MULTI-SOLITONS TO FOCUSING MASS-SUPERCRITICAL STOCHASTIC NONLINEAR SCHRÖDINGER EQUATIONS

MICHAEL RÖCKNER, YIMING SU, YANJUN SUN, AND DENG ZHANG

Abstract. We consider the stochastic nonlinear Schrödinger equation driven by linear multiplicative noise in the mass-supercritical case. Given arbitrary K solitary waves with distinct speeds, we construct stochastic multi-solitons pathwisely in the sense of controlled rough path, which behave asymptotically as the sum of the K prescribed solitons as time tends to infinity. In contrast to the mass-(sub)critical case in [42], the linearized Schrödinger operator around the ground state has more unstable directions in the supercritical case. Our pathwise construction utilizes the rescaling approach and the modulation method in [12]. We derive the quantitative decay rates dictated by the noise for the unstable directions, as well as the modulation parameters and remainder in the geometrical decomposition. They are important to close the key bootstrap estimates and to implement topological arguments to control the unstable directions. As a result, the temporal convergence rate of stochastic multi-solitons, which can be of either exponential or polynomial type, is related closely to the spatial decay rate of the noise and reflects the noise impact on soliton dynamics.

Contents

1991 *Mathematics Subject Classification.* 60H15, 35C08, 35Q51, 35Q55.

Key words and phrases. Multi-solitons, mass-supercriticality, stochastic nonlinear Schrödinger equations, controlled rough path.

1. Introduction and formulation of main results

1.1. **Introduction.** We consider the focusing stochastic nonlinear Schrödinger equations (SNLS for short) with linear multiplicative noise:

$$\begin{cases} dX(t) = i\Delta X(t)dt + i|X(t)|^{p-1}X(t)dt - \mu(t)X(t)dt + \sum\limits_{k=1}^{N} X(t)G_k(t)dB_k(t), \\ X(T_0) = X_0 \in H^1(\mathbb{R}^d). \end{cases} \tag{SNLS}$$

Here, $p \in (1+\frac{4}{d}, 1+\frac{4}{(d-2)_+})$, where $\frac{4}{(d-2)_+} = +\infty$ if $d = 1, 2$, or $\frac{4}{d-2}$ if $d \geq 3$, that is, the nonlinearity lies in the mass-supercritical regime. Note that $p = 1+\frac{4}{d}$ or $1+\frac{4}{d-2}$ correspond to the mass-critical or energy-critical case, respectively. Moreover, B_k, $1 \leq k \leq N$, are standard N-dimensional real-valued Brownian motions on a stochastic basis $(\Omega, \mathscr{F}, \{\mathscr{F}_t\}, \mathbb{P})$ with normal filtration $\{\mathscr{F}_t\}$, and $G_k(t,x) = i\phi_k(x)g_k(t)$, $x \in \mathbb{R}^d$, $t > 0$, where $\{\phi_k\} \subseteq C_b^\infty(\mathbb{R}^d, \mathbb{R})$, $\{g_k\} \subseteq C^\alpha(\mathbb{R}^+, \mathbb{R})$ with $\alpha \in (1/3, 1/2)$ are controlled rough path with respect to $\{B_k\}$, and $\{g_k\}$ and their Gubinelli's derivative are $\{\mathscr{F}_t\}$-adapted. The last term $X(t)G_k(t)dB_k(t)$ is taken in the sense of controlled rough paths (see Definition 1.1 below), and the term μ is the Stratonovich correction term, which is of the form

$$\mu(t,x) = \frac{1}{2}\sum_{k=1}^{N} \phi_k^2(x)g_k^2(t),\ x \in \mathbb{R}^d,\ t > 0, \tag{1.1}$$

to ensure the conservation law of mass as required in the physical context ([1, 2]). We note that the stochastic term $XG_kdB_k(t)$ can be viewed as a random potential acting on the quantum system. It coincides with the Itô integral when $\{X(t)\}$ is $\{\mathscr{F}_t\}$-adapted. In the special case where the noise is absent, i.e., $B_k \equiv 0$, $1 \leq k \leq N$, (SNLS) reduces to the classical nonlinear Schrödinger equation (NLS for short)

$$\begin{cases} i\partial_t u + \Delta u + |u|^{p-1}u = 0, \\ u(T_0) = u_0 \in H^1(\mathbb{R}^d). \end{cases} \tag{NLS}$$

The physical significance of SNLS is well-known. The 3D cubic NLS, which is a typical mass-supercritical model, is of physical importance in nonlinear optics and describes paraxial propagation of laser beams in a homogeneous Kerr medium, see [25]. In a crystal the noise corresponds to scattering of excitons by phonons, and the noise effect on the coherence of the ground state solitary waves was investigated in [2]. It also arises from the physical model of monolayer Scheibe aggregates [1]. Moreover, the noise effect on its collapse process was studied in [41]. We also refer to [19, 20] for numerical observations of noise effects on blow-up, and [38, 39] where stochastic stable blow-up solutions have been investigated.

Local well-posedness of (SNLS) and (NLS) is known in the energy space H^1, see, e.g., [4, 9, 14]. The main interest of this paper is to study the large-time dynamics, especially, the soliton dynamics of (SNLS) in the mass-supercritical case.

According to the celebrated soliton resolution conjecture, generic global solutions to NLS are expected to behave asymptotically as a superposition of solitons plus a dispersive decaying profile. In the last decades, significant progress has been achieved on the soliton resolution conjecture for energy-critical nonlinear wave equations, see [22, 33] and references therein. Multi-solitons to NLS were constructed initially in the mass-critical case [37]. Afterwards, multi-solitons have been constructed in various settings, including the mass-subcritical case [35], the mass-supercritical case [10, 12, 40], and the energy-critical case [32]. Non-pure multi-solitons (including their scattering profile), predicted by the soliton resolution conjecture, have been recently constructed in [43], and uniqueness was proved in the solution class with t^{-5-} decay rate.

It should be mentioned that soliton dynamics in the mass-supercritical setting is much more complicated than in the (sub)critical case. One major obstruction is that the linearized Schrödinger operator in the supercritical case has more unstable directions than in the (sub)critical case. In fact, the eigenvalue of the linearized operator around the soliton $e^{it}Q$ in the (sub)critical case is exactly zero, while in the supercritical case there exist two additional nonzero real eigenvalues (see [27, 44, 49]). As a result, the orthogonal conditions in the geometrical decomposition are insufficient to control all unstable directions. Moreover, in [23], Duyckaerts and Roudenko constructed global solutions $U(t)$ to the 3D focusing cubic NLS, such that $\lim_{t\to+\infty}\|U(t) - e^{it}Q\|_{H^1} = 0$ whereas $U(t) \neq e^{it}Q$. However, in the (sub)critical case, no such special solutions $U(t)$ can exist, due to the variational property of the ground state Q and the corresponding linearized operator. In this spirit, a family of multi-solitons have been constructed by Combet [10] with the same asymptotic behavior. This is in contrast to the (sub)critical case, where multi-solitons are believed to be asymptotically unique, see the conjecture raised by Martel [34]. Very recently, in the (sub)critical setting, the uniqueness of multi-solitons with polynomial asymptotic rate has been obtained by Côte and Friederich [11], Cao, Su and Zhang [8].

In the stochastic case, more difficulties occur in the study of large-time dynamics of SNLS. As a matter of fact, the presence of noise even destroys the basic conservation law of the energy. The energy of solutions to SNLS indeed evolves as a continuous semimartingale. Its evolution was carefully studied by numerical method in [38, 39]. Furthermore, because solitons are unstable with respect to H^1 perturbations, it is a priori unclear whether the input of noise destroys the soliton dynamics. This is very different from the scattering dynamics in [31, 52], which is stable under H^1 perturbations.

In [13, 15], it was first proved by de Bouard and Debussche that non-degenerate noise in the supercritical case can accelerate blow-up with positive probability. Afterwards, the small noise large deviation principle and the error in soliton transmission have been studied in [21]. Moreover, for the 2-D Gross-Pitaevskii equation perturbed by a random quadratic potential, it was proved in [18] that the solution with initial condition of a standing wave decomposes into the sum of a randomly modulated standing wave and a small remainder, and the first order of the remainder converges to a Gaussian process. See also [16, 17] for the soliton dynamics of stochastic Korteweg-de Vries equations. Recently, the quantitative construction of blow-up solutions have been obtained for the mass-critical SNLS. We refer to [46] for critical mass blow-up solutions, [24] for stochastic log-log blow-up solutions, and [43, 47] for multi-bubble (Bourgain-Wang type) blow-up solutions.

In addition to the lack of energy conservation, the pseudo-conformal symmetry of mass-critical NLS is also destroyed by the noise. As a result, unlike in the deterministic case, stochastic multi-solitons cannot be directly derived from the aforementioned stochastic blow-up solutions. This fact forces to construct stochastic multi-solitons on the soliton level. Recently, stochastic multi-solitons to mass-(sub)critical SNLS, i.e., (SNLS) with $1 < p \leq 1+4/d$, have been constructed in [42]. The construction of stochastic multi-solitons in the mass-supercritical case, however, remains open.

The aim of the present work is to address this problem for the mass-supercritical (SNLS). More precisely, we construct stochastic multi-solitons to (SNLS) in a pathwise way in the sense of controlled rough path. The constructed stochastic solutions behave asymptotically like a sum of solitary waves with distinct speeds, see Theorem 1.2 below. This provides new examples for the soliton resolution conjecture in the stochastic case. Our proof reveals that, though solitons are unstable under H^1 perturbation of initial data, the construction of multi-solitons in some sense has structural stability, that is, it is stable under perturbation of the NLS by first and zero order terms caused by the noise (see (RNLS) below). We construct the stochastic multi-solitons in two scenarios of noise, which correspond to the exponential and polynomial spatial decay rates of noise, respectively. Quantitative decay rates of unstable directions, and modulation parameters and the remainder in the geometrical decomposition are derived. Interestingly, the temporal decay rate of stochastic multi-solitons is dictated by the spatial decay rate of the noise, which reveals the noise effect on soliton dynamics.

1.2. **Main results.** Before formulating the main results, let us first recall some basic notions in the theory of controlled rough paths from [26, 28].

Fix $\alpha \in (1/3, 1/2)$. For $I = [S,T] \subseteq \mathbb{R}^+$, let $\mathscr{C}^\alpha(I, \mathbb{R}^N)$ denote the space of α-Hölder rough paths $(X, \mathbb{X})$, such that $X \in C^\alpha(I; \mathbb{R}^N)$, $\mathbb{X} \in C^{2\alpha}(I^2; \mathbb{R}^{N\times N})$, and the Chen relation holds

$$\mathbb{X}(s,t) - \mathbb{X}(s,u) - \mathbb{X}(u,t) = \delta X_{su} \delta X_{ut}$$

for all $S \leq s \leq u \leq t \leq T$. For simplicity we write

$$\|X\|_{\alpha,I} := \sup_{s,t\in I, s\neq t} \frac{|\delta X_{st}|}{|t-s|^\alpha} < \infty, \quad \|\mathbb{X}\|_{2\alpha,I} := \sup_{s,t\in I, s\neq t} \frac{|\mathbb{X}(s,t)|}{|t-s|^{2\alpha}} < \infty,$$

where $\delta X_{st} := X(t) - X(s)$. We also set $\dot{g} := \frac{dg}{dt}$ for any C^1 functions.

Given a path $X \in C^\alpha([S,T], \mathbb{R}^N)$, $0 \leq S < T < \infty$, we recall that a pair (Y, Y') is a controlled rough path with respect to X, if $Y \in C^\alpha([S,T], \mathbb{R}^N)$, $Y' \in C^\alpha([S,T], \mathbb{R}^{N\times N})$, and the remainder term R^Y, implicitly given by

$$\delta Y_{k,st} = \sum_{j=1}^{N} Y'_{kj}(s)\delta X_{j,st} + \delta R^Y_{k,st},$$

satisfies $\|R^Y_k\|_{2\alpha,[S,T]} < \infty$, $1 \le k \le N$. Y' is the so-called Gubinelli derivative of Y. Let $\mathscr{D}^{2\alpha}_X([S,T],\mathbb{R}^N)$ denote the space of controlled rough paths with respect to X.

For any $\alpha \in (1/3, 1/2)$, it is well-known that the N-dimensional Brownian motion $B = (B_j)_{j=1}^N$ can be enhanced to the α-Hölder rough path $\mathbf{B} = (B, \mathbb{B})$, where $\mathbb{B}(s,t) := \int_s^t \delta B_{sr} \otimes dB(r) \in \mathbb{R}^N \times \mathbb{R}^N$ is taken in the sense of Itô, $0 \le s < t < \infty$. For any $T \in (0,\infty)$, it holds that $\mathbf{B} \in \mathscr{C}^\alpha([0,T],\mathbb{R}^N)$ almost surely, see [26, Proposition 3.4].

Given a controlled rough path $(Y, Y') \in \mathscr{D}^{2\alpha}_B([S,T],\mathbb{R}^N)$, one can define the rough integration of Y against $\mathbf{B} = (B, \mathbb{B})$ as follows (see [26, Theorem 4.10], [28, Corollary 2]): For each $1 \le k \le N$,

$$\int_S^T Y_k(r)dB_k(r) := \lim_{|\mathcal{P}|\to 0} \sum_{l=1}^{N} \Big(Y_k(t_l)\delta B_{k,t_l t_{l+1}} + \sum_{j=1}^{N} Y'_{kj}(t_l)\mathbb{B}_{jk}(t_l, t_{l+1}) \Big),$$

where $\mathcal{P} := \{t_l\}_{l=0}^n$ is a partition of $[S,T]$ so that $t_0 = S$, $t_n = T$, $|\mathcal{P}| := \max_{0\le l\le n-1} |t_{l+1} - t_l|$.

As in [42], we assume that the noise in (SNLS) satisfies the following conditions:

$(A0)$: Asymptotic flatness. For every $1 \le k \le N$, $\phi_k \subseteq C_b^\infty(\mathbb{R}^d, \mathbb{R})$ such that

$$\lim_{|x|\to\infty} |x|^2 |\partial_x^\nu \phi_k(x)| = 0, \ \nu \neq 0. \tag{1.2}$$

$(A1)$: $\{g_k\}_{k=1}^N$ are $\{\mathscr{F}_t\}$-adapted continuous processes controlled by the Brownian motions $\{B_k\}$ with the Gubinelli derivative $\{g'_{kj}\}_{j,k=1}^N$, and $g_k \in L^2(\mathbb{R}^+)$, $1 \le k \le N$, $\mathbb{P}$-a.s.. Moreover, one of the following cases holds:

Case (I): For every $1 \le k \le N$, there exists $c_k > 0$ such that for $|x| > 0$,

$$\sum_{|\nu|\le 4} |\partial_x^\nu \phi_k(x)| \le C e^{-c_k|x|}. \tag{1.3}$$

Case (II): Let $\nu_* \in \mathbb{N}$. For every $1 \le k \le N$ and $|x| > 0$,

$$\sum_{|\nu|\le 4} |\partial_x^\nu \phi_k(x)| \le C|x|^{-\nu_*}. \tag{1.4}$$

In addition, there exists a random time σ_* and a deterministic constant $c^* > 0$ such that $\mathbb{P}$-a.s. $\sigma_* \in [0,\infty)$ and for any $t \ge \sigma_*$,

$$\int_t^\infty g_k^2 ds \log\left(\int_t^\infty g_k^2 ds\right)^{-1} \le \frac{c^*}{t^2}, \quad 1 \le k \le N. \tag{1.5}$$

We note that Case (I) and Case (II) correspond to the exponential and polynomial spatial decay rates of noise, respectively. Without loss of generality, we consider in Assumptions (A0) and (A1)

$$\sum_{|\nu|\le 4} |\partial_x^\nu \phi_k(x)| \le C\phi(|x|)$$

with a decay function ϕ of the following form

$$\phi(|x|) := \begin{cases} e^{-|x|}, & \text{in Case (I)}; \\ |x|^{-\nu_*}, & \text{in Case (II)}. \end{cases} \tag{1.6}$$

As we shall see later, these spatial decay rates of noise indeed affect the temporal decay rate of stochastic solitary waves, see (1.14) and (1.16) below.

The temporal condition (1.5) relates to the Levy Hölder continuity of Brownian motions, which permits to control the tail of the noise $B_*(t)$ in (3.1) below for t large enough. It is worth noting that the t^{-1} decay rate of $B_*(t)$ in (3.7) is essential to close the bootstrap estimate of $\|\varepsilon\|_{H^1}$ in Case (II), see, e.g., (4.30) below.

Let us also mention that the asymptotic flatness condition $(A0)$ ensures the local well-posedness of (SNLS), see [4, 52]. The smoothness condition of the spatial functions $\{\phi_k\}$ is assumed merely for simplicity. One can also treat infinitely many Brownian motions, i.e., $N = \infty$, and noise of low spatial regularity in certain Sobolev and lateral Strichartz spaces, as in the context of Zakharov system [29, 30].

The solutions to (SNLS) are defined in the following controlled rough path sense.

Definition 1.1. *Let $p \in (1 + \frac{4}{d}, 1 + \frac{4}{(d-2)_+})$, $d \geq 1$. We say that X is a solution to* (SNLS) *on $[T_0, \tau^*)$, where T_0, $\tau^* \in (0, \infty]$ are random variables, if for $\mathbb{P}$-a.e. $\omega \in \Omega$ and for any $\varphi \in C_c^\infty$, $t \mapsto \langle X(t,\omega), \varphi\rangle$ is continuous on $[T_0(\omega), \tau^*(\omega))$ and for any $T_0(\omega) \leq s < t \leq \tau^*(\omega)$,*

$$\langle X(t) - X(s), \varphi\rangle - \int_s^t \langle iX, \Delta\varphi\rangle + \langle i|X|^{p-1}X, \varphi\rangle - \langle \mu X, \varphi\rangle dr = \sum_{k=1}^N \int_s^t \langle i\phi_k g_k(r) X(r), \varphi\rangle dB_k(r). \tag{1.7}$$

Here the integral $\int_s^t \langle i\phi_k g_k X, \varphi\rangle dB_k(r)$ is taken in the sense of controlled rough paths with respect to the Brownian rough path $\{(B, \mathbb{B})\}$, and $\langle i\phi_k g_k X, \varphi\rangle \in \mathscr{D}_B^{2\alpha}([s,t], \mathbb{R})$, satisfying

$$\delta\langle i\phi_k g_k X, \varphi\rangle_{st} = \sum_{j=1}^N \langle -\phi_j\phi_k g_j(s) g_k(s) X(s) + i\phi_k g'_{kj}(s) X(s), \varphi\rangle \delta B_{j,st} + \delta R_{k,st}^{\langle i\phi_k g_k X, \varphi\rangle}, \tag{1.8}$$

and $\|\langle \phi_j\phi_k g_j g_k X, \varphi\rangle\|_{\alpha,[s,t]} + \|\langle \phi_k g'_{kj} X, \varphi\rangle\|_{\alpha,[s,t]} < \infty$, $\|R_k^{\langle i\phi_k g_k X, \varphi\rangle}\|_{2\alpha,[s,t]} < \infty$, $\alpha \in \left(\frac{1}{3}, \frac{1}{2}\right)$.

In the characterization of soliton dynamics, a key role is played by the ground state, which is the unique radial positive solution to the nonlinear elliptic equation

$$\Delta Q - Q + Q^p = 0. \tag{1.9}$$

It is known (see, e.g., [7]) that the ground state decays exponentially fast at infinity, i.e., there exist C, $\delta > 0$ such that for any $|\nu| \leq 3$,

$$|\partial_x^\nu Q(x)| \leq Ce^{-\delta|x|}, \quad x \in \mathbb{R}^d. \tag{1.10}$$

For any $w > 0$, let Q_w denote the rescaled ground state

$$Q_w(x) := w^{-\frac{2}{p-1}} Q(\frac{x}{w}), \quad x \in \mathbb{R}^d. \tag{1.11}$$

Note that by the ground state equation (1.9), Q_w satisfies the equation

$$\Delta Q_w - w^{-2} Q_w + Q_w^p = 0. \tag{1.12}$$

Given any $K \in \mathbb{N}$ and any $w_k \in \mathbb{R}^+$, $\alpha_k^0 \in \mathbb{R}^d$, $\theta_k^0 \in \mathbb{R}$, $1 \leq k \leq K$, our aim is to construct stochastic multi-solitons which behave asymptotically as a sum of solitary waves

$$R_k(t,x) := Q_{w_k}(x - v_k t - \alpha_k^0) e^{i(\frac{1}{2}v_k\cdot x - \frac{1}{4}|v_k|^2 t + (w_k)^{-2}t + \theta_k^0)} \tag{1.13}$$

with distinct velocities

$$v_k \neq v_{k'} \quad \text{for any } k \neq k'.$$

The main result of this paper is the following:

Theorem 1.2. *Consider* (SNLS) *with $p \in (1 + \frac{4}{d}, 1 + \frac{4}{(d-2)_+})$, $d \geq 1$. Assume (A0) and (A1) with $\nu_* \geq \nu_0$ in Case (II), where ν_0 is a deterministic constant given by (4.25) below. Then, there exists a positive random time T_0 such that for $\mathbb{P}$-a.e. $\omega \in \Omega$, there exist $X_*(\omega) \in H^1$ and an H^1-valued solution $X(t,\omega)$ to* (SNLS) *on $[T_0(\omega), \infty)$ satisfying $X(T_0, \omega) = X_*(\omega)$ and*

$$\|e^{-W_*(t)} X(t) - \sum_{k=1}^K R_k(t)\|_{H^1} \leq Ct\phi^{\frac{1}{2}}(\delta t), \ t \geq T_0. \tag{1.14}$$

Here, the random phase function W^* is given by

$$W_*(t,x) = -\sum_{k=1}^N \int_t^\infty i\phi_k(x) g_k(s) dB_k(s), \tag{1.15}$$

the solitary waves $\{R_k\}$ *are given by* (1.13)*,* ϕ *is the spatial decay function of the noise in (1.6) and* $C, \delta > 0$ *are deterministic positive constants. In particular, for* $t \geq T_0$*,*

$$\|X(t) - \sum_{k=1}^{K} R_k(t)\|_{H^1} \leq C \sum_{k=1}^{N} \left(\int_t^\infty g_k(s)^2 ds \log \left(\int_t^\infty g_k(s)^2 ds \right)^{-1} \right)^{\frac{1}{2}} + Ct\phi^{\frac{1}{2}}(\delta t). \tag{1.16}$$

Remark 1.3. Because the right-hand side of (1.16) tends to 0 as $t \to \infty$ due to $\{g_l\} \subseteq L^2(\mathbb{R}^+)$ in $(A1)$, the constructed stochastic solution in Theorem 1.2 converges to the given K solitary waves as time tends to infinity. As a result, Theorem 1.2 provides new examples for the soliton resolution conjecture in the stochastic mass-supercritical case.

We also note that the temporal decay rate of stochastic multi-solitons in (1.14) are dictated by the spatial decay rate of the noise. The decay rate can be of either exponential or polynomial type in two scenarios of noise Cases (I) and (II), respectively. This reflects the noise impact on soliton dynamics.

Remark 1.4. It should be mentioned that, compared to the (sub)critical case [42], many difficulties emerge in the present mass-supercritical case. One major difficulty is that the linearized Schrödinger operator has two extra unstable directions $\mathbf{a}^{\pm}$ (see (2.52) below) in the supercritical case, which are much harder to control.

The strategy here employs the modulation method inspired by [12] to modulate the final data of approximating solutions. As an immediate technical issue, the control of modulated final data to prescribed vectors requires more delicate analysis of the non-degeneracy of Jacobian matrices than in the (sub)critical case [42], see Proposition 2.7 below. We also derive that the radius of the modulation parameter and the constants in Proposition 2.7 can be chosen to be deterministic, which is important to take a large random time to close the bootstrap estimates of modulation parameters and remainder in the geometrical decomposition.

We remark that the bootstrap estimates are crucial to obtain uniform estimates of approximating solutions. In contrast to the (sub)critical case in [42], the bootstrap estimates in the supercritical case require an a-priori estimate of the unstable direction $\mathbf{a}^-$, which, however, cannot be closed by Gronwall's argument. In order to overcome this problem, a topological argument for $\mathbf{a}^-$, based on Brouwers fixed point theorem, is performed on a ball with radius dictated by the spatial decay rate of the noise. We refer to Subsection 1.3 below for more detailed explanations of the difference between the supercritical and (sub)critical cases, as well as the deterministic case.

Remark 1.5. We are not sure about the measurability of the solution X constructed in Theorem 1.2. The measurability issue arises from the choice of $\boldsymbol{a}_n^-(\omega) \in B_{\mathbb{R}^K}(\phi^{\frac{1}{2}+\frac{1}{4d}}(\widetilde{\delta} n))$ in Proposition 4.5 which is based on a topological argument, as well as from the compactness argument used to obtain a subsequence of approximating solutions in the proof of Theorem 1.7 in Section 5. It does not seem obvious here how to use measurable selection theorems for instance in [48] to select measurable versions of $\boldsymbol{a}_n^-$ and of X in our situation.

1.3. **Strategy of the proof.** Our proof utilizes the rescaling approach and the modulation analysis, including topological arguments to treat the new unstable directions of the linearized Schrödinger operator, which is different from [42] in the mass-(sub)critical case.

To be precise, we use the rescaling or Doss-Sussman type transformation

$$u(t) := e^{-W_*(t)} X(t), \tag{1.17}$$

where W_* is given by (1.15), to transfer (SNLS) to a random NLS

$$\begin{cases} i\partial_t u + (\Delta + b_* \cdot \nabla + c_*)u + |u|^{p-1}u = 0, \\ u(T_0) = e^{-W_*(T_0)} X_0, \end{cases} \tag{RNLS}$$

where the random coefficients b_* and c_* have the expressions

$$b_*(t,x) = 2\nabla W_*(t,x) = 2i \sum_{k=1}^{N} \int_t^\infty \nabla \phi_k(x) g_k(s) dB_k(s), \tag{1.18}$$

$$c_*(t,x) = \sum_{j=1}^{d} (\partial_j W_*(t,x))^2 + \Delta W_*(t,x)$$

$$= -\sum_{j=1}^{d}\left(\sum_{k=1}^{N}\int_t^\infty \partial_j\phi_k(x)g_k(s)dB_k(s)\right)^2 + i\sum_{k=1}^{N}\int_t^\infty \Delta\phi_k(x)g_k(s)dB_k(s). \tag{1.19}$$

The rescaled equation (RNLS) enables us to perform pathwise analysis in a sharp way, which is in general not possible for Itô integral. This approach works successfully for well-posedness and optimal control problems, see [3–5, 50, 52]. It is also comparable with the Fourier restriction method and permits to exploit the noise regularization effect on scattering. We refer the interested readers to [29–31, 45].

