



Approximation of (some) random FPUT lattices by KdV equations

Joshua A. McGinnis^a, J. Douglas Wright^{b,*}

^a University of Pennsylvania, Philadelphia, United States of America

^b Drexel University, Philadelphia, United States of America

ARTICLE INFO

Communicated by F. Bao-Feng

Keywords:

FPUT lattices

KdV approximation

Solitary waves

Stochastic homogenization

Random coefficient differential equations

ABSTRACT

We consider a Fermi–Pasta–Ulam–Tsingou lattice with randomly varying coefficients. We discover a relatively simple condition which when placed on the nature of the randomness allows us to prove that small amplitude/long wavelength solutions are almost surely rigorously approximated by solutions of Korteweg–de Vries equations for very long times. The key ideas combine energy estimates with homogenization theory and the technical proof requires a novel application of autoregressive processes.

1. Introduction

Consider a variable mass Fermi–Pasta–Ulam–Tsingou (FPUT) lattice¹:

$$\dot{q}(j, t) = \delta^+ p(j, t) \quad \text{and} \quad m(j) \dot{p}(j, t) = \delta^- [\mathcal{V}'(q)](j, t). \quad (1.1)$$

Here $t \in \mathbf{R}$, $j \in \mathbf{Z}$ and the unknowns $q(j, t)$ (the *relative displacement*) and $p(j, t)$ (the *velocity*) are real-valued. The *mass coefficients* $m(j)$ are strictly positive and

$$\mathcal{V}(q) := \frac{1}{2} q^2 + \frac{1}{3} q^3 \quad (1.2)$$

is the *spring potential*.² Lastly

$$\delta^+ f(j) = f(j+1) - f(j) \quad \text{and} \quad \delta^- f(j) = f(j) - f(j-1)$$

are the *right and left finite-difference operators*.

Models of this sort are ubiquitous in applications. A partial list: molecular dynamics, lamination, nondestructive testing, vehicular traffic, granular media, metamaterials, chemistry/biochemistry, and power generation [3]. The system (1.1) also plays a major role as a paradigm for the mathematical analysis of wave propagation—especially solitary waves—in nonlinear dispersive settings and it is the system's famous

connection to the Korteweg–de Vries (KdV) equation wherein our interest lies. Here are several important mathematical results about that connection:

- When $m(j)$ is constant, long-wavelength (say like $1/\epsilon$, where $0 < \epsilon \ll 1$), small-amplitude (order of ϵ^2) solutions are well-approximated over long time scales (order of $1/\epsilon^3$) by solutions of KdV equations. The relative ℓ^2 -error made by the approximation in this case is $\mathcal{O}(\epsilon)$. See [4] for the earliest formal derivation and [5] for the first rigorous result.
- The same sort of result holds when $m(j)$ is N -periodic (that is, $m(j+N) = m(j)$ for all j). Indeed, the spring potentials may also be taken to be N -periodic (for instance, replace $\mathcal{V}(q)$ with $\mathcal{V}_j(q) = \kappa(j)q^2/2 + \beta(j)q^3/3$ where $\kappa(j)$ and $\beta(j)$ are N -periodic). See [6,7].

While there have been a few derivations of KdV from random versions of the FPUT lattice previously (specifically [8,9]), all have been purely formal with no rigorous quantitative results. Even conjectures for the size of the error have been absent. We have been working for some time to remedy this. In our article [2] we discovered that if the $m(j)$ are independent identically distributed (i.i.d.) random variables

* Corresponding author.

E-mail address: jdoug@math.drexel.edu (J.D. Wright).

¹ We write the equations as a first order system as opposed to its possibly more familiar second order form

$$m(j) \ddot{x}(j) = \mathcal{V}'(x(j+1) - x(j)) - \mathcal{V}'(x(j) - x(j-1)).$$

The change of variables leading from this to (1.1) is $q(j) = x(j+1) - x(j)$ and $p(j) = \dot{x}(j)$.

² This choice of the spring potential—which is an instance of the “ α -potential” from [1]—is made mainly for simplicity. We could allow more complicated potentials and, so long as we had $\mathcal{V}'(0) > 0$ and $\mathcal{V}''(0) \neq 0$, only minor changes to our results would occur.

³ We also allow for heterogeneity in $\mathcal{V}$ as well as in m in [2]; our results apply to the case where $\mathcal{V}'(q)$ replaced by $\mathcal{V}'_j(q) = \kappa(j)q$ with $\kappa(j)$ another collection of i.i.d. random variables.

then the accuracy of long-wavelength approximations is substantially diminished and, consequently, only shorter time scales and the linear problem (that is when $\mathcal{V}'(q) = q$) are within reach. Precisely, we showed³ that long-wavelength solutions (again like $1/\epsilon$) converge almost surely and strongly, but rather slowly, to solutions of a wave equation on time scales on the order of $1/\epsilon$: the relative ℓ^2 -error made by the approximation is almost surely $\mathcal{O}(\sqrt{\epsilon \ln |\ln(\epsilon)|})$. Numerics indicate that our error estimate is close to sharp. In [10], McGinnis proved a similar result for a 2D lattice.

Furthermore, formal and numerical studies of random FPUT and other similar random lattice problems report that the waves in such systems experience a notable deterioration of their amplitude as time evolves (see, for instance, [11–14]). We have carried out our own simulations of the nonlinear problem (1.1) with i.i.d. random variables as coefficients in the long-wavelength/small amplitude regime. These simulations demonstrate that for time scales longer than $1/\epsilon$, solutions of (1.1) attenuate substantially; KdV-like dynamics (namely, resolution into solitons) is not observed. We include the results of our simulations below in Section 7, Fig. 1. *In short, we do not believe that when the coefficients are i.i.d. random variables a KdV approximation is appropriate or possible.*

However, there are more sorts of randomness than simply taking the coefficients to be i.i.d. In this paper we consider the random case, but we restrict the randomness in such a way that we can prove a fully rigorous KdV approximation. We believe that this is the first example of such a result involving randomness and nonlinear dispersive systems, though there are several earlier results which carefully derive—but do not fully justify—KdV as an effective equation for the evolution of long wave waves over randomly varying topography [15–18].

Here is our assumption on the masses:

Hypothesis 1.1. The masses are given by

$$m(j) = 1 + \delta^+ \delta^- \zeta(j) \quad (1.3)$$

where $\zeta(j)$, $j \in \mathbb{Z}$, are i.i.d. random variables with zero mean, variance σ^2 and support contained in $(-1/4, 1/4)$.

We refer to (1.3) as a *transparency condition* and we call (1.1) subject to *Hypothesis 1.1* the *transparent random mass FPUT lattice*. The use of “transparent” here is due to an observation from simulations: if the masses meet this condition then waves propagate relatively cleanly through the lattice without too much “back scattering” or “internal reflection”. Our idea for making this choice was inspired by the derivation of KdV as an effective equation for water waves over a random bottom in [15] where the topography is given as a perfect spatial derivative. The interested reader can jump ahead to Remark 4.3 for a counterfactual on what goes wrong in the approximation when the masses are merely i.i.d. The condition on the support of $\zeta(j)$ is there to ensure that the $m(j)$ are strictly positive, for if $|\zeta(j)| < 1/4$ then the triangle inequality tells us $m(j) > 0$. It also guarantees that $\sigma^2 < \infty$.

Our main result in a nutshell: for suitable initial conditions, solutions of the transparent random mass FPUT lattice almost surely satisfy

$$\begin{aligned} q(j, t) &= \epsilon^2 [A(\epsilon(j-t), \epsilon^3 t) + B(\epsilon(j+t), \epsilon^3 t)] + \mathcal{O}_{\ell^2}(\epsilon^2 \sqrt{|\ln(\epsilon)|}) \\ p(j, t) &= \epsilon^2 [-A(\epsilon(j-t), \epsilon^3 t) + B(\epsilon(j+t), \epsilon^3 t)] + \mathcal{O}_{\ell^2}(\epsilon^2 \sqrt{|\ln(\epsilon)|}) \end{aligned} \quad (1.4)$$

for $|t| \leq T_0/\epsilon^3$, where A and B solve the KdV equations

$$\begin{aligned} 2\partial_T A + \left(\frac{1}{12} + 2\sigma^2\right) \partial_w^3 A + \partial_w A^2 &= 0 \quad \text{and} \\ 2\partial_T B - \left(\frac{1}{12} + 2\sigma^2\right) \partial_t^3 B - \partial_t B^2 &= 0. \end{aligned}$$

The fully technical version of our result appears in Theorem 6.1 below.

Remark 1.2. To the uninitiated, it may look like the size of the error exceeds the size of the approximation. However, the long-wave scaling of the spatial coordinate gives $\|\epsilon^2 A(\epsilon(\cdot - t), \epsilon^3 t)\|_{\ell^2} = \mathcal{O}(\epsilon^{3/2})$ which indicates a relative ℓ^2 -error of $\mathcal{O}(\sqrt{\epsilon |\ln(\epsilon)|})$.

Our paper is organized as follows. Section 2 spells out some notation and other ground rules. Section 3 proves a general approximation theorem for (1.1). While motivated by KdV approximations, the result applies more broadly. Section 4 contains the derivation of KdV from (1.1) under Hypothesis 1.1; this is the heart of the paper. Section 5 contains a multitude of estimates which set the stage for the application of the general approximation theorem. It is in this section where probability plays a major role and where the technical guts of our result live. Section 6 ties everything together with the statement and proof of our main result, the technical version of (1.4). Then we present the result of supporting numerics in Section 7. We close out with a big list of open questions in Section 8.

2. Preliminaries

2.1. Function/sequence spaces

For a doubly infinite sequence $f : \mathbb{Z} \rightarrow \mathbb{R}$ we put, as per usual, $\|f\|_{\ell^2} := \sqrt{\sum_{j \in \mathbb{Z}} f^2(j)}$ and $\|f\|_{\ell^\infty} := \sup_{j \in \mathbb{Z}} |f(j)|$. Of course ℓ^2 and ℓ^∞ are the sets of all sequences where the associated norms are finite. If we write $\|f, g\|_{\ell^2}$ we mean $\|f\|_{\ell^2} + \|g\|_{\ell^2}$, the norm on $\ell^2 \times \ell^2$. The analogous convention applies to $\|f, g\|_{\ell^\infty}$. For functions $F : \mathbb{R} \rightarrow \mathbb{R}$ and non-negative integers n and r we put

$$\|F\|_{H^n(r)} := \sqrt{\int_{\mathbb{R}} (1 + X^2)^r F^2(X) dX + \int_{\mathbb{R}} (1 + X^2)^r (\partial_X^n F)^2(X) dX}$$

and $H^n(r)$ is the closure of the set of all smooth functions with respect to this norm. We define $H^n := H^n(0)$, $L^2(r) := H^0(r)$ and $L^2 := H^0(0)$. Next, $\|F\|_{W^{n,\infty}} := \sup_{X \in \mathbb{R}} |F(X)| + |\partial_X^n F(X)|$ and $W^{n,\infty}$ is the associated function space. By L^∞ we mean $W^{0,\infty}$. All of the spaces listed above are Banach spaces.

2.2. Probability

All probabilistic components in the paper descend through the random variables $\zeta(j)$. Associated probabilities are represented by $\mathbb{P}$ and expectations by $\mathbb{E}$.

2.3. $\mathcal{O}$, o And C notation

We use the following version of Landau’s “big $\mathcal{O}$ /little o ” notation. Given two real-valued functions, $f(\epsilon)$ and $g(\epsilon)$, we say $f(\epsilon) = \mathcal{O}(g(\epsilon))$ if for some $C_\star > 0$ and $\epsilon_\star > 0$, $|f(\epsilon)| \leq C_\star |g(\epsilon)|$ for $\epsilon \in (0, \epsilon_\star)$. We say $f(\epsilon) = o(g(\epsilon))$ if $\lim_{\epsilon \rightarrow 0^+} |f(\epsilon)/g(\epsilon)| = 0$. If Y is a Banach space and we have a family of elements $u_\epsilon \in Y$, we write $u_\epsilon = \mathcal{O}_Y(g(\epsilon))$ if $\|u_\epsilon\|_Y = \mathcal{O}(g(\epsilon))$ by the earlier definition. If we write $f(\epsilon) = g(\epsilon) + \mathcal{O}(h(\epsilon))$ (or similar) we mean $f(\epsilon) - g(\epsilon) = \mathcal{O}(h(\epsilon))$. We at times default to “big C ” notation: if we simply write $f(\epsilon) \leq Cg(\epsilon)$ and omit qualifiers then we mean $f(\epsilon) = \mathcal{O}(g(\epsilon))$.

Some quantities will depend on the random variables $\zeta(j)$. For such quantities, if we write $f(\epsilon) = \mathcal{O}(g(\epsilon))$ we mean this in an almost sure sense. Specifically, there exist constants $C_\star > 0$ (almost surely finite) and $\epsilon_\star$ (almost surely positive) such that $|f(\epsilon)| \leq C_\star |g(\epsilon)|$ for $\epsilon \in (0, \epsilon_\star)$. The values of $C_\star$ and $\epsilon_\star$ may depend upon the realization of the $\zeta(j)$. The same “almost sure” point of view holds for $\mathcal{O}_Y$ and o too.

To be clear: we always mean $\mathcal{O}$, o and their ilk rigorously, and we always mean them in the almost sure sense.

3. Approximation in general

We begin by proving a general approximation theorem using the strategy described in Section 5.3 of [7] (itself inspired by [5]). Suppose that for $\epsilon \in (0, 1)$ we have some functions $\tilde{q}_\epsilon(j, t)$ and $\tilde{p}_\epsilon(j, t)$ (the *approximators*) that we expect are good approximations to solutions of

(1.1) when ϵ is small. By this we mean that we know that $\tilde{q}_\epsilon(j, t)$ and $\tilde{p}_\epsilon(j, t)$ nearly solve (1.1) in the sense that the residuals

$$\text{Res}_1 := \delta^+ \tilde{p}_\epsilon - \partial_t \tilde{q}_\epsilon \quad \text{and} \quad \text{Res}_2 := \frac{1}{m} \delta^- [\mathcal{V}'(\tilde{q}_\epsilon)] - \partial_t \tilde{p}_\epsilon \quad (3.1)$$

are small relative to ϵ . To validate the approximation over the time scale $|t| \leq T_0/\epsilon^3$ we need information about

$$\begin{aligned} \alpha_1(\epsilon) &:= \sup_{|t| \leq T_0/\epsilon^3} \|\tilde{q}_\epsilon, \tilde{p}_\epsilon\|_{\ell^2}, \\ \beta_1(\epsilon) &:= \inf_{|t| \leq T_0/\epsilon^3} \|\tilde{q}_\epsilon, \tilde{p}_\epsilon\|_{\ell^2}, \\ \alpha_2(\epsilon) &:= \sup_{|t| \leq T_0/\epsilon^3} \|\partial_t \tilde{q}_\epsilon\|_{\ell^\infty} \quad \text{and} \\ \alpha_3(\epsilon) &:= \sup_{|t| \leq T_0/\epsilon^3} \|\text{Res}_1\|_{\ell^2} + \|\text{Res}_2\|_{\ell^2}. \end{aligned} \quad (3.2)$$

In particular, we require

$$\alpha_1(\epsilon) = o(1), \quad \alpha_2(\epsilon) = \mathcal{O}(\epsilon^3) \quad \text{and} \quad \alpha_3(\epsilon) = o(\beta_1(\epsilon)\epsilon^3). \quad (3.3)$$

We will verify these for the KdV approximation in Section 5 (specifically Proposition 5.3), but in this section these are to be interpreted as assumptions.

