Approximation of small-amplitude weakly coupled oscillators by discrete nonlinear Schrödinger equations

Dmitry Pelinovsky[1,2], Tiziano Penati[3], and Simone Paleari[3]

[1] Department of Mathematics, McMaster University, Hamilton, Ontario, Canada, L8S 4K1

[2] Department of Applied Mathematics, Nizhny Novgorod State Technical University, 24 Minin Street, Nizhny Novgorod, 603950, Russia

[3] Department of Mathematics "F.Enriques", Milano University, via Saldini 50, Milano, Italy, 20133

Abstract

Small-amplitude weakly coupled oscillators of the Klein–Gordon lattices are approximated by equations of the discrete nonlinear Schrödinger type. We show how to justify this approximation by two methods, which have been very popular in the recent literature. The first method relies on a priori energy estimates and multi-scale decompositions. The second method is based on a resonant normal form theorem. We show that although the two methods are different in the implementation, they produce equivalent results as the end product. We also discuss the applications of the discrete nonlinear Schrödinger equation in the context of existence and stability of breathers of the Klein–Gordon lattice.

Keywords: Klein–Gordon lattice, discrete nonlinear Schrödinger equations, existence and stability of breathers, small-amplitude approximations, energy method, normal forms.

AMS subject classification: 37K40, 37K55, 37K60, 70K45.

1 Introduction

We consider the one-dimensional discrete Klein–Gordon (dKG) equation with the hard quartic potential in the form

$$\ddot{x}_j + x_j + x_j^3 = \epsilon(x_{j+1} - 2x_j + x_{j-1}), \quad j \in \mathbb{Z}, \tag{1.1}$$

where $t \in \mathbb{R}$ is the evolution time, $x_j(t) \in \mathbb{R}$ is the horizontal displacement of the j-th particle in the one-dimensional chain, and $\epsilon > 0$ is the coupling constant of the linear interaction between neighboring particles. The dKG equation (1.1) is associated with the conserved-in-time energy

$$H = \frac{1}{2} \sum_{j \in \mathbb{Z}} \dot{x}_j^2 + x_j^2 + \epsilon(x_{j+1} - x_j)^2 + \frac{1}{4} \sum_{j \in \mathbb{Z}} x_j^4, \tag{1.2}$$

which is also the Hamiltonian function of the dKG equation (1.1) written in the canonical variables $\{x_j, \dot{x}_j\}_{j \in \mathbb{Z}}$. The initial-value problem for the dKG equation (1.1) is globally well-posed in the sequence space $\ell^2(\mathbb{Z})$, thanks to the coercivity of the energy H in (1.2) in $\ell^2(\mathbb{Z})$.

1

By using a scaling transformation

$$\tilde{x}_j(\tilde{t}) = (1 + 2\epsilon)^{-1/2} x_j(t), \quad \tilde{t} = (1 + 2\epsilon)^{1/2} t, \quad \tilde{\epsilon} = (1 + 2\epsilon)^{-1} \epsilon, \tag{1.3}$$

and dropping the tilde notations, the dKG equation (1.1) can be rewritten without the diagonal terms in the discrete Laplacian operator,

$$\ddot{x}_j + x_j + x_j^3 = \epsilon(x_{j+1} + x_{j-1}), \quad j \in \mathbb{Z}. \tag{1.4}$$

Note that the values of ϵ in (1.4) are now restricted to the range $\left(0, \frac{1}{2}\right)$, because the map $\epsilon \to (1 + 2\epsilon)^{-1} \epsilon$ is a diffeomorphism from $(0, \infty)$ to $\left(0, \frac{1}{2}\right)$. This restriction does not represent a limitation if we study the solutions of the dKG equation for sufficiently small values of ϵ.

We consider the Cauchy problem for the dKG equation (1.4) and we aim at giving an approximation of its solutions by means of equations of the discrete nonlinear Schrödinger type, up to suitable time scales. This approach can be useful in general, but it may have additional interest when particular classes of solutions of the dKG equation (1.4) are taken into account. In the case of systems of weakly coupled oscillators, the relevant objects are given by time-periodic and spatially localized solutions called breathers.

Existence and stability of breathers have been studied in the dKG equation in many recent works. In particular, exploring the limit of weak coupling between the nonlinear oscillators, existence [27] and stability [2, 4] of the fundamental (single-site) breathers were established (see also the recent works in [30, 31]). More complicated multi-breathers were classified from the point of their spectral stability in the recent works [1, 25, 33]. Nonlinear stability and instability of multi-site breathers were recently studied in [11].

If the oscillators have small amplitudes in addition to being weakly coupled, the stability of multi-breathers in the dKG equation is related to the stability of multi-solitons in the discrete nonlinear Schrödinger (dNLS) equation:

$$2i\dot{a}_j + 3|a_j|^2 a_j = a_{j+1} + a_{j-1}, \quad j \in \mathbb{Z}, \tag{1.5}$$

where $a_j(\epsilon t) \in \mathbb{C}$ is the envelope amplitude for the linear harmonic e^{it} supported by the linear dKG equation (1.4) with $\epsilon = 0$. The relation between the dKG and the dNLS equations (1.4) and (1.5) was observed in [29] and [33] based on numerical simulations and perturbation results, respectively.

The present contribution addresses the justification of the dNLS equation (1.5), and its generalizations, for the weakly coupled small-amplitude oscillators of the dKG equation (1.4). In fact, we are going to explore two alternative but complementary points of view on the justification process, which enables us to establish rigorous bounds on the error terms, over the time scale during which the dynamics of the dNLS equation (1.5) is observed.

The first method in the justification of the dNLS equation (1.5) for small-amplitude weakly coupled oscillators of the dKG equation (1.4) is based on a priori energy estimates and elementary continuation arguments. This method was used in the derivation of the dNLS equation [8] and the Korteweg–de Vries equation [5, 12, 13, 38] in a similar context of the Fermi–Pasta–Ulam lattice. The energy method is based on the decomposition of the solution into the leading-order multi-scale approximation and the error term. The error term is controlled by integrating the

dKG equation with a small residual term over the relevant time scale. The energy method is computationally efficient and simple enough for most practical applications.

The second method is based on the resonant normal form theorem, which transforms the given Hamiltonian of the dKG equation to a simpler form by means of near-identity canonical transformations [3, 16]. The normal form, once it is obtained in the sense of an abstract theorem, does not require any additional work for the derivation and the justification of both the dNLS equation and its generalizations, which appear immediately in the corresponding relevant regimes. Starting from the works [17, 18], the normal form approach for the dKG equation was recently elaborated in [30] and applied in [31] for a stability result.

We hope that the present discussion of the two equivalent methods can motivate the readers for the choice of a suitable analytical technique in the justification analysis of similar problems of lattice dynamics. It is our understanding that the two methods are equivalent with respect to the results (error estimates, time scales) but they have some differences in the way one proves such results.

Besides justifying the dNLS equation (1.5) on the time scale $\mathcal{O}(\epsilon^{-1})$, we also extend the error bounds on the longer time intervals of $\mathcal{O}(|\log(\epsilon)|\epsilon^{-1})$. Similar improvements were reported in various other contexts of the justification analysis [8, 22, 24, 26]. Within the context of breathers, we show how the known results on the existence and stability of multi-solitons in the dNLS equation (1.5) can be used for similar results for the dKG equation (1.4).

We finish the introduction with a review of related results. Small-amplitude breathers of the dKG and dNLS equations were approximated with the continuous nonlinear Schrödinger equation in the works [6, 7, 36]. An alternative derivation of the continuous nonlinear Schrödinger equation was discussed in the context of the Fermi–Pasta–Ulam lattice [19, 20, 21, 37]. In the opposite direction, the derivation and the justification of the dNLS equation from a continuous nonlinear Schrödinger equation with a periodic potential were developed in the works [34, 35]. The justification of the popular variational approximation for multi-solitons of the dNLS equation in the limit of weak coupling between the nonlinear oscillators is reported in [9]. Bifurcations of periodic traveling waves from the linear limit of coupled nonlinear oscillators was developed with the use of symmetries of the dNLS equation in [14, 15].

The paper is organized as follows. Section 2 reports the justification results obtained from the energy method and multi-scale expansions. Section 3 reports the justification results obtained from the normal form theorem. Section 4 discusses applications of these results for the existence and stability of breathers in the dKG equation.

2 Justification of the dNLS equation with the energy method

In what follows, we consider the limit of weak coupling between the nonlinear oscillators, where ϵ is a small positive parameter. We also consider the small-amplitude oscillations starting with small-amplitude initial data. Hence, we use the scaling transformation $x_j = \rho^{1/2}\xi_j$, where ρ is another small positive parameter. Incorporating both small parameters, we rewrite the dKG equation (1.4) in the equivalent form

$$\ddot{\xi}_j + \xi_j + \rho\xi_j^3 = \epsilon(\xi_{j+1} + \xi_{j-1}), \quad j \in \mathbb{Z}. \tag{2.1}$$

The standard approximation of multi-breathers in the dKG equation (2.1) with multi-solitons of the dNLS equation (1.5) corresponds to the balance $\rho = \epsilon$. In Sections 2.1–2.3, we generalize the standard dNLS approximation by assuming that $\epsilon^2 \ll \rho \leq \epsilon$. In Section 2.4, we discuss further generalizations when ρ belongs to the asymptotic range $\epsilon^3 \ll \rho \leq \epsilon^2$.

2.1 Preliminary estimates

To recall the standard dNLS approximation, we define the slowly varying approximate solution of the dKG equation (2.1) in the form

$$X_j(t) = a_j(\epsilon t)e^{it} + \bar{a}_j(\epsilon t)e^{-it}. \tag{2.2}$$

Substituting the leading-order solution (2.2) to the dKG equation (2.1) and removing the resonant terms $e^{\pm it}$ at the leading order of $\mathcal{O}(\epsilon)$, we obtain the dNLS equation in the form

$$2i\dot{a}_j + 3\nu|a_j|^2 a_j = a_{j+1} + a_{j-1}, \quad j \in \mathbb{Z}, \tag{2.3}$$

where the dot denotes the derivative with respect to the slow time $\tau = \epsilon t$ and the parameter $\nu = \rho/\epsilon$ is defined in the asymptotic range $\epsilon \ll \nu \leq 1$.

