

Growth of Sobolev norms for the quintic NLS on \mathbb{T}^2

E. Haus*, M. Procesi*

* Dipartimento di Matematica, Università di Roma "La Sapienza", Roma, I-00185, Italy

E-mail: haus@mat.uniroma1.it, mprocesi@mat.uniroma1.it

May 12, 2014

Abstract

We study the quintic Non Linear Schrödinger equation on a two dimensional torus and exhibit orbits whose Sobolev norms grow with time. The main point is to reduce to a sufficiently simple toy model, similar in many ways to the one discussed in [15] for the case of the cubic NLS. This requires an accurate combinatorial analysis.

1 Introduction

We consider the quintic defocusing NLS on the two-dimensional torus $\mathbb{T}^2 = \mathbb{R}^2/(2\pi\mathbb{Z})^2$

$$-i\partial_t u + \Delta u = |u|^4 u, \quad (1.1)$$

which is an infinite dimensional dynamical system with Hamiltonian

$$H = \int_{\mathbb{T}^2} |\nabla u|^2 + \frac{1}{3} \int_{\mathbb{T}^2} |u|^6 \quad (1.2)$$

having the mass (the L^2 norm) and the momentum

$$L = \int_{\mathbb{T}^2} |u|^2, \quad M = \int_{\mathbb{T}^2} \Im(u \cdot \nabla u) \quad (1.3)$$

as constants of motion. The well-posedness result of [8, 13] for data $u_0 \in H^s(\mathbb{T}^2)$, $s \geq 1$ gives the existence of a global-in-time smooth solution to (1.1) from smooth initial data and one would like to understand some qualitative properties of solutions.

A fruitful approach to this question is to apply the powerful tools of *singular perturbation theory*, such as KAM theory, Birkhoff Normal Form, Arnold diffusion, first developed in order to study finite-dimensional systems.

We are interested in the phenomenon of the growth of Sobolev norms, i.e. we look for solutions which initially oscillate only on scales comparable to the spatial period and eventually oscillate on arbitrarily short spatial scales. This is a natural extension of the results in [15] and [23] which prove similar results for the cubic NLS. In the strategy of the proof, we follow [15] as closely as possible; therefore our main result is the precise analogue of the one stated in [15] for the cubic NLS. Namely, we prove

Theorem 1.1. *Let $s > 1$, $K \gg 1$ and $0 < \delta \ll 1$ be given parameters. Then there exists a global smooth solution $u(t, x)$ to (1.1) and a time $T > 0$ with*

$$\|u(0)\|_{H^s(\mathbb{T}^2)} \leq \delta \quad \text{and} \quad \|u(T)\|_{H^s(\mathbb{T}^2)} \geq K.$$

Note that we are making no claim regarding the time T over which the growth of Sobolev norms occurs, this is the main difference between the approaches of [15] and [23].

1.1 Some literature

The growth of Sobolev norms for solutions of the Non Linear Schrödinger equation has been studied widely in the literature, but most of the results regard upper bounds on such growth. In the one dimensional case with an analytic non-linearity $\partial_{\bar{u}}P(|u|^2)$ Bourgain [10] and Staffilani [39] proved at most polynomial growth of Sobolev norms. In the same context Bourgain [12] proved a Nekhoroshev type theorem for a perturbation of the cubic NLS. Namely, for s large and a typical initial datum $u(0) \in H^s(T)$ of small size $\|u(0)\|_s \leq \varepsilon$ he proved

$$\sup_{t \leq T} \|u(t)\|_s \leq C\varepsilon, \quad |t| < T, \quad T \leq \varepsilon^{-A}$$

with $A = A(s) \rightarrow 0$ as $s \rightarrow \infty$. Similar upper bounds on the growth have been obtained also for the NLS equation on \mathbb{R} and \mathbb{R}^2 as well as on compact manifolds.

We finally mention the paper [17] which discusses the existence of stability regions for the NLS on tori.

Concerning instability results for the NLS on tori, we mention the papers by Kuksin [29] (see related works [26, 27, 28, 30]) who studied the growth of Sobolev norms for the equation

$$-i\partial_t u + \delta\Delta u = |u|^{2p}u, \quad p \in \mathbb{N}$$

and constructed solutions whose Sobolev norms grow by an inverse power of δ . Note that the solutions that he obtains (for $p = 2$) correspond to orbits of equation (1.1) with large initial data. A big progress appeared in the paper [15] where the authors prove Theorem 1.1 for cubic NLS. Note that the initial data are *small* in H^s . Finally the paper [23] follows the same general strategy of [15] and constructs orbits whose Sobolev norm grows (by an arbitrary factor) in a time which is polynomial in the growth factor. This is done by a careful analysis of the equation and using in a clever way various tools from diffusion in finite dimensional systems.

Note that these results do not imply the existence of solutions with diverging Sobolev norm, nor do they claim that the unstable behavior is typical. Recently, Hani [24] has achieved a remarkable progress towards the existence of unbounded Sobolev orbits: for a class of cubic NLS equations with non-polynomial nonlinearity, the combination of a result like Theorem 1.1 with some clever topological arguments leads to the existence of solutions with diverging Sobolev norm.

Regarding growth of Sobolev norms for other equations we mention the following papers: [10]– for the wave equation with a cubic nonlinearity but with a spectrally defined Laplacian, [19, 32]–for the Szegő equation, and [33]–for certain nonlinear wave equations. We also mention the long time stability results obtained in [1, 2, 3, 4, 20, 21, 40, 42].

A dual point of view to instability is to construct quasi-periodic orbits. These are non-generic solutions which are global in time and whose Sobolev norms are approximately constant. Among the relevant literature we mention [43, 34, 31, 11, 6, 16, 18, 5, 41, 38, 7]. Of particular interest are the recent results obtained through KAM theory which gives information on linear stability close to the quasi-periodic solutions. In particular the paper [37] proves the existence of both stable and unstable tori (of arbitrary finite dimension) for the cubic NLS.

In finite dimensional systems diffusive orbits are usually constructed by proving that the stable and unstable manifolds of a *chain of unstable tori* intersect. Usually this is done with tori of co-dimension one so that the manifolds should intersect for dimensional reasons. Unfortunately in the infinite dimensional case one is not able to prove the existence of co-dimension one tori. Actually the construction of almost-periodic orbits is an open problem except for very special cases such as integrable equations or equations with infinitely many external parameters (see for instance [35, 14, 9]).

In [15], [23] (and the present paper) this problem is avoided by taking advantage of the specific form of the equation. First one reduces to an approximate equation, i.e. the *first order Birkhoff normal form* see (1.5). Then for this dynamical system one proves directly the existence of chains of one dimensional unstable tori (periodic orbits) together with their heteroclinic connections. Next one proves the existence of a *slider solution* which shadows the heteroclinic chain in a finite time. Finally, one proves the persistence of the slider solution for the full NLS. In the next section we describe the strategy more in detail.

1.2 Informal description of the results

In order to understand the dynamics of (1.1) it is convenient to pass to a *moving frame* Fourier series representation:

$$u(t, x) = \sum_{j \in \mathbb{Z}^2} a_j(t) e^{ij \cdot x + i|j|^2 t},$$

so that the equations of motion become

$$-i\dot{a}_j = \sum_{\substack{j_1, j_2, j_3, j_4, j_5 \in \mathbb{Z}^2 \\ j_1 + j_2 + j_3 - j_4 - j_5 = j}} a_{j_1} a_{j_2} a_{j_3} \bar{a}_{j_4} \bar{a}_{j_5} e^{i\omega_6 t} \quad (1.4)$$

where $\omega_6 = |j_1|^2 + |j_2|^2 + |j_3|^2 - |j_4|^2 - |j_5|^2 - |j|^2$.

We define the *resonant truncation* of (1.4) as

$$-i\dot{\beta}_j = \sum_{\substack{j_1, j_2, j_3, j_4, j_5 \in \mathbb{Z}^2 \\ j_1 + j_2 + j_3 - j_4 - j_5 = j \\ |j_1|^2 + |j_2|^2 + |j_3|^2 - |j_4|^2 - |j_5|^2 = |j|^2}} \beta_{j_1} \beta_{j_2} \beta_{j_3} \bar{\beta}_{j_4} \bar{\beta}_{j_5}. \quad (1.5)$$

It is well known that the dynamics of (1.4) is well approximated by the one of (1.5) for finite but long times¹. Our aim is to first prove Theorem 1.1 for (1.5) and then

¹Actually, passing to the resonant truncation is equivalent to performing the first step of a Birkhoff normal form. However, since we follow closely the proof in [15], we chose to use similar notation.

extend the result to (1.4) by an approximation Lemma. The idea of the approximation Lemma roughly speaking is that, by integrating in time the l.h.s. of (1.4), one sees that the non-resonant terms (i.e., with $\omega_6 \neq 0$) give a contribution of order $O(a^9)$. By scaling $a^{(\lambda)}(t) = \lambda^{-1}a(\lambda^{-4}t)$, we see that the non-resonant terms are an arbitrarily small perturbation w.r.t. the resonant terms appearing in (1.5) and hence they can be ignored for arbitrarily long finite times.

We now outline the strategy used in order to prove Theorem 1.1 for the equations (1.5). The equations (1.5) are Hamiltonian with respect to the Hamiltonian function

$$\mathcal{H} = \frac{1}{3} \sum_{\substack{j_1, j_2, j_3, j_4, j_5, j_6 \in \mathbb{Z}^2 \\ j_1 + j_2 + j_3 = j_4 + j_5 + j_6 \\ |j_1|^2 + |j_2|^2 + |j_3|^2 = |j_4|^2 + |j_5|^2 + |j_6|^2}} \beta_{j_1} \beta_{j_2} \beta_{j_3} \bar{\beta}_{j_4} \bar{\beta}_{j_5} \bar{\beta}_{j_6} \quad (1.6)$$

and the symplectic form $\Omega = id\beta \wedge d\bar{\beta}$.

This is still a very complicated (infinite-dimensional) Hamiltonian system, but it has the advantage of having many invariant subspaces on which the dynamics simplifies significantly. Let us set up some notation.

Definition 1.1 (Resonance). *A sextuple $(k_1, k_2, k_3, k_4, k_5, k_6) \in (\mathbb{Z}^2)^6$, is a resonance if*

$$k_1 + k_2 + k_3 - k_4 - k_5 - k_6 = 0, \quad |k_1|^2 + |k_2|^2 + |k_3|^2 - |k_4|^2 - |k_5|^2 - |k_6|^2 = 0. \quad (1.7)$$

A resonance is trivial if it is of the form $(k_1, k_2, k_3, k_1, k_2, k_3)$ up to permutations of the last three elements.

Definition 1.2 (Completeness). *We say that a set $\mathcal{S} \subset \mathbb{Z}^2$ is complete if the following holds:*

for every quintuple $(k_1, k_2, k_3, k_4, k_5) \in \mathcal{S}^5$, if there exists $k_6 \in \mathbb{Z}^2$ s.t. $(k_1, k_2, k_3, k_4, k_5, k_6)$ is a resonance, then $k_6 \in \mathcal{S}$.

It is easily seen that, for any complete $\mathcal{S} \subset \mathbb{Z}^2$, the subspace defined by requiring $\beta_k = 0$ for all $k \notin \mathcal{S}$ is invariant.

Definition 1.3 (Action preserving). *A complete set $\mathcal{S} \subset \mathbb{Z}^2$ is said to be action preserving if all the resonances in \mathcal{S} are trivial.*

Remark that, for any complete and action preserving $\mathcal{S} \subset \mathbb{Z}^2$, the Hamiltonian restricted to \mathcal{S} is given by (see [36])

$$\mathcal{H}|_{\mathcal{S}} = \frac{1}{3} \left(\sum_{j \in \mathcal{S}} |\beta_j|^6 + 9 \sum_{\substack{j, k \in \mathcal{S} \\ j \neq k}} |\beta_j|^4 |\beta_k|^2 + 36 \sum_{\substack{j, k, m \in \mathcal{S} \\ j \prec k \prec m}} |\beta_j|^2 |\beta_k|^2 |\beta_m|^2 \right) \quad (1.8)$$

where \prec is any fixed total ordering of \mathbb{Z}^2 .

If \mathcal{S} is complete and action preserving, then $\mathcal{H}|_{\mathcal{S}}$ is function of the actions $|\beta_j|^2$ only, with non-vanishing twist: therefore, the corresponding motion is periodic, quasi-periodic or almost-periodic, depending on the initial data. In particular, if $\beta_j(0) = \beta_k(0)$ for all

$j, k \in \mathcal{S}$, then the motion is periodic. Finally, since all the actions are constants of motion, then so are the H^s -norms of the solution.

On the other hand, it is easy to give examples of sets \mathcal{S} that are complete but not action preserving. For instance, one can consider complete sets of the form $\mathfrak{S}^{(1)} = \{k_1, k_2, k_3, k_4\}$, where the k_j 's are the vertices of a non-degenerate rectangle in \mathbb{Z}^2 , or of the form $\mathfrak{S}^{(2)} = \{k_1, k_2, k_3, k_4, k_5, k_6\}$, where the $k_j \in \mathbb{Z}^2$ are all distinct and satisfy equations (1.7). Other examples are sets of the form $\mathfrak{S}^{(3)} = \{k_1, k_2, k_3, k_4\}$, with

$$k_1 + 2k_2 - 2k_3 - k_4 = 0, \quad |k_1|^2 + 2|k_2|^2 - 2|k_3|^2 - |k_4|^2 = 0 \quad (1.9)$$

studied in [22] or, more in general, the sets $\mathfrak{S}^{(4)} = \cup_j \mathfrak{S}_j^{(3)}$ studied in [25]². In all these cases, the variation of the H^s -norm of the solution is of order $O(1)$. Note that, while sets of the form $\mathfrak{S}^{(2)}, \mathfrak{S}^{(3)}, \mathfrak{S}^{(4)}$ exist in \mathbb{Z}^d for all d , the non-degenerate rectangles $\mathfrak{S}^{(1)}$ exist only in dimension $d \geq 2$. Let us briefly describe the dynamics on these sets. By writing the Hamiltonian in symplectic polar coordinates $\beta_j = \sqrt{I_j} e^{i\theta_j}$, one sees that all these systems are integrable. However, their phase portraits are quite different. In $\mathfrak{S}^{(1)}$ one can exhibit two periodic orbits $\mathbb{T}_1, \mathbb{T}_2$ that are linked by a heteroclinic connection. \mathbb{T}_1 is supported on the modes k_1, k_2 and \mathbb{T}_2 on k_3, k_4 . The H^s -norm of each periodic orbit is constant in time. By choosing $\mathfrak{S}^{(1)}$ appropriately one can ensure that these two values are different and this produces a growth of the Sobolev norms. Moreover, all the energy is transferred from \mathbb{T}_1 to \mathbb{T}_2 . In the other cases, i.e. $\mathfrak{S}^{(2)}, \mathfrak{S}^{(3)}, \mathfrak{S}^{(4)}$, there is no orbit transferring all the energy from some modes to others (see Appendix C).

These heteroclinic connections are the key to the energy transfer. In fact, assume that

$$\mathcal{S}_1 := \{\mathbf{v}_1, \dots, \mathbf{v}_n\}, \quad \mathcal{S}_2 := \{\mathbf{w}_1, \dots, \mathbf{w}_n\}$$

with n even, are two complete and action preserving sets. Assume moreover that, for all $1 \leq j \leq n/2$, $\{\mathbf{v}_{2j-1}, \mathbf{v}_{2j}, \mathbf{w}_{2j-1}, \mathbf{w}_{2j}\}$ are the vertices of a rectangle as in $\mathfrak{S}^{(1)}$. Finally, assume that $\mathcal{S}_1 \cup \mathcal{S}_2$ is complete and contains no non-trivial resonances except those of the form $(k, \mathbf{v}_{2j-1}, \mathbf{v}_{2j}, k, \mathbf{w}_{2j-1}, \mathbf{w}_{2j})$. As in the case of $\mathfrak{S}^{(1)}$ the periodic orbits:

$$\mathbb{T}_1: \quad \beta_{\mathbf{v}_j}(t) = b_1(t) \neq 0, \quad \beta_{\mathbf{w}_j}(t) = 0, \quad \forall j = 1, \dots, n$$

and

$$\mathbb{T}_2: \quad \beta_{\mathbf{w}_j}(t) = b_2(t) \neq 0, \quad \beta_{\mathbf{v}_j}(t) = 0, \quad \forall j = 1, \dots, n$$

are linked by a heteroclinic connection.

We iterate this procedure constructing a *generation set* $\mathcal{S} = \cup_{i=1}^N \mathcal{S}_i$ where each \mathcal{S}_i is complete and action preserving. The corresponding periodic orbit \mathbb{T}_i is linked by heteroclinic connections to \mathbb{T}_{i-1} and \mathbb{T}_{i+1} . There are two delicate points:

- (i) at each step, when adding a new generation \mathcal{S}_i , we need to ensure that the resulting generation set is still complete and contains no non-trivial resonances except for the prescribed ones. Such resonances are those of the form $(k, v_1, v_2, k, v_3, v_4)$ where $v_1, v_2 \in \mathcal{S}_i, v_3, v_4 \in \mathcal{S}_{i+1}$ for some $1 \leq i \leq N - 1$ and $\{v_1, v_2, v_3, v_4\}$ are the vertices of a rectangle.

²The papers [22, 25] actually consider the one-dimensional case, but of course the construction of complete sets can always be trivially extended to higher dimensions.

- (ii) we need to ensure that the Sobolev norms grow by an arbitrarily large factor K/δ , which requires taking n (the number of elements in each \mathcal{S}_j) and N (the number of generations) large.

The point (i) is a question of combinatorics. It requires some careful classification of the possible resonances and it turns out to be significantly more complicated than the cubic case. We discuss this in subsection 3.2.

The point (ii) is treated exactly in the same way as in [15], we discuss it for completeness in subsection 3.1, Remark 3.1.