It should be mentioned that, though the rescaling transform provides a nice way to reveal the structure of SNLS, it does not remove the difficulty in the analysis. One obstacle is to control the derivative term $b_*\cdot\nabla u$ caused by noise, which is in general difficult for Schrödinger equations, due to the lack of global regularity of Schrödinger groups. See, for instance, the case of the Schrödinger map [6]. This problematic term can be controlled here under the asymptotic flat condition $(A0)$ of the noise, by using the local smoothing estimates in [31, 51]. An alternative way to control this term is to use lateral Strichartz estimates, as in the context of stochastic Zakharov systems [29, 30], which help to weaken the regularity condition on the noise.

The H^1 local well-posedness of (RNLS) can be proved by using analogous arguments as in [4, 31]. Then the H^1 well-posedness of (SNLS) can be inherited from that of (RNLS) via Theorem 1.6 below. This fact has been proved in [42] in the mass-(sub)critical case. With slight modifications, based on the Sobolev embedding $H^1(\mathbb{R}^d)\hookrightarrow L^{p+1}(\mathbb{R}^d)$ with $p\in(1+\frac{4}{d},1+\frac{4}{(d-2)_+})$, the proof there also applies to the present mass-supercritical case.

Theorem 1.6. *Let $p\in(1+\frac{4}{d},1+\frac{4}{(d-2)_+})$, $d\geq 1$. Let u be the solution to* (RNLS) *on $[T_0,\tau^*)$ with $X_0\in H^1$, where $T_0,\tau^*\in(0,\infty]$ are random variables. Then, for $\mathbb{P}$-a.e. $\omega\in\Omega$, $X(\omega):=e^{W_*(\omega)}u(\omega)$ is the solution to* (SNLS) *on $[T_0(\omega),\tau^*(\omega))$ in the sense of Definition 1.1.*

As a result, the proof of Theorem 1.2 can be reduced to that of the following result for the rescaled random equation (RNLS).

Theorem 1.7. *Consider* (RNLS) *with $p\in(1+\frac{4}{d},1+\frac{4}{(d-2)_+})$, $d\geq 1$. Assume (A_0) and (A_1) with $\nu_*\geq\nu_0$ in Case (II), where ν_0 is a deterministic constant given by* (4.25). *Then, there exists a positive random time T_0 such that for $\mathbb{P}$-a.e. $\omega\in\Omega$, there exist $u_*(\omega)\in H^1$ and a unique solution $u\in C([T_0(\omega),\infty];H^1)$ to (RNLS) satisfying $u(\omega,T_0)=u_*(\omega)$ and*

$$\|u(t)-\sum_{k=1}^{K}R_k(t)\|_{H^1}\leq Ct\phi^{\frac{1}{2}}(\delta t),\ t\geq T_0. \tag{1.20}$$

where the solitary waves $\{R_k\}$ are given by (1.13), *ϕ is the decay function in (1.6) and $C,\delta>0$.*

In the sequel, we mainly prove Theorem 1.7. The proof utilizes the modulation method developed in [12, 35] and mainly proceeds in the following three steps. The impact of noise on the construction of stochastic multi-solitons is presented below as well.

• **Geometric decomposition:** First in Section 2, we derive the geometrical decomposition of approximating solutions into a sum of K solitons plus a remainder term $u=\sum_{k=1}^{K}\widetilde{R}_k+\varepsilon$, where $\widetilde{R}_k$ are modulated soliton profiles with modulation parameters $\{(\alpha_k,\theta_k)\}$, and ε is the remainder (see Proposition 2.4 below).

As mentioned above, unlike in the mass-(sub)critical case [42], the linearized Schrödinger operator $\mathscr{L}$ (see (2.1)) has four unstable directions, reflected by the following coercivity type estimate

$$(\mathscr{L}f,f)\geq C\|f\|_{H^1}^2-\frac{1}{C}\left(\left(\int\nabla Qf_1dx\right)^2+\left(\int Qf_2dx\right)^2+\left(\mathrm{Im}\int Y^+\bar{f}dx\right)^2+\left(\mathrm{Im}\int Y^-\bar{f}dx\right)^2\right)$$

for any $f=f_1+if_2\in H^1$, where $Y^\pm$ are two eigenfunctions of the linearized Schrödinger operator in the mass-supercritical case.

The new eigenfunctions give rise to two extra unstable directions $a_k^\pm:=\mathrm{Im}\int\widetilde{Y}_k^\pm\bar{\varepsilon}dx$, where $\widetilde{Y}_k^\pm$ are the modulated eigenfunctions of $Y^\pm$ with speed v_k, $1\leq k\leq N$. The extra unstable directions cannot be canceled by imposing orthogonal conditions in the geometric decomposition as in the (sub)critical case

[42]. Instead, they are controlled by modulating the final data with the eigenfunctions. That is, instead of equation (RNLS), we consider the following approximating random equation

$$\begin{cases} i\partial_t u + (\Delta + b_* \cdot \nabla + c_*)u + |u|^{p-1}u = 0, \\ u(T) = R(T) + i \sum_{k,\pm} b_k^{\pm} Y_k^{\pm}(T) \end{cases} \tag{1.21}$$

(see (2.35) below). One advantage to modulate the final data is, that it allows to steer the unstable directions $(a_k^+(T), a_k^-(T))$ at time T to any prescribed vector in $\mathbf{0} \times B_{\mathbb{R}^K}(r_0)$ with $r_0 \ll 1$. The proof of this fact requires careful analysis of the *non-degeneracy* of Jacobian matrices, which do not appear in the mass-(sub)critical case [42]. See the proof of Proposition 2.7 below.

Let us mention that a detailed proof for the *exponential decay* of the eigenfunctions $Y^{\pm}$ is given in Lemma 2.2 below. The exponential decay property of the eigenfunctions, as well as of the ground state, allow us to decouple the iterations between different soliton profiles and eigenfunctions, at the cost of exponential decay orders, which are favorable in the bootstrap estimates.

• **Bootstrap estimates:** The next step is to establish the crucial *bootstrap estimates* of the unstable direction a_k^+, and the modulation parameters (α_k, θ_k) and the remainder ε in the geometrical decomposition. The bootstrap estimates are important to derive uniform estimates for approximating solutions up to a universal time, and thus, allow to construct stochastic multi-solitons by using compactness arguments. This constitutes the main technical part of Sections 3 and 4.

The remainder term in the geometrical decomposition can be controlled by a coercivity type estimate of the Lyapunov functional, see Proposition 3.6 below.

The subtleness here is that, unlike in the deterministic case [12], the presence of noise, especially with the polynomial decay rate in Case (II), affects the decay rates in the bootstrap estimates.

As a matter of fact, in the deterministic case, because the ground state and eigenfunctions decay exponentially fast at infinity, the exponential decay rate is sufficient to close bootstrap estimates. However, in the stochastic Case (II), the above key quantities, that is, the modulation parameters, the remainder and the unstable directions, have merely polynomial decay rates, rather than the exponential decay rate in the deterministic case. As a result, the derivation of appropriate decay rates to close bootstrap estimates is much more delicate. Under the a-priori control of the unstable direction $|a_k^-(t)| \le \phi^{\frac12 + \frac{1}{4d}}(\widetilde{\delta} t)$, we derive the bootstrap estimates

$$|\alpha_k(t) - \alpha_k^0| + |\theta_k(t) - \theta_k^0| \le t\phi^{\frac12}(\widetilde{\delta} t), \quad \|\varepsilon(t)\|_{H^1} \le \phi^{\frac12}(\widetilde{\delta} t), \quad |a_k^+(t)| \le \phi^{\frac12}(\widetilde{\delta} t).$$

We note that the above decay rates of bootstrap estimates are dictated by the spatial decay rate of the noise. More technically, the t^{-1} decay rate of the tail of the noise $B_*(t)$ in (3.7) below is essential to close the bootstrap estimate of $\|\varepsilon\|_{H^1}$ in Case (II), see, e.g., estimate (4.30) below. These facts reflect the effect of noise on the soliton dynamics.

It is also worth noting that, because of the possible singularity at the origin of the second derivative of the supercritical nonlinearity in high dimensions, the extra unstable directions $\{a_k^{\pm}\}$ are controlled by $\|\varepsilon\|_{H^1}^{p\wedge 2}$, together with the noise decay rate and the negligible exponential decay rate. See Proposition 3.4 below.

• **Topological arguments:** In general, with the help of bootstrap estimates, one can obtain uniform estimates of approximating solutions by using standard continuity argument. See, e.g., [35, 42] in the mass-(sub)critical case.

In contrast to that, due to the unstable direction a_k^- in the supercritical case, the above bootstrap estimates only allow to refine the estimates of the modulation parameters, the remainder and the unstable direction a_k^+, but require a-priori control of the other unstable direction a_k^-.

In order to achieve the required a-priori control of a_k^-, we use a topological argument based on the Brouwer fixed point theorem inspired by [12]. Again, one needs to take into account the influence of noise, and the radius of the ball where the topological argument is performed is dictated by the decay rate of the noise. For instance, in Case (II), the topological arguments are performed on the ball $B_{\mathbb{R}^K}(\phi^{\frac12 + \frac{1}{4d}}(\widetilde{\delta} n))$, where ϕ has polynomial decay rate in (1.6), that is different from the exponential rate in the deterministic case. See Proposition 4.5 below.

Notations. *For any* $x = (x_1, x_2, \cdots, x_d) \in \mathbb{R}^d$ *and any multi-index* $\nu = (\nu_1, \nu_2, \cdots, \nu_d)$, *let* $|x| := \left(\sum_{l=1}^d x_l^2\right)^{\frac{1}{2}}$, $\partial_x^\nu := \partial_{x_1}^{\nu_1} \cdots \partial_{x_d}^{\nu_d}$ *and* $|\nu| := \nu_1 + \cdots + \nu_d$. *For* $s \in \mathbb{R}$, $1 \leq p \leq +\infty$, *let* $W^{s,p}(\mathbb{R}^d)$ *denote the standard Sobolev spaces, and* $H^s := W^{s,2}(\mathbb{R}^d)$. *In particular,* $L^p := W^{0,p}(\mathbb{R}^d)$ *is the space of* p*-integrable (complex-valued) functions endowed with the norm* $\|\cdot\|_{L^p}$. *When* $p = 2$, L^2 *is the Hilbert space endowed with the inner product* $\langle v, w\rangle = \int_{\mathbb{R}^d} v(x)\bar{w}(x)dx$.

Given any two Banach spaces $\mathcal{X}$ *and* $\mathcal{Y}$ *and any Fréchet differentiable map* $F : \mathcal{X} \to \mathcal{Y}$, *let* $dF(x) \in L(\mathcal{X}, \mathcal{Y})$ *(*dF *for short) denote the Fréchet derivative of* F *at* x. *For any* $h \in \mathcal{X}$, $dF.h$ *denotes the directional derivative of* dF *along the direction* h. *Moreover, let* $B_{\mathcal{X}}(x, r)$ *(resp.* $\mathring{B}_{\mathcal{X}}(x, r)$*) denote the closed (resp. open) ball of the Banach space* $\mathcal{X}$*, centered at* x *with radius* $r > 0$*, and* $S_{\mathcal{X}}(x, r)$ *the corresponding sphere. If* $x = 0$*, we simply write* $B_{\mathcal{X}}(r)$, $\mathring{B}_{\mathcal{X}}(r)$ *and* $S_{\mathcal{X}}(r)$.

Throughout this paper, positive constants C, η *and* δ *may change from line to line. Finally,* $f = \mathcal{O}(g)$ *means that* $|f/g|$ *stays bounded, and* $f = o(g)$ *means that* $|f/g|$ *converges to zero.*

2. Geometrical decomposition

This section is devoted to the geometrical decomposition of (RNLS) with modulated initial data. The crucial role here is played by the spectrum of linearized Schrödinger operators around the ground state.

2.1. **Linearized Schrödinger operators.** Let $\mathscr{L} = (\mathscr{L}_+, \mathscr{L}_-)$ be the linearized operator around the ground state, given by

$$\mathscr{L}_+ := -\Delta + I - pQ^{p-1}, \quad \mathscr{L}_- := -\Delta + I - Q^{p-1} \tag{2.1}$$

with $p \in (1 + \frac{4}{d}, 1 + \frac{4}{(d-2)_+})$. For any complex valued function $f = f_1 + if_2 \in H^1$, let

$$\mathscr{L}f := -\mathscr{L}_- f_2 + i\mathscr{L}_+ f_1, \tag{2.2}$$

and

$$(\mathscr{L}f, f) := \int f_1 \mathscr{L}_+ f_1 dx + \int f_2 \mathscr{L}_- f_2 dx. \tag{2.3}$$

It is known (see, e.g., [27, 44, 49]) that the linearized Schrödinger operator $\mathscr{L}$ has exactly one pair of real nonzero eigenvalues $e_0 (> 0)$ and $-e_0$, with the corresponding normalized eigenfunctions $Y^\pm$ satisfying

$$Y^\pm \in \mathcal{S}(\mathbb{R}^d), \tag{2.4}$$

$\|Y^\pm\|_{L^2} = 1$, $\bar{Y}^+ = Y^-$ and

$$\mathscr{L}Y^\pm = \pm e_0 Y^\pm. \tag{2.5}$$

Moreover, it has the following coercivity property, which is important to derive the geometrical decomposition and uniform estimates of solutions:

$$(\mathscr{L}f, f) \geq C\|f\|_{H^1}^2 - \frac{1}{C}\left(\left(\int \nabla Q f_1 dx\right)^2 + \left(\int Q f_2 dx\right)^2 + \left(\mathrm{Im}\int Y^+ \bar{f} dx\right)^2 + \left(\mathrm{Im}\int Y^- \bar{f} dx\right)^2\right) \tag{2.6}$$

for any $f = f_1 + if_2 \in H^1$, where $C > 0$ is a universal positive constant. See, e.g., [12, 23].

Remark 2.1. The above coercivity estimate reveals that the linearized operator $\mathscr{L}$ has four unstable directions. The first two can be controlled by the orthogonality conditions in the geometrical decomposition as in the mass-(sub)critical case [42], see Proposition 2.4 below. In contrast, the latter two unstable directions gives rise to the main difficulty in the soliton analysis in the supercritical case. In order to control these new unstable directions, we will further modulate the initial data and apply topological arguments.

The following result shows that the eigenfunctions $Y^\pm$ decay exponentially fast at infinity.

Lemma 2.2 (Exponential decay of eigenfunctions)**.** *There exist* $C, \delta > 0$ *such that*

$$|Y^+(x)| + |Y^-(x)| \leq Ce^{-\delta|x|}, \quad x \in \mathbb{R}^d. \tag{2.7}$$

Proof. The proof proceeds in two steps: we first prove that $Y^\pm$ are exponentially integrable, and then we show the pointwise exponential decay of $Y^\pm$.

(i) Exponential integrability: Let $Y_1 := \mathrm{Re}Y^+$, $Y_2 := \mathrm{Im}Y^+$. Let us first show that

$$\int (Y_1^2 + Y_2^2) e^{|x|} dx < \infty. \tag{2.8}$$

For this purpose, we let $f_\eta(x) := e^{\frac{|x|}{1+\eta|x|}}$, $\eta > 0$, $x \in \mathbb{R}^d$. It is easy to check that f_η is bounded, Lipschitz continuous and satisfies that

$$|\nabla f_\eta(x)| \leq f_\eta(x), \quad x \neq \mathbf{0}. \tag{2.9}$$

By (2.1), (2.2), (2.5) and $\bar{Y}^+ = Y^-$ one has

$$-\Delta Y_2 + Y_2 - Q^{p-1}Y_2 = -e_0 Y_1, \tag{2.10}$$

$$-\Delta Y_1 + Y_1 - pQ^{p-1}Y_1 = e_0 Y_2. \tag{2.11}$$

Taking the inner product of (2.10) and (2.11) with $f_\eta Y_2$ and $f_\eta Y_1$, respectively, and then summing up the results we obtain

$$\int \nabla Y_2 \cdot \nabla(f_\eta Y_2)dx + \int \nabla Y_1 \cdot \nabla(f_\eta Y_1)dx + \int f_\eta Y_2^2 dx + \int f_\eta Y_1^2 dx = \int Q^{p-1}\big(f_\eta Y_2^2 + pf_\eta Y_1^2\big)dx. \tag{2.12}$$

Next let us treat the left-hand side of (2.12). Using (2.9) and Hölder's inequality one has

$$\begin{aligned}\int \nabla Y_2 \cdot \nabla(f_\eta Y_2)dx =& \int f_\eta |\nabla Y_2|^2 dx + \int (\nabla f_\eta \cdot \nabla Y_2)Y_2 dx \\ \geq& \frac{1}{2}\int f_\eta |\nabla Y_2|^2 dx - \frac{1}{2}\int f_\eta Y_2^2 dx.\end{aligned} \tag{2.13}$$

Similarly, one has

$$\int \nabla Y_1 \cdot \nabla(f_\eta Y_1)dx \geq \frac{1}{2}\int f_\eta |\nabla Y_1|^2 dx - \frac{1}{2}\int f_\eta Y_1^2 dx. \tag{2.14}$$

Plugging (2.13) and (2.14) into (2.12) we then derive that

$$\text{L.H.S. of (2.12)} \geq \frac{1}{2}\int f_\eta (Y_1^2 + Y_2^2)dx. \tag{2.15}$$

Regarding the right-hand side of (2.12), by the exponential decay of the ground state (1.10), we see that $pQ^{p-1} < 1/4$ for any $|x| > C_0$ with C_0 large enough. This yields that for some $C > 0$ independent of η,

$$\begin{aligned}\text{R.H.S. of (2.12)} \leq& \frac{1}{4}\int_{|x|>C_0} f_\eta(Y_1^2 + Y_2^2)dx + \int_{|x|\leq C_0} Q^{p-1}(f_\eta Y_2^2 + pf_\eta Y_1^2)dx \\ \leq& \frac{1}{4}\int f_\eta(Y_1^2 + Y_2^2)dx + C.\end{aligned} \tag{2.16}$$

Thus, combining (2.15) and (2.16) together we get

$$\int (Y_1^2 + Y_2^2) f_\eta dx \leq C.$$

Letting $\eta \to 0$ and using Fatou's lemma we obtain (2.8).

(ii) Next, we shall prove the pointwise exponential decay

$$\sup_{x\in\mathbb{R}^d} (Y_1^{d+2} + Y_2^{d+2})e^{|x|} < \infty. \tag{2.17}$$

It follows from the following more general claim.

Claim: Let $d \geq 1$. If $g \in \mathcal{S}(\mathbb{R}^d)$ is nonnegative and there exists $m_0 \geq 2$ such that

$$\int g^{m_0}(x)e^{|x|}dx < \infty, \tag{2.18}$$

then

$$\sup_{x\in\mathbb{R}^d} g^{m_0+d}(x)e^{|x|} < \infty. \tag{2.19}$$

Applying this claim to the case where $m_0 = 2$ and $g = Y_1$, Y_2 in (2.19) we thus obtain (2.17).

It remains to prove the claim. For this purpose, we shall use the induction argument on dimensions. First when $d = 1$, by the continuity of $g(x)$ and $e^{|x|}$,

$$\sup_{|x|\leq 1} g^{m_0+1}(x)e^{|x|} < \infty. \tag{2.20}$$

Then, using (2.18) and $g \in \mathcal{S}(\mathbb{R})$ we have

$$\sup_{x>1} \int_1^x |(g^{m_0+1}(s)e^s)'| ds < \infty, \quad \text{and} \quad \sup_{y<-1} \int_y^{-1} |(g^{m_0+1}(s)e^{-s})'| ds < \infty,$$

which, via the fundamental theorem of calculus, yields that

$$\sup_{x>1} \left| g^{m_0+1}(x)e^{|x|} - g^{m_0+1}(1)e \right| < \infty \quad \text{and} \quad \sup_{y<-1} \left| g^{m_0+1}(y)e^{|y|} - g^{m_0+1}(-1)e \right| < \infty. \tag{2.21}$$

It follows from (2.20) and (2.21) that (2.19) is valid when $d = 1$.

Now, we assume that (2.18) implies (2.19) for all $d \leq k$ and consider the case where $d = k+1$.

Let $x = (x_1, x_2, \cdots, x_{k+1}) \in \mathbb{R}^{k+1}$. Denote the cube $O_{k+1}^c := [-1, 1]^{k+1}$ and its complement $O_{k+1} := \mathbb{R}^{k+1} \backslash O_{k+1}^c$. Using the continuity of $g(x)$ and $e^{|x|}$ again we have

$$\sup_{x \in O_{k+1}^c} g^{m_0+k+1}(x)e^{|x|} < \infty.$$

Hence, it suffices to prove

$$\sup_{x \in O_{k+1}} g^{m_0+k+1}(x)e^{|x|} < \infty. \tag{2.22}$$

Let

$$D_j := \{(l_1, l_2, \cdots, l_j) \in \mathbb{Z}^j : 1 \leq l_1 < l_2 < \cdots < l_j \leq k+1\} \quad \text{for} \quad 1 \leq j \leq k+1.$$

It is obvious that O_{k+1} is a union of the sets O_{k+1}^0 and $O_{k+1}^{(l_1, \cdots, l_j)}$, $\forall\, (l_1, \cdots, l_j) \in D_j$, where

$$O_{k+1}^0 := \{x \in O_{k+1} : x_m \geq 0, 1 \leq m \leq k+1\},$$

$$O_{k+1}^{(l_1, \cdots, l_j)} := \{x \in O_{k+1} : x_m \geq 0,\ m \neq l_1, l_2, \cdots, l_j,\ \text{and}\ x_{l_1}, x_{l_2}, \cdots, x_{l_j} \leq 0\}.$$

Thus, (2.22) follows immediately from

$$\sup_{x \in O_{k+1}^0} g^{m_0+k+1}(x)e^{|x|} < \infty, \tag{2.23}$$

and

$$\sup_{x \in O_{k+1}^{(l_1, \cdots, l_j)}} g^{m_0+k+1}(x)e^{|x|} < \infty, \quad \forall\, (l_1, \cdots, l_j) \in D_j,\ 1 \leq j \leq k+1. \tag{2.24}$$

Below, we prove (2.23), and the proof of (2.24) is similar. For any $x \in \mathbb{R}^{k+1}$ and $(l_1, \cdots, l_j) \in D_j$, we set

$$\begin{aligned}
x^{(l_1, \cdots, l_j)} &:= (x_1, \cdots, x_{l_1-1}, 1, x_{l_1+1}, \cdots, x_{l_2-1}, 1, x_{l_2+1}, \cdots, x_{l_j-1}, 1, x_{l_j+1}, \cdots, x_{k+1}), \\
\bar{x}^{(l_1, \cdots, l_j)} &:= (x_1, \cdots, x_{l_1-1}, 0, x_{l_1+1}, \cdots, x_{l_2-1}, 0, x_{l_2+1}, \cdots, x_{l_j-1}, 0, x_{l_j+1}, \cdots, x_{k+1}), \\
dx^{(l_1, \cdots, l_j)} &:= dx_1 \cdots dx_{l_1-1} dx_{l_1+1} \cdots dx_{l_2-1} dx_{l_2+1} \cdots dx_{l_j-1} dx_{l_j+1} \cdots dx_{k+1}.
\end{aligned}$$

Note that

$$|x^{(l_1, \cdots, l_j)}| \leq |\bar{x}^{(l_1, \cdots, l_j)}| + j, \quad |\bar{x}^{(l_1, \cdots, l_j)}| \leq |x|, \quad 1 \leq j \leq k+1, \tag{2.25}$$

and

$$|\bar{x}^{(l_1, \cdots, l_{j+1})}| \leq |\bar{x}^{(l_1, \cdots, l_j)}|, \quad 1 \leq j \leq k. \tag{2.26}$$

Since O_{k+1}^0 does not contain the singular point $\mathbf{0} \in \mathbb{R}^{k+1}$, by (2.18) and $g \in \mathcal{S}(\mathbb{R}^{k+1})$, there exists $C_{k+1} < \infty$ such that

$$\int_{O_{k+1}^0} \left| \partial_x^\nu \big(g^{m_0+k+1}(x)e^{|x|} \big) \right| dx \leq C_{k+1}$$

with $\nu = (1, 1, \cdots, 1)$ being a $(k+1)$-dimensional index. This along with (2.25) and the fundamental theorem of calculus yields that for any $x \in O_{k+1}^0$,

$$\begin{aligned}
& e^{|x|} g^{m_0+k+1}(x) \\
\leq & \sum_{l_1 \in D_1} g^{m_0+k+1}(x^{(l_1)}) e^{|x^{(l_1)}|} + \cdots + (-1)^{k+1} \sum_{(l_1, \cdots, l_{k+1}) \in D_{k+1}} g^{m_0+k+1}(x^{(l_1, \cdots, l_{k+1})}) e^{|x^{(l_1, \cdots, l_{k+1})}|} + C_{k+1} \\
\leq & \sum_{l_1 \in D_1} g^{m_0+k+1}(x^{(l_1)}) e^{|\bar{x}^{(l_1)}|+1} + \cdots + \sum_{(l_1, \cdots, l_{k+1}) \in D_{k+1}} g^{m_0+k+1}(x^{(l_1, \cdots, l_{k+1})}) e^{|\bar{x}^{(l_1, \cdots, l_{k+1})}|+k+1} + C_{k+1}
\end{aligned}$$

$$=:\sum_{j=1}^{k+1} I_j + C_{k+1}. \tag{2.27}$$

In order to obtain (2.23), we need to show that $\sum_{j=1}^{k+1} I_j$ has a universal upper bound for all $x \in O^0_{k+1}$. In view of the induction when $d \le k$, we just need to prove that for any $(l_1, \cdots, l_j) \in D_j$ with $1 \le j \le k$,

$$\int g^{m_0+j}(x^{(l_1,l_2,\cdots,l_j)}) e^{|\bar{x}^{(l_1,l_2,\cdots,l_j)}|} dx^{(l_1,\cdots,l_j)} < \infty. \tag{2.28}$$

To this end, we proceed by a further induction argument on the index j to obtain (2.28).

