ON AVERAGING AND MIXING FOR STOCHASTIC PDES

GUAN HUANG AND SERGEI KUKSIN

ABSTRACT. We examine the convergence in the Krylov–Bogolyubov averaging for nonlinear stochastic perturbations of linear PDEs with pure imaginary spectrum and show that if the involved effective equation is mixing, then the convergence is uniform in time.

1. INTRODUCTION

The Krylov–Bogolyubov averaging for stochastic PDEs which we are concerned with in this work, means the following. Starting with a linear PDE on a torus (or on a bounded domain) with pure imaginary discrete spectrum we consider its ε-small nonlinear stochastic perturbation. Then the above mentioned averaging describes the behaviour of the distributions of actions of solutions for the perturbed equation on time-intervals of order ε^{-1}. Here the actions of solutions are made by the halves of squared norms of their Fourier coefficients with respect to the basis, made by eigenfunctions of the original linear system. Description of the limit is made via an auxiliary *effective equation* which is another nonlinear stochastic equation whose nonlinearity is made from resonant terms of the nonlinear part of the perturbation. The effective equation may be mixing, and then as time goes to infinity its solutions converge in distributions to a statistical equilibrium, given by a measure in a function space. An observation which we make in this work is that in the mixing case the convergence in distribution of actions of solutions for the perturbed equation to those of solution for the effective equation, described by the Krylov–Bogolyubov averaging, is uniform in time.

The Krylov–Bogolyubov averaging under discussion applies to various classes of stochastic PDEs, depending on the type of the original unperturbed linear system. In Sections 2-3 we discuss in details the averaging for stochastic complex Ginzburg–Landau (CGL) equations, regarded as perturbations of linear Schrödinger equations, and in Section 4 briefly repeat the argument for stochastic nonlinear wave equations.

Notation. For a Banach space B and $R > 0$ we denote $\bar{B}_R(B) = \{b \in B : |b|_B \le R\}$; for a metric space M, $\mathcal{P}(M)$ stands for the space of probability Borel measures on M. By $\rightharpoonup$ we denote the weak convergence of measures and by $\mathcal{D}(\xi)$ – the distribution of a random variable ξ. For a function f and a measure μ we denote $\langle f, \mu\rangle = \int f\, d\mu$.

2. CGL: THE SETTING AND RESULT

We consider a stochastic CGL equation on a torus $T^D := \mathbb{R}/(L_1\mathbb{Z}) \times \mathbb{R}/(L_2\mathbb{Z}) \times \cdots \times \mathbb{R}/(L_D\mathbb{Z})$, $L_1, \dots, L_D > 0$,

$$u_t + i(-\Delta + V(x))u = \varepsilon\mu\Delta u + \varepsilon\mathscr{P}(\nabla u, u) + \sqrt{\varepsilon}\eta(t,x), \quad u = u(t,x),\ x \in T^D, \quad (2.1)$$

where $\mu \in \{0,1\}$, $\mathscr{P} : \mathbb{C}^{D+1} \to \mathbb{C}$ is a C^∞-smooth function, $\varepsilon \in (0,1]$ is a small parameter, the random force $\eta(t,x)$ is white in time and regular in x, and the potential $V(x)$ is a real smooth function. If $\mu = 0$, the nonlinearity $\mathscr{P}(\nabla u, u)$ should not depend on ∇u. For simplicity we assume that $\mu = 1$ (the case $\mu = 0$

can be treated similarly). Again only to simplify presentation we also assume that $V(x) > 0$ for all x.

For any $s \in \mathbb{R}$ we denote by H^s the Sobolev space of complex functions on T^D, provided with the norm $\|\cdot\|_s$,

$$\|u\|_s^2 = \langle(-\Delta)^s u, u\rangle + \langle u, u\rangle, \text{ if } s \geqslant 0,$$

where $\langle\cdot,\cdot\rangle$ is the real scalar product in $L^2(T^D;\mathbb{C})$,

$$\langle u, v\rangle = \Re \int_{T^d} u\bar{v}\, \mathrm{d}x, \quad u, v \in L^2(T^D;\mathbb{C}).$$

Let $\{\mathbf{e}_l(x), l \in \mathbb{N}\}$ be the usual trigonometric basis of the space $L^2(T^D)$, parametrized by natural numbers. Then $-\Delta\mathbf{e}_l = \kappa_l \mathbf{e}_l$, $\kappa_l \geq 0$, and we assume that $0 = \kappa_1 < \kappa_2 \leq \kappa_3, \dots$. We take the force term $\eta(t,x)$ in (2.1) to be of the form

$$\eta(t,x) = \frac{\partial}{\partial t}\xi(t,x), \quad \xi(t,x) := \sum_{l\geqslant 1} b_l \beta_l(t)\mathbf{e}_l(x). \tag{2.2}$$

Here $\beta_l(t) = \beta_l^R(t) + i\beta_l^I(t)$, where $\beta_l^R(t)$, $\beta_l^I(t)$, $l \geqslant 1$, are independent real-valued standard Brownian motions, defined on a complete probability space $(\Omega, \mathcal{F}, \mathbf{P})$ with a filtration $\{\mathcal{F}_t; t \geqslant 0\}$.[1] As a function of x, $\xi(t,x)$ is assumed to be smooth in the sense that the real numbers $b_l, l \geq 1$, decays to zero faster than any negative degree of l.

Introducing the slow time $\tau = \varepsilon t$, we rewrite eq. (2.1) as

$$\dot{u} + \varepsilon^{-1}(-\Delta + V(x))u = \Delta u + \mathscr{P}(\nabla u, u) + \dot{\xi}(\tau, x), \quad u = u(\tau, x), \ x \in T^D, \tag{2.3}$$

where the upper dot stands for $\partial/\partial\tau$, and $\xi(\tau,x)$ is as in (2.2) with $t := \tau$ and with another set of standard independent complex Brownian motions β_l. Here and below we write stochastic PDEs with additive noise as nonlinear PDEs with forcing terms of the form (2.2).

Definition 2.1. If L and E are Banach spaces with norms $|\cdot|_L$ and $|\cdot|_E$, then $\mathrm{Lip}_m(L,E)$, $m \geq 0$, is the collection of maps $F : L \to E$ such that for any $R \geqslant 1$,

$$\sup_{R>0}\Big((1+|R|)^{-m}\big(\mathrm{Lip}(F|_{\bar{B}_R(L)}) + \sup_{v\in\bar{B}_R(L)} |F(v)|_E\big)\Big) < \infty,$$

where $\mathrm{Lip}(f)$ is the Lipschitz constant of a mapping f.

We make the following assumption concerning the well-posedness of eq. (2.3). There and everywhere below in our paper

either always the indices s of involved Sobolev spaces H^s are integer,
or always they are any real numbers.

Assumption 2.2. There exist numbers $0 < s_1 < s_2 < +\infty$ and $\bar{m} \in \mathbb{N}$ such that for each $s \in (s_1, s_2)$,

(1) the mapping $H^s \to H^{s-1}$: $u \mapsto \mathscr{P}(\nabla u, u)$ belongs to $\mathrm{Lip}_{\bar{m}}(H^s, H^{s-1})$;
(2) for any $\varepsilon \in (0,1]$ and $u_0 \in H^s$ equation (2.3) has a unique strong solution $u^\omega(\tau; u_0)$, equal u_0 at $\tau = 0$, defined for $\tau \geqslant 0$, and

$$\mathbf{E} \sup_{\theta\leqslant\tau\leqslant\theta+1} \|u(\tau;u_0)\|_s^{2m'} \leqslant C_s(\|u_0\|_s), \quad \forall\theta \geqslant 0,$$

where m' is some number bigger than $\bar{m}$ and $C_s(\cdot)$ is a continuous non-decreasing function.
(3) If we work in the category of Sobolev spaces with integer indices, then the integer segment $(s_1, s_2) \cap \mathbb{Z}$ contains at least two points.

[1] So $\{\beta_l(t)\}$ are standard independent complex Brownian motions.

Under the above assumptions the family of solutions $\{u^\omega(\tau;u_0), u_0\in H^s\}$ defines in the spaces H^s, $s\in(s_1,s_2)$, Markov processes.

Assumption 2.2 is satisfied for many nonlinearities $\mathscr{P}$. In particular it holds if

$$\mathscr{P}(\nabla u,u)=-u+\mathfrak{z}f_p(|u|^2)u,\quad \mathfrak{z}\in\mathbb{C},\ |\mathfrak{z}|=1,\mathfrak{I}\mathfrak{z}\le 0,\ \mathfrak{R}\mathfrak{z}\le 0, \tag{2.4}$$

where $f_p(r)$ is a non-decreasing smooth function on $\mathbb{R}$, equal r^p for $r\ge 1$. The degree $p\ge 0$ is any if $D=1,2$, and $p<2/(D-2)$ if $D\ge 3$. See [4, Section 5].