Given a generation set \mathcal{S} as above we proceed in the following way: first we restrict to the finite-dimensional invariant subspace where $\beta_k = 0$ for all $k \notin \mathcal{S}$. To further simplify the dynamics we restrict to the invariant subspace:

$$\beta_v(t) = b_i(t), \quad \forall v \in \mathcal{S}_i, \quad \forall i = 1, \dots, n$$

this is the so called *toy model*. Note that the periodic solutions \mathbb{T}_i live in this subspace. The toy model is a Hamiltonian system, with Hamiltonian given by (2.2) and with the constant of motion $J = \sum_{i=1}^N |b_i|^2$. We work on the sphere $J = 1$, which contains all the \mathbb{T}_i with action $|b_i|^2 = 1$.

As discussed above, we construct a chain of heteroclinic connections going from \mathbb{T}_1 to \mathbb{T}_N . Then, we prove (see Proposition 2.1) the existence of a *slider solution* which “shadows” this chain, starting at time 0 from a neighborhood of \mathbb{T}_3 and³ ending at time T in a neighborhood of \mathbb{T}_{N-2} .

We proceed as follows: first, we perform a symplectic reduction that will allow us to study the local dynamics close to the periodic orbit \mathbb{T}_j , which puts the Hamiltonian in form (2.5). The new variables c_k are the ones obtained by synchronizing the b_k ($k \neq j$) with the phase of b_j . Then, we diagonalize the linear part of the vector field associated to (2.5). In particular, the eigenvalues are the Lyapunov exponents of the periodic orbit \mathbb{T}_j . As for the cubic case, one obtains that all the eigenvalues are purely imaginary, except four of them which, due to the symmetries of the problem, are of the form $\lambda, \lambda, -\lambda, -\lambda \in \mathbb{R}$. Note that these hyperbolic directions are directly related to the heteroclinic connections connecting \mathbb{T}_j to \mathbb{T}_{j-1} and to \mathbb{T}_{j+1} . It turns out that the heteroclinic connections are straight lines in the variables c_k . The equations of motion for the reduced system have the form (2.9) (which is very similar to the cubic case): this is crucial in order to be able to apply almost verbatim the proof given in [15]. Note that it is not obvious *a priori* that equations (2.9) hold true: for instance, this turns out to be false for the NLS of degree ≥ 7 .

The strategy of the proof, which is exactly the same as in [15], consists substantially of two parts:

- studying the linear dynamics close to \mathbb{T}_j , treating the non-linear terms as a small perturbation: one needs to prove that the flow associated to equations (2.9) maps

³One could ask why we construct a slider solution diffusing from the third mode b_3 to the third-to-last mode b_{N-2} , instead of diffusing from the first mode b_1 to the last mode b_N . The reason is that, since we rely on the proof given in [15], our statement is identical to the ones in Proposition 2.2 and Theorem 3.1 in [15]. As in [15], also in our case it would be possible to diffuse from the first to the last mode just by overcoming some very small notational issues.

points close to the *incoming* heteroclinic connection (from \mathbb{T}_{j-1}) to points close to the *outcoming* heteroclinic connection (towards \mathbb{T}_{j+1}) (note that, in order to take advantage of the linear dynamics close to \mathbb{T}_j , we need that almost all the energy is concentrated on \mathcal{S}_j);

- following closely the heteroclinic connection in order to flow from a neighborhood of \mathbb{T}_j to a neighborhood of \mathbb{T}_{j+1} .

The precise statement of these two facts requires the introduction of the notions of *targets* and *covering* and is summarized in Proposition 2.3. The main analytical tool for the proof are repeated applications of Gronwall's lemma. Our proof of Proposition 2.3 follows almost verbatim the proof of the analogue statement, given in Section 3 of [15]. However, the only way to check that the proof works also in our case, is to go through the whole proof in [15], which is rather long and technical, and make the needed adaptations. Therefore, for the convenience of the reader, in Appendix A we give a summary of the proof of Proposition 2.3, highlighting the points where there are significant differences with [15].

1.3 Comparison with the cubic case and higher order NLS equations

In the cubic NLS, the only resonant sets of frequencies are rectangles, which makes completely natural the choice of using rectangles as building blocks of the generation set \mathcal{S} . In the quintic and higher degree NLS many more resonant sets appear, which a priori gives much more freedom in the construction of \mathcal{S} . In particular, in the quintic case sets of the form $\mathfrak{S}^{(2)}$ are the most generic resonant sets, and therefore it would look reasonable to use them as building blocks. However (see Appendix C), such a choice does not allow full energy transfer from a generation to the next one and is therefore incompatible with our strategy. The same happens if one uses sets of the form $\mathfrak{S}^{(3)}$. This leads us to use rectangles for the construction of \mathcal{S} also in the quintic case.

It is worth remarking that, while non-degenerate rectangles do not exist in one space dimension, sets of the form $\mathfrak{S}^{(2)}$, $\mathfrak{S}^{(3)}$ already exist in one dimension. The equations of the toy model only depend on the combinatorics of the set \mathcal{S} . Therefore, if one were able to prove diffusion in a toy model built with resonant sets of the form $\mathfrak{S}^{(2)}$, $\mathfrak{S}^{(3)}$ (or other resonant sets that exist already in one dimension), then one could hope to prove the same type of result for some one-dimensional (non-cubic) NLS.

The use of rectangles as building blocks for the generation set of a quintic or higher order NLS makes things more complicated, since the rectangles induce many different resonant sets, see section 2. This leads to combinatorial problems which make it harder to prove the non-degeneracy and completeness of \mathcal{S} . The equations of the toy model also have a more complicated form than in the cubic case. Since this type of difficulties grows with the degree, dealing with the general case will most probably require some careful – and possibly complicated – combinatorics and one cannot expect to have a completely explicit formula for the toy model Hamiltonian of any degree.

In the quintic case the formula is explicit and relatively simple and we can explicitly perform the symmetry reduction. After some work, we still get equations of the form (2.9), which resembles the cubic case with some relevant differences: here the Lyapunov

exponent λ depends on n and tends to infinity as $n \rightarrow \infty$; moreover, the non-linear part of the vector field associated to (2.5) is not homogeneous in the variables c_k , as it contains both terms of order 3 and 5 (in the cubic case, it is homogeneous of order 3).

For the NLS of higher degree, not only the reduced Hamiltonian gets essentially unmanageable, but there also appears a further difficulty. Already for the NLS of degree 7, a toy model built using rectangles (after symplectic reduction and diagonalization) does not satisfy equations like (2.9), meaning that the heteroclinic connections are not straight lines. Such a problem can be probably overcome, but this requires a significant adaptation of the analytical techniques used in order to prove the existence of the slider solution (work in progress with M. Guardia).

1.4 Plan of the paper

In Section 2 We assume to have a generation set $\mathcal{S} = \cup_{i=1}^N \mathcal{S}_i$ which satisfies all the needed non-degeneracy properties and deduce the form of the toy model Hamiltonian. Then we study this Hamiltonian and prove the existence of slider solutions.

In Section 3 we prove the existence of non-degenerate generation sets such that the corresponding slider solution undergoes the required growth of Sobolev norms.

In Section 4 we prove, via the approximation Lemma 4.1 and a scaling argument, the persistence of solutions with growing Sobolev norm for the full NLS.

Since some of the proofs follow very closely the ones in [15], we move them to appendix.

2 The toy model

We now define a finite subset $\mathcal{S} = \cup_{i=1}^N \mathcal{S}_i \subset \mathbb{Z}^2$ which satisfies appropriate *non-degeneracy conditions* (Definition 2.5) as explained in the introduction. In the following we assume that such a set exists. This is not obvious and will be discussed in section 3.2.

For reasons that will be clear, and following [15], the \mathcal{S}_i 's will be called *generations*. In order to describe the resonances which connect different generations we introduce some notation.

Definition 2.1 (Family). *A family (of age $i \in \{1, \dots, N-1\}$) is a list $(v_1, v_2; v_3, v_4)$ of elements of \mathcal{S} such that the points form the vertices of a non-degenerate rectangle*

$$v_1 + v_2 = v_3 + v_4, \quad |v_1|^2 + |v_2|^2 = |v_3|^2 + |v_4|^2$$

and such that one has $v_1, v_2 \in \mathcal{S}_i$ and $v_3, v_4 \in \mathcal{S}_{i+1}$. Whenever $(v_1, v_2; v_3, v_4)$ form a family, we say that v_1, v_2 are the parents of v_3, v_4 and that v_3, v_4 are the children of v_1, v_2 . Moreover, we say that v_1 is the spouse of v_2 (and vice versa) and that v_3 is the sibling of v_4 (and vice versa). We denote (for instance) $v_1 = v_3^{par1}$, $v_2 = v_3^{par2}$, $v_1 = v_2^{sp}$, $v_4 = v_3^{sib}$, $v_3 = v_1^{ch1}$, $v_4 = v_1^{ch2}$.

Remark 2.1. *If $(v_1, v_2; v_3, v_4)$ is a family of age i , then the same holds for its trivial permutations $(v_2, v_1; v_3, v_4)$, $(v_1, v_2; v_4, v_3)$ and $(v_2, v_1; v_4, v_3)$.*

Definition 2.2. An integer vector $\lambda \in \mathbb{Z}^{|\mathcal{S}|}$ such that

$$\sum_i \lambda_i = 0, \quad |\lambda| := \sum_i |\lambda_i| \leq 6$$

is resonant for \mathcal{S} if

$$\sum_i \lambda_i v_i = 0, \quad \sum_i \lambda_i |v_i|^2 = 0$$

Note that to a family $\mathcal{F} = (v_1, v_2; v_3, v_4)$ we associate a special resonant vector $\lambda^{\mathcal{F}}$ with $|\lambda| = 4$, through $\sum_i \lambda_i^{\mathcal{F}} v_i = v_1 + v_2 - v_3 - v_4$. Similarly, to the couple of parents in the family \mathcal{F} we associate the vector $\lambda^{\mathcal{F}_p}$ through $\sum_i \lambda_i^{\mathcal{F}_p} v_i = v_1 + v_2$ and to the couple of children we associate $\lambda^{\mathcal{F}_c}$ through $\sum_i \lambda_i^{\mathcal{F}_c} v_i = v_3 + v_4$, so that $\lambda^{\mathcal{F}} = \lambda^{\mathcal{F}_p} - \lambda^{\mathcal{F}_c}$.

Definition 2.3 (Generation set). The set \mathcal{S} is said to be a generation set if it satisfies the following:

1. For all $i \in \{1, \dots, N-1\}$, every $v \in \mathcal{S}_i$ is a member of one and only one (up to trivial permutations) family of age i . We denote such a family by \mathcal{F}^v . (Note that $\mathcal{F}^v = \mathcal{F}^w$ if $v = w^{sp}$.)
2. For all $i \in \{2, \dots, N\}$, every $v \in \mathcal{S}_i$ is a member of one and only one (up to trivial permutations) family of age $i-1$. We denote such a family by \mathcal{F}_v . (Note that $\mathcal{F}_v = \mathcal{F}_w$ if $v = w^{sib}$.)
3. For all $v \in \cup_{i=2}^{N-1} \mathcal{S}_i$, one has $v^{sp} \neq v^{sib}$.

Remark 2.2. The vectors $\lambda^{\mathcal{F}}$ corresponding to the families of a generation set are linearly independent.

Note that, whenever two families \mathcal{F}_1 and \mathcal{F}_2 have a common member (which must be a child in one family and a parent in the other one), then $\lambda^{\mathcal{F}_1} + \lambda^{\mathcal{F}_2}$ is a non-trivial resonant vector whose support has cardinality exactly 6. This motivates the following definition:

Definition 2.4 (Resonant vector of type CF). A resonant vector λ is said to be of type CF (couple of families) if there exist two families $\mathcal{F}_1 \neq \mathcal{F}_2$ such that $\lambda = \pm(\lambda^{\mathcal{F}_1} + \lambda^{\mathcal{F}_2})$. (Note that, since $|\lambda| \leq 6$, the two families $\mathcal{F}_1, \mathcal{F}_2$ must have a common member.)

We say that a generation set is non-degenerate if the following condition is fulfilled.

Definition 2.5 (Non-degeneracy). Suppose that there exists $\lambda \in \mathbb{Z}^{|\mathcal{S}|}$, with $\sum_i \lambda_i = 1$ and $|\lambda| \leq 5$, such that

$$\sum_i \lambda_i |v_i|^2 - \left| \sum_i \lambda_i v_i \right|^2 = 0.$$

Then only four possibilities are allowed:

1. $|\lambda| = 1$.
2. $|\lambda| = 3$ and the support of λ consists exactly of three distinct elements of the same family and the two λ_i 's appearing with a positive sign correspond either to the two parents or to the two children of the family.

3. $|\lambda| = 5$ and there exist a family \mathcal{F} and an element $v \in \mathcal{S}$ such that $\lambda = \pm\lambda^{\mathcal{F}} + v$.
4. $|\lambda| = 5$ and there exists $v \in \mathcal{S}$ such that $\lambda - v$ is a resonant vector of type CF.

Note that, if \mathcal{S} is a non-degenerate generation set and λ is a resonant vector, then either $\lambda = \pm\lambda^{\mathcal{F}}$ for some family \mathcal{F} or λ is a resonant vector of type CF.

In what follows we will assume that \mathcal{S} is a non-degenerate generation set. This implies that \mathcal{S} is complete and all the subsets \mathcal{S}_i are pairwise disjoint, complete and action preserving. Finally the only resonances which appear are those induced by the family relations. Then, the Hamiltonian restricted to \mathcal{S} is

$$\begin{aligned}
\mathcal{H}|_{\mathcal{S}} = & \frac{1}{3} \left(\sum_{j \in \mathcal{S}} |\beta_j|^6 + 9 \sum_{\substack{j,k \in \mathcal{S} \\ j \neq k}} |\beta_j|^4 |\beta_k|^2 + 36 \sum_{\substack{j,k,m \in \mathcal{S} \\ j < k < m}} |\beta_j|^2 |\beta_k|^2 |\beta_m|^2 \right) + \quad (2.1) \\
& + 3 \sum_{i=1}^{N-1} \sum_{j \in \mathcal{S}_i} (\beta_j \beta_{j^{sp}} \bar{\beta}_{j^{ch_1}} \bar{\beta}_{j^{ch_2}} + \bar{\beta}_j \bar{\beta}_{j^{sp}} \beta_{j^{ch_1}} \beta_{j^{ch_2}}) \left(2 \sum_{\substack{k \in \mathcal{S} \\ k \notin \mathcal{F}^j}} |\beta_k|^2 + \sum_{m \in \mathcal{F}^j} |\beta_m|^2 \right) + \\
& + 12 \sum_{i=2}^{N-1} \sum_{j \in \mathcal{S}_i} \left(\beta_{j^{par_1}} \beta_{j^{par_2}} \beta_{j^{sp}} \bar{\beta}_{j^{sib}} \bar{\beta}_{j^{ch_1}} \bar{\beta}_{j^{ch_2}} + \beta_{j^{ch_1}} \beta_{j^{ch_2}} \beta_{j^{sib}} \bar{\beta}_{j^{sp}} \bar{\beta}_{j^{par_1}} \bar{\beta}_{j^{par_2}} \right).
\end{aligned}$$

We restrict to the invariant subspace $\mathcal{D} \subset \mathcal{S}$ where $\beta_k = b_i$ for all $k \in \mathcal{S}_i$ and $\forall i = 1, \dots, N$. Denote by n (which must be an even integer number) the cardinality of each generation. Following the construction in [15], one has $n = 2^{N-1}$. A straightforward computation (involving some easy combinatorics) of the Hamiltonian yields

$$\begin{aligned}
\frac{3}{n} \mathcal{H}|_{\mathcal{D}} = & \sum_{k=1}^N |b_k|^6 + 9 \left[(n-1) \sum_{k=1}^N |b_k|^6 + n \sum_{\substack{k,\ell=1 \\ k \neq \ell}}^N |b_k|^4 |b_\ell|^2 \right] + \\
& + 6 \left[(n-1)(n-2) \sum_{k=1}^N |b_k|^6 + 3n(n-1) \sum_{\substack{k,\ell=1 \\ k \neq \ell}}^N |b_k|^4 |b_\ell|^2 \right] + \\
& + 36n^2 \sum_{\substack{k,\ell,m=1 \\ k < \ell < m}}^N |b_k|^2 |b_\ell|^2 |b_m|^2 + \\
& + 18 \sum_{k=1}^{N-1} \left(-|b_k|^2 - |b_{k+1}|^2 + n \sum_{\ell=1}^N |b_\ell|^2 \right) (b_k^2 \bar{b}_{k+1}^2 + b_{k+1}^2 \bar{b}_k^2) + \\
& + 36 \sum_{k=2}^{N-1} |b_k|^2 (b_{k-1}^2 \bar{b}_{k+1}^2 + b_{k+1}^2 \bar{b}_{k-1}^2)
\end{aligned}$$

The equations of motion for the toy model can be deduced by considering the effective Hamiltonian $h(b, \bar{b}) := \frac{\mathcal{H}|_{\mathcal{D}(b, \bar{b})}}{n}$, endowed with the symplectic form $\Omega = idb \wedge d\bar{b}$.

Due to the conservation of the total mass L , the quantity

$$J := \frac{L}{n} = \sum_{k=1}^N |b_k|^2$$

is a constant of motion.

