Macroscopic wave propagation for 2D lattice with random masses

Joshua A. McGinnis

Department of Mathematics, Drexel University, Philadelphia, Pennsylvania, USA

Correspondence
Joshua A. McGinnis, Department of Mathematics, Drexel University, 3141 Chestnut St., Philadelphia, PA 19104, USA. Email: jam887@drexel.edu

Funding information
National Science Foundation, Grant/Award Number: DMS 2006172

Abstract
We consider a simple two-dimensional harmonic lattice with random, independent, and identically distributed masses. Using the methods of stochastic homogenization, we prove that solutions with initial data, which varies slowly relative to the lattice spacing, converge in an appropriate sense to solutions of an effective wave equation. The convergence is strong and almost sure. In addition, the role of the lattice’s dimension in the rate of convergence is discussed. The technique combines energy estimates with powerful classical results about sub-Gaussian random variables.

KEYWORDS
polyatomic lattice, random coefficients, stochastic homogenization

1 | INTRODUCTION

We prove an almost sure convergence result for the coarse-grained solutions of the following two-dimensional (2D), spatially discrete, harmonic lattice with random masses as the variation of the initial data becomes slower relative to the lattice spacing:

\[ m(j) \dddot{u}(j, t) = \Delta u(j, t), \]  

where \( j = (j_1, j_2) \in \mathbb{Z}^2 \). The discrete Laplacian is defined as

\[ \Delta u(j, t) := -4u(j, t) + \sum_{i=1}^{2} u(j + e_i, t) + u(j - e_i, t), \]
where \( \mathbf{e}_1 = (1,0) \) and \( \mathbf{e}_2 = (0,1) \). More specifically, our initial data have the scaling

\[
 u(\epsilon j) = \epsilon^{-1} \phi(\epsilon j) \quad \text{and} \quad \dot{u}(\epsilon j) = \psi(\epsilon j).
\]

We eventually require smoothness and algebraic decay of \( \phi \) and \( \psi \). The convergence holds for times proportional to the scale of the variation, that is, for all \( t \) s.t. \( |t| \leq T \epsilon^{-1} \), where \( T \) is a fixed constant. The \( m(j_1, j_2) \) are independent and identically distributed random variables (i.i.d.) contained almost surely in some interval \([a, b] \in \mathbb{R}^+\). Figure 1 depicts the lattice.

There is a long history of studying such lattices when the masses or constants of elasticity vary periodically but also when they vary randomly. Such lattices can provide simplified models for investigating physical phenomena arising in more complex systems, such as heat conduction, and have applications in solid-state physics. In particular, such models have been used to study the molecular vibrations in crystals which are typically thought of as having periodic structure. However, as pointed out in Bell, crystals may come in mixtures such as NaCl–KCl mixture where sodium and potassium atoms occupy “odd sites” randomly. Therefore, we are interested in understanding to what degree approximations to the wave equation are still valid, since random coefficients could in principle be used to model crystals with a mixture of elements, isotopes, or more general impurities.

Partial differential equation (PDE) approximations are a common avenue for investigating the behavior of lattices, and there are numerous examples of such approximations for deterministic 2D lattices. For periodic masses and slowly varying initial conditions, it has been shown that the behavior of a Bravais lattice of any dimension is approximated by a wave equation. For lattices where the interaction among neighboring masses is nonlinear, approximations by nonlinear PDEs have been used to investigate behavior for longer time scales than those considered here. The approximation of a 2D Fermi–Pasta–Ulam (FPU) lattice by the Kadomtsev–Petviashvili II (KP-II) equation has been justified in Hristov and Pelinovsky for waves of long length, traveling in certain directions. A justification of the KP-II approximation for general directions is given. The small parameter in those papers plays a similar role to \( \epsilon \) in this paper, and the approxima-
tions are good for $|t| \leq T\varepsilon^{-3}$. The resonant normal forms of the continuous approximations of similar lattices were shown to be given by various integrable PDEs, including KP-II, and these PDEs were used to investigate metastability for long times.\(^8\) Thus, the justified use of PDEs to explore the dynamics of 2D lattices is relatively well understood in the deterministic setting.

In comparison, the only rigorous results regarding wave propagation in disordered lattices are for weak randomness,\(^9\) and, otherwise, most of what is known comes from numerical experiments.\(^10\) Specifically, there has been a lot of interest in studying localization phenomena of disordered lattices.\(^11,12\) In particular, Pinski et al.\(^13\) look at a random mass lattice like the one we study here albeit in three dimensions (3D). In contrast to those studies, the main result of this paper is that, despite the fact that the eigenmodes may be localized, appropriately scaled initial conditions still evolve much like they would for a homogeneous lattice.

Much has already been said about the continuous analog of (1). For an extensive resource, see Fouque et al.\(^14\) Note that rates of convergence are rarely addressed. Furthermore, almost sure convergence results in the discrete setting require different techniques. In the continuous, setting convergence can be achieved more directly through the law of large numbers. In our setting, one must define what is meant by convergence, and this is why we prove convergence through coarse-graining, which is used in Mielke\(^5\) to prove convergence in the periodic problem. To achieve a rigorous rate of convergence, one must have bounds on the stochastic error terms. In McGinnis and Wright,\(^15\) the law of the iterated logarithm was used. Here we use the theory of sub-Gaussian random variables.

For initial conditions of the form (3), $\varepsilon$ a small positive number, we prove that the $\ell^2$ norm of the differences between true solutions and appropriately scaled solutions to the wave equation is almost surely $O(\varepsilon^{-1-\sigma})$, where $\sigma$ is any small positive number. While such an absolute error diverges as $\varepsilon \to 0^+$, it happens that this is enough to establish an almost sure convergence of the macroscopic dynamics within the coarse-graining setting.

The article\(^15\) studies a similar problem on a one-dimensional (1D) lattice. There, the constants of elasticity also vary randomly and the system is studied in the relative coordinates. The so-called multiscale method of homogenization, a by-now classical tool with a long history in PDE for deriving effective equations (see Cioranescu and Donato\(^16\)), is employed. Our results should be contrasted with those in 1D. In McGinnis and Wright,\(^15\) it is shown that the coarse-graining converges at a rate of $\sqrt{\varepsilon \log \log(\varepsilon^{-1})}$, and this is thought to be sharp based on numerical evidence. Below, we show that in 2D, the rate of convergence is no slower than $\varepsilon^{1-\sigma}$ when the masses are i.i.d. However, the rate of convergence is different for layered masses, that is, where masses are only random with respect to $j_1$ and constant with respect to $j_2$. We explore this in Section 7.2 and in the numerical experiments in Section 9. In such a case, we show that the convergence rate is comparable to the rate in the 1D lattice. We do not believe our estimates are sharp like those in the 1D setting because the analysis requires the use of a cutoff function, which demands slightly greater regularity and algebraic decay of the initial conditions. Still, numerical evidence suggests they are probably close to being sharp and likely only off by a logarithm and a vanishing power of $\varepsilon$ introduced with the cutoff.

Although large parts of the formal derivation are the same in 2D as they are in 1D, difficulty arises in the fact that, as far as we can tell, in 2D no known explicit formulation of a solution for (21) exists. It can, however, be solved on a finite domain. Therefore, a cutoff function involving $\varepsilon$ is introduced, and although it eventually disappears from the final Theorems 3 and 4, several new arguments need to be made, most important of which are almost sure estimates for solutions of (35) involving tails of sequences of sub-Gaussian random variables. In Section 2, we derive the effective equations. The core probability theory is in Section 3. A common energy estimate, akin
to those found in Chirilus-Bruckner et al., Gaison et al., and Schneider and Wayne is given in Section 4. These help bound the error in terms of the “residuals” defined in (4). Elementary theory of the energy of the effective wave equation, such as that in Evans, is given in Section 5. This theory is used to derive estimates of various norms needed in the following section, Section 6. Finally, Sections 7 and 8 contain the main estimate and convergence result. Section 7 also contains some discussion around the complications arising from including random springs in (1) as well as to what extent our methods may be extended to 3D. The appendix contains technical proofs of several statements that are believable enough to be skipped in a first read through.

2 DERIVING THE EFFECTIVE EQUATIONS

In this section, we homogenize the equation in a similar manner to what is done in McGinnis and Wright. Doing so we derive the effective wave equation which approximates (1). The arguments are purely formal, so, for example, the reader should be skeptical that \( O(\varepsilon^2) \) terms are really as small as they appear. It is not until the subsequent sections that we make the approximation rigorous.

2.1 Expansions

First, we define the “residual,” which, as we will discover in Section 4, quantifies how close some function is to a true solution. For any function \( \tilde{u}(j, t) \), put

\[
\text{Res} \tilde{u}(j, t) = m(j)\ddot{\tilde{u}}(j, t) - \Delta \tilde{u}(j, t).
\]

We look for an approximate solution to (1) of the form

\[
\tilde{u}(j, t) = \varepsilon^{-1} U_0(j, \varepsilon j, \varepsilon t) + \varepsilon U_2(j, \varepsilon j, \varepsilon t),
\]

where, for \( i \in \{1, 2\} \),

\[
U_i : \mathbb{Z}^2 \times \mathbb{R}^2 \times \mathbb{R} \to \mathbb{R}.
\]

\( U_2 \) is necessary in the ansatz because, as we see in the sequel, it allows us to bound the error resulting from the random inhomogeneity of the masses. We keep track of the arguments of \( U_i \) by letting \( X = \varepsilon j \) and \( \tau = \varepsilon t \), where \( X = (X_1, X_2) \in \mathbb{R}^2 \). The so-called macroscopic variable \( X \) helps us look for approximate solutions which vary slowly relative to the lattice spacing.

The goal of our derivation is to show that \( U_0 \) solves a constant coefficient wave equation in the macroscopic variables \( X \) and \( \tau \), which we call the effective wave equation.

We do this by finding the \( U_0 \) and \( U_2 \), which make the residual formally small, by plugging in (5) into (1). Thus we need to understand how \( \Delta \) acts on functions of the form \( U(j, \varepsilon j) \). Using the definition of \( \Delta \), we have

\[
\Delta U(j, \varepsilon j) = -4U(j, \varepsilon j) + \sum_{i=1}^{2} U(j + e_i, \varepsilon j + \varepsilon e_i) + U(j - e_i, \varepsilon j - \varepsilon e_i).
\]
Taylor expanding the terms in the sum in $\varepsilon$, we find that

$$\Delta U = \tilde{\Delta}_0 U + \varepsilon \tilde{\Delta}_1 U + \varepsilon^2 \tilde{\Delta}_2 U + O(\varepsilon^3),$$

where we define

$$\tilde{\Delta}_0 U(j, X) := -4U(j, X) + \sum_{i=1}^{2} U(j + e_i, X) + U(j - e_i, X),$$

$$\tilde{\Delta}_1 U(j, X) := \sum_{i=1}^{2} \partial X_i U(j + e_i, X) - \partial X_i U(j - e_i, X),$$

and

$$\tilde{\Delta}_2 U(j, X) := \frac{1}{2} \sum_{i=1}^{2} \partial X_i X_i U(j + e_i, X) + \partial X_i X_i U(j - e_i, X).$$

Now we calculate $\text{Res}\tilde{u}$ using (8). Grouping terms by powers of $\varepsilon$ yields

$$\text{Res}\tilde{u}(j, X, \tau) = -\varepsilon^{-1} \tilde{\Delta}_0 U_0(j, X, \tau) - \tilde{\Delta}_1 U_0(j, X, \tau) + \varepsilon \left( m(j) \partial \tau U_0(j, X, \tau) - \tilde{\Delta}_0 U_2(j, X, \tau) - \tilde{\Delta}_2 U_0(j, X, \tau) \right) + O(\varepsilon^2).$$

Since we want $\text{Res}\tilde{u}$ to be small, we solve for $U_0$ and $U_2$ so that formally the residual is $O(\varepsilon^2)$.

### 2.1.1 Solving the $O(\varepsilon^{-1})$ equation

We start by setting the $O(\varepsilon^{-1})$ term to 0 and discovering the implications. We have

$$\tilde{\Delta}_0 U_0(j, X, \tau) = 0.$$

If we demand $U_0$ be bounded in $j$, we can only meet (13) by putting

$$U_0(j, X, \tau) = U_0(X, \tau),$$

that is, $U_0$ does not depend on $j$, the microscopic variable. Using the definitions of $\tilde{\Delta}_1$ and $\tilde{\Delta}_2$, (14) implies

$$\tilde{\Delta}_1 U_0 = 0$$

and that

$$\tilde{\Delta}_2 U_0 = \Delta_X U_0 := \sum_{i=1}^{2} \partial X_i X_i U_0.$$
Note that $\Delta_X$ is the usual continuous Laplacian. From (15) and (16), the residual in (12) simplifies to

$$\text{Res} \tilde{u}(j, t) = \varepsilon m(j) \partial_{\tau \tau} U_0(X, \tau) - \varepsilon \Delta_X U_0(X, \tau) - \varepsilon \tilde{\Delta}_0 U_2(j, X, \tau) + O(\varepsilon^2).$$

(17)

Note, the $O(1)$ term has vanished.

### 2.1.2 Solving the $O(\varepsilon)$ equation

Now to make the $O(\varepsilon)$ terms vanish from the residual in (17), we need to solve

$$m(j) \partial_{\tau \tau} U_0(X, \tau) - \Delta_X U_0(X, \tau) = \tilde{\Delta}_0 U_2(j, X, \tau).$$

(18)

Let $z(j) := m(j) - \bar{m}$, and $\bar{m} := \mathbb{E}[m(j)]$, where $\mathbb{E}$ is the expected value with respect to the probability measure of the i.i.d. lattice of masses. One way to solve (18) is to write it as

$$\bar{m} \partial_{\tau \tau} U_0(X, \tau) - \Delta_X U_0(X, \tau) = \tilde{\Delta}_0 U_2(j, X, \tau) - z(j) \partial_{\tau \tau} U_0(X, \tau),$$

(19)

and then pick $U_0$ and $U_2$ to force the left-hand side and right-hand side to vanish independently. From the left-hand side, we find that $U_0$ solves the wave equation

$$\bar{m} \partial_{\tau \tau} U_0(X, \tau) - \Delta_X U_0(X, \tau) = 0.$$ (20)

This is the effective wave equation that, we later show, approximates the dynamics of (1). In this case, the constant wave speed, $\frac{1}{\sqrt{\bar{m}}}$, was guessed. A more probabilistic derivation is given in Subsection 2.2.

