

# Invariant Measures for Nonlinear Conservation Laws Driven by Stochastic Forcing\*

Gui-Qiang G. CHEN<sup>1</sup>     Peter H. C. PANG<sup>1</sup>

*Dedicated to Professor Andrew J. Majda on the occasion of his 70th birthday*

**Abstract** Some recent developments in the analysis of long-time behaviors of stochastic solutions of nonlinear conservation laws driven by stochastic forcing are surveyed. The existence and uniqueness of invariant measures are established for anisotropic degenerate parabolic-hyperbolic conservation laws of second-order driven by white noises. Some further developments, problems, and challenges in this direction are also discussed.

**Keywords** Stochastic solutions, Entropy solutions, Invariant measures, Existence, Uniqueness, Stochastic forcing, Anisotropic degenerate, Parabolic-hyperbolic equations, Long-time behavior

**2000 MR Subject Classification** 35B40, 35K65, 37-02, 37A50, 37C40, 60H15, 35Q35, 58J70, 60G51, 60J65

## 1 Introduction

The analysis of long-time behaviors of global solutions is the second pillar in the theory of partial differential equations (PDEs for short), after the analysis of well-posedness. For the analysis of solution behaviors in the asymptotic regime, we seek to understand the global properties of the solution map, such as attracting or repelling sets, stable and unstable fixed points, limiting cycles, or chaotic behaviors that are properly determined by the entire system rather than a given path.

The introduction of noises usually serves to model dynamics phenomenologically—dynamics too complicated to model from first principles, or dynamics only the statistics of which are accurately known, or dynamics almost inherently random such as the decision of many conscious agents—or a combination of such behaviors. Mathematically, noises introduce behaviors that differ from deterministic dynamics, displaying much richer phenomena such as effects of dissipation, ergodicity, among others (cf. [29, 37, 47, 54, 71, 100] and the references cited therein). These phenomena are of intrinsic interest.

In this paper, we focus our analysis mainly on white-in-time noises. Indeed, they are the most commonly studied class of noises, though space-time white noises (such as in [26]) and

---

Manuscript received August 14, 2019.

<sup>1</sup>Mathematical Institute, University of Oxford, Oxford, OX2 6GG, UK. E-mail: chengq@maths.ox.ac.uk  
ho.pang@jesus.ox.ac.uk

\*This work was supported by the UK Engineering and Physical Sciences Research Council Award EP/E035027/1, EP/L015811/1, the Royal Society-Wolfson Research Merit Award (UK) and an Oxford Croucher Scholarship.

more general rough fluxes (e.g. [78–79, 92]) have also been considered. The reason for the prevalence of white noises as a basic model is not difficult to understand. First, Brownian motion occupies the unusual position of being simultaneously a martingale and a Lévy process. More importantly, with increments that are not only independent but also normally distributed, it commands a level of universality by virtue of the central limit theorem. Some of the ideas, techniques, and approaches presented here can be applied to equations with more general or other forms of stochastic forcing.

The organization of this paper is as follows: In §2, the notion of invariant measures is first introduced, then the Krylov-Bogoliubov approach for the existence of invariant measures is presented, and some methods for the uniqueness of invariant measures including the strong Feller property and the coupling method are discussed. In §3, some recent developments in the analysis of long-time behaviors of solutions of nonlinear stochastic PDEs are discussed. In §4, we establish the existence of invariant measures for nonlinear anisotropic parabolic-hyperbolic equations driven by white noises. In §5, we establish the uniqueness of invariant measures for the stochastic anisotropic parabolic-hyperbolic equations. In §6, we present some further developments, problems, and challenges in this research direction.

## 2 Invariant Measures

In this section, we first introduce the notion of invariant measures for random dynamic systems, and then present several approaches to establish the existence and uniqueness of invariant measures.

### 2.1 Notion of invariant measures

The notion of invariant measures on a dynamical system is quite straightforward. Let  $(\mathfrak{X}, \Sigma, \mu)$  be a measure space, and let  $S : \mathfrak{X} \rightarrow \mathfrak{X}$  be a map. System  $(\mathfrak{X}, \Sigma, \mu, S)$  is a measure-preserving system if  $\mu(S^{-1}A) = \mu(A)$  for any  $A \in \Sigma$ . Then  $\mu$  is called an invariant measure of map  $S$ .

On a random dynamic system (RDS for short), there is an added layer of complexity. We follow the standard definitions in [1]; see also [23] for further references on RDSs and [50, 52] in a specifically parabolic SPDE context.

Let  $(\Omega, \mathcal{F}, \mathbb{P})$  be a probability space, and let  $\theta_t : \Omega \rightarrow \Omega$  be a collection of probability-preserving maps. A measurable RDS on a measurable space  $(\mathfrak{X}, \Sigma)$  over quadruple  $(\Omega, \mathcal{F}, \mathbb{P}, \theta_t)$  is a map:

$$\varphi : \mathbb{R} \times \Omega \times \mathfrak{X} \rightarrow \mathfrak{X},$$

satisfying the following:

- (i) Measurability:  $\varphi$  is  $(\mathcal{B}(\mathbb{R}) \otimes \mathcal{F} \otimes \Sigma, \Sigma)$ -measurable;
- (ii) Cocycle property:  $\varphi(t, \omega) = \varphi(t, \omega, \cdot) : \mathfrak{X} \rightarrow \mathfrak{X}$  is a cocycle over  $\theta$ :

$$\begin{aligned}\varphi(0, \omega) &= \text{id}_{\mathfrak{X}}, \\ \varphi(t+s, \omega) &= \varphi(s, \theta_t \omega) \varphi(t, \omega),\end{aligned}$$

where  $\mathcal{B}(\mathbb{R})$  denotes the collection of Borel sets in  $\mathbb{R}$ .

We think of  $\Omega \times \mathfrak{X} \rightarrow \Omega$  as a fibre bundle with fibre  $\mathfrak{X}$ . On the bundle, we have the skew product defined as  $\Theta_t = (\theta_t, \varphi)$ . Then the invariant measures can be defined as follows:

**Definition 2.1** (Invariant Measures) *An invariant measure on a RDS  $\varphi$  over  $\theta_t$  is a probability measure  $\mu$  on  $(\Omega \times \mathfrak{X}, \mathcal{F} \otimes \Sigma)$  satisfying*

$$(\Theta_t)_*\mu = \mu, \quad \mu(\cdot, \mathfrak{X}) = \mathbb{P},$$

where  $(\Theta_t)_*\mu := \mu \circ (\Theta_t)^{-1}$  is the pushforward measure.

Any probability measure  $\mu$  on  $\Omega \times \mathfrak{X}$  admits a disintegration:

$$\mu(\omega, u) = \nu_\omega(u) \mathbb{P}(\omega).$$

A measure  $\nu_\omega$  is stationary if

$$\varphi(t, \omega)_*\nu_\omega = \nu_{\theta_t \omega}.$$

Let  $\{\mathcal{F}_t\}_{t \geq 0}$  be a filtration associated with the RDS  $\varphi$ , i.e., an increasing sequence of  $\sigma$ -subalgebras of  $\mathcal{F}$  by which  $\varphi(t, \cdot, x)$  is measurable (adapted). A Markov invariant measure is an invariant measure for which map:  $\omega \mapsto \nu_\omega(\Gamma)$  is  $(\mathcal{F}_0, \mathcal{B}(\mathbb{R}))$ -measurable for any  $\Gamma \in \mathcal{B}(\mathfrak{X})$  (cf. [72, §4.2.1]).

The disintegration of measures is unique. There is a one-to-one correspondence between a Markov invariant measure and a stationary measure (cf. [22, 74–75]). Associated with an invariant measure is a random attracting set, which is generalized from the deterministic context (cf. [24–25]). In the context of dissipative PDEs perturbed by noises, it can be shown that the Hausdorff dimension of an attracting set is finite by the methods similar to those used in the deterministic case (cf. [98] and the references cited therein) of linearizing the flow and estimating the sums of global Lyapunov exponents (cf. [25, 31–32, 95]).

## 2.2 Approaches for the existence of invariant measures

There are several approaches to establish the existence of invariant measures. One of the approaches is the Krylov-Bogoliubov approach, as we are going to discuss here. Another approach is via Khasminskii's theorem (cf. [30]). Both of them are based on the compactness property provided by the Prohorov theorem.

We first recall that a sequence of probability measures  $\{\nu_n\}$  on a measure space  $\mathfrak{X}$  is tight if, for every  $\varepsilon > 0$ , there is a compact set  $K_\varepsilon \subseteq \mathfrak{X}$  such that

$$\nu_n(\mathfrak{X} \setminus K_\varepsilon) \leq \varepsilon \quad \text{uniformly in } n.$$

**Lemma 2.1** (Prohorov Theorem) *A tight sequence of probability measures  $\nu_n$  is weak\*-compact in the space of probability measures; that is, there exists a subsequence (still denoted)  $\nu_n$  and a probability measure  $\nu$  such that  $\nu_n \xrightarrow{*} \nu$ .*

**Theorem 2.1** (Krylov-Bogoliubov Theorem) *Let  $\mathcal{P}_t$  be a semigroup satisfying the Feller property that  $\phi \in C(\mathfrak{X})$  implies  $\mathcal{P}_s \phi \in C(\mathfrak{X})$  for any  $s > 0$ , and let  $\mu$  be a probability measure on  $\mathfrak{X}$  such that the measure sequence:*

$$\nu_T = \frac{1}{T} \int_0^T \mathcal{P}_t^* \mu \, dt \tag{2.1}$$

*is tight. Then there exists an invariant measure for  $\mathcal{P}_t$ .*

The key of its proof is that the invariant measure generated by the Krylov-Bogoliubov theorem is the weak\*-limit of  $\nu_T$  as  $T \rightarrow \infty$ . This can be seen as follows: By the Prohorov theorem, the tight sequence has a weakly converging subsequence (still denoted as)  $\{\nu_T\}$  for  $T$  ranging over a unbounded subset of  $\mathbb{R}$ , whose limit is  $\nu_*$ . Then

$$\begin{aligned}
 \langle \mathcal{P}_s^* \nu_*, \varphi \rangle_{\mathfrak{X}} &= \langle \nu_*, \mathcal{P}_s \varphi \rangle_{\mathfrak{X}} \\
 &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \langle \mathcal{P}_t^* \mu, \mathcal{P}_s \varphi \rangle_{\mathfrak{X}} dt \\
 &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \langle \mathcal{P}_{t+s}^* \mu, \varphi \rangle_{\mathfrak{X}} dt \\
 &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \langle \mathcal{P}_t^* \mu, \varphi \rangle_{\mathfrak{X}} dt + \lim_{T \rightarrow \infty} \frac{1}{T} \int_T^{T+s} \langle \mathcal{P}_t^* \mu, \varphi \rangle_{\mathfrak{X}} dt \\
 &\quad - \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^s \langle \mathcal{P}_t^* \mu, \varphi \rangle_{\mathfrak{X}} dt \\
 &= \lim_{T \rightarrow \infty} \int_{\mathfrak{X}} \varphi(u) d\nu_T.
 \end{aligned}$$

In the above, we require the Feller property to execute the first equality, as  $\mathcal{P}_s \varphi$  has to remain continuous. With this, the second and third terms after the fourth equality above tend to zero in the limit  $T \rightarrow \infty$ , as  $s$  is fixed.

The following lemma provides two sufficient conditions for the tightness of  $\{\nu_T\}$ .

**Lemma 2.2** *A measure sequence  $\{\nu_T\}$  is tight if one of the following conditions holds:*

- (i)  $\{\mathcal{P}_t^* \mu\}$  is tight;
- (ii)  $\{\mathcal{P}_t\}$  are compact for  $t > 0$  so that

$$\mathcal{P}_t(\mathfrak{X}) \subseteq \mathfrak{Y} \quad \text{for almost all } t > 0,$$

for a Banach space  $\mathfrak{Y}$  such that there is a compact embedding  $\mathfrak{Y} \hookrightarrow \mathfrak{X}$  and there exists  $C > 0$  independent of  $T$  so that

$$\frac{1}{T} \int_0^T \|\mathcal{P}_t u_0\|_{\mathfrak{Y}} dt \leq C \quad \text{for any } u_0 \in \mathfrak{X}. \quad (2.2)$$

In addition,  $\mu = \delta_{u_0}$  for some  $u_0 \in \mathfrak{X}$ .

**Proof** For (i), we know that, for any  $\varepsilon > 0$ , there is a compact set  $K_\varepsilon \subseteq \mathfrak{X}$  such that

$$\mathcal{P}_t^* \mu(\mathfrak{X} \setminus K_\varepsilon) \leq \varepsilon \quad \text{uniform in } t > 0.$$

Then

$$\nu_T(\mathfrak{X} \setminus K_\varepsilon) \leq \frac{1}{T} \int_0^T \mathcal{P}_t^* \mu(\mathfrak{X} \setminus K_\varepsilon) dt \leq \varepsilon.$$

For (ii), let  $K_R = \{u \in \mathfrak{X} : \|u\|_{\mathfrak{Y}} \leq R\}$ . Since  $\mathfrak{Y} \hookrightarrow \mathfrak{X}$  is compact,  $K_R$  is compact in  $\mathfrak{X}$ .

If  $u \in \mathfrak{X} \setminus K_R$ , then  $\|u\|_{\mathfrak{Y}} > R$ . Writing  $f(\cdot) = \|\cdot\|_{\mathfrak{Y}}$ , then

$$\nu_T(\mathfrak{X} \setminus K_R) \leq \nu_T(\{f(u) > R\}).$$

Applying the Markov inequality to  $f$ , we have

$$\begin{aligned}\nu_T(\mathfrak{X} \setminus K_R) &\leq \frac{1}{R} \int_{\mathfrak{X}} f(u) \, d\nu_T(u) \\ &= \frac{1}{RT} \int_0^T \int_{\mathfrak{X}} (\mathcal{P}_t f)(u) \, d\mu(u) \, dt.\end{aligned}$$

Since  $\mu = \delta_{u_0}$  for some  $u_0 \in \mathfrak{X}$ , then

$$\frac{1}{RT} \int_0^T \int_{\mathfrak{X}} (\mathcal{P}_t f)(u) \, d\mu(u) \, dt = \frac{1}{RT} \int_0^T (\mathcal{P}_t f)(u_0) \, dt = \frac{1}{RT} \int_0^T f(u(t)) \, dt.$$

Therefore, if the temporal average (2.2) is bounded, then, for any  $\varepsilon > 0$ , we can choose  $R > \frac{1}{\varepsilon}$  to conclude

$$\nu_T(\mathfrak{X} \setminus K_R) \leq \nu_T(\{f(u) > R\}) < \varepsilon.$$

In this way, a compact set  $K_R$  has been found such that  $\nu_T(\mathfrak{X} \setminus K_R) \leq \varepsilon$ , which implies that  $\{\nu_T\}$  is tight.

This framework can be further refined. An example of such an extension can be found in [21], in which the Feller property could not be proved in the context of the one-dimensional stochastic Navier-Stokes equations. Whilst the Feller condition is not available, the continuous dependence (without rates) can be shown. By using the continuous dependence, a class of functions,  $\mathcal{G} \supseteq C(\mathfrak{X})$ , is defined so that  $\mathcal{G}$  is continuous on the elements of the solution space with finite energy, though not necessarily the entire solution space. With these, it has been shown in [21] that  $\mathcal{P}_t$  is invariant under  $\mathcal{G}$ . Then the existence of invariant measures is proved in two steps: First, an energy bound is employed to yield the tightness, so that the existence of a limiting measure is shown to exist; then the limiting measure is shown to be invariant (without invoking the Feller property) by using the continuity condition imposed on  $\mathcal{G}$  and following the arguments as in the proof of Theorem 2.1.

### 2.3 Approaches for the proof of the uniqueness of invariant measures

It is well known that the invariant measures of a map form a convex set in the probability space on  $X$ . By the Krein-Milman theorem, a convex set is the closure of convex combinations of its extreme points. These extreme points  $\mu$  happen to be ergodic measures, which are characterized as the property that, for a measurable subset  $A \subseteq \mathfrak{X}$ ,

$$\mu((S^{-1}A)\Delta A) = 0 \Leftrightarrow \mu(A) = 0 \text{ or } \mu(A) = 1,$$

where  $A\Delta B := (A \setminus B) \cup (B \setminus A)$ .

Ergodic measures heuristically carve up the solution space into essentially disjoint subsets, since any two ergodic measures of a process either coincide or are singular with respect to one another. This is a simple consequence of the property stated above.

It also follows from the extremal property of ergodic measures that, if there are more than one invariant measure, then there are at least two ergodic measures.

There are several approaches to establish the uniqueness of invariant measures.

**The Strong Feller Property** This is one of the common conditions used to ensure the uniqueness.

**Definition 2.2** (Strong Feller Property) *A Markov transition semigroup  $\mathcal{P}_t$  is of the strong Feller property at time  $t$  if  $\mathcal{P}_t\varphi$  is continuous for every bounded measurable  $\varphi : \mathfrak{X} \rightarrow \mathbb{R}$ .*

The strong Feller property guarantees the uniqueness of invariant measures (cf. [42, 70]); see also [30, Theorem 5.2.1, 86], and the references cited therein.

The strong Feller property always holds for transition semigroups of processes associated with nonlinear stochastic evolution equations with Lipschitz nonlinear coefficients and nondegenerate diffusion (e.g. [91]).

**The Coupling Method** This method is a powerful tool in probability theory introduced in Doeblin-Fortet [39–40], which can be used to show the uniqueness of invariant measures.

The general argument proceeds as follows: Let  $X_t$  be a Markov process with initial distribution  $\mu_0$ , and let  $Y_t$  be an independent copy of the process with an initial distribution that is an invariant measure  $\mu$ . Then the first meeting time  $\mathcal{T}$  is a stopping time, and the process defined by

$$Z_t = \begin{cases} X_t & \text{for } t < \mathcal{T}, \\ Y_t & \text{for } t \geq \mathcal{T} \end{cases}$$

is also a copy of  $X_t$  by the strong Markov property.

