A stochastic thermalization of the Discrete Nonlinear Schrödinger Equation

Amirali Hannani, Stefano Olla

A stochastic thermalization of the Discrete Nonlinear Schrödinger Equation

Amirali Hannani[1] and Stefano Olla[1,2,3] *

September 3, 2021

Abstract

We introduce a mass conserving stochastic perturbation of the discrete nonlinear Schrödinger equation that models the action of a heat bath at a given temperature. We prove that the corresponding canonical Gibbs distribution is the unique invariant measure. In the one-dimensional cubic focusing case on the torus, we prove that in the limit for large time, continuous approximation, and low temperature, the solution converges to the steady wave of the continuous equation that minimizes the energy for a given mass.

1 Introduction

Consider the Nonlinear Schrödinger Equation in d space dimension:

$$\begin{aligned} &i\partial_t\psi(x,t) = -\Delta\psi(x,t) + \kappa|\psi|^{p-1}\psi(x,t); \qquad p>1,\\ &\psi : \Omega^d \times \mathbb{R}_+ \to \mathbb{C}; \quad \psi(x,0) := \psi_0(x), \end{aligned} \tag{1.1}$$

where $\Omega = \mathbb{R}$, or $\Omega = \mathbb{T}^1_L$, the circle of length L, for the periodic boundary conditions case, $\kappa = -1$ corresponds to the focusing case, and $\kappa = 1$ to the defocusing . This equation has many conserved quantities, in particular the most important are the energy and the mass:

$$\mathcal{H}(\psi) = \frac{1}{2}\int |\partial_x\psi|^2 dx + \frac{\kappa}{p+1}\int |\psi|^{p+1}dx, \qquad \mathcal{M}(\psi) = \int |\psi|^2 dx. \tag{1.2}$$

In some particular cases (like for $d=1$ and $p=3$), the dynamics is completely integrable.

We are particularly interested in the focusing case $\kappa=-1$, where the non-linearity contrast the dispersive effect of the Laplacian. Notice that, thanks to the Gagliardo-Nirenberg inequality (cf. (C.1)), $\mathcal{H}(\psi)$ is still bounded below if $\mathcal{M}(\psi)$ is fixed, and $p < 1+\frac{4}{d}$, also known as mass sub-critical case. In the one-dimensional mass sub-critical NLS with periodic

*[1]CEREMADE, UMR CNRS, Université Paris-Dauphine, PSL Research University,
[2]Institute Universitaire de France, [3]GSSI. hannani@ceremade.dauphine.fr, olla@ceremade.dauphine.fr.
This work was partially supported by ANR-15-CE40-0020-01 grant LSD. We thank Sourav Chatterjee for stimulating discussions at the early stage of this work. We also thank Nikolay Tzvetkov for suggesting reference [31] for the periodic NLSE.

boundary conditions ($d = 1$, $p < 5$, and $\Omega = \mathbb{T}^1_L$), it has been proven that the canonical Gibbs measure at temperature β^{-1}, formally defined as

$$\mathcal{Z}^{-1} \exp\{-\beta \mathcal{H}(\psi)\} \delta\left(\mathcal{M}(\psi) = m\right) \prod_x d\psi(x) \tag{1.3}$$

is invariant for the dynamics defined by (1.1). Rigorous definition of (1.3) can be found in [25], while its invariance for the dynamics is proven in [3], see also [4], [27], [29], [30], [28]. For $p = 3$, $d = 1$, (1.1) is completely integrable; hence it is obvious that (1.3) cannot be ergodic, not even conditioned to a value of the energy $\mathcal{H}$ (i.e., the *microcanonical Gibbs measure*) as there are other conserved quantities beyond energy and mass. A natural question is then how to define a stochastic perturbation of (1.1) such that acts as a *heat bath* at temperature β^{-1}, and such that the resulting stochastic dynamics has (1.3) as the unique stationary measure. This implies that the only conserved quantity of the dynamics should be the mass $\mathcal{M}$.

Formally, one way to define such stochastic dynamics is to consider the stochastic partial differential equation

$$i\partial_t \psi(x,t) = -\Delta\psi(x,t) + \kappa|\psi|^{p-1}\psi(x,t) - \gamma\psi(x,t)\left(i\beta^{-1} - \frac{\delta\mathcal{H}(\psi))}{\delta\theta(x)}\right) + \sqrt{2\gamma\beta^{-1}}\psi(x,t)W(x,t), \tag{1.4}$$

where $\theta(x)$ is the phase of $\psi(x)$ ($\psi(x) = |\psi(x)|e^{i\theta(x)}$), $W(x,t)$ is the standard space-time white noise, and $\gamma > 0$ is a parameter that regulates the intensity of the contact with the heat bath. Notice that $\frac{\delta\mathcal{H}(\psi))}{\delta\theta(x)} = \mathcal{I}m[\psi(x)^*\Delta\psi(x)]$, and that (1.4) should be intended in the Ito's sense. Consequently, the mass $\mathcal{M}(\psi)$ is still formally conserved by this dynamics. The heat bath acts with random but continuous rotations of the phase of $\psi(x)$ at each point x. Because of the singularity in space of the multiplicative white noise W and the non-linearities present in (1.4), it is very hard to give sense to the solution of this equation. There is an extensive literature on the NLSE with space correlated multiplicative noise (cf. [10], [11]), but it does not include non linearities like $\psi(x)\Delta\psi^*(x)$. Additive noises have also been studied (cf. [24], [5], [6]) but usually do not conserve the mass, and the corresponding dynamics have the Grand Canonical Gibbs measure as stationary.

We introduce instead a space discretization of (1.4), see (2.8), whose solution can be defined globally. The infinite temperature version of this stochastic evolution was introduced in (cf. [26]). This is a n^d (complex) dimensional stochastic evolution that conserves the mass, and for any given initial mass, the Gibbs measure on the corresponding complex sphere defined by (2.10), discrete analogous of (1.3), is well defined and invariant. We prove in Section 2 that this Gibbs measure is the unique invariant measure, and that the distribution of the process starting from an arbirtary initial condition converges exponentially in total variation to this stationary measure (cf. Theorem 2 and Proposition 2.3). These results on the ergodicity of the stochastic dynamics contained in Section 2 are general and are valid for any $d, p > 1, \kappa = \pm 1, n$ and more general non-linearity. Let us emphasize that, to the best of our knowledge, the novelty of this dynamics is that it is the first *mass conserving* perturbation of the DNLS (Discrete Non-Linear Schrödinger), such that the canonical Gibbs measure is the *unique invariant measure*, determining the dynamics long-time behavior. We should mention that in Section 6 of [5], a mass conserving noise is proposed such that the Canonical Gibbs measure remains invariant by the dynamics. However, this dynamics is not studied, and [5] mainly concerns another dynamics, which does not conserve the mass and converges to the Grand Canonical measure. Moreover, the above-mentioned mass conserving dynamics is different from ours; in fact, a straightforward analysis suggests that our dynamics is more degenerate.

From Section 3 and after, we concentrate on the one-dimensional focusing cubic case with periodic boundary conditions ($d = 1, p = 3, \kappa = -1$). For the continuous model, the minimizers of the energy $\mathcal{H}(\psi)$ under the mass constrain $\mathcal{M}(\psi) = m$ are known explicitely [15]. These minimizers, that we denote by $Q_{m,L}(x)$, are unique up to translations and multiplication by a constant phase. To these minimizers correspond a class of standing waves $\psi(x,t) = e^{i\omega t}Q_{m,L}(x)$, which are solutions of (1.1), where the frequency ω is determined by m and L. We call *solitons* these ground state standing waves, in analogy to the traveling solitary waves of the dynamics in $\mathbb{R}$. If $m \leq \frac{\pi^2}{L}$ these solitons are constant in space, while for $m > \frac{\pi^2}{L}$ are given by the dnoidal elliptic Jacobi functions (cf. Appendix D for the definition, and Chapter 2,3 of [23] for properties of these functions) properly rescaled. These non-trivial solitons catch the 0-temperature behavior of the dynamics. The purpose of our work is to show that the solution $\psi_n(x,t)$ of the stochastic discrete dynamics, for large time t, large n, and small temperature β^{-1}, is close, in an opportune norm, to the continuous soliton. The result is contained in Theorem 3.2, where it is first taken the limit $t \to \infty$ then $n \to \infty$, rescaling the temperature with n, i.e. $\beta_n \sim \infty$ faster than n. This is a way to interpret the *soliton resolution conjecture* (SRC) in the periodic case, where there is no possibility for the energy to escape to infinity. Intuitively, in the periodic case, our dynamics in the zero temperature limit dissipate the excess of the energy without losing any mass, forcing the system to approach the ground state as $t \to \infty$. This mechanism is somehow mimicking the dynamics of DNLS in [8], where energy disperse to infinity via a "radiating" part of the field carrying arbitrarily small mass. In fact, our dynamics is partially motivated by [8], where Chatterjee proves a "probabilistic" version of the SRC. In particular, in Theorem 3.1 in [8] it is proven that *almost every* ergodic invariant measure satisflythe SRC in the time average sense. Our stochastic dynamics provides the uniqueness of the invariant ergodic measure and the time mixing property.

In Theorem 3.2 the limit for $t \to \infty$ follows from the ergodic and time mixing properties of the dynamics proven in Section 2. Then we have to prove that the discrete Gibbs measure (finite n) concentrate fast enough in a small neighborhood of the corresponding lowest energy configurations, that we call *discrete solitons*, who converge to the continuous one as $n \to \infty$. This relies on large deviation properties of the discrete Gibbs measure, proven in Section 5. These large deviations estimates are based on some precise large deviations of the uniform probability measure μ_m^n on the complex $2n$-dimensional sphere S_m^n, that we prove in Appendix A, and the discrete version of the Gagliardo-Nirenberg inequality, proven in Appendix C. The Gibbs measure has a density $\exp(-\beta_n \mathcal{H}_n)$ with respect to the uniform measure μ_m^n. Splitting the energy $\mathcal{H}_n = G_n - V_n$, where G_n is the kinetic part, and V_n the potential part (3.8), one can observe that a "typical" configuration w.r.t μ_m^n has kinetic energy $G_n \sim n^2$. The large deviations estimates in Section A, in particular Lemma A.1, combined with Gagliardo-Nirenberg inequality (C.8) yields: for $0 \leq a < 2$, the "entropy factor" behaves as $\mu_m^n(G_n \sim n^a) \sim \mu_m^n(\mathcal{H}_n \sim n^a) \sim e^{-(2-a)n \ln n}$. Therefore, taking into account the Boltzmann factor $\exp(-\beta_n \mathcal{H}_n)$, we have for $0 \leq a < 2$: $\mu_{\beta_n,m}^n(\mathcal{H}_n \sim n^a) \sim e^{-\beta_n n^a} e^{-(2-a)n \ln n}$. Optimizing this estimate on $a \in [0,2)$, if $\beta_n \sim O(1)$, then $a = 1$ is the optimal value and the Gibbs measure concentrates on rather rough configurations with $|\psi(j) - \psi(j-1)| \sim \frac{1}{\sqrt{n}}$, so that $G_n \sim n$. This corresponds to the fact that Wiener measure is concentrated on configurations of Hölder regularity less than $\frac{1}{2}$. Instead, if $\beta_n \sim O(n)$ we have that $a = 0$ is the optimal value and this suggests that $\mu_{\beta_n,m}^n$ to concentrates on smooth configurations (i.e., with $|\psi(j) - \psi(j-1)| \sim \frac{1}{n}$) with $\mathcal{H}_n \sim O(1)$. Notice that minimal energy configurations (the discrete solitons), have energy of order one as well. However, this scaling is not enough for this measure to

concentrate on a small neighborhood of discrete solitons, and we need to go further. Finally, thanks to large deviation estimate (A.10), we deduce in Theorem 5.1 that scaling $\beta_n >> n$ is sufficient.

In the last step of the proof, we show in Proposition 4.1 that if ψ_n is a configuration with energy close to $E_n^0(m)$, then its linear interpolation $\bar{\psi}_n$ (see (3.17)) is close to the continuous soliton $Q_{m,L}$ in H^1 norm (up to a translation and multiplication by a phase, see (3.18)), for n sufficiently large. In that regard, first we observe that having energy close to $E_0^n(m)$ means the configuration is smooth $G_n \sim O(1)$, thanks to the discrete Gagliardo-Nirenberg inequality. Subsequently, since for smooth configurations $\mathcal{H}_n(\psi_n)$ is close to $\mathcal{H}(\bar{\psi}_n)$ (See Corollary (4.3.1)), one can conclude by compactness of the minimizing sequence corresponding to the continuous minimization problem characterizing solitons (3.6).

Appendix B contains the proof of the hypoellipticity of the discrete stochastic dynamics, necessary for the proof of the ergodicity of Section 2. Since the real and complex part of our field are somehow symmetric in the noise, this makes the proof of the hypoellipticity more complicated than usual, and computing three nested commutators is necessary (see (B.4)).

The Gibbs measure of DNLS have been studied both in Mathematics (cf. [9], [8]) and Physics community (cf. [21] and references therein: in particular: [33], [34], [20]; See also [14]). In the physics community, one usually takes the Kinetic energy with a negative sign and study the measure corresponding to Hamiltonian (2.2), by taking $h = 1$. Although this regime is substantially different from ours, and does not correspond to discretization of a continuous profile anymore, interesting phenomena such as discrete breathers is observed (cf. [12], [36]).

In mathematics community, we can mention most notably [9], and [8] (cf. [22], for a review). In [9], the Hamiltonian (2.2) is considered such that $Nh^2 \to 0$, as $h \to 0$, and $N \to \infty$, where N denotes the number of particles, and h is the interparticle distance. These assumptions only seems natural in $d \geq 3$. In this regime, certain phase transition happens: When $\beta m^2 < \theta_c$ the Gibbs measure concentrates on configurations such that $\psi_n(j) \sim o(n)$, whereas for $\beta m^2 > \theta_c$ breather-like structures appears, where a single site has macroscopic mass.

In [8], the model is defined on the box $[0, nh]^d$, such that $h \to 0$, $n \to \infty$, with $nh \to \infty$. In this regime, the *microcanonical measure* corresponding to energy E concentrates on soliton-like configurations in $\mathbb{R}^d$.

Comparing our result with [9], and [8], we highlight the fact that different scaling among the parameters h, n leads to substantially different phenomena: In [9], $Nh^2 \to 0$ makes the Gradient term negligible and phase transition is a consequence of competition among potential energy and mass constraint. In [8], $nh \to \infty$, kinetic and potential energy becomes comparable; however, the mass per particle goes to zero in the limit, demonstrating the macroscopic infinite volume, facilitating escape of the energy to infinity and resulting in soliton like behavior. In contrast, in our case we take $n \to \infty$, and $nh = 1$, representing the finite macroscopic volume, and positive mass per particle in the macroscopic limit. This scaling yields a dominant kinetic energy for typical configurations on the sphere of constant mass, and rescaling β_n makes the kinetic and potential energy comparable.

In particular, these different scaling change our large deviation estimates (A.1), and (A.10) comparing to estimates in [8] (See Section 10 of [8]).

2 Stochastic Dynamics

Fix $n \in \mathbb{N}$, let $\chi = \mathbb{C}^{n^d}$ be the configuration space, and denote a typical element of χ by $\{\psi(x)\}_{x\in\widetilde{\mathbb{T}}_n^d}$, where $\widetilde{\mathbb{T}}_n = \{1, 2, \dots, n\}$ is the discrete Torus of size n. Equivalently, one can see a function on $\widetilde{\mathbb{T}}_n^d$, $\psi : \widetilde{\mathbb{T}}_n^d \to \mathbb{C}$, as the discretization of a function u on the d-dimensional torus of length size nh, $u : \mathbb{T}_{nh}^d \to \mathbb{C}$, with mesh size $h > 0$, i.e., $\psi(x) = u(hx)$, for $x \in \widetilde{\mathbb{T}}_n^d$. Then the discrete nonlinear Schrödinger equation (DNLS) is the following system of ODEs:

$$i\frac{d\psi(x,t)}{dt} = -\Delta_d\psi(x,t) + \kappa|\psi(x,t)|^{p-1}\psi(x,t), \qquad x \in \widetilde{\mathbb{T}}_n^d \tag{2.1}$$

where Δ_d is the d-dimensional discrete Laplacian:

$$\Delta_d\psi(x) = h^{-2}\sum_{|y-x|=1}\big(\psi(y) - \psi(x)\big).$$

These equations conserve the energy, given by the Hamiltonian

$$\mathcal{H}_n(\psi) = s\sum_{\substack{x,y\in\widetilde{\mathbb{T}}_n^d,\\ |x-y|=1}}\frac{h^{-2}}{2}|\psi(x) - \psi(y)|^2 + \frac{s\kappa}{p+1}\sum_{x\in\widetilde{\mathbb{T}}_n^d}|\psi(x)|^{p+1}, \tag{2.2}$$

and the mass, given by the ℓ^2 norm:

$$\mathcal{M}_n(\psi) = s\sum_{x\in\widetilde{\mathbb{T}}_n^d}|\psi(x)|^2. \tag{2.3}$$

Here $s > 0$ is a scaling parameter that we will choose opportunely later.