We denote by A_V the Schrödinger operator

$$A_Vu:=-\Delta u+V(x)u.$$

Let $\{\lambda_l\}_{l\geqslant 1}$ be its eigenvalues, ordered in such a way that

$$0<\lambda_1\leqslant\lambda_2\leqslant\lambda_3\leqslant\cdots$$

(we recall that $V>0$) and let $\{\varphi_l,\ l\geqslant 1\}\subset L^2(T^D)$ be an orthonormal basis, formed by the corresponding eigenfunctions. We denote $\Lambda=(\lambda_1,\lambda_2,\dots)$ and call Λ the *frequency vector* of eq. (2.3). For a complex-valued function $u\in H^s$ we denote by

$$\Psi(u):=v=(v_1,v_2,\dots),\quad v_k\in\mathbb{C},$$

the vector of its Fourier coefficients with respect to the basis $\{\varphi_l\}_{l\geqslant 1}$: $u=\sum_{l\geqslant 1}v_l\varphi_l$. Note that Ψ is a real operator: it maps real functions $u(\cdot)$ to real vectors v. In the space of complex sequences $v=(v_1,v_2,\dots)$ we introduce the norms

$$|v|_s^2=\sum_{k=1}^{\infty}(|\lambda_k|^s+1)\,|v_k|^2,\quad s\in\mathbb{R},$$

and denote $h^s=\{v:|v|_s<\infty\}$. Then Ψ defines an isomorphism between the spaces H^s and h^s, for any s.

Now we write Eq. (2.3) in the v-variables:

$$\dot v_k+\varepsilon^{-1}i\lambda_kv_k=-\lambda_kv_k+P_k(v)+\sum_{l=1}^{\infty}\Psi_{kl}b_l\dot\beta_l,\quad k\in\mathbb{N}. \tag{2.5}$$

Here the k-th equation is obtained as the L^2-scalar product with φ_k of eq. (2.3), where $u=\Psi^{-1}v$. So

$$P(v)=(P_k(v),\ k\in\mathbb{N})=\Psi\Big(V(x)u+\mathscr{P}(\nabla u,u)\Big),\quad u=\Psi^{-1}v,$$

and $\Psi_{kl}=\langle\varphi_k,\mathbf{e}_l\rangle$ (thus $\big(\Psi_{kl}\big)$ is the matrix of the operator Ψ with respect to the trigonometric basis in $L^2(T^D)$ and the natural basis in h^0).

Our task is to study the dynamics of eq. (2.5) when $\varepsilon\ll 1$. An efficient way to deal with this problem is through the *interaction representation*, which means transition from variables $\{v_k(\tau)\}$ to variables $\{a_k(\tau)\}$, where

$$a_k(\tau)=e^{i\varepsilon^{-1}\lambda_k\tau}v_k(\tau),\quad k\ge 1.$$

In the a-variables equations (2.5) read

$$\dot a_k(\tau)=-\lambda_ka_k+e^{i\varepsilon^{-1}\lambda_k\tau}P_k(\Phi_{-\varepsilon^{-1}\Lambda\tau}a)+e^{i\varepsilon^{-1}\lambda_k\tau}\sum_{l=1}^{\infty}\Psi_{kl}b_l\dot\beta_l,\quad k\in\mathbb{N},\ \ \tau\ge 0, \tag{2.6}$$

where for a vector $\theta=(\theta_k,\ k\in\mathbb{N})\in\mathbb{R}^\infty$, Φ_θ stands for the rotation in h^s, defined by

$$\Phi_\theta v=v',\quad v_k'=e^{i\theta_k}v_k\quad\forall\,k. \tag{2.7}$$

Clearly operators Φ_θ define isometries of all spaces h^s. By $a^\varepsilon(\tau;v_0)=\big(a_k^\varepsilon(\tau;v_0),k\ge 1\big)$ we denote a solution $u(\tau;u_0)$, written in the a-variables. It solves system (2.6) with the initial data

$$a(0)=v_0:=\Psi(u_0),$$

and in view of Assumption 2.2.(2),

$$\mathbf{E}\sup_{\theta\leqslant\tau\leqslant\theta+1}|a^\varepsilon(\tau;v_0)|_s^{2m'}\leqslant C'_s(|v_0|_s),\quad \forall\theta\geqslant 0. \tag{2.8}$$

In order to describe the dynamics of eq. (2.6) with $\varepsilon\ll 1$ we introduce an *effective equation*:

$$\dot a_k = -\lambda_k a_k + R_k(a) + \sum_{l=1}^\infty B_{kl}\dot\beta_l,\quad k\in\mathbb{N},\quad \tau\geq 0. \tag{2.9}$$

Here $\{B_{kl}, k,l\geqslant 1\}$ is the principal square root of infinite matrix $\{A_{kl},\ k,l\geqslant 1\}$,[2]

$$A_{kl} = \begin{cases}\sum_{j\geqslant 1} b_j^2\Psi_{kj}\Psi_{lj}, & \text{if } \lambda_k=\lambda_l,\\ 0, & \text{else,}\end{cases}$$

and

$$R(a) := (R_k(a),\ k\in\mathbb{N}) = \lim_{T\to\infty}\frac{1}{T}\int_0^T \Phi_{\Lambda t}P(\Phi_{-\Lambda t}a)dt \tag{2.10}$$

(see [4, Lemma 4] concerning this limit). Under (1) of Assumption 2.2, for $s\in(s_1,s_2)$, the mapping $h^s\to h^{s-1}: v\mapsto R(v)$ belongs to $\mathrm{Lip}_{\bar m}(h^s,h^{s-1})$, see [4]. Therefore for any $a_0\in H^s$ with $s\in(s_1,s_2)$, a strong solution $a^0(\tau;a_0)$ of eq. (2.9), equal a_0 at $\tau=0$, is unique and exists at least locally. In [4, Theorem 2 and Proposition 1] we proved the following result:

Theorem 2.3. *If Assumption 2.2 holds, then for any $s_1<s_*<\bar s<s_2$ and any $v_0\in h^{\bar s}$ we have:*

i) eq. (2.9) has a unique strong solution $a^0(\tau;v_0)$, $\tau\geq 0$, equal to v_0 at $\tau=0$. It belongs to $C([0,\infty),h^{\bar s})$ a.s., and for any $\theta\geq 0$,

$$\mathbf{E}\sup_{\tau\in[\theta,\theta+1]}|a^0(\tau;v_0)|_{\bar s}^{2m'}\leqslant C'_{\bar s}(|v_0|_{\bar s}). \tag{2.11}$$

ii) For any $T>0$,

$$\mathcal{D}\big(a^\varepsilon(\tau;v_0)|_{[0,T]}\big)\rightharpoonup\mathcal{D}\big(a^0(\tau;v_0)|_{[0,T]}\big)\quad \textit{in } \mathcal{P}\big(C([0,T]);h^{s_*})\big)\ \textit{as } \varepsilon\to 0. \tag{2.12}$$

Remark 2.4. 1) In [4] it was proved only the second assertion of the theorem, implying that a^0 is a solution of eq. (2.9) in h^{s_*}. Then (2.8) and Fatou's lemma imply (2.11). Indeed, for $M\in\mathbb{N}$ let $\Pi_M(a_1,a_2,\dots)=(a_1,\dots,a_M,0,\dots)$. Then by (2.8) and convergence (2.12), for each $\theta\geq 0$ and any $M,N\in\mathbb{N}$

$$\mathbf{E}\sup_{\tau\in[\theta,\theta+1]}N\wedge|\Pi_M a^0(\tau;v_0)|_{\bar s}^{2m'}\leqslant C'_{\bar s}(|v_0|_{\bar s}).$$

Sending first $N\to\infty$ and then $M\to\infty$ and using Fatou's lemma we recover (2.11). Since a^0 is a solution in h^{s_*}, then dues to (2.11) it is a solution in $h^{\bar s}$. Uniqueness of a solution is obvious.

2) It follows immediately that (2.12) also holds for a solution $a^\varepsilon(\tau;v_0)$ if the initial data v_0 is a r.v. in $h^{\bar s}$, independent from the random field ξ. Moreover, a simple analysis of the proof in [4] implies that if $|v_0^\omega|_{\bar s}\leq M$ a.s., then the rate of convergence (2.12) depends only on M (and, of course, on $s_*<\bar s$).