Using the definition of  $Z_t$ , we can write

$$\begin{aligned} \mathcal{P}_t^* \mu_0 - \mu &= (Z_t)_* \mathbb{P} - (Y_t)_* \mathbb{P} \\ &= (\mathbf{1}_{\{t < \mathcal{T}\}} Z_t)_* \mathbb{P} + (\mathbf{1}_{\{t \geq \mathcal{T}\}} Z_t)_* \mathbb{P} - (\mathbf{1}_{\{t < \mathcal{T}\}} Y_t)_* \mathbb{P} - (\mathbf{1}_{\{t \geq \mathcal{T}\}} Y_t)_* \mathbb{P} \\ &= (\mathbf{1}_{\{t < \mathcal{T}\}} Z_t)_* \mathbb{P} - (\mathbf{1}_{\{t < \mathcal{T}\}} Y_t)_* \mathbb{P}. \end{aligned}$$

Then the total variation norm of  $\mathcal{P}_t^* \mu_0 - \mu$  can be estimated as

$$\begin{aligned} \|\mathcal{P}_t^* \mu_0 - \mu\|_{TV} &\leq \int (\mathbf{1}_{\{t < \mathcal{T}\}} Z_t)_* d\mathbb{P}(u) + \int (\mathbf{1}_{\{t < \mathcal{T}\}} Y_t)_* d\mathbb{P}(u) \\ &= \mathbb{P}(\{t < \mathcal{T}\}). \end{aligned}$$

Assume that  $\mathcal{T}$  can be shown to be almost surely finite. Then, as  $t \rightarrow \infty$ , we see that  $\mathcal{P}_t^* \mu_0 \rightarrow \mu$ , and there is only one invariant measure.

The coupling method has other applications in various different settings and can be implemented in qualitatively different ways; see also [77, 101] and the references cited therein.

In our applications for the uniqueness of invariant measures for stochastic anisotropic parabolic-hyperbolic equations in §5,  $\mathcal{T}$  will be slightly modified to be the time of entry into a small ball. Moreover, instead of the use of independent copies, we take two solutions starting at different initial data, since our Markov processes are solutions of the equations with pathwise uniqueness properties.

First, we show in §5.1 that the two solutions  $u$  and  $v$  enter a given ball in finite time, almost surely. This is a stopping time. From this, by the strong Markov property, we construct a sequence of increasing, almost surely finite stopping times in (5.2), which are spaced at least  $T$  apart, for some  $T > 0$  later to be fixed.

Then we show in §5.2 that, for a well-chosen  $T > 0$ , if a solution starts within the same given ball, and the noise is uniformly small in  $W_x^{1,\infty}$  over a duration of length  $T$ , then the temporal

average of  $\|u(t)\|_{L_x^1}$  over that temporal interval can be taken to be smaller than some  $\varepsilon$ . Since the noise is  $\sigma(x)W$ , the uniform smallness in  $W_x^{1,\infty}$  over an interval  $[\mathcal{T}, \mathcal{T} + T]$  depends entirely on the size of  $W$ .

We see that, for  $T > 0$ , the probability that the change in the noise remains small between  $\mathcal{T}$  and  $\mathcal{T} + T$  is strictly positive. By the strong Markov property, we can replace  $\mathcal{T}$  with any other stopping time (e.g. the one in the sequence constructed) spaced at least  $T$  apart. Using the  $L^1$ -contraction, we show finally in §5.3 that the probability that the difference between the two solutions remains large for all intervals  $[\mathcal{T}, \mathcal{T} + T]$ , with  $\mathcal{T}$  in the sequence of increasing stopping times, is bounded by the probability that the noise is large in  $W_x^{1,\infty}$  over all such sequences. This must be vanishingly small, as the probability is strictly less than one on each individual sequence.

### 3 Nonlinear Conservation Laws Driven by Stochastic Forcing

In this section, we discuss one strand of the recent developments in the analysis of long-time behaviors of global solutions of nonlinear conservation laws driven by stochastic forcing.

#### 3.1 The stochastic Burgers equation

The Burgers equation is the archetypal nonlinear transport equation in many ways. The stochastic Burgers equation has also been used in turbulence and interface dynamics modelling; see [28, 62, 66, 76] and the references cited therein.

The existence of a non-trivial invariant measure of the process associated to the one-dimensional Burgers equation driven by an additive spatially periodic white noise was first derived in Sinai [96].

The long-time behavior of global solutions of the Burgers equation in one spatial dimension driven by space-time white noise has also been considered in the form

$$\partial_t u + \partial_x \left( \frac{u^2}{2} \right) = \partial_{xx}^2 u + \partial_{xt}^2 \widetilde{W},$$

where  $\widetilde{W} := \widetilde{W}(x, t)$  is a zero-mean Gaussian process with a covariance function given by

$$\mathbb{E}[\widetilde{W}(x, t)\widetilde{W}(y, s)] = (x \wedge y)(t \wedge s).$$

Apart from the global well-posedness in  $L^2(\mathbb{R})$ , it is known that an invariant measure for the transition semigroup exists, for example, via an argument of [25, 51] by using the ergodic theorem (cf. [28, 62]). Similar techniques have also been applied to study the two-dimensional Navier-Stokes equations driven by space-time white noises (e.g. [26]).

Attention in the development of the stochastic Burgers equation with vanishing viscosity has also been turned to the question of additive (spatial) noise in an equation of the form

$$\partial_t u + \partial_x \left( \frac{u^2}{2} \right) = \sum_{k=0}^{\infty} \partial_x F_k(x) dW^k + \varepsilon \partial_{xx}^2 u. \quad (3.1)$$

The existence of invariant measures for equation (3.1) with  $\varepsilon = 0$  is known (e.g. [45]). One of the key points is that there is enough energy dissipation in the inviscid limiting solutions as  $\varepsilon \rightarrow 0$  (satisfying the Lax entropy condition) so that such an invariant measure exists.

The argument for the existence proof of invariant measures in [45] is not directly via the general methods discussed in §2 above. Instead, the structure of the equation is exploited to form a variational problem in [44]. The minimizers of the action functional

$$\mathcal{A}[y(t)] = \frac{1}{2} \int_{t_1}^{t_2} \dot{y}^2(s) \, ds + \int_{t_1}^{t_2} \sum_k F_k(y(s)) \, dW^k(s)$$

are the curves that satisfy Newton's equations for the characteristics. These minimizers have an existence and uniqueness property with probability one. Through this, a one force-one solution principle has been shown, in which the random attractor consists of a single trajectory almost surely, which in turn leads to the proof of the existence of an invariant measure for (3.1).

### 3.2 Kinetic formulation

The theory of kinetic formulation has been developed over the last three decades (cf. Perthame [90] and the references cited therein). In particular, the compactness of entropy solutions of multidimensional scalar hyperbolic conservation laws with a genuine nonlinearity was first established by Lions-Perthame-Tadmor in [80] via combining the kinetic formulation with corresponding velocity averaging. The velocity averaging is a technique whereby a genuine nonlinearity condition (i.e., a non-degeneracy condition on the nonlinearity) can be shown to imply the compactness (or even improved fractional regularity under a stronger condition) of solutions via the kinetic formulation, as seen in subsequent sections, especially in condition (3.14).

We discuss the kinetic formulation in the context of scalar hyperbolic conservation laws here.

One of the inspirations for a kinetic formulation originated from the kinetic theory of gases. One starts with a simple step function as the kinetic function:

$$\chi^r(\xi) := \chi(\xi, r) = \begin{cases} 1 & \text{for } 0 < \xi < r, \\ -1 & \text{for } r < \xi < 0, \\ 0 & \text{otherwise.} \end{cases}$$

Then, for any  $\eta \in C^1$ , the following representation formula holds:

$$\int_{\mathbb{R}} \eta'(\xi) \chi^u(\xi) \, d\xi = \eta(u) - \eta(0). \quad (3.2)$$

A simple combination of kinetic functions yields

$$|u - v| = \int (|\chi^u| + |\chi^v| - 2\chi^u \chi^v) \, d\xi. \quad (3.3)$$

This provides an approach to the derivation of the  $L^1$ -contraction between two solutions, by estimating the terms on the right.

There are several variations on the form of the kinetic function. Since  $|u - v| = (u - v)_+ + (v - u)_+$ , it suffices for a variation, or combinations, of the kinetic function to capture  $(u - v)_+$ , which is simpler than (3.3). This can be done by considering the following kinetic function:

$$\tilde{\chi}^u := \tilde{\chi}(\xi, u) = 1 - H(\xi - u) = H(u - \xi),$$

where  $H = \mathbf{1}_{[0,\infty)}$  is the Heaviside step function. We then have the representation formula:

$$\eta(u) = \int_{\mathbb{R}} \eta'(\xi) \tilde{\chi}^u(\xi) \, d\xi \quad \text{for } \eta \in C^1 \text{ with } \eta(-\infty) = 0.$$

In particular,

$$(u - v)_+ = \int \tilde{\chi}^u(\xi)(1 - \tilde{\chi}^v(\xi)) \, d\xi. \quad (3.4)$$

Such a kinetic function has been popularized by [80] and has been used, inter alia, in [12, 33–35], and even as far back as [61].

The usefulness of the kinetic function can be seen in the kinetic formulation of scalar conservation laws, in which the kinetic variable takes the place of the solution in the nonlinear coefficients so that a degree of linearity is restored for analysis. In this formulation, many powerful linear methods such as the Fourier transform become not only applicable, but also natural.

Following Chen-Pang [12], we now derive the kinetic formulation of nonlinear anisotropic parabolic-hyperbolic equations of second order:

$$\partial_t u + \nabla \cdot F(u) = \nabla \cdot (\mathbf{A}(u) \nabla u) + \sigma(u, x) \partial_t W, \quad (3.5)$$

where  $F$  is a locally Lipschitz vector flux function of polynomial growth,  $\mathbf{A}$  is a positive semi-definite matrix function with continuous entries of polynomial growth, and  $\nabla = \nabla_x := (\partial_{x_1}, \dots, \partial_{x_d})$ .

Consider the vanishing viscosity approximation to (3.5):

$$\partial_t u^\varepsilon + \nabla \cdot F(u^\varepsilon) = \nabla \cdot ((\mathbf{A}(u^\varepsilon) + \varepsilon \mathbf{I}) \nabla u^\varepsilon) + \sigma(u^\varepsilon, x) \partial_t W,$$

where  $I$  is the identity matrix. Let  $\eta \in C^1$  be an entropy with  $\eta(0) = 0$ . Using the Itô formula, we have

$$\begin{aligned} \partial_t \eta(u^\varepsilon) &= -\eta'(u^\varepsilon) \nabla \cdot F(u^\varepsilon) + \eta'(u^\varepsilon) \sigma(u^\varepsilon, x) \partial_t W + \frac{1}{2} \eta''(u^\varepsilon) \sigma^2(u^\varepsilon, x) \\ &\quad + \nabla \cdot (\eta'(u^\varepsilon) \mathbf{A}(u^\varepsilon) \nabla u^\varepsilon) - \eta''(u^\varepsilon) \mathbf{A}(u^\varepsilon) : (\nabla u^\varepsilon \otimes \nabla u^\varepsilon) \\ &\quad + \varepsilon \Delta \eta(u^\varepsilon) - \varepsilon \eta''(u^\varepsilon) |\nabla u^\varepsilon|^2, \end{aligned}$$

where we have used the notation:  $\mathbf{A} : \mathbf{B} = \sum_{i,j} \mathbf{a}_{ij} \mathbf{b}_{ij}$  for matrices  $\mathbf{A} = (\mathbf{a}_{ij})$  and  $\mathbf{B} = (\mathbf{b}_{ij})$  of the same size.

Applying the representation formula (3.2) yields

$$\begin{aligned} \partial_t \int \eta'(\xi) \chi^{u^\varepsilon} \, d\xi &= -\nabla \cdot \left( \int \eta'(\xi) F'(\xi) \chi^{u^\varepsilon} \, d\xi \right) + \langle \sigma(\cdot, x) \partial_t W(t) \delta(\cdot - u^\varepsilon), \eta'(\cdot) \rangle \\ &\quad + \nabla^2 : \left( \int \eta'(\xi) \mathbf{A}(\xi) \chi^{u^\varepsilon} \, d\xi \right) - \langle \mathbf{A}(\cdot) : (\nabla u^\varepsilon \otimes \nabla u^\varepsilon) \delta(\cdot - u^\varepsilon), \eta''(\cdot) \rangle \\ &\quad - \langle \varepsilon |\nabla u^\varepsilon|^2 \delta(\cdot - u^\varepsilon), \eta''(\cdot) \rangle + \frac{1}{2} \langle \sigma^2(\cdot, x) \delta(\cdot - u^\varepsilon), \eta''(\cdot) \rangle \\ &\quad + \varepsilon \Delta \left( \int \eta'(\xi) \chi^{u^\varepsilon} \, d\xi \right). \end{aligned}$$

Assume that  $u^\varepsilon(x, t) \rightarrow u(x, t)$  a.e. almost surely as  $\varepsilon \rightarrow 0$ . Then, taking  $\eta'(\xi)$  as a test function and letting  $\varepsilon \rightarrow 0$ , we arrive heuristically at the formulation:

$$\partial_t \chi^u + F'(\xi) \cdot \nabla \chi^u = \mathbf{A}(\xi) : \nabla^2 \chi^u + \sigma(\xi, x) \partial_t W(t) \delta(\xi - u) + \partial_\xi(m^u + n^u - p^u), \quad (3.6)$$

which holds in the distributional sense, where  $m^u$ ,  $n^u$  and  $p^u$  are Radon measures that are the limits of the following measure sequences:

$$\begin{aligned} \varepsilon |\nabla u^\varepsilon|^2 \delta(\xi - u^\varepsilon) &\rightharpoonup m^u, \\ \mathbf{A}(\xi) : (\nabla u^\varepsilon \otimes \nabla u^\varepsilon) \delta(\xi - u^\varepsilon) &\rightharpoonup n^u, \\ \frac{1}{2} \sigma^2(\xi, x) \delta(\xi - u^\varepsilon) &\rightharpoonup p^u. \end{aligned}$$

The Radon measure  $m^u$  is the kinetic dissipation measure and  $n^u$  is the parabolic defect measure, which capture the dissipation from the vanishing viscosity terms and the degenerate parabolic terms, respectively. In addition, the Radon measure

$$p^u = \frac{1}{2} \sigma^2(\xi, x) \delta(\xi - u)$$

arises from the Itô correction. As  $\mathbf{A}$  is positive semi-definite, it is manifest that  $m^u$ ,  $n^u$  and  $p^u$  are all non-negative.

More precisely, the parabolic defect measure  $n^u \geq 0$  is determined by the following: For any  $\varphi \in C_0(\mathbb{R} \times \mathbb{R}^d \times \mathbb{R}_+)$ ,

$$n^u(\varphi) = \int_{\mathbb{R}_+} \int_{\mathbb{R}^d} \varphi(u(x, t), x, t) \left| \nabla_x \cdot \left( \int_0^u \boldsymbol{\alpha}(\zeta) d\zeta \right) \right|^2 dx dt. \quad (3.7)$$

The kinetic dissipation measure  $m^u \geq 0$  satisfies the following:

(i) For  $B_R^c \subset \mathbb{R}$  as the complement of the ball of radius  $R$ ,

$$\lim_{R \rightarrow \infty} \mathbb{E}[(m^u + n^u)(B_R^c \times \mathbb{T}^d \times [0, T])] = 0. \quad (3.8)$$

(ii) For any  $\varphi \in C_0(\mathbb{R} \times \mathbb{R}^d)$ ,

$$\int_{\mathbb{R} \times \mathbb{R}^d \times [0, T]} \varphi(\xi, x) d(m^u + n^u)(\omega; \xi, x, t) \in L^2(\Omega) \quad (3.9)$$

admits a predictable representative (in the  $L^2$ -equivalence classes of functions).

Then, following Chen-Pang [12], we introduce the notion of kinetic solutions.

**Definition 3.1** (Stochastic Kinetic Solutions) *A function*

$$u \in L^p(\Omega \times [0, T]; L^p(\mathbb{R}^d)) \cap L^p(\Omega; L^\infty([0, T]; L^p(\mathbb{R}^d)))$$

is called a kinetic solution of (3.5) with initial data:  $u|_{t=0} = u_0$ , provided that  $u$  satisfies the following:

(i)  $\nabla \cdot \left( \int_0^u \boldsymbol{\alpha}(\xi) d\xi \right) \in L^2(\Omega \times \mathbb{R}^d \times [0, T]);$

(ii) for any bounded  $\varphi \in C(\mathbb{R})$ , the Chen-Perthame chain rule relation in [13] holds:

$$\nabla \cdot \left( \int_0^u \varphi(\xi) \boldsymbol{\alpha}(\xi) d\xi \right) = \varphi(u) \nabla \cdot \left( \int_0^u \boldsymbol{\alpha}(\xi) d\xi \right) \quad (3.10)$$

in  $\mathcal{D}'(\mathbb{T}^d)$  and almost everywhere in  $(t, \omega)$ .

(iii) There is a kinetic measure  $m^u \geq 0$   $\mathbb{P}$ -a.e. such that, given the parabolic defect measure  $n^u$ , the following holds almost surely: For any  $\varphi \in C_c^\infty(\mathbb{R}, \mathbb{R}^d \times [0, T))$ ,

$$\begin{aligned} & - \int_0^T \iint \chi(\xi, u) \partial_t \varphi \, d\xi \, dx \, dt - \iint \chi(\xi, u_0) \varphi(\xi, x, 0) \, d\xi \, dx \\ &= \int_0^T \iint \chi(\xi, u) F'(\xi) \cdot \nabla \varphi \, d\xi \, dx \, dt + \int_0^T \iint \chi(\xi, u) \mathbf{A}(\xi) : \nabla^2 \varphi \, d\xi \, dx \, dt \\ &+ \int_0^T \iint \varphi_\xi \, d(m^u + n^u)(\xi, x, t) - \frac{1}{2} \int_0^T \int \varphi_u(u, x, t) \sigma^2(u, x) \, dx \, dt \\ &- \int_0^T \int \varphi(u, x, t) \sigma(u, x) \, dx \, dW \quad \text{almost surely.} \end{aligned} \quad (3.11)$$

Equation (3.11) is obtained by testing (3.6) with  $\varphi$  and using the chain rule (3.10).

### 3.3 General scalar hyperbolic conservation laws driven by stochastic forcing

In Feng-Nualart [49], the well-posedness was studied for the one-dimensional scalar conservation laws driven by white noise

$$\partial_t u + \partial_x F(u) = \int_{z \in Z} \sigma(u, x; z) \, d_z W(t, z),$$

where  $Z$  is a metric space, and  $W$  is a space-time Gaussian noise martingale random measure with respect to a filtration  $\{\mathcal{F}_t\}_{t \geq 0}$  satisfying

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

for measurable sets  $A, B \subset Z$ , with a  $\sigma$ -finite Borel measure  $\mu$  on  $Z$ . The well-posedness theory was developed around the notion of strong stochastic entropy solutions introduced in Definition 2.6 in [49] when  $t \in [0, T)$  for any fixed  $T \in (0, \infty)$ . In addition to the usual definition of entropy solutions, the following further conditions on the solution,  $u = u(x, t)$ , for  $t \in [0, T]$  are required.