Solving for $U_2$ on the right-hand side of (19) would be easy if we had $\chi : \mathbb{Z}^2 \to \mathbb{R}$ s.t.

$$\Delta \chi(j) = z(j).$$

(21)

If such a $\chi$ were to exist, we would obtain

$$U_2(j, X, \tau) = \chi(j) \partial_{\tau \tau} U_0(X, \tau).$$

(22)

From this, the approximate solution would be

$$\bar{u}(j, t) = \varepsilon^{-1} U_0(\varepsilon j, \varepsilon t) + \varepsilon \chi(j) \partial_{\tau \tau} U_0(\varepsilon j, \varepsilon t).$$

(23)

This is how the 1D problem proceeds in McGinnis and Wright. However, we do not even know how to solve for such a $\chi$ in 2D or even obtain information on its asymptotics. Thus our formal derivation has been hindered for the moment.

We briefly preview how we can circumnavigate the issue. Instead of solving (21), where the support of $z$ is unbounded, we apply a cutoff to $z$ so that its support is on a finite domain. This domain needs to depend upon $\varepsilon$ since we are interested in the solution at times less than or equal $T \varepsilon^{-1}$. We then solve (21) with the cutoff in place and call the solution $\chi_{R_\varepsilon}$. Our claim is that, despite having to use $\chi_{R_\varepsilon}$ in our approximate solution, the residual still turns out to be nearly as
small as we could hope for. Much of the latter parts of the paper are dedicated to showing the many technical issues that arise due to introducing a cutoff. However, the main novelty of this paper is obtaining almost sure estimates on $\chi_{R, \varepsilon}$, so we show this sooner in Section 3.1.

Going forward, we set out to show the following informal proposition.

**Proposition 1.** For appropriately scaled and smooth initial conditions, like those in (3), and $\varepsilon$ small, the dynamics of (1) may be approximated by

$$
\tilde{u}(j, t) := \varepsilon^{-1}U(\varepsilon j, \varepsilon t) + \varepsilon \chi_{R, \varepsilon}(j) \partial_{\tau\tau} U(\varepsilon j, \varepsilon t),
$$

(24)

where $U$ solves

$$
mU_{\tau\tau} = \Delta_X U.
$$

(25)

Via $\tilde{u}$, one may show

$$
\hat{u} := \varepsilon^{-1}U
$$

(26)

approximates (1) just as well.

**Remark 1.** Our main theorems fill in the correct details, including questions about initial conditions.

Eventually, we are able to show that the first term in $\tilde{u}$, $\varepsilon^{-1}U$, is relatively large compared to $\varepsilon \chi_{R, \varepsilon}(j) \partial_{\tau\tau} U(X, \tau)$. Since $U$ solves the wave equation (25) depending on the mean of the masses, we are able to say the macroscopic dynamics of the lattice are approximated by an effective wave equation. In Section 3.1, we lay out the details of $\chi_{R, \varepsilon}$, but first we make an aside, which explains how we may find $U_0$ and $U_2$ from (18) with a more probabilistic argument.

### 2.2 Averaging

Another way to derive $U_0$ and $U_2$ from (18) is by averaging. It should be noted that the remainder of the section is purely expository and none of our future arguments hinge on the following derivation, but we also take this time to introduce some notation.

If $U_2$ grows too quickly in $j$, its dynamics play as large of a role as those of $U_0$. Thus we start by making an ad hoc assumption on $U_2$. The assumption captures how quickly we can allow $U_2$ to grow so as to still be relatively small compared to $U_0$. We require that for all $X$ and $\tau$

$$
\lim_{|j|_\infty \to \infty} \frac{U_2(j, X, \tau)}{|j|_\infty^2} = 0,
$$

(27)

where

$$
|j|_\infty := \max\{|j_1|, |j_2| \mid j = (j_1, j_2)\}.
$$

(28)

We have introduced the $\infty$ norm here for technical reasons. Henceforth, $\mathbf{0}$ is the vector $(0,0)$. We define disks of radius $R$ centered at $j_0$

$$
D(j_0, R) := \{j \in \mathbb{Z}^2 \mid |j - j_0|_\infty \leq R\}.
$$

(29)
We define the boundary of these disks, which become important to us in the next section, by

$$\delta D(j_0, R) := \{ j \mid |j - j_0|_\infty = R \}. \quad (30)$$

For any set $S \subset \mathbb{Z}^2$, $|S|$ means the number of elements in the set.

The assumption (27) gives a generous range for the possible growth of $U_2$ in $j$. From the assumption, one may argue that the spatial average of $\tilde{U}_0 U_2$ is $0$, that is,

$$\lim_{R \to \infty} \frac{1}{|D(0, R)|} \sum_{j \in D(0, R)} \tilde{U}_0 U_2(j, X, \tau) = 0 \quad (31)$$

for all $X$ and $\tau$. Since the right-hand side of (18) has average $0$, it is true for the left-hand side as well, so

$$\lim_{R \to \infty} \frac{1}{|D(0, R)|} \sum_{j \in D(0, R)} m(j) \partial_{\tau^2} U_0(X, \tau) - \Delta X U_0(X, \tau) = 0. \quad (32)$$

On the other hand, one may apply the law of large numbers instead to obtain

$$\lim_{R \to \infty} \frac{1}{|D(0, R)|} \sum_{j \in D(0, R)} m(j) \partial_{\tau^2} U_0(X, \tau) - \Delta X U_0(X, \tau) = \bar{m} \partial_{\tau^2} U_0(X, \tau) - \Delta X U_0(X, \tau). \quad (33)$$

Combining (32) with (33) yields the same wave equation as seen on the left-hand side of (19),

$$\bar{m} \partial_{\tau^2} U_0(X, \tau) - \Delta X U_0(X, \tau) = 0, \quad (34)$$

where we have recovered the effective wave speed from the law of large numbers.

### 3 Probabilistic Estimates on $\chi_R$

In this section, we write down a solution for $\chi_R$ using the fundamental solution for the lattice Laplacian, and then we derive almost sure estimates for $\chi_{R_\epsilon}$ in $R_\epsilon$ and $j$. This section is largely self-contained outside of some notation from Section 2.2 because we think the idea could be useful in other contexts. Therefore, we are going to drop $\epsilon$ from $R_\epsilon$, since $\epsilon$ is a parameter relevant to our specific problem. However, estimates proven here are informed by the terms that show up in the full calculation of the residual in Section 4.2.

#### 3.1 Poisson problem with random source

We find $\chi_R$ which solves

$$\Delta \chi_R(j) = 1_{D(0, R)}(j) z(j) := \begin{cases} z(j) & j \in D(0, R) \\ 0 & j \not\in D(0, R). \end{cases} \quad (35)$$
Here \( \mathbb{1}_S \) is the indicator function for a set \( S \). Recall that the \( z(j) \) are independent, mean 0 random variables contained almost surely in the interval \([a - \bar{m}, b - \bar{m}]\). Whether or not \( z(j) \) are identically distributed is irrelevant to us.

Let \( \iota : \mathbb{Z}^2 \to \{0, 1\} \) s.t.

\[
\iota(j) = \begin{cases} 
1 & j = 0 \\
0 & j \neq 0
\end{cases}
\]  

(36)

and \( \varphi : \mathbb{Z}^2 \to \mathbb{R} \) s.t.

\[
\Delta \varphi(j) = \iota(j).
\]  

(37)

Then \( \varphi \) is the fundamental solution. It is known that \( \varphi(0) = 0 \), and, for \( j \neq 0 \),

\[
\varphi(j) = \frac{1}{2\pi} \log |j| + C_0 + O(|j|^{-2}), \ |j| \to \infty.
\]  

(38)

Here we take \( |j| := \sqrt{j_1^2 + j_2^2} \). The proof of the existence and uniqueness of \( \varphi \) can be found in Duffin.\(^{21}\) In (38), log may be replaced with \( \log^+ \). It is defined by \( \log^+(x) = \max\{0, \log(x)\} \) and \( \log^+(0) = 0 \), so all logs in the sequel should be thought of as \( \log^+ \), even if we neglect the plus sign.

Define

\[
\chi_R(j) := \sum_{k \in D(0, R)} \varphi(j - k)z(k).
\]  

(39)

It is easily verified that \( \chi_R \) solves (35). Finally, we use below that for any generic linear operator \( \mathcal{L} \) acting on functions of \( \mathbb{Z}^2 \), we have

\[
\mathcal{L}(\chi_R)(j) = \sum_{k \in D(0, R)} \mathcal{L}(\varphi)(j - k)z(k).
\]  

(40)

3.2 Estimates on \( \chi_R \)

We make use of Hoeffding’s inequality, the justification of which can be found in Massart.\(^{22}\) We state the theorem here for reference.

**Theorem 1** (Hoeffding’s inequality \(^{27}\)). Let \( q_1, q_2, \ldots, q_n \) be independent random variables such that \( q_i \in [a_i, b_i] \) almost surely for all \( i \leq n \). Let

\[
S := \sum_{i=1}^n q_i - \mathbb{E}[q_i].
\]  

(41)

Then for any positive \( s \),

\[
P(S \geq s) \leq \exp \left( \frac{-2s^2}{\sum_{i=1}^n (a_i - b_i)^2} \right).
\]  

(42)
Recall that $z(k) \in [a - \bar{m}, b - \bar{m}]$ almost surely and is mean 0. Therefore, $z(k)\mathcal{L}(\varphi)(j - k) \in \mathcal{L}(\varphi)(j - k)(a - \bar{m}), \mathcal{L}(\varphi)(j - k)(b - \bar{m})$ and is mean 0. (It might be that $\mathcal{L}(\varphi)(j - k)$ is negative in which case the bounds of the interval are flipped.)

Define

$$\|\mathcal{L}\varphi\|_{D(j,R)}^2 := \sum_{k \in D(j,R)} (\mathcal{L}(\varphi)(k))^2 = \sum_{k \in D(0,R)} (\mathcal{L}(\varphi)(j - k))^2. \quad (43)$$

Application of Hoeffding’s inequality to $\mathcal{L}(\chi_R)$ and $-\mathcal{L}(\chi_R)$ as defined in (40) yields

$$P(|\mathcal{L}(\chi_R)(j)| \geq s) \leq 2 \exp \left( \frac{-2s^2}{(a - b)^2 \|\mathcal{L}\varphi\|_{D(j,R)}^2} \right). \quad (44)$$

The inequality (44) is telling us that each $\chi_R(j)$ is a sub-Gaussian random variable with a proxy variance dependent on $\|\mathcal{L}\varphi\|_{D(j,R)}^2$. It is thus possible to control the size of every $\chi_R(j)$ in $R$ and $j$ simultaneously and almost surely using the proxy variance. We now show how.

We wish to use (44) in applying a Borel–Cantelli lemma. To do so, we order the $\chi_R(j)$ into a sequence and restrict $R$ to take on only values $\mathbb{N}^+ / \{1\}$ for technical reasons. We map $R$ and $j$ to a set of indices using a bijection $I : \mathbb{Z}^2 \times \mathbb{N}^+ / \{1\} \rightarrow \mathbb{N}^+$, which satisfies the inequality

$$I(j, R) \leq C \max\{R^3, |j|^3\}. \quad (45)$$

To see if such a bijection is possible, consider the sets

$$P_m = \{(j, R) | \max\{|j|, R\} = m, R \geq 2\}, \quad (46)$$

where $m \geq 2$ is a positive integer. It is easy to see that we can bound the size of such sets with

$$|P_m| \leq C m^2, \quad (47)$$

where $C$ is a constant independent of $m$. Let $\mathcal{M}_m = |\bigcup_{k=2}^m P_k|$. Pick any bijection $I_2$ such that $I_2 : P_2 \rightarrow \{1, 2, \ldots, |P_2|\}$. For $m \geq 3$, we pick any bijection $I_m$ such that $I_m : P_m \rightarrow \{\mathcal{M}_{m-1} + 1, \mathcal{M}_{m-1} + 2, \ldots, \mathcal{M}_{m-1} + |P_m|\}$. Now define $I$ by setting $I(j, R)$ equal to $I_m(j, R)$ whenever $\max\{|j|, R\} = m$. By construction, this defines $I(j, R)$ for all $j$ and $R \geq 2$. If $j$ and $R$ are such that $\max\{|j|, R\} = m$, then $I(j, R) \leq \mathcal{M}_{m-1} + |P_m|$. Since $\mathcal{M}_{m-1} = \sum_{k=2}^{m-1} |P_k|, I(j, R) \leq C m^3$, which establishes (45).

Now that we can identify each $(j, R)$ with a natural number $n$. Let

$$\mathcal{L}(\chi)(n) := \mathcal{L}(\chi_R)(j) \text{ and } \|\mathcal{L}\varphi\|_{D(n)}^2 := \|\mathcal{L}\varphi\|_{D(j,R)}^2 \quad (48)$$

when $(j, R) = I^{-1}(n)$. One sees we still have (44) as

$$P(|\mathcal{L}(\chi)(n)| \geq s) \leq 2 \exp \left( \frac{-2s^2}{(a - b)^2 \|\mathcal{L}\varphi\|_{D(n)}^2} \right). \quad (49)$$
Using the Borel–Cantelli lemma, which one may find in Durrett,\textsuperscript{23} we have that for all but finitely many $n$,

\[
|\mathcal{L}(\chi)(n)| \leq \sqrt{(a - b)^2 \|\mathcal{L}\varphi\|_{D(n)}^2 \log(n)}
\]

almost surely. Together, (50) and (45) imply that there exists $C_\omega$ almost surely s.t. for all $\mathbf{j}$ and $R$

\[
|\mathcal{L}(\chi_R)(\mathbf{j})| \leq C_\omega \sqrt{\|\mathcal{L}\varphi\|_{D[j,R]}^2 (\log(|\mathbf{j}|) + \log(R))}.
\]

It is thus seen that the asymptotics of $\mathcal{L}(\chi_R)(\mathbf{j})$ could depend significantly on $\|\mathcal{L}\varphi\|_{D[j,R]}$, so we estimate it for some special operators relevant to the problem in this paper.

