Available online at www.sciencedirect.com
ScienceDirect
Stochastic Processes and their Applications 123 (2013) 4294–4336
stochastic processes and their applications
www.elsevier.com/locate/spa

Degenerate parabolic stochastic partial differential equations

Martina Hofmanová*

Department of Mathematical Analysis, Faculty of Mathematics and Physics, Charles University, Sokolovská 83, 186 75 Praha 8, Czech Republic
ENS Cachan Bretagne, IRMAR, CNRS, UEB, av. Robert Schuman, 35 170 Bruz, France
Institute of Information Theory and Automation of the ASCR, Pod Vodárenskou věží 4, 182 08 Praha 8, Czech Republic

Received 7 February 2012; received in revised form 26 June 2013; accepted 30 June 2013
Available online 11 July 2013

Abstract

We study the Cauchy problem for a scalar semilinear degenerate parabolic partial differential equation with stochastic forcing. In particular, we are concerned with the well-posedness in any space dimension. We adapt the notion of kinetic solution which is well suited for degenerate parabolic problems and supplies a good technical framework to prove the comparison principle. The proof of existence is based on the vanishing viscosity method: the solution is obtained by a compactness argument as the limit of solutions of nondegenerate approximations.

Keywords: Degenerate parabolic stochastic partial differential equation; Kinetic solution

1. Introduction

In this paper, we study the Cauchy problem for a scalar semilinear degenerate parabolic partial differential equation with stochastic forcing

$$\begin{aligned} &\mathrm{d}u + \operatorname{div}\big(B(u)\big)\mathrm{d}t = \operatorname{div}\big(A(x)\nabla u\big)\mathrm{d}t + \Phi(u)\,\mathrm{d}W, \quad x \in \mathbb{T}^N,\ t \in (0,T),\\ &u(0) = u_0, \end{aligned} \tag{1}$$

* Correspondence to: ENS Cachan Bretagne, av. Robert Schuman, 35 170 Bruz, France. Tel.: +33 2 99 05 93 41.
E-mail address: martina.hofmanova@bretagne.ens-cachan.fr.

0304-4149/$ - see front matter
http://dx.doi.org/10.1016/j.spa.2013.06.015

where W is a cylindrical Wiener process. Equations of this type are widely used in fluid mechanics since they model the phenomenon of convection–diffusion of an ideal fluid in porous media. Namely, the important applications including for instance two or three-phase flows can be found in petroleum engineering or in hydrogeology. For a thorough exposition of this area given from a practical point of view we refer the reader to [15] and to the references cited therein.

The aim of the present paper is to establish the well-posedness theory for solutions of the Cauchy problem (1) in any space dimension. Towards this end, we adapt the notion of kinetic formulation and kinetic solution which has already been studied in the case of hyperbolic scalar conservation laws in both deterministic (see e.g. [20,23,24,26], or [27] for a general presentation) and stochastic setting (see [9]), and also in the case of deterministic degenerate parabolic equations of second-order (see [5]). To the best of our knowledge, in the degenerate case, stochastic equations of type (1) have not been studied yet, neither by means of kinetic formulation nor by any other approach.

The concept of kinetic solution was first introduced by Lions, Perthame, Tadmor in [24] for deterministic scalar conservation laws and applies to more general situations than the one of entropy solution as considered for example in [4,12,21]. Moreover, it appears to be better suited particularly for degenerate parabolic problems since it allows us to keep the precise structure of the parabolic dissipative measure, whereas in the case of the entropy solution part of this information is lost and has to be recovered at some stage. This technique also supplies a good technical framework to prove the L^1-comparison principle which allows to prove uniqueness. Nevertheless, kinetic formulation can be derived only for smooth solutions, hence the classical result [17] giving L^p-valued solutions for the nondegenerate case has to be improved (see [18,12]).

In the case of hyperbolic scalar conservation laws, Debussche and Vovelle [9] defined a notion of generalized kinetic solution and obtained a comparison result showing that any generalized kinetic solution is actually a kinetic solution. Accordingly, the proof of existence was simplified since only weak convergence of approximate viscous solutions was necessary.

The situation is quite different in the case of parabolic scalar conservation laws. Indeed, due to the parabolic term, the approach of [9] is not applicable: the comparison principle can be proved only for kinetic solutions (not generalized ones) and therefore strong convergence of approximate solutions is needed in order to prove the existence. Moreover, the proof of the comparison principle itself is much more delicate than in the hyperbolic case.

We note that an important step in the proof of existence, identification of the limit of an approximating sequence of solutions, is based on a new general method of constructing martingale solutions of SPDEs (see Propositions 4.14, 4.15 and the sequel), that does not rely on any kind of martingale representation theorem and therefore holds independent interest especially in situations where these representation theorems are no longer available. First applications were already done in [3,25] and, in the finite-dimensional case, also in [19]. In the present work, this method is further generalized as the martingales to be dealt with are only defined for almost all times.

The exposition is organized as follows. In Section 2 we review the basic setting and define the notion of kinetic solution. Section 3 is devoted to the proof of uniqueness. We first establish a technical Proposition 3.2 which then turns out to be the keystone in the proof of the comparison principle in Theorem 3.3. We next turn to the proof of existence in Sections 4 and 5. First of all, in Section 4, we make an additional hypothesis upon the initial condition and employ the vanishing viscosity method. In particular, we study certain nondegenerate problems and establish suitable uniform estimates for the corresponding sequence of approximate solutions. The compactness argument then yields the existence of a martingale kinetic solution which together with the pathwise uniqueness gives the desired kinetic solution (defined on the original stochastic basis).

In Section 5, the existence of a kinetic solution is shown for general initial data. In the Appendix, we formulate and prove an auxiliary result concerning densely defined martingales.

2. Notation and the main result

We now give the precise assumptions on each of the terms appearing in Eq. (1). We work on a finite-time interval $[0, T]$, $T > 0$, and consider periodic boundary conditions: $x \in \mathbb{T}^N$ where $\mathbb{T}^N$ is the N-dimensional torus. The flux function

$$B = (B_1, \dots, B_N) : \mathbb{R} \longrightarrow \mathbb{R}^N$$

is supposed to be of class C^1 with a polynomial growth of its derivative, which is denoted by $b = (b_1, \dots, b_N)$. The diffusion matrix

$$A = (A_{ij})_{i,j=1}^N : \mathbb{T}^N \longrightarrow \mathbb{R}^{N\times N}$$

is of class C^∞, symmetric and positive semidefinite. Its square-root matrix, which is also symmetric and positive semidefinite, is denoted by σ.

Regarding the stochastic term, let $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geq 0}, \mathbb{P})$ be a stochastic basis with a complete, right-continuous filtration. Let $\mathcal{P}$ denote the predictable σ-algebra on $\Omega \times [0, T]$ associated to $(\mathscr{F}_t)_{t\geq 0}$. The initial datum may be random in general, i.e. $\mathscr{F}_0$-measurable, and we assume $u_0 \in L^p(\Omega; L^p(\mathbb{T}^N))$ for all $p \in [1, \infty)$. The process W is a cylindrical Wiener process: $W(t) = \sum_{k\geq 1} \beta_k(t)e_k$ with $(\beta_k)_{k\geq 1}$ being mutually independent real-valued standard Wiener processes relative to $(\mathscr{F}_t)_{t\geq 0}$ and $(e_k)_{k\geq 1}$ a complete orthonormal system in a separable Hilbert space $\mathfrak{U}$. In this setting, we can assume, without loss of generality, that the σ-algebra $\mathscr{F}$ is countably generated and $(\mathscr{F}_t)_{t\geq 0}$ is the filtration generated by the Wiener process and the initial condition. For each $z \in L^2(\mathbb{T}^N)$ we consider a mapping $\Phi(z) : \mathfrak{U} \to L^2(\mathbb{T}^N)$ defined by $\Phi(z)e_k = g_k(\cdot, z(\cdot))$. In particular, we suppose that $g_k \in C(\mathbb{T}^N \times \mathbb{R})$ and the following conditions

$$G^2(x,\xi) = \sum_{k\geq 1} \big|g_k(x,\xi)\big|^2 \leq L\big(1 + |\xi|^2\big), \tag{2}$$

$$\sum_{k\geq 1} \big|g_k(x,\xi) - g_k(y,\zeta)\big|^2 \leq L\big(|x-y|^2 + |\xi - \zeta| h(|\xi - \zeta|)\big), \tag{3}$$

are fulfilled for every $x, y \in \mathbb{T}^N$, $\xi, \zeta \in \mathbb{R}$, where h is a continuous nondecreasing function on $\mathbb{R}_+$ satisfying, for some $\alpha > 0$,

$$h(\delta) \leq C\delta^\alpha, \quad \delta < 1. \tag{4}$$

The conditions imposed on Φ, particularly assumption (2), imply that

$$\Phi : L^2(\mathbb{T}^N) \longrightarrow L_2(\mathfrak{U}; L^2(\mathbb{T}^N)),$$

where $L_2(\mathfrak{U}; L^2(\mathbb{T}^N))$ denotes the collection of Hilbert–Schmidt operators from $\mathfrak{U}$ to $L^2(\mathbb{T}^N)$. Thus, given a predictable process $u \in L^2(\Omega; L^2(0, T; L^2(\mathbb{T}^N)))$, the stochastic integral $t \mapsto \int_0^t \Phi(u)\mathrm{d}W$ is a well defined process taking values in $L^2(\mathbb{T}^N)$ (see [8] for detailed construction).

Finally, define the auxiliary space $\mathfrak{U}_0 \supset \mathfrak{U}$ via

$$\mathfrak{U}_0 = \left\{ v = \sum_{k\geq 1} \alpha_k e_k;\ \sum_{k\geq 1} \frac{\alpha_k^2}{k^2} < \infty \right\},$$

endowed with the norm

$$\|v\|_{\mathfrak{U}_0}^2 = \sum_{k\geq 1} \frac{\alpha_k^2}{k^2}, \qquad v = \sum_{k\geq 1} \alpha_k e_k.$$

Note that the embedding $\mathfrak{U} \hookrightarrow \mathfrak{U}_0$ is Hilbert–Schmidt. Moreover, trajectories of W are $\mathbb{P}$-a.s. in $C([0,T];\mathfrak{U}_0)$ (see [8]).

In the present paper, we use the brackets $\langle\cdot,\cdot\rangle$ to denote the duality between the space of distributions over $\mathbb{T}^N\times\mathbb{R}$ and $C_c^\infty(\mathbb{T}^N\times\mathbb{R})$. We denote similarly the integral

$$\langle F,G\rangle = \int_{\mathbb{T}^N}\int_{\mathbb{R}} F(x,\xi)G(x,\xi)\,\mathrm{d}x\,\mathrm{d}\xi, \quad F\in L^p(\mathbb{T}^N\times\mathbb{R}),\ G\in L^q(\mathbb{T}^N\times\mathbb{R}),$$

where $p,q\in[1,\infty]$ are conjugate exponents. The differential operators of gradient ∇, divergence div and Laplacian Δ are always understood with respect to the space variable x.

As the next step, we introduce the kinetic formulation of (1) as well as the basic definitions concerning the notion of kinetic solution. The motivation for this approach is given by the nonexistence of a strong solution and, on the other hand, the nonuniqueness of weak solutions, even in simple cases. The idea is to establish an additional criterion – the kinetic formulation – which is automatically satisfied by any strong solution to (1) and which permits to ensure the well-posedness.

Definition 2.1 (*Kinetic Measure*)**.** A mapping m from Ω to the set of nonnegative finite measures over $\mathbb{T}^N\times[0,T]\times\mathbb{R}$ is said to be a kinetic measure provided

(i) m is measurable in the following sense: for each $\psi\in C_0(\mathbb{T}^N\times[0,T]\times\mathbb{R})$ the mapping $m(\psi):\Omega\to\mathbb{R}$ is measurable,

(ii) m vanishes for large ξ: if $B_R^c=\{\xi\in\mathbb{R};\ |\xi|\geq R\}$ then

$$\lim_{R\to\infty}\mathbb{E}\,m\big(\mathbb{T}^N\times[0,T]\times B_R^c\big)=0, \tag{5}$$

(iii) for any $\psi\in C_0(\mathbb{T}^N\times\mathbb{R})$

$$\int_{\mathbb{T}^N\times[0,t]\times\mathbb{R}}\psi(x,\xi)\,\mathrm{d}m(x,s,\xi)\in L^2(\Omega\times[0,T])$$

admits a predictable representative.[1]

Definition 2.2 (*Kinetic Solution*)**.** Assume that, for all $p\in[1,\infty)$,

$$u\in L^p(\Omega\times[0,T],\mathcal{P},\mathrm{d}\mathbb{P}\otimes\mathrm{d}t;L^p(\mathbb{T}^N))$$

and

(i) there exists $C_p>0$ such that

$$\mathbb{E}\operatorname*{ess\,sup}_{0\leq t\leq T}\|u(t)\|_{L^p(\mathbb{T}^N)}^p\leq C_p, \tag{6}$$

(ii) $\sigma\,\nabla u\in L^2(\Omega\times[0,T];L^2(\mathbb{T}^N))$.

Let n_1 be a mapping from Ω to the set of nonnegative finite measures over $\mathbb{T}^N\times[0,T]\times\mathbb{R}$ defined for any Borel set $D\in\mathcal{B}(\mathbb{T}^N\times[0,T]\times\mathbb{R})$ as[2]

[1] Throughout the paper, the term *representative* stands for an element of a class of equivalence.

[2] We will write shortly $\mathrm{d}n_1(x,t,\xi)=|\sigma(x)\nabla u|^2\,\mathrm{d}\delta_{u(x,t)}(\xi)\,\mathrm{d}x\,\mathrm{d}t$.

$$n_1(D)=\int_{\mathbb{T}^N\times[0,T]}\left[\int_{\mathbb{R}}\mathbf{1}_D(x,t,\xi)\,\mathrm{d}\delta_{u(x,t)}(\xi)\right]\big|\sigma(x)\nabla u\big|^2\,\mathrm{d}x\,\mathrm{d}t,\qquad \mathbb{P}\text{-a.s.,}\tag{7}$$

and let

$$f=\mathbf{1}_{u>\xi}:\Omega\times\mathbb{T}^N\times[0,T]\times\mathbb{R}\longrightarrow\mathbb{R}.$$

Then u is said to be a kinetic solution to (1) with initial datum u_0 provided there exists a kinetic measure $m\geq n_1$ a.s., such that the pair $(f=\mathbf{1}_{u>\xi},m)$ satisfies, for all $\varphi\in C_c^\infty(\mathbb{T}^N\times[0,T)\times\mathbb{R})$, $\mathbb{P}$-a.s.,

$$\begin{aligned}&\int_0^T\big\langle f(t),\partial_t\varphi(t)\big\rangle\mathrm{d}t+\big\langle f_0,\varphi(0)\big\rangle+\int_0^T\big\langle f(t),b(\xi)\cdot\nabla\varphi(t)\big\rangle\mathrm{d}t\\&\quad+\int_0^T\big\langle f(t),\mathrm{div}\big(A(x)\nabla\varphi(t)\big)\big\rangle\mathrm{d}t\\&=-\sum_{k\geq1}\int_0^T\int_{\mathbb{T}^N}g_k\big(x,u(x,t)\big)\varphi\big(x,t,u(x,t)\big)\mathrm{d}x\,\mathrm{d}\beta_k(t)\\&\quad-\frac{1}{2}\int_0^T\int_{\mathbb{T}^N}G^2\big(x,u(x,t)\big)\partial_\xi\varphi\big(x,t,u(x,t)\big)\mathrm{d}x\,\mathrm{d}t+m(\partial_\xi\varphi).\end{aligned}\tag{8}$$

Remark 2.3. We emphasize that a kinetic solution is, in fact, a class of equivalence in $L^p(\Omega\times[0,T],\mathcal{P},\mathrm{d}\mathbb{P}\otimes\mathrm{d}t;L^p(\mathbb{T}^N))$ so not necessarily a stochastic process in the usual sense. Nevertheless, it will be seen later (see Corollary 3.4) that, in this class of equivalence, there exists a representative with good continuity properties, namely, $u\in C([0,T];L^p(\mathbb{T}^N))$, $\mathbb{P}$-a.s., and therefore, it can be regarded as a stochastic process.

Remark 2.4. Let us also make an observation which clarifies the point (ii) in the above definition: if $u\in L^2(\Omega\times[0,T];L^2(\mathbb{T}^N))$ then it can be shown that $\sigma\,\nabla u$ is well defined in $L^2(\Omega\times[0,T];H^{-1}(\mathbb{T}^N))$ since the square-root matrix σ belongs to $W^{1,\infty}(\mathbb{T}^N)$ according to [14,28].

By $f=\mathbf{1}_{u>\xi}$ we understand a real function of four variables, where the additional variable ξ is called velocity. In the deterministic case, i.e. corresponding to the situation $\Phi=0$, Eq. (8) in the above definition is the so-called kinetic formulation of (1)

$$\partial_t\mathbf{1}_{u>\xi}+b(\xi)\cdot\nabla\mathbf{1}_{u>\xi}-\mathrm{div}\big(A(x)\nabla\mathbf{1}_{u>\xi}\big)=\partial_\xi m$$

where the unknown is the pair $(\mathbf{1}_{u>\xi},m)$ and it is solved in the sense of distributions over $\mathbb{T}^N\times[0,T)\times\mathbb{R}$. In the stochastic case, we write formally[3]

$$\partial_t\mathbf{1}_{u>\xi}+b(\xi)\cdot\nabla\mathbf{1}_{u>\xi}-\mathrm{div}\big(A(x)\nabla\mathbf{1}_{u>\xi}\big)=\delta_{u=\xi}\,\Phi(u)\dot{W}+\partial_\xi\left(m-\frac{1}{2}G^2\delta_{u=\xi}\right).\tag{9}$$

It will be seen later that this choice is reasonable since for any u being a strong solution to (1) the pair $(\mathbf{1}_{u>\xi},n_1)$ satisfies (8) and consequently u is a kinetic solution to (1). The measure n_1

[3] Hereafter, we employ the notation which is commonly used in papers concerning the kinetic solutions to conservation laws and write $\delta_{u=\xi}$ for the Dirac measure centered at $u(x,t)$.

relates to the diffusion term in (1) and so is called parabolic dissipative measure. It gives us better regularity of solutions in the nondegeneracy zones of the diffusion matrix A which is exactly what one would expect according to the theory of (nondegenerate) parabolic SPDEs. Indeed, for the case of a nondegenerate diffusion matrix A, i.e. when the second order term defines a strongly elliptic differential operator, the kinetic solution u belongs to $L^2(\Omega; L^2(0,T; H^1(\mathbb{T}^N)))$ (cf. Definition 2.2(ii)). Thus, the measure $n_2 = m - n_1$ which takes account of possible singularities of solution vanishes in the nondegenerate case.

We now derive the kinetic formulation in the case of a sufficiently smooth u satisfying (1), namely, $u \in C([0,T]; C^2(\mathbb{T}^N))$, $\mathbb{P}$-a.s. Note, that also in this case, the measure n_2 vanishes. For almost every $x \in \mathbb{T}^N$, we aim at finding the stochastic differential of $\theta(u(x,t))$, where $\theta \in C^\infty(\mathbb{R})$ is an arbitrary test function. Such a method can be performed by the Itô formula since

$$\begin{aligned} u(x,t) &= u_0(x) - \int_0^t \operatorname{div}\big(B(u(x,s))\big)\,\mathrm{d}s + \int_0^t \operatorname{div}\big(A(x)\nabla u(x,s)\big)\,\mathrm{d}s \\ &\quad + \sum_{k\geq 1}\int_0^t g_k\big(x,u(x,s)\big)\,\mathrm{d}\beta_k(s), \quad \text{a.e. } (\omega,x)\in\Omega\times\mathbb{T}^N,\ \forall t\in[0,T]. \end{aligned} \tag{10}$$

In the following we denote by $\langle\cdot,\cdot\rangle_\xi$ the duality between the space of distributions over $\mathbb{R}$ and $C_c^\infty(\mathbb{R})$. Fix $x \in \mathbb{T}^N$ such that (10) holds true and consider $\mathbf{1}_{u(x,t)>\xi}$ as a (random) distribution on $\mathbb{R}$. Then

$$\langle \mathbf{1}_{u(x,t)>\xi}, \theta'\rangle_\xi = \int_{\mathbb{R}} \mathbf{1}_{u(x,t)>\xi}\theta'(\xi)\,\mathrm{d}\xi = \theta(u(x,t))$$

and the application of the Itô formula yields:

$$\begin{aligned} \mathrm{d}\langle \mathbf{1}_{u(x,t)>\xi}, \theta'\rangle_\xi &= \theta'(u(x,t))\Big[-\operatorname{div}\big(B(u(x,t))\big)\mathrm{d}t + \operatorname{div}\big(A(x)\nabla u(x,t)\big)\mathrm{d}t \\ &\quad + \sum_{k\geq 1} g_k(x,u(x,t))\,\mathrm{d}\beta_k(t)\Big] + \frac{1}{2}\theta''(u(x,t))G^2(u(x,t))\mathrm{d}t. \end{aligned}$$

Afterwards, we proceed term by term and employ the fact that all the necessary derivatives of u exist as functions

$$\begin{aligned} \theta'(u(x,t))\operatorname{div}\big(B(u(x,t))\big) &= \theta'(u(x,t))b(u(x,t))\cdot\nabla u(x,t) \\ &= \operatorname{div}\left(\int_{-\infty}^{u(x,t)} b(\xi)\theta'(\xi)\mathrm{d}\xi\right) = \operatorname{div}\big(\langle b\mathbf{1}_{u(x,t)>\xi}, \theta'\rangle_\xi\big) \\ \theta'(u(x,t))\operatorname{div}\big(A(x)\nabla u(x,t)\big) &= \sum_{i,j=1}^N \partial_{x_i}\big[A_{ij}(x)\theta'(u(x,t))\partial_{x_j}u(x,t)\big] \\ &\quad - \sum_{i,j=1}^N \theta''(u(x,t))\partial_{x_i}u(x,t)A_{ij}(x)\partial_{x_j}u(x,t) \\ &= \sum_{i,j=1}^N \partial_{x_i}\left(A_{ij}(x)\partial_{x_j}\int_{-\infty}^{u(x,t)}\theta'(\xi)\mathrm{d}\xi\right) \\ &\quad + \big\langle \partial_\xi n_1(x,t), \theta'\big\rangle_\xi \\ &= \operatorname{div}\Big(A(x)\nabla\langle \mathbf{1}_{u(x,t)>\xi}, \theta'\rangle_\xi\Big) + \big\langle \partial_\xi n_1(x,t), \theta'\big\rangle_\xi \end{aligned}$$

$$\theta'(u(x,t))g_k(x,u(x,t)) = \langle g_k(x,\xi)\delta_{u(x,t)=\xi},\theta'\rangle_\xi$$
$$\begin{aligned}\theta''(u(x,t))G^2(x,u(x,t)) &= \langle G^2(x,\xi)\delta_{u(x,t)=\xi},\theta''\rangle_\xi \\ &= -\big\langle \partial_\xi(G^2(x,\xi)\delta_{u(x,t)=\xi}),\theta'\big\rangle_\xi.\end{aligned}$$

Note, that according to the definition of the parabolic dissipative measure (7) it makes sense to write $\partial_\xi n_1(x,t)$, i.e. for fixed x,t we regard $n_1(x,t)$ as a random measure on $\mathbb{R}$: for any Borel set $D_1 \in \mathcal{B}(\mathbb{R})$

$$n_1(x,t,D_1) = \big|\sigma(x)\nabla u(x,t)\big|^2\delta_{u(x,t)}(D_1), \quad \mathbb{P}\text{-a.s.}$$

In the following, we distinguish between two situations. In the first case, we intend to use test functions independent on t. We set $\theta(\xi) = \int_{-\infty}^{\xi}\varphi_1(\zeta)\,\mathrm{d}\zeta$ for some test function $\varphi_1 \in C_c^\infty(\mathbb{R})$ and test the above against $\varphi_2 \in C^\infty(\mathbb{T}^N)$. Since linear combinations of the test functions $\psi(x,\xi) = \varphi_1(\xi)\varphi_2(x)$ form a dense subset of $C_c^\infty(\mathbb{T}^N\times\mathbb{R})$ we obtain for any $\psi \in C_c^\infty(\mathbb{T}^N\times\mathbb{R}),\ t\in[0,T],\ \mathbb{P}$-a.s.,

$$\begin{aligned}&\big\langle f(t),\psi\big\rangle - \big\langle f_0,\psi\big\rangle - \int_0^t\big\langle f(s),b(\xi)\cdot\nabla\psi\big\rangle\,\mathrm{d}s - \int_0^t\big\langle f(s),\operatorname{div}\big(A(x)\nabla\psi\big)\big\rangle\,\mathrm{d}s\\ &= \int_0^t\big\langle\delta_{u=\xi}\,\Phi(u)\,\mathrm{d}W,\psi\big\rangle + \frac12\int_0^t\big\langle\delta_{u=\xi}\,G^2,\partial_\xi\psi\big\rangle\,\mathrm{d}s - \big\langle n_1,\partial_\xi\psi\big\rangle([0,t)),\end{aligned}$$

where $\big\langle n_1,\partial_\xi\psi\big\rangle([0,t))_1\big(\partial_\xi\psi\mathbf{1}_{[0,t)}\big)$. In order to allow test functions from $C_c^\infty(\mathbb{T}^N\times[0,T)\times\mathbb{R})$, take $\varphi_3 \in C_c^\infty([0,T))$ and apply the Itô formula to calculate the stochastic differential of the product $\langle f(t),\psi\rangle\varphi_3(t)$. We have, $\mathbb{P}$-a.s.,

$$\begin{aligned}&\big\langle f(t),\psi\big\rangle\varphi_3(t) - \big\langle f_0,\psi\big\rangle\varphi_3(0) - \int_0^t\big\langle f(s),b(\xi)\cdot\nabla\psi\big\rangle\varphi_3(s)\,\mathrm{d}s\\ &\quad - \int_0^t\big\langle f(s),\operatorname{div}\big(A(x)\nabla\psi\big)\big\rangle\varphi_3(s)\,\mathrm{d}s\\ &= \int_0^t\big\langle\delta_{u=\xi}\,\Phi(u)\varphi_3(s)\,\mathrm{d}W,\psi\big\rangle + \frac12\int_0^t\big\langle\delta_{u=\xi}\,G^2,\partial_\xi\psi\big\rangle\varphi_3(s)\,\mathrm{d}s\\ &\quad - n_1\big(\partial_\xi\psi\mathbf{1}_{[0,t)}\varphi_3\big) + \int_0^t\big\langle f(s),\psi\big\rangle\partial_s\varphi_3(s)\,\mathrm{d}s.\end{aligned}$$

Evaluating this process at $t = T$ and setting $\varphi(x,t,\xi) = \psi(x,\xi)\varphi_3(t)$ yield Eq. (8); hence $f = \mathbf{1}_{u>\xi}$ is a distributional solution to the kinetic formulation (9) with $n_2 = 0$. Therefore any strong solution of (1) is a kinetic solution in the sense of Definition 2.2.