Denote $\psi(x) = \psi_r(x) + i\psi_i(x) = |\psi(x)|e^{i\theta(x)}$, the deterministic evolution equation (2.1) can be regarded as a Hamiltonian dynamics with the following generator:

$$\mathcal{A} = s^{-1}\sum_{x\in\widetilde{\mathbb{T}}_n^d}(\partial_{\psi_i(x)}\mathcal{H}_n)\partial_{\psi_r(x)} - (\partial_{\psi_r(x)}\mathcal{H}_n)\partial_{\psi_i(x)}. \tag{2.4}$$

Moreover, define the operator $\partial_{\theta(x)}$ acting on a suitable function $F : \chi \to \mathbb{C}$ as

$$\partial_{\theta(x)}F(\psi) = (\psi_r(x)\partial_{\psi_i(x)} - \psi_i(x)\partial_{\psi_r(x)})F(\psi). \tag{2.5}$$

Corresponding to a positive temperature $\beta^{-1} > 0$, define:

$$\mathcal{S} = \beta^{-1}\sum_{x\in\widetilde{\mathbb{T}}_n^d}e^{\beta\mathcal{H}_n}\partial_{\theta(x)}e^{-\beta\mathcal{H}_n}\partial_{\theta(x)}. \tag{2.6}$$

Fix $\beta > 0, \gamma > 0$, and consider the Markov process with values in χ, generated by

$$L = \mathcal{A} + \gamma\mathcal{S}, \tag{2.7}$$

where $\mathcal{S}$ and $\mathcal{A}$ are defined in (2.6) and (2.4). Since $\partial_{\theta(x)}\psi(x) = i\psi(x)$, we have

$$\mathcal{S}\psi(x) = -\psi(x)\left(\beta^{-1} + i\partial_{\theta(x)}\mathcal{H}_n(\psi)\right) = -\psi(x)\left(\beta^{-1} + is\ \mathrm{Im}[\psi^*(x)\Delta_d\psi(x)]\right)$$

gives us the explicit form of system of stochastic differential equations generated by (2.7):

$$\begin{aligned} d\psi(x,t) =& i[\Delta_d\psi(x,t) - \kappa|\psi(x,t)|^{p-1}\psi(x,t)]dt - \gamma\psi(x,t)(\beta^{-1} + i\partial_{\theta(x)}\mathcal{H}_n(\psi))dt \\ & - i\sqrt{2\gamma\beta^{-1}}\psi(x,t)dw(x,t), \qquad x \in \widetilde{\mathbb{T}}_n^d, \end{aligned} \tag{2.8}$$

where $\{w(x,t), x \in \widetilde{\mathbb{T}}_n\}$ are real independent Wiener processes.

We observed that $\mathcal{A}\mathcal{M}_n(\psi) = 0$, one can check that $\mathcal{S}\mathcal{M}_n(\psi) = 0$. Therefore, mass is a conserved quantity for the dynamics (2.7). Hence, if we assume the initial condition $\psi(0,t) = \psi_0 \in \mathbb{C}^{n^d}$, such that $\mathcal{M}_n(\psi_0) = m$, then our dynamics is confined in the compact manifold with $\mathcal{M}_n(\psi) = m$, which is a $(2n^d - 1)$-sphere. We denote this sphere by $S_{m,s}^n$:

$$S_{m,s}^n = \{\psi \in \mathbb{C}^{n^d} | \mathcal{M}_n(\psi) = m\}. \tag{2.9}$$

Proposition 2.1. *The generator L is hypoelliptic.*

The proof follows from Hörmander characterization, i.e., that the Lie algebra generated by $\{\mathcal{A}, \partial_{\theta(x)}, x \in \widetilde{\mathbb{T}}_n^d\}$ generates the tangent space of $S_{m,s}^n$. This is proven in Appendix B.

Let $d\mu_{m,s}^n$ be the uniform probability measure on $S_{m,s}^n$, one can define this measure as the projection of the Lebesgue measure on $S_{m,s}^n$, properly normalized. Define the canonical Gibbs measure with inverse temperature β on $S_{m,s}^n$ as

$$d\mu_{\beta,m,s}^n = \frac{1}{Z_n(\beta,m,s)} e^{-\beta\mathcal{H}_n(\psi)} d\mu_{m,s}^n, \tag{2.10}$$

Here $Z_n(\beta,m,s)$ is the partition function:

$$Z_n(\beta,m,s) = \int_{S_{m,s}^n} e^{-\beta\mathcal{H}_n(\psi)} d\mu_{m,s}^n. \tag{2.11}$$

Note that, since $\mathcal{H}_n$ is a smooth function on a compact set and therefore, bounded from below, $Z_n(\beta,m,s)$ is finite, and consequently, the existence of $d\mu_{\beta,m,s}^n$ is evident.

The observation that $\forall f \in C_b(S_{m,s}^n), \int_{\mathbb{C}^n} Lf d\mu_{\beta,m,s}^n = 0$, implies that $d\mu_{\beta,m,s}^n$ is an invariant measure for the dynamics (2.7)((2.8)). In fact, if we fix $m, \gamma, \beta > 0$ this measure is the unique invariant probability measure:

Theorem 2.1. *Fix the parameters $h, s, \gamma > 0$, the mass of the field m, and inverse temperature $\beta > 0$, the measure $d\mu_{\beta,m,s}^n$ is the unique invariant measure for the dynamics generated by (2.7).*

Proof. Without losing generality we can fix $h = s = 1$. Since the generator L is **hypoelliptic**, the stationary measure must have density w.r.t $d\mu_{m,s}^n$, and then also w.r.t $d\mu_{\beta,m,s}^n$. Denoting $f(\psi)$ the density w.r.t $d\mu_{\beta,m,s}^n$, it must satisfy the equation

$$0 = L^* f = (-\mathcal{A} + \gamma\mathcal{S})f, \tag{2.12}$$

where L^* denotes the adjoint of L in $L^2(d\mu_{\beta,m,s}^n)$. Since L is hypoelliptic, f is smooth and (2.12) is valid pointwise. Multiplying by f and integrating w.r.t $d\mu_{\beta,m,s}^n$, we have

$$0 = \gamma < f(-\mathcal{S})f >= \gamma \sum_{x \in \widetilde{\mathbb{T}}_n^d} < (\partial_{\theta(x)} f)^2 >, \tag{2.13}$$

where $<\cdot>$ denotes integration w.r.t $d\mu^n_{\beta,m,s}$. This means that $\partial_{\theta(x)} f = 0$ $d\mu^n_{\beta,m,s}$-a.e., and $\mathcal{A}f = 0$. We want to conclude that $f = 1$, $d\mu^n_{\beta,m,s}$-a.e.. Since $\partial_{\theta(x)} f = 0$ for any x, then $f = \tilde{f}(|\psi(x)|^2, x \in \widetilde{\mathbb{T}}^d_n)$. The operator $\mathcal{A}$ can be written as $\mathcal{A}^0 + \mathcal{A}^p$, with

$$\mathcal{A}^0 = \sum_{x \in \widetilde{\mathbb{T}}^d_n} \left\{ (\Delta_d \psi_i(x))\, \partial_{\psi_r(x)} - (\Delta_d \psi_r(x))\, \partial_{\psi_i(x)} \right\}, \tag{2.14}$$

and

$$\mathcal{A}^p = \kappa \sum_{x \in \widetilde{\mathbb{T}}^d_n} |\psi(x)|^{p-1} \left\{ \psi_i(x) \partial_{\psi_r(x)} - \psi_r(x) \partial_{\psi_i(x)} \right\} = \kappa \sum_{x \in \widetilde{\mathbb{T}}^d_n} |\psi(x)|^{p-1} \partial_{\theta(x)}. \tag{2.15}$$

It is immediate that $\mathcal{A}^p f = 0$, hence, $\mathcal{A}^0 f = 0$, pointwise. Let us denote $a(x) := |\psi(x)|^2$, and the canonical basis of $\mathbb{R}^d$ by $\{e_j\}_{j=1}^d$, then we have:

$$\begin{aligned}
0 = \mathcal{A}^0 f &= 2 \sum_{x \in \widetilde{\mathbb{T}}^d_n} \left\{ (\Delta_d \psi_i(x))\, \psi_r(x) - (\Delta_d \psi_r(x))\, \psi_i(x) \right\} \left[\partial_{a(x)} \tilde{f} \right] (|\psi(y)|^2, y \in \widetilde{\mathbb{T}}^d_n) \\
&= 2 \sum_{x \in \widetilde{\mathbb{T}}^d_n} \sum_{j=1}^d \nabla^j_+ (\psi_r(x)\psi_i(x - e_j) - \psi_i(x)\psi_r(x - e_j)) \left[\partial_{a(x)} \tilde{f} \right] (|\psi(y)|^2, y \in \widetilde{\mathbb{T}}^d_n) \\
&= 2 \sum_{x \in \widetilde{\mathbb{T}}^d_n} \sum_{j=1}^d [\psi_r(x)\psi_i(x - e_j) - \psi_i(x)\psi_r(x - e_j)] \left[(\partial_{a(x)} - \partial_{a(x-e_j)}) \tilde{f} \right] (|\psi(y)|^2, y \in \widetilde{\mathbb{T}}^d_n) \\
&= 2 \sum_{x \in \widetilde{\mathbb{T}}^d_n} \sum_{j=1}^d \sin\big(\theta_{x-e_j} - \theta_x\big) |\psi(x)||\psi(x - e_j)| \left[(\partial_{a(x)} - \partial_{a(x-e_j)}) \tilde{f} \right] (|\psi(y)|^2, y \in \widetilde{\mathbb{T}}^d_n),
\end{aligned} \tag{2.16}$$

where ∇^j_+ denotes the discrete gradient in the e_j direction $(\nabla^j_+ g)(x) = g(x+e_j) - g(x)$. Since this relation is true pointwise for any $\psi \in S^n_{m,s}$, by choosing a proper ψ (for example one can take θ_y equal to zero, for $y \in \widetilde{\mathbb{T}}^d_n$, except $\theta(x)$, and take $|\psi(y)| = 0$, for all $|y - x|_d = 1$, except $x - e_j$), we have that

$$(\partial_{a(x)} - \partial_{a(x-e_j)}) \tilde{f}(a(y), y \in \widetilde{\mathbb{T}}^d_n) = 0, \tag{2.17}$$

pointwise for every $x \in \widetilde{\mathbb{T}}^d_n$, and any $1 \le j \le d$.

From (2.17), we conclude that $(\partial_{a(x)} - \partial_{a(z)}) \tilde{f}(a(y), y \in \widetilde{\mathbb{T}}^d_n) = 0$ for any $x, z \in \widetilde{\mathbb{T}}^d_n$. This implies

$$\tilde{f}(a(y), y \in \widetilde{\mathbb{T}}^d_n) = F\left(\sum_{y \in \widetilde{\mathbb{T}}^d_n} a(y) \right) = F(m).$$

which yields the result. □

Remark 2.2. *Notice that the proof of Theorem 2, works for any other non-linearity of the form $F(|\psi|)$ with smooth F (at least C^2).*

By classical theorems in control theory, given the Hörmander condition, and the existence of a unique invariant measure with full support on $S^n_{m,s}$, it follows the strict positivity of the probability transition (cf. [16], proof of Theorem 2.1) and the following proposition:

Proposition 2.3. *Consider the dynamics which is generated by* (2.7)*, denote the law of this process by* $\mu_t^{\beta,n,m}$ *with initial condition* $\mu_0^{\beta,n,m} = \delta_{\psi_0}$*, where* ψ_0 *is an arbitrary element of* $S^n_{m,s}$*. There exist* $C(n,m,\psi_0)$ *and* γ_0*, such that*

$$\|\mu_t^{\beta,n,m} - d\mu^n_{\beta,m,s}\|_{TV} \leq Ce^{-\gamma_0 t}. \tag{2.18}$$

In particular, we have the weak convergence:

$$\mu_t^{\beta,n,m} \underset{t\to\infty}{\longrightarrow} \mu^n_{\beta,m,s}. \tag{2.19}$$

Proof. Since $\mu^n_{\beta,m,s}$ is the unique invariant measure (ergodicity), with full support (for any open set $A \subset S^n_{m,s}$, $\mu^n_{\beta,m}(A) > 0$), given the Hörmander condition we can use the result of [16], (proof of Theorem 2.1 in [16]) and deduce the strict positivity of the probability transition. Furthermore, having the strict positivity of the probability transition, compactness of the phase space, as well as the hypoellipticity of the generator, we can conclude by Theorem 8.9 of [2]. □

The novelty of the stochastic perturbation (2.6) can be described as follows: it's a mass-conserving white noise, such that the Gibbs measure is the unique invariant measure for the dynamics, and it provides good ergodic properties as in Theorem 2.19. This perturbation is quite "powerful" in the sense that its ergodic properties do not depend on the non-linearity, and we can consider either focusing or defocusing non-linearity. In either of these cases the long time behavior is given by the corresponding Gibbs measure. However, depending on the choice of parameters d, s, h, κ many interesting phenomena can be observed in the large scale limit. In the rest of this note, we focus on one particular case: one-dimensional focusing nonlinear Schrödinger Equation on the torus.

3 Large Scale Limit and Main Result

3.1 Preliminaries about periodic cubic nonlinear Schrödinger equation

In this section, we recall rather basic results about the focusing nonlinear Schrödinger equation (NLS) with periodic boundary conditions. Consider the following nonlinear cubic Schrödinger equation:

$$\begin{aligned} &i\partial_t\psi(x,t) = -\partial_{xx}\psi(x,t) - |\psi(x,t)|^2\psi(x,t),\ (t,x) \in \mathbb{R}_+ \times \mathbb{R}, \\ &\psi(x,0) = \psi_0(x),\ \psi_0 \in H^1(\mathbb{T}_L), \end{aligned} \tag{3.1}$$

where we assume the periodic boundary conditions by the definition of $H^1(\mathbb{T}_L)$ as:
$H^1(\mathbb{T}_L) = \{u \in H^1_{loc}(\mathbb{R},\mathbb{C}) |\ \forall x \in \mathbb{R}, \quad u(x+L) = u(x)\}$, with the following norms and inner product ($\bar{v}$ indicates the complex conjugate):

$$\|u\|_{L^p} = \left(\int_{\mathbb{T}_L} |u|^p dx\right)^{\frac{1}{p}},\ (u,v) = \int_{\mathbb{T}_L} u\bar{v}dx,\quad \|u\|_{H^1} = \left(\int_{\mathbb{T}_L} \left(|\partial_x u|^2 + |u|^2\right) dx\right)^{\frac{1}{2}}. \tag{3.2}$$

Global wellposedness of this problem is established in [4], [7]; in particular, $\forall t > 0$, $\psi(x,t) \in H^1(\mathbb{T}_L)$. Note that this equation has two important conserved quantities [1]: the energy or Hamiltonian $\mathcal{H}$, and L^2 norm or mass $\mathcal{M}$, defined by

$$\mathcal{H}(\psi) = \frac{1}{2}\int_{\mathbb{T}_L} |\partial_x \psi|^2 dx - \frac{1}{4}\int_{\mathbb{T}_L} |\psi|^4 dx, \qquad \mathcal{M}(\psi) = \int_{\mathbb{T}_L} |\psi|^2 dx. \tag{3.3}$$

One of the main features of this equation is the existence of a special class of solutions called the *"standing waves"* or *"periodic waves"*. These are *time periodic* solutions having the following form:

$$\psi(x,t) = e^{i\omega t} u(x). \tag{3.4}$$

If $\psi(x,t) = e^{i\omega t}u(x)$ be a solution of (3.1), then $u(x)$ should satisfy the following ODE, with periodic boundary condition:

$$u''(x) - \omega u(x) + |u(x)|^2 u(x) = 0. \tag{3.5}$$

Notice that the solution of (3.5) characterizes the minimum of the energy $\mathcal{H}(u)$, under the constrain $\mathcal{M}(u) = m$, where the frequency ω plays the role of Lagrange multiplier.

In general, we should consider complex valued solutions of (3.4). On the other hand, writing this solution as $u(x) = \rho(x)e^{i\theta(x)}$, the corresponding energy is given by

$$\mathcal{H}(u) = \frac{1}{2}\int_{\mathbb{T}_L} \left(|\rho'(x)|^2 + \rho(x)^2|\theta'(x)|^2\right) dx - \frac{1}{4}\int_{\mathbb{T}_L} |\rho(x)|^4 dx.$$

This shows that the minimum of the energy $\mathcal{H}(u)$, under the constrain $\mathcal{M}(u) = m$ is attained for $\theta(x) =$ constant. Consequently, this minimum are defined up to a constant phase and we can choose positive real solutions. Also notice that translations $u_y(x) = u(x+y)$ do not change energy and mass.

Here, if we fix the L, and assume u to be real-valued, and positive, and fix the mass of u to be $\mathcal{M}(u) = m$, then under these assumptions, (3.5) has a unique (up to a translation) smooth solution, this solution can be written in terms of Jacobi elliptic functions as $u(x) = \alpha\mathfrak{dn}(\lambda x, k)$, where $k \in (0,1), \alpha$, and $\lambda > 0$, $\omega > 0$ are uniquely determined by m, and L (cf. [15], [31], [13], cf. Appendix D for the definition of $\mathfrak{dn}$). We recall the following crucial result from [15], Proposition 3.2, which characterizes this solution as the minimizer of $\mathcal{H}(\psi)$ under the constraint that $\mathcal{M}(\psi) = m$.

