(|\triangle_{k,h}u|)^{\frac{1}{2}}\eta\|_{L^2(D)}^2 + 1\big) \\
&\qquad\qquad\qquad \times \big(1 + \|Du\|_{L^{2p\frac{n+2}{n}}(\mathrm{spt}\,\eta)}^{2p\frac{n+2}{n}} + \|u\|_{L^{2p(\frac{n+2}{n})^2}(\mathrm{spt}\,\eta)}^{2p(\frac{n+2}{n})^2} + \|f\|_{L^a(D)}^a\big).
\end{aligned}
$$

It still remains to handle the second term on the right-hand side of inequality (4.12). With the assumptions (2.2) on H and Young's inequality, we easily find

$$
\begin{aligned}
&\big\|T'_{q,K}(|\triangle_{k,h}u(s)|)^{\frac{1}{2}}|\triangle_{k,h}u(s)|^{-\frac{1}{2}}\triangle_{k,h}H(\cdot,s,Du)\eta\big\|_{L^2(D)}^2 \\
&\qquad \leq (L_H^2 + \varepsilon)\|T'_{q,K}(|\triangle_{k,h}u|)^{\frac{1}{2}}|\triangle_{k,h}u|^{-\frac{1}{2}}D\triangle_{k,h}u\eta\|_{L^2(D)}^2 \\
&\qquad\qquad + c(L,L_H,\varepsilon)\big(1 + \|Du\|_{L^{2p\frac{n+2}{n}}(\mathrm{spt}\,\eta)}^{2p\frac{n+2}{n}} + \|f_H\|_{L^a(D)}^a\big).
\end{aligned}
$$

For every $s \in (0,T)$ we now define

$$
G''(u,f)(s) := \frac{2q}{c''}G_p(s) + 1 + \|Du\|_{L^{2p\frac{n+2}{n}}(D')}^{2p\frac{n+2}{n}} + \|u\|_{L^{2p(\frac{n+2}{n})^2}(D')}^{2p(\frac{n+2}{n})^2} + \|f\|_{L^a(D)}^a, \tag{4.13}
$$

which is a $L^1(0,T)$ with probability one. Furthermore, we set $G_q := c''G''(u,f)/(2q) \geq G_p$ which immediately gives $Y_q \leq Y_p$. Then, taking into account the smallness condition (4.11), choosing ε sufficiently small and combining the previous estimates for the various terms arising in (4.12), we find a preliminary (though still K-depending) pathwise estimate

$$
\begin{aligned}
&\|T_{q,K}(|\triangle_{k,h}u(t)|)^{\frac{1}{2}}\eta\|_{L^2(D)}^2 + c^{-1}(L_H,\kappa,\nu)\int_0^t \|T'_{q,K}(|\triangle_{k,h}u|)^{\frac{1}{2}}|\triangle_{k,h}u|^{-\frac{1}{2}}D\triangle_{k,h}u\eta\|_{L^2(D)}^2\,ds \\
&\qquad \leq \|T_{q,K}(|\triangle_{k,h}u_0|)^{\frac{1}{2}}\eta\|_{L^2(D)}^2 + c''\int_0^t\big(\|T_{q,K}(|\triangle_{k,h}u(t)|)^{\frac{1}{2}}\eta\|_{L^2(D)}^2 + 1\big)G''(u,f)\,ds \\
&\qquad\qquad + c\int_0^t \|f_H(s)\|_{L^a(D)}^a\,ds + \int_0^t \langle V(\triangle_{k,h}u(x,t))\eta^2, \triangle_{k,h}H(\cdot,s,Du)\,dB_s\rangle_{L^2(D)}.
\end{aligned}
$$

Step 2. We may now apply in a first step Itô's formula in exactly the same way as before in the derivation of estimate (4.10), and we this get

$$
\begin{aligned}
&e^{-\int_0^t c''G''(u,f)\,ds}\|T_{q,K}(|\triangle_{k,h}u(t)|)^{\frac{1}{2}}\eta\|_{L^2(D)}^2 \\
&\qquad + c^{-1}\int_0^t e^{-\int_0^s c''G''(u,f)\,d\tilde{s}}\big\|T'_{q,K}(|\triangle_{k,h}u|)^{\frac{1}{2}}|\triangle_{k,h}u|^{-\frac{1}{2}}D\triangle_{k,h}u\eta\big\|_{L^2(D)}^2\,ds \\
&\qquad\qquad \leq \|T_{q,K}(|\triangle_{k,h}u_0|)^{\frac{1}{2}}\eta\|_{L^2(D)}^2 + 1 + c\int_0^t \|f_H(s)\|_{L^a(D)}^a\,ds \\
&\qquad\qquad\qquad + c\int_0^t e^{-\int_0^s c''G''(u,f)\,d\tilde{s}}\langle V(\triangle_{k,h}u(x,t))\eta^2, \triangle_{k,h}H(\cdot,s,Du)\,dB_s\rangle_{L^2(D)}.
\end{aligned}
$$