3) Theorem's assertion with the same proof remains true if in the r.h.s. of eq. (2.3) we replace the viscosity Δu by the hyperviscosity $-(-\Delta+1)^r u$, $r\in\mathbb{N}$ (provided that Assumption 2.2 holds for the equation). The equations with hyperviscosity are important for some applications, see below Example 2.11.

4) In [4] the bounds $\mathbf{E}\sup_{\theta\leqslant\tau\leqslant\theta+1}\|u(\tau;u_0)\|_s^m$ are assumed for all m-th moments of solutions, $m\in\mathbb{N}$. It was done only for simplicity, and under the additional restriction in Assumption 2.2.(1), which specifies the growth of the the nonlinearity

[2]The matrix (A_{kl}) defines a non-negative compact self-adjoint operator in the space l^2. So its principal square root (which defines another non-negative compact self-adjoint operator) exists.

$\mathscr{P}$, only bounds for the $2m'$-moments with some $m' > \bar{m}$ are needed for the proof. This difference is rather insignificant since using the standard techniques of exponential supermartingales it is usually easy to obtain bounds for all m-th moments after the second moments are estimated.

Again, solutions $a^0(\tau; v_0)$ of (2.9) define in the space h^{s_*} a Markov process.

Our goal in this work is to prove that if effective equation (2.9) is mixing, then convergence (2.12) is such that $\mathcal{D}(a^\varepsilon(\tau; v_0)) \rightharpoonup \mathcal{D}(a^0(\tau; v_0))$ uniformly in $\tau \geq 0$, with respect to the dual-Lipschitz distance. We recall

Definition 2.5. Let M be a complete and separable metric space. For any two measures μ_1, $\mu_2 \in \mathcal{P}(M)$ the dual-Lipschitz distance between them is

$$\|\mu_1 - \mu_2\|^*_{L,M} := \sup_{f \in C(M), |f|_{L,M} \leqslant 1} \left| \langle f, \mu_1 \rangle - \langle f, \mu_2 \rangle \right| \leqslant 2,$$

where $|f|_{L,M} = \mathrm{Lip} f + \|f\|_{C(M)}$.

For future usage we note that for $s \geq s'$ space $\mathcal{P}(h^s)$ is naturally embedded in $\mathcal{P}(h^{s'})$, and it easily follows from the definition that for $\mu_1, \mu_2 \in \mathcal{P}(h^s)$ the distance $\|\mu_1 - \mu_2\|^*_{L,h^{s'}}$ is a non-decreasing function of $s' \leq s$.

To proceed we have to assume that the effective equation is mixing, and the rate of mixing is uniform for initial data from bounded sets:

Assumption 2.6. For some $s_* \in (s_1, s_2)$ effective equation (2.9) is mixing in the space h^{s_*} with a stationary measure $\mu^0 \in \mathcal{P}(h^{s_*})$, and for each $M > 0$ and $v \in \bar{B}_M(h^{s_*})$, we have

$$\|\mathcal{D}(a^0(\tau; v)) - \mu^0\|^*_{L,h^{s_*}} \leqslant \mathfrak{g}_M(\tau), \tag{2.13}$$

where $\mathfrak{g}$ is a continuous function of (M, τ) which goes to zero when $\tau \to \infty$.

Relation (2.13) is a mild specification of the mixing in eq. (2.9), and a proof of the latter in fact usually establishes the former.

Our main result is the following:

Theorem 2.7. *Under Assumptions 2.2 and 2.6, for any $\bar{s} \in (s_*, s_2)$ and any $v_0 \in h^{\bar{s}}$*

$$\lim_{\varepsilon \to 0} \sup_{\tau \geqslant 0} \|\mathcal{D}(a^\varepsilon(\tau; v_0)) - \mathcal{D}(a^0(\tau; v_0))\|^*_{L,h^{s_*}} = 0,$$

where $a^\varepsilon(\tau; v_0)$ and $a^0(\tau; v_0)$ solve respectively equations (2.6) and (2.9) with initial conditions $a^\varepsilon(0; v_0) = a^0(0; v_0) = v_0$. Moreover, for any $M > 0$ the above convergence is uniform for $v_0 \in \bar{B}_M(h^{\bar{s}})$.

For $v = (v_1, v_2, \dots) \in h^s$ we introduce the vector of action variables $I(v) = (I_1(v), I_2(v), \dots)$, where $I_k(v) = \frac{1}{2}|v_k|^2$, $k = 1, 2, \dots$. Then $I(v) \in h^s_I \cap \mathbf{R}^\infty_+$, where h^s_I is the weighted l^1-space with the norm $|I|_{I,s} = \sum_{k=1}^\infty (|\lambda_k|^s + 1)|I_k|$. Since the interaction representation does not change the actions, then for the action variables of solutions for the original equations (2.5) we have

Corollary 2.8. *Under the assumptions of Theorem 2.7 actions of a solution $v^\varepsilon(\tau; v_0)$ for eq. (2.5) with $v_0 \in h^{\bar{s}}$, $\bar{s} > s_*$ satisfy*

$$\lim_{\varepsilon \to 0} \sup_{\tau \geqslant 0} \|\mathcal{D}\big(I(v^\varepsilon(\tau; v_0))\big) - \mathcal{D}\big((I(a^0(\tau; v_0))\big)\|^*_{L,h^{s_*}_I} = 0.$$

2.1. **Examples.** For a finite-dimensional stochastic equation

$$\dot a_k(\tau) = e^{i\varepsilon^{-1}\lambda_k\tau} P_k(\Phi_{-\varepsilon^{-1}\Lambda\tau} a) + e^{i\varepsilon^{-1}\lambda_k\tau}\sum_{l=1}^{N}\mathcal{M}_{kl}\dot\beta_l(\tau), \quad k=1,\dots,N, \tag{2.14}$$

where $a(\tau) = (a_1,\dots,a_N)(\tau) \in \mathbb{C}^N$ and $\Lambda = (\lambda_1,\dots,\lambda_N)$, satisfying a natural analogy of Assumptions 2.2, 2.6, a natural version of Theorem 2.7 holds. If the matrix $(\mathcal{M}_{kl})$ is non-degenerate and Assumption 2.2 is met, then a convenient sufficient condition for Assumption 2.6 follows from the Khasminski criterion for mixing (details will be given in paper on the finite-dimensional stochastic averaging under preparation).[3] For infinite-dimensional stochastic systems an instrumental criterion of mixing is not known yet. We believe that as in finite dimensions, Assumption 2.6 holds for effective equations for various equations (2.3), satisfying Assumption 2.2, at least if the random forces (2.2) is such that all b_l's are non-zero. Examples below are given to support this belief.

Example 2.9. Consider eq. (2.3), where $\mathscr{P}(\nabla u, u) = \mathscr{P}(u)$ has the form (2.4) with $\mathfrak{z} = -1$. Since the function f_p is monotone, then mapping $\mathscr{P}$ also is monotone in the sense that

$$\langle \mathscr{P}(u) - \mathscr{P}(v), u - v\rangle \ge 0 \quad \forall u, v \in L^\infty(T^D;\mathbb{C}).$$

From definition (2.10) we see that operator R is monotone as well. Taking any two vectors $v^1, v^2 \in h^s$ and denoting $w(\tau) = a^0(\tau; v^1) - a^0(\tau; v^2)$ we immediately derive from the equation for w that $(d/dt)|w^\omega(\tau)|_0^2 \le -\frac{1}{2}\lambda_1 |w^\omega(\tau)|_0^2$. So

$$|w^\omega(\tau)|_0 \le e^{-\lambda_1\tau}|v^1 - v^2|_0 \quad \forall\, \tau \ge 0,\ \forall \omega.$$

This very strong a-priori estimate implies Assumption 2.6 after some non-complicated work.

Example 2.10. Now let $\mathscr{P}(\nabla u, u) = \mathscr{P}(u)$ has the form (2.4), where $\Re\mathfrak{z} < 0$ and $\Im\mathfrak{z} < 0$. Then relation (2.13) with $\mathfrak{g}_M(\tau) = C(M)e^{-\kappa\tau}$, $\kappa > 0$, and with sufficiently large integer s_* follows from the abstract theorems in [10, Theorem 2.1] and [5, Theorem 3.1.7] since solutions a^0 inherit estimates for solutions a^ε via convergence (2.12). We leave to the reader details of this derivation (which are not quite trivial).