Concerning the point (ii) in Definition 2.2, it was already mentioned in Remark 2.4 that $\sigma\,\nabla u$ is well defined in $L^2(\Omega\times[0,T];H^{-1}(\mathbb{T}^N))$. As we assume more in Definition 2.2(ii) we obtain the following chain rule formula, which will be used in the proof of Theorem 3.3,

$$\sigma\,\nabla f = \sigma\,\nabla u\,\delta_{u=\xi} \quad \text{in } \mathcal{D}'(\mathbb{T}^N\times\mathbb{R}),\ \text{a.e. } (\omega,t)\in\Omega\times[0,T]. \tag{11}$$

It is a consequence of the next result.

Lemma 2.5. *Assume that $v \in L^2(\mathbb{T}^N)$ and $\sigma\,(\nabla v) \in L^2(\mathbb{T}^N)$. If $g = \mathbf{1}_{v>\xi}$ then it holds true*

$$\sigma\,\nabla g = \sigma\,\nabla v\,\delta_{v=\xi} \quad \textit{in } \mathcal{D}'(\mathbb{T}^N\times\mathbb{R}).$$

Proof. In order to prove this claim, we denote by σ^i the ith row of σ. Let us fix test functions $\psi_1 \in C^\infty(\mathbb{T}^N)$, $\psi_2 \in C_c^\infty(\mathbb{R})$ and define $\theta(\xi) = \int_{-\infty}^{\xi} \psi_2(\zeta)\,\mathrm{d}\zeta$. We denote by $\langle\cdot,\cdot\rangle_x$ the duality between the space of distributions over $\mathbb{T}^N$ and $C^\infty(\mathbb{T}^N)$. It holds

$$\begin{aligned}\big\langle \sigma^i\,\nabla g, \psi_1\psi_2\big\rangle &= -\left\langle \operatorname{div}(\sigma^i\psi_1), \int_{-\infty}^{v} \psi_2(\xi)\,\mathrm{d}\xi \right\rangle_x = -\big\langle \operatorname{div}\big(\sigma^i\psi_1\big), \theta(v)\big\rangle_x \\ &= \big\langle \sigma^i\,\nabla\theta(v), \psi_1\big\rangle_x.\end{aligned}$$

If the following was true

$$\sigma^i\,\nabla\theta(v) = \theta'(v)\,\sigma^i\,\nabla v \quad \text{in } \mathcal{D}'(\mathbb{T}^N), \tag{12}$$

we would obtain

$$\big\langle \sigma^i\,\nabla g, \psi_1\psi_2\big\rangle = \big\langle \theta'(v)\,\sigma^i\,\nabla v, \psi_1\big\rangle_x = \big\langle \sigma^i\,\nabla v\,\delta_{v=\xi}, \psi_1\psi_2\big\rangle$$

and the proof would be complete.

Hence it remains to verify (12). Towards this end, let us consider an approximation to the identity on $\mathbb{T}^N$, denoted by (ϱ_τ). To be more precise, let $\tilde{\varrho} \in C_c^\infty(\mathbb{R}^N)$ be a nonnegative symmetric function satisfying $\int_{\mathbb{R}^N} \tilde{\varrho} = 1$ and $\operatorname{supp}\tilde{\varrho} \subset B(0, 1/2)$. This function can be easily extended to become $\mathbb{Z}^N$-periodic, and let this modification be denoted by $\bar{\varrho}$. Now it is correct to define $\varrho = \bar{\varrho} \circ q^{-1}$, where q denotes the quotient mapping $q : \mathbb{R}^N \to \mathbb{T}^N = \mathbb{R}^N/\mathbb{Z}^N$, and finally

$$\varrho_\tau(x) = \frac{1}{\tau^N}\,\varrho\left(\frac{x}{\tau}\right).$$

Since the identity (12) is fulfilled by any sufficiently regular v, let us consider v^τ, the mollifications of v given by (ϱ_τ). We have

$$\sigma^i\,\nabla\theta(v^\tau) \longrightarrow \sigma^i\,\nabla\theta(v) \quad \text{in } \mathcal{D}'(\mathbb{T}^N).$$

In order to obtain convergence of the corresponding right hand sides, i.e.

$$\theta'(v^\tau)\,\sigma^i\,\nabla v^\tau \longrightarrow \theta'(v)\,\sigma^i\,\nabla v \quad \text{in } \mathcal{D}'(\mathbb{T}^N),$$

we employ similar arguments as in the commutation lemma of DiPerna and Lions (see [10, Lemma II.1]). Namely, since $\sigma^i\,(\nabla v) \in L^2(\mathbb{T}^N)$ it is approximated in $L^2(\mathbb{T}^N)$ by its mollifications $[\sigma^i\,\nabla v]^\tau$. Consequently,

$$\theta'(v^\tau)\big[\sigma^i\,\nabla v\big]^\tau \longrightarrow \theta'(v)\,\sigma^i\,\nabla v \quad \text{in } \mathcal{D}'(\mathbb{T}^N).$$

Thus, it is enough to show that

$$\theta'(v^\tau)\Big(\sigma^i\,\nabla v^\tau - \big[\sigma^i\,\nabla v\big]^\tau\Big) \longrightarrow 0 \quad \text{in } \mathcal{D}'(\mathbb{T}^N). \tag{13}$$

It holds

$$\begin{aligned}&\sigma^i(x)\nabla v^\tau(x) - \big[\sigma^i\,\nabla v\big]^\tau(x) \\ &\quad = \int_{\mathbb{T}^N} v(y)\,\sigma^i(x)(\nabla\varrho_\tau)(x-y)\,\mathrm{d}y + \int_{\mathbb{T}^N} v(y)\operatorname{div}_y\big(\sigma^i(y)\varrho_\tau(x-y)\big)\,\mathrm{d}y \\ &\quad = -\int_{\mathbb{T}^N} v(y)\big(\sigma^i(y) - \sigma^i(x)\big)(\nabla\varrho_\tau)(x-y)\mathrm{d}y + \int_{\mathbb{T}^N} v(y)\operatorname{div}\big(\sigma^i(y)\big)\varrho_\tau(x-y)\mathrm{d}y.\end{aligned}$$

The second term on the right hand side is the mollification of $v\mathrm{div}\sigma^i \in L^2(\mathbb{T}^N)$ hence converges in $L^2(\mathbb{T}^N)$ to $v\mathrm{div}\sigma^i$. We will show that the first term converges in $L^1(\mathbb{T}^N)$ to $-v\mathrm{div}\sigma^i$. Since $\tau|\nabla\varrho_\tau|(\cdot) \leq C\varrho_{2\tau}(\cdot)$ with a constant independent on τ, we obtain the following estimate

$$\left\| \int_{\mathbb{T}^N} v(y)\big(\sigma^i(y)-\sigma^i(x)\big)(\nabla\varrho_\tau)(x-y)\,\mathrm{d}y \right\|_{L^2(\mathbb{T}^N)} \leq C\|\sigma^i\|_{W^{1,\infty}(\mathbb{T}^N)}\|v\|_{L^2(\mathbb{T}^N)}.$$

Due to this estimate, it is sufficient to consider v and σ^i smooth and the general case can be concluded by a density argument. We infer[4]

$$\begin{aligned}
&-\int_{\mathbb{T}^N} v(y)\big(\sigma^i(y)-\sigma^i(x)\big)(\nabla\varrho_\tau)(x-y)\,\mathrm{d}y\\
&\quad= -\frac{1}{\tau^{N+1}}\int_{\mathbb{T}^N}\int_0^1 v(y)\,\mathrm{D}\sigma^i\big(x+r(y-x)\big)(y-x)\cdot(\nabla\varrho)\Big(\frac{x-y}{\tau}\Big)\,\mathrm{d}r\,\mathrm{d}y\\
&\quad= \int_{\mathbb{T}^N}\int_0^1 v(x-\tau z)\,\mathrm{D}\sigma^i(x-r\tau z)z\cdot(\nabla\varrho)(z)\,\mathrm{d}r\,\mathrm{d}z\\
&\quad\longrightarrow v(x)\,\mathrm{D}\sigma^i(x):\int_{\mathbb{T}^N} z\otimes(\nabla\varrho)(z)\,\mathrm{d}z,\qquad \forall x\in\mathbb{T}^N.
\end{aligned}$$

Integration by parts now yields

$$\int_{\mathbb{T}^N} z\otimes(\nabla\varrho)(z)\,\mathrm{d}z = -\mathrm{Id} \tag{14}$$

hence

$$v(x)\,\mathrm{D}\sigma^i(x):\int_{\mathbb{T}^N} z\otimes(\nabla\varrho)(z)\,\mathrm{d}z = -v(x)\mathrm{div}\big(\sigma^i(x)\big),\qquad \forall x\in\mathbb{T}^N,$$

and the convergence in $L^1(\mathbb{T}^N)$ follows by the Vitali convergence theorem from the above estimate. Employing the Vitali convergence theorem again, we obtain (13) and consequently also (12) which completes the proof. □

We proceed by two related definitions, which will be useful especially in the proof of uniqueness.

Definition 2.6 (*Young Measure*)**.** Let (X,λ) be a finite measure space. A mapping ν from X to the set of probability measures on $\mathbb{R}$ is said to be a Young measure if, for all $\psi\in C_b(\mathbb{R})$, the map $z\mapsto\nu_z(\psi)$ from X into $\mathbb{R}$ is measurable. We say that a Young measure ν vanishes at infinity if, for all $p\geq 1$,

$$\int_X\int_{\mathbb{R}} |\xi|^p\,\mathrm{d}\nu_z(\xi)\,\mathrm{d}\lambda(z) < \infty.$$

Definition 2.7 (*Kinetic Function*)**.** Let (X,λ) be a finite measure space. A measurable function $f: X\times\mathbb{R}\to[0,1]$ is said to be a kinetic function if there exists a Young measure ν on X vanishing at infinity such that, for λ-a.e. $z\in X$, for all $\xi\in\mathbb{R}$,

$$f(z,\xi) = \nu_z(\xi,\infty).$$

[4] By $:$ we denote the component-wise inner product of matrices and by $\otimes$ the tensor product.

Remark 2.8. Note, that if f is a kinetic function then $\partial_\xi f = -\nu$ for λ-a.e. $z \in X$. Similarly, let u be a kinetic solution of (1) and consider $f = \mathbf{1}_{u>\xi}$. We have $\partial_\xi f = -\delta_{u=\xi}$, where $\nu = \delta_{u=\xi}$ is a Young measure on $\Omega \times \mathbb{T}^N \times [0, T]$. Therefore, the expression (8) can be rewritten in the following form: for all $\varphi \in C_c^\infty(\mathbb{T}^N \times [0, T) \times \mathbb{R})$, $\mathbb{P}$-a.s.,

$$\begin{aligned}
&\int_0^T \big\langle f(t), \partial_t\varphi(t)\big\rangle \mathrm{d}t + \big\langle f_0, \varphi(0)\big\rangle + \int_0^T \big\langle f(t), b(\xi)\cdot\nabla\varphi(t)\big\rangle \mathrm{d}t \\
&\quad + \int_0^T \big\langle f(t), \operatorname{div}\big(A(x)\nabla\varphi(t)\big)\big\rangle \mathrm{d}t \\
&= -\sum_{k\geq 1}\int_0^T\int_{\mathbb{T}^N}\int_{\mathbb{R}} g_k(x,\xi)\varphi(x,t,\xi)\mathrm{d}\nu_{x,t}(\xi)\,\mathrm{d}x\,\mathrm{d}\beta_k(t) \\
&\quad - \frac{1}{2}\int_0^T\int_{\mathbb{T}^N}\int_{\mathbb{R}} G^2(x,\xi)\partial_\xi\varphi(x,t,\xi)\mathrm{d}\nu_{x,t}(\xi)\,\mathrm{d}x\,\mathrm{d}t + m(\partial_\xi\varphi).
\end{aligned} \tag{15}$$

For a general kinetic function f with corresponding Young measure ν, the above formulation leads to the notion of generalized kinetic solution as used in [9]. Although this concept is not established here, the notation will be used throughout the paper, i.e. we will often write $\nu_{x,t}(\xi)$ instead of $\delta_{u(x,t)=\xi}$.

Lemma 2.9. *Let (X, λ) be a finite measure space such that $L^1(X)$ is separable.*[5] *Let $\{f_n;\ n \in \mathbb{N}\}$ be a sequence of kinetic functions on $X \times \mathbb{R}$, i.e. $f_n(z, \xi) = \nu_z^n(\xi, \infty)$ where ν^n are Young measures on X. Suppose that, for some $p \geq 1$,*

$$\sup_{n\in\mathbb{N}}\int_X\int_{\mathbb{R}} |\xi|^p \mathrm{d}\nu_z^n(\xi)\,\mathrm{d}\lambda(z) < \infty.$$

Then there exists a kinetic function f on $X \times \mathbb{R}$ and a subsequence still denoted by $\{f_n;\ n \in \mathbb{N}\}$ such that

$$f_n \overset{w^*}{\longrightarrow} f, \quad \text{in } L^\infty(X\times\mathbb{R})\text{-weak}^*.$$

Proof. The proof can be found in [9, Corollary 6]. □

To conclude this section we state the main result of the paper.

Theorem 2.10. *Let $u_0 \in L^p(\Omega; L^p(\mathbb{T}^N))$, for all $p \in [1, \infty)$. Under the above assumptions, there exists a unique kinetic solution to the problem (1) and it has almost surely continuous trajectories in $L^p(\mathbb{T}^N)$, for all $p \in [1, \infty)$. Moreover, if u_1, u_2 are kinetic solutions to (1) with initial data $u_{1,0}$ and $u_{2,0}$, respectively, then for all $t \in [0, T]$*

$$\mathbb{E}\|u_1(t) - u_2(t)\|_{L^1(\mathbb{T}^N)} \leq \mathbb{E}\|u_{1,0} - u_{2,0}\|_{L^1(\mathbb{T}^N)}.$$

3. Uniqueness

We begin with the question of uniqueness. Due to the following proposition, we obtain an auxiliary property of kinetic solutions, which will be useful later on in the proof of the comparison principle in Theorem 3.3.

---

[5] According to [7, Proposition 3.4.5], it is sufficient to assume that the corresponding σ-algebra is countably generated.

Proposition 3.1 (*Left- and Right-Continuous Representatives*)**.** *Let u be a kinetic solution to (1). Then $f = \mathbf{1}_{u>\xi}$ admits representatives f^- and f^+ which are almost surely left- and right-continuous, respectively, at all points $t^* \in [0,T]$ in the sense of distributions over $\mathbb{T}^N \times \mathbb{R}$. More precisely, for all $t^* \in [0,T]$ there exist kinetic functions $f^{*,\pm}$ on $\Omega \times \mathbb{T}^N \times \mathbb{R}$ such that setting $f^\pm(t^*) = f^{*,\pm}$ yields $f^\pm = f$ almost everywhere and*

$$\big\langle f^\pm(t^* \pm \varepsilon), \psi\big\rangle \longrightarrow \big\langle f^\pm(t^*), \psi\big\rangle \quad \varepsilon \downarrow 0 \ \forall \psi \in C_c^2(\mathbb{T}^N \times \mathbb{R}) \ \mathbb{P}\text{-a.s.}$$

Moreover, $f^+ = f^-$ for all $t^* \in [0,T]$ except for some at most countable set.

Proof. As the space $C_c^2(\mathbb{T}^N \times \mathbb{R})$ (endowed with the topology of the uniform convergence on any compact set of functions and their first and second derivatives) is separable, let us fix a countable dense subset $\mathcal{D}_1$. Let $\psi \in \mathcal{D}_1$ and $\alpha \in C_c^1([0,T))$ and set $\varphi(x,t,\xi) = \psi(x,\xi)\alpha(t)$. Integration by parts and the stochastic version of Fubini's theorem applied to (15) yield

$$\int_0^T g_\psi(t)\alpha'(t)\mathrm{d}t + \langle f_0, \psi\rangle\alpha(0) = \langle m, \partial_\xi \psi\rangle(\alpha) \quad \mathbb{P}\text{-a.s.}$$

where

$$\begin{aligned} g_\psi(t) = \big\langle f(t), \psi\big\rangle &- \int_0^t \big\langle f(s), b(\xi)\cdot\nabla\psi\big\rangle \mathrm{d}s - \int_0^t \big\langle f(t), \operatorname{div}\big(A(x)\nabla\psi\big)\big\rangle \mathrm{d}s \\ &- \sum_{k\geq 1}\int_0^t \int_{\mathbb{T}^N}\int_{\mathbb{R}} g_k(x,\xi)\psi(x,\xi)\,\mathrm{d}\nu_{x,s}(\xi)\,\mathrm{d}x\,\mathrm{d}\beta_k(s) \\ &- \frac{1}{2}\int_0^t \int_{\mathbb{T}^N}\int_{\mathbb{R}} \partial_\xi\psi(x,\xi)G^2(x,\xi)\,\mathrm{d}\nu_{x,s}(\xi)\,\mathrm{d}x\,\mathrm{d}s. \end{aligned} \tag{16}$$

Hence $\partial_t g_\psi$ is a (pathwise) Radon measure on $[0,T]$ and by the Riesz representation theorem $g_\psi \in BV([0,T])$. Due to the properties of BV-functions [2, Theorem 3.28], we obtain that g_ψ admits left- and right-continuous representatives which coincide except for an at most countable set. Moreover, apart from the first one all terms in (16) are almost surely continuous in t. Hence, on a set of full measure, denoted by Ω_ψ, $\langle f, \psi\rangle$ also admits left- and right-continuous representatives which coincide except for an at most countable set. Let us denote them by $\langle f, \psi\rangle^\pm$ and set $\Omega_0 = \cap_{\psi\in\mathcal{D}_1}\Omega_\psi$. Note, that as $\mathcal{D}_1$ is countable, Ω_0 is also a set of full measure. Besides, for $\psi \in \mathcal{D}_1$, $(t,\omega) \mapsto \langle f(t,\omega), \psi\rangle^+$ has right continuous trajectories in time and is thus measurable with respect to (t,ω).

For $\psi \in C_c^2(\mathbb{T}^N \times \mathbb{R})$, we define $\langle f(t,\omega), \psi\rangle^+$ on $[0,T] \times \Omega_0$ as the limit of $\langle f(t,\omega), \psi_n\rangle^+$ for any sequence (ψ_n) in $\mathcal{D}_1$ converging to ψ. Then clearly $\langle f(\cdot,\cdot), \psi\rangle^+$ is also measurable in (t,ω) and has right continuous trajectories.

It is now straightforward to define f^+ by $\langle f^+, \psi\rangle = \langle f, \psi\rangle^+$. Then $f^+ : \Omega \times [0,T] \to L^\infty(\mathbb{T}^N \times \mathbb{R})$. Moreover, seen as a function $f^+ : \Omega \times [0,T] \to L^p_{\mathrm{loc}}(\mathbb{T}^N \times \mathbb{R})$, for some $p \in [1,\infty)$, it is weakly measurable and therefore measurable. According to the Fubini theorem f^+, as a function of four variables ω, x, t, ξ, is measurable.

Besides, f^+ is a.s. right-continuous in the required sense. Next, we show that f^+ is a representative (in time) of f, i.e. for a.e. $t^* \in [0,T)$ it holds that $f^+(t^*) = f(t^*)$, where the equality is understood in the sense of classes of equivalence in ω, x, ξ. Indeed, due to the Lebesgue differentiation theorem,

$$\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\int_{t^*}^{t^*+\varepsilon} f(\omega,x,t,\xi)\,\mathrm{d}t = f(\omega,x,t^*,\xi) \quad \text{a.e. } (\omega,x,t^*,\xi);$$

hence by the dominated convergence theorem

$$\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\int_{t^*}^{t^*+\varepsilon}\big\langle f(t,\omega),\psi\big\rangle\,\mathrm{d}t=\big\langle f(t^*,\omega),\psi\big\rangle\quad \text{a.e. }(\omega,t^*),$$

for any $\psi\in C_c^2(\mathbb{T}^N\times\mathbb{R})$. Since this limit is equal to $\big\langle f^+(t^*,\omega),\psi\big\rangle$ for $t^*\in[0,T]$ and $\omega\in\Omega_0$, the conclusion follows.

Now, it only remains to show that $f^+(t^*)$ is a kinetic function on $\Omega\times\mathbb{T}^N$ for all $t^*\in[0,T)$. Towards this end, we observe that for all $t^*\in[0,T)$

$$f_n(x,t^*,\xi):=\frac{1}{\varepsilon_n}\int_{t^*}^{t^*+\varepsilon_n}f(x,t,\xi)\,\mathrm{d}t$$

is a kinetic function on $X=\Omega\times\mathbb{T}^N$ and by (6) the assumptions of Lemma 2.9 are fulfilled. Accordingly, there exists a kinetic function $f^{*,+}$ and a subsequence (n_k^*) (which also depends on t^*) such that

$$f_{n_k^*}(t^*)\overset{w^*}{\longrightarrow}f^{*,+}\quad\text{in }L^\infty(\Omega\times\mathbb{T}^N\times\mathbb{R})\text{-}w^*.$$

Note, that the domain of definition of $f^{*,+}$ does not depend on ψ. Thus on the one hand we have

$$\big\langle f_{n_k^*}(t^*),\psi\big\rangle\overset{w^*}{\longrightarrow}\big\langle f^{*,+},\psi\big\rangle\quad\text{in }L^\infty(\Omega)\text{-}w^*,$$

and on the other hand, due to the definition of f^+,

$$\big\langle f_{n_k^*}(t^*),\psi\big\rangle\overset{w^*}{\longrightarrow}\big\langle f^{+}(t^*),\psi\big\rangle\quad\text{in }L^\infty(\Omega)\text{-}w^*.$$

The proof of existence of the left-continuous representative f^- can be carried out similarly and so will be left to the reader.

The fact that $f^+(t^*)$ and $f^-(t^*)$ coincide for all $t^*\in(0,T)\setminus I$, where $I\subset(0,T)$ is countable, follows directly from their definition since the representatives $\langle f(t^*,\omega),\psi\rangle^+$ and $\langle f(t^*,\omega),\psi\rangle^-$ coincide except for an at most countable set for $\psi\in\mathcal{D}_1$. □

From now on, we will work with these two fixed representatives of f and we can take any of them in an integral with respect to time or in a stochastic integral.

As the next step towards the proof of the comparison principle, we need a technical proposition relating two kinetic solutions of (1). We will also use the following notation: if $f:X\times\mathbb{R}\to[0,1]$ is a kinetic function, we denote by $\bar f$ the conjugate function $\bar f=1-f$.