First, in the case where $j = 1$, by (2.18) and (2.25), we have

$$\int g^{m_0+1}(x) e^{|\bar{x}^{(l_1)}|} dx \le \int g^{m_0+1}(x) e^{|x|} dx < \infty. \tag{2.29}$$

In particular, there exists $x^*_{l_1} > 1$ such that

$$\int g^{m_0+1}(x^{*,(l_1)}) e^{|\bar{x}^{(l_1)}|} dx^{(l_1)} < \infty, \tag{2.30}$$

where $x^{*,(l_1)} = (x_1, \cdots, x_{l_1-1}, x^*_{l_1}, x_{l_1+1}, \cdots, x_{k+1})$. Moreover, using $g \in \mathcal{S}(\mathbb{R}^{k+1})$, (2.18) and (2.25) once more we obtain

$$\Big| \int \partial_{x_{l_1}} \big(g^{m_0+1}(x) \big) e^{|\bar{x}^{(l_1)}|} dx \Big| \le \int \Big| \partial_{x_{l_1}} \big(g^{m_0+1}(x) \big) \Big| e^{|x|} dx < \infty.$$

Then, integrating over $x^{(l_1)}$ in $\mathbb{R}^k$ and over the x_{l_1}-coordinate from 1 to $x^*_{l_1}$, we get that there exists $C^{(l_1)} < \infty$ such that

$$\Big| \int g^{m_0+1}(x^{*,(l_1)}) e^{|\bar{x}^{(l_1)}|} dx^{(l_1)} - \int g^{m_0+1}(x^{(l_1)}) e^{|\bar{x}^{(l_1)}|} dx^{(l_1)} \Big| \le C^{(l_1)},$$

which along with (2.30) yields that

$$\int g^{m_0+1}(x^{(l_1)}) e^{|\bar{x}^{(l_1)}|} dx^{(l_1)} \le \int g^{m_0+1}(x^{*,(l_1)}) e^{|\bar{x}^{(l_1)}|} dx^{(l_1)} + C^{(l_1)} < \infty.$$

Thus, (2.28) holds when $j = 1$.

Next, we assume that (2.28) holds for $1 \le j \le n \le k-1$ and consider the case where $j = n+1$. According to $g \in \mathcal{S}(\mathbb{R}^{k+1})$, (2.26) and the induction hypothesis (2.28) when $j = n$, we have

$$\begin{aligned}
&\int g^{m_0+n+1}(x^{(l_1,l_2,\cdots,l_n)}) e^{|\bar{x}^{(l_1,l_2,\cdots,l_{n+1})}|} dx^{(l_1,l_2,\cdots,l_n)} \\
\le& C \int g^{m_0+n}(x^{(l_1,l_2,\cdots,l_n)}) e^{|\bar{x}^{(l_1,l_2,\cdots,l_n)}|} dx^{(l_1,l_2,\cdots,l_n)} < \infty,
\end{aligned} \tag{2.31}$$

and

$$\Big| \int \partial_{x_{l_{n+1}}} \big(g^{m_0+n+1}(x^{(l_1,l_2,\cdots,l_n)}) \big) e^{|\bar{x}^{(l_1,l_2,\cdots,l_{n+1})}|} dx^{(l_1,l_2,\cdots,l_n)} \Big| < \infty. \tag{2.32}$$

In particular, the finiteness in (2.31) implies that there exists $x^*_{l_{n+1}} > 1$ such that

$$\int g^{m_0+n+1}(x^{*,(l_1,l_2,\cdots,l_{n+1})}) e^{|\bar{x}^{(l_1,l_2,\cdots,l_{n+1})}|} dx^{(l_1,l_2,\cdots,l_{n+1})} < \infty, \tag{2.33}$$

where $x^{*,(l_1,l_2,\cdots,l_{n+1})} = (x_1, \cdots, x_{l_1-1}, 1, x_{l_1+1}, \cdots, x_{l_{n+1}-1}, x^*_{l_{n+1}}, x_{l_{n+1}+1}, \cdots, x_{k+1})$. Then, integrating the integrand of (2.32) for $x^{(l_1,l_2,\cdots,l_{n+1})}$ in $\mathbb{R}^{k-n}$ and for the $x_{l_{n+1}}$-coordinate from 1 to $x^*_{l_{n+1}}$ we get that there exists $C^{(l_1,l_2,\cdots,l_{n+1})} < \infty$ such that

$$\begin{aligned}
&\Big| \int g^{m_0+n+1}(x^{*,(l_1,l_2,\cdots,l_{n+1})}) e^{|\bar{x}^{(l_1,l_2,\cdots,l_{n+1})}|} dx^{(l_1,l_2,\cdots,l_{n+1})} \\
&- \int g^{m_0+n+1}(x^{(l_1,l_2,\cdots,l_{n+1})}) e^{|\bar{x}^{(l_1,l_2,\cdots,l_{n+1})}|} dx^{(l_1,l_2,\cdots,l_{n+1})} \Big| \le C^{(l_1,l_2,\cdots,l_{n+1})}.
\end{aligned} \tag{2.34}$$

Hence, it follows from (2.33) and (2.34) that

$$\int g^{m_0+n+1}(x^{(l_1,l_2,\cdots,l_{n+1})}) e^{|\bar{x}^{(l_1,l_2,\cdots,l_{n+1})}|} dx^{(l_1,l_2,\cdots,l_{n+1})}$$

$$\leq \int g^{m_0+n+1}(x^{*,(l_1,l_2,\cdots,l_{n+1})})e^{|\bar{x}^{(l_1,l_2,\cdots,l_{n+1})}|}dx^{(l_1,l_2,\cdots,l_{n+1})} + C^{(l_1,l_2,\cdots,l_{n+1})} < \infty,$$

which yields (2.28) in the case where $j = n+1$.

Thus, by induction over j, we obtain (2.28) for any $j \in \{1, \cdots, k\}$. This in turn implies that $\sum_{j=1}^{k} I_j$ in (2.27) is uniformly bounded, and so, (2.23) follows. Analogous arguments also lead to (2.24) for the remaining regimes $O_{k+1}^{(l_1,\cdots,l_j)}$, $(l_1, \cdots, l_j) \in D_j$, $1 \leq j \leq k+1$, and thus yield (2.19) for $d = k+1$. Therefore, using the induction argument again, we prove the claim for any $d \geq 1$ and finish the proof. □

2.2. **Geometrical decomposition.** In this subsection we derive the geometrical decomposition of solutions to the rescaled random NLS with modulated final data

$$\begin{cases} i\partial_t u + (\Delta + b_* \cdot \nabla + c_*)u + |u|^{p-1}u = 0, \\ u(T) = R(T) + i \sum\limits_{k,\pm} b_k^{\pm} Y_k^{\pm}(T). \end{cases} \tag{2.35}$$

Here, $p \in (1 + 4/d, 1 + 4/(d-2)_+)$, $T > 0$ is sufficiently large, b_* and c_* are the random coefficients given by (1.18) and (1.19), respectively. Moreover,

$$R = \sum_{k=1}^{K} R_k \tag{2.36}$$

with the soliton profiles R_k given by (1.13), $Y_k^{\pm}$ are the rescaled eigenfunctions of the form

$$Y_k^{\pm}(t,x) = Y_{w_k}^{\pm}(y_k) e^{i(\frac{1}{2}v_k \cdot x - \frac{1}{4}|v_k|^2 t + (w_k)^{-2}t + \theta_k^0)}, \tag{2.37}$$

with

$$Y_{w_k}^{\pm}(y_k) = (w_k)^{-\frac{2}{p-1}} Y^{\pm}\Big(\frac{y_k}{w_k}\Big), \quad \text{and} \quad y_k = x - v_k t - \alpha_k^0, \tag{2.38}$$

and $Y^{\pm}$ defined in (2.5). We also denote the vector of the perturbation parameters in the initial data by

$$\boldsymbol{b} := (b_1^+, \cdots, b_K^+, b_1^-, \cdots, b_K^-) \in \mathbb{R}^{2K}. \tag{2.39}$$

Remark 2.3. We note that, unlike equation (3.1) of [42], the rescaled random NLS (2.35) has the additional modulated term $i \sum_{k,\pm} b_k^{\pm} Y_k^{\pm}(T)$ in the final condition. It is introduced mainly to control the extra unstable directions of the linearized Schrödinger operator $\mathscr{L}$ in the mass-supercritical case.

For convenience, we set $\mathcal{P}_k := (\alpha_k, \theta_k) \in \mathbb{Y} := \mathbb{R}^d \times \mathbb{R}$, $1 \leq k \leq K$, and $\mathcal{P} := (\mathcal{P}_1, \mathcal{P}_2, \cdots, \mathcal{P}_K) \in \mathbb{Y}^K$.

Proposition 2.4 (Geometrical decomposition)**.** *Let u be a local solution to equation* (2.35)*. Then, there exist deterministic constants M, $\eta > 0$, such that the following holds:*

For any $T \geq M$ and any $\boldsymbol{b} \in B_{\mathbb{R}^{2K}}(\eta)$, there exist $T^* \in (0,T)$ and unique modulation parameters $\mathcal{P} \in C^1\left([T^*, T]; \mathbb{Y}^K\right)$, such that u admits the unique geometrical decomposition

$$u(t,x) = \sum_{k=1}^{K} \widetilde{R}_k(t,x) + \varepsilon(t,x) \ (=: \widetilde{R}(t,x) + \varepsilon(t,x)), \tag{2.40}$$

with the modulated soliton profile given by

$$\widetilde{R}_k(t,x) := Q_{w_k}\left(x - v_k t - \alpha_k(t)\right) e^{i\left(\frac{1}{2}v_k \cdot x - \frac{1}{4}|v_k|^2 t + (w_k)^{-2}t + \theta_k(t)\right)}, \tag{2.41}$$

and the following orthogonality conditions hold on $[T^*, T]$:

$$\mathrm{Re} \int \nabla \widetilde{R}_k(t) \bar{\varepsilon}(t) dx = 0, \quad \mathrm{Im} \int \widetilde{R}_k(t) \bar{\varepsilon}(t) dx = 0, \quad \forall 1 \leq k \leq K. \tag{2.42}$$

Moreover, the value of modulation parameters (α_k, θ_k) and remainder ε at time T satisfy

$$\|\varepsilon(T)\|_{H^1} + \sum_{k=1}^{K} \left(|\alpha_k(T) - \alpha_k^0| + |\theta_k(T) - \theta_k^0|\right) \leq C|\boldsymbol{b}|, \tag{2.43}$$

for some deterministic positive constant C independent of T.

Remark 2.5. (i) The orthogonality conditions in (2.42) allow to control the first two unstable directions arising from the coercive property (2.6) of the linearized Schrödinger operator. The remaining two unstable directions will be controlled in Proposition 2.7 and Proposition 3.4 below by selecting an appropriate vector $\boldsymbol{b}$ in the final condition of equation (2.35).

(ii) It is worth noting that the control of $\varepsilon(T), \alpha_k(T)$ and $\theta_k(T)$ in (2.43) is used to obtain the coercivity estimate of $\|\varepsilon(t)\|_{H^1}$ in Proposition 3.6, as well as the bootstrap estimates of $\sum_{k=1}^K(|\alpha_k(t)-\alpha_k^0|+|\theta_k(t)-\theta_k^0|)$ in Proposition 4.3. Moreover, estimate (2.43) is also used to derive the a-priori estimates (3.4) and (3.5) on $[T^*, T]$, which guarantee that all constants appearing in the estimates of modulation equations and functionals are deterministic and uniformly bounded.

In order to prove Proposition 2.4, let us first present the following result. It can be proved in an analogous manner as in the proof of [42, Lemma 6.4]. Thus, the proof is omitted here.

Lemma 2.6. *Given any $L, w_k \in \mathbb{R}^+$, α_k^0, $v_k \in \mathbb{R}^d$, $\theta_k^0 \in \mathbb{R}$, $1 \le k \le K$. Set*

$$R_L(x) := \sum_{k=1}^K R_{k,L}(x) = \sum_{k=1}^K Q_{w_k}(x - v_k L - \alpha_k^0) e^{i(\frac{1}{2}v_k\cdot x - \frac{1}{4}|v_k|^2 L + (w_k)^{-2}L + \theta_k^0)}. \tag{2.44}$$

Then, there exists a deterministic small constant $\delta_* > 0$ such that the following holds:

For any $0 < r, L^{-1} < \delta_*$ and for any $u \in H^1(\mathbb{R}^d)$ satisfying $\|u - R_L\|_{H^1} \le r$, there exist a unique C^1 function $\mathcal{P}(u) = (\alpha, \theta)$: $H^1 \to \mathbb{Y}^K$ with $\alpha := (\alpha_1, \alpha_2, \cdots, \alpha_K)$ and $\theta := (\theta_1, \theta_2, \cdots, \theta_K)$, such that u admits the decomposition

$$\begin{aligned} u &= \sum_{k=1}^K Q_{w_k}(x - v_k L - \alpha_k) e^{i(\frac{1}{2}v_k\cdot x - \frac{1}{4}|v_k|^2 L + (w_k)^{-2}L + \theta_k)} + \varepsilon_L \\ &=: \sum_{k=1}^K \widetilde{R}_{k,L} + \varepsilon_L, \end{aligned} \tag{2.45}$$

and $\widetilde{R}_{k,L}$, ε_L satisfy the orthogonality conditions:

$$Re\int \nabla \widetilde{R}_{k,L}\bar{\varepsilon}_L dx = \mathbf{0}, \quad Im\int \widetilde{R}_{k,L}\bar{\varepsilon}_L dx = 0, \quad 1 \le k \le K. \tag{2.46}$$

Moreover, there exists a deterministic constant $C > 0$ such that

$$\|\varepsilon_L\|_{H^1} + \sum_{k=1}^K \left(|\alpha_k - \alpha_k^0| + |\theta_k - \theta_k^0|\right) \le C\|u - R_L\|_{H^1}. \tag{2.47}$$

Now, Proposition 2.4 follows easily from Lemma 2.6.

Proof of Proposition 2.4. Let δ_* be as in Lemma 2.6 and $M = 2\delta_*^{-1}$. For any $T \ge M$, using (2.4) and the explicit expression (2.37) we estimate that for every $1 \le k \le K$,

$$\|Y_k^{\pm}(T)\|_{H^1} \le C(\|Y^{\pm}\|_{L^2} + \|\nabla Y^{\pm}\|_{L^2}) \le C, \tag{2.48}$$

which yields that for any $\boldsymbol{b} \in B_{\mathbb{R}^{2K}}(\eta)$ with η sufficiently small, $\|u(T) - R(T)\|_{H^1} \le \delta_*/2$. Then, by the local well posedness theory there exists $T^*(\ge \delta_*^{-1})$ close to T, such that $u(t) \in C([T^*, T]; H^1)$ and $\|u(t) - R(t)\|_{H^1} \le \delta_*$ for any $t \in [T^*, T]$.

Thus, by virtue of Lemma 2.6, there exist C^1 functions $(\alpha_k(t), \theta_k(t)) \in C^1([T^*, T]; \mathbb{Y})$, $1 \le k \le K$, such that for any $t \in [T^*, T]$, $u(t)$ admits the decomposition (2.45), and the orthogonality conditions in (2.46) hold with t replacing L, which verify (2.41) and (2.42). At last, (2.43) follows from (2.47) and (2.48). □

2.3. **Modulated final data.** Let $(\alpha_k, \theta_k) \in C^1([T^*, T]; \mathbb{Y})$ and ε be remainder from Proposition 2.4. Set

$$\widetilde{Y}_k^{\pm}(t, x) := Y_{w_k}^{\pm}(y_k(t, x)) e^{i\Phi_k(t,x)}, \tag{2.49}$$

with

$$y_k(t, x) := x - v_k t - \alpha_k(t) \tag{2.50}$$

and the phase function

$$\Phi_k(t,x) := \frac{1}{2}v_k \cdot x - \frac{1}{4}|v_k|^2 t + (w_k)^{-2}t + \theta_k(t). \tag{2.51}$$

Let

$$\mathbf{a}^{\pm}(t) := \left(a_k^{\pm}(t)\right)_{1\leq k\leq K} \quad \text{with} \;\; a_k^{\pm}(t) := \operatorname{Im}\int \widetilde{Y}_k^{\pm}(t,x)\bar{\varepsilon}(t,x)dx. \tag{2.52}$$

The following result permits to steer the unstable directions to prescribed vectors at the final time.

Proposition 2.7 (Modulated final data)**.** *There exist deterministic positive constants M, $r>0$, such that for any $T \geq M$ and any $\boldsymbol{a}^- \in B_{\mathbb{R}^K}(r)$, there exists a unique $\boldsymbol{b} \in B_{\mathbb{R}^{2K}}(\eta)$, where η is a deterministic positive constant depending on r, such that $\boldsymbol{b}$ depends continuously on $\boldsymbol{a}^-$, $|\boldsymbol{b}| \leq C|\boldsymbol{a}^-|$ and*

$$\mathbf{a}^+(T) = \mathbf{0} \quad \text{and} \quad \mathbf{a}^-(T) = \boldsymbol{a}^-, \tag{2.53}$$

where C is a deterministic positive constant.

Before proving Proposition 2.7, let us first introduce the following decoupling lemma that helps to decouple different soliton profiles and eigenfunctions $Y^{\pm}$, thanks to the exponential decay results (1.10) and Lemma 2.2. Its proof follows in an analogous manner as in the proof of Lemma 6.3 in [42].

Lemma 2.8 (Decoupling lemma)**.** *Let δ_0 be the minimum of the positive constants δ in* (1.10) *and* (2.7)*. Assume that $g_i \in C_b^2$, $i=1,2$, satisfy*

$$|g_i(x)| \leq C_1 e^{-\delta_0|x|}, \;\; x\in\mathbb{R}^d, \;\; i=1,2 \tag{2.54}$$

for some positive constant C_1. For every $1\leq k\leq K$, let

$$G_{i,k}(t,x) := (w_k)^{-\frac{2}{p-1}}\, g_i\left(\frac{x-v_kt-\alpha_k}{w_k}\right), \;\; i=1,2, \tag{2.55}$$

where $p \in \left(1+\frac{4}{d}, 1+\frac{4}{(d-2)_+}\right)$, $v_j \neq v_k$ for any $j\neq k$, and the parameters $w_k \in \mathbb{R}^+, v_k, \alpha_k \in \mathbb{R}^d$ satisfy

$$(w_k)^{-1} + w_k + |v_k| + |\alpha_k| \leq C_2, \tag{2.56}$$

for some positive constant C_2.

Then, we have that for any $j\neq k$ and p_1, $p_2>0$,

$$\int |G_{1,j}(t,x)|^{p_1}|G_{2,k}(t,x)|^{p_2}dx \leq Ce^{-\delta_2 t}, \tag{2.57}$$

where C and $\delta_2(>0)$ depend on δ_0, C_1, C_2, p_1 and p_2.

Now, we come to the proof of Proposition 2.7.

Proof of Proposition 2.7. Define the map

$$G_1 : \mathbb{R}^{2K} \mapsto H^1 \;\; \text{by} \;\; G_1(\boldsymbol{b}) = i\sum_{k,\pm} b_k^{\pm}Y_k^{\pm},$$

where $\boldsymbol{b}$ is given by (2.39) and $Y_k^{\pm}$ are the rescaled eigenfunctions given by (2.37).

Moreover, in view of Lemma 2.6, for any $v\in B_{H^1}(\delta_*)$, there exist unique $\alpha(v) = (\alpha_1(v), \alpha_2(v), \cdots, \alpha_K(v)) \in (\mathbb{R}^d)^K$ and $\theta(v) = (\theta_1(v), \theta_2(v), \cdots, \theta_K(v)) \in \mathbb{R}^K$, such that the geometrical decomposition (2.45) and the orthogonal conditions (2.46) hold, where u is replaced by $v+R(T)$, and $R(T)$ is as in (2.44) with L replaced by T. Then, we define the second map

$$G_2 : B_{H^1}(\delta_*) \mapsto H^1 \times (\mathbb{R}^d)^K \times \mathbb{R}^K,$$

by

$$G_2(v) = (\varepsilon, \alpha, \theta) := \left(v + R(T) - \sum_{k=1}^{K} Q_{w_k}(y_k(\alpha(v)))e^{i\Phi_k(\theta(v))}, \alpha(v), \theta(v)\right), \tag{2.58}$$

where

$$y_k(\alpha(v)) = x - v_kT - \alpha_k(v), \quad \Phi_k(\theta(v)) = \frac{1}{2}v_k\cdot x - \frac{1}{4}|v_k|^2T + (w_k)^{-2}T + \theta_k(v). \tag{2.59}$$

Let us further define the third map

$$G_3 : H^1 \times (\mathbb{R}^d)^K \times \mathbb{R}^K \mapsto \mathbb{R}^{2K},$$

by

$$G_3(\varepsilon, \alpha, \theta) := \Big(\mathrm{Im} \int \widetilde{Y}_k^{\pm} \bar{\varepsilon} dx\Big)_{1 \le k \le K, \pm}, \tag{2.60}$$

where $\widetilde{Y}_k^{\pm}$ is as in (2.49) with $y_k(t,x)$ and $\Phi_k(t,x)$ replaced by $y_k(\alpha)$ and $\Phi_k(\theta)$, respectively. For simplicity, we drop the parameter T in the following arguments.

Claim: $G = G_3 \circ G_2 \circ G_1$ is a diffeomorphism in a small neighborhood of $\mathbf{0} \in \mathbb{R}^{2K}$.

In order to prove this claim, by the chain rule, we shall compute the derivatives dG_i, $i = 1,2,3$, separately in the following. Before that, let us note that since by (2.48), $\|Y_k^{\pm}\|_{H^1}$ is bounded, there exists a deterministic constant $\eta' > 0$ such that for any $\boldsymbol{b} \in B_{\mathbb{R}^{2K}}(\eta')$, $\|G_1(\boldsymbol{b})\|_{H^1} \le \delta_*$. Hence, the map $G_2 \circ G_1$ is well-defined on $B_{\mathbb{R}^{2K}}(\eta')$.

(i) dG_1: We compute that

$$dG_1 = (iY_1^+, iY_2^+, \cdots, iY_K^+, iY_1^-, iY_2^-, \cdots, iY_K^-),$$

which yields that for any $\boldsymbol{b} \in \mathbb{R}^{2K}$,

$$dG_1.\boldsymbol{b} = i \sum_{k,\pm} b_k^{\pm} Y_k^{\pm}. \tag{2.61}$$

(ii) dG_2: Next, let us consider the second map G_2 given by (2.58). Set $F_1^j := (f_1^j, f_2^j, \cdots, f_d^j)^{\top}$ with $f_l^j(v,\alpha,\theta) := \mathrm{Re} \int \partial_l \widetilde{R}_j \bar{\varepsilon} dx$, and $F_2^j(v,\alpha,\theta) := \mathrm{Im} \int \widetilde{R}_j \bar{\varepsilon} dx$, $1 \le l \le d$, $1 \le j \le K$, where $\widetilde{R}_j$ is as in (2.41) with $t, \alpha_j(t), \theta_j(t)$ replaced by T, α_j, θ_j, respectively, $1 \le j \le K$.

By straightforward computations, we have that for any $h \in H^1$,

$$dG_2.h = \left(h + \sum_{k=1}^K \sum_{l=1}^d (\partial_l Q_{w_k}) e^{i\Phi_k(\theta)} \Big(\frac{d\alpha_{k,l}}{dv}.h\Big) - i \sum_{k=1}^K \widetilde{R}_k \Big(\frac{d\theta_k}{dv}.h\Big), \Big(\frac{d\alpha_k}{dv}.h\Big)_{1 \le k \le K}, \Big(\frac{d\theta_k}{dv}.h\Big)_{1 \le k \le K} \right), \tag{2.62}$$

where $\frac{d\alpha_k}{dv}.h := (\frac{d\alpha_{k,1}}{dv}.h, \frac{d\alpha_{k,2}}{dv}.h, \cdots, \frac{d\alpha_{k,d}}{dv}.h)^{\top}$.