For any smooth approximation function  $\beta(u)$  of function  $u_+$  on  $\mathbb{R}$  and any  $\varphi \in C^\infty(\mathbb{R}^d \times \mathbb{R}^d)$  with  $\varphi \geq 0$ , and for any  $\mathcal{F}_t$ -adapted function  $v$  satisfying  $\sup_{0 \leq t \leq T} \mathbb{E}[\|v\|_{L_x^p}^p] < \infty$ , there exists a deterministic function  $\{A(s, t) : 0 \leq s \leq t\}$  such that the functional

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

satisfies

$$\begin{aligned} & \mathbb{E} \left[ \int \int_{(s, t] \times Z} f(r, z, u(y, t), y) dW(r, z) \, dy \right] \\ & \leq \mathbb{E} \left[ \int_{(s, t] \times Z} \int \partial_v f(r, z, v(y, r), y) \sigma(y, u(y, r); z) \, dy \, dr d\mu(z) \right] + A(s, t), \end{aligned}$$

and that there is a sequence of partitions of  $[0, T]$  so that

$$\lim_{\max |t_{i+1} - t_i| \rightarrow 0} \sum_{i=1}^m A(t_i, t_{i+1}) = 0.$$

This notion of a solution addresses the problem that, in any direct adaptation of the deterministic notion of entropy solutions, one encounters the question of adaptiveness of the Itô integral in the noise-noise interaction. With this notion, in [49], the  $L^1$ -contraction and comparison estimates of strong stochastic entropy solutions in any spatial dimension were established, while the existence of solutions is limited to the one-dimensional case based on the compensated compactness argument in Chen-Lu [11].

In Chen-Ding-Karlsen [9], the existence theory for strong stochastic entropy solutions was established for any spatial dimension with the key observation that the following  $BV$  bound is a corollary from the  $L^1$ -contraction inequality:

$$\mathbb{E}[\|u(t)\|_{BV}] \leq \mathbb{E}[\|u_0\|_{BV}],$$

which provides the strong compactness required for the existence theory in any spatial dimension. More precisely, the following theorem holds.

**Theorem 3.1** *Consider the Cauchy problem of the equation*

$$\partial_t u + \nabla \cdot F(u) = \sigma(u) \partial_t W \quad (3.12)$$

*with initial condition*

$$u|_{t=0} = u_0, \quad (3.13)$$

*satisfying*

$$\mathbb{E}[\|u_0\|_{L^p}^p + \|u_0\|_{BV}] < \infty \quad \text{for } p > 1,$$

*where  $F$  is a locally Lipschitz function of polynomial growth and  $\sigma$  is a globally Lipschitz function. Then there exists a unique strong stochastic entropy solution of the Cauchy problem (3.12)–(3.13) satisfying*

$$\mathbb{E}[\|u(t)\|_{BV}] \leq \mathbb{E}[\|u_0\|_{BV}].$$

This existence theory in  $L^p \cap BV$  can also be extended to the second-order equations (3.5) as established in Chen-Pang [12], including the case with heterogeneous flux functions  $F = F(u, x)$  (i.e., the space-translational variant case).

A well-posedness theory can also be developed for kinetic solutions to the multidimensional scalar balance laws with stochastic force (3.12), by employing the Gyöngy-Krylov framework where the existence of a martingale solution with pathwise uniqueness guarantees the strong existence; see [34]. In particular, the existence of martingale solutions can be proved via the notion of kinetic solutions. These results can be extended (e.g. [33, 67]) to encompass degenerate parabolic equations

$$\partial_t u + \nabla \cdot F(u) - \nabla \cdot (\mathbf{A}(u) \nabla u) = \sigma(u) \partial_t W.$$

A well-posedness theory has also been established based on the viscosity solutions (such as in [3]). To achieve this, the difficulties caused by the noise-noise interaction that has a non-zero correlation for the multiplicative noise case are avoided by directly comparing two entropy solutions to a viscosity solution.

In Karlsen-Storrøsten [69], these different viewpoints have been partially reconciled via a Malliavin viewpoint, in which the constant in the Kruzhkov entropy is interpreted as a Malliavin differentiable variable.

Long-time asymptotic results concerning the existence and uniqueness of invariant measures have followed the well-posedness theory. Concerning the stochastic balance law

$$\partial_t u + \nabla \cdot F(u) = \Phi(x) \, dB,$$

with evolution on torus  $\mathbb{T}^d$ , where  $B = \sum_k e_k W_k$  is a cylindrical Wiener process,  $\{e_k\}$  is a complete orthonormal basis of a Hilbert space,  $\Phi$  is a Hilbert-Schmidt operator given by  $\Phi(x) = \sum_k g_k(x) e_k$ , and  $g_k(x)$  satisfies

$$\int_{\mathbb{T}^d} g_k(x) \, dx = 0,$$

the existence and uniqueness of invariant measures were shown in [35]. In this case, the noise is additive; that is, it depends only on the spatial variable, but is independent of the solution—a point to which we will return.

These results can be summarized as follows:

**Theorem 3.2** *Let  $F$  satisfy the non-degeneracy condition: For some  $b < 1$  and a constant  $C > 0$ ,*

$$\int_0^\infty e^{-t} \sup_{\tau \in \mathbb{R}, |\hat{k}|=1} \mathcal{L}^1(\{\xi : |F'(\xi) \cdot \hat{k} + \tau| \leq \varepsilon t\}) dt \leq C \varepsilon^b \quad (3.14)$$

*for the Lebesgue measure  $\mathcal{L}^1$  on  $\mathbb{R}$ , in addition to the condition that  $|F''(\xi)| \lesssim |\xi| + 1$ . Then there exists an invariant measure to the process. Furthermore, if  $|F''(\xi)| \lesssim 1$  is bounded, then the invariant measure is unique.*

The bounds for the spaces on which the invariant measures are supported have also been derived. This result has been obtained by employing the velocity averaging. It has been built also on the related ideas of kinetic solutions, which is first applied to the velocity averaging in the deterministic context. They avoided the question of the Fourier transforms of the Wiener process by introducing regularizing operators.

Similar results were also derived for

$$\partial_t u + \nabla \cdot (F(u) \circ dW) = 0,$$

by further employing the conservative form as considered in Lions-Perthame-Souganidis [78–79]; see Gess-Souganidis [59]. A generalization of this with a degenerate parabolic term  $\nabla \cdot (\mathbf{A}(u) \nabla u)$  has also been considered in [48, 60]. In particular, Fehrman-Gess [48] investigated the well-posedness and continuous dependence of the stochastic degenerate parabolic equations of porous medium type, including the cases with fast diffusion and heterogeneous fluxes.

By using the methods developed in [57, 64–65, 72] and developing the probabilistic Gronwall inequality based on delicate reasoning about a stopping time, such MHD equations driven by additive noise of zero spatial average in the vanishing Rossby number and vanishing magnetic Reynold’s number limit were also shown to have a unique invariant measure (that is necessarily ergodic) in [56].

## 4 Stochastic Anisotropic Parabolic-Hyperbolic Equations I: Existence of Invariant Measures

In this section, we present an approach for establishing the existence of invariant measures for nonlinear anisotropic parabolic-hyperbolic equations driven by stochastic forcing

$$\partial_t u + \nabla \cdot F(u) = \nabla \cdot (\mathbf{A}(u) \nabla u) + \sigma(x) \partial_t W, \quad (4.1)$$

where  $\mathbf{A}$  is positive semi-definite, and  $\sigma$  has zero average over  $\mathbb{T}^d$ . The main focus of this section is on the presentation of the approach, so we do not seek the optimality of the results, while the results presented below can be further improved by refining the arguments and technical estimates required for the approach which is out of the scope of this section. More precisely, we establish the following theorem.

**Theorem 4.1** *Let  $F$  and  $\mathbf{A}$  satisfy the nonlinearity-diffusivity condition: There exist  $\beta \in (1, 2)$ ,  $\kappa \in (0, 1)$  and  $C > 0$ , independent of  $\lambda$ , such that*

$$\sup_{\tau \in \mathbb{R}, |\widehat{k}|=1} \int \frac{\lambda(\mathbf{A}(\xi) : \widehat{k} \otimes \widehat{k} + \lambda)}{(\mathbf{A}(\xi) : \widehat{k} \otimes \widehat{k} + \lambda)^2 + \lambda^\beta |F'(\xi) \cdot \widehat{k} + \tau|^2} d\xi =: \eta(\lambda) \leq C\lambda^\kappa \rightarrow 0 \quad (4.2)$$

as  $\lambda \rightarrow 0$ . In addition, let  $F$  and  $\mathbf{A}$  satisfy the condition:

$$|F''(\xi)| \lesssim |\xi| + 1, \quad |\mathbf{A}'(\xi)| \lesssim |\xi| + 1. \quad (4.3)$$

Then there exists an invariant measure to the process associated with the solutions to (4.1).

The approach is motivated by Debussche-Vovelle [35] by extending the case from first-order scalar balance laws to the second-order degenerate parabolic-hyperbolic equation (4.1). The first-order case is handled in [35], based on the velocity averaging and built on Lemma 2.4 of Bouchut-Desvillettes [6]. In our approach, we require a modified version of this lemma, which is incorporated into the calculation that allows us to exploit the cancellations in an oscillatory integral in this more general case than the first-order case. We now proceed to prove the theorem as follows:

(i) First we incorporate regularizing operators into the equation in order to exploit the bounds that can be provided in the Duhamel representation of the solution.

(ii) We separate the Duhamel representation of the solution into four different summands, the  $W^{s,q}$  norms of which we estimate.

(iii) Adding these estimates together by the triangle inequality and using the compact inclusion of  $W^{s,q}$  into a suitable  $L^q$  norm allow us to invoke the Krylov-Bogoliubov machinery described in §2.2.

We expound on the nonlinearity condition (4.2) in a remark below. As the conditions in (4.3) are invoked along the way, we also explain their relevance.

Consider the kinetic formulation of equation (4.1):

$$\partial_t \chi^u + (F'(\xi) \cdot \nabla - \mathbf{A}(\xi) : \nabla \otimes \nabla) \chi^u = \partial_\xi (m^u + n^u - p^u) + \sigma(x) \delta(\xi - u) \partial_t W. \quad (4.4)$$

In order to handle the two measures:  $m^u + n^u - p^u$  and  $\sigma(x) \delta(\xi - u)$ , we need to regularize the operators as in [35], by adding  $\gamma(-\Delta)^\alpha + \theta I$  to each side:

$$\partial_t \chi^u + (F'(\xi) \cdot \nabla - \mathbf{A}(\xi) : \nabla \otimes \nabla + \gamma(-\Delta)^\alpha + \theta I) \chi^u$$

$$= (\gamma(-\Delta)^\alpha + \theta I)\chi^u + \partial_\xi(m^u + n^u - p^u) + \sigma(x)\delta(\xi - u)\partial_t W \quad (4.5)$$

for  $\alpha = \frac{\beta-1}{\beta} \in (0, \frac{1}{2})$  for some  $\beta \in (1, 2)$  required in the nonlinearity-diffusivity condition (4.2).

We adapt the semigroup approach. There are specific reasons to include these regularizing operators: In order to estimate the measure,  $\sigma(x)\delta(\xi - u)\partial_t W$ , we require a spatial regularization provided by  $(-\Delta)^\alpha$  and temporal decay provided by  $\theta I$ .

More specifically, let  $\mathcal{S}(t)$  be the semigroup of operator  $\partial_t + (F'(\xi) \cdot \nabla - \mathbf{A}(\xi) : \nabla \otimes \nabla + \gamma(-\Delta)^\alpha + \theta I)$ :

$$\begin{aligned} \mathcal{S}(t)f(x) &= e^{-(F'(\xi) \cdot \nabla - \mathbf{A}(\xi) : \nabla \otimes \nabla + \gamma(-\Delta)^\alpha + \theta I)t} f \\ &= e^{-\theta t} (e^{t\mathbf{A}(\xi) : \nabla \otimes \nabla - t\gamma(-\Delta)^\alpha} f)(x - F'(\xi)t) \quad \text{for any } f = f(x). \end{aligned} \quad (4.6)$$

Then we can express the solution,  $\chi^u$ , to the kinetic formulation in the mild formulation

$$\begin{aligned} \chi^u &= \mathcal{S}(t)\chi^u(\xi, x, 0) + \int_0^t \mathcal{S}(s)(\gamma(-\Delta)^\alpha - \theta I)\chi^u(\xi, x, t-s) \, ds \\ &\quad + \int_0^t \mathcal{S}(t-s)\partial_\xi(m^u + n^u - p^u)(\xi, x, s) \, ds + \int_0^t \mathcal{S}(t-s)\sigma(x)\delta(\xi - u(x, s)) \, dW_s. \end{aligned}$$

This leads to the decomposition

$$u = u^0 + u^\flat + M_1 + M_2, \quad (4.7)$$

where

$$u^0(x, t) = \int \mathcal{S}(t)\chi^u(\xi, x, 0) \, d\xi, \quad (4.8)$$

$$u^\flat(x, t) = \int \int_0^t \mathcal{S}(s)(\gamma(-\Delta)^\alpha - \theta I)\chi^u(\xi, x, t-s) \, ds \, d\xi, \quad (4.9)$$

$$\langle M_1, \varphi \rangle = \int_0^t \int \langle \partial_\xi(m^u + n^u - p^u)(\cdot, x, t-s), \mathcal{S}^*(s)\varphi \rangle \, dx \, ds, \quad (4.10)$$

$$\langle M_2, \varphi \rangle = \int_0^t \int \langle \delta(\cdot - u(x, s)), \mathcal{S}^*(t-s)\varphi \rangle \sigma(x) \, dx \, dW_s, \quad (4.11)$$

where  $\mathcal{S}^*(t)$  is the dual operator of the semigroup operator  $\mathcal{S}(t)$ , and  $\varphi \in C(\mathbb{T}^d)$ .

We now estimate each of the four terms separately in each subsection. The first two integrals are essentially “deterministic” parts and estimated by the velocity averaging method, and the final two integrals incorporate stochastic elements and are treated by a kernel estimate on semigroup  $\mathcal{S}(t)$ .

#### 4.1 Analysis of $u^0$

Notice that the local Fourier transform in  $x \in \mathbb{T}^d$  for any periodic function  $g(x, \cdot)$  in  $x$  with period  $P = (P_1, \dots, P_d)$  is

$$\widehat{g}(k, \cdot) = \frac{1}{|\mathbb{T}^d|} \int_{\mathbb{T}^d} g(x, \cdot) e^{-ix \cdot k} \, dx,$$

where frequencies  $k = (k_1, \dots, k_d)$  are discrete:

$$k_i = \frac{2\pi}{P_i} n_i, \quad n_i = 0, \pm 1, \pm 2, \dots, \quad i = 1, \dots, d.$$

Taking the Fourier transform in  $x$  and integrating in  $\xi$ , we have

$$\begin{aligned}\widehat{u^0}(k, t) &= \int \widehat{\mathcal{S}}(t) \widehat{\chi^u}(\xi, k, 0) \, d\xi \\ &= \int e^{-(iF'(\xi) \cdot \widehat{k} + \mathbf{A}(\xi) : (k \otimes k) + \omega_k |k|)t} \widehat{\chi^u}(\xi, k, 0) \, d\xi,\end{aligned}$$

where  $\omega_k = \gamma|k|^{2\alpha-1} + \theta|k|^{-1}$ .

For simplicity, we denote  $\widehat{k} = \frac{k}{|k|}$  and  $\mathcal{A} = \mathcal{A}(\xi, \widehat{k}) = \mathbf{A}(\xi) : \widehat{k} \otimes \widehat{k}$ . Then we square the above and integrate in  $t$  from 0 to  $T$  to obtain

$$\begin{aligned}\int_0^T |\widehat{u^0}(k, t)|^2 \, dt &= \int_0^T \left| \int e^{-(iF'(\xi) \cdot \widehat{k} + \mathcal{A}(\xi, \widehat{k})|k| + \omega_k)|k|t} \widehat{\chi^u}(\xi, k, 0) \, d\xi \right|^2 \, dt \\ &\leq \frac{1}{|k|} \int \left| \int \mathbb{1}_{\{s>0\}} e^{-(iF'(\xi) \cdot \widehat{k} + \mathcal{A}(\xi, \widehat{k})|k| + \omega_k)s} \widehat{\chi^u}(\xi, k, 0) \, d\xi \right|^2 \, ds.\end{aligned}\quad (4.12)$$

Notice that it is impossible to extract the entire non-oscillatory part of the exponential from the integral in  $\xi$ , as was done with the lemma of Bouchut-Desvillettes [6]. However, by extending the range of integration over all  $\mathbb{R}$  to make the function in  $s$  smoother so that its transform has better decay properties, we can partially exploit the cancellations later:

$$\int_0^T |\widehat{u^0}(k, t)|^2 \, dt \leq \frac{1}{|k|} \int_{-\infty}^{\infty} \left| \int e^{iF'(\xi) \cdot \widehat{k}s} e^{-(\omega_k + \mathcal{A}|k|)|s|} \widehat{\chi^u}(\xi, k, 0) \, d\xi \right|^2 \, ds. \quad (4.13)$$

We can evaluate the temporal Fourier transform of the integrand explicitly:

$$\mathcal{F}^{-1}\{e^{iF'(\xi) \cdot \widehat{k}s} e^{-(\omega_k + \mathcal{A}|k|)|s|}\}(\tau) = -\frac{2(\mathcal{A}|k| + \omega_k)}{(\mathcal{A}|k| + \omega_k)^2 + |F'(\xi) \cdot \widehat{k} + \tau|^2}.$$

Next, using the Parseval identity in the temporal variable and the Cauchy-Schwarz inequality, we have

$$\begin{aligned}\int_0^T |\widehat{u^0}(k, t)|^2 \, dt &\leq \frac{1}{|k|} \int_{-\infty}^{\infty} \left| \int e^{iF'(\xi) \cdot \widehat{k}s} e^{-(\omega_k + \mathcal{A}|k|)|s|} \widehat{\chi^u}(\xi, k, 0) \, d\xi \right|^2 \, ds \\ &= \frac{1}{|k|} \int_{-\infty}^{\infty} \left| \mathcal{F}^{-1}\left\{ \int e^{iF'(\xi) \cdot \widehat{k}s} e^{-(\omega_k + \mathcal{A}|k|)|s|} \widehat{\chi^u}(\xi, k, 0) \, d\xi \right\}(\tau) \right|^2 \, d\tau \\ &= \frac{4}{|k|} \int_{-\infty}^{\infty} \left| \int \frac{\mathcal{A}|k| + \omega_k}{(\mathcal{A}|k| + \omega_k)^2 + |F'(\xi) \cdot \widehat{k} + \tau|^2} \widehat{\chi^u}(\xi, k, 0) \, d\xi \right|^2 \, d\tau \\ &\leq \frac{4}{|k|} \int_{-\infty}^{\infty} \left( \int |\widehat{\chi^u}(\xi, k, 0)|^2 \frac{\mathcal{A}|k| + \omega_k}{(\mathcal{A}|k| + \omega_k)^2 + |F'(\xi) \cdot \widehat{k} + \tau|^2} \, d\xi \right) \\ &\quad \times \left( \int \frac{\mathcal{A}|k| + \omega_k}{(\mathcal{A}|k| + \omega_k)^2 + |F'(\xi) \cdot \widehat{k} + \tau|^2} \, d\xi \right) d\tau \\ &\leq \frac{4}{|k|\omega_k} \int |\widehat{\chi^u}(\xi, k, 0)|^2 \left( \int \frac{\mathcal{A}|k| + \omega_k}{(\mathcal{A}|k| + \omega_k)^2 + |F'(\xi) \cdot \widehat{k} + \tau|^2} \, d\tau \right) d\xi \\ &\quad \times \sup_{\tau} \int \frac{\omega_k(\mathcal{A}|k| + \omega_k)}{(\mathcal{A}|k| + \omega_k)^2 + |F'(\xi) \cdot \widehat{k} + \tau|^2} \, d\xi.\end{aligned}$$

Notice that the integral

$$\int \frac{\mathcal{A}|k| + \omega_k}{(\mathcal{A}|k| + \omega_k)^2 + |F'(\xi) \cdot \widehat{k} + \tau|^2} \, d\tau$$

is a constant for fixed  $\xi$  by the translation invariance of  $d\tau$ .