Theorem 3.1. *Fix $m, L \in \mathbb{R}_+$, and consider the following minimization problem:*

$$E_0(m,L) := \inf\{\mathcal{H}(u) | \mathcal{M}(u) = m,\ u \in H^1(\mathbb{T}_L)\}, \tag{3.6}$$

then we have: $-\infty < E_0(m,L) < 0$, and

1. *If $0 < m \leq \frac{\pi^2}{L}$, then the constant function $Q_{m,L}(x) = (\frac{m}{L})^{\frac{1}{2}}$ is the unique minimizer of (3.6). This uniqueness is up to a multiplication by a constant phase.*
2. *If $\frac{\pi^2}{L} < m$, then $Q_{m,L}(x) := \alpha\mathfrak{dn}(\lambda x, k)$ is the unique minimizer of (3.6), up to a translation and multiplication by a constant phase. Moreover, $\alpha, \lambda > 0, k \in (0,1)$ are determined uniquely by m, L.*

[1] In fact, since this equation is completely integrable, we have infinite conserved quantities. However, most of the results in this note can be generalized to the sub-critical non-linearities that are not integrable, i.e., we can change the nonlinearity term in (3.1) into $|\psi|^{p-1}\psi$ with $1 \leq p < 5$. Notice that if $p \neq 3$, w do not have the explicit characterization of the Solitions

Furthermore, we have compactness of the minimizing sequence up to a phase shift and translation in $H^1(\mathbb{T}_L)$, i.e., for any sequence u_n in $H^1(\mathbb{T}_L)$, such that $\mathcal{H}(u_n) \to E_0(m,L)$, as $n \to \infty$, there is a subsequence u_{n_k}, and sequences $\gamma_k \in [0, 2\pi)$, and $x_k \in \mathbb{T}_L$, where $e^{i\gamma_k} u_{n_k}(. + x_k) \to Q_{m,L}$, in $H^1(\mathbb{T}_L)$.

Since each solution of (3.5) (and consequently a solution to (3.1)) corresponds to the minimization problem (3.6), by abusing the terminology, we use the term "*standing wave*" or *Soliton* for $Q_{m,L}$.

Notice that multiplying (3.5) by $\bar{u}$ and integrating, we obtain the following relation

$$E_0(m,L) = \frac{1}{4}\int_{\mathbb{T}_L} u^4(x)dx - \frac{\omega m}{2}.$$

that implies $\omega \geq \frac{1}{2m}\int_{\mathbb{T}_L} u^4(x)dx + \frac{m}{2L^2}$.

3.2 Stochastic perturbation of discrete focusing NLS

In this section, we are going to perturb the NLS (3.1), with the stochastic heat bath, which we defined in Section 2, namely (2.6). Without loosing generality, in order to simplify notation, we fix the macroscopic length $L = 1$. This means that we fix the following parameters $h = \frac{1}{n}, s = \frac{1}{n}, d = 1, p = 3, \kappa = -1$. Here, we briefly recall the dynamics of Section 2 in this particular setup, in order to set the notations.
Fix $n \in \mathbb{N}$, the configuration space is $\chi = \mathbb{C}^n$ and denote a typical element of χ by $\{\psi(x)\}_{x\in\widetilde{\mathbb{T}}_n}$, with $\widetilde{\mathbb{T}}_n = \{1, 2, \ldots, n\}$ is the discrete torus of size n. Equivalently, a function ψ on $\widetilde{\mathbb{T}}_n$ can be seen as discretization of a function u on a unit torus, $u : \mathbb{T} \to \mathbb{C}$, with mesh size $\frac{1}{n}$, i.e., $\psi(x) = u(\frac{x}{n})$, for $x \in \widetilde{\mathbb{T}}_n$. Then the discrete cubic focusing nonlinear Schrödinger equation (DNLS) is the following system of ODEs:

$$i\frac{d\psi(x,t)}{dt} = -\Delta\psi(x,t) - |\psi(x,t)|^2\psi(x,t), \tag{3.7}$$

where $\Delta\psi(x,t) = n^2(\psi(x+1) - 2\psi(x) + \psi(x-1))$, and we imposed periodic boundary condition $\psi(0) \equiv \psi(n)$. Notice that we define Δ such that formally in the limit $n \to \infty$, this definition coincides with the continuous Laplacian on a unit torus.

Similar to the continuous case, we have the energy or Hamiltonian $\mathcal{H}_n : \mathbb{C}^n \to \mathbb{R}$ as a conserved quantity, that is defined by:

$$\mathcal{H}_n(\psi) = \frac{1}{n}\sum_{x\in\widetilde{\mathbb{T}}_n} \frac{n^2}{2}|\psi(x) - \psi(x-1)|^2 - \frac{1}{4n}\sum_{x\in\widetilde{\mathbb{T}}_n} |\psi(x)|^4 = G_n(\psi) - V_n(\psi), \tag{3.8}$$

where we have denoted the kinetic energy $G_n(\psi)$ and the potential energy $V_n(\psi)$ as:

$$G_n(\psi) = \frac{1}{n}\sum_{x\in\widetilde{\mathbb{T}}_n} \frac{n^2}{2}|\psi(x) - \psi(x-1)|^2, \qquad V_n(\psi) = \frac{1}{4n}\sum_{x\in\widetilde{\mathbb{T}}_n} |\psi(x)|^4. \tag{3.9}$$

The other conserved quantity is given by the mass $\mathcal{M}_n : \mathbb{C}^n \to \mathbb{R}$, defined by:

$$\mathcal{M}_n(\psi) = \frac{1}{n}\sum_{x\in\widetilde{\mathbb{T}}_n} |\psi(x)|^2. \tag{3.10}$$

Notice that we scaled (3.8) and (3.10), such that in the limit as $n \to \infty$, we recover $\mathcal{H}$, and $\mathcal{M}$ formally.

The stochastic perturbation we consider will only conserve the mass. Recall $\psi(x) = \psi_r(x) + i\psi_i(x) = |\psi(x)|e^{i\theta(x)}$, the generators of the Hamiltonian and stochastic noise at temperature β^{-1} read

$$A_n = n \sum_{x \in \widetilde{\mathbb{T}}_n} (\partial_{\psi_i(x)} \mathcal{H}_n) \partial_{\psi_r(x)} - (\partial_{\psi_r(x)} \mathcal{H}_n) \partial_{\psi_i(x)}, \tag{3.11}$$

$$\mathcal{S}_n = \beta^{-1} \sum_{x \in \widetilde{\mathbb{T}}_n} e^{\beta \mathcal{H}_n} \partial_{\theta(x)} e^{-\beta \mathcal{H}_n} \partial_{\theta(x)}. \tag{3.12}$$

Fix $\beta > 0, \gamma > 0$, then the generator of the dynamics and corresponding system of stochastic partial differential equations with values in χ, are as follows:

$$L_n = A_n + \gamma \mathcal{S}_n, \tag{3.13}$$

$$\begin{aligned} d\psi(x,t) =& i[\Delta\psi(x,t) + |\psi(x,t)|^2 \psi(x,t)]dt - \gamma\psi(x,t)(\beta^{-1} + i\partial_{\theta(x)}\mathcal{H}_n(\psi))dt \\ & - i\sqrt{2\gamma\beta^{-1}}\psi(x,t)dw(x,t), \end{aligned} \tag{3.14}$$

where $\{w(x,t), x \in \widetilde{\mathbb{T}}_n\}$ are real independent Wiener processes.

Due to the mass conservation, having an initial condition $\psi(0,t) = \psi_0 \in \mathbb{C}^n$ such that $\mathcal{M}_n(\psi_0) = m$, our dynamic will be confined in the sphere $S_m^n = \{\psi \in \mathbb{C}^n | \mathcal{M}_n(\psi) = m\}$. Denote the uniform probability measure on S_m^n by $d\mu_m^n$, and define the canonical Gibbs measure with inverse temperature β on S_m^n as

$$d\mu_{\beta,m}^n = \frac{1}{Z_n(\beta, m)} e^{-\beta \mathcal{H}_n(\psi)} d\mu_m^n, \tag{3.15}$$

Here $Z_n(\beta, m) = \int_{S_m^n} e^{-\beta \mathcal{H}_n(\psi)} d\mu_m^n$. As we observed, $Z_n(\beta, m)$ is finite, and consequently, the existence of $d\mu_{\beta,m}^n$ is evident, since $\mathcal{H}_n(\psi)$ is bounded from below in S_m^n. However, one can find a lower bound for $\mathcal{H}_n(\psi)$, which is uniform in n, using a version of Gagliardo-Nirenberg inequality in the discrete periodic setup. This will be discussed broadly in the Section 4 and Appendix C.

Applying the result of Section 2, we have the following results: By Theorem 2 we know that $d\mu_{\beta,m}^n$ is the unique invariant measure for the dynamics (3.13)((3.14)). Moreover, Proposition (2.3) states that if μ_t denotes the law of the process at time $t \geq 0$, generated by (3.13), with initial condition $\mu_0 = \delta_{\psi_0}$, where ψ_0 is an arbitrary element of S_m^n, then there exist $C(n, m, \psi_0)$ and γ_0 such that

$$\|\mu_t - \mu_{\beta,m}^n\|_{TV} \leq Ce^{-\gamma_0 t}. \tag{3.16}$$

If we run our dynamics for a long time, then take the limit of large n and small temperature β^{-1} properly, we end-up near *Solitons* or *standing waves* ($Q_{1,m}$ from Theorem 3.1), with probability one. Notice that here we can take the limit in β and n simultaneously, where we scale β by a factor of $\vartheta(n)$. In order to make these words rigorous, and connect the discrete setup to the continuous one, we need to introduce some notations. For any $\psi_n \in \mathbb{C}^n$, we define its linear interpolation $\bar{\psi}^n : \mathbb{T} \to \mathbb{C}$, on a unit torus by

$$\bar{\psi}_n(y) = \psi_n\big([ny]\big)\big([ny] + 1 - ny\big) + \psi_n\big([ny] + 1\big)\big(ny - [ny]\big), \quad \forall y \in \mathbb{T}, \tag{3.17}$$

where $[ny]$ denotes the greatest integer less than ny. Denote $H^1(\mathbb{T})$ by $H^1_{per}([0,1]) = H^1_{per}$. For $x \in \mathbb{T}$, let τ_x denotes the translation operator on H^1_{per}, i.e. $(\tau_x f)(y) = f(x+y)$, then, in order to deal with the phase multiplication and translation, define the following seminorm as in [8]:

$$\forall f, g \in H^1_{per}, \quad \|f-g\|_{\tilde{H}^1_{per}} := \inf_{\gamma\in[0,2\pi], x\in\mathbb{T}} \|e^{i\gamma}\tau_x f - g\|_{H^1_{per}}. \tag{3.18}$$

In the following we set $Q_{1,m} =: Q_m$. Now we can state the main theorem of this section:

Theorem 3.2. *Fix $m > 0$, $\gamma > 0$, and $\beta > 0$, let $\beta_n = \vartheta(n)\beta$, where $\vartheta(n) > 0$ is a scaling parameter, such that*

$$\lim_{n\to\infty} \frac{\vartheta(n)}{n} \to \infty. \tag{3.19}$$

Let $\mu_t^{\beta_n,n,m}$ be the law of the process given by its generator (3.13)*, with the initial condition $\mu_0^{n,m} = \delta_{\psi_0^{n,m}}$, where $\psi_0^{n,m}$ is a sequence of proper initial conditions, i.e., for all n, $\psi_0^{n,m} \in S^n_m$. Then $\forall\, \epsilon > 0$, we have:*

$$\lim_{n\to\infty}\lim_{t\to\infty} \mu_t^{\beta_n,n,m}\Big(\|\bar{\psi}_n - Q_m\|_{\tilde{H}^1_{per}} < \epsilon\Big) \to 1. \tag{3.20}$$

We briefly sketch the proof: we have already proved that $\mu^{\beta_n}_{n,m}$ is the limit in t of $\mu_t^{\beta_n,n,m}$. Consequently, all we have to prove is that

$$\lim_{n\to\infty} \mu^{\beta_n}_{n,m}\Big(\|\bar{\psi}_n - Q_m\|_{\tilde{H}^1_{per}} < \epsilon\Big) \to 1. \tag{3.21}$$

We can prove that the measure $\mu^{\beta_n}_{n,m}$ concentrates all its mass on the (discrete) configurations having close to the minimal energy, when we send temperature to zero with a proper speed. It turns out that the proper speed here is to scale β by $\vartheta(n)$, satisfying (3.19). Finally, we show that if a configuration has energy close to the minimal, it will be close to Q_m in the sense of (3.20). This can be done by adapting certain form of concentration compactness argument to the discrete setup.

Remark 3.1. ***About the exchange of limits in*** (3.20)***:*** *In the evolution equation* (3.14) *the drift term $\partial_{\theta(x)}\mathcal{H}_n(\psi) = \frac{1}{n}\mathcal{I}m[\psi(x)\Delta\psi^*(x)]$ would became very singular when $n \to \infty$ keeping the temperature positive. But with $\beta_n \to \infty$ fast enough the solution should became enough regular in space so that the corresponding limit as $n \to \infty$ should be given by the continuous deterministic NLS. This will be investigated in a future work [18]. The later suggests that one could study the joint limit $n, t \to \infty$, with $t_n = n^\alpha t$. We address the case $t_n \ll \beta_n$ in [18]. However, the case $t_n \gg \beta_n$ seems more challenging.*

4 Discrete "Soliton"

As we already observed in Theorem 3.1, the function Q_m (Solitons) can be characterized as the minimizer of a certain variational problem, where we have the compactness of the minimizing sequence. Therefore, one can observe that for a function $u \in H^1_{per}([0,1])$, with $\mathcal{M}(u) = m$, having "close to minimal" energies, means the function itself is close to Q_m in the following sense:

Lemma 4.1. *Assume $u \in H^1_{per}$ and $\mathcal{M}(u) = m$, then, $\forall\epsilon > 0$, $\exists\, \delta > 0$, such that if $\mathcal{H}(u) \le E_0(m) + \delta$, then there exists $\gamma \in [0, 2\pi]$, $x \in [0,1]$, such that $\|e^{i\gamma}u(.+x) - Q_m\|_{H^1_{per}} < \epsilon$, equivalently $\|u(x) - Q_m\|_{\tilde{H}^1_{per}} < \epsilon$.*

Proof. This is straightforward by the compactness of the minimizing sequence in Theorem 3.1. □

Similar to (3.6), fix $n > 1, m > 0$ and define $E_0^n(m)$ as follows:

$$E_0^n(m) := \inf\{\mathcal{H}_n(\psi_n) | \psi_n \in S_m^n\}. \tag{4.1}$$

Since $\mathcal{H}_n(\psi_n)$ is a continuous function from the compact set S_m^n to $\mathbb{R}$, the image of this function is compact, hence, $-\infty < E_0^n(m)$, and this infimum is achieved in a compact set, which will be called the set of "discrete Solitons" and denoted by $\emptyset \neq \mathcal{Q}_m^n \subset S_m^n$. By the same argument as in the continuous case, discrete solitons are real-valued and positive up to a constant phase.

For $\psi_n \in S_m^n$, we define $\|\psi_n\|^p_{\ell^p(\widetilde{\mathbb{T}}_n)} = \frac{1}{n}\sum_j |\psi_n(j)|^p$. Then we can write $\mathcal{H}_n(\psi_n) = G_n(\psi_n) - \frac{1}{4}\|\psi_n\|^4_{\ell^4(\widetilde{\mathbb{T}}_n)}$, and by using the discrete Gagliardo-Nirenberg inequality (C.11), we have:

$$-\theta(m) \leq E_0^n(m) < 0, \tag{4.2}$$

where $\theta(m) = \frac{C^2}{64}m^3 + \frac{C}{4}m^2$. First inequality is a direct consequence of (C.11), and the second one can be deduced by considering the constant function $\psi_n(x) = \sqrt{m}$, for all $x \in \widetilde{\mathbb{T}}_n$.

From (4.2) we establish a simple but useful lemma:

Lemma 4.2. *For every $\epsilon > 0$, there exists $C(m, \epsilon)$, such that for all $n \in \mathbb{N}$ and $\psi_n \in S_n^m$, with $\mathcal{H}_n(\psi_n) \leq E_0^n(m) + \epsilon$, we have $G_n(\psi_n) \leq C(m, \epsilon)$.*

Proof. Consider the inequality (4.2), and (C.11); denote $x = G_n(\psi_n)^{\frac{1}{2}}$ so $x \geq 0$. If $\mathcal{H}_n(\psi_n) \leq E_0^n + \epsilon$ then, thanks to (C.11) we have

$$E_0^n(m) + \epsilon \geq x^2 - c'm^{\frac{3}{2}} - c'm^2 \implies x \leq C^{\frac{1}{2}}(m, \epsilon). \tag{4.3}$$

where $C^{\frac{1}{2}}(m, \epsilon)$ is given by $C^{\frac{1}{2}}(m, \epsilon) = \frac{c'm^3 + \sqrt{c'm^3 + 4(c'm^2 + E_0^n(m) + \epsilon)}}{2}$, where the expression under the square root is clearly positive, thanks to the expression of $\theta(m)$. □

Lemma 4.2 states that if the energy is "small" ($O(1)$), then the configuration should be "smooth" i.e., $G_n \sim O(1)$.