In order to compute the directional derivatives $\frac{d\alpha_{k,l}}{dv}.h$ and $\frac{d\theta_k}{dv}.h$, we let $\frac{\partial F_1^j}{\partial \alpha_k}$, $\frac{\partial F_1^j}{\partial \theta_k}$ and $\frac{\partial F_2^j}{\partial \alpha_k}$ denote the following three Jacobian matrices, respectively,

$$\frac{\partial F_1^j}{\partial \alpha_k} := \begin{bmatrix} \frac{\partial f_1^j}{\partial \alpha_{k,1}} & \frac{\partial f_1^j}{\partial \alpha_{k,2}} & \cdots & \frac{\partial f_1^j}{\partial \alpha_{k,d}} \\ \frac{\partial f_2^j}{\partial \alpha_{k,1}} & \frac{\partial f_2^j}{\partial \alpha_{k,2}} & \cdots & \frac{\partial f_2^j}{\partial \alpha_{k,d}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_d^j}{\partial \alpha_{k,1}} & \frac{\partial f_d^j}{\partial \alpha_{k,2}} & \cdots & \frac{\partial f_d^j}{\partial \alpha_{k,d}} \end{bmatrix}, \quad \frac{\partial F_1^j}{\partial \theta_k} := \begin{bmatrix} \frac{\partial f_1^j}{\partial \theta_k} \\ \frac{\partial f_2^j}{\partial \theta_k} \\ \vdots \\ \frac{\partial f_d^j}{\partial \theta_k} \end{bmatrix}, \quad \frac{\partial F_2^j}{\partial \alpha_k} := \begin{bmatrix} \frac{\partial F_2^j}{\partial \alpha_{k,1}} & \frac{\partial F_2^j}{\partial \alpha_{k,2}} & \cdots & \frac{\partial F_2^j}{\partial \alpha_{k,d}} \end{bmatrix}. \tag{2.63}$$

By straightforward computations and the decoupling Lemma 2.8,

$$\begin{aligned} &\frac{\partial f_l^j}{\partial \alpha_{j,l}} = (w_j)^{-2} \|\partial_l Q_{w_j}\|_{L^2}^2 + \mathcal{O}(\|\varepsilon\|_{L^2}), \ \frac{\partial f_l^j}{\partial \theta_j} = -\frac{v_{j,l}}{2} \|Q_{w_j}\|_{L^2}^2 + \mathcal{O}(\|\varepsilon\|_{L^2}), \\ &\frac{\partial F_2^j}{\partial \theta_j} = \|Q_{w_j}\|_{L^2}^2 + \mathcal{O}(\|\varepsilon\|_{L^2}), \quad 1 \le l \le d, \quad 1 \le j \le K. \end{aligned} \tag{2.64}$$

The other terms in matrices (2.63) are of small order $\mathcal{O}(\|\varepsilon\|_{L^2} + e^{-\delta_2 T})$.

In view of the orthogonality conditions in (2.46), we have

$$F_1^j(v, \alpha(v), \theta(v)) = \mathbf{0}, \ F_2^j(v, \alpha(v), \theta(v)) = 0.$$

Then, differentiating the above identities with respect to v we get

$$\begin{aligned}
&\partial_v F_1^j + \sum_{k=1}^K \frac{\partial F_1^j}{\partial \alpha_k}\left(\frac{d\alpha_k}{dv}\right) + \sum_{k=1}^K \frac{\partial F_1^j}{\partial \theta_k}\frac{d\theta_k}{dv} = \mathbf{0}, \\
&\partial_v F_2^j + \sum_{k=1}^K \frac{\partial F_2^j}{\partial \alpha_k}\cdot\frac{d\alpha_k}{dv} + \sum_{k=1}^K \frac{\partial F_2^j}{\partial \theta_k}\frac{d\theta_k}{dv} = 0,
\end{aligned} \tag{2.65}$$

where $\frac{\partial F_1^j}{\partial \alpha_k}(\frac{d\alpha_k}{dv})$ means that the Jacobian matrix $\frac{\partial F_1^j}{\partial \alpha_k}$ acts on the vector $\frac{d\alpha_k}{dv}$, and $\frac{\partial F_2^j}{\partial \alpha_k}\cdot\frac{d\alpha_k}{dv}$ denotes the inner product between two vectors $\frac{\partial F_2^j}{\partial \alpha_k}$ and $\frac{d\alpha_k}{dv}$.

Plugging (2.63) and (2.64) into (2.65) and using (2.47) we compute that for any $h \in H^1$,

$$\begin{aligned}
&\frac{d\alpha_{k,l}}{dv}.h = -(w_k)^2\|\partial_l Q_{w_k}\|_{L^2}^{-2}\Big(\frac{1}{2}v_{k,l}\mathrm{Im}\int \widetilde{R}_k\bar{h}dx + \mathrm{Re}\int \partial_l\widetilde{R}_k\bar{h}dx\Big) + \mathcal{O}\big(\|h\|_{L^2}(\|v\|_{H^1} + e^{-\delta_2 T})\big), \\
&\frac{d\theta_k}{dv}.h = -\|Q_{w_k}\|_{L^2}^{-2}\mathrm{Im}\int \widetilde{R}_k\bar{h}dx + \mathcal{O}\big(\|h\|_{L^2}(\|v\|_{H^1} + e^{-\delta_2 T})\big), \quad 1 \le l \le d, \quad 1 \le k \le K.
\end{aligned} \tag{2.66}$$

(iii) dG_3: Regarding the third map G_3 given by (2.60), we compute that for any $h \in H^1$, $\beta = (\beta_j)_{1\le j\le K} \in \mathbb{R}^{d\times K}$ with $\beta_j \in \mathbb{R}^d$, $\gamma \in \mathbb{R}^K$,

$$\big(dG_3(\varepsilon,\alpha,\theta).(h,\beta,\gamma)\big)_{k,\pm} = \mathrm{Im}\int \widetilde{Y}_k^{\pm}\bar{h}dx + \sum_{j=1}^K\sum_{l=1}^d \beta_{j,l}\partial_{\alpha_{j,l}}\mathrm{Im}\int \widetilde{Y}_k^{\pm}\bar{\varepsilon}dx + \sum_{j=1}^K \gamma_j\partial_{\theta_j}\mathrm{Im}\int \widetilde{Y}_k^{\pm}\bar{\varepsilon}dx.$$

Note that

$$\begin{aligned}
&\partial_{\alpha_{j,l}}\mathrm{Im}\int \widetilde{Y}_k^{\pm}\bar{\varepsilon}dx = -\delta_{jk}\ \mathrm{Im}\int (w_k)^{-\frac{p+1}{p-1}}(\partial_l Y^{\pm})\Big(\frac{y_k(\alpha)}{w_k}\Big)e^{i\Phi_k(\theta)}\bar{\varepsilon}dx, \\
&\partial_{\theta_j}\mathrm{Im}\int \widetilde{Y}_k^{\pm}\bar{\varepsilon}dx = \delta_{jk}\ \mathrm{Re}\int \widetilde{Y}_k^{\pm}\bar{\varepsilon}dx,\ 1 \le j \le K,\ l \le l \le d.
\end{aligned}$$

Thus,

$$\big(dG_3(\varepsilon,\alpha,\theta).(h,\beta,\gamma)\big)_{k,\pm} = \mathrm{Im}\int \widetilde{Y}_k^{\pm}\bar{h}dx - \beta_k\cdot\mathrm{Im}\int (w_k)^{-\frac{p+1}{p-1}}(\nabla Y^{\pm})\Big(\frac{y_k(\alpha)}{w_k}\Big)e^{i\Phi_k(\theta)}\bar{\varepsilon}dx + \gamma_k\mathrm{Re}\int \widetilde{Y}_k^{\pm}\bar{\varepsilon}dx. \tag{2.67}$$

(iv) Non-degeneracy of dG: Now, for any $\boldsymbol{c} = \big(c_1^+,\cdots,c_K^+,c_1^-,\cdots,c_K^-\big) \in \mathbb{R}^{2K}$, let $v = G_1(\boldsymbol{b}) \in B_{H^1}(\delta_*) \subseteq H^1$, $h = G_1(\boldsymbol{c}) \in H^1$. Applying the chain rule we have

$$dG(\boldsymbol{b}).\boldsymbol{c} = dG_3\big(G_2(G_1(\boldsymbol{b}))\big).\big(dG_2(G_1(\boldsymbol{b})).(dG_1.\boldsymbol{c})\big). \tag{2.68}$$

By (2.62), we have

$$\begin{aligned}
dG_2(G_1(\boldsymbol{b})).G_1(\boldsymbol{c}) = \Bigg(&G_1(\boldsymbol{c}) + \sum_{k=1}^K\sum_{l=1}^d (\partial_l Q_{w_k})\,e^{i\Phi_k(\theta(G_1(\boldsymbol{b})))}\Big(\frac{d\alpha_{k,l}}{dv}.G_1(\boldsymbol{c})\Big) - i\sum_{k=1}^K \widetilde{R}_k\Big(\frac{d\theta_k}{dv}.G_1(\boldsymbol{c})\Big), \\
&\Big(\frac{d\alpha_k}{dv}.G_1(\boldsymbol{c})\Big)_{1\le k\le K}, \Big(\frac{d\theta_k}{dv}.G_1(\boldsymbol{c})\Big)_{1\le k\le K}\Bigg).
\end{aligned} \tag{2.69}$$

Set $Y_1 = \mathrm{Re}Y^+$, $Y_2 = \mathrm{Im}Y^+$. Note that, by (2.47), Lemma 2.8 and the algebraic identity

$$\mathrm{Re}\int QY^{\pm}dx = \int QY_1dx = -e_0^{-1}\int Q(L_-Y_2)dx = -e_0^{-1}\int (L_-Q)Y_2dx = 0, \tag{2.70}$$

we have

$$\begin{aligned}
\mathrm{Im}\int \widetilde{R}_k\overline{G_1(\boldsymbol{c})}dx =&\mathrm{Re}\int \widetilde{R}_k c_k^{\pm}\bar{\widetilde{Y}}_k^{\pm}dx + \mathrm{Re}\int \widetilde{R}_k c_k^{\pm}\big(\bar{Y}_k^{\pm} - \bar{\widetilde{Y}}_k^{\pm}\big)dx + \sum_{j\neq k,\pm}\mathrm{Re}\int \widetilde{R}_k c_j^{\pm}\bar{Y}_j^{\pm}dx \\
\le& C\mathrm{Re}\int QY^{\pm}dx + C|c_k^{\pm}|(|\alpha_k(G_1(\boldsymbol{b})) - \alpha_k^0| + |\theta_k(G_1(\boldsymbol{b})) - \theta_k^0|) + C|\mathbf{c}|e^{-\delta_2 T}
\end{aligned}$$

$$=\mathcal{O}\big(|\boldsymbol{c}|(e^{-\delta_2 T}+|\boldsymbol{b}|)\big), \tag{2.71}$$

where $\widetilde{Y}_k$ is as in (2.49) with $t, \alpha_k(t), \theta_k(t)$ replaced by T, $\alpha_k(G_1(\boldsymbol{b}))$, $\theta_k(G_1(\boldsymbol{b}))$, respectively, $1 \le k \le K$. Similarly, using the algebraic identity

$$\mathrm{Im}\int \nabla Q Y^{\pm} dx = \pm \int \nabla Q Y_2 dx = \pm e_0^{-1}\int \nabla Q (L_+ Y_1) dx = \pm e_0^{-1}\int (L_+\nabla Q) Y_1 dx = \mathbf{0}, \tag{2.72}$$

and (2.47), Lemma 2.8 again we obtain

$$\mathrm{Re}\int \partial_l \widetilde{R}_k \overline{G_1(\boldsymbol{c})} dx = \mathcal{O}\big(|\boldsymbol{c}|(e^{-\delta_2 T}+|\boldsymbol{b}|)\big), \quad l=1,2,\cdots,d. \tag{2.73}$$

Plugging the above estimates (2.71) and (2.73) into (2.66) and (2.69) we have

$$dG_2\left(G_1(\boldsymbol{b})\right).G_1(\boldsymbol{c}) = (G_1(\boldsymbol{c}), \mathbf{0}, \mathbf{0}) + \mathcal{O}\big(|\boldsymbol{c}|(e^{-\delta_2 T}+|\boldsymbol{b}|)\big).$$

Inserting this into (2.68) and using (2.67) we thus come to

$$dG(\boldsymbol{b}).\boldsymbol{c} = \big(A_1^+, \cdots, A_K^+, A_1^-, \cdots, A_K^-\big)^\top + \mathcal{O}\big(|\boldsymbol{c}|(e^{-\delta_2 T}+|\boldsymbol{b}|)\big)$$

with

$$A_k^+ = -\sum_{j,\pm} c_j^{\pm}\ \mathrm{Re}\int Y_k^+ \bar{Y}_j^{\pm} dx, \text{ and } A_k^- = -\sum_{j,\pm} c_j^{\pm}\ \mathrm{Re}\int Y_k^- \bar{Y}_j^{\pm} dx, \quad 1 \le k \le K.$$

This along with Lemma 2.8 yields that

$$dG(\boldsymbol{b}) = \begin{bmatrix} A & y_* A \\ y_* A & A \end{bmatrix} + \mathcal{O}\big(e^{-\delta_2 T}+|\boldsymbol{b}|\big), \tag{2.74}$$

where $A := \mathrm{diag}\Big(-(w_k)^{d-\frac{4}{p-1}}\Big)_{1\le k\le K}$ and $y_* := \int Y_1^2 - Y_2^2 dx$. Note that

$$y_*^2 = \left(\int Y_1^2 + Y_2^2 dx\right)^2 - 4\int Y_1^2 dx \int Y_2^2 dx = 1 - 4\int Y_1^2 dx \int Y_2^2 dx < 1.$$

We thus infer that there exist deterministic constants $M > 0$ and $\eta \in (0, \eta')$, such that for any $T \ge M$ and any $\boldsymbol{b} \in B_{\mathbb{R}^{2K}}(\eta)$, the determinant of $dG(\boldsymbol{b})$ is equal to

$$\big(\det(A)\big)^2 \big(1-y_*^2\big)^K + \mathcal{O}\big(e^{-\delta_2 M}+\eta\big) > 0$$

and is uniformly bounded from below, independent of ω. Taking into account $G(\mathbf{0}) = \mathbf{0}$ we obtain that for any $T \ge M$, G is a diffeomorphism from $B_{\mathbb{R}^{2K}}(\eta)$ to some neighborhood U of $\mathbf{0} \in \mathbb{R}^{2K}$, as claimed.

(v) Uniform deterministic ball inside the random image of G: We claim that, though the map G depends on T and the underlying probabilistic argument ω, the image U contains a deterministic small ball $B_{\mathbb{R}^{2K}}(r)$, where $r > 0$ is a sufficiently small deterministic constant.

Actually, by the differential mean value theorem, there exists $\xi = \lambda \boldsymbol{b}$ with $0 \le \lambda \le 1$ such that

$$G(\boldsymbol{b}) = dG(\xi)\cdot \boldsymbol{b}.$$

Using the matrix in (2.74) we estimate

$$\begin{aligned}\big|G(\boldsymbol{b})\big|^2 \ge& \sum_{k=1}^K \Big(\big(y_*^2+1\big)(w_k)^{2d-\frac{8}{p-1}}\big((b_k^+)^2+(b_k^-)^2\big) + 4y_*(w_k)^{2d-\frac{8}{p-1}} b_k^+ b_k^-\Big) - C\big(e^{-\delta_2 T}|\boldsymbol{b}|^2 + |\boldsymbol{b}|^3\big) \\ \ge& \Big(\big(y_*-1\big)^2 (w_*)^{2d-\frac{8}{p-1}} - Ce^{-\delta_2 T}\Big)|\boldsymbol{b}|^2 - C|\boldsymbol{b}|^3,\end{aligned}$$

where $w_* := \min_{1\le k\le K}\{w_k\} > 0$, and C is a positive constant independent of T. Then, for M possibly larger and η possibly even smaller, we get that for any $T \ge M$ and $\boldsymbol{b} \in B_{\mathbb{R}^K}(\eta)$,

$$\big|G(\boldsymbol{b})\big|^2 \ge \frac{1}{2}\big(y_*-1\big)^2 (w_*)^{2d-\frac{8}{p-1}}|\boldsymbol{b}|^2 - C|\boldsymbol{b}|^3 \ge \frac{1}{4}\big(y_*-1\big)^2 (w_*)^{2d-\frac{8}{p-1}}|\boldsymbol{b}|^2,$$

Thus, letting $r := \frac{1}{2}(1-y_*)(w_*)^{d-\frac{4}{p-1}}\eta$ and taking into account $G(\partial B_{\mathbb{R}^K}(\eta)) = \partial U$ we obtain that U contains the deterministic small ball $B_{\mathbb{R}^{2K}}(r)$, as claimed.

Now, for any $\boldsymbol{a}^- \in B_{\mathbb{R}^K}(r)$, $(\boldsymbol{0}, \boldsymbol{a}^-) \in B_{\mathbb{R}^{2K}}(r)$, by the inverse of the diffeomorphism G, there exists a unique $\boldsymbol{b} = \boldsymbol{b}(\boldsymbol{a}^-) \in B_{\mathbb{R}^{2K}}(\eta)$ such that $G(\boldsymbol{b}(\boldsymbol{a}^-)) = (\boldsymbol{0}, \boldsymbol{a}^-)$ and $|\boldsymbol{b}(\boldsymbol{a}^-)| \leq C|\boldsymbol{a}^-|$, where C is a deterministic positive constant as the determinant of $dG(\boldsymbol{b})$ has a deterministic uniform lower bound. Therefore, we obtain (2.53) and finish the proof of Proposition 2.7. □

3. Modulation equations and remainder

In this section, we aim to control the modulation parameters and remainder in the geometrical decomposition (2.40).

To begin with, let us first control the noise appearing in the rescaling transform (1.17).

3.1. **Control of noise.** Set $B_{*,k}(t) := \int_t^\infty g_k(s)dB_k(s)$, $1 \leq k \leq N$, and

$$B_*(t) := \sup_{t\leq s<\infty} \sum_{k=1}^N |B_{*,k}(s)|, \quad t > 0. \tag{3.1}$$

Since $g_k \in L^2(\mathbb{R}^+)$, one has

$$\lim_{t\to+\infty} B_*(t) = 0, \ \mathbb{P}\text{-a.s..} \tag{3.2}$$

This yields that there exists a large random time σ_1, $\sigma_1 \in (0,\infty)$, $\mathbb{P}$-a.s., such that

$$\sup_{t\geq\sigma_1} B_*(t) \leq 1, \quad \mathbb{P} - a.s. \tag{3.3}$$

In view of Lemma 3.1 and (2.47), for $|\boldsymbol{b}|$ small enough, we can choose $T^* \geq \sigma_1$ such that for any $t \in [T^*, T]$,

$$\sup_{T^*\leq t\leq T} \|\varepsilon(t)\|_{H^1} < 1, \tag{3.4}$$

and for every $1 \leq k \leq K$,

$$B_*(t) + |\alpha_k(t) - \alpha_k^0| \leq \frac{1}{10}\min\{1, w_k, \alpha_k^0\}. \tag{3.5}$$

In particular, $B_*(t)$, $|\alpha_k(t)|$ and $\|\varepsilon(t)\|_{H^1}$ are bounded by a universal deterministic constant on $[T^*, T]$.

We remark that in Section 4 below, thanks to the bootstrap estimates (4.10)-(4.12), both estimates (3.4) and (3.5) are valid after a large random time.

In Case (II), we have the following refined decay estimate of $B_*(t)$. It has been used in [42], for the convenience of reader, we include its proof here.

Lemma 3.1. *In Case (II), there exists a positive random time σ_2 such that $\mathbb{P}$-a.s. $\sigma_2 \geq \sigma_*$, where σ_* is the random time as in* (1.5)*, and for any $t \geq \sigma_2$,*

$$|B_{*,k}(t)| \leq 2\bigg(2\int_t^\infty g_k^2(s)ds \log\Big(\int_t^\infty g_k^2(s)ds\Big)^{-1}\bigg)^{\frac{1}{2}} \leq \frac{2\sqrt{c^*}}{t}, \quad 1 \leq k \leq N. \tag{3.6}$$

In particular, there exists a deterministic constant $C > 0$ suc that $\mathbb{P}$-a.s. for any $t \geq \sigma_2$,

$$B_*(t) \leq \frac{C}{t}. \tag{3.7}$$

Proof. In the following we fix $1 \leq k \leq N$. Let $\sigma_{k,\infty} := \int_0^\infty g_k^2(s)ds$. Then, $\sigma_{k,\infty} \in (0,\infty)$, $\mathbb{P}$-a.s., due to the L^2-integrability of g_k.

In view of the theorem on time-change for martingales, there exists a Brownian motion W_k such that $B_{*,k}(t) = W_k(\sigma_{k,\infty}) - W_k(\int_0^t g_k^2(r)dr)$, $\mathbb{P}$-a.s.. Moreover, by the Levy Hölder continuity of Brownian motions and the invariance under time shift of the law of Brownian motions, i.e., for every $n \in \mathbb{N}$, $W_k(n+\cdot) - W_k(\cdot)$ has the same law as the standard Brownian motion, we note that $\mathbb{P}$-a.s.

$$\lim_{h\to 0} \sup_{\substack{|t-t'|\leq h \\ t,t'\in[n-1,n+1]}} \frac{|W_k(t') - W_k(t)|}{\sqrt{2h\log(1/h)}} = 1 \tag{3.8}$$

for every $n \in \mathbb{N}$. Then, set

$$h_n := \inf \left\{ h \in [0, \frac{1}{10}], \sup_{\substack{|t-t'|\leq h \\ t,t' \in [n-1,n+1]}} \frac{|W_k(t') - W_k(t)|}{\sqrt{2h\log(1/h)}} > 2 \right\} \wedge \frac{1}{10}.$$

We see that h_n is $\mathscr{F}_\infty$-measurable, $0 < h_n \leq 1/10$ $\mathbb{P}$-a.s. due to (3.8), and

$$\sup_{\substack{|t-t'|\leq h \\ t,t' \in [n-1,n+1]}} \frac{|W_k(t') - W_k(t)|}{\sqrt{2h\log(1/h)}} \leq 2, \quad \forall h \leq h_n. \tag{3.9}$$

Let $[\sigma_{k,\infty}]$ be the largest integer less than $\sigma_{k,\infty}$, and define $h_{[\sigma_{k,\infty}]}$ by $h_{[\sigma_{k,\infty}]}(\omega) := (h_{[\sigma_{k,\infty}(\omega)]})(\omega)$ for $\omega \in \Omega$. Then, $h_{[\sigma_{k,\infty}]}$ is $\mathscr{F}_\infty$-measurable. Moreover, by (3.9), we infer that $\mathbb{P}$-a.s.

$$\begin{aligned} |W_k(\sigma_{k,\infty}) - W_k(\sigma_{k,\infty} - h)| \leq & \sup_{\substack{|t-t'|\leq h \\ t,t' \in [[\sigma_{k,\infty}]-1, [\sigma_{k,\infty}]+1]}} |W_k(t') - W_k(t)| \\ \leq & 2\sqrt{2h\log(1/h)}, \quad \forall h \leq h_{[\sigma_{k,\infty}]} \wedge \sigma_{k,\infty}. \end{aligned} \tag{3.10}$$

Now, let

$$\sigma_{k,0} := \inf \left\{ t > 0 : \int_t^\infty g_k^2(s)ds \leq h_{[\sigma_{k,\infty}]} \right\}.$$

Since by the L^2-integrability of g_k, $\int_t^\infty g_k^2(s)ds \to 0$ as $t \to \infty$, $\mathbb{P}$-a.s., one has $0 \leq \sigma_{k,0} < \infty$, $\mathbb{P}$-a.s.. We also see from the definition of $\sigma_{k,0}$ that $\mathbb{P}$-a.s.

$$\int_t^\infty g_k^2(s)ds \leq \int_{\sigma_{k,0}}^\infty g_k^2(s)ds = h_{[\sigma_{k,\infty}]}, \quad \forall t \geq \sigma_{k,0}.$$

Thus, taking into account (3.10) we derive that $\mathbb{P}$-a.s. for any $t \geq \sigma_{k,0}$,

$$|B_{*,k}(t)| = |W_k(\sigma_{k,\infty}) - W_k(\sigma_{k,\infty} - \int_t^\infty g_k^2(s)ds)| \leq 2\bigg(2\int_t^\infty g_k^2(s)ds \log\Big(\int_t^\infty g_k^2(s)ds\Big)^{-1}\bigg)^{\frac{1}{2}}. \tag{3.11}$$

Therefore, setting

$$\sigma_2 = \max\{\sigma_*, \sigma_{k,0},\ 1 \leq k \leq N\} \tag{3.12}$$

and using (1.5) in $(A1)$ we obtain (3.6) and finish the proof. $\square$

3.2. **Control of modulation equations.** We have the following control of modulation equations.

Proposition 3.2 (Control of modulation equations)**.** *Let $|\boldsymbol{b}|$ be sufficiently small, T large enough, and T^* close to T such that Proposition 2.4,* (3.4) *and* (3.5) *hold. Then, we have*

$$\sum_{k=1}^K (|\dot{\alpha}_k(t)| + |\dot{\theta}_k(t)|) \leq C \left(\|\varepsilon(t)\|_{H^1} + B_*(t)\phi(\delta_1 t) + e^{-\delta_2 t} \right), \quad \forall t \in [T^*, T], \tag{3.13}$$

where ϕ is the spatial decay function of noise in (1.6)*, and C, δ_1, δ_2 are deterministic constants, depending on w_k, v_k, α_k^0 and δ_0.*

Remark 3.3. By using random NLS equation (2.35) and the geometrical decomposition (2.40), we obtain the equation of the remainder ε in (3.22) below. Then, taking the inner product of equation (3.22) with two unstable directions, using the orthogonality conditions (2.42) and applying Lemma 2.8 one can get the control of modulation equations $|\dot{\alpha}_k(t)|$ and $|\dot{\theta}_k(t)|$. For more details, we refer to analogous arguments in the proof of Proposition 3.3 in [42].

It is worth noting that, in the derivation of (3.13), the condition $\varepsilon(T) = 0$ was used in the mass-(sub)critical case [42] to obtain a-priori control of ε on $[T^*, T]$. But in the mass-supercritical case here, we do not have this condition. Instead, the a-priori control of ε on $[T^*, T]$ is derived from (2.43) and the continuity of ε by taking $|\boldsymbol{b}|$ small enough.

For the extra unstable directions $\{a_k^\pm\}$, we have the following estimate.

Proposition 3.4 (Control of extra unstable directions)**.** *Assume the conditions of Proposition 3.2 to hold. Then, for every $1 \leq k \leq K$, we have*

$$\left|\dot{a}_k^{\pm}(t) \mp e_0(w_k)^{-2} a_k^{\pm}(t)\right| \leq C\big(\|\varepsilon\|_{H^1}^{p\wedge 2} + B_*(t)\phi(\delta_1 t) + e^{-\delta_2 t}\big), \quad \forall t \in [T^*, T], \tag{3.14}$$

where C, δ_1, δ_2 are universal deterministic constants, depending on w_k, v_k, α_k^0 and δ_0.

Remark 3.5. We note that the exponent $p\wedge 2$ is due to the singularity of the second derivative of supercritical nonlinearity at the origin when $p < 2$.