Our goal is to show that if we have approximators with these features then the true solution of (1.1) whose initial conditions are consistent with those of the approximators remains close over the long time scale. The result we prove here is specialized to FPUT lattices where the spring potentials are homogeneous and of the form (1.2), but requires only the following non-degeneracy condition on the masses:

$$\inf_{j \in \mathbb{Z}} m(j) > 0 \quad \text{and} \quad \sup_{j \in \mathbb{Z}} m(j) < \infty. \quad (3.4)$$

The condition on the support of $\zeta(j)$ in Hypothesis 1.1 implies the above, though we do not require all of that hypothesis in this section.

Here is the result:

Theorem 3.1. Suppose that the mass coefficients satisfy (3.4), the approximators $\tilde{q}_\epsilon(j, t)$ and $\tilde{p}_\epsilon(j, t)$ meet (3.3) and the initial conditions for (1.1) satisfy

$$\|q(0) - \tilde{q}_\epsilon(0), p(0) - \tilde{p}_\epsilon(0)\|_{\ell^2} = \mathcal{O}\left(\frac{\alpha_3(\epsilon)}{\epsilon^3}\right).$$

Then the solution $(q(t), p(t))$ of (1.1) satisfies the absolute error estimate

$$\sup_{|t| \leq T_0/\epsilon^3} \|q(t) - \tilde{q}_\epsilon(t), p(t) - \tilde{p}_\epsilon(t)\|_{\ell^2} = \mathcal{O}\left(\frac{\alpha_3(\epsilon)}{\epsilon^3}\right)$$

as well as the relative error estimate

$$\sup_{|t| \leq T_0/\epsilon^3} \frac{\|q(t) - \tilde{q}_\epsilon(t), p(t) - \tilde{p}_\epsilon(t)\|_{\ell^2}}{\|q(t), p(t)\|_{\ell^2}} = \mathcal{O}\left(\frac{\alpha_3(\epsilon)}{\beta_1(\epsilon)\epsilon^3}\right).$$

Remark 3.2. Note that this theorem gives error estimates in terms of $\alpha_3(\epsilon)$ and so improvements in this term improve the error. Tighter estimates on $\alpha_1(\epsilon)$ or $\alpha_2(\epsilon)$ do not have that effect.

Proof. We introduce the errors:

$$\eta := q - \tilde{q}_\epsilon \quad \text{and} \quad \xi := p - \tilde{p}_\epsilon$$

where $q(j, t)$ and $p(j, t)$ solve (1.1). Time differentiation of these expressions together with (1.1) and some algebra get us:

$$\dot{\eta} = \delta^+ \xi + \text{Res}_1 \quad \text{and} \quad \dot{\xi} = \frac{1}{m} \delta^- (\mathcal{W}'(\eta, \tilde{q}_\epsilon)) + \text{Res}_2. \quad (3.5)$$

In the above

$$\mathcal{W}(a, b) := \mathcal{V}(b+a) - \mathcal{V}(b) - \mathcal{V}'(b)a = \frac{1}{2}(1+2b)a^2 + \frac{1}{3}a^3$$

so that

$$\mathcal{W}'(a, b) := \partial_a \mathcal{W}(a, b) = \mathcal{V}'(b+a) - \mathcal{V}'(b) = (1+2b)a + a^2.$$

Now we define the energy functional:

$$\mathcal{H}(u, v; t) := \sum_{j \in \mathbb{Z}} \left(\frac{1}{2} m_j v(j)^2 + \mathcal{W}(u(j), \tilde{q}_\epsilon(j, t)) \right).$$

In the above $(u, v) = (u(j), v(j))$ is in $\ell^2 \times \ell^2$. Under our assumptions, the square root of this quantity is equivalent to the $\ell^2 \times \ell^2$ norm in the following sense: there exist $\epsilon_* \in (0, 1)$ and $C_* > 1$ such that

$$0 < \epsilon < \epsilon_*, \quad \|u\|_{\ell^2} \leq 1, \quad |t| \leq T_0/\epsilon^3 \implies C_*^{-1} \|u, v\|_{\ell^2} \leq \sqrt{\mathcal{H}(u, v; t)} \leq C_* \|u, v\|_{\ell^2}. \quad (3.6)$$

Here are the details. First of all, simple estimation gives

$$\frac{1}{2} \inf_{j \in \mathbb{Z}} m(j) \|v\|_{\ell^2}^2 \leq \sum_{j \in \mathbb{Z}} \frac{1}{2} m_j v(j)^2 \leq \frac{1}{2} \sup_{j \in \mathbb{Z}} m(j) \|v\|_{\ell^2}^2$$

and thus (3.4) tells us that $\sqrt{\sum_{j \in \mathbb{Z}} \frac{1}{2} m_j v(j)^2}$ is equivalent to $\|v\|_{\ell^2}$. This gives the “ v ” part of (3.6).

For the “ u ” part, recall that $\|f\|_{\ell^\infty} \leq \|f\|_{\ell^2}$ and so the assumption $\alpha_1(\epsilon) = o(1)$ implies that $\|\tilde{q}_\epsilon\|_{\ell^\infty} = o(1)$ as well. Thus for ϵ sufficiently small we have $\|\tilde{q}_\epsilon\|_{\ell^\infty} \leq 1/30$. This, in conjunction with the assumption that $\|u\|_{\ell^2} \leq 1$, gives us

$$\left| \sum_{j \in \mathbb{Z}} \left(\tilde{q}_\epsilon(j, t) + \frac{1}{3} u(j) \right) u(j)^2 \right| \leq \frac{11}{30} \|u\|_{\ell^2}^2.$$

In turn the triangle inequality gives:

$$\frac{2}{15} \|u\|_{\ell^2}^2 \leq \sum_{j \in \mathbb{Z}} \underbrace{\frac{1}{2} (1 + 2\tilde{q}_\epsilon(j, t)) u(j)^2 + \frac{1}{3} u^3(j)}_{\mathcal{W}(u(j), \tilde{q}_\epsilon(j, t))} \leq \frac{13}{15} \|u\|_{\ell^2}^2.$$

So we have (3.6).

For the next step, we suppose that $\eta(j, t)$ and $\xi(j, t)$ solve (3.5) and put $H(t) := \mathcal{H}(\eta(t), \xi(t); t)$. Differentiation of $H(t)$ with respect to t gives:

$$\begin{aligned} \dot{H}(t) &= \sum_{j \in \mathbb{Z}} \left(m_j \xi(j, t) \dot{\xi}(j, t) + \mathcal{W}'(\eta(j, t), \tilde{q}_\epsilon(j, t)) \dot{\eta}(j, t) \right. \\ &\quad \left. + \partial_b \mathcal{W}(\eta(j, t), \tilde{q}_\epsilon(j, t)) \partial_t \tilde{q}_\epsilon(j, t) \right). \end{aligned}$$

Using (3.5) (and suppressing some dependencies) results in:

$$\begin{aligned} \dot{H} &= \sum_{j \in \mathbb{Z}} \left(\xi(\delta^- (\mathcal{W}'(\eta, \tilde{q}_\epsilon)) + \text{Res}_2) + \mathcal{W}'(\eta, \tilde{q}_\epsilon)(\delta^+ \xi + \text{Res}_1) \right. \\ &\quad \left. + \partial_b \mathcal{W}(\eta, \tilde{q}_\epsilon) \partial_t \tilde{q}_\epsilon \right). \end{aligned}$$

We sum by parts and terms cancel:

$$\dot{H} = \sum_{j \in \mathbb{Z}} \left(\xi \text{Res}_2 + \mathcal{W}'(\eta, \tilde{q}_\epsilon) \text{Res}_1 + \partial_b \mathcal{W}(\eta, \tilde{q}_\epsilon) \partial_t \tilde{q}_\epsilon \right).$$

Subsequently, Cauchy-Schwarz and the like get us:

$$\dot{H} \leq \|\xi\|_{\ell^2} \|\text{Res}_2\|_{\ell^2} + \|\mathcal{W}'(\eta, \tilde{q}_\epsilon)\|_{\ell^2} \|\text{Res}_1\|_{\ell^2} + \|\partial_b \mathcal{W}(\eta, \tilde{q}_\epsilon)\|_{\ell^1} \|\partial_t \tilde{q}_\epsilon\|_{\ell^\infty}.$$

One easily computes that $\partial_b \mathcal{W}(\eta, \tilde{q}_\epsilon) = \eta^2$. In which case we conclude, using the earlier formula for $\mathcal{W}'$ and routine estimates, that

$$\begin{aligned} \dot{H} &\leq \|\xi\|_{\ell^2} \|\text{Res}_2\|_{\ell^2} + \left((1 + 2\|\tilde{q}_\epsilon\|_{\ell^\infty}) \|\eta\|_{\ell^2} + \|\eta\|_{\ell^2}^2 \right) \|\text{Res}_1\|_{\ell^2} \\ &\quad + \|\eta\|_{\ell^2}^2 \|\partial_t \tilde{q}_\epsilon\|_{\ell^\infty}. \end{aligned}$$

Next we use (3.3) to get:

$$\dot{H} \leq 2\alpha_3(\epsilon) (\|\eta\|_{\ell^2} + \|\xi\|_{\ell^2}) + 2\alpha_2(\epsilon) \|\eta\|_{\ell^2}^2.$$

Then we use (3.6):

$$\dot{H} \leq 2C_*^2 (\alpha_3(\epsilon) H^{1/2} + \alpha_2(\epsilon) H).$$

Since $\dot{H} = 2H^{1/2} \frac{d}{dt} H^{1/2}$ the above can be recast as

$$\frac{d}{dt} H^{1/2} \leq C_*^2 (\alpha_3(\epsilon) + \alpha_2(\epsilon) H^{1/2}).$$

We have assumed $\alpha_2(\epsilon) = \mathcal{O}(\epsilon^3)$ so the above implies

$$\frac{d}{dt} H^{1/2} \leq C_*^2 (\alpha_3(\epsilon) + C_2 \epsilon^3 H^{1/2})$$

for a constant $C_2 > 0$.

An application of Grönwall's inequality gets us:

$$H^{1/2}(t) \leq C_2^{-1} \left(e^{C_2^* C_2 \epsilon^3 t} - 1 \right) \frac{\alpha_3(\epsilon)}{\epsilon^3} + e^{C_2^* C_2 \epsilon^3 t} H^{1/2}(0).$$

Using (3.6) again:

$$\|\eta(t), \xi(t)\|_{\ell^2} \leq C_2^* C_2^{-1} \left(e^{C_2^* C_2 \epsilon^3 t} - 1 \right) \frac{\alpha_3(\epsilon)}{\epsilon^3} + C_2^* e^{C_2^* C_2 \epsilon^3 t} \|\eta(0), \xi(0)\|_{\ell^2}.$$

We take the supremum of this over $|t| \leq T_0/\epsilon^3$ and get

$$\sup_{|t| \leq T_0/\epsilon^3} \|\eta(t), \xi(t)\|_{\ell^2} \leq C_* \left(\frac{\alpha_3(\epsilon)}{\epsilon^3} + \|\eta(0), \xi(0)\|_{\ell^2} \right).$$

The constant $C_* > 0$ is independent of ϵ .

In conclusion, if we assume that

$$\|\eta(0), \xi(0)\|_{\ell^2} = \mathcal{O} \left(\frac{\alpha_3(\epsilon)}{\epsilon^3} \right)$$

then we have shown

$$\sup_{|t| \leq T_0/\epsilon^3} \|\eta(t), \xi(t)\|_{\ell^2} = \mathcal{O} \left(\frac{\alpha_3(\epsilon)}{\epsilon^3} \right).$$

This is the absolute error estimate. As for the relative error a standard reverse triangle inequality argument shows that

$$\sup_{|t| \leq T_0/\epsilon^3} \frac{\|\eta(t), \xi(t)\|_{\ell^2}}{\|q(t), p(t)\|_{\ell^2}} = \mathcal{O} \left(\frac{\alpha_3(\epsilon)}{\beta_1(\epsilon)\epsilon^3} \right). \quad \square$$

4. Derivation of the effective equations

Now that we have Theorem 3.1, we can move on to deriving the KdV equations from (1.1). The procedure for the derivation is a multiple scales expansion, inspired by [19]. We assume the following form of our approximators:

$$\tilde{q}_\epsilon(j, t) = \sum_{n=0}^3 \epsilon^{n+2} Q_n(j, \epsilon j, \epsilon t, \epsilon^3 t) \quad \text{and} \quad \tilde{p}_\epsilon(j, t) = \sum_{n=0}^3 \epsilon^{n+2} P_n(j, \epsilon j, \epsilon t, \epsilon^3 t) \quad (4.1)$$

where the $Q_n = Q_n(j, X, \tau, T)$ and $P_n = P_n(j, X, \tau, T)$ are maps

$$\mathbf{Z} \times \mathbf{R} \times \mathbf{R} \times \mathbf{R} \rightarrow \mathbf{R}.$$

Of course we are viewing ϵ as being small. Given that we put $X = \epsilon j$ in $\tilde{q}_\epsilon$ and $\tilde{p}_\epsilon$, we think of X as being the *long-wave length scale* and j being the *microscopic length scale*.

For expansions of the sort we are carrying out, it pays to be organized at the outset. First we define the following operators for functions $U = U(j, X)$:

$$S_j^\pm U(j, X) := U(j \pm 1, X) \quad \text{and} \quad \delta_j^\pm U(j, X) := \pm (U(j \pm 1, X) - U(j, X)).$$

These are *partial shifts* and *partial finite-differences* with respect to j . Next, for $\epsilon > 0$ put

$$D^\pm U(j, X) := \pm (U(j \pm 1, X \pm \epsilon) - U(j, X)).$$

If $u(j) = U(j, \epsilon j)$ then $\delta^\pm u(j) = D^\pm U(j, \epsilon j)$. That is to say, $D^\pm$ are the *total finite-difference operators*.

Expanding the right-hand sides of $D^\pm U(j, X)$ in (formal) Taylor series with respect to ϵ gives $D^\pm U(j, X) = \delta_j^\pm U(j, X) + \sum_{n=1}^{\infty} \epsilon^n \frac{(\pm 1)^{n+1}}{n!} S_j^\pm \partial_X^n U(j, X)$. Truncating the sum at $n = M$ would give a formal error on the order of ϵ^{M+1} and so we define

$$\epsilon^{M+1} E_M^\pm := D^\pm - \delta_j^\pm - \sum_{n=1}^M \epsilon^n \frac{(\pm 1)^{n+1}}{n!} S_j^\pm \partial_X^n.$$

These operators give the exact error made by such a truncation. Note that for $M = 0$ we just ignore the sum, i.e. $\epsilon E_0^\pm := D^\pm - \delta_j^\pm$.