In the rest of this section, we prove that $\bar{\psi}_n$, the linear interpolation of a configuration ψ_n, is arbitrarily close to Q_m in $\tilde{H}^1_{per}$, if we take n sufficiently large, and the energy of ψ_n, $\mathcal{H}_n(\psi_n)$, sufficiently close to $E_0^n(m)$. The proof relies on the fact that the configurations with close to minimal energies are smooth in the sense that their linear interpolation's norm (L^p, H^1 or even the energy) is close to the corresponding discrete norms. This result heavily depends on the inequality of Appendix C. We begin by stating this result:

Proposition 4.1. *Fix $m > 0$, for any $\epsilon > 0$, there exists $\eta(\epsilon)$ and $N_0(\epsilon)$, such that for $n > N_0(\epsilon)$, if $\mathcal{H}_n(\psi_n) \leq E_0^n(m) + \eta$, then we have: $\|\bar{\psi}_n - Q_m\|_{\tilde{H}^1_{per}} < \epsilon$.*

We divide the proof of (4.1), into a couple of simple lemmas. The advantage of the linear interpolation (3.17) is that it conserves the kinetic energy, i.e., $G_n(\psi_n) = \frac{1}{2}\int_0^1 |\partial_x \bar{\psi}_n|^2$. But unfortunately, in general, we have $\|\psi_n\|_{\ell^p(\widetilde{\mathbb{T}}_n)} \geq \|\bar{\psi}_n\|_{L^p}$ for $p \geq 1$, thanks to the Jensen inequality. Consequently, in general we have $\mathcal{H}_n(\psi_n) \leq \mathcal{H}(\bar{\psi}_n)$ and $\mathcal{M}_n(\psi_n) \geq \mathcal{M}(\bar{\psi}_n)$. However, the following lemma helps to establish the fact that these quantities are "close", for configurations with near minimal energies.

Lemma 4.3. *For all $n \in \mathbb{N}$, if $\psi_n \in S_m^n$ we have:*

$$\left| \|\bar{\psi}_n\|^p_{L^p(\mathbb{T})} - \|\psi_n\|^p_{\ell^p(\widetilde{\mathbb{T}}_n)} \right| \leq \frac{p(2G_n(\psi_n))^{1/2}\left(m^{1/2} + G_n(\psi_n)^{1/2}\right)^{p-1}}{n},$$

Proof. We have $\psi_n \in S_n^m$, and define $\ell_n = \min\{|\psi(x)| \,\big|\, x \in \widetilde{\mathbb{T}}_n\}$, clearly $\ell_n \leq \sqrt{m}$. Moreover, for any $x \in \widetilde{\mathbb{T}}_n$, we have:

$$|\ell_n - \psi_n(x)| \leq \sum_{j=1}^{n} |\psi_n(j) - \psi_n(j-1)| \leq \sqrt{n}\left(\sum_{j=1}^{n} |\psi_n(j) - \psi_n(j-1)|^2\right)^{1/2} = \sqrt{2G_n(\psi_n)},$$

where we used a Cauchy Schwartz inequality. Therefore, we can deduce that

$$\sup_x |\psi(x)| \leq c_1 = m^{\frac{1}{2}} + (2G_n(\psi_n))^{\frac{1}{2}}.$$

Moreover, thanks to the definition of $\bar{\psi}_n(y)$, we have $|\bar{\psi}_n(y)| \leq c_1$, for all $y \in \mathbb{T}$. Then we can simply compute:

$$\begin{aligned}
\left| \|\bar{\psi}_n\|^p_{L^p(\mathbb{T})} - \|\psi_n\|^p_{\ell^p(\widetilde{\mathbb{T}}_n)} \right| \leq & \sum_{x=0}^{n-1} \int_{\frac{x}{n}}^{\frac{x+1}{n}} \left| |\psi_n(x)|^p - |\bar{\psi}_n(y)|^p \right| dy \\
\leq & p c_1^{p-1} \sum_{x=1}^{n} \int_{\frac{x-1}{n}}^{\frac{x}{n}} \left| |\psi_n(x)| - |\bar{\psi}_n(y)| \right| dy \qquad (4.4) \\
\leq & \frac{p c_1^{p-1}}{n} \sum_{x=1}^{n} |\psi_n(x) - \psi_n(x-1)| \leq \frac{p c_1^{p-1} (2G_n(\psi_n))^{\frac{1}{2}}}{n},
\end{aligned}$$

where the first inequality comes from the definition, in the second inequality we used the fact that $\psi_n(x)$ and $\bar{\psi}_n(y)$ are bounded uniformly in x and y, and in the third inequality we used the definition of $\bar{\psi}_n(y)$:

$$\big| |\psi_n(x)| - |\bar{\psi}_n(y)| \big| \leq |\psi_n(x) - \bar{\psi}_n(y)| \leq |\psi_n(x) - \psi_n(x+1)|.$$

Notice that the last inequality in (4.4) is obtained as above. □

As a straightforward consequence of Lemma 4.3, we can deduce the following corollaries:

Corollary 4.3.1. *For any $c > 0$, there exist $C_1(c, m)$, such that for every $n \in \mathbb{N}$, and $\psi_n \in S_m^n$, such that $G_n(\psi_n) < c$, then $|\mathcal{H}(\bar{\psi}_n) - \mathcal{H}_n(\psi_n)| \leq \frac{C_1(c,m)}{n}$.*

Proof. Thanks to the definition of $\bar{\psi}_n$ (3.17), the weak derivative of $\bar{\psi}_n$ is given as follows: for any $y \in [0,1]$, if $\frac{x}{n} \leq y < \frac{x+1}{n}$ with $x \in \widetilde{\mathbb{T}}_n$, then $\partial_y \bar{\psi}_n(y) = n(\psi_n(x+1) - \psi_n(x))$. Therefore, we have: $\frac{1}{2}\int_0^1 |\partial_y \bar{\psi}_n|^2 = \frac{n}{2}\sum_{x=1}^{n} |\psi(x) - \psi(x-1)|^2 = G_n(\psi_n)$. Hence, we have:

$$\left|\mathcal{H}_n(\psi_n) - \mathcal{H}(\bar{\psi}_n)\right| = \frac{1}{4}\left| \|\bar{\psi}_n\|^4_{L^4(\mathbb{T})} - \|\psi_n\|^4_{\ell^4(\tilde{\widetilde{\mathbb{T}}}_n)} \right|,$$

and we can conclude thanks to Lemma 4.3. □

Corollary 4.3.2. *For any $\delta > 0$, there exist $\eta > 0$ and $N_0(\delta)$, such that for $n > N_0(\delta)$ if $\mathcal{H}_n(\psi_n) \leq E_0^n(m) + \eta$ then $\mathcal{H}(\bar{\psi}_n) \leq E_0^n(m) + \delta$.*

Proof. It follows immediately from Corollary 4.3.1 and Lemma 4.2. □

Proposition 4.2.

$$\lim_{n\to\infty} E_0^n(m) \to E_0(m). \tag{4.5}$$

Proof. Before proceeding, we emphasize the fact that all the constant $c, c1, c2, c', \dots$ are independent of n in this proof.
Recall the definition of Q_m as the minimizer of (3.6). Moreover, recall the definition of the set of discrete Solitions $\mathcal{Q}_m^n$, as the set of mininizer of (4.1). Take $q_n \in \mathcal{Q}_m^n$, notice that thanks to the inequality $|\psi_n(x) - \psi_n(x-1)| \geq ||\psi_n(x)| - |\psi_n(x-1)||$, we can take q_n to be real-values and positive. Then we have: $\mathcal{H}(Q_m) = E_0(m)$, and for all n, $\mathcal{H}_n(q_n) = E_0^n(m)$.
thanks to Lemma 4.2 there exists $c > 0$ uniform in n, such that $G_n(q_n) \leq c$. Therefore, we can use the result of Corollary 4.3.1, and deduce that there exists C_1 independent of n, such that:

$$|\mathcal{H}(\bar{q}_n) - \mathcal{H}_n(q_n)| \leq \frac{C_1}{n}. \tag{4.6}$$

For any $\psi \in H^1_{per}([0,1])$, define $\lambda_n(\psi)$ as follows:

$$\lambda_n(\psi) = \left(\frac{m}{\mathcal{M}(\psi)}\right)^{\frac{1}{2}}. \tag{4.7}$$

In particular, let $\lambda_n = \lambda(\bar{q}_n)$ and observe that for n sufficiently large, $|\lambda_n^2 - 1| \leq \frac{c_0}{n}$, with c_0 independent of n, thanks to Lemma 4.3. More precisely, we can take $c_0 = \frac{2\tilde{c}}{m}$, for n sufficiently large, where $\tilde{c}$ is given by Lemma 4.3. Now, if we use the definition of $\mathcal{H}$, for n sufficiently large we obtain:

$$\begin{aligned} |\mathcal{H}(\lambda_n\bar{q}_n) - \mathcal{H}(\bar{q}_n)| &\leq |\lambda_n^2 - 1| \int_0^1 \frac{|\partial_y \bar{q}_n(y)|^2}{2} dy + \frac{|\lambda_n^4 - 1|}{4} \int_0^1 |\bar{q}_n(y)|^4 dy \\ &\leq \frac{c_1}{n}, \end{aligned} \tag{4.8}$$

where c_1 is independent of n, and we used the estimate $|\lambda_n^2 - 1| \leq \frac{c_0}{n}$; moreover, in order to treat the first term, we take advantage of the fact that $G_n(q_n) = \int_0^1 \frac{|\partial_y \bar{q}_n(y)|^2}{2} dy \leq c$. Lastly, the second term is bounded as follows: we used the bound $\|q_n\|_{\ell^4(\widetilde{\mathbb{T}}_n)} \leq c'$ (thanks to Lemma 4.2 and (4.2)), then we conclude by using the fact $|\|q_n\|_{\ell^4(\widetilde{\mathbb{T}}_n)} - \|\bar{q}_n\|_{L^4(\mathbb{T})}| \leq \frac{\tilde{c}'}{n}$ which is a direct consequence of Lemma 4.3.
Notice that $\mathcal{M}(\lambda_n\bar{q}_n) = m$; therefore, $E_0(m) \leq \mathcal{H}(\lambda_n\bar{q}_n)$. Combining this fact with (4.6) and (4.8), for n large enough we have:

$$E_0(m) \leq E_0^n(m) + \frac{c''}{n}, \tag{4.9}$$

where c'' is a constant independent of n, and we used the fact that $\mathcal{H}_n(q_n) = E_0^n(m)$.

On the other hand, recall that Q_m is smooth, real-valued and non-negative thanks to Theorem 3.1. Define $Q_m^n : \widetilde{\mathbb{T}}_n \to \mathbb{C}$ as $Q_m^n(x) = Q_m(\frac{x}{n})$, for $x \in \widetilde{\mathbb{T}}_n$. Let

$$\tilde{\lambda}_n := \left(\frac{m}{\mathcal{M}_n(Q_m^n)}\right)^{\frac{1}{2}}.$$

Thank to the properties of Q_m (in particular the fact that Q_m is smooth with bounded H^1 and L^4 norm), for n large enough we have:

$$|\tilde{\lambda}_n^2 - 1| \leq \frac{c_2}{n}, \tag{4.10}$$

where one can take $c_2 = \frac{4\|Q_m\|_{L^\infty}\|Q_m'\|_{L^\infty}}{m}$ (Q' denotes the derivative of Q). Moreover, since Q_m is smooth, $G_n(Q_m^n)$ and $V_n(Q_m^n)$ are bounded uniformly in n by $\frac{\|Q_m'\|_{L^\infty}}{2}$, and $\frac{\|Q_m\|_{L^\infty}^4}{4}$, respectively. Hence, thanks to (4.10) for n sufficiently large we have:

$$|\mathcal{H}_n(\tilde{\lambda}_n Q_m^n) - \mathcal{H}_n(Q_m^n)| \leq \frac{c_3}{n}. \tag{4.11}$$

Again, since Q_m is at least C^3, by a simple computation we get for n sufficiently large:

$$\begin{aligned} &|\mathcal{H}_n(Q_m^n) - \mathcal{H}(Q_m)| \leq \frac{1}{2}\sum_{x=1}^{n}\int_{\frac{x-1}{n}}^{\frac{x}{n}} \left| n^2\Big|Q_m\big(\frac{x}{n}\big) - Q_m\big(\frac{x-1}{n}\big)\Big|^2 - |\partial_y Q_m(y)|^2\right| dy + \\ &\frac{1}{4}\sum_{x=1}^{n}\int_{\frac{x-1}{n}}^{\frac{x}{n}} \Big||Q_m\big(\frac{x}{n}\big)|^4 - |Q_m(y)|^4\Big| dy \leq \frac{\|Q_m''\|_{L^\infty}\|Q_m'\|_{L^\infty}}{n} + \frac{\|Q_m'\|_{L^\infty}\|Q_m\|_{L^\infty}^3}{n} \\ &\leq \frac{c_4}{n}. \end{aligned} \tag{4.12}$$

Therefore, combining the estimates (4.11) and (4.12), and recalling the fact that $\mathcal{H}(Q_m) = E_0(m)$, we have for n large enough:

$$|\mathcal{H}_n(\tilde{\lambda} Q_m^n) - \mathcal{H}(Q_m)| \leq \frac{c}{n} \implies E_0^n(m) \leq E_0(m) + \frac{c}{n}, \tag{4.13}$$

where we used the fact that $\mathcal{M}(\tilde{\lambda}_n Q_m^n) = m$, hence $E_0^n(m) \leq \mathcal{H}_n(\tilde{\lambda}_n Q_m^n)$. Finally, taking the limit of $n \to \infty$ in (4.13) and (4.9), properly (lim sup and lim inf, respectively), we deduce the result (4.5). □

We finish this section by proving the Proposition 4.1:

Proof of Proposition 4.1. In consequence of corollary 4.3.2 and Proposition 4.5 we have that for any $\delta > 0$, there exist $\eta > 0$ and $N_0(\delta)$, such that for $n > N_0(\delta)$ if $\mathcal{H}_n(\psi_n) \leq E_0^n(m) + \eta$ then $\mathcal{H}(\bar{\psi}_n) \leq E_0(m) + 2\delta$. Define

$$\lambda_{\bar{\psi}_n} = \big(\frac{m}{\mathcal{M}(\bar{\psi}_n)}\big)^{\frac{1}{2}} \geq 1,$$

so that $\mathcal{M}(\lambda_{\bar{\psi}_n}\bar{\psi}_n) = m$. Furthermore by Lemma 4.3 $\lambda_{\bar{\psi}_n} \to 1$. We also have that

$$\begin{aligned} \mathcal{H}\left(\lambda_{\bar{\psi}_n}\bar{\psi}_n\right) &= \mathcal{H}(\bar{\psi}_n) + (\lambda_{\bar{\psi}_n}^2 - 1)G_n(\psi_n) - (\lambda_{\bar{\psi}_n}^4 - 1)V(\bar{\psi}_n) \\ &\leq E_0(m) + 2\delta + (\lambda_{\bar{\psi}_n}^2 - 1)C, \end{aligned} \tag{4.14}$$

where we bounded G_n thanks to Lemma 4.2. By lemma 4.1, we have $\|\lambda_{\bar{\psi}_n}\bar{\psi}_n - Q_m\|_{\tilde{H}^1_{per}} < \epsilon/2$, and since

$$\|\lambda_{\bar{\psi}_n}\bar{\psi}_n - \bar{\psi}_n\|_{\tilde{H}^1_{per}} \leq |\lambda_{\bar{\psi}_n}^2 - 1|^{1/2}\|\bar{\psi}_n\|_{\tilde{H}^1_{per}} < \epsilon/2$$

for n large enough, we conclude the proof. □

5 Large Deviation Estimates

In Proposition 4.1, we proved that if the energy $\mathcal{H}_n(\psi_n)$ is sufficiently close to the minimal energy $E_n^0(m)$ for n sufficiently large, then the linear interpolation of a configuration ψ_n is close to Q_m in $\tilde{H}^1$-norm. In this section, we prove that the measure $\mu^n_{\beta_n,m}$ (3.15) concentrate on configurations with minimal energy as $n \to \infty$, if we set $\beta_n = \vartheta(n)\beta$, where $\vartheta(n)$ satisfies (3.19). As m is fixed in this section, we will drop it from the notations.

Theorem 5.1. *For any $\epsilon > 0$, we have:*

$$\lim_{n\to\infty} \mu^n_{\beta_n}(\mathcal{H}_n(\psi_n) \leq E_0^n + \epsilon) = 1, \tag{5.1}$$

The proof of Theorem (5.1) depends on two large deviation estimates for the uniform probability measure $d\mu^n_m$ that are proven in appendix A:

1. For any n and any $0 < \mathfrak{g}$ we have:

$$\mu^n_m(G_n(\psi_n) < \mathfrak{g}) \leq \exp(-2n\ln n)\left(\frac{2\mathfrak{g}}{m}\right)^{n-1} 2^n, \tag{5.2}$$

   This bound is proven in Lemma A.1, following the same spirit as in [8], Section 10. However, because of our special scaling in G_n, one should follow the dependence of the rate function on n carefully, in contrast to the estimate in [8]. This lemma provides the aforementioned upper bounds. Combining (5.2) with Gagliardo-Nirenberg inequality, we can deduce a suitable upper bound for $\mathcal{H}_n$.