With the account of the dNLS equation (2.3), the leading-order solution (2.2) substituted into the dKG equation (2.1) produces the residual terms in the form

$$\mathrm{Res}_j(t) := \rho \left(a_j^3 e^{3it} + \bar{a}_j^3 e^{-3it}\right) + \epsilon^2 \left(\ddot{a}_j e^{it} + \ddot{\bar{a}}_j e^{-it}\right). \tag{2.4}$$

The second residual term is resonant but occurs in the higher order $\mathcal{O}(\epsilon^2)$, which is not an obstacle in the justification analysis. The first residual term is non-resonant but it occurs at the leading order of $\mathcal{O}(\rho) \gg \mathcal{O}(\epsilon^2)$. Therefore, the first term needs to be removed, which is achieved with the standard near-identity transformation. Namely, we extend the leading-order approximation (2.2) to the form

$$X_j(t) = a_j(\epsilon t)e^{it} + \bar{a}_j(\epsilon t)e^{-it} + \frac{1}{8}\rho \left(a_j^3(\epsilon t)e^{3it} + \bar{a}_j^3(\epsilon t)e^{-3it}\right). \tag{2.5}$$

For simplicity, we do not mention that X_j depends on ϵ and ρ. Substituting the approximation (2.5) into the dKG equation (2.1), we obtain the new residual terms in the form

$$
\begin{aligned}
\mathrm{Res}_j(t) \quad := \quad & \epsilon^2 \left(\ddot{a}_j e^{it} + \ddot{\bar{a}}_j e^{-it}\right) - \frac{1}{8}\epsilon\rho \left((a_{j+1}^3 + a_{j-1}^3)e^{3it} + (\bar{a}_{j+1}^3 + \bar{a}_{j-1}^3)e^{-3it}\right) \\
& + \frac{3}{8}\rho^2 \left(a_j e^{it} + \bar{a}_j e^{-it}\right)^2 \left(a_j^3 e^{3it} + \bar{a}_j^3 e^{-3it}\right) + \frac{9}{4}\epsilon\rho \left(ia_j^2 \dot{a}_j e^{3it} - i\bar{a}_j^2 \dot{\bar{a}}_j e^{-3it}\right) \\
& + \frac{3}{64}\rho^3 \left(a_j e^{it} + \bar{a}_j e^{-it}\right) \left(a_j^3 e^{3it} + \bar{a}_j^3 e^{-3it}\right)^2 + \frac{1}{8}\epsilon^2\rho \left(\ddot{a}_j^3 e^{3it} + \ddot{\bar{a}}_j^3 e^{-3it}\right) \\
& + \frac{1}{512}\rho^4 \left(a_j^3 e^{3it} + \bar{a}_j^3 e^{-3it}\right)^3.
\end{aligned}
\tag{2.6}
$$

Note that all the time derivatives of a_j in the residual term (2.6) can be eliminated from the dNLS equation (2.3) provided that $\{a_j\}_{j \in \mathbb{Z}}$ is a twice differentiable sequence with respect to time. For all purposes we need, it is sufficient to consider the sequence space $\ell^2(\mathbb{Z})$. Hence we denote the sequence $\{a_j\}_{j \in \mathbb{Z}}$ in $l^2(\mathbb{Z})$ by $\mathbf{a}$.

The next results give preliminary estimates on global solutions of the dNLS equation (2.3), the leading-order approximation (2.5), and the residual term (2.6).

4

Lemma 1 *For every $\mathbf{a}_0 \in \ell^2(\mathbb{Z})$ and every $\nu \in \mathbb{R}$, there exists a unique global solution $\mathbf{a}(t)$ of the dNLS equation (2.3) in $\ell^2(\mathbb{Z})$ for every $t \in \mathbb{R}$ such that $\mathbf{a}(0) = \mathbf{a}_0$. Moreover, the solution $\mathbf{a}(t)$ is smooth in t and $\|\mathbf{a}(t)\|_{\ell^2} = \|\mathbf{a}_0\|_{\ell^2}$.*

Proof. Local well-posedness and smoothness of the local solution $\mathbf{a}$ with respect to the time variable t follow from the contraction principle applied to an integral version of the dNLS equation (2.3). The contraction principle can be applied because the discrete Laplacian operator is a bounded operator on $\ell^2(\mathbb{Z})$, whereas $\ell^2(\mathbb{Z})$ is a Banach algebra with respect to pointwise multiplication and the $\ell^2(\mathbb{Z})$ norm is an upper bound for the $\ell^\infty(\mathbb{Z})$ norm of a sequence. Global continuation of the local solution $\mathbf{a}$ follows from the $\ell^2(\mathbb{Z})$ conservation of the dNLS equation (2.3). ∎

Lemma 2 *For every $\mathbf{a}_0 \in \ell^2(\mathbb{Z})$, there exists a positive constant $C_X(\|\mathbf{a}_0\|_{\ell^2})$ (that depends on $\|\mathbf{a}_0\|_{\ell^2}$) such that for every $\rho \in (0, 1]$ and every $t \in \mathbb{R}$, the leading-order approximation (2.5) is estimated by*

$$\|\mathbf{X}(t)\|_{\ell^2} + \|\dot{\mathbf{X}}(t)\|_{\ell^2} \leq C_X(\|\mathbf{a}_0\|_{\ell^2}). \tag{2.7}$$

Proof. The result follows from the Banach algebra property of $\ell^2(\mathbb{Z})$ and the global existence result of Lemma 1. ∎

Lemma 3 *Assume that $\rho \leq \epsilon$. For every $\mathbf{a}_0 \in \ell^2(\mathbb{Z})$, there exists a positive ϵ-independent constant $C_R(\|\mathbf{a}_0\|_{\ell^2})$ (that depends on $\|\mathbf{a}_0\|_{\ell^2}$) such that for every $\epsilon \in (0, 1]$ and every $t \in \mathbb{R}$, the residual term in (2.6) is estimated by*

$$\|\mathbf{Res}(t)\|_{\ell^2} \leq C_R(\|\mathbf{a}_0\|_{\ell^2})\epsilon^2. \tag{2.8}$$

Proof. The result follows from the Banach algebra property of $\ell^2(\mathbb{Z})$, as well as from the global existence and smoothness of the solution $\mathbf{a}(t)$ of the dNLS equation (2.3) in Lemma 1. ∎

2.2 Justification of the dNLS equation on the dNLS time scale

The main result of this section is the following justification theorem.

Theorem 1 *Assume that ρ is defined in the asymptotic range $\epsilon^2 \ll \rho \leq \epsilon$. For every $\tau_0 > 0$, there is a small $\epsilon_0 > 0$ and positive constants C_0 and C such that for every $\epsilon \in (0, \epsilon_0)$, for which the initial data satisfies*

$$\|\xi(0) - \mathbf{X}(0)\|_{l^2} + \|\dot{\xi}(0) - \dot{\mathbf{X}}(0)\|_{l^2} \leq C_0 \rho^{-1} \epsilon^2, \tag{2.9}$$

the solution of the dKG equation (2.1) satisfies for every $t \in [-\tau_0 \rho^{-1}, \tau_0 \rho^{-1}]$,

$$\|\xi(t) - \mathbf{X}(t)\|_{l^2} + \|\dot{\xi}(t) - \dot{\mathbf{X}}(t)\|_{l^2} \leq C \rho^{-1} \epsilon^2. \tag{2.10}$$

Remark 1 *If $\rho = \epsilon$, the justification result of Theorem 1 guarantees that the dynamics of small-amplitude oscillators follows closely the dynamics of the dNLS equation (1.5) on the dNLS time scale $[-\tau_0, \tau_0]$ for the variable $\tau = \epsilon t$.*

Remark 2 *If $\rho = \epsilon^{8/5}$, the error term in (2.10) satisfies the $\mathcal{O}_{\ell^2}(\epsilon^{2/5})$ bound. The error term is controlled on the longer time scale $[-\tau_0 \epsilon^{-3/5}, \tau_0 \epsilon^{-3/5}]$ for the variable $\tau = \epsilon t$ of the dNLS equation (2.3) with $\nu = \epsilon^{3/5}$.*

To develop the justification analysis, we write $\xi(t) = \mathbf{X}(t) + \mathbf{y}(t)$, where $\mathbf{X}(t)$ is the leading-order approximation (2.5) and $\mathbf{y}(t)$ is the error term. Substituting the decomposition into the lattice equation (2.1), we obtain the evolution problem for the error term:

$$\ddot{y}_j + y_j + 3\rho X_j^2 y_j + 3\rho X_j y_j^2 + \rho y_j^3 - \epsilon(y_{j+1} + y_{j-1}) + \mathrm{Res}_j = 0, \quad j \in \mathbb{Z}, \qquad (2.11)$$

where the residual term $\mathbf{Res}(t)$ is given by (2.6) if $\mathbf{a}(t)$ satisfies the dNLS equation (2.3). Associated with the evolution equation (2.11), we also define the energy of the error term

$$E(t) := \frac{1}{2} \sum_{j \in \mathbb{Z}} \left[\dot{y}_j^2 + y_j^2 + 3\rho X_j^2 y_j^2 - 2\epsilon y_j y_{j+1} \right]. \qquad (2.12)$$

For every $\epsilon \in \left(0, \frac{1}{4}\right)$, the energy $E(t)$ is coercive and controls the $\ell^2(\mathbb{Z})$ norm of the solution in the sense

$$\|\dot{\mathbf{y}}(t)\|_{\ell^2}^2 + \|\mathbf{y}(t)\|_{\ell^2}^2 \leq 4E(t), \qquad (2.13)$$

for every t, for which the solution $\mathbf{y}(t)$ is defined. The rate of change for the energy (2.12) is found from the evolution problem (2.11):

$$\frac{dE}{dt} = \sum_{j \in \mathbb{Z}} \left[-\dot{y}_j \mathrm{Res}_j + 3\rho X_j \dot{X}_j y_j^2 - 3\rho X_j y_j^2 \dot{y}_j - \rho y_j^3 \dot{y}_j \right]. \qquad (2.14)$$

Thanks to the coercivity (2.13), the Cauchy–Schwarz inequality, and the continuous embedding of $\ell^2(\mathbb{Z})$ to $\ell^\infty(\mathbb{Z})$, we obtain

$$\left| \frac{dE}{dt} \right| \leq 2E^{1/2} \left[\|\mathbf{Res}(t)\|_{\ell^2} + 6\rho E^{1/2} \|\mathbf{X}(t)\|_{\ell^2} \|\dot{\mathbf{X}}(t)\|_{\ell^2} + 12\rho E \|\mathbf{X}(t)\|_{\ell^2} + 8\rho E^{3/2} \right]. \qquad (2.15)$$

To simplify the analysis, it is better to introduce the parametrization $E = Q^2$ and rewrite (2.15) in the equivalent form

$$\left| \frac{dQ}{dt} \right| \leq \|\mathbf{Res}(t)\|_{\ell^2} + 6\rho Q \|\mathbf{X}(t)\|_{\ell^2} \|\dot{\mathbf{X}}(t)\|_{\ell^2} + 12\rho Q^2 \|\mathbf{X}(t)\|_{\ell^2} + 8\rho Q^3. \qquad (2.16)$$

The energy estimate (2.16) is the starting point for the proof of Theorem 1.