Some simple operators we need to consider are the identity and shift operators

\[
S^\pm_\mathbf{i}(\xi)(\mathbf{j}) := \xi(\mathbf{j} \pm \mathbf{e}_\mathbf{i}).
\]

For $\varphi$ given by (38), we find that there exists a constant $C$ s.t.

\[
\max\left\{\left\|S^\pm_\mathbf{i} \varphi\right\|_{D[j,R]}^2, \|\varphi\|_{D[j,R]}^2\right\} \leq CR^2 \log(|\mathbf{j}| + R)^2.
\]

Another linear operator we need to consider is the center difference, $\delta_i$, which we define as

\[
\delta_i(\varphi)(\mathbf{j}) := \varphi(\mathbf{j} + \mathbf{e}_i) - \varphi(\mathbf{j} - \mathbf{e}_i).
\]

From Lemma A3, there exists another constant $C$ s.t.

\[
\|\delta_i \varphi\|_{D[j,R]}^2 \leq C \log(|\mathbf{j}| + R).
\]

Combining (51) with (55) and (53), we get the concluding theorem of this section.

**Theorem 2.** Let $\{z(\mathbf{j})\}_{\mathbf{j} \in \mathbb{Z}^2}$ be independent, mean zero random variables with $z(\mathbf{j}) \in [a - \bar{m}, b - \bar{m}]$ a.s. Let $R \geq 2$ be any real number, and $\chi_R(\mathbf{j})$ be as defined in (39) with $\varphi$ defined by (37). Then, there almost surely exists a constant $C_\omega$ s.t.

\[
\max\{|\chi_R(\mathbf{j})|, |S^\pm_\mathbf{i}(\chi_R)(\mathbf{j})|\} \leq C_\omega R \left(\log(|\mathbf{j}|)^{\frac{3}{2}} + \log(R)^{\frac{3}{2}}\right),
\]

and

\[
|\delta_i(\chi_R)(\mathbf{j})| \leq C_\omega (\log(|\mathbf{j}|) + \log(R)).
\]

**Remark 2.** Note that the definition in (39) implies $\chi_R = \chi_{\lfloor R \rfloor}$, and thus the theorem holds for all real $R \geq 2$, even though $R$ was originally a positive integer in the bijection in (45).

**Remark 3.** The $\omega$ indicates that the constant depends on the realization of the masses.
Proof. Insert (53) and (55) into (51). After sweeping any constants into $C_\alpha$, one has
\[
\max\{|X_R(j)|, |S^\pm_i(X_R)(j)|\} \leq C_\alpha R \log(|j| + R) \sqrt{\log(|j|) + \log(R)}
\] (58)
and
\[
|\delta_i(X_R)(j)| \leq C_\alpha \sqrt{\log(|j| + R)(\log(|j|) + \log(R))}.
\] (59)

An elementary argument in Lemma A1 yields
\[
\max\{|X_R(j)|, |S^\pm_i(X_R)(j)|\} \leq C_\alpha R (\log(|j|) + \log(R))^{3/2}
\] (60)
and
\[
|\delta_i(X_R)(j)| \leq C_\alpha (\log(|j|) + \log(R)).
\] (61)

A typical argument using convexity of $(\cdot)^{3/2}$ and sweeping constants into $C_\alpha$ now gives the result.

We store these estimates for future use, turning now to the more deterministic parts of the argument.

4 | LATTICE ENERGY ARGUMENT

Next, we follow a process similar to that laid out in the 1D case in McGinnis and Wright, which shows how the residual may be used to estimate the approximation error. Then, we calculate the residual from our approximate solution. The main technical difference in our argument comes from the use of the cutoff function in $X_{R_\epsilon}$, which only appears when we calculate the residual.

4.1 | Error estimate

We must first define the “approximation error,” which is
\[
\xi(j, t) := \tilde{u}(j, t) - u(j, t).
\] (62)

Using that $u$ is a solution to (1), we find
\[
m(j)\tilde{\xi}(j, t) = \Delta\xi(j, t) + \text{Res} \tilde{u}(j, t),
\] (63)
where $\text{Res} \tilde{u}$, defined in (4), is simply $m\tilde{u}_{tt} - \Delta \tilde{u}$.

It is notationally useful to define some operators similar to $\delta_i$ given in (54). These are the two (forward and backward) partial difference operators as well as the second partial difference
operators
\[
\delta^\pm_\epsilon(\xi) (j) := \pm (\xi(j + e_i) - \xi(j))
\]
\[
\Delta_\epsilon \xi := \delta^+_\epsilon \delta^-_\epsilon \xi.
\]

The energy of these error terms is given by
\[
H(t) := \frac{1}{2} \sum_{j \in \mathbb{Z}^2} m(j) \dot{\xi}(j, t)^2 + (\delta^+_1 \xi(j, t))^2 + (\delta^+_2 \xi(j, t))^2.
\]

An argument made by differentiating \(H\) gives the inequality,
\[
\sup_{|t| \leq \varepsilon^{-1} T} \left\| \dot{\xi}(-, t), \delta^+_1 \xi(-, t), \delta^+_2 \xi(-, t) \right\|_{\ell^2} \leq C \left\| \dot{\xi}(-, 0), \delta^+_1 \xi(-, 0), \delta^+_2 \xi(-, 0) \right\|_{\ell^2} + \varepsilon^{-1} \sup_{|t| \leq \varepsilon^{-1} T} \| \text{Res} \tilde{u}(-, t) \|_{\ell^2}.
\]

Since the argument follows nearly identical steps as its 1D counterpart found in McGinnis and Wright,\textsuperscript{15} we do not repeat it here. The inequality does not depend on the actual form of the approximate solution so the use of a cutoff does not affect this part of the argument. The only two ingredients important to our setup that go into the argument are that the approximation should be valid for \(|t| \leq T \varepsilon^{-1}\), and that \(m(j) \in [a, b] \subset \mathbb{R}^+\) almost surely, which is also the case for the 1D argument.

Note, the first term on the right-hand side of the inequality (66) quantifies the discrepancy between the initial condition of the true solution and the approximate solution. The issue of initial conditions is a bit technical given that our approximate solution has the form (24). We elaborate on this point further in Section 5 and eventually show in Theorem 3 that
\[
\left\| \dot{\xi}(-, 0), \delta^+_1 \xi(-, 0), \delta^+_2 \xi(-, 0) \right\|_{\ell^2} \sim \varepsilon^{-1} CT \sup_{|t| \leq \varepsilon^{-1} T} \| \text{Res} \tilde{u}(-, t) \|_{\ell^2},
\]
where \(\sim\) means they are the same size in \(\varepsilon\). We refer to the left-hand side of (67) as \(\text{Res}_0\).

We want to obtain an inequality for \(\xi\) as well because we wish to know how close \(\tilde{u}\) is to \(u\). In the 1D case, the result is written in terms of the velocities and the relative displacements, defined as follows:
\[
p(j, t) := \dot{u}(j, t) \text{ and } r_i(j, t) := \delta^+_i u(j, t),
\]
but this can be extended to positions in the same way we do here. Given we have analogous definitions of \(\tilde{p}\) and \(\tilde{r}_i\), we have so far the following lemma.

**Lemma 1.** For some constant \(C\) dependent on \(T, a, \) and \(b,\)
\[
\sup_{|t| \leq \varepsilon^{-1} T} \| p - \tilde{p}, r_1 - \tilde{r}_1, r_2 - \tilde{r}_2 \|_{\ell^2} \leq C \left( \text{Res}_0 + \varepsilon^{-1} \sup_{|t| \leq \varepsilon^{-1} T} \| \text{Res} \tilde{u} \|_{\ell^2} \right),
\]
whereby straightforward integration of $\bar{p}$ and $p$ and use of the triangle inequality yields

$$
\sup_{|t| \leq \varepsilon^{-1}T} \|u - \bar{u}\|_{\ell^2} \leq \varepsilon^{-1}C \left( \text{Res}_0 + \varepsilon^{-1} \sup_{|t| \leq \varepsilon^{-1}T} \|\text{Res} \bar{u}\|_{\ell^2} \right).
$$

Here $\text{Res}_0 := \|p(0) - \bar{p}(0), r_1(0) - \bar{r}_1(0), r_2(0) - \bar{r}_2(0)\|_{\ell^2}$.

Thus we have shown the $\ell^2$ discrepancy between the true and approximate solutions is controlled by the residual. Formally, in Section 2, the residual is $O(\varepsilon^2)$, suggesting the error is $O(1)$ and relatively small compared to $u$. However, as was discovered in the 1D case, we cannot trust the formal argument to give the correct size of the error. Thus we must calculate the residual exactly.

### 4.2 Calculating the residual

Now we calculate $\text{Res} \bar{u}$ for our approximate solution given in (24) as

$$
\bar{u}(j, t) = \varepsilon^{-1}U(\varepsilon j, \tau) + \varepsilon \chi \mathcal{R}_\varepsilon U_{\tau\tau}(\varepsilon j, \tau).
$$

Recall that $X = \varepsilon j$ and $\tau = \varepsilon t$. Recall the shifts, center differences, and partial differences defined in (52), (54), and (64), respectively. With these, we calculate, by trying to emulate the product rule for ordinary derivatives, that

$$
\Delta \bar{u} = \varepsilon^{-1} \Delta U + \varepsilon \Delta (\chi \mathcal{R}_\varepsilon)U_{\tau\tau} + \varepsilon \sum_{i=1}^{2} \delta_i(\chi \mathcal{R}_\varepsilon)\delta_i^{-}(U_{\tau\tau}) + S_i^+(\chi \mathcal{R}_\varepsilon)\Delta_i(U_{\tau\tau}).
$$

In a moment, we see how the cutoff function we used earlier causes an additional term to appear in the residual. First, we need some notation to write down this term. For $S \subset \mathbb{Z}^2$, we denote the indicator of $S$ as $\mathbb{I}_S$. The complement of $D(0, R_\varepsilon)$ is denoted by $D(0, R_\varepsilon)^c$.

Using (35), that is, the fact that $\Delta X = z \mathbb{I}_{D(0, R_\varepsilon)}$, in (72), we obtain

$$
\Delta \bar{u} = \varepsilon^{-1} \Delta U - \varepsilon \Delta X U + \varepsilon \sum_{i=1}^{2} \delta_i(\chi \mathcal{R}_\varepsilon)\delta_i^{-}(U_{\tau\tau}) + S_i^+(\chi \mathcal{R}_\varepsilon)\Delta_i(U_{\tau\tau}).
$$

Recall that $z = m - \bar{m}$ and that $\bar{m}U_{\tau\tau} = \Delta X U$. With these, (73) becomes

$$
\Delta \bar{u} = \varepsilon^{-1} \Delta U - \varepsilon \Delta X U + \varepsilon \left( mU_{\tau\tau} - z \mathbb{I}_{D(0, R_\varepsilon)^c}U_{\tau\tau} + \sum_{i=1}^{2} \delta_i(\chi \mathcal{R}_\varepsilon)\delta_i^{-}(U_{\tau\tau}) + S_i^+(\chi \mathcal{R}_\varepsilon)\Delta_i(U_{\tau\tau}) \right).
$$

Plugging our computation of $\Delta \bar{u}$ into $\text{Res} \bar{u} = m\bar{u}_{\tau\tau} - \Delta \bar{u}$ yields

$$
\begin{align*}
\text{Res} \bar{u} &= \varepsilon^{-1} (\varepsilon^2 \Delta X U - \Delta U) + \varepsilon \left( z \mathbb{I}_{D(0, R_\varepsilon)^c}U_{\tau\tau} - \sum_{i=1}^{2} \delta_i(\chi \mathcal{R}_\varepsilon)\delta_i^{-}(U_{\tau\tau}) + S_i^+(\chi \mathcal{R}_\varepsilon)\Delta_i(U_{\tau\tau}) \right) \\
&\quad + \varepsilon^3 m\chi \mathcal{R}_\varepsilon U_{\tau\tau\tau\tau}.
\end{align*}
$$
Simple use of the triangle inequality yields several terms that we need to estimate in $\epsilon$. The norms in the following inequality are $\ell^2$ norms:

$$
\| \text{Res} \tilde{u} \| \leq \epsilon^{-1} \| \epsilon^2 \Delta_X U - \Delta U \| 
+ \epsilon \left( \left\| L_{D(0,R)} \chi_{R}^\epsilon \Delta_U \right\| + \sum_{i=1}^{2} \left\| S_i \chi_{R}^\epsilon \right\| + \left\| S_i^+ \chi_{R}^\epsilon \Delta_i(U) \right\| \right) 
+ \epsilon^3 \left\| m \chi_{R}^\epsilon \right\|.
$$

(76)

Since $U$ is a solution to a wave equation, it would be useful to us to convert these norms into norms of smooth functions on $\mathbb{R}^2$. Once converted, we can use properties of the wave equation to bound the norms in terms of the initial data. Controlling the residual is essential to proving our main theorem. Thus we are concerned with proving the following proposition throughout the next couple of sections.

**Proposition 2.** Suppose the $m(j)$ are i.i.d. random variables contained in some interval $[a, b] \subset \mathbb{R}^+$ almost surely. Let $\sigma$ be any small positive number. Then almost surely there exists a constant $C_\omega(T, a, b, \tilde{m})$ s.t.

$$
\sup_{|t| \leq \epsilon^{-1}T} \| \text{Res} \tilde{u} \| \leq \epsilon^{1-\sigma} \log(\epsilon^{-1-\sigma})^3 C_\omega \| \nabla \phi, \psi \|_{H^5}.
$$

(77)

The norm, $\| \nabla \phi, \psi \|_{H^5}$, is a norm on the initial conditions of $U(X, \tau)$ given in the next section. Crucially, it does not depend on $\epsilon$.

5. THE EFFECTIVE WAVE

Eventually, we want the $\ell^2$ norms in (75) to be bounded by some norm of the initial conditions, but importantly, this norm needs to be of functions on $\mathbb{R}^2$. Doing so will exhibit how exactly these $\ell^2$ norms depend on $\epsilon$. This process requires two steps. The first is to bound the $\ell^2$ norms by norms on smooth functions. This may be done assuming $U$ is smooth since $U$ is a function of $X = \epsilon j$. The second step is to bound the norm of $U$, by norms on the initial conditions. The main assumption we are making is that the appropriate initial conditions for the lattice equation are equal to some smooth functions in $X$. To obtain the initial conditions for the lattice, we sample these smooth functions at $\epsilon j$. In this section, we specify the initial conditions for the lattice problem and work through the second step of bounding norms of $U$ as a smooth function by norms on the initial conditions.