Evidencing the dependence on J , we get

$$\begin{aligned} 3h &= 4 \sum_{k=1}^N |b_k|^6 - 9n \sum_{h=1}^N |b_h|^2 \left[\sum_{k=1}^N |b_k|^4 - 2 \sum_{k=1}^{N-1} (b_k^2 \bar{b}_{k+1}^2 + b_{k+1}^2 \bar{b}_k^2) \right] \\ &+ 18 \sum_{k=1}^{N-1} (-|b_k|^2 - |b_{k+1}|^2) (b_k^2 \bar{b}_{k+1}^2 + b_{k+1}^2 \bar{b}_k^2) + \\ &+ 36 \sum_{k=2}^{N-1} |b_k|^2 (b_{k-1}^2 \bar{b}_{k+1}^2 + b_{k+1}^2 \bar{b}_{k-1}^2). \end{aligned} \quad (2.2)$$

2.1 Invariant subspaces

Since J is a constant of motion the dynamics is confined to its level sets. For simplicity, we will restrict to $J = 1$, i.e. to

$$\Sigma := \{b \in \mathbb{C}^N : \sum_{k=1}^N |b_k|^2 = 1\}.$$

All the monomials in the toy model Hamiltonian have even degree in each of the modes (b_j, \bar{b}_j) , which implies that

$$\text{Supp}(b) := \{1 \leq j \leq N | b_j \neq 0\}$$

is invariant in time. This automatically produces many invariant subspaces some of which will play a specially important role, namely: (i) the subspaces M_j corresponding to $\text{Supp}(b) = \{j\}$ for some $1 \leq j \leq N$. In this case the dynamics is confined to the circle $|b_j|^2 = J$, with

$$b_j(t) = \sqrt{J} \exp \left[-i \left(3n - \frac{4}{3} \right) J^2 t \right]. \quad (2.3)$$

The intersection of M_j with Σ is a single periodic orbit, which we denote by \mathbb{T}_j .

(ii) The subspaces generated by M_j and M_{j+1} (corresponding to $\text{Supp}(b) = \{j, j+1\}$) for some $1 \leq j \leq N-1$. Here the Hamiltonian becomes

$$\begin{aligned} 3h_{2g} &= 4 (|b_j|^6 + |b_{j+1}|^6) - 9n (|b_j|^2 + |b_{j+1}|^2) [|b_j|^4 + |b_{j+1}|^4 - 2(b_j^2 \bar{b}_{j+1}^2 + b_{j+1}^2 \bar{b}_j^2)] \\ &- 18 (|b_j|^2 + |b_{j+1}|^2) (b_j^2 \bar{b}_{j+1}^2 + b_{j+1}^2 \bar{b}_j^2) \end{aligned} \quad (2.4)$$

Passing to symplectic polar coordinates:

$$b_j = \sqrt{I_1} e^{i\theta_1}, \quad b_{j+1} = \sqrt{I_2} e^{i\theta_2},$$

we have

$$3h_{2g} = (4 - 9n)(I_1 + I_2)^3 + 6(I_1 + I_2)I_1I_2(3n - 2 + 6(n - 1)\cos(2(\theta_1 - \theta_2))) ,$$

since $J = I_1 + I_2$ is a conserved quantity the dynamics is integrable and easy to study.

We pass to the symplectic variables

$$J, I_1, \theta_2, \varphi = \theta_2 - \theta_1$$

and obtain the Hamiltonian

$$3h_{2g} = (4 - 9n)J^3 + 6JI_1(J - I_1)(3n - 2 + 6(n - 1)\cos(2\varphi)) .$$

The phase portrait (ignoring the evolution of the cyclic variable θ_2) restricted to Σ is described in Figure 1.

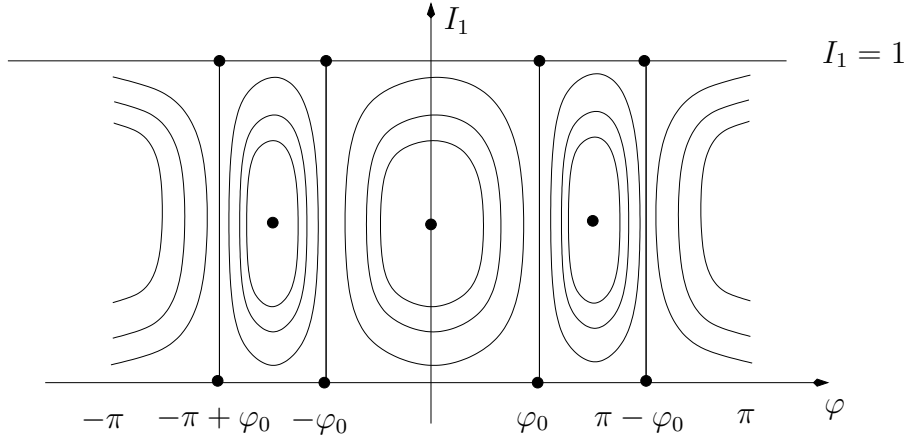


Figure 1: Phase portrait of the two generation Hamiltonian h_{2g} on Σ .

Remark 2.3. The coordinates I_1, φ and the domain given by the cylinder $(\varphi, I_1) \in \mathbb{S}^1 \times [0, 1]$ are singular since the angle $\varphi = \theta_2 - \theta_1$ is ill-defined when $I_1 = 0$ or $I_1 = 1$. In the correct picture for the reduced dynamics, each of the lines $I_1 = 0$ and $I_1 = 1$ should be shrunk to a single point, thus obtaining (topologically) a two dimensional sphere (see Figure 2).

This can also be seen in the following way. The level set $J = 1$ is a three dimensional sphere \mathbb{S}^3 , with the gauge symmetry group \mathbb{S}^1 acting freely on it. Due to the Hopf fibration, the topology of the quotient space is \mathbb{S}^2 .

Note that, as for the case of the cubic NLS (see [15, 23]), there exist heteroclinic connections linking \mathbb{T}_j to \mathbb{T}_{j+1} . Again as in the cubic case the orbits have fixed angle

$$\varphi(t) = \varphi_0 = \frac{1}{2} \arccos\left(-\frac{3n-2}{6(n-1)}\right), \quad I_1(t) = \frac{e^{2\lambda t}}{1 + e^{2\lambda t}}$$

where $\lambda = 2\sqrt{(9n-8)(3n-4)}$. Our aim will be to construct slider solutions that are very concentrated on the mode b_3 at the time $t = 0$ and very concentrated on the mode b_{N-2} at the time $t = T$. These solutions will start very close to the periodic orbit \mathbb{T}_3 and then use the heteroclinic connections in order to slide from \mathbb{T}_3 to \mathbb{T}_4 and so on until \mathbb{T}_{N-2} .

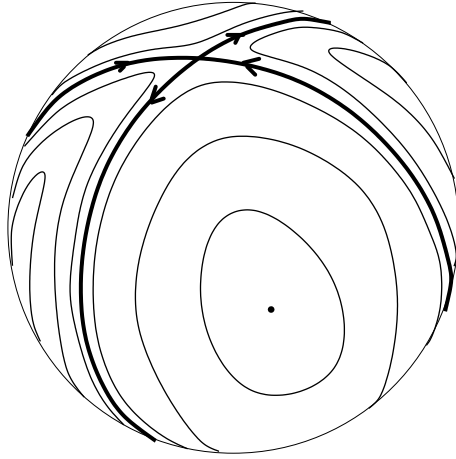


Figure 2: A sketch of the phase portrait of the two generation Hamiltonian h_{2g} on Σ in the correct topology.

2.2 Symplectic reduction

Now, since we are interested in studying the dynamics close to the j -th periodic orbit \mathbb{T}_j , we introduce a set of coordinates which are in phase with it and give a symplectic reduction with respect to the constant of motion J . This procedure is the same that was carried out, for the cubic NLS, in [23] and, substantially, already in [15].

Let $\vartheta^{(j)}$ be the phase of the complex number b_j . Then, for $k \neq j$, let $c_k^{(j)}$ the variable obtained by conjugating b_k with the phase $\vartheta^{(j)}$, i.e.

$$c_k^{(j)} = b_k e^{-i\vartheta^{(j)}} .$$

Then, the change of coordinates (well defined on $\{b_j \neq 0\}$) given by

$$\begin{aligned} & (b_1, \dots, b_N, \bar{b}_1, \dots, \bar{b}_N) \mapsto \\ & \mapsto (c_1^{(j)}, \dots, c_{j-1}^{(j)}, J, c_{j+1}^{(j)}, \dots, c_N^{(j)}, \bar{c}_1^{(j)}, \dots, \bar{c}_{j-1}^{(j)}, \vartheta^{(j)}, \bar{c}_{j+1}^{(j)}, \dots, \bar{c}_N^{(j)}) \end{aligned}$$

is symplectic. Namely, in the new coordinates the symplectic form is given by

$$\Omega = i dc^{(j)} \wedge d\bar{c}^{(j)} + dJ \wedge d\vartheta^{(j)} .$$

Then, we rewrite the Hamiltonian h in terms of the new coordinates (from now on, in order to simplify the notation, we will omit the superscript (j) in the $c^{(j)}$ variables and in their complex conjugates $\bar{c}^{(j)}$ and in the phase $\vartheta^{(j)}$). Thus, we get the expression

$$\begin{aligned}
3h &= 4 \sum_{k \neq j} |c_k|^6 - 4 \left(\sum_{k \neq j} |c_k|^2 \right)^3 + (18n - 12)J^2 \sum_{k \neq j} |c_k|^2 - 9nJ \sum_{k \neq j} |c_k|^4 + \\
&- (9n - 12)J \left(\sum_{k \neq j} |c_k|^2 \right)^2 + 18 \sum_{\substack{k=1 \\ k \neq j-1, j}}^{N-1} (-|c_k|^2 - |c_{k+1}|^2 + nJ) (c_k^2 \bar{c}_{k+1}^2 + c_{k+1}^2 \bar{c}_k^2) + \\
&+ 18 \left[\sum_{\substack{k=1 \\ k \neq j-1, j}}^N |c_k|^2 + (n-1)J \right] \left(J - \sum_{\substack{\ell=1 \\ \ell \neq j}}^N |c_\ell|^2 \right) (c_{j-1}^2 + \bar{c}_{j-1}^2) + \\
&+ 18 \left[\sum_{\substack{k=1 \\ k \neq j, j+1}}^N |c_k|^2 + (n-1)J \right] \left(J - \sum_{\substack{\ell=1 \\ \ell \neq j}}^N |c_\ell|^2 \right) (c_{j+1}^2 + \bar{c}_{j+1}^2) + \tag{2.5} \\
&+ 36 \sum_{\substack{k=2 \\ k \neq j-1, j, j+1}}^{N-1} |c_k|^2 (c_{k-1}^2 \bar{c}_{k+1}^2 + c_{k+1}^2 \bar{c}_{k-1}^2) + 36 |c_{j-1}|^2 \left(J - \sum_{\substack{k=1 \\ k \neq j}}^N |c_k|^2 \right) (c_{j-2}^2 + \bar{c}_{j-2}^2) + \\
&+ 36 \left(J - \sum_{\substack{k=1 \\ k \neq j}}^N |c_k|^2 \right) (c_{j-1}^2 \bar{c}_{j+1}^2 + c_{j+1}^2 \bar{c}_{j-1}^2) + 36 |c_{j+1}|^2 \left(J - \sum_{\substack{k=1 \\ k \neq j}}^N |c_k|^2 \right) (c_{j+2}^2 + \bar{c}_{j+2}^2) .
\end{aligned}$$

Observe that the Hamiltonian h does not depend on ϑ . Since J is a constant of motion, the terms depending only on J can be erased from the Hamiltonian. Up to those constant terms, one has

$$h = h_2 + r_4 , \tag{2.6}$$

where h_2 is the part of order 2 in (c, \bar{c}) (which corresponds to the linear part of the vector field) and r_4 is of order at least 4 in (c, \bar{c}) . By an explicit computation, one obtains

$$h_2 = 2J^2 \left[(3n - 2) \sum_{\substack{k=1 \\ k \neq j}}^N |c_k|^2 + 3(n - 1)(c_{j-1}^2 + \bar{c}_{j-1}^2 + c_{j+1}^2 + \bar{c}_{j+1}^2) \right] . \tag{2.7}$$

It is easily seen that the dynamics associated to the vector field generated by h_2 is elliptic in the modes c_k with $1 \leq k \leq j - 2$ or $j + 2 \leq k \leq N$, while it is hyperbolic in the modes c_{j-1} and c_{j+1} . In order to put in evidence the hyperbolic dynamics, we perform a change of coordinates which diagonalizes the linear part of the vector field. Namely, for $k = j - 1, j + 1$, we set

$$\begin{aligned}
c_k &= \frac{1}{\sqrt{2\Im(\omega^2)}} (\bar{\omega} c_k^- + \omega c_k^+) \\
\bar{c}_k &= \frac{1}{\sqrt{2\Im(\omega^2)}} (\omega c_k^- + \bar{\omega} c_k^+)
\end{aligned}$$

where $\omega = e^{i\varphi_0}$ with

$$\varphi_0 = \frac{1}{2} \arccos \left(-\frac{3n-2}{6(n-1)} \right).$$

Note that this change of variables affects only the hyperbolic modes, which are expressed in terms of the new variables $(c_{j-1}^+, c_{j-1}^-, c_{j+1}^+, c_{j+1}^-)$. This transformation is symplectic, writing h_2 as a function of the new variables we get

$$h_2 = 2J^2 \left[(3n-2) \sum_{\substack{k=1 \\ k \neq j-1, j, j+1}}^N |c_k|^2 + \sqrt{(9n-8)(3n-4)} (c_{j-1}^+ c_{j-1}^- + c_{j+1}^+ c_{j+1}^-) \right]. \quad (2.8)$$

We have proved that the periodic orbit (2.3) is hyperbolic and we have explicitly written the quadratic part of the Hamiltonian in the local variables. Similarly to the case of the cubic NLS these local variables are actually well adapted to describing also the global dynamics connecting two periodic orbits, as discussed in the previous section.

To this purpose we study the integrable *two generation* Hamiltonian (2.4) after all the changes of variables described in this section, i.e. in the variables c_{j+1}^+, c_{j+1}^- . Direct substitution shows that the Hamiltonian is given by

$$h_{2g} = 2J\sqrt{(9n-8)(3n-4)} c_{j+1}^+ c_{j+1}^- \left\{ J - \frac{1}{2\Im(\omega^2)} [(c_{j+1}^+)^2 + (c_{j+1}^-)^2 + 2\Re(\omega^2) c_{j+1}^+ c_{j+1}^-] \right\}.$$

It is important to note that all the monomials in h_{2g} contain both c_{j+1}^+ and c_{j+1}^- , so the subspaces $c_{j+1}^+ = 0$ and $c_{j+1}^- = 0$ (which correspond to the heteroclinic connections) are invariant for the 2-generation dynamics. It is useful to denote by $c^* = \{c_h\}_{h \neq j-1, j, j+1}$ so that the dynamical variables of the Hamiltonian (2.5) become $(c_{j-1}^+, c_{j-1}^-, c_{j+1}^+, c_{j+1}^-, c^*, \bar{c}^*)$.

Now, since

$$h_{2g} = h|_{c_{j-1}^+ = c_{j-1}^- = q_1 = 0, c^* = 0},$$

exploiting also the symmetry between (c_{j-1}^+, c_{j-1}^-) and (c_{j+1}^+, c_{j+1}^-) , this implies that also in h none of the monomials in $(c_{j-1}^+, c_{j-1}^-, c_{j+1}^+, c_{j+1}^-, c^*, \bar{c}^*)$ depends only on one of the variables $c_{j-1}^+, c_{j-1}^-, c_{j+1}^+, c_{j+1}^-$.

Finally, we recall that all the monomials in $h(c_{j-1}^+, c_{j-1}^-, c_{j+1}^+, c_{j+1}^-, c^*)$ have even degree in each of the couples (c_k^*, \bar{c}_k^*) and in both couples (c_k^+, c_k^-) .

From these observations, and from the bound $O(c^2) \lesssim J = O(1)$, we immediately deduce the following relations about the Hamilton equations associated to h :

$$\begin{aligned} \dot{c}_{j-1}^- &= -2J^2 \sqrt{(9n-8)(3n-4)} c_{j-1}^- + O(c^2 c_{j-1}^-) + O(c_{\neq j-1}^2 c_{j-1}^+) & (2.9) \\ \dot{c}_{j-1}^+ &= 2J^2 \sqrt{(9n-8)(3n-4)} c_{j-1}^+ + O(c^2 c_{j-1}^+) + O(c_{\neq j-1}^2 c_{j-1}^-) \\ \dot{c}_{j+1}^- &= -2J^2 \sqrt{(9n-8)(3n-4)} c_{j+1}^- + O(c^2 c_{j+1}^-) + O(c_{\neq j+1}^2 c_{j+1}^+) \\ \dot{c}_{j+1}^+ &= 2J^2 \sqrt{(9n-8)(3n-4)} c_{j+1}^+ + O(c^2 c_{j+1}^+) + O(c_{\neq j+1}^2 c_{j+1}^-) \\ \dot{c}^* &= 2J^2(3n+2)ic^* + O(c^2 c^*), \end{aligned}$$

where we denote $c = (c_{j-1}^+, c_{j-1}^-, c_{j+1}^+, c_{j+1}^-, c^*)$, $c_{\neq j-1} = (c_{j+1}^+, c_{j+1}^-, c^*)$, $c_{\neq j+1} = (c_{j-1}^+, c_{j-1}^-, c^*)$. These relations are the precise analogue of Proposition 3.1 in [15], where the factor $2J^2 \sqrt{(9n-8)(3n-4)}$ here replaces the factor $\sqrt{3}$ in [15].

