ELSEVIER

Available online at www.sciencedirect.com

ScienceDirect

Journal of Functional Analysis 255 (2008) 313–373

JOURNAL OF Functional Analysis

www.elsevier.com/locate/jfa

# Stochastic scalar conservation laws

Jin Feng [*,1], David Nualart [2]

*Department of Mathematics, University of Kansas, Lawrence, KS 66045, United States*

Received 11 February 2008; accepted 17 February 2008

Available online 9 April 2008

Communicated by Paul Malliavin

**Abstract**

We introduce a notion of stochastic entropic solution *à la* Kruzkov, but with Ito's calculus replacing deterministic calculus. This results in a rich family of stochastic inequalities defining what we mean by a solution. A uniqueness theory is then developed following a stochastic generalization of $L^1$ contraction estimate. An existence theory is also developed by adapting compensated compactness arguments to stochastic setting. We use approximating models of vanishing viscosity solution type for the construction. While the uniqueness result applies to any spatial dimensions, the existence result, in the absence of special structural assumptions, is restricted to one spatial dimension only.
Published by Elsevier Inc.

*Keywords:* Stochastic analysis; Scalar conservation law; Stochastic compensated compactness

## 1. Introduction

We are interested in the well-posedness (existence and uniqueness) for first order nonlinear stochastic partial differential equation (SPDE) of the following type

$$\partial_t u(t,x) + \operatorname{div}_x F\big(u(t,x)\big) = \int_{z\in Z} \sigma\big(x, u(t,x); z\big)\, \partial_t W(t,dz). \tag{1}$$

[*] Corresponding author.
*E-mail address:* jfeng@math.ku.edu (J. Feng).

[1] Supported in part by US ARO W911NF-08-1-0064.
[2] Supported in part by US NSF DMS-0604207.

0022-1236/$ – see front matter Published by Elsevier Inc.
doi:10.1016/j.jfa.2008.02.004

In the above, $x \in R^d$, $t \geqslant 0$, $u(t,x)$ is a random scalar-valued function, $F = (F_1, \ldots, F_d) : R \to R^d$ is a vector field (*the flux*). Regarding the random term on the right-hand side of the equation, $Z$ is a metric space, and $W(t,dz)$ is a space–time Gaussian white noise martingale random measure with respect to a filtration $\{\mathcal{F}_t\}$ (e.g. Walsh [18], Kurtz and Protter [13]) with

$$E\big[W(t,A) \cap W(t,B)\big] = \mu(A \cap B)t \tag{2}$$

for measurable $A, B \subset Z$, where $\mu$ is a (deterministic) $\sigma$-finite Borel measure on the metric space $Z$. In addition, $\sigma : R^d \times R \times Z \mapsto R$.

In the case of $\sigma = 0$, (1) reduces to a *deterministic* partial differential equation known as the scalar conservation law

$$\partial_t u(t,x) + \operatorname{div}_x F\big(u(t,x)\big) = 0, \tag{3}$$

which has been extensively studied in nonlinear partial differential equation theory literature (e.g. Dafermos [3]). A well-known difficulty for (3) is that solutions cannot be interpreted in classical sense: non-differentiability in $x$ for $u(t,x)$ develops in finite time, even if $u(0,x)$ is chosen to be smooth [3, Theorem 5.1.1]. On the other hand, because of the nonlinearity in $F$, (Schwartz) distributional weak solution will generally not be unique (e.g. Section 4.2 of [3]). Kruzkov [9,10] introduced a method for selecting a weak solution motivated by physical consideration (the entropic solution). Well-posedness of (3) in the entropic solution sense can be proved for $u(t) \in L^1 \cap L^\infty$, $t \geqslant 0$. There are also other methods of selecting weak solutions. Most of these different approaches can be shown to be equivalent, at least in one space dimension $d = 1$. It is worth mentioning that, from a physical point of view, vector-valued $u$ version of (3) is ultimately more interesting. However, little is known about well-posedness in that case. A detailed exposition about deterministic conservation law, for scalar- as well as vector-valued $u$, is given by Dafermos [3]. See also Chen [2] for a survey. Chapter 11 of Evans [7] contains a brief but informative introduction to the scalar case.

The goal of this article is to introduce a proper generalization of entropic solution to the stochastic case (1) (Definition 2.5). Such notion will enable us to prove uniqueness of solution under mild assumptions on $F$ and $\sigma$ (Theorem 3.5). We will also give existence result for slightly more restrictive situations in one space dimension in Section 4.

The following example gives us a feel on the scope of application that model (1) covers.

**Example 1.1.** Let $Z = \{1, 2, \ldots, m\}$ and $\mu$ be a counting measure on $Z$, (1) reduces to

$$\partial_t u(t,x) + \operatorname{div}_x F\big(u(t,x)\big) = \sum_{k=1}^m \sigma_k\big(x, u(t,x)\big)\partial_t W_k(t), \tag{4}$$

where $W_1, \ldots, W_m$ are independent standard Brownian motions.

In particular, taking $d = 1$ and $F(u) = |u|^2/2$, the equation reduces to the stochastic Burgers' equation

$$\partial_t u(t,x) + u(t,x)u_x(t,x) = \sum_{k=1}^m \sigma_k\big(x, u(t,x)\big)\partial_t W_k(t).$$

The $W(t,dz)$ term can be extended to general semi-martingale random measure, and the theory developed here is expected to hold as well. We do not pursue this direction in this article.

There is a well-known connection between Hamilton–Jacobi and conservation law equations. Such connection can be transferred to the stochastic case as well. Let scalar function $\phi = \phi(t,x) : [0,\infty) \times R^d \mapsto R$ be a solution to

$$\partial_t \phi(t,x) + F\big(\nabla_x \phi(t,x)\big) = \int_Z \sigma\big(x, \nabla_x \phi(t,x); z\big)\, \partial_t W(t,dz). \tag{5}$$

Let vector-valued function $u(t,x) = \nabla_x \phi(t,x)$, then

$$\partial_t u + \nabla_x F(u) = \int_Z \big(\nabla_x \sigma(x,u;z) + \partial_u \sigma(x,u;z)(\nabla_x \cdot u)\big)\, \partial_t W(t,dz).$$

The case of $d = 1$ and $\sigma = \sigma(x;z)$ independent of $u$ gives scalar conservation law as considered in (1). In a series of publications [14], Lions and Souganidis consider equations related to (5):

$$\partial_t \phi + F\big(\nabla \phi, D^2 \phi\big) = \sum_{k=1}^m \sigma_k\big(\nabla \phi(t,x)\big) \circ dW_k(t),$$

where $D^2$ is the Hessian operator and $\circ$ stands for Stratonovich type integral. Stochastic generalizations of viscosity solution are used.

## 2. Stochastic entropic solution—definition and main result

### 2.1. Definitions

**Definition 2.1.** $(\Phi, \Psi)$ is called an entropy–entropy flux pair if $\Phi \in C^1(R)$ and $\Psi = (\Psi_1, \dots, \Psi_d) : R^d \mapsto R^d$ is a vector field satisfying

$$\Psi_k'(r) = \Phi'(r)(F_k)'(r), \quad k = 1, \dots, d. \tag{6}$$

**Remark 2.2.** $\Psi_k$ can be chosen as

$$\Psi_k(r) = \int_v^r \Phi'(s)(F_k)'(s)\, ds, \quad \text{for some fixed } v \in R.$$

Note that, unlike the usual definition, we do not require $\Phi$ to be convex in this definition.

A special class of entropy–entropy flux pairs will play a major role in later analysis. We define it next. For each $\varepsilon > 0$, let $\beta = \beta_\varepsilon \in C^\infty(R)$ be convex,

$$\beta(r) = 0, \quad r \leqslant 0, \qquad \beta(r) = C_\varepsilon + r, \quad C_\varepsilon > 0,\ r \geqslant \varepsilon.$$

$$\begin{aligned} \mathcal{K} = \big\{&(\Phi, \Psi) \text{ is an entropy–entropy flux pair:} \\ &\Phi(r) = \Phi^u(r) = \beta(u - r), \text{ or } \Phi(r) = \Phi_v(r) = \beta(r - v),\ u, v \in R\big\}. \end{aligned} \tag{7}$$

Throughout this article, we assume the following regularities.

**Condition 2.3.**

(1) $F_k \in C^2(R)$, and $F_k''(s)$ have at most polynomial growth in $s$, for all $k = 1, \dots, d$;
(2) For each compact subset $K \subset R^d \times R^d$, there exists $M_K : Z \mapsto R$ and nonnegative, nondecreasing, continuous function $\rho_K : R \mapsto R$ with $\rho_K(0) = 0$ such that

$$\left|\sigma(y, v; z) - \sigma(x, u; z)\right| \leqslant \left(|u - v|^{1/2}\rho_K\left(|u - v|\right) + |x - y|\right)M_K(z),$$
$$\forall (x, y) \in K,\ z \in Z,$$

where

$$C_K \equiv \int_{z \in Z} M_K^2(z)\,\mu(dz) < \infty.$$

**Example 2.4.** Let $\sigma_k : R^d \times R \mapsto R$ be Lipschitz for each $k = 1, \dots, m$ and consider (4), then the second part of the above conditions is satisfied.

**Definition 2.5** *(Stochastic entropic solution).* Let $(\Omega, \{\mathcal{F}_t\colon t \geqslant 0\} \subset \mathcal{F}, P)$ be a filtered probability space where $W(t, \cdot)$ is adapted space–time Gaussian white noise martingale random measure satisfying (2).

We call an $L^2(R^d)$-valued $\{\mathcal{F}_t\}$ -adapted stochastic process $u = u(t) = u(t, x)$ a stochastic entropic solution of (1), provided

(1) for each $T > 0$, $p = 2, 3, 4, \dots,$

$$\sup_{0 \leqslant t \leqslant T} E\left[\left\|u(t)\right\|_p^p\right] < \infty, \tag{8}$$

and for each $N = 1, 2, \dots$ fixed,

$$\int_0^T E\left[\int_{z \in Z}\int_{|x| \leqslant N} \left|\sigma\left(x, u(r, x); z\right)\right|^4 dx\,\mu(dz)\right] dr < \infty. \tag{9}$$

(2) For each $0 \leqslant s \leqslant t$, each $0 \leqslant \varphi \in C_{\mathrm{c}}^2(R^d)$, and each $(\Phi, \Psi) \in \mathcal{K}$,

$$\begin{aligned}
&\left\langle \Phi\left(u(t, \cdot)\right), \varphi\right\rangle - \left\langle \Phi\left(u(s, \cdot)\right), \varphi\right\rangle \\
&\quad\leqslant \int_s^t \left\langle \Psi\left(u(r, \cdot)\right), \nabla_x \varphi\right\rangle dr + \int_{(s,t] \times Z} \frac{1}{2}\left\langle \Phi''\left(u(r, \cdot)\right)\sigma^2\left(\cdot, u(r, \cdot); z\right), \varphi\right\rangle \mu(dz)\,dr \\
&\qquad + \int_{(s,t] \times Z} \left\langle \Phi'\left(u(r, \cdot)\right) \cdot \sigma\left(\cdot, u(r, \cdot); z\right), \varphi\right\rangle W(dr, dz).
\end{aligned} \tag{10}$$

Perhaps a more revealing way to re-state integral inequality (10) is to say that $u$ is a (Schwartz distributional) weak solution to

$$\begin{aligned} &\partial_t \Phi\big(u(t,x)\big) + \operatorname{div}_x \Psi\big(u(t,x)\big) \\ &\quad \leqslant \frac{1}{2}\int_Z \Phi''\big(u(t,x)\big)\sigma^2\big(x,u(t,x);z\big)\,\mu(dz) \\ &\qquad + \int_Z \Phi'\big(u(t,x)\big)\sigma\big(x,u(t,x);z\big)\,\frac{\partial W(t,dz)}{\partial t}. \end{aligned} \tag{11}$$

When $\sigma = 0$, the right-hand side of the above inequality drops to zero. (11) reduces to exactly the defining differential inequality in deterministic entropic solution initially introduced by Kruzkov [9].

Some explanation on the meaning of (10) is necessary: $(\Phi,\Psi) \in \mathcal{K}$ implies that $\Phi''$ and $\Phi'$ are bounded and $\Psi$ has at most polynomial growth. Together with (8) and (9), each term in (10) is well defined. A significant special yet common case satisfying (9) is when $\sigma$ is uniformly bounded $\sup_{x,u,z}|\sigma(x,u,z)| < +\infty$, and $\mu(Z) < \infty$.

By an interpolation argument, to verify that (8) holds for $p = 2, 3, \ldots,$ it is good enough to show for even positive integer valued cases of $p = 2, 4, 6, \ldots.$ Moreover, both imply that (8) holds for all $p \in [2,\infty)$.

Unlike deterministic scalar conservation law (i.e. the case $\sigma = 0$), to prove path-wise uniqueness, we also need to capture more explicitly "noise–noise interaction" between any two possibly different stochastic solutions. We strengthen the definition of solution as follows.

**Definition 2.6** *(Stochastic strong entropic solution).* We call an $L^2(R^d)$-valued, $\{\mathcal{F}_t\}$-adapted process $v = v(t) = v(t,x)$ a stochastic strong entropic solution of (1) if the following holds:

(1) it is a stochastic entropic solution (i.e. (8), (9) and (10) hold for $u$ replaced by $v$);
(2) for each $L^2(R^d)$-valued, $\mathcal{F}_t$-adapted process $\tilde{u}(t)$ satisfying

$$\sup_{0\leqslant t\leqslant T} E\big[\|\tilde{u}(t)\|_p^p\big] < \infty, \quad T > 0,\ p = 2,3,\ldots,$$

and for each $\beta \in C^\infty(R)$ of the form (17), $0 \leqslant \varphi \in C_c^\infty(R^d \times R^d)$, and

$$f(r,z;v,y) = \int_{x\in R^d} \beta'\big(\tilde{u}(r,x) - v\big)\sigma\big(x,\tilde{u}(r,x);z\big)\varphi(x,y)\,dx,$$

there exists a deterministic function $\{A(s,t)\colon\ 0 \leqslant s \leqslant t\}$ such that

$$\begin{aligned} &E\bigg[\int_y \int_{(s,t]\times Z} f\big(r,z;v(t,y),y\big)\,W(dr,dz)\,dy\bigg] \\ &\quad \leqslant E\bigg[\int_{(s,t]\times Z}\int_y \frac{\partial}{\partial v} f\big(r,z;\tilde{u}(r,y),y\big)\sigma\big(y,v(r,y);z\big)\,\mu(dz)\,dy\,dr\bigg] + A(s,t) \end{aligned} \tag{12}$$

with the following property: for each $T > 0$, there exist partitions $0 = t_1 < t_2 < \cdots < t_m = T$ satisfying

$$\lim_{\max_i |t_{i+1}-t_i| \to 0+} \sum_{i=1}^{m} A(t_i, t_{i+1}) = 0, \quad t \geqslant 0.$$

### 2.2. Main results

We list another set of conditions stronger than those in Condition 2.3.

**Condition 2.7.**

(1) $d = 1$,
(2) $F \in C^2(R)$ and the set $\{r \in R\colon F''(r) \neq 0\}$ is dense in $R$,
(3) there exist $f \in L^\infty(R^d) \cap L^2(R^d)$, deterministic constant $C > 0$, and $M : Z \mapsto R$ such that $\int_Z M^2(z)\,\mu(dz) < \infty$,

$$\left|\sigma(x, u; z)\right| \leqslant f(x)\big(1 + |u|\big)M(z), \tag{13}$$

and

$$\left|\sigma(x, u; z) - \sigma(y, v; z)\right| \leqslant C\big(|u - v| + |x - y|\big)\big(M(z) + \left|\sigma(x, u; z)\right|\big). \tag{14}$$

The main result of this article is the following.

**Theorem 2.8.** *Assume Condition* 2.3 *holds, and that* $\bigcap_{p=1,2,\ldots} L^p(R^d)$*-valued random variable* $u_0$ *satisfies*

$$E\big[\|u_0\|_p^p + \|u_0\|_2^p\big] < \infty, \quad p = 1, 2, \ldots.$$

(Uniqueness) *Suppose that* $u, v$ *are two stochastic entropic solutions of* (1) *with the same initial condition* $u(0) = u_0 = v(0)$*, and that one of* $u, v$ *is a strong stochastic entropic solution. Then almost surely* $u(t) = v(t)$ *for* $t \geqslant 0$.

(Existence) *Assume furthermore that Condition* 2.7 *holds, then there exists a strong stochastic entropic solution* (*hence also entropic solution*) *for* (1) *with initial value* $u_0$.

### 2.3. Notations

Throughout, we denote the space of smooth, rapidly decreasing functions

$$\mathcal{S}\big(R^d\big) = \Big\{f \in C^\infty\colon \sup_x \left|x^m D_x^n f(x)\right| < \infty,\ m, n = 1, 2, \ldots\Big\}. \tag{15}$$

Let $J \in C_c^\infty(R^d)$ be the standard mollifier defined by

$$J(x) = \begin{cases} C\exp\{\frac{1}{|x|^2-1}\} & \text{if } |x| < 1, \\ 0 & \text{if } |x| \geqslant 1, \end{cases} \tag{16}$$

where constant $C > 0$ is selected so that $\int_{R^d} J(z)\,dz = 1$. For each $\varepsilon > 0$, we set $J_\varepsilon(z) = \varepsilon^{-d} J(\varepsilon^{-1}|z|)$. $J_\varepsilon \in C_c^\infty(R^d)$ with $\mathrm{supp}(J_\varepsilon) \subset [-\varepsilon, \varepsilon]^d$. For each $f \in L_{\mathrm{loc}}(R^d)$, we define its mollification

$$f_\varepsilon(x) = J_\varepsilon * f(x) = \int_{R^d} J_\varepsilon(x-y) f(y)\,dy = \int_{R^d} J_\varepsilon(y) f(x-y)\,dy.$$

Let $A \subset R^d$, then function $\chi_A(x) = 1$ if $x \in A$ and $\chi_A(x) = 0$ if $x \notin A$. To simplify, with a slight abuse of notation, we denote $\chi(x) = \chi_{[0,+\infty)}(x)$.

For $a \in R$, we denote $a_+ = \max\{a, 0\}$. Then $|a| = a_+ + (-a)_+$. We need smooth functions approximating $\beta(r) = r_+ \in C(R)$. We consider $J$ in the special case of $d = 1$ and define

$$\rho_\varepsilon(r) = \int_{-\infty}^{r-\varepsilon} J_\varepsilon(s)\,ds, \qquad \beta_\varepsilon(r) = \int_{-\infty}^{r} \rho_\varepsilon(s)\,ds, \quad r \in R. \tag{17}$$

Then by direct verification, we have the following.

**Lemma 2.9.** *The above constructed $\rho_\varepsilon, \beta_\varepsilon \in C^\infty(R)$ have the following properties: $\beta_\varepsilon' = \rho_\varepsilon$, $\beta_\varepsilon''(r) = J_\varepsilon(r-\varepsilon)$; $\rho_\varepsilon$ is a nondecreasing function and*

$$\beta_\varepsilon'(r) = \rho_\varepsilon(r) = \begin{cases} 0 & \text{if } r \leqslant 0, \\ 1 & \text{if } r \geqslant 2\varepsilon; \end{cases} \tag{18}$$

*and $\beta_\varepsilon$ is convex and*

$$\beta_\varepsilon(r) = \begin{cases} 0 & \text{if } r \leqslant 0, \\ \varepsilon\hat{C} + (r - 2\varepsilon) & \text{if } r \geqslant 2\varepsilon, \end{cases} \tag{19}$$

*where $\hat{C} = \int_{-1}^{1} (\int_{t=-1}^{s} J(t)\,dt)\,ds < 2$. Furthermore,*

$$0 \leqslant \beta_\varepsilon''(r) = J_\varepsilon(r-\varepsilon) \leqslant \varepsilon^{-1} C, \quad 0 \leqslant r \leqslant 2\varepsilon,$$

*implying*

$$0 \leqslant r\beta_\varepsilon''(r) \leqslant 2C, \quad \text{for } 0 \leqslant r \leqslant 2\varepsilon.$$

## 3. Uniqueness

### 3.1. A doubling lemma

Let $u$ be a stochastic entropic solutions and $v$ be a stochastic strong entropic solution. We estimate the evolution of $\|(u(t) - v(t))_+\|_1 = \|(u(t) - v(t))_+\|_{L^1}$.

First, let $\beta = \beta_\varepsilon(r)$ as constructed in Lemma 2.9. We approximate $\|(u(t) - v(t))_+\|_1$ by $\int_{R^d \times R^d} \beta(u(t,x) - v(t,y))\varphi(x,y)\,dx\,dy$ (by considering limits $\beta_\varepsilon(r) \to r_+$ and $\varphi(x,y)\,dx\,dy \to \delta_x(dy)\,dx$). Then we develop estimate for time evolution of such approximate. In the deterministic scalar conservation law setting (i.e. $\sigma = 0$), Kruzkov [9] appears to be the first who introduced such an argument for a uniqueness proof. Our goal in this section is to generalize such arguments properly to a stochastic setting (see Lemma 3.2).

Let $0 \leqslant \varphi \in C_c^\infty(R^d \times R^d)$. For $t > s \geqslant 0$,

$$\int_{R^d \times R^d} \beta\big(u(t,x) - v(t,y)\big)\varphi(x,y)\,dx\,dy - \int_{R^d \times R^d} \beta\big(u(s,x) - v(s,y)\big)\varphi(x,y)\,dx\,dy$$
$$= \left(\int \beta\big(u(t,x) - v(t,y)\big)\varphi(x,y)\,dx\,dy - \int \beta\big(u(s,x) - v(t,y)\big)\varphi(x,y)\,dx\,dy\right)$$
$$+ \left(\int \beta\big(u(s,x) - v(t,y)\big)\varphi(x,y)\,dx\,dy - \int \beta\big(u(s,x) - v(s,y)\big)\varphi(x,y)\,dx\,dy\right)$$
$$\equiv I_1 + I_2.$$

First, we estimate $I_2$. We introduce notation

$$\alpha(u,v) = \big(\alpha_1(u,v), \ldots, \alpha_d(u,v)\big),$$

where (noting $\beta(r) = 0$ for $r < 0$)

$$\alpha_k(u,v) = \int_v^\infty \beta'(u-w)F_k'(w)\,dw = \int_v^u \beta'(u-w)F_k'(w)\,dw. \tag{20}$$

**Lemma 3.1.**

$$I_2 \leqslant \int_s^t \int_{R^d \times R^d} \alpha\big(u(s,x), v(r,y)\big) \cdot \nabla_y \varphi(x,y)\,dx\,dy\,dr$$
$$+ \frac{1}{2}\int_s^t \int_Z \int_{x,y} \beta''\big(u(s,x) - v(r,y)\big)\sigma^2\big(y, v(r,y); z\big)\varphi(x,y)\,dx\,dy\,\mu(dz)\,dr$$
$$- \int_{(s,t]\times Z} \int_{x,y} \beta'\big(u(s,x) - v(r,y)\big)\sigma\big(y, v(r,y); z\big)\varphi(x,y)\,dx\,dy\,W(dr,dz). \tag{21}$$

**Proof.** Let $u \in R$ be fixed. We take $\Phi(v) = \beta(u-v)$, $\Psi_k(v) = \int_u^v (-\beta')(u-w)F_k'(w)\,dw$, and apply (10) to $v(t,y)$. Therefore, for each $x \in R^d$,

$$\int_y \beta\big(u - v(t,y)\big)\varphi(x,y)\,dy - \int_y \beta\big(u - v(s,y)\big)\varphi(x,y)\,dy$$
$$\leqslant \int_s^t \int_{y\in R^d} \int_u^{v(r,y)} \left(-\sum_{k=1}^d \beta'(u-w)F_k'(w)\,dw \frac{\partial}{\partial y_k}\varphi(x,y)\right)dy\,dr$$

$$+\frac{1}{2}\int_{(s,t]\times Z}\int_{y}\beta''\big(u-v(r,y)\big)\sigma^2\big(y,v(r,y);z\big)\varphi(x,y)\,dy\,\mu(dz)\,dr$$

$$+\int_{(s,t]\times Z}\int_{y}-\beta'\big(u-v(r,y)\big)\sigma\big(y,v(r,y);z\big)\varphi(x,y)\,dy\,W(dr,dz).$$

Taking $u=u(s,x)$ inside the above inequality and integrating $x$, and applying Fubini's theorem, we arrive at (21). □

We estimate $I_1$ next. We introduce notation

$$\hat\alpha(u,v)=\big(\hat\alpha_1(u,v),\dots,\hat\alpha_d(u,v)\big),$$

where (noting $\beta'(r)=0$ for $r\leqslant 0$)

$$\hat\alpha_k(u,v)=\int_{-\infty}^{u}\beta'(w-v)F_k'(w)\,dw=\int_{v}^{u}\beta'(w-v)F_k'(w)\,dw. \tag{22}$$

For each $v\in R^d$ fixed, taking $\Phi(u)=\beta(u-v)$ and $\Psi(u)=\hat\alpha(u,v)$, we apply (10) for $u(t,x)$, then we take $v=v(t,y)$,

$$\int_{R^d\times R^d}\beta\big(u(t,x)-v(t,y)\big)\varphi(x,y)\,dx\,dy-\int_{R^d\times R^d}\beta\big(u(s,x)-v(t,y)\big)\varphi(x,y)\,dx\,dy$$

$$\leqslant\int_s^t\int_{R^d\times R^d}\hat\alpha\big(u(r,x),v(t,y)\big)\cdot\nabla_x\varphi(x,y)\,dx\,dy\,dr$$

$$+\frac{1}{2}\int_s^t\int_Z\int_{R^d\times R^d}\beta''\big(u(r,x)-v(t,y)\big)\sigma^2\big(x,u(r,x);z\big)\varphi(x,y)\,dx\,dy\,\mu(dz)\,dr$$

$$+\int_{y\in R^d}\int_{(s,t]\times Z}\int_{x\in R^d}\beta'\big(u(r,x)-v(t,y)\big)\sigma\big(x,u(r,x);z\big)\varphi(x,y)\,dx\,W(dr,dz)\,dy$$

$$\equiv I_3+I_4+I_5.$$

To achieve simplicity in exposition, we have slightly abused notation for the term $I_5$ as this should not be understood as an Ito's integral (the integrand contains anticipative term $v(t,y)$). The rigorous meaning of it should be understood in the following sense. Let

$$f(r,z,v,y)=\int_{R^d}\beta'\big(u(r,x)-v\big)\sigma\big(x,u(r,x);z\big)\varphi(x,y)\,dx$$

and

$$G_5(s,t;v,y)=\int_{(s,t]\times Z} f(r,z,v,y)\,W(dr,dz). \tag{23}$$

For each $v$, $y$ fixed, the above is an Ito's integral. Then, we define

$$I_5=I_5(s,t)=\int_{y\in R^d} G_5\big(s,t;v(t,y),y\big)\,dy.$$

A key defining property (12) in stochastic strong entropic solution gives us

$$E[I_5]\leqslant E\bigg[-\int_{R^d\times R^d}\bigg(\int_{(s,t]\times Z}\beta''\big(u(r,x)-v(r,y)\big)\sigma\big(y,v(r,y);z\big)\\ \times\sigma\big(x,u(r,x);z\big)\,\mu(dz)\,dr\bigg)\varphi(x,y)\,dx\,dy\bigg]+A(s,t).$$

Together with the estimate on $E[I_3]$ and $E[I_4]$, we have

$$\begin{aligned}
&E[I_1]\\
&\quad\leqslant E\bigg[\int_s^t\int_{R^d\times R^d}\hat{\alpha}\big(u(r,x),v(r,y)\big)\cdot\nabla_x\varphi(x,y)\,dx\,dr\,dy\bigg]\\
&\qquad+\frac{1}{2}E\bigg[\int_{(s,t]\times Z}\int_{R^d\times R^d}\beta''\big(u(r,x)-v(r,y)\big)\sigma^2\big(x,u(r,x);z\big)\varphi(x,y)\,dx\,dy\,\mu(dz)\,dr\bigg]\\
&\qquad-E\bigg[\int_{(s,t]\times Z}\int_{R^d\times R^d}\beta''\big(u(\bar r,x)-v(\bar r,y)\big)\sigma\big(y,v(\bar r,y);z\big)\\
&\qquad\times\sigma\big(x,u(\bar r,x);z\big)\varphi(x,y)\,dx\,dy\,\mu(dz)\,d\bar r\bigg]\\
&\qquad+A(s,t).
\end{aligned}$$

Combine this with the estimate on $I_2$ in (21), by arbitrariness of $0\leqslant s\leqslant t$, we can arrive at a stochastic version of the doubling of variable estimate first introduced by Kruzkov [9] in deterministic context (see for instance Evans [7, Theorem 3, p. 608]).