2. For any $\epsilon > 0$, there exists $d = d(\epsilon)$ and $N(\epsilon)$, such that for $n \geq N(\epsilon)$:

$$\mu^n_m(\mathcal{H}_n(\psi_n) < E_0^n + \epsilon) \geq d^n e^{-2n\ln n}. \tag{5.3}$$

   This is proven in Lemma A.2.

We will proceed as follows: first, we state a proof of (5.1), when $\beta_n = \beta n \ln n$. This proof is quite simple and illustrates how does the above estimates are involved. Finally, we prove the general case $\beta_n = \vartheta(n)\beta$.

Proof of Theorem 5.1 with $\beta_n = \beta n \ln n$. Assume $0 < \epsilon < 1$, in order to prove (5.1), it is sufficient to prove that:

$$p_n := \frac{\int_{S^n_m} 1_{\{\mathcal{H}_n(\psi_n) - E_n^0 \geq \epsilon\}} e^{-\beta_n \mathcal{H}_n(\psi_n)} d\mu^n_m}{\int_{S^n_m} 1_{\{\mathcal{H}_n(\psi_n) - E_n^0 < \frac{\epsilon}{2}\}} e^{-\beta_n \mathcal{H}_n(\psi_n)} d\mu^n_m} \underset{n\to\infty}{\longrightarrow} 0. \tag{5.4}$$

Thanks to the lower bound (5.3), there exists $d > 0$ and N_1, such that for $n > N_1$ we have:

$$\begin{aligned}\left(\int_{S^n_m} 1_{\{\mathcal{H}_n(\psi_n) - E_n^0 < \frac{\epsilon}{2}\}} e^{-\beta_n \mathcal{H}_n(\psi_n)} d\mu^n_m\right)^{-1} &\leq e^{\beta_n(E_0^n + \frac{\epsilon}{2})}\left[\mu^n_m\left(\mathcal{H}_n(\psi_n) - E_n^0 \leq \frac{\epsilon}{2}\right)\right]^{-1} \\ &\leq e^{n\ln n\left(\beta(E_0^n + \frac{\epsilon}{2}) + 2\right) - n\ln d}.\end{aligned} \tag{5.5}$$

Let $c > \epsilon + \frac{2}{\beta}$; recall Lemma 4.2 and let $c' = C(m,c) > 0$, which is given by this lemma. By using (5.2) there exists N_2, such that for $n > N_2$:

$$\begin{aligned}
&\int_{S_m^n} 1_{\{\mathcal{H}_n - E_n^0(m) \geq \epsilon\}} e^{-\beta_n \mathcal{H}_n} d\mu_m^n = \\
&\int_{S_m^n} 1_{\{c > \mathcal{H}_n - E_n^0(m) \geq \epsilon\}} e^{-\beta_n \mathcal{H}_n} d\mu_m^n + \int_{S_m^n} 1_{\{\mathcal{H}_n - E_n^0(m) \geq c\}} e^{-\beta_n \mathcal{H}_n} d\mu_m^n \\
&\leq e^{-\beta n \ln n (E_0^n(m) + \epsilon)} \mu_n^m(G_n < c') + e^{-\beta n \ln n (E_0^n(m) + c)} \\
&\leq e^{-n \ln n \left(\beta(E_n^0(m) + \epsilon) + 2\right) + n \ln(4c'/m)} + e^{-\beta n \ln n (E_0^n(m) + c)},
\end{aligned} \tag{5.6}$$

where in the second line we used the fact that $\{c > \mathcal{H}_n - E_0^n(m) \geq \epsilon\} \subset \{G_n \leq c'\}$, thanks to the choice of c', see Lemma 4.2. Finally, taking $N > \{N_1, N_2\}$, and combining (5.5) and (5.6), gives us the following:

$$0 \leq p_n \leq e^{-n\left(\ln n \frac{\beta\epsilon}{2} - \ln(4c'/dm)\right)} + e^{-n\left(\ln n \left(c\beta - \frac{\beta\epsilon}{2} - 2\right) - \ln d\right)} \underset{n\to\infty}{\longrightarrow} 0, \tag{5.7}$$

thanks to the choice of c. □

Now we prove Theorem 5.1 in the general situation with $\beta_n = \beta\vartheta(n)$ satisfying (3.19):

Proof of Theorem 5.1 with $\beta_n = \vartheta(n)\beta$. Fix $0 < \epsilon < 1$, (the other cases will be straightforward). In order to prove (5.1), it is sufficient to prove (5.4). As for (5.5), there exist $d > 0$ and N_1 such that for any $n > N_1$:

$$\left(\int_{S_m^n} 1_{\{\mathcal{H}_n - E_n^0 < \frac{\epsilon}{2}\}} e^{-\beta_n \mathcal{H}_n} d\mu_m^n \right)^{-1} \leq e^{\beta\vartheta(n)(E_0^n + \frac{\epsilon}{2})} (d)^{-n} e^{2n \ln n}. \tag{5.8}$$

Let us decompose the numerator of (5.4) into two parts and denote them by q_n and q_n':

$$\begin{aligned}
\int_{S_m^n} 1_{\{\mathcal{H}_n - E_n^0 \geq \epsilon\}} e^{-\beta_n \mathcal{H}_n} d\mu_m^n = &\underbrace{\int_{S_m^n} 1_{\{\ln n > \mathcal{H}_n - E_n^0 \geq \epsilon\}} e^{-\beta_n \mathcal{H}_n} d\mu_m^n}_{q_n} + \\
&\underbrace{\int_{S_m^n} 1_{\{\mathcal{H}_n - E_n^0 \geq \ln n\}} e^{-\beta_n \mathcal{H}_n} d\mu_m^n}_{q_n'}.
\end{aligned} \tag{5.9}$$

We simply bound $q_n' \leq e^{-\beta_n(E_n^0 + \ln n)}$ and observe that:

$$\left(\int_{S_m^n} 1_{\{\mathcal{H}_n - E_n^0(m) < \frac{\epsilon}{2}\}} e^{-\beta_n \mathcal{H}_n} d\mu_m^n \right)^{-1} q_n' \leq e^{-\ln n(\beta\vartheta(n) - 2n)} e^{\beta\vartheta(n)\frac{\epsilon}{2} - n \ln d} \underset{n\to\infty}{\longrightarrow} 0, \tag{5.10}$$

as $n \to \infty$, where we used the fact that d is a constant independent of n, as well as the condition $\lim_{n\to\infty} \frac{\vartheta(n)}{n} = \infty$.
Now we treat the term corresponding to q_n, thanks to (5.2). First, observe that for any $E_n^0(m) < a \leq n$, thanks to the inequality (C.11), if we have $\mathcal{H}_n(\psi_n) \leq a$, we can deduce $G_n(\psi_n) \leq c_1(m) + 2a$, where $c_1(m)$ is a constant independent of n. (in fact, $c_1(m) = (\tilde{c}^2 m^3 + 2\tilde{c}m^2)$, with $\tilde{c} = \frac{C}{4}$ and C is the constant in (C.11)); consequently, we have for any $E_n^0(m) < a \leq n$:

$$\mu_m^n(\mathcal{H}_n \leq a) \leq \mu_m^n\big(G_n \leq 2a + c_1(m)\big). \tag{5.11}$$

Recall (5.2): for any $0 < \alpha < 2n$ denote $\alpha_o = \frac{2\alpha}{m}$, then

$$\mu_m^n\big(G_n \le \alpha\big) \le 2^n \alpha_o^{n-1} e^{-2n\ln n} \tag{5.12}$$

holds. Therefore, for large n thanks to (5.11), (5.12), we have for any $E_n^0(m) < a \le n$:

$$\mu_m^n(\mathcal{H}_n \le a) \le 2^n e^{-2n\ln n}\Big(\frac{4a + 2c_1(m)}{m}\Big)^{n-1}. \tag{5.13}$$

Take $h > 0$ independent of n, let $N = \frac{\ln n}{h}$. Then we have for n sufficiently large:

$$\begin{aligned}
&q_n = \sum_{j=0}^{N-1} \int_{S_m^n} 1_{\{E_0^n+\epsilon+jh \le \mathcal{H}_n < E_0^n+\epsilon+(j+1)h\}} e^{-\beta_n \mathcal{H}_n} d\mu_m^n \le \\
&\sum_{j=0}^{N-1} e^{-\beta_n(E_0^n+\epsilon+jh)} \mu_m^n\Big(E_0^n+\epsilon+jh \le \mathcal{H}_n < E_0^n+\epsilon+(j+1)h\Big) \le \\
&2^n e^{-2n\ln n} \sum_{j=0}^{N-1} e^{-\beta_n(E_0^n+\epsilon+jh)} \Big(\frac{4}{m}\Big)^{n-1} e^{(n-1)\ln\left(E_0^n+\epsilon+(j+1)h+\frac{c_1(m)}{2}\right)} = \\
&2^n \Big(\frac{4}{m}\Big)^{n-1} e^{-2n\ln n} e^{\beta_n(h+\frac{c_1(m)}{2})} \times \\
&\sum_{j=1}^{N} \exp\bigg(-\beta_n(E_0^n+\epsilon+jh+\frac{c_1(m)}{2}) + (n-1)\ln\Big(E_0^n+\epsilon+jh+\frac{c_1(m)}{2}\Big)\bigg),
\end{aligned} \tag{5.14}$$

where we take advantage of the estimate (5.13) in the second line. Notice that the term $E_0^n + \frac{c_1(m)}{2} > 0$ thanks to the lower bound (4.2). Recall that $\beta_n = \vartheta(n)\beta$, with $\lim_{n\to\infty} \frac{\vartheta(n)}{n} \to \infty$. Therefore, for n sufficiently large $-\beta_n + \frac{n-1}{x} < 0$, for any $x \in [h, 2\ln n]$. However, the later expression is the derivative of $-\beta_n x + (n-1)\ln(x)$, hence, this function is decreasing on the interval $[h + E_0^n + \epsilon + \frac{c_1(m)}{2}, 2\ln n]$ for any n sufficiently large, and $-\beta_n x + (n-1)\ln(x)$ achieves its minimum at $x = h + E_0^n + \epsilon + \frac{c_1(m)}{2}$ in the aforementioned interval. Combining this fact with (5.14) we get:

$$\begin{aligned}
&q_n \le 2^n \Big(\frac{4}{m}\Big)^{n-1} e^{-2n\ln n} e^{\beta_n(h+\frac{c_1(m)}{2})} \times \\
&N \exp\bigg(-\beta_n(E_0^n+\epsilon+h+\frac{c_1(m)}{2}) + \ln\Big(E_0^n+\epsilon+h+\frac{c_1(m)}{2}\Big)\bigg) \\
&= 2^n \Big(\frac{4}{m}\Big)^{n-1} e^{-2n\ln n} N \exp\bigg(-\beta_n(E_0^n+\epsilon) + (n-1)\ln\Big(E_0^n+\epsilon+h+\frac{c_1(m)}{2}\Big)\bigg).
\end{aligned} \tag{5.15}$$

Notice that $0 < (E_0^n + \epsilon + h + \frac{c_1(m)}{2}) < (\epsilon + h + \frac{c_1(m)}{2}) =: c'$, and c' is a constant independent of n. Combining the later estimate (5.15), with (5.8) we get for n sufficiently large:

$$\bigg(\int_{S_m^n} 1_{\{\mathcal{H}_n - E_n^0(m) < \frac{\epsilon}{2}\}} e^{-\beta_n \mathcal{H}_n} d\mu_m^n\bigg)^{-1} q_n \le e^{-\beta_n \frac{\epsilon}{2}} \bigg(d^{-n} 2^n (\frac{4}{m})^{n-1} (c')^{n-1}\bigg) \frac{\ln n}{h} \to 0, \tag{5.16}$$

as $n \to \infty$. Notice that (5.16) is evident, since the first term $e^{-\beta_n n}$ is super-exponentially small thanks to the assumption $\beta_n = \beta\vartheta(n)$ with $\lim_{n\to\infty} \frac{\vartheta(n)}{n} = \infty$ and the second term is bounded by $e^{n\tilde{c}}$, where $\tilde{c}$ is a constant independent of n. Finally, recalling the decomposition (5.9) and combining (5.16) with (5.10) gives us (5.4) and finishes the proof. □

Finally, the proof of Theorem 3.2 is a direct consequence of Proposition 4.1, and Theorem 5.1:

Proof of Theorem 3.2. Fix $\epsilon > 0$, thanks to the Proposition (2.3), in particular (2.19), we have:

$$\lim_{n\to\infty}\lim_{t\to\infty}\mu_t^{\beta_n,n,m}\Big(\|\bar{\psi}_n - Q_m\|_{\tilde{H}^1_{per}} < \epsilon\Big) = \lim_{n\to\infty}\mu^n_{\beta_n,m}\Big(\|\bar{\psi}_n - Q_m\|_{\tilde{H}^1_{per}} < \epsilon\Big).$$

On the other hand let us take $\delta = \delta(\epsilon)$, which is given by Proposition 4.1, then for all $n > N_0(\epsilon)$ thank to this proposition we have:

$$1 \geq \mu^n_{\beta_n,m}\Big(\|\bar{\psi}_n - Q_m\|_{\tilde{H}^1_{per}} < \epsilon\Big) \geq \mu^n_{\beta_n,m}\Big(|\mathcal{H}_n(\psi_n) - E^n_0(m)| < \delta\Big). \tag{5.17}$$

However, notice that $\lim_{n\to\infty}\mu^n_{\beta_n,m}\Big(|\mathcal{H}_n(\psi_n) - E^n_0(m)| < \delta\Big) = 1$, thanks to Theorem 5.1, in particular (5.1), and this finishes the proof of Theorem 3.2, i.e.,(3.20). □

Appendix A Some large deviations for the uniform probability on the sphere

We collect here some large deviation estimates concerning μ^n_m, the uniform probability on the complex n-dimensional sphere S^n_m, and in particular the estimates (5.2) and (5.3). Note that in this appendix we slightly change our notations and denote the elements of $\mathbb{C}^n$ by z or $\underline{z}$ instead of ψ.

Lemma A.1. *For any $n \in \mathbb{N}$, let $0 < \mathfrak{g}$. For any $0 < \delta < 1$ we have:*

$$\mu^n_m(G_n(\psi_n) < \mathfrak{g}) \leq \frac{1}{\delta(1-\delta)^{n-1}}\left(\frac{2\mathfrak{g}}{m}\right)^{n-1}\exp(-2n\ln n). \tag{A.1}$$

Proof. Let $\{Z_j\}_{j=1}^{\infty}$, be a sequence of i.i.d standard complex normal random variables on $(\Omega, \mathcal{F}, \mathbb{P})$, i.e, for any $n > 0$, the probability density function of $(Z_1, \ldots, Z_n)$ is given by:

$$f(\underline{z}) = \prod_{j=1}^{n}\frac{e^{-|z_j|^2}}{\pi}, \qquad \underline{z} := (z_1, \ldots, z_n) \in \mathbb{C}^n. \tag{A.2}$$

Consequently, the random vector $\{\Psi_n(j) = \frac{\sqrt{mn}Z_j}{\left(\sum_\ell |Z_\ell|^2\right)^{1/2}}, \quad j = 1, \ldots, n\}$ is distributed uniformly on S^m_n. For $k \in \widetilde{\mathbb{T}}_n$, let the random variable $\hat{Z}_k$ be defined as the Fourier transform of $Z_1, \ldots, Z_n$:

$$\hat{Z}_k = \frac{1}{\sqrt{n}}\sum_{j=1}^{n} e^{-2\pi j\frac{k}{n}}Z_j. \tag{A.3}$$

Notice that $(\hat{Z}_1, \ldots, \hat{Z}_n)$, has the same distribution as $(Z_1, \ldots, Z_n)$. Moreover, we have the following identities thanks to the properties of discrete Fourier transform:

$$\sum_{j=1}^{n}|Z_j|^2 = \sum_{k=1}^{n}|\hat{Z}_k|^2, \tag{A.4}$$

$$\sum_{j=1}^{n} |Z_j - Z_{j-1}|^2 = \sum_{k=1}^{n} \omega_k^2 |\hat{Z}_k|^2, \tag{A.5}$$

where $\omega_k = 2|\sin\big(\pi\frac{k}{n}\big)|$. Denote

$$\mathfrak{g}_o := \frac{2\mathfrak{g}}{n^2 m}, \tag{A.6}$$

and take $0 < \lambda$ such that, $0 < 1 - \mathfrak{g}_o\lambda$. By using Chebyshev's inequality, as well as (A.4) and (A.5), we have:

$$\begin{aligned}
&\mu_n^m(G_n(\psi_n) \le \mathfrak{g}) = \mathbb{P}\bigg(n^2 m \frac{\sum_{j=1}^n |Z_j - Z_{j-1}|^2}{\sum_{j=1}^n |Z_j|^2} \le 2\mathfrak{g}\bigg) \\
&= \mathbb{P}\Big(\sum_{k=1}^{n} \omega_k^2 |\hat{Z}_k|^2 \le \sum_{k=1}^{n} |\hat{Z}_k|^2 \mathfrak{g}_o\Big) = \mathbb{P}\bigg(\exp\Big(-\sum_{k=1}^{n} \lambda\big(\omega_k^2 - \mathfrak{g}_o\big)|\hat{Z}_k|^2\Big) \ge 1\bigg) \\
&\le \mathbb{E}\bigg(\exp\Big(-\sum_{k=1}^{n} \lambda\big(\omega_k^2 - \mathfrak{g}_o\big)|\hat{Z}_k|^2\Big)\bigg) = \prod_{k=1}^{n} \mathbb{E}\left(\exp\left(-\lambda(\omega_k^2 - \mathfrak{g}_o)|\hat{Z}_1|^2\right)\right) \\
&= \prod_{k=1}^{n} \frac{1}{\lambda\big(\omega_k^2 - \mathfrak{g}_o\big) + 1}.
\end{aligned} \tag{A.7}$$