From the equations of motion (2.9), we deduce that

$$-i\dot{c}_{j+1} = \frac{\partial h_{2g}}{\partial \bar{c}_{j+1}} + O(c_{j+1}c_{\neq j+1}^2).$$

We have

$$h_{2g} = 2J\sqrt{(9n-8)(3n-4)}c_{j+1}^+c_{j+1}^-(J-|c_{j+1}|^2)$$

where c_{j+1}^+, c_{j+1}^- can be thought of as functions of (c_{j+1}, \bar{c}_{j+1}) . Then

$$\dot{c}_{j+1} = 2iJ\sqrt{(9n-8)(3n-4)}(J-|c_{j+1}|^2)\frac{\partial(c_{j+1}^+c_{j+1}^-)}{\partial \bar{c}_{j+1}} + O(c_{j+1}c_{j+1}^+c_{j+1}^-) + O(c_{j+1}c_{\neq j+1}^2).$$

We compute

$$2i\frac{\partial(c_{j+1}^+c_{j+1}^-)}{\partial \bar{c}_{j+1}} = \sqrt{\frac{2}{\Im(\omega^2)}}(\omega c_{j+1}^+ - \bar{\omega}c_{j+1}^-)$$

from which we deduce

$$\dot{c}_{j+1} = J\sqrt{\frac{2(9n-8)(3n-4)}{\Im(\omega^2)}}(\omega c_{j+1}^+ - \bar{\omega}c_{j+1}^-)(J-|c_{j+1}|^2) + O(c_{j+1}c_{j+1}^+c_{j+1}^-) + O(c_{j+1}c_{\neq j+1}^2) \quad (2.10)$$

which is the analogue for c_{j+1} of equation (3.19) in [15]. In the same way, one deduces

$$\dot{c}_{j-1} = J\sqrt{\frac{2(9n-8)(3n-4)}{\Im(\omega^2)}}(\omega c_{j-1}^+ - \bar{\omega}c_{j-1}^-)(J-|c_{j-1}|^2) + O(c_{j-1}c_{j-1}^+c_{j-1}^-) + O(c_{j-1}c_{\neq j-1}^2) \quad (2.11)$$

which is the analogue of equation (3.19) in [15] for the evolution of c_{j-1} .

2.3 Existence of a “slider solution”

In this section, we are going to prove the following proposition (which is the analogue of Proposition 2.2 in [15]), that establishes the existence of a slider solution.

Proposition 2.1. *For all $\epsilon > 0$ and $N \geq 6$, there exist a time $T_0 > 0$ and an orbit of the toy model such that*

$$\begin{aligned} |b_3(0)| &\geq 1 - \epsilon, & |b_j(0)| &\leq \epsilon & j &\neq 3 \\ |b_{N-2}(T_0)| &\geq 1 - \epsilon, & |b_j(T_0)| &\leq \epsilon & j &\neq N-2. \end{aligned}$$

Furthermore, one has $\|b(t)\|_{\ell^\infty} \sim 1$ for all $t \in [0, T_0]$.

More precisely there exists a point x_3 within $O(\epsilon)$ of \mathbb{T}_3 (using the usual metric on Σ), a point x_{N-2} within $O(\epsilon)$ of \mathbb{T}_{N-2} and a time $T_0 \geq 0$ such that $S(T_0)x_3 = x_{N-2}$, where $S(t)x$ is the dynamics at time t of the toy model Hamiltonian with initial datum x .

In order to prove Proposition 2.1 of our paper, we completely rely on the proof of the analogue Proposition 2.2 in [15]. In order to keep our notations as close as possible to those of [15] we rescale the time $t = 2\sqrt{(9n-8)(n-4/3)}\tau$ in our toy model, this means rescaling $h \rightarrow \sqrt{3}h/2\sqrt{(9n-8)(3n-4)}$, where h is defined in (2.2), so that the Lyapunov exponents of the linear dynamics are $\sqrt{3}$. We hence prove Proposition 2.1 for the rescaled toy model. By formulæ(2.9), (2.10), (2.11) we have the analogue of Proposition 3.1 and of eq. (3.19) of [15].

Proposition 2.2. *Let $3 \leq j \leq N-2$ and let $b(\tau)$ be a solution of the rescaled toy model living on Σ and with $b_j(\tau) \neq 0$. We have the system of equations:*

$$\dot{c}_{j-1}^- = -\sqrt{3}c_{j-1}^- + O(c^2c_{j-1}^-) + O(c_{\neq j-1}^2c_{j-1}^+) \quad (2.12a)$$

$$\dot{c}_{j-1}^+ = \sqrt{3}c_{j-1}^+ + O(c^2c_{j-1}^+) + O(c_{\neq j-1}^2c_{j-1}^-) \quad (2.12b)$$

$$\dot{c}_{j+1}^- = -\sqrt{3}c_{j+1}^- + O(c^2c_{j+1}^-) + O(c_{\neq j+1}^2c_{j+1}^+) \quad (2.12c)$$

$$\dot{c}_{j+1}^+ = \sqrt{3}c_{j+1}^+ + O(c^2c_{j+1}^+) + O(c_{\neq j+1}^2c_{j+1}^-) \quad (2.12d)$$

$$\dot{c}^* = i\kappa c^* + O(c^2c^*), \quad \kappa = \frac{\sqrt{3}(3n-2)}{\sqrt{(9n-8)(3n-4)}} \quad (2.12e)$$

Moreover,

$$\dot{c}_{j+1} = \sqrt{\frac{3}{2\Im(\omega^2)}}(\omega c_{j+1}^+ - \bar{\omega}c_{j+1}^-)(J - |c_{j+1}|^2) + O(c_{j+1}c_{j+1}^+c_{j+1}^-) + O(c_{j+1}c_{\neq j+1}^2) \quad (2.13)$$

and

$$\dot{c}_{j-1} = \sqrt{\frac{3}{2\Im(\omega^2)}}(\omega c_{j-1}^+ - \bar{\omega}c_{j-1}^-)(J - |c_{j-1}|^2) + O(c_{j-1}c_{j-1}^+c_{j-1}^-) + O(c_{j-1}c_{\neq j-1}^2) \quad (2.14)$$

Finally, since the equations (2.12) come from the Hamiltonian (2.5) which is an even polynomial of degree six, one has that all the symbols $O(c^3)$ are actually⁴ $O(c^3) + O(c^5)$. For instance

$$O(c^2c_{j-1}^-) = O(c^2c_{j-1}^-) + O(c^4c_{j-1}^-), \quad O(c_{\neq j-1}^2c_{j-1}^+) = O(c_{\neq j-1}^2c_{j-1}^+) + O(c^2c_{\neq j-1}^2c_{j-1}^+) \quad (2.15)$$

Note that the only difference with [15] is that our remainder terms (of type $O(c^2c_{j-1}^-)$, $O(c_{\neq j-1}^2c_{j-1}^+)$, etc.) are not homogeneous of degree three but have also a term of degree five (which is completely irrelevant in the analysis).

We now introduce some definitions and notations of [15].

Definition 2.6 (Targets). *A target is a triple (M, d, R) , where M is a subset of Σ , d is a semi-metric on Σ and $R > 0$ is a radius. We say that a point $x \in \Sigma$ is within a*

⁴ as in [15], we use the *schematic notation* $\mathcal{O}(\cdot)$. The symbol $\mathcal{O}(y)$ indicates a linear combination of terms that resemble y up to the presence of multiplicative constants and complex conjugations. So for instance a term like $2i\bar{c}_{j+1}|c_{j+2}|^2c_{j+3}^2 - 3c_{j+1}|c_{j+2}|^4$ is of the form $\mathcal{O}(c^5)$ and more precisely $\mathcal{O}(c_{j+1}c_{\neq j+1}^4)$

target (M, d, R) if we have $d(x, y) < R$ for some $y \in M$. Given two points $x, y \in \Sigma$, we say that x hits y , and write $x \mapsto y$, if we have $y = S(t)x$ for some $t \geq 0$. Given an initial target (M_1, d_1, R_1) and a final target (M_2, d_2, R_2) , we say that (M_1, d_1, R_1) can cover (M_2, d_2, R_2) , and write $(M_1, d_1, R_1) \rightarrow (M_2, d_2, R_2)$, if for every $x_2 \in M_2$ there exists an $x_1 \in M_1$, such that for any point $y_1 \in \Sigma$ with $d(x_1, y_1) < R_1$ there exists a point $y_2 \in \Sigma$ with $d_2(x_2, y_2) < R_2$ such that y_1 hits y_2 .

We refer the reader to pages 64-66 of [15] for a presentation of the main properties of targets.

We need a number of parameters. First, an increasing set of exponents

$$1 \ll A_3^0 \ll A_3^+ \ll A_4^- \ll \cdots \ll A_{N-2}^- \ll A_{N-2}^0.$$

For sake of concreteness, we will take these to be consecutive powers of 10. Next, we shall need a small parameter $0 < \sigma \ll 1$ depending on N and the exponents A which basically measures the distance to \mathbb{T}_j at which the quadratic Hamiltonian dominates the quartic terms. Then we need a set of scale parameters

$$1 \ll r_{N-2}^0 \ll r_{N-2}^- \ll r_{N-2}^+ \ll r_{N-3}^- \ll \cdots \ll r_3^+ \ll r_3^0$$

where each parameter is assumed to be sufficiently large depending on the preceding parameters and on σ and the A 's. These parameters represent a certain shrinking of each target from the previous one (in order to guarantee that each target can be covered by the previous). Finally, we need a very large time parameter $T \gg 1$ that we shall assume to be as large as necessary depending on all the previous parameters.

Setting

$$\{c_1, \dots, c_h\} := c_{\leq h}, \quad \{c_h, c_{h+1}, \dots, c_N\} = c_{\geq h}$$

we call $c_{\leq j-1}$ the *trailing modes*, $c_{\geq j+1}$ the *leading modes*, $c_{\leq j-2}$ the *trailing peripheral modes* and finally $c_{\geq j+2}$ the *leading peripheral modes*. We construct a series of targets:

- An incoming target (M_j^-, d_j^-, R_j^-) (located near the stable manifold of \mathbb{T}_j) defined as follows:

M_j^- is the subset of Σ where

$$c_{\leq j-2}, c_{j-1}^+ = 0, \quad c_{j-1}^- = \sigma, \quad |c_{\geq j+1}| \leq r_j^- e^{-2\sqrt{3}T},$$

$R_j^- = T^{A_j^-}$ and the semi-metric is

$$d_j^-(x, \tilde{x}) := e^{2\sqrt{3}T} |c_{\leq j-2} - \tilde{c}_{\leq j-2}| + e^{\sqrt{3}T} |c_{j-1}^- - \tilde{c}_{j-1}^-| + e^{4\sqrt{3}T} |c_{j-1}^+ + \tilde{c}_{j-1}^+| + e^{3\sqrt{3}T} |c_{\geq j+1} - \tilde{c}_{\geq j+1}|$$

- A ricochet target (M_j^0, d_j^0, R_j^0) (located very near \mathbb{T}_j itself), defined as follows:

M_j^0 is the subset of Σ where

$$c_{\leq j-1}, c_{j+1}^- = 0, \quad |c_{j+1}^+| \leq r_j^0 e^{-\sqrt{3}T}, \quad |c_{\geq j+2}| \leq r_j^0 e^{-2\sqrt{3}T},$$

$R_j^- = T^{A_j^0}$ and the semi-metric is

$$d_j^0(x, \tilde{x}) := e^{2\sqrt{3}T} \left(|c_{\leq j-2} - \tilde{c}_{\leq j-2}| + |c_{j+1}^+ + \tilde{c}_{j+1}^+| \right) + e^{\sqrt{3}T} |c_{j-1}^- - \tilde{c}_{j-1}^-| + e^{3\sqrt{3}T} \left(|c_{j-1}^+ + \tilde{c}_{j-1}^+| + |c_{j+1}^- + \tilde{c}_{j+1}^-| + |c_{\geq j+2} - \tilde{c}_{\geq j+2}| \right)$$

- An outgoing target (M_j^+, d_j^+, R_j^+) (located near the unstable manifold of \mathbb{T}_j) defined as follows:

M_j^+ is the subset of Σ where

$$c_{\leq j-1}, c_{j+1}^- = 0, \quad c_{j+1}^+ = \sigma, \quad |c_{\geq j+2}| \leq r_j^+ e^{-2\sqrt{3}T},$$

$R_j^- = T^{A_j^+}$ and the semi-metric is

$$d_j^+(x, \tilde{x}) := e^{2\sqrt{3}T} |c_{\leq j-1} - \tilde{c}_{\leq j-1}| + e^{4\sqrt{3}T} |c_{j+1}^- - \tilde{c}_{j+1}^-| + e^{\sqrt{3}T} |c_{j+1}^+ + \tilde{c}_{j+1}^+| + e^{3\sqrt{3}T} |c_{\geq j+2} - \tilde{c}_{\geq j+2}|$$

By Section 3.5 of [15] Proposition 2.1 follows from

Proposition 2.3. $(M_j^-, d_j^-, R_j^-) \rightarrow (M_j^0, d_j^0, R_j^0)$ for all $3 < j \leq N-2$, $(M_j^0, d_j^0, R_j^0) \rightarrow (M_j^+, d_j^+, R_j^+)$ for all $3 \leq j < N-2$, $(M_j^+, d_j^+, R_j^+) \rightarrow (M_{j+1}^-, d_{j+1}^-, R_{j+1}^-)$ for all $3 \leq j < N-2$.

Proof. See Appendix A. □

Proof of Proposition 2.1. By [15] Lemma 3.1 we deduce the covering relations

$$(M_3^0, d_3^0, R_3^0) \rightarrow (M_{N-2}^0, d_{N-2}^0, R_{N-2}^0), \quad (2.16)$$

in turn this implies that there is at least one solution $b(t)$ to (3.1) which starts within the ricochet target (M_3^0, d_3^0, R_3^0) at some time t_0 and ends up within the ricochet target $(M_{N-2}^0, d_{N-2}^0, R_{N-2}^0)$ at some later time $t_1 > t_0$. But from the definition of these targets, we thus see that $b(t_0)$ lies within a distance $O(r_3^0 e^{-\sqrt{3}T})$ of \mathbb{T}_3 , while $b(t_1)$ lies within a distance $O(r_{N-2}^0 e^{-\sqrt{3}T})$ of \mathbb{T}_{N-2} . The claim follows. □

3 Construction of the set \mathcal{S}

3.1 The density argument and the norm explosion property

The perturbative argument for the construction of the frequency set \mathcal{S} works exactly as in [15], Section 4. However, for the convenience of the reader, we recall here the main points.

A convenient way to construct a generation set is to first fix a “genealogical tree”, i.e. an abstract combinatorial model of the parenthood and brotherhood relations, and then to choose a placement function, embedding this abstract combinatorial model in \mathbb{R}^2 . Our choice of the abstract combinatorial model is the one described in [15] pp. 99-100. Then, once the combinatorial model is fixed, the choice of the embedding in \mathbb{R}^2 is equivalent to the choice of the following free parameters:

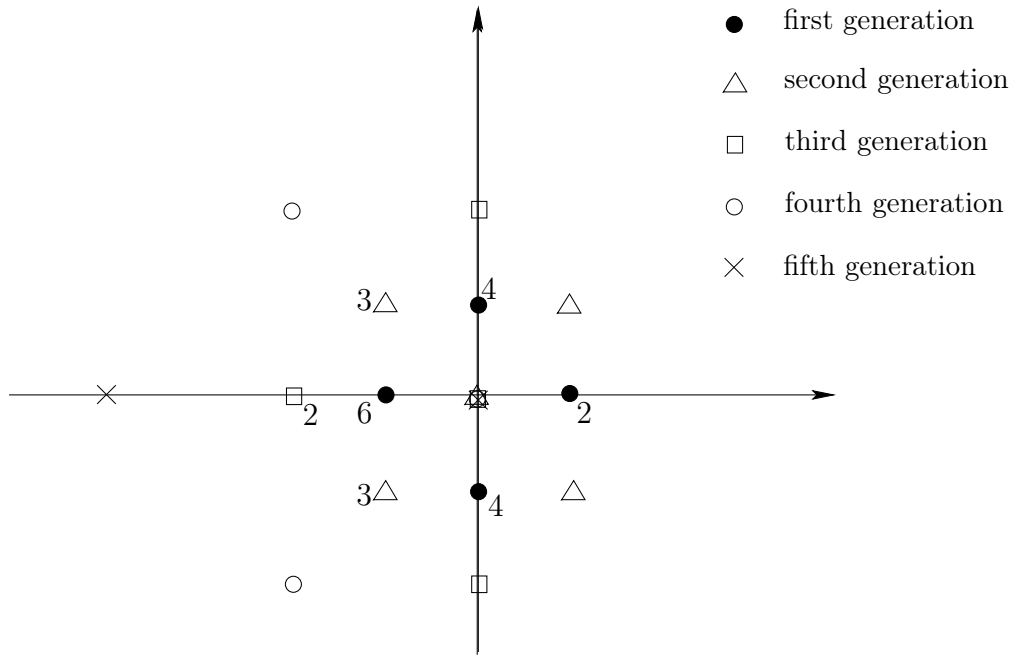


Figure 3: The prototype embedding with five generations. Note that this is a highly degenerate realization of the abstract combinatorial model of [15]. Since $N = 5$, each generation contains 16 points; we have explicitly written the multiplicity of each point when it is not one. In zero there are: 0 points of the first generation, 8 points of the second, 12 of the third, 14 of the fourth and 15 points of the fifth generation.

- the placement of the first generation \mathcal{S}_1 (which implies the choice of a parameter in \mathbb{R}^{2^N});
- the choice of a procreation angle $\vartheta^{\mathcal{F}}$ for each family of the generation set (which globally implies the choice of a parameter in $\mathbb{T}^{(N-1)2^{N-2}}$, since $(N-1)2^{N-2}$ is the number of families).

We denote the corresponding generation set by $\mathcal{S}(\mathcal{S}_1, \vartheta^{\mathcal{F}})$ and the space of parameters by $\mathcal{X} := \mathbb{R}^{2^N} \times \mathbb{T}^{(N-1)2^{N-2}}$.

In Section 3.2 we will prove that the set of parameters producing *degenerate* generation sets is contained in a closed set of null measure in \mathcal{X} .

We claim that the set of $(\mathcal{S}_1, \vartheta^{\mathcal{F}}) \in \mathcal{X}$ such that $\mathcal{S}(\mathcal{S}_1, \vartheta^{\mathcal{F}}) \subset \mathbb{Q}^2 \setminus \{0\}$ is dense in \mathcal{X} . This is a consequence of two facts:

- the density of $\mathbb{Q}^2 \setminus \{0\}$ in \mathbb{R}^2 (for the placement of the first generation);
- the density of (non-zero) rational points on circles having a diameter with rational endpoints.

These two remarks imply that the set of $(\mathcal{S}_1, \vartheta^{\mathcal{F}}) \in \mathcal{X}$ such that $\mathcal{S}(\mathcal{S}_1, \vartheta^{\mathcal{F}})$ is non-degenerate and $\mathcal{S}(\mathcal{S}_1, \vartheta^{\mathcal{F}}) \subset \mathbb{Q}^2 \setminus \{0\}$ is dense in \mathcal{X} .