**Lemma 3.2.** *For each $t > 0$, we have*

$$E\Bigg[\int_{R^d\times R^d}\beta\big(u(t,x)-v(t,y)\big)\varphi(x,y)\,dx\,dy-\int_{R^d\times R^d}\beta\big(u(0,x)-v(0,y)\big)\varphi(x,y)\,dx\,dy\Bigg]$$

$$\leqslant E\Bigg[\int_0^t\int_{R^d\times R^d}\big(\alpha\big(u(r,x),v(r,y)\big)\cdot\nabla_y\varphi(x,y)+\hat\alpha\big(u(r,x),v(r,y)\big)\cdot\nabla_x\varphi(x,y)\big)\,dx\,dy\,dr\Bigg]$$

$$+\frac{1}{2}E\Bigg[\int_{(0,t]\times Z}\int_{R^d\times R^d}\beta''\big(u(r,x)-v(r,y)\big)$$

$$\times\big(\sigma\big(y,v(r,y);z\big)-\sigma\big(x,u(r,x),z\big)\big)^2\varphi(x,y)\,dx\,dy\,\mu(dz)\,dr\Bigg].$$

**Proof.** We select the sequence of partitions of $[0,t]$ appearing in the defining relation of strong entropic solution (Definition 2.6): $0=t_1\leqslant\cdots\leqslant t_m=t<\infty$. Using the above estimate on $I_1$ and the estimate (21) on $I_2$ (set the $s=t_i, t=t_{i+1}$ there),

$$E\Bigg[\int\beta\big(u(t_{i+1},x)-v(t_{i+1},y)\big)\varphi(x,y)\,dx\,dy-\int\beta\big(u(t_i,x)-v(t_i,y)\big)\varphi(x,y)\,dx\,dy\Bigg]$$

$$\leqslant E\Bigg[\int_{t_i}^{t_{i+1}}\int_{R^d\times R^d}\big(\alpha\big(u(r,x),v(r,y)\big)\cdot\nabla_y\varphi(x,y)+\hat\alpha\big(u(r,x),v(r,y)\big)\cdot\nabla_x\varphi(x,y)\big)\,dx\,dy\,dr\Bigg]$$

$$+\frac{1}{2}E\Bigg[\int_{(t_i,t_{i+1}]\times Z}\int_{R^d\times R^d}\beta''\big(u(r,x)-v(r,y)\big)$$

$$\times\big(\sigma\big(y,v(r,y);z\big)-\sigma\big(x,u(r,x),z\big)\big)^2\varphi(x,y)\,dx\,dy\,\mu(dz)\,dr\Bigg]$$

$$+A(t_i,t_{i+1}).$$

Summing over $i$,

$$E\Bigg[\int_{R^d\times R^d}\beta\big(u(t,x)-v(t,y)\big)\varphi(x,y)\,dx\,dy-\int_{R^d\times R^d}\beta\big(u(0,x)-v(0,y)\big)\varphi(x,y)\,dx\,dy\Bigg]$$

$$\leqslant E\Bigg[\int_0^t\int_{R^d\times R^d}\big(\alpha\big(u(r,x),v(r,y)\big)\cdot\nabla_y\varphi(x,y)+\hat\alpha\big(u(r,x),v(r,y)\big)\cdot\nabla_x\varphi(x,y)\big)\,dx\,dy\,dr\Bigg]$$

$$+\frac{1}{2}E\Bigg[\int_{(0,t]\times Z}\int_{R^d\times R^d}\beta''\big(u(r,x)-v(r,y)\big)$$

$$\times \big(\sigma\big(y, v(r,y); z\big) - \sigma\big(x, u(r,x); z\big)\big)^2 \varphi(x,y)\,dx\,dy\,\mu(dz)\,dr\Bigg]$$

$$+ \sum_{i=1}^{m} A(t_i, t_{i+1}).$$

Taking $\lim_{m\to\infty}$, we arrive at the desired inequality. □

*3.2. Uniqueness*

Using Lemma 3.2 as the point of departure, we now let $\varphi(x,y)\,dx\,dy \to \delta_x(dy)\,dx$ and $\beta(r) \to r_+$ to arrive at an $L^1$ type estimate for $u - v$.

First, we select test function $\varphi$ in the following manner. Let $J$ be a one-dimensional standard mollifier as defined by (16), and let $0 \leqslant \psi \in C_c^\infty(R^d)$. We choose

$$\varphi_\delta(x,y) = \left(\delta^{-d} \prod_{k=1}^{d} J\left(\frac{x_k - y_k}{2\delta}\right)\right) \psi\left(\frac{x+y}{2}\right) \in C_c^\infty\big(R^d \times R^d\big). \tag{24}$$

Then

$$\partial_{x_j}\varphi_\delta(x,y) = \frac{1}{2\delta}\left(\delta^{-d} J'\left(\frac{x_j - y_j}{2\delta}\right) \prod_{k\neq j} J\left(\frac{x_k - y_k}{2\delta}\right)\right)\psi\left(\frac{x+y}{2}\right) + \frac{1}{2}\left(\delta^{-d} \prod_{k=1}^{d} J\left(\frac{x_k - y_k}{2\delta}\right)\right)\partial_j\psi\left(\frac{x+y}{2}\right), \tag{25}$$

$$\partial_{y_j}\varphi_\delta(x,y) = -\frac{1}{2\delta}\left(\delta^{-d} J'\left(\frac{x_j - y_j}{2\delta}\right) \prod_{k\neq j} J\left(\frac{x_k - y_k}{2\delta}\right)\right)\psi\left(\frac{x+y}{2}\right) + \frac{1}{2}\left(\delta^{-d} \prod_{k=1}^{d} J\left(\frac{x_k - y_k}{2\delta}\right)\right)\partial_j\psi\left(\frac{x+y}{2}\right), \tag{26}$$

and

$$(\partial_{x_j} + \partial_{y_j})\varphi_\delta(x,y) = \left(\delta^{-d} \prod_{k=1}^{d} J\left(\frac{x_k - y_k}{2\delta}\right)\right)\partial_j\psi\left(\frac{x+y}{2}\right) \in C_c^\infty\big(R^d \times R^d\big). \tag{27}$$

We let $\beta_\varepsilon$ be defined according to (17), and take $\beta = \beta_\varepsilon$, $\varphi = \varphi_\delta$ in Lemma 3.2. We note that $\beta_\varepsilon(r) \to r_+$ uniformly in $r \in R$ as $\varepsilon \to 0+$. Recall the definition of $\alpha_k, \hat{\alpha}_k$ in (20) and (22), using Lemma 2.9, for each $k = 1, 2, \ldots$, each $u, v \in R^d$ fixed and $\alpha = (\alpha_1, \ldots, \alpha_d)$,

$$\lim_{\varepsilon\to 0+}\alpha_k(u,v)=\lim_{\varepsilon\to 0+}\int_v^\infty \beta_\varepsilon'(u-w)F_k'(w)\,dw$$
$$=\int_v^\infty \chi(u-w)F_k'(w)\,dw=\chi(u-v)\big(F_k(u)-F_k(v)\big)$$
$$=\lim_{\varepsilon\to 0+}\hat\alpha_k(u,v).$$

We recall that by earlier convention on notations, $\chi(r)=\chi_{[0,+\infty)}(r)$.

The right-hand side of the inequality in Lemma 3.2 consists of two terms. In view of the above limit for $\lim_{\varepsilon\to 0+}\alpha_k$ and $\lim_{\varepsilon\to 0+}\hat\alpha_k$, the first term is easier to be controlled in the $\lim_{\delta\to 0+}\lim_{\varepsilon\to 0+}$ limit. However, the second term is easier to be controlled in the $\lim_{\varepsilon\to 0+}\lim_{\delta\to 0+}$ limit. Therefore, we need more careful estimates by considering $\varepsilon\to 0+$, $\delta\to 0+$ at the same time with appropriate speeds.

**Lemma 3.3.** *Suppose that $\varepsilon\to 0+$, $\delta\to 0+$ and $\varepsilon\delta^{-1}\to 0+$ (e.g. let $\delta=\varepsilon^{2/3}$), then*

$$\limsup_{\varepsilon\to 0+,\,\delta\to 0+,\,\varepsilon\delta^{-1}\to 0+} E\left[\int_0^t\left|\int_{R^d\times R^d}\Big(\alpha\big(u(r,x),v(r,y)\big)\cdot\nabla_y\varphi(x,y)\right.\right.$$
$$\left.\left.+\hat\alpha\big(u(r,x),v(r,y)\big)\cdot\nabla_x\varphi(x,y)\Big)\,dx\,dy\right|dr\right]$$
$$\leqslant E\left[\int_0^t\left|\int_{R^d}\sum_{k=1}^d\chi\big(u(r,x)-v(r,x)\big)\big(F_k\big(u(r,x)\big)-F_k\big(v(r,x)\big)\big)\partial_k\psi(x)\,dx\right|dr\right].$$

**Proof.** We need to estimate the difference between $\alpha_k(u,v)$ and $\chi(u-v)(F_k(u)-F_k(v))$ more precisely. When $u\leqslant v$, for $w\geqslant v$, $\beta_\varepsilon'(u-w)=0$, therefore

$$\alpha_k(u,v)=0=\chi(u-v)\big(F_k(u)-F_k(v)\big).$$

When $u>v$, then by Lemma 2.9,

$$\alpha_k(u,v)=\int_v^u\beta_\varepsilon'(u-w)F_k'(w)\,dw$$
$$=\int_v^{v\vee(u-2\varepsilon)}F_k'(w)\,dw+\int_{v\vee(u-2\varepsilon)}^u\beta_\varepsilon'(u-w)F_k'(w)\,dw$$
$$=\chi(u-v)\big(F_k(u)-F_k(v)\big)+\int_{v\vee(u-2\varepsilon)}^u\big(\beta_\varepsilon'(u-w)-1\big)F_k'(w)\,dw.$$

This, together with the at most polynomial growth assumption on $F'$ in Condition 2.3, implies

$$\left|\alpha_k(u,v)-\chi(u-v)\big(F_k(u)-F_k(v)\big)\right|\leqslant 2\int_{u-\varepsilon}^{u}\left|F'(w)\right|dw\leqslant \varepsilon C_p\big(1+|u|^p\big)$$

for some $p\geqslant 1$, where $C_p$ is independent of $\varepsilon$ and $u,v$ and can be chosen to be independent of $k=1,2,\dots,d$ as well. Similar conclusion holds for

$$\hat{\alpha}_k(u,v)=\chi(u-v)\big(F_k(u)-F_k(v)\big)+\int_{v}^{u\wedge(v+2\varepsilon)}\big(\beta'(w-v)-1\big)F_k'(w)\,dw.$$

Combine the above estimate with (25), (26) and (27),

$$\begin{aligned}
&E\left[\int_0^t\left|\int_{R^d\times R^d}\Big(\alpha\big(u(r,x),v(r,y)\big)\cdot\nabla_y\varphi(x,y)\right.\right.\\
&\qquad\left.\left.+\hat{\alpha}\big(u(r,x),v(r,y)\big)\cdot\nabla_x\varphi(x,y)\Big)\,dx\,dy\right|dr\right]\\
&\leqslant E\left[\int_0^t\left|\int_{R^d\times R^d}\sum_{k=1}^{d}\chi\big(u(r,x)-v(r,y)\big)\Big(F_k\big(u(r,x)\big)-F_k\big(v(r,y)\big)\Big)\right.\right.\\
&\qquad\left.\left.\times\big(\partial_{x_k}\varphi_\delta(x,y)+\partial_{y_k}\varphi_\delta(x,y)\big)\,dx\,dy\right|dr\right]\\
&\qquad+\frac{\varepsilon}{\delta}C\sum_{j=1}^{d}E\left[\int_0^t\int_{R^d\times R^d}\big(2+\left|u(r,x)\right|^p+\left|v(r,y)\right|^p\big)\right.\\
&\qquad\left.\times\left(\delta^{-d}\left|J'\left(\frac{x_j-y_j}{2\delta}\right)\right|\prod_{k\neq j}J\left(\frac{x_k-y_k}{2\delta}\right)\right)\psi\left(\frac{x+y}{2}\right)dx\,dy\,dr\right]\\
&\qquad+\varepsilon C\sum_{j=1}^{d}E\left[\int_0^t\int_{R^d\times R^d}\big(2+\left|u(r,x)\right|^p+\left|v(r,y)\right|^p\big)\right.\\
&\qquad\left.\times\left(\delta^{-d}\prod_{k=1}^{d}J\left(\frac{x_k-y_k}{2\delta}\right)\right)\partial_j\psi\left(\frac{x+y}{2}\right)dx\,dy\,dr\right].
\end{aligned}\tag{28}$$

This gives the conclusion of the lemma. □

Next, we estimate the second term on the right-hand side of the inequality in Lemma 3.2.

**Lemma 3.4.** *Suppose that $\varepsilon \to 0+$, $\delta \to 0+$ and $\delta^2\varepsilon^{-1} \to 0+$ (e.g. let $\delta = \varepsilon^{2/3}$), then*

$$\limsup_{\varepsilon\to 0+,\,\delta\to 0+,\,\delta^2\varepsilon^{-1}\to 0+} E\bigg[\int_{(0,t]\times Z}\int_{R^d\times R^d} \beta_\varepsilon''\big(u(r,x)-v(r,y)\big) \times \big|\sigma\big(y,v(r,y);z\big)-\sigma\big(x,u(r,x),z\big)\big|^2\varphi_\delta(x,y)\,dx\,dy\,\mu(dz)\,dr\bigg] \leqslant 0.$$

**Proof.** Select compact $K = K_\psi \subset R^d \times R^d$ to be such that $\operatorname{supp}(\varphi_\delta) \subset K$ for all $0 < \delta < 1$.
Under Condition 2.3 on $\sigma(x,u;z)$, and by the estimate $0 \leqslant \beta_\varepsilon'' \leqslant \varepsilon^{-1}C$,

$$0 \leqslant r\beta_\varepsilon''(r) \leqslant 2C\chi_{[0,2\varepsilon]}(r)$$

(e.g. Lemma 2.9) and that $\beta_\varepsilon''(r) = 0$ for $r \geqslant 2\varepsilon$ or $r \leqslant 0$, for $(x,y) \in K$,

$$\begin{aligned}
&\beta''(u-v)\big|\sigma(y,v;z)-\sigma(x,u,z)\big|^2\\
&\quad\leqslant \big(2\beta''(u-v)|u-v|\rho_K^2\big(|u-v|\big)+2\beta''(u-v)|x-y|^2\big)M_K^2(z)\\
&\quad\leqslant 4C\rho_K^2(2\varepsilon)M_K^2(z)+2C\varepsilon^{-1}|x-y|^2M_K^2(z).
\end{aligned}$$

We have therefore

$$\begin{aligned}
&E\bigg[\int_{(0,t]\times Z}\int_{R^d\times R^d}\beta''\big(u(r,x)-v(r,y)\big)\big(\sigma\big(y,v(r,y);z\big)-\sigma\big(x,u(r,x),z\big)\big)^2\\
&\qquad\times\varphi_\delta(x,y)\,dx\,dy\,\mu(dz)\,dr\bigg]\\
&\quad\leqslant\bigg(4C\rho_K^2(2\varepsilon)\bigg(\int_K\varphi_\delta(x,y)\,dx\,dy\bigg)\int_Z M_K^2(z)\,\mu(dz)+2(C_\psi)\varepsilon^{-1}\delta^2\\
&\qquad\times\int_Z M_K^2(z)\,\mu(dz)\|\psi\|_\infty\bigg)t.
\end{aligned}\tag{29}$$

The conclusion follows. □

**Theorem 3.5.** *Suppose $u$ is a stochastic entropic solution of* (1) *and $v$ is a stochastic strong entropic solution. Then*

(1) (*$L^1$ contraction*)

$$E\big[\big\|\big(u(t)-v(t)\big)_+\big\|_1\big] \leqslant E\big[\big\|\big(u(0)-v(0)\big)_+\big\|_1\big].\tag{30}$$

(2) (*Comparison principle*) *Suppose that $v(0,x) \leqslant u(0,x)$ a.e. in $x$ holds almost surely, and that $E[\|(u(0,\cdot)-v(0,\cdot))_+\|_1] < \infty$, then almost surely*

$$v(t,x) \leqslant u(t,x) \quad \text{a.e. in } x.$$

**Remark 3.6.** The following proof can be simplified considerably if we assume

$$\sup_{0\leqslant t\leqslant T} E\big[\|u(t)\|_1 + \|v(t)\|_1\big] < \infty.$$

However, we adapted a definition of entropic solution which does not assume the above. We have no effective way of establishing such estimates in the absence of additional structural assumptions on $\sigma$.

**Proof.** We define $\varphi_\delta$ according to (24) where $\delta = \varepsilon^{2/3}$. Taking $\varepsilon \to 0+$ limit to the inequality in Lemma 3.2, by Lemmas 3.3 and 3.4,

$$E\bigg[\int_{R^d} \big(u(t,x) - v(t,x)\big)_+ \psi(x)\,dx - \int_{R^d} \big(u(0,x) - v(0,x)\big)_+ \psi(x)\,dx\bigg]$$

$$\leqslant E\Bigg[\int_0^t \Bigg|\int_{R^d} \sum_{k=1}^d \chi\big(u(r,x) - v(r,x)\big)\big(F_k\big(u(r,x)\big) - F_k\big(v(r,x)\big)\big)\partial_k\psi(x)\,dx\Bigg|\,dr\Bigg]. \tag{31}$$

Let $0 \leqslant \psi_N(x) = e^{-N^{-1}|x|} \in W^{1,p}(R^d)$, $p = 1, 2, \ldots, \infty$. Then

$$\partial_k\psi_N(x) = -\frac{1}{N}\frac{x_k}{|x|}\psi_N(x), \qquad x \neq 0,$$

and $\|\partial_k\psi_N\|_\infty \leqslant N^{-1}$. Noting estimate (8) in the definition of entropic solution, by standard approximation by truncation and mollification arguments, (31) holds with $\psi$ replaced by $\psi_N$.

By the at most polynomial growth assumption on $F'$ in Condition 2.3, there exists $p_0 \geqslant 1$, such that for any integer $p > p_0 \geqslant 1$,

$$\chi(u - v)\big|F_k(u) - F_k(v)\big| \leqslant \chi(u - v)\int_v^u \big|F_k'(r)\big|\,dr$$

$$\leqslant C_1\big((u - v)_+ + |u|^{1+p} + |v|^{1+p}\big). \tag{32}$$

By (8) for $u$ and $v$,

$$\bigg|\int_0^t E\bigg[\int_x \chi\big(u(r,x) - v(r,x)\big)\big(F_k\big(u(r,x)\big) - F_k\big(v(r,x)\big)\big)\partial_k\psi_N(x)\,dx\bigg]\,dr\bigg|$$

$$\leqslant C_1 N^{-1}\int_0^t E\bigg[\int_{R^d} \big(u(r,x) - v(r,x)\big)_+ \psi_N(x)\,dx\bigg]\,dr + \theta_N(t), \tag{33}$$

where

$$\theta_N(t) = C_1 N^{-1}\sup_{0\leqslant r\leqslant t} E\big[\|u(r)\|_{1+p}^{1+p} + \|v(r)\|_{1+p}^{1+p}\big].$$

Let $w_N(t) = E[\int_{R^d}(u(t,x) - v(t,x))_+ \psi_N(x)\,dx]$ and $w(t) = E[\int_{R^d}(u(t,x) - v(t,x))_+\,dx]$. Then

$$w_N(t) \leqslant w_N(0) + \theta_N(T) + C_1 N^{-1}\int_0^t w_N(r)\,dr, \quad 0 \leqslant t \leqslant T.$$

By (8) for $u$, $v$, $\sup_{0\leqslant t\leqslant T} w_N(t) < \infty$. Therefore by the Gronwall's inequality,

$$\sup_{0\leqslant t\leqslant T} w_N(t) \leqslant \big(w_N(0) + \theta_N(T)\big)e^{C_1 N^{-1} T}.$$

Send $N \to \infty$. By the monotone convergence theorem, $w_N(t) \to w(t)$ for every $t \geqslant 0$. We arrive at (30).

From (30) it follows that, if $v(0,x) \leqslant u(0,x)$ a.e. in $x$ almost surely, then $v(t,x) \leqslant u(t,x)$ a.e. in $x$ almost surely. □

## 4. A constructive existence theory

### *4.1. Heuristic outlines*

We refer to Dafermos [3] (in particular Chapter VI) for background discussions and references on physical motivation of deterministic conservation laws. The stochastic case can be considered similarly.