Proof of Theorem 1. Let $\tau_0 > 0$ be fixed arbitrarily but independently of ϵ and assume that the initial norm of the perturbation term satisfies the following bound

$$Q(0) \leq C_0 \rho^{-1} \epsilon^2, \qquad (2.17)$$

6

where C_0 is a positive ϵ-independent constant and $\epsilon \in \left(0, \frac{1}{4}\right)$ is sufficiently small. Note that the bound (2.17) follows from the assumption (2.9) and the energy (2.12) subject to the choice of the constant C_0.

To justify the dNLS equation (2.3) on the time scale $[-\tau_0\rho^{-1}, \tau_0\rho^{-1}]$ for t, we define

$$T_0 := \sup\left\{ t_0 \in [0, \tau_0\rho^{-1}] : \sup_{t \in [-t_0, t_0]} Q(t) \leq C_Q \rho^{-1} \epsilon^2 \right\}, \tag{2.18}$$

where $C_Q > C_0$ is a positive ϵ-independent constant to be determined below. By the continuity of the solution in the $\ell^2(\mathbb{Z})$ norm, it is clear that $T_0 > 0$.

By using Lemmas 2 and 3, as well as the definition (2.18), we write the energy estimate (2.16) for every $t \in [-T_0, T_0]$ in the form

$$\left| \frac{dQ}{dt} \right| \leq C_R \epsilon^2 + \rho \left(6C_X^2 + 12C_X C_Q \rho^{-1} \epsilon^2 + 8C_Q^2 \rho^{-2} \epsilon^4 \right) Q. \tag{2.19}$$

If $\epsilon > 0$ is sufficiently small and $\epsilon^2 \ll \rho$, for every $t \in [-T_0, T_0]$, one can always find a positive ϵ-independent k_0 such that

$$6C_X^2 + 12C_X C_Q \rho^{-1} \epsilon^2 + 8C_Q^2 \rho^{-2} \epsilon^4 \leq k_0. \tag{2.20}$$

Integrating (2.19), we obtain

$$Q(t) e^{-\rho k_0 |t|} - Q(0) \leq \int_0^{|t|} C_R \epsilon^2 e^{-\rho k_0 t'} dt' \leq \frac{C_R \epsilon^2}{\rho k_0}. \tag{2.21}$$

By using (2.17), we obtain for every $t \in [-T_0, T_0]$:

$$Q(t) \leq \rho^{-1} \epsilon^2 \left(C_0 + k_0^{-1} C_R \right) e^{k_0 \tau_0}. \tag{2.22}$$

Hence, we can define $C_Q := \left(C_0 + k_0^{-1} C_R \right) e^{k_0 \tau_0}$ and extend the time interval in (2.18) by elementary continuation arguments to the full time span with $T_0 = \tau_0 \rho^{-1}$. This completes the justification of the dNLS equation (2.3) in Theorem 1. $\qquad\square$

2.3 Justification of the dNLS equation on the extended time scale

Next, we justify the dNLS equation (2.3) on the extended time scale

$$[-A|\log(\rho)|\rho^{-1}, A|\log(\rho)|\rho^{-1}], \tag{2.23}$$

for the variable t, where the positive constant A is fixed independently of ϵ. The main result of this section is the following justification theorem.

Theorem 2 *Assume that there is $\alpha \in (0, 1)$ such that ρ is defined in the asymptotic range*

$$\epsilon^{\frac{2}{1+\alpha}} \ll \rho \leq \epsilon.$$

For every $A \in \left(0, k_0^{-1}\alpha\right)$, where k_0 is defined in (2.28) below, there is a small $\epsilon_0 > 0$ and positive constants C_0 and C such that for every $\epsilon \in (0, \epsilon_0)$, for which the initial data satisfies

$$\|\xi(0) - \mathbf{X}(0)\|_{l^2} + \|\dot{\xi}(0) - \dot{\mathbf{X}}(0)\|_{l^2} \leq C_0 \rho^{-1} \epsilon^2, \tag{2.24}$$

the solution of the dKG equation (2.1) satisfies for every t in the time span (2.23),

$$\|\xi(t) - \mathbf{X}(t)\|_{l^2} + \|\dot{\xi}(t) - \dot{\mathbf{X}}(t)\|_{l^2} \leq C \rho^{-1-\alpha} \epsilon^2. \tag{2.25}$$

Remark 3 If $\rho = \epsilon$, the extended time scale (2.23) corresponds to the interval $[-A|\log(\epsilon)|, A|\log(\epsilon)|]$ for the variable $\tau = \epsilon t$ in the dNLS equation (2.3), hence it extends to all times τ as $\epsilon \to 0$.

Remark 4 If $\rho = \epsilon^{8/5}$, then the error term in (2.25) satisfies the $\mathcal{O}_{\ell^2}(\epsilon^{2(1-4\alpha)/5})$ bound, which is small if $\alpha \in \left(0, \frac{1}{4}\right)$. The error term is controlled on the longer time scale

$$\left[-\tau_0|\log(\epsilon)|\epsilon^{-3/5}, \tau_0|\log(\epsilon)|\epsilon^{-3/5}\right]$$

for the variable $\tau = \epsilon t$ of the dNLS equation (2.3) with $\nu = \epsilon^{3/5}$.

Proof of Theorem 2. We use the same assumption (2.17) on the initial norm of the perturbation term. To justify the dNLS equation (2.3) on the time scale (2.23) for t, we define

$$T_0^* := \sup \left\{ t_0 \in \left[0, A|\log(\rho)|\rho^{-1}\right] : \sup_{t \in [-t_0, t_0]} Q(t) \leq C_Q \rho^{-1-\alpha} \epsilon^2 \right\}, \tag{2.26}$$

where C_Q is a positive ϵ-independent constant to be determined below.

By using Lemmas 2 and 3, as well as the definition (2.26), we write the energy estimate (2.16) for every $t \in [-T_0^*, T_0^*]$ in the form

$$\left| \frac{dQ}{dt} \right| \leq C_R \epsilon^2 + \rho \left(6C_X^2 + 12 C_X C_Q \rho^{-1-\alpha} \epsilon^2 + 8 C_Q^2 \rho^{-2(1+\alpha)} \epsilon^4 \right) Q. \tag{2.27}$$

If $\epsilon > 0$ is sufficiently small and $\epsilon^2 \ll \rho^{1+\alpha}$, then for every $t \in [-T_0^*, T_0^*]$, one can always find a positive ϵ-independent k_0 such that

$$6C_X^2 + 12 C_X C_Q \rho^{-1-\alpha} \epsilon^2 + 8 C_Q^2 \rho^{-2(1+\alpha)} \epsilon^4 \leq k_0. \tag{2.28}$$

By integrating the energy estimate (2.27) in the same way as is done in (2.21), we obtain for every $t \in [-T_0^*, T_0^*]$:

$$\begin{aligned} Q(t) &\leq \rho^{-1} \epsilon^2 \left(C_0 + k_0^{-1} C_R \right) e^{k_0 A |\log(\rho)|} \\ &\leq \rho^{-1-\alpha} \epsilon^2 \left(C_0 + k_0^{-1} C_R \right), \end{aligned} \tag{2.29}$$

where the last bound holds because $k_0 A \in (0, \alpha)$. Hence, we can define $C_Q := C_0 + k_0^{-1} C_R$ and extend the time interval in (2.26) by elementary continuation arguments to the full time span with $T_0^* = A|\log(\rho)|\rho^{-1}$. This completes the justification of the dNLS equation (2.3) on the time scale (2.23). $\qquad \square$

2.4 Approximations with the generalized dNLS equation

Extensions of the justification analysis are definitely possible by including more ϵ-dependent terms into the dNLS equation (2.3) and the leading-order approximation (2.5), which makes the residual term (2.6) to be as small as $\mathcal{O}(\epsilon^n)$ for any $n \geq 2$. These extensions are not so important if $\epsilon^2 \ll \rho \leq \epsilon$ but they become crucial to capture the correct balance between linear and nonlinear effects on the dynamics of small-amplitude oscillators if $\rho \leq \epsilon^2$.

To illustrate these extensions, we show how to modify the justification analysis in the asymptotic range $\epsilon^3 \ll \rho \leq \epsilon^2$. We use the same leading-order approximation (2.5) in the form

$$X_j(t) = a_j(\epsilon t)e^{it} + \bar{a}_j(\epsilon t)e^{-it} + \frac{1}{8}\rho\left(a_j^3(\epsilon t)e^{3it} + \bar{a}_j^3(\epsilon t)e^{-3it}\right), \tag{2.30}$$

but assume that $\mathbf{a}(\tau)$ with $\tau = \epsilon t$ satisfy the generalized dNLS equation

$$2i\dot{a}_j + 3\epsilon\delta|a_j|^2 a_j = a_{j+1} + a_{j-1} + \frac{\epsilon}{4}\left(a_{j+2} + 2a_j + a_{j-2}\right), \quad j \in \mathbb{Z}. \tag{2.31}$$

Here we have introduced the parameter $\delta = \rho/\epsilon^2$ in the asymptotic range $\epsilon \ll \delta \leq 1$. Substituting (2.30) and (2.31) into the dKG equation (2.1), we obtain the modifications of the residual terms (2.6) in the form

$$\begin{aligned}
\mathrm{Res}_j(t) \;:=\; & \frac{1}{4}\epsilon^2\left(4\ddot{a}_j + a_{j+2} + 2a_j + a_{j-2}\right)e^{it} + \frac{1}{4}\epsilon^2\left(4\ddot{\bar{a}}_j + \bar{a}_{j+2} + 2\bar{a}_j + \bar{a}_{j-2}\right)e^{-it} \\
& -\frac{1}{8}\epsilon\rho\left((a_{j+1}^3 + a_{j-1}^3)e^{3it} + (\bar{a}_{j+1}^3 + \bar{a}_{j-1}^3)e^{-3it}\right) + \frac{1}{8}\epsilon^2\rho\left(\dddot{a}_j^3 e^{3it} + \dddot{\bar{a}}_j^3 e^{-3it}\right) \\
& +\frac{3}{8}\rho^2\left(a_j e^{it} + \bar{a}_j e^{-it}\right)^2\left(a_j^3 e^{3it} + \bar{a}_j^3 e^{-3it}\right) + \frac{9}{4}\epsilon\rho\left(ia_j^2\dot{a}_j e^{3it} - i\bar{a}_j^2\dot{\bar{a}}_j e^{-3it}\right) \\
& +\frac{3}{64}\rho^3\left(a_j e^{it} + \bar{a}_j e^{-it}\right)\left(a_j^3 e^{3it} + \bar{a}_j^3 e^{-3it}\right)^2 + \frac{1}{512}\rho^4\left(a_j^3 e^{3it} + \bar{a}_j^3 e^{-3it}\right)^3. \tag{2.32}
\end{aligned}$$

By using the extended dNLS equation (2.31), we realize that the residual terms of the $\mathcal{O}_{\ell^2}(\epsilon^2)$ order are canceled and the residual term in (2.32) enjoys the improved estimate

$$\|\mathbf{Res}(t)\|_{\ell^2} \leq C_R(\|\mathbf{a}_0\|_{\ell^2})\epsilon^3, \tag{2.33}$$

compared with the previous estimate (2.8). As a result, the justification analysis developed in the proof of Theorems 1 and 2 holds verbatim and results in the following theorems.