In order to prove the growth of Sobolev norms, we require a further property on the generation set \mathcal{S} , i.e. the *norm explosion property*

$$\sum_{k \in \mathcal{S}_{N-2}} |k|^{2s} > \frac{1}{2} 2^{(s-1)(N-5)} \sum_{k \in \mathcal{S}_3} |k|^{2s} . \quad (3.1)$$

Given $N \gg 1$, our aim is to prove the existence of a non degenerate generation set $\mathcal{S} \subset \mathbb{Q}^2 \setminus \{0\}$ satisfying (3.1). The fact that (3.1) is an open condition on the space of parameters \mathcal{X} , together with the above remarks, implies that it is enough to prove the existence of a (possibly degenerate) generation set $\mathcal{S} \subset \mathbb{R}^2$ satisfying (3.1), which is achieved by the *prototype embedding* described in [15], pp. 101–102 (see Figure 3).

Remark 3.1. *Note that, for any given positive integer ℓ , the function $F : \mathbb{S}^{\ell-1} \rightarrow \mathbb{R}$, where*

$$\mathbb{S}^{\ell-1} = \left\{ (x_1, \dots, x_\ell) \in \mathbb{R}^\ell \left| \sum_{i=1}^{\ell} x_i^2 = 1 \right. \right\} ,$$

defined by

$$F(x_1, \dots, x_\ell) = \sum_{i=1}^{\ell} x_i^{2s}$$

attains its minimum (since $s > 1$) at

$$(x_1, \dots, x_\ell) = (\ell^{-1/2}, \dots, \ell^{-1/2})$$

and its maximum at

$$(x_1, x_2, \dots, x_\ell) = (1, 0, \dots, 0) .$$

From this one deduces that for each family \mathcal{F} with parents v_1, v_2 and children v_3, v_4 one must have

$$\frac{|v_3|^{2s} + |v_4|^{2s}}{|v_1|^{2s} + |v_2|^{2s}} \leq 2^{s-1}$$

and therefore, for all $1 \leq i \leq N-1$,

$$\frac{\sum_{k \in \mathcal{S}_{i+1}} |k|^{2s}}{\sum_{k \in \mathcal{S}_i} |k|^{2s}} \leq 2^{s-1} .$$

which implies

$$\frac{\sum_{k \in \mathcal{S}_j} |k|^{2s}}{\sum_{k \in \mathcal{S}_i} |k|^{2s}} \leq 2^{(s-1)(j-i)} .$$

for all $1 \leq i \leq j \leq N$. This means that we have to choose N large if we want the ratio

$$\frac{\sum_{k \in \mathcal{S}_{N-2}} |k|^{2s}}{\sum_{k \in \mathcal{S}_3} |k|^{2s}}$$

to be large.

Moreover, since

$$F(\ell^{-1/2}, \dots, \ell^{-1/2}) = \ell^{-s+1} , \quad F(1, 0, \dots, 0) = 1$$

we have for all $1 \leq i, j \leq N$

$$\frac{\sum_{k \in \mathcal{S}_j} |k|^{2s}}{\sum_{k \in \mathcal{S}_i} |k|^{2s}} \leq n^{s-1} .$$

which implies that also n has to be chosen large enough.

In this sense the prototype embedding and the choice $n = 2^{N-1}$ are optimal, because they attain the maximum possible growth of the quantity $\sum_{k \in \mathcal{S}_i} |k|^{2s}$ both at each step and between the first and the last generation.

Then, once we are given a non-degenerate generation set contained in $\mathbb{Q}^2 \setminus \{0\}$ and satisfying (3.1), it is enough to multiply by any integer multiple of the least common denominator of its elements in order to get a non-degenerate generation set $\mathcal{S} \in \mathbb{Z}^2 \setminus \{0\}$ and satisfying (3.1) (note that (3.1) is invariant by dilations of the set \mathcal{S}). Note that we can dilate \mathcal{S} as much as we wish, so we can make $\min_{k \in \mathcal{S}} |k|$ as large as desired.

These considerations are summarized by the following proposition (the analogue of Proposition 2.1 in [15]).

Proposition 3.1. *For all $K, \delta, \mathcal{R} > 0$, there exist $N \gg 1$ and a non-degenerate generation set $\mathcal{S} \subset \mathbb{Z}^2$ such that*

$$\frac{\sum_{k \in \mathcal{S}_{N-2}} |k|^{2s}}{\sum_{k \in \mathcal{S}_3} |k|^{2s}} \gtrsim \frac{K^2}{\delta^2} \quad (3.2)$$

and such that

$$\min_{k \in \mathcal{S}} |k| \geq \mathcal{R} . \quad (3.3)$$

3.2 Genericity of non-degenerate generation sets

The main result of this section is the following.

Proposition 3.2. *There exists a closed set of zero measure $\mathcal{D} \subset \mathcal{X}$ such that the generation set $\mathcal{S}(\mathcal{S}_1, \vartheta^{\mathcal{F}})$ is non-degenerate for all $(\mathcal{S}_1, \vartheta^{\mathcal{F}}) \in \mathcal{X} \setminus \mathcal{D}$.*

First, we need a lemma ensuring that any linear relation among the elements of the generation set that is not a linear combination of the family relations is generically not fulfilled.

Lemma 3.1. *Let $\mu \in \mathbb{Z}^{N2^{N-1}}$, $i = 1, \dots, M$ be an integer vector, linearly independent from the subspace of $\mathbb{R}^{N2^{N-1}}$ generated by all the vectors $\lambda^{\mathcal{F}}$ associated to the families. Then, there for an open set of full measure $\mathfrak{S} \subset \mathcal{X}$, one has that if $(\mathcal{S}_1, \vartheta^{\mathcal{F}}) \in \mathfrak{S}$, then $\mathcal{S}(\mathcal{S}_1, \vartheta^{\mathcal{F}})$ is such that*

$$\sum_{j=1}^{N2^{N-1}} \mu_j v_j \neq 0 . \quad (3.4)$$

Proof. We denote the elements of \mathcal{S} by $v_1, \dots, v_{|\mathcal{S}|}$, with $|\mathcal{S}| = N2^{N-1}$. For simplicity and without loss of generality, we order the v_j 's so that couples of siblings always have consecutive subindices.

For each family \mathcal{F} , both the linear and the quadratic relations

$$\sum_j \lambda_j^{\mathcal{F}} v_j = 0, \quad \sum_j \lambda_j^{\mathcal{F}} |v_j|^2 = 0$$

are satisfied. The coefficients of the linear relations can be collected in a matrix $\Lambda^{\mathcal{F}}$ with $(N-1)2^{N-2}$ rows (as many as the number of families) and $N2^{N-1}$ columns (as many as the elements of \mathcal{S}), so that the linear relations become

$$\Lambda^{\mathcal{F}} v = 0.$$

We choose to order the rows of $\Lambda^{\mathcal{F}}$ so that the matrix is in *lower row echelon* form (see figure).

$$\begin{array}{cccccccccccc} w_1 & w_2 & w_3 & w_4 & w_5 & p_5 & w_6 & p_6 & w_7 & p_7 & w_8 & p_8 \\ \hline 1 & 1 & 0 & 0 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & -1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & -1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & -1 & -1 \end{array}$$

Each row of a matrix in lower row echelon form has a *pivot*, i.e. the first nonzero coefficient of the row starting from the right. Being in lower row echelon form means that the pivot of a row is always strictly to the right of the pivot of the row above it. In the matrix $\Lambda^{\mathcal{F}}$, the pivots are all equal to -1 and they correspond to one and only one of the child from each family. In order to use this fact, we accordingly rename the elements of the generation set by writing $v = (p, w) \in \mathbb{R}^{2a} \times \mathbb{R}^{2b}$, with $a = (N-1)2^{N-2}$, $b = N2^{N-1} - a = (N+1)2^{N-2}$, where the $p_j \in \mathbb{R}^2$ are the elements of the generation set corresponding to the pivots and the $w_\ell \in \mathbb{R}^2$ are all the others, i.e. all the elements of the first generation and one and only one child (the non-pivot) from each family. Here, the index ℓ ranges from 1 to b , while the index j ranges from $2^{N-1} + 1$ to b (note that $a + 2^{N-1} = b$), so that a couple (p_j, w_ℓ) corresponds to a couple of siblings if and only if $j = \ell$. Then, the linear relations $\Lambda^{\mathcal{F}} v = 0$ can be used to write each p_j as a linear combination of the w_ℓ 's with $\ell \leq j$ only:

$$p_j = \sum_{\ell \leq j} \eta_\ell w_\ell, \quad \eta_\ell \in \mathbb{Q}. \quad (3.5)$$

Finally, the quadratic relations $\Lambda^{\mathcal{F}} |v|^2 = 0$ constrain each w_ℓ with $\ell > 2^{N-1}$ (i.e. not in the first generation) to a circle depending on the w_j with $j < \ell$; note that this circle has positive radius provided that the parents of w_ℓ are distinct. Then, eq. (3.5) implies that the l.h.s. of (3.4) can be rewritten in a unique way as a linear combination of the w_ℓ 's only, so we have

$$\sum_{\ell=1}^b \nu_\ell w_\ell = 0. \quad (3.6)$$

Hence, the assumption that μ is linearly independent from the space generated by the $\lambda^{\mathcal{F}}$'s is equivalent to the fact that $\nu \in \mathbb{R}^{2b}$ does not vanish.

Now, let

$$\bar{\ell} := \max \{ \ell \mid \nu_\ell \neq 0 \}$$

so that (3.6) is equivalent to

$$w_{\bar{\ell}} = -\frac{1}{\nu_{\bar{\ell}}} \sum_{\ell < \bar{\ell}} \nu_\ell w_\ell \quad (3.7)$$

If $\bar{\ell} \leq 2^{N-1}$, then $w_{\bar{\ell}}$ is in the first generation. Since there are no restrictions (either linear or quadratic) on the first generation, the statement is trivial. Hence assume $\bar{\ell} > 2^{N-1}$. We can assume (by removing from \mathcal{X} a closed subset of zero measure) that $v_h \neq v_k$ for all $h \neq k$. Then the quadratic constraint on $w_{\bar{\ell}} \in \mathbb{R}^2$ gives a circle of positive radius. By excluding at most one point this circle, we can ensure that the relation (3.7) is not fulfilled, which proves the thesis of the lemma. \square

In view of Lemma 3.1, those vectors $\mu \in \mathbb{Z}^{N2^{N-1}}$ that are linear combinations of the family vectors assume a special importance, since that is the only case in which the relation $\sum \mu_i v_i = 0$ cannot be excluded when constructing the set \mathcal{S} . In that case, we will refer to $\mu \sim 0$ as a *formal identity*. In general, we will write $\mu \sim \lambda$ whenever the vector $\mu - \lambda$ is a linear combination of the family relations.

We introduce some more notation: given a vector $\lambda \in \mathbb{Z}^{N2^{N-1}}$, we denote by $\pi_j \lambda$ the projection of λ on the space generated by the members of the j -th generation. Now, let $R_\alpha = \sum_i \alpha_i \lambda^{\mathcal{F}_i}$ be a linear combination with integer coefficients of the family vectors. We denote by n_{R_α} the number of families on which the linear combination is supported, i.e. the cardinality of $\{i \mid \alpha_i \neq 0\}$. Moreover, we denote by $n_{R_\alpha}^k$ the number of families of age k on which R_α is supported, i.e. the cardinality of

$$\{i \mid \alpha_i \neq 0 \text{ and } \mathcal{F}_i \text{ is a family of age } k\} .$$

Finally, we denote respectively by m_{R_α} and M_{R_α} the minimal and the maximal age of families on which R_α is supported. Then, we make the two following simple remarks.

Remark 3.2. *If $n_{R_\alpha}^k = n_{R_\alpha}^{k+1} = 1$, then $\pi_{k+1} R_\alpha$ is supported on at least two distinct elements.*

Remark 3.3. *If $n_{R_\alpha}^k \neq n_{R_\alpha}^{k+1}$, then $\pi_{k+1} R_\alpha$ is supported on at least two distinct elements.*

Before proving the main result of this section, we need some lemmas.

Lemma 3.2. *If $n_{R_\alpha} \geq 3$, then R_α is supported on at least 8 distinct elements.*

Proof. For simplicity of notation, here we put $m := m_{R_\alpha}$ and $M := M_{R_\alpha}$. First, observe that $\pi_m R_\alpha$ is supported on $2n_{R_\alpha}^m$ elements and that $\pi_{M+1} R_\alpha$ is supported on $2n_{R_\alpha}^M$ elements. So, if $n_{R_\alpha}^m + n_{R_\alpha}^M \geq 4$, the thesis is trivial.

Up to symmetry between parents and children, we may choose $n_{R_\alpha}^m \leq n_{R_\alpha}^M$. So, the only non-trivial cases to consider are $(n_{R_\alpha}^m, n_{R_\alpha}^M) = (1, 1)$ and $(n_{R_\alpha}^m, n_{R_\alpha}^M) = (1, 2)$.

Case (1,1). We must have $M \geq m + 2$, since there must be at least three families in R_α . Now, let $C := \max_i n_{R_\alpha}^i$. If $C = 1$, then by Remark 3.2 the support of R_α involves at least 4 generations and at least 2 elements for each generation, so it includes at least

8 elements. If $C > 1$, then there exist $m \leq i, j < M$ with $i \neq j$ such that $n_{R_\alpha}^i < n_{R_\alpha}^{i+1}$ and $n_{R_\alpha}^{j+1} < n_{R_\alpha}^j$. Then, by Remark 3.3, $\pi_{i+1}R_\alpha$ and $\pi_{j+1}R_\alpha$ are supported on at least 2 elements each. Since $\pi_m R_\alpha$ and $\pi_{M+1}R_\alpha$ are supported on exactly 2 elements and since the four indices $m, i+1, j+1, M+1$ are all distinct, then we have the thesis.

Case (1,2). Here, $\pi_m R_\alpha$ is supported on 2 elements and $\pi_M R_\alpha$ is supported on 4 elements. Moreover, there exists $m \leq i < M$ such that $n_{R_\alpha}^{i+1} < n_{R_\alpha}^i$, which by Remark 3.3 gives us at least 2 elements in the support of $\pi_{i+1}R_\alpha$. Thus, we have the thesis. \square

From Lemma 3.2, the next corollary follows immediately.

Corollary 3.1. *If R_α is supported on at most 7 elements, then R_α is an integer multiple of either a family vector or a resonant vector of type CF.*

Lemma 3.3. *Let $A, B, C \in \mathbb{R}$, $R > 0$ and $p, q \in \mathbb{R}^2 \simeq \mathbb{C}$ be fixed. Let*

$$c_1(\vartheta) := p + Re^{i\vartheta} \quad c_2(\vartheta) := p - Re^{i\vartheta}.$$

Then, the function $F : \mathbb{S}^1 \rightarrow \mathbb{R}$ defined by

$$F(\vartheta) := A|c_1(\vartheta)|^2 + B|c_2(\vartheta)|^2 + C - |Ac_1(\vartheta) + Bc_2(\vartheta) + q|^2$$

is an analytic function of ϑ , and it is a constant function only if $A = B$ or if $(A + B - 1)p + q = 0$.

Proof. An explicit computation yields

$$F(\vartheta) = 2R(B - A)\langle (A + B - 1)p + q, e^{i\vartheta} \rangle + K$$

where K is a suitable constant that does not depend on ϑ . \square

Corollary 3.2. *If $A \neq B$ and $(A + B - 1)p + q \neq 0$, then the zeros of F are isolated.*

Lemma 3.4. *Let \mathcal{F} be a family of age i in \mathcal{S} and let $\lambda^{\mathcal{F}_p}$ be the vector corresponding to the sum of the parents of the family \mathcal{F} . Moreover, let $\mu \in \mathbb{Z}^{|\mathcal{S}|}$ be another vector with $|\mu| \leq 5$, such that $\pi_j \mu = 0$ for all $j > i + 1$ and such that the support of μ and the support of the vector $\lambda^{\mathcal{F}_c}$ corresponding to the children of \mathcal{F} are disjoint. Finally, let $h, k \in \mathbb{Z} \setminus \{0\}$. Assume that the formal identity $h\mu + k\lambda^{\mathcal{F}_p} \sim 0$ holds. Then, only two possibilities are allowed:*

1. $h\mu + k\lambda^{\mathcal{F}_p} = 0$;
2. $h\mu + k\lambda^{\mathcal{F}_p}$ is an integer multiple of $\lambda^{\tilde{\mathcal{F}}}$, where $\tilde{\mathcal{F}}$ is a family of age $i - 1$, one of whose children is a parent in \mathcal{F} .

Proof. We first remark that $h\mu + k\lambda^{\mathcal{F}_p}$ is supported on at most 7 elements. Moreover, since it is a linear combination of some family vectors (because of the formal identity $h\mu + k\lambda^{\mathcal{F}_p} \sim 0$), we are in a position to apply Corollary 3.1 and conclude that $h\mu + k\lambda^{\mathcal{F}_p}$ must be an integer multiple of either a family vector or a resonant vector of type CF.

Now, assume by contradiction that $h\mu + k\lambda^{\mathcal{F}_p}$ is a nonzero integer multiple of a resonant vector of type CF. Then, the support of $h\mu + k\lambda^{\mathcal{F}_p}$ cannot include both parents of the family \mathcal{F} , since the support of a CF vector including a couple of parents of age i should

include also a couple of children of age $i + 2$, but we know by the assumptions of this lemma that the support of $h\mu + k\lambda^{\mathcal{F}_p}$ does not include elements of age greater than $i + 1$. Therefore, at least one of the elements in λ must cancel out with one of the elements in $\lambda^{\mathcal{F}_p}$, but then the support of $h\mu + k\lambda^{\mathcal{F}_p}$ can include at most 5 elements, and therefore it cannot be a vector of type CF.