Proof of Proposition 3.4. We first note from (2.52) that

$$\dot{a}_k^{\pm}(t) = \mathrm{Im} \int \partial_t \bar{\varepsilon}\, \widetilde{Y}_k^{\pm} dx + \mathrm{Im} \int \bar{\varepsilon}\, \partial_t \widetilde{Y}_k^{\pm} dx. \tag{3.15}$$

Using the explicit expression (2.41) and (2.49) we compute

$$\partial_t \widetilde{R}_k(t,x) = i\left(-\frac{|v_k|^2}{4} + (w_k)^{-2} + \dot{\theta}_k(t)\right)\widetilde{R}_k(t,x) - (v_k + \dot{\alpha}_k(t)) \cdot (\nabla Q_{w_k})(y_k(t))e^{i\Phi_k(t,x)}, \tag{3.16}$$

$$\nabla \widetilde{R}_k(t,x) = (\nabla Q_{w_k})(y_k(t))e^{i\Phi_k(t,x)} + \frac{i}{2} v_k \widetilde{R}_k(t,x), \tag{3.17}$$

$$\Delta \widetilde{R}_k(t,x) = \left(\Delta Q_{w_k} + i v_k \cdot (\nabla Q_{w_k}) - \frac{|v_k|^2}{4} Q_{w_k}\right)(y_k(t))e^{i\Phi_k(t,x)}, \tag{3.18}$$

and

$$\partial_t \widetilde{Y}_k^{\pm}(t,x) = i\left(-\frac{|v_k|^2}{4} + (w_k)^{-2} + \dot{\theta}_k(t)\right)\widetilde{Y}_k^{\pm}(t,x) - (v_k + \dot{\alpha}_k(t)) \cdot \left(\nabla Y_{w_k}^{\pm}\right)(y_k(t))e^{i\Phi_k(t,x)}, \tag{3.19}$$

$$\nabla \widetilde{Y}_k^{\pm}(t,x) = \left(\nabla Y_{w_k}^{\pm}\right)(y_k(t))e^{i\Phi_k(t,x)} + \frac{i}{2} v_k \widetilde{Y}_k^{\pm}(t,x), \tag{3.20}$$

$$\Delta \widetilde{Y}_k^{\pm}(t,x) = \left(\Delta Y_{w_k}^{\pm} + i v_k \cdot (\nabla Y_{w_k}^{\pm}) - \frac{|v_k|^2}{4} Y_{w_k}^{\pm}\right)(y_k(t))e^{i\Phi_k(t,x)}, \tag{3.21}$$

where y_k is given by (2.50) and the phase function $\Phi_k(t,x)$ is given by (2.51). Taking into account equation (1.12) of the rescaled ground state we infer that $\widetilde{R}_k(t,x)$ satisfies the equation

$$i\partial_t \widetilde{R}_k + \Delta \widetilde{R}_k + |\widetilde{R}_k|^{p-1}\widetilde{R}_k = -i\dot{\alpha}_k(t) \cdot (\nabla Q_{w_k})(y_k(t))e^{i\Phi_k} - \dot{\theta}_k(t)\widetilde{R}_k.$$

Then, using the rescaled random NLS (2.35) and the geometrical decomposition (2.40) we obtain

$$\begin{aligned}
\partial_t \varepsilon =& i\Delta\varepsilon + \sum_{k=1}^{K} \dot{\alpha}_k \cdot (\nabla Q_{w_k})(y_k(t))e^{i\Phi_k} - i\sum_{k=1}^{K} \dot{\theta}_k \widetilde{R}_k + i|\widetilde{R} + \varepsilon|^{p-1}(\widetilde{R} + \varepsilon) \\
&- i\sum_{k=1}^{K} |\widetilde{R}_k|^{p-1}\widetilde{R}_k + ib_* \cdot (\nabla\widetilde{R} + \nabla\varepsilon) + ic_*(\widetilde{R} + \varepsilon).
\end{aligned} \tag{3.22}$$

Plugging (3.19)-(3.22) into (3.15) we obtain

$$\begin{aligned}
\dot{a}_k^{\pm}(t) =& \mathrm{Re} \int \left((w_k)^{-2} Y_{w_k}^{\pm} - \Delta Y_{w_k}^{\pm}\right)(y_k(t))e^{i\Phi_k(t)}\bar{\varepsilon} dx \\
&+ \left(\mathrm{Re} \int \dot{\theta}_k \widetilde{Y}_k^{\pm} \bar{\varepsilon} dx - \mathrm{Im} \int \dot{\alpha}_k \cdot \left(\nabla Y_{w_k}^{\pm}\right)(y_k(t))e^{i\Phi_k(t)}\bar{\varepsilon} dx\right) \\
&+ \sum_{j=1}^{K} \left(\mathrm{Re} \int \dot{\theta}_j \bar{\widetilde{R}}_j \widetilde{Y}_k^{\pm} dx + \mathrm{Im} \int \dot{\alpha}_j \cdot (\nabla Q_{w_j})(y_j(t))e^{-i\Phi_j(t)}\widetilde{Y}_k^{\pm} dx\right) \\
&- \mathrm{Re} \int \left(|\widetilde{R} + \varepsilon|^{p-1}\big(\bar{\widetilde{R}} + \bar{\varepsilon}\big) - \sum_{j=1}^{K} |\widetilde{R}_j|^{p-1}\bar{\widetilde{R}}_j\right)\widetilde{Y}_k^{\pm} dx \\
&- \mathrm{Re} \int \left(\bar{b}_* \cdot \big(\nabla\bar{\widetilde{R}} + \nabla\bar{\varepsilon}\big) + \bar{c}_*\big(\bar{\widetilde{R}} + \bar{\varepsilon}\big)\right)\widetilde{Y}_k^{\pm} dx
\end{aligned}$$

$$=: \sum_{m=1}^{5} I_m. \tag{3.23}$$

Note that, by identities (2.10) and (2.11),

$$Y^{\pm} - \Delta Y^{\pm} = \mp i e_0 Y^{\pm} + Q^{p-1} Y^{\pm} + (p-1) Q^{p-1} \mathrm{Re} Y^{\pm}. \tag{3.24}$$

It follows that

$$\begin{aligned} I_1 &= \pm\, e_0\, (w_k)^{-2}\, \mathrm{Im} \int \widetilde{Y}_k^{\pm} \bar{\varepsilon} dx + \left(\frac{p+1}{2} \mathrm{Re} \int |\widetilde{R}_k|^{p-1} \widetilde{Y}_k^{\pm} \bar{\varepsilon} dx + \frac{p-1}{2} \mathrm{Re} \int |\widetilde{R}_k|^{p-3} \bar{\widetilde{R}}_k^2 \widetilde{Y}_k^{\pm} \varepsilon dx \right) \\ &=: \pm e_0\, (w_k)^{-2}\, \mathrm{Im} \int \widetilde{Y}_k^{\pm} \bar{\varepsilon} dx + I_6. \end{aligned} \tag{3.25}$$

Hence, plugging (3.25) into (3.23) we come to

$$\dot{a}_k^{\pm}(t) \mp e_0\, (w_k)^{-2}\, a_k^{\pm}(t) = \sum_{m=2}^{6} I_m. \tag{3.26}$$

Next, we estimate each term I_m, $2 \le m \le 6$, separately.

(i) Estimate of I_2. First, by Hölder's inequality, the estimate of modulation equation (3.13) and the fact that $Y^{\pm} \in \mathcal{S}(\mathbb{R}^d)$, we have

$$\begin{aligned} |I_2| \le & C(|\dot{\alpha}_k| + |\dot{\theta}_k|) \|\varepsilon\|_{L^2} \big(\|Y^{\pm}\|_{L^2} + \|\nabla Y^{\pm}\|_{L^2} \big) \\ \le & C(\|\varepsilon\|_{H^1}^2 + B_*(t)\phi(\delta_1 t) + e^{-\delta_2 t}). \end{aligned} \tag{3.27}$$

(ii) Estimate of I_3. Using the identities (2.70) and (2.72), the modulation estimate (3.13) and the decoupling Lemma 2.8 we derive

$$\begin{aligned} |I_3| \le & C|\dot{\theta}_k| \left| \mathrm{Re} \int Q Y^{\pm} dx \right| + C|\dot{\alpha}_k| \left| \mathrm{Im} \int \nabla Q Y^{\pm} dx \right| \\ & + \sum_{j \ne k} \left(|\dot{\theta}_j| \int \left| \bar{\widetilde{R}}_j \widetilde{Y}_k^{\pm} \right| dx + |\dot{\alpha}_j| \int \left| (\nabla Q_{w_j})(y_j(t)) e^{-i\Phi_j(t)} \widetilde{Y}_k^{\pm} \right| dx \right) \\ \le & C(\|\varepsilon\|_{H^1}^2 + B_*(t)\phi(\delta_1 t) + e^{-\delta_2 t}). \end{aligned} \tag{3.28}$$

(iii) Estimate of I_4 and I_6. In order to estimate I_4 and I_6, when $p \ge 2$, from direct computations and Lemma 2.8, one has

$$\begin{aligned} \mathrm{Re} \int |\widetilde{R} + \varepsilon|^{p-1} (\bar{\widetilde{R}} + \bar{\varepsilon}) \widetilde{Y}_k^{\pm} dx = & \mathrm{Re} \int |\widetilde{R}_k + \varepsilon|^{p-1} (\bar{\widetilde{R}}_k + \bar{\varepsilon}) \widetilde{Y}_k^{\pm} dx + \mathcal{O}\big(e^{-\delta_2 t}\big) \\ = & \mathrm{Re} \int |\widetilde{R}_k|^{p-1} \bar{\widetilde{R}}_k \widetilde{Y}_k^{\pm} dx + \frac{p+1}{2} \mathrm{Re} \int |\widetilde{R}_k|^{p-1} \widetilde{Y}_k^{\pm} \bar{\varepsilon} dx \\ & + \frac{p-1}{2} \mathrm{Re} \int |\widetilde{R}_k|^{p-3} \bar{\widetilde{R}}_k^2 \widetilde{Y}_k^{\pm} \varepsilon dx + \mathcal{O}\left(\|\varepsilon\|_{H^1}^2 + e^{-\delta_2 t} \right). \end{aligned} \tag{3.29}$$

Plugging (3.29) into I_4 we derive that for $p \ge 2$,

$$|I_4 + I_6| \le C(\|\varepsilon\|_{H^1}^2 + e^{-\delta_2 t}). \tag{3.30}$$

When $p < 2$, using Lemma 2.8 one can decouple different soliton profiles and eigenfucntions to get

$$\begin{aligned} I_4 = & - \mathrm{Re} \int \left(|\widetilde{R}_k + \varepsilon|^{p-1} (\bar{\widetilde{R}}_k + \bar{\varepsilon}) - |\widetilde{R}_k|^{p-1} \bar{\widetilde{R}}_k \right) \widetilde{Y}_k^{\pm} dx + \mathcal{O}\big(e^{-\delta_2 t}\big) \\ = & - \mathrm{Re} \left(\int_{|\widetilde{R}_k| > 2|\varepsilon|} + \int_{|\widetilde{R}_k| \le 2|\varepsilon|} \right) |\widetilde{R}_k + \varepsilon|^{p-1} (\bar{\widetilde{R}}_k + \bar{\varepsilon}) \widetilde{Y}_k^{\pm} - |\widetilde{R}_k|^{p-1} \bar{\widetilde{R}}_k \widetilde{Y}_k^{\pm} dx + \mathcal{O}\big(e^{-\delta_2 t}\big). \end{aligned} \tag{3.31}$$

Note that, since $p \in (1 + \frac{4}{d}, 1 + \frac{4}{d-2})$, we may take $\rho(> 1)$ close to 1 such that $2 \le \rho p \le \frac{2d}{d-2}$ if $d \ge 3$, $2 \le \rho p < +\infty$ if $d = 1, 2$. Then, using the Sobolev embedding $H^1(\mathbb{R}^d) \hookrightarrow L^{\rho p}(\mathbb{R}^d)$ and the Hölder inequality we estimate

$$\left| \mathrm{Re} \int_{|\widetilde{R}_k| \le 2|\varepsilon|} |\widetilde{R}_k + \varepsilon|^{p-1} (\bar{\widetilde{R}}_k + \bar{\varepsilon}) \widetilde{Y}_k^{\pm} - |\widetilde{R}_k|^{p-1} \bar{\widetilde{R}}_k \widetilde{Y}_k^{\pm} dx \right|$$

$$\leq C\int |\varepsilon|^p|\widetilde{Y}_k^{\pm}|dx \leq C\|\varepsilon\|^p_{L^{\rho p}}\|\widetilde{Y}_k^{\pm}\|_{L^{\rho'}} \leq C\|\varepsilon\|^p_{H^1}. \tag{3.32}$$

Moreover, one has the expansion

$$|\widetilde{R}_k+\varepsilon|^{p-1}(\bar{\widetilde{R}}_k+\bar{\varepsilon}) = |\widetilde{R}_k|^{p-1}\bar{\widetilde{R}}_k + \frac{p-1}{2}|\widetilde{R}_k|^{p-3}\bar{\widetilde{R}}_k^2\varepsilon + \frac{p+1}{2}|\widetilde{R}_k|^{p-1}\bar{\varepsilon} + Er \tag{3.33}$$

with the error term

$$Er = \int_0^1 (1-s)\Big[\frac{\partial^2 f}{\partial z^2}(\widetilde{R}_k+s\varepsilon)\varepsilon^2 + 2\frac{\partial^2 f}{\partial z\bar{z}}(\widetilde{R}_k+s\varepsilon)|\varepsilon|^2 + \frac{\partial^2 f}{\partial \bar{z}^2}(\widetilde{R}_k+s\varepsilon)\bar{\varepsilon}^2\Big]ds,$$

where f is defined by $f(z) := |z|^{p-1}\bar{z}$, $z\in\mathbb{C}$. When $p<2$ and $x\in\{x : |\widetilde{R}_k| > 2|\varepsilon|\}$, we have

$$|Er| \leq C\int_0^1 |\widetilde{R}_k+s\varepsilon|^{p-2}|\varepsilon|^2 ds \leq C|\varepsilon|^p,$$

and so, as in (3.32),

$$\Big|\mathrm{Re}\int_{|\widetilde{R}_k|>2|\varepsilon|}\widetilde{Y}_k^{\pm}Erdx\Big| \leq C\int |\varepsilon|^p|\widetilde{Y}_k^{\pm}|dx \leq C\|\varepsilon\|^p_{H^1}. \tag{3.34}$$

Thus, plugging (3.32)-(3.34) into (3.31) we obtain

$$I_4 = -\,\mathrm{Re}\int_{|\widetilde{R}_k|>2|\varepsilon|}\frac{p+1}{2}|\widetilde{R}_k|^{p-1}\widetilde{Y}_k^{\pm}\bar{\varepsilon} + \frac{p-1}{2}|\widetilde{R}_k|^{p-3}\bar{\widetilde{R}}_k^2\widetilde{Y}_k^{\pm}\varepsilon dx + \mathcal{O}(\|\varepsilon\|^p_{H^1}+e^{-\delta_2 t}). \tag{3.35}$$

Similarly, as in the proof of (3.32), one has

$$\begin{aligned} I_6 =&\mathrm{Re}\int \frac{p+1}{2}|\widetilde{R}_k|^{p-1}\widetilde{Y}_k^{\pm}\bar{\varepsilon}dx + \mathrm{Re}\int \frac{p-1}{2}|\widetilde{R}_k|^{p-3}\bar{\widetilde{R}}_k^2\widetilde{Y}_k^{\pm}\varepsilon dx \\ =&\mathrm{Re}\int_{|\widetilde{R}_k|>2|\varepsilon|}\frac{p+1}{2}|\widetilde{R}_k|^{p-1}\widetilde{Y}_k^{\pm}\bar{\varepsilon}dx + \mathrm{Re}\int_{|\widetilde{R}_k|>2|\varepsilon|}\frac{p-1}{2}|\widetilde{R}_k|^{p-3}\bar{\widetilde{R}}_k^2\widetilde{Y}_k^{\pm}\varepsilon dx + \mathcal{O}(\|\varepsilon\|^p_{H^1}). \end{aligned} \tag{3.36}$$

Combining (3.35) and (3.36) together, we thus obtain that when $p<2$,

$$|I_4+I_6| \leq C(\|\varepsilon\|^p_{H^1}+e^{-\delta_2 t}). \tag{3.37}$$

Therefore, we conclude from (3.30) and (3.37) that for $p\in(1+\frac{4}{d}, 1+\frac{4}{(d-2)_+})$,

$$|I_4+I_6| \leq C(\|\varepsilon\|^{p\wedge 2}_{H^1}+e^{-\delta_2 t}). \tag{3.38}$$

(iv) Estimate of I_5. It remains to treat the I_5 term involving the random coefficients. By Hölder's inequality, expressions (1.18) and (1.19), and the change of variables, I_5 can be bounded by

$$\begin{aligned} |I_5| \leq& CB_*(t)\sum_{l=1}^N\int |\nabla\phi_l(y+v_kt+\alpha_k(t))|(|\nabla Q_{w_k}(y)|+|Q_{w_k}(y)|)dy \\ &+ CB_*(t)\sum_{l=1}^N\int \big(|\nabla\phi_l(y+v_kt+\alpha_k(t))|^2+|\Delta\phi_l(y+v_kt+\alpha_k(t))|\big)|Q_{w_k}(y)|dy \\ &+ CB_*(t)\|\nabla\varepsilon\|_{L^2}\left(\sum_{l=1}^N\int |\nabla\phi_l(y+v_kt+\alpha_k(t))|^2|Y^{\pm}(w_k^{-1}y)|^2dy\right)^{\frac{1}{2}} \\ &+ CB_*(t)\|\varepsilon\|_{L^2}\left(\sum_{l=1}^N\int \big(|\nabla\phi_l(y+v_kt+\alpha_k(t))|^2+|\Delta\phi_l(y+v_kt+\alpha_k(t))|\big)^2|Y^{\pm}(w_k^{-1}y)|^2dy\right)^{\frac{1}{2}} + Ce^{-\delta_2 t}. \end{aligned} \tag{3.39}$$

We note that the spatial functions of the noise travel with the speeds $\{v_k\}$. Hence, intuitively, after a large time, they shall be separated sufficiently far away from the ground state.

In order to capture this fact, we split the integration into two regimes $|y|\leq |v_k|t/2$ and $|y|>|v_k|t/2$. The key observation is that, for $|y|\leq|v_k|t/2$, where t is large enough such that $t\geq 8|\alpha_k^0|/|v_k|$, by (3.5), $|y+v_kt+\alpha_k(t)|\geq|v_k|t-|y|-|\alpha_k(t)|\geq|v_k|t/4$. Thus, in view of Assumption (A$_1$), the integration in this regime can be controlled by the spatial decay rate of $\nabla\phi_l$ in the noise. Moreover, in the outer large regime

$|y| > |v_k|t/2$, the integration can be bounded by the exponential decay of the ground state Q. As a result, the first and second integrations on the right-hand side of (3.39) above can be bounded by

$$
\begin{aligned}
& CB_*(t) \sum_{l=1}^{N} \Bigg(\int_{|y|\leq \frac{|v_k|}{2}t} |\nabla \phi_l(y+v_kt+\alpha_k)|(|\nabla Q_{w_k}(y)| + |Q_{w_k}(y)|) + |\nabla \phi_l(y+v_kt+\alpha_k)|^2 |Q_{w_k}(y)| \\
& \qquad\qquad + |\Delta \phi_l(y+v_kt+\alpha_k)||Q_{w_k}(y)|dy + \int_{|y|\geq \frac{|v_k|}{2}t} |Q_{w_k}(y)| + |\nabla Q_{w_k}(y)|dy \Bigg) \\
\leq & CB_*(t)\phi(\frac{|v_k|}{4}t) \left(\int |Q_{w_k}(y)| + |\nabla Q_{w_k}(y)|dy + \int_{|y|\geq \frac{|v_k|}{2}t} e^{-\frac{2\delta_0}{w_k}|y|}dy \right) \\
\leq & CB_*(t)(\phi(\delta_1 t) + e^{-\delta_2 t})
\end{aligned} \tag{3.40}
$$

for some positive deterministic constants C, δ_1 and δ_2, depending on w_k, α_k^0, v_k and δ_0 from Lemma 2.8.

Similarly, in view of the exponential decay of $Y^{\pm}$ in Lemma 2.2 and the spatial decay of noise in Assumption (A_1), the remaining two terms on the right-hand side of (3.39) also can be bounded by

$$C(\|\varepsilon\|_{H^1}^2 + B_*(t)\phi(\delta_1 t) + e^{-\delta_2 t}). \tag{3.41}$$

Finally, combining (3.26)-(3.28), (3.38), (3.40) and (3.41) altogether we obtain (3.14) and finish the proof. □

3.3. **Control of the remainder.** In this subsection, we control the remainder ε in the geometrical decomposition. The key role is played by the Lyapunov functional $\mathcal{G}$ defined in (3.45) below.

Because the velocities $\{v_k\}$ of soliton profiles are different, without loss of generality, we may assume that $v_{1,1} < v_{2,1} < \cdots < v_{K,1}$. Set $A_0 := \frac{1}{4}\min_{2\leq k\leq K}\{v_{k,1} - v_{k-1,1}\}$, $A_k := \frac{1}{2}(v_{k-1,1} + v_{k,1})$, $2 \leq k \leq K$. Let $\Psi(x)$ be a smooth non-decreasing function on $\mathbb{R}$ such that $0 \leq \Psi \leq 1$, $\Psi(x) = 0$ for $x \leq -A_0$, $\Psi(x) = 1$ for $x \geq A_0$, and for some $C > 0$,

$$(\Psi'(x))^2 \leq C\Psi(x), \quad (\Psi''(x))^2 \leq C\Psi'(x). \tag{3.42}$$

Define the localization functions by

$$
\begin{aligned}
&\varphi_1(t,x) = 1 - \Psi\left(\frac{x_1 - A_2 t}{t}\right), \quad \varphi_K(t,x) = \Psi\left(\frac{x_1 - A_K t}{t}\right), \\
&\varphi_k(t,x) = \Psi\left(\frac{x_1 - A_k t}{t}\right) - \Psi\left(\frac{x_1 - A_{k+1} t}{t}\right), \quad 2 \leq k \leq K-1
\end{aligned} \tag{3.43}
$$

for any $x = (x_1, x_2, \cdots, x_d) \in \mathbb{R}^d$. Note that $\sum_{k=1}^K \varphi_k(t,x) = 1$, and

$$|\partial_x \varphi_k(t,x)| + |\partial_x^3 \varphi_k(t,x)| + |\partial_t \varphi_k(t,x)| \leq \frac{C}{t}. \tag{3.44}$$

Proposition 3.6 below provides the main control of the remainder in the geometrical decomposition.

Proposition 3.6 (Coercivity type control of remainder)**.** *There exist a deterministic constant $r(>0)$ and a positive random variable T, such that for any $\boldsymbol{a}^- \in B_{\mathbb{R}^K}(r)$ and T^* close to T, there exist universal deterministic constants C, δ_1, δ_2, depending on w_k, v_k, α_k^0 and δ_0,*

$$
\begin{aligned}
\|\varepsilon(t)\|_{H^1}^2 \leq & C\int_t^\infty \Big(\frac{1}{s} + B_*(s)\Big)\|\varepsilon(s)\|_{H^1}^2 ds + C\Big(\int_t^\infty \|\varepsilon(s)\|_{H^1}^{p\wedge 2} ds\Big)^2 + C\int_t^\infty B_*(s)\phi(\delta_1 s)ds \\
& + C\Big(\int_t^\infty B_*(s)\phi(\delta_1 s)ds\Big)^2 + C\left(|\mathbf{a}^-(t)|^2 + |\boldsymbol{a}^-|^2 + e^{-\delta_2 t}\right) + \beta\|\varepsilon(t)\|_{H^1}^2, \ \forall t \in [T^*, T],
\end{aligned} \tag{3.45}
$$

where $\beta \to 0$ as $\|\varepsilon(t)\|_{H^1} \to 0$.

The key role in the proof of Proposition 3.6 is played by the following Lyapunov type functional

$$\mathcal{G}(t) := \|\nabla u\|_{L^2}^2 - \frac{2}{p+1}\|u\|_{L^{p+1}}^{p+1} + \sum_{k=1}^K \Bigg\{ \left((w_k)^{-2} + \frac{|v_k|^2}{4}\right) \int |u(t,x)|^2 \varphi_k(t,x)dx$$

$$- v_k \cdot \operatorname{Im} \int \nabla u(t,x) \bar{u}(t,x) \varphi_k(t,x) dx \Big\}. \tag{3.46}$$

where u is the solution to the rescaled random NLS (2.35). As in the proof of Propositions 4.1, 4.3 and 4.4 in [42], we have the following control of the Lyapunov functional.