We would like to view (1) as limit of some microscopic stochastic system behaving *effectively* like

$$\partial_t u(t,x) + \operatorname{div}_x F\big(u(t,x)\big) = \int_{z\in Z} \sigma\big(x, u(t,x); z\big)\,\partial_t W(t,dz) + \varepsilon\Delta_{xx}u, \quad u(0) = u_0, \quad (34)$$

with asymptotically vanishing $\varepsilon$. The $\varepsilon\Delta_{xx}$ term (with $\varepsilon > 0$) has a smoothing effect on solution $u$. For now, let us pretend that $u = u_\varepsilon$ is a solution of (34) which is sufficiently smooth so that spatial derivatives up to the second order exist in classical sense and are continuous (Lemma 4.10). Let $\Phi \in C^2(R)$ and $\Psi = (\Psi_1,\ldots,\Psi_d)$ be an entropy–entropy flux pair (Definition 2.1). Then by Ito's formula, at least formally,

$$\begin{aligned} &\partial_t \Phi\big(u(t,x)\big) + \operatorname{div}_x \Psi\big(u(t,x)\big) \\ &\quad= \int_{z\in Z} \Phi'\big(u(t,x)\big)\sigma\big(x,u(t,x);z\big)\,\partial_t W(t,dz) + \varepsilon\Phi'\big(u(t,x)\big)\Delta_{xx}u(t,x) \\ &\qquad+ \frac{1}{2}\Phi''\big(u(t,x)\big)\int_Z \sigma^2\big(x,u(t,x);z\big)\,\mu(dz). \end{aligned} \quad (35)$$

It is tempting to send $\varepsilon \to 0$ and arrive at a limit. This is not correct. With each $\varepsilon > 0$ fixed, we can establish second order derivative information about $u$ by exploiting the smoothing effect of $\varepsilon\Delta$. However, we do not know the magnitude of fluctuation for the nonlinear term

$\varepsilon\Phi'(u(t,x))\Delta_{xx}u(t,x)$ as $\varepsilon$ vanishes (note that $u = u_\varepsilon$). This is the exact same difficulty as in deterministic conservation law case ($\sigma = 0$) which is handled as follows. First, we observe

$$\varepsilon\Phi'\big(u(t,x)\big)\Delta_{xx}u(t,x) = \varepsilon\Delta_{xx}\Phi\big(u(t,x)\big) - \varepsilon\Phi''\big(u(t,x)\big)\big|\nabla_x u(t,x)\big|^2. \tag{36}$$

We can always view $\Delta_{xx}\Phi(u(t,x))$ in (Schwartz) distributional sense, as far as $u$ is locally integrable. We do not have control over $|\nabla_x u(t,x)|$ uniformly over $\varepsilon > 0$, and it is not reasonable to hope so (this quantity can blow up in finite time, even if $u(0) \in C^\infty \cap C_\mathrm{b}$, in the case of $\sigma = 0$). However, since $\Phi$ is convex (i.e. $\Phi'' \geqslant 0$) for $\varphi \geqslant 0$, we have a *one-sided* trivial bound

$$\big\langle \varepsilon\Phi''(u)|\nabla_x u|^2, \varphi\big\rangle \geqslant 0.$$

In summary, we have now for $0 \leqslant \varphi \in C_\mathrm{c}^2$,

$$\begin{aligned}
&\big\langle \Phi\big(u(t,\cdot)\big), \varphi\big\rangle - \big\langle \Phi\big(u(s,\cdot)\big), \varphi\big\rangle \\
&\quad = \int_s^t \big\langle \Psi\big(u(r,\cdot)\big), \nabla_x\varphi\big\rangle\, dr + \int_{(s,t]\times Z} \frac{1}{2}\big\langle \Phi''\big(u(r,\cdot)\big)\sigma^2\big(\cdot, u(r,\cdot); z\big), \varphi\big\rangle\, \mu(dz)\, dr \\
&\qquad + \varepsilon \int_s^t \Big(\big\langle \Phi\big(u(r,\cdot)\big), \Delta\varphi\big\rangle - \big\langle \Phi''\big(u(r,\cdot)\big)\big|\nabla_x u(r,\cdot)\big|^2, \varphi\big\rangle\Big)\, dr \\
&\qquad + \int_{(s,t]\times Z} \big\langle \Phi'\big(u(r,\cdot)\big)\sigma\big(\cdot, u(r,\cdot); z\big), \varphi\big\rangle\, W(dr,dz) \\
&\quad \leqslant \int_s^t \big\langle \Psi\big(u(r,\cdot)\big), \nabla_x\varphi\big\rangle\, dr + \int_{(s,t]\times Z} \frac{1}{2}\big\langle \Phi''\big(u(r,\cdot)\big)\sigma^2\big(\cdot, u(r,\cdot); z\big), \varphi\big\rangle\, \mu(dz)\, dr \\
&\qquad + o(\varepsilon) + \int_{(s,t]\times Z} \big\langle \Phi'\big(u(r,\cdot)\big)\sigma\big(\cdot, u(r,\cdot); z\big), \varphi\big\rangle\, W(dr,dz).
\end{aligned} \tag{37}$$

Both the left-hand and right-hand sides of the inequality are stable under $\varepsilon \to 0+$ limit, provided we have $L^p_{\mathrm{loc}}$ type stability of $u = u_\varepsilon$ (which is a lot easier and possible to estimate than $\nabla u_\varepsilon$ or $\Delta u_\varepsilon$). Sending $\varepsilon \to 0$, (10) follows.

As in the deterministic scalar conservation law case, we face two main issues in order to make the above rigorous. One, we need regularity estimates on the solution $u$ for the approximate equation (34) so that the Ito's formula can be applied to the transformation of $u(t,x) = u_\varepsilon(t,x)$. Two, we also need to verify relative compactness on $u = u_\varepsilon$ (in some appropriate topology) as $\varepsilon$ goes to zero. The first issue is more or less known in various different contexts for slightly different models in stochastic analysis literature. We will adapt existing methods and discuss the issue more carefully (because of the generality here) in the first subsection below. The second issue, however, has never been considered in its current generality (i.e. with the stochastic term). Kim [8] modifies deterministic arguments to construct a very special SPDE which is essentially reformulated as a randomness in coefficients type of conservation law. To handle the general situation here, we have to introduce non-trivial new technical ideas. In particular, we derive stochastic versions of the (originally deterministic) compensated compactness results. As in the deterministic case, the compensated compactness argument will eventually restrict our consideration to one space dimension (i.e. $d = 1$) only. However, most of our estimates are not restricted by dimensionality.

In deterministic theory, one can also obtain compactness in $C([0,\infty), L^p)$ (e.g. $p = 1$ or $p = \infty$) by Ascoli–Arzela type argument by estimating corresponding modulus of continuities in time and in space variables (e.g. [3, Section 6.3]). In particular, the spatial modulus of continuity estimate is usually achieved by perturbing conservation law equation through a spatial translation of the solution and then prove a comparison result. Such procedure does not generalize well to the type of SPDEs as in (1). Because of the $\sigma$ term, spatial translation of solution will generally not be another solution or even an approximate solution in some well controlled sense. Similar observation was also made by Kim [8] in a simpler model context.

**Theorem 4.1.** *Suppose that Condition* 2.7 *holds. Then there exists a stochastic strong entropic solution* (*hence also a stochastic entropic solution, see Definitions* 2.6 *and* 2.5) $u = u(t,x)$ *for* (1).

We divide proof into several parts below.

### *4.2. Existence and regularity of approximate equation (34)*

Throughout this subsection, we assume that $F = (F_1, \ldots, F_d)$ satisfies $F_k \in C^\infty$ for each $k$, and that the $m$th order derivative of $F_k$ satisfies $|F_k^{(m)}(r)| \leqslant C_m < \infty$, $m = 0, 1, 2, \ldots.$ We also assume that $\sigma(x,u;z)$, $D_x^m \sigma(x,u;z)$, $\partial_u^m \sigma(x,u;z)$ exist and are continuous and uniformly bounded and $D_x^m \sigma(\cdot,u;z) \in \mathcal{S}(R^d)$ (see (15)), for all $m = 1, 2, \ldots.$ Finally, $\int_Z \sup_{x,u} |\sigma^2(x,u;z)| \mu(dz) < \infty$.

#### *4.2.1. Existence of solution when $\varepsilon > 0$ is held fixed*

Let the fundamental solution of the heat equation be denoted by

$$G(t,x) = G_\varepsilon(t,x) = \frac{1}{(4\pi\varepsilon t)^{d/2}} e^{-|x|^2/(4\varepsilon t)}, \quad t > 0.$$

Let $E[\|u_0\|_2^2] < \infty$. First, we define successive approximates to (34): let $u^0(t,x) = u_0(x)$, $u^n(0,x) = u_0(x)$ and

$$du^n(t,x) + \nabla \cdot F\big(u^{n-1}(t,x)\big)\,dt = \varepsilon \Delta u^n(t,x)\,dt + \int_Z \sigma\big(x, u^{n-1}(t,x); z\big)\, W(dt,dz). \tag{38}$$

We consider the mild solution for the above equation given by

$$\begin{aligned} u^n(t,x) = {} & \int_y G(t, x-y) u_0(y)\,dy - \int_0^t \int_y G(t-s, x-y) \sum_{i=1}^d \partial_{y_i} F_i\big(u^{n-1}(s,y)\big)\,dy\,ds \\ & + \int_{(0,t]\times Z} \int_y G(t-s, x-y) \sigma\big(y, u^{n-1}(s,y); z\big)\,dy\, W(ds,dz). \end{aligned} \tag{39}$$

The above right-hand side needs some explanation. The first term is always well defined because of $E[\|u_0\|_2^2] < \infty$, the second term is well defined provided $D_y u^{n-1} \in L^\infty([0,T]; L^2(R^d))$ for each $T > 0$; the third term is always well defined because of our earlier assumptions on $\sigma$. At this point, it is only clear that $u^1$ is defined. We need regularity information on $u^{n-1}$ to conclude that $u^n$ is defined for $n \geqslant 2$. We claim that all $u^n(t,x)$ are well defined and $u^n(t,\cdot) \in \mathcal{S}(R^d)$. To verify this claim, we need the following properties.

**Lemma 4.2.** *Let $h = h(s) = h(s,x)$ be an adapted process in $C([0,T]; H^p(R^d))$, $p = 1,2,\ldots,$ $T > 0$ and $h(s,\cdot) \in \mathcal{S}(R^d)$. Let*

$$V(t,x) = \int_{(0,t]\times R^d} G(t-s, x-y)h(s,y)\,dy\,ds.$$

*Then $V = V(t) = V(t,x)$ is an adapted process in $C([0,T]; H^p(R^d))$, $T > 0$ and $V(t,\cdot) \in \mathcal{S}(R^d)$ and in particular,*

$$\partial_{x_k} \int_{(0,t]\times R^d} G(t-s, x-y)h(s,y)\,dy\,ds = -\int_{(0,t]\times R^d} G(t-s, x-y)\partial_{y_k} h(s,y)\,dy\,ds.$$

Similarly, if for each $z \in Z$ fixed, $f(\cdot,\cdot;\ z) = f(t,y;z) \in C([0,T]; H^p(R^d))$ is an $\mathcal{F}_t$-adapted process in $t$ for $p = 1,2,\ldots, T > 0$, and

$$E\left[\int_0^T \int_{R^d} \int_Z \left(f^2(s,y;z) + \left|D_y^m f(s,y;z)\right|^2\right) dy\,\mu(dz)\,ds\right] < \infty,$$

where $m = 1,2,\ldots,$ and $f(s,\cdot;\ z) \in \mathcal{S}(R^d)$, then

$$N(t,x) = \int_{(0,t]\times Z} \int_y G(t-s, x-y) f(s,y;z)\,dy\,W(ds,dz), \tag{40}$$

has the following property for each $T > 0$.

**Lemma 4.3.** *$N(t) \in C([0,T]; H^p(R^d))$, $p = 1,2,\ldots,$ and in particular,*

$$\partial_{x_k} N(t,x) = -\int_{(0,t]\times Z} \int_y G(t-s, x-y)\partial_{y_k} f(s,y;z)\,dy\,W(ds,dz), \quad a.s.$$

*and $N(t,\cdot) \in C^\infty(R^d)$.*

**Proof.** Continuity of $N$ as an $L^2$-valued process in $t$ can be handled as in Proposition 7.3 of Da Prato and Zabczyk [4].

Regarding $\partial_{x_k} N$, all we need is to show that (Schwartz) distributional derivative of $N$ agrees with the right-hand side. Then, since the right-hand side is continuous in $x$, the identify is established. For each $\varphi \in C_c^\infty(R^d)$,

$$E\Big[\Big|\int_{[0,t]\times Z}\int_x \big\{\big(G(t-s)*_x f\big)(s,x;z)\partial_{x_k}\varphi(x)$$

$$-G(t-s)*_x(-\partial_k f)(s,x;z)\varphi(x)\big\}\,dx\,W(ds,dz)\Big|^2\Big]$$

$$=\int_{[0,t]\times Z} E\Big[\Big|\int\big(G(t-s)*_x f(s,x;z)\partial_k\varphi(x)$$

$$-G(t-s)*_x(-\partial_k f)(s,x;z)\varphi(x)\big)\,dx\Big|^2\Big]\mu(dz)\times ds=0.$$

In the above, $*_x$ means convolution with respect to spatial variable $x$ only. By such representation, it follows that $\partial_{x_k}N$ has trajectory in $C([0,T];L^2(R^d))$.

Replace $f$ by $\partial_{x_k}f$ and repeat the above arguments, we conclude that $N$ has trajectory in $C([0,T];H^p(R^d))$ and $N(t,\cdot)\in C^p(R^d)$ for $p=1,2,\ldots.$ □

**Lemma 4.4.** *$N(t,\cdot)\in\mathcal{S}(R^d)$ almost surely for each $t$ fixed.*

**Proof.** From Lemma 4.3, we already know that $N(t,\cdot)\in C^\infty(R^d)$. Therefore, we only need to show

$$\sup_{x\in R^d}\big(|x|^m\big|N(t,x)\big|\big)<\infty,\quad \text{a.s.}$$

On the one hand, by a Sobolev (Morrey's) inequality (e.g. [7, (23), p. 268]), there exists deterministic constant $C>0$ when $p\geqslant d$,

$$\sup_{x\in R^d}|x|^m\big|N(t,x)\big|\leqslant C\big\||x|^m N(t,\cdot)\big\|_{W^{1,p}(R^d)}.$$

On the other hand, direct computation shows that for $t>0$, there exist (constant coefficient) $m$th order polynomials of $t$, denoted by $C_i(t)$, $i=0,1,2,\ldots,m$, ($C_0\neq 0$ is a constant), such that

$$t^m\partial^m_{x_k}G(t,x-y)=\big(C_0x_k^m+C_1(t)x_k^{m-1}+\cdots+C_m(t)\big)G(t,x-y).$$

Therefore, by induction, it is sufficient to prove that for all $j=0,1,\ldots,m$,

$$\Big\|\int_{(0,t]\times Z}\int_y |t-s|^j\partial^m_{x_k}G(t-s,x-y)f(s,y;z)\,dy\,W(ds,dz)\Big\|_{W_x^{1,p}}$$

$$=\Big\|\int_{(0,t]\times Z}\int_y |t-s|^j G(t-s,x-y)\partial^m_{y_k}f(s,y;z)\,dy\,W(ds,dz)\Big\|_{W_x^{1,p}}<\infty,\quad \text{a.s.}$$

The above holds because of another Sobolev embedding

$$\|\cdot\|_{W^{1,p}}\leqslant C\|\cdot\|_{H^{1+m}},\quad p\geqslant 2,\ 2m>d,$$

and that because for $i = 0, 1, \ldots,$

$$E\Bigg[\int_x \bigg| \int_{(0,t]\times Z} \int_y |t-s|^j G(t-s, x-y)\partial_{y_k}^{n+i} f(s, y; z)\, dy\, W(ds, dz)\bigg|^2 dx\Bigg]$$

$$\leqslant C \int_x E\Bigg[\int_{(0,t]\times Z} \big(G(t-s) *_x \partial_{y_k}^{n+i} f\big)^2(s, x; z)\, \mu(dz)\, ds\Bigg] dx$$

$$\leqslant CE\Bigg[\int_{[0,t]\times Z} \big\|\partial_{y_k}^{n+i} f(s, \cdot; z)\big\|_2^2 \mu(dz)\, ds\Bigg] < \infty,$$

where the first inequality follows from Burkholder–Davis–Gundy inequality, and the second follows from Young inequality for convolutions. □

In view of Lemmas 4.2–4.4, by induction, we conclude the following.

**Lemma 4.5.** *For each* $n = 1, 2, \ldots,$ $u^n(t) \in C([0, T]; H^p(R^d)),$ $p = 1, 2, \ldots,$ $u^n(t, \cdot) \in \mathcal{S}(R^d),$ $t > 0.$

It is well known that, in the context of stochastic semi-linear equation, under moderate conditions, a mild solution is also a weak solution. The following is a statement of this kind in our present context. Its proof follows, for instance, from a straightforward adaptation of [4, Proposition 6.4].

**Lemma 4.6.** *For each* $\varphi \in C_c^\infty(R^d)$,

$$\big\langle u^n(t), \varphi\big\rangle - \big\langle u^n(0), \varphi\big\rangle = \int_0^t \big\langle \nabla\varphi, F\big(u^{n-1}(s)\big)\big\rangle\, ds$$
$$+ \int_{[0,t]\times Z} \int_x \sigma\big(x, u^{n-1}(r, x); z\big)\varphi(x)\, dx\, W(dr, dz)$$
$$+ \varepsilon \int_0^t \big\langle \Delta\varphi, u^{n-1}(r)\big\rangle\, dr.$$

We define energy functionals $e_{2m} : L^2(R^d) \mapsto [0, \infty]$:

$$e_{2m}(u) = \frac{1}{2}\big\|\Delta^m u\big\|_2^2, \quad m = 0, 1, 2, \ldots. \tag{41}$$

**Lemma 4.7.** *There exist finite constants* $C_{\varepsilon,m,T} > 0$ *which is independent of* $n$ *such that*

$$E\big[e_{2m}\big(u^n(t)\big)\big] \leqslant C_{\varepsilon,m,T}\Bigg(1 + \sum_{k=0}^m E\big[e_{2k}(u_0)\big]\Bigg), \quad t \leqslant T. \tag{42}$$

**Proof.** Apply Ito's formula to $e_{2m}$, there exists finite constant $\tilde{C}_{\varepsilon,m,T} > 0$,

$$E\big[e_{2m}\big(u^n(t)\big)\big] = E\left[e_{2m}\big(u^n(0)\big) + \int_0^t \big(\big\langle \Delta^m F\big(u^{n-1}(r)\big), \Delta^m \nabla u^n(r)\big\rangle - \varepsilon \big\|\Delta^m \nabla u^n(r)\big\|_2^2\big)\,dr\right]$$

$$+ \frac{1}{2} E\left[\int_{[0,t]\times Z} \int_x \big|\Delta_x^m \sigma\big(x, u^{n-1}(r,x); z\big)\big|^2\,dx\,\mu(dz)\,dr\right]$$

$$\leqslant E\big[e_{2m}(u_0)\big] + \tilde{C}_{\varepsilon,m,T} \int_0^t \big(E\big[\big\|\Delta^m\big(F\big(u^{n-1}(r)\big)\big)\big\|_2^2\big]$$

$$+ 1 + E\big[e_0\big(u^{n-1}(r)\big) + \cdots + e_{2m}\big(u^{n-1}(r)\big)\big]\big)\,dr.$$

Denote

$$M_n(t) = \left(1 + \sum_{k=0}^m E\big[e_{2k}\big(u^n(t)\big)\big]\right), \qquad M(0) = \left(1 + \sum_{k=0}^m E\big[e_{2k}(u_0)\big]\right).$$

Then

$$M_n(t) \leqslant cM(0) + c\int_0^t M_{n-1}(s)\,ds$$

where the constant $c > 0$ is independent of $n$. Choose $K$ so large that $c\int_0^T e^{-Kt}\,dt < 1$. Then it follows inductively (in $n$) that

$$M_n(t) \leqslant cM(0)e^{Kt}. \qquad \square$$

We now show that $u^n$ converges in appropriate sense to a limiting process. We adapt a well-known fixed point argument which can be found in proof of part two of Theorem 7.4 in [4], for instance. Because of the term $\mathrm{div}_x\, F(u(t,x))$, the adaptation requires explanation.

**Lemma 4.8.** *There exists a $L^p(R^d)$-valued ($p \geqslant 2$), $\mathcal{F}_t$-adapted process $u$ satisfying*

$$\lim_{n\to\infty} \sup_{0\leqslant t\leqslant T} E\big[\big\|u(t) - u^n(t)\big\|_p^p\big] = 0, \tag{43}$$

*and for $m = 0, 1, 2, \ldots,$*

$$E\big[e_{2m}\big(u(t)\big)\big] \leqslant C_{\varepsilon,m,T}\left(1 + \sum_{k=0}^m E\big[e_{2k}(u_0)\big]\right), \quad t \leqslant T. \tag{44}$$

*In addition,*

$$\sup_{0\leqslant t\leqslant T} E\big[\big\|u(t)\big\|_p^p\big] < \infty. \tag{45}$$

*Moreover, u is a mild solution to* (34) *in the sense that*

$$u(t,x)=\int_y G(t,x-y)u(0,y)\,dy-\int_0^t\int_y G(t-s,x-y)\sum_{i=1}^d \partial_{y_i}F_i\big(u(s,y)\big)\,dy\,ds$$
$$+\int_{(0,t]\times Z}\int_y G(t-s,x-y)\sigma\big(y,u(s,y);z\big)\,dy\,W(ds,dz). \tag{46}$$

**Proof.** First, by direct integration

$$\big\|\partial_{x_i}G(t,\cdot)\big\|_1=\int_x\big|\partial_{x_i}G(t,x)\big|\,dx=Ct^{-1/2},\quad t>0. \tag{47}$$

We denote

$$\big(\mathcal{L}_1\big(u^n\big)\big)(t,x)=\int_0^t\int_{y\in R^d}\sum_{i=1}^d\partial_i G(t-s,x-y)F_i\big(u^n(s,y)\big)\,dy\,ds.$$

For $p\geqslant 1$, $t\geqslant 0$, we define a deterministic measure on $[0,t]$ by

$$m(ds)=m_t(ds)=\big\|\partial_i G(t-s,\cdot)\big\|_1\,ds=2Cd(\sqrt{t}-\sqrt{t-s}\,).$$

Then

$$E\big[\big\|\mathcal{L}_1\big(u^n\big)(t,\cdot)-\mathcal{L}_1\big(u^m\big)(t,\cdot)\big\|_p^p\big]$$
$$\leqslant c_1\sum_{i=1}^d E\left[\left\|\int_0^t\big(\partial_i G(t-s)*\big(u^n(s)-u^m(s)\big)\big)(\cdot)\,ds\right\|_p^p\right]$$
$$\leqslant c_2\sum_{i=1}^d E\left[\left|\int_0^t\big\|\big(\partial_i G(t-s)\big)*\big(u^n(s)-u^m(s)\big)\big\|_p\,ds\right|^p\right]$$
$$\leqslant c_3E\left[\left|\int_0^t\big\|u^n(s)-u^m(s)\big\|_p m(ds)\right|^p\right]$$
$$\leqslant c_4E\left[\int_0^t\big\|u^n(s)-u^m(s)\big\|_p^p m(ds)\right]=c_4\int_0^t E\big[\big\|u^n(s)-u^m(s)\big\|_p^p\big]m(ds)$$
$$\leqslant c_5t^{1/2}\sup_{0\leqslant s\leqslant t}E\big[\big\|u^n(s)-u^m(s)\big\|_p^p\big],$$

where the second inequality follows from the Minkowski's inequality, the third one follows from the Young's inequality for convolutions, the fourth one follows from Jensen's inequality. Therefore,

$$\sup_{0\leqslant s\leqslant t} E\big[\big\|\mathcal{L}_1\big(u^n\big)(s)-\mathcal{L}_1\big(u^m\big)(s)\big\|_p^p\big]\leqslant ct^{1/2}\sup_{0\leqslant s\leqslant t} E\big[\big\|u^n(s)-u^m(s)\big\|_p^p\big].$$

Denote

$$\big(\mathcal{L}_2\big(u^n\big)\big)(t,x)=\int_{(0,t]\times Z}\int_{y\in R^d} G(t-s,x-y)\sigma\big(y,u^n(s,y);z\big)\,dy\,W(ds,dz).$$

By properties of stochastic integral for $p\geqslant 2$,

$$\begin{aligned}
&E\big[\big\|\mathcal{L}_2\big(u^n\big)(t,\cdot)-\mathcal{L}_2\big(u^m\big)(t,\cdot)\big\|_p^p\big]\\
&\quad\leqslant c_6\int_x E\bigg[\bigg(\int_{(0,t]\times Z}\bigg|\int_y G(t-s,x-y)\big(\sigma\big(y,u^n(s,y);z\big)\\
&\qquad -\sigma\big(y,u^m(s,y);z\big)\big)\,dy\bigg|^2\mu(dz)\,ds\bigg)^{p/2}\bigg]\,dx\\
&\quad\leqslant c_7E\bigg[\int_0^t\int_x\big|G(t-s,\cdot)*\big(u^n(s)-u^m(s)\big)(x)\big|^{2(p/2)}\,dx\,ds\bigg]\\
&\quad\leqslant c_8E\bigg[\int_0^t\big(\big\|G(t-s,\cdot)\big\|_1\big\|u^n(s)-u^m(s)\big\|_p\big)^p\,ds\bigg]\\
&\quad\leqslant c_9t\sup_{0\leqslant s\leqslant t}E\big[\big\|u^n(s)-u^m(s)\big\|_p^p\big].
\end{aligned}$$

Combine the above, and apply them to (39), there exist $\rho\in(0,1)$ and $T_0>0$ which are independent of the initial conditions $u^n(0)$, $n=1,2,\ldots,$ such that

$$\big|\big\|u^n-u^m\big\|\big|\equiv\sup_{0\leqslant t\leqslant T_0}\big(E\big[\big\|u^n(t)-u^m(t)\big\|_p^p\big]\big)^{1/p}\leqslant\rho\big|\big\|u^{n-1}-u^{m-1}\big\|\big|.$$

By a fixed point argument and by "pasting" short time existence result to obtain global existence result, we have existence of $u$ satisfying (43).