Now invoking (4.2) and setting  $\lambda = \frac{\omega_k}{|k|}$ , we have

$$\int_0^T |\widehat{u^0}(k, t)|^2 dt \leq \frac{C}{|k|\omega_k} \eta\left(\frac{\omega_k}{|k|}\right) \int |\widehat{\chi^u}(\xi, k, 0)|^2 d\xi$$

for some constant  $C$  depending on  $\gamma$  and  $\theta$ . That is,

$$\int_0^T |k|^{1+\kappa} \omega_k^{1-\kappa} |\widehat{u^0}(k, t)|^2 dt \leq C \int |\widehat{\chi^u}(\xi, k, 0)|^2 d\xi.$$

Since  $u_0$  has null average over  $\mathbb{T}^d$ ,

$$\begin{aligned} \widehat{u^0}(0, t) &= \int_{\mathbb{T}^d} u^0(x, t) dx = \int \widehat{\chi^u}(\xi, 0, 0) d\xi \\ &= \iint \chi^u(\xi, x, 0) d\xi dx = \int_{\mathbb{T}^d} u_0(x) dx = 0. \end{aligned} \quad (4.14)$$

Then, summing over all the discrete frequencies  $k$  with  $|k| \neq 0$ , using the Plancherel theorem again — in space this time — and noting that  $\omega_k \geq \gamma|k|^{2\alpha-1}$ , we have the estimate:

$$\int_0^T \|u\|_{H_x^{(1-\alpha)\kappa+\alpha}}^2 dt \leq C \|u_0\|_{L_x^1}. \quad (4.15)$$

## 4.2 Analysis of $u^b$

The calculation is similar:

$$\begin{aligned} &\int_0^T |\widehat{u^b}(k, t)|^2 dt \\ &= \int_0^T \left| \int_0^t \widehat{\mathcal{S}}(s)(\gamma|k|^{2\alpha} + \theta I) \widehat{\chi^u}(\xi, k, t-s) ds d\xi \right|^2 dt \\ &= \int_0^T \left| \int_0^T \mathbf{1}_{\{t-s \geq 0\}} \int e^{-(iF'(\xi) \cdot k + \omega_k|k| + \mathbf{A}(\xi):k \otimes k)s} \omega_k|k| \widehat{\chi^u}(\xi, k, t-s) d\xi ds \right|^2 dt \\ &\leq \left( \int_0^\infty \omega_k|k| e^{-\omega_k|k|s} ds \right) \\ &\quad \times \int_0^T \left( \int_0^T \left| \int e^{-(iF'(\xi) \cdot k + \frac{\omega_k|k|}{2} + \mathbf{A}(\xi):k \otimes k)s} \sqrt{\omega_k|k|} \widehat{\chi^u}(\xi, k, t) d\xi \right|^2 ds \right) dt, \end{aligned}$$

where we have used the Cauchy-Schwarz inequality and extended the domain of the inner temporal integration to  $[0, \infty)$ .

This leaves us in the exact position of equation (4.12) with an additional temporal integral in  $t$  (applied only to the kinetic function  $\widehat{\chi^u}$ ) and an additional factor of  $|k|\omega_k$ . Therefore, we can conclude as in equation (4.15) by using the zero-spatial average property (4.14) and  $\omega_k \leq (\gamma + \theta)|k|^{2\alpha-1}$  that

$$\int_0^T \|u^b(t)\|_{H_x^{(1-\alpha)\kappa}}^2 dt \leq C \int_0^T \|u(t)\|_{L_x^1} dt. \quad (4.16)$$

**Remark 4.1** Condition (4.2) is reminiscent of the nonlinearity condition given in the deterministic setting by Chen-Perthame [14]. If we discard the regularising operator  $(-\Delta)^{2\alpha}$  in (4.5), i.e., by setting  $\alpha = 0$ , then  $\beta = 1$  in (4.2). On the other hand, we can choose  $\beta$  sufficiently close to 2 so that (4.2) holds, by selecting  $\alpha$  close to  $\frac{1}{2}$ . For both cases, we are able to conclude that the  $u^b$ -part of the solution operator is compact. However, as we will see below, the regularizing effect of  $(-\Delta)^{2\alpha}$  is crucial in estimating (4.10)–(4.11) in the way as we do, via (4.19), in the next subsections. As the two terms (4.10) and (4.11) arise from the martingale and the Itô approximation, respectively, this decay requirement beyond  $o(1)$  does not appear in the deterministic setting.

### 4.3 Analysis of $M_1$

Next we turn to the analysis of the two measures  $M_1$  and  $M_2$ . For this, we follow [35] closely, since the only difference is the parabolic defect measure, which has the same sign as the kinetic dissipation measure, and the magnitude of the kinetic dissipation measure is never invoked in [35]. In this and the following sections, we repeatedly apply bound (4.19) in order to pursue the compactness estimates.

From (4.6), we see

$$\begin{aligned} & \partial_\xi(\mathcal{S}^*(t-s)h(\xi, x)) \\ &= (t-s)F''(\xi) \cdot \nabla(\mathcal{S}^*(t-s)h) + \mathbf{A}'(\xi) : \nabla^2(\mathcal{S}^*(t-s)h) + \mathcal{S}^*(t-s)\partial_\xi h. \end{aligned} \quad (4.17)$$

Then we have

$$\begin{aligned} \langle M_1, \varphi \rangle &= - \int_0^t \iint \partial_\xi(\mathcal{S}^*(s)\varphi) \, d(m^u + n^u - p^u)(\xi, x, t-s) \\ &= \int_0^t \iint (t-s)F''(\xi) \cdot \nabla(\mathcal{S}^*(t-s)\varphi) \, d(m^u + n^u - p^u)(\xi, x, t-s) \\ &\quad + \int_0^t \iint \mathbf{A}'(\xi) : \nabla^2(\mathcal{S}^*(t-s)\varphi) \, d(m^u + n^u - p^u)(\xi, x, t-s). \end{aligned} \quad (4.18)$$

Now we show the following total variation estimate.

**Lemma 4.1** *Let  $u : \mathbb{T}^d \times [0, T] \times \Omega$  be a solution with initial data  $u_0$ . Let  $\psi \in C_c(\mathbb{R})$  be any nonnegative and compactly supported continuous function, and  $\Psi = \int_0^s \int_0^T \psi(t) \, dt \, dr$ . Then*

$$\mathbb{E} \left[ \int_{\mathbb{T}^d \times [0, T] \times \mathbb{R}} \psi(\xi) \, d|m^u + n^u - p^u|(\xi, x, t) \right] \leq D_0 \mathbb{E}[\|\psi(u)\|_{L^1_{x,t}}] + \mathbb{E}[\|\Psi(u_0)\|_{L^1_x}],$$

where  $D_0 := \|\sigma^2\|_{L^\infty(\mathbb{T})}$ .

**Proof** The proof is the same as that in [35] and involves bounding  $|m^u + n^u - p^u| \leq m^u + n^u + p^u$ , so that

$$\begin{aligned} & \mathbb{E} \left[ \int_0^T \iint \psi(\xi) \, d|m^u + n^u - p^u|(\xi, x, t) \right] \\ & \leq \mathbb{E} \left[ \int_0^T \iint \psi(\xi) \, d(m^u + n^u - p^u)(\xi, x, t) \right] + 2\mathbb{E} \left[ \int_0^T \iint \psi(\xi) \, dp^u(\xi, x, t) \right] \\ & = \mathbb{E} \left[ - \int \Psi(u) \, dx \Big|_0^T \right] + \mathbb{E} \left[ \int_0^T \int \sigma^2(x) \psi(u) \, dx \, dt \right], \end{aligned}$$

by using the kinetic equation in the sense of (3.11). Now, using the non-negativity of  $\psi$ , we have

$$\mathbb{E} \left[ \int_0^T \iint \psi(\xi) \, d|m^u + n^u - p^u|(\xi, x, t) \right] \leq \mathbb{E} \left[ \int \Psi(u_0) \, dx \right] + D_0 \mathbb{E} \left[ \int_0^T \int \psi(u) \, dx \, dt \right].$$

This estimate is quite crude, as one does not take the cancellation between measures  $m^u + n^u$  and  $p^u$ , both non-negative, into account. Since there is no available way to quantify  $m^u + n^u$ , this is the best possible at the moment.

In addition to a total variation estimate, we also require the kernel estimate

$$\|(-\Delta)^{\frac{\hat{\beta}}{2}} e^{(\mathbf{A}:\nabla \otimes \nabla - \gamma(-\Delta)^\alpha)t} \|_{L_{x,\xi}^p \rightarrow L_{x,\xi}^q} \leq C(\gamma t)^{-\frac{d}{2\alpha}(\frac{1}{p}-\frac{1}{q})-\frac{\hat{\beta}}{2\alpha}}. \quad (4.19)$$

The reason for the no additional improvement over the estimate for operator  $e^{t\mathbf{A}:\nabla \otimes \nabla}$  is that we have not specified how degenerate  $\mathbf{A}$  is — it may well be simply the zero matrix. It is the use of this kernel estimate that necessitates the inclusion of the regularizations  $\gamma(-\Delta)^\alpha + \theta I$ .

By the kernel estimate (4.19), we have

$$\begin{aligned} \|(-\Delta)^{\frac{\hat{\beta}}{2}} \nabla(\mathcal{S}^*(t)\varphi) \|_{L^\infty} &\leq C(\gamma t)^{-\mu} e^{-\theta t} \|\varphi\|_{L^p}, \\ \|(-\Delta)^{\frac{\hat{\beta}}{2}} \nabla^2(\mathcal{S}^*(t)\varphi) \|_{L^\infty} &\leq C(\gamma t)^{-\mu-\frac{1}{2\alpha}} e^{-\theta t} \|\varphi\|_{L^p}, \end{aligned}$$

where  $\mu := \frac{\hat{\beta}+1}{2\alpha} + d(\frac{1}{2\alpha} - \frac{1}{2\alpha p'})$  for  $p' > 1$ , and the universal constant  $C$  is independent of  $\gamma$  and  $\theta$ .

Inserting these estimates into (4.18), we have the estimate

$$\begin{aligned} &\mathbb{E} \left[ \int_0^T \langle (-\Delta)^{\frac{\hat{\beta}}{2}} M_1, \varphi \rangle \, dt \right] \\ &= \mathbb{E} \left[ \int_0^T \int_0^t \iint (-\Delta)^{\frac{\hat{\beta}}{2}} F''(\xi) \cdot \nabla(\mathcal{S}^*(t-s)\varphi) \, d(m^u + n^u - p^u)(\xi, x, t-s) \, dt \right. \\ &\quad \left. + \int_0^T \int_0^t \iint (-\Delta)^{\frac{\hat{\beta}}{2}} \mathbf{A}'(\xi) : \nabla^2(\mathcal{S}^*(t-s)\varphi) \, d(m^u + n^u - p^u)(\xi, x, t-s) \, dt \right] \\ &\leq \mathbb{E} \left[ \int_0^T \int \|(-\Delta)^{\frac{\hat{\beta}}{2}} \nabla(\mathcal{S}^*(t-s)\varphi) \|_{L_{x,\xi}^\infty} |F''(\xi)| (t-s) \, d|m^u + n^u - p^u|(\xi, x, s) \, dt \right] \\ &\quad + \mathbb{E} \left[ \int_0^T \int \|(-\Delta)^{\frac{\hat{\beta}}{2}} \nabla^2(\mathcal{S}^*(t-s)\varphi) \|_{L_{x,\xi}^\infty} |\mathbf{A}'(\xi)| \, d|m^u + n^u - p^u|(\xi, x, s) \, dt \right]. \end{aligned}$$

By the presence of factor  $e^{-\theta(t-s)}$ , we can also bound the outer temporal integral by using the definition of the Gamma function

$$\Gamma(z) = \int_0^\infty x^{z-1} e^{-x} \, dx,$$

so that, taking  $\langle \cdot, \cdot \rangle$  as the  $L^{p'}(\mathbb{T}^d)$ – $L^p(\mathbb{T}^d)$  pairing,

$$\begin{aligned} &\mathbb{E} \left[ \int_0^T \langle (-\Delta)^{\frac{\hat{\beta}}{2}} M_1, \varphi \rangle \, dt \right] \\ &\leq \int_0^T (\gamma \tau)^{-\mu} e^{-\theta \tau} \, d\tau \, \mathbb{E} \left[ \int_{\mathbb{R} \times \mathbb{T}^d \times [0, T]} \|\varphi\|_{L_x^p} |F''(\xi)| \, d|m^u + n^u - p^u|(\xi, x, s) \right] \end{aligned}$$

$$\begin{aligned}
& + \int_0^T (\gamma\tau)^{-\mu-\frac{1}{2\alpha}} e^{-\theta\tau} d\tau \mathbb{E} \left[ \int_{\mathbb{R} \times \mathbb{T}^d \times [0, T]} \|\varphi\|_{L_x^p} |\mathbf{A}'(\xi)| d|m^u + n^u - p^u|(\xi, x, s) \right] \\
& \leq C\theta^{\mu+1}\gamma^{-\mu}|\Gamma(-\mu+1)| \mathbb{E} \left[ \int_{\mathbb{R} \times \mathbb{T}^d \times [0, T]} \|\varphi\|_{L_x^p} |F''(\xi)| d|m^u + n^u - p^u|(\xi, x, s) \right] \\
& \quad + C\theta^{\mu-1-\frac{1}{2\alpha}}\gamma^{-\mu-\frac{1}{2\alpha}} \left| \Gamma(-\mu+1-\frac{1}{2\alpha}) \right| \\
& \quad \times \mathbb{E} \left[ \int_{\mathbb{R} \times \mathbb{T}^d \times [0, T]} \|\varphi\|_{L_x^p} |\mathbf{A}'(\xi)| d|m^u + n^u - p^u|(\xi, x, s) \right].
\end{aligned}$$

By duality, the total variation estimate, and the sublinearity of  $F''$  and  $\mathbf{A}'$ , we have

$$\begin{aligned}
& \mathbb{E}[\|M_1\|_{L_t^1 W_x^{\beta, p'}}] \\
& \leq C\theta^{\mu+1}\gamma^{-\mu}|\Gamma(-\mu+1)| \mathbb{E} \left[ \int_{\mathbb{R} \times \mathbb{T}^d \times [0, T]} |F''(\xi)| d|m^u + n^u - p^u|(\xi, x, s) \right] \\
& \quad + C\theta^{\mu-1-\frac{1}{2\alpha}}\gamma^{-\mu-\frac{1}{2\alpha}} \left| \Gamma(-\mu+1-\frac{1}{2\alpha}) \right| \mathbb{E} \left[ \int_{\mathbb{R} \times \mathbb{T}^d \times [0, T]} |\mathbf{A}'(\xi)| d|m^u + n^u - p^u|(\xi, x, s) \right] \\
& \leq C \left( \theta^{\mu+1}\gamma^{-\mu}|\Gamma(-\mu+1)| + \theta^{\mu-1-\frac{1}{2\alpha}}\gamma^{-\mu-\frac{1}{2\alpha}} \left| \Gamma(-\mu+1-\frac{1}{2\alpha}) \right| \right) \\
& \quad \times \left( 1 + \int_0^T \mathbb{E}[\|u(t)\|_{L_x^1}] dt + \mathbb{E}[\|u_0\|_{L_x^3}^3] \right), \tag{4.20}
\end{aligned}$$

where we have chosen  $\gamma$  and  $\theta$  such that

$$C \left( \theta^{\mu+1}\gamma^{-\mu}|\Gamma(-\mu+1)| + \theta^{\mu-1-\frac{1}{2\alpha}}\gamma^{-\mu-\frac{1}{2\alpha}} \left| \Gamma(-\mu+1-\frac{1}{2\alpha}) \right| \right) \leq \frac{\varepsilon_0}{2}$$

for sufficiently small  $\varepsilon_0$  to be determined later.

#### 4.4 Analysis of $M_2$

$$\langle M_2, \varphi \rangle = \int_0^t \int \langle \delta(\cdot - u(x, s)), \varphi(\mathcal{S}(t-s)\sigma(x)) \rangle dx dW_s.$$

We again invoke the kernel estimate. In fact, it is here that the kernel estimate becomes indispensable. In the stochastic setting, with a forcing term given by  $\sigma(x)\delta(\xi - u(x, t))\partial_t W$ , which does not easily lend itself to the space-time Fourier transform, one may not simply take the Fourier transform on both sides so that the factor,  $i(\tau + F'(\xi) \cdot k) + \mathbf{A}(\xi) : (k \otimes k)$ , on the left side can simply be divided out, with a certain genuine nonlinearity (i.e., the non-degeneracy condition; cf. [10, 80, 97]). Thus, we have to find a different way to handle the forcing term.

Expanding the effect of the semigroup, we have

$$\langle M_2, \varphi \rangle = \int_0^t \int_{\mathbb{T}^d} e^{-\theta(t-s)} \varphi e^{-(\mathbf{A}(u):\nabla \otimes \nabla + \gamma(-\Delta)^\alpha)(t-s)} \sigma(x - F'(u(x, s))(t-s)) dx dW_s.$$

Since  $\sigma$  is bounded in  $\mathbb{T}^d$ , we see that  $\sigma(\cdot - F'(u(\cdot, s))(t-s))$  is bounded in  $x$ .