Lemma 3.7 (Control of Lyapunov functional)**.** *For any $t \in [T^*, T]$ one has*

$$\left| \frac{d}{dt} \mathcal{G}(t) \right| \leq \frac{C}{t} \left(\|\varepsilon(t)\|_{H^1}^2 + e^{-\delta_2 t} \right) + C B_*(t) \left(\|\varepsilon(t)\|_{H^1}^2 + \phi(\delta_1 t) + e^{-\delta_2 t} \right), \tag{3.47}$$

where C, δ_1, δ_2 are deterministic constants, depending on w_k, v_k, α_k^0 and δ_0. Moreover, the following expansion holds:

$$\begin{aligned} \mathcal{G}(t) = &\sum_{k=1}^{K} \left(\|\nabla Q_{w_k}\|_{L^2}^2 - \frac{2}{p+1} \|Q_{w_k}\|_{L^{p+1}}^{p+1} + (w_k)^{-2} \|Q_{w_k}\|_{L^2}^2 \right) + H(\varepsilon(t)) \\ &+ \mathcal{O}(e^{-\delta_2 t}) + \beta \|\varepsilon(t)\|_{H^1}^2, \end{aligned} \tag{3.48}$$

where $\beta \to 0$ as $\|\varepsilon(t)\|_{H^1} \to 0$, and $H(\varepsilon)$ contains the quadratic terms of the remainder ε, i.e.,

$$\begin{aligned} H(\varepsilon) = &\int |\nabla \varepsilon|^2 dx - \sum_{k=1}^{K} \int |\widetilde{R}_k|^{p-1} |\varepsilon|^2 + (p-1) |\widetilde{R}_k|^{p-3} \left(\operatorname{Re} \widetilde{R}_k \bar{\varepsilon} \right)^2 dx \\ &+ \sum_{k=1}^{K} \left\{ \left((w_k)^{-2} + \frac{|v_k|^2}{4} \right) \int |\varepsilon|^2 \varphi_k dx - v_k \cdot \operatorname{Im} \int \nabla \varepsilon \bar{\varepsilon} \varphi_k dx \right\}. \end{aligned} \tag{3.49}$$

Remark 3.8. The quadratic term $H(\varepsilon(t))$ has the crucial coercivity type estimate

$$\|\varepsilon(t)\|_{H^1}^2 \leq C H(\varepsilon(t)) + C \big(|\mathbf{a}^+(t)|^2 + |\mathbf{a}^-(t)|^2 \big), \ t \in [T^*, T] \tag{3.50}$$

for some $C > 0$. This can be proved by using the coercivity of the linearized operator in (2.6), the orthogonal conditions in (2.42), and analogous arguments as in the proof of the 1D case in [36, Appendix B].

We are now ready to prove Proposition 3.6.

Proof of Proposition 3.6. In view of Proposition 2.7, we can take a deterministic small constant $r (> 0)$ such that for any $\boldsymbol{a}^- \in B_{\mathbb{R}^K}(r)$, there exists a unique $\boldsymbol{b} \in \mathbb{R}^{2K}$ such that

$$\mathbf{a}^+(T) = \mathbf{0}, \quad \mathbf{a}^-(T) = \boldsymbol{a}^-, \quad \text{and} \quad |\boldsymbol{b}| \leq C |\boldsymbol{a}^-| \leq C r. \tag{3.51}$$

Then, we take r possibly smaller, a random time T large enough and T^* close to T, such that the geometrical decomposition in Proposition 2.4, (3.4) and (3.5) hold.

Using the coercivity estimate (3.50), for any $t \in [T^*, T]$, one has

$$\|\varepsilon(t)\|_{H^1}^2 \leq C |H(\varepsilon(T))| + C |H(\varepsilon(t)) - H(\varepsilon(T))| + C \big(|\mathbf{a}^+(t)|^2 + |\mathbf{a}^-(t)|^2 \big). \tag{3.52}$$

We shall estimate the right-hand side of (3.52). By Proposition 3.4,

$$\left| \dot{a}_k^+(t) - e_0 (w_k)^{-2} a_k^+(t) \right| \leq C (\|\varepsilon(t)\|_{H^1}^{p \wedge 2} + B_*(t) \phi(\delta_1 t) + e^{-\delta_2 t}). \tag{3.53}$$

Then, it follows from (3.53), $\mathbf{a}^+(T) = \mathbf{0}$ and Gronwall's inequality that for any $t \in [T^*, T]$,

$$|a_k^+(t)| \leq C \int_t^\infty \Big(\|\varepsilon(s)\|_{H^1}^{p \wedge 2} + B_*(s) \phi(\delta_1 s) \Big) ds + C e^{-\delta_2 t}. \tag{3.54}$$

Moreover, by (2.43) and (3.51),

$$|H(\varepsilon(T))| \leq C \|\varepsilon(T)\|_{H^1}^2 \leq C |\boldsymbol{b}|^2 \leq C |\boldsymbol{a}^-|^2. \tag{3.55}$$

Regarding the $|H(\varepsilon(t)) - H(\varepsilon(T))|$ term, by the expansion (3.48), for any $t \in [T^*, T]$,

$$|H(\varepsilon(t)) - H(\varepsilon(T))| \leq C e^{-\delta_2 t} + \beta \|\varepsilon(t)\|_{H^1}^2 + |\mathcal{G}(t) - \mathcal{G}(T)|, \tag{3.56}$$

where $\beta \to 0$ as $\|\varepsilon\|_{H^1}^2 \to 0$. Integrating both sides of (3.47) on $[t, T]$ we have

$$|\mathcal{G}(t) - \mathcal{G}(T)| \leq C \int_t^T \frac{1}{s} \left(\|\varepsilon(s)\|_{H^1}^2 + e^{-\delta_2 s} \right) ds + C \int_t^T B_*(s) \left(\|\varepsilon(s)\|_{H^1}^2 + \phi(\delta_1 s) + e^{-\delta_2 s} \right) ds. \tag{3.57}$$

Hence, combining (3.56) and (3.57) together and using (3.3) we derive that

$$|H(\varepsilon(t))-H(\varepsilon(T))| \leq C \int_t^{\infty} \Big(\frac{1}{s}+B_*(s)\Big)\|\varepsilon(s)\|_{H^1}^2 ds + C\int_t^{\infty} B_*(s)\phi(\delta_1 s)ds + Ce^{-\delta_2 t} + \beta\|\varepsilon(t)\|_{H^1}^2. \quad (3.58)$$

Therefore, plugging (3.54), (3.55) and (3.58) into (3.52) we obtain (3.45). □

4. Uniform estimates of approximating solutions

For every $n \in \mathbb{N}$, we consider the approximating equation

$$\begin{cases} i\partial_t u_n + (\Delta + b_* \cdot \nabla + c_*)u_n + |u_n|^{p-1}u_n = 0, \\ u_n(n) = R(n) + i\sum\limits_{k,\pm} b_{k,n}^{\pm} Y_k^{\pm}(n). \end{cases} \quad (4.1)$$

Let δ_1, δ_2 be as in Proposition 3.2, and

$$\delta_3 := \frac{1}{2}\min_{1\leq k\leq K}\{e_0(w_k)^{-2}\}, \quad (4.2)$$

with $e_0(>0)$ being the eigenvalue of the linearized Schrödinger operator in (2.5), and w_k the frequency of the soliton (1.13). Set

$$\widetilde{\delta} := \begin{cases} \frac{1}{2}(\delta_1 \wedge \delta_2 \wedge \delta_3), & \text{in Case (I);} \\ \delta_1, & \text{in Case (II).} \end{cases} \quad (4.3)$$

The main result of this section is the following crucial uniform estimate, which shows that the approximating equation (4.1) can be solved backward up to a universal time T_0, uniformly in n.

Theorem 4.1 (Uniform estimate)**.** *Let $\widetilde{\delta}$ be as in* (4.3) *and ν_0 as in* (4.25) *below. Assume* (A_0) *and* (A_1) *with $\nu_* \geq \nu_0$ in Case (II). Then, there exist a positive random time T_0 and a deterministic constant $C>0$, such that for $\mathbb{P}$-a.e. $\omega \in \Omega$ there exists $N_0(\omega) \geq 1$ such that for any $n \geq N_0(\omega)$, $T_0(\omega) < n$ and the following holds:*

There exists $\boldsymbol{b}_n(\omega) \in B_{\mathbb{R}^{2K}}(C\phi^{\frac{1}{2}+\frac{1}{4d}}(\widetilde{\delta}n))$ such that the solution $u_n(\omega)$ to equation (4.1) admits the geometrical decomposition (2.40) *on $[T_0(\omega), n]$ and satisfies*

$$\|u_n(t,\omega) - R(t,\omega)\|_{H^1} \leq Ct\phi^{\frac{1}{2}}(\widetilde{\delta}t), \quad \forall t \in [T_0(\omega), n], \quad (4.4)$$

where ϕ is the spatial decay function of noise given by (1.6).

Remark 4.2. We note that the temporal convergence rate of the approximating solution u_n, as well as the smallness of the modulated parameter $\mathbf{b}$, are dictated by the spatial decay function of the noise.

The proof of Theorem 4.1 mainly proceeds in two steps. First in Subsection 4.1, we prove the bootstrap estimates of the reminder ε_n, the modulation parameters (α_n, θ_n), and the unstable direction $\mathbf{a}_n^+$ under an a-priori control of $\mathbf{a}_n^-$. Then, in Subsection 4.2, we control the remaining parameter $\mathbf{a}_n^-$ by using topological arguments.

Let us mention that, in the sequel, we mainly consider $\boldsymbol{a}_n^- \in B_{\mathbb{R}^K}(\phi^{\frac{1}{2}+\frac{1}{4d}}(\widetilde{\delta}n))$ with n large enough such that $\phi^{\frac{1}{2}+\frac{1}{4d}}(\widetilde{\delta}n) \leq r$ and $\boldsymbol{b} \in B_{\mathbb{R}^{2K}}(\eta)$, where r, η are small deterministic constants from Propositions 2.7 and 2.4, respectively, so that the geometrical decomposition and the final condition in two propositions hold. In particular, for any $\boldsymbol{a}_n^- \in B_{\mathbb{R}^K}(\phi^{\frac{1}{2}+\frac{1}{4d}}(\widetilde{\delta}n)))$, there exists a unique small vector $\boldsymbol{b}_n \in \mathbb{R}^{2K}$ such that

$$\mathbf{a}_n^+(n) = \mathbf{0}, \quad \mathbf{a}_n^-(n) = \boldsymbol{a}_n^-, \quad (4.5)$$

$$|\boldsymbol{b}_n| \leq C|\boldsymbol{a}_n^-| \leq C\phi^{\frac{1}{2}+\frac{1}{4d}}(\widetilde{\delta}n) \leq C\phi^{\frac{1}{2}+\frac{1}{4d}}(\widetilde{\delta}t), \ \forall t \leq n, \quad (4.6)$$

where $C>0$ is a deterministic positive constant independent of n.

4.1. **Bootstrap estimates.** The main bootstrap estimates of this subsection are stated as follows.

Proposition 4.3 (Bootstrap estimates of ε_n, α_n, θ_n and $\mathbf{a}_n^+$)**.** *Assume the conditions of Theorem 4.1 to hold. Then, there exists a random time $\tau^*(>0)$, such that $\mathbb{P}$-a.e. $\omega\in\Omega$ there exists $N_0(\omega)\geq 1$ such that for any $n\geq N_0(\omega)$, $\tau^*(\omega)<n$ and the following holds:*

Let $t^*(\omega)\in(\tau^*(\omega),n]$ be such that for any $t\in[t^*,n]$, $u_n(\omega)$ satisfies

$$\|u_n(t,\omega)-R(t,\omega)\|_{H^1}\leq\frac{1}{2}\delta_* \tag{4.7}$$

with δ_* being the small number from Lemma 2.6, and the following estimates hold:

$$\|\varepsilon_n(t,\omega)\|_{H^1}\leq\phi^{\frac12}(\widetilde{\delta}t),\ |\mathbf{a}_n^+(t,\omega)|\leq\phi^{\frac12}(\widetilde{\delta}t),\ |\mathbf{a}_n^-(t,\omega)|\leq\phi^{\frac12+\frac{1}{4d}}(\widetilde{\delta}t), \tag{4.8}$$

$$\sum_{k=1}^K(|\alpha_{n,k}(t,\omega)-\alpha_k^0|+|\theta_{n,k}(t,\omega)-\theta_k^0|)\leq t\phi^{\frac12}(\widetilde{\delta}t). \tag{4.9}$$

Then, there exists a smaller time $t_*\in(\tau^*,t^*)$, such that u_n admits the geometrical decomposition (2.40) *on the larger time interval $[t_*,n]$, and the improved estimates hold:*

$$\|u_n(t,\omega)-R(t,\omega)\|_{H^1}\leq\frac{1}{4}\delta_*, \tag{4.10}$$

$$\|\varepsilon_n(t,\omega)\|_{H^1}\leq\frac12\phi^{\frac12}(\widetilde{\delta}t),\ |\mathbf{a}_n^+(t,\omega)|\leq\frac12\phi^{\frac12}(\widetilde{\delta}t), \tag{4.11}$$

$$\sum_{k=1}^K(|\alpha_{n,k}(t,\omega)-\alpha_k^0|+|\theta_{n,k}(t,\omega)-\theta_k^0|)\leq\frac12 t\phi^{\frac12}(\widetilde{\delta}t),\quad\forall t\in[t_*,n]. \tag{4.12}$$

Remark 4.4. We remark that the constants in estimates (4.10)-(4.12) are smaller than those in (4.7)-(4.9). This is because in the following analysis one can gain small factors $o(1)$ before the deterministic constants C. The small factors are contributed by the exponential and polynomial decay of time, the tail of noise $B_*(t)$ and ν_*^{-1} with ν_* large enough.

We also note that the bootstrap estimates above require a-priori control of $\mathbf{a}_n^-(t)$, which cannot be improved in estimates (4.10)-(4.12). Later in Proposition 4.5, we shall use topological arguments to choose a suitable final data $\boldsymbol{a}_n^-$ to obtain the required a-priori control.

Proof of Proposition 4.3. We define the random time τ^* by

$$\tau^*:=\begin{cases}\max\{M,\sigma_1,\tau_j,j=1,2,3\}, & \text{in Case (I);}\\ \max\{M,\sigma_1,\sigma_2,\tau_j,j=1,4,5\}, & \text{in Case (II),}\end{cases} \tag{4.13}$$

and let $N_0=[\tau^*]+1$, where M is the deterministic large time from the geometrical decomposition in Proposition 2.4, σ_1 and σ_2 are the random times in Subsection 3.1, and the random times τ_j, $1\leq j\leq 5$, will be determined below. To ease notations, we omit the dependence of ω in the sequel.

Note that, under the above bootstrap estimates, the estimates in Section 3 are all valid after τ^* and the corresponding constants are deterministic. To be precise, in view of (4.8), (4.9), and the continuity of solutions in H^1, one can take $t_*\in(\tau^*,t^*)$ close to t^*, such that the geometrical decomposition (2.40) and the following estimates hold on $[t_*,n]$:

$$\|\varepsilon_n(t)\|_{H^1}\leq 2\phi^{\frac12}(\widetilde{\delta}t),\quad|\mathbf{a}_n^+(t)|\leq 2\phi^{\frac12}(\widetilde{\delta}t),\quad|\mathbf{a}_n^-(t)|\leq 2\phi^{\frac12+\frac{1}{4d}}(\widetilde{\delta}t) \tag{4.14}$$

$$\sum_{k=1}^K(|\alpha_{n,k}(t)-\alpha_k^0|+|\theta_{n,k}(t)-\theta_k^0|)\leq 2t\phi^{\frac12}(\widetilde{\delta}t). \tag{4.15}$$

Then, let

$$\tau_1:=\inf\Big\{t>0:2\phi^{\frac12}(\widetilde{\delta}t)\leq 1,\ B_*(t)+2t\phi^{\frac12}(\widetilde{\delta}t)\leq\frac{1}{10}\min\{1,w_k,\alpha_k^0\}\Big\}. \tag{4.16}$$

By the definition (4.13), $t_*>\tau^*\geq\tau_1$. We thus infer from estimates (4.14) and (4.15) that the upper bounds in (3.4) and (3.5) are valid on $[t_*,n]$. One also can apply Proposition 3.2 to obtain

$$\sum_{k=1}^K(|\dot{\alpha}_{n,k}(t)|+|\dot{\theta}_{n,k}(t)|)\leq C\left(\|\varepsilon_n(t)\|_{H^1}+B_*(t)\phi(\delta_1 t)+e^{-\delta_2 t}\right),\quad\forall t\in[t_*,n]. \tag{4.17}$$

Below we consider Case (I) and Case (II) separately to derive the bootstrap estimates (4.10)-(4.12).

Case (I): First, by (4.14) and (4.15), for any $t \in [t_*, n]$,

$$\begin{aligned}\|u_n(t)-R(t)\|_{H^1} \leq& \|R(t)-\sum_{k=1}^K \widetilde{R}_{n,k}(t)\|_{H^1}+\|\varepsilon_n\|_{H^1}\\ \leq& C\sum_{k=1}^K(|\alpha_{n,k}(t)-\alpha_k^0|+|\theta_{n,k}(t)-\theta_k^0|)+\|\varepsilon_n\|_{H^1}\\ \leq& 2Cte^{-\frac{1}{2}\widetilde{\delta}t}+2e^{-\frac{1}{2}\widetilde{\delta}t},\end{aligned} \tag{4.18}$$

where $\widetilde{R}_{n,k}$ is the approximating soliton profile from the geometrical decomposition (2.41) with $\alpha_k(t)$, $\theta_k(t)$ replaced by $\alpha_{n,k}(t)$, $\theta_{n,k}(t)$. Setting

$$\tau_2 := \inf\Big\{t \geq \frac{2}{\widetilde{\delta}}:\ 2Cte^{-\frac{1}{2}\widetilde{\delta}t}+2e^{-\frac{1}{2}\widetilde{\delta}t} \leq \frac{1}{4}\delta_*\Big\}, \tag{4.19}$$

where δ_* the small number from Lemma 2.6. Then, since $\tau^* \geq \tau_1 \vee \tau_2$, estimate (4.10) is verified on $[t_*, n]$.

Regarding the remaining estimates (4.11) and (4.12), using (3.54), (4.14) and the inequality

$$p\wedge 2 > 1+\frac{1}{2d},\ \forall d \geq 1, \tag{4.20}$$

we have that for every $1 \leq k \leq K$ and any $t \in [t_*, n]$,

$$\begin{aligned}|a^+_{n,k}(t)| \leq& C\int_t^{+\infty} e^{-\frac{p\wedge 2}{2}\widetilde{\delta}s}+B_*(s)e^{-\delta_1 s}ds+Ce^{-\delta_2 t}\\ \leq& C\Big(\frac{2}{(p\wedge 2)\widetilde{\delta}}e^{\frac{1-p\wedge 2}{2}\widetilde{\delta}t}+\frac{1}{\widetilde{\delta}}e^{-\frac{1}{2}\widetilde{\delta}t}+e^{-\frac{1}{2}\widetilde{\delta}t}\Big)e^{-\frac{1}{2}\widetilde{\delta}t}\\ \leq& C\Big(\frac{3}{\widetilde{\delta}}e^{-\frac{1}{4d}\widetilde{\delta}t}+e^{-\frac{1}{4d}\widetilde{\delta}t}\Big)e^{-\frac{1}{2}\widetilde{\delta}t}.\end{aligned} \tag{4.21}$$

Moreover, for any $t \in [t_*, n]$, integrating (4.17) on $[t, n]$ and using (2.43), (3.3) (4.6), (4.14) we derive

$$\begin{aligned}&\sum_{k=1}^K(|\alpha_{n,k}(t)-\alpha_k^0|+|\theta_{n,k}(t)-\theta_k^0|)\\ \leq& \sum_{k=1}^K(|\alpha_{n,k}(t)-\alpha_{n,k}(n)|+|\alpha_{n,k}(n)-\alpha_k^0|+|\theta_{n,k}(t)-\theta_{n,k}(n)|+|\theta_{n,k}(n)-\theta_k^0|)\\ \leq& C\int_t^{+\infty} e^{-\frac{1}{2}\widetilde{\delta}s}+B_*(s)e^{-\delta_1 s}+e^{-\delta_2 s}ds+C|\boldsymbol{b}_n|\\ \leq& C\Big(\frac{2}{\widetilde{\delta}t}+\frac{1}{t}e^{-\frac{1}{4d}\widetilde{\delta}t}\Big)te^{-\frac{1}{2}\widetilde{\delta}t}.\end{aligned} \tag{4.22}$$

At last, in view of Proposition 3.6, estimates (4.6), (4.14) and (4.20), we derive that for any $t \in [t_*, n]$,

$$\begin{aligned}\|\varepsilon_n(t)\|^2_{H^1} \leq& C\int_t^{+\infty}\Big(\frac{1}{s}+B_*(s)\Big)e^{-\widetilde{\delta}s}ds+C\Big(\int_t^{+\infty}e^{-\frac{p\wedge 2}{2}\widetilde{\delta}s}ds\Big)^2+C\Big(\int_t^{+\infty}B_*(s)e^{-\widetilde{\delta}s}ds\Big)^2\\ &+Ce^{-(1+\frac{1}{2d})\widetilde{\delta}t}+Ce^{-\delta_2 t}+\beta\|\varepsilon_n(t)\|^2_{H^1}\\ \leq& C\Big(\frac{1}{\widetilde{\delta}t}+\frac{1}{\widetilde{\delta}}B_*(t)+\frac{4}{(p\wedge 2)^2\widetilde{\delta}^2}e^{(1-p\wedge 2)\widetilde{\delta}t}+\frac{1}{\widetilde{\delta}^2}e^{-\widetilde{\delta}t}+e^{-\frac{1}{2d}\widetilde{\delta}t}\Big)e^{-\widetilde{\delta}t}+\beta\|\varepsilon_n(t)\|^2_{H^1}\\ \leq& C\Big(\frac{1}{\widetilde{\delta}t}+\frac{1}{\widetilde{\delta}}B_*(t)+\frac{5}{\widetilde{\delta}^2}e^{-\frac{1}{4d}\widetilde{\delta}t}+e^{-\frac{1}{4d}\widetilde{\delta}t}\Big)e^{-\widetilde{\delta}t}+\beta\|\varepsilon_n(t)\|^2_{H^1}.\end{aligned} \tag{4.23}$$

Since $\beta \to 0$ as $\|\varepsilon_n(t)\|_{H^1} \to 0$, there exists a deterministic small constant $0 < \varepsilon^* < 1$, such that $\beta \leq \frac{1}{2}$ if $\|\varepsilon_n(t)\|_{H^1} \leq \varepsilon^*$.

Thus, let

$$\tau_3 := \inf\Big\{t \geq 1 + \frac{2}{\widetilde{\delta}}\ln\big(\frac{2}{\varepsilon^*}\big) : \ C\Big(\frac{2}{\widetilde{\delta}t} + \frac{1}{\widetilde{\delta}}B_*(t) + \big(\frac{5}{\widetilde{\delta}^2} + \frac{3}{\widetilde{\delta}} + 1\big)e^{-\frac{\widetilde{\delta}t}{4d}}\Big) \leq \frac{1}{8}\Big\}. \tag{4.24}$$

Note that τ_3 is finite almost surely, due to the vanishing of the noise B_* at infinity in (3.2). In view of the above estimates (4.21)-(4.23) and the definition of τ^*, we consequently get the improved estimates (4.11) and (4.12) on $[t_*, n]$.

Case (II): Let us first choose a deterministic large constant ν_0 such that

$$\nu_0 \geq 4d, \quad \text{and} \quad C\Big(\frac{2}{\nu_0 - 2} + \frac{\widetilde{\delta}}{\nu_0 - 1} + \frac{2}{(\nu_0 - 2)\widetilde{\delta}} + \frac{4}{(\nu_0 - 2)^2\widetilde{\delta}^2}\Big) \leq \frac{1}{16}, \tag{4.25}$$

where C is the deterministic constant in estimates (4.28)-(4.30) below. In the following we consider any $\nu_* \geq \nu_0$ fixed.

We derive from (4.14) and (4.15) that, as in (4.18), for any $t \in [t_*, n]$,

$$\|u_n(t) - R(t)\|_{H^1} \leq \|R(t) - \sum_{k=1}^{K} \widetilde{R}_{n,k}(t)\|_{H^1} + \|\varepsilon_n(t)\|_{H^1} \leq C(t+1)(\widetilde{\delta}t)^{-\frac{\nu_*}{2}}. \tag{4.26}$$

Then, let

$$\tau_4 := \inf\Big\{t > 0 : C(t+1)(\widetilde{\delta}t)^{-\frac{\nu_*}{2}} \leq \frac{1}{4}\delta_*\Big\}, \tag{4.27}$$

where δ_* is as in Lemma 2.6. Since by the definition (4.13), $\tau^* \geq \tau_1 \vee \tau_4$, we infer that estimate (4.10) holds on $[t_*, n]$.