Theorem 3 *Assume that ρ is defined in the asymptotic range $\epsilon^3 \ll \rho \leq \epsilon^2$. For every $\tau_0 > 0$, there is a small $\epsilon_0 > 0$ and positive constants C_0 and C such that for every $\epsilon \in (0, \epsilon_0)$, for which the initial data satisfies*

$$\|\xi(0) - \mathbf{X}(0)\|_{l^2} + \|\dot{\xi}(0) - \dot{\mathbf{X}}(0)\|_{l^2} \leq C_0\rho^{-1}\epsilon^3, \tag{2.34}$$

the solution of the dKG equation (2.1) satisfies for every $t \in [-\tau_0\rho^{-1}, \tau_0\rho^{-1}]$,

$$\|\xi(t) - \mathbf{X}(t)\|_{l^2} + \|\dot{\xi}(t) - \dot{\mathbf{X}}(t)\|_{l^2} \leq C\rho^{-1}\epsilon^3. \tag{2.35}$$

Theorem 4 *Assume that there is $\alpha \in \left(0, \frac{1}{2}\right)$ such that ρ is defined in the asymptotic range*

$$\epsilon^{\frac{3}{1+\alpha}} \ll \rho \leq \epsilon^2.$$

There is $A_0 > 0$ such that for every $A \in (0, A_0)$, there is a small $\epsilon_0 > 0$ and positive constants C_0 and C such that for every $\epsilon \in (0, \epsilon_0)$, for which the initial data satisfies

$$\|\xi(0) - \mathbf{X}(0)\|_{l^2} + \|\dot{\xi}(0) - \dot{\mathbf{X}}(0)\|_{l^2} \leq C_0 \rho^{-1} \epsilon^3, \tag{2.36}$$

the solution of the dKG equation (2.1) satisfies for every t in the time span (2.23),

$$\|\xi(t) - \mathbf{X}(t)\|_{l^2} + \|\dot{\xi}(t) - \dot{\mathbf{X}}(t)\|_{l^2} \leq C \rho^{-1-\alpha} \epsilon^3. \tag{2.37}$$

We note that $\mathbf{X}$ in Theorems 3 and 4 is defined by the leading-order approximation (2.30), whereas $\mathbf{a}$ satisfies the generalized dNLS equation (2.31). The time scales in Theorems 3 and 4 are appropriate for the generalized dNLS equation (2.31) because $\delta \leq 1$ and $\epsilon \rho^{-1} \geq \epsilon^{-1}$.

3   Justification of the dNLS equation with the normal form method

The purpose of this section is to show that the results of Theorems 1, 2, 3, and 4 can be obtained with a different method relying on the resonant normal form theorem, mainly working at the level of the Hamiltonians. This slightly different point of view, as we stressed in the introduction, moves the main difficulties in the early steps of this approach, in terms of definitions and theorems to get the normal form established. But after this effort, it is straightforward to get the desired results of justification of the dNLS equation in many different regimes.

Another difference between this section and the previous one is in the dimension of the chain, infinite for the energy method, and finite for the normal form method, but with estimates uniform in the size of the chain. The main reason for this asymmetry in the presentation is that the normal form theorems from the previous works [17, 18, 30] were developed for finite chains, and their extension to the infinite case is beyond the scope of the present paper.

In what follows, we consider the dKG equation (1.4) on a finite chain of $2N + 1$ oscillators under periodic boundary conditions, where N is arbitrary large but finite. The finite dKG chain is associated with the Hamiltonian $H = H_0 + H_1$, where

$$H_0 := \frac{1}{2} \sum_{j=-N}^{N} \left[y_j^2 + x_j^2 - 2\epsilon x_{j+1} x_j \right], \qquad H_1 := \frac{1}{4} \sum_{j=-N}^{N} x_j^4, \tag{3.1}$$

subject to the periodic boundary conditions $x_{-N} = x_{N+1}$ and $y_{-N} = y_{N+1}$. It is quite clear from the expression above that H is an extensive quantity, i.e. roughly speaking proportional to N, and more precisely, it is translation invariant and with a short interaction range (see Section 2 in [30] for details). By preserving the extensivity, via a suitable normal form construction, we are able to get uniform estimates with respect to N.

To be more definite, it was proven in [30] that for any small coupling ϵ, there exists a canonical transformation $T_\mathcal{X}$ which puts the Hamiltonian $H = H_0 + H_1$, with H_0 and H_1 in (3.1), into an extensive resonant normal form of order r

$$H^{(r)} = H_\Omega + \mathcal{Z} + P^{(r+1)} , \qquad \{H_\Omega, \mathcal{Z}\} = 0 , \qquad (3.2)$$

where H_Ω is the Hamiltonian for the system of $2N+1$ identical oscillators of frequency Ω (which is the average of the linear frequencies [18]), $\mathcal{Z}$ is a non-homogeneous polynomial of order $2r+2$, $P^{(r+1)}$ is a remainder of order $2r+4$ and higher, and r grows as an inverse power of ϵ. Such a normal form was shown to be well defined in a small ball $B_{\rho^{1/2}}(0) \subset \mathcal{P}$ of the phase space $\mathcal{P}$, endowed with the Euclidean norm (which becomes the $\ell^2(\mathbb{Z})$ norm in the limit $N \to \infty$), provided $r\rho^{1/2} \ll 1$. The linear part of the Hamiltonian $H_\Omega = \Omega\rho$ is equivalent to the selected squared norm (uniformly with N), thus the almost invariance of H_Ω over times $|t| \sim (r^2\rho)^{-r-1}$ is easily derived since $\dot{H}_\Omega = \{H_\Omega, P^{(r+1)}\}$.

Looking at the structure of $\mathcal{Z}$, the normal form $H_\Omega + \mathcal{Z}$ produces a generalized dNLS equation, where all the oscillators are coupled to all neighbors and the coupling coefficients both for linear and nonlinear terms decay exponentially with the distance between sites. To be more specific, $\mathcal{Z}$ can be split as the sum of homogeneous polynomials $Z_0, Z_1, ..., Z_r$, where Z_j is of the order $2j+2$, and $r \geq 1$. Each of these homogenous polynomials can be developed in powers of the coupling coefficient ϵ, where the term of order ϵ^m is responsible for the coupling between lattice sites separated by the distance m. The key ingredient to obtain the normal form is the preservation of the translation invariance (called cyclic symmetry in [18, 30]), which also allows us to produce estimates that are uniform with N.

If we limit to $r = 1$, the transformed Hamiltonian (3.2) reads

$$H^{(1)} = \mathcal{K} + P^{(2)} , \qquad \mathcal{K} := H_\Omega + Z_0 + Z_1 ,$$

where the quadratic and quartic polynomials Z_0 and Z_1 include all-to-all interactions, exponentially decaying with ϵ. Hence, $\mathcal{K}$ represents the Hamiltonian of the generalized dNLS equation. If we truncate both Z_0 and Z_1 at the leading order in ϵ, we recover the Hamiltonian of the usual dNLS equation.

The linear transformation is analyzed in Section 3.1. The nonlinear normal form transformation is performed in Section 3.2. Approximations with the usual dNLS equation are obtained in Section 3.3. Approximations with the generalized dNLS equation are discussed in Section 3.4.

3.1 Linear transformation

We start with the definitions of cyclic symmetry, interaction range, centered alignments and exponential decay (see also [18, 30]).

The translational invariance of the model (3.1) is formalized by using the idea of *cyclic symmetry*. The *cyclic permutation* operator τ is defined as

$$\tau(x_{-N}, \ldots, x_N) = (x_{-N+1}, \ldots, x_N, x_{-N}). \qquad (3.3)$$

This operator can be applied separately to the variables x and y. We extend the action of this operator on the space of functions as $(\tau f)(x, y) = f(\tau x, \tau y)$.

Definition 1 (Cyclic symmetry) *We say that a function F is* cyclically symmetric *if $\tau F = F$.*

We introduce an operator, indicated by an upper symbol $\oplus$, acting on functions: given a function f, a new function $F = f^{\oplus}$ is constructed as

$$F = f^{\oplus} := \sum_{l=-N}^{N} \tau^l f \ . \tag{3.4}$$

We say that $f^{\oplus}(x, y)$ is generated by the *seed* $f(x, y)$. Our convention is to denote the cyclically symmetric functions by capital letters and their seeds by the corresponding lower case letters. It is worth to note that the Hamiltonian $H = H_0 + H_1$ defined by (3.1) is clearly of the form $H = h^{\oplus}$, generated by the seed $h(x, y) = \frac{1}{2}(y_0^2 + x_0^2) - \epsilon x_1 x_0 + \frac{1}{4} x_0^4$.

Definition 2 (Interaction range) *Given the exponents (j, k), we define the support $S(x^j y^k)$ of the monomials $x^j y^k$ and the interaction distance $\ell(x^j y^k)$ as follows:*

$$S(x^j y^k) = \{l \ : \ j_l \neq 0 \ \text{or} \ k_l \neq 0\} \ , \quad \ell(x^j y^k) = \mathrm{diam}\big(S(x^j y^k)\big) \ . \tag{3.5}$$

We stress that, differently from what has been developed in [17, 18], it is possible to impose that the seeds of all the functions are *centrally aligned*, according to the following definition [30].