The kernel estimate then gives

$$\|e^{-\theta(t-s)} \varphi e^{-(\mathbf{A}(u):\nabla \otimes \nabla + \gamma(-\Delta)^\alpha)(t-s)} \sigma(\cdot - F'(u(\cdot, s))(t-s))\|_{H_x^{\beta}}$$

$$\leq C(\gamma(t-s))^{-\frac{\widehat{\beta}}{2\alpha}} \|\sigma\|_{L_x^2},$$

just as in [35]. In the same way, we have

$$\mathbb{E}\left[\left\|\int_0^t \langle \delta(\cdot - u(x, s)), \mathcal{S}(t-s)\sigma(x) \rangle dW_s \right\|_{H_x^{\widehat{\beta}}}^2\right] \leq C\gamma^{-\frac{\widehat{\beta}}{\alpha}} \theta^{\frac{\widehat{\beta}}{\alpha}-1} |\Gamma(1 - \frac{\widehat{\beta}}{\alpha})|.$$

Then we have

$$\mathbb{E}[\|M_2\|_{H_x^{\widehat{\beta}}}^2] \leq C\gamma^{-\frac{\widehat{\beta}}{\alpha}} \theta^{\frac{\widehat{\beta}}{\alpha}-1} |\Gamma(1 - \frac{\widehat{\beta}}{\alpha})|.$$

#### 4.5 Completion of the existence proof

From (4.15)–(4.16), we have

$$\mathbb{E}[\|u^0 + u^b + M_2\|_{L_t^2 W_x^{s,q}}^2] \leq \mathbb{E}[\|u(0)\|_{L_x^1}] + \mathbb{E}[\|u\|_{L^1([0,T], L_x^1)}] + CT,$$

where  $q > 1$  and  $s > 0$ .

By the standard Jensen and Young inequalities, we obtain

$$\frac{1}{T} \mathbb{E}^2[\|u^0 + u^b + M_2\|_{L_t^1 W_x^{s,q}}] \leq \mathbb{E}[\|u^0 + u^b + M_2\|_{L_t^2 W_x^{s,q}}^2],$$

so that

$$\mathbb{E}^2[\|u^0 + u^b + M_2\|_{L_t^1 W_x^{s,q}}] \leq CT(\mathbb{E}[\|u(0)\|_{L_x^1}] + \mathbb{E}[\|u\|_{L^1([0,T], L_x^1)}] + T).$$

Then we have

$$\mathbb{E}[\|u^0 + u^b + M_2\|_{L_t^1 W_x^{s,q}}] \leq C(\mathbb{E}[\|u(0)\|_{L_x^1}] + T) + \frac{\varepsilon_0}{2} \mathbb{E}[\|u\|_{L^1([0,T], L_x^1)}].$$

From (4.20), we further have

$$\mathbb{E}[\|M_1\|_{L^1([0,T], W_x^{\widehat{\beta}, p'})}] \leq \frac{\varepsilon_0}{2} (1 + \mathbb{E}[\|u\|_{L^1([0,T], L_x^1)}] + \mathbb{E}[\|u_0\|_{L_x^3}^3]).$$

By the continuous embedding  $W_x^{s,q} \hookrightarrow L_x^1$ ,

$$\mathbb{E}[\|u\|_{L^1([0,T], W_x^{s,q})}] \leq C(\alpha, \widehat{\beta}, \gamma, \theta) (1 + \mathbb{E}[\|u(0)\|_{L_x^3}^3] + T). \quad (4.21)$$

Since  $W^{s,q}$  is compactly embedded in  $L^1$  for  $q \geq 1$ , the Krylov-Bogoliubov mechanism (§2.2) leads to the existence of an invariant measure.

## 5 Stochastic Anisotropic Parabolic-Hyperbolic Equations II: Uniqueness of Invariant Measures

In this section, we prove the uniqueness of invariant measures for the second-order nonlinear stochastic equations (4.1).

**Theorem 5.1** *Let  $F$  and  $\mathbf{A}$  satisfy the non-degeneracy condition (4.2) and the boundedness condition:*

$$|F''(\xi)| \lesssim 1, \quad |\mathbf{A}'(\xi)| \lesssim 1. \quad (5.1)$$

*Then the invariant measure established in Theorem 4.1 is unique.*

To show the uniqueness, we first show that the solutions enter a certain ball in  $L_x^1$  in finite time almost surely. Then we show that the solutions, starting on a fixed ball, enter arbitrarily small balls almost surely, if the noise is sufficiently small in  $W_x^{1,\infty}$ . This allows us to conclude that any pair of balls enters an arbitrarily small ball of one another, since the noise is sufficiently small for any given duration with positive probability. This is the property of recurrence discussed in the coupling method in §2, which implies the uniqueness of invariant measures. In showing the recurrence, we follow §4 of [35] quite closely.

### 5.1 Uniqueness I: Finite time to enter a ball

The following lemma is proved in the same way as in [35], via a Borel-Cantelli argument.

**Lemma 5.1** *There are both a radius  $\widehat{\kappa}$  (depending on the initial conditions) and an almost surely finite stopping time  $\mathcal{T}$  such that a solution enters  $B_{\widehat{\kappa}}(0) \subseteq L^1(\mathbb{T}^d)$  in time  $\mathcal{T}$ .*

The proof uses the coupling method, where  $v$  is another solution to the same equation with initial condition  $v(0) = v_0$ . It furnishes us with the recursively defined sequence of stopping times, with  $\mathcal{T}_0 = 0$  and

$$\mathcal{T}_l = \inf\{t \geq \mathcal{T}_{l-1} + T : \|u(t)\|_{L_x^1} + \|v(t)\|_{L_x^1} \leq 2\widehat{\kappa}\}, \quad (5.2)$$

which are also almost surely finite.

### 5.2 Uniqueness II: Bounds with small noise

We now prove the following key lemma for the pathwise solutions.

**Lemma 5.2** *For any  $\varepsilon > 0$ , there are  $T > 0$  and  $\widetilde{\kappa} > 0$  such that, for the initial conditions  $u_0$  satisfying*

$$\|u_0\|_{L_x^1} \leq 2\widehat{\kappa},$$

*and the noise satisfying*

$$\sup_{t \in [0, T]} \|\sigma W\|_{W_x^{1,\infty}} \leq \widetilde{\kappa},$$

*then*

$$\oint_0^T \|u(t)\|_{L_x^1} dt \leq \varepsilon,$$

*where we have used the symbol  $\oint$  to denote the averaged integral.*

**Proof** One of the differences in our estimates from [35] is that a kernel estimate is used on  $v_F^\# + v_A^\#$ , instead of velocity averaging techniques, since the extra derivatives are required to be handled here. Of course, this method can also be applied to the first-order case so that the need to estimate the average term  $\oint v^\# dx$  in [35] can be eliminated. We divide the proof into nine steps.

(1) Let  $u$  be a solution of

$$\partial_t u + \nabla \cdot F(u) + \nabla \cdot (\mathbf{A}(u) \nabla u) = \sigma(x) \partial_t W$$

with initial condition  $u(0) = u_0$ , and let  $\tilde{u}$  be the solution to the same equation with initial condition  $\tilde{u}_0$  satisfying

$$\|u_0 - \tilde{u}_0\|_{L_x^1} \leq \frac{\varepsilon}{8}, \quad \|\tilde{u}_0\|_{L_x^2} \leq C\hat{\kappa}\varepsilon^{-\frac{d}{2}},$$

which can be found by convolving  $u_0$  with a mollifying kernel, where  $\hat{\kappa}$  is the radius constant of Lemma 5.1.

(2) Consider the difference between solution  $\tilde{u}$  and noise  $\sigma(x)W$ :  $v = \tilde{u} - \sigma(x)W$ , which is a kinetic solution to

$$\partial_t v = -\nabla \cdot F(v + \sigma(x)W) + \nabla \cdot (\mathbf{A}(v + \sigma(x)W)\nabla(v + \sigma(x)W)).$$

The kinetic formulation for this equation can be derived as in (3.6):

$$\begin{aligned} & \partial_t \chi^v + F'(\xi) \cdot \nabla \chi^v - \mathbf{A}(\xi) : \nabla^2 \chi^v \\ &= (F'(\xi) - F'(\xi + \sigma(x)W)) \cdot \nabla \chi^v - \nabla \cdot ((\mathbf{A}(\xi) - \mathbf{A}(\xi + \sigma(x)W))\nabla \chi^v) \\ & \quad - F'(\xi + \sigma(x)W)\delta(\xi - v) \cdot \nabla(\sigma(x)W) + \nabla \cdot (\mathbf{A}(\xi + \sigma(x)W)\delta(\xi - v)\nabla(\sigma(x)W)) \\ & \quad - \partial_\xi(\delta(\xi - v)\mathbf{A}(\xi + \sigma(x)W) : (\nabla(\sigma(x)W) \otimes \nabla(\sigma(x)W))) \\ & \quad + \partial_\xi(m^v + N^v). \end{aligned} \tag{5.3}$$

A notable difference here is that the parabolic defect measure  $N^v$  is not the limit of

$$\delta(\xi - (v^\varepsilon + \sigma(x)W))\mathbf{A}(\xi) : (\nabla(v^\varepsilon + \sigma W) \otimes \nabla(v^\varepsilon + \sigma(x)W)),$$

but rather the limit of

$$\begin{aligned} N_\varepsilon^u &= \delta(\xi - v^\varepsilon)\mathbf{A}(\xi + \sigma(x)W) : (\nabla v^\varepsilon \otimes \nabla v^\varepsilon) \\ & \quad + \delta(\xi - v^\varepsilon)\mathbf{A}(\xi + \sigma(x)W) : (\nabla(\sigma(x)W) \otimes \nabla(\sigma(x)W)) \\ & \quad + \delta(\xi - v^\varepsilon)\mathbf{A}(\xi + \sigma(x)W) : (\nabla v^\varepsilon \otimes \nabla(\sigma(x)W)). \end{aligned} \tag{5.4}$$

The asymmetry in the cross term in failing to contain both  $\nabla v \otimes \nabla(\sigma(x)W)$  and  $\nabla(\sigma(x)W) \otimes \nabla v$  arises from the fact that the convex entropy used is  $\Phi(v)$ , instead of  $\Phi(v + \sigma W)$ . One of the key insights in [13] is that, using the symmetry and nonnegativity of  $\mathbf{A}$ ,  $\mathbf{A}$  can be written as the square of another symmetric, positive semi-definite matrix so that (5.4) is non-negative. The limit of  $N_\varepsilon^u$  is the non-negative parabolic defect measure  $N^u$ .

(3) As before, we insert the regularizing operators:  $\gamma(-\Delta)^\alpha + \theta I$  (with fixed  $\gamma$  and  $\theta$  in this case) on both sides. Again, we can decompose the solution into the following components:

$$\langle v(t), \varphi \rangle = \langle v^0 + v^b + v_F^\sharp + v_A^\sharp + M_F + M_A + M_1 + M_2, \varphi \rangle,$$

with

$$\begin{aligned} v^0(x, t) &= \int \mathcal{S}(t)\chi^v(\xi, x, 0) \, d\xi, \\ v^b(x, t) &= \int \int_0^t \mathcal{S}(s)(\gamma(-\Delta)^\alpha + \theta I)\chi^v(\xi, x, t-s) \, ds \, d\xi, \end{aligned}$$

$$\begin{aligned}
v_F^\sharp(x, t) &= \int \int_0^t \mathcal{S}(t-s) (F'(\xi) - F'(\xi + \sigma(x)W)) \cdot \nabla \chi^v(\xi, x, s) \, ds \, d\xi, \\
v_A^\sharp(x, t) &= - \int \int_0^t \mathcal{S}(t-s) \nabla \cdot ((\mathbf{A}(\xi) - \mathbf{A}(\xi + \sigma(x)W)) \nabla \chi^v(\xi, x, s)) \, ds \, d\xi, \\
\langle M_F, \varphi \rangle &= - \int \int_0^t F'(v + \sigma(x)W) \cdot \nabla(\sigma(x)W) (\mathcal{S}^*(t-s)\varphi)(x, v(x, s)) \, ds \, dx, \\
\langle M_A, \varphi \rangle &= - \int \int_0^t \mathbf{A}(v + \sigma(x)W) : (\nabla(\sigma(x)W) \otimes \nabla(\mathcal{S}^*(t-s)\varphi)(v(x, s), x)) \, ds \, dx, \\
\langle M_1, \varphi \rangle &= - \iint \int_0^t \partial_\xi(\mathcal{S}^*(t-s)\varphi) \, d(m^v + N^v)(\xi, x, s), \\
\langle M_2, \varphi \rangle &= \int \int_0^t \partial_\xi(\mathcal{S}^*(t-s)\varphi)(v(x, s), x) \mathbf{A}(v + \sigma(x)W) : (\nabla(\sigma(x)W) \otimes \nabla(\sigma(x)W)) \, ds \, dx.
\end{aligned}$$

Now we estimate each of these integrals, with some variations from [35] especially for the terms involving  $\mathbf{A}$ . For this,  $C > 0$  is a universal constant, independent of  $\varepsilon, \tilde{\kappa}$ , and  $T$ .

(4) We first have the familiar estimates:

$$\int_0^T \|v^0(t)\|_{H_x^\alpha}^2 \, dt \leq C\gamma^r \|u_0\|_{L_x^1}$$

and

$$\int_0^T \|v^b(t)\|_{L_x^2}^2 \, dt \leq C\gamma^{r+1} \int_0^T \|v(t)\|_{L_x^1} \, dt$$

from the velocity averaging arguments, where  $|r| < 1$  (we see that there is an extra power of  $\gamma$  in the second estimate from those arguments, no matter what  $r$  might be).

These imply

$$\int_0^T \|v^0\|_{L_x^1} \, dt \leq CT^{-\frac{1}{2}} \gamma^{\frac{r}{2}} \|u_0\|_{L_x^1}^{\frac{1}{2}}, \quad (5.5)$$

$$\int_0^T \|v^b(t)\|_{L_x^1} \, dt \leq C\gamma^{\frac{r+1}{2}} \left( \int_0^T \|v(t)\|_{L_x^1} \, dt \right)^{\frac{1}{2}}. \quad (5.6)$$

(5) For  $v_F^\sharp$  and  $v_A^\sharp$ , we use the fact that

$$\begin{aligned}
& (F'(\xi) - F'(\xi + \sigma(x)W)) \cdot \nabla \chi^v(\xi, x, s) \\
&= \nabla \cdot ((F'(\xi) - F'(\xi + \sigma(x)W)) \chi^v(\xi, x, s)) - (F''(\xi + \sigma(x)W) \cdot \nabla \sigma(x)W) \chi^v(\xi, x, s), \\
& (\mathbf{A}(\xi) - \mathbf{A}(\xi + \sigma(x)W)) \nabla \chi^v(\xi, x, s) \\
&= \nabla \cdot ((\mathbf{A}(\xi) - \mathbf{A}(\xi + \sigma(x)W)) \chi^v(\xi, x, s)) - (\mathbf{A}'(\xi + \sigma(x)W) \nabla \sigma(x)W) \chi^v(\xi, x, s).
\end{aligned}$$

Now we apply the kernel estimates. Let  $\varphi \in L^2$  be any test function, and let  $\langle \cdot, \cdot \rangle$  be the pairing in  $L^2$ . Then

$$\begin{aligned}
\langle v_F^\sharp(t), \varphi \rangle &= \iint \int_0^t \varphi \mathcal{S}(t-s) \nabla \cdot ((F'(\xi) - F'(\xi + \sigma(x)W)) \chi^v(\xi, x, s)) \, ds \, d\xi \, dx \\
&\quad - \iint \int_0^t \varphi \mathcal{S}(t-s) ((F''(\xi + \sigma(x)W) \cdot \nabla \sigma(x)W) \chi^v(\xi, x, s)) \, ds \, d\xi \, dx
\end{aligned}$$

$$\begin{aligned}
&= \iint \int_0^t \nabla(\mathcal{S}^*(t-s)\varphi) \cdot (F'(\xi) - F'(\xi + \sigma(x)W)) \chi^v(\xi, x, s) \, ds \, d\xi \, dx \\
&\quad - \iint \int_0^t \mathcal{S}^*(t-s)\varphi (F''(\xi + \sigma(x)W) \cdot \nabla\sigma(x)W) \chi^v(\xi, x, s) \, ds \, d\xi \, dx. \quad (5.7)
\end{aligned}$$

Similarly, we have

$$\begin{aligned}
\langle v_A^\sharp(t), \varphi \rangle &= \iint \int_0^t \varphi \mathcal{S}(t-s) \nabla^2 : ((\mathbf{A}(\xi) - \mathbf{A}(\xi + \sigma(x)W)) \chi^v(\xi, x, s)) \, ds \, d\xi \, dx \\
&\quad - \iint \int_0^t \varphi \mathcal{S}(t-s) \nabla \cdot (\mathbf{A}'(\xi + \sigma(x)W) \nabla(\sigma(x)W) \chi^v(\xi, x, s)) \, ds \, d\xi \, dx \\
&= \iint \int_0^t \nabla^2(\mathcal{S}^*(t-s)\varphi) : (\mathbf{A}(\xi) - \mathbf{A}(\xi + \sigma(x)W)) \chi^v(\xi, x, s) \, ds \, d\xi \, dx \\
&\quad - \iint \int_0^t (\nabla(\mathcal{S}^*(t-s)\varphi) \otimes \nabla\sigma(x)W) : \mathbf{A}'(\xi + \sigma(x)W) \chi^v(\xi, x, s) \, ds \, d\xi \, dx. \quad (5.8)
\end{aligned}$$