Notice that in the first line, we used the fact that Ψ_n is uniformly distributed on S_n^m, and in the last line we used the fact that $\hat{Z}_k$ are independent complex Gaussian variable with the same distribution as Z_i, as well as the choice of λ, which permits us to compute the last expectation. We emphasize the fact that the last bound holds for any $0 < \lambda < \mathfrak{g}_o^{-1} = \frac{n^2 m}{2\mathfrak{g}}$, which can depend on n. In fact, our choice of λ depends on n.
Before proceeding, let us recall the following trigonometric identity:

$$\prod_{k=1}^{n-1} \sin\left(\frac{\pi k}{n}\right) = \frac{n}{2^{n-1}}, \implies \frac{1}{\prod_{k=1}^{n-1} \omega_k^2} = \frac{1}{n^2}. \tag{A.8}$$

For any $0 < \delta < 1$, let us take $\lambda = \frac{(1-\delta)}{\mathfrak{g}_o}$. Notice that we have $1 - \lambda\mathfrak{g}_o = \delta$. Thanks to the choice of λ, by using (A.7) and (A.8), we obtain

$$\begin{aligned}
\mu_n^m(G_n(\psi_n) < \mathfrak{g}) \le \prod_{k=1}^{n} \frac{1}{\lambda\big(\omega_k^2 - \mathfrak{g}_o\big) + 1} &\le \frac{1}{\delta} \prod_{k=1}^{n-1} \frac{1}{\lambda\omega_k^2} \le \frac{1}{\delta} \frac{1}{(1-\delta)^{n-1}} \mathfrak{g}_o^{n-1} \frac{1}{n^2} \\
&= \frac{1}{\delta(1-\delta)^{n-1}} \exp(-2n\ln n) \left(\frac{2\mathfrak{g}}{m}\right)^{n-1}.
\end{aligned} \tag{A.9}$$

□

Notice that the bound (5.2) corresponds to the choice $\delta = 1/2$.

We obtain now the lower bound (5.3), indicating that set of configurations with close to minimal energy is "large enough".

Lemma A.2. *For any $\epsilon > 0$, there exist $N(\epsilon)$, and a constant $\mathrm{c} = \mathrm{c}(\epsilon)$, independent of n, such that for $n > N(\epsilon)$ we have:*

$$\mu_m^n(\mathcal{H}_n < E_0^n + \epsilon) \ge \mathrm{c}^n e^{-2n\ln n}. \tag{A.10}$$

Proof. Denote by Q a discrete Soliton, we have that $E_0^n = \mathcal{H}_n(Q) = G_n(Q) - V_n(Q)$. we know from the results of section 4 that Q is uniformly bounded in n as well as $G_n(Q)$ and $V_n(Q)$. Observe that

$$\begin{aligned}\{\psi \in S_m^n : \mathcal{H}_n(\psi) < E_0^n + \epsilon\} \ \supset \\ \{\psi \in S_m^n : |G_n(\psi) - G_n(Q)| \le \epsilon/2, |V_n(\psi) - V_n(Q)| \le \epsilon/2\}\,.\end{aligned} \tag{A.11}$$

Consequently, we need to construct a neighborhood $\tilde{\mathcal{A}} \subset S_m^n$ of Q that is contained in the set on the RHS of (A.11), and such that $\mu_m^n(\tilde{\mathcal{A}}) \ge \mathrm{c}^n e^{-2n\ln n}$ for some constant depending on ε.

Let us identify $\mathbb{C}^n \sim \mathbb{R}^{2n}$, and denote the corresponding real components of Q by $(q_1, \dots, q_{2n})$, and the components of a generic $\psi \in S_m^n \sim \mathbb{S}^{2n}_{\sqrt{nm}}$ as $(x_1, \dots, x_{2n})$. We can choose the discrete Soliton Q, such that $q_{2n} \ge q_j \ge 0$.

For any small $\delta > 0$, define the set $\tilde{A}_\delta \subset \mathbb{R}^{2n-1}$ as follows:

$$\tilde{A}_\delta = \left\{ \underline{\xi} \in \Big[-\frac{\delta}{2n}, \frac{\delta}{2n} \Big]^{2n-1} \Big| \quad \Big| \sum_{j=1}^{2n-1} \xi_j q_j \Big| \le \frac{q_{2n}\delta}{2\sqrt{n}} \right\}. \tag{A.12}$$

The volume of this set can be easily estimated by

$$\mathrm{vol}(\tilde{A}_\delta) \ge \frac{1}{3}\left(\frac{\delta}{n}\right)^{2n-1}. \tag{A.13}$$

We postpone the proof of (A.13) later.

We now define our neighborhood of Q as

$$\tilde{\mathcal{A}}_\delta = \left\{ \underline{x} \in \mathbb{S}^{2n}_{\sqrt{nm}} : x_j = q_j + \xi_j, j = 1 \dots, 2n-1;\ \underline{\xi} \in \tilde{A}_\delta \right\}. \tag{A.14}$$

Notice that if $\underline{x} \in \tilde{\mathcal{A}}_\delta$, we have automatically that $x_{2n} = (nm - \sum_{j=1}^{2n-1}(q_j + \xi_j)^2)^{\frac{1}{2}}$. Furthermore, we have that

$$|x_{2n} - q_{2n}| \le \frac{2\delta}{\sqrt{n}}. \tag{A.15}$$

It is easy to check that if $\underline{x} \in \tilde{\mathcal{A}}_\delta$ then $|V_n(\underline{x}) - V_n(Q)| \le 2c\delta/n$, where c is a constant independent of n. About the gradients term, denoting $\xi_{2n} = x_{2n} - q_{2n}$, we have

$$\begin{aligned} G_n(\underline{\xi}) = \frac{n}{2}\sum_{j=1}^{n-1}(\xi_{i+1} - \xi_i)^2 + \frac{n}{2}(\xi_n - \xi_1)^2 \\ + \frac{n}{2}\sum_{j=1}^{n-2}(\xi_{n+i+1} - \xi_{n+i})^2 + \frac{n}{2}(\xi_{2n} - \xi_{2n-1})^2 + \frac{n}{2}(\xi_{2n} - \xi_{n+1})^2 \le 6\delta^2, \end{aligned} \tag{A.16}$$

and

$$\begin{aligned} |G_n(\underline{x}) - G_n(Q)| &= \left| n\sum_j (q_j - q_{j-1})(\xi_j - \xi_{j-1}) + G_n(\underline{\xi}) \right| \\ &\le (2G_n(Q))^{1/2}\left(2G_n(\underline{\xi})\right)^{1/2} + G_n(\underline{\xi}) \\ &\le C\left(2G_n(\underline{\xi})\right)^{1/2} + G_n(\underline{\xi}) \le C'\delta. \end{aligned} \tag{A.17}$$

It follows that, choosing $\delta < \epsilon/2C'$, the set $\tilde{\mathcal{A}}_\delta$ is contained in the set defined in (A.11).

In order to compute $\mu_m^n(\tilde{\mathcal{A}}_\delta)$ we use the following change of variable formula for any measurable $f : \mathbb{S}_r^{2n} \to \mathbb{R}$: (cf. Appendix A of [1])

$$\int_{\mathbb{S}_r^{2n}} f(\underline{x}) d\sigma_r^{2n}(\underline{x}) = \frac{r^2}{2r^{2n} n V(\mathbb{B}_1^{2n})} \int_{\mathbb{B}_r^{2n-1}} \frac{f(\underline{y}, \sqrt{r^2 - \|\underline{y}^2\|}) + f(\underline{y}, -\sqrt{r^2 - \|\underline{y}^2\|})}{\sqrt{r^2 - \|\underline{y}\|^2}} dy_1 \dots dy_{2n-1}, \tag{A.18}$$

where $\|.\|$ denotes the Euclidean norm in $\mathbb{R}^{2n-1}$, and $V(\mathbb{B}_1^{2n}) = \frac{\pi^n}{n!}$ denotes the volume of the unit ball. Applying the above formula and noticing that $nm \geq \sqrt{nm - \|\underline{y}\|^2}$, we have

$$\begin{aligned}
\mu_m^n(\tilde{\mathcal{A}}_\delta) &= \frac{mn!}{2(nm)^n \pi^n} \int_{\mathbb{B}_{\sqrt{nm}}^{2n-1}} \frac{1_{\tilde{\mathcal{A}}_\delta}(\underline{y}, \sqrt{nm - \|\underline{y}\|^2}) + 1_{\tilde{\mathcal{A}}_\delta}(\underline{y}, -\sqrt{nm - \|\underline{y}\|^2})}{\sqrt{nm - \|\underline{y}\|^2}} dy_1 \dots dy_{2n-1} \\
&\geq \frac{n!}{2n(nm)^n \pi^n} \int_{\mathbb{B}_{\sqrt{nm}}^{2n-1}} \left[1_{\tilde{\mathcal{A}}_\delta}(\underline{y}, \sqrt{nm - \|\underline{y}\|^2}) + 1_{\tilde{\mathcal{A}}_\delta}(\underline{y}, -\sqrt{nm - \|\underline{y}\|^2}) \right] dy_1 \dots dy_{2n-1} \\
&= \frac{n!}{n(nm)^n \pi^n} \int_{-\frac{\delta}{n}}^{\frac{\delta}{n}} \cdots \int_{-\frac{\delta}{n}}^{\frac{\epsilon}{n}} 1_{\tilde{A}_\delta}(\underline{\xi}) d\xi_1 \dots d\xi_{2n-1} \geq \frac{n!}{n(nm)^n \pi^n} \frac{1}{3} \left(\frac{\delta}{n} \right)^{2n-1}
\end{aligned} \tag{A.19}$$

and by Stirling approximation we have the desired lower bound.

□

Proof of (A.13). Let $\{\xi_j\}_{j=1}^\infty$ be a sequence of i.i.d random variables uniformly distributed on $[-\frac{\delta}{2n}, \frac{\delta}{2n}]$. Thanks to Chebyshev's inequality we get:

$$\begin{aligned}
&\mathbb{P}\Big(\Big| \sum_{j=1}^{2n-1} \xi_j q_j \Big| \leq \frac{q_{2n}\delta}{2\sqrt{n}} \Big) = 1 - \mathbb{P}\Big(\Big| \sum_{j=1}^{2n-1} \xi_j q_j \Big|^2 > \Big(\frac{q_{2n}\delta}{2\sqrt{n}} \Big)^2 \Big) \geq \\
&1 - \frac{4n}{q_{2n}^2 \delta^2} \mathbb{E}\Big(\Big| \sum_{j=1}^{2n-1} \xi_j q_j \Big|^2 \Big) \geq 1 - \frac{4n}{\delta^2} \mathbb{E}(\xi_1^2) \sum_{j=1}^{2n-1} \left(\frac{q_j}{q_{2n}} \right)^2 = 1 - \frac{1}{3n} \sum_{j=1}^{2n-1} \left(\frac{q_j}{q_{2n}} \right)^2 \geq 1 - \frac{2}{3},
\end{aligned} \tag{A.20}$$

where we used our choice of the discrete Soliton $0 \leq q_j \leq q_{2n}$. □

We conclude this section mentioning some more precise limits on the large deviations for the uniform measure on the sphere, with a matching lower bound for Large deviation estimate (A.1). *These results are not used for proving theorem 5.1 and theorem 3.2, so their proof would be published in a future work [18].* For $0 \leq a < 2$ we have:

$$\limsup_{n\to\infty} \frac{1}{n} \ln\Big(e^{(2-a)n \ln n} \mu_m^n(G_n < cn^a) \Big) \leq \ln\Big(\frac{2c}{m} \Big). \tag{A.21}$$

$$\liminf_{n\to\infty} \frac{1}{n} \ln\Big(e^{(2-a)n \ln n} \mu_m^n(G_n \leq cn^a) \Big) \geq \ln\Big(\frac{2c}{m} \Big). \tag{A.22}$$

Appendix B Hypoellipticity

In this section, we prove that the generator (2.7), is hypoelliptic, and therefore the invariant measure has a smooth density. Notice that we add the subscript n, to emphasize the dependence on n.

Lemma B.1. *Recall the operator $L_n = A_n + S_n$, where $A_n = \mathcal{A}$ and $S_n = \mathcal{S}$ are given by (2.4), (2.6), respectively. Then L_n is hypoelliptic. Consequently, the invariant measure has smooth density with respect to $d\mu_n^m$.*

Proof. Let us the fix the parameters $h = s = \gamma = 1$, the proof for other cases is similar. In order to prove this lemma, it is sufficient to show that L_n satisfies the so-called Hörmander's condition. Then the hypoellipticity, and smoothness of the invariant measure follow by the Hörmander's Theorem (hypoellipticity follows from Thorem 22.2.1 of [19], for a general review one can also see [2], and [17]). We prove this condition in the case $d = 1$ in details, the generalization to higher dimensions is a matter of messier algebra (We comment on this at the end of the proof). Let us denote $Y_0 = A_n$ and $Y_x = \partial_{\theta(x)}$ for $x \in \widetilde{\mathbb{T}}_n$. L_n satisfies the Hörmander's condition if the Lie algebra generated by the family

$$\{Y_x\}_{x=1}^n, \quad \{[Y_x, Y_y]\}_{x,y=0}^n, \quad \{[[Y_x, Y_y], Y_z]\}_{x,y,z=0}^n, \dots,$$

has full rank (here $2n-1$) at every point $\psi \in S_n^m$.
Let us define the following notation: for $x, y \in \widetilde{\mathbb{T}}_n$ and symbols i, r, we define $\mathcal{R}_{x^i,y^r}$, $\mathcal{R}_{x^i,y^i}$, $\mathcal{R}_{x^r,y^r}$, and $\mathcal{R}_{x^r,y^i}$ as the following rotations:

$$\begin{aligned} &\mathcal{R}_{x^r,y^r} = \psi_r(x)\partial_{\psi_r(y)} - \psi_r(y)\partial_{\psi_r(x)}, \quad \mathcal{R}_{x^r,y^i} = \psi_r(x)\partial_{\psi_i(y)} - \psi_i(y)\partial_{\psi_r(x)}, \\ &\mathcal{R}_{x^i,y^r} = \psi_i(x)\partial_{\psi_r(y)} - \psi_r(y)\partial_{\psi_i(x)}, \quad \mathcal{R}_{x^i,y^i} = \psi_i(x)\partial_{\psi_i(y)} - \psi_i(y)\partial_{\psi_i(x)}. \end{aligned} \tag{B.1}$$

We can rewrite $\partial_{\theta(x)}$, and the Hamiltonian operator A_n in terms of these rotations:

$$\begin{aligned} &\partial_{\theta(x)} = \mathcal{R}_{x^r,x^i}, \\ &A_n = \sum_{x \in \widetilde{\mathbb{T}}_n} \mathcal{R}_{(x+1)^r,x^i} + \mathcal{R}_{x^r,(x+1)^i} - 2\mathcal{R}_{x^r,x^i} + \kappa|\psi(x)|^{p-1}\mathcal{R}_{x^r,x^i}. \end{aligned} \tag{B.2}$$