Then, if $h\mu + k\lambda^{\mathcal{F}_p}$ is a nonzero integer multiple of a single family vector $\tilde{\mathcal{F}}$, observe that its support must contain one and only one of the parents of \mathcal{F} . In fact, if both canceled out, then the support of $h\mu + k\lambda^{\mathcal{F}_p}$ could contain at most 3 elements, which is absurd. If none of them canceled out, then we should have $\tilde{\mathcal{F}} = \mathcal{F}$, which in turn is absurd since, by the assumptions of this lemma, the support of $h\mu + k\lambda^{\mathcal{F}_p}$ cannot include any of the children of the family \mathcal{F} . This concludes the proof of the lemma. \square

We can now prove the main proposition.

Proof of Proposition 3.2. The proof is based on the following induction procedure. At each step, we assume to have already fixed i generations and say $h < 2^{N-2}$ families with children in the $i + 1$ -th generation. Our induction hypothesis is that the non-degeneracy condition is satisfied for the vectors μ whose support involves only the elements that we have already fixed. Then, our aim is to show that the non-degeneracy condition holds true also for the set for the vectors supported on the already fixed elements plus the two children of a new family (whose procreation angle has to be accordingly chosen) with children in the $i + 1$ -th generation, up to removing from \mathcal{X} a closed set of null measure.

First, we observe that, at the inductive step zero, i.e. when placing the first generation \mathcal{S}_1 , the set of parameters that satisfy both non-degeneracy and non-vanishing of any fixed finite number of linear relations that are not formal identities is obviously open and of full measure.

Then, we have to study what happens when choosing a procreation angle, i.e. when generating the children of a family $\mathcal{F} = (p_1, p_2; c_1, c_2)$ whose parents have already been fixed. Let $\lambda_1, \lambda_2 \in \mathbb{Z}^{|\mathcal{S}|}$ be defined by $\sum_j (\lambda_i)_j v_j = c_i$ for $i = 1, 2$. We need to study the non-degeneracy condition associated to the vector $\lambda(A, B, \mu) \in \mathbb{Z}^{|\mathcal{S}|}$ given by

$$\lambda(A, B, \mu) := A\lambda_1 + B\lambda_2 + \mu ,$$

where μ satisfies the same properties as in the assumptions of Lemma 3.4 and

$$|A| + |B| + |\mu| \leq 5$$

$$A + B + \sum_j \mu_j = 1 .$$

If $A \neq B$ and if $(A + B - 1)(p_1 + p_2) + 2\sum_j \mu_j v_j \neq 0$, then we are done, because, thanks to Corollary 3.2, the non-degeneracy condition is satisfied for any choice of the generation angle except at most a finite number. Therefore, we have to study separately the case $A = B$ and, for $A \neq B$, we have to prove that $(A + B - 1)\lambda^{\mathcal{F}_p} + 2\mu \sim 0$ holds as a formal identity only in the cases allowed by Definition 2.5. Whenever the formal identity $(A + B - 1)\lambda^{\mathcal{F}_p} + 2\mu \sim 0$ does not hold, we can impose $(A + B - 1)(p_1 + p_2) + 2\sum_j \mu_j v_j \neq 0$ by just removing from \mathcal{X} a closed set of measure zero, thanks to Lemma 3.1.

Case $A = B$. If $(A, B) = (0, 0)$ there is nothing to prove, thanks to the induction hypothesis. Then we have to study $(A, B) = \pm(1, 1)$. In this case, thanks to the linear relation defining the family \mathcal{F} , we have the formal identity

$$\lambda(A, B, \mu) \sim \pm \lambda^{\mathcal{F}_p} + \mu =: \nu^\pm$$

with $|\mu| \leq 3$, $\sum_j \nu_j^+ = -1$ and $\sum_j \nu_j^- = 3$. The good point is that ν^\pm is entirely supported on the elements of the generation set that have already been fixed, so we can apply the induction hypothesis of non-degeneracy to ν^\pm and distinguish the 4 cases given by Definition 2.5: we have to verify that $\lambda(A, B, \mu)$ accordingly falls into one of the allowed cases.

- ν^\pm satisfies case 1 of Definition 2.5. Then one readily verifies that $\lambda(A, B, \mu)$ satisfies either case 2 or case 3 of Definition 2.5.
- ν^\pm satisfies case 2 of Definition 2.5. Observe that the family involved by the statement of case 2 cannot be \mathcal{F} , since ν^\pm cannot be supported on either child of the family \mathcal{F} . Then μ must cancel out one of the two parents appearing in $\pm \lambda^{\mathcal{F}_p}$. It cannot be supported on both parents because that would not be consistent with $|\mu| \leq 3$ and $|\nu^\pm| = 3$. Then one verifies that $\lambda(A, B, \mu)$ satisfies case 4 of Definition 2.5.
- ν^\pm satisfies case 3 of Definition 2.5. Since $|\nu^\pm| = 1$, then nothing cancels out, so the support of ν^\pm includes both parents of \mathcal{F} . But this is absurd, so this case cannot happen.
- ν^\pm satisfies case 4 of Definition 2.5. This case is again absurd, since ν^\pm should be supported on 5 of the 6 elements of a CF vector, including the two parents of the family \mathcal{F} .

Case $A \neq B$. By symmetry, we may impose $|A| > |B|$. Assume that $(A + B - 1)\lambda^{\mathcal{F}_p} + 2\mu \sim 0$ holds as a formal identity: we must prove that this can be true only in the cases allowed by Definition 2.5. First, we consider the case $A + B - 1 = 0$: then, we must have the formal identity $\mu \sim 0$ with $|\mu| \leq 5$: so, by Corollary 3.1 either μ is (up to the sign) a family vector (which may happen only if $(A, B) = (1, 0)$ due to the constraint $|A| + |B| + |\mu| \leq 5$) or $\mu = 0$. Consider the case $(A, B) = (1, 0)$: if μ is a family vector, then $\lambda(A, B, \mu)$ falls into case 3 of Definition 2.5; if $\mu = 0$, then $\lambda(A, B, \mu)$ falls into case 1 of Definition 2.5. If $(A, B) = (2, -1)$ or $(A, B) = (3, -2)$, then $\mu = 0$. Then, in both cases, from

$$\sum_j \lambda_j(A, B, \mu) |v_j|^2 - \left| \sum_j \lambda_j(A, B, \mu) v_j \right|^2$$

with some explicit computations one deduces $|c_1 - c_2|^2 = 0$ which is absurd, since the induction hypothesis implies $p_1 \neq p_2$ and since the endpoints of a diameter of a circle with positive radius are distinct.

Now, if $A + B - 1 \neq 0$ we can apply Lemma 3.4 and deduce that $(A + B - 1)\lambda^{\mathcal{F}_p} + 2\mu$ is either zero or an integer multiple of the vector of a family where one of the parents of \mathcal{F} appears as a child. Suppose first $(A + B - 1)\lambda^{\mathcal{F}_p} + 2\mu = 0$. Then $A + B - 1$ must be even. If $(A, B) = (-1, 0)$, then $\mu = \lambda^{\mathcal{F}_p}$ and $\lambda(A, B, \mu)$ falls into case 2 of Definition 2.5. If $(A, B) = (2, 1)$, then $\mu = -\lambda^{\mathcal{F}_p}$ and $\lambda(A, B, \mu)$ falls into case 3 of Definition 2.5. These are

the only possible cases if $(A+B-1)\lambda^{\mathcal{F}_p} + 2\mu = 0$. Finally, assume that $(A+B-1)\lambda^{\mathcal{F}_p} + 2\mu$ is an integer multiple of the vector of a family where one of the parents of \mathcal{F} appears as a child. Then μ must be such that the other parent of \mathcal{F} is canceled out, so $A+B-1$ again has to be even. If $(A, B) = (-1, 0)$, then $\mu - \lambda^{\mathcal{F}_p}$ is the vector of a family where one of the parents of \mathcal{F} appears as a child and $\lambda(A, B, \mu)$ falls into case 4 of Definition 2.5. This is the only possible case, since the support of μ must include one parent of \mathcal{F} and the other three members of the family where the other parent of \mathcal{F} appears as a child. This also concludes the proof of the proposition. \square

4 Proof of Theorem 1.1

In the previous sections we have proved the existence of non-degenerate sets \mathcal{S} on which the Hamiltonian is (2.1) and the existence of a *slider solution* for its dynamics. We now turn to the NLS equation 1.4 with the purpose of proving the persistence of this type of solution.

As in [15], one can easily prove that (1.4) is locally well-posed in $\ell^1(\mathbb{Z}^2)$: to this end, one first introduces the multilinear operator

$$\mathcal{N}(t) : \ell^1(\mathbb{Z}^2) \times \ell^1(\mathbb{Z}^2) \times \ell^1(\mathbb{Z}^2) \times \ell^1(\mathbb{Z}^2) \times \ell^1(\mathbb{Z}^2) \rightarrow \ell^1(\mathbb{Z}^2)$$

defined by

$$(\mathcal{N}(t)(a, b, c, d, f))_j := \sum_{\substack{j_1, j_2, j_3, j_4, j_5 \in \mathbb{Z}^2 \\ j_1 + j_2 + j_3 - j_4 - j_5 = j}} a_{j_1} b_{j_2} c_{j_3} \bar{d}_{j_4} \bar{f}_{j_5} e^{i\omega_6 t} \quad (4.1)$$

so that (1.4) can be rewritten as

$$-i\dot{a}_j = (\mathcal{N}(t)(a, a, a, a, a))_j .$$

Then, in order to obtain local well-posedness it is enough to observe that the following multilinear estimate holds

$$\|\mathcal{N}(t)(a, b, c, d, f)\|_{\ell^1} \lesssim \|a\|_{\ell^1} \|b\|_{\ell^1} \|c\|_{\ell^1} \|d\|_{\ell^1} \|f\|_{\ell^1} . \quad (4.2)$$

Lemma 4.1. *Let $0 < \sigma < 1$ be an absolute constant (all implicit constants in this lemma may depend on σ). Let $B \gg 1$, and let $T \ll B^4 \log B$. Let $g(t) := \{g_j(t)\}_{j \in \mathbb{Z}^2}$ be a solution of the equation*

$$\dot{g}(t) = i(\mathcal{N}(t)(g(t), g(t), g(t), g(t), g(t)) + \mathcal{E}(t)) \quad (4.3)$$

for times $0 \leq t \leq T$, where $\mathcal{N}(t)$ is defined in (4.1) and the initial data $g(0)$ is compactly supported. Assume also that the solution $g(t)$ and the error term $\mathcal{E}(t)$ obey the bounds of the form

$$\|g(t)\|_{\ell^1(\mathbb{Z}^2)} \lesssim B^{-1} \quad (4.4)$$

$$\left\| \int_0^t \mathcal{E}(\tau) d\tau \right\|_{\ell^1(\mathbb{Z}^2)} \lesssim B^{-1} \quad (4.5)$$

We conclude that if $a(t)$ denotes the solution to the NLS (1.4) with initial data $a(0) = g(0)$ then we have

$$\|a(t) - g(t)\|_{\ell^1(\mathbb{Z}^2)} \lesssim B^{-1-\sigma/2} \quad (4.6)$$

for all $0 \leq t \leq T$.

Proof. The proof is the transposition to the quintic case of the proof of Lemma 2.3 of [15] and is postponed to Appendix B. \square

Given δ, K , construct \mathcal{S} as in Proposition 3.1. Note that we are free to specify $\mathcal{R} = \mathcal{R}(\delta, K)$ (which measures the inner radius of the frequencies involved in \mathcal{S}) as large as we wish. This construction fixes $N = N(\delta, K)$ (the number of generations). We introduce a further parameter ϵ (which we are free to specify as a function of δ, K) and construct the slider solution $b(t)$ to the toy model concentrated at scale ϵ according to Proposition 2.1 above. This proposition also gives us a time $T_0 = T_0(K, \delta)$. Note that the toy model has the following scaling

$$b^{(\lambda)}(t) := \lambda^{-1} b\left(\frac{t}{\lambda^4}\right).$$

We choose the initial data for NLS by setting

$$a_j(0) = b_i^{(\lambda)}(0), \quad \text{for all } j \in \mathcal{S}_i \quad (4.7)$$

and $a_j(t) = 0$ when $j \notin \mathcal{S}$. We want to apply the Approximation Lemma 4.1 with a parameter B chosen large enough so that

$$B^4 \log B \gg \lambda^4 T_0. \quad (4.8)$$

We set $g(t) = \{g_j(t)\}_{j \in \mathbb{Z}^2}$ defined by the slider solution as

$$g_j(t) = b_i^{(\lambda)}(t) \forall j \in \mathcal{S}_i$$

$g_j(t) = 0$ otherwise. Then we set $\mathcal{E}(t) := \{\mathcal{E}_j(t)\}_{j \in \mathbb{Z}^2}$ with

$$\mathcal{E}_j(t) = - \sum_{\substack{k_i \in \mathcal{S}: k_1 + k_2 + k_3 - k_4 - k_5 = j \\ \omega_6 \neq 0}} g_{k_1} g_{k_2} g_{k_3} \bar{g}_{k_4} \bar{g}_{k_5} e^{i\omega_6 t}$$

where $\omega_6 = |k_1|^2 + |k_2|^2 + |k_3|^2 - |k_4|^2 - |k_5|^2 - |j|^2$. We recall that the frequency support of $g(t)$ is in \mathcal{S} for all times. We choose $B = C(N)\lambda$ and then show that for large enough λ the required conditions (4.4), (4.5) hold true. Observe that (4.8) holds true with this choice for large enough λ . Note first that simply by considering its support, the fact that $|\mathcal{S}| = C(N)$, and the fact that $\|b(t)\|_{\ell^\infty} \sim 1$, we can be sure that $\|b(t)\|_{\ell^1(\mathbb{Z})} \sim C(N)$ and therefore

$$\|b^{(\lambda)}(t)\|_{\ell^1(\mathbb{Z})}, \|g(t)\|_{\ell^1(\mathbb{Z}^2)} \leq \lambda^{-1} C(N). \quad (4.9)$$

Thus, (4.4) holds with the choice $B = C(N)\lambda$. For the second condition (4.5), we claim

$$\left\| \int_0^t \mathcal{E}(\tau) d\tau \right\|_{\ell^1} \lesssim C(N)(\lambda^{-5} + \lambda^{-9}T). \quad (4.10)$$

This implies (4.5) since $B = \lambda C(N)$ and $T = \lambda^4 T_0$.

We now prove (4.10). Since ω_6 does not vanish in the sum defining \mathcal{E} , we can replace $e^{i\omega_6\tau}$ by $\frac{d}{d\tau} \left[\frac{e^{i\omega_6\tau}}{i\omega_6} \right]$ and then integrate by parts. Three terms arise: the boundary terms at $\tau = 0, T$ and the integral term involving

$$\frac{d}{d\tau} [g_{k_1} g_{k_2} g_{k_3} \bar{g}_{k_4} \bar{g}_{k_5}].$$

For the boundary terms, we use (4.9) to obtain an upper bound of $C(N)\lambda^{-5}$. For the integral term, the τ derivative falls on one of the g factors. We replace this differentiated term using the equation to get an expression that is 9-linear in g and bounded by $C(N)\lambda^{-9}T$. Once λ has been chosen as above, we choose \mathcal{R} sufficiently large so that the initial data $g(0) = a(0)$ has the right size:

$$\left(\sum_{j \in \mathcal{S}} |g_j(0)|^2 |j|^{2s} \right)^{\frac{1}{2}} \sim \delta. \quad (4.11)$$

This is possible since the quantity on the left scales like λ^{-1} and \mathcal{R}^s respectively in the parameters λ, \mathcal{R} . The issue here is that our choice of frequencies \mathcal{S} only gives us a large factor (that is $\frac{K}{\delta}$) by which the Sobolev norm of the solution will grow. If our initial datum is much smaller than δ in size, the Sobolev norm of the solution will not grow to be larger than K . It remains to show that we can guarantee

$$\left(\sum_{j \in \mathbb{Z}^2} |a_j(\lambda^4 T_0)|^2 |j|^{2s} \right)^{\frac{1}{2}} \geq K, \quad (4.12)$$

where $a(t)$ is the evolution of the initial datum $g(0)$ under the NLS. We do this by first establishing

$$\left(\sum_{j \in \mathcal{S}} |g_j(\lambda^4 T_0)|^2 |j|^{2s} \right)^{\frac{1}{2}} \gtrsim K, \quad (4.13)$$

and then

$$\sum_{j \in \mathcal{S}} |g_j(\lambda^4 T_0) - a_j(\lambda^4 T_0)|^2 |j|^{2s} \lesssim 1. \quad (4.14)$$

In order to prove (4.13), consider the ratio:

$$\mathcal{Q} := \frac{\sum_{j \in \mathcal{S}} |g_j(\lambda^4 T_0)|^2 |j|^{2s}}{\sum_{j \in \mathcal{S}} |g_j(0)|^2 |j|^{2s}} = \frac{\sum_{i=1}^N |b_i^{(\lambda)}(\lambda^4 T_0)|^2 \sum_{j \in \mathcal{S}_i} |j|^{2s}}{\sum_{i=1}^N |b_i^{(\lambda)}(0)|^2 \sum_{j \in \mathcal{S}_i} |j|^{2s}}. \quad (4.15)$$

Set $\mathfrak{J}_i := \sum_{j \in \mathcal{S}_i} |j|^{2s}$, by construction $\frac{\mathfrak{J}_i}{\mathfrak{J}_j} \sim 2^{i-j}$ and by the choice of N one has $\frac{\mathfrak{J}_3}{\mathfrak{J}_{N-2}} \lesssim \delta^2 K^{-2}$. Then one has

$$\frac{\sum_{i=1}^N |b_i^{(\lambda)}(\lambda^4 T_0)|^2 \mathfrak{J}_i}{\sum_{i=1}^N |b_i^{(\lambda)}(0)|^2 \mathfrak{J}_i} \gtrsim \frac{\mathfrak{J}_{N-2}(1-\epsilon)}{\epsilon \sum_{i \neq 3} \mathfrak{J}_i + (1-\epsilon)\mathfrak{J}_3} =$$

$$= \frac{(1 - \epsilon)}{\epsilon \sum_{i \neq 3} \frac{\hat{J}_i}{\hat{J}_{N-2}} + (1 - \epsilon) \frac{\hat{J}_3}{\hat{J}_{N-2}}} = \frac{1}{\frac{\hat{J}_3}{\hat{J}_{N-2}} + O(\epsilon)} \gtrsim \frac{K^2}{\delta^2}.$$

provided that $\epsilon = \epsilon(N, K, \delta)$ is sufficiently small.