5.1 Initial conditions

We specify the initial conditions for the lattice in (1). For smooth enough functions $\phi, \psi : \mathbb{R}^2 \to \mathbb{R}$, we suppose

$$
u(j, 0) = \epsilon^{-1} \phi(\epsilon j) \text{ and } \dot{\nu}(j, 0) = \psi(\epsilon j).
$$

(78)
Now it is apparent that the approximate solution, \( \tilde{u} \), we have derived above may not be able to take the initial conditions we prescribed for \( u \) and \( \hat{u} \) because \( \tilde{u}(j, t) = \varepsilon^{-1}U(j, \varepsilon t) + \varepsilon \chi(j)U_{\tau\tau}(j, \varepsilon t) \). However, recalling that \( \tilde{m}U_{\tau\tau} = \Delta_X U \), we still may specify

\[
U(X, 0) = \phi(X) \quad \text{and} \quad \partial_x U(X, 0) = \psi(X).
\]  

(79)

Thus at \( t = 0 \), we have

\[
\tilde{u}(j, 0) = \varepsilon^{-1}\phi(j) + \varepsilon \tilde{m}^{-1} \chi(j) \Delta_X \phi(j) \quad \text{and} \quad \hat{u}(j, 0) = \psi(j) + \tilde{m}^{-1} \varepsilon^2 \chi(j) \Delta_X \psi(j),
\]  

(80)

where \( \Delta_X \phi(j) \) is \( \Delta_X \phi(X) \) evaluated at \( j \). Even though the initial conditions of \( \tilde{u} \) and \( u \) do not match, they are not too large so (67) holds. Indeed, when \( t = 0 \),

\[
\| \delta^+ \tilde{u} - \delta^+_i u \|_{\ell_2} = \varepsilon m^{-1} \| \delta^+ (\chi \Delta_X \phi) \|_{\ell_2} \quad \text{and} \quad \| \hat{u} - \tilde{u} \|_{\ell_2} = \varepsilon^2 m^{-1} \| \chi \Delta_X \psi \|_{\ell_2},
\]  

(81)

and terms \( \| \delta^+ (\chi \Delta_X \phi) \|_{\ell_2} \) and \( \| \chi \Delta_X \psi \|_{\ell_2} \) are very similar to terms found in the residual (76) at \( t = 0 \) and may be bounded by the same kind of arguments we eventually use in Section 7. In a similar vein, if \( \tilde{u} \) is a good approximate solution, then \( \varepsilon^{-1}U(j, \varepsilon t) \) is just as good since \( \| \tilde{u} - \varepsilon^{-1}U \| = \| \varepsilon \chi U_{\tau\tau} \| \) is no larger in \( \varepsilon \) than the right-hand side of (70). This is again because \( \| \chi U_{\tau\tau} \| \) is similar to the largest terms found in the residual. We give this alternative approximate solution the name

\[
\hat{u}(j, t) := \varepsilon^{-1}U(j, \varepsilon t).
\]  

(82)

The approximate solution \( \hat{u} \) is the one in which we are actually interested because its evolution is described by a wave equation only. One consequence of the fact that \( \varepsilon \chi(j)U_{\tau\tau} \) is as large as the residual, is that \( \tilde{u} \) does not describe the dynamics any better than \( \hat{u} \), at least in the sense of approximation we are using. Therefore, our primary interest in \( \tilde{u} \) is only as a gateway to \( \hat{u} \). Now that we have given \( U \) initial conditions, we need to bound it in terms of these initial conditions. The norms in (76) have yet to be converted to a norm of smooth functions, but we can already see from terms like \( \| z D(0, R_\varepsilon)^p U_{\tau\tau} \| \) and \( \| m \chi R_{\varepsilon} U_{\tau\tau\tau\tau} \| \) that we need to consider weighted norms of derivatives of \( U \) and a norm of \( U \) restricted to the exterior of a disk. This is our next task, and it enforces regularity and algebraic decay of the initial conditions.

### 5.2 The energy

In order to bound the terms in (76), we need to control \( L^2 \) norms of partial derivatives of \( U \) in terms of the initial conditions. This can be done using arguments involving the energy of \( U \). Let \( c^2 = \tilde{m}^{-1} \). For \( V(X, \tau) : \mathbb{R}^2 \times \mathbb{R} \to \mathbb{R}^k \), define the \( k \)-dimensional vector of energies by

\[
E(V)(\tau) := \frac{1}{2} \int_{\mathbb{R}^2} \left( \partial_x V_i \right)^2 + c^2 | \nabla(V_i) |^2 dX.
\]  

(83)

\( V_i \) is the \( i \)th element of \( V \). Note that \( E(V) \) is a vector of energies and \( | E(V) | \) is the 2-norm of this vector. Finally, let \( D^n V \) be the \( n \)th total derivative of \( V \) with respect to \( X \). When \( V \) is scalar valued, \( DV = \nabla V \).
For $U$ satisfying the wave equation in (25) with initial conditions (79), it is well known that, for all $\tau$,

$$
\left| E(D^k \partial_t^j U)(\tau) \right| \leq C \left( \| \nabla \phi \|^2_{H^{i+k}} + \| \psi \|^2_{H^{i+k}} \right).
$$

(84)

(See, for instance, Craig\textsuperscript{24} or Evans.\textsuperscript{20})

Since the energy of derivatives of $U$ is equivalent to $L^2$ norms of derivatives of $U$, we have the following lemma.

**Lemma 2.** For $U$ satisfying $\bar{m} U_{\tau \tau} = \Delta_X U$ with smooth initial conditions given by (78), we have for all $\tau$ that

$$
\sum_{j=1}^{k} \left\| D^j \partial_t^j U \right\|^2_{L^2} \leq C \sum_{j=0}^{k-1} \left| E(D^j \partial_t^j U)(\tau) \right| \leq C \left( \| \nabla \phi \|^2_{H^{i+k-1}} + \| \psi \|^2_{H^{i+k-1}} \right).
$$

(85)

The smoothness of the initial conditions is dictated by the norm on the right-hand side of the inequality. The constant depends on $\bar{m}$, and we see that $\nabla \phi \in H^{i+k-1}$ and $\psi \in H^{i+k-1}$ are needed. This final inequality is essential to us in Section 6.

### 5.3 Weighted energy

In (76), there are terms where $U$ is weighted by $\chi_{R_\varepsilon}$ or an operator applied to $\chi_{R_\varepsilon}$. By (56), these versions of $\chi_{R_\varepsilon}$ may be bounded by functions involving logarithms in $j$. For now, assume we have dealt with the problem of converting $\log(|j|)$ into $\log(|X|)$. We can obviously find a smooth function $w(X) : \mathbb{R}^2 \to \mathbb{R}$ s.t.

$$
w(X) = \log(|X| + 1)^{\frac{3}{2}} + 1 \ \forall \ |X| \geq 1 \text{ and } w(X) \geq 1,
$$

(86)

and for all $n \in \mathbb{N}^+$, there exists a constant $W$ s.t.

$$
|D^n w| \leq W \ \text{ and } \ |D(w^2)| \leq W.
$$

(87)

With the same $V$ we used in defining the energy in (83), the “weighted energy” is given by

$$
E_w(V_i)(\tau) := \frac{1}{2} \int_{\mathbb{R}^2} w^2(\partial_t^j V_i)^2 + c^2 w^2 |\nabla(V_i)|^2 dX.
$$

(88)

The weight in the energy amplifies the solution as it travels away from the origin. This suggests the weighted energy may grow in time, and it is shown in Lemma A4, that for all $\tau$,

$$
\left| E_w(D^k \partial_t^j U)(\tau) \right| \leq (|\tau| + 1)C \left( \| \nabla \phi \|^2_{H^{i+k}} + \| \psi \|^2_{H^{i+k}} \right),
$$

(89)

where the norms $H^k_w$ on the right-hand side are weighted Sobolev norms. More specifically, we have

$$
\| \psi \|^2_{H^k_w} := \sum_{j=0}^{k} \| w D^j \psi \|_{L^2}.
$$

(90)
Assumptions (86) and (87) on $w$ also give us

\[ \left\| D^k (w \partial_t^i U) \right\|_{L^2} \leq C \sum_{j=0}^k \left\| wD^j \partial_t^i U \right\|_{L^2} \leq C \sum_{j=0}^k \sqrt{E_w (D^j \partial_t^{i-1} U)}. \tag{91} \]

From (89) and (91), we obtain the following lemma.

**Lemma 3.** Let $U$ satisfy $\bar{m} U_{\tau\tau} = \Delta_X U$ with smooth enough initial conditions given by (78). Then, for $i \geq 1$ and all $\tau$,

\[ \left\| w \partial_t^i U \right\|_{H^k} \leq C \sqrt{\left| \tau \right| + 1} \left( \left\| \nabla \phi \right\|_{H^{i+k}} + \left\| \psi \right\|_{H^{i+k}} \right). \tag{92} \]

This final inequality is also important to us in Section 6. We note, that for the inequality to hold, we enforce that $i \geq 1$ since if both $i$ and $k$ are 0, we have just $U$, which we cannot estimate in terms of the energy. This is not an issue, since all the terms in the residual that also contain $\chi_{R_c}$ have time derivatives.

### 5.4 Tail energy

In (76), there is one term which has its support outside a disk. Eventually, this term will be bounded by a norm on $U$ which treats $U$ as a smooth function on $\mathbb{R}^2$, and so we provide the definition of a ball in $\mathbb{R}^2$

\[ B(R) = \{ X \in \mathbb{R}^2 \mid |X| < R \}. \tag{93} \]

Letting $V$ be the same as we had in defining the energies above, we now define the “tail energy” of the function as

\[ \tilde{E}(V_j)(\tau) := \int_{B(\epsilon|\tau|+\epsilon^{-\sigma})} (\partial_t^i V_j)^2 + c^2 \left| \nabla (V_j) \right|^2 dX. \tag{94} \]

Here $\sigma$ is a small parameter that plays an important role. The radius of the ball is slightly larger than the light cone of the effective wave equation because of the addition of $\epsilon^{-\sigma}$. This allows us to extract an extra power of $\epsilon$ assuming the initial conditions decay fast enough, as we see below. Without this, the term arising from the cutoff in (76) would be too large.

Lemma A5 shows that, for $i \geq 1$,

\[ \left\| \tilde{E} (D^k \partial_t^i U)(\tau) \right\| \leq C \left( \left\| \nabla \phi \right\|_{H^{i+k-1}(\beta(\epsilon^{-\sigma})\cdot)}^2 + \left\| \psi \right\|_{H^{i+k-1}(\beta(\epsilon^{-\sigma})\cdot)}^2 \right), \tag{95} \]

where $C$ depends only on $\bar{m}$. Therefore, we have

\[ \left\| \partial_t^i U \right\|_{H^k(B(\epsilon\tau+\epsilon^{-\sigma})\cdot)} \leq C \left( \left\| \nabla \phi \right\|_{H^{i+k-1}(\beta(\epsilon^{-\sigma})\cdot)} + \left\| \psi \right\|_{H^{i+k-1}(\beta(\epsilon^{-\sigma})\cdot)} \right). \tag{96} \]

It follows from Lemma A6, that for any $f : \mathbb{R}^2 \to \mathbb{R}, \sigma > 0$,

\[ \left\| f \right\|_{H^k(\beta(\epsilon^{-\sigma})\cdot)} \leq \epsilon \left\| f \right\|_{H^k} : = \epsilon \sum_{j=0}^k \left( (1 + |j|)^{-\sigma} \right) \left\| D^j f \right\|_{L^2}. \tag{97} \]

Therefore, we obtain the following lemma.
Lemma 4. Let $U$ satisfy $\bar{m} U_{\tau \tau} = \Delta_X U$ with smooth enough initial conditions given by (78). Then, for $i \geq 1$ and all $\tau$,

$$
\left\| \partial^i \tau U \right\|_{H^k(B(\epsilon |\rho| + \epsilon^{-\sigma})^c)} \leq \epsilon C \left( \| \nabla \phi \|_{H^{i+k-1}} + \| \psi \|_{H^{i+k-1}} \right).
$$

(98)

We now use inequalities (85), (92), and (98) in the next section in estimating the terms in (76).

6 | RESIDUAL ESTIMATES

We bound the terms from left to right in (76) by the initial conditions given by (78). The smoothness or algebraic decay of the initial conditions required for each bound is dictated by the norm on the right-hand side of the inequality in the lemmas below. Each lemma requires us to first bound the $\ell^2$ norms by norms of smooth functions, that is, Sobolev norms. This is achieved by various technical lemmas in the appendix, but the idea is that, because of the parameter $\epsilon$, these are just like Riemann sums. Then, these Sobolev norms are bounded by initial conditions using the inequalities (84), (92), and (94) given in the previous section. For the terms where $\chi_{R_{\epsilon}}$ is present, we need the inequalities (56) and (57) found at the end of Section 3.

We begin by bounding the first-term norm in (76), which would be present regardless of how the masses are chosen, for example, even if the masses were all constant.

Lemma 5. For $U$ satisfying (25) with sufficiently smooth initial conditions given in (79), there exists a constant $C$ depending only on $\bar{m}$ s.t.

$$
\sup_{|t| \leq \epsilon^{-1} T} \epsilon^{-1} \left\| \Delta U(\epsilon \cdot, \epsilon t) - \epsilon^2 \Delta_X U(\epsilon \cdot, \epsilon t) \right\|_{\ell^2} \leq \epsilon^2 C \left( \| \nabla \phi \|_{H^5} + \| \psi \|_{H^5} \right).
$$

(99)

Proof. Recall the definition of $\Delta_i U$ in (64). For each $i = 1, 2$, by Taylor’s theorem with remainder, there exists a $\hat{j} = (\hat{j}_1, \hat{j}_2)$ with $\hat{j}_i \in [j_i - 1, j_i + 1]$ s.t

$$
\Delta_i U(\epsilon \hat{j}) - \epsilon^2 U_{X_iX_i}(\epsilon \hat{j}) = \epsilon^4 U_{X_iX_iX_iX_i}(\epsilon \hat{j}).
$$

(100)

By convexity of $(\cdot)^2$,

$$
(\Delta U(\epsilon \hat{j}) - \epsilon^2 \Delta_X(U)(\epsilon \hat{j}))^2 \leq 2 \sum_{i=1}^{2} (\Delta_i U(\epsilon \hat{j}) - \epsilon^2 U_{X_iX_i}(\epsilon \hat{j}))^2.
$$

(101)

Combining (100) and (101) yields

$$
(\Delta U(\epsilon \hat{j}) - \epsilon^2 \Delta_X(U)(\epsilon \hat{j}))^2 \leq \epsilon^8 2 \sum_{i=1}^{2} \left( U_{X_iX_iX_iX_i}(\epsilon \hat{j}) \right)^2.
$$

(102)

Summing over $\hat{j}$ and using Corollary A1, we find

$$
\left\| \Delta U(\epsilon \cdot, \epsilon t) - \epsilon^2 \Delta_X U(\epsilon \cdot, \epsilon t) \right\|_{\ell^2} \leq \epsilon^3 4 \sum_{i=1}^{2} \left\| U_{X_iX_iX_iX_i} \right\|_{H^2}.
$$

(103)
Now we apply (85) to get a constant $C$ which depends only on $\bar{m}$ s.t.

$$\|\Delta U(\varepsilon, \varepsilon t) - \varepsilon^2 \Delta_X U(\varepsilon, \varepsilon t)\|_{L^2} \leq \varepsilon^3 8C\left( \|\nabla \phi\|_{H^5} + \|\psi\|_{H^5} \right).$$

(104)

The cutoff plays an important role in bounding the remaining terms. Thus, it is necessary to finally define a sufficiently large radius. Let $c^2 = \bar{m}^{-1}$ as before, and let

$$R_\varepsilon := \frac{cT + \varepsilon - \sigma}{\varepsilon} + 1,$$

(105)

where $\sigma$ is any small positive number. The exact radius for the cutoff we have chosen is partly due to making some technical pieces to fit together more nicely. However, the basic idea is to choose a radius slightly larger than the light cone of the effective wave equation.