Step 3. Similarly to the proof of Lemma 4.2, we introduce the random time

$$\tau_R := \inf\Big\{ t \in [0,T] : \int_0^t \big\|\,|\triangle_{k,h}u(s)|^{2q-1}\eta^2|\triangle_{k,h}H(\cdot,s,Du)|\big\|_{L^1(D)}^2\,ds > R \Big\}$$

with $\tau_R = T$ when the set is empty. Differently from Lemma 4.2, the property

$$P\Big(\int_0^T \big\|\,|\triangle_{k,h}u(s)|^{2q-1}\eta^2|\triangle_{k,h}H(\cdot,s,Du)|\big\|_{L^1(D)}^2\,ds < \infty \Big) = 1$$

which is needed to have $P(\lim_{R\to\infty}\tau_R = T) = 1$ is not clear a priori. We shall prove it a posteriori.

Notice that, by Lemma 3.11,

$$\int_0^{t\wedge\tau_R} e^{-2\int_0^s c''G''(u,f)\,d\tilde{s}}\Big(\int_D |V(\triangle_{k,h}u(x,s))|\eta^2|\triangle_{k,h}H(s,Du)|\,dx \Big)^2\,ds$$

$$\leq \int_0^{t\wedge\tau_R} \big\|\,|\triangle_{k,h}u(s)|^{2q-1}\eta^2|\triangle_{k,h}H(\cdot,s,Du)|\big\|_{L^1(D)}^2\,ds \leq R\,.$$

Step 3a. The last calculation shows that the stochastic integral from Step 2, stopped at τ_R, is a martingale (and thus it has zero expectation). Therefore (as in Lemma 4.2)

$$E\Big[e^{-\int_0^{t\wedge\tau_R} c''G''(u,f)\,ds}\|T_{q,K}(|\triangle_{k,h}u(t\wedge\tau_R)|)^{\frac{1}{2}}\eta\|_{L^2(D)}^2 \Big]$$

$$+ c^{-1}E\Big[\int_0^{t\wedge\tau_R} e^{-\int_0^s c''G''(u,f)\,d\tilde{s}}\big\|T'_{q,K}(|\triangle_{k,h}u|)^{\frac{1}{2}}|\triangle_{k,h}u|^{-\frac{1}{2}}D\triangle_{k,h}u\eta\big\|_{L^2(D)}^2\,ds \Big]$$

$$\leq \|T_{q,K}(|\triangle_{k,h}u_0|)^{\frac{1}{2}}\eta\|_{L^2(D)}^2 + 1 + cE\Big[\int_0^T \|f_H(s)\|_{L^a(D)}^a\,ds \Big].$$

At this stage we may pass to the limit $K \to \infty$ via Fatou's Lemma on the left-hand side and monotone convergence on the right-hand side, and we obtain

$$E\Big[e^{-\int_0^{t\wedge\tau_R} c''G''(u,f)\,ds}\|W_q(\triangle_{k,h}u(t\wedge\tau_R))\eta\|_{L^2(D)}^2 \Big]$$

$$+ c^{-1}E\Big[\int_0^{t\wedge\tau_R} e^{-\int_0^s c''G''(u,f)\,d\tilde{s}}\big\|\,|\triangle_{k,h}u(s)|^{q-1}D\triangle_{k,h}u(s)\eta\big\|_{L^2(D)}^2\,ds \Big]$$

$$\leq \|W_q(\triangle_{k,h}u_0)\eta\|_{L^2(D)}^2 + 1 + cE\Big[\int_0^T \|f_H(s)\|_{L^a(D)}^a\,ds \Big]. \tag{4.14}$$