Proposition 3.2. *Let u_1, u_2 be two kinetic solutions to (1) and denote $f_1=\mathbf{1}_{u_1>\xi}$, $f_2=\mathbf{1}_{u_2>\xi}$. Then for $t\in[0,T]$ and any nonnegative functions $\varrho\in C^\infty(\mathbb{T}^N)$, $\psi\in C_c^\infty(\mathbb{R})$ we have*

$$\begin{aligned}&\mathbb{E}\int_{(\mathbb{T}^N)^2}\int_{\mathbb{R}^2}\varrho(x-y)\psi(\xi-\zeta)f_1^\pm(x,t,\xi)\bar f_2^\pm(y,t,\zeta)\,\mathrm{d}\xi\,\mathrm{d}\zeta\,\mathrm{d}x\,\mathrm{d}y\\&\quad\le\mathbb{E}\int_{(\mathbb{T}^N)^2}\int_{\mathbb{R}^2}\varrho(x-y)\psi(\xi-\zeta)f_{1,0}(x,\xi)\bar f_{2,0}(y,\zeta)\,\mathrm{d}\xi\,\mathrm{d}\zeta\,\mathrm{d}x\,\mathrm{d}y+\mathrm{I}+\mathrm{J}+\mathrm{K},\end{aligned}\tag{17}$$

where

$$\mathrm{I}=\mathbb{E}\int_0^t\int_{(\mathbb{T}^N)^2}\int_{\mathbb{R}^2}f_1\bar f_2\big(b(\xi)-b(\zeta)\big)\cdot\nabla_x\alpha(x,\xi,y,\zeta)\,\mathrm{d}\xi\,\mathrm{d}\zeta\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}s,$$

$$\begin{aligned}
\mathrm{J} = \ & \mathbb{E}\int_0^t \int_{(\mathbb{T}^N)^2}\int_{\mathbb{R}^2} f_1\bar{f}_2 \sum_{i,j=1}^N \partial_{y_j}\big(A_{ij}(y)\partial_{y_i}\alpha\big)\,\mathrm{d}\xi\,\mathrm{d}\zeta\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}s\\
&+\mathbb{E}\int_0^t \int_{(\mathbb{T}^N)^2}\int_{\mathbb{R}^2} f_1\bar{f}_2 \sum_{i,j=1}^N \partial_{x_j}\big(A_{ij}(x)\partial_{x_i}\alpha\big)\,\mathrm{d}\xi\,\mathrm{d}\zeta\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}s\\
&-\mathbb{E}\int_0^t \int_{(\mathbb{T}^N)^2}\int_{\mathbb{R}^2} \alpha(x,\xi,y,\zeta)\,\mathrm{d}\nu^1_{x,s}(\xi)\,\mathrm{d}x\,\mathrm{d}n_{2,1}(y,s,\zeta)\\
&-\mathbb{E}\int_0^t \int_{(\mathbb{T}^N)^2}\int_{\mathbb{R}^2} \alpha(x,\xi,y,\zeta)\,\mathrm{d}\nu^2_{y,s}(\zeta)\,\mathrm{d}y\,\mathrm{d}n_{1,1}(x,s,\xi),\\
\mathrm{K} = \ & \frac{1}{2}\mathbb{E}\int_0^t \int_{(\mathbb{T}^N)^2}\int_{\mathbb{R}^2} \alpha(x,\xi,y,\zeta)\sum_{k\geq 1}\big|g_k(x,\xi)-g_k(y,\zeta)\big|^2\mathrm{d}\nu^1_{x,s}(\xi)\mathrm{d}\nu^2_{y,s}(\zeta)\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}s,
\end{aligned}$$

and the function α is defined as $\alpha(x,\xi,y,\zeta)=\varrho(x-y)\psi(\xi-\zeta)$.

Proof. Let us denote by $\langle\!\langle\cdot,\cdot\rangle\!\rangle$ the scalar product in $L^2(\mathbb{T}^N_x\times\mathbb{T}^N_y\times\mathbb{R}_\xi\times\mathbb{R}_\zeta)$. In order to prove the statement in the case of $f_1^+,\ \bar{f}_2^+$, we employ similar calculations as in [9, Proposition 9] to obtain

$$\begin{aligned}
\mathbb{E}\langle\!\langle f_1^+(t)\bar{f}_2^+(t),\alpha\rangle\!\rangle = \ & \mathbb{E}\langle\!\langle f_{1,0}\bar{f}_{2,0},\alpha\rangle\!\rangle\\
&+\mathbb{E}\int_0^t \int_{(\mathbb{T}^N)^2}\int_{\mathbb{R}^2} f_1\bar{f}_2\big(b(\xi)-b(\zeta)\big)\cdot\nabla_x\alpha\,\mathrm{d}\xi\,\mathrm{d}\zeta\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}s\\
&+\mathbb{E}\int_0^t \int_{(\mathbb{T}^N)^2}\int_{\mathbb{R}^2} f_1\bar{f}_2 \sum_{i,j=1}^N \partial_{y_j}\big(A_{ij}(y)\partial_{y_i}\alpha\big)\,\mathrm{d}\xi\,\mathrm{d}\zeta\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}s\\
&+\mathbb{E}\int_0^t \int_{(\mathbb{T}^N)^2}\int_{\mathbb{R}^2} f_1\bar{f}_2 \sum_{i,j=1}^N \partial_{x_j}\big(A_{ij}(x)\partial_{x_i}\alpha\big)\,\mathrm{d}\xi\,\mathrm{d}\zeta\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}s\\
&+\frac{1}{2}\mathbb{E}\int_0^t \int_{(\mathbb{T}^N)^2}\int_{\mathbb{R}^2} \bar{f}_2\partial_\xi\alpha\, G_1^2\,\mathrm{d}\nu^1_{x,s}(\xi)\,\mathrm{d}\zeta\,\mathrm{d}y\,\mathrm{d}x\,\mathrm{d}s\\
&-\frac{1}{2}\mathbb{E}\int_0^t \int_{(\mathbb{T}^N)^2}\int_{\mathbb{R}^2} f_1\partial_\zeta\alpha\, G_2^2\,\mathrm{d}\nu^2_{y,s}(\zeta)\,\mathrm{d}\xi\,\mathrm{d}y\,\mathrm{d}x\,\mathrm{d}s\\
&-\mathbb{E}\int_0^t \int_{(\mathbb{T}^N)^2}\int_{\mathbb{R}^2} G_{1,2}\alpha\,\mathrm{d}\nu^1_{x,s}(\xi)\,\mathrm{d}\nu^2_{y,s}(\zeta)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}s\\
&-\mathbb{E}\int_0^t \int_{(\mathbb{T}^N)^2}\int_{\mathbb{R}^2} \bar{f}_2^-\partial_\xi\alpha\,\mathrm{d}m_1(x,s,\xi)\,\mathrm{d}\zeta\,\mathrm{d}y\\
&+\mathbb{E}\int_0^t \int_{(\mathbb{T}^N)^2}\int_{\mathbb{R}^2} f_1^+\partial_\zeta\alpha\,\mathrm{d}m_2(y,s,\zeta)\,\mathrm{d}\xi\,\mathrm{d}x. \qquad (18)
\end{aligned}$$

In particular, since $\alpha\geq 0$, the last term in (18) satisfies

$$\begin{aligned}
&\mathbb{E}\int_0^t \int_{(\mathbb{T}^N)^2}\int_{\mathbb{R}^2} f_1^+\partial_\zeta\alpha\,\mathrm{d}m_2(y,s,\zeta)\,\mathrm{d}\xi\,\mathrm{d}x\\
&\quad=-\mathbb{E}\int_0^t \int_{(\mathbb{T}^N)^2}\int_{\mathbb{R}^2} \alpha\,\mathrm{d}\nu^1_{x,s}(\xi)\,\mathrm{d}x\,\mathrm{d}n_{2,1}(y,s,\zeta)
\end{aligned}$$

$$-\mathbb{E}\int_0^t\int_{(\mathbb{T}^N)^2}\int_{\mathbb{R}^2}\alpha\,\mathrm{d}\nu^1_{x,s}(\xi)\,\mathrm{d}x\,\mathrm{d}n_{2,2}(y,s,\zeta)$$
$$\leq -\mathbb{E}\int_0^t\int_{(\mathbb{T}^N)^2}\int_{\mathbb{R}^2}\alpha\,\mathrm{d}\nu^1_{x,s}(\xi)\,\mathrm{d}x\,\mathrm{d}n_{2,1}(y,s,\zeta)$$

and by symmetry

$$-\mathbb{E}\int_0^t\int_{(\mathbb{T}^N)^2}\int_{\mathbb{R}^2}\bar{f}_2^-\partial_\xi\alpha\,\mathrm{d}m_1(x,s,\xi)\,\mathrm{d}\zeta\,\mathrm{d}y$$
$$\leq -\mathbb{E}\int_0^t\int_{(\mathbb{T}^N)^2}\int_{\mathbb{R}^2}\alpha\,\mathrm{d}\nu^2_{y,s}(\zeta)\,\mathrm{d}y\,\mathrm{d}n_{1,1}(x,s,\xi).$$

Thus, the desired estimate (17) follows.

In the case of $f_1^-,\ \bar{f}_2^-$ we take $t_n\uparrow t$, write (17) for $f_1^+(t_n),\ \bar{f}_2^+(t_n)$ and let $n\to\infty$. □

Theorem 3.3 (*Comparison Principle*)**.** *Let u be a kinetic solution to (1). Then there exist u^+ and u^-, representatives of u, such that, for all $t\in[0,T]$, $f^\pm(x,t,\xi)=\mathbf{1}_{u^\pm(x,t)>\xi}$ for a.e. (ω,x,ξ). Moreover, if u_1,u_2 are kinetic solutions to (1) with initial data $u_{1,0}$ and $u_{2,0}$, respectively, then for all $t\in[0,T]$*

$$\mathbb{E}\|u_1^\pm(t)-u_2^\pm(t)\|_{L^1(\mathbb{T}^N)}\leq\mathbb{E}\|u_{1,0}-u_{2,0}\|_{L^1(\mathbb{T}^N)}. \tag{19}$$

Proof. Denote $f_1=\mathbf{1}_{u_1>\xi},\ f_2=\mathbf{1}_{u_2>\xi}$. Let $(\psi_\delta),\ (\varrho_\tau)$ be approximations to the identity on $\mathbb{R}$ and $\mathbb{T}^N$, respectively. Namely, let $\psi\in C_c^\infty(\mathbb{R})$ be a nonnegative symmetric function satisfying $\int_\mathbb{R}\psi=1$, $\operatorname{supp}\psi\subset(-1,1)$ and set

$$\psi_\delta(\xi)=\frac{1}{\delta}\,\psi\left(\frac{\xi}{\delta}\right).$$

For the space variable $x\in\mathbb{T}^N$, we employ the approximation to the identity defined in Lemma 2.5. Then we have

$$\mathbb{E}\int_{\mathbb{T}^N}\int_\mathbb{R} f_1^\pm(x,t,\xi)\bar{f}_2^\pm(x,t,\xi)\,\mathrm{d}\xi\,\mathrm{d}x$$
$$=\mathbb{E}\int_{(\mathbb{T}^N)^2}\int_{\mathbb{R}^2}\varrho_\tau(x-y)\psi_\delta(\xi-\zeta)f_1^\pm(x,t,\xi)\bar{f}_2^\pm(y,t,\zeta)\,\mathrm{d}\xi\,\mathrm{d}\zeta\,\mathrm{d}x\,\mathrm{d}y+\eta_t(\tau,\delta),$$

where $\lim_{\tau,\delta\to0}\eta_t(\tau,\delta)=0$. With regard to Proposition 3.2, we need to find suitable bounds for terms I, J, K.

Since b has at most polynomial growth, there exist $C>0,\ p>1$ such that

$$\big|b(\xi)-b(\zeta)\big|\leq\Gamma(\xi,\zeta)|\xi-\zeta|,\qquad \Gamma(\xi,\zeta)\leq C\big(1+|\xi|^{p-1}+|\zeta|^{p-1}\big).$$

Hence

$$|\mathrm{I}|\leq\mathbb{E}\int_0^t\int_{(\mathbb{T}^N)^2}\int_{\mathbb{R}^2}f_1\bar{f}_2\Gamma(\xi,\zeta)|\xi-\zeta|\psi_\delta(\xi-\zeta)\,\mathrm{d}\xi\,\mathrm{d}\zeta\,\big|\nabla_x\varrho_\tau(x-y)\big|\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}s.$$

As the next step we apply integration by parts with respect to $\zeta,\ \xi$. Focusing only on the relevant integrals we get

$$\int_\mathbb{R}f_1(\xi)\int_\mathbb{R}\bar{f}_2(\zeta)\Gamma(\xi,\zeta)|\xi-\zeta|\psi_\delta(\xi-\zeta)\mathrm{d}\zeta\,\mathrm{d}\xi$$

$$
\begin{aligned}
&= \int_{\mathbb{R}} f_1(\xi) \int_{\mathbb{R}} \Gamma(\xi,\zeta')|\xi-\zeta'|\psi_\delta(\xi-\zeta')\mathrm{d}\zeta'\,\mathrm{d}\xi \\
&\quad - \int_{\mathbb{R}^2} f_1(\xi) \int_{-\infty}^{\zeta} \Gamma(\xi,\zeta')|\xi-\zeta'|\psi_\delta(\xi-\zeta')\mathrm{d}\zeta'\,\mathrm{d}\xi\,\mathrm{d}\nu^2_{y,s}(\zeta) \\
&= \int_{\mathbb{R}^2} f_1(\xi) \int_{\zeta}^{\infty} \Gamma(\xi,\zeta')|\xi-\zeta'|\psi_\delta(\xi-\zeta')\mathrm{d}\zeta'\,\mathrm{d}\xi\,\mathrm{d}\nu^2_{y,s}(\zeta) \\
&= \int_{\mathbb{R}^2} \Upsilon(\xi,\zeta)\mathrm{d}\nu^1_{x,s}(\xi)\mathrm{d}\nu^2_{y,s}(\zeta)
\end{aligned}
\tag{20}
$$

where

$$
\Upsilon(\xi,\zeta) = \int_{-\infty}^{\xi}\int_{\zeta}^{\infty} \Gamma(\xi',\zeta')|\xi'-\zeta'|\psi_\delta(\xi'-\zeta')\mathrm{d}\zeta'\,\mathrm{d}\xi'.
$$

Therefore, we find

$$
|\mathrm{I}| \le \mathbb{E}\int_0^t \int_{(\mathbb{T}^N)^2}\int_{\mathbb{R}^2} \Upsilon(\xi,\zeta)\,\mathrm{d}\nu^1_{x,s}(\xi)\mathrm{d}\nu^2_{y,s}(\zeta)\,\big|\nabla_x \varrho_\tau(x-y)\big|\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}s.
$$

The function Υ can be estimated using the substitution $\xi''=\xi'-\zeta'$

$$
\begin{aligned}
\Upsilon(\xi,\zeta) &= \int_{\zeta}^{\infty}\int_{|\xi''|<\delta,\,\xi''<\xi-\zeta'} \Gamma(\xi''+\zeta',\zeta')|\xi''|\psi_\delta(\xi'')\,\mathrm{d}\xi''\,\mathrm{d}\zeta' \\
&\le C\delta \int_{\zeta}^{\xi+\delta} \max_{|\xi''|<\delta,\,\xi''<\xi-\zeta'} \Gamma(\xi''+\zeta',\zeta')\,\mathrm{d}\zeta' \\
&\le C\delta \int_{\zeta}^{\xi+\delta} \big(1+|\xi|^{p-1}+|\zeta'|^{p-1}\big)\,\mathrm{d}\zeta' \\
&\le C\delta\big(1+|\xi|^p+|\zeta|^p\big);
\end{aligned}
$$

hence, since $\nu^1,\ \nu^2$ vanish at infinity,

$$
|\mathrm{I}| \le Ct\delta \int_{\mathbb{T}^N} \big|\nabla_x \varrho_\tau(x)\big|\,\mathrm{d}x \le Ct\delta\tau^{-1}.
$$

We recall that $f_1 = \mathbf{1}_{u_1(x,t)>\xi},\ f_2 = \mathbf{1}_{u_2(y,t)>\zeta}$; hence

$$
\partial_\xi f_1 = -\nu^1 = -\delta_{u_1(x,t)=\xi}, \qquad \partial_\zeta f_2 = -\nu^2 = -\delta_{u_2(y,t)=\zeta}
$$

and as both u_1, u_2 possess some regularity in the nondegeneracy zones of A due to Definition 2.2(ii), we obtain as in (11)

$$
\sigma\,\nabla f_1 = \sigma\,\nabla u_1\,\delta_{u_1(x,s)=\xi}, \qquad \sigma\,\nabla \bar{f}_2 = -\sigma\,\nabla u_2\,\delta_{u_2(y,s)=\zeta}
$$

in the sense of distributions over $\mathbb{T}^N\times\mathbb{R}$. The first term in J can be rewritten in the following manner using integration by parts (and considering only relevant integrals)

$$
\begin{aligned}
&\int_{\mathbb{T}^N} f_1 \int_{\mathbb{T}^N} \bar{f}_2\,\partial_{y_j}\big(A_{ij}(y)\partial_{y_i}\varrho_\tau(x-y)\big)\mathrm{d}y\,\mathrm{d}x \\
&\quad = \int_{(\mathbb{T}^N)^2} f_1(x,s,\xi)A_{ij}(y)\partial_{y_j}\bar{f}_2(y,s,\zeta)\partial_{x_i}\varrho_\tau(x-y)\mathrm{d}x\,\mathrm{d}y
\end{aligned}
$$

and similarly for the second term. Let us define

$$\Theta_\delta(\xi)=\int_{-\infty}^{\xi}\psi_\delta(\zeta)\,\mathrm{d}\zeta.$$

Then we have $\mathrm{J}=\mathrm{J}_1+\mathrm{J}_2+\mathrm{J}_3$ with

$$\mathrm{J}_1=-\mathbb{E}\int_0^t\int_{(\mathbb{T}^N)^2}(\nabla_x u_1)^*\sigma(x)\sigma(x)(\nabla\varrho_\tau)(x-y)\,\Theta_\delta\big(u_1(x,s)-u_2(y,s)\big)\mathrm{d}x\mathrm{d}y\mathrm{d}s,$$

$$\mathrm{J}_2=-\mathbb{E}\int_0^t\int_{(\mathbb{T}^N)^2}(\nabla_y u_2)^*\sigma(y)\sigma(y)(\nabla\varrho_\tau)(x-y)\,\Theta_\delta\big(u_1(x,s)-u_2(y,s)\big)\mathrm{d}x\mathrm{d}y\mathrm{d}s,$$

$$\begin{aligned}\mathrm{J}_3 = -\mathbb{E}\int_0^t\int_{(\mathbb{T}^N)^2}&\big[|\sigma(x)\nabla_x u_1|^2+|\sigma(y)\nabla_y u_2|^2\big]\varrho_\tau(x-y)\psi_\delta\big(u_1(x,s)\\ &-u_2(y,s)\big)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}s.\end{aligned}$$

Let

$$\mathrm{H}=\mathbb{E}\int_0^t\int_{(\mathbb{T}^N)^2}(\nabla_x u_1)^*\sigma(x)\sigma(y)(\nabla_y u_2)\varrho_\tau(x-y)\psi_\delta\big(u_1(x,s)-u_2(y,s)\big)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}s.$$

We intend to show that $\mathrm{J}_1=\mathrm{H}+o(1),\ \mathrm{J}_2=\mathrm{H}+o(1)$, where $o(1)\to 0$ as $\tau\to 0$ uniformly in δ, and consequently

$$\begin{aligned}\mathrm{J}=-\mathbb{E}\int_0^t\int_{(\mathbb{T}^N)^2}&\big|\sigma(x)\nabla_x u_1-\sigma(y)\nabla_y u_2\big|^2\varrho_\tau(x-y)\\ &\times\psi_\delta\big(u_1(x,s)-u_2(y,s)\big)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}s+o(1)\le o(1).\end{aligned}\tag{21}$$

We only prove the claim for J_1 since the case of J_2 is analogous. Let us define

$$g(x,y,s)=(\nabla_x u_1)^*\sigma(x)\,\Theta_\delta\big(u_1(x,s)-u_2(y,s)\big).$$

Here, we employ again the assumption (ii) in Definition 2.2. Recall, that it gives us some regularity of the solution in the nondegeneracy zones of the diffusion matrix A and hence $g\in L^2(\Omega\times\mathbb{T}_x^N\times\mathbb{T}_y^N\times[0,T])$. It holds

$$\begin{aligned}\mathrm{J}_1 &= -\mathbb{E}\int_0^t\int_{(\mathbb{T}^N)^2}g(x,y,s)\Big(\sigma(x)-\sigma(y)\Big)(\nabla\varrho_\tau)(x-y)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}s\\ &\quad-\mathbb{E}\int_0^t\int_{(\mathbb{T}^N)^2}g(x,y,s)\sigma(y)(\nabla\varrho_\tau)(x-y)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}s,\\ \mathrm{H} &= \mathbb{E}\int_0^t\int_{(\mathbb{T}^N)^2}g(x,y,s)\mathrm{div}_y\Big(\sigma(y)\varrho_\tau(x-y)\Big)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}s\\ &= \mathbb{E}\int_0^t\int_{(\mathbb{T}^N)^2}g(x,y,s)\mathrm{div}\big(\sigma(y)\big)\varrho_\tau(x-y)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}s\\ &\quad-\mathbb{E}\int_0^t\int_{(\mathbb{T}^N)^2}g(x,y,s)\sigma(y)(\nabla\varrho_\tau)(x-y)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}s,\end{aligned}$$

where divergence is applied row-wise to a matrix-valued function. Therefore, it is enough to show that the first terms in J_1 and H have the same limit value if $\tau\to 0$. For H, we obtain easily

$$\mathbb{E}\int_0^t\int_{(\mathbb{T}^N)^2}g(x,y,s)\mathrm{div}\big(\sigma(y)\big)\varrho_\tau(x-y)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}s$$

$$\longrightarrow \mathbb{E}\int_0^t \int_{\mathbb{T}^N} g(y,y,s)\mathrm{div}\big(\sigma(y)\big)\,\mathrm{d}y\,\mathrm{d}s$$

so it remains to verify

$$-\mathbb{E}\int_0^t \int_{(\mathbb{T}^N)^2} g(x,y,s)\Big(\sigma(x)-\sigma(y)\Big)(\nabla\varrho_\tau)(x-y)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}s$$
$$\longrightarrow \mathbb{E}\int_0^t \int_{\mathbb{T}^N} g(y,y,s)\mathrm{div}\big(\sigma(y)\big)\,\mathrm{d}y\,\mathrm{d}s.$$

Here, we employ again the arguments of the commutation lemma of DiPerna and Lions (see [10, Lemma II.1], cf. Lemma 2.5). Let us denote by g^i the ith element of g and by σ^i the ith row of σ. Since $\tau|\nabla\varrho_\tau|(\cdot)\leq C\varrho_{2\tau}(\cdot)$ with a constant independent of τ, we obtain the following estimate

$$\mathbb{E}\int_0^t \int_{\mathbb{T}^N}\left|\int_{\mathbb{T}^N} g^i(x,y,s)\Big(\sigma^i(x)-\sigma^i(y)\Big)(\nabla\varrho_\tau)(x-y)\,\mathrm{d}x\right|\mathrm{d}y\,\mathrm{d}s$$
$$\leq C \operatorname*{ess\,sup}_{\substack{x',y'\in\mathbb{T}^N\\|x'-y'|\leq\tau}}\left|\frac{\sigma^i(x')-\sigma^i(y')}{\tau}\right|\mathbb{E}\int_0^T\int_{(\mathbb{T}^N)^2}\big|g^i(x,y,s)\big|\varrho_{2\tau}(x-y)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}s.$$

Note that according to [14,28], the square-root matrix of A is Lipschitz continuous and therefore the essential supremum can be estimated by a constant independent of τ. Next

$$\mathbb{E}\int_0^T\int_{(\mathbb{T}^N)^2}\big|g^i(x,y,s)\big|\varrho_{2\tau}(x-y)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}s$$
$$\leq\left(\mathbb{E}\int_0^T\int_{(\mathbb{T}^N)^2}\big|g^i(x,y,s)\big|^2\varrho_{2\tau}(x-y)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}s\right)^{\frac12}\left(\int_{(\mathbb{T}^N)^2}\varrho_{2\tau}(x-y)\,\mathrm{d}x\,\mathrm{d}y\right)^{\frac12}$$
$$\leq\left(\mathbb{E}\int_0^T\int_{\mathbb{T}^N}\big|(\nabla_x u_1)^*\sigma(x)\big|^2\int_{\mathbb{T}^N}\varrho_{2\tau}(x-y)\,\mathrm{d}y\,\mathrm{d}x\,\mathrm{d}s\right)^{\frac12}$$
$$\leq\big\|(\nabla_x u_1)^*\sigma(x)\big\|_{L^2(\Omega\times\mathbb{T}^N\times[0,T])}.$$

So we get an estimate which is independent of τ and δ. It is sufficient to consider the case when g^i and σ^i are smooth. The general case follows by the density argument from the above bound. It holds

$$-\mathbb{E}\int_0^t\int_{(\mathbb{T}^N)^2} g^i(x,y,s)\Big(\sigma^i(x)-\sigma^i(y)\Big)(\nabla\varrho_\tau)(x-y)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}s$$
$$=-\frac{1}{\tau^{N+1}}\mathbb{E}\int_0^t\int_{(\mathbb{T}^N)^2}\int_0^1 g^i(x,y,s)\,\mathrm{D}\sigma^i\big(y+r(x-y)\big)(x-y)$$
$$\cdot(\nabla\varrho)\Big(\frac{x-y}{\tau}\Big)\,\mathrm{d}r\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}s$$
$$=-\mathbb{E}\int_0^t\int_{(\mathbb{T}^N)^2}\int_0^1 g^i(y+\tau z,y,s)\,\mathrm{D}\sigma^i(y+r\tau z)z\cdot(\nabla\varrho)(z)\,\mathrm{d}r\,\mathrm{d}z\,\mathrm{d}y\,\mathrm{d}s$$
$$\longrightarrow -\mathbb{E}\int_0^t\int_{(\mathbb{T}^N)^2} g^i(y,y,s)\,\mathrm{D}\sigma^i(y)z\cdot(\nabla\varrho)(z)\,\mathrm{d}z\,\mathrm{d}y\,\mathrm{d}s.$$

Moreover, by (14),

$$-\mathbb{E}\int_0^t\int_{(\mathbb{T}^N)^2} g^i(y,y,s)\,\mathrm{D}\sigma^i(y)z\cdot(\nabla\varrho)(z)\,\mathrm{d}z\,\mathrm{d}y\,\mathrm{d}s$$
$$=\mathbb{E}\int_0^t\int_{\mathbb{T}^N} g^i(y,y,s)\mathrm{div}\big(\sigma^i(y)\big)\,\mathrm{d}y\,\mathrm{d}s$$

and accordingly (21) follows.