Definition 3 (Centered alignment) *Let $F = f^{\oplus}$ be a cyclically symmetric function, with f depending on $2N + 1$ variables, $f = f(x_{-N}, \ldots, x_0, \ldots, x_N)$. The seed f is said centrally aligned if it admits the decomposition*

$$f = \sum_{m=0}^{N} f^{(m)} \ , \qquad S(f^{(m)}) \subseteq [-m, \ldots, m] \ . \tag{3.6}$$

In order to formalize and control the interaction range, we introduce one more definition.

Definition 4 (Exponential decay) *The seed f of a function F is said to be of class $\mathcal{D}(C_f, \mu)$ if there exist two positive constants C_f and $\mu < 1$ such that for any centrally aligned component $f^{(m)}$ it holds*

$$\left\| f^{(m)} \right\| \leq C_f \mu^m \ , \qquad m = 0, \ldots, N \ ,$$

where $\|\cdot\|$ is a standard polynomial norm[1].

Remark 5 *If we are dealing with an Hamiltonian, Definition 4 encodes, when the constant C_f does not depend on N, the short range nature of the interaction; this, together with the translation invariance given by means of the cyclic symmetry constitute the extensivity of the Hamiltonian.*

---

[1]Given a homogeneous polynomial $f(x, y) = \sum_{|j|+|k|=s} f_{j,k} x^j y^k$ of degree s in x, y and a positive radius R, we define the polynomial norm of f by $\|f\|_R := R^s \sum_{|j|+|k|=s} |f_{j,k}|$; we often drop the subscript R.

We can now focus on the harmonic part H_0 of the Hamiltonian H. From (3.1), H_0 can be written as the quadratic form

$$H_0(x,y) = \frac{1}{2}y \cdot y + \frac{1}{2}Ax \cdot x \tag{3.7}$$

where A is a circulant and symmetric matrix given by

$$A := \mathbb{I} - \epsilon(\tau + \tau^\top) . \tag{3.8}$$

Here $\tau = (\tau_{ij})$ is the matrix representing the cyclic permutation (3.3), i.e. with $\tau_{ij} = \delta_{i,j+1 \,(\mathrm{mod}\, 2N+1)}$ using the Kronecker's delta notation. The following proposition reduces the quadratic part H_0 of the Hamiltonian to the quadratic normal form and preserves the extensivity of H_0.

Proposition 1 *For every $\epsilon \in (0, \frac{1}{2})$ the canonical linear transformation $q = A^{1/4}x$, $p = A^{-1/4}y$ transforms the quadratic Hamiltonian H_0 to the quadratic normal form*

$$H^{(0)} = H_\Omega + Z_0 , \qquad \{H_\Omega, Z_0\} = 0 , \tag{3.9}$$

where $H_\Omega = h_\Omega^\oplus$ and $Z_0 = \zeta_0^\oplus$ are cyclically symmetric polynomials, with centrally aligned seeds h_Ω and ζ_0 of the form

$$h_\Omega = \frac{\Omega}{2}(q_0^2 + p_0^2) \tag{3.10}$$

and

$$\zeta_0 = \sum_{m=1}^N \zeta_0^{(m)}, \quad \zeta_0^{(m)} = b_m[q_0(q_m + q_{-m}) + p_0(p_m + p_{-m})]. \tag{3.11}$$

Here Ω and b_m are defined by

$$\Omega := \frac{1}{2N+1}\sum_{j=-N}^{N+1} \omega_j , \quad b_m := \left(A^{1/2}\right)_{1,m+1} , \tag{3.12}$$

where ω_j are the frequencies of the normal modes of H_0. Moreover, there exists a suitable positive constant C_{ζ_0} such that each component $\zeta_0^{(m)}$ satisfies the exponential decay

$$\left\|\zeta_0^{(m)}\right\| \le C_{\zeta_0}(2\epsilon)^m ,$$

hence $\zeta_0 \in \mathcal{D}\big(C_{\zeta_0}, 2\epsilon\big)$.

Proof. We give here only few ideas to grasp the exponential decay of the all-to-all interactions due to the linear transformation. After applying $q = A^{1/4}x$, $p = A^{-1/4}y$, we have

$$H_0 = \frac{1}{2}p^\top A^{1/2}p + \frac{1}{2}q^\top A^{1/2}q. \tag{3.13}$$

By defining $T := \tau + \tau^\top$, one can rewrite $A^{1/2}$ as

$$A^{1/2} = (\mathbb{I} - \epsilon T)^{1/2} = \sum_{l=0}^\infty \binom{1/2}{l}(-\epsilon)^l T^l .$$

13

In order to obtain the decomposition (3.9), we separate the diagonal part from the off-diagonal part $A^{1/2} = \Omega\, \mathbb{I} + B$ and insert this decomposition into (3.13). The exponential decay $(2\epsilon)^m$ comes from the observation that $\left(T^l\right)_{1,m+1} = 0$ for all $0 \le l < m$ and from the estimate $\left|(T^m)_{1,m+1}\right| \le 2^m$. One can restrict to consider only the first raw due to the circulant nature of all the matrices involved (for all details see Appendix 6.1.1 in [18]). ∎

The following proposition shows how the linear transformation in Proposition 1 changes the quartic part H_1 of the Hamiltonian and preserves the extensivity of H_1.

Proposition 2 *Under the linear transformation in Proposition 1, the quartic part H_1 given in (3.1) is cyclically symmetric ($H_1 = h_1^{\oplus}$) with a centrally aligned seed given by*

$$h_1 = \sum_{m=0}^{N} h_1^{(m)} \ . \tag{3.14}$$

Moreover, there exists a suitable positive constant C_{h_1} such that each component $h_1^{(m)}$ satisfies the exponential decay

$$\left\| h_1^{(m)} \right\| \le C_{h_1} (2\epsilon)^m \ ,$$

hence $h_1 \in \mathcal{D}\big(C_{h_1}, 2\epsilon\big)$.

Proof. The proof of this proposition includes some technical steps, similar to those in the proof of Proposition 2 in [30] and Lemma 3.4 in [18]. We only stress here that there is no loss in the decay rate (2ϵ) between the seeds of Z_0 and H_1 thanks to the different choice of alignment, as proven in Lemma 5 of [30]. ∎

We can clarify the statements of Propositions 1 and 2 by saying that in a suitable set of coordinates, the coupling part of the quadratic Hamiltonian H_0 shows all-to-all linear interactions, with an exponentially decaying strength with respect to the distance between the sites. Such a linear transformation introduces similar all-to-all interactions also in the quartic Hamiltonian H_1. Moreover, in the new coordinates q_j, the seed h_1 of the quartic term has the same exponential decay as the seed ζ_0 of the quadratic term.

3.2   First-order nonlinear normal form transformation

By using Propositions 1 and 2, the Hamiltonian H in (3.1) is transformed into the form

$$H = H_\Omega + Z_0 + H_1. \tag{3.15}$$

We are now ready to state the (first-order) normal form theorem. This first-order theorem represents the easiest formulation of the more generic Theorem 1 of [30]. The idea is to perform, by using the Lie transform algorithm explained in [16], one normalizing step, provided ϵ is small enough. Moreover, the normalizing canonical transformation is well defined in a (small) neighborhood $B_{\rho^{1/2}}$ of the origin, where ρ is sufficiently small.

Theorem 5 *Consider the Hamiltonian $H = h_\Omega^\oplus + \zeta_0^\oplus + h_1^\oplus$ with seeds h_Ω, ζ_0, h_1, in (3.10), (3.11), and (3.14). There exist positive γ, $\epsilon_* < \frac{1}{2}$ and C_* such that for every $\epsilon \in (0, \epsilon_*)$, there exists a generating function $\mathcal{X}_1 = \chi_1^\oplus$ of a Lie transform such that $T_{\mathcal{X}_1} H^{(1)} = H$, where $H^{(1)}$ is a cyclically symmetric function of the form*

$$H^{(1)} = H_\Omega + Z_0 + Z_1 + P^{(2)} , \tag{3.16}$$

with $0 = \{H_\Omega, Z_0\} = \{H_\Omega, Z_1\}$, whereas $Z_1 = \zeta_1^\oplus$ is a polynomial of degree four whose seed ζ_1 is of class $\mathcal{D}(C_{h_1}, 2\epsilon)$, and $P^{(2)}$ is a remainder that includes terms of degree equal or bigger than six. Moreover, if the smallness condition on the energy

$$\rho < \rho_* := \frac{1}{96(1+e)C_*} , \tag{3.17}$$

is satisfied, then the following statements hold true:

1. *$\mathcal{X}_1$ defines an analytic canonical transformation on the domain $B_{\frac{2}{3}\rho^{1/2}}$ such that*

$$B_{\frac{1}{3}\rho^{1/2}} \subset T_{\mathcal{X}_1} B_{\frac{2}{3}\rho^{1/2}} \subset B_{\rho^{1/2}} \qquad B_{\frac{1}{3}\rho^{1/2}} \subset T_{\mathcal{X}_1}^{-1} B_{\frac{2}{3}\rho^{1/2}} \subset B_{\rho^{1/2}} .$$

   Moreover, the deformation of the domain $B_{\frac{2}{3}\rho^{1/2}}$ is controlled by

$$z \in B_{\frac{2}{3}\rho^{1/2}} \qquad \Rightarrow \qquad \|T_{\mathcal{X}_1}(z) - z\| \le 4^4 C_* \rho^{3/2} , \qquad \left\|T_{\mathcal{X}_1}^{-1}(z) - z\right\| \le 4^4 C_* \rho^{3/2} . \tag{3.18}$$

2. *the remainder is an analytic function on $B_{\frac{2}{3}\rho^{1/2}}$, and it is represented by a series of cyclically symmetric homogeneous polynomials $H_s^{(1)}$ of degree $2s + 2$*

$$P^{(2)} = \sum_{s=2}^\infty H_s^{(1)} \qquad H_s^{(1)} = \left(h_s^{(1)}\right)^\oplus , \qquad h_s^{(1)} \in \mathcal{D}(2\tilde{C}_*^{s-1} C_{h_1}, \sqrt{2\epsilon}) . \tag{3.19}$$