Step 3b. Next we apply Burkholder-Davis-Gundy inequality to the inequality above from Step 2, stopped at τ_R and raised to the power $1/q$. Taking the limit $K \to \infty$ as in (4.14), we get

$$E\Big[\sup_{t\in[0,T]} e^{-\frac{1}{q}\int_0^{t\wedge\tau_R} c''G''(u,f)\,ds}\|W_q(\triangle_{k,h}u(t\wedge\tau_R))\eta\|_{L^2(D)}^{\frac{2}{q}} \Big]$$

$$\leq \|W_q(\triangle_{k,h}u_0)\eta\|_{L^2(D)}^{\frac{2}{q}} + 1 + cE\Big[\Big(\int_0^T \|f_H(s)\|_{L^a(D)}^a\,ds \Big)^{\frac{1}{q}} \Big]$$

$$+ CE\Big[\Big(\int_0^{T\wedge\tau_R} e^{-2\int_0^s c''G''(u,f)\,d\tilde{s}}\big\|\,|\triangle_{k,h}u(s)|^{2q-1}\eta^2|\triangle_{k,h}H(\cdot,s,Du)|\big\|_{L^1(D)}^2\,ds \Big)^{\frac{1}{2q}} \Big].$$

Since due to Hölder's inequality we have

$$\big\|\,|\triangle_{k,h}u(s)|^{2q-1}\eta^2|\triangle_{k,h}H(\cdot,s,Du)|\big\|_{L^1(D)}^2$$

$$\leq \big\|\,|\triangle_{k,h}u(s)|^q\eta\big\|_{L^2(D)}^2 \big\|\,|\triangle_{k,h}u(s)|^{q-1}\eta|\triangle_{k,h}H(\cdot,s,Du)|\big\|_{L^2(D)}^2\,,$$

the last term of the previous inequality, similarly to the proof of Lemma 4.2, is bounded by

$$CE\big[I_1^{1/2}I_2^{1/2}\big] \leq \frac{1}{2}E[I_1] + \frac{C^2}{2}E[I_2]$$

where

$$I_1 = \sup_{t\in[0,T]} e^{-\frac{1}{q}\int_0^{t\wedge\tau_R} c''G''(u,f)\,d\tilde{s}} \|W_q(\triangle_{k,h}u(t\wedge\tau_R))\eta\|_{L^2(D)}^{\frac{2}{q}},$$

$$I_2 = \Big(\int_0^{T\wedge\tau_R} e^{-\int_0^s c''G''(u,f)\,d\tilde{s}} \big\||\triangle_{k,h}u(s)|^{q-1}\eta|\triangle_{k,h}H(\cdot,s,Du)|\big\|_{L^2(D)}^2\,ds\Big)^{\frac{1}{q}}.$$

Hence, we have proved that

$$\frac{1}{2}E[I_1] \le \|W_q(\triangle_{k,h}u_0)\eta\|_{L^2(D)}^{\frac{2}{q}} + 1 + cE\Big[\Big(\int_0^T \|f_H(s)\|_{L^a(D)}^a\,ds\Big)^{\frac{1}{q}}\Big] + \frac{C^2}{2}E[I_2].$$

Now, by the assumptions (2.2) on H, Young's inequality and the bound on q (which in particular guarantees $q2a/(a-2)\le a$), we have

$$\frac{C^2}{2}E[I_2] \le CE\Big[\Big(\int_0^{T\wedge\tau_R} e^{-\int_0^s c''G''(u,f)\,d\tilde{s}} \big\||\triangle_{k,h}u(s)|^{q-1}D\triangle_{k,h}u(s)\eta\big\|_{L^2(D)}^2\,ds\Big)^{\frac{1}{q}}\Big]$$

$$+ C + CE\Big[\Big(\int_0^T \|f_H(s)\|_{L^a(D)}^a\,ds\Big)^{\frac{1}{q}}\Big] + \frac{1}{4}E[I_1]$$

$$+ CE\Big[\Big(\int_0^T e^{-\int_0^s c''G''(u,f)\,d\tilde{s}} \|W_q(Du(s))\|_{L^2(D')}^2\,ds\Big)^{\frac{1}{q}}\Big].$$

We observe that the last term remains bounded, due to the assumption of the lemma on the average and the choice of $G''(u,f)$ (which ensures that $c''G''(u,f) \ge 2qG_p$). Thus, by inequality (4.14) proved above, we find

$$\frac{1}{4}E[I_1] \le c\|W_q(\triangle_{k,h}u_0)\eta\|_{L^2(D)}^{\frac{2}{q}} + c + cE\Big[\int_0^T \|f_H(s)\|_{L^a(D)}^a\,ds\Big]^{\frac{1}{q}}$$

with a new constant.