In order to prove (4.14), we use the Approximation Lemma 4.1 we obtain that

$$\sum_{j \in \mathcal{S}} |g_j(\lambda^4 T_0) - a_j(\lambda^4 T_0)|^2 |j|^{2s} \lesssim \lambda^{-2-\sigma} \sum_{j \in \mathcal{S}} |j|^{2s} \leq \frac{1}{2}. \quad (4.16)$$

The last inequality is obtained by scaling λ by some (big) parameter C and \mathcal{R} by $C^{1/s}$ so that the bound (4.11) still holds while $\lambda^{-2-\sigma} \sum_{n \in \mathcal{S}} |j|^{2s}$ scales as $C^{-\sigma}$.

A Proof of Proposition 2.3

This proof is in fact exactly the same as in [15], however in that paper all the results are stated for the cubic case (even though they are clearly more general) and so we give a schematic overview of the main steps.

Lemma A.1. *Suppose that $[0, \tau]$ is a time interval on which we have the smallness condition*

$$\int_0^\tau |c(s)|^2 ds \lesssim 1$$

then we have the estimates:

$$\begin{aligned} |c_{j \pm 1}^-(\tau)| &\lesssim e^{-\sqrt{3}\tau} |c_{j \pm 1}^-(0)| + \int_0^\tau e^{-\sqrt{3}(\tau-s)} |c_{j \pm 1}^+(s)| |c_{\neq j \pm 1}|^2, \\ |c_{j \pm 1}^+(\tau)| &\lesssim e^{\sqrt{3}\tau} |c_{j \pm 1}^+(0)| + \int_0^\tau e^{\sqrt{3}(\tau-s)} |c_{j \pm 1}^-(s)| |c_{\neq j \pm 1}|^2, \\ |c_{j \pm 1}(\tau)| &\lesssim e^{\sqrt{3}\tau} |c_{j \pm 1}(0)|, \\ |c^*(\tau)| &\lesssim |c^*(0)|. \end{aligned}$$

Proof. As in [15] this Lemma follows from equations (2.12) by Gronwall's inequality and the definition of $O(\cdot)$. \square

We now prove that the incoming target covers the ricochet target. We start from some basic upper bounds on the flow.

Proposition A.1. *Let $b(\tau)$ be a solution to toy model such that $b(0)$ is within (M_j^-, d_j^-, R_j^-) . Let $c(\tau)$ denote the coordinates of $b(\tau)$ as in (2.12) Then, for all $0 \leq \tau \leq T$ we have the bounds:*

$$\begin{aligned} |c^*(\tau)| &= O(T^{A_j^-} e^{-2\sqrt{3}T}) \\ |c_{j-1}^-(\tau)| &= O(\sigma e^{-\sqrt{3}\tau}) \\ |c_{j-1}^+(\tau)| &= O(T^{2A_j^-+1} e^{-4\sqrt{3}T+\sqrt{3}\tau}) \\ |c_{j+1}^-(\tau)| &= O(r_j^-(1+\tau) e^{-2\sqrt{3}T-\sqrt{3}\tau}) \\ |c_{j+1}^+(\tau)| &= O(r_j^- e^{-2\sqrt{3}T+\sqrt{3}\tau}) \end{aligned} \quad (A.1)$$

Proof. This is Proposition 3.2 of [15]. The proof is an application of the continuity method and of Lemma A.1. \square

Now from these basic upper bounds and from the equations of motion (2.12), (2.15) we deduce improved upper bounds on the dynamical variables. We first consider c_{j-1}^- , we have

$$\dot{c}_{j-1}^- = -\sqrt{3}c_{j-1}^- + \mathcal{O}((c_{j-1}^-)^3) + \mathcal{O}((c_{j-1}^-)^5) + O(T^{A_j^-} e^{-2\sqrt{3}T})$$

for some explicit expression $\mathcal{O}((c_{j-1}^-)^3) + \mathcal{O}((c_{j-1}^-)^5)$. Now let g be the solution to the corresponding equation

$$\dot{g} = -\sqrt{3}g + \mathcal{O}(g^3) + \mathcal{O}(g^5),$$

with the same initial datum $g(0) = \sigma$. One has the bound

$$g(\tau) = O(\sigma e^{\sqrt{3}\tau}), \quad (\text{A.2})$$

which is formula (3.35) of [15]. Then by estimating the *error term* $E_{j-1}^- := c_{j-1}^- - g$ one has

$$c_{j-1}^-(\tau) = g(\tau) + O(T^{A_j^-+1} e^{-2\sqrt{3}T}) \quad (\text{A.3a})$$

$$\mathcal{O}(c^2) = \mathcal{O}(g^2) + O(T^{A_j^-+1} e^{-2\sqrt{3}T}) \quad (\text{A.3b})$$

$$\mathcal{O}(c_{\neq j+1}^2) = \mathcal{O}(g^2) + O(T^{A_j^-+1} e^{-2\sqrt{3}T-\sqrt{3}\tau}) \quad (\text{A.3c})$$

which are respectively formulæ(3.36)-(3.38) of [15]. Now we control the leading peripheral modes. Inserting (A.3b) in (2.12e) we see that

$$\dot{c}_{\geq j+2} = i\kappa c_{\geq j+2} + \mathcal{O}(c_{\geq j+2}g^2) + \mathcal{O}(c_{\geq j+2}g^4) + O(T^{A_j^-} e^{-2\sqrt{3}T} |c_{\geq j+2}|)$$

We approximate this by the corresponding linear equation

$$\dot{u} = i\kappa u + \mathcal{O}(ug^2) + \mathcal{O}(ug^4)$$

where $u(\tau) \in \mathbb{C}^{N-j-1}$. This equation has a fundamental solution $G_{\geq j+2}(\tau) : \mathbb{C}^{N-j-1} \rightarrow \mathbb{C}^{N-j-1}$. From (A.2) and Gronwall's inequality we have

$$\int_0^T g^2(\tau) d\tau = O(1) \quad (\text{A.4})$$

and

$$|G_{\geq j+2}|, |G_{\geq j+2}^{-1}| = O(1). \quad (\text{A.5})$$

Setting $c_{\geq j+2}(0) = e^{-2\sqrt{3}T} a_{\geq j+2} + O(T^{A_j^-} e^{-3\sqrt{3}T})$, we define

$$E_{\geq j+2} := c_{\geq j+2} - e^{-2\sqrt{3}T} G_{\geq j+2} a_{\geq j+2}.$$

Applying the bound on $c_{\geq j+2}$ from Proposition A.1 and Gronwall's inequality, we conclude

$$|E_{\geq j+2}(\tau)| = O(T^{A_j^-} e^{-3\sqrt{3}T})$$

for all $0 \leq \tau \leq T$ and thus

$$c_{\geq j+2}(\tau) = e^{-2\sqrt{3}T} G_{\geq j+2}(\tau) a_{\geq j+2} + O(T^{A_j^-} e^{-3\sqrt{3}T}). \quad (\text{A.6})$$

This is formula (3.41) of [15].

Now we consider the two leading secondary modes c_{j+1}^+, c_{j+1}^- simultaneously. From (2.12), (A.3) and Proposition A.1, we have the system

$$\begin{pmatrix} \dot{c}_{j+1}^- \\ \dot{c}_{j+1}^+ \end{pmatrix} = \sqrt{3} \begin{pmatrix} -c_{j+1}^- \\ c_{j+1}^+ \end{pmatrix} + M(\tau) \begin{pmatrix} c_{j+1}^- \\ c_{j+1}^+ \end{pmatrix} + \begin{pmatrix} O(T^{A_j^-+1} e^{-4\sqrt{3}T}) \\ O(T^{A_j^-+1} e^{-4\sqrt{3}T+\sqrt{3}\tau}) \end{pmatrix}$$

Here $M(\tau)$ is a two by two matrix with entries all $\mathcal{O}(g^2) + \mathcal{O}(g^4)$. Passing to the variables

$$\tilde{a}_{j+1}(\tau) := \begin{pmatrix} \tilde{a}_{j+1}^-(\tau) \\ \tilde{a}_{j+1}^+(\tau) \end{pmatrix}$$

where

$$\tilde{a}_{j+1}^-(\tau) = e^{2\sqrt{3}T+\sqrt{3}\tau} c_{j+1}^-(\tau), \quad \tilde{a}_{j+1}^+(\tau) = e^{2\sqrt{3}T-\sqrt{3}\tau} c_{j+1}^+(\tau),$$

we get the equation

$$\begin{cases} \partial_\tau \tilde{a}_{j+1}(\tau) &= A(\tau) \tilde{a}_{j+1}(\tau) + O(T^{A_j^-+1} e^{-2\sqrt{3}T+\sqrt{3}\tau}) \\ \tilde{a}_{j+1}(0) &= a_{j+1} + O(T^{A_j^-} e^{-\sqrt{3}T}) \end{cases} \quad (\text{A.7})$$

where $A(\tau)$ is some known matrix which (by (A.2)) has bounds

$$A(\tau) = \sigma^2 \begin{pmatrix} O(e^{-2\sqrt{3}\tau}) & O(1) \\ O(e^{-4\sqrt{3}\tau}) & O(e^{-2\sqrt{3}\tau}) \end{pmatrix}.$$

We have obtained formula (3.42) of [15]. Hence, following verbatim the proof given in [15], we get

$$\begin{pmatrix} e^{2\sqrt{3}T+\sqrt{3}\tau} c_{j+1}^- \\ e^{2\sqrt{3}T-\sqrt{3}\tau} c_{j+1}^+ \end{pmatrix} = G_{j+1}(\tau) a_{j+1} + O(T^{A_j^-+2} e^{-\sqrt{3}T}) \quad (\text{A.8})$$

which is formula (3.45) of [15].

Then, following Section 3.7 of [15] verbatim, we deduce that the incoming target covers the ricochet target.

Then, one has to prove that the ricochet target covers the outgoing target. In order to do this, one should adapt sections 3.8-3.9 of [15] exactly as we have done in the previous section. Since this is completely straightforward, we will not write it down.

The last step consists in proving that the outgoing target (M_j^+, d_j^+, r_j^+) covers the next incoming target $(M_{j+1}^-, d_{j+1}^-, r_{j+1}^-)$. An initial datum in the outgoing target has the form

$$\begin{aligned} c_{\leq j-1}(0) &= O(T^{A_j^+} e^{-2\sqrt{3}T}) \\ c_{j+1}^-(0) &= O(T^{A_j^+} e^{-4\sqrt{3}T}) \\ c_{j+1}^+(0) &= \sigma + O(T^{A_j^+} e^{-\sqrt{3}T}) \\ c_{\geq j+2}(0) &= e^{-2\sqrt{3}T} a_{\geq j+2} + O(T^{A_j^+} e^{-3\sqrt{3}T}) \end{aligned}$$

for some $a_{\geq j+2}$ of magnitude at most r_j^+ . From (2.12e), (2.13), (2.14) we deduce

$$\dot{c}_{\neq j+1} = O(|c_{\neq j+1}|) .$$

Thus, for all $0 \leq \tau \leq 10 \log \frac{1}{\sigma}$, Gronwall's inequality gives

$$c_{\neq j+1}(\tau) = O\left(\frac{1}{\sigma^{O(1)}} T^{A_j^+} e^{-2\sqrt{3}T}\right) . \quad (\text{A.9})$$

The stable leading mode c_{j+1}^- can be controlled by (2.12c), which by (A.9) becomes

$$\dot{c}_{j+1}^- = O(|c_{j+1}^-|) + O\left(\frac{1}{\sigma^{O(1)}} T^{2A_j^+} e^{-4\sqrt{3}T}\right) .$$

By Gronwall's inequality we conclude

$$c_{j+1}^-(\tau) = O\left(\frac{1}{\sigma^{O(1)}} T^{2A_j^+} e^{-4\sqrt{3}T}\right) . \quad (\text{A.10})$$

Then, taking the c_{j+1}^+ component of (2.10) we obtain, by (A.9) and (A.10)

$$\dot{c}_{j+1}^+ = \sqrt{3}(1 - |c_{j+1}^+|^2)c_{j+1}^+ + O\left(\frac{1}{\sigma^{O(1)}} T^{2A_j^+} e^{-4\sqrt{3}T}\right) .$$

As in [15], we define \hat{g} to be the solution to the ODE

$$\partial_\tau \hat{g} = \sqrt{3}(1 - |\hat{g}|^2)\hat{g} \quad (\text{A.11})$$

with initial datum $\hat{g}(0) = \sigma$. Such solution can be easily computed and is given by

$$\hat{g}(\tau) = \frac{1}{\sqrt{1 + e^{-2\sqrt{3}(\tau - \tau_0)}}}$$

where τ_0 is defined by

$$\frac{1}{\sqrt{1 + e^{2\sqrt{3}\tau_0}}} = \sigma .$$

We note that

$$\hat{g}(2\tau_0) = \frac{1}{\sqrt{1 + e^{-2\sqrt{3}\tau_0}}} = \sqrt{1 - \sigma^2}$$

and that $2\tau_0 \leq 10 \log \frac{1}{\sigma}$ if σ is small enough. Then, estimating as in [15] (via Gronwall's inequality) the error

$$E_{j+1}^+ := c_{j+1}^+ - \hat{g} ,$$

we get

$$c_{j+1}^+(\tau) = \hat{g}(\tau) + O\left(\frac{1}{\sigma^{O(1)}} T^{A_j^+} e^{-\sqrt{3}T}\right) . \quad (\text{A.12})$$

This (together with (A.9) and (A.10)) implies

$$\mathcal{O}(c^2) + \mathcal{O}(c^4) = \mathcal{O}(\hat{g}^2) + \mathcal{O}(\hat{g}^4) + O\left(\frac{1}{\sigma^{O(1)}} T^{A_j^+} e^{-\sqrt{3}T}\right) . \quad (\text{A.13})$$

Now, from (2.12c), (A.9) and (A.13), we have

$$\dot{c}_{\geq j+2} = i\kappa c_{\geq j+2} + \mathcal{O}(\hat{g}^2 c_{\geq j+2}) + \mathcal{O}(\hat{g}^4 c_{\geq j+2}) + O\left(\frac{1}{\sigma O(1)} T^{2A_j^+} e^{-3\sqrt{3}T}\right).$$

We approximate this flow by the linear equation

$$\dot{u} = i\kappa u + \mathcal{O}(u\hat{g}^2) + \mathcal{O}(u\hat{g}^4)$$

where $u(\tau) \in \mathbb{C}^{N-j-1}$. This equation has a fundamental solution $\hat{G}_{\geq j+2}(\tau) : \mathbb{C}^{N-j-1} \rightarrow \mathbb{C}^{N-j-1}$ for all $\tau \geq 0$; from the boundedness of \hat{g} and Gronwall's inequality we get

$$|\hat{G}_{\geq j+2}(\tau)|, |\hat{G}_{\geq j+2}^{-1}(\tau)| \lesssim \frac{1}{\sigma O(1)}. \quad (\text{A.14})$$

As in [15], a Gronwall estimate of the error

$$E_{\geq j+2}(\tau) := c_{\geq j+2}(\tau) - e^{-2\sqrt{3}T} \hat{G}_{\geq j+2}(\tau) a_{\geq j+2}$$

gives

$$c_{\geq j+2}(\tau) = e^{-2\sqrt{3}T} \hat{G}_{\geq j+2}(\tau) a_{\geq j+2} + O\left(\frac{1}{\sigma O(1)} T^{2A_j^+} e^{-3\sqrt{3}T}\right) \quad (\text{A.15})$$

which is equation (3.62) of [15]. Then, at the time $\tau = 2\tau_0 \leq 10 \log \frac{1}{\sigma}$, the estimates become

$$\begin{aligned} c_{\leq j-1}(2\tau_0) &= O\left(\frac{1}{\sigma O(1)} T^{A_j^+} e^{-2\sqrt{3}T}\right) \\ c_{j+1}^-(2\tau_0) &= O\left(\frac{1}{\sigma O(1)} T^{2A_j^+} e^{-4\sqrt{3}T}\right) \\ c_{j+1}^+(2\tau_0) &= \sqrt{1-\sigma^2} + O\left(\frac{1}{\sigma O(1)} T^{A_j^+} e^{-\sqrt{3}T}\right) \\ c_{\geq j+2}(2\tau_0) &= e^{-2\sqrt{3}T} \hat{G}_{\geq j+2}(2\tau_0) a_{\geq j+2} + O\left(\frac{1}{\sigma O(1)} T^{2A_j^+} e^{-3\sqrt{3}T}\right). \end{aligned}$$

From this, we deduce

$$|b_j| = \left(1 - \sum_{k \neq j} |c_k|^2\right)^{\frac{1}{2}} = \sigma + O\left(\frac{1}{\sigma O(1)} T^{A_j^+} e^{-\sqrt{3}T}\right).$$

Moving back to the coordinates b_1, \dots, b_N , this means that we have

$$\begin{aligned} b_{\leq j-1}(2\tau_0) &= O\left(\frac{1}{\sigma O(1)} T^{A_j^+} e^{-2\sqrt{3}T}\right) \\ b_j(2\tau_0) &= \left[\sigma + \Re O\left(\frac{1}{\sigma O(1)} T^{A_j^+} e^{-\sqrt{3}T}\right)\right] e^{i\vartheta^{(j)}(2\tau_0)} \\ b_{j+1}(2\tau_0) &= \left[\sqrt{1-\sigma^2} + \Re O\left(\frac{1}{\sigma O(1)} T^{A_j^+} e^{-\sqrt{3}T}\right)\right] \bar{\omega} e^{i\vartheta^{(j)}(2\tau_0)} + O\left(\frac{1}{\sigma O(1)} T^{2A_j^+} e^{-4\sqrt{3}T}\right) \\ b_{\geq j+2}(2\tau_0) &= e^{i\vartheta^{(j)}(2\tau_0)} e^{-2\sqrt{3}T} \hat{G}_{\geq j+2}(2\tau_0) a_{\geq j+2} + O\left(\frac{1}{\sigma O(1)} T^{2A_j^+} e^{-3\sqrt{3}T}\right). \end{aligned}$$

where the notation $f = \Re O(\cdot)$ means that both $f = O(\cdot)$ and $f \in \mathbb{R}$. We now have to recast this in terms of the variables $c_1^{(j+1)}, \dots, c_N^{(j+1)}$ in phase with \mathbb{T}_{j+1} . Following [15], we denote these variables by $\tilde{c}_1, \dots, \tilde{c}_N$. We first note that

$$\vartheta^{(j+1)}(2\tau_0) = \vartheta^{(j)}(2\tau_0) + \bar{\omega} + O\left(\frac{1}{\sigma^{O(1)}} T^{2A_j^+} e^{-4\sqrt{3}T}\right).$$

Then, we deduce our final estimates

$$\begin{aligned} \tilde{c}_{\leq j-1}(2\tau_0) &= O\left(\frac{1}{\sigma^{O(1)}} T^{A_j^+} e^{-2\sqrt{3}T}\right) \\ \tilde{c}_j^-(2\tau_0) &= \sigma + O\left(\frac{1}{\sigma^{O(1)}} T^{A_j^+} e^{-\sqrt{3}T}\right) \\ \tilde{c}_j^+(2\tau_0) &= O\left(\frac{1}{\sigma^{O(1)}} T^{2A_j^+} e^{-4\sqrt{3}T}\right) \\ \tilde{c}_{\geq j+2}(2\tau_0) &= \omega e^{-2\sqrt{3}T} \hat{G}_{\geq j+2}(2\tau_0)_{a_{\geq j+2}} + O\left(\frac{1}{\sigma^{O(1)}} T^{2A_j^+} e^{-3\sqrt{3}T}\right). \end{aligned}$$

This, together with (A.14), shows that the outgoing target (M_j^+, d_j^+, r_j^+) covers the next incoming target $(M_{j+1}^-, d_{j+1}^-, r_{j+1}^-)$ (it is enough to choose $a_{\geq j+2}$ appropriately).