If we plug (4.1) into the residuals (3.1) and carry out some substantial algebra we find that

$$\text{Res}_1 = \epsilon^2 Z_{12} + \epsilon^3 Z_{13} + \epsilon^4 Z_{14} + \epsilon^5 Z_{15} + \epsilon^6 W_1 \quad \text{and}$$

$$\text{Res}_2 = \frac{1}{m} (\epsilon^2 Z_{22} + \epsilon^3 Z_{23} + \epsilon^4 Z_{24} + \epsilon^5 Z_{25} + \epsilon^6 W_2)$$

where

$$Z_{12} := \delta_j^+ P_0$$

$$Z_{22} := \delta_j^- Q_0$$

$$Z_{13} := \delta_j^+ P_1 + S_j^+ \partial_X P_0 - \partial_\tau Q_0$$

$$Z_{23} := \delta_j^- Q_1 + S_j^- \partial_X Q_0 - m \partial_\tau P_0$$

$$Z_{14} := \delta_j^+ P_2 + S_j^+ \partial_X P_1 + \frac{1}{2} S_j^+ \partial_X^2 P_0 - \partial_\tau Q_1$$

$$Z_{24} := \delta_j^- Q_2 + S_j^- \partial_X Q_1 - \frac{1}{2} S_j^- \partial_X^2 Q_0 + \delta_j^- Q_0^2 - m \partial_\tau P_1$$

$$Z_{15} := \delta_j^+ P_3 + S_j^+ \partial_X P_2 + \frac{1}{2} S_j^+ \partial_X^2 P_1 + \frac{1}{6} S_j^+ \partial_X^3 P_0 - \partial_\tau Q_2 - \partial_\tau Q_0$$

$$Z_{25} := \delta_j^- Q_3 + S_j^- \partial_X Q_2 - \frac{1}{2} S_j^- \partial_X^2 Q_1 + \frac{1}{6} S_j^- \partial_X^3 Q_0 + 2 \delta_j^- (Q_0 Q_1) + \partial_X Q_0^2 - m \partial_\tau P_2 - m \partial_\tau P_0$$

and

$$W_1 := \sum_{n=0}^3 E_{3-n}^+ P_n - \partial_\tau Q_3 - \sum_{n=1}^3 \epsilon^{n-1} \partial_\tau Q_n$$

$$W_2 := \sum_{n=0}^3 E_{3-n}^- Q_n + E_1^- Q_0^2 + 2 E_0^- (Q_0 Q_1)$$

$$+ D^- \left(2 Q_0 Q_2 + (Q_1 + \epsilon Q_2 + \epsilon^2 Q_3)^2 \right) - m \partial_\tau P_3 - m \sum_{n=1}^3 \epsilon^{n-1} \partial_\tau P_n. \quad (4.2)$$

The usual way to proceed is to select the $Q_0, P_0, \dots, Q_3, P_3$ so that each $Z_{1k} = Z_{2k} = 0$. In this case we would have $\text{Res}_1 = \epsilon^6 W_1$ and $\text{Res}_2 = \frac{1}{m} \epsilon^6 W_2$ which we can then estimate using the formulas for the Q_n and P_n . This strategy works perfectly well in the homogeneous and periodic problems as all the terms are rigorously the size they formally appear to be, modulo an annoying factor of $\epsilon^{-1/2}$ caused by the long-wave scaling. But it fails in the random problem; the randomness leads to terms which are much larger than they appear.

Our modified strategy is to solve

$$Z_{12} = Z_{22} = Z_{13} = Z_{23} = Z_{14} = Z_{24} = 0$$

(which will largely determine $Q_0, P_0, \dots, Q_2, P_2$) and then to do “something else” for Z_{15} and Z_{25} . At the end of this, we find that $\text{Res}_1 = \epsilon^5 Z_{15} + \epsilon^6 W_1$ and $\text{Res}_2 = \frac{1}{m} (\epsilon^5 Z_{25} + \epsilon^6 W_2)$. In Section 5 we show that these are $\mathcal{O}_{\ell^2}(\epsilon^5 \sqrt{|\ln(\epsilon)|})$. This is enough to apply Theorem 3.1 and get the error estimates shown in (1.4).

4.1. A tutorial on solving $Z_{1k} = Z_{2k} = 0$

As we proceed, we will discover that each pair of equations $Z_{1k} = Z_{2k} = 0$ can be written in the form

$$\begin{aligned} \delta_j^+ P_{k-2} &= \bar{F}_0(X, \tau, T) + \sum_{n=1}^N f_n(j) \bar{F}_n(X, \tau, T) \\ \delta_j^- Q_{k-2} &= \bar{G}_0(X, \tau, T) + \sum_{n=1}^N g_n(j) \bar{G}_n(X, \tau, T). \end{aligned} \quad (4.3)$$

The sequences $f_n(j)$ and $g_n(j)$ are mean-zero random variables which come, in one way or another, from $\zeta(j)$; they depend only on the microscale coordinate. The $\bar{F}_n$ and $\bar{G}_n$ functions do not depend on the microscale coordinate at all. They will be made up of pieces of the

various P_i and Q_i where $i < k-2$. In this way (4.3) allows us to figure out P_{k-2} and Q_{k-2} from the earlier functions.

We decompose (4.3) into a “long-wave” part (those pieces that do not depend on the microscale coordinate j at all) and a “microscale” part (those that do). The long-wave part just consists of the terms $\bar{F}_0$ and $\bar{G}_0$ and so we set

$$\bar{F}_0 = 0 \quad \text{and} \quad \bar{G}_0 = 0. \quad (4.4)$$

This is a sort of solvability condition that will wind up giving us the long-wave dynamics; how it all plays out will be seen when we get in the weeds below.

The microscale part is what is left over:

$$\delta_j^+ P_{k-2} = \sum_{n=1}^N f_n(j) \bar{F}_n(X, \tau, T) \quad \text{and} \quad \delta_j^- Q_{k-2} = \sum_{n=1}^N g_n(j) \bar{G}_n(X, \tau, T). \quad (4.5)$$

We can just write down a solution for this:

$$P_{k-2}(j, X, \tau, T) = \bar{P}_{k-2}(X, \tau, T) + \sum_{n=1}^N \chi_n(j) \bar{F}_n(X, \tau, T) \quad \text{and} \quad (4.6)$$

$$Q_{k-2}(j, X, \tau, T) = \bar{Q}_{k-2}(X, \tau, T) + \sum_{n=1}^N \kappa_n(j) \bar{G}_n(X, \tau, T)$$

where we select χ_n and κ_n so that

$$\delta^+ \chi_n = f_n \quad \text{and} \quad \delta^- \kappa_n = g_n.$$

Solving these equations for χ_n and κ_n from f_n and g_n is one of the key steps in the whole procedure and as we shall show the transparency condition makes this a relatively easy affair...at least at first. The functions $\bar{P}_{k-2}(X, \tau, T)$ and $\bar{Q}_{k-2}(X, \tau, T)$ are “constants of integration”; in most cases we determine these from (4.4) at a later point in the derivation.

Now we get into actually solving the equations.

$$4.2. \quad Z_{12} = Z_{22} = 0$$

These read $\delta_j^+ P_0 = 0$ and $\delta_j^- Q_0 = 0$ which tells us that

$$Q_0(j, X, \tau, T) = \bar{Q}_0(X, \tau, T) \quad \text{and} \quad P_0(j, X, \tau, T) = \bar{P}_0(X, \tau, T). \quad (4.7)$$

Remark 4.1. In this section any function with a “bar” on top will not depend on j . We make this convention so that we do not need to perpetually clutter up our algebra with functional dependencies. For the same reason it is helpful to keep in mind that m and ζ depend only on j and not on the other variables.

$$4.3. \quad Z_{13} = Z_{23} = 0$$

Using (4.7) these equations become

$$\delta_j^+ P_1 = \partial_\tau \bar{Q}_0 - \partial_X \bar{P}_0 \quad \text{and} \quad \delta_j^- Q_1 = m \partial_\tau \bar{P}_0 - \partial_X \bar{Q}_0.$$

Using the transparency condition (1.3) converts these to

$$\delta_j^+ P_1 = \partial_\tau \bar{Q}_0 - \partial_X \bar{P}_0 \quad \text{and} \quad \delta_j^- Q_1 = \partial_\tau \bar{P}_0 - \partial_X \bar{Q}_0 + \delta^+ \delta^- \zeta \partial_\tau \bar{P}_0.$$

Following the steps from the tutorial in Section 4.1 we see that the long-wave part (4.4) of these equations is

$$\partial_\tau \bar{Q}_0 - \partial_X \bar{P}_0 = 0 \quad \text{and} \quad \partial_\tau \bar{P}_0 - \partial_X \bar{Q}_0 = 0. \quad (4.8)$$

This is the wave equation wearing a fake mustache and glasses and we readily solve it:

$$\bar{Q}_0 = A(X - \tau, T) + B(X + \tau, T) \quad \text{and} \quad \bar{P}_0 = -A(X - \tau, T) + B(X + \tau, T). \quad (4.9)$$

Remark 4.2. We use the convention that

$$w = X - \tau \quad \text{and} \quad l = X + \tau$$

so that $A = A(w, T)$ and $B = B(l, T)$. Note that A and B do not depend on j . It is these functions that will ultimately solve KdV equations.

After (4.8) we are left with the microscale part

$$\delta_j^+ P_1 = 0 \quad \text{and} \quad \delta_j^- Q_1 = \delta^+ \delta^- \zeta \partial_\tau \bar{P}_0.$$

The solution formula (4.6) gives $Q_1 = \bar{Q}_1 + \chi \partial_\tau \bar{P}_0$ where we want $\delta^- \chi = \delta^+ \delta^- \zeta$. Finding χ is easily done as we can simply cancel a δ^- from both sides and put $\chi = \delta^+ \zeta$. This is so simple because of the transparency condition (1.3) and this is one of the reasons we have assumed it. Likewise (4.6) says that we should put $P_1 = \bar{P}_1$, but it will turn out that $\bar{P}_1$ will be zero so we just enforce that now. In short we have

$$Q_1 = \bar{Q}_1 + \delta_j^+ \zeta \partial_\tau \bar{P}_0 \quad \text{and} \quad P_1 = 0. \quad (4.10)$$

Note that $\delta_j^+ \zeta$ is bounded in j because of the compact support assumption in Hypothesis 1.1.

Remark 4.3. What if we had not made the transparency assumption but instead assumed that $m(j) = 1 + z(j)$ where $z(j)$ are i.i.d. mean zero random variables? The long-wave part is the same as above but now the microscale part is $\delta_j^- Q_1 = z \partial_\tau \bar{P}_0$. To use the solution formula (4.6) we would want to find χ so that $\delta^- \chi = z$, or rather

$$\chi(j) = \chi(j-1) + z(j).$$

This equation tells us that $\chi(j)$ is a random walk with steps given by $z(j)$ and as such we expect χ to grow like $\sqrt{j}$.

To see why this is an issue, notice that Q_1 would include the term $\chi(j) A_w(X - \tau, T)$, which then would show up in the approximator (4.1) as $\epsilon^3 \chi(j) A_w(\epsilon(j - t), \epsilon^3 t)$. The term A_w is propagating to the right with roughly unit speed and thus when $t \sim 1/\epsilon^3$ will be located at $j \sim 1/\epsilon^3$. In turn this indicates $\chi(j) \sim \epsilon^{-3/2}$ towards the end of the approximation time interval. Hence the term $\epsilon^3 \chi A_w$ would be substantially larger than it appears: the techniques from [2] show that almost surely

$$\sup_{|t| \leq T_0/\epsilon^3} \|\epsilon^3 \chi(\cdot) A_w(\epsilon(\cdot - t), \epsilon^3 t)\|_{\ell^2} \leq C \epsilon \sqrt{\ln |\ln(\epsilon)|}.$$

Were $\chi = \mathcal{O}_{\ell^\infty}(1)$ the right-hand side of the preceding estimate would be $C \epsilon^{5/2}$ (see Lemma 5.6 below). And so we find that the “ ϵ^3 term” in the approximator is more than an order of magnitude larger than it should be, bigger in fact than the leading order term in the approximation. Disaster!

The lesson learned: if a term in our approximation involves a random walk it will ultimately be at least $\epsilon^{-3/2}$ larger than it formally appears to be. We call this difficulty a **random walk disaster**.

$$4.4. \quad Z_{14} = Z_{24} = 0$$

The relations (1.3), (4.7), (4.8), (4.10) and a little algebra convert these equations to

$$\delta_j^+ P_2 = \partial_\tau \bar{Q}_1 - \frac{1}{2} \partial_{X\tau}^2 \bar{Q}_0 + \delta_j^+ \zeta \partial_X^2 \bar{P}_0 \quad \text{and}$$

$$\delta_j^- Q_2 = -\partial_X \bar{Q}_1 + \frac{1}{2} \partial_X^2 \bar{Q}_0 - \delta_j^- \zeta \partial_X^2 \bar{Q}_0.$$

The long-wave part of this is

$$0 = \partial_\tau \bar{Q}_1 - \frac{1}{2} \partial_{X\tau}^2 \bar{Q}_0 \quad \text{and} \quad 0 = -\partial_X \bar{Q}_1 + \frac{1}{2} \partial_X^2 \bar{Q}_0$$

which can be solved by putting

$$\bar{Q}_1 = \frac{1}{2} \partial_X \bar{Q}_0 = \frac{1}{2} (\partial_w A + \partial_l B). \quad (4.11)$$

This leaves the microscale part

$$\delta_j^+ P_2 = \delta_j^+ \zeta \partial_X^2 \bar{P}_0 \quad \text{and} \quad \delta_j^- Q_2 = -\delta_j^- \zeta \partial_X^2 \bar{Q}_0. \quad (4.12)$$

Which as per (4.6) we solve by taking

$$P_2 = \bar{P}_2 + \zeta \partial_X^2 \bar{P}_0 \quad \text{and} \quad Q_2 = \bar{Q}_2 - \zeta \partial_X^2 \bar{Q}_0. \quad (4.13)$$

Once again, the transparency condition (1.3) made finding this solution a simple matter of cancellation; it is the reason why the transparency condition has two finite-differences on ζ . If we had put only one finite-difference in (1.3) then another random walk disaster as described in Remark 4.3 would occur when we solve (4.12).

4.5. Something else for Z_{15} and Z_{25}

The relations (1.3), (4.7), (4.8), (4.10), (4.11), (4.13) and quite a bit of algebra get us:

$$\begin{aligned} Z_{15} &= -\partial_T \bar{Q}_0 - \partial_\tau \bar{Q}_2 + \partial_X \bar{P}_2 + \frac{1}{6} \partial_X^3 \bar{P}_0 \\ &\quad + \delta_j^+ P_3 + \left(\zeta + S_j^+ \zeta \right) \partial_X^3 \bar{P}_0 \\ Z_{25} &= -\partial_T \bar{P}_0 - \partial_\tau \bar{P}_2 + \partial_X \bar{Q}_2 - \left(\frac{1}{12} - 2\sigma^2 \right) \partial_X^3 \bar{Q}_0 + \partial_X (\bar{Q}_0^2) \\ &\quad + \delta_j^- Q_3 - \left(\zeta + S_j^- \zeta + \frac{1}{2} \delta_j^- \zeta + \zeta \delta_j^+ \delta_j^- \zeta + 2\sigma^2 \right) \partial_X^3 \bar{Q}_0 \\ &\quad + \delta_j^- \delta_j^+ \zeta \partial_X (\bar{Q}_0^2) - \delta_j^+ \delta_j^- \zeta (\partial_T \bar{P}_0 + \partial_\tau \bar{P}_2). \end{aligned}$$

Recall that σ^2 is the variance of $\zeta(j)$. We need Z_{15} and Z_{25} to be small relative to ϵ , but zeroing them out completely happens to be too restrictive; we will need to modify the microscale part of the decomposition described in Section 4.1. But before we get there we deal with the long-wave part.