The same type estimates can be used to show (45). Then (44) follows from (42) and Fatou's lemma, and the mild solution property (46) follows from the fixed point argument applied to (39). □

Taking limit $n\to\infty$ to the equality in Lemma 4.6 we have

**Lemma 4.9.** *For each $\varphi \in C_c^\infty(R^d)$,*

$$\langle u(t),\varphi\rangle - \langle u(0),\varphi\rangle = \int_0^t \langle \nabla\varphi, F(u(s))\rangle\, ds + \int_{[0,t]\times Z}\int_x \sigma(x,u(r,x);z)\varphi(x)\,dx\,W(dr,dz) + \varepsilon\int_0^t \langle \Delta\varphi, u(r)\rangle\, dr.$$

*4.2.2. Regularity of solution when $\varepsilon > 0$ is fixed*

As in the proof of Proposition 7.3 in [4], we can show that

$$U(t,x) = \int_{(0,t]\times Z}\int_y G(t-s,x-y)\sigma(y,u(s,y);z)\,dy\,W(ds,dz)$$

has a continuous modification as $L^2(R^d)$-valued process. Therefore, suppose that $E[e_{2m}(u_0)] < \infty$, then not only do we have (44), it can also be shown *a posterior* that $u \in C([0,\infty); L^2(R^d))$ for all $d = 1, 2, \ldots$.

In the rest of this subsection, we denote $u = u_\varepsilon$ to emphasize its dependence on $\varepsilon > 0$.

**Lemma 4.10.** *Suppose that $E[e_{2m}(u_0)] < \infty$ for $2m \geqslant [d/2] + 3$. Then there exists an $\mathcal{F}_t$-adapted process $u = u(t) \in C([0,\infty); L^2(R^d))$ satisfying almost surely that*

(1) $e_{2m}(u(t)) < \infty$, *for all $t > 0$;*
(2) $\partial_{ij}u = \partial_{x_i,x_j}u(t,\cdot) \in C(R^d)$ *for all $i,j = 1,\ldots,d$.*

*Therefore,* (34) *holds in the classical strong sense. That is, for each $x$ fixed,* (34) *holds as a finite-dimensional stochastic differential equation.*

**Proof.** The conclusions then follow from (44) and from a Sobolev inequality—see Evans [7, Theorem 6, p. 270]. □

Apply Ito's formula to (34), we obtain the following.

**Lemma 4.11.** *Let $\Phi \in C^2(R)$ and convex. If $E[e_{2m}(u_0)] < \infty$ for $2m > [d/2] + 3$, then there exists $\mathcal{F}_t$-adapted solution with properties listed as in Lemma* 4.10 *such that* (37) *holds.*

*4.2.3. Uniform in $\varepsilon$ estimate for solutions of (34)*

We now denote $u_\varepsilon, F_\varepsilon, \sigma_\varepsilon$ to emphasize their dependence on $\varepsilon$. Throughout this subsection, we assume that those conditions on $F_\varepsilon, \sigma_\varepsilon$ at the beginning of Section 4.2 still holds. Next, we derive some estimates which are uniform in $\varepsilon$. For such purpose, we require initial conditions satisfy, for some $2m > [d/2] + 3$,

$$E[e_{2m}(u_\varepsilon(0))] < \infty, \quad \varepsilon > 0,$$

and

$$\sup_{\varepsilon} E\big[\big\|u_\varepsilon(0)\big\|_p^p + \big\|u_\varepsilon(0)\big\|_2^p\big] < \infty, \quad p = 1, 2, \ldots. \tag{48}$$

Finally, we require that

$$\sup_{\varepsilon>0}\big|\sigma_\varepsilon(x, u; z)\big| \leqslant f(x)\big(1 + |u|\big)M(z), \tag{49}$$

where $\int_Z M^2(z)\,\mu(dz) < \infty$ and $f \in L^\infty(R^d) \cap L^2(R^d)$.

**Lemma 4.12.** *For even positive integers $p = 2, 4, 6, \ldots,$*

$$\sup_{\varepsilon} \sup_{0\leqslant t\leqslant T} E\big[\big\|u_\varepsilon(t, \cdot)\big\|_p^p\big] < \infty. \tag{50}$$

**Proof.** By (45), we already know that $\sup_{0\leqslant t\leqslant T} E[\|u_\varepsilon(t)\|_p^p] < \infty$ for every $p \geqslant 2$ and $T \geqslant 0$.

Let $\Phi(u) = (p)^{-1}|u|^p$ and $\Psi(u) = (\Psi_1(u), \ldots, \Psi_d(u))$ be $\Psi_k(u) = \int_0^u \Phi'(r)(F_\varepsilon)_k'(r)\,dr$. Then for each $x \in R^d$ fixed, we apply (35), (36), Lemma 4.10 and integration with respect to $x$ to arrive at

$$\begin{aligned} &E\big[\big\|u_\varepsilon(t)\big\|_p^p\big] - E\big[\big\|u_\varepsilon(0)\big\|_p^p\big] \\ &\quad\leqslant p(p-1)\int_0^t E\bigg[\int_Z\int_x u_\varepsilon^{p-2}(s, x)\sigma_\varepsilon^2\big(x, u_\varepsilon(s, x); z\big)\,dx\,\mu(dz)\bigg]\,ds. \end{aligned}$$

Gronwall inequality (noting (49)) then implies

$$\sup_{0\leqslant t\leqslant T} E\big[\big\|u_\varepsilon(t)\big\|_p^p\big] \leqslant C_T \sup_{\varepsilon>0} E\big[\big\|u_\varepsilon(0)\big\|_p^p\big]. \qquad \square \tag{51}$$

In the case of $p = 2$, since by (37),

$$\begin{aligned} &\big\|u_\varepsilon(t)\big\|_2^2 - \big\|u_\varepsilon(0)\big\|_2^2 \\ &\quad= \int_0^t\int_Z\int_x \sigma_\varepsilon^2\big(x, u_\varepsilon(s, x); z\big)\,dx\,\mu(dz)\,ds - 2\varepsilon\int_{[0,t]\times R^d}\big|\nabla_x u_\varepsilon(s, x)\big|^2\,dx\,ds \\ &\qquad+ \int_{(0,t]\times Z}\int_x u_\varepsilon(s, x)\sigma\big(x, u_\varepsilon(s, x); z\big)\,dx\,W(ds, dz). \end{aligned} \tag{52}$$

This leads to the following estimates:

**Lemma 4.13.** *For each* $p = 1, 2, \ldots$

$$\sup_{\varepsilon>0} E\left[\left(\varepsilon \int_0^T \|\nabla u_\varepsilon(t)\|_2^2\, dt\right)^p\right] < \infty. \tag{53}$$

**Proof.** Apply Ito's formula to $\|u_\varepsilon(t)\|_2^{2p}$ using (52), by (50), (48),

$$\sup_{\varepsilon>0} E\big[\|u_\varepsilon(T)\|_2^{2p} + \|u_\varepsilon(0)\|_2^{2p}\big] < \infty.$$

Note that, (49) and (50) give

$$\sup_{\varepsilon>0} E\left[\left(\int_0^T \int_z \int_x \sigma_\varepsilon^2\big(x, u_\varepsilon(s,x); z\big)\, dx\, \mu(dz)\, ds\right)^p\right] < \infty$$

and by Burkholder–Davis–Gundy inequality and (49),

$$\begin{aligned} &\sup_{\varepsilon>0} E\left[\left|\int_{(0,T]\times Z} \int_x u_\varepsilon(s,x)\sigma\big(x, u_\varepsilon(s,x); z\big)\, dx\, W(ds,dz)\right|^p\right] \\ &\quad \leqslant c \sup_{\varepsilon>0} E\left[\left|\int_{[0,T]\times Z} u_\varepsilon^2(s,x)\sigma^2\big(x, u_\varepsilon(s,x); z\big)\, dx\, \mu(dz)\, ds\right|^{p/2}\right] < \infty. \end{aligned}$$

In view of (52), the conclusion follows from the above estimates. □

More generally, we have the following useful estimate.

**Lemma 4.14.** *Let* $\Phi \in C^2(R)$ *with* $\Phi, \Phi', \Phi''$ *having at most polynomial growth.* $\Phi$ *needs not be convex. Then*

$$\sup_{\varepsilon>0} E\left[\left|\varepsilon \int_0^T \int_{R^d} \Phi''\big(u_\varepsilon(t,x)\big)\big|\nabla_x u_\varepsilon(t,x)\big|^2\, dx\, dt\right|^p\right] < \infty, \quad p = 1, 2, \ldots,\ T > 0. \tag{54}$$

**Proof.** Let $(\Phi, \Psi)$ be an entropy–entropy flux pair. The equality in (37) holds when $u$ is replaced by $u_\varepsilon$, $\varphi = 1$, for general (possibly non-convex) $\Phi \in C^2$. Using (50), the rest of the proof follows that of the previous Lemma 4.13. □

### 4.3. *Convergence of* $\{u_\varepsilon(t,x)\colon \varepsilon > 0\}$ *as measure-valued processes*

We generalize L.C. Young's relaxed measure approach to treat convergence of nonlinear PDEs in this stochastic setting. We identify $u_\varepsilon(t,x)$ with a random measure-valued function

$$\nu_\varepsilon(t,x,du) = \delta_{u_\varepsilon(t,x)}(du).$$

With a slight abuse of notation, we also denote

$$\nu_\varepsilon(t) = \nu_\varepsilon(t, dx, du) = \nu_\varepsilon(t, x, du)\,dx,$$

and view it as a random measure-valued process in the following sense.

Let $\mathcal{M}_0 = \mathcal{M}(R^d \times R)$ be the space of nonnegative Radon measures $\nu$ on $R^d \times R$ with $\nu(dx, R) = dx$. We endow $\mathcal{M}_0$ with a topology $\tau_0$ so that $\nu_n \to \nu \in \mathcal{M}_0$ if and only if $\langle f, \nu_n\rangle \to \langle f, \nu\rangle$ for all $f \in C_b(R^d \times R)$ satisfying $f(x,u) = 0$ when $|x| > k$ for some $k > 0$. $(\mathcal{M}_0, \tau_0)$ is metrizable as follows. We denote

$$\begin{aligned}\Pi^{\nu_1,\nu_2} = \bigl\{\pi \in \mathcal{M}\bigl(R^d \times R \times R^d \times R\bigr)\colon \pi\bigl(dx, du; R^d \times R\bigr) = \nu_1(dx, du);\\ \pi\bigl(R^d \times R; dy, dv\bigr) = \nu_2(dy, dv)\bigr\}, \quad \nu_1, \nu_2 \in \mathcal{M}_0,\end{aligned} \tag{55}$$

and introduce

$$r(\nu_1, \nu_2) = \sum_{k=1}^\infty \frac{1}{2^k} \frac{q_k(\nu_1, \nu_2)}{1 + q_k(\nu_1, \nu_2)},$$

where

$$q_k^2(\nu_1, \nu_2) = \inf\Biggl\{ \int_{|x|\leqslant k,\, |y|\leqslant k} \bigl(|x-y|^2 + |u-v|^2 \wedge 1\bigr)\pi(dx, du; dy, dv)\colon \pi \in \Pi^{\nu_1,\nu_2}\Biggr\}.$$

Note that on each subspace $A_k = \{(x,u); |x| \leqslant k, u \in R\} \subset R^d \times R$, $\nu(A_k) = C_k$ is a finite constant which only depends on $k$; $\sqrt{|x-y|^2 + |u-v|^2 \wedge 1}$ defines a metric on $A_k$ which gives the same topology as the one induced by usual Euclidean distance. Consequently $q_k$ is just a 2-Wasserstein metric on space of measures on $A_k$ with fixed total mass $C_k$. It induces the usual weak convergence topology on such a sub-space of finite measures. See Ambrosio, Gigli and Savaré [1, Chapter 7] for some properties of such metric.

It follows that $(\mathcal{M}_0, r)$ is a complete separable metric space. It can be shown that each $\nu_\varepsilon(t)$ has continuous trajectories in $C([0,\infty), \mathcal{M}_0) \subset M([0,\infty); \mathcal{M}_0)$. Here and below, we write $M([0,\infty); \mathcal{M}_0)$ to denote the space of Borel-measurable, $\mathcal{M}_0$-valued processes on $[0,\infty)$ topologized by a metric

$$d\bigl(\nu_1(\cdot), \nu_2(\cdot)\bigr) = \int_0^\infty e^{-t}\bigl(1 \wedge r\bigl(\nu_1(t), \nu_2(t)\bigr)\bigr)\,dt. \tag{56}$$

$(M([0,\infty); \mathcal{M}_0), d)$ is a complete separable metric space. For properties of such type of space, see Kurtz [12, Section 4]. We have trouble establishing convergence in probability (even along subsequences) of $\{\nu_\varepsilon(\cdot)\colon \varepsilon > 0\}$ in $C([0,\infty); \mathcal{M}_0)$. We will prove convergence in $M([0,\infty); \mathcal{M}_0)$ instead.

By existence of slicing measure (e.g. [6, Theorem 10, p. 14]), for each $\nu \in \mathcal{M}_0$, there exists a probability measure-valued function $\nu(x; \cdot) = \nu(x; du) \in \mathcal{P}(R)$ such that for each $f \in C_b(R^d \times R)$,

(1) $x \mapsto \int_u f(x,u)\nu(x;du)$ is Lebesgue measurable;
(2) $\int_{R^d\times R} f(x,u)\nu(dx,du) = \int_{R^d}\int_R f(x,u)\nu(x;du)\,dx$.

Therefore, each process $\nu(t) \in M([0,\infty), \mathcal{M}_0)$ also admits a representation

$$\nu(t,dx,du) = \nu(t,x;du)\,dx.$$

From this point on, we assume Condition 2.7 holds for $\sigma$ and for $F$, except the condition $d = 1$ (i.e. $d$ may still be any positive integer). We also assume that

$$E\big[\|u_0\|_p^p\big] < \infty, \quad p = 1,2,\ldots.$$

We take $\sigma_\varepsilon$ and $F_\varepsilon$ to be the following particular approximation of $\sigma$ and $F$. Let $J \in C_c^\infty(R)$ be the one-dimensional mollifier in (16), and let $\phi \in C_c^\infty(R)$ be such that $0 \leqslant \phi \leqslant 1$, $\phi(r) = 1$ for $|r| < 1$ and $\phi(r) = 0$ for $|r| > 2$ and $|\phi'(r)| \leqslant 2$. Let $F_\varepsilon = (F_{1,\varepsilon},\ldots,F_{d,\varepsilon})$ with

$$F_{k,\varepsilon}(r) = \big(\phi\big(\varepsilon|r|^2\big)F_k(r)\big) * J_\varepsilon(r),$$

$$\sigma_\varepsilon(x,u;z) = \int_y\int_v \left(\prod_{k=1}^d J_\varepsilon(x_k-y_k)J_\varepsilon(u-v)\right)\big(\phi\big(\varepsilon\big(|y|^2+|v|^2\big)\big)\sigma(y,v;z)\big)\,dy\,dv,$$

where $J_\varepsilon(r) = \varepsilon^{-1}J(r/\varepsilon)$. Then $F_\varepsilon, \sigma_\varepsilon$ satisfy the conditions required at the beginning of Section 4.2. Under (14),

$$\begin{aligned}
&\big|\sigma_\varepsilon(x,u;z) - \sigma(x,u,z)\big| \\
&\quad\leqslant \left(\int_y\int_v L_{\phi,\sigma}\big(|y|+|v|\big)\prod_{k=1}^d J_\varepsilon(y_k)J_\varepsilon(v)\,dy_k\,dv\right)\big(M(z)+\big|\sigma(x,u;z)\big|\big) \\
&\quad\leqslant \varepsilon C\big(M(z)+\big|\sigma(x,u;z)\big|\big),
\end{aligned} \tag{57}$$

where $L_{\phi,\sigma} > 0$ is a constant. Similarly, we can estimate the error for approximating $F$ by $F_\varepsilon$. By part one of Condition 2.3, there exist constants $C > 0$ and $p_0 \in \{1,2,\ldots\}$,

$$\big|F_k'(r+s) - F_k'(r)\big| \leqslant |s|C\big(1+|r|^{p_0}\big), \quad r \in R,\ |s| < 1.$$

Therefore

$$\big|F_{k,\varepsilon}'(r) - F_k'(r)\big| \leqslant \varepsilon C_1\big(1+|r|^{p_0}\big). \tag{58}$$

We now construct a smooth approximation of $u_0$. Let

$$u_\varepsilon(0,x) = \int_{y\in R^d} J_\varepsilon(x-y)\big(u_0(y)\phi\big(\varepsilon|y|^2\big)\big)\,dy \in C_c^\infty\big(R^d\big). \tag{59}$$

Then for each $\varepsilon > 0$ and $m = 1, 2, \dots$ fixed,

$$E\big[e_{2m}\big(u_\varepsilon(0)\big)\big] = \frac{1}{2}E\big[\big\|\Delta^m J_\varepsilon * \big(u_0\phi\big(\varepsilon|y|^2\big)\big)\big\|_2^2\big] \leqslant C_{m,\varepsilon}E\big[\|u_0\|_2^2\big] < \infty,$$

and

$$\sup_{\varepsilon>0} E\big[\big\|u_\varepsilon(0)\big\|_p^p\big] \leqslant E\big[\|u_0\|_p^p\big] < \infty, \quad p = 1, 2, \dots. \tag{60}$$

Using the above error estimates, we can derive the following main result of this section.

**Lemma 4.15.** *There exists an $\mathcal{F}_t$-adapted process $\nu_0(\cdot)$ with trajectory in $M([0,\infty); \mathcal{M}_0)$ such that*

$$\lim_{\varepsilon\to 0+} E\big[r\big(\nu_\varepsilon(t), \nu_0(t)\big)\big] = 0, \quad t \geqslant 0.$$

*This implies that $\{\nu_\varepsilon(\cdot)\colon \varepsilon > 0\}$, as metric space $(M([0,\infty); \mathcal{M}_0), d)$-valued random variables, converges in probability to $\nu_0(\cdot)$. Therefore, for each $0 \leqslant t_1 \leqslant \dots \leqslant t_m$,*

$$\lim_{\varepsilon\to 0+}\big(\nu_\varepsilon(t_1), \dots, \nu_\varepsilon(t_m)\big) = \big(\nu_0(t_1), \dots, \nu_0(t_m)\big), \quad \textit{in probability}.$$

**Proof.** Let $0 \leqslant \varphi_\delta \in C_c^\infty(R^d \times R^d)$ be of the form as in (24), and $\beta_\varepsilon$ be of the form as in Lemma 2.9. We derive an approximate version of the inequality appearing in Lemma 3.2.

Notice that

$$\begin{aligned}
&\Delta_{xx}\beta_\varepsilon\big(u_\theta(t,x) - u_\kappa(t,y)\big) \\
&\quad = \beta_\varepsilon'\big(u_\theta(t,x) - u_\kappa(t,y)\big)\Delta_{xx}u_\theta + \beta_\varepsilon''\big(u_\theta(t,x) - u_\kappa(t,y)\big)\big|\nabla_x u_\theta(t,x)\big|^2.
\end{aligned}$$

By Ito's formula,

$$\begin{aligned}
&\int_{R^d\times R^d} \beta_\varepsilon\big(u_\theta(t,x) - u_\kappa(t,y)\big)\varphi_\delta(x,y)\,dx\,dy \\
&\quad \leqslant \int_{R^d\times R^d} \beta_\varepsilon\big(u_\theta(0,x) - u_\kappa(0,y)\big)\varphi_\delta(x,y)\,dx\,dy + M(t) + A^1(t) + A^2(t) + A^3(t),
\end{aligned}$$

with non-decreasing processes

$$\begin{aligned}
A^1(t) = A^1_{\varepsilon,\delta,\theta,\kappa}(t) = \int_0^t \bigg| \int_{R^d\times R^d} &\beta_\varepsilon'\big(u_\theta(r,x) - u_\kappa(r,y)\big)\big[-\operatorname{div}_x F_\theta\big(u_\theta(r,x)\big) \\
&+ \operatorname{div}_y F_\kappa\big(u_\kappa(r,y)\big)\big]\varphi_\delta(x,y)\,dx\,dy\bigg|\,dr,
\end{aligned}$$

$$A^2(t) = A^2_{\varepsilon,\delta,\theta,\kappa}(t) = \frac{1}{2}\int_0^t \int_{R^d\times R^d} \left|\beta''_\varepsilon\big(u_\theta(r,x) - u_\kappa(r,y)\big)\right| \int_z \left|\sigma_\theta\big(x, u_\theta(r,x); z\big) - \sigma_\kappa\big(y, u_\kappa(r,y); z\big)\right|^2 \mu(dz)\varphi_\delta(x,y)\,dx\,dy\,dr,$$

$$A^3(t) = A^3_{\varepsilon,\delta,\theta,\kappa}(t) = \int_0^t \int_{R^d\times R^d} \beta_\varepsilon\big(u_\theta(r,x) - u_\kappa(r,y)\big)\left|(\theta\Delta_{xx} - \kappa\Delta_{yy})\varphi_\delta(x,y)\right|dx\,dy\,dr,$$

and martingale

$$M(t) = M_{\varepsilon,\delta,\theta,\kappa}(t) = \int_{R^d\times R^d} \int_{(0,t]\times Z} \beta'_\varepsilon\big(u_\theta(r,x) - u_\kappa(r,y)\big) \times \big(\sigma_\theta\big(x, u_\theta(r,x); z\big) - \sigma_\kappa\big(y, u_\kappa(r,y); z\big)\big)\, W(dr,dz)\varphi_\delta(x,y)\,dx\,dy.$$

We can invoke the *a priori* estimates in Lemma 4.12 to estimate the $A^k$s. First, since $\varphi_\delta$ has compact support in $x, y$, uniformly in $\delta > 0$, and since

$$0 \leqslant \beta_\varepsilon(r) \leqslant r, \quad \sup_{x,y}\left|(\theta\Delta_{xx} - \kappa\Delta_{yy})\varphi_\delta(x,y)\right| \leqslant C(\theta+\kappa)\delta^{-2},$$

by (50), for each $t > 0$,

$$\lim_{\theta\to 0+,\,\kappa\to 0+,\,\delta\to 0+,\,\varepsilon\to 0+,\,\delta^{-2}(\theta+\kappa)\to 0+} E\big[A^3(t)\big] = 0.$$

Noting $0 \leqslant \beta''(r) \leqslant \varepsilon^{-1}C$ (Lemma 2.9),

$$E\big[A^2(t)\big] \leqslant I + II + III$$

where

$$I = \frac{1}{2}\int_0^t E\bigg[\int_{R^d\times R^d} \left|\beta''_\varepsilon\big(u_\theta(r,x) - u_\kappa(r,x)\big)\right| \times \int_z \left|\sigma_\theta\big(x, u_\theta(r,x); z\big) - \sigma_\kappa\big(y, u_\kappa(r,y); z\big)\right|^2 \mu(dz)\varphi_\delta(x,y)\,dx\,dy\bigg]dr,$$

$$II = \varepsilon^{-1}C\int_0^t E\bigg[\int_{R^d\times R^d} \int_z \left|\sigma_\theta\big(x, u_\theta(r,x); z\big) - \sigma\big(x, u_\theta(r,x); z\big)\right|^2 \mu(dz)\varphi_\delta(x,y)\,dx\,dy\bigg]dr,$$

$$III = \varepsilon^{-1}C\int_0^t E\bigg[\int_{R^d\times R^d} \int_z \left|\sigma\big(y, u_\kappa(r,y); z\big) - \sigma_\kappa\big(y, u_\kappa(r,y); z\big)\right|^2 \mu(dz)\varphi_\delta(x,y)\,dx\,dy\bigg]dr.$$

By (57),

$$\lim_{\varepsilon\to 0+,\,\delta\to 0+,\,\kappa\to 0+,\,\theta\to 0+;\,\varepsilon^{-1}(\theta^2+\kappa^2)\to 0+}(II+III)=0.$$

By identical arguments as in the proof of Lemma 3.4, we have (29) holds when $u(r,x)$ and $v(r,y)$ are replaced by $u_\theta(r,x)$ and $u_\kappa(r,y)$, respectively. Therefore

$$\lim_{\varepsilon\to 0+,\,\delta\to 0+,\,\theta\to 0+,\,\kappa\to 0+,\,\delta^2\varepsilon^{-1}\to 0+} I=0.$$

In summary

$$\lim_{\theta\to 0+,\,\kappa\to 0+,\,\delta\to 0+,\,\varepsilon\to 0+,\,\varepsilon^{-1}(\theta^2+\kappa^2)\to 0+,\,\varepsilon^{-1}\delta^2\to 0+} E\big[A^2(t)\big]=0.$$