Step 3c. In Step 3a and Step 3b we have almost proved the two bounds claimed by the lemma since the previous inequality along with (4.14) gives us

$$E\Big[\sup_{t\in[0,T]} e^{-\frac{1}{q}\int_0^{t\wedge\tau_R} c''G''(u,f)\,ds} \|W_q(\triangle_{k,h}u(t\wedge\tau_R))\eta\|_{L^2(D)}^{\frac{2}{q}}\Big]$$

$$+ E\Big[\Big(\int_0^{T\wedge\tau_R} e^{-\int_0^s c''G''(u,f)\,d\tilde{s}} \big\||\triangle_{k,h}u(s)|^{q-1}D\triangle_{k,h}u(s)\eta\big\|_{L^2(D)}^2\,ds\Big)^{\frac{1}{q}}\Big]$$

$$\le c\Big(\|W_q(\triangle_{k,h}u_0)\eta\|_{L^2(D)}^2 + c + cE\Big[\int_0^T \|f_H(s)\|_{L^a(D)}^a\,ds\Big]\Big)^{\frac{1}{q}}.$$

It now remains to justify (as already observed above) the limit $\tau_R \to T$ as $R \to \infty$ with probability one. Indeed, since $R \mapsto \tau_R$ is non-decreasing and bounded above by T, there exists the a.s. limit

$$\tau := \lim_{R\to\infty} \tau_R$$

and $\tau(\omega) \in [0,T]$. By Fatou's lemma and monotone convergence,

$$E\Big[\Big(\sup_{t\in[0,\tau]} e^{-\int_0^t c''G''(u,f)\,ds} \|W_q(\triangle_{k,h}u(t))\eta\|_{L^2(D)}^2$$

$$+ \int_0^\tau e^{-\int_0^s c''G''(u,f)\,d\tilde{s}} \big\||\triangle_{k,h}u(s)|^{q-1}D\triangle_{k,h}u(s)\eta\big\|_{L^2(D)}^2\,ds\Big)^{\frac{1}{q}}\Big]$$

is finite, hence the argument of the expectation is finite with probability one. Since $\int_0^T c''G''(u,f)\,ds$ is finite with probability one, we get

$$\sup_{t\in[0,\tau]} \|W_q(\triangle_{k,h}u(t))\eta\|_{L^2(D)}^2 + \int_0^\tau \big\||\triangle_{k,h}u(s)|^{q-1}D\triangle_{k,h}u(s)\eta\big\|_{L^2(D)}^2\,ds < \infty$$

with probability one. Thus (with the same inequalities used above)

$$\int_0^\tau \big\| |\triangle_{k,h} u(s)|^{2q-1} \eta^2 |\triangle_{k,h} H(\cdot, s, Du)| \big\|_{L^1(D)}^2 \, ds$$

$$\leq C \Big(1 + \sup_{t \in [0,\tau]} \| W_q(\triangle_{k,h} u(t)) \eta \|_{L^2(D)}^2 + \int_0^\tau \big\| |\triangle_{k,h} u(s)|^{q-1} D\triangle_{k,h} u(s) \eta \big\|_{L^2(D)}^2 \, ds$$

$$+ \int_0^T \| W_q(Du(s)) \|_{L^2(D')}^2 \, ds + \int_0^T \| f_H(s) \|_{L^a(D)}^a \, ds \Big)^2 < \infty$$

with probability one. If $\tau(\omega) < T$, by definition of τ_R we have

$$\int_0^\tau \big\| |\triangle_{k,h} u(s)|^{2q-1} \eta^2 |\triangle_{k,h} H(\cdot, s, Du)| \big\|_{L^1(D)}^2 \, ds = \infty$$

which is false, hence $P(\tau = T) = 1$. Having this basic fact, the same estimates just proved give us the result of the lemma, by taking into account the inequality $|DW(\triangle_{k,h} u)| \leq q |\triangle_{k,h} u|^{q-1} |D\triangle_{k,h} u|$ and the definition of G_q (and hence of Y_q) given after (4.13). □

5 Proof of the regularity result

Having the previous lemma at hand, we may now proceed to the main result of our paper.

Theorem 5.1. *Let $u \in V^2(D_T, \mathbb{R}^N)$ be a weak solution to the initial boundary value problem to (2.3) with initial values $u(\cdot, 0) = u_0(\cdot) \in W^{1,a-2}(D, \mathbb{R}^N)$. Assume further the assumptions (2.1) with $\nu > (n-2)/n$ such that*

$$L_H^2 < (L_H^*)^2(n) := \frac{2}{\kappa(n-1)} \Big(\Big[1 - \Big(\frac{n-2}{n} \Big)^2 \Big]^{\frac{1}{2}} - \Big[1 - \nu^2 \Big]^{\frac{1}{2}} \Big).$$