We now bound the second term appearing in (76). It contains the random variables $z$, but these do not play a role. More importantly, it is the term that arises due to the cutoff, so it could be regarded as artificial. It also requires algebraic decay of the initial conditions to bound. This decay is dictated by the $H^3_\sigma$ norm appearing in the following lemma.

**Lemma 6.** For $U$ satisfying (25) with sufficiently smooth initial conditions given by (79) and for any $\sigma > 0$, there exists a constant $C$ depending on $a, b$, and $\bar{m}$ s.t.

$$\sup_{|t| \leq \varepsilon^{-1}T} \varepsilon \|z\|_{D(D(0, R_\varepsilon))\cap H^2(B(\bar{R}(T))\cap \mathbb{R}^n)} \leq \varepsilon C\left( \|\nabla \phi\|_{H^5_\sigma} + \|\psi\|_{H^5_\sigma} \right).$$

(106)

**Proof.** First, recall that $z(j) \in [a - \bar{m}, b - \bar{m}]$. There exists a $C$ that depends on $a, b$ or $\bar{m}$ s.t.

$$\|z\|_{D(D(0, R_\varepsilon))\cap H^2(B(\bar{R}(\tau)))} \leq C\|z\|_{D(D(0, R_\varepsilon))\cap H^2(B(\bar{R}(\tau)))}.$$

(107)

Recall the definition of $R_\varepsilon$ in (105). Note that $R_\varepsilon$ has the form $\varepsilon^{-1} \bar{R}(T) + 1$, where $\bar{R}(\tau) := c\tau + \varepsilon^{-\sigma}$. According to Lemma A8,

$$\|D(0, R_\varepsilon)\|_{H^2(B(\bar{R}(\tau)))} \leq 2\varepsilon^{-1}\|U\|_{H^2(B(\bar{R}(\tau)))}.$$

(108)

Since $|\tau| \leq T$,

$$\|U\|_{H^2(B(\bar{R}(\tau)))} \leq \|U\|_{H^2(B(\bar{R}(T)))}.$$

(109)

It follows from (98) that

$$\|U\|_{H^2(B(\bar{R}(\tau)))} \leq \varepsilon C\left( \|\nabla \phi\|_{H^5_\sigma} + \|\psi\|_{H^5_\sigma} \right).$$

(110)

Stringing these inequalities together and taking sup yields the result.

Now we bound the terms involving $\chi_{R_\varepsilon}$. Crucially, we require the inequalities proven at the end of Section 3 to make them small enough.
Lemma 7. Let $U$ satisfy (25) with sufficiently smooth initial conditions given by (79) and $\varepsilon < 1$. There exists a constant $C_\omega$ a.s. s.t.

$$
\sup_{|t| \leq \varepsilon^{-1}T} \varepsilon \left\| \delta_i(\chi_{R\varepsilon}) \delta_i^-(U_{\tau\tau}) \right\|_{\ell^2} \leq \varepsilon \log(R\varepsilon) C_\omega \sqrt{T + 1} \left( \left\| \nabla \phi \right\|_{H^5_w} + \left\| \psi \right\|_{H^5_w} \right).
$$

(111)

Proof. Recall the bound given by (57). We, therefore, have that

$$
\left\| \delta_i(\chi_{R\varepsilon}) \delta_i^-(U_{\tau\tau}) \right\|_{\ell^2} \leq C_\omega \left( \left\| \log^+(\cdot) \delta_i^-(U_{\tau\tau}) \right\|_{\ell^2} + \log(R\varepsilon) \left\| \delta_i^-(U_{\tau\tau}) \right\|_{\ell^2} \right).
$$

(112)

Recalling the definition of $w$, the weight in (86), we have

$$
\log^+(\cdot) \leq \log(\varepsilon^{-1}) \log(\varepsilon \cdot + 1) \leq \log(R\varepsilon) w_\varepsilon(\cdot).
$$

(113)

where

$$
w_\varepsilon(\cdot) := w(\varepsilon \cdot).
$$

(114)

Then

$$
\left\| \delta_i^-(U_{\tau\tau}) \right\|_{\ell^2} \leq \left\| w_\varepsilon \delta_i^-(U_{\tau\tau}) \right\|_{\ell^2}
$$

and

$$
\left\| \log^+(\cdot) \delta_i^-(U_{\tau\tau}) \right\|_{\ell^2} \leq \log(R\varepsilon) \left\| w_\varepsilon \delta_i^-(U_{\tau\tau}) \right\|_{\ell^2}.
$$

(115)

Applying Corollary 11.7, we have

$$
\left\| w_\varepsilon \delta_i^-(U_{\tau\tau}) \right\|_{\ell^2} \leq \varepsilon^{-1} \left\| w(\cdot)(U_{\tau\tau}(\cdot) - U_{\tau\tau}(\cdot - \varepsilon e_i)) \right\|_{H^2}.
$$

(116)

Note that $\delta_i^-(U)(\cdot, \tau) := U(\cdot, \tau) - U(\cdot - \varepsilon e_i, \tau)$ solves the same wave equation $U$ does with initial conditions

$$
\delta_i^-(U)(X, 0) = \delta_i^-(\phi)(X) \text{ and } \delta_i^-(U)(X, 0) = \delta_i^-(\phi)(X).
$$

(117)

Hence, we can apply (92)

$$
\left\| w(\cdot)(U_{\tau\tau}(\cdot) - U_{\tau\tau}(\cdot - \varepsilon e_i)) \right\|_{H^2} \leq C \sqrt{|\tau| + 1} \left( \left\| \delta_i^-(\nabla \phi) \right\|_{H^5_w} + \left\| \delta_i^-(\psi) \right\|_{H^5_w} \right).
$$

(118)

By Lemma A9,

$$
\left\| w(\cdot)(U_{\tau\tau}(\cdot) - U_{\tau\tau}(\cdot - \varepsilon e_i)) \right\|_{H^2} \leq \varepsilon C \sqrt{|\tau| + 1} \left( \left\| \nabla \phi \right\|_{H^5_w} + \left\| \psi \right\|_{H^5_w} \right).
$$

(119)

Stringing the inequalities together and taking sup gives us the result.

Lemma 8. Let $U$ satisfy (25) with sufficiently smooth initial conditions given by (79) and $\varepsilon < 1$. There exists a constant $C_\omega$ a.s. s.t.

$$
\sup_{|t| \leq \varepsilon^{-1}T} \varepsilon \left\| \Delta_i^+(\chi_{R\varepsilon}) \Delta_i(U_{\tau\tau}) \right\|_{\ell^2} \leq \varepsilon^2 R\varepsilon \log(R\varepsilon)^3 \left( \left\| \nabla \phi \right\|_{H^5_w} + \left\| \psi \right\|_{H^5_w} + 1 \right).
$$

(120)
**Proof.** The proof is essentially the same as the proof of the previous theorem, but now we use Corollary A2 instead of Lemma A9.

**Lemma 9.** Let $U$ satisfy (25) with sufficiently smooth initial conditions given by (79) and $\epsilon < 1$. There exists a constant $C_\omega$ a.s. s.t.

$$
\sup_{|t| \leq \epsilon^{-1}T} \epsilon^3 \left\| m \chi_{R_\epsilon} U_{\tau \tau \tau} \right\|_{\ell^2} \leq \epsilon^2 R_\epsilon \log(R_\epsilon)^{3/2} C_\omega \sqrt{T + 1} \left( \| \nabla \phi \|_{H^5_w} + \| \psi \|_{H^5_w} \right).
$$

(121)

**Proof.** Since $m(j) \in [a, b]$, there exists a constant depending on $a$, $b$ or $m$ s.t.

$$
\left\| m \chi_{R_\epsilon} U_{\tau \tau \tau} \right\|_{\ell^2} \leq C \left\| \chi_{R_\epsilon} U_{\tau \tau \tau} \right\|_{\ell^2}.
$$

(122)

The steps are now very similar to those found in Lemma 7 but using the inequality in (56) instead of the one in (57).

7 | RESIDUAL BOUND AND DISCUSSION

Lemma 5 bounds a completely deterministic term that would appear no matter how the masses are chosen. Lemma 6 bounds a term that arises due to our use of a cutoff function. Recall we need to use a cutoff in order to work around solving (21). In the case where this can be solved, say, for example, where the masses vary periodically, then this term would not appear. The norm, $H_3^2$, in this bound is the largest norm. As $\sigma$ becomes smaller, this norm becomes larger. The term in Lemma 7 is the first term that must be bounded using probabilistic arguments. The constant $C_\omega$ exists almost surely but depends on the actual realization of masses. It, therefore, could be arbitrarily large since there is always a small probability that (50) holds only for $n$ extremely large. Thus $C_\omega$ may be worthy of statistical quantification in a follow-up work. Recalling the definition of $R_\epsilon$ in (105), the bound for the term in Lemma 8 is the dominant one in $\epsilon$. It is $O(\epsilon^{1-\sigma} \log(\epsilon^{-1-\sigma})^{3/2})$. This term also requires the most smoothness and decay of the initial conditions. The final term, bounded in Lemma 9, is also $O(\epsilon^{1-\sigma} \log(\epsilon^{-1-\sigma})^{3/2})$ for similar reasons. It may be conjectured that $\sigma$ here is artificial as it arises from our inability to solve (21). We could analyze more; for example, we could find the dependence of $T$, $m$, $a$ or $b$, but what we are most interested in is tracking $\epsilon$.

7.1 | Proof of Proposition 2

**Proof.** Recall the calculation for $\text{Res} \bar{u}$ in (76). We have bounded each of the terms, in order, using Lemmas 5–9. We obtain Proposition 2 by using the largest parts of each of the bounds.

7.2 | Other masses and 3D

The method we have employed is robust enough to consider other ways of realizing the masses. For example, we may consider periodic masses by which we mean there exists a positive integer $k$ s.t. $m(j_1, j_2) = m(j_1 + k, j_2 + k)$. Then $\chi$ is periodic and bounded, so the analysis of the terms
appearing in the residual becomes simpler. For instance, we no longer need the term with the
cutoff. In this case, one of the largest terms in $\varepsilon$ is the one given in Lemma 7. A quick count shows
that an $\varepsilon$ is lost from converting the $\ell^2$ norm to an $H$ norm, but an $\varepsilon$ is picked up on account of the
finite difference. Thus the term is $O(\varepsilon)$, which produces roughly the same size residual in $\varepsilon$ as we
obtained for the i.i.d. masses. The only difference is that for the i.i.d. case, the residual is slightly
larger due to the logarithms and the use of the cutoff.

We also may consider masses which are all identical. In such a case, the only term which
appears in the residual is the one bounded in Lemma 5. This improves the accuracy of the
approximate solution substantially as the residual would be $O(\varepsilon^2)$.

Another generalization we can make is that masses need not be identically distributed, so long
as they all fall into some interval $[a, b] \in \mathbb{R}^+$ and have the same expected value. Since our meth-
ods did not use any other features of the masses being identically distributed, for example, equal
variances or probabilities, our result extends to this case without modification.

Another important example is layered media. Suppose that

$$\{m(j_1, j_2)\}_{j_1 = -\infty}^{\infty}.$$ (123)

is random i.i.d. sequence of masses and that for all $j_1$

$$\cdots = m(j_1, -1) = m(j_1, 0) = m(j_1, 1) = \cdots$$ (124)

In this case, it is actually possible to solve (21) and thereby not use a cutoff, but we proceed
using the same tools we have developed since such tools can be utilized in higher dimensions.
In this case, Hoeffding’s inequality does not immediately hold since we do not have complete
independence. Reconsider (40)

$$L(\chi_{R_\varepsilon})(j_1, j_2) = \sum_{|k_1| \leq R_\varepsilon} \sum_{|k_2| \leq R_\varepsilon} L(\varphi)(j_1 - k_2, j_2 - k_2)z(k_1, k_2)$$

$$= \sum_{k_1 = j_1 - R_\varepsilon}^{j_1 + R_\varepsilon} \sum_{k_2 = j_2 - R_\varepsilon}^{j_2 + r} L(\varphi)(k_1, k_2)z(j_1 - k_1, j_2 - k_2).$$ (125)

Let $z(k_1) := z(k_1, \cdot)$. This is well defined because of (124). Thus

$$L(\chi_{R_\varepsilon})(j_1, j_2) = \sum_{k_1 = j_1 - R_\varepsilon}^{j_1 + R_\varepsilon} z(j_1 - k_1) \sum_{k_2 = j_2 - R_\varepsilon}^{j_2 + r} L(\varphi)(k_1, k_2).$$ (126)

Let

$$L(R_\varepsilon)(j_2, k_1) := \sum_{k_2 = j_2 - r}^{j_2 + r} L(\varphi)(k_1, k_2)$$ (127)

so

$$L(\chi_{R_\varepsilon})(j_1, j_2) = \sum_{k_1 = j_1 - R_\varepsilon}^{j_1 + R_\varepsilon} z(j_1 - k_1)L(R_\varepsilon)(j_2, k_1).$$ (128)
This is a sum involving independent mean zero random variables which are contained in some interval, so we may apply Hoeffding’s inequality. Let

\[ \| \mathcal{L}(R_e)(j_1, j_2) \|^2 := \sum_{k_1 = j_1 - R_e}^{j_1 + r} \mathcal{L}(R_e)(j_2, k_1)^2. \]  