Moreover, one has, via (3.54), (4.14) and $p \wedge 2 > 1$,

$$\begin{aligned} |a^+_{n,k}(t)| \leq & C\Big(\int_t^{+\infty} (\widetilde{\delta}s)^{-\frac{p\wedge 2}{2}\nu_*} + B_*(s)(\widetilde{\delta}s)^{-\nu_*}ds + e^{-\delta_2 t}\Big) \\ \leq & C\Big(\frac{2}{(\nu_* - 2)\widetilde{\delta}}(\widetilde{\delta}t)^{\frac{1-p\wedge 2}{2}\nu_* + 1} + \frac{1}{\nu_*}(\widetilde{\delta}t)^{-\frac{\nu_*}{2}} + (\widetilde{\delta}t)^{\frac{\nu_*}{2}}e^{-\delta_2 t}\Big)(\widetilde{\delta}t)^{-\frac{\nu_*}{2}}. \end{aligned} \tag{4.28}$$

Estimating as in the proof of (4.22), for any $t \in [t_*, n]$, integrating (4.17) on $[t, n]$ we get

$$\begin{aligned} \sum_{k=1}^{K}(|\alpha_{n,k}(t) - \alpha^0_k| + |\theta_{n,k}(t) - \theta^0_k|) \leq & C\Big(\int_t^{+\infty} (\widetilde{\delta}s)^{-\frac{\nu_*}{2}} + B_*(s)(\widetilde{\delta}s)^{-\nu_*} + e^{-\delta_2 s}ds + (\widetilde{\delta}t)^{-(\frac{1}{2}+\frac{1}{4d})\nu_*}\Big) \\ \leq & C\Big(\frac{2}{\nu_* - 2} + \frac{\widetilde{\delta}}{\nu_* - 1}(\widetilde{\delta}t)^{-\frac{\nu_*}{2}-1} + (\delta_2 t)^{-1}(\widetilde{\delta}t)^{\frac{\nu_*}{2}}e^{-\delta_2 t} + t^{-1}(\widetilde{\delta}t)^{-\frac{\nu_*}{4d}}\Big)t(\widetilde{\delta}t)^{-\frac{\nu_*}{2}}. \end{aligned} \tag{4.29}$$

An application of Proposition 3.6, (3.7), (4.6), (4.14) and (4.20) also gives that for any $t \in [t_*, n]$,

$$\begin{aligned} \|\varepsilon_n(t)\|^2_{H^1} \leq & C\int_t^{+\infty} s^{-1}(\widetilde{\delta}s)^{-\nu_*}ds + C\Big(\int_t^{+\infty}(\widetilde{\delta}s)^{-\frac{p\wedge 2}{2}\nu_*}ds\Big)^2 + C\Big(\int_t^{+\infty}(\widetilde{\delta}s)^{-\nu_*-1}ds\Big)^2 \\ & + C(\widetilde{\delta}t)^{-(1+\frac{1}{2d})\nu_*} + Ce^{-\delta_2 t} + \beta\|\varepsilon_n(t)\|^2_{H^1} \\ \leq & C\Big(\frac{1}{\nu_*} + \frac{1}{\nu_*^2}(\widetilde{\delta}t)^{-\nu_*} + \frac{4}{(\nu_* - 2)^2\widetilde{\delta}^2}(\widetilde{\delta}t)^{(1-(p\wedge 2))\nu_* + 2} + (\widetilde{\delta}t)^{-\frac{1}{2d}\nu_*} + (\widetilde{\delta}t)^{\nu_*}e^{-\delta_2 t}\Big)(\widetilde{\delta}t)^{-\nu_*} + \beta\|\varepsilon_n(t)\|^2_{H^1}. \end{aligned} \tag{4.30}$$

Thus, let

$$\tau_5 := \inf\Big\{t \geq 1 + \frac{\nu_*}{\delta_2} + \frac{1}{\widetilde{\delta}}\Big(\frac{2}{\varepsilon^*}\Big)^{\frac{2}{\nu_*}} : \ C\Big(\big(\widetilde{\delta}t\big)^{\nu_*}e^{-\delta_2 t} + \big(\widetilde{\delta}t\big)^{-\frac{\nu_*}{4d}}\Big) \leq \frac{1}{16}\Big\}, \tag{4.31}$$

where C is the deterministic constant from the above estimates (4.28)-(4.30), and $\widetilde{\delta}$ is given by (4.3). In view of estimates (4.28)-(4.30), the definition of τ^* in (4.13) and the choice of ν_* in (4.25), we verify estimates (4.11) and (4.12) on $[t_*, n]$. The proof is consequently complete. □

4.2. **Topological arguments.** As mentioned above, Proposition 4.3 requires the a-priori control of $\mathbf{a}_n^-$ so that the estimates of ε_n, α_n, θ_n and $\mathbf{a}_n^+$ can be improved.

In order to control the unstable direction $\mathbf{a}_n^-$, we use a topological argument as in [12]. It is necessary to keep valid the estimate (4.8) of $\mathbf{a}_n^-$, this motivates the following definition

$$T_0(\boldsymbol{a}_n^-) := \inf\{T \geq \tau^* : \text{ estimates } (4.7) - (4.9) \text{ hold on } [T, n]\}. \tag{4.32}$$

Let $\nu_* \geq \nu_0$ with ν_0 satisfying (4.25). We define a universal random time T_0, independent of n, by

$$T_0 := \begin{cases} \max\{M, \sigma_1, \tau_j, j = 1, 2, 3, 6\}, & \text{in Case (I);} \\ \max\{M, \sigma_1, \sigma_2, \tau_j, j = 1, 4, 5, 7\}, & \text{in Case (II),} \end{cases} \tag{4.33}$$

where M is the large deterministic time from the geometrical decomposition in Proposition 2.4, σ_1 and σ_2 are the random times in Subsection 3.1, τ_j, $1 \leq j \leq 5$, are defined as in the previous subsection,

$$\tau_6 := \inf\Big\{t > 0 : C_1 e^{(\frac{1}{2} + \frac{1}{4d} - \frac{p\wedge 2}{2})\widetilde{\delta} t} \leq \frac{\widetilde{\delta}}{4}\Big\}, \tag{4.34}$$

and

$$\tau_7 := \inf\Big\{t \geq \frac{3\nu_*}{2\delta_3} : C_2\Big((\widetilde{\delta} t)^{(\frac{1}{2} + \frac{1}{4d} - \frac{p\wedge 2}{2})\nu_*} + (\widetilde{\delta} t)^{(\frac{1}{2} + \frac{1}{4d})\nu_*} e^{-\delta_2 t}\Big) \leq \frac{\delta_3}{4}\Big\} \tag{4.35}$$

with C_1, C_2 being the deterministic constants in (4.44) and (4.39) below, respectively.

We aim to show that there exists an appropriate vector $\boldsymbol{a}_n^- \in B_{\mathbb{R}^K}(\phi^{\frac{1}{2}+\frac{1}{4d}}(\widetilde{\delta} n))$ such that $T_0(\boldsymbol{a}_n^-)$ is less than the universal time T_0, i.e., $T_0(\boldsymbol{a}_n^-) \leq T_0$. This is important in the next section to pass to the limit of the approximating solutions to construct the desired stochastic multi-solitons.

Proposition 4.5 (Uniform backward time). *Let $\widetilde{\delta}$, v_*, τ^* be as in Proposition 4.3, and $T_0(\boldsymbol{a}_n^-)$, T_0 defined as above. Then, for $\mathbb{P}$-a.e. $\omega \in \Omega$ there exists $N_0(\omega) \geq 1$ such that for any $n \geq N_0(\omega)$, there exists $\boldsymbol{a}_n^-(\omega) \in B_{\mathbb{R}^K}(\phi^{\frac{1}{2}+\frac{1}{4d}}(\widetilde{\delta} n))$ such that $T_0(\boldsymbol{a}_n^-(\omega)) \leq T_0(\omega)$.*

Proof. We shall prove this by contradiction. Suppose that there exists a measurable set $\Omega' \subseteq \Omega$ with positive probability, such that for every $\omega \in \Omega'$ and for any N_0 large enough, there exists $n_0(\omega) \geq N_0$, such that $T_0(\boldsymbol{a}_{n_0}^-)(\omega) > T_0(\omega)$ for all $\boldsymbol{a}_{n_0}^- \in B_{\mathbb{R}^K}(\phi^{\frac{1}{2}+\frac{1}{4d}}(\widetilde{\delta} n_0(\omega)))$. Below we fix $\omega \in \Omega'$ and take $N_0(\omega) > T_0(\omega)$. We omit the dependence of ω to simplify the notations.

In view of Proposition 2.7, for N_0 possibly larger such that $\phi^{\frac{1}{2}+\frac{1}{4d}}(\widetilde{\delta} N_0) \leq r$, we infer that for any $\boldsymbol{a}_{n_0}^- \in B_{\mathbb{R}^K}\big(\phi^{\frac{1}{2}+\frac{1}{4d}}(\widetilde{\delta} n_0)\big)$, there exists a unique vector $\boldsymbol{b}_{n_0} \in \mathbb{R}^{2K}$ such that

$$\mathbf{a}_{n_0}^+(n_0) = \mathbf{0}, \quad \mathbf{a}_{n_0}^-(n_0) = \boldsymbol{a}_{n_0}^-, \quad \text{and} \quad |\boldsymbol{b}_{n_0}| \leq C\phi^{\frac{1}{2}+\frac{1}{4d}}(\widetilde{\delta} n_0). \tag{4.36}$$

By the definition of $T_0(\boldsymbol{a}_{n_0}^-)$, u_{n_0} admits the geometrical decomposition (2.40) and estimates (4.7)-(4.9) on $(T_0(\boldsymbol{a}_{n_0}^-), n_0]$. It is still valid on the closed interval $[T_0(\boldsymbol{a}_{n_0}^-), n_0]$, due to the continuity of solutions and modulation parameters.

Moreover, by Proposition 4.3 and the definition of T_0 in (4.33), one has the improved estimates (4.10)-(4.12) on $[T_0(\boldsymbol{a}_{n_0}^-), n_0]$. In particular,

$$\begin{aligned}
&\|u_{n_0}(T_0(\boldsymbol{a}_{n_0}^-)) - R(T_0(\boldsymbol{a}_{n_0}^-))\|_{H^1} \leq \frac{\delta_*}{4}, \quad \|\varepsilon_{n_0}(T_0(\boldsymbol{a}_{n_0}^-))\|_{H^1} \leq \frac{1}{2}\phi^{\frac{1}{2}}(\widetilde{\delta} T_0(\boldsymbol{a}_{n_0}^-)), \\
&\sum_{k=1}^{K}(|\alpha_{n_0,k}(T_0(\boldsymbol{a}_{n_0}^-)) - \alpha_k^0| + |\theta_{n_0,k}(T_0(\boldsymbol{a}_{n_0}^-)) - \theta_k^0|) \leq \frac{1}{2} T_0(\boldsymbol{a}_{n_0}^-)\phi^{\frac{1}{2}}(\widetilde{\delta} T_0(\boldsymbol{a}_{n_0}^-)), \\
&|\mathbf{a}_{n_0}^+(T_0(\boldsymbol{a}_{n_0}^-))| \leq \frac{1}{2}\phi^{\frac{1}{2}}(\widetilde{\delta} T_0(\boldsymbol{a}_{n_0}^-)).
\end{aligned}$$

Note that

$$|\mathbf{a}_{n_0}^-(T_0(\boldsymbol{a}_{n_0}^-))| = \phi^{\frac{1}{2}+\frac{1}{4d}}(\widetilde{\delta} T_0(\boldsymbol{a}_{n_0}^-)). \tag{4.37}$$

That is, $\mathbf{a}_{n_0}^-(T_0(\boldsymbol{a}_{n_0}^-))$ is in the sphere $S_{\mathbb{R}^K}\big(\phi^{\frac{1}{2}+\frac{1}{4d}}(\widetilde{\delta} T_0(\boldsymbol{a}_{n_0}^-))\big)$. In fact, if $|\mathbf{a}_{n_0}^-(T_0(\boldsymbol{a}_{n_0}^-))| < \phi^{\frac{1}{2}+\frac{1}{4d}}(\widetilde{\delta} T_0(\boldsymbol{a}_{n_0}^-))$, then by the above estimates and the continuity of modulation parameters and remainder, one can find a small time $\eta > 0$ such that estimates (4.7)-(4.9) hold on $[T_0(\boldsymbol{a}_{n_0}^-) - \eta, n_0]$, which contradicts the definition of $T_0(\boldsymbol{a}_{n_0}^-)$.

Now, we define the map by

$$\begin{aligned}\Lambda: \quad & B_{\mathbb{R}^K}\big(\phi^{\frac12+\frac{1}{4d}}(\widetilde{\delta}n_0)\big) \to S_{\mathbb{R}^K}\big(\phi^{\frac12+\frac{1}{4d}}(\widetilde{\delta}n_0)\big),\\ & \qquad\qquad \boldsymbol{a}^-_{n_0} \mapsto \phi^{\frac12+\frac{1}{4d}}(\widetilde{\delta}n_0)\phi^{-(\frac12+\frac{1}{4d})}(\widetilde{\delta}T_0(\boldsymbol{a}^-_{n_0}))\mathbf{a}^-_{n_0}(T_0(\boldsymbol{a}^-_{n_0})).\end{aligned}$$

It follows from (4.37) that Λ is well defined.

Below we show that Λ is continuous, and it is the identity map when restricted to $S_{\mathbb{R}^K}(\phi^{\frac12+\frac{1}{4d}}(\widetilde{\delta}n_0))$. Assuming these to hold, we then infer that the continuous map $\widetilde{\Lambda} := -\Lambda$ from $B_{\mathbb{R}^K}\big(\phi^{\frac12+\frac{1}{4d}}(\widetilde{\delta}n_0)\big)$ to $S_{\mathbb{R}^K}\big(\phi^{\frac12+\frac{1}{4d}}(\widetilde{\delta}n_0)\big)$ has no fixed point, which contradicts the Brouwer fixed point theorem. In fact, since the image of $\widetilde{\Lambda}$ is in $S_{\mathbb{R}^K}\big(\phi^{\frac12+\frac{1}{4d}}(\widetilde{\delta}n_0)\big)$, it is clear that

$$\widetilde{\Lambda}(\boldsymbol{a}^-_{n_0}) \neq \boldsymbol{a}^-_{n_0}, \ \forall \boldsymbol{a}^-_{n_0} \in \mathring{B}_{\mathbb{R}^K}\big(\phi^{\frac12+\frac{1}{4d}}(\widetilde{\delta}n_0)\big).$$

Moreover, since Λ is the identity map when restricted to $S_{\mathbb{R}^K}(\phi^{\frac12+\frac{1}{4d}}(\widetilde{\delta}n_0))$, we have

$$\widetilde{\Lambda}(\boldsymbol{a}^-_{n_0}) = -\Lambda(\boldsymbol{a}^-_{n_0}) = -\boldsymbol{a}^-_{n_0} \neq \boldsymbol{a}^-_{n_0}, \ \forall \boldsymbol{a}^-_{n_0} \in S_{\mathbb{R}^K}\big(\phi^{\frac12+\frac{1}{4d}}(\widetilde{\delta}n_0)\big).$$

The above two facts together then yield that $\widetilde{\Lambda}$ has no fixed point, as claimed.

Below let us start with the proof of the continuity of the map Λ.

(i)**. Continuity of** Λ: It suffices to prove that the map $\boldsymbol{a}^-_{n_0} \mapsto T_0(\boldsymbol{a}^-_{n_0})$ is continuous. For this purpose, we fix $\boldsymbol{a}^-_{n_0} \in B_{\mathbb{R}^K}(\phi^{(\frac12+\frac{1}{4d})}(\widetilde{\delta}n_0))$ and let $\widehat{T}(\boldsymbol{a}^-_{n_0}) \in (T_0, T_0(\boldsymbol{a}^-_{n_0}))$ be close enough to $T_0(\boldsymbol{a}^-_{n_0})$ such that estimates (4.14) and (4.15) with n replaced by n_0 hold on $[\widehat{T}(\boldsymbol{a}^-_{n_0}), n_0]$. Let

$$\mathcal{N}(t, \boldsymbol{a}^-_{n_0}) := |\phi^{-(\frac12+\frac{1}{4d})}(\widetilde{\delta}t)\mathbf{a}^-_{n_0}(t)|^2, \quad t \in [\widehat{T}(\boldsymbol{a}^-_{n_0}), n_0]. \tag{4.38}$$

Below we mainly consider the polynomial decay rate in the stochastic Case (II), as Case (I) is easier and can be treated in an analogous manner.

Case (II): Recall that $\phi(x) = |x|^{-\nu_*}$ is the spatial decay function of noise given by (1.6) in Case (II). By straightforward computations, (3.14) and (4.2), we have for any $t \in [\widehat{T}(\boldsymbol{a}^-_{n_0}), n_0]$,

$$\begin{aligned}\frac{d\mathcal{N}}{dt}(t, \boldsymbol{a}^-_{n_0}) =& \frac{d}{dt}\bigg((\widetilde{\delta}t)^{(1+\frac{1}{2d})\nu_*}|\mathbf{a}^-_{n_0}(t)|^2\bigg)\\ =&\big(1+(2d)^{-1}\big)\widetilde{\delta}\nu_*\big(\widetilde{\delta}t\big)^{(1+\frac{1}{2d})\nu_*-1}|\mathbf{a}^-_{n_0}(t)|^2 + 2\big(\widetilde{\delta}t\big)^{(1+\frac{1}{2d})\nu_*}\sum_{k=1}^K\Big(a^-_{n_0,k}(t)\frac{d}{dt}a^-_{n_0,k}(t)\Big)\\ \leq&\big(1+(2d)^{-1}\big)\nu_* t^{-1}\mathcal{N}(t, \boldsymbol{a}^-_{n_0}) + 2(\widetilde{\delta}t)^{(1+\frac{1}{2d})\nu_*}\sum_{k=1}^K\Big(-e_0(w_k)^{-2}\big(a^-_{n_0,k}(t)\big)^2\\ &\quad + C\big|a^-_{n_0,k}(t)\big|\big((\widetilde{\delta}t)^{-\frac{p\wedge 2}{2}\nu_*} + t^{-1}(\widetilde{\delta}t)^{-\nu_*} + e^{-\delta_2 t}\big)\Big)\\ \leq&\Big(\frac32\nu_* t^{-1} - 2\delta_3\Big)\mathcal{N}(t, \boldsymbol{a}^-_{n_0}) + C_2\big((\widetilde{\delta}t)^{(\frac12+\frac{1}{4d}-\frac{p\wedge2}{2})\nu_*} + (\widetilde{\delta}t)^{(\frac12+\frac{1}{4d})\nu_*}e^{-\delta_2 t}\big)\sqrt{\mathcal{N}(t, \boldsymbol{a}^-_{n_0})},\end{aligned} \tag{4.39}$$

which along with (4.33) and (4.35) yields that for any $t \in [\widehat{T}(\boldsymbol{a}^-_{n_0}), n_0]$,

$$\frac{d\mathcal{N}}{dt}(t, \boldsymbol{a}^-_{n_0}) \leq -\delta_3\mathcal{N}(t, \boldsymbol{a}^-_{n_0}) + \frac{\delta_3}{4}\sqrt{\mathcal{N}(t, \boldsymbol{a}^-_{n_0})}. \tag{4.40}$$

We next claim that for $\eta > 0$ small enough, there exists $\delta > 0$ such that

$$\mathcal{N}(t, \boldsymbol{a}^-_{n_0}) < 1 - \delta, \quad \forall t \in [T_0(\boldsymbol{a}^-_{n_0}) + \eta, n_0], \tag{4.41}$$

and

$$\mathcal{N}(t, \boldsymbol{a}^-_{n_0}) > 1 + \delta, \quad \forall t \in [\widehat{T}(\boldsymbol{a}^-_{n_0}), T_0(\boldsymbol{a}^-_{n_0}) - \eta]. \tag{4.42}$$

In the case where $T_0(\boldsymbol{a}^-_{n_0}) + \eta > n_0$ we only consider (4.42), which may happen if $\boldsymbol{a}^-_{n_0} \in S_{\mathbb{R}^K}((\widetilde{\delta}n_0)^{-(\frac12+\frac{1}{4d})\nu_*})$, see the next step (ii) below.

Below we prove (4.41), and the proof of (4.42) is similar. To this end, we argue by contradiction and assume that there exists $\eta_* > 0$ such that for any $m \geq 1$, there exists $t_m \in [T_0(\boldsymbol{a}^-_{n_0}) + \eta_*, n_0]$, such that

$$\mathcal{N}(t_m, \boldsymbol{a}^-_{n_0}) \geq 1 - \frac{1}{m}. \tag{4.43}$$

By compactness, there exists a subsequence (still denoted by $\{m\}$) such that $t_m \to t_0$ as $m \to +\infty$ and $t_0 \in [T_0(\boldsymbol{a}^-_{n_0}) + \eta_*, n_0]$. Then passing to the limit $m \to +\infty$ in (4.43) we get $\mathcal{N}(t_0, \boldsymbol{a}^-_{n_0}) \geq 1$. But by the definition of $T_0(\boldsymbol{a}^-_{n_0})$ in (4.32), $\mathcal{N}(t, \boldsymbol{a}^-_{n_0}) \leq 1$ for any $t \geq T_0(\boldsymbol{a}^-_{n_0})$. It thus follows that $\mathcal{N}(t_0, \boldsymbol{a}^-_{n_0}) = 1$. Then, taking $t = t_0$ in (4.40) we obtain

$$\frac{d\mathcal{N}}{dt}(t_0, \boldsymbol{a}^-_{n_0}) \leq -\frac{3}{4}\delta_3 < 0.$$

This yields that $\mathcal{N}(\widetilde{t}_0, \boldsymbol{a}^-_{n_0}) > 1$ for some $\widetilde{t}_0 \in [T_0(\boldsymbol{a}^-_{n_0}), t_0]$, which however contradicts the fact that $\mathcal{N}(\widetilde{t}_0, \boldsymbol{a}^-_{n_0}) \leq 1$, thereby proving (4.41), as claimed.

Now, since by (4.38), $\mathcal{N}(t, \boldsymbol{a}^-_{n_0})$ is continuous in $\mathbf{a}^-_{n_0}(t)$, and for all $t \in [\widehat{T}(\boldsymbol{a}^-_{n_0}), n_0]$, $\mathbf{a}^-_{n_0}(t)$ is continuous in $\boldsymbol{a}^-_{n_0}$ due to Proposition 2.7 and the continuity of the flow of (4.1). It follows that $\mathcal{N}(t, \boldsymbol{a}^-_{n_0})$ is continuous in $\boldsymbol{a}^-_{n_0}$. Moreover, using the contradiction assumption that $T_0(\boldsymbol{a}^-_{n_0}) > T_0$, Proposition 2.7 and the continuity of the flow of (4.1) again we get that there exists $\zeta (> 0)$ small enough, such that for any $\widetilde{\boldsymbol{a}}^-_{n_0} \in B_{\mathbb{R}^K}(\boldsymbol{a}^-_{n_0}, \zeta)$, one has $\widehat{T}(\widetilde{\boldsymbol{a}}^-_{n_0}) \leq T_0(\boldsymbol{a}^-_{n_0})$. Thus, $\mathcal{N}(T_0(\boldsymbol{a}^-_{n_0}), \widetilde{\boldsymbol{a}}^-_{n_0})$ is well-defined. Then, by the continuity of the map $\widetilde{\boldsymbol{a}}^-_{n_0} \mapsto \mathcal{N}(T_0(\boldsymbol{a}^-_{n_0}), \widetilde{\boldsymbol{a}}^-_{n_0})$, there exists $\zeta = \zeta(\delta, T_0(\boldsymbol{a}^-_{n_0}))$ possibly smaller with δ as in (4.41) and (4.42), such that for any $\widetilde{\boldsymbol{a}}^-_{n_0} \in B_{\mathbb{R}^K}(\boldsymbol{a}^-_{n_0}, \zeta)$, one has $|\mathcal{N}(T_0(\boldsymbol{a}^-_{n_0}), \widetilde{\boldsymbol{a}}^-_{n_0}) - \mathcal{N}(T_0(\boldsymbol{a}^-_{n_0}), \boldsymbol{a}^-_{n_0})| = |\mathcal{N}(T_0(\boldsymbol{a}^-_{n_0}), \widetilde{\boldsymbol{a}}^-_{n_0}) - 1| \leq \frac{\delta}{2}$. Taking into account (4.41) and (4.42) with $\boldsymbol{a}^-_{n_0}$ replaced by $\widetilde{\boldsymbol{a}}^-_{n_0}$, we thus derive that $|T_0(\widetilde{\boldsymbol{a}}^-_{n_0}) - T_0(\boldsymbol{a}^-_{n_0})| < \eta$ for any $\widetilde{\boldsymbol{a}}^-_{n_0} \in B_{\mathbb{R}^K}(\boldsymbol{a}^-_{n_0}, \zeta)$. This gives the continuity of the map $\boldsymbol{a}^-_{n_0} \mapsto T_0(\boldsymbol{a}^-_{n_0})$ in Case (II).

Case (I): In this case, we have from (1.6) that $\phi(x) = e^{-|x|}$. For any $t \in [\widehat{T}(\boldsymbol{a}^-_{n_0}), n_0]$, one can replace (4.39) by the following estimate

$$\begin{aligned}
\frac{d\mathcal{N}}{dt}(t, \boldsymbol{a}^-_{n_0}) =& \frac{d}{dt}\left(e^{(1+\frac{1}{2d})\widetilde{\delta}t}|\mathbf{a}^-_{n_0}(t)|^2\right) \\
=& \Big(1 + \frac{1}{2d}\Big)\widetilde{\delta}e^{(1+\frac{1}{2d})\widetilde{\delta}t}|\mathbf{a}^-_{n_0}(t)|^2 + 2e^{(1+\frac{1}{2d})\widetilde{\delta}t}\sum_{k=1}^{K}\Big(a^-_{n_0,k}(t)\frac{d}{dt}a^-_{n_0,k}(t)\Big) \\
\leq& \Big(1 + \frac{1}{2d}\Big)\widetilde{\delta}\mathcal{N}(t, \boldsymbol{a}^-_{n_0}) + 2e^{(1+\frac{1}{2d})\widetilde{\delta}t}\sum_{k=1}^{K}\Big(-e_0\,(w_k)^{-2}\,\big(a^-_{n_0,k}(t)\big)^2 \\
&\quad + C\big|a^-_{n_0,k}(t)\big|\big(e^{-\frac{p\wedge 2}{2}\widetilde{\delta}t} + e^{-\delta_1 t} + e^{-\delta_2 t}\big)\Big) \\
\leq& -\widetilde{\delta}\mathcal{N}(t, \boldsymbol{a}^-_{n_0}) + C_1 e^{(\frac{1}{2}+\frac{1}{4d}-\frac{p\wedge 2}{2})\widetilde{\delta}t}\sqrt{\mathcal{N}(t, \boldsymbol{a}^-_{n_0})} \\
\leq& -\widetilde{\delta}\mathcal{N}(t, \boldsymbol{a}^-_{n_0}) + \frac{\widetilde{\delta}}{4}\sqrt{\mathcal{N}(t, \boldsymbol{a}^-_{n_0})}.
\end{aligned} \tag{4.44}$$

Then, similar arguments as in Case (II) lead to the continuity of the map $\boldsymbol{a}^-_{n_0} \mapsto T_0(\boldsymbol{a}^-_{n_0})$.

(ii)**. Identity of Λ when restricted to sphere**: It remains to prove that Λ is the identity map when restricted to $S_{\mathbb{R}^K}(\phi^{\frac{1}{2}+\frac{1}{4d}}(\widetilde{\delta}n_0))$.