Then there exists $\alpha > 0$ depending only on n, ν and a such that for every subset $D_c \Subset D$ we have

$$P \big(\| u \|_{C^{0,\alpha}(D_c \times [0,T], \mathbb{R}^N)} < \infty \big) = 1.$$

Proof. For the proof of this result we want to apply Proposition 3.9. Therefore, the crucial point is to show higher integrability of Du for "great" powers with probability one, in order that hypothesis (3.6) of the proposition is satisfied. This will be established by an iterative improvement of the integrability exponent. We start by defining a sequence $(\tilde{q}_j)$ via

$$\tilde{q}_0 := 1,$$

$$\tilde{q}_{j+1} := \min \Big\{ \tilde{q}_j \frac{n+2}{n}, 1 + \tilde{q}_j \frac{n+2}{n} \frac{a-4}{a}, \frac{a-2}{2}, \tilde{q}_j + 1 \Big\} \text{ for } j \geq 1.$$

Before defining a further sequence (q_j) in order to perform the iteration, we make some observations on $L_H^*(s)$ as a function in $s \in [n, 2/(1-\nu)]$ (we note that $L_H^*(2q)$ already appeared in hypothesis (4.11) which gave an upper bound for q in the iteration). Clearly, $L_H^*(s)$ is continuous and strictly decreasing in s, with $L_H^*(2/(1-\nu)) = 0$.

We now set $q_j = \tilde{q}_j$ as long as $\tilde{q}_j > \tilde{q}_{j-1}$ and $L_H^*(2\tilde{q}_j) > L_H$, and for the first index j which doesn't satisfy these assumptions any more we set $q_j = q^*$ for a number $q^* > n/2$ (which is determined below). In what follows we shall denote this set of indices by $J \subset N_0$. We first study some properties of the sequence $\tilde{q}$ and give a definition of the final member q^* of the sequence $(q_j)_{j \in J}$: the first and the forth term in the rewritten formula for $\tilde{q}$ are strictly increasing in j and diverge for $j \to \infty$, whereas the the monotonicity properties of the second term depend on both the values of a and the size of q_j. More precisely, if $a \geq 2(n+2)$, then the second term increases with j and diverges for $j \to \infty$, but for every $a \in (n+2, 2(n+2))$ it increases only up to $q_{max}(a,n) = na/(4(n+2) - 2a) > n/2$. Observing $L_H^*(n) > L_H$ by assumption and $L_H^*(2/(1-\nu)) = 0$, by continuity of $L_H^*(s)$ we hence can determine a number $\tilde{s} \in (n, 2/(1-\nu)]$ such that $L_H^*(\tilde{s}) = L_H$. We can thus define

$$q^* := \text{arbitrary number in } \Big(\frac{n}{2}, \min \Big\{ \frac{\tilde{s}}{2}, \frac{a-2}{2}, q_{max} \Big\} \Big).$$

It is easy to calculate that this number q^* is reached after a finite number of steps (depending only on n, ν, a and the difference $q_{max} - q^*$ (in the sense that the number of steps diverges as $q^* \nearrow q_{max}$), hence $|J| < \infty$, i.e. $(q_j)_{j \in J}$ is a finite sequence.

We are now going to establish by induction that for every $j \in J$ we have

$$
(i) \quad \sup_{|h| < \mathrm{dist}(D_j, \partial D_{j-1})} E\big[\big\|Y_{q_j}^{q_j} W_{q_j}(\triangle_{k,h} u)\big\|_{V^2(D_j \times (0,T))}^{2/q_j}\big] \leq C_j \text{ for all } k \in \{1, \ldots, n\},
$$

$$
(ii) \quad \sup_{|h| < \mathrm{dist}(D_j, \partial D_{j-1})} E\big[\big\|Y_{q_j}^{q_j} W_{q_j}(\triangle_{k,h} u)\big\|_{L^{2\frac{n+2}{n}}(D_j \times (0,T))}^{2/q_j}\big] \leq c(n, D_j) C_j \text{ for all } k \in \{1, \ldots, n\},
$$

$$
(iii) \quad E\big[\big\|Y_{q_j}^{q_j} W_{q_j}(Du)\big\|_{L^{2\frac{n+2}{n}}(D_j \times (0,T))}^{2/q_j}\big] \leq \widetilde{C}_j,
$$

$$
(iv) \quad Du \in L^\infty(0,T; L^{2q_j}(D_j, \mathbb{R}^{nN})) \text{ with probability one.}
$$

Here $(Y_{q_j})_{j \in J}$ is a sequence of random variables given by $Y_{q_j}(t, \omega) = \exp(-\int_0^t G_{q_j}(s, \omega)\, ds)$ for each $j \in J$, for a sequence of functions $(G_{q_j})_{j \in J}$ which are in $L^1(0,T)$ with probability one and which will be determined later, and $(D_j)_{j \in J}$ is a monotone decreasing sequence of open sets satisfying $D_c \subset D_j \subset D_{j-1} \subset \ldots \subset D_0 \subset D_{-1} = D$.