4.5.1. Kill the long-wave part with KdV equations

As per normal, we zero out the long-wave parts of Z_{15} and Z_{25} . We have conveniently arranged all such terms in the first line on the right of the preceding formulas for Z_{15} and Z_{25} and so we put

$$\begin{aligned} 0 &= -\partial_T \bar{Q}_0 - \partial_\tau \bar{Q}_2 + \partial_X \bar{P}_2 + \frac{1}{6} \partial_X^3 \bar{P}_0 \\ 0 &= -\partial_T \bar{P}_0 - \partial_\tau \bar{P}_2 + \partial_X \bar{Q}_2 - \left(\frac{1}{12} - 2\sigma^2 \right) \partial_X^3 \bar{Q}_0 + \partial_X (\bar{Q}_0^2). \end{aligned} \quad (4.14)$$

Within (4.14) lurk the KdV equations; here is how we coax them into the daylight. Let

$$\bar{Q}_2(X, \tau, T) = A_2(X - \tau, X + \tau, T) + B_2(X - \tau, X + \tau, T)$$

$$\text{and} \quad \bar{P}_2(X, \tau, T) = -A_2(X - \tau, X + \tau, T) + B_2(X - \tau, X + \tau, T)$$

and use (4.9) in (4.14) to get

$$\begin{aligned} 0 &= \partial_T A + \partial_T B + 2\partial_l A_2 - 2\partial_w B_2 - \frac{1}{6} (-\partial_w^3 A + \partial_l^3 B) \\ 0 &= -\partial_T A + \partial_T B - 2\partial_l A_2 - 2\partial_w B_2 + \left(\frac{1}{12} - 2\sigma^2 \right) (\partial_w^3 A + \partial_l^3 B) \\ &\quad - (\partial_w A^2 + 2\partial_w AB + 2A\partial_l B + \partial_l B^2). \end{aligned}$$

Subtracting these gives:

$$\begin{aligned} 0 &= 2\partial_T A + \left(\frac{1}{12} + 2\sigma^2 \right) \partial_w^3 A + \partial_w A^2 \\ &\quad + 4\partial_l A_2 - \left(\frac{1}{4} - 2\sigma^2 \right) \partial_l^3 B + (2\partial_w AB + 2A\partial_l B + \partial_l B^2). \end{aligned} \quad (4.15)$$

If we let B be an l -antiderivative of B (specifically $B(l, T) := \int_0^l B(y, T) dy$) and set

$$A_2 = \frac{1}{4} \left[\left(\frac{1}{4} - 2\sigma^2 \right) \partial_l^2 B - (2\partial_w AB + 2AB + B^2) \right] \quad (4.16)$$

many terms in (4.15) die. What survives is

$$0 = 2\partial_T A + \left(\frac{1}{12} + 2\sigma^2 \right) \partial_w^3 A + \partial_w A^2. \quad (4.17)$$

This is a KdV equation! A parallel argument (after adding instead subtracting equations a few steps above) shows we should take

$$B_2 = \frac{1}{4} \left[\left(\frac{1}{4} - 2\sigma^2 \right) \partial_w^2 A - (A^2 + 2AB + 2A\partial_l B) \right] \quad (4.18)$$

with A a w -antiderivative of A (specifically $A(w, T) := \int_0^w A(y, T) dy$). In which case we get that B solves another KdV equation:

$$0 = 2\partial_T B - \left(\frac{1}{12} + 2\sigma^2 \right) \partial_l^3 B - \partial_l B^2. \quad (4.19)$$

To summarize: taking A , B , A_2 and B_2 as we have just described means that (4.14) is satisfied.

4.5.2. Handle with the microscopic part using autoregressive processes

The next step in dealing with Z_{15} and Z_{25} is to control the microscopic parts that are left over after (4.14):

$$\begin{aligned} Z_{15} &= \delta_j^+ P_3 + \left(\zeta + S_j^+ \zeta \right) \partial_X^3 \bar{P}_0 \\ Z_{25} &= \delta_j^- Q_3 - \left(\zeta + S_j^- \zeta + \frac{1}{2} \delta_j^- \zeta + \zeta \delta_j^+ \delta_j^- \zeta + 2\sigma^2 \right) \partial_X^3 \bar{Q}_0 \\ &\quad + \delta_j^- \delta_j^+ \zeta \partial_X (\bar{Q}_0^2) - \delta_j^+ \delta_j^- \zeta (\partial_T \bar{P}_0 + \partial_\tau \bar{P}_2). \end{aligned} \quad (4.20)$$

Many, but not all, of these terms in Z_{25} can be eliminated with the same cancellation tricks that worked earlier. To see this, we let

$$\begin{aligned} P_3 &= \gamma_1 \partial_X^3 \bar{P}_0 \quad \text{and} \\ Q_3 &= \left(\gamma_2 + \frac{1}{2} \zeta \right) \partial_X^3 \bar{Q}_0 - \delta_j^+ \zeta \partial_X (\bar{Q}_0^2) + \delta_j^+ \zeta (\partial_T \bar{P}_0 + \partial_\tau \bar{P}_2) \end{aligned}$$

where for the moment we leave $\gamma_1 = \gamma_1(j)$ and $\gamma_2 = \gamma_2(j)$ unspecified. Substituting the above into (4.20) gives

$$\begin{aligned} Z_{15} &= \left[\delta_j^+ \gamma_1 + \zeta + S_j^+ \zeta \right] \partial_X^3 \bar{P}_0 \\ Z_{25} &= \left[\delta_j^- \gamma_2 - \left(\zeta + S_j^- \zeta + \zeta \delta_j^+ \delta_j^- \zeta + 2\sigma^2 \right) \right] \partial_X^3 \bar{Q}_0. \end{aligned} \quad (4.21)$$

If we followed the strategy from the tutorial in Section 4.1, we would put $\delta_j^+ \gamma_1 = -\zeta - S_j^+ \zeta$ and $\delta_j^- \gamma_2 = \zeta + S_j^- \zeta + \zeta \delta_j^+ \delta_j^- \zeta + 2\sigma^2$ and get $Z_{15} = Z_{25} = 0$. Since the $\zeta(j)$ are i.i.d. random variables we would then find that $\gamma_1(j)$ and $\gamma_2(j)$ are random walks, which leads us to another disaster as described in Remark 4.3 (this time in the residual terms). Why not just stack another finite-difference on ζ in the transparency condition? This would help in Z_{15} but would not be useful in handling the parts stemming from $\zeta \delta_j^+ \delta_j^- \zeta$ in Z_{25} .

To avoid these problematic random walks we take γ_1 and γ_2 to solve

$$\begin{aligned} \delta_j^+ \gamma_1 &= -\epsilon \operatorname{sgn}(j) \gamma_1 - (\zeta + S_j^+ \zeta) \quad \text{and} \\ \delta_j^- \gamma_2 &= -\epsilon \operatorname{sgn}(j) \gamma_2 + (\zeta + S_j^- \zeta + \zeta \delta_j^+ \delta_j^- \zeta + 2\sigma^2). \end{aligned} \quad (4.22)$$

In which case we find that (4.21) becomes

$$Z_{15} = -\epsilon \operatorname{sgn}(j) \gamma_1 \partial_X^3 \bar{P}_0 \quad \text{and} \quad Z_{25} = -\epsilon \operatorname{sgn}(j) \gamma_2 \partial_X^3 \bar{Q}_0.$$

The extra factors of ϵ on the right-hand sides here means that our choices of P_3 and Q_3 are formally as good as putting $Z_{15} = Z_{25} = 0$. Estimates for P_3 and Q_3 (and consequently Z_{15} , Z_{25} and the residuals) ultimately require us to understand γ_1 and γ_2 . The equations in (4.22) are examples of *autoregressive processes* [20]. These are dissipative cousins of random walks and with classical probabilistic methods we will show that they roughly cost us a factor of $\epsilon^{-1/2}$ (see Lemma 5.10 below) instead of the $\epsilon^{-3/2}$ we get from using random walks. This is big but not too big for our estimates to handle.

4.6. Summing up

At this point we have completely determined all the functions $P_0, \dots, Q_3$ in the approximation. As it can be challenging to sort through it all, we close out this section by summarizing the derivation.

Definition 4.4. Suppose $A(w, T)$ and $B(l, T)$ solve the KdV equations (4.17) and (4.19) and $\gamma_1(j)$ and $\gamma_2(j)$ solve the autoregressive processes (4.22). Take $A_2(w, l, T)$ and $B_2(w, l, T)$ as in (4.16) and (4.18). Define

$Q_k(j, X, \tau, T)$ and $P_k(j, X, \tau, T)$ via

$$\left\{ \begin{array}{l} Q_0 = A + B, \\ Q_1 = \frac{1}{2} \partial_X Q_0 + \delta_j^+ \zeta \partial_\tau P_0, \\ Q_2 = A_2 + B_2 - \zeta \partial_X^2 Q_0 \\ Q_3 = \left(\gamma_2 + \frac{1}{2} \zeta \right) \partial_X^3 Q_0 - \delta_j^+ \zeta \partial_X (Q_0^2) \\ \quad + \delta_j^+ \zeta (\partial_T P_0 - \partial_\tau A_2 + \partial_\tau B_2) \end{array} \right\} \left\{ \begin{array}{l} P_0 = -A + B \\ P_1 = 0 \\ P_2 = -A_2 + B_2 + \zeta \partial_X^2 P_0 \\ P_3 = \gamma_1 \partial_X^3 P_0 \end{array} \right.$$

where it is understood that $w = X - \tau$ and $l = X + \tau$. Then we call $\tilde{q}_\epsilon(j, t)$ and $\tilde{p}_\epsilon(j, t)$, as defined in (4.1), the **extended KdV approximators**.

In this section we have proven:

Lemma 4.5. *The extended KdV approximators have*

$$\text{Res}_1 = \epsilon^6 (-\text{sgn}(j) \gamma_1 \partial_X^3 P_0 + W_1) \quad \text{and} \quad \text{Res}_2 = \frac{\epsilon^6}{m} (-\text{sgn}(j) \gamma_2 \partial_X^3 Q_0 + W_2)$$

with W_1 and W_2 given at (4.2).

We move on to proving many estimates related to the extended KdV approximators.

5. Estimates on the approximators and residuals

To streamline some of the forthcoming statements we put forth the following convention:

Definition 5.1. We say A and B are **good solutions of KdV** on $[-T_0, T_0]$ if they satisfy (4.17) and (4.19) along with the estimate

$$0 < \sup_{|T| \leq T_0} \|A(\cdot, T)\|_{H^7(1)} + \|B(\cdot, T)\|_{H^7(1)} < \infty.$$

Remark 5.2. The existence of good solutions of KdV on intervals of arbitrary length is by now classical (see [21]). The lower bound is just to guarantee that the approximation is not trivial.

In this section we prove:

Proposition 5.3. *Assume Hypothesis 1.1. Let $\tilde{q}_\epsilon(j, t)$ and $\tilde{p}_\epsilon(j, t)$ be the extended KdV approximators as in Definition 4.4 where we further assume that A and B are good solutions of KdV on $[-T_0, T_0]$. Then almost surely the quantities defined at (3.2) satisfy*

$$\alpha_1(\epsilon) = \mathcal{O}(\epsilon^{3/2}), \quad \alpha_2(\epsilon) = \mathcal{O}(\epsilon^3), \quad \alpha_3(\epsilon) = \mathcal{O}(\epsilon^5 \sqrt{|\ln(\epsilon)|}) \quad \text{and} \\ \beta_1^{-1}(\epsilon) = \mathcal{O}(\epsilon^{-3/2}).$$

Estimates on terms which do not involve γ_1 or γ_2 can be handled using well-understood techniques found in previous works, whereas the rest require new ideas. All dependence on γ_1 and γ_2 enters through P_3 and Q_3 , the latter of which has some terms without them. And so we put

$$Q_{3\gamma} := \gamma_2 \partial_X^3 Q_0 \quad \text{and} \quad Q_{30} := Q_3 - \gamma_2 \partial_X^3 Q_0. \quad (5.1)$$

To be clear, Q_{30} has no instances of a γ within.

Similarly if in the formulas for W_1 and W_2 we eliminate any term with a γ in it we get:

$$W_{10} := \sum_{n=0}^2 E_{3-n}^+ P_n - \sum_{n=1}^2 \epsilon^{n-1} \partial_T Q_n - \epsilon^2 \partial_T Q_{30} \\ W_{20} := \sum_{n=0}^2 E_{3-n}^- Q_n + E_0^- Q_{30} + E_1^- Q_0^2 + 2E_0^- (Q_0 Q_1) \\ + D^- \left(2Q_0 Q_2 + (Q_1 + \epsilon Q_2 + \epsilon^2 Q_{30})^2 \right) - m \sum_{n=1}^2 \epsilon^{n-1} \partial_T P_n.$$

Thus the terms with a γ are:

$$W_{1\gamma} := W_1 - W_{10} \quad \text{and} \quad W_{2\gamma} := W_2 - W_{20}. \quad (5.2)$$

5.1. Terms without γ_1 and γ_2

In this part we prove:

Lemma 5.4. *Assume Hypothesis 1.1. Let $\tilde{q}_\epsilon(j, t)$ and $\tilde{p}_\epsilon(j, t)$ be the extended KdV approximators as in Definition 4.4 where we further assume that A and B are good solutions of KdV on $[-T_0, T_0]$. Then*

$$\sum_{n=0}^2 \sup_{|t| \leq T_0/\epsilon^3} (\|P_n(\cdot, \epsilon, \epsilon t, \epsilon^3 t)\|_{\ell^2} + \|Q_n(\cdot, \epsilon, \epsilon t, \epsilon^3 t)\|_{\ell^2}) = \mathcal{O}(\epsilon^{-1/2}), \\ \sup_{|t| \leq T_0/\epsilon^3} \|Q_{30}(\cdot, \epsilon, \epsilon t, \epsilon^3 t)\|_{\ell^2} = \mathcal{O}(\epsilon^{-1/2}),$$

and

$$\sup_{|t| \leq T_0/\epsilon^3} (\|W_{10}(\cdot, \epsilon, \epsilon t, \epsilon^3 t)\|_{\ell^2} + \|W_{20}(\cdot, \epsilon, \epsilon t, \epsilon^3 t)\|_{\ell^2}) = \mathcal{O}(\epsilon^{-1/2}).$$

Remark 5.5. Note that in Hypothesis 1.1 we assumed that $|\zeta(j)| < 1/4$ for all j . A consequence of this is that none of the estimates in Lemma 5.4 depend on the realization of the $\zeta(j)$. That is to say there is no probability needed to understand this lemma.