Example 2.11. Consider eq. (2.3) with $\mathscr{P}(\nabla u, u) = -i\rho|u|^2 u$, with $V = 0$ and with a hyperviscosity instead of viscosity:

$$\dot u - i\varepsilon^{-1}\Delta u = -(-\Delta + 1)^r u - i\rho|u|^2 u + \dot\xi(\tau, x), \quad x \in \mathbb{T}^D_L = \mathbb{R}^D/(L\mathbb{Z}^D). \tag{2.15}$$

Here $r \in \mathbb{N}$, $\rho > 0$ is a scaling factor and ξ has the form (2.2), where all b_l's are non-zero. Now $A_V = -\Delta$, so its eigenfunctions are exponents $e^{i2\pi s\cdot x}$, $s \in L^{-1}\mathbb{Z}^D$. The action variables $I_s(\tau)$ of solutions $u(\tau, x)$ are naturally parametrised by $s \in L^{-1}\mathbb{Z}^D$. The theory of wave turbulence (WT), among other things, examine the behaviour of the expectations of the actions $\mathbf{E}I_s(\tau)$ under the limit of WT:

$$\varepsilon \to 0, \quad L \to \infty,$$

when ρ is properly scaled with ε and L. See [9] and the introduction in [3]. For any dimension D solutions of eq. (2.15) with r sufficiently large in terms of D satisfy Assumption 2.2 by the same straightforward proof as for eq. (2.3) with $\mathscr{P}$ as in (2.4), and the equation also meets Assumption 2.6 with sufficiently large integer s_* in view of the abstract theorems, mentioned in Example 2.10. So Corollary 2.8 applies to eq. (2.15) with ρ and L fixed, under the limit $\varepsilon \to 0$ and imply that the expectations of the actions of solutions for (2.15) converge to those of solutions

[3] In [2, Theorem 2.9] the uniform in time convergence as above is proved for a class of systems (2.14). The proof in [2] is based on the observation that the corresponding v-equations are mixing with a rate of mixing, independent from ε.

for the corresponding effective equation, uniformly in time. The limit $\varepsilon\to 0$ in eq. (2.15) and in similar equations is known in nonlinear physics as the limit of *discrete turbulence*. See [9], [6] and [3, Sections 1.2, 12.1].

3. Proof of Theorem 2.7

Below we always assume Assumptions 2.2 and 2.6. We fix $\bar s > s_*$ and $v_0\in h^{\bar s}$ as in Theorem 2.7. We abbreviate $\|\cdot\|^*_{L,h^{s_*}}$ to $\|\cdot\|^*_L$ and $a^\epsilon(\tau;v_0)$ to $a^\epsilon(\tau)$. For any $T'\geq 0$ we denote by $a^0_{T'}(\tau)$ a weak solution of the effective equation (2.9) such that

$$\mathcal{D}(a^0_{T'}(0))=\mathcal{D}(a^\varepsilon(T')).$$

Note that solution $a^0_{T'}(\tau)$ depends on ε and that $a^0_0(\tau)=a^0(\tau;v_0)$.

The following lemma follows from Theorem 2.3, Assumption 2.2.(2) and Remark 2.4.2).

Lemma 3.1. *For any $\delta>0$ and $T>0$ there exists $\varepsilon_1=\varepsilon_1(\delta,T)>0$ such that if $\varepsilon\le\varepsilon_1$, then*

$$\sup_{\tau\in[0,T]}\|\mathcal{D}(a^\varepsilon(T'+\tau))-\mathcal{D}(a^0_{T'}(\tau))\|^*_{L,h^{s_*}}\leqslant\delta/2,\quad \forall\, T'\geqslant 0. \tag{3.1}$$

Since by Assumption 2.6, $\mathcal{D}(a^0(\tau;0))\to\mu^0$ in $\mathcal{P}(h^{s_*})$ as $\tau\to\infty$, then from estimate (2.11) and Fatou's lemma we derive that

$$\langle |v|^{2m'}_{\bar s},\mu^0\rangle\leqslant C'_{\bar s}(0)\coloneqq C_0. \tag{3.2}$$

We need the following statement for the mixing in eq. (2.9).

Lemma 3.2. *For any solution $a^0(\tau;\mu)\in h^{s*}$, $\tau\geqslant 0$, of effective equation (2.9) such that $\mathcal{D}(a^0(0;\mu))=:\mu$, where the measure μ satisfies*

$$\langle |a|^{2m'}_{s_*},\mu(da)\rangle=\mathbf{E}|a^0(0;\mu)|^{2m'}_{s_*}\leqslant M\qquad \text{for some } M>0,$$

we have

$$\|\mathcal{D}\big(a^0(\tau;\mu)\big)-\mu^0\|^*_L\le g_M(\tau,d)\quad \text{if } \|\mu-\mu^0\|^*_L\le d\le 2. \tag{3.3}$$

Here the function $g:\mathbb{R}^3_+\to\mathbb{R}_+$, $(M,\tau,d)\to g_M(\tau,d)$ is continuous, vanishes with d, converges to zero when $\tau\to\infty$ and is such that for each fixed $M\geqslant 0$ the function $(\tau,d)\to g_M(\tau,d)$ is uniformly continuous in d for $(\tau,d)\in[0,\infty)\times[0,2]$.

A proof of the lemma is rather straightforward but lengthy. It is given below in Subsection 3.1.

We denote

$$M_*\coloneqq C'_{\bar s}(|v_0|_{\bar s}),$$

where $C'_{\bar s}(|v_0|_{\bar s})$ is as in (2.8) and (2.11). Constants in the estimates below depend on M_*, but this dependence usually is not indicated.

Lemma 3.3. *Take any $\delta>0$ and choose a $T^*=T^*(\delta)>0$, satisfying*

$$g_{M_*}(T,2)\le\delta/4\qquad\forall\, T\ge T^*.$$

Then there exists $\varepsilon_2=\varepsilon_2(\delta)>0$ such that if $\varepsilon\le\varepsilon_2$ and $\|\mathcal{D}(a^\varepsilon(T'))-\mu^0\|^*_L\leqslant\delta$ for some $T'\geq 0$, then

$$\|\mathcal{D}(a^\varepsilon(T'+T^*))-\mu^0\|^*_L\leqslant\delta, \tag{3.4}$$

and

$$\sup_{\tau\in[T',T'+T^*]}\|\mathcal{D}(a^\varepsilon(\tau))-\mu^0\|^*_L\leqslant\tfrac{\delta}{2}+\sup_{\tau\ge 0}g_{M_*}(\tau,\delta). \tag{3.5}$$

Proof. Let us choose $\varepsilon_2(\delta) = \varepsilon_1(\frac{\delta}{2}, T^*(\delta))$, where $\varepsilon_1(\cdot)$ is as in Lemma 3.1. Then we get from (3.1), (2.11), (3.3) with $M = M_*$ and the definition of T^* that for $\varepsilon \le \varepsilon_2$,

$$\|\mathcal{D}(a^\varepsilon(T'+T^*)) - \mu^0\|_L^* \leqslant \|\mathcal{D}(a^\varepsilon(T'+T^*)) - \mathcal{D}(a^0_{T'}(T^*))\|_L^* + \|\mathcal{D}(a^0_{T'}(T^*)) - \mu^0\|_L^* \leqslant \delta.$$

This proves (3.4). Next, in view of (3.1) and (3.3),

$$\begin{aligned}&\sup_{\theta\in[0,T^*]} \|\mathcal{D}(a^\varepsilon(T'+\theta)) - \mu^0\|_L^* \\ &\leqslant \sup_{\theta\in[0,T^*]} \|\mathcal{D}(a^\varepsilon(T'+\theta)) - \mathcal{D}(a^0_{T'}(\theta))\|_L^* + \sup_{\theta\in[0,T^*]} \|\mathcal{D}(a^0_{T'}(\theta)) - \mu^0\|_L^* \\ &\leqslant \frac{\delta}{2} + \max_{\theta\in[0,T^*]} g_{M_*}(\theta,\delta).\end{aligned}$$

This implies (3.5). □

We are now ready to prove Theorem 2.7.