We now estimate $A^1$. First, we approximate $\alpha=\alpha_\varepsilon, \hat\alpha=\hat\alpha_\varepsilon$ in (20) and (22) by $\alpha_{\varepsilon,\theta}=(\alpha_{\varepsilon,\theta;1},\dots)$ and $\hat\alpha_{\varepsilon,\theta}=(\hat\alpha_{\varepsilon,\theta;1},\dots)$

$$\alpha_{\varepsilon,\theta;k}(u,v)=\int_v^\infty \beta'_\varepsilon(u-w)F'_{\theta,k}(w)\,dw=\int_v^u \beta'_\varepsilon(u-w)F'_{\theta,k}(w)\,dw,$$

$$\hat\alpha_{\varepsilon,\theta;k}(u,v)=\int_{-\infty}^u \beta'_\varepsilon(w-v)F'_{\theta,k}(w)\,dw=\int_v^u \beta'_\varepsilon(w-v)F'_{\theta,k}(w)\,dw.$$

By (58), there exists $p>1$ such that

$$\big|\alpha_{\varepsilon,\theta;k}(u,v)-\alpha_{\varepsilon;k}(u,v)\big|\leqslant C\theta\big(1+|u|^p+|v|^p\big).$$

Similar estimate holds for $\hat\alpha_{\varepsilon,\theta;k}-\hat\alpha_{\varepsilon;k}$. Therefore

$$\begin{aligned}
A^1(t)&=\int_0^t\bigg|E\bigg[\int_{R^d\times R^d}\Big(\alpha_{\varepsilon,\theta}\big(u_\theta(r,x),u_\kappa(r,y)\big)\nabla_y\varphi_\delta(x,y)\\
&\qquad+\hat\alpha_{\varepsilon,\kappa}\big(u_\theta(r,x),u_\kappa(r,y)\big)\nabla_x\varphi_\delta(x,y)\Big)\,dx\,dy\bigg]\bigg|\,dr\\
&\leqslant\int_0^t\bigg|E\bigg[\int_{R^d\times R^d}\Big(\alpha_\varepsilon\big(u_\theta(r,x),u_\kappa(r,y)\big)\nabla_y\varphi_\delta(x,y)\\
&\qquad+\hat\alpha_\varepsilon\big(u_\theta(r,x),u_\kappa(r,y)\big)\nabla_x\varphi_\delta(x,y)\Big)\,dx\,dy\bigg]\bigg|\,dr\\
&\qquad+C\Big(\frac{\theta}{\delta}+\frac{\kappa}{\delta}\Big)E\bigg[\int_s^t\int_{R^d\times R^d}\Big(1+\big|u_\theta(t,x)\big|^p+\big|v_\kappa(t,y)\big|^p\Big)\\
&\qquad\times\bigg(\delta^{-d}\bigg|J'\Big(\frac{x_j-y_j}{2\delta}\Big)\bigg|\prod_{k\neq j}J\Big(\frac{x_k-y_k}{2\delta}\Big)\bigg)\psi\Big(\frac{x+y}{2}\Big)\,dx\,dy\,dr\bigg].
\end{aligned}$$

We note that the estimates in (28) still hold with $u, v$ replaced by $u_\theta, v_\kappa$. We write $\lim = \limsup_{\varepsilon\to 0+,\delta\to 0+,\kappa\to 0+,\theta\to 0+,\frac{\theta+\kappa}{\delta}\to 0+,\frac{\varepsilon}{\delta}\to 0+}$. By (27) and (32),

$$\begin{aligned}
&\lim A^1(t) \\
&\quad\leqslant \lim E\Bigg[\int_0^t \Bigg|\int_{R^d\times R^d} \sum_{j=1}^d \chi\big(u_\theta(r,x)-u_\kappa(r,y)\big)\big(F_j\big(u_\theta(r,x)\big)-F_j\big(u_\kappa(r,y)\big)\big) \\
&\qquad\times J_\delta(x-y)\partial_j\psi\left(\frac{x+y}{2}\right)dx\,dy\Bigg|\,dr\Bigg] \\
&\quad\leqslant C\lim E\Bigg[\int_0^t \int_{R^d\times R^d} \big|u_\theta(r,x)-u_\kappa(r,y)\big| J_\delta(x-y)\partial_j\psi\left(\frac{x+y}{2}\right)dx\,dy\,dr\Bigg],
\end{aligned}$$

where $C$ is a constant independent of the choice of $\psi$.

By symmetry, we also have similar estimates when the roles of $u_\theta(t,x)$ and $u_\kappa(t,y)$ are reversed.

Now, we let $\theta,\kappa\to 0+$ and take $\varepsilon=\sqrt{\theta}\vee\sqrt{\kappa}$, $\delta=\varepsilon^{2/3}$. Then $(\theta+\kappa)\delta^{-2}\to 0+$, $(\theta^2+\kappa^2)\varepsilon^{-1}\to 0+$, and $\delta^2\varepsilon^{-1}\to 0+$, $\varepsilon\delta^{-1}\to 0+$. From the construction of $\beta_\varepsilon$, there exists a constant $C_0>0$ such that

$$\big|\big(\beta_\varepsilon(r)+\beta_\varepsilon(-r)\big)-|r|\big|\leqslant \varepsilon C_0.$$

Denote

$$m_\psi(t)=\lim E\Bigg[\int_{R^d\times R^d} \big|u_\theta(t,x)-u_\kappa(t,y)\big| J_\delta(x-y)\psi\left(\frac{x+y}{2}\right)dx\,dy\Bigg].$$

It follows from the above estimates that

$$m_\psi(t)-m_\psi(s)\leqslant C\int_s^t \sum_{j=1}^d m_{\partial_j\psi}(r)\,dr,\quad 0\leqslant s\leqslant t.$$

By simple approximation, the above still holds when $\psi=\psi_N(x)=e^{-N^{-1}|x|}\in W^{1,p}(R^d)$, $p=1,2,\ldots,\infty$. As in the proof of Theorem 3.5, it follows (by Gronwall inequality) then

$$\lim_{N\to\infty} m_{\psi_N}(t)\leqslant \lim_{N\to\infty} m_{\psi_N}(0).$$

We now recall the way $u_\theta(0,x)$ is constructed in (59), by the integrability estimates in (60),

$$\lim_{N\to\infty} m_{\psi_N}(0) \leqslant \lim_{N\to\infty}\left(\lim_{\theta\to0+} E\left[\int_x \left|u_\theta(0,x)-u_0(x)\right|\psi_N(x)\,dx\right]\right.$$
$$\left.+\lim_{\kappa\to0+} E\left[\int_y \left|u_\kappa(0,y)-u_0(y)\right|\psi_N(y)\,dy\right]\right)=0.$$

We now estimate $\lim_N m_{\psi_N}(t)$ using $q_K(\nu_\theta(t),\nu_\kappa(t))$ for $K=1,2,\ldots$. Let stochastic measure

$$\pi_t(dx,du;dy,dv)=\delta_{u_\kappa(t,y)}(dv)\big(J_\delta(x-y)\,dy\big)\delta_{u_\theta(t,x)}(du)\,dx.$$

Then $\pi_t\in\Pi^{\nu_\theta(t),\nu_\kappa(t)}$ (see definition in (55)), for $N\geqslant K$,

$$\int_{R^d\times R^d} \left|u_\theta(t,x)-u_\kappa(t,y)\right|J_\delta(x-y)\psi_N\left(\frac{x+y}{2}\right)dx\,dy$$
$$\geqslant e^{-1}\int_{u,v\in R;|x|\leqslant K,|y|\leqslant K} |u-v|\pi_t(dx,du;dy,dv)$$
$$\geqslant e^{-1}q_K^2\big(\nu_\theta(t),\nu_\kappa(t)\big)-e^{-1}\int_{|x|<K,|y|<K} |x-y|^2J_\delta(x-y)\,dx\,dy.$$

This implies that, for each $K$ fixed,

$$\lim_{\theta,\kappa\to0+} E\big[q_K^2\big(\nu_\theta(t),\nu_\kappa(t)\big)\big]$$
$$\leqslant e\lim_{N\to\infty}\lim_{\kappa\to0+,\,\theta\to0+} E\left[\int_{(x,y)\in R^d\times R^d} \left|u_\theta(0,x)-u_\kappa(0,y)\right|J_\delta(x-y)\psi_N\left(\frac{x+y}{2}\right)dx\,dy\right]$$
$$\leqslant \lim_{N\to\infty} m_{\psi_N}(0)\leqslant 0.$$

By definition of $d$ in (56) and by dominated convergence theorem

$$\lim_{\theta,\kappa\to0+} E\big[d\big(\nu_\theta(\cdot),\nu_\kappa(\cdot)\big)\big]=\int_0^\infty e^{-t}\lim_{\theta,\kappa\to0+} E\big[1\wedge r\big(\nu_\theta(t),\nu_\kappa(t)\big)\big]\,dt=0.$$

There exists a process $\nu_0(\cdot)\in M([0,\infty);\mathcal{M}_0)$ satisfying the conclusion of the lemma. □

**Remark 4.16.** We suspect that $\lim_{\varepsilon\to0+}\nu_\varepsilon=\nu_0$ in probability as $C([0,\infty);\mathcal{M}_0)$-valued random variables. However, due to a lack of good control of $E[\sup_{0\leqslant t\leqslant T}|M(t)|]$ by $r(\nu_\theta(t),\nu_\kappa(t))$, and due to a lack of uniform estimate on modulus of continuity (in space $C([0,\infty);\mathcal{M}_0)$) for $\nu_\varepsilon$, we cannot confirm this.

### 4.4. Stochastic compensated compactness and existence of function-valued limit process

We generalize the deterministic compensated compactness argument to a stochastic setting, showing that $\nu_0(t,x,du)$ as a (random) measure in $u$ has to be point mass. Consequently it can be identified with a random function $u_0(t,x)$. Unlike the rest of the paper, we assume an extra assumption $d=1$ throughout this section.

Let

$$\bar{u}(t,x) = \int_{u\in R} u\,\nu_0(t,x;du).$$

The main result of this section is

**Lemma 4.17.** *The following holds*:

$$\int_{u\in R} F(u)\nu_0(t,x,du)\,dt\,dx = F\big(\bar{u}(t,x)\big)\,dt\,dx, \quad \text{a.s.}$$

*Furthermore, if the set of $r$ such that $F''(r)\neq 0$ is dense in $r\in R$, then*

$$\nu_0(t,dx,du) = \delta_{\bar{u}(t,x)}(du)\,dx, \quad \text{almost surely}. \tag{61}$$

We divide the proof into several parts which are proved in subsections that follows.

First, it is useful to further relax our view point by viewing $\nu_\varepsilon(dt,dx,du) = \nu_\varepsilon(t,dx,du)\,dt$ (where $\varepsilon \geqslant 0$) as random measures on $[0,\infty)\times R^d \times R$. Let $\mathcal{M} = \mathcal{M}([0,\infty)\times R^d\times R)$ be the space of nonnegative Radon measures $\nu$ on $[0,\infty)\times R^d\times R$ satisfying $\nu(dt,dx,R) = dt\times dx$. We endow $\mathcal{M}$ with a variant of weak topology so that $\nu_n\to\nu$ if and only if $\langle f,\nu_n\rangle \to \langle f,\nu\rangle$ for all $f = f(t,x,u)\in F\subset C_{\mathrm{b}}([0,\infty)\times R^d\times R)$. The set $F$ consists of bounded continuous functions $f$ with compact support in $t,x$ uniformly in $u$. That is, there exists $C = C_f$ such that $f(t,x;u) = 0$ once $t+|x|>C$, for all $u\in R$. As in the $\mathcal{M}_0$ introduced earlier, there is a metrizable topology on $\mathcal{M}$ which coincide with the above notion of sequential convergence and turning $(\mathcal{M},\tau)$ into a Polish space. Therefore, $\{\nu_\varepsilon\colon \varepsilon>0\}$ is a sequence of $\mathcal{M}$-valued random variables.

Note that $M([0,\infty),\mathcal{M}_0)$ can be continuously embedded into $\mathcal{M}$.

Let $(\Phi,\Psi)$ be a given entropy–entropy flux pair with $\Phi,\Phi',\Phi''$ having at most polynomial growth. We define Ito's integral

$$M_\varepsilon(t,x) = \int_{[0,t]\times Z} \Phi'\big(u_\varepsilon(r,x)\big)\sigma_\varepsilon\big(x,u_\varepsilon(r,x);z\big)\,W(dr,dz)$$

and let

$$\Phi_\varepsilon(t,x) = \Phi\big(u_\varepsilon(t,x)\big), \qquad \Psi_\varepsilon(t,x) = \Psi\big(u_\varepsilon(t,x)\big),$$

and

$$\chi_\varepsilon(t,x) = \chi_{\varepsilon,1}+\chi_{\varepsilon,2}, \qquad \psi_\varepsilon(t,x) = \psi_{\varepsilon,1}+\psi_{\varepsilon,2},$$

where

$$\chi_{\varepsilon,1}(t,x)=\varepsilon\partial_x^2\Phi\big(u_\varepsilon(t,x)\big)=\varepsilon\partial_x\big(\Phi'\big(u_\varepsilon(t,x)\big)\partial_x u_\varepsilon(t,x)\big),$$
$$\chi_{\varepsilon,2}(t,x)=\partial_t M_\varepsilon(t,x);$$

and

$$\psi_{\varepsilon,1}(t,x)=\frac{1}{2}\int_Z \Phi''\big(u_\varepsilon(t,x)\big)\sigma_\varepsilon^2\big(x,u_\varepsilon(t,x);z\big)\,\mu(dz),$$
$$\psi_{\varepsilon,2}(t,x)=-\varepsilon\Phi''\big(u_\varepsilon(t,x)\big)\big|\partial_x u_\varepsilon(t,x)\big|^2.$$

The meaning of $\chi_{\varepsilon,2}$ is given as follows. $M_\varepsilon(t)=M_\varepsilon(t,x,\omega)$ is a continuous function in $t$ for each $x,\omega\in\Omega$ fixed. We can take Schwartz distributional derivative $\partial_t$ in $t$ of $M_\varepsilon$ and such derivative is $\chi_{\varepsilon,2}$.

The equality in (37) (Lemma 4.11) is a statement that,

$$\partial_t\Phi_\varepsilon(t,x,\omega)+\partial_x\Psi_\varepsilon(t,x,\omega)=\chi_\varepsilon(t,x,\omega)+\psi_\varepsilon(t,x,\omega) \tag{62}$$

holds $\omega$-wise. Note again, $\partial_t$, $\partial_x$ above should all be understood in Schwartz distributional sense.

Let $T>0$ be an arbitrarily given but fixed constant. We denote $\mathcal{O}=(0,T)\times R$.

*4.4.1. A priori estimates for several sequences of random fields*

The main result of this subsection is the following.

**Lemma 4.18.** $\{\partial_t\Phi_\varepsilon+\partial_x\Psi_\varepsilon\colon \varepsilon>0\}$ *is a sequence of* $H^{-1}_{\mathrm{loc}}(\mathcal{O})$*-valued random variables. As such random variables, the sequence is tight.*

**Proof.** We apply a stochastic generalization of the Murat lemma (see Lemma A.3) to show that, as $H^{-1}_{\mathrm{loc}}(\mathcal{O})$-valued random variables, the left-hand side of (62) is tight. This only requires verifying conditions of Lemma A.3.

By the integrability conditions on $u_\varepsilon$ in Lemma 4.12, $\{\Phi_\varepsilon\colon \varepsilon>0\}$ and $\{\Psi_\varepsilon\colon \varepsilon>0\}$ are both stochastically bounded as $L^p_{\mathrm{loc}}(\mathcal{O})$-valued random variables, $2\leqslant p<\infty$. Therefore, the left-hand side of (62) is a stochastically bounded sequence in $W^{-1,p}_{\mathrm{loc}}(\mathcal{O})$.

By the moment estimates (50) on $u_\varepsilon$ in Lemma 4.13, $\{\psi_{\varepsilon,1}\colon \varepsilon>0\}$ is stochastically bounded in $L^2_{\mathrm{loc}}(\mathcal{O})$, hence it is stochastically bounded in as random variable in $\mathcal{M}_{\mathrm{loc}}(\mathcal{O})$ (space of Radon signed-measures on each fixed bounded open subset of $\mathcal{O}$) with total variation norm. By (54), $\{\psi_{\varepsilon,2}\colon \varepsilon>0\}$ is also stochastically bounded (in total variation norm), as $\mathcal{M}_{\mathrm{loc}}(\mathcal{O})$-valued random variables.

By (53), $\lim_{\varepsilon\to 0+}\chi_{\varepsilon,1}=0$ in probability as sequence of $H^{-1}_{\mathrm{loc}}(\mathcal{O})$-valued random variables, and is therefore tight. Finally, we claim that the set of $H^{-1}_{\mathrm{loc}}(\mathcal{O})$-valued random variables $\{\chi_{\varepsilon,2}\colon \varepsilon>0\}$ is tight (which we will prove in Lemma 4.20 below). □

We call a deterministic function $0 \leqslant \varphi(t,x) \leqslant 1$ a cutoff function, if $\varphi(t,x) \in C_c^\infty([0,\infty) \times R^d)$. We use it to discuss local properties of functions such as local integrability and so on. We localize $M_\varepsilon$ by

$$M_\varepsilon^\varphi(t,x) = \int_0^t \varphi(r,x)\, M_\varepsilon(dr,x).$$

We set $M_\varepsilon^\varphi(t,x) = 0$ for $t < 0$.

In order to estimate some fractional derivatives of $M_\varepsilon$ locally, it will be convenient for us to introduce the notion of Marchaud fractional derivative for $\alpha \in (0,1)$ (Samko, Kilbas, Marichev [16, Section 5.4]). We define

$$\big(D_\pm^\alpha \phi\big)(t) = \frac{-\alpha}{\Gamma(1-\alpha)} \int_0^\infty \frac{\phi(t \mp s) - \phi(t)}{s^{1+\alpha}}\, ds,$$

for those $\phi$ where the integrand above is $L^1$ integrable. At least for $\phi \in C_c^\infty(R)$ [16, Section 5.7],

$$\left(-\frac{d^2}{dt^2}\right)^{\alpha/2} \phi = \left(\frac{d}{dt}\right)^\alpha \phi = D_+^\alpha \phi.$$

Provided $f \in L_{\mathrm{loc}}^s(R)$, $D_-^\alpha f \in L_{\mathrm{loc}}^p(R)$ for some $p > 1$ and $s^{-1} = p^{-1} - \alpha$,

$$\int_{-\infty}^\infty f(t) D_+^\alpha \phi(t)\, dt = -\int_{-\infty}^\infty \phi(t) D_-^\alpha f(t)\, dt, \quad \phi \in C_c^\infty(R).$$

See [16, Corollary 2 of Theorem 6.2]. Therefore, in such cases, Schwartz distributional derivative $\partial^\alpha f = D_-^\alpha f$.

Recall that $M_\varepsilon^\varphi(t,x)$ is a continuous (in time) local martingale with each $x$ fixed. Hence it is almost surely Hölder continuous in $t$ (almost everywhere) with exponent $0 < \beta < 1/2$ when $x$ is fixed (Revuz, Yor [15, Exercise 1.20, p. 187]). Consequently, for $0 < \alpha < \beta < 1/2$, $\partial_t^\alpha M_\varepsilon^\varphi = D_-^\alpha M_\varepsilon^\varphi$.

**Lemma 4.19.** *Assume that* (49) *holds. Let $\varphi$ be a cutoff function. Then there exists an $0 < \alpha < 1/2$ such that, as $H_{\mathrm{loc}}^{-1+\alpha}(\mathcal{O})$-valued random variables, $\{\partial_t M_\varepsilon^\varphi \colon \varepsilon > 0\}$ are stochastically bounded. That is, for each $\delta > 0$, there exists a constant $C_\delta > 0$ with*

$$\sup_{\varepsilon > 0} P\big(\big\| \partial_t M_\varepsilon^\varphi \big\|_{-1+\alpha} > C_\delta\big) < \delta.$$

**Proof.** First, the integrability estimates in (50) imply that

$$\sup_{\varepsilon>0} E\big[\big\| M_\varepsilon^\varphi \big\|_2^2\big] < \infty. \tag{63}$$

Therefore, we only need to prove that, for each $\delta > 0$, there exists a constant $C_\delta > 0$ satisfying

$$\sup_{\varepsilon>0} P\big(\big\|\partial_t^\alpha M_\varepsilon^\varphi\big\|_2 > C_\delta\big) < \delta. \tag{64}$$

We verify this next.

Recall that we assume $\Phi(u)$ is of at most polynomial growth as $u \to \infty$. Take $\gamma > 6$, then for $0 \leqslant s < t \leqslant T$, there exists $p_0 > 2$ such that

$$\begin{aligned}
&E\big[\big\|M_\varepsilon^\varphi(t,\cdot) - M_\varepsilon^\varphi(s,\cdot)\big\|_\gamma^\gamma\big] \\
&\leqslant C_1 \int_{|x|<c_\varphi} E\left[\left|\int_s^t (\Phi')^2\big(u_\varepsilon(r,x)\big)\varphi^2(r,x)\int_Z \sigma_\varepsilon^2\big(x, u_\varepsilon(r,x); z\big)\,\mu(dz)\,dr\right|^{\gamma/2}\right] dx \\
&\leqslant C_2 E\left[\int_{|x|<c_\varphi} \left|\int_s^t \varphi^2(r,x)\big(1 + \big|u_\varepsilon(r,x)\big|^{p_0}\big)\,dr\right|^{\gamma/2} dx\right] \\
&\leqslant C_3|t-s|^{\gamma/2-1} E\left[\int_0^T \big(1 + \big\|u_\varepsilon(r)\big\|_{p_0\gamma/2}^{p_0\gamma/2}\big)\,dr\right] \leqslant C_4|t-s|^{\gamma/2-1}.
\end{aligned}$$

The first inequality above follows from martingale inequalities, the second one follows from (49) and the third one from Jensen's inequality. The last inequality follows from (50).

From the above,

$$E\left[\int_0^T\int_0^T \left(\frac{\|M_\varepsilon^\varphi(t) - M_\varepsilon^\varphi(s)\|_\gamma}{|t-s|^{1/p-1/\gamma}}\right)^\gamma ds\,dt\right] \leqslant C_5 < \infty.$$

By Chebychev's inequality,

$$P\left(\int_0^T\int_0^T \left(\frac{\|M_\varepsilon^\varphi(t) - M_\varepsilon^\varphi(s)\|_\gamma}{|t-s|^{1/p-1/\gamma}}\right)^\gamma ds\,dt > \lambda\right) \leqslant C_6\lambda^{-1}, \quad \lambda > 0. \tag{65}$$

By a normed space version of Garsia's inequality (e.g. Stroock, Varadhan [17, Exercise 2.4.1, p. 60]), if

$$\int_0^T\int_0^T \left(\frac{\|M_\varepsilon^\varphi(t) - M_\varepsilon^\varphi(s)\|_{L^\gamma}}{|t-s|^{1/2-1/\gamma}}\right)^\gamma ds\,dt \leqslant \lambda$$

then

$$\big\|M_\varepsilon^\varphi(t) - M_\varepsilon^\varphi(s)\big\|_\gamma \leqslant C_\lambda|t-s|^{-3/\gamma+1/2},$$

for some deterministic constant $C_\lambda$. This implies that, for $0 < \alpha < \beta \equiv 1/2 - 3/\gamma$,

$$\left\| \partial_t^\alpha M_\varepsilon^\varphi \right\|_\gamma^\gamma = \left\| D_{-,t}^\alpha M_\varepsilon^\varphi \right\|_\gamma^\gamma \leqslant C_7 \int\limits_t \int\limits_x \left| \int\limits_s \frac{|M_\varepsilon^\varphi(t+s,x) - M_\varepsilon^\varphi(t,x)|}{s^{(1+\alpha)}}\, ds \right|^\gamma dx\, dt$$

$$\leqslant C_8 \int\limits_t \left| \int\limits_{s=0}^{\infty} \frac{\|M_\varepsilon^\varphi(t+s) - M_\varepsilon^\varphi(t)\|_\gamma}{s^{(1+\alpha)}}\, ds \right|^\gamma dt \leqslant C_9 C_\lambda < \infty, \quad \varepsilon > 0, \tag{66}$$

where $D_{-,t}^\alpha$ emphasizes that the Marchaud fractional derivative is taken with respect to $t$.

Combining of (65) and (66) gives (64). □

**Lemma 4.20.** *Assume that* (49) *holds. Then for each* $\delta > 0$ *and* $T > 0$, *there exists a compact set* $K = K(\delta, T) \Subset H_{\mathrm{loc}}^{-1}(\mathcal{O})$ *such that*

$$\inf_{\varepsilon > 0} P(\chi_{\varepsilon,2} \in K) > 1 - \delta.$$

**Proof.** The conclusion follows from the compact embedding of $H_{\mathrm{loc}}^{-1+\alpha}(\mathcal{O})$ to $H_{\mathrm{loc}}^{-1}(\mathcal{O})$ and from the results in Lemma 4.19. □

*4.4.2. Identifying $\nu_0(t, x; du)$ as a function*

To simplify notation, for any function $f = f(u)$, we denote

$$\overline{f} = \overline{f}(t,x) = \int\limits_{u \in R} f(u) \nu_0(t,x,du)$$

whenever the integral makes sense. In particular, $\bar{u}(t,x) = \int_{u\in R} u \nu_0(t,x;du)$. We also write $X \overset{\mathrm{D}}{=} Y$ for random variables $X, Y$ having identical probability law/distribution.