Proof. The proof is similar to that of Proposition 4.2 of [7], though there are a few small, but substantive, differences. The main tool we need is:

Lemma 5.6. *Let $M \geq 0$ be an integer. Suppose that $f(j) \in \ell^\infty$ and $F(X) \in H^{M+1}$. If $u_\epsilon(j) := f(j)F(\epsilon j)$ then*

$$\|u\|_{\ell^\infty} \leq \|f\|_{\ell^\infty} \|F\|_{L^\infty},$$

$$\|u\|_{\ell^2} \leq C \epsilon^{-1/2} \|f\|_{\ell^\infty} \|F\|_{H^1},$$

$$\|E_M^\pm u\|_{\ell^2} \leq C \epsilon^{-1/2} \|f\|_{\ell^\infty} \|F\|_{H^{M+1}}$$

and

$$\|D^\pm u\|_{\ell^2} \leq C \epsilon^{-1/2} \|f\|_{\ell^\infty} \|F\|_{H^1}.$$

The constants $C > 0$ depend only on M .

Proof. Lemma 4.3 of [7] is nearly identical to this, but has the requirement that $f(j)$ be N -periodic. Still we can piggyback the proof of our result on that one. The first estimate is all but obvious. For the second we have the easy estimate $\|u\|_{\ell^2} \leq \|f\|_{\ell^\infty} \|F_\epsilon\|_{\ell^2}$ where $F_\epsilon(j) := F(\epsilon j)$, $j \in \mathbb{Z}$. But then the second estimate of Lemma 4.3 of [7] applies and shows $\|F_\epsilon\|_{\ell^2} \leq C \epsilon^{-1/2} \|F\|_{H^1}$. For the third, a direct computation shows that $E_M^+ u(j) = f(j+1)(E_M^+ F_\epsilon)(j)$ which implies $\|E_M^+ u\|_{\ell^2} \leq \|f\|_{\ell^\infty} \|E_M^+ F_\epsilon\|_{\ell^2}$. The third estimate from Lemma 4.3 of [7] implies that $\|E_M^+ F_\epsilon\|_{\ell^2} \leq C \epsilon^{-1/2} \|F\|_{H^{M+1}}$. To get an estimate for E_M^- is similar. The final estimate, for $D^\pm u$, follows from the definition of $D^\pm$, the triangle inequality and the second estimate in this lemma. $\square$

We also need the following, to control the antiderivatives in A_2 and B_2 :

Lemma 5.7. Suppose that $F(X) \in L^2(1)$ then $F(X) := \int_0^X F(y)dy$ is in L^∞ and $\|F\|_{L^\infty} \leq \sqrt{\pi}\|F\|_{L^2(1)}$.

Proof. We use Cauchy–Schwarz and the fact that $\int_{\mathbb{R}} (1+y^2)^{-1} = \pi$. To wit:

$$\begin{aligned} |F(X)| &\leq \int_0^X (1+y^2)^{-1/2} (1+y^2)^{1/2} |F(y)| dy \\ &\leq \sqrt{\int_{\mathbb{R}} (1+y^2)^{-1} dy} \sqrt{\int_{\mathbb{R}} (1+y^2) |F(y)|^2 dy} = \sqrt{\pi} \|F\|_{L^2(1)}. \end{aligned}$$

Taking the supremum over X seals the deal. $\square$

Armed with Lemmas 5.6 and 5.7 we can get into proving the estimates in Lemma 5.4. There are many terms and handling each would inflate this paper like a bounce house. So we do not do that. Instead we show how to estimate a “prototype” term which captures the nuances. That term is

$$g = E_0^- (\delta_j^+ \zeta \mathcal{A} \partial_l^2 B)$$

which some digging will show appears in Res_2 . Using the estimate for E_0^- from Lemma 5.6 we have

$$\|g\|_{\ell^2} \leq C e^{-1/2} \|\delta_j^+ \zeta\|_{\ell^\infty} \|\mathcal{A} \partial_l^2 B\|_{H^1}.$$

By the triangle inequality and the definition of δ_j^+ we have $\|\delta_j^+ \zeta\|_{\ell^\infty} \leq 2\|\zeta\|_{\ell^\infty}$ and the supposition that the support of $\zeta(j)$ is in $(-1/4, 1/4)$ ultimately gives $\|\delta_j^+ \zeta\|_{\ell^\infty} \leq 1/2$. Also classical Sobolev–Hölder inequalities tell us that $\|\mathcal{A} \partial_l^2 B\|_{H^1} \leq \|\mathcal{A}\|_{W^{1,\infty}} \|\partial_l^2 B\|_{H^1} \leq \|\mathcal{A}\|_{W^{1,\infty}} \|B\|_{H^3}$.

Since $\|\mathcal{A}\|_{W^{1,\infty}} = \|\mathcal{A}\|_{L^\infty} + \|\mathcal{A}_w\|_{L^\infty}$ and $\mathcal{A}$ is an antiderivative of A we can use Lemma 5.7 to conclude that $\|\mathcal{A}\|_{L^\infty} \leq \sqrt{\pi}\|A\|_{L^2(1)}$. Likewise, Sobolev’s inequality tells us that $\|\mathcal{A}_w\|_{L^\infty} = \|A\|_{L^\infty} \leq C\|A\|_{H^1}$. So all together we have

$$\|g\|_{\ell^2} \leq C e^{-1/2} (\|A\|_{L^2(1)} + \|A\|_{H^1}) \|B\|_{H^3}.$$

Since A and B are assumed to be good solutions of KdV on $[-T_0, T_0]$ we get $\sup_{|j| \leq T_0/e^3} \|g\|_{\ell^2} \leq C e^{-1/2}$, which is the targeted estimate.

All the other terms are handled using the same sorts of steps used above. We close the proof with a comment on the regularity needed. The most smoothness required for A and B comes from the terms in $\partial_T Q_{30}$. As in [5–7], one finds that $\partial_w^6 A$ and $\partial_l^6 B$ make an appearance and so, to deploy estimates like in Lemma 5.6, we need A and B to be in H^7 . $\square$

5.2. The autoregressive part

Now we need to put bounds on terms where γ_1 and γ_2 appear. The first question: how big are these sequences? The equations in (4.22) which these satisfy are examples of autoregressive models, specifically AR(1) processes [20]. We have the following almost sure estimate for solutions of such processes:

Lemma 5.8. Suppose that $z(n)$, $n \geq 0$, are i.i.d. random variables with zero mean and compact support. Fix $\theta \in (-1, 1)$ and let

$$\chi(n) := \sum_{k=0}^{n-1} \theta^k z(n-k). \quad (5.3)$$

Then there exists a constant $C > 0$ so that

$$\sup_{n>0} \frac{|\chi(n)|}{\sqrt{\ln(e+n)}} \leq C \sqrt{\frac{1}{1-\theta^2}}.$$

The constant C depends on the realization of $z(n)$ but does not depend on θ ; it is almost surely finite.

Proof. The result is a consequence of Hoeffding’s inequality, whose proof can be found in [22]:

Theorem 5.9. Let $w(0), \dots, w(n-1)$ be mean-zero, independent random variables with $-b_k \leq w(k) \leq b_k$ almost surely and $\chi(n) = \sum_{k=0}^{n-1} w(k)$. Then for any $\mu \geq 0$

$$\mathbb{P}(|\chi(n)| \geq \mu) \leq 2 \exp\left(-\frac{\mu^2}{2 \sum_{k=0}^{n-1} b_k^2}\right).$$

We apply this to (5.3); let $w_n(k) := \theta^k z(n-k)$. Since $\mathbb{E}[z(j)] = 0$ we have $\mathbb{E}[w_n(k)] = 0$ for all choices of n and k . Since the $z(j)$ are independent it follows that, for fixed n , the $w_n(k)$ are independent with respect to k . The support of $z(j)$ is compact so there is $a \geq 0$ for which the support lies in $[-a, a]$. Then the support of $\theta^k z(n-k)$ is in $[-a\theta^k, a\theta^k]$. Thus $w_n(0), \dots, w_n(n-1)$ pass the hypotheses of Theorem 5.9 with $b_k = a\theta^k$ and we have:

$$\mathbb{P}(|\chi(n)| \geq \mu) \leq 2 \exp\left(-\frac{\mu^2}{2a^2 \sum_{k=0}^{n-1} \theta^{2k}}\right) = 2 \exp\left(-\frac{\mu^2(1-\theta^2)}{2a^2(1-\theta^{2n})}\right).$$

Now let $\mu(n) := \sqrt{\ln(e+n) \frac{4a^2(1-\theta^{2n})}{1-\theta^2}}$ so that

$$\mathbb{P}(|\chi(n)| \geq \mu(n)) \leq 2 \exp\left(-\frac{\mu^2(1-\theta^2)}{2a^2(1-\theta^{2n})}\right) = 2 \exp(-2 \ln(e+n)) = \frac{2}{(e+n)^2}.$$

Since $\sum_{n \geq 0} 2/(e+n)^2$ is finite, the Borel–Cantelli Lemma [23] tells us that, almost surely, $|\chi(n)| \geq \mu(n)$ happens for at most finitely many n . For a given realization of $z(n)$ let N_ω be the largest value of n at which $|\chi(n)| \geq \mu(n)$ and put $c_\omega := \max_{1 \leq n \leq N_\omega} |\chi(n)|/\mu(n)$. Thus we have

$$|\chi(n)| \leq 2ac_\omega \sqrt{\ln(e+n) \frac{1-\theta^{2n}}{1-\theta^2}} \leq 2ac_\omega \sqrt{\frac{\ln(e+n)}{1-\theta^2}}$$

for all n . Putting $C = 2ac_\omega$ completes the proof. $\square$

With Lemma 5.8 we can prove

Lemma 5.10. Take Hypothesis 1.1 as given. Suppose that $\gamma_1(j)$ and $\gamma_2(j)$ solve (4.22) and $\gamma_1(0) = \gamma_2(0) = 0$. Then there exists a constant $C > 0$ such that for all $\epsilon \in (0, 1)$ we have

$$\sup_{j \in \mathbb{Z}} \frac{|\gamma_1(j)| + |\gamma_2(j)|}{\sqrt{\ln(e+|j|)}} \leq C e^{-1/2}.$$

The constant C depends on the realization of the $\zeta(j)$ but is almost surely finite.

Proof. We prove the estimate for γ_2 as the one for γ_1 is similar but easier. Taking $j > 0$ in the second equation in (4.22) gives

$$\gamma_2(j) - \gamma_2(j-1) = -\epsilon \gamma_2(j) + (\zeta(j) + S^- \zeta(j) + \zeta(j) \delta^+ \delta^- \zeta(j) + 2\sigma^2)$$

or rather

$$\gamma_2(j) = \frac{1}{1+\epsilon} \gamma_2(j-1) + \frac{1}{1+\epsilon} (\zeta(j) + S^- \zeta(j) + \zeta(j) \delta^+ \delta^- \zeta(j) + 2\sigma^2).$$

If we take $\gamma_2(0) = 0$ then the we can find $\gamma_2(j)$ (for $j > 0$) from the above by iteration. In particular we have

$$\begin{aligned} \gamma_2(j) &= \frac{1}{1+\epsilon} \sum_{k=0}^{j-1} \theta_\epsilon^k \zeta(j-k) + \frac{1}{1+\epsilon} \sum_{k=0}^{j-1} \theta_\epsilon^k S^{-1} \zeta(j-k) \\ &\quad + \frac{1}{1+\epsilon} \sum_{k=0}^{j-1} \theta_\epsilon^k (\zeta(j-k) \delta^+ \delta^- \zeta(j-k) + 2\sigma^2) \\ &=: \frac{1}{1+\epsilon} (\gamma_{21}(j) + \gamma_{22}(j) + \gamma_{23}(j)) \end{aligned}$$

where we have put $\theta_\epsilon := 1/(1+\epsilon)$. To be clear $\gamma_{21}(j)$, $\gamma_{22}(j)$ and $\gamma_{23}(j)$ correspond to the three sums in the order of their appearance.

The random variables $\zeta(j)$ meet the hypotheses of [Lemma 5.8](#) and so we can apply the results to $\gamma_{21}(j)$ and $\gamma_{22}(j)$ forthwith to get:

$$\sup_{j \in \mathbb{Z}} \frac{|\gamma_{21}(j)|}{\sqrt{\ln(e + |j|)}} \leq C \sqrt{\frac{1}{1 - \theta_\epsilon^2}} \quad \text{and} \quad \sup_{j \in \mathbb{Z}} \frac{|\gamma_{22}(j)|}{\sqrt{\ln(e + |j|)}} \leq C \sqrt{\frac{1}{1 - \theta_\epsilon^2}}$$

for some $C > 0$ which is almost surely finite. An easy calculation shows that $1/(1 - \theta_\epsilon^2) = (1 + \epsilon^2)/(2\epsilon + \epsilon^2) < 1/\epsilon$ when $\epsilon \in (0, 1)$. Thus we have

$$\sup_{j \geq 0} \frac{|\gamma_{21}(j)| + |\gamma_{22}(j)|}{\sqrt{\ln(e + |j|)}} \leq C\epsilon^{-1/2}.$$

Dealing with $\gamma_{23}(j)$ is a bit more complicated because the summands are not independent. We have $\gamma_{23}(j) = \sum_{k=0}^{j-1} \theta_\epsilon^k v(j - k)$ where

$$v(j) = \zeta(j)\zeta(j+1) + \zeta(j)\zeta(j-1) - 2\zeta(j)^2 + 2\sigma^2.$$

From this we see that $v(j)$ and $v(j+1)$ are dependent. As are $v(j)$ and $v(j+2)$, since $\zeta(j+1)$ appears in both. But $v(j+3)$ and $v(j)$ have no terms in common and it follows that they are independent. Thus $\{v(3l)\}_{l \in \mathbb{Z}}$ is an i.i.d. collection of random variables. As are $\{v(3l+1)\}_{l \in \mathbb{Z}}$ and $\{v(3l+2)\}_{l \in \mathbb{Z}}$. We break up $\gamma_{23}(k)$ accordingly:

$$\gamma_{23}(j) = \sum_{\substack{0 \leq k \leq j-1 \\ k=0 \bmod 3}} \theta_\epsilon^k v(j - k) + \sum_{\substack{0 \leq k \leq j-1 \\ k=1 \bmod 3}} \theta_\epsilon^k v(j - k) + \sum_{\substack{0 \leq k \leq j-1 \\ k=2 \bmod 3}} \theta_\epsilon^k v(j - k).$$

Each of the three sums passes the hypotheses of [Lemma 5.8](#), though there are some small subtleties. We estimate the first as the others are all but the same. Put $k = 3l$ to find

$$\sum_{\substack{0 \leq k \leq j-1 \\ k=0 \bmod 3}} \theta_\epsilon^k v(j - k) = \sum_{l=0}^{\lfloor (j-1)/3 \rfloor} \theta_\epsilon^{3l} v(j - 3l).$$

Then we have from [Lemma 5.8](#):

$$\sum_{l=0}^{\lfloor (j-1)/3 \rfloor} \theta_\epsilon^{3l} v(j - 3l) \leq C \sqrt{\ln(e + \lfloor (j-1)/3 \rfloor)} \sqrt{\frac{1}{1 - \theta_\epsilon^6}}.$$

As $\theta_\epsilon^2 > 0$ we find $1/(1 - \theta_\epsilon^6) = 1/(1 - \theta_\epsilon^2)(1 + \theta_\epsilon^2 + \theta_\epsilon^4) \leq 1/(1 - \theta_\epsilon^2) < 1/\epsilon$. This, along with the fact that $\ln$ is an increasing function, gives

$$\sum_{l=0}^{\lfloor (j-1)/3 \rfloor} \theta_\epsilon^{3l} v(j - 3l) \leq C \sqrt{\ln(e + |j|)} \epsilon^{-1/2}$$

which in turn leads to the estimate we are after.