Proof. [of Theorem 2.7.] Let us fix arbitrary $\delta > 0$ and take some $0 < \delta_1 \le \delta/4$. Below in the proof the functions ε_1, ε_2 and T^* are as in Lemmas 3.1 and 3.3.

i) By the definition of $T^* = T^*(\delta_1)$, (2.11) and (3.3),

$$\|\mathcal{D}\big(a^0_{T'}(\tau)\big) - \mu^0\|_L^* \le \delta_1/4 \quad \forall\, \tau \ge T^*, \tag{3.6}$$

for any $T' \ge 0$.

ii) By Lemma 3.1, if $\varepsilon \le \varepsilon_1 = \varepsilon_1(\frac{\delta_1}{2}, T^*) > 0$, then

$$\sup_{0\le\tau\le T^*} \|\mathcal{D}\big(a^\varepsilon(\tau)\big) - \mathcal{D}\big(a^0(\tau;v_0)\big)\|_L^* \le \tfrac{\delta_1}{2}. \tag{3.7}$$

In particular, in view of (3.6) with $T' = 0$,

$$\|\mathcal{D}\big(a^\varepsilon(T^*)\big) - \mu^0\|_L^* < \delta_1. \tag{3.8}$$

iii) By (3.8) and (3.4) with $\delta = \delta_1$ and with $T' = nT^*$, $n = 1, 2, \dots$ we get inductively that

$$\|\mathcal{D}\big(a^\varepsilon(nT^*)\big) - \mu^0\|_L^* \le \delta_1 \quad \forall\, n \in \mathbb{N}, \tag{3.9}$$

if $\varepsilon \le \varepsilon_2 = \varepsilon_2(\delta_1)$.

iv) Now by (3.9) and (3.5) with $\delta = \delta_1$, for any $n \in \mathbb{N}$ and $0 \le \theta \le T^*$,

$$\|\mathcal{D}\big(a^\varepsilon(nT^*+\theta)\big) - \mu^0\|_L^* \le \delta_1/2 + \sup_{\theta\ge0} g_{M_*}(\theta,\delta_1), \tag{3.10}$$

if $\varepsilon \le \varepsilon_2(\delta_1)$.

v) Finally, if $\varepsilon \le \varepsilon_\#(\delta_1) = \min\{\varepsilon_1, \varepsilon_2\}$, then by (3.7) if $\tau \le T^*$ and by (3.6)+(3.10) if $\tau \ge T^*$ we have that

$$\|\mathcal{D}\big(a^\varepsilon(\tau)\big) - \mathcal{D}\big(a^0(\tau;v_0)\big)\|_L^* \le \delta_1 + \sup_{\theta\ge0} g_{M_*}(\theta,\delta_1) \qquad \forall\, \tau \ge 0.$$

By Lemma 3.2, for M_* fixed the function $g_{M_*}(\theta,d)$ is uniformly continuous in d and vanishes at $d = 0$. So there exists $\delta^* = \delta^*(\delta)$, which we may assume to be $\le \delta/4$, such that if $\delta_1 = \delta^*$, then $g_{M_*}(\theta,\delta_1) \le \delta/2$ for every $\theta \geqslant 0$. Then by the estimate above,

$$\|\mathcal{D}\big(a^\varepsilon(\tau)\big) - \mathcal{D}\big(a^0(\tau;v_0)\big)\|_L^* \le \delta \quad \text{if} \quad \varepsilon \le \varepsilon_*(\delta) := \varepsilon_\#\big(\delta^*(\delta)\big) > 0,$$

for every positive δ. This proves the theorem's assertion. □

3.1. **Proof of Lemma 3.2.** In this proof we write solutions of effective equation (2.9) as $a(\tau)$ (rather than $a^0(\tau)$).

i) At this step, for any non-random $v^1, v^2 \in \bar B_M(h^{s_*})$ we denote $a_j^\omega(\tau) = a^\omega(\tau, v^j)$, $j = 1, 2$, and examine the distance $\|\mathcal{D}(a_1(\tau)) - \mathcal{D}(a_2(\tau))\|_L^*$ as a function of τ and $|v^1 - v^2|_{s_*}$. We assume that $|v^1 - v^2|_{s_*} \le \bar d$ for some $\bar d \ge 0$ and set $w^\omega(\tau) = a_1^\omega(\tau) - a_2^\omega(\tau)$. We have

$$\dot w^\omega = -A_V w^\omega + R(a_1^\omega) - R(a_2^\omega).$$

So by Duhamel's principle,

$$w^\omega(\tau) = e^{-A_V\tau} w(0) + \int_0^\tau e^{-A_V(\tau - t)} (R(a_1^\omega(t)) - R(a_2^\omega(t)))dt.$$

Here $|w(0)|_{s_*} \le \bar d$, and by (1) of Assumption 2.2

$$|R(a_1^\omega(\tau)) - R(a_2^\omega(\tau))|_{s_*-1} \le C|w^\omega(\tau)|_{s_*} X^\omega(\tau), \quad X^\omega(\tau) = 1 + |a_1^\omega(\tau)|_{s_*}^{\bar m} + |a_2^\omega(\tau)|_{s_*}^{\bar m}.$$

For $\theta > 0$ the norm of the operator $e^{-A_V\theta} : H^{s_*-1} \to H^{s_*}$ is bounded by $\chi(\theta)$, where

$$\chi(\theta) = \begin{cases} C\theta^{-1/2}, & 0 < \theta \le 1, \\ Ce^{-c\theta}, & \theta > 1, \end{cases}$$

with some $C, c > 0$. So if $\max_{0\le t\le\tau} X^\omega(t) \le K$, then

$$|w^\omega(\tau)|_{s_*} \le \bar d e^{-\lambda_1 \tau} + KC \int_0^\tau \chi(\tau - l)|w^\omega(l)|_{s_*} dl.$$

Applying Gronwall's lemma we derive from here that

$$|w^\omega(\tau)|_{s_*} \le \bar d(1 + CKe^{c_1 K}), \quad \forall \tau \ge 0. \tag{3.11}$$

for some positive constants C, c_1.

Denote $Y(T) = \sup_{0\le t\le T} |X^\omega(t)|$. By (2.11),

$$\mathbf{E} Y(T) \leqslant 2(2 + C'_{s_*}(M))(T + 1).$$

For $K > 0$ let $\Omega_K(T)$ be the event $\{Y(T) \ge K\}$. Then $\mathbf{P}(\Omega_K(T)) \le 2(2 + C'_{s_*}(M))(T+1)K^{-1}$, and $|a_1(\tau) - a_2(\tau)|_{s_*} = |w^\omega(\tau)|_{s_*}$ satisfies (3.11) for $\omega \notin \Omega_K(T)$. From here we see that if $f \in C_b(h^{s_*})$ is such that $|f|_{C_b(h^{s_*})} \le 1$ and $\mathrm{Lip}\, f \le 1$, then

$$\begin{aligned} \mathbf{E}\big(f(a_1(\tau)) - f(a_2(\tau))\big) &\le 2\mathbf{P}(\Omega_K(\tau)) + \bar d(1 + CKe^{c_1K}) \\ &\leqslant 4(2 + C'_{s_*}(M))(\tau + 1)K^{-1} + \bar d(1 + CKe^{c_1K}) \quad \forall K > 0. \end{aligned} \tag{3.12}$$

Let us denote by $g_M^1(\tau, \bar d)$ the function in the r.h.s. above with $K = \ln\ln\big(\bar d^{-1} \vee 3\big)$. This is a continuous function of $(M, \tau, \bar d) \in \mathbb{R}_+^3$, vanishing when $\bar d = 0$. Due to (2.13) and (3.12),

$$\begin{aligned} \|\mathcal{D}(a(\tau; v^1)) - \mathcal{D}(a(\tau; v^2))\|_L^* &= \|\mathcal{D}(a_1(\tau)) - \mathcal{D}(a_2(\tau))\|_L^* \\ &\le (2\mathfrak{g}_M(\tau)) \wedge g_M^1(\tau, \bar d) \wedge 2 =: g_M^2(\tau, \bar d), \end{aligned} \tag{3.13}$$

if $v_1, v_2 \in \bar B_M(h^{s_*})$ and $|v^1 - v^2|_{s_*} \le \bar d$, for any $M, \bar d > 0$. The function g^2 is continuous in the variables $(M, \tau, \bar d)$, vanishes with $\bar d$ and goes to zero when $\tau \to \infty$ since $\mathfrak{g}_M(\tau)$ does.

ii) At this step we consider a solution $a^0(\tau; \mu) =: a(\tau; \mu)$ of (2.9) as in the lemma and examine the l.h.s. of (3.3) as a function of τ. For any $K > 0$ consider the conditional probabilities $\mu_K = \mu(\cdot \mid \bar B_K(h^{s_*}))$ and $\bar\mu_K = \mu(\cdot \mid h^{s_*} \smallsetminus \bar B_K(h^{s_*}))$. Then $\mu = A_K\mu_K + \bar A_K \bar\mu_K$, where $A_K = \mu(\bar B_K(h^{s_*}))$ and $\bar A_K = \mathbf{P}\{|a(0)|_{s_*} > K\} \leqslant M/K^{2m'}$ as $\mathbf{E}|a(0)|_{s_*}^{2m'} \le M$. So

$$\mathcal{D}(a(\tau, \mu)) = A_K \mathcal{D}(a(\tau; \mu_K)) + \bar A_K \mathcal{D}(a(\tau; \bar\mu_K)). \tag{3.14}$$