(129)

Then we have by Hoeffding’s inequality that

\[ P( | \mathcal{L}(\chi_{R_e})(j_1, j_2) | \geq t ) \leq 2P\left( \frac{-2t^2}{(a - b)^2 \| \mathcal{L}(R_e)(j_1, j_2) \|^2} \right). \]

(130)

Now the same argument works as is made in Section 3.2, but the relevant quantity to calculate is the square root of (129). Take \( \mathcal{L} \) to be any of the operators in Section 3.2. Then

\[ \| \mathcal{L}(R_e)(j_1, j_2) \|^2 = \sum_{k_1 = j_1 - R_e}^{j_1 + r} \left( \sum_{k_2 = j_2 - R_e}^{j_2 + R_e} \mathcal{L}(\varphi)(k_1, k_2) \right)^2. \]

(131)

It is possible to use Jensen’s inequality to obtain

\[ \| \mathcal{L}(R_e)(j_1, j_2) \|^2 \leq (2R_e + 1) \sum_{k_1 = j_1 - R_e}^{j_1 + R_e} \sum_{k_2 = j_2 - R_e}^{j_2 + R_e} \mathcal{L}(\varphi)(k_1, k_2)^2 = (2R_e + 1)\| \mathcal{L}\varphi \|_{D(j_R)}^2. \]

(132)

One notices that \( \| \mathcal{L}\varphi \|_{D(j_R)}^2 \) are norms we have already computed in Section 3.2 (see (53) for instance). The \( r \) out front provides an extra half power of \( \epsilon^{-1 - \sigma} \) when one considers (105) and after taking square roots as in (51). Therefore, in contrast with (77), we have

\[ \sup_{|t| \leq \epsilon^{-1}T} \| \text{Res} \tilde{u} \|_{L^2} \leq \epsilon^{\frac{3}{2}} C\omega \log(\epsilon^{-1 - \sigma}) \frac{3}{2} \| \phi, \psi \|_{H^3}. \]

(133)

This example sheds some light on the complexity introduced when considering random masses. In contrast, for periodic masses, even if (124) holds, then (21) is solvable and the solution is periodic so most importantly bounded. Therefore the residual is always \( O(\epsilon) \).

Our methods are also extendable to 3D, and we could show that the residual is roughly \( O(\epsilon^{1/2}) \) in 3D, which is greater than it is in 2D which is roughly \( O(\epsilon) \). Roughly means ignoring logarithms. The difference in a half power of \( \epsilon \) is expected because of the scaling of the initial conditions, and the increase would also occur for constant or periodic mass lattices. It is known that the fundamental solution for the lattice Laplacian for the 3D cubic lattice is \( \varphi_3(j) = \frac{C}{|j|} + O(\frac{1}{|j|^4}) \) as \( j \to \infty \) with \( \varphi_3(0) = C.^{25} \) From this, one can show that the 3D version of the inequality (53) is a half-power smaller in \( R_e \), which would seemingly contribute to improving the residual. However, we cannot make the 3D version of (55) any better. At best, we can say \( \| \delta_3 \varphi_3 \|_{D(j, R_e)} \) is bounded by a constant independent of \( j \) and \( R_e \), which is only a marginal improvement over the 2D version of the inequality. We emphasize that we are not certain there is only a marginal improvement. It could be that the residual is roughly the same size in \( \epsilon \) in 3D as it is in 2D and that our methods
simply fail to show it. On the other hand, if it were the case that the residual were roughly the same size for the random mass lattice in 3D, then it would mean it is also smaller than the residual of the periodic mass lattice.

7.3 Random springs

The 1D argument in McGinnis and Wright\textsuperscript{15} is able to handle random springs in addition to random masses since the equations of motion are able to be written in terms of $r$, the relative displacements, and $p$, the velocities. This change of coordinates nicely separates the variable spring constants from the discrete Laplacian. Let us consider our problem with random springs. Let $\alpha(j)$ be the spring connecting the mass at $j$ and the mass at $j + e_1$, and $\beta(j)$ the spring connecting the mass at $j$ and the mass at $j + e_2$. Equations of motion, analogous to (1), can now be written as

$$m\ddot{u} = \delta^-_1 \alpha \delta^-_1 u + \delta^+_2 \beta \delta^-_2 u.$$  \hspace{1cm} (134)

One could also consider more complicated interactions.

Note the operators $\delta^\pm_i$ are given in (64). We need to be able to separate $\alpha$ and $\beta$ from the partial differences, which is possible. Let $\dot{u} = p$, $q_1 = \alpha \delta^-_1 u$, and $q_2 = \beta \delta^-_2 u$. We obtain the system of lattice differential equations

$$m \dot{p} = \delta^+_1 q_1 + \delta^+_2 q_2$$
$$\frac{1}{\alpha} q_1 = \delta^-_1 p \text{ and } \frac{1}{\beta} q_2 = \delta^-_2 p.$$  \hspace{1cm} (135)

Despite its simple appearance, it is not obvious how to homogenize the system. If we follow the usual technique from the 1D setting and take the ansatz as $\bar{p}(j) = P_0(\varepsilon j) + \varepsilon \chi_\sigma(j)P_1(\varepsilon j)$ and $\bar{q}(j) = Q_{1,0}(\varepsilon j) + \varepsilon \delta^-_1 \chi_\sigma(j)eQ_{1,1}$, we would need to have that $\delta^-_1 \chi_\sigma = \frac{1}{\alpha} - E[\frac{1}{\alpha}]$ and $\delta^+_2 \chi_\sigma = \frac{1}{\beta} - E[\frac{1}{\beta}]$ simultaneously ($s$ for spring). Assuming the $\alpha(j)$ and $\beta(j)$ are i.i.d., we find that $\delta^-_2 \frac{1}{\alpha} = \delta^-_1 \frac{1}{\beta}$. We could enforce this restriction, which would make us redefine what random means in this context, but it is clear, it takes a new approach to proceed with fully random spring constants.

7.4 Main estimate result

**Theorem 3.** Suppose $\psi, \nabla \phi \in H^5_\sigma$, where $\sigma$ is any small positive real number. Let $\varepsilon < 1$. Let

$$u(j, 0) = \varepsilon^{-1} \phi(\varepsilon j) \text{ and } \dot{u}(j, 0) = \psi(\varepsilon j),$$  \hspace{1cm} (136)

and suppose it solves (1) with independent masses s.t. $m(j) \in [a, b]$ for positive numbers $a$ and $b$. Let $U$ satisfy

$$mU_{\tau\tau} = \Delta_X U$$  \hspace{1cm} (137)

with

$$U(X, 0) = \phi(X) \text{ and } \dot{\partial}_\tau U(X, 0) = \psi(X).$$  \hspace{1cm} (138)
Set
\[ \hat{u}(j, t) = \epsilon^{-1} U(\epsilon j, \epsilon t). \] (139)

Then there exists a \( C_\omega(a, b, \bar{m}, T) \) a.s. s.t.
\[ \sup_{|t| \leq \epsilon^{-1} T} \| u - \hat{u} \|_{L^2} \leq \epsilon^{-1 - \sigma} C_\omega \log(\epsilon^{-1 - \sigma})^{\frac{3}{2}} \| \phi, \psi \|_{H^\frac{5}{3}_\sigma}. \] (140)

and
\[ \sup_{|t| \leq \epsilon^{-1} T} \| \dot{u} - \dot{\hat{u}} \|_{L^2} \leq \epsilon^{-\sigma} C_\omega \log(\epsilon^{-1 - \sigma})^{\frac{3}{2}} \| \phi, \psi \|_{H^\frac{5}{3}_\sigma}. \] (141)

Remark 4. Although the right-hand side of (140) and (141) appear to grow, the size of \( u \) and \( \dot{u} \) is roughly \( \epsilon^{-2} \) and \( \epsilon^{-1} \), so the error is relatively small.

Proof. Let
\[ \tilde{u}(j, t) = \epsilon^{-1} U(\epsilon j, \epsilon t) + \epsilon \chi_{\bar{R}}(j) U_{\tau\tau}(\epsilon j, \epsilon t). \] (142)
This is what has typically been our \( \tilde{u} \) as defined in (24). Then we have by (69) and (70) that, for a constant depending on \( a, b \), and \( T \),
\[ \sup_{|t| \leq \epsilon^{-1} T} \| \dot{u} - \dot{\tilde{u}} \|_{L^2} \leq C \left( \text{Res}_0 + \epsilon^{-1} \sup_{|t| \leq \epsilon^{-1} T} \| \text{Res} \tilde{u} \|_{L^2} \right) \] (143)
and
\[ \sup_{|t| \leq \epsilon^{-1} T} \| u - \tilde{u} \|_{L^2} \leq \epsilon^{-1} C \left( \text{Res}_0 + \epsilon^{-1} \sup_{|t| \leq \epsilon^{-1} T} \| \text{Res} \tilde{u} \|_{L^2} \right). \] (144)
Recall that \( \text{Res}_0 = \| \dot{u}(0) - \dot{\tilde{u}}(0), \delta^+_{i1} u(0) - \delta^+_{i} \tilde{u}(0), \delta^+_{i} u(0) - \delta^+_{i} \tilde{u}(0) \| \) is the discrepancy between the initial conditions of \( \tilde{u} \) and \( u \), and as noted in (67), we need it to be roughly the same size in \( \epsilon \) as \( \epsilon^{-1} \sup_{|t| \leq \epsilon^{-1} T} \| \text{Res} \tilde{u} \| \). Specifically, we just need to know that \( \text{Res}_0 \) is \( O(\epsilon^{-\sigma} \log(\epsilon^{-1 - \sigma})^{3/2}) \) because that is our bound for \( \epsilon^{-1} \sup_{|t| \leq \epsilon^{-1} T} \| \text{Res} \tilde{u} \| \) given by Proposition 2. The terms in \( \text{Res}_0 \) are \( \epsilon^2 \chi_{\bar{R}} \ U_{\tau\tau}(0), \epsilon \delta^+_{i1} (\chi_{\bar{R}} \ U_{\tau\tau}(0)), \) and \( \epsilon \delta^+_{i} (\chi_{\bar{R}} \ U_{\tau\tau}(0)) \) and work out to be very similar to the terms appearing in the calculated residual in (76). Thus they may be bounded using the same ideas as those used in the lemmas in Section 7. For instance, \( \epsilon \delta^+_{i1} (\chi_{\bar{R}} \ U_{\tau\tau}) = \epsilon S^+_{i1} (\chi_{\bar{R}}) \delta^+_{i1} U + \epsilon \delta^+_{i1} (\chi_{\bar{R}}) U_{\tau\tau} \). Then \( \epsilon \delta^+_{i1} (\chi_{\bar{R}}) U_{\tau\tau} \) may be bounded by the arguments in Lemma 7. However, since we have \( U_{\tau\tau} \) and not \( \delta^+_{i1} (\chi_{\bar{R}}) U_{\tau\tau} \) as in the lemma, \( \epsilon \| \delta^+_{i1} (\chi_{\bar{R}}) U_{\tau\tau} \| \) is \( O(\epsilon^{-\sigma} \log(\epsilon^{-1 - \sigma})) \) which is larger, but still small enough. We find similarly that \( \epsilon S^+_{i1} (\chi_{\bar{R}}) \delta^+_{i1} U_{\tau\tau} \) can also be bounded by arguments appearing in Lemma 8. Again it turns out to be slightly larger, that is, \( O(\epsilon^{-\sigma} \log(\epsilon^{-1 - \sigma})^{3/2}) \) because here we have \( \delta^+_{i} U \) as opposed to \( \Delta^+_{i} U \). Note these terms also require less regularity.
Therefore, we use (77) in Proposition 2 to obtain (140) and (141) but for \( \tilde{u} \) instead of \( \hat{u} \). Thus, it remains to analyze the difference between the two. We have

\[
\sup_{|t| \leq \varepsilon^{-1} T} \| \hat{u} - \tilde{u} \|_{\ell^2} \leq \sup_{|t| \leq \varepsilon^{-1} T} \| \varepsilon^2 \chi_{R\varepsilon}(\cdot) U_{\varepsilon \varepsilon}(\cdot) \|_{\ell^2}
\]  

(145)

and

\[
\sup_{|t| \leq \varepsilon^{-1} T} \| \hat{u} - \tilde{u} \|_{\ell^2} \leq \sup_{|t| \leq \varepsilon^{-1} T} \| \varepsilon^2 \chi_{R\varepsilon}(\cdot) U_{\varepsilon \varepsilon}(\cdot) \|_{\ell^2}.
\]  

(146)

Both of these can be bounded using essentially the same steps as those in Lemma 9. One finds

\[
\sup_{|t| \leq \varepsilon^{-1} T} \| \hat{u} - \tilde{u} \|_{\ell^2} \leq \varepsilon C_{\omega R \varepsilon} \left( (\sqrt{T} + \log(R \varepsilon)^{3/2})(\|\phi\|_{H^1_w} + \varepsilon\|\psi\|_{H^2_w}) \right) 
\]  

(147)

\[
\sup_{|t| \leq \varepsilon^{-1} T} \| \hat{u} - \tilde{u} \|_{\ell^2} \leq C_{\omega R \varepsilon} \left( (\sqrt{T} + \log(R \varepsilon)^{3/2})(\|\phi\|_{H^2_w} + \varepsilon\|\psi\|_{H^1_w}) \right) 
\]  

(148)

Recalling the definition of \( R\varepsilon \) in (105), these bounds have the correct power of \( \varepsilon \). Thus, with the appropriate constant \( C_{\omega \varepsilon} \), everything can be dominated by the right-hand side in (140) and (141).