We need estimates for $\gamma_2(j)$ when $j < 0$ too. If we take $j < 0$ in the second equation of [\(4.22\)](#) we get

$$\gamma_2(j) - \gamma_2(j-1) = \epsilon \gamma_2(j) + (\zeta(j) + S^- \zeta(j) + \zeta(j) \delta^+ \delta^- \zeta(j) + 2\sigma^2).$$

We rearrange this:

$$\gamma_2(j-1) = (1 - \epsilon) \gamma_2(j) - (\zeta(j) + S^- \zeta(j) + \zeta(j) \delta^+ \delta^- \zeta(j) + 2\sigma^2).$$

As we have taken $\gamma_2(0) = 0$ the above formula gives us $\gamma_2(-1)$ and, more generally, $\gamma_2(j)$, $j < 0$, by iteration. For $j = -l < 0$ we obtain:

$$\begin{aligned} \gamma_2(-l) &= - \sum_{k=0}^{l-1} \theta_\epsilon^k \zeta(-l + k + 1) - \sum_{k=0}^{l-1} \theta_\epsilon^k S^{-1} \zeta(-l + k + 1) \\ &\quad - \sum_{k=0}^{l-1} \theta_\epsilon^k (\zeta(-l + k + 1) \delta^+ \delta^- \zeta(-l + k + 1) + 2\sigma^2) \end{aligned}$$

where $\theta_\epsilon := 1 - \epsilon$. The first two sums pass the hypotheses of [Lemma 5.8](#) and since $1/(1 - \theta_\epsilon^2) = 1/(2\epsilon - \epsilon^2) < 1/\epsilon$ when $\epsilon \in (0, 1)$ we can bound both as we did for γ_{21} and γ_{22} earlier. And the same skulduggery about independence that worked for γ_{23} works for the third sum. All together we get

$$\sup_{j \leq 0} \frac{|\gamma_2(j)|}{\sqrt{\ln(e + |j|)}} \leq C\epsilon^{-1/2}.$$

That completes the proof of [Lemma 5.10](#). $\square$

Next we prove the main workhorse lemma for controlling γ terms in our approximation:

Lemma 5.11. *If $F = F(w, T)$ has $\sup_{|T| \leq T_0} \|F(\cdot, T)\|_{H^1(1)} < \infty$ then*

$$\sup_{|t| \leq T_0/\epsilon^3} \|\gamma_k(\cdot) F(\epsilon(\cdot \pm t), \epsilon^3 t)\|_{\ell^2} \leq C\epsilon^{-1} \sqrt{|\ln(\epsilon)|} \sup_{|T| \leq T_0} \|F(\cdot, T)\|_{H^1(1)} \quad (5.4)$$

for $k = 1, 2$.

If in addition $\sup_{|T| \leq T_0} \|F(\cdot, T)\|_{H^2(1)} < \infty$ then

$$\sup_{|t| \leq T_0/\epsilon^3} \|E_0^\pm (\gamma_k(\cdot) F(\epsilon(\cdot \pm t), \epsilon^3 t))\|_{\ell^2} \leq C\epsilon^{-1} \sqrt{|\ln(\epsilon)|} \sup_{|T| \leq T_0} \|F(\cdot, T)\|_{H^2(1)} \quad (5.5)$$

for $k = 1, 2$. (The choices for $+$ or $-$ in $E_0^\pm$ and $F(\epsilon(\cdot \pm t), \epsilon^3 t)$ are not linked.)

The constant $C > 0$ is almost surely finite.

Proof. First we tackle [\(5.4\)](#). We handle $k = 1$ and the “ $-$ ” sign. The other cases are no different. First

$$\|\gamma_1(\cdot) F(\epsilon(\cdot - t), \epsilon^3 t)\|_{\ell^2}^2 = \sum_{j \in \mathbb{Z}} \gamma_1^2(j) F(\epsilon(j - t), \epsilon^3 t)^2.$$

Using the estimate from [Lemma 5.10](#) gets

$$\|\gamma_1(\cdot) F(\epsilon(\cdot - t), \epsilon^3 t)\|_{\ell^2}^2 \leq C\epsilon^{-1} \sum_{j \in \mathbb{Z}} \ln(e + |j|) F(\epsilon(j - t), \epsilon^3 t)^2.$$

A simple estimate leads us to

$$\begin{aligned} \|\gamma_1(\cdot) F(\epsilon(\cdot - t), \epsilon^3 t)\|_{\ell^2}^2 &\leq C\epsilon^{-1} \sup_{j \in \mathbb{Z}} \frac{\ln(e + |j|)}{1 + (\epsilon(j - t))^2} \sum_{j \in \mathbb{Z}} (1 + (\epsilon(j - t))^2) F(\epsilon(j - t), \epsilon^3 t)^2. \end{aligned}$$

We can apply the second estimate in [Lemma 5.6](#) to the sum (with $u = (1 + (\epsilon(j - t))^2) F(\epsilon(j - t), \epsilon^3 t)^2$) and find

$$\begin{aligned} \|\gamma_1(\cdot) F(\epsilon(\cdot - t), \epsilon^3 t)\|_{\ell^2}^2 &\leq C\epsilon^{-2} \sup_{j \in \mathbb{Z}} \frac{\ln(e + |j|)}{1 + (\epsilon(j - t))^2} \|\sqrt{1 + (\cdot)^2} F(\cdot, \epsilon^3 t)\|_{H^1}^2 \\ &\leq C\epsilon^{-2} \sup_{j \in \mathbb{Z}} \frac{\ln(e + |j|)}{1 + (\epsilon(j - t))^2} \|F(\cdot, \epsilon^3 t)\|_{H^1(1)}^2 \end{aligned}$$

Thus

$$\begin{aligned} \sup_{|t| \leq T_0/\epsilon^3} \|\gamma_1(\cdot) F(\epsilon(\cdot - t), \epsilon^3 t)\|_{\ell^2} &\leq C\epsilon^{-1} \sqrt{\sup_{|t| \leq T_0/\epsilon^3} \sup_{j \in \mathbb{Z}} \frac{\ln(e + |j|)}{1 + (\epsilon(j - t))^2}} \sup_{|T| \leq T_0} \|F(\cdot, T)\|_{H^1(1)}. \end{aligned}$$

From this we see that the proof of [\(5.4\)](#) will be complete once we show that there is $C = C(T_0) > 0$ such that $0 < \epsilon < 1$ implies

$$\sup_{|t| \leq T_0/\epsilon^3} \sup_{j \in \mathbb{Z}} \frac{\ln(e + |j|)}{1 + (\epsilon(j - t))^2} \leq C |\ln(\epsilon)|.$$

The proof is mainly elementary Calculus, but that does not mean it is obvious. Here are the details. Let $f_\epsilon(y, t) := \frac{\ln(e + |y|)}{1 + (\epsilon(y - t))^2}$. We show $\sup_{|t| \leq T_0/\epsilon^3} \sup_{y \in \mathbb{R}} f_\epsilon(y, t) \leq C |\ln(\epsilon)|$. Since $f_\epsilon(y, t) = f_\epsilon(-y, -t)$ we have $\sup_{|t| \leq T_0/\epsilon^3} \sup_{y \in \mathbb{R}} f_\epsilon(y, t) = \sup_{0 \leq t \leq T_0/\epsilon^3} \sup_{y \in \mathbb{R}} f_\epsilon(y, t)$. If $t \geq 0$ and $y \geq 0$ then $|y - t| \leq |y|$ which implies $f_\epsilon(-y, t) \leq f_\epsilon(y, t)$. Thus $\sup_{|t| \leq T_0/\epsilon^3} \sup_{y \in \mathbb{R}} f_\epsilon(y, t) = \sup_{0 \leq t \leq T_0/\epsilon^3} \sup_{y \geq 0} f_\epsilon(y, t)$.

Next we argue that $f_\epsilon(y, t)$ achieves its supremum at a point in $(0, \infty)$. Clearly $f_\epsilon(y, t)$ is non-negative and $f_\epsilon(y, t) \rightarrow 0$ as $y \rightarrow \infty$. It is easy enough to show that $\lim_{y \rightarrow 0^+} \partial_y f_\epsilon(y, t) > 0$ when $t \geq 0$. Since f_ϵ is smooth (except at $y = 0$), these considerations imply the existence of $y_\epsilon(t) \in (0, \infty)$ for which $f_\epsilon(y_\epsilon(t), t) = \sup_{y \in \mathbb{R}} f_\epsilon(y, t)$ and $\partial_y f_\epsilon(y_\epsilon(t), t) = 0$.

So we search for solutions of $\partial_y f_\epsilon(y, t) = 0$ with $y \geq 0$. We claim that for $t \geq 0$ and $\epsilon \in (0, 1)$ that

$$\partial_y f_\epsilon(y, t) = 0 \text{ and } y \geq 0 \implies t \leq y \leq t + \frac{1}{\epsilon}. \quad (5.6)$$

Given the claim, $t \leq y_\epsilon(t) \leq t + \epsilon^{-1}$ follows and as such:

$$f_\epsilon(y_\epsilon(t), t) = \frac{\ln(e + y_\epsilon(t))}{1 + (\epsilon(y_\epsilon(t) - t))^2} \leq \ln(e + t + \epsilon^{-1}).$$

In turn we have $\sup_{0 \leq t \leq T_0/\epsilon^3} \sup_{y \geq 0} f_\epsilon(y, t) \leq \ln(e + T_0\epsilon^{-3} + \epsilon^{-1}) \leq C|\ln(\epsilon)|$ for a constant depending only on T_0 .

So we will be done if we establish the claim. Routine computations show that $\partial_y f_\epsilon(y, t) = 0$ if and only if

$$\frac{1 + \epsilon^2(y - t)^2}{2\epsilon^2(y - t)} = (e + y)\ln(e + y). \quad (5.7)$$

Note that if $0 \leq y < t$ then the left-hand side of (5.7) is negative whereas the right-hand side is positive. So there can be no solutions with $y < t$ and this implies the left-hand inequality in (5.6).

For the right-hand inequality, let us assume that $y - t > \epsilon^{-1}$ and $t \geq 0$. This gives

$$\frac{1 + \epsilon^2(y - t)^2}{2\epsilon^2(y - t)} = \frac{1}{2\epsilon^2(y - t)} + \frac{y - t}{2} < \frac{1}{2\epsilon} + \frac{y - t}{2}.$$

Next, since $t \geq 0$ then $y - t \leq y$ which implies $\frac{1 + \epsilon^2(y - t)^2}{2\epsilon^2(y - t)} < \frac{1}{2\epsilon} + \frac{y}{2}$.

Since $y - t > \epsilon^{-1}$ and $t \geq 0$ we have $y > \epsilon^{-1}$. Thus $\frac{1 + \epsilon^2(y - t)^2}{2\epsilon^2(y - t)} < y$. Also we clearly have $y < (e + y)\ln(e + y)$ and so all told

$$\frac{1 + \epsilon^2(y - t)^2}{2\epsilon^2(y - t)} < (e + y)\ln(e + y)$$

when $y - t > \epsilon^{-1}$ and $t \geq 0$. This precludes $\partial_y f_\epsilon(y, t) = 0$ and the right inequality in (5.6) follows. Thus we are done with the proof of (5.4).

The estimate (5.5) follows from (5.4) with a few tricks. First we have by direct calculation $E_0^+(\gamma_1 F) = \gamma_1(j + 1)(E_0^+ F)$. Second, if we let $I_\epsilon G(X) = \epsilon^{-1} \int_X^{X+\epsilon} G(Y) dY$ then the Fundamental Theorem of Calculus and the definition of E_0^+ tell us $E_0^+ F = I_\epsilon \partial_w F$. One can show (see the argument that leads to equation (3.4) in [2]) that $\|I_\epsilon G\|_{H^n(r)} \leq C\|G\|_{H^n(r)}$. Putting it all together we get

$$\begin{aligned} & \sup_{|t| \leq T_0/\epsilon^3} \|E_0^+(\gamma_k(\cdot)F(\epsilon(\cdot \pm t), \epsilon^3 t))\|_{\ell^2} \\ &= \sup_{|t| \leq T_0/\epsilon^3} \|\gamma_k(\cdot + 1)I_\epsilon \partial_w F(\epsilon(\cdot \pm t), \epsilon^3 t)\|_{\ell^2} \\ &\leq C\epsilon^{-1} \sqrt{|\ln(\epsilon)|} \sup_{|T| \leq T_0} \|I_\epsilon \partial_w F(\cdot, T)\|_{H^1(1)} \\ &\leq C\epsilon^{-1} \sqrt{|\ln(\epsilon)|} \sup_{|T| \leq T_0} \|F(\cdot, T)\|_{H^2(1)}. \end{aligned}$$

Now we can control all the γ dependent terms.

Lemma 5.12. Assume Hypothesis 1.1. Let $\tilde{q}_\epsilon(j, t)$ and $\tilde{p}_\epsilon(j, t)$ be the extended KdV approximators as in Definition 4.4 where we further assume that A and B are good solutions of KdV on $[-T_0, T_0]$. Then almost surely

$$\sup_{|t| \leq T_0\epsilon^{-3}} (\|P_3(\cdot, \epsilon, \epsilon t, \epsilon^3 t)\|_{\ell^2} + \|Q_{3\gamma}(\cdot, \epsilon, \epsilon t, \epsilon^3 t)\|_{\ell^2}) = \mathcal{O}(\epsilon^{-1} \sqrt{|\ln(\epsilon)|})$$

and

$$\sup_{|t| \leq T_0\epsilon^{-3}} (\|W_{1\gamma}(\cdot, \epsilon, \epsilon t, \epsilon^3 t)\|_{\ell^2} + \|W_{2\gamma}(\cdot, \epsilon, \epsilon t, \epsilon^3 t)\|_{\ell^2}) = \mathcal{O}(\epsilon^{-1} \sqrt{|\ln(\epsilon)|}).$$

Proof. The estimates for P_3 and $Q_{3\gamma}$ are immediate from Lemma 5.11 and their definitions. A direct calculation shows that

$$W_{1\gamma} = E_0^+ P_3 - \partial_\tau Q_{3\gamma} - \epsilon^2 \partial_T Q_{3\gamma}.$$

Each of these can be estimated with Lemma 5.11 as well. Another calculation gives

$$\begin{aligned} W_{2\gamma} &= E_0^- Q_{3\gamma} - m \partial_\tau P_3 - m \epsilon^2 \partial_T P_3 \\ &\quad + 2\epsilon^2 D^-(Q_{3\gamma} Q_1) + 2\epsilon^3 D^-(Q_{3\gamma} Q_2) + 2\epsilon^4 D^-(Q_{3\gamma} Q_3) + \epsilon^4 D^- Q_{3\gamma}^2. \end{aligned}$$

The first line of the above we estimate with Lemma 5.11. The ones in the second line all hinge on estimating terms of the form $D^-(Q_{3\gamma} Q_l)$ for different choices of l . The definition of D^- and the triangle inequality give $\|D^-(Q_{3\gamma} Q_l)(\cdot, \epsilon, \epsilon t, \epsilon^3 t)\|_{\ell^2} \leq 2\|Q_{3\gamma}(\cdot, \epsilon, \epsilon t, \epsilon^3 t) Q_l(\cdot, \epsilon, \epsilon t, \epsilon^3 t)\|_{\ell^2}$. Then we use the fact that $\|fg\|_{\ell^2} \leq \|f\|_{\ell^2} \|g\|_{\ell^2}$ to get $\|D^-(Q_{3\gamma} Q_l)(\cdot, \epsilon, \epsilon t, \epsilon^3 t)\|_{\ell^2} \leq 2\|Q_{3\gamma}(\cdot, \epsilon, \epsilon t, \epsilon^3 t)\|_{\ell^2} \|Q_l(\cdot, \epsilon, \epsilon t, \epsilon^3 t)\|_{\ell^2}$. At this point the remainder of the estimates

follow from earlier estimates on the component Q_k and bookkeeping. $\square$

5.3. Finishing up

We are now in position to prove Proposition 5.3.