Observe that for any $\alpha_1, \alpha_2, \alpha_3, \alpha_4 \in \{x^\mu | x \in \widetilde{\mathbb{T}}_n, \mu \in \{r, i\}\}$, (these indices are of the form x^i,x^r), we have (recall $[a, b] = ab - ba$):

$$\begin{aligned} [\mathcal{R}_{\alpha_1,\alpha_2}, \mathcal{R}_{\alpha_3,\alpha_4}] =& \delta_{\alpha_1,\alpha_4}\mathcal{R}_{\alpha_2,\alpha_3} + \delta_{\alpha_2,\alpha_3}\mathcal{R}_{\alpha_1,\alpha_4} - \delta_{\alpha_1,\alpha_3}\mathcal{R}_{\alpha_2,\alpha_4} - \delta_{\alpha_2,\alpha_4}\mathcal{R}_{\alpha_1,\alpha_3} = \\ &\sum_{i,j=1}^{4} \delta_{\{i+j-5\}}\mathcal{R}_{\alpha_i\alpha_j}\delta_{\alpha_k,\alpha_l}(-1)^{i+j+1}, \end{aligned} \tag{B.3}$$

where $\{k, l\} := \{1, 2, 3, 4\} \setminus \{i, j\}$.
We rewrite the following commutators in terms of these rotations for every $x \in \widetilde{\mathbb{T}}_n$:

$$\begin{aligned} &\partial_{\theta(x)} = \mathcal{R}_{x^r,x^i}, \\ &\mathcal{A}_x := [A_n, \partial_{\theta(x)}] = \mathcal{R}_{x^r,(x+1)^r} + \mathcal{R}_{x^i,(x+1)^i} - \mathcal{R}_{(x-1)^r,x^r} - \mathcal{R}_{(x-1)^i,x^i}, \\ &\mathcal{A}_{x,x+1} := [[A_n, \partial_{\theta(x)}], \partial_{\theta(x+1)}] = \mathcal{R}_{x^r,(x+1)^i} - \mathcal{R}_{x^i,(x+1)^r}, \\ &\mathcal{A}_{x,x+1,x} := [[[A_n, \partial_{\theta(x)}], \partial_{\theta(x+1)}], \partial_{\theta(x)}] = \mathcal{R}_{x^r,(x+1)^r} + \mathcal{R}_{x^i,(x+1)^i}. \end{aligned} \tag{B.4}$$

We can compute the following commutators: $\mathcal{A}^{(2)}_{x,x+1} := [\mathcal{A}_{x,x+1}, \mathcal{A}_{x+1,x+2,x+1}]$ and $[\mathcal{A}^{(2)}_{x,x+1}, \partial_{\theta(x+2)}]$ thanks to (B.3), and observe that $\mathcal{R}_{x^r,(x+2)^i} - \mathcal{R}_{x^i,(x+2)^r}$ and $\mathcal{R}_{x^r,(x+2)^r} + \mathcal{R}_{x^i,(x+2)^i}$ belong to our Lie algebra. Repeating this process, following an induction, we observe that for $x, y \in \widetilde{\mathbb{T}}_n$ the following terms are in the Lie algebra generated by $\{Y_x\}_{x=1}^n$, $\{[Y_x, Y_y]\}_{x,y=0}^n$, $\{[[Y_x, Y_y], Y_z]\}_{x,y,z=0}^n, \ldots$:

$$\mathcal{G}_n^o := \{\mathcal{R}_{x^r,x^i}, \mathcal{R}_{x^r,y^i} - \mathcal{R}_{x^i,y^r}, \mathcal{R}_{x^r,y^r} + \mathcal{R}_{x^i,y^i} \big| x, y \in \widetilde{\mathbb{T}}_n\}. \tag{B.5}$$

Notice that in the linear case (absence of non-linearity i.e., $p = 2$), the terms appeared in (B.5) represent a basis for the Lie algebra. (All the elements are linear combination of these terms).

In the following, we observe that $\mathcal{G}_n^o$ (B.5), has rank $2n-1$ for any $\psi \in S^n = \{\psi \in \mathbb{C}^n | \sum_{x=1}^n |\psi(x)|^2 = 1\}$, notice that we consider S^n as a $2n-1$ real sphere $\mathbb{S}^{2n-1}$ (the case where we replace S^n by S^n_m can be treated similarly). Let us proof by an induction. The case $n = 1$ is trivial, since $\partial_{\theta_1} = \psi_r(1)\partial_{\psi_i(1)} - \psi_i(1)\partial_{\psi_r(1)}$ has rank one for any $\psi(1) \in \mathbb{S}^1$ ($|\psi_r(1)|^2 + |\psi_i(1)|^2 = 1$).

Assume $\mathcal{G}_n^o$ has rank $2n-1$ at every point of $\mathbb{S}^{2n-1}$, we prove that $\mathcal{G}_{n+1}^o$ has rank $2n+1$ at every point of $\mathbb{S}^{2n+1}$. We split the proof into two cases:

Case 1.

Take $\psi \in S^{n+1}$, and assume that there exists at least one point x, such that $|\psi(x)| = 0$, we can take $x = n+1$, since we are in the periodic setup. We have $\psi_r(n+1) = \psi_i(n+1) = 0$; therefore, $\hat{\psi} = (\psi_1, \ldots, \psi_n) \in S^n$, and $\mathcal{G}_n^o$ has rank $2n-1$ by induction hypothesis. On the other hand, since $\hat{\psi} \in S^n$, there exists $y \in \widetilde{\mathbb{T}}_n$, such that $|\psi(y)| \neq 0$. First, observe that

$$\begin{aligned} B :=& \{\psi_r(y)\partial_{\psi_i(n+1)} - \psi_i(y)\partial_{\psi_r(n+1)}, \psi_r(y)\partial_{\psi_r(n+1)} + \psi_i(y)\partial_{\psi_i(n+1)}\} = \\ & \{\mathcal{R}_{y^r,(n+1)^i} - \mathcal{R}_{y^i,(n+1)^r}, \mathcal{R}_{y^r,(n+1)^r} - \mathcal{R}_{y^i,(n+1)^i}\} \subset \mathcal{G}_{n+1}^o, \end{aligned} \tag{B.6}$$

has rank two (this is straightforward, since $(\psi_r(y), \psi_i(y)) \neq 0$, and one can see a linear combination of elements of B is zero iff $|\psi(y)| = 0$). Then the result follows from the induction hypothesis, as well as the fact that B is orthogonal to $\mathcal{G}_n^o$.

Case2.

Take $\psi \in S^{n+1}$ and assume $|\psi(x)| \neq 0$ for all $x \in \widetilde{\mathbb{T}}_{n+1}$. In this case, we claim the set

$$\mathcal{G}^1_{n+1} := \{\mathcal{R}_{(n+1)^r,(n+1)^i}, \mathcal{R}_{(n+1)^r,x^i} - \mathcal{R}_{(n+1)^i,x^r}, \mathcal{R}_{(n+1)^r,x^r} + \mathcal{R}_{(n+1)^i,x^i} \big| x \in \widetilde{\mathbb{T}}_n\} \subset \mathcal{G}_n^o, \tag{B.7}$$

has rank $2n+1$. In fact, this set has $2n+1$ elements, where we observe that they are linearly independent. Take real [2] coefficients $\{a_x, b_x, c\}_{x=1}^n$ such that

$$c\mathcal{R}_{(n+1)^r,(n+1)^i} + \sum_{x=1}^n a_x\Big(\mathcal{R}_{(n+1)^r,x^i} - \mathcal{R}_{(n+1)^i,x^r}\Big) + b_x\Big(\mathcal{R}_{(n+1)^r,x^r} + \mathcal{R}_{(n+1)^i,x^i}\Big) = 0.$$

Computing the coefficients of $\partial_{\psi_r(x)}$ and $\partial_{\psi_i(x)}$, for any $x \in \widetilde{\mathbb{T}}_n$ we get:

$$\begin{aligned} \Big(b_x\psi_r(n+1) - a_x\psi_i(n+1)\Big)\partial_{\psi_r(x)} = 0, \\ \Big(a_x\psi_r(n+1) + b_x\psi_i(n+1)\Big)\partial_{\psi_i(x)} = 0. \end{aligned} \tag{B.8}$$

[2] Notice that we are considering the Field $\mathbb{R}$ here, by decomposing ψ into real and imaginary parts.

Notice that if $(a_x, b_x) \neq 0$, then $\det \begin{pmatrix} b_x & -a_x \\ a_x & b_x \end{pmatrix} > 0$. However, in order to (B.8) holds, the later cannot happen, since $(\psi_r(n+1), \psi_i(n+1)) \neq 0$; therefore $a_x = b_x = 0$ for all $x \in \widetilde{\mathbb{T}}_n$, and we can deduce $c = 0$, which yields the result in the case $d = 1$.
In order to prove the result for $d > 1$, for any $x, y \in \widetilde{\mathbb{T}}_n^d$, and any $\mu, \nu \in \{r, i\}$, we define $\mathcal{R}_{x^\mu, y^\nu}$, similar to (B.1). Recall $\{e_j\}_{j=1}^d$ as the canonical basis of $\mathbb{R}^d$, then (B.2) will be modified as:

$$\begin{aligned} &\partial_{\theta(x)} = \mathcal{R}_{x^r, x^i}, \\ &A_n = \sum_{x \in \widetilde{\mathbb{T}}_n} \sum_{j=1}^{d} \Big(\mathcal{R}_{(x+e_j)^r, x^i} + \mathcal{R}_{x^r, (x+e_j)^i} - 2\mathcal{R}_{x^r, x^i} \Big) + \sum_{x \in \widetilde{\mathbb{T}}_n} \kappa |\psi(x)|^{p-1} \mathcal{R}_{x^r, x^i}. \end{aligned} \tag{B.9}$$

The identity (B.3) remains true by taking $\alpha_1, \alpha_2, \alpha_3, \alpha_4 \in \{x^\mu | x \in \widetilde{\mathbb{T}}_n^d, \mu \in \{r, i\}\}$. This leads to the following modification of (B.4), for any $x \in \widetilde{\mathbb{T}}_n^d$, and $1 \leq k \leq d$:

$$\begin{aligned} &\mathcal{A}_x := [A_n, \partial_{\theta(x)}] = \sum_{j=1}^{d} \mathcal{R}_{x^r, (x+e_j)^r} + \mathcal{R}_{x^i, (x+e_j)^i} - \mathcal{R}_{(x-e_j)^r, x^r} - \mathcal{R}_{(x-e_j)^i, x^i}, \\ &\mathcal{A}_{x, x+e_k} := [[A_n, \partial_{\theta(x)}], \partial_{\theta(x+e_k)}] = \mathcal{R}_{x^r, (x+e_k)^i} - \mathcal{R}_{x^i, (x+e_k)^r}, \\ &\mathcal{A}_{x, x+e_k, x} := [[[A_n, \partial_{\theta(x)}], \partial_{\theta(x+e_k)}], \partial_{\theta(x)}] = \mathcal{R}_{x^r, (x+e_k)^r} + \mathcal{R}_{x^i, (x+e_k)^i}. \end{aligned} \tag{B.10}$$

Following the exact same strategy as in the previous case, by an induction we observe that all the terms of the form $\mathcal{R}_{x^r, (x+l_k e_k)^i} - \mathcal{R}_{x^i, (x+l_k e_k)^r}$ and $\mathcal{R}_{x^r, (x+l_k e_k)^r} + \mathcal{R}_{x^i, (x+l_k e_k)^i}$, for any $x \in \widetilde{\mathbb{T}}_n^d$, any $1 \leq k \leq d$, and any $l_k \in \widetilde{\mathbb{T}}_n$, belong to our Lie algebra. Notice that thanks to (B.3), we have:

$$\begin{aligned} &[\mathcal{R}_{x^r, (x+l_k e_k)^i} - \mathcal{R}_{x^i, (x+l_k e_k)^r}, \mathcal{R}_{(x+l_k e_k)^r, (x+l_k e_k+l_{k'} e_{k'})^i} - \mathcal{R}_{(x+l_k e_k)^i, (x+l_k e_k+l_{k'} e_{k'})}] = \\ &- \Big(\mathcal{R}_{x^r, (x+l_k e_k+l_{k'} e_{k'})^r} + \mathcal{R}_{x^i, (x+l_k e_k+l_{k'} e_{k'})^i} \Big) \\ &[\mathcal{R}_{x^r, (x+l_k e_k+l_{k'} e_{k'})^r} + \mathcal{R}_{x^i, (x+l_k e_k+l_{k'} e_{k'})^i}, \mathcal{R}_{x^r, x^i}] = \mathcal{R}_{x^i, (x+l_k e_k+l_{k'} e_{k'})^r} - \mathcal{R}_{x^i, (x+l_k e_k+l_{k'} e_{k'})^r}. \end{aligned} \tag{B.11}$$

Repeating the above procedure for $d-1$ times, we can deduce the following set is included in our Lie algebra:

$$\mathcal{G}_n^{o,d} := \{\mathcal{R}_{x^r, x^i}, \mathcal{R}_{x^r, y^i} - \mathcal{R}_{x^i, y^r}, \mathcal{R}_{x^r, y^r} + \mathcal{R}_{x^i, y^i} | x, y \in \widetilde{\mathbb{T}}_n^d\}. \tag{B.12}$$

Recall that we observed that the rank of $\mathcal{G}_n^o$ is $2n-1$. However, due to symmetry one can observe that $\mathcal{G}_{n^d}^o$ and $\mathcal{G}_n^{o,d}$ has the same rank and this finishes the proof. □

Remark B.1. *Notice that the proof of Lemma B.1 can be adapted to any other non-linearity of the form $f(|\psi(x)|)$, where f is smooth.*

Appendix C Discrete Gagliardo-Nirenberg Inequality

We present different versions of the Gagliardo-Nirenberg inequality. This inequality is crucial in the study of the sub-critical nonlinear focusing Schrödinger equation, for proving

the well-posedness and characterization of the Solitons (cf. [7],[32],[35],[15]). In particular, this inequality has been used in the the proof of Theorem 3.1 in [15]. We take advantage of the discrete version of this inequality, so we can establish properties of configurations with minimal or close to minimal energy.
Gagliardo-Nirenberg inequality states that for every $u \in H^1(\mathbb{R}^d)$, and $1 < p < 1 + \frac{4}{d}$, there exists a constant $C(p,d)$, such that (cf. [7],[32],[35]):

$$\|u\|_{L^{p+1}}^{p+1} = \int |u|^{p+1} \leq C(d,p)\Big(\int |\nabla u|^2\Big)^{\frac{d(p-1)}{4}} \Big(\int |u|^2\Big)^{\frac{p+1}{2} - \frac{d(p-1)}{4}}. \tag{C.1}$$

While we are focusing on the case where $d = 1$, and $p = 3 < 1 + \frac{4}{d} = 5$ and the domain is periodic, we state the following version from [[15] Section3.2, [25] Lemma 4.1]. For all $u \in H^1_{per} = H^1(\mathbb{T})$, there exists a constant $C > 0$:

$$\begin{aligned} \|u\|_{L^4}^4 &= \int_0^1 |u|^4 \leq C(\|\partial_x u\|_{L^2}\|u\|_{L^2}^3 + \|u\|_{L^2}^4) \\ &= C\left(\Big(\int_0^1 |u|^2\Big)^{\frac{3}{2}}\Big(\int_0^1 |\partial_x u|^2\Big)^{\frac{1}{2}} + \Big(\int_0^1 |u|^2\Big)^2\right). \end{aligned} \tag{C.2}$$

We need counterparts of these inequalities in the discrete setting, in order to obtain these inequalities, we generalize results from [[8] section 17]. First, we define: Fix $n > 0$, and consider a function $f : \widetilde{\mathbb{T}}_n \to \mathbb{C}$, define the discrete $\ell^p(\widetilde{\widetilde{\mathbb{T}}}_n)$ norm of f, for $p \geq 1$, as:

$$\|f\|_{\ell^p(\widetilde{\mathbb{T}}_n)} = \Big(\frac{1}{n}\sum_{x \in \widetilde{\mathbb{T}}_n} |f(x)|^p\Big)^{\frac{1}{p}}. \tag{C.3}$$

Notice that our definition differs from the conventional one by a factor $n^{-\frac{1}{p}}$. This difference is motivated by the fact that in the limit as $n \to \infty$, we can recover the continuous L^p norm, formally. Define the $H^1_{per}(\widetilde{\mathbb{T}}_n)$ norm of f as follows:

$$\|f\|_{H^1(\widetilde{\mathbb{T}}_n)} := \Big(\frac{1}{n}\sum_{x \in \widetilde{\mathbb{T}}_n} n^2|f(x) - f(x-1)|^2 + \frac{1}{n}\sum_{x \in \widetilde{\mathbb{T}}_n} |f(x)|^2\Big)^{\frac{1}{2}}. \tag{C.4}$$

We can also define the space $\ell^p(\mathbb{Z})$, with the following norm: For $f : \mathbb{Z} \to \mathbb{C}$ and $p \geq 1$ define:

$$\|f\|_{\ell^p(\mathbb{Z})} = \Big(\sum_{x \in \mathbb{Z}} |f(x)|^p\Big)^{\frac{1}{p}}. \tag{C.5}$$

As usual we have: $\ell^p(\mathbb{Z}) = \{f : \mathbb{Z} \to \mathbb{C} | \|f\|_{\ell^p(\mathbb{Z})} < \infty\}$. We denote the discrete gradient of $f : \mathbb{Z} \to \mathbb{C}$ by $G(f)$, and define it as:

$$G(f) := \frac{1}{2}\sum_{x \in \mathbb{Z}} |f(x) - f(x-1)|^2. \tag{C.6}$$

Note the difference between $G(f)$ and $G_n(f)$ in (3.9), where we scale the second definition by n^2 in order to get the continuous counterpart, formally.