Notice that

$$\begin{aligned}
&\int_0^T \int_0^t \|\nabla(\mathcal{S}^*(t-s)\varphi) \cdot (F'(\cdot) - F'(\cdot + \sigma(\cdot)W)) \chi^v(\cdot, \cdot, s)\|_{L_{x,\xi}^1} \, ds \, dt \\
&\leq \int_0^T \int_0^t \|\nabla(\mathcal{S}^*(t-s)\varphi)\|_{L_{x,\xi}^\infty} \|F'(\cdot) - F'(\cdot + \sigma(\cdot)W)\|_{L_{x,\xi}^\infty} \|\chi^v(\cdot, \cdot, s)\|_{L_{x,\xi}^1} \, ds \, dt \\
&\leq \int_0^T \int_0^t \|\nabla\mathcal{S}^*(t-s)\|_{L^2 \rightarrow L^\infty} \|\varphi\|_{L_x^2} \|F'(\cdot) - F'(\cdot + \sigma(\cdot)W)\|_{L_{x,\xi}^\infty} \|\chi^v(\cdot, \cdot, s)\|_{L_{x,\xi}^1} \, ds \, dt \\
&\leq C\tilde{\kappa} \|\varphi\|_{L_x^2} \sup_{s \in [0,T]} \left( \int_0^T e^{\theta(t-s)} (\gamma t)^{-\frac{d+2}{4\alpha}} \, dt \right) \int_0^T \|v(s)\|_{L_x^1} \, ds \\
&\leq C\tilde{\kappa} \|\varphi\|_{L_x^2} \gamma^{-\frac{d+2}{4\alpha}} \theta^{\frac{d+2}{4\alpha}-1} \int_0^\infty e^{-t} t^{-\frac{d+2}{4\alpha}} \, dt \int_0^T \|v(s)\|_{L_x^1} \, ds \\
&= C\tilde{\kappa} \|\varphi\|_{L_x^2} \gamma^{-\frac{d+2}{4\alpha}} \theta^{\frac{d+2}{4\alpha}-1} \left| \Gamma\left(1 - \frac{d+2}{4\alpha}\right) \right| \int_0^T \|v(s)\|_{L_x^1} \, ds; \quad (5.9)
\end{aligned}$$

$$\begin{aligned}
&\int_0^T \int_0^t \|(\mathcal{S}^*(t-s)\varphi) \nabla(\sigma(x)W) \cdot (F''(\cdot + \sigma(\cdot)W) \chi^v(\cdot, \cdot, s))\|_{L_{x,\xi}^1} \, ds \, dt \\
&\leq \int_0^T \int_0^t \|\mathcal{S}^*(t-s)\varphi\|_{L_{x,\xi}^\infty} \|F''(\cdot + \sigma(\cdot)W)\|_{L_{x,\xi}^\infty} \|\sigma W\|_{W_x^{1,\infty}} \|\chi^v(\cdot, \cdot, s)\|_{L_{x,\xi}^1} \, ds \, dt \\
&\leq C\tilde{\kappa} \|\varphi\|_{L_x^2} \gamma^{-\frac{d}{4\alpha}} \theta^{\frac{d}{4\alpha}-1} \left| \Gamma\left(1 - \frac{d}{4\alpha}\right) \right| \int_0^T \|v(s)\|_{L_x^1} \, ds; \quad (5.10)
\end{aligned}$$

$$\begin{aligned}
&\int_0^T \int_0^t \|\nabla^2(\mathcal{S}^*(t-s)\varphi) : (\mathbf{A}(\cdot) - \mathbf{A}(\cdot + \sigma(\cdot)W)) \chi^v(\cdot, \cdot, s)\|_{L_{x,\xi}^1} \, ds \, dt \\
&\leq \int_0^T \int_0^t \|\nabla^2(\mathcal{S}^*(t-s)\varphi)\|_{L_{x,\xi}^\infty} \|\mathbf{A}(\cdot) - \mathbf{A}(\cdot + \sigma(\cdot)W)\|_{L_{x,\xi}^\infty} \|\chi^v(\cdot, \cdot, s)\|_{L_{x,\xi}^1} \, ds \, dt \\
&\leq C\tilde{\kappa} \|\varphi\|_{L_x^2} \gamma^{-\frac{d+4}{4\alpha}} \theta^{\frac{d+4}{4\alpha}-1} \left| \Gamma\left(1 - \frac{d+4}{4\alpha}\right) \right| \int_0^T \|v(s)\|_{L_x^1} \, ds; \quad (5.11)
\end{aligned}$$

$$\begin{aligned}
& \int_0^T \int_0^t \|(\nabla(\mathcal{S}^*(t-s)\varphi) \otimes \nabla(\sigma(\cdot)W)) : \mathbf{A}'(\cdot + \sigma(\cdot)W) \chi^v(\cdot, \cdot, s)\|_{L_{x,\xi}^1} \, ds \, dt \\
& \leq \int_0^T \int_0^t \|\nabla(\mathcal{S}^*(t-s)\varphi)\|_{L_{x,\xi}^\infty} \|\sigma W\|_{W_x^{1,\infty}} \|\mathbf{A}'(\cdot + \sigma(\cdot)W)\|_{L_{x,\xi}^\infty} \|\chi^v(\cdot, \cdot, s)\|_{L_{x,\xi}^1} \, ds \, dt \\
& \leq C\tilde{\kappa} \|\varphi\|_{L_x^2} \gamma^{-\frac{d+2}{4\alpha}} \theta^{\frac{d+2}{4\alpha}-1} |\Gamma(1 - \frac{d+2}{4\alpha})| \int_0^T \|v(s)\|_{L_x^1} \, ds. \tag{5.12}
\end{aligned}$$

Now, by (5.1), we have assumed that

$$|F''(\xi)| \lesssim 1, \quad |\mathbf{A}'(\xi)| \lesssim 1,$$

and  $\|\sigma(x)W\|_{W_x^{1,\infty}} \leq \tilde{\kappa}$ , so that we can use the estimates (the second from the first by the Poincaré-Wirtinger inequality, since  $\int_{\mathbb{T}^d} \sigma(x) \, dx = 0$ ):

$$\begin{aligned}
|F'(\xi) - F'(\xi + \sigma(x)W)| + |F''(\xi + \sigma(x)W) \cdot \nabla(\sigma W)| &\leq C\tilde{\kappa}, \\
|\mathbf{A}(\xi) - \mathbf{A}(\xi + \sigma(x)W)| + |\mathbf{A}'(\xi + \sigma(x)W) \nabla(\sigma W)| &\leq C\tilde{\kappa}.
\end{aligned}$$

Putting these estimate (5.7)–(5.12) back into the bound:  $\|v^\sharp(t)\|_{L_x^2} = \sup_{\|\varphi\|_{L_x^2}=1} \langle v^\sharp(t), \varphi \rangle$ , we have

$$\begin{aligned}
& \int_0^T \|v_A^\sharp(t) + v_F^\sharp(t)\|_{L_x^2} \, dt \\
& \leq C\tilde{\kappa} \left( \gamma^{-\frac{d+2}{4\alpha}} \theta^{\frac{d+2}{4\alpha}-1} |\Gamma(1 - \frac{d+2}{4\alpha})| + \gamma^{-\frac{d}{4\alpha}} \theta^{\frac{d}{4\alpha}-1} |\Gamma(1 - \frac{d}{4\alpha})| \right. \\
& \quad \left. + \gamma^{-\frac{d+4}{4\alpha}} \theta^{\frac{d+4}{4\alpha}-1} |\Gamma(1 - \frac{d+4}{4\alpha})| \right) \int_0^T \|v(t)\|_{L_x^1} \, dt. \tag{5.13}
\end{aligned}$$

(6) For  $M_F$  and  $M_A$ , we employ the kernel estimate and  $\|\sigma W\|_{W_x^{1,\infty}} \leq \tilde{\kappa}$  to obtain

$$\begin{aligned}
|\langle M_F, \varphi \rangle| &\leq \int_0^t \|F'(v + \sigma W)\|_{L_x^1} \|\mathcal{S}\varphi\|_{L_x^\infty} \|\nabla \sigma W\|_{L_x^\infty} \, ds, \\
|\langle M_A, \varphi \rangle| &\leq \int_0^t \|\mathbf{A}(v + \sigma W)\|_{L_x^1} \|\nabla(\mathcal{S}\varphi)\|_{L_x^\infty} \|\nabla \sigma W\|_{L_x^\infty} \, ds.
\end{aligned}$$

Now, by (4.3), we have

$$\|F'(v + \sigma W)\|_{L_x^1} + \|\mathbf{A}(v + \sigma W)\|_{L_x^1} \leq C(1 + \|v(t)\|_{L_x^1} + \|\sigma\|_{L_x^1} |W|).$$

These give

$$\begin{aligned}
& \int_0^T \|M_F(t) + M_A(t)\|_{L_x^1} \, dt \\
& \leq C\tilde{\kappa} \int_0^T \int_0^t (1 + \|v(s)\|_{L_x^1}) e^{-\theta(t-s)} (1 + (\gamma(t-s))^{-\frac{1}{2\alpha}}) \, ds \, dt \\
& \leq C\tilde{\kappa} \left( \theta^{-1} + \gamma^{-\frac{1}{2\alpha}} \theta^{\frac{1}{2\alpha}-1} |\Gamma(1 - \frac{1}{2\alpha})| \right) \int_0^T (1 + \|v(s)\|_{L_x^1}) \, ds. \tag{5.14}
\end{aligned}$$

(7) For  $M_2$ , we have

$$\langle M_2, \varphi \rangle = \int_0^t \int_0^s \partial_\xi (\mathcal{S}^*(t-s)\varphi)(v(x, s), x) \mathbf{A}(v + \sigma(x)W) : (\nabla(\sigma(x)W) \otimes \nabla(\sigma(x)W)) ds dx.$$

We notice that

$$\partial_\xi (\mathcal{S}(t-s)\varphi)(v(x, s), x) = (t-s)F''(v(x, s)) \cdot \nabla(\mathcal{S}^*(t-s)\varphi) + \mathbf{A}'(v(x, s)) : \nabla^2(\mathcal{S}^*(t-s)\varphi),$$

as explained in (4.17). By (5.1), we have assumed that

$$|F''(\xi)| \lesssim 1, \quad |\mathbf{A}'(\xi)| \lesssim 1.$$

Again we have

$$\|\mathbf{A}(v + \sigma(x)W)\|_{L_x^1} \leq C(1 + \|v(s)\|_{L_x^1} + \|\sigma\|_{L_x^1}|W|).$$

Finally, using the kernel estimate yields

$$\begin{aligned} & |\langle M_2, \varphi \rangle| \\ & \leq C \int_0^t (t-s) \|F''\|_{L^\infty} \|\nabla \mathcal{S}^*(t-s)\varphi\|_{L_{x,\xi}^\infty} \|\nabla \sigma W\|_{L_x^\infty}^2 (1 + \|v(s)\|_{L_x^1} + \|\sigma\|_{L_x^1}|W|) ds \\ & \quad + C \int_0^t (t-s) \|\mathbf{A}'\|_{L^\infty} \|\nabla^2 \mathcal{S}^*(t-s)\varphi\|_{L_{x,\xi}^\infty} \|\nabla \sigma W\|_{L_x^\infty}^2 (1 + \|v(s)\|_{L_x^1} + \|\sigma\|_{L_x^1}|W|) ds \\ & \leq C\tilde{\kappa}^2 \int_0^t (t-s) (\|\nabla \mathcal{S}^*(t-s)\| + \|\nabla^2 \mathcal{S}^*(t-s)\|) \|\varphi\|_{L_x^\infty} (1 + \|v(s)\|_{L_x^1} + \|\sigma\|_{L_x^1}|W|) ds. \end{aligned}$$

Therefore, we have

$$\begin{aligned} & \int_0^T \|M_2(t)\|_{L_x^1} dt \\ & \leq C\tilde{\kappa}^2 \int_0^T \int_0^t (t-s) (\|\nabla \mathcal{S}^*(t-s)\| + \|\nabla^2 \mathcal{S}^*(t-s)\|) (1 + \|v(s)\|_{L_x^1} + \|\sigma\|_{L_x^1}|W|) ds dt \\ & \leq C\tilde{\kappa}^2 \int_0^T \int_0^t (t-s) ((\gamma(t-s))^{-\frac{1}{2\alpha}} + (\gamma(t-s))^{-\frac{1}{\alpha}}) e^{-\theta(t-s)} (1 + \|v(s)\|_{L_x^1} + \|\sigma\|_{L_x^1}|W|) ds dt, \end{aligned}$$

and

$$\begin{aligned} & \int_0^T \|M_2(t)\|_{L_x^1} dt \\ & \leq C\tilde{\kappa}^2 \left( \gamma^{-\frac{1}{2\alpha}} \theta^{\frac{1}{2\alpha}-2} \left| \Gamma(2 - \frac{1}{2\alpha}) \right| + \gamma^{-\frac{1}{\alpha}} \theta^{\frac{1}{\alpha}-2} \left| \Gamma(2 - \frac{1}{\alpha}) \right| \right) \int_0^T (1 + \|v(s)\|_{L_x^1}) ds. \end{aligned} \quad (5.15)$$

(8) For the kinetic measure  $M_1$ , we use the total variation estimate again. First, with  $\varphi \in L_x^\infty$ ,

$$\begin{aligned} & |\langle M_1, \varphi \rangle| \\ & = \left| \iint \int_0^t \partial_\xi (\mathcal{S}^*(t-s)\varphi) d(m^v + N^v)(x, \xi, s) \right| \end{aligned}$$

$$\begin{aligned}
&= \left| \iint \int_0^t ((t-s)F''(\xi) \cdot \nabla(\mathcal{S}^*(t-s)\varphi) + \mathbf{A}'(\xi) : \nabla^2(\mathcal{S}^*(t-s)\varphi)) \, d(m^v + N^v)(x, \xi, s) \right| \\
&\leq C\|\varphi\|_{L_x^\infty} \iint \int_0^t (\gamma^{-\frac{1}{2\alpha}}(t-s)^{1-\frac{1}{2\alpha}} + \gamma^{-\frac{1}{\alpha}}(t-s)^{1-\frac{1}{\alpha}}) e^{-\theta(t-s)} d|m^v + N^v|(x, \xi, s),
\end{aligned}$$

so that

$$\begin{aligned}
&\int_0^T \|M_1(t)\|_{L_x^1} \, dt \\
&\leq C \left( \gamma^{-\frac{1}{2\alpha}} \theta^{\frac{1}{2\alpha}-2} \left| \Gamma(2 - \frac{1}{2\alpha}) \right| + \gamma^{-\frac{1}{\alpha}} \theta^{\frac{1}{\alpha}-2} \left| \Gamma(2 - \frac{1}{\alpha}) \right| \right) |m^v + N^v|(\mathbb{R} \times \mathbb{T}^d \times [0, T]).
\end{aligned}$$

As in Lemma 4.1, we test equation (5.3) against  $\xi$  to find

$$\begin{aligned}
&\frac{1}{2} \|v(t)\|_{L_x^2}^2 + |m^v + N^v|(\mathbb{R} \times \mathbb{T}^d \times [0, t]) \\
&\leq \frac{1}{2} \|\tilde{u}_0\|_{L_x^2}^2 + \left| \int_0^t \iint \xi (F'(\xi) - F'(\xi + \sigma(x)W)) \cdot \nabla \chi^v \, d\xi \, dx \, ds \right| \\
&\quad + \left| \int_0^t \int v F'(v + \sigma(x)W) \cdot \nabla(\sigma(x)W) \, dx \, ds \right| \\
&\quad + \left| \int_0^t \int \mathbf{A}(v + \sigma(x)W) : (\nabla(\sigma W) \otimes \nabla(\sigma W)) \, dx \, ds \right| \\
&\leq \frac{1}{2} \|\tilde{u}_0\|_{L_x^2}^2 + \left| \int_0^t \iint \xi F''(\xi + \sigma(x)W) \cdot \nabla(\sigma(x)W) \chi^v \, d\xi \, dx \, ds \right| \\
&\quad + C\tilde{\kappa}(1 + \tilde{\kappa}) \left| \int_0^t \int v(1 + |v| + |\sigma(x)W|) \, dx \, ds \right| \\
&\leq \frac{1}{2} \|\tilde{u}_0\|_{L_x^2}^2 + C\tilde{\kappa} \int_0^t (1 + \|v(s)\|_{L_x^2}^2) \, ds.
\end{aligned}$$

Then Gronwall's inequality implies

$$|m^v + N^v| \leq C e^{C\tilde{\kappa}t} (\|\tilde{u}_0\|_{L_x^2}^2 + 1) \leq C e^{C\tilde{\kappa}t} (\tilde{\kappa}^2 \varepsilon^{-d} + 1).$$

Therefore, we have

$$\begin{aligned}
&\int_0^T \|M_1(t)\|_{L_x^1} \, dt \\
&\leq C \left( \gamma^{-\frac{1}{2\alpha}} \theta^{\frac{1}{2\alpha}-2} \left| \Gamma(2 - \frac{1}{2\alpha}) \right| + \gamma^{-\frac{1}{\alpha}} \theta^{\frac{1}{\alpha}-2} \left| \Gamma(2 - \frac{1}{\alpha}) \right| \right) e^{C\tilde{\kappa}T} (\tilde{\kappa}^2 \varepsilon^{-d} + 1). \quad (5.16)
\end{aligned}$$

(9) Completion of the estimates. First, we set  $\alpha \leq \frac{1}{2}$  so that the instances of  $|\Gamma|$  are never evaluated at a negative integer, where it is infinite. With the finite bound of all the values of  $|\Gamma|$  and finitely many instances of  $\Gamma$  in estimates (5.5)–(5.16) above, we can write those estimates as

$$\begin{aligned}
&\int_0^T \|v^0(t)\|_{L_x^1} \, dt \leq CT^{-\frac{1}{2}} \gamma^{\frac{r}{2}} \|u_0\|_{L_x^1}^{\frac{1}{2}}, \\
&\int_0^T \|v^b(t)\|_{L_x^1} \, dt \leq C\gamma^{\frac{r+1}{2}} \left( \int_0^T \|v\|_{L_x^1} \, dt \right)^{\frac{1}{2}} \leq C\gamma^{\frac{r+1}{2}} \int_0^T (1 + \|v\|_{L_x^1}) \, dt,
\end{aligned}$$

$$\begin{aligned}
\int_0^T \|v_F^\# + v_A^\#\|_{L_x^1} dt &\leq C\tilde{\kappa}(\gamma^{-\frac{d+2}{4\alpha}}\theta^{\frac{d+2}{4\alpha}-1} + \gamma^{-\frac{d}{4\alpha}}\theta^{\frac{d}{4\alpha}-1} + \gamma^{-\frac{d+4}{4\alpha}}\theta^{\frac{d+4}{4\alpha}-1}) \int_0^T \|v(t)\|_{L_x^1} dt, \\
\int_0^T \|M_F(t) + M_A(t)\|_{L_x^1} dt &\leq C\tilde{\kappa}(\theta^{-1} + \gamma^{-\frac{1}{2\alpha}}\theta^{\frac{1}{2\alpha}-1}) \int_0^T (1 + \|v(t)\|_{L_x^1}) dt, \\
\int_0^T \|M_2(t)\|_{L_x^1} dt &\leq C\tilde{\kappa}^2(\gamma^{-\frac{1}{2\alpha}}\theta^{\frac{1}{2\alpha}-2} + \gamma^{-\frac{1}{\alpha}}\theta^{\frac{1}{\alpha}-2}) \int_0^T (1 + \|v(t)\|_{L_x^1}) dt, \\
\int_0^T \|M_1\|_{L_x^1} dt &\leq \frac{C}{T}(\gamma^{-\frac{1}{2\alpha}}\theta^{\frac{1}{2\alpha}-2} + \gamma^{-\frac{1}{\alpha}}\theta^{\frac{1}{\alpha}-2})e^{C_0\tilde{\kappa}T}(\tilde{\kappa}^2\varepsilon^{-d} + 1).
\end{aligned}$$

Combining all the estimates together yields

$$\begin{aligned}
\int_0^T \|v(t)\|_{L_x^1} dt &\leq C_1 T^{-\frac{1}{2}}\gamma^{\frac{1}{2}}\|u_0\|_{L_x^1}^{\frac{1}{2}} + (C_2\gamma^{\frac{r+1}{2}} + C_3(\gamma, \theta)(\tilde{\kappa} + \tilde{\kappa}^2))\left(1 + \int_0^T \|v\|_{L_x^1} ds\right) \\
&\quad + C_4(\gamma, \theta)T^{-1}e^{c\tilde{\kappa}T}(\tilde{\kappa}^2\varepsilon^{-d} + 1).
\end{aligned}$$

We can choose  $\gamma$ ,  $\theta$ ,  $T$  and  $\tilde{\kappa}$  in that order so that, for some  $q$  to be determined,

$$C_2\gamma^{\frac{r+1}{2}} \leq q\varepsilon.$$

For  $\alpha < \frac{1}{4}$ , we see that every  $\theta$  has positive power above, except in  $C_3(\gamma, \theta)$  for the estimate of  $\|M_F + M_A\|_{L_{x,t}^1}$ , so that  $C(\rho, \theta)$  involves  $\theta$  with positive power. Therefore, we choose  $\theta$  such that

$$C_4(\gamma, \theta) < 1,$$

so that we can choose  $T$  sufficiently large such that

$$C_1 T^{-\frac{1}{2}}\|u_0\|_{L_x^1}^{\frac{1}{2}} + C_4(\gamma, \theta)T^{-1}(\tilde{\kappa}^2\varepsilon^{-d} + 1) \leq q\varepsilon.$$

Finally, we choose  $\tilde{\kappa}$  such that

$$C_3(\gamma, \theta)\tilde{\kappa}(1 + \tilde{\kappa}) \leq q\varepsilon,$$

and

$$C_0\tilde{\kappa}T \leq q\varepsilon.$$

Taking  $q$  sufficiently small, we have

$$\int_0^T \|v(t)\|_{L_x^1} dt \leq \frac{\varepsilon}{4}, \quad \int_0^T \|\tilde{u}(t)\|_{L_x^1} dt \leq \frac{3\varepsilon}{8},$$

which leads to

$$\int_0^T \|u(t)\|_{L_x^1} dt \leq \frac{\varepsilon}{2}$$

by the following  $L^1$ -contraction property of the pathwise solutions.