We start by setting
$$
Y_1 := e^{-\frac{1}{2}\int_0^t c' G'(u,f)\, ds},
$$
where $G'(u,f)$ was defined in (4.8). It is obvious from its definition that $Y_1 \colon [0,T] \times \Omega \to (0,1]$ satisfies $P(\inf_{t \in [0,T]} Y_1 > 0) = 1$. We then observe from Lemma 4.2 that

$$
\sup_{|h| < d} E\Big[\sup_{t \in (0,T)} \|Y_1 \triangle_{k,h} u\|_{L^2(D_0)}^2 + \int_0^T \|Y_1 D\triangle_{k,h} u\|_{L^2(D_0)}^2\, dt\Big]
$$
$$
\leq c'\Big(\|D_k u_0\|_{L^2(D)}^2 + 1 + E\big[\|f_H\|_{L^{\frac{2a}{a-2}}(D_T)}^{\frac{2a}{a-2}}\big]\Big) =: C_0
$$

is satisfied for every open set D_0 compactly supported in D. By definition of the space V^2, this establishes the statement $(i)_0$. Furthermore, $(ii)_0$ follows immediately from the Sobolev embedding (2.4), applied for P-almost every ω to the functions $Y_1 \triangle_{k,h} u$, for $k \in \{1, \ldots, n\}$. To conclude the first step of the iteration it only remains to justify the statements $(iii)_0$ and $(iv)_0$. To this end we take advantage of Theorem 3.6 twice, in the way as explained in Remark 3.7 (and actually as already performed in Corollary 4.4). First we apply it with the choices $p = q = 2q_0(n+2)/n$ to the inequality from $(ii)_0$ (for all $k \in \{1, \ldots, n\}$), leading to the existence of Du in the Lebesgue space $L^{2(n+2)/n}(D_0 \times (0,T), \mathbb{R}^{nN})$ with the required estimate for the average of $Y_1 Du$; secondly, we apply it with the choice $p = 2q_0$ and $q = \infty$ to $(i)_0$ – more precisely to the first term in the V^2-norm – and, keeping in mind the pathwise strict positivity of Y_1, we end up with the existence of Du in $L^\infty(0,T; L^2(D_0, \mathbb{R}^{nN}))$ with probability one.

We now proceed to the inductive step. Assume for a given $j \in J$ that $(i)_\ell$–$(iv)_\ell$ are valid on open sets $D_\ell \subset D_{\ell-1}$ with random variables $Y_{q_\ell} \colon [0,T] \times \Omega \to (0,1]$ of the required form for all $\ell \in \{0, \ldots, j-1\}$. Then, keeping in mind $(iii)_{j-1}$ and the definition of the number q^*, we note that the assumptions of Lemma 4.5 are satisfied (for p, D' replaced by q_{j-1}, D_{j-1}), and we hence deduce (with the admissible choice $q = q_j$) the estimate

$$
\sup_{|h| < d_j} E\Big[\Big(\sup_{t \in (0,T)} \|Y_{q_j}^{q_j} W_{q_j}(\triangle_{k,h} u)\|_{L^2(D_j)}^2 + \int_0^T \|Y_{q_j}^{q_j} D W_{q_j}(\triangle_{k,h} u)\|_{L^2(D_j)}^2\, dt\Big)^{\frac{1}{q_j}}\Big]
$$
$$
\leq c\Big(\|W_{q_{j-1}}(D_k u(x,0))\|_{L^2(D)}^2 + 1 + E\big[\|f_H(s)\|_{L^a(D_T)}^a\big]\Big)^{\frac{1}{q_j}} =: C_j
$$

for every $k \in \{1, \ldots, n\}$, a domain $D_j \subset D_{j-1}$ satisfying $d_j := \mathrm{dist}(D_j, \partial D_{j-1}) > 0$ and a random variable Y_{q_j} defined via G_{q_j} given in Lemma 4.5 and satisfying in particular $P(\inf_{t \in [0,T]} Y_{q_j} > 0) = 1$. This shows $(i)_j$, and $(ii)_j$ in turn is an immediate consequence after the application of the Sobolev embedding as above. Moreover, the statements $(iii)_j$ and $(iv)_j$ again follow from $(ii)_j$ and $(i)_j$, respectively, after the application of Theorem 3.6 with the choices $p = q = 2q_j(n+2)/n$ and $p = 2q_j, q = \infty$, respectively. This finishes the proof of the induction.