B Proof of Lemma 4.1

Proof. First note that since $a(0) = g(0)$ is assumed to be compactly supported, the solution $a(t)$ to (1.4) exists globally in time, is smooth with respect to time, and is in $\ell^1(\mathbb{Z}^2)$ in space. Write

$$F(t) := -i \int_0^t \mathcal{E}(\tau) d\tau, \quad \text{and} \quad d(t) := g(t) + F(t).$$

Observe that

$$-i\dot{d} = \mathcal{N}(d - F, d - F, d - F, d - F, d - F),$$

Observe that $g = O_{\ell^1}(B^{-1})$ and $F = O_{\ell^1}(B^{-1-\sigma})$ by hypothesis. In particular we have $d = O_{\ell^1}(B^{-1})$. By multilinearity and (4.2) we thus have

$$-i\dot{d} = \mathcal{N}(d, d, d, d, d) + O_{\ell^1}(B^{-5-\sigma}). \quad (\text{B.1})$$

Now write $e := a - d$, recall that a is the solution of the NLS. Then we have

$$-i(\dot{d} + \dot{e}) = \mathcal{N}(d + e, d + e, d + e, d + e, d + e). \quad (\text{B.2})$$

Subtracting (B.2) from (B.1) (and using (4.2)) we get

$$i\dot{e} = O_{\ell^1}(B^{-5-\sigma}) + O_{\ell^1}(B^{-4}\|e\|_1) + O_{\ell^1}(\|e\|_1^5),$$

so taking the ℓ^1 norm and differentiating in time we have:

$$\frac{d}{dt}\|e\|_1 \lesssim B^{-5-\sigma} + B^{-4}\|e\|_1 + \|e\|_1^5.$$

We make the bootstrap assumption that $\|e\|_1 = O(B^{-1})$ for all $t \in [0, T]$, so that one can absorb the third term on the right hand side in the second. Gronwall's inequality then gives:

$$\|e\|_1 \leq B^{-1-\sigma} \exp(CB^{-4}t)$$

for all $t \in [0, T]$. Since $T \ll B^4 \log B$ we have $\|e\|_1 \ll B^{-1-\sigma/2}$ and the result follows by the bootstrap argument. \square

The result of Lemma 4.1 is that $g(t)$ is a good approximation of a solution to (1.4) on a time interval of approximate length $B^4 \log B$, a factor $\log B$ larger than the interval $[0, B^4]$ for which the solution is controlled by a straightforward local-in-time argument. We choose the exponent $\sigma/2$ for concreteness, but it could be replaced by any exponent between 0 and σ .

C Two-generation sets without full energy transmission

We describe the dynamics associated to the sets $\mathfrak{S}^{(2)}$, $\mathfrak{S}^{(3)}$ given in the introduction.

In $\mathfrak{S}^{(2)}$ we have six complex variables β_k , $k \in \mathfrak{S}^{(2)}$ and correspondingly six constants of motion, so that the system is integrable. Passing to symplectic polar coordinates $\beta_k = \sqrt{J_k} e^{i\theta_k}$ we find that $J_{k_1} - J_{k_2}$, $J_{k_1} - J_{k_3}$, $J_{k_4} - J_{k_5}$, $J_{k_4} - J_{k_6}$ are constant in time. Then one can study the dynamics reduced to the invariant subspace where all these constants are zero. We are left with four degrees of freedom denoted by $I_1, I_2, \theta_1, \theta_2$, and the Hamiltonian:

$$H = 31(I_1 + I_2)^3 - 66I_1I_2(I_1 + I_2) + 24I_1^{3/2}I_2^{3/2} \cos(3(\theta_1 - \theta_2))$$

Then we reduce to the subspace⁵ where $I_1 + I_2 = 1$ and get the phase portrait of Figure 4. It is evident from the picture that there is no orbit connecting $I_1 = 0$ to $I_1 = 1$. One could argue that this is due to our choice of invariant subspace. However, if we set for instance $J_{k_1} \neq J_{k_2}$ then we cannot transfer all the mass to k_4, k_5, k_6 since this would imply $J_{k_1} = J_{k_2} = J_{k_3} = 0$.

The case of $\mathfrak{S}^{(3)}$ is discussed in detail in [22]. Proceeding as above one gets the phase portrait of Figure 5. One could generalize this approach by taking two complete and action preserving sets $\mathcal{S}_1, \mathcal{S}_2$ and connecting them with resonances as $\mathfrak{S}^{(2)}$ or $\mathfrak{S}^{(3)}$ as we have discussed in introduction for $\mathfrak{S}^{(1)}$. However the dynamics is in fact qualitatively the same and one does not have full energy transfer.

We have experimented also with higher order NLS equations. We have not performed a complete classification but it appears that the sets $\mathfrak{S}^{(2)}$, $\mathfrak{S}^{(3)}$ never give full energy transfer while $\mathfrak{S}^{(1)}$ does.

⁵this subspace is invariant due to the conservation of mass

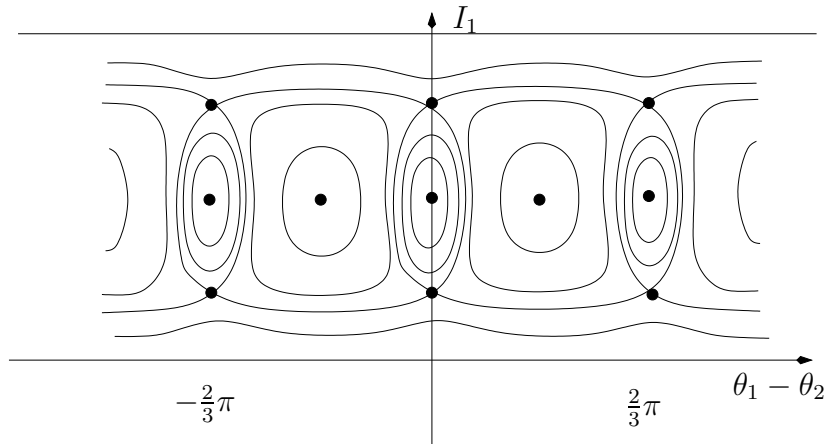


Figure 4: The phase portrait of H on the subspace $I_1 + I_2 = 1$, the dynamical variables are clearly $I_1, \theta_1 - \theta_2$.

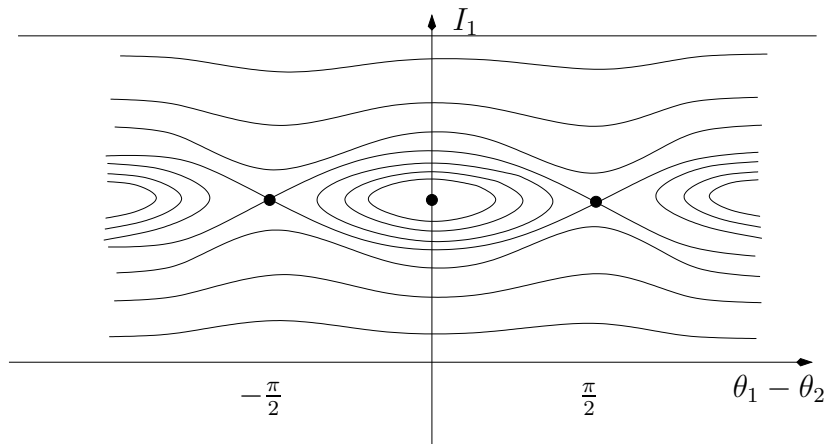


Figure 5: The phase portrait of $\mathfrak{S}^{(3)}$.

References

- [1] D. Bambusi. Long time stability of some small amplitude solutions in nonlinear Schrödinger equations. *Comm. Math. Phys.*, 189(1):205–226, 1997.
- [2] D. Bambusi. Nekhoroshev theorem for small amplitude solutions in nonlinear Schrödinger equations. *Math. Z.*, 230(2):345–387, 1999.
- [3] D. Bambusi. Birkhoff normal form for some nonlinear PDEs. *Comm. Math. Phys.*, 234(2):253–285, 2003.
- [4] D. Bambusi and B. Grébert. Birkhoff normal form for partial differential equations with tame modulus. *Duke Math. J.*, 135(3):507–567, 2006.

- [5] M. Berti and L. Biasco. Branching of Cantor manifolds of elliptic tori and applications to PDEs. *Comm. Math. Phys.*, 305(2011), no. 3, 741–796.
- [6] M. Berti and P. Bolle. Quasi-periodic solutions with Sobolev regularity of NLS on \mathbb{T}^d with a multiplicative potential. *Journal European Math. Society*, 15 (2013), 229–286.
- [7] M. Berti, L. Corsi, and M. Procesi. An abstract Nash-Moser theorem and quasi-periodic solutions for NLW and NLS on compact Lie groups and homogeneous manifolds . to appear in *Comm. Math. Phys.*
- [8] J. Bourgain. Fourier transform restriction phenomena for certain lattice subsets and applications to nonlinear evolution equations. I. Schrödinger equations. *Geom. Funct. Anal.* 3 (1993), no. 2, 107–156.
- [9] J. Bourgain. Construction of approximative and almost periodic solutions of perturbed linear Schrödinger and wave equations. *Geom. Funct. Anal.*, 2:201–230, 1996.
- [10] J. Bourgain. On the growth in time of higher Sobolev norms of smooth solutions of Hamiltonian PDE. *Internat. Math. Res. Notices*, 6:277–304, 1996.
- [11] J. Bourgain. Quasi-periodic solutions of Hamiltonian perturbations of 2D linear Schrödinger equations. *Annals of Math.* 148 (1998) 363–439.
- [12] J. Bourgain. On diffusion in high-dimensional Hamiltonian systems and PDE. *J. Anal. Math.* 80 (2000), 1–35.
- [13] N. Burq, P. Gérard, and N. Tzvetkov. Strichartz inequalities and the nonlinear Schrödinger equation on compact manifolds. *Amer. J. Math.* 126 (2004), no. 3, 569–605.
- [14] L. Chierchia and P. Perfetti. Second order Hamiltonian equations on \mathbb{T}^∞ and almost-periodic solutions. *J. Differential Equations*, 116(1):172–201, 1995.
- [15] J. Colliander, M. Keel, G. Staffilani, H. Takaoka, and T. Tao. Transfer of energy to high frequencies in the cubic defocusing nonlinear Schrödinger equation. *Invent. Math.* 181 (2010), no. 1, 39–113.
- [16] H. Eliasson and S. Kuksin. KAM for the nonlinear Schrödinger equation. *Ann. of Math. (2)* 172 (2010), no. 1, 371–435.
- [17] E. Faou, L. Gauckler, and C. Lubich. Sobolev Stability of Plane Wave Solutions to the Cubic Nonlinear Schrödinger Equation on a Torus. *Comm. Partial Differential Equations* 38 (2013), no. 7, 1123–1140.
- [18] J. Geng, X. Xu, and J. You. An infinite dimensional KAM theorem and its application to the two dimensional cubic Schrödinger equation. *Adv. in Math.* 226 (2011), no. 6, 5361–5402.
- [19] P. Gérard and S. Grellier. The cubic Szegő equation. *Ann. Sci. Éc. Norm. Supér. (4)*, 43(5):761–810, 2010.
- [20] B. Grébert, R. Imekraz, and E. Paturel. Normal forms for semilinear quantum harmonic oscillators. *Comm. Math. Phys.*, 291(3):763–798, 2009.
- [21] B. Grébert, T. Kappeler, and J. Pöschel. Normal form theory for the NLS equation. Preprint available at <http://arxiv.org/abs/0907.39381>, 2009.

- [22] B. Grébert and L. Thomann. Resonant dynamics for the quintic nonlinear Schrödinger equation. *Ann. Inst. H. Poincaré Anal. Non Linéaire* 29 (2012), no. 3, 455–477.
- [23] M. Guardia and V. Kaloshin. Growth of Sobolev norms in the cubic defocusing nonlinear Schrödinger equation. To appear in *Journal of the European Mathematical Society*.
- [24] Z. Hani. Long-time strong instability and unbounded orbits for some periodic nonlinear Schrödinger equation. To appear in *Archive for Rational Mechanics and Analysis*.
- [25] E. Haus and L. Thomann. Dynamics on resonant clusters for the quintic nonlinear Schrödinger equation. *Dyn. Partial Differ. Equ.* 10 (2013), no. 2, 157–169.
- [26] S. B. Kuksin. On squeezing and flow of energy for nonlinear wave equations. *Geom. Funct. Anal.*, 5(4):668–701, 1995.
- [27] S. B. Kuksin. Growth and oscillations of solutions of nonlinear Schrödinger equation. *Comm. Math. Phys.*, 178(2):265–280, 1996.
- [28] S. B. Kuksin. On turbulence in nonlinear Schrödinger equations. *Geom. Funct. Anal.*, 7(4):783–822, 1997.
- [29] S. B. Kuksin. Oscillations in space-periodic nonlinear Schrödinger equations. *Geom. Funct. Anal.*, 7(2):338–363, 1997.
- [30] S. B. Kuksin. Spectral properties of solutions for nonlinear PDEs in the turbulent regime. *Geom. Funct. Anal.*, 9(1):141–184, 1999.
- [31] S. B. Kuksin and J. Pöschel. Invariant Cantor manifolds of quasi-periodic oscillations for a nonlinear Schrödinger equation. *Ann. of Math. (2)* 143 (1996), no. 1, 149–179.
- [32] O. Pocovnicu. Explicit formula for the solution of the Szegő equation on the real line and applications. *Discrete Contin. Dyn. Syst.*, 31(3):607–649, 2011.
- [33] O. Pocovnicu. First and second order approximations for a nonlinear wave equation. *J. Dynam. Differential Equations*, 2:305–333, 2013.
- [34] J. Pöschel. A KAM-Theorem for some nonlinear PDEs. *Ann. Sc. Norm. Pisa*, 23 (1996) 119–148.
- [35] J. Pöschel. On the construction of almost periodic solutions for a nonlinear Schrödinger equation. *Ergodic Theory Dynam. Systems*, 22(5):1537–1549, 2002.
- [36] C. Procesi and M. Procesi. A normal form for the Schrödinger equation with analytic non-linearities. *Communications in Mathematical Physics* 312 (2012), no. 2, 501–557.
- [37] C. Procesi and M. Procesi. A KAM algorithm for the resonant non-linear Schrödinger equation. Preprint available at <http://arxiv.org/abs/1211.4242>, 2013.
- [38] Procesi M., Xu X., Quasi-Töplitz Functions in KAM Theorem, *SIAM J.Math. Anal.* 45 (2013) no. 4, 2148–2181.
- [39] G. Staffilani. On the growth of high Sobolev norms of solutions for KdV and Schrödinger equations. *Duke Math Journal*, 86(1):109–142, 1997.
- [40] W.-M. Wang. Logarithmic bounds on Sobolev norms for time dependent linear Schrödinger equations. *Comm. Partial Differential Equations*, 33(10–12):2164–2179, 2008.

- [41] W.-M. Wang. Supercritical Nonlinear Schrödinger equations I: Quasi-Periodic Solutions. Preprint available at <http://arxiv.org/abs/1007.0156>, 2010.
- [42] W.-M. Wang. Supercritical Nonlinear Schrödinger Equations II: Almost Global Existence. Preprint available at <http://arxiv.org/abs/1007.0154>, 2010.
- [43] E. Wayne. Periodic and quasi-periodic solutions of nonlinear wave equations via KAM theory. *Comm. Math. Phys.* *127* (1990), 479–528.

This research was supported by the European Research Council under FP7, project “Hamiltonian PDEs and small divisor problem: a dynamical systems approach” (HamPDEs).