The interval $(0, \epsilon_*)$ with $\epsilon_* < \frac{1}{2}$ comes from the inequality

$$f(\epsilon) := \left(\frac{3\Omega}{64 C_{\zeta_0}}\right) \frac{(1 - 2\epsilon)\left[1 - (2\epsilon)^{\frac{3}{4}}\right]}{\sqrt{2\epsilon}} > 1$$

(see for reference formula (33) in [31]), and the constants C_* and γ can be written as

$$C_* = \frac{4 C_{h_1}}{3\gamma(1 - 2\epsilon)\left[1 - (2\epsilon)^{\frac{3}{4}}\right]} \tag{3.20}$$

and

$$\gamma = 2\Omega \left(1 - \frac{1}{2f(\epsilon)}\right) \qquad \Rightarrow \qquad \Omega < \gamma < 2\Omega. \tag{3.21}$$

Since ϵ is sufficiently smaller than $\frac{1}{2}$, the constants C_* is essentially independent on ϵ, i.e.

$$C_* = \mathcal{O}\left(\frac{C_{h_1}}{\Omega}\right),$$

which implies that the same holds true for the threshold ρ_* so that

$$\rho_* \approx \frac{2\Omega}{3 C_{h_1}(1 + e)} . \tag{3.22}$$

15

3.3 Approximation with the dNLS equation

We apply here the normal form transformation of Theorem 5 in order to approximate the Cauchy problem $\dot{z} = \{H, z\}$ of the finite dKG equation (1.4) with a small initial datum z_0. Let us denote with $\mathcal{K} := H_\Omega + Z_0 + Z_1$ the normal form part of the Hamiltonian $H^{(1)} = \mathcal{K} + P^{(2)}$ in formula (3.16). Since Z_0 and Z_1 have centrally aligned seeds with the exponential decay (see decompositions (3.11) and (3.14)), we have

$$Z_0 = \sum_{m=1}^{N} Z_0^{(m)} \, , \qquad Z_0^{(m)} := \left(\zeta_0^{(m)} \right)^{\oplus} \tag{3.23}$$

and

$$Z_1 = \sum_{m=0}^{N} Z_1^{(m)} \, , \qquad Z_1^{(m)} := \left(\zeta_1^{(m)} \right)^{\oplus} \, . \tag{3.24}$$

Note that the expansion for Z_0 starts at $m = 1$, while Z_1 starts with $m = 0$. By truncating the ϵ expansion of each normal form term Z_j at their leading orders, we define the *effective normal form Hamiltonian* $\mathcal{K}_{\text{eff}}$ as

$$\mathcal{K}_{\text{eff}} := H_\Omega + Z_0^{(1)} + Z_1^{(0)} \, , \qquad \mathcal{K}_{\text{res}} := \mathcal{K} - \mathcal{K}_{\text{eff}} \, . \tag{3.25}$$

As already stressed in [30], the truncated normal form $\mathcal{K}_{\text{eff}}$ represents the Hamiltonian of the dNLS equation. In complex coordinates $\psi_j = (q_j + \mathrm{i}p_j)/\sqrt{2}$, the Hamiltonian $\mathcal{K}_{\text{eff}}$ reads as

$$\mathcal{K}_{\text{eff}} = (\Omega + 2b_1) \sum_j |\psi_j|^2 - b_1 \sum_j |\psi_{j+1} - \psi_j|^2 + \frac{3}{8} \sum_j |\psi_j|^4 \, , \tag{3.26}$$

where $b_1 = \mathcal{O}(\epsilon) < 0$ is the same as in the expression (3.12) of Proposition 1. The corresponding dNLS equation is

$$\mathrm{i}\dot{\psi}_j = \frac{\partial \mathcal{K}_{\text{eff}}}{\partial \overline{\psi}_j} = \Omega \psi_j + b_1 (\psi_{j+1} + \psi_{j-1}) + \frac{3}{4} \psi_j |\psi_j|^2 \, , \tag{3.27}$$

and it has the same structure as the dNLS equation (2.3).

We denote with $z(t)$ the evolution of the dKG transformed Hamiltonian $\mathcal{K} + P^{(2)}$, with $z_a(t)$ the evolution of the dNLS model $\mathcal{K}_{\text{eff}}$ and consequently with $\delta(t)$ the error

$$\delta(t) := z(t) - z_a(t) \, . \tag{3.28}$$

The two time scales over which we control the error of the approximation are given by

$$T_0 := \frac{1}{\rho} \, , \qquad T_0^* := \frac{\alpha}{\kappa_0 \rho} \ln \left(\frac{1}{\rho} \right) \, , \tag{3.29}$$

where $\alpha \in (0, 1)$ is an arbitrary parameter, and $\kappa_0 = \mathcal{O}(C_{h_1})$ is given in (3.43). Similar definitions are used in (2.18) and (2.26), in the proof of Theorems 1 and 2.

Theorem 6 *Let us take ρ fulfilling (3.17) and $\epsilon \in (0, \epsilon_*)$ as in Theorem 5. Let us first consider the two independent parameters ρ and ϵ in the regime $\epsilon^2 \ll \rho \leq \epsilon$. Then, there exists a positive constant C independent of ρ and ϵ such that for any initial datum $z_0 \in B_{\frac{2}{3}\rho^{1/2}}$ with $\|\delta_0\| \leq \rho^{-1/2}\epsilon^2$, the following holds true:*

$$\|\delta(t)\| \leq \mathsf{C}\rho^{-1/2}\epsilon^2 \ , \qquad |t| \leq T_0 \ . \tag{3.30}$$

Let us now consider the two independent parameters ρ and ϵ in the regime $\epsilon^{\frac{2}{1+\alpha}} \ll \rho \leq \epsilon$, where $\alpha \in (0, 1)$ is arbitrary. Then, there exists a positive constant C independent of ρ and ϵ such that for any initial datum $z_0 \in B_{\frac{2}{3}\rho^{1/2}}$ with $\|\delta_0\| \leq \rho^{-1/2}\epsilon^2$, the following holds true:

$$\|\delta(t)\| \leq \rho^{-1/2-\alpha}\epsilon^2, \quad |t| \leq T_0^* \ . \tag{3.31}$$

Remark 6 *The upper bound for the error δ given in (3.30) and (3.31) refers to the time evolution of the normal form (3.26) in the transformed variables ψ, which are near-identity deformations of the original variables (x, y). Since the transformation T_χ is Lipschitz, with a Lipschitz constant L of order $L = \mathcal{O}(1)$, the same bound of the error holds also in the original coordinates. Thus, from the analytic point of view, the nonlinear deformation of the variables does not affect the dependence of the estimates on ρ and ϵ: only the constant C is changed by the Lipschitz factor L.*

Remark 7 *The above estimates are equivalent, both in terms of error smallness and time scale, to the ones obtained in Theorems 1 and 2, once the original variables $x_j = \rho^{1/2}\xi_j$ are recovered.*

Remark 8 *The requirement $\epsilon^2 \ll \rho$ on the time scale T_0 is needed in order to provide a meaningful approximation, which means that the error is much smaller than the leading approximation $z_a(t)$*

$$\|\delta(t)\| \leq \rho^{-1/2}\epsilon^2 \ll \rho^{1/2} \sim \|z_a(t)\| \ .$$

The same reason lies behind the requirement $\epsilon^{\frac{2}{1+\alpha}} \ll \rho$ on the extended time scale T_0^*.

Before entering into the proof of Theorem 6, we need a further definition in order to control the norm of vector fields. Given F an extensive Hamiltonian with seed f, we will make use of the notation X_F to indicate the associated Hamiltonian vector field $J\nabla F$, with J given by the standard Poisson structure. The Hamiltonian vector field inherits, in a particular form, the cyclic symmetry. Indeed, it was proved in [30, 31] that

$$\partial_{x_j}F = \tau^j\partial_{x_0}F, \quad \partial_{y_j}F = \tau^j\partial_{y_0}F, \quad j = -N, \ldots, N \ . \tag{3.32}$$

As a result, a possible (but not unique) choice for the seed of X_F is given by the couple $(\partial_{y_0}F, -\partial_{x_0}F)$. This fact allows us to define in a reasonable and consistent way the following norm

$$\left\|X_F\right\|_R^\oplus := \|\partial_{y_0}F\|_R + \|\partial_{x_0}F\|_R \ . \tag{3.33}$$

As is shown in Proposition 1 of [30], the norm (3.33) allows us to control a natural operator norm. Moreover, as is stated in Lemma 4 of [30], if $F = f^\oplus$ with f of class $\mathcal{D}(C_f, \mu)$, then $\left\|X_F\right\|_R^\oplus$ is controlled by C_f. Both these properties will be used in the forthcoming estimates (3.38) and (3.41).

Proof of Theorem 6. Following a standard approach, we first decompose the Hamiltonian $H = H_L + H_N$ in its quadratic and quartic parts

$$H_L := H_\Omega + Z_0, \quad H_N := Z_1 + P^{(2)},$$

so that $\mathcal{K}_{\text{eff}} = H_L + H_N - P^{(2)} - \mathcal{K}_{\text{res}}$. Correspondingly, the vector field is decomposed as $X_H = X_{H_L} + X_N$. Denote the linear operator for X_{H_L} by $\mathcal{L}$. The equation of motions for $z(t)$ and $z_a(t)$ reads

$$\begin{cases} \dot{z} = \mathcal{L}z + X_N(z) , \\ \dot{z}_a = \mathcal{L}z_a + X_N(z_a) - \text{Res}(t) , \end{cases} \quad \text{with} \quad \text{Res}(t) := X_{P^{(2)}}(z_a(t)) + X_{\mathcal{K}_{\text{res}}}(z_a(t)) . \quad (3.34)$$

The error $\delta(t)$ defined by (3.28) satisfies the equation

$$\dot{\delta} = \mathcal{L}\delta + [X_N(z_a + \delta) - X_N(z_a)] + \text{Res}(t) , \quad (3.35)$$

whose solution, with the initial value δ_0, is given by Duhamel formula

$$\delta(t) = e^{\mathcal{L}t}\delta_0 + e^{\mathcal{L}t} \int_0^t e^{-\mathcal{L}s}[X_N(z_a + \delta) - X_N(z_a) + \text{Res}(s)]ds . \quad (3.36)$$

Now, since $\{H_L, H_\Omega\} = 0$, one has that $\mathcal{L}$ is an isometry. This allows to estimate

$$\|\delta(t)\| \leq \|\delta_0\| + \int_0^t [\|X_N(z_a(s) + \delta(s)) - X_N(z_a(s))\| + \|\text{Res}(s)\|]ds . \quad (3.37)$$