The first version of the discrete Gagliardo-Nirenberg inequality can be recalled from Proposition 17.6 of [8] with a small modification: For every $1 < p \leq \infty$, let $\theta = \frac{1}{2} - \frac{1}{p+1}$, obviously $\theta \in (0,1)$, we have: $\forall f \in \ell^p(\mathbb{Z}) \cap \ell^2(\mathbb{Z})$, there exists a constant $C(p)$ such that:

$$\|f\|_{\ell^{p+1}(\mathbb{Z})} \leq C(p) \left(\|f\|_{\ell^2(\mathbb{Z})}\right)^{1-\theta} (G(f))^{\frac{\theta}{2}}. \tag{C.7}$$

In particular, for $p = 3$, we have $\forall f \in \ell^4(\mathbb{Z}) \cap \ell^2(\mathbb{Z})$, there exists a constant C such that:

$$\|f\|^4_{\ell^4(\mathbb{Z})} \leq C(\|f\|_{\ell^2(\mathbb{Z})})^3 (G(f))^{\frac{1}{2}}. \tag{C.8}$$

We can deduce the following lemma from the later, which is crucial for our purposes.

Lemma C.1. *Recall the definition of G_n (3.9), and $\|.\|_{\ell^p(\widetilde{\mathbb{T}}_n)}$ (C.3), for every $f : \widetilde{\mathbb{T}}_n \to \mathbb{C}$, there exist a constant C independent of n such that:*

$$\|f\|^4_{\ell^4(\widetilde{\mathbb{T}}_n)} \leq C\big((\|f\|_{\ell^2(\widetilde{\mathbb{T}}_n)})^3 (G_n(f))^{\frac{1}{2}} + \|f\|^2_{\ell^2(\widetilde{\mathbb{T}}_n)}\big), \tag{C.9}$$

we write this inequality in this open form:

$$\frac{1}{n}\sum_{x\in\widetilde{\mathbb{T}}_n} |f(x)|^4 \leq C\left(\Big(\frac{1}{n}\sum_{x\in\widetilde{\mathbb{T}}_n} \frac{n^2}{2}|f(x)-f(x-1)|^2\Big)^{\frac{1}{2}}\Big(\frac{1}{n}\sum_{x\in\widetilde{\mathbb{T}}_n}|f(x)|^2\Big)^{\frac{3}{2}} + \Big(\frac{1}{n}\sum_{x\in\widetilde{\mathbb{T}}_n}|f(x)|^2\Big)^2\right). \tag{C.10}$$

Usually we have $\frac{1}{n}\sum_{x\in\widetilde{\mathbb{T}}_n}|f(x)|^2 = m$, hence, we have:

$$\frac{1}{n}\sum_{x\in\widetilde{\mathbb{T}}_n} |f(x)|^4 \leq C\left(m^{\frac{3}{2}} G_n(f)^{\frac{1}{2}} + m^2\right). \tag{C.11}$$

Proof. We prove this lemma by constructing a function $\tilde{f} \in \ell^4(\mathbb{Z}) \cap \ell^2(\mathbb{Z})$, from f as follows: Translate f such that $|f(x)|^2$ achieves its minimum at $x = n$. By this construction, we have $|f(n)|^2 \leq \frac{1}{n}\sum_{x\in\widetilde{\mathbb{T}}_n}|f(x)|^2 = m$. Define $\tilde{f}$ on $\mathbb{Z}$ as:

$$\tilde{f}(x) = \begin{cases} f(x), \ \ \forall x \in \{1,\dots,n\}, \\ f(n)(2-\frac{x}{n}), \ \ \forall x \in \{n+1,\dots,2n\}, \\ f(n)(1+\frac{x}{n}), \ \ \forall x \in \{-n,\dots,-1\}, \\ f(n), \quad \text{if} \quad x = 0, \\ 0 \ \ \text{otherwise.} \end{cases} \tag{C.12}$$

By the definition of $\tilde{f}$, for every $p \geq 1$ we have:

$$\|\tilde{f}\|^p_{\ell^p(\mathbb{Z})} = \sum_{x\in\widetilde{\mathbb{T}}_n} |f(x)|^p + |f(n)|^p \sum_{x=0}^{n-1} 2(\frac{x}{n})^p + |f(n)|^p. \tag{C.13}$$

By estimating $\frac{1}{n}\sum_{x=1}^{n}(\frac{x}{n})^p$ with its integral value, we have $c_1(p) > 0$, $c_2(p) > 0$ independent of n, such that:

$$\|f\|^p_{\ell^p(\widetilde{\mathbb{T}}_n)} + c_1(p)|f(n)|^p \leq \frac{1}{n}\|\tilde{f}\|^p_{\ell^p(\mathbb{Z})} \leq \|f\|^p_{\ell^p(\widetilde{\mathbb{T}}_n)} + c_2(p)|f(n)|^p. \tag{C.14}$$

Moreover, we can compute $G(\tilde{f})$:

$$\begin{aligned} G(\tilde{f}) =& \sum_{x\in\mathbb{Z}} \frac{1}{2}|\tilde{f}(x) - \tilde{f}(x-1)|^2 = \frac{1}{2}\sum_{x=1}^{n}|f(x)-f(x-1)|^2 + \\ & |f(n)|^2\Big(\sum_{x=-n}^{0}\frac{1}{2n^2} + \sum_{x=n+1}^{2n}\frac{1}{2n^2}\Big) = \frac{1}{n}G_n(f) + \frac{1}{n}|f(n)|^2. \end{aligned} \tag{C.15}$$

Since we fix n, by (C.14) $\tilde{f} \in \ell^2(\mathbb{Z}) \cap \ell^4(\mathbb{Z})$; therefore, we can apply the inequality (C.8). By using the fact that $|f(n)|^2 \leq \frac{1}{n}\sum_{x\in\widetilde{\mathbb{T}}_n} |f(x)|^2 = \|f\|^2_{\ell^2(\widetilde{\mathbb{T}}_n)}$, and estimates (C.15) and (C.14), we get:

$$\begin{aligned}
\|f\|^4_{\ell^4(\widetilde{\mathbb{T}}_n)} &\leq \frac{1}{n}\|\tilde{f}\|_{\ell^4(\mathbb{Z})} \leq C\big(\frac{1}{n}\|\tilde{f}\|^2_{\ell^2(\mathbb{Z})}\big)^{\frac{3}{2}} \big(nG(\tilde{f})\big)^{\frac{1}{2}} \\
&\leq C\big(\|f\|^2_{\ell^2(\widetilde{\mathbb{T}}_n)} + c_2(2)|f(n)|^2\big)^{\frac{3}{2}}\big(G_n(f) + |f(n)|^2\big)^{\frac{1}{2}} \\
&\leq C'\big(\|f\|^2_{\ell^2(\widetilde{\mathbb{T}}_n)}\big)^{\frac{3}{2}}\big(G_n(f) + |f(n)|^2\big)^{\frac{1}{2}} \\
&\leq C'\big(\|f\|^2_{\ell^2(\widetilde{\mathbb{T}}_n)}\big)^{\frac{3}{2}} G_n(f)^{\frac{1}{2}} + C'\big(\|f\|^2_{\ell^2(\widetilde{\mathbb{T}}_n)}\big)^{\frac{3}{2}}\|f\|^4_{\ell^2(\widetilde{\mathbb{T}}_n)}.
\end{aligned} \tag{C.16}$$

This inequality proves the lemma, since $C' = C(1 + c_2(2))^{\frac{3}{2}}$ is a constant independent of n. In the last line, we used the inequality $\sqrt{a+b} \leq \sqrt{a} + \sqrt{b}$ for $a, b > 0$. □

Appendix D Jacobi Elliptic Functions

Given $k \in (0, 1)$, the incomplete elliptic integral of the first kind, for any $\phi \in \mathbb{R}$ is defined as:

$$x = F(\phi; k) := \int_0^\phi \frac{d\theta}{\sqrt{1 - k^2\sin^2(\theta)}}.$$

Consequently, one can define $\mathfrak{cn}(\cdot), \mathfrak{sn}(\cdot), \mathfrak{dn}(\cdot)$ via the inverse of $F(\cdot, k)$:

$$\mathfrak{sn}(x, k) := \sin(\phi), \quad \mathfrak{cn}(x, k) := \cos(\phi), \quad \mathfrak{dn}(x, k) := \sqrt{1 - k^2\sin^2(\phi)}. \tag{D.1}$$

From (D.1), it is straightforward to see for all x

$$\mathfrak{sn}^2(x, k) + \mathfrak{cn}^2(x, k) = k^2\mathfrak{sn}^2(x, k) + \mathfrak{dn}^2(x, k) = 1. \tag{D.2}$$

Moreover, the derivative (w.r.t x) of these functions can be obtained directly from the definition:

$$\begin{aligned}
\partial_x\mathfrak{sn}(x, k) = \mathfrak{cn}(x, k)\mathfrak{dn}(x, k), \quad &\partial_x\mathfrak{cn}(x, k) = -\mathfrak{sn}(x, k)\mathfrak{dn}(x, k), \\
&\partial_x\mathfrak{dn}(x, k) = -k^2\mathfrak{cn}(x, k)\mathfrak{sn}(x, k).
\end{aligned} \tag{D.3}$$

Moreover, the period of these functions is given via the following complete elliptic integral:

$$K(k) := F(\frac{\pi}{2}; k), \tag{D.4}$$

where $\mathfrak{dn}$ is $2K$ periodic and even, $\mathfrak{sn}$ and $\mathfrak{cn}$ are $4K$ periodic, where $\mathfrak{sn}$ is $2K$ anti periodic and odd, and $\mathfrak{cn}$ is $2K$ anti periodic and even.
Notice the limiting cases: $K(k) \to \frac{\pi}{2}$ as $k \to 0$, and $K(k) \to \infty$ as $k \to 1$. Moreover, as for $k = 0$, $\mathfrak{sn}(x, 0) = \sin(x)$, $\mathfrak{cn}(x, 0) = \cos(x)$, $\mathfrak{dn}(x, 0) = 1$. Furthermore, $\mathfrak{sn}(x, 1) = \tanh(x)$, $\mathfrak{cn}(x, 1) = \mathfrak{dn}(x, 1) = \mathrm{sech}(x)$.
Finally, notice that from (D.3) one can deduce that $\frac{1}{\alpha}\mathfrak{dn}(\frac{x}{\beta}, k)$, $\frac{1}{\alpha}\mathfrak{cn}(\frac{x}{\beta}, k)$, and $\frac{1}{\alpha}\mathfrak{sn}(\frac{x}{\beta}, k)$ are solutions to (3.5), where α, β, k are determined by ω, L in each case, respectively.

References

[1] S. Axler, P. Bourdon, and R. Wade. *Harmonic Function Theory*. Graduate Texts in Mathematics. Springer New York, 2013.

[2] Luc Rey Bellet. *Ergodic Properties of Markov Processes*, pages 1–39. Springer Berlin Heidelberg, Berlin, Heidelberg, 2006.

[3] J. Bourgain. Periodic nonlinear schrödinger equation and invariant measures. *Communications in Mathematical Physics*, 166(1):1–26, Dec 1994.

[4] J. Bourgain. *Global Solutions of Nonlinear Schrödinger Equations*. Number v. 46 in American Mathematical Society colloquium publications, Providence, RI. American Mathematical Society Providence RI, 1999.

[5] Eric A Carlen, Jürg Fröhlich, and Joel Lebowitz. Exponential relaxation to equilibrium for a one-dimensional focusing non-linear schrödinger equation with noise. *Communications in mathematical physics*, 342(1):303–332, 2016.

[6] Eric A Carlen, Jürg Fröhlich, Joel Lebowitz, and Wei-Min Wang. Quantitative bounds on the rate of approach to equilibrium for some one-dimensional stochastic nonlinear schrödinger equations. *Nonlinearity*, 32(4):1352–1374, mar 2019.

[7] T. Cazenave, American Mathematical Society, and Courant Institute of Mathematical Sciences. *Semilinear Schrodinger Equations*. Courant lecture notes in mathematics. American Mathematical Society, Providence, RI, 2003.

[8] Sourav Chatterjee. Invariant measures and the soliton resolution conjecture. *Communications on Pure and Applied Mathematics*, 67(11):1737–1842, 2014.

[9] Sourav Chatterjee and Kay Kirkpatrick. Probabilistic methods for discrete nonlinear schrödinger equations. *Communications on Pure and Applied Mathematics*, 65(5):727–757, 2012.

[10] A. de Bouard and A. Debussche. A stochastic nonlinear Schrödinger equation with multiplicative noise. *Communications in Mathematical Physics*, 205(1):161–181, Aug 1999.

[11] A. de Bouard and A. Debussche. The stochastic nonlinear Schrödinger equation in H^1. *Stochastic Analysis and Applications*, 21(1):97–126, 2003.

[12] S. Flach, K. Kladko, and R. S. MacKay. Energy thresholds for discrete breathers in one-, two-, and three-dimensional lattices. *Phys. Rev. Lett.*, 78:1207–1210, Feb 1997.

[13] Thierry Gallay and Mariana Hărăguş. Stability of small periodic waves for the nonlinear Schrödinger equation. *Journal of Differential Equations*, 234(2):544–581, 2007.

[14] Giacomo Gradenigo, Stefano Iubini, Roberto Livi, and Satya N Majumdar. Localization transition in the discrete nonlinear Schrödinger equation: ensembles inequivalence and negative temperatures. *Journal of Statistical Mechanics: Theory and Experiment*, 2021(2):023201, feb 2021.

[15] Stephen Gustafson, Stefan Le Coz, and Tai-Peng Tsai. Stability of Periodic Waves of 1D Cubic Nonlinear Schrödinger Equations. *Applied Mathematics Research eXpress*, 2017(2):431–487, 06 2017.

[16] Martin Hairer. A probabilistic argument for the controllability of conservative systems. *arXiv preprint math-ph/0506064*, 2005.

[17] Martin Hairer. On Malliavin's proof of Hörmander's theorem. *Bulletin des Sciences Mathématiques*, 135(6):650–666, 2011. Special issue in memory of Paul Malliavin.

[18] A. Hannani and S. Olla. In preparation.

[19] L. Hörmander. *The Analysis of Linear Partial Differential Operators III: Pseudo-Differential Operators*. Classics in Mathematics. Springer Berlin Heidelberg, 2007.

[20] Magnus Johansson and Kim Ø. Rasmussen. Statistical mechanics of general discrete nonlinear Schrödinger models: Localization transition and its relevance for Klein-Gordon lattices. *Phys. Rev. E*, 70:066610, Dec 2004.

[21] Panayotis G Kevrekidis. *The discrete nonlinear Schrödinger equation: mathematical analysis, numerical computations and physical perspectives*, volume 232. Springer Science & Business Media, 2009.

[22] K. Kirkpatrick. Solitons and gibbs measures for nonlinear schrödinger equations. *Mathematical Modelling of Natural Phenomena*, 7(2):95–112, 2 2012.

[23] Derek F Lawden. *Elliptic functions and applications*, volume 80. Springer Science & Business Media, 2013.

[24] J. L. Lebowitz, Ph. Mounaix, and W.-M. Wang. Approach to equilibrium for the stochastic nls. *Communications in Mathematical Physics*, 321(1):69–84, Jul 2013.

[25] Joel L. Lebowitz, Harvey A. Rose, and Eugene R. Speer. Statistical mechanics of the nonlinear schrödinger equation. *Journal of Statistical Physics*, 50(3):657–687, Feb 1988.

[26] Viviana Letizia. *Microscopic models for Fourier's law*. PhD thesis, Université Paris-Dauphine – PSL Research University, 2017.

[27] H. P. McKean. Statistical mechanics of nonlinear wave equations (4): Cubic Schrödinger. *Communications in Mathematical Physics*, 168(3):479–491, Apr 1995.

[28] H. P. McKean and K. L. Vaninsky. Brownian motion with restoring drift: The petit and micro-canonical ensembles. *Communications in Mathematical Physics*, 160(3):615–630, Mar 1994.

[29] H. P. McKean and K. L. Vaninsky. Action-angle variables for the cubic Schrödinger equation. *Communications on Pure and Applied Mathematics*, 50(6):489–562, 1997.

[30] H. P. McKean and K. L. Vaninsky. Cubic Schrödinger: The petit canonical ensemble in action-angle variables. *Communications on Pure and Applied Mathematics*, 50(7):593–622, 1997.

[31] J.A. Pava. *Nonlinear Dispersive Equations: Existence and Stability of Solitary and Periodic Travelling Wave Solutions*. Mathematical surveys and monographs. American Mathematical Society, Providence, RI, 2009.

[32] P. Raphaël. Stability and blow up for the non linear Schrödinger equation. *Lecture notes from the Clay Math 2008 Summer School on Evolution Equations.Clay Mathematics Proceedings, 17. American Mathematical Society, Providence, R.I.; Clay Mathematics Institute, Cambridge, Mass.*, pages 269–323, 2014.

[33] K. Ø. Rasmussen, T. Cretegny, P. G. Kevrekidis, and Niels Grønbech-Jensen. Statistical mechanics of a discrete nonlinear system. *Phys. Rev. Lett.*, 84:3740–3743, Apr 2000.

[34] Benno Rumpf. Simple statistical explanation for the localization of energy in nonlinear lattices with two conserved quantities. *Phys. Rev. E*, 69:016618, Jan 2004.

[35] T. Tao, Conference Board of the Mathematical Sciences, National Science Foundation (U.S.)., and American Mathematical Society. *Nonlinear Dispersive Equations: Local and Global Analysis*. Conference Board of the Mathematical Sciences. Regional conference series in mathematics. American Mathematical Society, 2006.

[36] M I Weinstein. Excitation thresholds for nonlinear localized modes on lattices. *Nonlinearity*, 12(3):673–691, jan 1999.