The last term K is, due to (3), bounded as follows

$$\begin{aligned}\mathrm{K}&\le\frac{L}{2}\mathbb{E}\int_0^t\int_{(\mathbb{T}^N)^2}\varrho_\tau(x-y)|x-y|^2\int_{\mathbb{R}^2}\psi_\delta(\xi-\zeta)\,\mathrm{d}\nu^1_{x,s}(\xi)\,\mathrm{d}\nu^2_{y,s}(\zeta)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}s\\&\quad+\frac{L}{2}\mathbb{E}\int_0^t\int_{(\mathbb{T}^N)^2}\varrho_\tau(x-y)\int_{\mathbb{R}^2}\psi_\delta(\xi-\zeta)|\xi-\zeta|h(|\xi-\zeta|)\mathrm{d}\nu^1_{x,s}(\xi)\mathrm{d}\nu^2_{y,s}(\zeta)\mathrm{d}x\mathrm{d}y\mathrm{d}s\\&\le\frac{Lt}{2\delta}\int_{(\mathbb{T}^N)^2}|x-y|^2\varrho_\tau(x-y)\,\mathrm{d}x\,\mathrm{d}y+\frac{LtC_\psi h(\delta)}{2}\int_{(\mathbb{T}^N)^2}\varrho_\tau(x-y)\,\mathrm{d}x\,\mathrm{d}y\\&\le\frac{Lt}{2}\delta^{-1}\tau^2+\frac{LtC_\psi h(\delta)}{2},\end{aligned}$$

where $C_\psi=\sup_{\xi\in\mathbb{R}}|\xi\psi(\xi)|$. Finally, we set $\delta=\tau^{4/3}$, let $\tau\to0$ and deduce

$$\mathbb{E}\int_{\mathbb{T}^N}\int_{\mathbb{R}}f_1^\pm(t)\bar{f}_2^\pm(t)\,\mathrm{d}\xi\,\mathrm{d}x\le\mathbb{E}\int_{\mathbb{T}^N}\int_{\mathbb{R}}f_{1,0}\bar{f}_{2,0}\,\mathrm{d}\xi\,\mathrm{d}x.$$

Let us now consider $f_1=f_2=f$. Since $f_0=\mathbf{1}_{u_0>\xi}$ we have the identity $f_0\bar{f}_0=0$ and therefore $f^\pm(1-f^\pm)=0$ a.e. (ω,x,ξ) and for all t. The fact that $f^\pm$ is a kinetic function and Fubini's theorem then imply that, for any $t\in[0,T]$, there exists a set $\Sigma_t\subset\Omega\times\mathbb{T}^N$ of full measure such that, for $(\omega,x)\in\Sigma_t$, $f^\pm(\omega,x,t,\xi)\in\{0,1\}$ for a.e. $\xi\in\mathbb{R}$. Therefore, there exist $u^\pm:\Omega\times\mathbb{T}^N\times[0,T]\to\mathbb{R}$ such that $f^\pm=\mathbf{1}_{u^\pm>\xi}$ for a.e. (ω,x,ξ) and all t. In particular, $u^\pm=\int_{\mathbb{R}}(f^\pm-\mathbf{1}_{0>\xi})\mathrm{d}\xi$ for a.e. (ω,x) and all t. It follows now from Proposition 3.1 and the identity

$$|\alpha-\beta|=\int_{\mathbb{R}}|\mathbf{1}_{\alpha>\xi}-\mathbf{1}_{\beta>\xi}|\,\mathrm{d}\xi,\quad\alpha,\ \beta\in\mathbb{R},$$

that $u^+=u^-=u$ for a.e. $t\in[0,T]$. Since

$$\int_{\mathbb{R}}\mathbf{1}_{u_1^\pm>\xi}\overline{\mathbf{1}_{u_2^\pm>\xi}}\,\mathrm{d}\xi=(u_1^\pm-u_2^\pm)^+$$

we obtain the comparison property

$$\mathbb{E}\big\|\big(u_1^\pm(t)-u_2^\pm(t)\big)^+\big\|_{L^1(\mathbb{T}^N)}\le\mathbb{E}\big\|(u_{1,0}-u_{2,0})^+\big\|_{L^1(\mathbb{T}^N)}.\quad\square$$

As a consequence of Theorem 3.3, namely from the comparison property (19), the uniqueness part of Theorem 2.10 follows. Furthermore, we obtain the continuity of trajectories in $L^p(\mathbb{T}^N)$.

Corollary 3.4 (*Continuity in Time*). *Let u be a kinetic solution to (1). Then there exists a representative of u which has almost surely continuous trajectories in $L^p(\mathbb{T}^N)$, for all $p\in[1,\infty)$.*

Proof. Remark, that due to the construction of $f^{\pm}$ it holds, for all $p \in [1, \infty)$,

$$\mathbb{E} \sup_{0 \le t \le T} \int_{\mathbb{T}^N} |u^{\pm}(x,t)|^p \, \mathrm{d}x = \mathbb{E} \sup_{0 \le t \le T} \int_{\mathbb{T}^N} \int_{\mathbb{R}} |\xi|^p \, \mathrm{d}\nu^{\pm}_{x,t}(\xi) \, \mathrm{d}x \le C. \tag{22}$$

Now, we are able to prove that the modification u^+ is right-continuous in the sense of $L^p(\mathbb{T}^N)$. According to Proposition 3.1 applied to the solution f^+, we obtain

$$\langle f^+(t+\varepsilon), \psi \rangle \longrightarrow \langle f^+(t), \psi \rangle, \quad \varepsilon \downarrow 0, \ \forall \psi \in L^1(\mathbb{T}^N \times \mathbb{R}).$$

Setting $\psi(x,\xi) = \psi_1(x) \partial_\xi \psi_2(\xi)$ for some functions $\psi_1 \in L^1(\mathbb{T}^N)$ and $\partial_\xi \psi_2 \in C_c^\infty(\mathbb{R})$, it reads

$$\int_{\mathbb{T}^N} \psi_1(x) \psi_2\big(u^+(x,t+\varepsilon)\big) \, \mathrm{d}x \longrightarrow \int_{\mathbb{T}^N} \psi_1(x) \psi_2\big(u^+(x,t)\big) \, \mathrm{d}x. \tag{23}$$

In order to obtain that $u^+(t+\varepsilon) \overset{w}{\to} u^+(t)$ in $L^p(\mathbb{T}^N)$, $p \in [1,\infty)$, we set $\psi_2^\delta(\xi) = \xi \chi_\delta(\xi)$ where (χ_δ) is a truncation on $\mathbb{R}$, i.e. we define $\chi_\delta(\xi) = \chi(\delta\xi)$, where χ is a smooth function with bounded support satisfying $0 \le \chi \le 1$ and

$$\chi(\xi) = \begin{cases} 1, & \text{if } |\xi| \le \dfrac{1}{2}, \\ 0, & \text{if } |\xi| \ge 1, \end{cases}$$

and deduce

$$\begin{aligned} &\left| \int_{\mathbb{T}^N} \psi_1(x)\, u^+(x,t+\varepsilon) \, \mathrm{d}x - \int_{\mathbb{T}^N} \psi_1(x)\, u^+(x,t) \, \mathrm{d}x \right| \\ &\quad \le \int_{\mathbb{T}^N} \big| \psi_1(x)\, u^+(x,t+\varepsilon) \big| \, \mathbf{1}_{|u^+(x,t+\varepsilon)| > 1/2\delta} \, \mathrm{d}x \\ &\qquad + \left| \int_{\mathbb{T}^N} \psi_1(x) \psi_2^\delta\big(u^+(x,t+\varepsilon)\big) - \psi_1(x) \psi_2^\delta\big(u^+(x,t)\big) \, \mathrm{d}x \right| \\ &\qquad + \int_{\mathbb{T}^N} \big| \psi_1(x)\, u^+(x,t) \big| \, \mathbf{1}_{|u^+(x,t)| > 1/2\delta} \, \mathrm{d}x \longrightarrow 0, \quad \varepsilon \downarrow 0, \end{aligned}$$

since the first and the third term on the right hand side tend to zero as $\delta \to 0$ uniformly in ε due to the uniform estimate (22) and the second one vanishes as $\varepsilon \to 0$ for any δ by (23).

The strong convergence in $L^2(\mathbb{T}^N)$ then follows easily as soon as we verify the convergence of the $L^2(\mathbb{T}^N)$-norms. This can be done by a similar approximation procedure, using $\psi_1(x) = 1$ and $\psi_2^\delta(\xi) = \xi^2 \chi_\delta(\xi)$. For the strong convergence in $L^p(\mathbb{T}^N)$ for general $p \in [1,\infty)$ we employ the Hölder inequality and the uniform bound (22).

A similar approach then shows that the modification u^- is left-continuous in the sense of $L^p(\mathbb{T}^N)$. The rest of the proof, showing that $u^-(t) = u^+(t)$ for all $t \in [0,T]$ can be carried out similarly to [9, Corollary 12]. □

4. Existence—smooth initial data

In this section we prove the existence part of Theorem 2.10 under an additional assumption upon the initial condition: $u_0 \in L^p(\Omega; C^\infty(\mathbb{T}^N))$, for all $p \in [1,\infty)$. We employ the vanishing viscosity method, i.e. we approximate Eq. (1) by certain nondegenerate problems, while using also some appropriately chosen approximations Φ^ε, B^ε of Φ and B, respectively.

These equations have smooth solutions and consequent passage to the limit gives the existence of a kinetic solution to the original equation. Nevertheless, the limit argument is quite technical and has to be done in several steps. It is based on the compactness method: the uniform energy estimates yield tightness of a sequence of approximate solutions and thus, on another probability space, this sequence converges almost surely due to the Skorokhod representation theorem. The limit is then shown to be a martingale kinetic solution to (1). Combining this fact and the pathwise uniqueness with the Gyöngy–Krylov characterization of convergence in probability, we finally obtain the desired kinetic solution.

4.1. Nondegenerate case

Consider a truncation (χ_ε) on $\mathbb{R}$ and approximations to the identity (φ_ε), (ψ_ε) on $\mathbb{T}^N \times \mathbb{R}$ and $\mathbb{R}$, respectively. To be more precise concerning the case of $\mathbb{T}^N \times \mathbb{R}$, we make use of the same notation as at the beginning of the proof of Theorem 3.3 and define

$$\varphi_\varepsilon(x,\xi) = \frac{1}{\varepsilon^{N+1}}\varrho\Big(\frac{x}{\varepsilon}\Big)\psi\Big(\frac{\xi}{\varepsilon}\Big).$$

The regularizations of Φ, B are then defined in the following way

$$B_i^\varepsilon(\xi) = \big((B_i * \psi_\varepsilon)\chi_\varepsilon\big)(\xi), \qquad i = 1,\dots,N,$$
$$g_k^\varepsilon(x,\xi) = \begin{cases}\big((g_k * \varphi_\varepsilon)\chi_\varepsilon\big)(x,\xi), & \text{if } k \le \lfloor 1/\varepsilon \rfloor,\\ 0, & \text{if } k > \lfloor 1/\varepsilon \rfloor,\end{cases}$$

where $x \in \mathbb{T}^N$, $\xi \in \mathbb{R}$. Consequently, we set $B^\varepsilon = (B_1^\varepsilon,\dots,B_N^\varepsilon)$ and define the operator Φ^ε by $\Phi^\varepsilon(z)e_k = g_k^\varepsilon(\cdot, z(\cdot))$, $z \in L^2(\mathbb{T}^N)$. Clearly, the approximations B^ε, g_k^ε are of class C^∞ with a compact support therefore Lipschitz continuous. Moreover, the functions g_k^ε satisfy (2), (3) uniformly in ε and the following Lipschitz condition holds true

$$\forall x \in \mathbb{T}^N\ \forall \xi,\ \zeta \in \mathbb{R} \quad \sum_{k\ge1} |g_k^\varepsilon(x,\xi) - g_k^\varepsilon(x,\zeta)|^2 \le L_\varepsilon |\xi - \zeta|^2. \tag{24}$$

From (2) we conclude that $\Phi^\varepsilon(z)$ is Hilbert–Schmidt for all $z \in L^2(\mathbb{T}^N)$. Also the polynomial growth of B remains valid for B^ε and holds uniformly in ε. Suitable approximation of the diffusion matrix A is obtained as its perturbation by εI, where I denotes the identity matrix. We denote $A^\varepsilon = A + \varepsilon \mathrm{I}$.

Consider an approximation of problem (1) by a nondegenerate equation

$$\begin{aligned} &\mathrm{d}u^\varepsilon + \operatorname{div}\big(B^\varepsilon(u^\varepsilon)\big)\mathrm{d}t = \operatorname{div}\big(A^\varepsilon(x)\nabla u^\varepsilon\big)\mathrm{d}t + \Phi^\varepsilon(u^\varepsilon)\,\mathrm{d}W,\\ &u^\varepsilon(0) = u_0. \end{aligned} \tag{25}$$

Theorem 4.1. *Assume that $u_0 \in L^p(\Omega; C^\infty(\mathbb{T}^N))$ for all $p \in (2,\infty)$. For any $\varepsilon > 0$, there exists a $C^\infty(\mathbb{T}^N)$-valued process which is the unique strong solution to (25). Moreover, it belongs to*

$$L^p(\Omega; C([0,T]; W^{l,q}(\mathbb{T}^N))) \quad \textit{for every } p \in (2,\infty),\ q \in [2,\infty),\ l \in \mathbb{N}.$$

Proof. For any fixed $\varepsilon > 0$, the assumptions of [18, Theorem 2.1, Corollary 2.2] are satisfied and therefore the claim follows. □

Let m^ε be the parabolic dissipative measure corresponding to the diffusion matrix $A + \varepsilon I$. To be more precise, set

$$\mathrm{d}n_1^\varepsilon(x,t,\xi) = \big|\sigma(x)\nabla u^\varepsilon\big|^2 \,\mathrm{d}\delta_{u^\varepsilon(x,t)}(\xi)\,\mathrm{d}x\,\mathrm{d}t,$$
$$\mathrm{d}n_2^\varepsilon(x,t,\xi) = \varepsilon|\nabla u^\varepsilon|^2 \,\mathrm{d}\delta_{u^\varepsilon(x,t)}(\xi)\,\mathrm{d}x\,\mathrm{d}t,$$

and define $m^\varepsilon = n_1^\varepsilon + n_2^\varepsilon$. Then, using the same approach as in Section 2, one can verify that the pair $(f^\varepsilon = \mathbf{1}_{u^\varepsilon>\xi},\ m^\varepsilon)$ satisfies the kinetic formulation of (25): let $\varphi \in C_c^\infty(\mathbb{T}^N\times\mathbb{R}),\ t\in[0,T]$, then it holds true $\mathbb{P}$-a.s.

$$\begin{aligned}&\big\langle f^\varepsilon(t),\varphi\big\rangle - \big\langle f_0,\varphi\big\rangle - \int_0^t \big\langle f^\varepsilon(s), b^\varepsilon(\xi)\cdot\nabla\varphi\big\rangle\,\mathrm{d}s\\ &\quad - \int_0^t \Big\langle f^\varepsilon(s), \operatorname{div}\big(A(x)\nabla\varphi\big)\Big\rangle\,\mathrm{d}s - \varepsilon\int_0^t \big\langle f^\varepsilon(s),\Delta\varphi\big\rangle\,\mathrm{d}s\\ &= \int_0^t \big\langle \delta_{u^\varepsilon=\xi}\ \Phi^\varepsilon(u^\varepsilon)\,\mathrm{d}W,\varphi\big\rangle + \frac12\int_0^t \big\langle \delta_{u^\varepsilon=\xi}\,G^2, \partial_\xi\varphi\big\rangle\,\mathrm{d}s - \big\langle m^\varepsilon,\partial_\xi\varphi\big\rangle([0,t)).\end{aligned} \tag{26}$$

Note, that by taking limit in ε we lose this precise structure of n_2.

4.2. Energy estimates

In this subsection we shall establish the so-called energy estimate that makes it possible to find uniform bounds for approximate solutions and that will later on yield a solution by invoking a compactness argument.

Lemma 4.2. *For all $\varepsilon\in(0,1)$, for all $t\in[0,T]$ and for all $p\in[2,\infty)$, the solution u^ε satisfies the inequality*

$$\mathbb{E}\|u^\varepsilon(t)\|^p_{L^p(\mathbb{T}^N)} \le C\big(1+\mathbb{E}\|u_0\|^p_{L^p(\mathbb{T}^N)}\big). \tag{27}$$

Proof. According to Theorem 4.1, the process u^ε is an $L^p(\mathbb{T}^N)$-valued continuous semimartingale so we can apply the infinite-dimensional Itô formula [8, Theorem 4.17] for the function $f(v) = \|v\|^p_{L^p(\mathbb{T}^N)}$. If q is the conjugate exponent to p then $f'(v) = p|v|^{p-2}v \in L^q(\mathbb{T}^N)$ and

$$f''(v) = p(p-1)|v|^{p-2}\,\mathrm{Id} \in \mathscr{L}\big(L^p(\mathbb{T}^N), L^q(\mathbb{T}^N)\big).$$

Therefore

$$\begin{aligned}\|u^\varepsilon(t)\|^p_{L^p(\mathbb{T}^N)} &= \|u_0\|^p_{L^p(\mathbb{T}^N)} - p\int_0^t\int_{\mathbb{T}^N} |u^\varepsilon|^{p-2}u^\varepsilon \operatorname{div}\big(B^\varepsilon(u^\varepsilon)\big)\,\mathrm{d}x\,\mathrm{d}s\\ &\quad + p\int_0^t\int_{\mathbb{T}^N} |u^\varepsilon|^{p-2}u^\varepsilon \operatorname{div}\big(A(x)\nabla u^\varepsilon\big)\,\mathrm{d}x\,\mathrm{d}s\\ &\quad + \varepsilon p\int_0^t\int_{\mathbb{T}^N} |u^\varepsilon|^{p-2}u^\varepsilon\,\Delta u^\varepsilon\,\mathrm{d}x\,\mathrm{d}s\\ &\quad + p\sum_{k\ge1}\int_0^t\int_{\mathbb{T}^N} |u^\varepsilon|^{p-2}u^\varepsilon g_k^\varepsilon(x,u^\varepsilon)\,\mathrm{d}x\,\mathrm{d}\beta_k(s)\\ &\quad + \frac12 p(p-1)\int_0^t\int_{\mathbb{T}^N} |u^\varepsilon|^{p-2}G_\varepsilon^2(x,u^\varepsilon)\,\mathrm{d}x\,\mathrm{d}s.\end{aligned} \tag{28}$$

If we define $H^\varepsilon(\xi)=\int_0^\xi|\zeta|^{p-2}B^\varepsilon(\zeta)\,\mathrm{d}\zeta$ then the second term on the right hand side vanishes due to the boundary conditions

$$-p\int_0^t\int_{\mathbb{T}^N}|u^\varepsilon|^{p-2}u^\varepsilon\mathrm{div}\big(B(u^\varepsilon)\big)\mathrm{d}x\,\mathrm{d}s=p\int_0^t\int_{\mathbb{T}^N}\mathrm{div}\big(H^\varepsilon(u^\varepsilon)\big)\,\mathrm{d}x\,\mathrm{d}s=0.$$

The third term is nonpositive as the matrix A is positive-semidefinite

$$\begin{aligned}p&\int_0^t\int_{\mathbb{T}^N}|u^\varepsilon|^{p-2}u^\varepsilon\mathrm{div}\big(A(u^\varepsilon)\nabla u^\varepsilon\big)\mathrm{d}x\,\mathrm{d}s\\&=-p\int_0^t\int_{\mathbb{T}^N}|u^\varepsilon|^{p-2}\big(\nabla u^\varepsilon\big)^*A(x)\big(\nabla u^\varepsilon\big)\,\mathrm{d}x\,\mathrm{d}s\le 0\end{aligned}$$

and the same holds for the fourth term as well since A is only replaced by εI. The last term is estimated as follows

$$\begin{aligned}\frac{1}{2}p(p-1)\int_0^t\int_{\mathbb{T}^N}|u^\varepsilon|^{p-2}G_\varepsilon^2(x,u^\varepsilon)\,\mathrm{d}x\,\mathrm{d}s&\le C\int_0^t\int_{\mathbb{T}^N}|u^\varepsilon|^{p-2}\big(1+|u^\varepsilon|^2\big)\,\mathrm{d}x\,\mathrm{d}s\\&\le C\Big(1+\int_0^t\|u^\varepsilon(s)\|^p_{L^p(\mathbb{T}^N)}\mathrm{d}s\Big).\end{aligned}$$

Finally, expectation and application of the Gronwall lemma yield (27). □

Corollary 4.3. *The set $\{u^\varepsilon;\ \varepsilon\in(0,1)\}$ is bounded in $L^p(\Omega;C([0,T];L^p(\mathbb{T}^N)))$, for all $p\in[2,\infty)$.*

Proof. Continuity of trajectories follows from Theorem 4.1. To verify the claim, a uniform estimate of $\mathbb{E}\sup_{0\le t\le T}\|u^\varepsilon(t)\|^p_{L^p(\mathbb{T}^N)}$ is needed. We repeat the approach from the preceding lemma, only for the stochastically forced term we apply the Burkholder–Davis–Gundy inequality. We have

$$\begin{aligned}\mathbb{E}\sup_{0\le t\le T}\|u^\varepsilon(t)\|^p_{L^p(\mathbb{T}^N)}&\le\mathbb{E}\|u_0\|^p_{L^p(\mathbb{T}^N)}+C\bigg(1+\int_0^T\mathbb{E}\|u^\varepsilon(s)\|^p_{L^p(\mathbb{T}^N)}\mathrm{d}s\bigg)\\&\quad+p\,\mathbb{E}\sup_{0\le t\le T}\bigg|\sum_{k\ge1}\int_0^t\int_{\mathbb{T}^N}|u^\varepsilon|^{p-2}u^\varepsilon g_k^\varepsilon(x,u^\varepsilon)\,\mathrm{d}x\,\mathrm{d}\beta_k(s)\bigg|\end{aligned}$$

and using the Burkholder–Davis–Gundy and the Schwartz inequality, the assumption (2) and the weighted Young inequality in the last step yield

$$\begin{aligned}\mathbb{E}&\sup_{0\le t\le T}\bigg|\sum_{k\ge1}\int_0^t\int_{\mathbb{T}^N}|u^\varepsilon|^{p-2}u^\varepsilon g_k^\varepsilon(x,u^\varepsilon)\,\mathrm{d}x\,\mathrm{d}\beta_k(s)\bigg|\\&\le C\,\mathbb{E}\bigg(\int_0^T\sum_{k\ge1}\bigg(\int_{\mathbb{T}^N}|u^\varepsilon|^{p-1}|g_k^\varepsilon(x,u^\varepsilon)|\,\mathrm{d}x\bigg)^2\mathrm{d}s\bigg)^{\frac12}\\&\le C\,\mathbb{E}\bigg(\int_0^T\Big\||u^\varepsilon|^{\frac p2}\Big\|^2_{L^2(\mathbb{T}^N)}\sum_{k\ge1}\Big\||u^\varepsilon|^{\frac{p-2}{2}}|g_k^\varepsilon(\cdot,u^\varepsilon(\cdot))|\Big\|^2_{L^2(\mathbb{T}^N)}\mathrm{d}s\bigg)^{\frac12}\\&\le C\,\mathbb{E}\bigg(\int_0^T\|u^\varepsilon\|^p_{L^p(\mathbb{T}^N)}\Big(1+\|u^\varepsilon\|^p_{L^p(\mathbb{T}^N)}\Big)\,\mathrm{d}s\bigg)^{\frac12}\end{aligned}$$

$$\leq C\,\mathbb{E}\bigg(\sup_{0\leq t\leq T}\|u^\varepsilon(t)\|^p_{L^p(\mathbb{T}^N)}\bigg)^{\frac{1}{2}}\bigg(1+\int_0^T\|u^\varepsilon(s)\|^p_{L^p(\mathbb{T}^N)}\,\mathrm{d}s\bigg)^{\frac{1}{2}}$$
$$\leq \frac{1}{2}\,\mathbb{E}\sup_{0\leq t\leq T}\|u^\varepsilon(t)\|^p_{L^p(\mathbb{T}^N)}+C\bigg(1+\int_0^T\mathbb{E}\|u^\varepsilon(s)\|^p_{L^p(\mathbb{T}^N)}\,\mathrm{d}s\bigg).$$

Therefore

$$\mathbb{E}\sup_{0\leq t\leq T}\|u^\varepsilon(t)\|^p_{L^p(\mathbb{T}^N)}\leq C\bigg(1+\mathbb{E}\|u_0\|^p_{L^p(\mathbb{T}^N)}+\int_0^T\mathbb{E}\|u^\varepsilon(s)\|^p_{L^p(\mathbb{T}^N)}\,\mathrm{d}s\bigg)$$

and the corollary follows from (27). □

4.3. Compactness argument

To show that there exists $u:\Omega\times\mathbb{T}^N\times[0,T]\to\mathbb{R}$, a kinetic solution to (1), one needs to verify the strong convergence of the approximate solutions u^ε. This can be done by combining tightness of their laws with the pathwise uniqueness, which was proved above.