The second term in the r.h.s. can be estimated with the definition of the residual and using the information that $z_a(t)$ preserves the norm, as a consequence of the conservation of H_Ω

$$\|X_{P^{(2)}}(z_a(s))\| \leq \mathsf{C}\frac{C_{h_1}C_*\rho^{5/2}}{(1 - \sqrt[4]{2\epsilon})^2} , \qquad \|X_{\mathcal{K}_{\text{res}}}(z_a(t))\| \leq \mathsf{C}\frac{\left[C_{\zeta_0}\rho^{1/2}\epsilon^2 + C_{h_1}\rho^{3/2}\epsilon\right]}{(1 - 2\epsilon)^2} , \quad (3.38)$$

where the two contributions in the second inequality come from the truncation of Z_0 and Z_1 respectively. Thus, we obtain

$$\|\text{Res}(s)\| \leq \mathsf{C}\frac{\rho^{1/2}}{(1 - \sqrt[4]{2\epsilon})^2}\left[C_{\zeta_0}\epsilon^2 + C_{h_1}\rho\epsilon + C_{h_1}C_*\rho^2\right] . \quad (3.39)$$

On the other hand, if

$$\|\delta\| \ll \|z_a\| \sim \rho^{1/2} , \quad (3.40)$$

then the increment of the nonlinear field can be well approximated by

$$\|X_N(z_a(s) + \delta(s)) - X_N(z_a(s))\| \leq \left\|X_N'(\zeta_a)\right\| \|\delta\| ,$$

where

$$\zeta_a := z_a + \lambda\delta , \qquad \lambda \in (0, 1) .$$

If the smallness condition (3.40) for δ holds, then $\|\zeta_a\| \sim \rho^{1/2}$, which implies

$$\|X_N(z_a(s) + \delta(s)) - X_N(z_a(s))\| \leq \|X_N'\|_\rho \|\delta\| .$$

By using the decomposition $X_N' = X_{Z_1}' + X_{P^{(2)}}'$ it is possible to obtain

$$\|X_N'\|_{\rho^{1/2}} \leq \mathsf{C}_1 \frac{C_{h_1}}{(1 - \sqrt[4]{2\epsilon})^2} \rho . \tag{3.41}$$

By inserting (3.39) and (3.41) into (3.37), one gets a typical Gronwall-like integral inequality, which provides the time-dependent upper bound

$$\begin{aligned}
\|\delta(t)\| &\leq e^{\kappa_0 \rho t} \|\delta_0\| + \mathsf{C}\rho^{1/2} \left[\frac{\epsilon^2}{\rho} + \epsilon + C_*\rho\right] \left(e^{\kappa_0 \rho t} - 1\right) \\
&\leq e^{\kappa_0 \rho t} \rho^{-1/2} \epsilon^2 + \mathsf{C}\rho^{-1/2} \left(e^{\kappa_0 \rho t} - 1\right) \left[\epsilon^2 + \rho\epsilon + C_*\rho^2\right] ,
\end{aligned} \tag{3.42}$$

where κ_0 provides an upper bound for $\|X_N'\|_1$ in (3.41)

$$\kappa_0 := \mathsf{C}_1 \frac{C_{h_1}}{(1 - \sqrt[4]{2\epsilon_*})^2} = \mathcal{O}(C_{h_1}) \tag{3.43}$$

and C depends only on ϵ_*, C_{ζ_0}, C_{h_1}. Then, the bound (3.30) follows from the assumption $\rho \leq \epsilon$.

The bound (3.31) is obtained similarly, just replacing the time span T_0^* in the above (3.42), which easily provides the factor $\rho^{-\alpha}$ in front of the estimate. $\qquad\square$

3.4   Approximations with the generalized dNLS equation

The standard dNLS approximation is no more valid when $\epsilon^2 \sim \rho$. Indeed, in such a case, the contribution $\epsilon^2 \rho^{-1}$ coming from the truncation of the linear field X_{H_L} in (3.42) is of order one, hence the error $\delta(t)$ can be comparable with the approximation $z_a(t)$

$$\|\delta(t)\| \leq \mathsf{C}\rho^{1/2} \sim \|z_a(t)\| .$$

In such a regime, it is then necessary to include in the Hamiltonian $\mathcal{K}_{\mathrm{eff}}$ at least the term $Z_0^{(2)}$, responsible for the next-nearest neighbourhood linear interaction:

$$\mathcal{K}_{\mathrm{eff}} := H_\Omega + Z_0^{(1)} + Z_0^{(2)} + Z_1^{(0)} . \tag{3.44}$$

Following the same steps as in the proof of Theorem 6, it is possible to prove the following result, which is fully equivalent to Theorems 3 and 4.

Theorem 7 *Let us take ρ fulfilling (3.17) and $\epsilon \in (0, \epsilon_*)$ as in Theorem 5. Let us first consider the two independent parameters ρ and ϵ in the regime $\epsilon^3 \ll \rho \leq \epsilon^2$. Then, there exists a positive constant C independent on ρ and ϵ such that for any initial datum $z_0 \in B_{\frac{2}{3}\rho}$ with $\|\delta_0\| \leq \rho^{-1/2}\epsilon^3$, it holds true*

$$\|\delta(t)\| \leq \mathsf{C}\rho^{-1/2}\epsilon^3 , \qquad |t| \leq T_0 . \tag{3.45}$$

Let us now consider the two independent parameters ρ and ϵ in the regime $\epsilon^{\frac{3}{1+\alpha}} \ll \rho \leq \epsilon$, where $\alpha \in (0, 1)$ is arbitrary. Then, there exists a positive constant C independent of ρ and ϵ such that for any initial datum $z_0 \in B_{\frac{2}{3}\rho^{1/2}}$ with $\|\delta_0\| \leq \rho^{-1/2}\epsilon^2$, the following holds true:

$$\|\delta(t)\| \leq \rho^{-1/2-\alpha}\epsilon^3 \ , \qquad |t| \leq T_0^* \ . \tag{3.46}$$

The result of Theorem 7 yields the Hamiltonian for the generalized dNLS equation:

$$\mathcal{K}_{\text{eff}} = (\Omega + 2b_1 + 2b_2)\sum_j |\psi_j|^2 - b_1 \sum_j |\psi_{j+1} - \psi_j|^2 - b_2 \sum_j |\psi_{j+2} - \psi_j|^2 + \frac{3}{8}\sum_j |\psi_j|^4 \ , \tag{3.47}$$

where $b_2 = \mathcal{O}(\epsilon^2) < 0$ is the same as in the expression (3.12) of Proposition 1. The corresponding generalized dNLS equation is

$$\mathrm{i}\dot{\psi}_j = \Omega\psi_j + b_1(\psi_{j+1} + \psi_{j-1}) + b_2(\psi_{j+2} + \psi_{j-2}) + \frac{3}{4}\psi_j|\psi_j|^2 \ , \tag{3.48}$$

which has the same structure as the generalized dNLS equation (2.31). Indeed, remembering that Ω in (3.48) also has an expansion in ϵ, and that the time variable is rescaled with ϵ in (2.31), we can rewrite the right-hand-side of the generalized dNLS equation (2.31) as follows:

$$\frac{\epsilon}{2}a_j + (a_{j+1} + a_{j-1}) + \frac{\epsilon}{4}(a_{j+2} + a_{j-2}) \ .$$

This shows an ϵ correction to the nearest neighbour coefficient, which in the normal form approach is embedded in the ϵ-dependence of Ω, b_1, b_2 and of the transformed coordinates.

More generally, within the normal form approach, different regimes of parameters can be treated with no efforts: once the requested scaling between ϵ and ρ is chosen, one easily derives the minimal, and also the optimal, number of terms in the expansions of Z_0 and Z_1 to be included. The estimates follows as easily as before. Here we give the estimates for a general choice of truncation:

$$\mathcal{K}_{\text{eff}} = H_\Omega + \sum_{j=1}^{l-1} Z_0^{(j)} + \sum_{j=0}^{n-1} Z_1^{(j)} \ , \tag{3.49}$$

where $N \geq l \geq 2$ and $N \geq n \geq 1$. The error term δ is now estimated similarly to (3.42) as follows:

$$\|\delta(t)\| \leq e^{\kappa_0 \rho t}\|\delta_0\| + \mathsf{C}\rho^{1/2}\left[\frac{\epsilon^l}{\rho} + \epsilon^n + C_*\rho\right]\left(e^{\kappa_0 \rho t} - 1\right) \ , \qquad l \geq 2 \ , \qquad n \geq 1 \ . \tag{3.50}$$

Hence one can deal with all the regimes and with the desired error precision in a compact and flexible way. The extension to higher order terms in the nonlinearity would require further steps of the normal form transformations, thus modifying thresholds ϵ_* and ρ_*, following the general version of Theorem 5 given in [30].

4 Applications of the dNLS equation

We conclude the paper with a brief account of possible applications of the dNLS equations [see (2.3) and (3.27)], and their generalizations [see (2.31) and (3.48)], in the context of small-amplitude weakly coupled oscillators of the dKG equation (1.4).

Existence of breathers. Breathers of the dKG equation (time-periodic solutions localized on the lattice) can be constructed approximately by looking at the discrete solitons of the dNLS equation (1.5) in the form $a_j(\tau) = A_j e^{i\Omega\tau}$, where $\Omega \in \mathbb{R}\backslash[-1, 1]$ is defined outside the spectral band of the linearized dNLS equation and $\mathbf{A} \in \ell^2(\mathbb{R})$ is time-independent.

The limit $\epsilon \to 0$ is referred to as the anti-continuum limit of the dKG equation (1.4), when the breathers at a fixed energy are continued uniquely from the limiting configurations supported on few lattice sites [27, 33]. Compared to the anti-continuum limit, the dNLS approximation is very different, because the discrete solitons of the dNLS equation (1.5) are not nearly compactly supported due to the fact that the dNLS equation (1.5) has no small parameter. Indeed, the continuation arguments in [27, 33] are no longer valid in the small-amplitude approximation, when the breather period T is defined near the linear limit 2π, because the inverse linearized operators become unbounded in the linear oscillator limit as $T \to 2\pi$. As a result, approximate breathers obtained from Theorem 1 are no longer compactly supported.