To this end, for any $\boldsymbol{a}^-_{n_0} \in S_{\mathbb{R}^K}(\phi^{\frac{1}{2}+\frac{1}{4d}}(\widetilde{\delta}n_0))$, using (4.36) we have

$$\mathcal{N}(n_0, \boldsymbol{a}^-_{n_0}) = |\phi^{-(\frac{1}{2}+\frac{1}{4d})}(\widetilde{\delta}n_0)\mathbf{a}^-_{n_0}(n_0)|^2 = |\phi^{-(\frac{1}{2}+\frac{1}{4d})}(\widetilde{\delta}n_0)\boldsymbol{a}^-_{n_0}|^2 = 1. \tag{4.45}$$

Moreover, letting $t = n_0$ in (4.44) and (4.40) we obtain

$$\frac{d\mathcal{N}}{dt}(n_0, \boldsymbol{a}^-_{n_0}) < 0 \tag{4.46}$$

in both Case (I) and Case (II). Suppose that $T_0(\boldsymbol{a}^-_{n_0}) < n_0$, then (4.45) and (4.46) imply that there exists $t \in (T_0(\boldsymbol{a}^-_{n_0}), n_0)$ such that $\mathcal{N}(t, \boldsymbol{a}^-_{n_0}) > 1$. But by the definition of $T_0(\boldsymbol{a}^-_{n_0})$ in (4.32), $\mathcal{N}(t, \boldsymbol{a}^-_{n_0}) \leq 1$ for any $t \in [T_0(\boldsymbol{a}^-_{n_0}), n_0]$. This leads to a contradiction. Thus, we get $T_0(\boldsymbol{a}^-_{n_0}) = n_0$.

Consequently, by the definition of Λ and (4.36), $\Lambda(\boldsymbol{a}_{n_0}^-) = \boldsymbol{a}_{n_0}^-$ for all $\boldsymbol{a}_{n_0}^- \in S_{\mathbb{R}^K}(\phi^{\frac{1}{2}+\frac{1}{4d}}(\widetilde{\delta}n_0))$, which shows that the map Λ restricted to $S_{\mathbb{R}^K}(\phi^{\frac{1}{2}+\frac{1}{4d}}(\widetilde{\delta}n_0))$ is the identity map.

Therefore, the proof of Proposition 4.5 is complete. $\square$

4.3. **Proof of uniform estimate.** Now, we are in position to prove the uniform estimate in Theorem 4.1.

Proof of Theorem 4.1. By Propositions 2.7 and 4.5, for $\mathbb{P}$-a.e. $\omega \in \Omega$ and for $n = n(\omega)$ large enough, there exist $\boldsymbol{a}_n^-(\omega) \in B_{\mathbb{R}^K}(\phi^{\frac{1}{2}+\frac{1}{4d}}(\widetilde{\delta}n))$ and a unique $\boldsymbol{b}_n(\omega) = \boldsymbol{b}_n(\boldsymbol{a}_n^-(\omega))$, such that $T_0(\boldsymbol{a}_n^-(\omega)) \leq T_0(\omega)$, where $T_0(\omega)$ is independent of n. Thus, by the definition of $T_0(\boldsymbol{a}_n^-(\omega))$ in (4.32), $u_n(\omega)$ admits the geometrical decomposition (2.40) and estimates (4.8), (4.9) on $[T_0(\omega), n]$.

Regarding the error estimate (4.4), we see that in Case (I), for any $t \in [T_0(\omega), n]$, by (4.8) and (4.9),

$$\begin{aligned}\|u_n(t,\omega) - R(t,\omega)\|_{H^1} \leq& \|R(t,\omega) - \widetilde{R}_n(t,\omega)\|_{H^1} + \|\varepsilon_n(t,\omega)\|_{H^1} \\ \leq& C\sum_{k=1}^{K}(|\alpha_{n,k}(t,\omega) - \alpha_k^0| + |\theta_{n,k}(t,\omega) - \theta_k^0|) + \|\varepsilon_n(t,\omega)\|_{H^1} \\ \leq& Cte^{-\frac{1}{2}\widetilde{\delta}t}.\end{aligned}$$

Moreover, in Case (II), for any $t \in [T_0(\omega), n]$,

$$\|u_n(t,\omega) - R(t,\omega)\|_{H^1} \leq C\sum_{k=1}^{K}(|\alpha_{n,k}(t,\omega) - \alpha_k^0| + |\theta_{n,k}(t,\omega) - \theta_k^0|) + \|\varepsilon_n(t,\omega)\|_{H^1} \leq Ct(\widetilde{\delta}t)^{-\frac{\nu_*}{2}}.$$

Thus, estimate (4.4) is verified in both Case(I) and Case (II). The proof of Theorem 4.1 is complete. $\square$

5. Proof of main results

We are now ready to prove the main results in Theorems 1.2 and 1.7.

Proof of Theorem 1.7. Let us fix $\omega \in \Omega$ as in Theorem 4.1 and omit it in the following arguments to ease notations. Let $\{u_n\}$ be the approximating solutions to (4.1). By virtue of Theorem 4.1 and the expressions of $R(t)$ and $R_k(t)$ in (2.36) and (1.13), respectively, we derive that

$$\|u_n(t)\|_{H^1} \leq \|u_n(t) - R(t)\|_{H^1} + \|R(t)\|_{H^1} \leq Ct\phi^{\frac{1}{2}}(\widetilde{\delta}t) + \sum_{k=1}^{K}\|R_k(t)\|_{H^1} \leq C, \quad \forall t \in [T_0, n], \tag{5.1}$$

where C depends on w_k and $\widetilde{\delta}$ and is independent of n and t. It follows that $\{u_n(t)\}$ is uniformly bounded in $H^1(\mathbb{R}^d)$. In particular, letting $t = T_0$ we obtain a subsequence (still denoted by $\{u_n(T_0)\}$) such that

$$u_n(T_0) \rightharpoonup u_0 \text{ in } H^1(\mathbb{R}^d), \quad \text{as } n \to \infty, \tag{5.2}$$

for some $u_0 \in H^1(\mathbb{R}^d)$.

Claim: One has the strong convergence of $u_n(T_0)$ in $L^2(\mathbb{R}^d)$, i.e.,

$$u_n(T_0) \to u_0 \text{ in } L^2(\mathbb{R}^d), \quad \text{as } n \to \infty. \tag{5.3}$$

To this end, for any $\eta > 0$, since $u_0 \in H^1(\mathbb{R}^d)$, we can take A_0 large enough such that

$$\int_{|x|\geq A_0} |u_0(x)|^2 dx \leq \frac{\eta}{8}. \tag{5.4}$$

By Proposition 4.5, estimates (4.7)-(4.9) hold on the time interval $[T_0, n]$. For n large enough, we fix $\widetilde{T}_0 \in (T_0, n]$ independent of n such that

$$\|\varepsilon_n(\widetilde{T}_0)\|_{H^1}^2 \leq \phi(\widetilde{\delta}\widetilde{T}_0) \leq \frac{\eta}{16}. \tag{5.5}$$

Then, we take A_1 large enough such that for $|x| \geq A_1$ and every $1 \leq k \leq K$,

$$\inf_{n\geq 1} |x - v_k\widetilde{T}_0 - \alpha_{n,k}(\widetilde{T}_0)| \geq |x| - |v_k|\widetilde{T}_0 - \sup_{n\geq 1, t\geq T_0} |\alpha_{n,k}(t)| \geq A_1 - \frac{1}{2}A_1 = \frac{1}{2}A_1,$$

and, via the exponential decay of the ground state in (1.10), we may take A_1 larger such that

$$\sup_{n\geq 1}\int_{|x|\geq A_1}|\widetilde{R}_n(\widetilde{T}_0)|^2dx \leq C\sum_{k=1}^{K}\int_{|x|\geq \frac{A_1}{2w_k}} e^{-2\delta|x|}dx \leq C\sum_{k=1}^{K} e^{-\frac{\delta A_1}{w_k}} \leq \frac{\eta}{16}, \tag{5.6}$$

where C depends on A_1, w_k and δ.

Combining (5.5) and (5.6) we obtain

$$\sup_{n\geq 1}\int_{|x|\geq A_1}|u_n(\widetilde{T}_0)|^2dx \leq 2\sup_{n\geq 1}\int_{|x|\geq A_1}|\widetilde{R}_n(\widetilde{T}_0)|^2dx + 2\|\varepsilon_n(\widetilde{T}_0)\|_{H^1}^2 \leq \frac{\eta}{4}. \tag{5.7}$$

Moreover, let $g(x)\in C^1(\mathbb{R})$ be such that $0\leq g(x)\leq 1$, $g(x)=0$ for $|x|\leq \frac{1}{2}$, $g(x)=1$ for $|x|\geq 1$, and $|g'(x)|\leq 2$ for $x\in\mathbb{R}$. Let $g_{A(\eta)} := g(\frac{|x|}{A(\eta)})$, where $A(\eta)$ is a constant to be determined later.

By the integration-by-parts formula and the boundness of $B_*(t)$, $\|u_n\|_{H^1}$ in (3.3) and (5.1), there exists a positive constant C_1 such that for any $t\in[T_0,n]$,

$$\left|\frac{d}{dt}\int g_{A(\eta)}|u_n(t)|^2dx\right| = \left|2\mathrm{Im}\int g'_{A(\eta)}(\partial_1 u_n)\bar{u}_n dx + 2\mathrm{Re}\sum_{l=1}^{N}\int_t^{+\infty} g_l(s)dB_l(s)\int g'_{A(\eta)}(\partial_1\phi_l)|u_n|^2dx\right| \leq \frac{C_1}{A(\eta)}.$$

This implies that

$$\begin{aligned}
\int_{|x|\geq A(\eta)}|u_n(T_0)|^2dx &\leq \int_{\mathbb{R}^d}|u_n(T_0)|^2 g_{A(\eta)}dx\\
&\leq \int_{\mathbb{R}^d}|u_n(\widetilde{T}_0)|^2 g_{A(\eta)}dx + \int_{T_0}^{\widetilde{T}_0}\left|\frac{d}{dt}\int_{\mathbb{R}^d}|u_n(t)|^2 g_{A(\eta)}dx\right|dt\\
&\leq \int_{|x|\geq\frac{1}{2}A(\eta)}|u_n(\widetilde{T}_0)|^2dx + \frac{C_1}{A(\eta)}(\widetilde{T}_0-T_0).
\end{aligned} \tag{5.8}$$

Thus, setting $A(\eta)=\max\left\{A_0,\ 2A_1,\ \frac{8C_1(\widetilde{T}_0-T_0)}{\eta}\right\}$ and combining (5.4), (5.7) and (5.8) together we obtain

$$\begin{aligned}
\int|u_n(T_0)-u_0|^2dx &= \int_{|x|\leq A(\eta)}|u_n(T_0)-u_0|^2dx + \int_{|x|>A(\eta)}|u_n(T_0)-u_0|^2dx\\
&\leq \int_{|x|\leq A(\eta)}|u_n(T_0)-u_0|^2dx + 2\int_{|x|>A(\eta)}|u_n(T_0)|^2dx + 2\int_{|x|>A(\eta)}|u_0|^2dx\\
&\leq \int_{|x|\leq A(\eta)}|u_n(T_0)-u_0|^2dx + \eta.
\end{aligned} \tag{5.9}$$

By the compact embedding $H^1(B_{\mathbb{R}^d}(A(\eta)))\hookrightarrow L^2(B_{\mathbb{R}^d}(A(\eta)))$,

$$\lim_{n\to+\infty}\int_{|x|\leq A(\eta)}|u_n(T_0)-u_0|^2dx = 0,$$

which along with (5.9) yields that $\lim_{n\to+\infty}\int|u_n(T_0)-u_0|^2dx\leq\eta$, thereby proving (5.3) due to the arbitrariness of $\eta>0$.

Now, we consider equation

$$\begin{cases} i\partial_t u + (\Delta + b_*\cdot\nabla + c_*)u + |u|^{p-1}u = 0,\\ u(T_0)=u_0. \end{cases} \tag{5.10}$$

By (4.4) and (5.3), the standard well-posedness theory shows that there exists a unique L^2-solution u to (5.10) on $[T_0,+\infty)$, where T_0 is the universal time in Proposition 4.5, such that

$$\lim_{n\to+\infty}\|u_n(t)-u(t)\|_{L^2}=0,\quad \forall t\in[T_0,\infty). \tag{5.11}$$

The preservation of H^1-regularity also yields $u(t)\in H^1$ for $t\in[T_0,+\infty)$.

Moreover, by the uniform estimates in Theorem 4.1, for any $t\in[T_0,\infty)$ and for n large enough,

$$\|u_n(t)-R(t)\|_{H^1}\leq Ct\phi^{\frac{1}{2}}(\widetilde{\delta}t). \tag{5.12}$$

Together with (5.11), this yields that up to a subsequence (still denoted by {n} which may depend on t),

$$u_n(t) - R(t) \rightharpoonup u(t) - R(t) \ \text{ weakly in } \ H^1, \text{ as } n \to \infty. \tag{5.13}$$

Letting $n \to +\infty$ in (5.12) we thus obtain (1.20) and finish the proof of Theorem 1.7. $\square$

Proof of Theorem 1.2. (1.14) follows directly from (1.17) and (1.20). To prove (1.16), it suffices to prove that

$$\|X(t) - e^{-W_*(t)}X(t)\|_{H^1} \le C \sum_{k=1}^{N} \left(\int_t^\infty g_k^2 ds \log \left(\int_t^\infty g_k^2 ds \right)^{-1} \right)^{\frac{1}{2}} =: CL(t). \tag{5.14}$$

To this end, using the inequality $|1 - e^{ix}| \le 2|x|$ for any $x \in \mathbb{R}$, the explicit formula (1.15) and the mass conservation of $X(t)$, we compute that

$$\begin{aligned} \|X(t) - e^{-W_*(t)}X(t)\|_{H^1} \le & \|(1 - e^{-W_*(t)})X(t)\|_{L^2} + \|\nabla(1 - e^{-W_*(t)})X(t)\|_{L^2} + \|(1 - e^{-W_*(t)})\nabla X(t)\|_{L^2} \\ \le & C\|W_*(t)\|_{W^{1,\infty}}(\|X(t)\|_{L^2} + \|\nabla X(t)\|_{L^2}) \\ \le & CL(t)(1 + \|\nabla X(t)\|_{L^2}), \end{aligned} \tag{5.15}$$

where we also used the Levy Hölder continuity estimate of Brownian motions in the last step.

Note that, by (1.17), (1.20) and the mass conservation law of $X(t)$,

$$\begin{aligned} \|\nabla X(t)\|_{L^2} \le & \|e^{W_*(t)}\nabla u(t)\|_{L^2} + \|\nabla W_*(t)X(t)\|_{L^2} \\ \le & \|u(t) - R(t)\|_{H^1} + \|R(t)\|_{H^1} + \|X(t)\|_{L^2} \le C, \end{aligned} \tag{5.16}$$

where C is independent of t, It follows that $\{X(t)\}$ is uniformly bounded in the energy space.

Therefore, plugging (5.16) into (5.15) we obtain (5.14) and finish the proof of Theorem 1.2. $\square$

Acknowledgments

Y. Sun and D. Zhang would like to thank Professor Yingchao Xie for many valuable discussions to improve this paper. We gartefully acknowledge the funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - Project-ID 317210226 - SFB 1283. Y. Su is supported by NSFC grant (No. 12371122). D. Zhang is also grateful for the NSFC grants (No. 12271352, 12322108) and Shanghai Frontiers Science Center of Modern Analysis.

References

[1] O. Bang, P.L. Christiansen, F. If, K.O. Rasmussen, Temperature effects in a nonlinear model of monolayer Scheibe aggregates. *Phys. Rev. E* **49** (1994), 4627-4636.

[2] O. Bang, P.L. Christiansen, F. If, K.O. Rasmussen, Y.B. Gaididei, White noise in the two-dimensional nonlinear Schrödinger equation. *Appl. Anal.* **57** (1995), no. 1-2, 3-15.

[3] V. Barbu, M. Röckner, D. Zhang, Stochastic nonlinear Schrödinger equations with linear multiplicative noise: rescaling approach. *J. Nonlinear Sci.* **24** (2014), no. 3, 383-409.

[4] V. Barbu, M. Röckner, D. Zhang, Stochastic nonlinear Schrödinger equations. *Nonlinear Anal.* **136** (2016), 168-194.

[5] V. Barbu, M. Röckner, D. Zhang, Optimal bilinear control of nonlinear stochastic Schrödinger equations driven by linear multiplicative noise. *Ann. Probab.* **46** (2018), no. 4, 1957-1999.

[6] I. Bejenaru, A.D. Ionescu, C.E. Kenig, D. Tataru, Global Schrödinger maps in dimensions $d \ge 2$: small data in the critical Sobolev spaces. *Ann. of Math. (2)* **173** (2011), no. 3, 1443-1506.

[7] H. Berestycki, P. Lions, Nonlinear scalar field equations. I. Existence of a ground state. *Arch. Rational Mech. Anal.* **82** (1983), no. 4, 313-345.

[8] D. Cao, Y. Su, D. Zhang, On uniqueness of multi-bubble blow-up solutions and multi-solitons to L^2-critical nonlinear Schrödinger equations. *Arch. Ration. Mech. Anal.* **247** (2023), no. 1, Paper No. 4, 81 pp.

[9] T. Cazenave, Semilinear Schrödinger equations. Courant Lecture Notes in Mathematics, 10. New York University, Courant Institute of Mathematical Sciences, New York; American Mathematical Society, Providence, RI, 2003. xiv+323 pp.

[10] V. Combet, Multi-existence of multi-solitons for the supercritical nonlinear Schrödinger equation in one dimension. *Discrete Contin. Dyn. Syst.* **34** (2014), no. 5, 1961-1993.
[11] R. Côte, X. Friederich, On smoothness and uniqueness of multi-solitons of the non-linear Schrödinger equations. *Comm. Partial Differential Equations* **46** (2021), no. 12, 2325-2385.
[12] R. Côte, Y. Martel, F. Merle, Construction of multi-soliton solutions for the L^2-supercritical gKdV and NLS equations. *Rev. Mat. Iberoam.* **27** (2011), no. 1, 273-302.
[13] A. de Bouard, A. Debussche, On the effect of a noise on the solutions of the focusing supercritical nonlinear Schrödinger equation. *Probab. Theory Related Fields* **123** (2002), no. 1, 76-96.
[14] A. de Bouard, A. Debussche, The stochastic nonlinear Schrödinger equation in H^1. *Stochastic Anal. Appl.* **21** (2003), no. 1, 97-126.
[15] A. de Bouard, A. Debussche, Blow-up for the stochastic nonlinear Schrödinger equation with multiplicative noise. *Ann. Probab.* **33** (2005), no. 3, 1078-1110.
[16] A. de Bouard, A. Debussche, Random modulation of solitons for the stochastic Korteweg-de Vries equation. *Ann. Inst. H. Poincaré C Anal. Non Linéaire* **24** (2007), no. 2, 251-278.
[17] A. de Bouard, A. Debussche, Soliton dynamics for the Korteweg-de Vries equation with multiplicative homogeneous noise. *Electron. J. Probab.* **14** (2009), no. 58, 1727-1744.
[18] A. de Bouard, R. Fukuizumi, Modulation analysis for a stochastic NLS equation arising in Bose-Einstein condensation. *Asymptot. Anal.* **63** (2009), no. 4, 189-235.
[19] A. Debussche, L.D. Menza, Numerical simulation of focusing stochastic nonlinear Schrödinger equations. *Phys. D* **162** (2002), no. 3-4, 131-154.
[20] A. Debussche, L.D. Menza, Numerical resolution of stochastic focusing NLS equations. *Appl. Math. Lett.* **15** (2002), no. 6, 661-669.
[21] A. Debussche, E. Gautier, Small noise asymptotic of the timing jitter in soliton transmission. *Ann. Appl. Probab.* **18** (2008), no. 1, 178-208.
[22] T. Duyckaerts, C.E. Kenig, F. Merle, Soliton resolution for the radial critical wave equation in all odd space dimensions. *Acta Math.* **230** (2023), no. 1, 1-92.
[23] T. Duyckaerts, S. Roudenko, Threshold solutions for the focusing 3D cubic Schrödinger equation. *Rev. Mat. Iberoam.* **26** (2010), no. 1, 1-56.
[24] C. Fan, Y. Su, D. Zhang, A note on log-log blow up solutions for stochastic nonlinear Schrödinger equations. *Stoch. Partial Differ. Equ. Anal. Comput.* **10** (2022), no. 4, 1500-1514.
[25] G. Fibich, The nonlinear Schrödinger equation: Singular solutions and optical collapse. *Appl. Math. Sci.* **192** Springer, Cham, 2015. xxxii+862 pp.
[26] P. Friz, M. Hairer, A Course on Rough Paths. With an Introduction to Regularity Structures. *Universitext.* Springer, Cham, 2014. xiv+251 pp.
[27] M. Grillakis, Analysis of the linearization around a critical point of an infinite-dimensional Hamiltonian system. *Comm. Pure Appl. Math.* **43** (1990), no. 3, 299-333.
[28] M. Gubinelli, Controlling rough paths. *J. Funct. Anal.* **216** (2004), no. 1, 86-140.
[29] S. Herr, M. Röckner, M. Spitz, D. Zhang, The energy-critical stochastic Zakharov system. arXiv:2410.05034
[30] S. Herr, M. Röckner, M. Spitz, D. Zhang, The three-dimensional stochastic Zakharov system. *Ann. Probab.* **53** (2025), no. 3, 848-905.
[31] S. Herr, M. Röckner, D. Zhang, Scattering for stochastic nonlinear Schrödinger equations. *Comm. Math. Phys.* **368** (2019), no. 2, 843-884.
[32] J. Jendrej, Construction of two-bubble solutions for the energy-critical NLS. *Anal. PDE* **10** (2017), no. 8, 1923-1959.
[33] J. Jendrej, A. Lawrie, Soliton resolution for the energy-critical nonlinear wave equation in the radial case. *Ann. PDE* **9** (2023), no. 2, Paper No. 18, 117 pp.
[34] Y. Martel, Interaction of solitons from the PDE point of view. *Proceedings of the International Congress of Mathematicians—Rio de Janeiro 2018. Vol. III. Invited lectures*, 2439-2466, World Sci. Publ., Hackensack, NJ, 2018.
[35] Y. Martel, F. Merle, Multi solitary waves for nonlinear Schrödinger equations. *Ann. Inst. H. Poincaré C Anal. Non Linéaire* **23** (2006), no. 6, 849-864.

[36] Y. Martel, F. Merle, T.P. Tsai, Stability in H^1 of the sum of K solitary waves for some nonlinear Schrödinger equations. *Duke Math. J.* **133** (2006), no. 3, 405-466.

[37] F. Merle, Construction of solutions with exactly k blow-up points for the Schrödinger equation with critical nonlinearity. *Comm. Math. Phys.* **129** (1990), no. 2, 223-240.

[38] A. Millet, A.D. Rodriguez, S. Roudenko, K. Yang, Behavior of solutions to the 1D focusing stochastic nonlinear Schrödinger equation with spatially correlated noise. *Stoch. Partial Differ. Equ. Anal. Comput.* **9** (2021), no. 4, 1031-1080.

[39] A. Millet, S. Roudenko, K. Yang, Behaviour of solutions to the 1D focusing stochastic L^2-critical and supercritical nonlinear Schrödinger equation with space-time white noise. *IMA J. Appl. Math.* **86** (2021), no. 6, 1349-1396.

[40] N.T. Vinh, Existence of multi-solitary waves with logarithmic relative distances for the NLS equation. *C. R. Math. Acad. Sci. Paris* **357** (2019), no. 1, 13-58.

[41] K.O. Rasmussen, Y.B. Gaididei, O. Bang, P.L. Chrisiansen, The influence of noise on critical collapse in the nonlinear Schrödinger equation. *Phys. Letters A* **204** (1995), 121-127.

[42] M. Röckner, Y. Su, D. Zhang, Multi solitary waves to stochastic nonlinear Schrödinger equations. *Probab. Theory Related Fields* **186** (2023), no. 3-4, 813-876.

[43] M. Röckner, Y. Su, D. Zhang, Multi-bubble Bourgain-Wang solutions to nonlinear Schrödinger equations. *Trans. Amer. Math. Soc.* **377** (2024), no. 1, 517-588.

[44] W. Schlag, Spectral theory and nonlinear partial differential equations: a survey. *Discrete Contin. Dyn. Syst.* **15** (2006), no. 3, 703-723.

[45] M. Spitz, D. Zhang, Z. Zhao, Regularization by noise for the energy- and mass-critical nonlinear Schrödinger equations. arXiv:2505.05421

[46] Y. Su, D. Zhang, Construction of minimal mass blow-up solutions to rough nonlinear Schrödinger equations. *J. Funct. Anal.* **284** (2023), no. 5, Paper No. 109796, 61 pp.

[47] Y. Su, D. Zhang, On the multi-bubble blow-up solutions to focusing mass-critical stochastic nonlinear Schrödinger equations in dimensions one and two. *Probab. Theory Related Fields* **193** (2025), no 1-2, 163-246.

[48] D.H. Wagner, Survey of measurable selection theorems: an update. *Lecture Notes in Math.* **794** Springer, Berlin, 1980, pp. 176-219.

[49] M. Weinstein, Modulational stability of ground states of nonlinear Schrödinger equations. *SIAM J. Math. Anal.* **16** (1985), no. 3, 472-491.

[50] D. Zhang, Optimal bilinear control of stochastic nonlinear Schrödinger equations: mass-(sub)critical case. *Probab. Theory Related Fields* **178** (2020), no. 1-2, 69-120.

[51] D. Zhang, Strichartz and local smoothing estimates for stochastic dispersive equations with linear multiplicative noise. *SIAM J. Math. Anal.* **54** (2022), no. 6, 5981–6017.

[52] D. Zhang, Stochastic nonlinear Schrödinger equations in the defocusing mass and energy critical cases. *Ann. Appl. Probab.* **33** (2023), no. 5, 3652-3705.

Fakultät für Mathematik, Universität Bielefeld, D-33501 Bielefeld, Germany, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, China.
E-mail address: roeckner@math.uni-bielefeld.de

School of Mathematics, Hangzhou Normal University, 311121 Hangzhou, China.
E-mail address, Yiming Su: yimingsu@amss.ac.cn

School of Mathematical Sciences, Shanghai Jiao Tong University, China.
E-mail address, Yanjun Sun: 0830-syj@sjtu.edu.cn

School of Mathematical Sciences, MOE-LSC, CMA-Shanghai, Shanghai Jiao Tong University, China.
E-mail address, Deng Zhang: dzhang@sjtu.edu.cn