First, we need to prove a better spatial regularity of the approximate solutions. Towards this end, we introduce two seminorms describing the $W^{\lambda,1}$-regularity of a function $u\in L^1(\mathbb{T}^N)$. Let $\lambda\in(0,1)$ and define

$$p^\lambda(u)=\int_{\mathbb{T}^N}\int_{\mathbb{T}^N}\frac{|u(x)-u(y)|}{|x-y|^{N+\lambda}}\,\mathrm{d}x\,\mathrm{d}y,$$
$$p^\lambda_\varrho(u)=\sup_{0<\tau<2D_N}\frac{1}{\tau^\lambda}\int_{\mathbb{T}^N}\int_{\mathbb{T}^N}|u(x)-u(y)|\varrho_\tau(x-y)\,\mathrm{d}x\,\mathrm{d}y,$$

where (ϱ_τ) is the approximation to the identity on $\mathbb{T}^N$ (as introduced in the proof of Lemma 2.5) that is radial, i.e. $\varrho_\tau(x)=1/\tau^N\varrho(|x|/\tau)$; and by D_N we denote the diameter of $[0,1]^N$. The fractional Sobolev space $W^{\lambda,1}(\mathbb{T}^N)$ is defined as a subspace of $L^1(\mathbb{T}^N)$ with finite norm

$$\|u\|_{W^{\lambda,1}(\mathbb{T}^N)}=\|u\|_{L^1(\mathbb{T}^N)}+p^\lambda(u).$$

According to [9], the following relations hold true between these seminorms. Let $s\in(0,\lambda)$; there exists a constant $C=C_{\lambda,\varrho,N}$ such that for all $u\in L^1(\mathbb{T}^N)$

$$p^\lambda_\varrho(u)\leq Cp^\lambda(u),\qquad p^s(u)\leq\frac{C}{\lambda-s}\,p^\lambda_\varrho(u).\tag{29}$$

Theorem 4.4 ($W^{\varsigma,1}$*-Regularity*). *Set $\varsigma=\min\{\frac{2\alpha}{\alpha+1},\frac{1}{2}\}$, where α was introduced in (4). Then for all $s\in(0,\varsigma)$ there exists a constant $C_{T,s}>0$ such that for all $t\in[0,T]$ and all $\varepsilon\in(0,1)$*

$$\mathbb{E}\,p^s\big(u^\varepsilon(t)\big)\leq C_{T,s}\big(1+\mathbb{E}\,p^\varsigma(u_0)\big).\tag{30}$$

In particular, there exists a constant $C_{T,s,u_0}>0$ such that for all $t\in[0,T]$

$$\mathbb{E}\|u^\varepsilon(t)\|_{W^{s,1}(\mathbb{T}^N)}\leq C_{T,s,u_0}\big(1+\mathbb{E}\|u_0\|_{W^{\varsigma,1}(\mathbb{T}^N)}\big).\tag{31}$$

Proof. Proof of this statement is based on Proposition 3.2. We have

$$\mathbb{E}\int_{(\mathbb{T}^N)^2}\int_{\mathbb{R}}\varrho_\tau(x-y)f^\varepsilon(x,t,\xi)\bar{f}^\varepsilon(y,t,\xi)\,\mathrm{d}\xi\,\mathrm{d}x\,\mathrm{d}y$$

$$\le \mathbb{E}\int_{(\mathbb{T}^N)^2}\int_{\mathbb{R}^2}\varrho_\tau(x-y)\psi_\delta(\xi-\zeta)f^\varepsilon(x,t,\xi)\bar{f}^\varepsilon(y,t,\zeta)\,\mathrm{d}\xi\,\mathrm{d}\zeta\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}t+\delta$$
$$\le \mathbb{E}\int_{(\mathbb{T}^N)^2}\int_{\mathbb{R}^2}\varrho_\tau(x-y)\psi_\delta(\xi-\zeta)f_0(x,\xi)\bar{f}_0(y,\zeta)\mathrm{d}\xi\mathrm{d}\zeta\mathrm{d}x\mathrm{d}y+\delta+\mathrm{I}^\varepsilon+\mathrm{J}^\varepsilon+\mathrm{K}^\varepsilon$$
$$\le \mathbb{E}\int_{(\mathbb{T}^N)^2}\int_{\mathbb{R}}\varrho_\tau(x-y)f_0(x,\xi)\bar{f}_0(y,\xi)\,\mathrm{d}\xi\,\mathrm{d}x\,\mathrm{d}y+2\delta+\mathrm{I}^\varepsilon+\mathrm{J}^\varepsilon+\mathrm{K}^\varepsilon,$$

where $\mathrm{I}^\varepsilon,\mathrm{J}^\varepsilon,\mathrm{K}^\varepsilon$ are defined correspondingly to $\mathrm{I},\mathrm{J},\mathrm{K}$ in Proposition 3.2 but using the approximated coefficients B^ε, A^ε, Φ^ε instead. From the same estimates as the ones used in the proof of Theorem 3.3, we conclude

$$\begin{aligned}&\mathbb{E}\int_{(\mathbb{T}^N)^2}\varrho_\tau(x-y)\big(u^\varepsilon(x,t)-u^\varepsilon(y,t)\big)^+\,\mathrm{d}x\,\mathrm{d}y\\&\quad\le \mathbb{E}\int_{(\mathbb{T}^N)^2}\varrho_\tau(x-y)\big(u_0(x)-u_0(y)\big)^+\,\mathrm{d}x\mathrm{d}y+2\delta+Ct\big(\delta^{-1}\tau+\delta^{-1}\tau^2+\delta^\alpha\big)+\mathrm{J}^\varepsilon.\end{aligned}$$

In order to control the term J^ε, recall that (keeping the notation from Theorem 3.3)

$$\begin{aligned}\mathrm{J}^\varepsilon&=-\mathbb{E}\int_0^t\int_{(\mathbb{T}^N)^2}(\nabla_x u^\varepsilon)^*\sigma(x)\sigma(x)(\nabla\varrho_\tau)(x-y)\,\Theta_\delta\big(u^\varepsilon(x)-u^\varepsilon(y)\big)\mathrm{d}x\mathrm{d}y\mathrm{d}r,\\&\quad-\mathbb{E}\int_0^t\int_{(\mathbb{T}^N)^2}(\nabla_y u^\varepsilon)^*\sigma(y)\sigma(y)(\nabla\varrho_\tau)(x-y)\,\Theta_\delta\big(u^\varepsilon(x)-u^\varepsilon(y)\big)\mathrm{d}x\mathrm{d}y\mathrm{d}r,\\&\quad-\mathbb{E}\int_0^t\int_{(\mathbb{T}^N)^2}\big[|\sigma(x)\nabla_x u^\varepsilon|^2+|\sigma(y)\nabla_y u^\varepsilon|^2\big]\varrho_\tau(x-y)\psi_\delta\big(u^\varepsilon(x)-u^\varepsilon(y)\big)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}r\\&\quad-\varepsilon\,\mathbb{E}\int_0^t\int_{(\mathbb{T}^N)^2}\big|\nabla_x u^\varepsilon-\nabla_y u^\varepsilon\big|^2\varrho_\tau(x-y)\psi_\delta\big(u^\varepsilon(x)-u^\varepsilon(y)\big)\,\mathrm{d}x\mathrm{d}y\mathrm{d}r\\&=\mathrm{J}_1+\mathrm{J}_2+\mathrm{J}_3+\mathrm{J}_4.\end{aligned}$$

The first three terms on the above right hand side correspond to the diffusion term $\operatorname{div}(A(x)\nabla u^\varepsilon)$. Since all u^ε are smooth and hence the chain rule formula is not an issue here, J_4 is obtained after integration by parts from similar terms corresponding to $\varepsilon\Delta u^\varepsilon$. Next, we have

$$\begin{aligned}\mathrm{J}_1&=-\mathbb{E}\int_0^t\int_{(\mathbb{T}^N)^2}(\nabla_x u^\varepsilon)^*\sigma(x)\operatorname{div}_y\Big(\sigma(y)\,\Theta_\delta\big(u^\varepsilon(x)-u^\varepsilon(y)\big)\Big)\varrho_\tau(x-y)\mathrm{d}x\mathrm{d}y\mathrm{d}r\\&\quad+\mathbb{E}\int_0^t\int_{(\mathbb{T}^N)^2}(\nabla_x u^\varepsilon)^*\sigma(x)\Big(\sigma(y)-\sigma(x)\Big)(\nabla\varrho_\tau)(x-y)\\&\quad\times\Theta_\delta\big(u^\varepsilon(x)-u^\varepsilon(y)\big)\mathrm{d}x\mathrm{d}y\mathrm{d}r\end{aligned}$$

and

$$\begin{aligned}\mathrm{J}_2&=\mathbb{E}\int_0^t\int_{(\mathbb{T}^N)^2}(\nabla_y u^\varepsilon)^*\sigma(y)\operatorname{div}_x\Big(\sigma(x)\,\Theta_\delta\big(u^\varepsilon(x)-u^\varepsilon(y)\big)\Big)\varrho_\tau(x-y)\mathrm{d}x\mathrm{d}y\mathrm{d}r\\&\quad+\mathbb{E}\int_0^t\int_{(\mathbb{T}^N)^2}(\nabla_y u^\varepsilon)^*\sigma(y)\Big(\sigma(x)-\sigma(y)\Big)(\nabla\varrho_\tau)(x-y)\\&\quad\times\Theta_\delta\big(u^\varepsilon(x)-u^\varepsilon(y)\big)\mathrm{d}x\mathrm{d}y\mathrm{d}r;\end{aligned}$$

hence $\mathrm{J}_1 = \mathrm{H} + \mathrm{R}_1$ and $\mathrm{J}_2 = \mathrm{H} + \mathrm{R}_2$ where

$$\mathrm{H} = \mathbb{E}\int_0^t \int_{(\mathbb{T}^N)^2} (\nabla_x u^\varepsilon)^* \sigma(x)\sigma(y)(\nabla_y u^\varepsilon)\varrho_\tau(x-y)\psi_\delta\big(u^\varepsilon(x) - u^\varepsilon(y)\big)\mathrm{d}x\mathrm{d}y\mathrm{d}r$$

$$\begin{aligned}\mathrm{R}_1 \ =\ &\mathbb{E}\int_0^t \int_{(\mathbb{T}^N)^2} (\nabla_x u^\varepsilon)^* \sigma(x)\Theta_\delta\big(u^\varepsilon(x) - u^\varepsilon(y)\big)\\ &\times\Big(\big(\sigma(y)-\sigma(x)\big)(\nabla\varrho_\tau)(x-y) - \mathrm{div}\big(\sigma(y)\big)\varrho_\tau(x-y)\Big)\mathrm{d}x\mathrm{d}y\mathrm{d}r\end{aligned}$$

$$\begin{aligned}\mathrm{R}_2 \ =\ &\mathbb{E}\int_0^t \int_{(\mathbb{T}^N)^2} (\nabla_y u^\varepsilon)^* \sigma(y)\Theta_\delta\big(u^\varepsilon(x) - u^\varepsilon(y)\big)\\ &\times\Big(\big(\sigma(x)-\sigma(y)\big)(\nabla\varrho_\tau)(x-y) + \mathrm{div}\big(\sigma(x)\big)\varrho_\tau(x-y)\Big)\mathrm{d}x\mathrm{d}y\mathrm{d}r.\end{aligned}$$

As a consequence, we see that $\mathrm{J}^\varepsilon = \mathrm{J}_4 + \mathrm{J}_5 + \mathrm{R}_1 + \mathrm{R}_2$ where

$$\mathrm{J}_5 = -\mathbb{E}\int_0^t \int_{(\mathbb{T}^N)^2} \big|\sigma(x)\nabla_x u^\varepsilon - \sigma(y)\nabla_y u^\varepsilon\big|^2 \varrho_t(x-y)\psi_\delta\big(u^\varepsilon(x) - u^\varepsilon(y)\big)\,\mathrm{d}x\mathrm{d}y\mathrm{d}r$$

and therefore $\mathrm{J}^\varepsilon \le \mathrm{R}_1 + \mathrm{R}_2$. Let us introduce an auxiliary function

$$T_\delta(\xi) = \int_0^\xi \Theta_\delta(\zeta)\,\mathrm{d}\zeta.$$

With this in hand we obtain

$$\begin{aligned}\mathrm{R}_1 &= \mathbb{E}\int_0^t \int_{(\mathbb{T}^N)^2} \sigma(x)\nabla_x T_\delta\big(u^\varepsilon(x) - u^\varepsilon(y)\big)\\ &\quad\times\Big(\big(\sigma(y)-\sigma(x)\big)(\nabla\varrho_\tau)(x-y) - \mathrm{div}\big(\sigma(y)\big)\varrho_\tau(x-y)\Big)\mathrm{d}x\mathrm{d}y\mathrm{d}r\\ &= -\mathbb{E}\int_0^t \int_{(\mathbb{T}^N)^2} T_\delta\big(u^\varepsilon(x) - u^\varepsilon(y)\big)\Big[\mathrm{div}\big(\sigma(x)\big)\big(\sigma(y)-\sigma(x)\big)(\nabla\varrho_\tau)(x-y)\\ &\quad-\sigma(x)\mathrm{div}\big(\sigma(x)\big)(\nabla\varrho_\tau)(x-y) + \sigma(x)\big(\sigma(y)-\sigma(x)\big)(\nabla^2\varrho_\tau)(x-y)\\ &\quad-\mathrm{div}\big(\sigma(x)\big)\mathrm{div}\big(\sigma(y)\big)\varrho_\tau(x-y) - \sigma(x)\mathrm{div}\big(\sigma(y)\big)(\nabla\varrho_\tau)(x-y)\Big]\mathrm{d}x\mathrm{d}y\mathrm{d}r\end{aligned}$$

and similarly

$$\begin{aligned}\mathrm{R}_2 &= \mathbb{E}\int_0^t \int_{(\mathbb{T}^N)^2} T_\delta\big(u^\varepsilon(x) - u^\varepsilon(y)\big)\Big[\mathrm{div}\big(\sigma(y)\big)\big(\sigma(x)-\sigma(y)\big)(\nabla\varrho_\tau)(x-y)\\ &\quad-\sigma(y)\mathrm{div}\big(\sigma(y)\big)(\nabla\varrho_\tau)(x-y) - \sigma(y)\big(\sigma(x)-\sigma(y)\big)(\nabla^2\varrho_\tau)(x-y)\\ &\quad+\mathrm{div}\big(\sigma(y)\big)\mathrm{div}\big(\sigma(x)\big)\varrho_\tau(x-y) - \sigma(y)\mathrm{div}\big(\sigma(x)\big)(\nabla\varrho_\tau)(x-y)\Big]\mathrm{d}x\mathrm{d}y\mathrm{d}r;\end{aligned}$$

hence

$$\begin{aligned}\mathrm{R}_1 + \mathrm{R}_2 = \ &\mathbb{E}\int_0^t \int_{(\mathbb{T}^N)^2} T_\delta\big(u^\varepsilon(x) - u^\varepsilon(y)\big)\\ &\times\Big[2\big(\mathrm{div}\big(\sigma(x)\big) + \mathrm{div}\big(\sigma(y)\big)\big)\big(\sigma(x)-\sigma(y)\big)(\nabla\varrho_\tau)(x-y)\\ &+\big(\sigma(x)-\sigma(y)\big)^2(\nabla^2\varrho_\tau)(x-y) + 2\mathrm{div}\big(\sigma(x)\big)\mathrm{div}\big(\sigma(y)\big)\varrho_\tau(x-y)\Big]\mathrm{d}x\mathrm{d}y\mathrm{d}r.\end{aligned}$$

Since $|T_\delta(\xi)| \leq |\xi|, \tau|\nabla\varrho_\tau|(\cdot) \leq C\varrho_{2\tau}(\cdot)$ and $\tau^2|\nabla^2\varrho_\tau|(\cdot) \leq C\varrho_{2\tau}(\cdot)$ with constants independent on τ, we deduce that

$$\mathrm{J}^\varepsilon \leq \mathrm{R}_1 + \mathrm{R}_2 \leq C_\sigma\, \mathbb{E}\int_0^t \int_{(\mathbb{T}^N)^2} \varrho_{2\tau}(x-y)\big|u^\varepsilon(x) - u^\varepsilon(y)\big|\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}r$$

and therefore

$$\begin{aligned}
&\mathbb{E}\int_{(\mathbb{T}^N)^2} \varrho_\tau(x-y)\big|u^\varepsilon(x,t) - u^\varepsilon(y,t)\big|\,\mathrm{d}x\,\mathrm{d}y \\
&\quad\leq \mathbb{E}\int_{(\mathbb{T}^N)^2} \varrho_\tau(x-y)\big|u_0(x) - u_0(y)\big|\,\mathrm{d}x\,\mathrm{d}y + C_T\big(\delta + \delta^{-1}\tau + \delta^{-1}\tau^2 + \delta^\alpha\big) \\
&\qquad + C_\sigma\, \mathbb{E}\int_0^t \int_{(\mathbb{T}^N)^2} \varrho_{2\tau}(x-y)\big|u^\varepsilon(x,s) - u^\varepsilon(y,s)\big|\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}r.
\end{aligned}$$

By optimization in δ, i.e. setting $\delta = \tau^\beta$, we obtain

$$\sup_{0<\tau<2D_N} \frac{C_T\big(\delta + \delta\tau^{-1} + \delta^{-1}\tau^2 + \delta^\alpha\big)}{\tau^\varsigma} \leq C_T,$$

where the maximal choice of the parameter ς is $\min\big\{\frac{2\alpha}{\alpha+1}, \frac{1}{2}\big\}$. As a consequence,

$$\begin{aligned}
&\mathbb{E}\int_{(\mathbb{T}^N)^2} \varrho_\tau(x-y)\big|u^\varepsilon(x,t) - u^\varepsilon(y,t)\big|\,\mathrm{d}x\,\mathrm{d}y \\
&\quad\leq C_T\left(\tau^\varsigma + \tau^\varsigma\, \mathbb{E}\, p^\varsigma(u_0) + \mathbb{E}\int_0^t \int_{\mathbb{T}^N} \varrho_{2\tau}(x-y)\big|u^\varepsilon(x,r) - u^\varepsilon(y,r)\big|\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}r\right).
\end{aligned}$$

Let us multiply the above by τ^{-1-s}, $s \in (0,\varsigma)$, and integrate with respect to $\tau \in (0, 2D_N)$. As $|x-y| \leq \tau$ on the left hand side, we can estimate from below

$$\int_{|x-y|}^{2D_N} \frac{1}{\tau^{1+s}}\,\varrho_\tau(x-y)\mathrm{d}\tau = \frac{1}{|x-y|^{N+s}} \int_{|x-y|/2D_N}^1 \lambda^{N+s-1}\varrho(\lambda)\mathrm{d}\lambda \geq \frac{C_s}{|x-y|^{N+s}}$$

and similarly for the last term on the right hand side we estimate from above

$$\int_{|x-y|/2}^{2D_N} \frac{1}{\tau^{1+s}}\,\varrho_{2\tau}(x-y)\,\mathrm{d}\tau \leq \frac{C_s}{|x-y|^{N+s}}.$$

Accordingly,

$$\mathbb{E}\, p^s\big(u^\varepsilon(t)\big) \leq C_{T,s}\left(1 + \mathbb{E}\, p^\varsigma(u_0) + \mathbb{E}\int_0^t p^s\big(u^\varepsilon(r)\big)\,\mathrm{d}r\right)$$

and (30) follows by the Gronwall lemma. Furthermore, due to (27)

$$\mathbb{E}\|u^\varepsilon(t)\|_{L^1(\mathbb{T}^N)} \leq \mathbb{E}\|u^\varepsilon(t)\|_{L^2(\mathbb{T}^N)} \leq C\Big(1 + \big(\mathbb{E}\|u_0\|^2_{L^2(\mathbb{T}^N)}\big)^{\frac{1}{2}}\Big);$$

hence we obtain (31). As a consequence of the previous estimate, the constant in (31) depends on the $L^2(\Omega; L^2(\mathbb{T}^N))$-norm of the initial condition. □

Corollary 4.5. *For all $\gamma\in(0,\varsigma)$ and $q>1$ satisfying $\gamma q<\sigma$, there exists a constant $C>0$ such that for all $\varepsilon\in(0,1)$*

$$\mathbb{E}\|u^\varepsilon\|^q_{L^q(0,T;W^{\gamma,q}(\mathbb{T}^N))}\leq C. \tag{32}$$

Proof. The claim is a consequence of the bounds (27) and (31). Indeed, fix $\gamma\in(0,\varsigma)$ and $q\in(1,\infty)$. We will use an interpolation inequality:

$$\|\cdot\|_{W^{\gamma,q}(\mathbb{T}^N)}\leq C\|\cdot\|^{1-\theta}_{W^{\gamma_0,q_0}(\mathbb{T}^N)}\|\cdot\|^{\theta}_{W^{\gamma_1,q_1}(\mathbb{T}^N)}, \tag{33}$$

where $\gamma_0,\gamma_1\in\mathbb{R}$, $q_0,q_1\in(0,\infty)$, $\gamma=(1-\theta)\gamma_0+\theta\gamma_1$, $\frac{1}{q}=\frac{1-\theta}{q_0}+\frac{\theta}{q_1}$, $\theta\in(0,1)$, which follows from a more general result [31, Theorem 1.6.7] or [30, Theorem 2.4.1]. Fix $s\in(\gamma q,\varsigma)$ and set $\gamma_0=s$, $\gamma_1=0$, $q_0=1$, $q_1=p$. Then we obtain $\theta=\frac{s-\gamma}{s}$, $p=\frac{(s-\gamma)q}{s-\gamma q}$ and

$$\begin{aligned}\mathbb{E}\int_0^T\|u^\varepsilon(t)\|^q_{W^{\gamma,q}(\mathbb{T}^N)}\,\mathrm{d}t&\leq C\,\mathbb{E}\int_0^T\Big(\|u^\varepsilon(t)\|^{(1-\theta)q}_{W^{s,1}(\mathbb{T}^N)}\|u^\varepsilon(t)\|^{\theta q}_{L^p(\mathbb{T}^N)}\Big)\,\mathrm{d}t\\&\leq C\Big(\mathbb{E}\|u^\varepsilon(t)\|_{L^1(0,T;W^{s,1}(\mathbb{T}^N))}\Big)^{(1-\theta)q}\Big(\mathbb{E}\|u^\varepsilon(t)\|^p_{L^p(0,T;L^p(\mathbb{T}^N))}\Big)^{1-(1-\theta)q}\leq C.\quad\square\end{aligned}$$

Also a better time regularity is needed.