As an immediate consequence of the induction, we can now conclude the desired higher integrability result to a great power, via the following observation. Via (iv) we find in the limit

$$Du \in L^\infty(0, T; L^{2q^*}(D_c, \mathbb{R}^{nN}))$$

with probability one, and by definition the exponent $2q^*$ is greater than the space dimension n. Hence, assumption (3.6) of Proposition 3.9 is guaranteed. For its application we still need to check the integrability condition on $a(x, s), b(x, s)$ given by

$$a(x, s) := \operatorname{div} A(x, s, u, Du) \qquad \text{and} \qquad b(x, s) := H(x, s, Du) \,.$$

Since $A(x, t, u, z)$ is differentiable in x, u, and z with bounds (2.1), we obtain $a \in L^2(D_c \times (0, T), \mathbb{R}^N)$ with probability one as a direct consequence of $Du \in V^2(D_c \times (0, T), \mathbb{R}^{nN})$ and $f \in L^a(D_T) \subset L^4(D_T)$. Furthermore, the growth of H according to (2.2) with $f_H \in L^a(D_T \times \Omega)$ implies $b \in L^{2+\varepsilon}(0, T; L^2(D_c, \mathbb{R}^{n'N}))$ with probability one. Thus, Proposition 3.9 yields the asserted Hölder continuity of u with probability one and finishes the proof of the theorem. $\qquad\qquad\square$

References

[1] D. G. Aronson and J. Serrin, *Local behavior of solutions of quasilinear parabolic equations*, Arch. Rational Mech. Anal. **25** (1967), 81–122.

[2] S. Attanasio and F. Flandoli, *Renormalized solutions for stochastic transport equations and the regularization by bilinear multiplicative noise*, Commun. Partial Differ. Equations (to appear), arXiv:1007.4102.

[3] S. Campanato, *Equazioni paraboliche del secondo ordine e spazi $\mathcal{L}^{2,\theta}(\Omega, \delta)$*, Ann. Mat. Pura Appl. (4) **73** (1966), 55–102.

[4] S. Campanato, *On the nonlinear parabolic systems in divergence form. Hölder continuity and partial Hölder continuity of the solutions*, Ann. Mat. Pura Appl. (4) **137** (1984), 83–122.

[5] S. Campanato, *Elliptic systems with non-linearity q greater or equal to two. Regularity of the solution of the Dirichlet problem*, Ann. Mat. Pura Appl. Ser. 4 **147** (1987), 117–150.

[6] H. O. Cordes, *Über die erste Randwertaufgabe bei quasilinearen Differentialgleichungen zweiter Ordnung in mehr als zwei Variablen*, Math. Ann. **131** (1956), 278–312.

[7] G. Da Prato and J. Zabczyk, *Stochastic equations in infinite dimensions*, Encyclopedia of Mathematics and its Applications, vol. 44, Cambridge University Press, Cambridge, 1992.

[8] E. De Giorgi, *Sulla differenziabilità e l'analiticità delle estremali degli integrali multipli regolari*, Mem. Accad. Sci. Torino Ser. III **3** (1957), 25–43.

[9] E. De Giorgi, *Un esempio di estremali discontinue per un problema variazionale di tipo ellittico*, Boll. Unione Mat. Ital., IV. **1** (1968), 135–137.

[10] L. Denis, A. Matoussi, and L. Stoica, *Maximum principle and comparison theorem for quasi-linear stochastic PDE's*, Electron. J. Probab. **14** (2009), no. 19, 500–530.

[11] E. DiBenedetto, *Degenerate parabolic equations*, Universitext. New York, NY: Springer-Verlag. xv, 387 p., 1993.

[12] F. Duzaar and G. Mingione, *Second order parabolic systems, optimal regularity, and singular sets of solutions*, Ann. Inst. Henri Poincaré Anal. Non Linéaire **22** (2005), no. 6, 705–751.

[13] F. Flandoli, *Random perturbation of PDEs and fluid dynamic models*, Saint Flour summer school lectures 2010, Lecture Notes in Math., vol. **2015**, Springer, Berlin, 2011, p. 176.

[14] F. Flandoli, M. Gubinelli, and E. Priola, *Well-posedness of the transport equation by stochastic perturbation*, Invent. Math. **180** (2010), no. 1, 1–53.