Let $(\Phi_i, \Psi_i)$, $i = 1, 2$, be two choices of entropy–entropy flux pairs, where $\Phi_i$ has at most polynomial growth (therefore $\Psi_i$ will have at most polynomial growth as well).

**Lemma 4.21.** *For every deterministic* $\varphi \in C_c^\infty(\mathcal{O})$,

$$\langle \varphi, \overline{\Psi_1 \Phi_2} - \overline{\Phi_1 \Psi_2} \rangle \overset{\mathrm{D}}{=} \langle \varphi, \overline{\Psi_1} \cdot \overline{\Phi_2} - \overline{\Phi_1} \cdot \overline{\Psi_2} \rangle. \tag{67}$$

**Proof.** On the one hand, by Lemma 4.15 and the uniform in $\varepsilon$ moment estimates in (50), for each $\varphi \in C_c^\infty((0,T) \times R^d)$, the following convergence in probability result holds

$$\lim_{\varepsilon \to 0+} \int \varphi(t,x) \big( \Psi_1\big(u_\varepsilon(t,x)\big) \Phi_2\big(u_\varepsilon(t,x)\big) - \Phi_1\big(u_\varepsilon(t,x)\big) \Psi_2\big(u_\varepsilon(t,x)\big) \big)\, dx\, dt$$

$$= \lim_{\varepsilon \to 0+} \int\limits_{(t,x) \in [0,T] \times R} \varphi(t,x) \int\limits_{u \in R} \big( \Psi_1(u) \Phi_2(u) - \Phi_1(u) \Psi_2(u) \big) \nu_\varepsilon(t,x,du)\, dt\, dx$$

$$= \langle \varphi, \overline{\Psi_1 \Phi_2 - \Phi_1 \Psi_2} \rangle.$$

On the other hand, we can apply Theorem A.2 by choosing

$$G_\varepsilon(t,x) = \big(\Phi_1\big(u_\varepsilon(t,x)\big), \Psi_1\big(u_\varepsilon(t,x)\big)\big), \qquad H_\varepsilon(t,x) = \big(-\Psi_2\big(u_\varepsilon(t,x)\big), \Phi_2\big(u_\varepsilon(t,x)\big)\big).$$

In view of Lemmas 4.12 and 4.18, we have the following convergence in probability law result:

$$\lim_{\varepsilon\to 0+} \int \varphi(t,x)\big(\Psi_1\big(u_\varepsilon(t,x)\big)\Phi_2\big(u_\varepsilon(t,x)\big) - \Phi_1\big(u_\varepsilon(t,x)\big)\Psi_2\big(u_\varepsilon(t,x)\big)\big)\,dt\,dx$$
$$\stackrel{\mathrm{D}}{=} \langle \varphi, \overline{\Psi_1}\cdot\overline{\Phi_2} - \overline{\Phi_1}\cdot\overline{\Psi_2}\rangle.$$

We conclude the proof. □

We now finish the proof of the main result of this section.

**Proof of Lemma 4.17.** Let deterministic $0 \leqslant \varphi \in C_c^\infty(\mathcal{O})$. Take $\Phi_1(u) = u$, $\Psi_1(u) = F(u)$ and $\Phi_2(u) = F(u)$, $\Psi_2(u) = \int_0^u (F')^2(r)\,dr$. Eq. (67) gives

$$\big\langle \varphi, \overline{F^2} - \overline{u\Psi_2}\big\rangle \stackrel{\mathrm{D}}{=} \big\langle \varphi, (\overline{F})^2 - \bar{u}\cdot\overline{\Psi_2}\big\rangle. \tag{68}$$

On the other hand, for $t > 0$, $x \in R$, $\omega \in \Omega$ fixed, and $u \in R$, by Schwartz's inequality,

$$\big(F(u) - F\big(\bar{u}(t,x)\big)\big)^2 = \left(\int_{\bar{u}(t,x)}^{u} F'(v)\,dv\right)^2 \leqslant \big(u - \bar{u}(t,x)\big)\big(\Psi_2(u) - \Psi_2\big(\bar{u}(t,x)\big)\big). \tag{69}$$

Integrate the above as a function of $u$ against $\nu_0(t,x,du)$,

$$\overline{F^2} + \big(F(\bar{u})\big)^2 - 2\overline{F}F(\bar{u}) \leqslant \overline{u\Psi_2} - \bar{u}\cdot\overline{\Psi_2}. \tag{70}$$

Take expectation on both (68) and (70) and combine them,

$$\int \varphi(t,x)E\big[\overline{F} - F(\bar{u})\big]^2\,dt\,dx \leqslant 0.$$

By the arbitrariness of $\varphi$, the following holds almost surely:

$$\overline{F}\,dt\,dx = \left(\int_R F(u)\nu_0(t,x,du)\right)dt\,dx = \big(F\big(\bar{u}(t,x)\big)\big)\,dt\,dx = F(\bar{u})\,dt\,dx. \tag{71}$$

Therefore

$$\int \varphi(t,x)(dt\,dx)\int_{\omega\in\Omega}\int_{u\in R}\big(F(u) - F\big(\bar{u}(t,x,\omega)\big)\big)^2\nu_0(t,x,du;\omega)P(d\omega)$$
$$= \int \varphi(t,x)(dt\,dx)E\big[\overline{F^2} - \big(F(\bar{u})\big)^2\big] = \int \varphi(t,x)(dt\,dx)E[\overline{u\Psi_2} - \bar{u}\cdot\overline{\Psi_2}]$$

$$= \int \varphi(t,x)(dt\,dx) \int\limits_{\omega} \int\limits_{u} \big(\big(u-\bar{u}(t,x,\omega)\big)\big(\Psi_2(u)-\Psi_2\big(\bar{u}(t,x;\omega)\big)\big)\big)\nu_0(t,x,du;\omega)P(d\omega).$$

In the above, the second equality follows from (71) and the third one follows from (68). We conclude then, almost surely (69) holds as an equality for all $u$ on the support of random measure $\nu_0(t,x;du)$. On the other hand, from the property of Schwartz inequality, this cannot be true unless $F'$ is constant during $u,\bar{u}$. With the condition on $F$, the support of $u$ has to be on single point mass $\bar{u}$ almost surely. □

By (50) and Fatou's lemma

$$\sup_{0\leqslant t\leqslant T} E\bigg[\int\limits_{x}\int\limits_{u} |u|^p \nu_0(t;x,du)\,dx\bigg] < \infty, \quad p=2,4,\ldots.$$

If (61) holds, then

$$\sup_{0\leqslant t\leqslant T} E\bigg[\int\limits_{x} \big|\bar{u}(t,x)\big|^p\,dx\bigg] < \infty, \quad p=2,4,\ldots. \tag{72}$$

*4.5. Existence of stochastic entropic solution*

*4.5.1. Existence of measure-valued solution*

Let $F_\varepsilon(r)=(F_{\varepsilon,1},\ldots,F_{\varepsilon,d})$ be as defined in last section. Let convex $\Phi\in C^2(R)$ have at most polynomial growth. Define $\Psi_\varepsilon=(\Psi_{\varepsilon,1},\ldots,\Psi_{\varepsilon,d})$ with

$$\Psi_{\varepsilon,k}(r)=\int\limits_0^r \Phi'(s)(F_{\varepsilon,k})'(s)\,ds.$$

Then (37) can be written as (Lemma 4.11)

$$\begin{aligned}
\big\langle\Phi\varphi,\nu_\varepsilon(t)\big\rangle-\big\langle\Phi\varphi,\nu_\varepsilon(s)\big\rangle \leqslant & \int\limits_{(s,t]\times R^d\times R} (\Psi_\varepsilon\cdot\nabla\varphi)\nu_\varepsilon(r,x,du)\,dx\,dr \\
& + \frac{1}{2}\int\limits_{Z}\int\limits_{(s,t]\times R^d\times R} \big(\Phi''\sigma_\varepsilon^2\varphi\big)\nu_\varepsilon(r,x,du)\,dx\,dr\,\mu(dz) \\
& + \varepsilon\int\limits_{(s,t]\times R^d\times R} (\Phi\,\Delta\varphi)\nu_\varepsilon(r,x,du)\,dx\,dr \\
& + \int\limits_{(s,t]\times Z} \big\langle\Phi'\sigma_\varepsilon\varphi,\nu_\varepsilon(r)\big\rangle\,W(dr,dz),
\end{aligned} \tag{73}$$

where shorthand notation $\Phi\varphi=\Phi(u)\varphi(x)$ and so on are used.

The main result of this section is

**Lemma 4.22.** *Let $u(0), F, \sigma$ satisfy conditions in Theorem* 4.1. *Then the $\mathcal{F}_t$-adapted process $\nu_0(t)$ has trajectory in $C([0,\infty),\mathcal{M}_0)$, and it satisfies*

$$\begin{aligned}\langle \Phi\varphi, \nu_0(t)\rangle - \langle \Phi\varphi, \nu_0(s)\rangle \leqslant{} & \int_{(s,t]\times R^d\times R} (\Psi\cdot\nabla\varphi)\nu_0(r,x,du)\,dx\,dr \\ & + \frac{1}{2}\int_Z \int_{(s,t]\times R^d\times R} \big(\Phi''\sigma^2\varphi\big)\nu_0(r,x,du)\,dx\,dr\,\mu(dz) \\ & + \int_{(s,t]\times Z} \langle \Phi'\sigma\varphi, \nu_0(r)\rangle W(dr,dz). \end{aligned} \tag{74}$$

We show the proof in two steps. First, we establish the following.

**Lemma 4.23.** *Assume that* (61) *holds, then*

$$\lim_{t\to s+} E\big[r\big(\nu_0(t),\nu_0(s)\big)\big] = 0, \quad s\geqslant 0.$$

*In particular, this implies that $\nu_0(\cdot)\in C([0,\infty);\mathcal{M}_0)$.*

**Proof.** Let $y\in R^d$. Take $\Phi(u) = |u-u_\varepsilon(s,y)|^2$ and $\varphi(x) = J_\delta(x-y)$ and apply (37) (Lemma 4.11). We notice that

$$\Psi_{\varepsilon,k}(u) = 2\int_0^u F'_{\varepsilon,k}(r)\big(r-u_\varepsilon(s,y)\big)\,dr = 2\int_0^r F'_{\varepsilon,k}(r)r\,dr - 2u_\varepsilon(s,y)\big(F_{\varepsilon,k}(u)-F_{\varepsilon,k}(0)\big),$$

$k=1,2,\dots,d$, have at most polynomial growth, and that

$$\big|\partial_{x_k}J_\delta(x-y)\big| \leqslant c_1\delta^{-1}, \qquad \big|\Delta_x J_\delta(x-y)\big| \leqslant c_2\delta^{-2}.$$

For $\psi\geqslant 0$ and $\psi\in C_c^\infty(R^d)$, we denote

$$\mathcal{O} = \operatorname{int}\big(\operatorname{supp}(\psi)\big), \qquad \mathcal{O}^\delta = \big\{x\in R^d\colon \operatorname{dist}(x,\mathcal{O})<\delta\big\}.$$

Then

$$\begin{aligned} & E\bigg[\int_{y\in\mathcal{O}}\int_{x\in\mathcal{O}^\delta} \big|u_\varepsilon(t,x)-u_\varepsilon(s,y)\big|^2 J_\delta(x-y)\,dx\,\psi(y)\,dy\,\bigg|\mathcal{F}_s\bigg] \\ &\quad\leqslant \int_{y\in\mathcal{O}}\int_{x\in\mathcal{O}^\delta} \big|u_\varepsilon(s,x)-u_\varepsilon(s,y)\big|^2 J_\delta(x-y)\,dx\,\psi(y)\,dy \\ &\qquad + \frac{c_3}{\delta}\int_s^t E\bigg[\int_{y\in\mathcal{O}}\int_{x\in\mathcal{O}^\delta} \big(1+\big|u_\varepsilon(s,y)\big|^2+\big|u_\varepsilon(r,x)\big|^{p_1}\big)\,dx\,\psi(y)\,dy\,\bigg|\mathcal{F}_s\bigg]\,dr \end{aligned}$$

$$+\int_s^t E\left[\int_{y\in\mathcal{O}}\int_{x\in\mathcal{O}^\delta}\int_z \sigma^2\big(x,u_\varepsilon(r,x);z\big)J_\delta(x-y)\,dx\,\psi(y)\,dy\,\mu(dz)\Big|\mathcal{F}_s\right]dr$$

$$+\frac{\varepsilon c_4}{\delta^2}\int_s^t E\left[\int_{y\in\mathcal{O}}\int_{x\in\mathcal{O}^\delta}\big(1+\big|u_\varepsilon(r,x)\big|^{p_2}\big)\,dx\,\psi(y)\,dy\Big|\mathcal{F}_s\right]dr$$

for some $p_1, p_2 \geqslant 2$.

Let stochastic measure

$$\pi(dx,du;dy,dv)=\delta_{u_\varepsilon(s,y)}(dv)\big(J_\delta(x-y)\,dy\big)\delta_{u_\varepsilon(t,x)}(du)\,dx.$$

Then direct verification shows that $\pi\in\Pi^{\nu_\varepsilon(t),\nu_\varepsilon(s)}$ (see definition in (55)). Furthermore, for each $K>0$ fixed such that $\{x\colon |x|<K\}\subset\mathcal{O}$, and for all $\delta>0$,

$$\begin{aligned}
&\int_{(x,y)\in\mathcal{O}\times\mathcal{O}^\delta}\big|u_\varepsilon(t,x)-u_\varepsilon(s,y)\big|^2 J_\delta(x-y)\,dx\,dy\\
&\quad=\int_{(x,y)\in\mathcal{O}\times\mathcal{O}^\delta}|u-v|^2\pi(dx,du;dy,dv)\\
&\quad\geqslant q_K\big(\nu_\varepsilon(t),\nu_\varepsilon(s)\big)-\int_{\mathcal{O}\times\mathcal{O}^\delta}|x-y|^2 J_\delta(x-y)\,dx\,dy\\
&\quad\geqslant q_K\big(\nu_\varepsilon(t),\nu_\varepsilon(s)\big)-\delta^2 c_5.
\end{aligned}$$

In view of the convergence in probability result in Lemma 4.15,

$$\begin{aligned}
&E\big[q_K\big(\nu_0(t),\nu_0(s)\big)\big]\\
&\quad\leqslant c_5\delta^2+E\left[\int_y\int_x|u-v|^2 J_\delta(x-y)\nu_0(s;dx,du)\nu_0(s;dy,dv)\right]\\
&\qquad+\frac{c_6}{\delta}\int_s^t\sup_{\varepsilon>0}E\left[\int_{y\in\mathcal{O}}\int_{x\in\mathcal{O}^\delta}\big(1+\big|u_\varepsilon(s,y)\big|^2+\big|u_\varepsilon(r,x)\big|^{p_1}\big)\,dx\,dy\right]dr\\
&\qquad+c_7\int_s^t\sup_{\varepsilon>0}E\left[\int_{y\in\mathcal{O}}\int_{x\in\mathcal{O}^\delta}\int_z\sigma^2\big(x,u_\varepsilon(r,x);z\big)J_\delta(x-y)\,dx\,dy\,\mu(dz)\right]dr.
\end{aligned}$$

Noting (50) and (61),

$$\limsup_{t\to s+}E\big[q_K\big(\nu_0(t),\nu_0(s)\big)\big]\leqslant c_5\delta^2+E\left[\int_{y\in\mathcal{O}}\int_{x\in\mathcal{O}^\delta}\big|\bar u(s,x)-\bar u(s,y)\big|^2 J_\delta(x-y)\,dx\,dy\right].$$

By the arbitrariness of $\delta > 0$ and integrability estimates (72) on $\bar{u}(s)$ we arrive at

$$\lim_{t\to s+} E\big[q_K\big(\nu_0(t), \nu_0(s)\big)\big] = 0, \quad s \geqslant 0,\ K = 1, 2, \ldots.$$

Hence conclude the lemma. □

Next, we show that

**Lemma 4.24.**

$$\lim_{\varepsilon\to 0+} E\bigg[\bigg|\int_{(s,t]\times Z} \big\langle \Phi'\sigma_\varepsilon\varphi, \nu_\varepsilon(r)\big\rangle W(dr, dz) - \int_{(s,t]\times Z} \big\langle \Phi'\sigma\varphi, \nu_0(r)\big\rangle W(dr, dz)\bigg|^2\bigg] = 0.$$

**Proof.**

$$E\bigg[\bigg|\int_{(s,t]\times Z} \big\langle \Phi'\sigma_\varepsilon\varphi, \nu_\varepsilon(r)\big\rangle W(dr, dz) - \int_{(s,t]\times Z} \big\langle \Phi'\sigma\varphi, \nu_0(r)\big\rangle W(dr, dz)\bigg|^2\bigg]$$
$$\leqslant \int_{(s,t]} E\bigg[\int_Z \big(\big\langle \Phi'\sigma_\varepsilon\varphi, \nu_\varepsilon(r)\big\rangle - \big\langle \Phi'\sigma\varphi, \nu_0(r)\big\rangle\big)^2 \mu(dz)\bigg] dr.$$

Therefore, the result follows from Lemma 4.15 and from (50). □

Assuming (61) holds, then (74) becomes (10). Combined with estimates (72) $\bar{u}(t, x)$ is a stochastic entropic solution.

### *4.6. Existence of stochastic strong entropic solution*

To be consistent with the notations in the definition of strong entropic solution (as well as the uniqueness proof), we write $v_\varepsilon = u_\varepsilon$ where $u_\varepsilon$ is constructed as in Lemma 4.10, and $v = v(t, y) = \bar{u}(t, y) = \int u\nu_0(t, y; dv)$. We assume all the conditions at the beginning of Section 4.2.3 regarding $v_\varepsilon(0)$, $F_\varepsilon$, and $\sigma_\varepsilon$. We also assume that $\sigma$ satisfies (13). We assume that (61), a conclusion of Lemma 4.17, holds and consider general dimensions $d = 1, 2, \ldots$. (72) translates into

$$\sup_{0\leqslant t\leqslant T} E\big[\big\|v(t)\big\|_p^p\big] < \infty, \quad p = 2, 4, \ldots. \tag{75}$$

Let $\tilde{u}(t) = \tilde{u}(t, x)$ be an arbitrary $\mathcal{F}_t$-adapted $L^p(R^d)$-valued process with

$$\sup_{0\leqslant t\leqslant T} E\big[\big\|\tilde{u}(t)\big\|_p^p\big] < \infty, \quad \forall T > 0,\ p = 2, 4, \ldots. \tag{76}$$

Let $\beta$ be of the form as in (17). We prove that (12) holds in the following lemma.

**Lemma 4.25.** *For each $T > 0$, there exists a deterministic function $\{A(s,t)\colon\ 0 \leqslant s \leqslant t < \infty\}$ such that*

$$E\bigg[\int\limits_{(s,t]\times Z}\int\limits_{y}\int\limits_{x} \beta'\big(\tilde{u}(r,x)-v(t,y)\big)\sigma\big(x,\tilde{u}(r,x);z\big)\varphi(x,y)\,dx\,dy\,W(dr,dz)\bigg]$$

$$\leqslant E\bigg[-\int\limits_{R^d\times R^d}\bigg(\int\limits_{Z\times(s,t]} \beta''\big(\tilde{u}(r,x)-v(r,y)\big)\sigma\big(y,v(r,y);z\big)$$

$$\times\,\sigma\big(x,\tilde{u}(r,x);z\big)\,\mu(dz)\,dr\bigg)\varphi(x,y)\,dx\,dy\bigg]+A(s,t)$$

*with the property that, for each sequence of partitions $0 \leqslant t_1 \leqslant \cdots \leqslant t_m \leqslant T$,*

$$\lim_{\max_i|t_{i+1}-t_i|\to 0+}\sum_i A(t_i,t_{i+1})=0.$$

The proof consists of several involved estimates. We present them in further subsections.

*4.6.1. A special martingale $N$, and its estimates*

For each $\alpha \in C^2$ such that $\alpha, \alpha', \alpha'' \in C_{\mathrm{b}}(R)$, and each $\varphi \in C_{\mathrm{c}}^{\infty}(R^d \times R^d)$, we denote

$$N(\alpha,\varphi)(s,t;y,v)=\int\limits_{(s,t]\times Z}\int\limits_{x}\alpha\big(\tilde{u}(r,x)-v\big)\sigma\big(x,\tilde{u}(r,x);z\big)\varphi(x,y)\,dx\,W(dr,dz), \tag{77}$$

where $0 \leqslant s \leqslant t$, $(y,v) \in R^d \times R$, and the integral is defined in Ito's sense. For each $s$ fixed, the above is a martingale in $t \geqslant s$. Next, we derive some useful properties and a priori estimates regarding $N$.

We note that $N(\alpha,\varphi)(s,t;y,v)=0$ whenever $|y| > C$ for some large $C = C_\varphi$ depending only on the support of $\varphi$.

**Lemma 4.26.** *The following identities hold almost surely for each $(y,v) \in R^d \times R$ fixed*:

$$\partial_v N(\alpha,\varphi)(s,t;y,v)=N(-\alpha',\varphi)(s,t;y,v),$$

$$\partial_{y_i} N(\alpha,\varphi)(s,t;y,v)=N(\alpha,\partial_{y_i}\varphi)(s,t;y,v).$$

**Proof.** The proof of Theorem 7.6 of Kunita [11, p. 180] can be modified to show this. □

**Lemma 4.27.** *Suppose that $\alpha \in C_{\mathrm{c}}(R)$. Then for each $T > 0$, $p > 5$, there exist $a > 0$, $C_2 > 0$, for any $\delta > 0$,*

$$E\Big[\sup_{s,t\in[0,T],\,|t-s|<\delta}\big\|N(\alpha,\varphi)(s,t;\cdot,\cdot)\big\|_p^p\Big] < C_2\delta^a.$$

**Proof.** First,

$$
\begin{aligned}
&E\big[\big\|N(\alpha,\varphi)(s,t;\cdot,\cdot)\big\|_p^p\big] \\
&= \int_{v\in R}\int_{|y|<c_\varphi} E\bigg[\bigg|\int_{(s,t]\times Z}\int_x \alpha\big(\tilde u(r,x)-v\big)\sigma\big(x,\tilde u(r,x);z\big)\varphi(x,y)\,dx\,W(dr,dz)\bigg|^p\bigg]\,dy\,dv \\
&\leqslant c_1\int_v\int_y E\Bigg[\bigg(\int_s^t\int_z\bigg|\int_x \alpha\big(\tilde u(r,x)-v\big)\sigma\big(x,\tilde u(r,x);z\big)\varphi(x,y)\,dx\bigg|^2\mu(dz)\,dr\bigg)^{p/2}\Bigg]\,dy\,dv \\
&\leqslant c_2|t-s|^{(p/2-1)}\int_v\int_{|y|<c_\varphi} E\bigg[\int_0^T\int_{|x|<c_\varphi}\big|\alpha\big(\tilde u(r,x)-v\big)\big|^p\big(1+\big|\tilde u(r,x)\big|^p\big)\,dx\,dr\bigg]\,dy\,dv \\
&\leqslant c_3|t-s|^{(p/2-1)}E\bigg[\int_0^T\int_{|y|<c_\varphi}\int_{|x|<c_\varphi}\bigg(\int_{|v|\leqslant\tilde u(r,x)|+c_\alpha}\|\alpha\|_\infty^p\big(1+\big|\tilde u(r,x)\big|^p\big)\,dv\bigg)\,dx\,dy\,dr\bigg] \\
&\leqslant c_4|t-s|^{(p/2-1)}E\bigg[\int_0^T\bigg(1+\int_x\big|\tilde u(r,x)\big|^{p+1}\,dx\bigg)\,dr\bigg]\leqslant c_5|t-s|^{(p/2-1)}.
\end{aligned}
$$

In the above, the first inequality follows from martingale inequality, the second one follows from Jensen's inequality and (13), the third inequality follows from the compact support assumption on $\alpha$, and the last inequality from (76). The above implies

$$
E\bigg[\int_0^T\int_0^T\bigg(\frac{\|N(\alpha,\varphi)(s,t;\cdot,\cdot)\|_p}{|t-s|^{\frac12-\frac1{2p}}}\bigg)^p\,ds\,dt\bigg]\leqslant c_6<\infty. \tag{78}
$$