Lemma 4.6. *Suppose that $\lambda\in(0,1/2)$, $q\in[2,\infty)$. There exists a constant $C>0$ such that for all $\varepsilon\in(0,1)$*

$$\mathbb{E}\|u^\varepsilon\|^q_{C^\lambda([0,T];H^{-2}(\mathbb{T}^N))}\leq C. \tag{34}$$

Proof. Let $q\in[2,\infty)$. Recall that the set $\{u^\varepsilon;\,\varepsilon\in(0,1)\}$ is bounded in $L^q(\Omega;C(0,T;L^q(\mathbb{T}^N)))$. Since all B^ε have the same polynomial growth we conclude, in particular, that

$$\{\operatorname{div}(B^\varepsilon(u^\varepsilon))\},\qquad\{\operatorname{div}(A(x)\nabla u^\varepsilon)\},\qquad\{\varepsilon\Delta u^\varepsilon\}$$

are bounded in $L^q(\Omega;C(0,T;H^{-2}(\mathbb{T}^N)))$ and consequently

$$\mathbb{E}\left\|u^\varepsilon-\int_0^\cdot\Phi^\varepsilon(u^\varepsilon)\,\mathrm{d}W\right\|^q_{C^1([0,T];H^{-2}(\mathbb{T}^N))}\leq C.$$

In order to deal with the stochastic integral, let us recall the definition of the Riemann–Liouville operator: let X be a Banach space, $p\in(1,\infty]$, $\alpha\in(1/p,1]$ and $f\in L^p(0,T;X)$, then we define

$$\big(R_\alpha f\big)(t)=\frac{1}{\Gamma(\alpha)}\int_0^t(t-s)^{\alpha-1}f(s)\,\mathrm{d}s,\qquad t\in[0,T].$$

It is well known that R_α is a bounded linear operator from $L^p(0,T;X)$ to the space of Hölder continuous functions $C^{\alpha-1/p}([0,T];X)$ (see e.g. [29, Theorem 3.6]). Assume now that $q\in(2,\infty)$, $\alpha\in(1/q,1/2)$. Then according to the stochastic Fubini theorem [8, Theorem 4.18]

$$\int_0^t\Phi^\varepsilon\big(u^\varepsilon(s)\big)\,\mathrm{d}W(s)=\big(R_\alpha Z\big)(t),$$

where

$$Z(s) = \frac{1}{\Gamma(1-\alpha)} \int_0^s (s-r)^{-\alpha}\, \Phi^\varepsilon\big(u^\varepsilon(r)\big)\, \mathrm{d}W(r).$$

Therefore using the Burkholder–Davis–Gundy and the Young inequality and the estimate (2), we have

$$\begin{aligned}
\mathbb{E}\left\| \int_0^{\cdot} \Phi^\varepsilon(u^\varepsilon)\, \mathrm{d}W \right\|^q_{C^{\alpha-1/q}([0,T];L^2(\mathbb{T}^N))} &\le C\, \mathbb{E}\|Z\|^q_{L^q(0,T;L^2(\mathbb{T}^N))} \\
&\le C \int_0^T \mathbb{E}\left(\int_0^t \frac{1}{(t-s)^{2\alpha}} \|\Phi^\varepsilon(u^\varepsilon)\|^2_{L_2(\mathfrak{U};L^2(\mathbb{T}^N))} \mathrm{d}s \right)^{\frac{q}{2}} \mathrm{d}t \\
&\le C T^{\frac{q}{2}(1-2\alpha)} \mathbb{E} \int_0^T \Big(1 + \|u^\varepsilon(s)\|^q_{L^2(\mathbb{T}^N)}\Big) \mathrm{d}s \\
&\le C T^{\frac{q}{2}(1-2\alpha)} \Big(1 + \|u^\varepsilon\|^q_{L^q(\Omega;L^q(0,T;L^2(\mathbb{T}^N)))}\Big) \le C
\end{aligned}$$

and the claim follows. □

Corollary 4.7. *For all $\vartheta > 0$ there exist $\beta > 0$ and $C > 0$ such that for all $\varepsilon \in (0,1)$*

$$\mathbb{E}\|u^\varepsilon\|_{C^\beta([0,T];H^{-\vartheta}(\mathbb{T}^N))} \le C. \tag{35}$$

Proof. If $\vartheta > 2$, the claim follows easily from (34) by the choice $\beta = \lambda$. If $\vartheta \in (0,2)$ the proof follows easily from interpolation between $H^{-2}(\mathbb{T}^N)$ and $L^2(\mathbb{T}^N)$. Indeed,

$$\begin{aligned}
\mathbb{E} \sup_{0\le t\le T} \|u^\varepsilon(t)\|_{H^{-\vartheta}(\mathbb{T}^N)} &\le C\, \mathbb{E}\left(\sup_{0\le t\le T} \|u^\varepsilon(t)\|^{1-\theta}_{H^{-2}(\mathbb{T}^N)} \sup_{0\le t\le T} \|u^\varepsilon(t)\|^{\theta}_{L^2(\mathbb{T}^N)} \right) \\
&\le C \left(\mathbb{E}\left(\sup_{0\le t\le T} \|u^\varepsilon(t)\|^{1-\theta}_{H^{-2}(\mathbb{T}^N)} \right)^p \right)^{\frac{1}{p}} \left(\mathbb{E}\left(\sup_{0\le t\le T} \|u^\varepsilon(t)\|^{\theta}_{L^2(\mathbb{T}^N)} \right)^q \right)^{\frac{1}{q}} \\
&\le C \left(1 + \mathbb{E} \sup_{0\le t\le T} \|u^\varepsilon(t)\|^{(1-\theta)p}_{H^{-2}(\mathbb{T}^N)} \right)^{\frac{1}{p}} \left(1 + \mathbb{E} \sup_{0\le t\le T} \|u^\varepsilon(t)\|^{\theta q}_{L^2(\mathbb{T}^N)} \right)^{\frac{1}{q}}
\end{aligned}$$

where the exponent p for the Hölder inequality is chosen in order to satisfy $(1-\theta)p = 1$, i.e. since $\theta = \frac{2-\vartheta}{2}$, we have $p = \frac{2}{\vartheta}$. The first parenthesis can be estimated using (34) while the second one using (27). Similar computations yield the second part of the norm of $C^\beta([0,T];H^{-\vartheta}(\mathbb{T}^N))$. Indeed,

$$\begin{aligned}
&\mathbb{E} \sup_{\substack{0\le s,t\le T\\ s\ne t}} \frac{\|u^\varepsilon(t) - u^\varepsilon(s)\|_{H^{-\vartheta}(\mathbb{T}^N)}}{|t-s|^\beta} \\
&\le C\, \mathbb{E}\left(\sup_{\substack{0\le s,t\le T\\ s\ne t}} \frac{\|u^\varepsilon(t) - u^\varepsilon(s)\|^{1-\theta}_{H^{-2}(\mathbb{T}^N)}}{|t-s|^\beta} \sup_{\substack{0\le s,t\le T\\ s\ne t}} \|u^\varepsilon(t) - u^\varepsilon(s)\|^{\theta}_{L^2(\mathbb{T}^N)} \right) \\
&\le C \left(1 + \mathbb{E} \sup_{\substack{0\le s,t\le T\\ s\ne t}} \frac{\|u^\varepsilon(t) - u^\varepsilon(s)\|^{(1-\theta)p}_{H^{-2}(\mathbb{T}^N)}}{|t-s|^{\beta p}} \right)^{\frac{1}{p}} \left(1 + \mathbb{E} \sup_{0\le t\le T} \|u^\varepsilon(t)\|^{\theta q}_{L^2(\mathbb{T}^N)} \right)^{\frac{1}{q}}
\end{aligned}$$

where the same choice $p = \frac{2}{\vartheta}$ and the condition $\beta p \in (0, \frac{1}{2})$, which is needed for (34), gives (35) for $\beta \in (0, \frac{\vartheta}{4})$. □

Corollary 4.8. *Suppose that* $\kappa \in (0, \frac{\varsigma}{2(4+\varsigma)})$*. There exists a constant* $C > 0$ *such that for all* $\varepsilon \in (0, 1)$

$$\mathbb{E}\|u^\varepsilon\|_{H^\kappa(0,T;L^2(\mathbb{T}^N))} \le C. \tag{36}$$

Proof. It follows from Lemma 4.6 that

$$\mathbb{E}\|u^\varepsilon\|^q_{H^\lambda(0,T;H^{-2}(\mathbb{T}^N))} \le C, \tag{37}$$

where $\lambda \in (0, 1/2)$, $q \in [1, \infty)$. Let $\gamma \in (0, \varsigma/2)$. If $\kappa = \theta\lambda$ and $0 = -2\theta + (1-\theta)\gamma$ then it follows by the interpolation (see [1, Theorem 3.1]) and the Hölder inequality that

$$\begin{aligned}\mathbb{E}\|u^\varepsilon\|_{H^\kappa(0,T;L^2(\mathbb{T}^N))} &\le C\,\mathbb{E}\Big(\|u^\varepsilon\|^\theta_{H^\lambda(0,T;H^{-2}(\mathbb{T}^N))}\|u^\varepsilon\|^{1-\theta}_{L^2(0,T;H^\gamma(\mathbb{T}^N))}\Big)\\ &\le C\Big(\mathbb{E}\|u^\varepsilon\|^{\theta p}_{H^\lambda(0,T;H^{-2}(\mathbb{T}^N))}\Big)^{\frac{1}{p}}\Big(\mathbb{E}\|u^\varepsilon\|^{(1-\theta)r}_{L^2(0,T;H^\gamma(\mathbb{T}^N))}\Big)^{\frac{1}{r}},\end{aligned}$$

where the exponent r is chosen in order to satisfy $(1-\theta)r = 2$. The proof now follows from (32) and (37). □

Now, we have all in hand to conclude our compactness argument by showing tightness of a certain collection of laws. First, let us introduce some notation which will be used later on. If E is a Banach space and $t \in [0, T]$, we consider the space of continuous E-valued functions and denote by ϱ_t the operator of restriction to the interval $[0, t]$. To be more precise, we define

$$\begin{aligned}\varrho_t : C([0, T]; E) &\longrightarrow C([0, t]; E)\\ k &\longmapsto k|_{[0,t]}.\end{aligned} \tag{38}$$

Plainly, ϱ_t is a continuous mapping. Let us define the path space

$$\mathcal{X}_u = \Big\{u \in L^2\big(0, T; L^2(\mathbb{T}^N)\big) \cap C\big([0, T]; H^{-1}(\mathbb{T}^N)\big);\ \varrho_0 u \in L^2(\mathbb{T}^N)\Big\}$$

equipped with the norm

$$\|\cdot\|_{\mathcal{X}_u} = \|\cdot\|_{L^2(0,T;L^2(\mathbb{T}^N))} + \|\cdot\|_{C([0,T];H^{-1}(\mathbb{T}^N))} + \|\varrho_0\cdot\|_{L^2(\mathbb{T}^N)}.$$

Next, we set $\mathcal{X}_W = C([0, T]; \mathfrak{U}_0)$ and $\mathcal{X} = \mathcal{X}_u \times \mathcal{X}_W$. Let μ_{u^ε} denote the law of u^ε on $\mathcal{X}_u$, $\varepsilon \in (0, 1)$, and μ_W the law of W on $\mathcal{X}_W$. Their joint law on $\mathcal{X}$ is then denoted by μ^ε.

Theorem 4.9. *The set* $\{\mu^\varepsilon;\ \varepsilon \in (0, 1)\}$ *is tight and therefore relatively weakly compact in* $\mathcal{X}$.

Proof. First, we employ an Aubin–Dubinskii type compact embedding theorem which, in our setting, reads (see [22] for a general exposition; the proof of the following version can be found in [13]):

$$L^2(0, T; H^\gamma(\mathbb{T}^N)) \cap H^\kappa(0, T; L^2(\mathbb{T}^N)) \overset{c}{\hookrightarrow} L^2(0, T; L^2(\mathbb{T}^N)).$$

For $R > 0$ we define the set

$$\begin{aligned}B_{1,R} = \{u \in L^2(0, T; H^\gamma(\mathbb{T}^N)) \cap H^\kappa(0, T; L^2(\mathbb{T}^N));\\ \|u\|_{L^2(0,T;H^\gamma(\mathbb{T}^N))} + \|u\|_{H^\kappa(0,T;L^2(\mathbb{T}^N))} \le R\}\end{aligned}$$

which is thus relatively compact in $L^2(0,T;L^2(\mathbb{T}^N))$. Moreover, by (32) and (36)

$$\mu_{u^\varepsilon}\big(B^C_{1,R}\big)\leq\mathbb{P}\left(\|u^\varepsilon\|_{L^2(0,T;H^\gamma(\mathbb{T}^N))}>\frac{R}{2}\right)+\mathbb{P}\left(\|u^\varepsilon\|_{H^\kappa(0,T;L^2(\mathbb{T}^N))}>\frac{R}{2}\right)$$
$$\leq\frac{2}{R}\Big(\mathbb{E}\|u^\varepsilon\|_{L^2(0,T;H^\gamma(\mathbb{T}^N))}+\mathbb{E}\|u^\varepsilon\|_{H^\kappa(0,T;L^2(\mathbb{T}^N))}\Big)\leq\frac{C}{R}.$$

In order to prove tightness in $C([0,T];H^{-1}(\mathbb{T}^N))$ we employ the compact embedding

$$C^\beta([0,T];H^{-\vartheta}(\mathbb{T}^N))\overset{c}{\hookrightarrow}C^{\tilde\beta}([0,T];H^{-1}(\mathbb{T}^N))\hookrightarrow C([0,T];H^{-1}(\mathbb{T}^N)),$$

where $\tilde\beta<\beta,\ 0<\vartheta<1$. Define

$$B_{2,R}=\{u\in C^\beta([0,T];H^{-\vartheta}(\mathbb{T}^N));\ \|u\|_{C^\beta([0,T];H^{-\vartheta}(\mathbb{T}^N))}\leq R\}$$

then by (35)

$$\mu_{u^\varepsilon}\big(B^C_{2,R}\big)\leq\frac{1}{R}\mathbb{E}\|u^\varepsilon\|_{C^\beta([0,T];H^{-\vartheta}(\mathbb{T}^N))}\leq\frac{C}{R}.$$

Tightness for the initial value is guaranteed as well since $u^\varepsilon(0)=u_0$ is smooth. As a consequence, the set

$$B_R=\Big\{u\in B_{1,R}\cap B_{2,R};\ \|\varrho_0 u\|_{H^1(\mathbb{T}^N)}\leq R\Big\}$$

is relatively compact in $\mathcal{X}_u$ and if $\eta>0$ is given then for some suitably chosen $R>0$ it holds true

$$\mu_{u^\varepsilon}(B_R)\geq 1-\eta,$$

we obtain the tightness of $\{\mu_{u^\varepsilon};\ \varepsilon\in(0,1)\}$. Since also the laws μ_0 and μ_W are tight as being Radon measures on the Polish spaces $\mathcal{X}_0$ and $\mathcal{X}_W$, respectively, we conclude that also the set of their joint laws $\{\mu^\varepsilon;\ \varepsilon\in(0,1)\}$ is tight and Prokhorov's theorem therefore implies that it is also relatively weakly compact. □

Passing to a weakly convergent subsequence $\mu^n=\mu^{\varepsilon_n}$ (and denoting by μ the limit law) we now apply the Skorokhod representation theorem to infer the following proposition.

Proposition 4.10. *There exists a probability space $(\tilde\Omega,\tilde{\mathscr{F}},\tilde{\mathbb{P}})$ with a sequence of $\mathcal{X}$-valued random variables $(\tilde u^n,\tilde W^n)$, $n\in\mathbb{N}$, and $(\tilde u,\tilde W)$ such that*

(i) *the laws of $(\tilde u^n,\tilde W^n)$ and $(\tilde u,\tilde W)$ under $\tilde{\mathbb{P}}$ coincide with μ^n and μ, respectively,*
(ii) *$(\tilde u^n,\tilde W^n)$ converges $\tilde{\mathbb{P}}$-almost surely to $(\tilde u,\tilde W)$ in the topology of $\mathcal{X}$.*

Remark 4.11. Note, that we can assume without loss of generality that the σ-algebra $\tilde{\mathscr{F}}$ is countably generated. This fact will be used later on for the application of the Banach–Alaoglu theorem. It should be also noted that the energy estimates remain valid also for the candidate solution $\tilde u$. Indeed, for any $p\in[1,\infty)$, it follows

$$\tilde{\mathbb{E}}\operatorname*{ess\,sup}_{0\leq t\leq T}\|\tilde u(t)\|^p_{L^p(\mathbb{T}^N)}\leq\liminf_{n\to\infty}\tilde{\mathbb{E}}\sup_{0\leq t\leq T}\|\tilde u^n(t)\|^p_{L^p(\mathbb{T}^N)}$$
$$=\liminf_{n\to\infty}\mathbb{E}\sup_{0\leq t\leq T}\|u^n(t)\|^p_{L^p(\mathbb{T}^N)}\leq C.$$

Let us define functions

$$f^n = \mathbf{1}_{u^n>\xi} : \Omega \times \mathbb{T}^N \times [0,T] \times \mathbb{R} \longrightarrow \mathbb{R},$$
$$\tilde{f}^n = \mathbf{1}_{\tilde{u}^n>\xi},\ \tilde{f} = \mathbf{1}_{\tilde{u}>\xi} : \tilde{\Omega} \times \mathbb{T}^N \times [0,T] \times \mathbb{R} \longrightarrow \mathbb{R},$$

and measures

$$\mathrm{d}m^n(x,t,\xi) = \mathrm{d}n_1^n(x,t,\xi) + \mathrm{d}n_2^n(x,t,\xi),$$
$$\mathrm{d}\tilde{m}^n(x,t,\xi) = \mathrm{d}\tilde{n}_1^n(x,t,\xi) + \mathrm{d}\tilde{n}_2^n(x,t,\xi),$$

where

$$\mathrm{d}n_1^n(x,t,\xi) = \left|\sigma(x)\nabla u^n\right|^2 \mathrm{d}\delta_{u^n(x,t)}(\xi)\,\mathrm{d}x\,\mathrm{d}t,$$
$$\mathrm{d}n_2^n(x,t,\xi) = \varepsilon_n\left|\nabla u^n\right|^2 \mathrm{d}\delta_{u^n(x,t)}(\xi)\,\mathrm{d}x\,\mathrm{d}t,$$
$$\mathrm{d}\tilde{n}_1^n(x,t,\xi) = \left|\sigma(x)\nabla \tilde{u}^n\right|^2 \mathrm{d}\delta_{u^n(x,t)}(\xi)\,\mathrm{d}x\,\mathrm{d}t,$$
$$\mathrm{d}\tilde{n}_2^n(x,t,\xi) = \varepsilon_n\left|\nabla \tilde{u}^n\right|^2 \mathrm{d}\delta_{u^n(x,t)}(\xi)\,\mathrm{d}x\,\mathrm{d}t.$$

Note that all the above measures are well defined. Indeed, Theorem 4.1 implies, in particular, that $u^n \in C([0,T]; H^1(\mathbb{T}^N))$, $\mathbb{P}$-a.s., with $C([0,T]; H^1(\mathbb{T}^N))$ being a Borel subset of $\mathcal{X}_u$ since the embedding $C([0,T]; H^1(\mathbb{T}^N)) \hookrightarrow \mathcal{X}_u$ is continuous. Thus, it follows from Proposition 4.10 that $\tilde{u}^n \in C([0,T]; H^1(\mathbb{T}^N))$, $\tilde{\mathbb{P}}$-a.s., consequently $\tilde{m}^n(\psi) : \Omega \to \mathbb{R}$ is measurable and

$$\tilde{m}^n(\psi) \overset{d}{\sim} m^n(\psi), \quad \forall \psi \in C_0(\mathbb{T}^N \times [0,T] \times \mathbb{R}).$$

Let $\mathcal{M}_b(\mathbb{T}^N \times [0,T] \times \mathbb{R})$ denote the space of bounded Borel measures on $\mathbb{T}^N \times [0,T] \times \mathbb{R}$ whose norm is given by the total variation of measures. It is the dual space to the space of all continuous functions vanishing at infinity $C_0(\mathbb{T}^N \times [0,T] \times \mathbb{R})$ equipped with the supremum norm. This space is separable, so the following duality holds for $q, q^* \in (1,\infty)$ being conjugate exponents (see [11, Theorem 8.20.3]):

$$L_w^q(\tilde{\Omega}; \mathcal{M}_b(\mathbb{T}^N \times [0,T] \times \mathbb{R})) \simeq \left(L^{q^*}(\tilde{\Omega}; C_0(\mathbb{T}^N \times [0,T] \times \mathbb{R}))\right)^*,$$

where the space on the left hand side contains all weak*-measurable mappings $n : \tilde{\Omega} \to \mathcal{M}_b(\mathbb{T}^N \times [0,T] \times \mathbb{R})$ such that

$$\tilde{\mathbb{E}}\|n\|_{\mathcal{M}_b}^q < \infty.$$

Lemma 4.12. *It holds true (up to subsequences) that*

(i) *there exists a set of full Lebesgue measure* $\mathcal{D} \subset [0,T]$ *which contains* $t = 0$ *such that*

$$\tilde{f}^n(t) \overset{w^*}{\longrightarrow} \tilde{f}(t) \quad \text{in } L^\infty(\tilde{\Omega} \times \mathbb{T}^N \times \mathbb{R})\text{-weak}^*,\ \forall t \in \mathcal{D},$$

(ii) *there exists a kinetic measure* $\tilde{m}$ *such that*

$$\tilde{m}^n \overset{w^*}{\longrightarrow} \tilde{m} \quad \text{in } L_w^2(\tilde{\Omega}; \mathcal{M}_b(\mathbb{T}^N \times [0,T] \times \mathbb{R}))\text{-weak}^*. \tag{39}$$

Moreover, $\tilde{m}$ *can be rewritten as* $\tilde{n}_1 + \tilde{n}_2$, *where*

$$\mathrm{d}\tilde{n}_1(x,t,\xi) = \left|\sigma(x)\nabla\tilde{u}\right|^2 \mathrm{d}\delta_{\tilde{u}(x,t)}(\xi)\,\mathrm{d}x\,\mathrm{d}t$$

and $\tilde{n}_2$ *is almost surely a nonnegative measure over* $\mathbb{T}^N \times [0,T] \times \mathbb{R}$.

Proof. According to Proposition 4.10, there exists a set $\Sigma \subset \tilde{\Omega} \times \mathbb{T}^N \times [0,T]$ of full measure and a subsequence still denoted by $\{\tilde{u}^n;\ n \in \mathbb{N}\}$ such that $\tilde{u}^n(\omega,x,t) \to \tilde{u}(\omega,x,t)$ for all $(\omega,x,t) \in \Sigma$. We infer that

$$\mathbf{1}_{\tilde{u}^n(\omega,x,t)>\xi} \longrightarrow \mathbf{1}_{\tilde{u}(\omega,x,t)>\xi} \tag{40}$$

whenever

$$\Big(\tilde{\mathbb{P}} \otimes \mathcal{L}_{\mathbb{T}^N} \otimes \mathcal{L}_{[0,T]}\Big)\Big\{(\omega,x,t) \in \Sigma;\ \tilde{u}(\omega,x,t) = \xi\Big\} = 0,$$

where by $\mathcal{L}_{\mathbb{T}^N}$, $\mathcal{L}_{[0,T]}$ we denoted the Lebesgue measure on $\mathbb{T}^N$ and $[0,T]$, respectively. However, the set

$$D = \Big\{\xi \in \mathbb{R};\ \Big(\tilde{\mathbb{P}} \otimes \mathcal{L}_{\mathbb{T}^N} \otimes \mathcal{L}_{[0,T]}\Big)\big(\tilde{u} = \xi\big) > 0\Big\}$$

is at most countable since we deal with finite measures. To obtain a contradiction, suppose that D is uncountable and denote

$$D_k = \left\{\xi \in \mathbb{R};\ \Big(\tilde{\mathbb{P}} \otimes \mathcal{L}_{\mathbb{T}^N} \otimes \mathcal{L}_{[0,T]}\Big)\big(\tilde{u} = \xi\big) > \frac{1}{k}\right\}, \quad k \in \mathbb{N}.$$

Then $D = \cup_{k\in\mathbb{N}} D_k$ is a countable union, so there exists $k_0 \in \mathbb{N}$ such that D_{k_0} is uncountable. Hence

$$\begin{aligned}\Big(\tilde{\mathbb{P}} \otimes \mathcal{L}_{\mathbb{T}^N} \otimes \mathcal{L}_{[0,T]}\Big)\big(\tilde{u} \in D\big) &\geq \Big(\tilde{\mathbb{P}} \otimes \mathcal{L}_{\mathbb{T}^N} \otimes \mathcal{L}_{[0,T]}\Big)\big(\tilde{u} \in D_{k_0}\big)\\ &= \sum_{\xi \in D_{k_0}} \Big(\tilde{\mathbb{P}} \otimes \mathcal{L}_{\mathbb{T}^N} \otimes \mathcal{L}_{[0,T]}\Big)\big(\tilde{u} = \xi\big) > \sum_{\xi \in D_{k_0}} \frac{1}{k_0} = \infty\end{aligned}$$

and the desired contradiction follows. We conclude that the convergence in (40) holds true for a.e. (ω,x,t,ξ) and obtain by the dominated convergence theorem

$$\tilde{f}^n \overset{w^*}{\longrightarrow} \tilde{f} \quad \text{in } L^\infty(\tilde{\Omega} \times \mathbb{T}^N \times [0,T] \times \mathbb{R})\text{-weak}^*; \tag{41}$$

hence (i) follows for a subsequence and the convergence at $t = 0$ follows by a similar approach.