[15] F. Flandoli, M. Gubinelli, and E. Priola, *Full well-posedness of point vortex dynamics corresponding to stochastic 2D Euler equations*, Stoch. Proc. Appl. **121** (2011), no. 7, 1445–1463.

[16] M. Giaquinta and E. Giusti, *Partial regularity for the solutions to nonlinear parabolic systems*, Ann. Mat. Pura Appl. (4) **97** (1973), 253–266.

[17] M. Giaquinta and G. Modica, *Almost-everywhere regularity results for solutions of non linear elliptic systems*, Manuscr. Math. **28** (1979), 109–158.

[18] E. Giusti and M. Miranda, *Sulla regolarità delle soluzioni deboli di una classe di sistemi ellitici quasi-lineari*, Arch. Ration. Mech. Anal. **31** (1968), 173–184.

[19] E. Giusti and M. Miranda, *Un esempio di soluzioni discontinue per un problema di minimo relativo ad un integrale regolare del calcolo delle variazioni*, Boll. Unione Mat. Ital., IV. Ser. **1** (1968), 219–226.

[20] A. V. Ivanov, *The Harnack inequality for generalized solutions of second order quasilinear parabolic equations*, Trudy Mat. Inst. Steklov. **102** (1967), 51–84.

[21] P.-A. Ivert, *Regularitätsuntersuchungen von Lösungen elliptischer Systeme von quasilinearen Differentialgleichungen zweiter Ordnung*, Manuscr. Math. **30** (1979), 53–88.

[22] E. Kalita, *On the Hölder continuity of solutions of nonlinear parabolic systems*, Commentat. Math. Univ. Carol. **35** (1994), no. 4, 675–680.

[23] A. Koshelev, *Regularity of the solutions of quasilinear elliptic systems*, Uspekhi Mat. Nauk **33** (1978), no. 4(202).

[24] A. Koshelev, *Regularity of solutions for some quasilinear parabolic systems*, Math. Nachr. **162** (1993), 59–88.

[25] A. Koshelev, *Regularity problem for quasilinear elliptic and parabolic systems*, Lecture Notes in Math., vol. **1614**, Springer, Berlin, 1995.

[26] N. V. Krylov and B. L. Rozovskii, *Stochastic evolution equations*, Current problems in mathematics, Vol. 14 (Russian), Akad. Nauk SSSR, Vsesoyuz. Inst. Nauchn. i Tekhn. Informatsii, Moscow, 1979, pp. 71–146, 256.

[27] M. Kurihara, *On a Harnack inequality for nonlinear parabolic equations*, Publ. Res. Inst. Math. Sci. Ser. A **3** (1967/1968), 211–241.

[28] G. Mingione, *Regularity of minima: an invitation to the Dark Side of the Calculus of Variations*, Appl. Math. **51** (2006), no. 4, 355–425.

[29] J. Moser, *A new proof of De Giorgi's theorem concerning the regularity problem for elliptic differential equations*, Commun. Pure Appl. Math. **13** (1960), 457–468.

[30] J. Nash, *Continuity of solutions of parabolic and elliptic equations*, Amer. J. Math. **80** (1958), 931–954.

[31] J. Necas and V. Šverák, *On regularity of solutions of nonlinear parabolic systems*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **18** (1991), no. 1, 1–11.

[32] B. Øksendal, *Stochastic differential equations. an introduction with applications. 5th ed.*, Universitext. Berlin: Springer. xix, 324 p., 1998.

[33] E. Pardoux, *Equations aux dérivées partielles stochastiques non linéaires monotones. Etude de solutions fortes de type Itô*, Ph.D. thesis.

[34] D. Revuz and M. Yor, *Continuous martingales and Brownian motion*, second ed., Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. **293**, Springer-Verlag, Berlin, 1994.

[35] B.L. Rozovskii, *Stochastic evolution systems. Linear theory and applications to non-linear filtering*, Mathematics and Its Applications (Soviet Series), 35. Dordrecht etc.: Kluwer Academic Publishers. xviii, 315 p., 1990.

[36] B. Schmalfuss, *Qualitative properties for the stochastic Navier-Stokes equation*, Nonlinear Anal. **28** (1997), no. 9, 15451563.

[37] J. Stará and O. John, *Some (new) counterexamples of parabolic systems*, Commentat. Math. Univ. Carol. **36** (1995), no. 3, 503–510.

[38] K. Uhlenbeck, *Regularity for a class of nonlinear elliptic systems*, Acta Math. **138** (1977), 219–240.