Note $N(\alpha,\varphi)(s,t;\cdot,\cdot)=N(\alpha,\varphi)(0,t;\cdot,\cdot)-N(\alpha,\varphi)(0,s;\cdot,\cdot)$. By a normed linear space version of Garsia inequality [17, Exercise 2.4.1, p. 60], when $p>8$,

$$
\sup_{s,t\in[0,T];|s-t|<\delta}\big\|N(\alpha,\varphi)(s,t;\cdot,\cdot)\big\|_p\leqslant c_7\bigg(\int_0^T\int_0^T\frac{\|N(\alpha,\varphi;\cdot,\cdot)\|_p}{|t-s|^{\frac12-\frac1{2p}}}\,ds\,dt\bigg)\delta^{\frac{p-5}{2p}}.
$$

Taking the expectation, the conclusion follows from (78). □

**Lemma 4.28.** *Suppose that $\alpha,\alpha'\in C_c(R)$, and $p>8$, then*

$$
E\Big[\sup_{0\leqslant s,t\leqslant T,|t-s|<\delta}\big\|N(\alpha,\varphi)(s,t;\cdot,\cdot)\big\|_\infty^p\Big]\leqslant C_{2,\alpha,\varphi}\delta^a
$$

*for some $a>0$.*

**Proof.** By a Sobolev inequality (e.g. [7, (23), p. 268]), there exist deterministic constants $c_{1,p}, c_{2,p} > 0$

$$\begin{aligned}
&\bigl\|N(\alpha,\varphi)(s,t;\cdot,\cdot)\bigr\|_\infty \\
&\quad\leqslant c_{1,p}\bigl\|N(\alpha,\varphi)(s,t;\cdot,\cdot)\bigr\|_{W^{1,p}(R^d\times R)} \\
&\quad\leqslant c_{2,p}\Biggl(\bigl\|N(\alpha,\varphi)(s,t;\cdot,\cdot)\bigr\|_p+\bigl\|N(\alpha',\varphi)(s,t;\cdot,\cdot)\bigr\|_p+\sum_{k=1}^d\bigl\|N(\alpha,\partial_{y_k}\varphi)(s,t;\cdot,\cdot)\bigr\|_p\Biggr).
\end{aligned}$$

In view of the above, the conclusions follow from Lemma 4.27. □

*4.6.2. $\Gamma_{i,\varepsilon}$ and their estimates*

Let $\beta \in C^\infty(R)$ satisfy the conditions of Lemma 2.9. We will need to estimate the following random variables. Let $k = 1, 2, \dots, d$ be fixed and

$$\Gamma_{1,\varepsilon}(s,t)=\int_y\int_s^t \bigl(N\bigl(\beta'',\varphi\bigr)\bigl(s,r;y,v_\varepsilon(r,y)\bigr)\bigr)\bigl(-\partial_{y_k}F_\varepsilon\bigl(v_\varepsilon(r,y)\bigr)\bigr)\,dr\,dy,$$

$$\Gamma_{2,\varepsilon}(s,t)=\int_s^t\int_y N(\beta'',\varphi)\bigl(s,r;y,v_\varepsilon(r,y)\bigr)\varepsilon\Delta_{yy}v_\varepsilon(r,y)\,dy\,dr.$$

**Lemma 4.29.** *Let*

$$A(s,t)=A_1(s,t)+A_2(s,t),$$

*where*

$$A_k(s,t)=\liminf_{\varepsilon>0}E\bigl[\bigl|\Gamma_{k,\varepsilon}(s,t)\bigr|\bigr],\quad k=1,2.$$

*Then for every sequence of partitions of* $[0,T]$, $T>0$, $0\leqslant t_1\leqslant\cdots\leqslant t_m\leqslant T$,

$$\lim_{\max_i|t_{i+1}-t_i|\to 0+}\sum_i A(t_i,t_{i+1})=0.$$

We divide the proof into several lemmas. First, we estimate $\Gamma_{1,\varepsilon}(s,t)$. Let

$$G_\varepsilon(u,v)=\int_0^v \beta''(u-r)F'_{\varepsilon,k}(r)\,dr,\quad u,v\in R.$$

$\beta''$ has compact support, implying that the above integration can be restricted to $0\leqslant r\leqslant |u|$. Therefore,

$$\sup_{\varepsilon>0}\bigl|G_\varepsilon(u,v)\bigr|\leqslant C_\beta\bigl(1+|u|^p\bigr),\quad \forall u,v\in R,$$

for some integer $p \in \{1, 2, \ldots\}$. For each $\varphi \in C_c^\infty(R^d \times R^d)$, let

$$X_\varepsilon(\varphi)(s,t;y,v) = \int\limits_{(s,t]\times Z} \int\limits_x G_{\varepsilon,k}\big(\tilde u(r,x), v\big)\sigma\big(x, \tilde u(r,x); z\big)\varphi(x,y)\,dx\,W(dr,dz).$$

As in the proof of Lemma 4.26,

$$\partial_v X_\varepsilon(\varphi)(s,t;y,v) = \int\limits_{(s,t]\times Z} \int\limits_x \partial_v G_\varepsilon\big(\tilde u(r,x), v\big)\sigma\big(x, \tilde u(r,x); z\big)\varphi(x,y)\,dx\,W(dr,dz),$$

$$\partial_{y_k} X_\varepsilon(\varphi)(s,t;y,v) = X_\varepsilon(\partial_{y_k}\varphi)(s,t;y,v).$$

As in the proof of Lemmas 4.27 and 4.28, we have

**Lemma 4.30.** *For each $p > 2$, there exist a constant $C = C(\beta, \varphi, p)$ independent of $\varepsilon$, and a constant $a > 0$ such that*

$$E\big[\big\|X_\varepsilon(\varphi)(s,t;\cdot,\cdot)\big\|_\infty^p\big] \leqslant C|t-s|^a.$$

By a formal application of integration by parts (note that we are handling integral with anticipating integrand), we have (79). We justify this rigorously next.

**Lemma 4.31.** *The following representation holds almost surely*:

$$\Gamma_{1,\varepsilon}(s,t) = \int\limits_s^t \int\limits_y X_\varepsilon(\partial_{y_k}\varphi)\big(s,r;y,v_\varepsilon(r,y)\big)\,dy\,dr. \tag{79}$$

**Proof.** Let $J_\delta$ be a one-dimensional mollifier as defined before (smooth and have compact support). First, through integration by parts,

$$\begin{aligned}
&\int\limits_y \bigg(\int\limits_v N(\beta'', \varphi)(s,r;y,v)F_\varepsilon'(v)J_\delta\big(v - v_\varepsilon(r,y)\big)\,dv\bigg)\partial_{y_k} v_\varepsilon(r,y)\,dy \\
&\quad = \int\limits_y \int\limits_v \partial_v X_\varepsilon(\varphi)(s,r;y,v)J_\delta\big(v - v_\varepsilon(r,y)\big)\,dv\,\partial_{y_k} v(r,y)\,dy \\
&\quad = \int\limits_y \int\limits_v X_\varepsilon(\varphi)(s,r;y,v)J_\delta'\big(v - v_\varepsilon(r,y)\big)\,dv\,\partial_{y_k} v(r,y)\,dy \\
&\quad = -\int\limits_v \int\limits_y X_\varepsilon(\varphi)(s,r;y,v)\partial_{y_k} J_\delta\big(v - v_\varepsilon(r,y)\big)\,dy\,dv \\
&\quad = \int\limits_v \int\limits_y \partial_{y_k} X_\varepsilon(\varphi)(s,r;y,v)J_\delta\big(v - v_\varepsilon(r,y)\big)\,dy\,dv
\end{aligned}$$

$$= \int_y \int_v X_\varepsilon (\partial_{y_k} \varphi)(s, r; y, v) J_\delta \big(v - v_\varepsilon(r, y)\big)\, dv\, dy.$$

Sending $\delta \to 0+$ and noting the estimates in Lemmas 4.28, 4.30, we arrive at the conclusion. □

Therefore, by Lemma 4.30,

$$E\big[\big|\Gamma_{1,\varepsilon}(s,t)\big|\big] \leqslant \int_s^t \int_{|y| \leqslant c_\varphi} C_1 |r-s|^a\, dy\, dr \leqslant C_2 |t-s|^{1+a} \equiv A_1(s,t)$$

satisfies that for all partitions of $[0, T]$, $\lim_{\max_i |t_{i+1}-t_i| \to 0+} \sum_i |A_1(t_i, t_{i+1})| = 0$.

Now we estimate $\Gamma_{2,\varepsilon}$. First, we note that (36) holds even if $\Phi$ is not convex. In particular,

$$\begin{aligned} &\beta''\big(\tilde{u}(\bar{r}, x) - v_\varepsilon(t, y)\big) \Delta_{yy} v_\varepsilon(r, y) \\ &\quad = \Delta_{yy} \beta'\big(\tilde{u}(\bar{r}, x) - v_\varepsilon(r, y)\big) - \beta'''\big(\tilde{u}(\bar{r}, x) - v_\varepsilon(r, y)\big) \big|\nabla_y v_\varepsilon(r, y)\big|^2. \end{aligned}$$

Similar to the last lemma,

**Lemma 4.32.**

$$\begin{aligned} \Gamma_{2,\varepsilon}(s,t) &= \int_s^t \int_{|y|<c_\varphi} N(\beta'', \varphi)\big(s, r; y, v_\varepsilon(r, y)\big) \varepsilon \Delta_{yy} v_\varepsilon(r, y)\, dy\, dr \\ &= \varepsilon \int_s^t \int_{|y|<c_\varphi} N(\beta'', \Delta_{yy}\varphi)\big(s, r; y, v_\varepsilon(r, y)\big)\, dy\, dr \\ &\quad - \int_s^t \int_{|y|<c_\varphi} N(\beta''', \varphi)\big(s, r; y, v_\varepsilon(r, y)\big) \varepsilon \big|\nabla_y v_\varepsilon(r, y)\big|^2\, dy\, dr. \end{aligned}$$

The first terms can be handled using Lemma 4.28. Estimating the second term in the $\Gamma_{2,\varepsilon}$ is more involved. We formulate details in the following

**Lemma 4.33.** *There exists a deterministic function $A = A(s,t)$, $0 \leqslant s \leqslant t$, such that*

$$\limsup_{\varepsilon \to 0+} E\left[\int_s^t \int_y \big|N(\beta''', \varphi)\big(s, r; y, v_\varepsilon(r, y)\big)\big| \varepsilon \big|\nabla_y v_\varepsilon(r, y)\big|^2\, dy\, dr\right] \leqslant A(s,t);$$

*and for each sequence of partitions $0 \leqslant t_1 \leqslant \cdots \leqslant t_m \leqslant T$,*

$$\lim_{\max_i |t_{i+1}-t_i| \to 0+} \sum_i A(t_i, t_{i+1}) = 0.$$

**Proof.** We define

$$\Lambda_\varepsilon(t) = \int_0^t \varepsilon \left\| \nabla_y v_\varepsilon(r) \right\|_2^2 dr, \qquad Y_\delta = \sup_{s,t \in [0,T];\ |t-s|<\delta} \left\| N(\beta''', \varphi)(s,t;\cdot,\cdot) \right\|_\infty,$$

and let

$$A(s,t) = \limsup_{\varepsilon \to 0+} E\left[ Y_{|t-s|} \big( \Lambda_\varepsilon(t) - \Lambda_\varepsilon(s) \big) \right].$$

By Lemma 4.13,

$$\sup_{\varepsilon>0} E\left[ \left| \Lambda_\varepsilon(T) \right|^p \right] < \infty, \quad p = 1, 2, \ldots,\ T > 0. \tag{80}$$

Combined with Lemma 4.28, $A(s,t)$ is finite.

Let $T > 0$ and partition $0 = t_1 < \cdots < t_m = T$ with $\delta_m = \max_i |t_{i+1} - t_i|$. It follows then

$$\begin{aligned}
\sum_{i=1}^m A(t_i, t_{i+1}) &\leqslant \limsup_{\varepsilon \to 0+} \sum_{i=1}^m E\left[ Y_{\delta_m} \big( \Lambda_\varepsilon(t_{i+1}) - \Lambda_\varepsilon(t_i) \big) \right] = \limsup_{\varepsilon \to 0+} E\left[ Y_{\delta_m} \Lambda_\varepsilon(T) \right] \\
&\leqslant \big( E\left[ |Y_{\delta_m}|^p \right] \big)^{1/p} \Big( \sup_{\varepsilon>0} E\left[ \left| \Lambda_\varepsilon(T) \right|^q \right] \Big)^{1/q} \\
&\leqslant C \delta_m^a,
\end{aligned}$$

for some $a > 0$, $p^{-1} + q^{-1} = 1$ with $p > 8$. In the above, we invoked (80) and Lemma 4.28 for the last inequality.

The conclusion follows. $\square$

We conclude that Lemma 4.29 holds.

*4.6.3. Proof of Lemma 4.25*

With the above estimates, we are ready to prove the main result of this section.

**Proof.** Let $J$, $J_\delta$ be mollifiers defined as in (16) in the special case of one dimension. Recall notation (77), we first let

$$\begin{aligned}
Z_{\varepsilon,\delta}(t) &= \int_{|y|<c_\varphi} \int_v \big( N(\beta', \varphi)\big(s,t;y,v - v_\varepsilon(t,y)\big)\big) J_\delta(v)\, dv\, dy \\
&= \int_{|y|<c_\varphi} \int_v \big( N(\beta', \varphi)(s,t;y,v) \big) J_\delta\big(v - v_\varepsilon(t,y)\big)\, dv\, dy.
\end{aligned} \tag{81}$$

Then $Z_{\varepsilon,\delta}$ is a semi-martingale in $t \geqslant s$. We recall that $N(\beta', \varphi)(s,t;y,v) = 0$ whenever $|y| > c_\varphi$ for some large $c = c_\varphi$ depending only on the support of $\varphi$. Therefore, the integration of $y$ above can be restricted to a bounded set $|y| < c_\varphi$.

The following approximation result holds:

$$\lim_{\delta\to 0+} E\big[Z_{\varepsilon,\delta}(t)\big] = E\Big[\lim_{\delta\to 0+} Z_{\varepsilon,\delta}(t)\Big] = E\bigg[\int_y \big(N(\beta',\varphi)\big(s,t;y,v_\varepsilon(t,y)\big)\big)\,dy\bigg].$$

In the above, we interchanged the order of expectation and $\lim_{\delta\to 0+}$. This follows from dominated convergence theorem, which is justified by observation

$$\big|Z_{\varepsilon,\delta}(t)\big| \leqslant \tilde{C}_\varphi \big\|N(\beta',\varphi)(s,t;\cdot,\cdot)\big\|_\infty$$

and by estimate in Lemma 4.28 and by assumption in (76). Similarly,

$$\lim_{\varepsilon\to 0+}\lim_{\delta\to 0+} E\big[Z_{\varepsilon,\delta}(t)\big] = E\bigg[\int_y N(\beta',\varphi)\big(s,t;y,v(t,y)\big)\,dy\bigg].$$

On the other hand, by Ito's formula (i.e. $d(XY) = X\,dY + Y\,dX + d[X,Y]$ for two semimartingales $X, Y$) and by integration by parts applied to (81),

$$\begin{aligned}
Z_{\varepsilon,\delta}(t) = {} & \int_y\int_v\int_{r=s}^{t} J_\delta\big(v - v_\varepsilon(r,y)\big) d_r N(\beta',\varphi)(s,r;y,v)\,dy\,dv \\
& + \int_y\int_s^t\bigg(\int_v N(\beta'',\varphi)(s,r;y,v) J_\delta\big(v - v_\varepsilon(r,y)\big)\,dv\bigg) \\
& \times\bigg(-\operatorname{div}_y F_\varepsilon\big(v_\varepsilon(r,y)\big)\,dr + \varepsilon\Delta_{yy} v_\varepsilon(r,y)\,dr + \int_z \sigma_\varepsilon\big(y, v_\varepsilon(r,y); z\big)\,W(dr,dz)\bigg)\,dy \\
& + \int_y\int_{(s,t]\times Z}\bigg(\int_v N(\beta''',\varphi)(s,r;y,v)\frac{1}{2} J_\delta\big(v - v_\varepsilon(r,y)\big)\,dv\bigg) \\
& \times \sigma_\varepsilon^2\big(y, v_\varepsilon(r,y); z\big)\,\mu(dz)\,dr\,dy \\
& - \int_x\int_y\int_{(s,t]\times Z}\bigg(\int_v \beta''\big(\tilde{u}(r,x) - v\big) J_\delta\big(v - v_\varepsilon(r,y)\big)\,dv\bigg)\sigma\big(x,\tilde{u}(r,x);z\big)\varphi(x,y) \\
& \times \sigma_\varepsilon\big(y, v_\varepsilon(r,y); z\big)\,\mu(dz)\,dr\,dx\,dy.
\end{aligned}$$

Therefore,

$$E\big[Z_{\varepsilon,\delta}(t)\big] = I_{\varepsilon,\delta} + II_{\varepsilon,\delta} + III_{\varepsilon,\delta} + IV_{\varepsilon,\delta},$$

where

$$I_{\varepsilon,\delta} = E\left[\int_y \int_s^t \left(\left\{N(\beta'', \varphi)(s,r;y,\cdot) * J_\delta(\cdot)\right\}\left(v_\varepsilon(r,y)\right)\right)\left(-\operatorname{div}_y F_\varepsilon\left(v_\varepsilon(r,y)\right)\right) dr\, dy\right],$$

$$II_{\varepsilon,\delta} = E\left[\int_y \int_s^t \left(\left\{N(\beta'', \varphi)(s,r;y,\cdot) * J_\delta(\cdot)\right\}\left(v_\varepsilon(r,y)\right)\right)\left(\varepsilon \Delta_{yy} v_\varepsilon(r,y)\right) dr\, dy\right],$$

$$III_{\varepsilon,\delta} = \frac{1}{2} E\left[\int_y \int_{(s,t]\times Z} \left(\left\{N(\beta''', \varphi)(s,r;y,\cdot) * J_\delta(\cdot)\right\}\left(v_\varepsilon(r,y)\right)\right)\sigma_\varepsilon^2\left(y, v_\varepsilon(r,y); z\right) \mu(dz)\, dr\, dy\right],$$

$$IV_{\varepsilon,\delta} = -E\left[\int_x \int_y \int_{(s,t]\times Z} \left(\int_v \beta''\left(\tilde{u}(r,x) - v\right) J_\delta\left(v - v_\varepsilon(r,y)\right) dv\right) \times \sigma\left(x, \tilde{u}(r,x); z\right)\sigma_\varepsilon\left(y, v_\varepsilon(r,y); z\right) \mu(dz)\, \varphi(x,y)\, dx\, dy\, dr\right].$$

Since $\|\beta''(\cdot)\|_\infty < \infty$, by dominated convergence theorem,

$$\lim_{\delta\to 0+} IV_{\varepsilon,\delta} = -E\left[\int_x \int_y \int_{(s,t]\times Z} \beta''\left(\tilde{u}(r,x) - v_\varepsilon(r,y)\right)\sigma\left(x, \tilde{u}(r,x); z\right) \times \sigma_\varepsilon\left(y, v_\varepsilon(r,y); z\right) \mu(dz)\, \varphi(x,y)\, dx\, dy\, dr\right].$$

Furthermore, noting (57), by the estimates (50), (76), a uniform (in $\varepsilon$) integrability arguments gives

$$\lim_{\varepsilon\to 0+} \lim_{\delta\to 0+} IV_{\varepsilon,\delta} = -E\left[\int_x \int_y \int_{(s,t]\times Z} \beta''\left(\tilde{u}(r,x) - v(r,y)\right)\sigma\left(x, \tilde{u}(r,x); z\right) \times \sigma\left(y, v(r,y); z\right) \mu(dz)\, \varphi(x,y)\, dx\, dy\, dr\right].$$

Regarding $III_{\varepsilon,\delta}$, by (49) and (50), we have

$$\liminf_{\varepsilon\to 0+} \liminf_{\delta\to 0+} III_{\varepsilon,\delta} \leqslant CE\left[\int_s^t \left\|N(\beta''', \varphi)(s,r;\cdot,\cdot)\right\|_\infty \left(1 + \int_{|y|\leqslant c_\varphi} \left|v(r,y)\right|^2 dy\right) dr\right] \equiv A_3(s,t).$$

In view of Lemma 4.28 and (75), for all partitions of $[0,T]$,

$$\lim_{\max_i |t_{i+1}-t_i|\to 0+} \sum_{i=1}^m A_3(t_i, t_{i+1}) = 0.$$

We now estimate $I_{\varepsilon,\delta}$. By Lemma 4.28 and the dominated convergence theorem,

$$\lim_{\delta\to 0+} I_{\varepsilon,\delta} = E\left[\int_y \int_s^t \big(N(\beta'',\varphi)\big(s,r;y,v_\varepsilon(r,y)\big)\big)\big(-\operatorname{div}_y F_\varepsilon\big(v_\varepsilon(r,y)\big)\big)\,dr\,dy\right].$$

By Lemma 4.29,

$$\liminf_{\varepsilon\to 0+} \lim_{\delta\to 0+} I_{\varepsilon,\delta} \leqslant A_1(s,t)$$

with $A_1$ satisfying the requirement of the lemma.

Similarly, we can handle the $II_{\varepsilon,\delta}$ term.

Take $A(s,t) = A_1(s,t) + A_2(s,t) + A_3(s,t)$, then the lemma follows. □

## Acknowledgments

We thank Yaozhong Hu for useful discussions, and Elton Hsu for a careful reading and suggestions.

## Appendix A. A probabilistic generalization of Div–curl lemma

In weak convergence/Young measure approach to construction of deterministic nonlinear PDEs (e.g. [3,6]), it is important to extract information on limit of some nonlinear functionals. The well-known div–curl lemma is such a useful device. Below, we generalize it to a stochastic setting.

Throughout this appendix, we assume bounded open set $\mathcal{O} \subset R^m$ has a smooth $C^\infty$ boundary $\partial\mathcal{O}$. For $F = (F_1,\ldots,F_m) \in L^2(\mathcal{O};R^m)$, we identify $\operatorname{curl} F = \nabla \times F$ with an $m\times m$-matrix-valued function with the $(i,j)$th component defined by

$$(\nabla \times F)_{ij} = \partial_{x_j} F_i - \partial_{x_i} F_j \in H^{-1}(\mathcal{O}).$$

Let $X_\varepsilon, X$ be Polish space $S$-valued random variables. By $\lim_{\varepsilon\to 0+} X_\varepsilon \stackrel{\mathrm{D}}{=} X$, we mean $X_\varepsilon$ converges in probability law/distribution to $X$. By tightness of $\{X_\varepsilon\colon \varepsilon > 0\}$, we mean the family of probability distributions (on $S$) of the random variables is tight.

**Lemma A.1.** *Let $\{(F_\varepsilon, G_\varepsilon)\colon \varepsilon > 0\}$ and $(\bar F, \bar G)$ be a sequence of $H^p(\mathcal{O};R^m) \times H^q(\mathcal{O};R^m)$-valued random variables, where $p = -q \in \{0,\pm1,\pm2,\ldots\}$.*

*Suppose the following conditions hold.*

(1) *$\{F_\varepsilon\colon \varepsilon > 0\}$ is stochastically bounded in $H^p(\mathcal{O};R^m)$. That is, for each $\delta > 0$, there exists a deterministic constant $C_\delta \in (0,\infty)$ such that*

$$\sup_{\varepsilon>0} P\big(\|F_\varepsilon\|_{H^p} > C_\delta\big) < \delta.$$

(2) *Let $(\cdot,\cdot)_{H^p}$ denote the inner product in $H^p(\mathcal{O}; R^m)$. For each finite collection of deterministic $\phi_1, \ldots, \phi_k \in H^p(\mathcal{O}; R^m)$, $k = 1, 2, \ldots$, sequence of $R^k \times H^q(\mathcal{O}; R^m)$-valued random variables converges*:

$$\lim_{\varepsilon\to 0+}\big((\phi_1, F_\varepsilon)_{H^p}, \ldots, (\phi_k, F_\varepsilon)_{H^p}, G_\varepsilon\big) \stackrel{\mathrm{D}}{=} \big((\phi_1, \bar{F})_{H^p}, \ldots, (\phi_k, \bar{F})_{H^p}, \bar{G}\big). \tag{A.1}$$

*Let $\langle\cdot,\cdot\rangle$ be the continuous bilinear pairing between $H^p(\mathcal{O}; R^m)$ and $H^q(\mathcal{O}; R^m)$, $p = -q$. Then*

$$\lim_{\varepsilon\to 0+}\langle F_\varepsilon, G_\varepsilon\rangle \stackrel{\mathrm{D}}{=} \langle \bar{F}, \bar{G}\rangle,$$

*where the above is joint convergence with that in* (A.1)*. That is,*

$$\lim_{\varepsilon\to 0+}\big((\phi_1, F_\varepsilon)_{H^p}, \ldots, (\phi_k, F_\varepsilon)_{H^p}; G_\varepsilon; \langle F_\varepsilon, G_\varepsilon\rangle\big) \stackrel{\mathrm{D}}{=} \big((\phi_1, \bar{F})_{H^p}, \ldots, (\phi_k, \bar{F})_{H^p}; \bar{G}; \langle \bar{F}, \bar{G}\rangle\big).$$

**Proof.** By Skorokhod representation theorem (e.g. [5, Theorem 3.18]), without loss of generality, we assume that all random variables are defined on the same probability space and all convergence in probability distribution is convergence almost surely.