8 | COARSE GRAINING

Theorem 3 says that the macroscopic behavior of the system evolves according to an effective wave equation. We formalize this notion by proving a convergence result in the macroscopic setting. We have a number of operators to introduce. The lattice Fourier transform, \( F : \ell(\mathbb{Z}^2) \rightarrow L^2(\mathbb{R}^2) \), is given by

\[
F[f](y) = \frac{1}{(2\pi)^2} \sum_{j \in \mathbb{Z}^2} \exp(-ij \cdot y)f(j).
\]  

(149)

Here \( y \in \mathbb{R}^2 \). Its inverse is

\[
F^{-1}[g](j) = \int_{[-\pi,\pi]^2} \exp(iy \cdot j)g(y)dy.
\]  

(150)

Let \( F : L^2(\mathbb{R}^2) \rightarrow L^2(\mathbb{R}^2) \) be the typical Fourier transform and \( F^{-1} \) be its inverse. The sampling operator is

\[
S(u)(j) = u(j).
\]  

(151)

A cutoff operator is

\[
\vartheta_{\gamma}(y) = \begin{cases} 
1 & y \in [-\gamma, \gamma]^2 \\
0 & \text{else}
\end{cases}.
\]  

(152)
Finally, we define a low-pass interpolator. For a continuous variable $\mathbf{x} \in \mathbb{R}^2$

$$\mathcal{L}[f](\mathbf{x}) = F^{-1}[\partial_\pi(\cdot)F[f](\cdot)](\mathbf{x}).$$

(153)

The following theorem says that the abstract diagram found in Mielke\textsuperscript{5} holds in the setting of i.i.d. masses almost surely. The diagram depicts how time evolution commutes with the coarse-graining operator $\mathcal{L}$ in the limit as $\epsilon \to 0$, meaning that one can first evolve the system according to the (1), and then apply $\mathcal{L}$, or one can apply $\mathcal{L}$ initially to obtain macroscopic initial conditions and then evolve those according to the effective wave equation and arrive at the same result.

### 8.1 Main convergence result

**Theorem 4.** Let $U$ solve (137) with initial conditions given by (138). Let $u$ solve (1) with the same assumptions on the masses as in Theorem 3 and with initial conditions given by (136). Put

$$U_\epsilon(\mathbf{X}, \tau) = \epsilon \mathcal{L}(u(\cdot, \tau/\epsilon))(\mathbf{X}/\epsilon)$$

(154)

so that

$$\partial_\tau U_\epsilon(\mathbf{X}, \tau) = \mathcal{L}(\dot{u}(\cdot, \tau/\epsilon))(\mathbf{X}/\epsilon).$$

(155)

Then

$$\lim_{\epsilon \to 0} \sup_{\tau \leq T} ||U_\epsilon - U, \partial_\tau U_\epsilon - \partial_\tau U||_{L^2} = 0$$

(156)

almost surely.

**Remark 5.** The rate of convergence is no worse than $\epsilon^{1-\sigma} \log(\epsilon^{-1-\sigma})^3$ as seen in the proof.

**Proof.**

$$||U_\epsilon - U||_{L^2} \leq ||\epsilon \mathcal{L}(u(\cdot, \tau/\epsilon))(\cdot/\epsilon) - \mathcal{L}S(U(\cdot, \tau))(\cdot/\epsilon)||_{L^2} + ||\mathcal{L}S(U(\cdot, \tau))(\cdot/\epsilon) - U(\cdot, \tau)||_{L^2}.\quad (157)$$

According to Lemma A10, the second term goes to 0 uniformly in $\tau$ at a rate of $\epsilon$. From the first term, we have

$$||\epsilon \mathcal{L}(u(\cdot, \tau/\epsilon))(\cdot/\epsilon) - \mathcal{L}S(U(\cdot, \tau))(\cdot/\epsilon)||_{L^2} = \epsilon ||\epsilon \mathcal{L}(u(\cdot, \tau/\epsilon))(\cdot) - \mathcal{L}S(U(\cdot, \tau))(\cdot)||_{L^2}.\quad (158)$$

By Plancherel

$$\epsilon ||\epsilon \mathcal{L}(u(\cdot, \tau/\epsilon))(\mathbf{X}) - \mathcal{L}S(U(\cdot, \tau))(\cdot)||_{L^2} \leq \epsilon ||\epsilon u(\cdot, \tau/\epsilon) - SU(\cdot, \tau)||_{\ell^2}.\quad (159)$$

Now we use that $SU(\cdot, \tau) = \epsilon \hat{u}(\cdot, \epsilon t)$ as in Theorem 3 and the result follows from that same theorem. Recall $\tau = \epsilon t$. Similarly

$$\epsilon ||\mathcal{L}(\dot{u}(\cdot, \tau/\epsilon))(\mathbf{X}) - \mathcal{L}S\dot{u}(\cdot, \tau)(\cdot)||_{L^2} \leq \epsilon ||\dot{u}(\cdot, \tau/\epsilon) - S\dot{u}(\cdot, \tau)||_{\ell^2}.\quad (160)$$

Again we use Theorem 3 where $U_\dot{\tau}(\cdot, \tau) = \hat{\dot{u}}(\cdot, t)$. 
FIGURE 2 The left (right) panel shows a log-log graph of a.e.d. (a.e.v) versus $\epsilon$. Since the masses are chosen randomly, 10 realization of masses were tested. The distribution of the error is plotted with a box and whisker plot at each value of $\epsilon$. The slope of the line of best fit numerically approximates the power of $\epsilon$ in (140) ((141)). a.e.d., absolute error of the displacement; a.e.v., absolute error of the velocity.

9 | NUMERICAL RESULTS

Our numerical results focus on confirming the upper bounds found in Theorem 3. All experiments simulated (1) using the fourth-order Runge–Kutta (RK4) on a suitably large domain of radius $8/\epsilon$. Thus the solutions interacted minimally with the boundary. The effective wave equation was also computed numerically using a finite difference approximation with RK4 as the time stepper on a domain of radius 8. Grid spacing size was chosen to be $1/32$, which allows for frequent enough sampling to obtain an approximate solution at every lattice point when $\epsilon = 1/32$. The time step used for the lattice dynamics was $1/32$ but $1/1024$ for the wave equation. This ensures an approximate solution could be obtained for all $t \leq 1/32$.

We refer to the left-hand side of (140) as the absolute error of the displacement (a.e.d.) and the left-hand side of (141) as the absolute error of the velocity (a.e.v). For the next two experiments, we have chosen

$$
\phi(X_1, X_2) = \text{sech}\left(\frac{1}{2}(X_1 - 1)^2 + (X_2 - 1)^2\right), \quad \psi(X_1, X_2) = \text{sech}\left((X_1 + 1)^2 + \frac{1}{2}(X_2 + 1)^2\right)
$$

and looked at $\epsilon$ over $\{1/2, 1/4, 1/8, 1/16, 1/32\}$. Every $m_j$ is sampled from $\{1/2, 3/2\}$, and for the first experiment the masses are chosen to be i.i.d. As $\epsilon$ varies, this grid of randomly chosen masses remains fixed.

The upper bound for the a.e.v. obtained in Theorem 3 is $\epsilon^{-\sigma} \log(\epsilon^{-1-\sigma})^{3/2}$ for the a.e.v. (Recall that $\sigma$, is any arbitrarily small positive number.) The slopes seen in Figure 2 are in agreement with the bounds obtained in the theorem, that is, the slopes reflect to what power of $\epsilon$ the error depends. In fact the estimate is close to sharp. We can thus think of such bounds as giving an analytic prediction on the size of the error in many cases.

For the second experiment, we use the setup for the masses given by (123) and (124), that is, the masses are layered. According to (133), we expect the slopes to be no more than a half power less than those seen in Figure 2. This is indeed what we see in Figure 3. Again, the numerical error is close to the predicted error.

Finally, we compare these results to what happens in three deterministic cases seen in Figure 4. In one case, we choose the masses to be constant in which case we would expect the slope for the
These panels reflect the same measurements as those in Figure 2, but now the masses have been chosen according to (123) and (124). a.e.d., absolute error of the displacement; a.e.v., absolute error of the velocity.

Masses have been chosen deterministically in three different ways. a.e.d., absolute error of the displacement; a.e.v., absolute error of the velocity.

a.e.v and the a.e.d. to be no worse than 1 and 0, respectively. In a second, the masses are periodically layered, that is, they only vary periodically with a period of 2 along one of the coordinate axis and along the other, (124) holds. For the third case, masses are chosen to vary periodically along both coordinate axes with a period of 2. In both cases, we expect the slope of the a.e.v. and a.e.d. to be no worse than 0 and 1. In both cases, the upper bound holds; however, unlike in all the previous experiments, numerically computed a.e.d. seems to be substantially better than the analytic upper bound, since the slope for both periodic cases is closer to 0 than to 1.

One important observation is that when the randomly chosen masses are layered, the approximation to the wave equation does not converge as quickly as it does when they are not layered. The main physical explanation we propose for the difference in the observed slope between Figures 2 and 3 is that, in the second experiment, reflections caused by the random masses manifest as long ripples transverse to the direction in which the masses are randomly changing. This is in contrast to the first experiment where reflections manifest as localized disturbances. Figure 5 gives some empirical evidence for this phenomenon.

One possible heuristic explanation for why we do not get improvement for periodic masses is that there is always a direction in $\mathbb{R}^2$ along which the averages of masses in lines transverse to that direction are varying (unless the masses are all constant). This produces a kind of layering that cannot occur if all the masses are chosen to be i.i.d. since masses along any line average to the same value. Hence, in this sense, the homogenization is more uniform.
The left panel shows a snapshot of a wave traveling through masses which are i.i.d. for some \( t > 0 \), and the right shows a wave with the same initial conditions at the same point in time traveling through layered random masses. One can spot the long transverse ripples in the right panel, whereas in the left panel deviations appear more localized.

This table summarizes the epsilon dependence of the absolute errors. The numbers indicate the power of \( \epsilon \). \( \sigma \) is a small and positive constant. Results have been rounded.

| Mass/Dimension          | a.e.v. Predicted | a.e.v. Simulated | a.e.d. Predicted | a.e.d. Simulated |
|-------------------------|------------------|------------------|------------------|------------------|
| 1D constant coefficient | 1.5              | 1.5              | 0.5              | 0.5              |
| 1D periodic coefficient | 0.5              | 0.7              | \(-0.5\)         | 0.5              |
| 1D random coefficient   | \(-\sigma\)      | 0.2              | \(-1 - \sigma\) | \(-0.8\)        |
| 2D constant coefficient | 1                | 1.1              | 0                | 0.0              |
| 2D periodic coefficient | 0                | 0.1              | \(-1\)           | 0.0              |
| 2D random coefficient   | \(-\sigma\)      | 0.1              | \(-1 - \sigma\) | \(-1.0\)        |
| 2D layered coefficient  | \(-0.5 - \sigma\)| \(-0.4\)         | \(-1.5 - \sigma\)| \(-1.5\)        |

Abbreviations: 1D, one dimensional; 2D, two dimensional; a.e.d., absolute error of the displacement; a.e.v., absolute error of the velocity.

Finally, we have performed a number of similar experiments in 1D, the results of which are summarized in Table 1. The predicted values can all be proven using essentially the same method as what has been demonstrated or discussed in Section 7.2 and throughout the rest of the paper. Even though our predicted values are upper bounds, we see that in most cases our estimates are close to sharp. The exception is when the masses are chosen periodically, where the predicted a.e.d. often overshoots by close to a full power of \( \epsilon \). This indicates that integrating \( p \), and then applying a triangle inequality to obtain an upper bound on \( u \) is not efficient in the periodic case.

For both the constant coefficient and periodic cases, a half power of \( \epsilon \) is lost with each increase in dimension. This is due to the length scaling and is neglected when one considers the coarse-graining limit. We see no reason this trend should not continue into higher dimensions. On the other hand, there is not a decrease in the power of \( \epsilon \) in the i.i.d. case. Heuristically, one can see this by comparing the variance of \( \chi \) in 1D and in 2D. The fundamental solution \( \phi \) of \( \Delta \) plays an important role in the growth rate. In 1D, this fundamental solution is given by \( \phi(\cdot) = \frac{1}{2} | \cdot | \). Taking the expectation of \( \chi_R^2 \) and then taking the square root, where \( \chi_R \) is defined by (39), yields that \( \chi_R \) is approximately the size of \( R^3 \) in \( \ell^\infty \). The same procedure in 2D yields \( R \log(R) \). This accounts for an additional half-power loss \( \epsilon \) in the 1D case. The argument using the ideas of sub-Gaussian random variables is one way to formalize this intuition and obtain an almost sure estimate. Ultimately, this half power is negated by the increase in dimension, and since the random term in the residual is
still the dominant one we find that the size of the absolute error (as dependent upon $\varepsilon$) roughly does not change from 1D to 2D. Therefore, the coarse-graining limit converges faster in 2D.

10 CONCLUSION

We have rigorously justified the wave equation as a descriptor for the macroscopic linearized dynamics of a 2D square lattice composed of i.i.d. masses. We have given analytic as well as numerical evidence that this description is more accurate in 2D than it is in 1D. We think such a result is modest evidence that waves propagate more easily in a disordered lattice in higher dimensions. An important exception that was seen to this occurs in the case of layered random masses, where the error became larger from 2D to 1D in the same way it did for periodic or constant masses. Also, we believe our methods can be readily extended to the 3D case to show the validity of the approximation by a 3D wave equation, but our methods do not indicate that the approximation becomes more accurate in 3D as it did from 1D to 2D.

Although there are results in the continuous setting that are similar to ours, as far as we can tell, this is the first result that provides a rigorous, almost sure bound on the rate of convergence of (1) to a wave equation, and we think that such a rate of convergence provides insight into the effects of dimensionality on the dynamics of disordered lattices. Finally, the techniques introduced, especially the use of sub-Gaussian random variables can probably be used to access similar results for various other discrete systems with different kinds of disorder.

DATA AVAILABILITY STATEMENT
Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

REFERENCES
1. Brillouin L. Wave Propagation in Periodic Structures. Dover Books on Engineering and Engineering Physics. Dover Publications; 1953.
2. Lee LW, Dhar A. Heat conduction in a two-dimensional harmonic crystal with disorder. Phys Rev Lett E. 2005;95:094302.
3. Maradudin A, Weiss GH. The disordered lattice problem: a review. Journal of the Society for Industrial and Applied Mathematics. 1958;6:302-319.
4. Bell R J. The dynamics of disordered lattices. Rep Prog Phys. 1972;35:1315.
5. Mielke A. Macroscopic behavior of microscopic oscillations in harmonic lattices via Wigner–Husimi transforms. Arch Rational Mech Anal. 2006;181:401-448.
6. Hristov N, Pelinovsky DE. Justification of the KP-II approximation in dynamics of two-dimensional FPU systems. Z Angew Math Phys. 2022;73:1-26.
7. Pelinovsky DE, Schneider G. KP-II approximation for a scalar FERMI-PASTA-Ulam system on a 2D square lattice. SIAM J Appl Math. 2023;83:79-98.
8. Gallone M, Pasquali S. Metastability phenomena in two-dimensional rectangular lattices with nearest-neighbour interaction. Nonlinearity. 2021;34:4983-5044.
9. Lukkarinen J, Spohn H. Kinetic limit for wave propagation in a random media. Arch Rational Mech Anal. 2007;183:93-162.
10. Martinez MJ, Kevrekidis PG, Porter MA. Superdiffusive transport and energy localization in disordered granular crystals. Phys Rev E. 2016;93:022902.
11. Chaudhuri A, Kundu A, Roy D, Dhar A, Lebowitz JL, Spohn H. Heat transport and phonon localization in mass-disordered harmonic crystals. Phys Rev B. 2010;81:064301.
12. de Moura FABF. Vibrational modes in a two-dimensional aperiodic harmonic lattice. *J Phys: Condens Matter.* 2010;22:435401.