Proof. We begin with $\alpha_1(\epsilon)$. From its definition, (4.1), (5.1) and the triangle inequality we have

$$\begin{aligned} \alpha_1(\epsilon) &\leq \sum_{n=0}^2 \epsilon^{n+2} \sup_{|t| \leq T_0/\epsilon^3} (\|Q_n(\cdot, \epsilon, \epsilon t, \epsilon^3 t)\|_{\ell^2} + \|P_n(\cdot, \epsilon, \epsilon t, \epsilon^3 t)\|_{\ell^2}) \\ &\quad + \epsilon^5 \sup_{|t| \leq T_0/\epsilon^3} \|Q_{30}(\cdot, \epsilon, \epsilon t, \epsilon^3 t)\|_{\ell^2} \\ &\quad + \epsilon^5 \sup_{|t| \leq T_0/\epsilon^3} (\|Q_{3\gamma}(\cdot, \epsilon, \epsilon t, \epsilon^3 t)\|_{\ell^2} + \|P_3(\cdot, \epsilon, \epsilon t, \epsilon^3 t)\|_{\ell^2}) \end{aligned}$$

Then Lemmas 5.4 and 5.12 gives us $\alpha_1(\epsilon) = \mathcal{O}(\epsilon^{3/2})$.

For $\alpha_3(\epsilon)$, we use its definition, Lemma 4.5, (5.2) and the triangle inequality to obtain

$$\begin{aligned} \alpha_3(\epsilon) &\leq \epsilon^6 \sup_{|t| \leq T_0/\epsilon^3} (\|\gamma_1(\cdot) \partial_X^3 P_0(\cdot, \epsilon, \epsilon t, \epsilon^3 t)\|_{\ell^2} + \|\gamma_2(\cdot) \partial_X^3 Q_0(\cdot, \epsilon, \epsilon t, \epsilon^3 t)\|_{\ell^2}) \\ &\quad + \epsilon^6 \sup_{|t| \leq T_0/\epsilon^3} (\|W_{10}(\cdot, \epsilon, \epsilon t, \epsilon^3 t)\|_{\ell^2} + \|W_{20}(\cdot, \epsilon, \epsilon t, \epsilon^3 t)\|_{\ell^2}) \\ &\quad + \epsilon^6 \sup_{|t| \leq T_0/\epsilon^3} (\|W_{1\gamma}(\cdot, \epsilon, \epsilon t, \epsilon^3 t)\|_{\ell^2} + \|W_{2\gamma}(\cdot, \epsilon, \epsilon t, \epsilon^3 t)\|_{\ell^2}). \end{aligned}$$

Then Lemmas 5.4, 5.11 and 5.12 give $\alpha_3(\epsilon) = \mathcal{O}(\epsilon^5 \sqrt{|\ln(\epsilon)|})$.

To prove the estimate for $\alpha_2(\epsilon)$, from (4.1) we have

$$\partial_t \tilde{q}_\epsilon = \epsilon^3 \partial_\tau Q_0 + \epsilon^5 \partial_T Q_0 + \epsilon^4 \underbrace{\sum_{k=1}^3 (\epsilon^k \partial_\tau Q_k + \epsilon^{k+2} \partial_T Q_k)}_{\epsilon^4 \tilde{h}_\epsilon}.$$

Since $\|f\|_{\ell^\infty} \leq \|f\|_{\ell^2}$ we have $\|\tilde{h}_\epsilon\|_{\ell^\infty} \leq \|\tilde{h}_\epsilon\|_{\ell^2}$. All terms appearing in $\tilde{h}_\epsilon$ have been estimated in one place or another previously and each is $\mathcal{O}_{\ell^2}(\epsilon^{-1/2})$ at worst so that we get $\sup_{|t| \leq T_0/\epsilon^3} \|\epsilon^4 \tilde{h}_\epsilon\|_{\ell^\infty} = \mathcal{O}(\epsilon^{7/2})$. On the other hand using the first estimate in Lemma 5.6 shows that

$$\begin{aligned} & \sup_{|t| \leq T_0/\epsilon^3} \|\epsilon^3 \partial_\tau Q_0(\cdot, \epsilon, \epsilon t, \epsilon^3 t)\|_{\ell^\infty} \\ &\leq \sup_{|t| \leq T_0/\epsilon^3} \epsilon^3 (\|A_w(\cdot, \epsilon^3 t)\|_{W^{1,\infty}} + \|B_l(\cdot, \epsilon^3 t)\|_{W^{1,\infty}}) \leq C\epsilon^3. \end{aligned}$$

So all told we have $\alpha_2(\epsilon) = \mathcal{O}(\epsilon^3)$.

Next, if we let

$$\tilde{q}_\epsilon(j, t) := \sum_{k=1}^3 \epsilon^{k+2} Q_k(j, \epsilon j, \epsilon t, \epsilon^3 t) \quad \text{and} \quad \tilde{p}_\epsilon(j, t) := \sum_{k=1}^3 \epsilon^{k+2} P_k(j, \epsilon j, \epsilon t, \epsilon^3 t) \quad (5.8)$$

then the estimates from Lemmas 5.4 and 5.12 lead to

$$\sup_{|t| \leq T_0/\epsilon^3} \|\tilde{q}_\epsilon, \tilde{p}_\epsilon\|_{\ell^2} = \mathcal{O}(\epsilon^{5/2}). \quad (5.9)$$

So long A and B are not both identically zero it is easy using the conservation laws of KdV to find that $\inf_{|T| \leq T_0} \|A(\cdot, T)\|_{H^7(1)} + \|B(\cdot, T)\|_{H^7(1)} \geq b > 0$, for some b . This leads, by the triangle inequality, to $\beta_1(\epsilon) = \inf_{|t| \leq T_0/\epsilon^3} \|\tilde{q}_\epsilon, \tilde{p}_\epsilon\|_{\ell^2} \geq C\epsilon^{3/2}$. This completes the proof. $\square$

6. The main event

Now we can state and prove our main theorem in full detail.

Theorem 6.1. Let $m(j)$ be a realization of the mass coefficients subject to Hypothesis 1.1. Fix $T_0 > 0$ and $\Phi, \Psi \in H^7(1)$. Let $(q(j, t), p(j, t))$ be the solution of the transparent random mass FPUT lattice (1.1) with initial data $q(j, 0) = \epsilon^2 \Phi(\epsilon j)$ and $p(j, 0) = \epsilon^2 \Psi(\epsilon j)$.

Let $A(w, T)$ and $B(l, t)$ be the solutions of the KdV equations (4.17) and (4.19) with initial data

$$A(w, 0) = \frac{1}{2}\Phi(w) - \frac{1}{2}\Psi(w) \quad \text{and} \quad B(l, 0) = \frac{1}{2}\Phi(l) + \frac{1}{2}\Psi(l).$$

Then there exists $\epsilon_\star = \epsilon_\star(m, T_0, \Phi, \Psi)$ (almost surely positive) and $C_\star = C_\star(m, T_0, \Phi, \Psi) > 0$ (almost surely finite) such that, for all $\epsilon \in (0, \epsilon_\star)$, we have the absolute ℓ^2 -error estimates

$$\begin{aligned} \sup_{|l| \leq T_0/\epsilon^3} \left\| q(\cdot, t) - \epsilon^2 [A(\epsilon(\cdot - t), \epsilon^3 t) + B(\epsilon(\cdot + t), \epsilon^3 t)] \right\|_{\ell^2} \\ \leq C_\star \epsilon^2 \sqrt{|\ln(\epsilon)|} \quad \text{and} \\ \sup_{|l| \leq T_0/\epsilon^3} \left\| p(\cdot, t) - \epsilon^2 [-A(\epsilon(\cdot - t), \epsilon^3 t) + B(\epsilon(\cdot + t), \epsilon^3 t)] \right\|_{\ell^2} \\ \leq C_\star \epsilon^2 \sqrt{|\ln(\epsilon)|}. \end{aligned}$$

If at least one of Φ or Ψ is non-zero then the associated relative ℓ^2 -error estimates are $\mathcal{O}(\sqrt{\epsilon} |\ln(\epsilon)|)$.

Proof. Take A and B with the initial data as in the statement and form $\tilde{q}_\epsilon$ and $\tilde{p}_\epsilon$ as in Definition 4.4. Then we have the estimates for $\alpha_1(\epsilon)$, $\alpha_2(\epsilon)$, $\alpha_3(\epsilon)$ and $\beta_1(\epsilon)$ as in Proposition 5.3, which is to say we have met condition (3.3) from the statement of Theorem 3.1.

Note that

$$q(j, t) - \tilde{q}_\epsilon(j, t) = q(j, t) - \epsilon^2 [A(\epsilon(j - t), \epsilon^3 t) + B(\epsilon(j + t), \epsilon^3 t)] - \tilde{q}_\epsilon(j, t)$$

where $\tilde{q}_\epsilon$ is given in (5.8). In (5.9) we showed that $\tilde{q}_\epsilon = \mathcal{O}_{\ell^2}(\epsilon^{5/2})$ for $|t| \leq T_0/\epsilon^3$. The initial conditions for p , q , A and B are arranged so that

$$q(j, 0) - \tilde{q}_\epsilon(j, 0) = -\tilde{q}_\epsilon(j, 0)$$

And so we have $\|q(0) - \tilde{q}_\epsilon(0)\|_{\ell^2} = \|\tilde{q}_\epsilon(0)\|_{\ell^2} = \mathcal{O}(\epsilon^{5/2})$. Similarly we have

$$p(j, t) - \tilde{p}_\epsilon(j, t) = p(j, t) - \epsilon^2 [-A(\epsilon(j - t), \epsilon^3 t) + B(\epsilon(j + t), \epsilon^3 t)] - \tilde{p}_\epsilon(j, t)$$

and $\|p(0) - \tilde{p}_\epsilon(0)\|_{\ell^2} = \|\tilde{p}_\epsilon(0)\|_{\ell^2} = \mathcal{O}(\epsilon^{5/2})$.

Since $\alpha_3(\epsilon)/\epsilon^3 = \mathcal{O}(\epsilon^2 \sqrt{|\ln(\epsilon)|})$ and $\epsilon^{5/2} = o(\epsilon^2 \sqrt{|\ln(\epsilon)|})$, these estimates imply that we meet the condition on the initial data in the statement of Theorem 3.1. Thus we have the conclusion

$$\sup_{|t| \leq T_0/\epsilon^3} \|q(t) - \tilde{q}_\epsilon(t), p(t) - \tilde{p}_\epsilon(t)\|_{\ell^2} = \mathcal{O}(\epsilon^2 \sqrt{|\ln(\epsilon)|}).$$

Then the triangle inequality plus (5.8) give

$$\sup_{|t| \leq T_0/\epsilon^3} \left\| q(\cdot, t) - \epsilon^2 [A(\epsilon(\cdot - t), \epsilon^3 t) + B(\epsilon(\cdot + t), \epsilon^3 t)] \right\|_{\ell^2} = \mathcal{O}(\epsilon^2 \sqrt{|\ln(\epsilon)|})$$

and

$$\sup_{|t| \leq T_0/\epsilon^3} \left\| p(\cdot, t) - \epsilon^2 [-A(\epsilon(\cdot - t), \epsilon^3 t) + B(\epsilon(\cdot + t), \epsilon^3 t)] \right\|_{\ell^2} = \mathcal{O}(\epsilon^2 \sqrt{|\ln(\epsilon)|}).$$

This is the absolute error estimate in the theorem. The relative error estimate follows from the estimate on $\beta_1(\epsilon)$. $\square$

7. Numerics

In this section we report the outcomes of a variety of numerical simulations of solutions of (1.1). In all cases our methodology is to truncate (1.1) to $|j| \leq M$ where $M \gg 1$ and enforce periodic boundary conditions (M is always taken to be so incredibly vast that the solutions are never large anywhere near the edges of the computational domain). The resulting system is a large finite-dimensional ODE which we solve with a standard RK4 algorithm. This is essentially the same method as used in [2,7]. The calculations were performed in MATLAB.

7.1. Amplitude attenuation

The first experiment simulates (1.1) with a number of choices for $m(j)$. These are:

- $m(j) = 1$ for all j , that is, they are constant.
- $m(j) = 1 + (-1)^j/4$, i.e. 2-periodic.
- $m(j)$ meet the transparency condition (1.3) where $\zeta(j)$ are drawn from the uniform distribution on $[-1/8, 1/8]$.
- $m(j)$ are i.i.d. random variables, drawn uniformly from $[1/2, 3/2]$.

For all these cases, we choose as initial conditions:

$$q(j, 0) = 3\epsilon^2 \operatorname{sech}^2(\sqrt{6}\epsilon j) \quad \text{and} \quad p(j, 0) = -3\epsilon^2 \operatorname{sech}^2(\sqrt{6}\epsilon j). \quad (7.1)$$

We take $\epsilon = 1/2, 1/4, 1/8$ and $1/16$ and simulate from $t = 0$ out to $t = 3/\epsilon^3$.

Famously, solutions of KdV equations with smooth and localized initial data will, over time, resolve into the sum of separated solitary waves of fixed amplitude [24]. Thus, if the solution of the FPUT lattice is well-approximated by a KdV equation we expect the ℓ^∞ -norm to at least roughly stabilize over long time periods. And so in Fig. 1 we plot $\|q(\cdot, T/\epsilon^3), p(\cdot, T/\epsilon^3)\|_{\ell^\infty}/\epsilon^2$ vs T , for $0 \leq T \leq 3$. (The scaling here is to be consistent with the long wave scaling so that we may compare various choices of ϵ on the same plot.) We see exactly this stabilization in the plots for the constant, 2-periodic and transparent cases. Furthermore, the stabilization becomes more pronounced as ϵ decreases, which is consistent with the rigorous KdV approximation theorems here and in [5–7]. But when the masses are taken to be i.i.d., there is an obvious, pronounced decay of the amplitude; this attenuation (up to the scaling) becomes stronger as ϵ decreases. This is why we said in the Introduction that a KdV approximation for the i.i.d. problem is not appropriate.