First, for each $h, \delta > 0$, by condition (1) and the assumption that $\lim_{\varepsilon\to 0+} G_\varepsilon \stackrel{\mathrm{D}}{=} \bar{G}$,

$$\begin{aligned}\lim_{\varepsilon\to 0+} P\big(\big|\langle F_\varepsilon, G_\varepsilon - \bar{G}\rangle\big| > h\big) &\leqslant \lim_{\varepsilon\to 0+} P\big(\|F_\varepsilon\|_{H^p}\|G_\varepsilon - \bar{G}\|_{H^q} > h\big)\\ &\leqslant \lim_{\varepsilon\to 0+}\big(P\big(\|G_\varepsilon - \bar{G}\|_{H^q} > hC_\delta^{-1}\big) + P\big(\|F_\varepsilon\|_{H^p} > C_\delta\big)\big) < \delta.\end{aligned}$$

Therefore, to conclude the lemma, it is sufficient to prove that $\lim_{\varepsilon\to 0+}\langle F_\varepsilon, \bar{G}\rangle \stackrel{\mathrm{D}}{=} \langle \bar{F}, \bar{G}\rangle$.

Let $\{f_1, \ldots, f_k, \ldots\}$ and $\{g_1, \ldots, g_k, \ldots\}$ be a dual system of (deterministic) complete orthonormal bases for $H^p(\mathcal{O}; R^m)$ and $H^q(\mathcal{O}; R^m)$, respectively. That is,

$$\langle \bar{f}, \bar{g}\rangle = \sum_{k=1}^{\infty}(\bar{f}, f_k)_{H^p}(\bar{g}, g_k)_{H^q}, \quad \forall \bar{f} \in H^p,\ \bar{g} \in H^q.$$

For every $h, \delta > 0$, by condition (1),

$$\begin{aligned}&\lim_{N\to\infty}\sup_{\varepsilon>0} P\left(\left|\sum_{k=N+1}^{\infty}(F_\varepsilon, f_k)_{H^p}(\bar{G}, g_k)_{H^q}\right|^2 > h\right)\\ &\leqslant \lim_{N\to\infty}\sup_{\varepsilon>0} P\left(\|F_\varepsilon\|_{H^p}^2 \sum_{k=N+1}^{\infty}(\bar{G}, g_k)_{H^q}^2 > h\right)\\ &\leqslant \lim_{N\to\infty}\left(P\left(\sum_{k=N+1}^{\infty}(\bar{G}, g_k)_{H^q}^2 > hC_\delta^{-2}\right) + \sup_{\varepsilon>0} P\big(\|F_\varepsilon\|_{H^p} > C_\delta\big)\right) \leqslant \delta.\end{aligned}$$

By the above uniform in $\varepsilon$ estimate, and by condition (2),

$$\lim_{\varepsilon\to 0+}\langle F_\varepsilon,\bar G\rangle = \lim_{\varepsilon\to\infty}\lim_{N\to\infty}\left(\sum_{k=1}^{\infty}(F_\varepsilon,f_k)_{H^p}(\bar G,g_k)_{H^q} - \sum_{k=N+1}^{\infty}(F_\varepsilon,f_k)_{H^p}(\bar G,g_k)_{H^q}\right)$$
$$\stackrel{\mathrm{D}}{=} \lim_{N\to\infty}\lim_{\varepsilon\to 0+}\sum_{k=1}^{N}(F_\varepsilon,f_k)_{H^p}(\bar G,g_k)_{H^q}$$
$$\stackrel{\mathrm{D}}{=} \lim_{N\to\infty}\sum_{k=1}^{N}(\bar F,f_k)_{H^p}(\bar G,g_k)_{H^q} = \langle \bar F,\bar G\rangle. \qquad \square$$

**Theorem A.2** *(Div–curl). Let* $(G_\varepsilon,H_\varepsilon),\varepsilon>0$, $(\bar G,\bar H)$ *be a sequence of* $L^2(\mathcal{O};R^m)\times L^2(\mathcal{O};R^m)$*-valued random variables. Suppose the following holds*:

(1) $\{G_\varepsilon\colon \varepsilon>0\}$ *and* $\{H_\varepsilon\colon \varepsilon>0\}$ *are both stochastically bounded as* $L^2(\mathcal{O};R^m)$*-valued random variables. That is, for each* $\delta>0$*, there exists a deterministic constant* $C_\delta\in(0,\infty)$ *such that*

$$\sup_{\varepsilon>0}P\big(\|G_\varepsilon\|_{L^2}+\|H_\varepsilon\|_{L^2}>C_\delta\big)<\delta.$$

(2) *For each finite collection of deterministic* $\phi_1,\dots,\phi_k\in L^2(\mathcal{O};R^m)$*,*

$$\lim_{\varepsilon\to 0+}\big(\langle\phi_1,G_\varepsilon\rangle,\dots,\langle\phi_k,G_\varepsilon\rangle;\langle\phi_1,H_\varepsilon\rangle,\dots,\langle\phi_k,H_\varepsilon\rangle\big)$$
$$\stackrel{\mathrm{D}}{=}\big(\langle\phi_1,\bar G\rangle,\dots,\langle\phi_k,\bar G\rangle;\langle\phi_1,\bar H\rangle,\dots,\langle\phi_k,\bar H\rangle\big).$$

(3) *Both* $\{\nabla\cdot G_\varepsilon\colon \varepsilon>0\}$ *and* $\{\nabla\times H_\varepsilon\colon \varepsilon>0\}$ *are tight as sequences of* $H^{-1}(\mathcal{O})$*-valued random variables.*

*Then for each finite collection of deterministic* $\varphi_1,\dots,\varphi_k\in C_c^\infty(\mathcal{O})$*,*

$$\lim_{\varepsilon\to 0+}\big(\langle\varphi_1,G_\varepsilon\cdot H_\varepsilon\rangle,\dots,\langle\varphi_k,G_\varepsilon\cdot H_\varepsilon\rangle\big)\stackrel{\mathrm{D}}{=}\big(\langle\varphi_1,\bar G\cdot\bar H\rangle,\dots,\langle\varphi_k,\bar G\cdot\bar H\rangle\big),$$

*where the convergence is joint convergence in probability law/distribution with that in condition* (2)*.*

**Proof.** Let $H^2(\mathcal{O},R^m)$-valued random variables $h_\varepsilon$ be defined as (weak) solution to

$$-\Delta h_\varepsilon = H_\varepsilon,\quad x\in\mathcal{O},\qquad h_\varepsilon=0,\quad x\in\partial\mathcal{O}. \tag{A.2}$$

Condition (1) of the theorem implies that $\{h_\varepsilon\colon \varepsilon>0\}$ is stochastically bounded as $H^2(\mathcal{O};R^m)$-valued random variables. Since any bounded set in $H^2(\mathcal{O};R^m)$ is a compact set in $L^2(\mathcal{O};R^m)$, $\{h_\varepsilon\colon \varepsilon>0\}$ is a tight sequence as $L^2(\mathcal{O};R^m)$-valued random variable. Selecting subsequence if necessary, there exists a $L^2(\mathcal{O};R^m)$-valued random variable $h_0$ such that

$$\lim_{\varepsilon\to 0+}h_\varepsilon\stackrel{\mathrm{D}}{=}h_0.$$

Indeed, $h_0$ has to be $H^2(\mathcal{O}; R^m)$-valued. To see this, we invoke the Skorohod representation and assume $\lim_{\varepsilon\to 0+}\|h_\varepsilon - h_0\|_{L^2} = 0$ almost surely. Then by lower semicontinuity of $\|\cdot\|_{H^2}$,

$$\big\|h_0(\cdot,\omega)\big\|_{H^2} \leqslant \liminf_{\varepsilon\to 0+}\big\|h_\varepsilon(\cdot,\omega)\big\|_{H^2}.$$

Consequently for each $\delta > 0$, there exists $C_\delta > 0$,

$$\begin{aligned}P\big(\|h_0\|_{H^2} > C_\delta\big) &\leqslant P\Big(\liminf_{\varepsilon\to 0+}\|h_\varepsilon\|_{H^2} > C_\delta\Big) = P\bigg(\bigcup_{\kappa>0}\bigcap_{0<\varepsilon<\kappa}\big\{\|h_\varepsilon\|_{H^2} > C_\delta\big\}\bigg)\\ &= \lim_{\kappa\to 0+} P\bigg(\bigcap_{0<\varepsilon<\kappa}\big\{\|h_\varepsilon\|_{H^2} > C_\delta\big\}\bigg) \leqslant \limsup_{\kappa\to 0+} P\big(\|h_\kappa\|_{H^2} > C_\delta\big) < \delta.\end{aligned}$$

Let

$$f_\varepsilon = -\nabla\cdot h_\varepsilon, \qquad N_\varepsilon = H_\varepsilon - \nabla f_\varepsilon.$$

Then $\{f_\varepsilon\colon \varepsilon > 0\}$ is stochastically bounded as $H^1(\mathcal{O})$-valued random variables; and is a tight sequence as $L^2(\mathcal{O})$-valued random variables. Selecting a subsequence if necessary, we have $f_0$, a $L^2(\mathcal{O}; R^m)$-valued random variable, such that, as $L^2(\mathcal{O}; R^m)$-valued random variables

$$\lim_{\varepsilon\to 0+} f_\varepsilon \overset{\mathrm{D}}{=} f_0.$$

Furthermore, by $H^1(\mathcal{O})$ stochastic boundedness of $\{f_\varepsilon\colon \varepsilon > 0\}$, $f_0$ is also a $H^1(\mathcal{O})$-valued random variable.

Since for each $1 \leqslant i \leqslant m$,

$$\begin{aligned}N_\varepsilon^i = H_\varepsilon^i - \partial_{x_i} f_\varepsilon &= \sum_{j=1}^m -\partial_{x_j}^2 h_\varepsilon^i + \partial_{x_i}\partial_{x_j} h_\varepsilon^j\\ &= \sum_{j=1}^m \partial_{x_j}\big(\partial_{x_i} h_\varepsilon^j - \partial_{x_j} h_\varepsilon^i\big) = \sum_{j=1}^m \partial_{x_j}\big((\nabla\times h_\varepsilon)_{ij}\big),\end{aligned} \tag{A.3}$$

and since by condition (3) of the theorem, $\{\nabla\times h_\varepsilon\colon \varepsilon > 0\}$ is tight as $H^1(\mathcal{O}; R^m)$-valued random variables, it follows that $\{N_\varepsilon\colon \varepsilon > 0\}$ is tight as $L^2(\mathcal{O}; R^m)$-valued random variables. Selecting subsequence if necessary, there exists $L^2(\mathcal{O}, R^m)$-valued random variable $N_0$ such that, as $L^2(\mathcal{O}; R^m)$-valued random variables,

$$\lim_{\varepsilon\to 0+} N_\varepsilon \overset{\mathrm{D}}{=} N_0.$$

To summarize, we can find subsequence so that

$$\lim_{\varepsilon\to 0+}(h_\varepsilon, f_\varepsilon, N_\varepsilon) \overset{\mathrm{D}}{=} (h_0, f_0, N_0). \tag{A.4}$$

Note that such convergence holds in the sense that the triplet $(h_\varepsilon, f_\varepsilon, N_\varepsilon)$ are viewed as metric space $E$-valued random variables, where

$$E = L^2(\mathcal{O}; R^m) \times L^2(\mathcal{O}) \times L^2(\mathcal{O}; R^m)$$

with corresponding norm topology. Note that $(-\Delta): L^2(\mathcal{O}; R^m) \mapsto H^{-2}(\mathcal{O}; R^m)$ is a continuous map, by continuous mapping theorem and (A.2), joint with the convergence in (A.4), as a sequence of $H^{-2}(\mathcal{O}; R^m)$-valued random variables, $H_\varepsilon$ converges in law/probability distribution to $H^{-2}(\mathcal{O}; R^m)$-valued $H_0$

$$\lim_{\varepsilon \to 0+} H_\varepsilon \stackrel{\mathrm{D}}{=} H_0,$$

where $H_0$ is defined through $-\Delta h_0 = H_0, x \in \mathcal{O}$ (note also that $h_0 = 0, x \in \partial\mathcal{O}$). Since $h_0$ is $H^2(\mathcal{O}; R^m)$-valued, we conclude that $H_0$ is indeed $L^2(\mathcal{O}; R^m)$-valued. By condition (2) of the theorem, $\bar{H}$ and $H_0$ has to be of the same probability law/distribution $\bar{H} \stackrel{\mathrm{D}}{=} H_0$.

By Skorohod representation, we may assume without lose of generality that all random variables live in the same reference probability space and selecting subsequence if necessary, convergences are convergences in almost sure sense. Then

$$f_0 = -\nabla \cdot h_0, \qquad N_0 = H_0 - \nabla f_0.$$

Finally, for each $\varphi \in C_c^\infty(\mathcal{O})$,

$$\int G_\varepsilon \cdot H_\varepsilon \varphi\, dx = \int G_\varepsilon \cdot (N_\varepsilon + \nabla f_\varepsilon)\varphi\, dx = \langle G_\varepsilon, N_\varepsilon \varphi\rangle - \langle G_\varepsilon, (\nabla\varphi) f_\varepsilon\rangle - \langle \nabla \cdot G_\varepsilon, f_\varepsilon \varphi\rangle,$$

where $\langle\cdot,\cdot\rangle$ is the continuous bilinear pairing between $H^{-p}(\mathcal{O})$ and $H_0^p(\mathcal{O})$, $p = 0, 1$. Assuming $\varphi$ is deterministic, by Lemma A.1,

$$\begin{aligned}\lim_{\varepsilon \to \infty} \langle \varphi, G_\varepsilon \cdot H_\varepsilon\rangle &\stackrel{\mathrm{D}}{=} \langle \bar{G}, N_0\varphi\rangle - \langle \bar{G}, (\nabla\varphi) f_0\rangle - \langle \nabla \cdot \bar{G}, f_0\varphi\rangle \\ &= \int \bar{G} \cdot (N_0 + \nabla f_0)\varphi\, dx = \langle \varphi, \bar{G} \cdot \bar{H}\rangle. \qquad \square\end{aligned}$$

The following lemma offers a practical way of verifying condition (3) of Theorem A.2, by exploring structural information in (stochastic) scalar conservation law equations.

**Lemma A.3** *(Murat's lemma). Suppose that*:

(1) $\{\phi_\varepsilon\colon \varepsilon > 0\}$ *is a stochastically bounded sequence in* $W^{-1,p}(\mathcal{O})$ *for some* $p > 2$. *That is, for each* $\delta > 0$, *there exists a* $C_\delta \in (0, \infty)$ *such that*

$$\sup_{\varepsilon > 0} P\left(\|\phi_\varepsilon\|_{W^{-1,p}} > C_\delta\right) < \delta.$$

(2) $\phi_\varepsilon = \chi_\varepsilon + \psi_\varepsilon$.
(3) $\{\chi_\varepsilon\colon \varepsilon > 0\}$ *is tight as* $H^{-1}(\mathcal{O})$-*valued random variables.*

(4) $\{\psi_\varepsilon\colon\ \varepsilon > 0\}$, *as signed Radon measure* $\mathcal{M}(\mathcal{O})$*-valued random variables, is stochastically bounded in total variation norm* $\|\cdot\|_{\mathcal{M}(\mathcal{O})}$. *That is, for each* $\delta > 0$, *there exists* $C_\delta \in (0, \infty)$ *such that*

$$\sup_{\varepsilon>0} P\big(\|\psi_\varepsilon\|_{\mathcal{M}(\mathcal{O})} > C_\delta\big) < \delta.$$

*Then the sequence of random fields* $\{\phi_\varepsilon\colon\ \varepsilon > 0\}$ *is tight as a sequence of* $H^{-1}(\mathcal{O})$*-valued random variables.*

**Proof.** First of all, by conditions of the lemma, for each $\delta > 0$, we can find constants $C_{1,\delta}, C_{2,\delta} > 0$ and a deterministic compact set $K_{1,\delta} \Subset H^{-1}(\mathcal{O})$ such that

$$\inf_{\varepsilon} P(\Omega_{\varepsilon,\delta}) > 1 - \delta,$$

where

$$\Omega_{\varepsilon,\delta} = \big\{\omega \in \Omega\colon\ \big\|\phi_\varepsilon(\cdot,\omega)\big\|_{W^{-1,p}} \leqslant C_{1,\delta}\big\} \cap \big\{\omega\colon\ \chi_\varepsilon(\cdot,\omega) \in K_{1,\delta}\big\} \\ \cap \big\{\omega\colon\ \big\|\psi_\varepsilon(\cdot,\omega)\big\|_{\mathcal{M}(\mathcal{O})} \leqslant C_{2,\delta}\big\}.$$

We repeat proof for the deterministic version of such lemma (e.g. [3, Lemma 15.2.1] or [6, Corollary 1]) in this stochastic context to ensure that choice of certain compact sets does not depend on $\omega$. First, we recall that the embedding from $\mathcal{M}(\mathcal{O}) \mapsto W^{-1,q}(\mathcal{O})$ is compact for any $q \in (1, \frac{m}{m-1})$ (e.g. Evans [6, Theorem 6]). Therefore there exists a deterministic compact $K_{2,\delta} = K_{2,\delta}(C_{2,\delta}) \Subset W^{-1,q}(\mathcal{O})$ such that

$$\big\{\omega\colon\ \|\psi_\varepsilon\|_{\mathcal{M}(\mathcal{O})} \leqslant C_{2,\delta}\big\} \subset \big\{\omega\colon\ \psi_\varepsilon(\cdot,\omega) \in K_{2,\delta}\big\}.$$

We now define new random fields $g_\varepsilon, h_\varepsilon \in W_0^{1,2}(\mathcal{O})$ as the (unique) weak solutions of

$$-\Delta g_\varepsilon = \chi_\varepsilon, \qquad g_\varepsilon = 0 \quad \text{on } \partial\mathcal{O}, \quad \text{and} \quad -\Delta h_\varepsilon = \psi_\varepsilon, \qquad h_\varepsilon = 0 \quad \text{on } \partial\mathcal{O},$$

and denote $f_\varepsilon = g_\varepsilon + h_\varepsilon$. Then by elliptic theory, there exist deterministic compact set $K_{3,\delta} = K_{3,\delta}(K_{1,\delta}) \Subset W_0^{1,2}(\mathcal{O})$ and deterministic compact set $K_{4,\delta} = K_{4,\delta}(K_{2,\delta}) \Subset W_0^{1,q}(\mathcal{O})$ such that

$$\big\{\omega\colon\ \chi_\varepsilon(\cdot,\omega) \in K_{1,\delta}\big\} \subset \big\{\omega\colon\ g_\varepsilon(\cdot,\omega) \in K_{3,\delta}\big\}, \qquad \big\{\omega\colon\ \psi_\varepsilon(\cdot,\omega) \in K_{2,\delta}\big\} \subset \big\{\omega\colon\ h_\varepsilon(\cdot,\omega) \in K_{4,\delta}\big\}.$$

Consequently, there exist deterministic compact sets

$$K_{5,\delta} \Subset W_0^{1,q}(\mathcal{O}), \qquad K_{6,\delta} = K_{6,\delta}(K_{5,\delta}) \Subset W_0^{-1,q}(\mathcal{O}),$$

such that

$$\{\omega\colon\ \chi_\varepsilon \in K_{1,\delta}\} \cap \big\{\omega\colon\ \|\psi_\varepsilon\|_{\mathcal{M}(\mathcal{O})} \leqslant C_{2,\delta}\big\} \subset \{\omega\colon\ f_\varepsilon \in K_{5,\delta}\} \subset \{\omega\colon\ \phi_\varepsilon \in K_{6,\delta}\},$$

where the last inclusion follows from

$$-\Delta f_\varepsilon = \phi_\varepsilon, \quad x \in \mathcal{O}, \qquad f_\varepsilon(x) = 0, \quad x \in \partial\mathcal{O}.$$

By interpolation between $W_0^{-1,q}(\mathcal{O})$ and $W_0^{-1,p}(\mathcal{O})$, there exists a deterministic compact set $K_{7,\delta} = K_{7,\delta}(K_{6,\delta}, C_{1,\delta}) \Subset H^{-1}(\mathcal{O})$ such that

$$\begin{aligned}\Omega_{\varepsilon,\delta} &= \bigl\{\omega\colon \|\phi_\varepsilon\|_{W^{-1,p}} \leqslant C_{1,\delta}\bigr\} \cap \{\omega\colon \chi_\varepsilon \in K_{1,\delta}\} \cap \bigl\{\omega\colon \|\psi_\varepsilon\|_{\mathcal{M}(\mathcal{O})} \leqslant C_{2,\delta}\bigr\} \\ &\subset \bigl\{\omega\colon \|\phi_\varepsilon\|_{W^{-1,p}} \leqslant C_{1,\delta}\bigr\} \cap \{\omega\colon \phi_\varepsilon \in K_{6,\delta}\} \subset \{\omega\colon \phi_\varepsilon \in K_{7,\delta}\}.\end{aligned}$$

Consequently,

$$\inf_{\varepsilon>0} P(\phi_\varepsilon \in K_{7,\delta}) > 1 - \delta.$$

Conclusion of the lemma follows. □

## References


[1] Luigi Ambrosio, Nicola Gigli, Giuseppe Savaré, Gradient Flows in Metric Spaces and in the Space of Probability Measures, Lectures Math. ETH Zürich, Birkhäuser, Basel, 2005.
[2] Gui-Qiang Chen, in: C.M. Dafermos, E. Feireisl (Eds.), The Handbook of Differential Equations, vol. 2, Elsevier Science, Amsterdam, 2005, pp. 1–104.
[3] Constantine Dafermos, Hyperbolic Conservation Laws in Continuum Physics, Springer, New York, 2000.
[4] Giuseppe Da Prato, Jerzy Zabczyk, Stochastic Equations in Infinite Dimensions, Encyclopedia Math. Appl., vol. 44, Cambridge Univ. Press, Cambridge, 1992.
[5] Stewart N. Ethier, Thomas G. Kurtz, Markov Processes, Wiley, New York, 1986.
[6] Lawrence G. Evans, Weak Convergence Methods for Nonlinear Partial Differential Equations, CBMS Reg. Conf. Ser. Math., vol. 74, Amer. Math. Soc., Providence, RI, 1990, published for the Conference Board of the Mathematical Sciences, Washington, DC.
[7] Lawrence C. Evans, Partial Differential Equations, Amer. Math. Soc., Providence, RI, 1998.
[8] Jong Uhn Kim, On a stochastic scalar conservation law, Indiana Univ. Math. J. 52 (1) (2003) 227–255.
[9] S. Kruzkov, First order quasilinear equations with several space variables, Mat. Sb. 123 (1970) 228–255; English transl. in: Math. USSR Sb. 10 (1970) 217–273.
[10] S. Kruzkov, First order quasilinear equations in several independent variables, Math. USSR Sb. 10 (1972) 217–243.
[11] H. Kunita, Stochastic differential equations and stochastic flows of diffeomorphisms, in: P.L. Hennequin (Ed.), École d'Été de Probabilités de Saint-Flour XII—1982, in: Lecture Notes in Math., vol. 1097, Springer, Berlin, 1982.
[12] Thomas G. Kurtz, Random time changes and convergence in distribution under the Meyer–Zheng conditions, Ann. Probab. 19 (3) (1991) 1010–1034.
[13] Thomas G. Kurtz, E. Philip Protter, Weak convergence of stochastic integrals and differential equations II: Infinite-dimensional case, in: CIME School in Probability, in: Lecture Notes in Math., vol. 1627, Springer, Berlin, 1996, pp. 197–285.
[14] Pierre-Louis Lions, Panagiotis E. Souganidis, Fully nonlinear stochastic partial differential equations, C. R. Acad. Sci. Paris Sér. I Math. 326 (9) (1998) 1085–1092; Fully nonlinear stochastic partial differential equations: Non-smooth equations and applications, C. R. Acad. Sci. Paris Sér. I Math. 327 (8) (1998) 735–741; Fully nonlinear stochastic PDE with semilinear stochastic dependence, C. R. Acad. Sci. Paris Sér. I Math. 331 (8) (2000) 617–624; Uniqueness of weak solutions of fully nonlinear stochastic partial differential equations, C. R. Acad. Sci. Paris Sér. I Math. 331 (10) (2000) 783–790.
[15] Daniel Revuz, Marc Yor, Continuous Martingales and Brownian Motion, third ed., Grundlehren Math. Wiss., vol. 293, Springer, Berlin, 1999.
[16] S.G. Samko, A.A. Kilbas, O.I. Marichev, Fractional Integrals and Derivatives. Theory and Applications, Gordon & Breach, Yverdon, 1993; translated from the 1987 Russian original, revised by the authors.

[17] Daniel W. Stroock, S.R. Srinivasa Varadhan, Multidimensional Diffusion Processes, Grundlehren Math. Wiss., vol. 233, Springer, Berlin, 1979.
[18] John B. Walsh, An introduction to stochastic partial differential equations, in: École dété de probabilités de Saint-Flour, XIV—1984, in: Lecture Notes in Math., vol. 1180, Springer, Berlin, 1986, pp. 265–439.