The approximation of Theorem 1 and the construction of truly periodic solutions to the dKG equation (1.4) can be extended to all times. To do so, we can use the Fourier series in time and eliminate all but the first Fourier harmonic by a Lyapunov–Schmidt reduction procedure. Then, the components of the first Fourier component satisfies a stationary dNLS-type equation, where the dNLS equation (1.5) is the leading equation. In this way, similarly to the work [34], one can justify the continuation of discrete solitons of the dNLS equation (1.5) as approximations of the truly periodic breathers in the dKG equation (1.4).

Within the same scheme of Lyapunov–Schmidt decomposition, another equivalent route to prove the existence of breathers in the dKG equation (1.4) is obtained by means of Theorem 5. Indeed, the discrete solitons of the dNLS equation (1.5) can be characterized as constrained critical points of the energy, which are continued, under non-degeneracy conditions, to critical points of the true energy of the dKG equation (1.4), see [31] for details.

Spectral stability of breathers. The spectral stability of breathers in the dKG equation (1.4) can be related to the spectral stability of solitons in the dNLS equation (1.5). By Theorem 1 with $\rho = \epsilon$, we are not able to relate stable or unstable eigenvalues of the dNLS solitons with the Floquet multipliers of the dKG breathers, because the error term also grows exponentially at the time scale $\mathcal{O}(\epsilon^{-1})$ (the same problem is discussed in [12, 24] in the context of stability of the travelling waves in FPU lattices). However, by Theorem 2 with $\rho = \epsilon$, the approximation result is extended to the time scale $\mathcal{O}(\epsilon^{-1}|\log(\epsilon)|)$. Therefore, we can conclude that all the unstable eigenvalues of the dNLS solitons persist as unstable Floquet multipliers of the dKG breathers within the $\mathcal{O}(\epsilon)$ distance from the unit circle.

If the unstable eigenvalues of the dNLS solitons do not exist, we only obtain approximate spectral stability of the dKG breathers, because the unstable Floquet multipliers of the dKG breathers may still exist on the distance smaller than $\mathcal{O}(\epsilon)$ to the unit circle. On the other hand,

if the spectrally stable dNLS solitons are known to have internal modes [32], then the Floquet multipliers of the dKG breathers persist on the unit circle by known symmetries of the Floquet multipliers [33].

Long time stability of breathers. By means of the normal form approach, it is possible to prove the long time stability result for single-site (fundamental) breather solutions of the dKG equation (1.4). Indeed, the variational characterization of the existence problem for such breathers in the normal form essentially implies an orbital stability in the normal form, which is translated into a long time stability in the original dKG equation [31].

In the case of multi-site dNLS solitons, nonlinear instability is induced by isolated internal modes of negative Krein signature, which are coupled with the continuous spectrum by the non-linearity [23]. By using the extended time scale $\mathcal{O}(\epsilon^{-1}|\log(\epsilon)|)$ of Theorem 2 with $\rho = \epsilon$, we can predict persistence of this instability for small-amplitude dKG breathers. This nonlinear instability was recently confirmed for multi-site dKG breathers in [11].

Also quasi-periodic localized solutions were constructed for the dNLS equation, in the situation when the internal mode of the dNLS soliton occurs on the other side of the spectral band of the continuous spectrum [10, 28]. These solutions correspond approximately to quasi-periodic dKG breathers. It is still an open question to consider true quasi-periodic breather solutions of the dKG equation (1.4).

References

[1] J.F.R. Archilla, J. Cuevas, B. Sänchez-Rey, and A. Alvarez, "Demonstration of the stability or instability of multibreathers at low coupling", Physica D **180** (2003), 235–255.

[2] D. Bambusi, "Exponential stability of breathers in Hamiltonian networks of weakly coupled oscillators", Nonlinearity **9** (1996), 433–457.

[3] D. Bambusi, "Galerkin averaging method and Poincaré normal form for some quasilinear PDEs", Ann. Sc. Norm. Super. Pisa Cl. Sci. **(5) 4** (2005), 669–702.

[4] D. Bambusi, "Asymptotic stability of breathers in some Hamiltonian networks of weakly coupled oscillators", Comm. Math. Phys. **324** (2013), 515–547.

[5] D. Bambusi and A. Ponno. "On metastability in FPU", Comm. Math. Phys. **264** (2006), 539-561.

[6] D. Bambusi, S. Paleari, and T. Penati, "Existence and continuous approximation of small amplitude breathers in 1D and 2D Klein-Gordon lattices", Appl. Anal. **89** (2010), 1313–1334.

[7] D. Bambusi and T. Penati, "Continuous approximation of breathers in 1D and 2D DNLS lattices", Nonlinearity **23** (2010), 143–157.

[8] B. Bidegaray–Fesquet, E. Dumas, and G. James, "From Newton's cradle to the discrete p-Schrödinger equation", SIAM J. Math. Anal. **45** (2013), 3404–3430.

[9] C. Chong, D.E. Pelinovsky, and G. Schneider, "On the validity of the variational approximation in discrete nonlinear Schrodinger equations", Physica D **241** (2012), 115–124.

[10] S. Cuccagna, "Orbitally but not asymptotically stable ground states for the discrete NLS", Discrete Contin. Dyn. Syst. **26** (2010), 105–134.

[11] J. Cuevas–Maraver, P.G. Kevrekidis, and D.E. Pelinovsky, "Nonlinear instabilities of multi-site breathers in Klein–Gordon lattices", Stud. Appl. Math. **137** (2016), 214–237.

[12] E. Dumas and D.E. Pelinovsky, "Justification of the log-KdV equation in granular chains: the case of precompression", SIAM J. Math. Anal. **46** (2014), 4075–4103.

[13] G. Friesecke and R.L. Pego, "Solitary waves on FPU lattices : I. Qualitative properties, renormalization and continuum limit", Nonlinearity **12** (1999), 1601-1627.

[14] C. Garcia–Azpeitia, "Periodic traveling and standing waves in a circular chain of coupled pendula and Newton's cradle", J. Nonlinear Sci. (2016), DOI: 10.1007/s00332-016-9318-5.

[15] C. Garcia–Azpeitia, "Global bifurcation of periodic traveling waves in discrete nonlinear Schrödinger equations", arXiv:1508.00832 (16 pages).

[16] A. Giorgilli, "Exponential stability of Hamiltonian systems", Dynamical systems. Part I, Pubbl. Cent. Ric. Mat. Ennio Giorgi, Scuola Norm. Sup., Pisa (2003), 87–198.

[17] A. Giorgilli, S. Paleari, and T. Penati, "Extensive adiabatic invariants for nonlinear chains", J. Statist. Phys., **148** (2012), 1106–1134.

[18] A. Giorgilli, S. Paleari, and T. Penati, "An extensive adiabatic invariant for the Klein-Gordon model in the thermodynamic limit", Annales Henri Poincaré, **16** (2015), 897–959.

[19] J. Giannoulis and A. Mielke, "The nonlinear Schrödinger equation as a macroscopic limit for an oscillator chain with cubic nonlinearities", Nonlinearity **17** (2004), 551–565.

[20] J. Giannoulis and A. Mielke, "Dispersive evolution of pulses in oscillator chains with general interaction potentials", Discrete Contin. Dyn. Syst. Ser. B **6** (2006), 493–523.

[21] J. Giannoulis, M. Herrmann, and A. Mielke, "Lagrangian and Hamiltonian two-scale reduction", J. Math. Phys. **49** (2008), 103505, 42 pp.

[22] J. Holmer and M. Zworski, "Soliton interaction with slowly varying potentials", Int. Math. Res. Not. **2008**, rnn026, 36 pages.

[23] P.G. Kevrekidis, D.E. Pelinovsky, and A. Saxena, "When linear stability does not exclude nonlinear instability", Phys. Rev. Lett. **114** (2015), 214101 (6 pages).

[24] A. Khan and D.E. Pelinovsky, "Long-time stability of small FPU solitary waves", arXiv:1603.01289 (2016) (12 pages)

[25] V. Koukouloyannis and P.G. Kevrekidis, "On the stability of multibreathers in Klein–Gordon chains", Nonlinearity **22** (2009), 2269–2285.

[26] D. Lannes and J. Rauch, "Validity of nonlinear geometric optics with times growing logarithmically", Proc. AMS **129** (2000), 1087–1096.

[27] R.S. MacKay and S. Aubry, "Proof of existence of breathers for time-reversible or Hamiltonian networks of weakly coupled oscillators", Nonlinearity **7** (1994), 1623-1643.

[28] M. Maeda, "Existence and asymptotic stability of quasi-periodic solution of discrete NLS with potential in $\mathbb{Z}$", arXiv: 1412.3213 (2014)

[29] A.M. Morgante, M. Johansson, G. Kopidakis, and S. Aubry, "Standing wave instabilities in a chain of nonlinear coupled oscillators", Physica D **162** (2002), 53–94.

[30] S. Paleari and T. Penati, "An extensive resonant normal form for an arbitrary large Klein-Gordon model", Annali Matematica Pura ed Applicata, **240** (2016), 133–165.

[31] S. Paleari and T. Penati, "Long time stability of small amplitude breathers in a mixed FPU-KG model", ZAMP, to appear (2016).

[32] D. Pelinovsky and A. Sakovich, "Internal modes of discrete solitons near the anti-continuum limit of the dNLS equation", Physica D **240** (2011), 265–281.

[33] D.E. Pelinovsky and A. Sakovich, "Multi-site breathers in Klein-Gordon lattices: stability, resonances, and bifurcations", Nonlinearity **25** (2012), 3423–3451.

[34] D. Pelinovsky, G. Schneider, and R. MacKay, Justification of the lattice equation for a nonlinear elliptic problem with a periodic potential, Comm. Math. Phys. **284** (2008), 803–831.

[35] D. Pelinovsky and G. Schneider, "Bounds on the tight-binding approximation for the Gross–Pitaevskii equation with a periodic potential", J. Diff. Eqs. **248** (2010), 837–849.

[36] T. Penati and S. Paleari, "Breathers and Q-Breathers: Two Sides of the Same Coin", SIAM J. Applied Dynamical Systems, **11** (2012), 1–30.

[37] G. Schneider, "Bounds for the nonlinear Schrödinger approximation of the Fermi-Pasta-Ulam system", Appl. Anal. **89** (2010), 1523–1539.

[38] G. Schneider and C.E. Wayne, "Counter-propagating waves on fluid surfaces and the continuum limit of the Fermi-Pasta-Ulam model", In *International Conference on Differential Equations (Berlin, 1999), vol. 1 (eds B Fiedler, K Gröger, J Sprekels)*, pp. 390–404 (World Sci. Publishing, River Edge, NJ, USA, 2000).