7.2. Numerical computation of optimal error bound

In the second experiment we aim to corroborate the conclusions of our main result, Theorem 6.1. We simulate (1.1) with $m(j)$ subject to the transparency condition (1.3) where $\zeta(j)$ are drawn from the uniform distribution on $[-1/8, 1/8]$. In this case, $\sigma^2 = 1/192$. We choose the initial data so that $B(l, T)$ is zero and $A(w, T)$ is an exact solitary wave solution of (4.17), namely $3 \operatorname{sech}^2\left(\sqrt{\frac{6}{1+24\sigma^2}}(w - T)\right)$. That is to say, we take

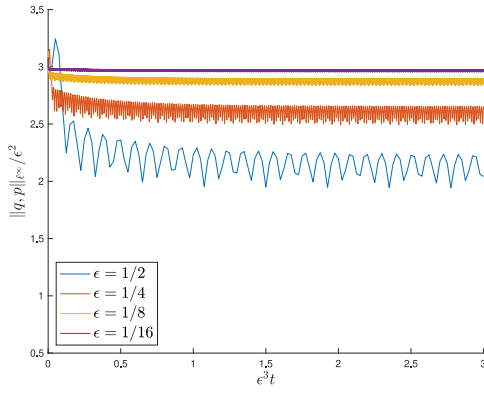
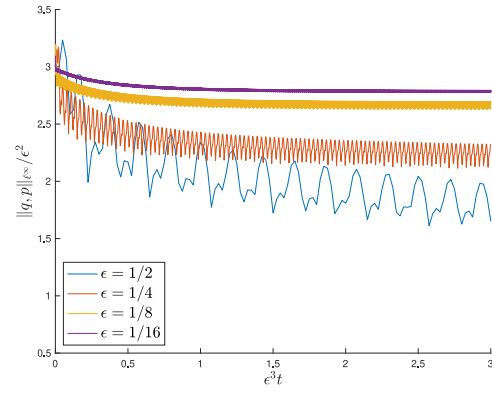
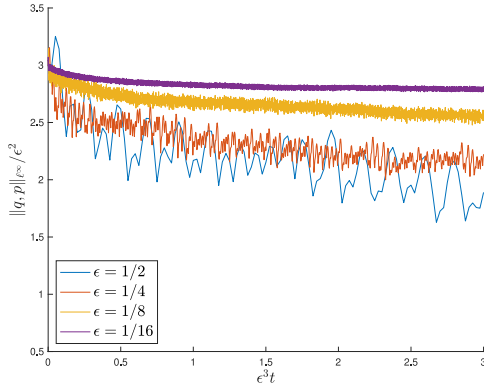
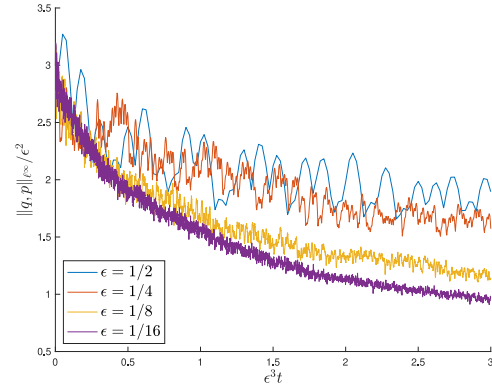
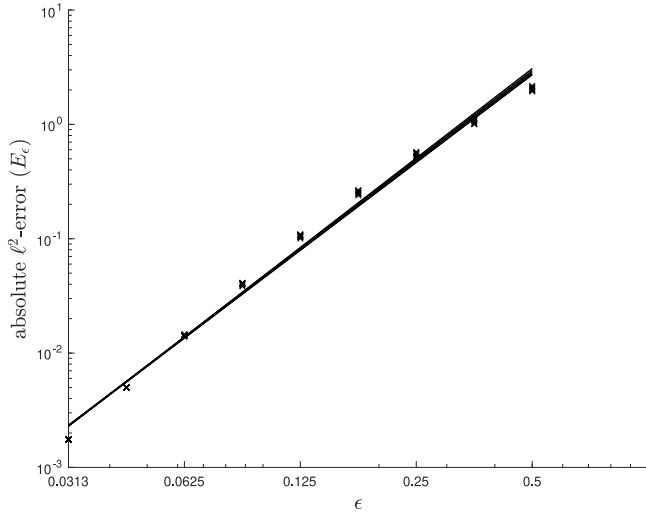
$$\begin{aligned} q(j, 0) &= 3\epsilon^2 \operatorname{sech}^2\left(\sqrt{\frac{6}{1+24\sigma^2}}\epsilon j\right) \quad \text{and} \\ p(j, 0) &= -3\epsilon^2 \operatorname{sech}^2\left(\sqrt{\frac{6}{1+24\sigma^2}}\epsilon j\right). \end{aligned} \quad (7.2)$$

We simulate for $\epsilon = 2^{-l/2}$ where $l = 2, \dots, 10$ and run the simulations from $t = 0$ to $t = 3/\epsilon^3$. (When $\epsilon = 1/32$ this takes a very long time!) To be clear, we fix a realization and then vary ϵ as stated with the same realization used throughout. Then we compute the overall absolute error, specifically:

$$\begin{aligned} E_\epsilon &:= \sup_{0 \leq t \leq T_0/\epsilon^3} \left\| q(\cdot, t) - 3\epsilon^2 \operatorname{sech}^2\left(\sqrt{\frac{6}{1+24\sigma^2}}(\epsilon(\cdot - t - \epsilon^2 t))\right) \right\|_{\ell^2} \\ &+ \sup_{0 \leq t \leq T_0/\epsilon^3} \left\| p(\cdot, t) + 3\epsilon^2 \operatorname{sech}^2\left(\sqrt{\frac{6}{1+24\sigma^2}}(\epsilon(\cdot - t - \epsilon^2 t))\right) \right\|_{\ell^2}. \end{aligned}$$

Then we repeat for another realization (ten different realizations all together). If we plot E_ϵ vs ϵ on a loglog plot, Theorem 6.1 tells us the best fit line to the data should have slope somewhere around 2, or larger. The results are shown in Fig. 2; all ten realizations on the same graph. The line of best fit has slope exceeding 2.5 in each case; they are 2.5715, 2.5982, 2.5677, 2.5445, 2.5778, 2.5948, 2.5510, 2.5760, 2.5499, 2.5563.

This numerically computed slope is over .5 larger than what we expect from our rigorous estimate. That is to say, the numerics indicate that the approximation of the transparent mass FPUT lattice by KdV is a fair bit better than what our results from Theorem 6.1 show.

(A) $m(j)$ are constant(B) $m(j)$ are 2-periodic(C) $m(j)$ are transparent(D) $m(j)$ are i.i.d.Fig. 1. Scaled ℓ^∞ -amplitude vs scaled time for solutions of (1.1) with long-wave data as in (7.1).Fig. 2. Ten realizations of absolute ℓ^2 -error vs ϵ , loglog plot, for solutions of (1.1) with “solitary wave-like” data as in (7.2).

We repeated the same experiment for many different realizations and the numerically computed slope was always near to 2.5. Thus we conjecture that the absolute ℓ^2 -error is at worst $\mathcal{O}(\epsilon^{5/2})$, which is the same size as the error for the constant and periodic problems. Of course we do not know how to prove such a thing at this time.

8. What is next

Our results are the first piece of much larger program aimed at bringing stochastic homogenization to nonlinear dispersive problems. Here are a number of open problems, some of which should be relatively straightforward given our results here and others of which will require substantial new technical ideas.

- (1) Prove a result analogous to Theorem 6.1 but in expectation instead of in the almost sure sense. In our work on the linear i.i.d. lattice [2] we proved approximation results in both senses and the strategy for the expectation result is almost surely transferable.
- (2) Study (1.1) but allow spatial heterogeneity in the spring potentials as well. We expect an analogous transparency condition can be used to achieve a result similar to the one here.
- (3) Confirm (or reject!) the conjecture that the sharp order of the KdV approximation error for the transparent mass FPUT lattice is smaller than $\mathcal{O}(\epsilon^{5/2})$. The error estimate we prove here is due entirely to our use of the autoregressive processes in the extended approximation. But perhaps a yet more clever option exists to handle the terms which we encountered at Z_{15} and Z_{25} .
- (4) If one can get the conjectured sharp error estimate, it opens the door to replacing the transparency condition (1.3) with

$$m(j) = 1 + \delta^- \zeta(j).$$

Here there is only one finite-difference on $\zeta(j)$ (which are still i.i.d.) instead of two. Some preliminary simulations indicate this condition, which we call the *translucent random mass FPUT lattice*, should have a valid KdV approximation.

- (5) The transparency condition is hardly natural or obvious. This raises the question: are there conditions that one can place on $m(j)$ that are less rigid than the transparency (or translucency) conditions? Note that the transparency condition implies that the $m(j)$ are correlated with one another; perhaps strictures placed upon correlation lengths in $m(j)$ can be used to get results similar to our results here.
- (6) How can one rigorously capture the amplitude attenuation seen in the numerics for the i.i.d. random mass problem in Fig. 1? Clearly KdV is the wrong approach, but perhaps there is some other modulation equation that can capture the dynamics. The articles [12,13] suggest the use of nonlinear diffusion equations. On the other hand, an analysis of a spectral problem associated with linear random lattices by [25] indicates a connection to Anderson localization and the authors of [26] utilize Boltzmann models in a high-dimensional linear version of (1.1).
- (7) Can our approach be carried over to the problem of long water waves over random bathymetry? Our transparency condition is inspired by a similar condition on the bathymetry proposed in [15]. Can a continuous version of an autoregressive process (that is, an *Ornstein–Uhlenbeck process*) be used to control the residuals in that problem and prove a rigorous approximation?
- (8) How about nonlinear wave equations with random coefficients? Nonlinear Schrödinger equations (discrete or continuous)? Any of the multitude of equations named for Joseph Valentin Boussinesq? Or problems in higher spatial dimensions?

CRedit authorship contribution statement

Joshua A. McGinnis: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. **J. Douglas Wright:** Conceptualization, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Software, Visualization, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

The authors would like to thank Amanda French, C. Eugene Wayne and Atilla Yilmaz for helpful conversations related to this project. Also, JDW would like to recognize the National Science Foundation, United States who supported this research with grant DMS-2006172.

References

- [1] Enrico Fermi, John Pasta, Stanislaw Ulam, *Collected Paper of Enrico Fermi*, vol. 2, University of Chicago Press, Chicago, 1965.
- [2] Joshua A. McGinnis, J. Douglas Wright, Using random walks to establish wavelike behavior in a linear FPUT system with random coefficients, *Discrete Contin. Dyn. Syst. Ser. S* 15 (9) (2022) 2581–2607.
- [3] Alexander Pankov, *Travelling Waves and Periodic Oscillations in Fermi-Pasta-Ulam Lattices*, Imperial College Press, London, 2005, p. xvi+194.
- [4] N.J. Zabusky, M.D. Kruskal, Interaction of “Solitons” in a collisionless plasma and the recurrence of initial states, *Phys. Rev. Lett.* 15 (1965) 240–243.
- [5] Guido Schneider, C. Eugene Wayne, Counter-propagating waves on fluid surfaces and the continuum limit of the Fermi-Pasta-Ulam model, in: *International Conference on Differential Equations*, vol. 1, (no. 2) World Sci. Publ., River Edge, NJ, 2000, pp. 390–404, (Berlin, 1999).
- [6] Martina Chirilus-Brukner, Christopher Chong, Oskar Prill, Guido Schneider, Rigorous description of macroscopic wave packets in infinite periodic chains of coupled oscillators by modulation equations, *Discrete Contin. Dyn. Syst. Ser. S* 5 (5) (2012) 879–901.
- [7] Jeremy Gaison, Shari Moskow, J. Douglas Wright, Qimin Zhang, Approximation of polyatomic FPU lattices by KdV equations, *Multiscale Model. Simul.* 12 (3) (2014) 953–995.
- [8] Takeshi Iizuka, Takenobu Nakao, Miki Wadati, Propagation of solitons in random lattices, *J. Phys. Soc. Japan* 60 (12) (1991) 4167–4174.
- [9] Miki Wadati, Deformation of solitons in random media, *J. Phys. Soc. Japan* 59 (12) (1990) 4201–4203.
- [10] Joshua A. McGinnis, Macroscopic wave propagation for 2D lattice with random masses, *Stud. Appl. Math.* 151 (2) (2023) 752–790.
- [11] J. Garnier, F. Kh. Abdullaev, Soliton dynamics in a random Toda chain, *Phys. Rev. E* 67 (2003) 026609.
- [12] T.V. Lapyteva, J.D. Bodyfelt, S. Flach, Do nonlinear waves in random media follow nonlinear diffusion equations? *Physica D* 256/257 (2013) 1–6.
- [13] T.V. Lapyteva, M.V. Ivanchenko, S. Flach, Nonlinear lattice waves in heterogeneous media, *J. Phys. A* 47 (2014).
- [14] Alejandro J. Martínez, P.G. Kevrekidis, Mason A. Porter, Superdiffusive transport and energy localization in disordered granular crystals, *Phys. Rev. E* 93 (2) (2016) 022902, 21.
- [15] Rodolfo R. Rosales, George C. Papanicolaou, Gravity waves in a channel with a rough bottom, *Stud. Appl. Math.* (2) (1983) 89–102.
- [16] W. Craig, P. Guyenne, C. Sulem, Water waves over a random bottom, *J. Fluid Mech.* 640 (2009) 79–107.
- [17] Anne de Bouard, Walter Craig, Oliver Díaz-Espinoza, Philippe Guyenne, Catherine Sulem, Long wave expansions for water waves over random topography, *Nonlinearity* 21 (9) (2008) 2143–2178.
- [18] Walter Craig, Catherine Sulem, Asymptotics of surface waves over random bathymetry, *Quart. Appl. Math.* 68 (1) (2010) 91–112.
- [19] Doina Cioranescu, Patrizia Donato, An introduction to homogenization, in: *Oxford Lecture Series in Mathematics and its Applications*, vol. 17, The Clarendon Press Oxford University Press, New York, 1999, p. x+262.
- [20] Geoffrey Grimmett, David Stirzaker, *Probability and Random Processes*, third ed., Oxford University Press, Oxford, United Kingdom, 2001.
- [21] Terence Tao, Nonlinear dispersive equations, in: *Conference Board of the Mathematical Sciences*, in: CBMS Regional Conference Series in Mathematics, vol. 106, Washington, DC; by the American Mathematical Society, Providence, RI, 2006, p. xvi+373, Local and global analysis.
- [22] Wassily Hoeffding, Probability inequalities for sums of bounded random variables, *J. Amer. Statist. Assoc.* 58 (301) (1963) 13–30.
- [23] Rick Durrett, *Probability: Theory and examples*, fourth ed., in: *Cambridge Series in Statistical and Probabilistic Mathematics*, Cambridge University Press, New York, USA, 2010.
- [24] P.G. Drazin, R.S. Johnson, *Solitons: an Introduction*, in: *Cambridge Texts in Applied Mathematics*, Cambridge University Press, Cambridge, 1989, p. xii+226.
- [25] Kyle E. Besing, *Spectral Properties and Localization of Two Random Laplacians on Graphs*, (Thesis (Ph.D.)), ProQuest LLC, Ann Arbor, MI, 2015, p. 166, The University of Alabama at Birmingham.
- [26] Jani Lukkarinen, Herbert Spohn, Kinetic limit for wave propagation in a random medium, *Arch. Ration. Mech. Anal.* 183 (1) (2007) 93–162.