In view of Assumption 2.6,

$$\|\mathcal{D}(a(\tau;\mu_K))-\mu^0\|_L^* = \|\int \big[\mathcal{D}(a(\tau;v))\big]\mu_K(dv)-\mu^0\|_L^*$$
$$\leq \int \|\mathcal{D}(a(\tau;v))-\mu^0\|_L^*\mu_K(dv)\leq \mathfrak{g}_K(\tau).$$

Therefore due to (3.14),

$$\|\mathcal{D}(a(\tau,\mu))-\mu^0\|_L^* \leqslant A_K\|\mathcal{D}(a(\tau,\mu_K))-\mu^0\|_L^* + \bar{A}_K\|\mathcal{D}(a(\tau,\bar{\mu}_K))-\mu^0\|_L^*$$
$$\leq \|\mathcal{D}(a(\tau,\mu_K))-\mu^0\|_L^* + 2\bar{A}_K \leq \mathfrak{g}_K(\tau)+2\frac{M}{K^{2m'}} \quad \text{for } K>0 \text{ and } \tau\geq 0.$$

Let $K_1(\tau)>0$ be a continuous non-decreasing function such that $K_1(\tau)\to\infty$ and $\mathfrak{g}_{K_1(\tau)}(\tau)\to 0$ as $\tau\to\infty$ (it exists since $\mathfrak{g}_K(\tau)$ is a continuous function of (K,τ), going to 0 as $\tau\to\infty$ for each fixed K). Then choosing in the estimate above $K=K_1$ we get

$$\|\mathcal{D}(a(\tau;\mu))-\mu^0\|_L^* \leqslant \mathfrak{g}_{K_1(\tau)}(\tau)+\frac{2M}{K_1(\tau)^{2m'}} =: \hat{\mathfrak{g}}_M(\tau). \tag{3.15}$$

Obviously $\hat{\mathfrak{g}}_M(\tau)\geq 0$ is a continuous function on $\mathbb{R}^2_+$, converging to 0 as $\tau\to\infty$.

iii) Now we examine the l.h.s. of (3.3) as a function of τ and d. Recall that the Kantorovich distance between measures ν_1,ν_2 on h^{s_*} is

$$\|\nu_1-\nu_2\|_K = \sup_{\mathrm{Lip} f\leq 1}\langle f,\nu_1\rangle - \langle f,\nu_2\rangle \leq \infty.$$

Obviously $\|\nu_1-\nu_2\|_L^* \leq \|\nu_1-\nu_2\|_K$. Since the $2m'$-th moments of μ and μ^0 are bounded by $M\vee C_0$ by (3.2) and the assumption on μ and since $\|\mu-\mu^0\|_L^*\leq d$, then

$$\|\mu-\mu^0\|_K \leq \tilde{C}(M\vee C_0)^{\gamma_1}d^{\gamma_2} := \tilde{d}, \quad \gamma_1=\tfrac{1}{2m'},\ \gamma_2=\tfrac{2m'-1}{2m'}; \tag{3.16}$$

see [1, Section 11.4] and [11, Chapter 7]. By the Kantorovich–Rubinstein theorem (see [11, 1]) there exist r.v.'s ξ and ξ_0, defined on a new probability space $(\Omega',\mathcal{F}',\mathbf{P}')$, such that $\mathcal{D}(\xi)=\mu$, $\mathcal{D}(\xi_0)=\mu^0$ and

$$\mathbf{E}|\xi_1-\xi_0|_{s_*} = \|\mu-\mu^0\|_K. \tag{3.17}$$

Then using (3.13) and denoting by $a_{st}(\tau)$ a stationary solution of (2.9), $\mathcal{D}(a_{st}(\tau))\equiv\mu^0$, we have:

$$\|\mathcal{D}(a(\tau))-\mu^0\|_L^* = \|\mathcal{D}a(\tau;\mu^0)-\mathcal{D}(a_{st}(\tau))\|_L^* \leq \mathbf{E}^{\omega'}\|\mathcal{D}(a(\tau;\xi^{\omega'}))-\mathcal{D}(a(\tau;\xi_0^{\omega'}))\|_L^*$$
$$\leq \mathbf{E}^{\omega'}g^2_{\bar{M}}(\tau,|\xi^{\omega'}-\xi_0^{\omega'}|_{s_*}), \qquad \bar{M}=\bar{M}^{\omega'}=|\xi^{\omega'}|_{s_*}\vee|\xi_0^{\omega'}|_{s_*}.$$

As $\mathbf{E}^{\omega'}\bar{M}^{2m'}\leq 2(M\vee C_0)$, then for any $K>0$,

$$\mathbf{P}^{\omega'}(Q'_K)\leq 2K^{-2m'}(M\vee C_0), \qquad Q'_K=\{\bar{M}\geq K\}\subset\Omega'.$$

Since $g^2\leq 2$ and for $\omega'\notin Q'_K$ we have $|\xi^{\omega'}|_{s_*},|\xi_0^{\omega'}|_{s_*}\leq K$, then

$$\|\mathcal{D}(a(\tau))-\mu^0\|_L^* \leq 4K^{-2m'}(M\vee C_0)+\mathbf{E}^{\omega'}g^2_K(\tau,|\xi^{\omega'}-\xi_0^{\omega'}|_{s_*}) \quad \forall K>0.$$

For an $r>0$ let us denote $\Omega'_r=\{|\xi^{\omega'}-\xi_0^{\omega'}|_{s_*}\geq r\}$. Then by (3.17) and (3.16), $\mathbf{P}^{\omega'}\Omega'_r\leq\tilde{d}r^{-1}$. So

$$\|\mathcal{D}(a(\tau))-\mu^0\|_L^* \leq 4K^{-2m'}(M\vee C_0)+2\tilde{d}r^{-1}+g^2_K(\tau,r), \qquad \text{for any } K,r>0. \tag{3.18}$$

Let $g_0(l)$ be a positive continuous function on $\mathbb{R}_+$ such that $g_0(l)\to\infty$ as $l\to+\infty$ in such a way that $|C'_{s_*}\big(g_0(l)\big)(\ln\ln l)^{-1/2}|\leqslant 2C'_{s_*}(0)$ for $l\geqslant 3$. With $r=\tilde{d}^{1/2}$ and

$K = g_0(r^{-1})$, we denote the r.h.s of (3.18) as $g^3_M(\tau, r)$ (so we substitute in (3.18) $\tilde d = r^2$ and $K = g_0(r^{-1})$). By (3.18) and the definition of g^2 (see (3.13)), we have

$$\|\mathcal{D}(a(\tau)) - \mu^0\|^*_L \le g^3_M(\tau, r) \leqslant 4(g_0(r^{-1}))^{-2m'}(M \vee C_0) + 2r$$
$$+ 4(\tau+1)\big(1 + C'_{s_*}\big(g_0(r^{-1})\big)\big)\big(\ln\ln(r^{-1}\vee 3)\big)^{-1} + rC\ln\ln(r^{-1}\vee 3)\exp\big(c_1\ln\ln(r^{-1}\vee 3)\big).$$

As $r \to 0$ the second and fourth terms converge to zero. By the choice of g_0, the first term clearly converges to zero with r, so does the third term, which is $\leqslant 8(1+\tau)(1+C'_{s_*}(0))(\ln\ln(r^{-1}))^{-1/2}$ for $r \leqslant \frac{1}{3}$. Hence $g^3_M(\tau, r)$ defines a continuous function on $\mathbb{R}^3_+$, vanishing with r. Using (3.16) let us write $r = \tilde d^{1/2}$ as $r = R_M(d)$, where R is a continuous function $\mathbb{R}^2_+ \to \mathbb{R}_+$, non-decreasing in d and vanishing with d. Setting $g^4_M(\tau, d) = g^3_M(\tau, R_M(d \wedge 2))$ and using that $\|\mu - \mu^0\|^*_L \le 2$, we get from the above that

$$\|\mathcal{D}(a(\tau)) - \mu^0\|^*_L \le g^4_M(\tau, \|\mu - \mu^0\|^*_L).$$

Finally, evoking (3.15) we arrive at (3.3) with $g_M = g^5_M$, where

$$g^5_M(\tau, d) = g^4_M(\tau, d) \wedge \hat{\mathfrak{g}}_M(\tau) \wedge 2, \quad 0 \le d \le 2.$$

The function g^5 is continuous, vanishes with d and converges to zero as $\tau \to \infty$. For any fixed $M > 0$ this convergence is uniform in d due to the term $\hat{\mathfrak{g}}_M(\tau)$. So for a fixed $M > 0$ the function $(\tau, d) \mapsto g^5_M(\tau, d)$ extends to a continuous function on the compact set $[0, \infty] \times [0, 2]$ (where it vanishes when $\tau = \infty$). Thus g^5_M is uniformly continuous in d, and the lemma is proved.