**Lemma 5.3** *Let  $\mathbf{A}$  be symmetric positive-semi-definite, and let both  $\mathbf{A}(\xi)$  and  $F(\xi)$  be Hölder continuous and of polynomial growth. Then, for each initial data function  $u_0$ , there exists a unique measurable  $u : \mathbb{T}^d \times [0, T] \times \Omega \rightarrow \mathbb{R}$  solving (3.5) in the sense of Definition 3.1. Moreover, for  $u_0, \tilde{u}_0 \in L^1(\mathbb{T}^d)$ ,*

$$\|u(t) - \tilde{u}(t)\|_{L_x^1} \leq \|u_0 - \tilde{u}_0\|_{L_x^1} \quad \text{almost surely.}$$

For our case,

$$\|u(t) - \tilde{u}(t)\|_{L_x^1} \leq \|u_0 - \tilde{u}_0\|_{L_x^1} \leq \frac{\varepsilon}{8} \quad \text{almost surely.}$$

This completes the proof.

**Remark 5.1** This almost sure  $L^1$ -contraction property is not available in the multiplicative case. In fact, it is not available in many other situations, such as in systems or for non-conservative equations where the  $L^1$ -contraction is not ready to provide a stability condition. It is of interest to study the uniqueness and ergodicity properties of invariant measures for equations without this property.

### 5.3 Uniqueness III: Conclusion

Let  $u_0^1$  and  $u_0^2$  be in  $L_x^1$ . For a given  $\varepsilon > 0$ , let  $\tilde{u}_0^1$  and  $\tilde{u}_0^2$  be in  $L_x^3$  such that

$$\|u_0^i - \tilde{u}_0^i\|_{L_x^1} \leq \frac{\varepsilon}{4}.$$

Denote their corresponding solutions  $u^1, u^2, \tilde{u}^1$  and  $\tilde{u}^2$ , respectively. Let us now put  $\tilde{u}^1$  and  $\tilde{u}^2$ , in place of  $u$  and  $v$  in §5.1, and the corresponding sequence of stopping times constructed recursively in (5.2):

$$\mathcal{T}_l = \inf\{t \geq \mathcal{T}_{l-1} + T : \|\tilde{u}^1(t)\|_{L_x^1} + \|\tilde{u}^2(t)\|_{L_x^1} \leq 2\tilde{\kappa}\}.$$

As in [35], choosing  $T$  and  $\tilde{\kappa}$  as above, we obtain by the  $L^1$ -contraction (for the additive noise, there is the  $L^1$ -contraction almost sure):

$$\begin{aligned} & \mathbb{P}\left\{\int_{\mathcal{T}_l}^{\mathcal{T}_l+T} \|u^1(s) - u^2(s)\|_{L_x^1} ds \leq \varepsilon \mid \mathcal{F}_{\mathcal{T}_l}\right\} \\ & \geq \mathbb{P}\left\{\int_{\mathcal{T}_l}^{\mathcal{T}_l+T} \|\tilde{u}^1(s) - \tilde{u}^2(s)\|_{L_x^1} ds \leq \frac{\varepsilon}{2} \mid \mathcal{F}_{\mathcal{T}_l}\right\} \\ & \geq \mathbb{P}\left\{\sup_{t \in [\mathcal{T}_l, \mathcal{T}_l+T]} \|\sigma W(t) - \sigma W(\mathcal{T}_l)\|_{W_x^{1,\infty}} \leq \tilde{\kappa} \mid \mathcal{F}_{\mathcal{T}_l}\right\}. \end{aligned}$$

Since  $\tilde{\kappa} > 0$ , and  $\sigma$  is Lipschitz, we can denote the positive probability of the event as  $\lambda$ . By the strong Markov property, we know that it does not change with  $l$ .

This allows us to write

$$\mathbb{P}\left\{\int_{\mathcal{T}_l}^{\mathcal{T}_l+T} \|u^1(s) - u^2(s)\|_{L_x^1} ds \geq \varepsilon \quad \text{for } l = l_0, l_0 + 1, \dots, l_0 + k\right\} \leq (1 - \lambda)^k,$$

so that

$$\begin{aligned} & \mathbb{P}\left\{\lim_{l \rightarrow \infty} \int_{\mathcal{T}_l}^{\mathcal{T}_l+T} \|u^1(s) - u^2(s)\|_{L_x^1} ds \geq \varepsilon\right\} \\ & = \mathbb{P}\left\{\exists l_0, \forall l \geq l_0 : \int_{\mathcal{T}_l}^{\mathcal{T}_l+T} \|u^1(s) - u^2(s)\|_{L_x^1} ds \geq \varepsilon\right\} \\ & = 0. \end{aligned}$$

This limit exists as  $s \mapsto \|u^1(s) - u^2(s)\|_{L_x^1}$  is non-increasing, by the  $L^1$ -contraction property. Then, by the same property,

$$\mathbb{P}\left\{\lim_{t \rightarrow \infty} \|u^1(t) - u^2(t)\|_{L_x^1} \geq \varepsilon\right\} = 0.$$

Therefore, almost surely,

$$\lim_{t \rightarrow \infty} \|u^1(t) - u^2(t)\|_{L_x^1} = 0,$$

which implies the uniqueness of the invariant measure.

## 6 Further Developments, Problems and Challenges

In this section, we discuss some further developments, problems and challenges in this direction.

### 6.1 Further problems

There are several natural problems that follow from the analysis discussed above. We restrict ourselves again to nonlinear conservation laws driven by stochastic forcing.

One of the problems is the long-time behavior problem for solutions of nonlinear conservation laws driven by multiplicative noises. The noises of form  $\nabla \cdot (F(u) \circ dW)$  have been considered in [59–60], in which the dynamics remains in the zero-spatial-average subspace of  $L^1(\mathbb{T}^d)$ .

The well-posedness for nonlinear conservation laws driven by multiplicative noises is quite well understood from several different perspectives—the strong entropy stochastic solutions of Feng-Nualart [49] and of Chen-Ding-Karlsen [9], the viscosity solution methods of Bauzet-Vallet-Wittbold [3], and the kinetic approach of Debussche-Hofmanová-Vovelle [33–34], as we have mentioned above. Nevertheless, the long-time behavior problem for solutions is wide open, since there is no effective way to control  $\|u(t)\|_{L_x^1}$ .

We remark on two aspects of the noises that can affect qualitative long-time behaviors of solutions.

(i) The question seems to depend heavily on the roots and growth of the noise coefficient function  $\sigma(u)$ —if the noise is degenerate (not cylindrical), say  $\sigma(u) = 0$  for certain  $u = r \in \mathbb{R}$ , then  $u \equiv r$  is a fixed point of the evolution. By the  $L^1$ -contraction, it is possible to prove certain long-time behavior results for the solutions for the unbounded noise coefficient function with one root. Both the growth of  $\sigma$  and how many roots it possesses affect the long-time behaviors of solutions, as is evident also in the analysis of other equations such as the KPP equation (discussed below). In the case that  $\sigma$  has no roots, there are no fixed points. It is possible that the nonlinear conservation laws driven by bounded noises with no roots have non-trivial invariant measures.

(ii) If the noise is  $\sigma(u)dB = \sum_k g_k(u)dW^k$ , where  $B = W^k e_k$  is a cylindrical Wiener process, the behaviors are expected to be very different from the case that the noise is simply  $\sigma(u)dW$ .

Another natural direction to consider is the case of nonlinear systems of balance laws. For this case, such as for the isentropic Euler system, the kinetic formulation is not “pure”—it contains the instances of the solution mixed with the kinetic operator (cf. [81]). At present, it seems that the methods discussed above are not directly applicable to the systems.

### 6.2 The Navier-Stokes equations

The two-dimensional incompressible Navier-Stokes equations (INSEs for short) driven by stochastic forcing has been a subject of intense interest. We focus on the analysis of asymptotic behaviors of solutions to keep ourselves from getting sidetracked.

The existence of invariant measures for the 2-D INSEs on a regular bounded domain with a general noise that is a Gaussian random field and white-in-time has been known at least since [51]. The uniqueness and ergodicity for the 2-D INSEs have also been established; see [55] and the references therein for such results and further existence results of invariant measures under different conditions. These results have subsequently been improved, including for the noises that are localized in time and Gaussian in space, in [7–8, 87–88], and in some references cited in this paper.

We remark particularly that the corresponding existence questions for the 2-D INSEs with multiplicative noises have been established, for example in [53], via the Skorohod embedding and a Faedo-Galerkin procedure, which have shown the existence of martingale solutions and stationary martingale solutions, from which in turn the existence of an invariant measure can be derived.

The asymptotic behaviors of 2-D INSEs driven by white-in-time noises or Poisson distributed unbounded kick noises have been explored, and the existence and uniqueness of invariant measures for these systems are known. See also [72–73, 89] for the related references.

There are also more recent results on INSEs driven by space-time white noises in 2-D or 3-D; see [26–27, 102] and the references therein. For example, it is known that the transition semigroup of the Kolmogorov equation associated to the 3-D stochastic INSEs driven by a cylindrical white noise has a unique (and hence ergodic) invariant measure.

The existence of invariant measures for the compressible Navier-Stokes equations, even in the 2-D case, is wide open.

See [46, 83], and discussions in [82, Chp. 3] as well as references contained there for further treatments on ergodicity results; also see [84].

### 6.3 The asymptotic strong Feller property

Using the 2-D INSEs as a springboard, the notion of the asymptotic strong Feller property has been introduced in Hairer-Mattingly [63] as a weaker and more natural replacement of the sufficient “strong Feller” property (Definition 2.2) in dissipative infinite-dimensional systems, the possession of which guarantees the uniqueness of an invariant measure. In the finite-dimensional case of SDEs, there is a related notion of eventual strong Feller property, for which sufficient conditions are given in [5].

The definition of asymptotic strong Feller property depends on a preliminary definition.

**Definition 6.1** (Totally Separating System) *A pseudo-metric is a function  $d : \mathfrak{X}^2 \rightarrow \mathbb{R}_0^+$  for which  $d(x, x) = 0$  and the triangle inequality is satisfied, and  $d_1 \geq d_2$  if the inequality holds for all arguments  $(x, y) \in \mathfrak{X}^2$ . Let  $\{d_k\}_{k=0}^\infty$  be an increasing sequence of pseudo-metrics on a Polish space  $\mathfrak{X}$ . Then  $\{d_k\}_{k=0}^\infty$  is a totally separating system of pseudo-metrics if*

$$\lim_{k \rightarrow \infty} d_k = 1$$

*pointwise everywhere off the diagonal on  $\mathfrak{X}^2$ .*

Then the asymptotic strong Feller property is defined as follows.

**Definition 6.2** (Asymptotic Strong Feller Property) *A Markov transition semigroup  $\mathcal{P}_t$  on a Polish space  $X$  is asymptotically strong Feller at  $x$  if there exist both a totally separating*

system of pseudo-metrics  $\{d_k\}_{k=0}^\infty$  and an increasing sequence of times  $t_k$  such that

$$\inf_{U \in \text{nb}(x)} \limsup_{k \rightarrow \infty} \sup_{y \in U} \|P_{t_k}(x, \cdot) - P_{t_k}(y, \cdot)\|_{d_k} = 0,$$

where  $\text{nb}(x)$  is the collection of open sets containing  $x$ ,  $P$  is the transition probabilities associated to  $\mathcal{P}$ , and  $\|P_1 - P_2\|_{d_k}$  is the norm given by

$$\|P_1 - P_2\|_{d_k} = \inf \int_{\mathfrak{X}^2} d_k(w, z) \Pi(dw, dz),$$

the infimum being taken over all positive measures on  $\mathfrak{X}^2$  with marginals  $P_1$  and  $P_2$ .

The idea behind the asymptotic strong Feller condition is that ergodicity is preserved even if the stochastic forcing is restricted to a few unstable modes, and dissipated in the others. Using this idea, the ergodicity of the 2-D stochastic INSE with degenerate noise has been established (cf. [63]). Some results of ergodicity for the 3-D INSEs driven by mildly degenerate noise relying on the strong asymptotic Feller property have also established (cf. [93–94] and the references cited therein).

#### 6.4 The KPP equation and multiplicative noises

The Kolmogorov-Petrovsky-Piskunov equation (KPP for short) is given by

$$\begin{aligned} \partial_t u &= \nabla \cdot (\mathbf{A}(x, t) \nabla u) + h(u) + g(u) \partial_t W, \\ u|_{t=0} &= \varphi. \end{aligned}$$

Attention is often restricted to the case in which  $g$  and  $h$  both vanish at the two points  $a, b \in \mathbb{R}$  and  $g, h > 0$  on  $(a, b)$ . In this way, the asymptotic size is controlled in  $L^1$ .

It has been shown in Chueshov-Vuillermot [15–20] that, for the semilinear equation with  $h(u) = sg(u)$ , evolution on a bounded, open domain with zero Neumann boundary condition is bounded in space. Moreover, the notion of stability in probability has also been introduced in [18].

**Definition 6.3** (Stability in Probability) *A function  $u_f$  is stable in probability if, for every  $\varepsilon > 0$ , the following relation holds:*

$$\lim_{\|\varphi - u_f\|_{L^\infty} \rightarrow 0} \mathbb{P} \left\{ \omega \in \Omega : \sup_{t \in \mathbb{R}^+ \setminus \{0\}} \|u_\varphi(s, \cdot, t, \omega) - u_f\|_{L^\infty} > \varepsilon \right\} = 0.$$

Otherwise,  $u_f$  is unstable in probability.

A function  $u_f$  is globally asymptotically stable in probability if it is stable in probability and

$$\mathbb{P} \left\{ \omega \in \Omega : \lim_{t \rightarrow \infty} \|u_\varphi(s, \cdot, t, \omega) - u_f\|_{L^\infty} = 0 \right\} = 1.$$

By considering the moments of the spatial average, it has been shown that the constant functions  $u_1 = a$  and  $u_2 = b$  are fixed points whose stability in probability depends on the values of  $s$ . The results of [20] have been refined, say in [4], and the properties of the global attractor, including the computation of exact Lyapunov exponents in a decay scenario have been derived.

As we have remarked, the main reason that multiplicative noises complicate the analysis of stochastic PDEs is that one fails to have much control over the spatial average, except when additional restrictions on the noise and initial conditions are specified. When the noise has a root, that constant is immediately a fixed point. This cannot be avoided even when working over the non-compact domain  $\mathbb{R}$  because the  $L^p$  boundedness often relies on the space that is compact, and is a difficulty we have to overcome in order to gain a deeper understanding of the asymptotic behaviors of solutions.

## 6.5 Large deviation principles

Beyond the existence and uniqueness of invariant measures, large deviation principles touch on their specific properties. Whilst it goes some way outside the scope of this survey even to introduce the theory of large deviations, which attempts to characterize the limiting behavior of a family of probability measures (in our case, invariant measures) depending on some parameter by using a rate function, we should be remiss to neglect mentioning it altogether; two vintage references to the subject are [38, 58]. More modern treatments can be found in [36, 43, 68, 99] and the references cited therein. Of particular interest has been the “zero-noise” limit of stochastic equations in which one looks at the stochastic equations with a small parameter  $\varepsilon$  multiplied to the noise. Questions of large deviation type also arise in stochastic homogenization theory. Each of these subjects can justify an independent survey. Pertaining specifically to stochastic conservation laws, the literature is, however, more sparse. Going some way outside the classical Freidlin-Wentzell theory, some results have been announced pertaining to large deviation estimates for stochastic conservation laws. Specifically, in [85], large deviation principles have been investigated and derived in the limit of jointly vanishing noise and viscosity by using delicate scaling arguments. Notably, in [2], the bounds for the rate function have also been derived in the vanishing viscosity limit only, so that the noise is allowed in the limit, and in the multidimensional setting. Finally, we mention the more recent work (cf. [41]) and the references cited therein, large deviation principles have been derived for the first-order scalar conservation laws with small multiplicative noise on  $\mathbb{T}^d$  in the zero-noise limit by using the Freidlin-Wentzell theory. Much still remains to be explored in this direction.