As the next step we shall show that the set $\{\tilde{m}^n;\ n \in \mathbb{N}\}$ is bounded in $L^2_w(\tilde{\Omega}; \mathcal{M}_b(\mathbb{T}^N \times [0,T] \times \mathbb{R}))$. With regard to the computations used in the proof of the energy inequality, we get from (28)

$$\begin{aligned}&\int_0^T \int_{\mathbb{T}^N} \big|\sigma(x)\nabla u^n\big|^2 \mathrm{d}x\,\mathrm{d}t + \varepsilon_n \int_0^T \int_{\mathbb{T}^N} \big|\nabla u^n\big|^2 \mathrm{d}x\,\mathrm{d}t \leq C\|u_0\|^2_{L^2(\mathbb{T}^N)}\\ &\quad + C\sum_{k\geq 1} \int_0^T \int_{\mathbb{T}^N} u^n g^n_k(x,u^n)\,\mathrm{d}x\,\mathrm{d}\beta_k(t) + C\int_0^T \int_{\mathbb{T}^N} G^2_n(x,u^n)\,\mathrm{d}x\,\mathrm{d}s.\end{aligned}$$

Taking square and expectation and finally by the Itô isometry, we deduce

$$\begin{aligned}\tilde{\mathbb{E}}\big|\tilde{m}^n(\mathbb{T}^N \times [0,T] \times \mathbb{R})\big|^2 &= \mathbb{E}\big|m^n(\mathbb{T}^N \times [0,T] \times \mathbb{R})\big|^2\\ &= \mathbb{E}\bigg|\int_0^T \int_{\mathbb{T}^N} \big|\sigma(x)\nabla u^n\big|^2 \mathrm{d}x\,\mathrm{d}t\\ &\quad + \varepsilon_n \int_0^T \int_{\mathbb{T}^N} \big|\nabla u^n\big|^2 \mathrm{d}x\,\mathrm{d}t\bigg|^2 \leq C.\end{aligned}$$

Thus, according to the Banach–Alaoglu theorem, (39) is obtained (up to subsequence). However, it still remains to show that the weak* limit $\tilde{m}$ is actually a kinetic measure. The first point of Definition 2.1 is straightforward as it corresponds to the weak*-measurability of $\tilde{m}$. The second one giving the behavior for large ξ follows from the uniform estimate (28). Indeed, let (χ_δ) be a truncation on $\mathbb{R}$, then it holds, for $p \in [2, \infty)$, that

$$\begin{aligned}\tilde{\mathbb{E}} \int_{\mathbb{T}^N \times [0,T] \times \mathbb{R}} |\xi|^{p-2} \,\mathrm{d}\tilde{m}(x, t, \xi) &\le \liminf_{\delta \to 0} \tilde{\mathbb{E}} \int_{\mathbb{T}^N \times [0,T] \times \mathbb{R}} |\xi|^{p-2} \chi_\delta(\xi) \,\mathrm{d}\tilde{m}(x, t, \xi) \\ &= \liminf_{\delta \to 0} \lim_{n \to \infty} \tilde{\mathbb{E}} \int_{\mathbb{T}^N \times [0,T] \times \mathbb{R}} |\xi|^{p-2} \\ &\quad \times \chi_\delta(\xi) \,\mathrm{d}\tilde{m}^n(x, t, \xi) \le C,\end{aligned}$$

where the last inequality follows from (28) and the sequel. As a consequence, $\tilde{m}$ vanishes for large ξ. The remaining requirement of Definition 2.1 follows from [6, Theorem 3.7] since for any $\psi \in C_0(\mathbb{T}^N \times \mathbb{R})$

$$t \longmapsto \int_{\mathbb{T}^N \times [0,t] \times \mathbb{R}} \psi(x, \xi) \,\mathrm{d}\tilde{m}(s, x, \xi)$$

is $\tilde{\mathscr{F}} \otimes \mathcal{B}([0, T])$-measurable and $(\tilde{\mathscr{F}}_t)$-adapted for the filtration introduced below after this proof.

Finally, by the same approach as above, we deduce that there exist kinetic measures $\tilde{o}_1$, $\tilde{o}_2$ such that

$$\tilde{n}_1^n \xrightarrow{w^*} \tilde{o}_1, \qquad \tilde{n}_2^n \xrightarrow{w^*} \tilde{o}_2 \quad \text{in } L^2_w(\tilde{\Omega}; \mathcal{M}_b(\mathbb{T}^N \times [0, T] \times \mathbb{R}))\text{-weak}^*.$$

Then from (28) we obtain

$$\tilde{\mathbb{E}} \int_0^T \int_{\mathbb{T}^N} \left|\sigma(x) \nabla \tilde{u}^n\right|^2 \mathrm{d}x \,\mathrm{d}t \le C;$$

hence the application of the Banach–Alaoglu theorem yields that, up to subsequence, $\sigma \nabla \tilde{u}^n$ converges weakly in $L^2(\tilde{\Omega} \times \mathbb{T}^N \times [0, T])$. On the other hand, from the strong convergence given by Proposition 4.10 and the fact that $\sigma \in W^{1,\infty}(\mathbb{T}^N)$, we conclude using integration by parts, for all $\psi \in C^1(\mathbb{T}^N \times [0, T])$, that

$$\int_0^T \int_{\mathbb{T}^N} \sigma(x) \nabla \tilde{u}^n \psi(x, t) \,\mathrm{d}x \,\mathrm{d}t \longrightarrow \int_0^T \int_{\mathbb{T}^N} \sigma(x) \nabla \tilde{u} \psi(x, t) \,\mathrm{d}x \,\mathrm{d}t, \quad \tilde{\mathbb{P}}\text{-a.s.}$$

Therefore

$$\sigma \nabla \tilde{u}^n \xrightarrow{w} \sigma \nabla \tilde{u}, \quad \text{in } L^2(\mathbb{T}^N \times [0, T]),\ \tilde{\mathbb{P}}\text{-a.s.}$$

Since any norm is weakly sequentially lower semicontinuous, it follows for all $\varphi \in C_0(\mathbb{T}^N \times [0, T] \times \mathbb{R})$ and fixed $\xi \in \mathbb{R}$, $\tilde{\mathbb{P}}$-a.s.,

$$\int_0^T \int_{\mathbb{T}^N} \left|\sigma(x) \nabla \tilde{u}\right|^2 \varphi^2(x, t, \xi) \,\mathrm{d}x \,\mathrm{d}t \le \liminf_{n \to \infty} \int_0^T \int_{\mathbb{T}^N} \left|\sigma(x) \nabla \tilde{u}^n\right|^2 \varphi^2(x, t, \xi) \mathrm{d}x \mathrm{d}t$$

and by the Fatou lemma

$$\int_0^T \int_{\mathbb{T}^N} \int_{\mathbb{R}} \left|\sigma(x) \nabla \tilde{u}\right|^2 \varphi^2(x, t, \xi) \,\mathrm{d}\delta_{\tilde{u}=\xi} \,\mathrm{d}x \,\mathrm{d}t$$

$$\leq \liminf_{n\to\infty} \int_0^T \int_{\mathbb{T}^N} \int_{\mathbb{R}} \big|\sigma(x)\nabla\tilde{u}^n\big|^2 \varphi^2(x,t,\xi)\,\mathrm{d}\delta_{\tilde{u}^n=\xi}\,\mathrm{d}x\,\mathrm{d}t, \qquad \tilde{\mathbb{P}}\text{-a.s.}$$

In other words, this gives $\tilde{n}_1 = |\sigma\,\nabla\tilde{u}|^2\delta_{\tilde{u}=\xi} \leq \tilde{o}_1$ $\tilde{\mathbb{P}}$-a.s. hence $\tilde{n}_2 = \tilde{o}_2 + (\tilde{o}_1 - \tilde{n}_1)$ is $\tilde{\mathbb{P}}$-a.s. a nonnegative measure and the proof is complete. □

Finally, let us define the following filtration generated by $\tilde{u}, \tilde{W}, \tilde{m}$

$$\hat{\mathscr{F}}_t = \sigma\big(\boldsymbol{\varrho}_t\tilde{u}, \boldsymbol{\varrho}_t\tilde{W}, \tilde{m}(\theta\psi), \theta \in C([0,T]), \operatorname{supp}\theta \subset [0,t], \psi \in C_0(\mathbb{T}^N \times \mathbb{R})\big)$$

and let $(\tilde{\mathscr{F}}_t)$ be its augmented filtration, i.e. the smallest complete right-continuous filtration that contains $(\hat{\mathscr{F}}_t)$. Then $\tilde{u}$ is $(\tilde{\mathscr{F}}_t)$-predictable $H^{-1}(\mathbb{T}^N)$-valued processes since it has continuous trajectories. Furthermore, by the embeddings $L^p(\mathbb{T}^N) \hookrightarrow H^{-1}(\mathbb{T}^N), p \in [2,\infty)$, and $L^2(\mathbb{T}^N) \hookrightarrow L^p(\mathbb{T}^N)$, $p \in [1,2)$, we conclude that, for all $p \in [1,\infty)$,

$$\tilde{u} \in L^p(\tilde{\Omega} \times [0,T], \tilde{\mathcal{P}}, \mathrm{d}\mathbb{P} \otimes \mathrm{d}t; L^p(\mathbb{T}^N)),$$

where $\tilde{\mathcal{P}}$ denotes the predictable σ-algebra associated to $(\tilde{\mathscr{F}}_t)_{t\geq 0}$. Remark, that $\tilde{f}$, a Borel function of $\tilde{u}$ and ξ, is measurable with respect to $\tilde{\mathcal{P}} \otimes \mathcal{B}(\mathbb{T}^N) \otimes \mathcal{B}(\mathbb{R})$.

4.4. Passage to the limit

In this paragraph we provide the technical details of the identification of the limit process with a kinetic solution. The technique performed here will be used also in the proof of existence of a pathwise kinetic solution.

Theorem 4.13. *The triple $\big((\tilde{\Omega}, \tilde{\mathscr{F}}, (\tilde{\mathscr{F}}_t), \tilde{\mathbb{P}}), \tilde{W}, \tilde{u}\big)$ is a martingale kinetic solution to the problem (1).*

Note, that as the set $\mathcal{D}$ from Lemma 4.12 is a complement of a set with zero Lebesgue measure, it is dense in $[0,T]$. Let us define for all $t \in \mathcal{D}$ and some fixed $\varphi \in C_c^\infty(\mathbb{T}^N \times \mathbb{R})$

$$\begin{aligned}
M^n(t) &= \big\langle f^n(t), \varphi\big\rangle - \big\langle f_0, \varphi\big\rangle - \int_0^t \big\langle f^n(s), b^n(\xi)\cdot\nabla\varphi\big\rangle\,\mathrm{d}s \\
&\quad - \int_0^t \Big\langle f^n(s), \operatorname{div}\big(A(x)\nabla\varphi\big)\Big\rangle\,\mathrm{d}s - \varepsilon_n \int_0^t \big\langle f^n(s), \Delta\varphi\big\rangle\,\mathrm{d}s \\
&\quad - \frac{1}{2}\int_0^t \big\langle \delta_{u^n=\xi} G_n^2, \partial_\xi\varphi\big\rangle\,\mathrm{d}s + \big\langle m^n, \partial_\xi\varphi\big\rangle([0,t)), \quad n \in \mathbb{N}, \\
\tilde{M}^n(t) &= \big\langle \tilde{f}^n(t), \varphi\big\rangle - \big\langle \tilde{f}^n(0), \varphi\big\rangle - \int_0^t \big\langle \tilde{f}^n(s), b^n(\xi)\cdot\nabla\varphi\big\rangle\,\mathrm{d}s \\
&\quad - \int_0^t \Big\langle \tilde{f}^n(s), \operatorname{div}\big(A(x)\nabla\varphi\big)\Big\rangle\,\mathrm{d}s - \varepsilon_n \int_0^t \big\langle \tilde{f}^n(s), \Delta\varphi\big\rangle\,\mathrm{d}s \\
&\quad - \frac{1}{2}\int_0^t \big\langle \delta_{\tilde{u}^n=\xi} G_n^2, \partial_\xi\varphi\big\rangle\,\mathrm{d}s + \big\langle \tilde{m}^n, \partial_\xi\varphi\big\rangle([0,t)), \quad n \in \mathbb{N}, \\
\tilde{M}(t) &= \big\langle \tilde{f}(t), \varphi\big\rangle - \big\langle \tilde{f}(0), \varphi\big\rangle - \int_0^t \big\langle \tilde{f}(s), b(\xi)\cdot\nabla\varphi\big\rangle\,\mathrm{d}s \\
&\quad - \int_0^t \Big\langle \tilde{f}(s), \operatorname{div}\big(A(x)\nabla\varphi\big)\Big\rangle\,\mathrm{d}s - \frac{1}{2}\int_0^t \big\langle \delta_{\tilde{u}=\xi} G^2, \partial_\xi\varphi\big\rangle\,\mathrm{d}s + \big\langle \tilde{m}, \partial_\xi\varphi\big\rangle([0,t)).
\end{aligned}$$

The proof of Theorem 4.13 is a consequence of the following two propositions.

Proposition 4.14. *The process $\tilde{W}$ is an $(\tilde{\mathscr{F}}_t)$-cylindrical Wiener process, i.e. there exists a collection of mutually independent real-valued $(\tilde{\mathscr{F}}_t)$-Wiener processes $\{\tilde{\beta}_k\}_{k\geq 1}$ such that $\tilde{W} = \sum_{k\geq 1} \tilde{\beta}_k e_k$.*

Proof. Hereafter, fix $K \in \mathbb{N}$, times $0 \leq s_1 < \cdots < s_K \leq s \leq t$, $s, t \in \mathcal{D}$, continuous functions

$$\gamma : C\big([0, s]; H^{-1}(\mathbb{T}^N)\big) \times C\big([0, s]; \mathfrak{U}_0\big) \longrightarrow [0, 1], \qquad g : \mathbb{R}^K \longrightarrow [0, 1]$$

and test functions $\psi_1, \ldots, \psi_K \in C_0(\mathbb{T}^N \times \mathbb{R})$ and $\theta_1, \ldots, \theta_K \in C([0, T])$ such that $\operatorname{supp}\theta_i \subset [0, s_i]$, $i = 1, \ldots, K$. For notational simplicity, we write $\boldsymbol{g}(\tilde{m})$ instead of

$$g\big(\tilde{m}(\theta_1\psi_1), \ldots, \tilde{m}(\theta_K\psi_K)\big)$$

and similarly $\boldsymbol{g}(\tilde{m}^n)$ and $\boldsymbol{g}(m^n)$. By $\boldsymbol{\varrho}_s$ we denote the operator of restriction to the interval $[0, s]$ as introduced in (38).

Obviously, $\tilde{W}$ is a $\mathfrak{U}_0$-valued cylindrical Wiener process and is $(\tilde{\mathscr{F}}_t)$-adapted. According to the Lévy martingale characterization theorem, it remains to show that it is also a $(\tilde{\mathscr{F}}_t)$-martingale. It holds true

$$\tilde{\mathbb{E}}\,\gamma\big(\boldsymbol{\varrho}_s\tilde{u}^n, \boldsymbol{\varrho}_s\tilde{W}^n\big)\boldsymbol{g}(\tilde{m}^n)\big[\tilde{W}^n(t) - \tilde{W}^n(s)\big] = \mathbb{E}\,\gamma\big(\boldsymbol{\varrho}_s u^n, \boldsymbol{\varrho}_s W\big)\boldsymbol{g}(m^n)\big[W(t) - W(s)\big] = 0$$

since W is a martingale and the laws of $(\tilde{u}^n, \tilde{W}^n)$ and (u^n, W) coincide. Next, the uniform estimate

$$\sup_{n\in\mathbb{N}} \tilde{\mathbb{E}}\|\tilde{W}^n(t)\|^2_{\mathfrak{U}_0} = \sup_{n\in\mathbb{N}} \mathbb{E}\|W(t)\|^2_{\mathfrak{U}_0} < \infty$$

and the Vitali convergence theorem yields

$$\tilde{\mathbb{E}}\,\gamma\big(\boldsymbol{\varrho}_s\tilde{u}, \boldsymbol{\varrho}_s\tilde{W}\big)\boldsymbol{g}(\tilde{m})\big[\tilde{W}(t) - \tilde{W}(s)\big] = 0$$

which finishes the proof. □

Proposition 4.15. *The processes*

$$\tilde{M}(t), \qquad \tilde{M}^2(t) - \sum_{k\geq 1}\int_0^t \big\langle \delta_{\tilde{u}=\xi}\, g_k, \varphi\big\rangle^2 \mathrm{d}r, \qquad \tilde{M}(t)\tilde{\beta}_k(t) - \int_0^t \big\langle \delta_{\tilde{u}=\xi}\, g_k, \varphi\big\rangle \mathrm{d}r,$$

indexed by $t \in \mathcal{D}$, are $(\tilde{\mathscr{F}}_t)$-martingales.

Proof. All these processes are $(\tilde{\mathscr{F}}_t)$-adapted as they are Borel functions of $\tilde{u}$ and $\tilde{\beta}_k$, $k \in \mathbb{N}$, up to time t. For the rest, we use the same approach and notation as the one used in the previous lemma. Let us denote by $\tilde{\beta}^n_k$, $k \geq 1$ the real-valued Wiener processes corresponding to $\tilde{W}^n$, that is $\tilde{W}^n = \sum_{k\geq 1} \tilde{\beta}^n_k e_k$. For all $n \in \mathbb{N}$, the process

$$M^n = \int_0^{\cdot} \big\langle \delta_{u^n=\xi}\, \Phi^n(u^n)\mathrm{d}W, \varphi\big\rangle = \sum_{k\geq 1}\int_0^{\cdot} \big\langle \delta_{u^n=\xi}\, g^n_k, \varphi\big\rangle \mathrm{d}\beta_k(r)$$

is a square integrable $(\mathscr{F}_t)$-martingale by (2) and by the fact that the set $\{u^n;\ n \in \mathbb{N}\}$ is bounded in $L^2(\Omega; L^2(0, T; L^2(\mathbb{T}^N)))$. Therefore

$$(M^n)^2 - \sum_{k\geq 1}\int_0^{\cdot} \big\langle \delta_{u^n=\xi}\, g^n_k, \varphi\big\rangle^2 \mathrm{d}r, \qquad M^n\beta_k - \int_0^{\cdot} \big\langle \delta_{u^n=\xi}\, g^n_k, \varphi\big\rangle \mathrm{d}r$$

are $(\mathscr{F}_t)$-martingales and this implies together with the equality of laws that

$$\begin{aligned}
&\tilde{\mathbb{E}}\,\gamma\big(\varrho_s\tilde{u}^n,\varrho_s\tilde{W}^n\big)\boldsymbol{g}(\tilde{m}^n)\big[\tilde{M}^n(t)-\tilde{M}^n(s)\big]\\
&\quad=\mathbb{E}\,\gamma\big(\varrho_s u^n,\varrho_s W\big)\boldsymbol{g}(m^n)\big[M^n(t)-M^n(s)\big]=0, \qquad (42)\\
&\tilde{\mathbb{E}}\,\gamma\big(\varrho_s\tilde{u}^n,\varrho_s\tilde{W}^n\big)\boldsymbol{g}(\tilde{m}^n)\left[(\tilde{M}^n)^2(t)-(\tilde{M}^n)^2(s)-\sum_{k\geq1}\int_s^t\big\langle\delta_{\tilde{u}^n=\xi}\,g_k^n,\varphi\big\rangle^2\,\mathrm{d}r\right]\\
&\quad=\mathbb{E}\,\gamma\big(\varrho_s u^n,\varrho_s W\big)\boldsymbol{g}(m^n)\left[(M^n)^2(t)-(M^n)^2(s)-\sum_{k\geq1}\int_s^t\big\langle\delta_{u^n=\xi}\,g_k^n,\varphi\big\rangle^2\,\mathrm{d}r\right]\\
&\quad=0, \qquad (43)\\
&\tilde{\mathbb{E}}\,\gamma\big(\varrho_s\tilde{u}^n,\varrho_s\tilde{W}^n\big)\boldsymbol{g}(\tilde{m}^n)\left[\tilde{M}^n(t)\tilde{\beta}_k^n(t)-\tilde{M}^n(s)\tilde{\beta}_k^n(s)-\int_s^t\big\langle\delta_{\tilde{u}^n=\xi}\,g_k^n,\varphi\big\rangle\,\mathrm{d}r\right]\\
&\quad=\mathbb{E}\,\gamma\big(\varrho_s u^n,\varrho_s W\big)\boldsymbol{g}(m^n)\left[M^n(t)\beta_k(t)-M^n(s)\beta_k(s)-\int_s^t\big\langle\delta_{u^n=\xi}\,g_k^n,\varphi\big\rangle\,\mathrm{d}r\right]\\
&\quad=0. \qquad (44)
\end{aligned}$$

Moreover, for any $s,t\in\mathcal{D}, s\leq t$, the expectations in (42)–(44) converge by the Vitali convergence theorem. Indeed, all terms are uniformly integrable by (2) and (27) and converge $\tilde{\mathbb{P}}$-a.s. (after extracting a subsequence) due to Lemma 4.12, (40), (41), Proposition 4.10 and the construction of $\Phi^\varepsilon,\ B^\varepsilon$. Hence

$$\begin{aligned}
&\tilde{\mathbb{E}}\,\gamma\big(\varrho_s\tilde{u},\varrho_s\tilde{W}\big)\boldsymbol{g}(\tilde{m})\big[\tilde{M}(t)-\tilde{M}(s)\big]=0,\\
&\tilde{\mathbb{E}}\,\gamma\big(\varrho_s\tilde{u},\varrho_s\tilde{W}\big)\boldsymbol{g}(\tilde{m})\left[\tilde{M}^2(t)-\tilde{M}^2(s)-\sum_{k\geq1}\int_s^t\big\langle\delta_{\tilde{u}=\xi}\,g_k,\varphi\big\rangle^2\mathrm{d}r\right]=0,\\
&\tilde{\mathbb{E}}\,\gamma\big(\varrho_s\tilde{u},\varrho_s\tilde{W}\big)\boldsymbol{g}(\tilde{m})\left[\tilde{M}(t)\tilde{\beta}_k(t)-\tilde{M}(s)\tilde{\beta}_k(s)-\int_s^t\big\langle\delta_{\tilde{u}=\xi}\,g_k,\varphi\big\rangle\mathrm{d}r\right]=0,
\end{aligned}$$

which gives the $(\tilde{\mathscr{F}}_t)$-martingale property. □

Proof of Theorem 4.13. If all the processes in 4.15 were continuous-time martingales then it would hold true

$$\left\langle\!\!\left\langle\tilde{M}-\int_0^{\cdot}\big\langle\delta_{\tilde{u}=\xi}\,\Phi(\tilde{u})\,\mathrm{d}\tilde{W},\varphi\big\rangle\right\rangle\!\!\right\rangle=0,$$

where by $\langle\!\langle\cdot\rangle\!\rangle$ we denote the quadratic variation process, and therefore, for every $\varphi\in C_c^\infty(\mathbb{T}^N\times\mathbb{R}),\ t\in[0,T],\ \tilde{\mathbb{P}}$-a.s.,

$$\begin{aligned}
&\big\langle\tilde{f}(t),\varphi\big\rangle-\big\langle\tilde{f}_0,\varphi\big\rangle-\int_0^t\big\langle\tilde{f}(s),b(\xi)\cdot\nabla\varphi\big\rangle\mathrm{d}s-\int_0^t\big\langle\tilde{f}(s),\operatorname{div}\big(A(x)\nabla\varphi\big)\big\rangle\mathrm{d}s\\
&\quad=\int_0^t\big\langle\delta_{\tilde{u}=\xi}\,\Phi(\tilde{u})\,\mathrm{d}\tilde{W},\varphi\big\rangle+\frac{1}{2}\int_0^t\big\langle\delta_{\tilde{u}=\xi}G^2,\partial_\xi\varphi\big\rangle\,\mathrm{d}s-\big\langle\tilde{m},\partial_\xi\varphi\big\rangle([0,t)) \qquad (45)
\end{aligned}$$

and the proof would be completed with $\tilde{u}$ satisfying the kinetic formulation even in a stronger sense than required by Definition 2.2.

In the case of martingales indexed by $t \in \mathcal{D}$, we employ Proposition A.1 to conclude the validity of (45) for all $\varphi \in C_c^\infty(\mathbb{T}^N \times \mathbb{R})$, $t \in \mathcal{D}$, $\tilde{\mathbb{P}}$-a.s., and we need to allow a formulation which is weak also in time. Mimicking the technique developed in order to derive the kinetic formulation in Section 2, let us define

$$N(t) = \big\langle \tilde{f}_0, \varphi\big\rangle + \int_0^t \big\langle \tilde{f}(s), b(\xi)\cdot\nabla\varphi\big\rangle \mathrm{d}s + \int_0^t \big\langle \tilde{f}(s), \mathrm{div}\big(A(x)\nabla\varphi\big)\big\rangle \mathrm{d}s$$
$$+ \int_0^t \big\langle \delta_{\tilde{u}=\xi}\, \Phi(\tilde{u})\,\mathrm{d}\tilde{W}, \varphi\big\rangle + \frac{1}{2}\int_0^t \big\langle \delta_{\tilde{u}=\xi} G^2, \partial_\xi\varphi\big\rangle\,\mathrm{d}s.$$

Note, that N is a continuous real-valued semimartingale and

$$N(t) = \big\langle \tilde{f}(t), \varphi\big\rangle + \big\langle \tilde{m}, \partial_\xi\varphi\big\rangle([0,t)), \qquad \forall t \in \mathcal{D}.$$

Next, we apply the Itô formula to calculate the stochastic differential of the product $N(t)\varphi_1(t)$, where $\varphi_1 \in C_c^\infty([0,T))$. After the application of the Fubini theorem to the term including the kinetic measure $\tilde{m}$, we obtain exactly the formulation (8). □

4.5. *Pathwise solutions*

In order to finish the proof, we make use of the Gyöngy–Krylov characterization of convergence in probability introduced in [16]. It is useful in situations when the pathwise uniqueness and the existence of at least one martingale solution imply the existence of a unique pathwise solution.