4. NLW: THE SETTING AND RESULT

In this section we briefly discuss stochastic nonlinear wave (NLW) equations. Following [7, 8] we consider the following equations on a smooth bounded domain $\mathfrak{D} \subset \mathbb{R}^3$:

$$\partial_t^2 u + \gamma\partial_t u - \Delta u = -\gamma f(u) + \gamma h(x) + \sqrt{\gamma}\,\eta(t, x), \quad x \in \mathfrak{D}, \tag{4.1}$$

supplemented with the Dirichlet boundary condition on $\partial\mathfrak{D}$. Here $\gamma \in (0, 1]$ is a small parameter, $h(x)$ is a function in $H^1_0(\mathfrak{D}; \mathbb{R})$ and the nonlinearity f is C^2-smooth. The random force $\eta(t, x)$ is a white noise of time of the form

$$\eta(t, x) = \frac{\partial}{\partial t}\sum_{j=1}^{\infty} b_j \beta^R_j(t)\mathfrak{e}_j(x).$$

Here $\{\beta^R_j(t), j \geqslant 1\}$ is a sequence of independent standard real Brownian motions, $\{\mathfrak{e}_j(x), j \geqslant 1\}$ is an orthonormal basis in $L^2(\mathfrak{D}; \mathbb{R})$ composed of eigenfunctions of the Laplacian operator; that is $-\Delta\mathfrak{e}_j = \lambda_j\mathfrak{e}_j$ with $0 < \lambda_1 \leqslant \lambda_2 \leqslant \dots$. The set $\{b_j, j \geqslant 1\}$ is a sequence of positive real numbers, satisfying

$$\mathcal{B} := \sum_{j \ge 1} \lambda_j b_j^2 < +\infty. \tag{4.2}$$

The nonlinear term f meets the following growth condition,

$$|f''(u)| \leqslant C(|u|^{\rho-1} + 1), \quad u \in \mathbb{R}, \tag{4.3}$$

where C and $\rho < 2$ are positive constants, as well as the dissipativity conditions

$$F(u) \geqslant -\kappa u^2 - C, \quad u \in \mathbb{R}, \tag{4.4}$$

$$f(u)u - F(u) \geqslant -\kappa u^2 - C, \; u \in \mathbb{R}, \tag{4.5}$$

where F is the primitive of f and κ is a positive constant.

Lemma 4.1. *([7, Lemma 4.5]). Under the condition (4.3) we have*

$$\|f(u) - f(v)\|_0^2 \leqslant C(\|u\|^{2\rho}_{1-s_\rho} + \|v\|^{2\rho}_{1-s_\rho} + 1)\|u - v\|^2_{1-s_\rho}, \quad s_\rho = \tfrac{2-\rho}{2(\rho+1)}.$$

Let us denote $\partial_t u = Lv$, where $L = (-\Delta)^{1/2}$, and set $\xi = u + iv$. Then in terms of ξ equation (4.1) reads

$$\partial_t \xi = iL\xi - \gamma i \Im\xi - \gamma i L^{-1} f(\Re\xi) - \gamma i L^{-1}\eta(t,x).$$

Introducing the slow time $\tau = \gamma t$, we have

$$\partial_\tau \xi = \gamma^{-1} iL\xi - i\Im\xi - iL^{-1} f(\Re\xi) - i\tilde\eta(\tau,x), \tag{4.6}$$

where $\tilde\eta(\tau,x) = \sum_{j=1}^{\infty} \tilde b_j \dot\beta_j^R(\tau)\mathfrak{e}_j(x)$ with $\tilde b_j = b_j\lambda_j^{-1/2}$, $j \geqslant 1$.

The NLW equation (4.1) (without the γ-factor in the r.h.s) was studied in [7, 8], where the global well-posedness and estimates for the norms of solutions are established in the Sobolev space $H^s \times H^{s-1} = \{(u,\dot u)\}$, $s \in [1, 2-\frac{\varrho}{2}]$. A simple analysis of the proof of [7, Proposition 3.4] or [8, Proposition 5.4] implies the following statement on the global well-posedness of eq. (4.6):

Theorem 4.2. *Assume conditions (4.3), (4.4) and (4.5). Then for any $s \in [1, 2-\frac{\varrho}{2})$ and $\xi_0 \in H^s(\mathfrak{D};\mathbb{C})$, equation (4.6) has a unique strong solution $\xi^\gamma(\tau;\xi_0)$, equal to ξ_0 at $\tau = 0$ and defined for $\tau \geqslant 0$. Uniformly in $\gamma \in (0,1]$ this solution satisfies*

$$\mathbf{E} \sup_{\tau\in[\theta,\theta+1]} \|\xi^\gamma(\tau;\xi_0)\|_s^{2m} \leqslant C_s(m, \|\xi_0\|_s, \mathcal{B}), \quad \forall \theta \geqslant 0,\ m \in \mathbb{N}, \tag{4.7}$$

where C_s is a continuous function, non-decreasing in all its arguments.

Now let us write eq. (4.6) in terms of Fourier coefficients with respect to the basis $\{\mathfrak{e}_j(x)\}$: $\xi = \sum_{j\geqslant 1} v_j \mathfrak{e}_j$. As in Section 2, we use h^s to denote the space of sequences of complex Fourier coefficients and denote by Ψ the map $H^s(\mathfrak{D};\mathbb{C}) \to h^s$, $\xi \mapsto v := (v_j, j \geqslant 1)$. Then we have

$$\dot v_k = \gamma^{-1} i\lambda_k^{1/2} v_k - i\Im v_k + i\hat P_k(v) + i\tilde b_k \beta_k^R(\tau), \quad k \in \mathbb{N}, \tag{4.8}$$

where $\hat P(v) = (\hat P_k(v), k \in \mathbb{N}) = \Psi(L^{-1} f(\Re\xi))$ with $\xi = \Psi^{-1} v$. Passing to the interaction representation

$$a_k(\tau) = e^{-i\gamma^{-1}\lambda_k^{1/2}\tau} v_k(\tau), \quad k \geqslant 1,$$

we obtain the following system of equations for the a-variables

$$\dot a_k(\tau) = ie^{-i\gamma^{-1}\lambda_k^{1/2}\tau}\Big(-\Im\big(e^{i\gamma^{-1}\lambda_k^{1/2}\tau} a_k\big) + \hat P_k(\Phi_{\gamma^{-1}\hat\Lambda\tau} a) + \tilde b_k \dot\beta_k^R\Big), \quad k \in \mathbb{N}, \tag{4.9}$$

where $\hat\Lambda = (\lambda_k^{1/2}, k \geqslant 1)$ and operator Φ is defined in (2.7). Let us calculate an effective equation for (4.9), following [4].

i) To calculate the effective diffusion we decomplexify components $a_k = a_k^R + ia_k^I$ of the a-vector as $(a_k^R + a_k^I) \in \mathbb{R}^2$ and write the dispersion matrix of the equation as a block-diagonal real matrix with the blocks $\tilde b_k \begin{pmatrix} \cos\varphi_k & -\sin\varphi_k \\ 0 & 0 \end{pmatrix}$, where $\varphi_k = \varphi^k(\tau) = \gamma^{-1}\lambda_k^{1/2}\tau$ for $k = 1, 2, \dots$. So the diffusion matrix of eq. (4.9) – let us call it $A(\tau)$ – is formed by blocks $\tilde b_k^2 \begin{pmatrix} \cos^2\varphi_k & -\cos\varphi_k\sin\varphi_k \\ -\cos\varphi_k\sin\varphi_k & \sin^2\varphi_k \end{pmatrix}$. Consider the limit $\mathcal{A}(\tau) = \lim_{\gamma\to 0} \frac{1}{\tau}\int_0^\tau A(l)\,dl$. This is a τ–independent block-diagonal real matrix with the blocks $\frac{1}{2}\tilde b_k^2\,\mathrm{id}$, and this is the diffusion matrix of the effective equation. Corresponding dispersion matrix has blocks $\frac{1}{\sqrt2}\tilde b_k\,\mathrm{id}$. Coming back to the complex coordinates we see that the noise in the effective equation is $\frac{1}{\sqrt2}\sum \tilde b_k\,\dot{\tilde\beta}_k(\tau)$, where $\{\tilde\beta_k(\tau)\}$ are standard independent complex Brownian motions.