## References

- [1] Arnold, L., Random Dynamical Systems, Springer-Verlag, Berlin Heidelberg, 1998.
- [2] Barré, J., Bernardin, C. and Chertrite, R., Density large deviations for multidimensional stochastic hyperbolic conservation laws, *J. Stat. Phys.*, **170**(3), 2017, 466–491.
- [3] Bauzet, C., Vallet, G. and Wittbold, P., The Cauchy problem for conservation laws with a multiplicative stochastic perturbation, *J. Hyper. Diff. Eq.*, **9**(4), 2012, 661–709.
- [4] Bergé, B. and Saussereau, B., On the long-time behavior of a class of parabolic SPDEs: Monotonicity methods and exchange of stability, *ESAIM: Probab. Stat.*, **9**, 2005, 254–276.
- [5] Birkens, J., Sufficient conditions for the eventual strong Feller property for degenerate stochastic evolutions, *J. Math. Anal. Appl.*, **379**(2), 2011, 469–481.
- [6] Bouchut, F. and Desvillettes, L., Averaging lemmas without time Fourier transform and application to discretized kinetic equations, *Proc. Royal Soc. Edin. Sec. A*, **129**, 1999, 19–36.
- [7] Bricmont, J., Kupiainen, A. and Lefevre, R., Ergodicity of the 2D Navier-Stokes equations with random forcing, *Comm. Math. Phys.*, **224**(1), 2001, 65–81.

- [8] Bricmont, J., Kupiainen, A. and Lefevere, R., Exponential mixing of the 2D Navier-Stokes dynamics, *Comm. Math. Phys.*, **224**(1), 2002, 87–132.
- [9] Chen, G.-Q., Ding, Q. and Karlsen, K., On nonlinear stochastic balance laws, *Arch. Ration. Mech. Anal.*, **204**(3), 2012, 707–743.
- [10] Chen, G.-Q. and Frid, H., Large time behavior of entropy solutions in  $L^\infty$  for multidimensional conservation laws, *Advances in Nonlinear PDEs and Related Areas*, 28–44, World Sci. Publ., River Edge, NJ, 1998.
- [11] Chen, G.-Q. and Lu, Y.-G., A study of approaches to applying the theory of compensated compactness, *Chinese Sci. Bull.*, **34**, 1989, 15–19.
- [12] Chen, G.-Q. and Pang, P., On nonlinear anisotropic degenerate parabolic-hyperbolic equations with stochastic forcing, 2019, arXiv:1903.02693.
- [13] Chen, G.-Q. and Perthame, B., Well-posedness for non-isotropic degenerate parabolic-hyperbolic equations, *Ann. I.H.P. Anal. Non-Linéaires*, **20**(4), 2003, 645–668.
- [14] Chen, G.-Q. and Perthame, B., Large-time behavior of periodic entropy solutions to anisotropic degenerate parabolic-hyperbolic equations, *Proc. A.M.S.*, **137**(9), 2009, 3003–3011.
- [15] Chueshov, I. and Vuillermot, P., On the large-time dynamics of a class of random parabolic equations, *C. R. Acad. Sci. Paris*, **322**(1), 1996, 1181–1186.
- [16] Chueshov, I. and Vuillermot, P., On the large-time dynamics of a class of parabolic equations subjected homogeneous white noise: Stratonovitch’s case, *C. R. Acad. Sci. Paris*, **323**(1), 1996, 29–33.
- [17] Chueshov, I. and Vuillermot, P., Long-time behavior of solutions to a class of quasilinear parabolic equations with random coefficients, *Ann. I.H.P. Anal. Non-Linéaires*, **15**(2), 1998, 191–232.
- [18] Chueshov, I. and Vuillermot, P., Long-time behavior of solutions to a class of stochastic parabolic equations with homogeneous white noise: Straonovitch’s case, *Probab. Theory Related Fields*, **112**(2), 1998, 149–202.
- [19] Chueshov, I. and Vuillermot, P., On the large-time dynamics of a class of parabolic equations subjected to homogeneous white noise: Itô’s case, *C. R. Acad. Sci. Paris*, **326**(1), 1998, 1299–1304.
- [20] Chueshov, I. and Vuillermot, P., Long-time behavior of solutions to a class of stochastic parabolic equations with homogeneous white noise: Itô’s case, *Stoch. Anal. Appl.*, **18**(4), 2000, 581–615.
- [21] Coti-Zelati, M., Glatt-Holtz, N. and Trivisa, K., Invariant measures for the stochastic one-dimensional compressible Navier-Stokes equations, 2018, arXiv:1802.04000v1.
- [22] Crauel, H., Markov measures for random dynamical systems, *Stochastics and Stochastic Reports*, **37**(3), 1991, 153–173.
- [23] Crauel, H., *Random Probability Measures on Polish Spaces*, Taylor and Francis, London, 2002.
- [24] Crauel, H., Debussche, A. and Flandoli, F., Random attractors, *J. Dynam. Diff. Eq.*, **9**(2), 1997, 307–341.
- [25] Crauel, H. and Flandoli, F., Attractors for random dynamical systems, *Probab. Theory Relat. Fields*, **100**(3), 1994, 365–393.
- [26] Da Prato, G. and Debussche, A., Two-dimensional Navier-Stokes equations driven by a space-time white noise, *J. Funct. Anal.*, **196**(1), 2002, 180–210.
- [27] Da Prato, G. and Debussche, A., Ergodicity for the 3D stochastic Navier-Stokes equations, *J. Math. Pures Appl.*, **82**(8), 2003, 877–947.
- [28] Da Prato, G., Debussche, A., and Temam, R., Stochastic Burger’s equation, *No.D.E.A.*, **1**, 1994, 389–402.
- [29] Da Prato, G., Flandoli, F., Priola, E., and Röckner, M., Strong uniqueness for stochastic evolution equations in Hilbert spaces perturbed by a bounded measurable drift, *Ann. Probab.*, **41**(5), 2013, 3306–3344.
- [30] Da Prato, G. and Zabczyk, J., *Stochastic Equations in Infinite Dimensions*, Cambridge University Press, Cambridge, 2008.
- [31] Debussche, A., On the finite dimensionality of random attractors, *Stoch. Anal. Appl.*, **15**(4), 1997, 473–492.
- [32] Debussche, A., Hausdorff dimension of a random invariant set, *J. Math. Pures Appl.*, **77**(10), 1998, 967–988.
- [33] Debussche, A., Hofmanová, M. and Vovelle, J., Degenerate parabolic stochastic partial differential equations: Quasilinear case, *Ann. Probab.*, **44**(3), 2016, 1916–1955.
- [34] Debussche, A. and Vovelle, J., Scalar conservation laws with stochastic forcing, *J. Funct. Anal.*, **259**(4), 2010, 1014–1042.
- [35] Debussche, A. and Vovelle, J., Invariant measure of scalar first-order conservation laws with stochastic forcing, *Probab. Theory Relat. Fields*, **163**(3), 2015, 575–611.

- [36] Dembo, A. and Zeitouni, O., Large Deviations Techniques and Applications, 2nd ed., Springer-Verlag, New York, 1997.
- [37] Delarue, F., Flandoli, F. and Vincenzi, D., Noise prevents collapse of Vlasov-Poisson point charges, *Comm. Pure Appl. Math.*, **67**(10), 2014, 1700–1736.
- [38] Deuschel, J.-D., and Stroock, D., Large Deviations, Academic Press, London, 1989.
- [39] Doeblin, W., Éléments d’une théorie générale des chaînes simples constantes de Markoff, *Ann. Sci. École Normale Supérieure*, **57**, 1940, 61–111.
- [40] Doeblin, W. and Fortet, R., Sur le chaînes à liaisons complète, *Bull. Soc. Math. France*, **65**, 1937, 132–148.
- [41] Dong, Z., Wu, J.-L., Zhang, R.-R. and Zhang, T.-S., Large deviation principles for first-order scalar conservation laws with stochastic forcing, 2018, arXiv:1806.02955v1.
- [42] Doob, J., Asymptotic property of Markoff transition probability, *Trans. Amer. Math. Soc.*, **64**(3), 1948, 393–421.
- [43] Dupuis, P. and Ellis, R. S., A Weak Convergence Approach to the Theory of Large Deviations, New York, Wiley, 1997.
- [44] E, Weinan, Khanin, K., Mazel, A. and Sinai, Ya., Probability distribution functions for the random forced Burgers equation, *Phys. Rev. Lett.*, **78**(10), 1997, 1904–1907.
- [45] E, Weinan, Khanin, K., Mazel, A. and Sinai, Ya., Invariant measures for Burgers equation with stochastic forcing, *Ann. Math.*, **151**(3), 2000, 877–960.
- [46] E, Weinan, Mattingly, J. and Sinai, Ya., Gibbsian dynamics and ergodicity for the stochastically forced Navier-Stokes equation, *Comm. Math. Phys.*, **224**(1), 2001, 83–106.
- [47] Fedrizzi, E., Flandoli, F., Priola, E. and Vovelle, J., Regularity of stochastic kinetic equations, *Electron. J. Probab.*, **22**(48), 2017, 1–42.
- [48] Fehrman, B. and Gess, B., Well-posedness of nonlinear diffusion equations with nonlinear, conservative noise, 2019, arXiv:1712.05775.
- [49] Feng, J. and Nualart, D., Stochastic scalar conservation laws, *J. Funct. Anal.*, **255**(2), 2008, 313–373.
- [50] Flandoli, F., Stochastic flows and Lyapunov exponents for abstract stochastic PDEs of parabolic type, Lyapunov Exponents Proceedings, Arnold, L., Crauel, H., Eckmann, J.-P. (eds), LNM **1486**, 1991, 196–205.
- [51] Flandoli, F., Dissipativity and invariant measures for stochastic Navier-Stokes equations, *No.D.E.A.*, **1**(4), 1994, 403–423.
- [52] Flandoli, F., Regularity Theory and Stochastic Flows for Parabolic SPDEs, Gordon and Breach Science Publishers, Singapore, 1995.
- [53] Flandoli, F. and Gatarek, D., Martingale and stationary solutions for stochastic Navier-Stokes equations, *Probab. Theory Relat. Fields*, **102**(3), 1995, 367–391.
- [54] Flandoli, F., Gubinelli, M. and Priola, E., Well-posedness of the transport equation by stochastic perturbation, *Invent. Math.*, **180**(1), 2010, 1–53.
- [55] Flandoli, F. and Maslowski, B., Ergodicity of the 2-D Navier-Stokes equations under random perturbations, *Comm. Math. Phys.*, **172**(1), 1995, 119–141.
- [56] Földes, J., Friedlander, S., Glatt-Holtz, N. and Richards, G., Asymptotic analysis for randomly forced MHD, 2016, arXiv:1604.06352.
- [57] Földes, J., Glatt-Holtz, N., Richards, G. and Whitehead, J. P., Ergodicity in randomly forced Rayleigh-Bénard convection, 2015, arXiv:1511.01247.
- [58] Freidlin, M. I. and Wentzell, A. D., Random Perturbation of Dynamical Systems, 2nd ed., Springer-Verlag, New York, 1998.
- [59] Gess, B. and Souganidis, P., Long-time behavior, invariant measures and regularizing effects for stochastic scalar conservation laws, 2014, arXiv:1411.3939.
- [60] Gess, B. and Souganidis, P., Stochastic non-isotropic degenerate parabolic-hyperbolic equations, 2016, arXiv:1611.01303.
- [61] Giga, Y. and Miyakawa, T., A kinetic construction of global solutions of first order quasilinear equations, *Duke Math. J.*, **50**(2), 1989, 505–515.
- [62] Gyöngy, I. and Nualart, D., On the stochastic Burgers equation in the real line, *Ann. Probab.*, **27**(2), 1999, 782–802.

- [63] Hairer, M. and Mattingly, J., Ergodicity of the 2D Navier-Stokes equations with degenerate stochastic forcing, *Ann. Math.*, **164**, 2006, 993–1032.
- [64] Hairer, M. and Mattingly, J., Spectral gaps in Wasserstein distances and the 2D stochastic Navier-Stokes equations, *Ann. Probab.*, **36**(6), 2008, 2050–2091.
- [65] Hairer, M. and Mattingly, J., A theory of hypoellipticity and unique ergodicity for semilinear stochastic PDEs, *Electr. J. Probab.*, **16**(23), 2011, 658–738.
- [66] Hairer, M. and Voss, J., Approximations to the stochastic Burgers equation, *J. Nonlin. Sci.*, **21**(6), 2011, 897–920.
- [67] Hofmanová, M., Degenerate parabolic stochastic partial differential equations, *Stochastic Process. Appl.*, **123**(12), 2013, 4294–4336.
- [68] Hollander den, F., Large Deviations, Fields Institute Monographs, A.M.S., Providence, RI, 2000.
- [69] Karlsen, K. H. and Storrøsten, E., On stochastic conservation laws and Malliavin calculus, 2015, arXiv:1507.05518v2.
- [70] Khas'minskii, R., Ergodic properties of recurrent diffusion processes and stabilization of the solutions to the Cauchy problem for parabolic equations, *Theory Probab. Appl.*, **5**(2), 1960, 179–196.
- [71] Krylov, N. and Röckner, M., Strong solutions of stochastic equations with singular time dependent drift, *Probab. Theory Relat. Fields*, **131**(2), 2005, 154–196.
- [72] Kuksin, S. and Shirikyan, A., Mathematics of Two-Dimensional Turbulence, Cambridge University Press, Cambridge, 2002.
- [73] Kuksin, S. and Shirikyan, A., Randomly Forced Nonlinear PDEs and Statistical Hydrodynamics in 2 Space Dimensions, European Mathematical Society, Zürich, 2006.
- [74] Le Jan, Y., Équilibre statistique pour les produits de difféomorphismes aléatoires indépendants, *Ann. l'I.H.P. Probab. Stat.*, **23**(1), 1987, 111–120.
- [75] Ledrappier, F., Positivity of the exponent for stationary sequences of matrices, Lyapunov Exponents Proceedings, Arnold, L., Wihstutz, V. (eds.), LNM **1186**, 1986, 56–73.
- [76] Leòn, J., Nualart, D. and Pettersson, R., The stochastic burgers equation: Finite moments and smoothness of the density, *Infinite Dimensional Analysis*, **3**, 2000, 363–385.
- [77] Lindvall, T., Lectures on the Coupling Method, John Wiley & Sons, New York, 1992.
- [78] Lions, P.-L., Perthame, B. and Souganidis, P., Scalar conservation laws with rough (stochastic) fluxes, *Stochastic PDEs: Anal. Comput.*, **1**, 2013, 664–686.
- [79] Lions, P.-L., Perthame, B. and Souganidis, P., Scalar conservation laws with rough (stochastic) fluxes, the spatially dependent case, *Stochastic PDEs: Anal. Comput.*, **2**(4), 2014, 517–538.
- [80] Lions, P.-L., Perthame, B. and Tadmor, E., A kinetic formulation of multidimensional scalar conservation laws and related equations, *J. Amer. Math. Soc.*, **7**(1), 1994, 169–191.
- [81] Lions, P.-L., Perthame, B. and Tadmor, E., Kinetic formulation of the isentropic gas dynamics and  $p$ -systems, *Comm. Math. Phys.*, **163**, 1994, 415–431.
- [82] Majda, A., Introduction to Turbulent Dynamical Systems in Complex Systems, Springer-Verlag, Cham, 2016.
- [83] Majda, A., and Tong, X.-T., Ergodicity of truncated stochastic Navier Stokes with deterministic forcing and dispersion, *J. Nonlin. Sci.*, **26**(5), 2016, 1483–1506.
- [84] Majda, A. and Wang, X.-M., Nonlinear Dynamics and Statistical Theories for Basic Geophysical Flows, Cambridge University Press, Cambridge, 2006.
- [85] Mariani, M., Large deviations principles for stochastic scalar conservation laws, *Probab. Theory Relat. Fields*, **147**, 2010, 607–648.
- [86] Maslowski, B. and Seidler, J., Invariant measures for nonlinear SPDE's: Uniqueness and stability, *Archivum Mathematicum*, **34**(1), 1998, 153–172.
- [87] Masmoudi, N. and Young, L.-S., Ergodic theory of infinite dimensional systems with applications to dissipative parabolic PDEs, *Comm. Math. Phys.*, **227**(3), 2002, 461–481.
- [88] Mattingly, J. C., Exponential convergence for the stochastically forced Navier-Stokes equations and other partially dissipative dynamics, *Comm. Math. Phys.*, **230**(3), 2002, 421–462.
- [89] Mattingly, J. C., Recent progress for the stochastic Navier-Stokes equations, *Journéé ÉDP.*, **11**, 2003, 52 pages, DOI: 10.5802/jedp.625.
- [90] Perthame, B., Kinetic Formulation of Conservation Laws, Oxford University Press, Oxford, 2002.

- [91] Peszat, S. and Zabczyk, J., Strong Feller property and irreducibility for diffusions on Hilbert spaces, *Ann. Probab.*, **23**(1), 1995, 157–172.
- [92] Peszat, S. and Zabczyk, J., *Stochastic Partial Differential Equations with Lévy Noise: An Evolution Equation Approach*, Cambridge University Press, Cambridge, 2007.
- [93] Röckner, M. and Zhang, X.-C., Stochastic tamed 3D Navier-Stokes equations: existence, uniqueness and ergodicity, *Prob. Theory Relat. Fields*, **145**, 2009, 211–267.
- [94] Romito, M. and Xu, L.-H., Ergodicity of the 3D stochastic Navier-Stokes equations driven by mildly degenerate noise, *Stoc. Proc. Appl.*, **121**, 2011, 673–700.
- [95] Schmalfuss, B., The random attractor of the stochastic Lorenz system, *ZAMP*, **48**(6), 1997, 951–975.
- [96] Sinai, Ya., Two results concerning asymptotic behavior of solutions of the Burgers equation with force, *J. Stat. Phys.*, **64**(1), 1991, 1–12.
- [97] Tadmor, E. and Tao, T., Velocity averaging, kinetic formulations and regularizing effects in quasi-linear pdes, *Comm. Pure Appl. Math.*, **60**(10), 2007, 1488–1521.
- [98] Temam, R., *Infinite Dimensional Dynamical Systems in Mechanics and Physics*, Springer-Verlag, New York, 1997.
- [99] Varadhan, S. R. S., *Large Deviations*, Courant Lecture Notes, A.M.S., Providence, RI, 2016.
- [100] Veretennikov, A., On strong solutions and explicit formulas for solutions of stochastic integral equations, *Math. USSR Sbornik*, **39**(3), 1981.
- [101] Villani, C., *Optimal Transport*, Grundlehren der Mathematischen Wissenschaften, **338**, Springer-Verlag, 2009.
- [102] Zhu, R. and Zhu, X., Three-dimensional Navier-Stokes equations driven by space-time white noise, *J. Diff. Eq.*, **259**(9), 2015, 4443–4508.