Proposition 4.16. *Let X be a Polish space equipped with the Borel σ-algebra. A sequence of X-valued random variables $\{Y_n;\ n \in \mathbb{N}\}$ converges in probability if and only if for every subsequence of joint laws, $\{\mu_{n_k,m_k};\ k \in \mathbb{N}\}$, there exists a further subsequence which converges weakly to a probability measure μ such that*

$$\mu\big((x,y) \in X \times X;\ x = y\big) = 1.$$

We consider the collection of joint laws of (u^n, u^m) on $\mathcal{X}_u \times \mathcal{X}_u$, denoted by $\mu_u^{n,m}$. For this purpose we define the extended path space

$$\mathcal{X}^J = \mathcal{X}_u \times \mathcal{X}_u \times \mathcal{X}_W.$$

As above, denote by μ_W the law of W and set $\nu^{n,m}$ to be the joint law of (u^n, u^m, W). Similarly to Theorem 4.9 the following fact holds true. The proof is nearly identical and so will be left to the reader.

Proposition 4.17. *The collection $\{\nu^{n,m};\ n, m \in \mathbb{N}\}$ is tight on $\mathcal{X}^J$.*

Let us take any subsequence $\{\nu^{n_k,m_k};\ k \in \mathbb{N}\}$. By Prokhorov's theorem, it is relatively weakly compact hence it contains a weakly convergent subsequence. Without loss of generality we may assume that the original sequence $\{\nu^{n_k,m_k};\ k \in \mathbb{N}\}$ itself converges weakly to a measure ν. According to the Skorokhod representation theorem, we infer the existence of a probability space $(\bar{\Omega}, \bar{\mathscr{F}}, \bar{\mathbb{P}})$ with a sequence of random variables $(\hat{u}^{n_k}, \check{u}^{m_k}, \bar{W}^k)$, $k \in \mathbb{N}$, converging almost surely in $\mathcal{X}^J$ to a random variable $(\hat{u}, \check{u}, \bar{W})$ and

$$\bar{\mathbb{P}}\big((\hat{u}^{n_k}, \check{u}^{m_k}, \bar{W}^k) \in \ \cdot\big) = \nu^{n_k,m_k}(\cdot), \qquad \bar{\mathbb{P}}\big((\hat{u}, \check{u}, \bar{W}) \in \ \cdot\big) = \nu(\cdot).$$

Observe that in particular, $\mu_u^{n_k,m_k}$ converges weakly to a measure μ_u defined by

$$\mu_u(\cdot) = \bar{\mathbb{P}}\big((\hat{u}, \check{u}) \in \ \cdot\big).$$

As the next step, we should recall the technique established in the previous section. Analogously, it can be applied to both $(\hat{u}^{n_k}, \bar{W}^k)$, $(\hat{u}, \bar{W})$ and $(\check{u}^{m_k}, \bar{W}^k)$, $(\check{u}, \bar{W})$ in order to show that $(\hat{u}, \bar{W})$ and $(\check{u}, \bar{W})$ are martingale kinetic solutions of (1) with corresponding kinetic measures $\hat{m}$ and $\check{m}$, respectively, defined on the same stochastic basis $(\bar{\Omega}, \bar{\mathscr{F}}, (\bar{\mathscr{F}}_t), \bar{\mathbb{P}})$, where $(\bar{\mathscr{F}}_t)$ is the augmented filtration to

$$\sigma\Big(\varrho_t\hat{u}, \varrho_t\check{u}, \varrho_t\bar{W}, \hat{m}(\theta_1\psi_1), \check{m}(\theta_2\psi_2);$$
$$\theta_i \in C([0,T]), \operatorname{supp}\theta_i \subset [0,t], \psi_i \in C_0(\mathbb{T}^N \times \mathbb{R}), i = 1,2\Big), \quad t \in [0,T].$$

Since $\hat{u}(0) = \check{u}(0) = \bar{u}_0$, $\bar{\mathbb{P}}$-a.s., we infer from Theorem 3.3 that $\hat{u} = \check{u}$ in $\mathcal{X}_u$, $\bar{\mathbb{P}}$-a.s., hence

$$\mu_u\big((x,y) \in \mathcal{X}_u \times \mathcal{X}_u;\ x = y\big) = \bar{\mathbb{P}}\big(\hat{u} = \check{u} \text{in } \mathcal{X}_u\big) = 1.$$

Now, we have all in hand to apply Proposition 4.16. It implies that the original sequence u^n defined on the initial probability space $(\Omega, \mathscr{F}, \mathbb{P})$ converges in probability in the topology of $\mathcal{X}_u$ to a random variable u. Without loss of generality, we assume that u^n converges to u almost surely in $\mathcal{X}_u$ and again by the method from Section 4.4 we finally deduce that u is a pathwise kinetic solution to (1). Actually, identification of the limit is more straightforward here since in this case all the work is done for the initial setting and only one fixed driving Wiener process W is considered.

5. Existence—general initial data

In this final section we provide an existence proof in the general case of $u_0 \in L^p(\Omega; L^p(\mathbb{T}^N))$, for all $p \in [1,\infty)$. It is a straightforward consequence of the previous section. We approximate the initial condition by a sequence $\{u_0^\varepsilon\} \subset L^p(\Omega; C^\infty(\mathbb{T}^N))$, $p \in [1,\infty)$, such that $u_0^\varepsilon \to u_0$ in $L^1(\Omega; L^1(\mathbb{T}^N))$. That is, the initial condition u_0^ε can be defined as a pathwise mollification of u_0 so that it holds true

$$\|u_0^\varepsilon\|_{L^p(\Omega;L^p(\mathbb{T}^N))} \leq \|u_0\|_{L^p(\Omega;L^p(\mathbb{T}^N))}, \quad \varepsilon \in (0,1),\ p \in [1,\infty). \tag{46}$$

According to the previous section, for each $\varepsilon \in (0,1)$, there exists a kinetic solution u^ε to (1) with initial condition u_0^ε. By the application of the comparison principle (19),

$$\mathbb{E}\|u^{\varepsilon_1}(t) - u^{\varepsilon_2}(t)\|_{L^1(\mathbb{T}^N)} \leq \mathbb{E}\|u_0^{\varepsilon_1} - u_0^{\varepsilon_2}\|_{L^1(\mathbb{T}^N)}, \quad \varepsilon_1, \varepsilon_2 \in (0,1);$$

hence $\{u^\varepsilon;\ \varepsilon \in (0,1)\}$ is a Cauchy sequence in $L^1(\Omega \times [0,T], \mathcal{P}, \mathrm{d}\mathbb{P} \otimes \mathrm{d}t; L^1(\mathbb{T}^N))$. Consequently, there exists $u \in L^1(\Omega \times [0,T], \mathcal{P}, \mathrm{d}\mathbb{P} \otimes \mathrm{d}t; L^1(\mathbb{T}^N))$ such that

$$u^\varepsilon \longrightarrow u \quad \text{in } L^1(\Omega \times [0,T], \mathcal{P}, \mathrm{d}\mathbb{P} \otimes \mathrm{d}t; L^1(\mathbb{T}^N)).$$

By (46) and Remark 4.11, we still have the uniform energy estimates, $p \in [1,\infty)$,

$$\mathbb{E} \operatorname*{ess\,sup}_{0\leq t\leq T} \|u^\varepsilon(t)\|_{L^p(\mathbb{T}^N)}^p \leq C_{T,u_0} \tag{47}$$

as well as (using the usual notation)

$$\mathbb{E}\big|m^\varepsilon(\mathbb{T}^N \times [0,T] \times \mathbb{R})\big|^2 \leq C_{T,u_0}.$$

Thus, using these observations as in Lemma 4.12, one finds that there exists a subsequence $\{u^n;\ n \in \mathbb{N}\}$ such that

(i) $f^n \xrightarrow{w^*} f$ in $L^\infty(\Omega \times \mathbb{T}^N \times [0,T] \times \mathbb{R})$-weak*,

(ii) there exists a kinetic measure m such that

$$m^n \xrightarrow{w^*} m \quad \text{in } L^2_w(\Omega; \mathcal{M}_b(\mathbb{T}^N \times [0,T] \times \mathbb{R}))\text{-weak}^*$$

and $m_1 + n_2$, where

$$\mathrm{d}n_1(x,t,\xi) = \left|\sigma(x)\nabla u\right|^2 \mathrm{d}\delta_{u(x,t)}(\xi)\,\mathrm{d}x\,\mathrm{d}t$$

and n_2 is almost surely a nonnegative measure over $\mathbb{T}^N \times [0,T] \times \mathbb{R}$.

With these facts in hand, we are ready to pass to the limit in (8) and conclude that u satisfies the kinetic formulation in the sense of distributions. Note, that (47) remains valid also for u so (6) follows and, according to the embedding $L^p(\mathbb{T}^N) \hookrightarrow L^1(\mathbb{T}^N)$, for all $p \in [1,\infty)$, we deduce

$$u \in L^p(\Omega \times [0,T], \mathcal{P}, \mathrm{d}\mathbb{P} \otimes \mathrm{d}t; L^p(\mathbb{T}^N)).$$

The proof of Theorem 2.10 is complete.

Acknowledgments

The author wishes to thank Arnaud Debussche and Jan Seidler for many stimulating discussions and valuable suggestions. Last but not least, many thanks go to the referees for their comments helping to improve the manuscript.

This research was supported in part by the GA ČR Grant no. P201/10/0752 and the GA UK Grant no. 556712.

Appendix. Densely defined martingales

In this section, we present an auxiliary result which is used in the proof of existence of a martingale kinetic solution in Theorem 4.13. To be more precise, it solves the following problem: it is needed to show equality of a certain martingale M and a stochastic integral $\int_0^t \sigma\,\mathrm{d}W$ but the process M is only defined on a dense subset $\mathcal{D}$ of $[0,T]$ containing zero and no continuity property is a priori known. Therefore, one cannot just prove that the quadratic variation of their difference vanishes as it is not well defined.

To begin with, let us fix some notation. Let H, U be separable Hilbert spaces with orthonormal bases $(g_j)_{j\geq 1}$ and $(f_k)_{k\geq 1}$, respectively, and inner products $\langle\cdot,\cdot\rangle$ and $\langle\cdot,\cdot\rangle_U$, respectively. For simplicity, we will work on a finite-time interval $[0,T]$, $T \in \mathcal{D}$.

Proposition A.1. *Assume that $W(t) = \sum_{k\geq 1}\beta_k(t) f_k$ is a cylindrical Wiener process in U defined on a stochastic basis $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geq 0}, \mathbb{P})$ with a complete, right-continuous filtration. If $(M(t); t \in \mathcal{D})$ is a square integrable $(\mathscr{F}_t)$-adapted H-valued stochastic process such that, for any $s, t \in \mathcal{D}$, $s \leq t$, $j, k \geq 1$, $\mathbb{P}$-a.s.,*

$$\begin{aligned}
&\mathbb{E}\big[\langle M(t) - M(s), g_j\rangle | \mathscr{F}_s\big] = 0,\\
&\mathbb{E}\left[\langle M(t), g_j\rangle^2 - \langle M(s), g_j\rangle^2 - \int_s^t \|\sigma^* g_j\|_U^2\,\mathrm{d}r \,\middle|\, \mathscr{F}_s\right] = 0,\\
&\mathbb{E}\left[\beta_k(t)\langle M(t), g_j\rangle - \beta_k(s)\langle M(s), g_j\rangle - \int_s^t \langle f_k, \sigma^* g_j\rangle_U\,\mathrm{d}r \,\middle|\, \mathscr{F}_s\right] = 0,
\end{aligned} \tag{A.1}$$

where σ is an $(\mathscr{F}_t)$-progressively measurable $L_2(U;H)$-valued stochastically integrable process, i.e.

$$\mathbb{E}\int_0^T \|\sigma\|^2_{L_2(U;H)}\,\mathrm{d}r < \infty, \tag{A.2}$$

then

$$M(t) = \int_0^t \sigma\,\mathrm{d}W, \quad \forall t\in\mathcal{D},\ \mathbb{P}\text{-a.s.}$$

In particular, M can be defined for all $t\in[0,T]$ such that it has a modification which is a continuous $(\mathscr{F}_t)$-martingale.

Proof. The crucial point to be shown here is the following: for any $(\mathscr{F}_t)$-progressively measurable $L_2(U;H)$-valued process ϕ satisfying (A.2) and any $s,t\in\mathcal{D}, s\le t, j\ge 1$, it holds, $\mathbb{P}$-a.s.,

$$\mathbb{E}\left[\big\langle M(t)-M(s),g_j\big\rangle\left\langle\int_s^t\phi\,\mathrm{d}W,g_j\right\rangle-\int_s^t\langle\sigma^*g_j,\phi^*g_j\rangle_U\,\mathrm{d}r\,\middle|\,\mathscr{F}_s\right]=0. \tag{A.3}$$

We consider simple processes first. Let ϕ be an $(\mathscr{F}_t)$-adapted simple process with values in finite-dimensional operators of $L(U;H)$ that satisfies (A.2), i.e.

$$\phi(t)=\phi_0\mathbf{1}_{\{0\}}(t)+\sum_{i=0}^{I}\phi_i\mathbf{1}_{(t_i,t_{i+1}]}(t),\quad t\in[0,T],$$

where $\{0=t_0<t_1<\cdots<t_I=T\}$ is a division of $[0,T]$ such that $t_i\in\mathcal{D}, i=0,\dots,I$. Then the stochastic integral in (A.3) is given by

$$\begin{aligned}\int_s^t\phi\,\mathrm{d}W &= \phi_{m-1}\big(W(t_m)-W(s)\big)+\sum_{i=m}^{n-1}\phi_i\big(W(t_{i+1})-W(t_i)\big)+\phi_n\big(W(t)-W(t_n)\big)\\ &=\sum_{k\ge1}\Bigg(\phi^k_{m-1}\big(\beta_k(t_m)-\beta_k(s)\big)+\sum_{i=m}^{n-1}\phi^k_i\big(\beta_k(t_{i+1})-\beta_k(t_i)\big)\\ &\quad+\phi^k_n\big(\beta_k(t)-\beta_k(t_n)\big)\Bigg)\end{aligned}$$

provided $t_{m-1}\le s<t_m, t_n\le t<t_{n+1}, \phi^k_i=\phi_i f_k$. Next, we write

$$M(t)-M(s)=\big(M(t_m)-M(s)\big)+\sum_{i=m}^{n-1}\big(M(t_{i+1})-M(t_i)\big)+\big(M(t)-M(t_n)\big)$$

and conclude

$$\begin{aligned}&\mathbb{E}\left[\big\langle M(t)-M(s),g_j\big\rangle\left\langle\int_s^t\phi\,\mathrm{d}W,g_j\right\rangle\middle|\,\mathscr{F}_s\right]\\ &\quad=\mathbb{E}\Big[\big\langle\phi_{m-1}\big(W(t_m)-W(s)\big),g_j\big\rangle\big\langle M(t_m)-M(s),g_j\big\rangle\end{aligned}$$

$$+\sum_{i=m}^{n-1}\big\langle\phi_i\big(W(t_{i+1})-W(t_i)\big),g_j\big\rangle\big\langle M(t_{i+1})-M(t_i),g_j\big\rangle$$

$$+\big\langle\phi_n\big(W(t)-W(t_n)\big),g_j\big\rangle\big\langle M(t)-M(t_n),g_j\big\rangle\Big|\,\mathscr{F}_s\Big] \tag{A.4}$$

as one can neglect all the mixed terms due to the martingale property of $\beta_k,\ k\geq 1$, and (A.1). Indeed, let $i\in\{m,\dots,n-1\}$ then

$$\mathbb{E}\Big[\big\langle\phi_i(W(t_{i+1})-W(t_i)),g_j\big\rangle\big\langle M(t_m)-M(s),g_j\big\rangle|\mathscr{F}_s\Big]$$
$$=\mathbb{E}\bigg[\mathbb{E}\bigg[\sum_{k\geq 1}\big\langle\phi_i^k\big(\beta_k(t_{i+1})-\beta_k(t_i)\big),g_j\big\rangle\big\langle M(t_m)-M(s),g_j\big\rangle\bigg|\,\mathscr{F}_{t_i}\bigg]\bigg|\,\mathscr{F}_s\bigg]$$
$$=\mathbb{E}\bigg[\big\langle M(t_m)-M(s),g_j\big\rangle\sum_{k\geq 1}\langle\phi_i^k,g_j\rangle\,\mathbb{E}\Big[\beta_k(t_{i+1})-\beta_k(t_i)|\mathscr{F}_{t_i}\Big]\bigg|\,\mathscr{F}_s\bigg]=0,$$

where the interchange of summation with scalar product and expectation, respectively, is justified by the fact that

$$\sum_{k\geq 1}\phi_i^k\big(\beta_k(t_{i+1})-\beta_k(t_i)\big)=\int_{t_i}^{t_{i+1}}\phi_i\,\mathrm{d}W$$

is convergent in $L^2(\Omega;H)$.

As the next step, we proceed with (A.4). If $i\in\{m,\dots,n-1\}$ then we obtain using again the martingale property of $\beta_k,\ k\geq 1$, and (A.1)

$$\mathbb{E}\Big[\big\langle\phi_i\big(W(t_{i+1})-W(t_i)\big),g_j\big\rangle\big\langle M(t_{i+1})-M(t_i),g_j\big\rangle|\mathscr{F}_s\Big]$$
$$=\mathbb{E}\bigg[\mathbb{E}\bigg[\sum_{k\geq 1}\big\langle\phi_i^k,g_j\big\rangle\big(\beta_k(t_{i+1})-\beta_k(t_i)\big)\big\langle M(t_{i+1})-M(t_i),g_j\big\rangle\bigg|\,\mathscr{F}_{t_i}\bigg]\bigg|\,\mathscr{F}_s\bigg]$$
$$=\mathbb{E}\bigg[\sum_{k\geq 1}\langle\phi_i^k,g_j\rangle\,\mathbb{E}\Big[\beta_k(t_{i+1})\big\langle M(t_{i+1}),g_j\big\rangle-\beta_k(t_i)\big\langle M(t_i),g_j\big\rangle|\mathscr{F}_{t_i}\Big]\bigg|\,\mathscr{F}_s\bigg]$$
$$=\mathbb{E}\bigg[\sum_{k\geq 1}\langle\phi_i^k,g_j\rangle\int_{t_i}^{t_{i+1}}\big\langle f_k,\sigma^*g_j\big\rangle_U\,\mathrm{d}r\bigg|\,\mathscr{F}_s\bigg]$$
$$=\mathbb{E}\bigg[\sum_{k\geq 1}\langle f_k,\phi_i^*g_j\rangle_U\int_{t_i}^{t_{i+1}}\big\langle f_k,\sigma^*g_j\big\rangle_U\,\mathrm{d}r\bigg|\,\mathscr{F}_s\bigg]$$
$$=\mathbb{E}\bigg[\int_{t_i}^{t_{i+1}}\langle\sigma^*g_j,\phi^*g_j\rangle_U\,\mathrm{d}r\bigg|\,\mathscr{F}_s\bigg].$$

The remaining terms being dealt with similarly. As a consequence, we see that (A.3) holds true for simple processes and the general case follows by classical arguments using approximation.

Now, we have all in hand to complete the proof. Let $t \in \mathcal{D}$ and set $s = 0$ in (A.1), (A.3), then

$$\mathbb{E}\left(\langle M(t), g_j\rangle - \left\langle \int_0^t \sigma\, \mathrm{d}W, g_j \right\rangle\right)^2$$
$$= \mathbb{E}\langle M(t), g_j\rangle^2 - 2\,\mathbb{E}\langle M(t), g_j\rangle \left\langle \int_0^t \sigma\, \mathrm{d}W, g_j \right\rangle + \mathbb{E}\left\langle \int_0^t \sigma\, \mathrm{d}W, g_j \right\rangle^2 = 0, \quad j \geq 1,$$

and the claim follows. □

References

[1] H. Amann, Compact embeddings of vector-valued Sobolev and Besov spaces, Glas. Mat. Ser. III 35 (55) (2000) 161–177.

[2] L. Ambrosio, N. Fusco, D. Pallara, Functions of Bounded Variation and Free Discontinuity Problems, in: Oxford Mathematical Monographs, The Clarendon Press, Oxford University Press, New York, 2000.

[3] Z. Brzeźniak, M. Ondreját, Strong solutions to stochastic wave equations with values in Riemannian manifolds, J. Funct. Anal. 253 (2007) 449–481.

[4] J. Carrillo, Entropy solutions for nonlinear degenerate problems, Arch. Ration. Mech. Anal. 147 (1999) 269–361.

[5] G.Q. Chen, B. Perthame, Well-posedness for non-isotropic degenerate parabolic–hyperbolic equations, Ann. Inst. H. Poincaré Anal. Non Linéaire 20 (4) (2003) 645–668.

[6] K.L. Chung, R.J. Williams, Introduction to Stochastic Integration, Birkhäuser, Boston, Basel, Berlin, 1990.

[7] D.L. Cohn, Measure Theory, Birkhäuser, Boston, Basel, Stuttgart, 1980.

[8] G. Da Prato, J. Zabczyk, Stochastic Equations in Infinite Dimensions, in: Encyclopedia Math. Appl., vol. 44, Cambridge University Press, Cambridge, 1992.

[9] A. Debussche, J. Vovelle, Scalar conservation laws with stochastic forcing, J. Funct. Anal. 259 (2010) 1014–1042.

[10] R.J. DiPerna, P.L. Lions, Ordinary differential equations, transport theory and Sobolev spaces, Invent. Math. 98 (1989) 511–547.

[11] R.E. Edwards, Functional Analysis, Theory and Applications, Holt, Rinehart and Winston, 1965.

[12] J. Feng, D. Nualart, Stochastic scalar conservation laws, J. Funct. Anal. 255 (2) (2008) 313–373.

[13] F. Flandoli, D. Gątarek, Martingale and stationary solutions for stochastic Navier–Stokes equations, Probab. Theory Related Fields 102 (3) (1995) 367–391.

[14] M.I. Freidlin, On the factorization of nonnegative definite matrices, Theory Probab. Appl. 13 (1968) 354–356.

[15] G. Gagneux, M. Madaune-Tort, Analyse Mathématique de Modèles Non Linéaires de L'ingénierie Pétrolière, Springer-Verlag, 1996.

[16] I. Gyöngy, N. Krylov, Existence of strong solutions for Itô's stochastic equations via approximations, Probab. Theory Related Fields 105 (2) (1996) 143–158.

[17] I. Gyöngy, C. Rovira, On L^p-solutions of semilinear stochastic partial differential equations, Stochastic Process. Appl. 90 (2000) 83–108.

[18] M. Hofmanová, Strong solutions of semilinear stochastic partial differential equations, NoDEA Nonlinear Differential Equations Appl. 20 (3) (2013) 757–778.

[19] M. Hofmanová, J. Seidler, On weak solutions of stochastic differential equations, Stoch. Anal. Appl. 30 (1) (2012) 100–121.

[20] C. Imbert, J. Vovelle, A kinetic formulation for multidimensional scalar conservation laws with boundary conditions and applications, SIAM J. Math. Anal. 36 (1) (2004) 214–232.

[21] S.N. Kružkov, First order quasilinear equations with several independent variables, Mat. Sb. 81 (123) (1970) 228–255.

[22] J.L. Lions, Quelques Méthodes de Résolution des Problèmes aux Limites non Linéaires, Dunod, Paris, 1969.

[23] P.L. Lions, B. Perthame, E. Tadmor, Formulation cinétique des lois de conservation scalaires multidimensionnelles, C. R. Acad. Sci., Paris Sér. I (1991) 97–102.

[24] P.L. Lions, B. Perthame, E. Tadmor, A kinetic formulation of multidimensional scalar conservation laws and related equations, J. Amer. Math. Soc. 7 (1) (1994) 169–191.

[25] M. Ondreját, Stochastic nonlinear wave equations in local Sobolev spaces, Electron. J. Probab. 15 (2010) 1041–1091.

[26] B. Perthame, Uniqueness and error estimates in first order quasilinear conservation laws via the kinetic entropy defect measure, J. Math. Pures Appl.(9) 77 (1998) 1055–1064.

[27] B. Perthame, Kinetic Formulation of Conservation Laws, in: Oxford Lecture Ser. Math. Appl., vol. 21, Oxford University Press, Oxford, 2002.
[28] R.S. Philips, L. Sarason, Elliptic–parabolic equations of the second order, J. Math. Mech. 17 (1968) 891–917.
[29] S.S. Samko, A.A. Kilbas, O.I. Marichev, Fractional Integrals and Derivatives, Gordon and Breach, Yverdon, 1993.
[30] H. Triebel, Interpolation Theory, Function Spaces, Differential Operators, North-Holland Publishing Company, Amsterdam, New York, Oxford, 1978.
[31] H. Triebel, Theory of Function Spaces II, Birkhäuser, Basel, 1992.