ii) The drift in the effective equation is a sum of two terms. The second one is

$$\tilde R(a) = (\tilde R_k(a), k \geqslant 1) := \lim_{T\to\infty} \frac{1}{T}\int_0^T \Phi_{-\hat\Lambda t}\hat P(\Phi_{\hat\Lambda t} a)dt$$

(cf. (2.10)). The k-th component of the first term is

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T e^{-i\lambda_k^{1/2}t}\big(-i\Im(e^{i\lambda_k^{1/2}t}a_k)\big)dt = -i\tfrac{1}{2}(\Im a_k - i\Re a_k) = -\tfrac{1}{2}a_k.$$

iii) So the effective equation for (4.9) reads

$$\dot a_k(\tau) = -\tfrac{1}{2}a_k + i\tilde R_k(a) + \tfrac{1}{\sqrt{2}}\tilde b_k d\tilde\beta_k, \qquad k\in\mathbb{N}. \tag{4.10}$$

By Lemma 4.1 and [4, Lemma 4], we have

$$|\tilde R(a_1) - \tilde R(a_2)|_1 \leqslant C\big(|a_1|^\rho_{1-s_\rho}\wedge|a_2|^\rho_{1-s_\rho}+1\big)|a_1-a_2|_{1-s_\rho}, \quad a_1,a_2\in h^1. \tag{4.11}$$

In particular, the mapping $a\mapsto\tilde R(a)$ belongs to $\mathrm{Lip}_{\rho+1}(h^s,h^s)$, $s\in[s_\rho,1]$. Therefore effective equation (4.10) is at least locally well-posed in h^s, $s\in[s_\rho,1]$. Let us denote $a^0(\tau;a_0)$ its solution with an initial condition $a^0(0;a_0)=a_0$. Similarly let $a^\gamma(\tau;v_0)$ be a solution of eq. (4.9), equal a_0 at $\tau=0$. Then using (4.7) and arguing as in [4] we get

Theorem 4.3. *Under the assumptions of Theorem 4.2 we have:*

i) for any $a_0\in h^1$, eq. (4.10) has a unique strong solution $a^0(\tau;a_0)$, $\tau\geqslant 0$. It belongs to $C([0,\infty),h^1)$ a.s., and for any $\theta\geqslant 0$ satisfies

$$\sup\nolimits_{\tau\in[\theta,\theta+1]}|a^0(\tau;a_0)|_1^{2m}\leqslant C(m,|a|_1,\mathcal{B}), \quad \forall m\in\mathbb{N}, \tag{4.12}$$

where C is a continuous function, increasing in all its arguments.

ii) For any $s>1$, $a_0\in h^s$ and $T>0$ we have

$$\mathcal{D}\big(a^\gamma(\tau;a_0)|_{[0,T]}\big) \rightharpoonup \mathcal{D}\big(a^0(\tau;a_0)|_{[0,T]}\big) \quad \text{in } \mathcal{P}\big(C([0,T]);h^1)\big) \text{ as } \gamma\to 0. \tag{4.13}$$

Moreover, for any $M>0$ the above convergence is uniform for $a_0\in\bar B_M(h^s)$.

Repeating the argument in Section 3 we get the following analogy to Theorem 2.7:

Theorem 4.4. *Assume in addition to the assumption in Theorem 4.3 that the effective equation (4.10) is mixing in the space h^1 and that the requirement as in Assumption 2.6 is met. Then for any $s>1$ and $v_0\in h^s$ we have the following convergence*

$$\lim_{\gamma\to 0}\sup_{\tau\geqslant 0}\|\mathcal{D}(a^\gamma(\tau;v_0)) - \mathcal{D}(a^0(\tau;v_0))\|^*_{L,h^1} = 0.$$

Moreover, for any $M>0$ the above convergence is uniform for $v_0\in\bar B_M(h^s)$.

For the original equation (4.1) we introduce the vector of action variables for its solution $\mathfrak{u}=(u,u_t)$ as $I(\mathfrak{u})=(I_1(\mathfrak{u}),\dots)$, where $I_k(\mathfrak{u})=\frac{1}{2}|\langle u,\mathfrak{e}_k\rangle|^2+\frac{1}{2}\lambda_k^{-1}|\langle u_t,\mathfrak{e}_k\rangle|^2$, $k=1,\dots$. Then we have the following corollary in analogy to Corollary 2.8:

Corollary 4.5. *Under the assumption of Theorem 4.4, for any $s>1$ and $\mathfrak{u}_0=[u_1,u_2]\in H^s\times H^{s-1}$, the action-vector of a solution $\mathfrak{u}^\gamma(t;\mathfrak{u}_0)$ for equation (4.1), equal $\mathfrak{u}_0$ at $t=0$, satisfies*

$$\lim_{\gamma\to 0}\sup_{\tau\geqslant 0}\|\mathcal{D}\big(I(\mathfrak{u}^\gamma(\tau\gamma^{-1};\mathfrak{u}_0))\big) - \mathcal{D}\big((I(a^0(\tau;v_0))\big)\|^*_{L,h^1_I} = 0,$$

where $v_0=\Psi(u_1+iL^{-1}u_2)$.

The result of Theorem 4.4 is conditional since it requires that eq. (4.9) is mixing. It is not our goal in this short section to check the mixing property. But we mention that since the effective equation (4.10) is similar to the NLW equation (4.1), written in the form (4.6), and since by (4.13) solutions for (4.10) inherit the estimates on solutions for (4.1), obtained in [7, 8], then most likely the proof of the exponential mixing in eq. (4.1), given in [7] (and based on an abstract theorem from [5]), applies to establish the exponential mixing for eq. (4.10) and thus verify for it the analogy of Assumption 2.6, required in Theorem 4.4.

Acknowledgment

We are thankful to Armen Shirikyan for discussion. GH was supported by National Natural Science Foundation of China (Significant project No.11790273).

References

[1] A. Boritchev and S. B. Kuksin. *One-dimensional turbulence and the stochastic Burgers equation*. Mathematical Surveys and Monographs. AMS Publications, Providence, 2021.
[2] A. Dymov. Nonequilibrium statistical mechanics of weakly stochastically perturbed system of oscillators. *Ann. Henri Poincare*, 17:1825–1882, 2016.
[3] A. Dymov and S. B. Kuksin. Formal expansions in stochastic model for wave turbulence 1: kinetic limit. *Comm. Math. Phys.*, 382:951–1014, 2021.
[4] G. Huang, S. B. Kuksin, and A. Maiocchi. Time-averaging for weakly nonlinear CGL equations with arbitrary potential. *Fields Inst. Comm.*, 75:323–349, 2015.
[5] S. Kuksin and A. Shirikyan. *Mathematics of Two-Dimensional Turbulence*. Cambridge University Press, Cambridge, 2012.
[6] S. B. Kuksin and A. Maiocchi. Resonant averaging for small-amplitude solutions of stochastic nonlinear Schrödinger equations. *Proc. A of the Royal Soc. Edinburgh*, 147:357–394, 2017.
[7] D. Martirosyan. Exponential mixing for the white-forced damped nonlinear wave equation. *Evol. Equ. Control Theory*, 3(4):645–670, 2014.
[8] D. Martirosyan and V. Nersesyan. Local large deviations principle for occupation measures of the stochastic damped nonlinear wave equation. *Ann. IHP Prob. Stat.*, 54(4):2002–2041, 2018.
[9] S. Nazarenko. *Wave Turbulence*. Springer, 2011.
[10] C. Odasso. Exponential mixing for stochastic PDEs: the non-additive case. *Prob. Theory Relat. Fields*, 140:41–82, 2008.
[11] C. Villani. *Optimal Transport*. Springer-Verlag, Berlin, 2009.

Yau Mathematical Sciences Center, Tsinghua University, Beijing, China
Email address: huangguan@tsinghua.edu.cn

Institut de Mathématiques de Jussieu–Paris Rive Gauche, CNRS, Université Paris Diderot, UMR 7586, Sorbonne Paris Cité, F-75013, Paris, France & School of Mathematics, Shandong University, Jinan, Shandong, China
Email address: sergei.kuksin@imj-prg.fr