13. Pinski SD, Schirmacher W, Whall T, Römer RA. Localization-delocalization transition for disordered cubic harmonic lattices. *J Phys: Condens Matter.* 2012;24:405401.

14. Fouque J-P, Garnier J, Papanicolaou G, Sølna K. Wave Propagation and Time Reversal in Randomly Layered Media. Stochastic Modeling and Applied Probability, vol. 56. Springer Science+ Business Media; 2007.

15. McGinnis JA, Wright JD. Using random walks to establish wavelike behavior in an FPUT system with random coefficients. *Discrete Contin Dyn Syst Ser S.* 2022;15:2581–2607.

16. Cioranescu D, Donato P. An Introduction to Homogenization. Oxford Lecture Series in Mathematics and Its Applications, 17. Oxford University Press; 1999.

17. Chirilus-Bruckner M, Chong C, Prill O, Schneider G. Rigorous description of macroscopic wave packets in infinite periodic chains of coupled oscillators by modulation equations. *Discrete Contin Dyn Syst Ser S.* 2012;5:879-901.

18. Gaison J, Moskow S, Wright JD, Zhang Q. Approximation of polyatomic FPU lattices by KdV equations. *Mult Scale Model Simul.* 2014;12:953-995.

19. Schneider G, Wayne CE. Counter-propagating waves on fluid surfaces and the continuum limit of the Fermi-Pasta-Ulam model. *International Conference on Differential Equations. Vols 1, 2.* (Berlin, 1999): World Scientific; 2000:390-404.

20. Evans LC. *Partial Differential Equations.* 2nd ed. Graduate Studies in Math. Vol 19. American Mathematical Society; 2010.

21. Duffin RJ. Basic properties of discrete analytic functions. *Duke Math J.* 1956;23:335-363.

22. Massart P. Concentration inequalities and model selection. In: *Ecole d’Eté de Probabilités de Saint-Flour XXXIII-2003.* Springer Verlag; 2007:21-23.

23. Durrett R. *Probability, Theory and Examples.* Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press; 2010:65.

24. Craig W. *A Course on Partial Differential Equations.* Graduate Studies in Math 197. American Mathematical Society; 2018:122-123.

25. Lawler GF, Limic V. *Random Walk: A Modern Introduction.* Cambridge Studies in Advanced Mathematics. Cambridge University Press; 2010:87-122.

26. McNamee J, Stenger F, Whitney EL. Whittaker’s cardinal function in retrospect. *Math Comput.* 1971;25:141-154.

27. Hoeffding W. Probability Inequalities for Sums of Bounded Random Variables. *J Am Stat Assoc.* 1963;58(301):13–30. https://doi.org/10.1080/01621459.1963.10500830

How to cite this article: McGinnis JA. Macroscopic wave propagation for 2D lattice with random masses. *Stud Appl Math.* 2023;151:752–790. https://doi.org/10.1111/sapm.12612

A.1 Probabilistic estimates A

Lemma A1. For $R \geq 2$, and $j \in \mathbb{Z}^2$, there exists a constant $C$ s.t.

\[ \log^+ (|j| + R) \leq C (\log^+ (|j|) + \log^+ (R)) \]  \hspace{1cm} (A1)

Proof. The inequality is trivial for $|j| = 0$. For $|j| \neq 0$, we can prove the inequality for log. By the concavity of log, log$(\cdot + 1)$ is subadditive.

\[ \log(|j| + R - 1 + 1) \leq \log(|j| + 1) + \log(R) \leq C (\log(|j|) + \log(R)). \]  \hspace{1cm} (A2)

Lemma A2. For $\varphi$ given by (38) and $\delta_i$ defined in (54) we have for some $C$

\[ |\delta_i(\varphi)(j)| \leq \frac{C}{|j| + 1}. \]  \hspace{1cm} (A3)
Proof. By (38), we have
\[
\delta_i(\varphi)(j) = \frac{1}{2\pi} \log^+ |j + e_i| - \frac{1}{2\pi} \log^+ |j - e_i| + O(|j|^{-2}). \tag{A4}
\]
Note that when $|j| = 0$, we are left with only small $O(|j|^{-2})$ terms. For $|j| = 1$, we have
\[
\left| \log^+ |j + e_i| - \log^+ |j - e_i| \right| \leq \log(2). \tag{A5}
\]
For $|j| > 1$, it can be shown that
\[
\left| \log |j + e_i| - \log |j - e_i| \right| \leq \frac{4}{|j| + 1}. \tag{A6}
\]
Thus we obtain the result.

**Lemma A3.** For $\varphi$ given by (38) and $\delta_i$ defined in (54), then when $R \geq 2$, we have for some $C$
\[
\|\delta_i \varphi\|^2_{D(j,R)} \leq C \log(|j| + R). \tag{A7}
\]

Proof. By Lemma A2,
\[
\|\delta_i \varphi\|^2_{D(j,R)} \leq C \sum_{k \in D(j,R)} \left( \frac{1}{|k| + 1} \right)^2. \tag{A8}
\]
The largest magnitude of $|k|$ in $D(j,R)$ is $|j| + R$. Also note that there is some constant $C$ s.t. the number of elements in $D(j,R)$ of some magnitude $R'$ is less than $C(R' + 1)$. Thus there exists a constant $C$ s.t.
\[
\sum_{k \in D(j,R)} \left( \frac{1}{|k| + 1} \right)^2 \leq C \sum_{R' = 0}^{R+1} \left( \frac{1}{R' + 1} \right)^2 = C \sum_{R' = 0}^{R+2} \left( \frac{1}{R' + 1} \right) \tag{A9}
\]
A common bound on the harmonic series yields
\[
\sum_{k \in D(j,R)} \left( \frac{1}{|k| + 1} \right)^2 \leq C \log(|j| + R + 2), \tag{A10}
\]
which yields the result.

**A.2 The effective wave $A$**

**Lemma A4.** Let $U$ satisfy (25) with initial conditions given by (79). Suppose that $w$ satisfies (86) and (87). Then
\[
\left| E_w(D^k \delta_i U)(\tau) \right| \leq (|\tau| + 1)C \left( \|\nabla \varphi\|^2_{H^{2+k}_w} + \|\psi\|^2_{H^{2+k}_w} \right). \tag{A11}
\]
The constant $C$ depends upon $\bar{m}$ as well as the bound $W$ in (87).
**Proof.** Note that $D^k U(X, \tau) : \mathbb{R}^2 \times \mathbb{R} \to \mathbb{R}^{2^k}$. Let the jth component of the $2^k$-dimensional vector of $D^k U$ be denoted by $D^k_j U$, that is, $D^k_j$ is some mixed partial of $X_1$ and $X_2$ of order $k$. Consider a ball of radius $R$ in $\mathbb{R}^d$ about the origin denoted $B(R)$.

$$E_w(\partial^i_j D^k_j U)(\tau) = \lim_{R \to \infty} \int_{B(R)} w^2 \left( \partial^{i+1}_j D^k_j U \right) \left( \partial^{i+2}_j D^k_j U \right) + c^2 w^2 \nabla \left( \partial^i_j D^k_j U \right) \cdot \nabla \left( \partial^{i+1}_j D^k_j U \right) dX.$$  
(A12)

Using integration by parts, we find

$$\dot{E}_w(\partial^i_j D^k_j U)(\tau) = \lim_{R \to \infty} \int_{B(R)} w^2 \left( \partial^{i+1}_j D^k_j U \right) \left( \partial^{i+2}_j D^k_j U \right) - c^2 \nabla \cdot \left( w^2 \nabla \left( \partial^i_j D^k_j U \right) \right) \left( \partial^{i+1}_j D^k_j U \right) dX + \lim_{R \to \infty} \int_{\partial B(R)} w^2 c^2 \left( \frac{\partial}{\partial \nu} \left( \partial^i_j D^k_j U \right) \right) \left( \partial^{i+1}_j D^k_j U \right) dS.$$  
(A13)

The boundary term vanishes in the limit due to the finite propagation speed of the wave. We need to calculate

$$\nabla \cdot \left( w^2 \nabla \left( \partial^i_j D^k_j U \right) \right) = (\nabla w^2) \cdot \left( \partial^i_j D^k_j U \right) + w^2 \Delta \left( \partial^i_j D^k_j U \right).$$  
(A14)

Since $\partial^i_j D^k_j U$ satisfies (25), we are left with

$$\dot{E}_w(\partial^i_j D^k_j U)(\tau) = \lim_{R \to \infty} \int_{B(R)} \nabla w^2 \cdot \nabla \left( \partial^i_j D^k_j U \right) \left( \partial^{i+1}_j D^k_j U \right) dX.$$  
(A15)

Using the assumption on $w$ in (87) and Cauchy–Schwarz, we have

$$\left| \nabla w^2 \cdot \nabla \left( \partial^i_j D^k_j U \right) \right| \leq C \left| \nabla \left( \partial^i_j D^k_j U \right) \right|.$$  
(A16)

Therefore,

$$\dot{E}_w(\partial^i_j D^k_j U)(\tau) \leq \lim_{R \to \infty} C \int_{B(R)} \left| \nabla \left( \partial^i_j D^k_j U \right) \right| \left| \partial^{i+1}_j D^k_j U \right| dX.$$  
(A17)

Using Cauchy–Schwarz once more and swapping derivatives

$$\dot{E}_w(\partial^i_j D^k_j U)(\tau) \leq C \left\| D^{k+1}_j \partial^i_j U \right\|_{L^2} \left\| D^k_j \partial^{i+1}_j U \right\|_{L^2}.$$  
(A18)

From (85), we know how to bound such beasts.

$$\dot{E}_w(\partial^i_j D^k_j U)(\tau) \leq C \left( \left\| \nabla \phi \right\|^2_{H^{i+k}} + \left\| \psi \right\|^2_{H^{i+k}} \right).$$  
(A19)

Integrating yields

$$E_w(\partial^i_j D^k_j U)(\tau) \leq |\tau| C \left( \left\| \phi \right\|^2_{H^{i+k+1}} + \left\| \psi \right\|^2_{H^{i+k}} \right) + E_w(D^k_j \partial^i_j U)(0).$$  
(A20)
Note that at $\tau = 0$ and when $i$ is even,
\[
\partial_i^j D_j^k U = c^2 D_j^k \Delta^\frac{i}{2} U = c^2 D_j^k \Delta^\frac{i}{2} \phi
\]  
(A21)

while, when $i$ is odd,
\[
\partial_i^j D_j^k U = c^2 D_j^k \Delta^\frac{i-1}{2} \partial_\tau U = c^2 D_j^k \Delta^\frac{i-1}{2} \psi.
\]  
(A22)

Using (A21) or (A22) and the definition (90)
\[
E_w \left( \partial_i^j D_j^k U \right)(0) \leq C \left( \|\nabla \phi\|_{H_{\bar{m}}^{i+k}}^2 + \|\psi\|_{H_{\bar{m}}^{i+k}}^2 \right).
\]  
(A23)

This holds for all $j \in \{1, 2, 3, \ldots, 2^k\}$. Thus
\[
\left| E_w \left( D^k \partial_\tau^i U \right)(\tau) \right| \leq (|\tau| + 1)C \left( \|\nabla \phi\|_{H_{\bar{m}}^{i+k}(B(\varepsilon-\sigma)c)}^2 + \|\psi\|_{H_{\bar{m}}^{i+k}(B(\varepsilon-\sigma)c)}^2 \right).
\]  
(A24)

Again, the constant $C$ depends upon $k$, $\bar{m}$, and also $W$.

**Lemma A5.** Let $U$ satisfy (25) with initial conditions given by (79). Then
\[
\left| E \left( D^k \partial_\tau^i U \right)(\tau) \right| \leq C \left( \|\nabla \phi\|_{H_{\bar{m}}^{i+k}(B(\varepsilon-\sigma)c)}^2 + \|\psi\|_{H_{\bar{m}}^{i+k}(B(\varepsilon-\sigma)c)}^2 \right).
\]  
(A25)

The constant $C$ depends only on $\bar{m}$.

**Proof.** Without loss of generality, let $\tau$ be positive. Note that $D^k U(X, \tau) : \mathbb{R}^2 \times \mathbb{R} \to \mathbb{R}^{2^k}$. Consider the $j$th component of the $2^k$-dimensional vector of $D^k U$ denoted by $D^k_j U$. We apply Leibniz’s rule
\[
\dot{E} \left( \partial_\tau^i D_j^k U \right)(\tau) = \int_{B(c|\tau|+\varepsilon-\sigma)c} \left( \partial_\tau^{i+1} D_j^k U \right) \left( \partial_\tau^{i+2} D_j^k U \right) + c^2 \nabla \left( \partial_\tau^i D_j^k U \right) \cdot \nabla \left( \partial_\tau^{i+1} D_j^k U \right) dX
\]
\[
- \frac{1}{2} \int_{\partial B(c\tau+\varepsilon-\sigma)c} c \left( \partial_\tau^{i+1} D_j^k U \right)^2 + c^3 \left| \nabla \left( \partial_\tau^i D_j^k U \right) \right|^2 dS.
\]  
(A26)

Let $\nu$ be the unit normal pointing out of the ball. Since $D_j^k U$ satisfies (25), integration by parts yields
\[
\dot{E}^{(i)} \left( \partial_\tau^i D_j^k U \right)(\tau) = - \int_{\partial B(c\tau+\varepsilon-\sigma)c} c^2 \frac{\partial \left( \partial_\tau^i D_j^k U \right)}{\partial \nu} \left( \partial_\tau^{i+1} D_j^k U \right) \]  
\[
+ \left( \frac{1}{2} c \left( \partial_\tau^{i+1} D_j^k U \right)^2 + \frac{1}{2} c^3 \left| \nabla \left( \partial_\tau^i D_j^k U \right) \right|^2 \right) dS.
\]  
(A27)
The first term in the integrand is bounded as
\[
c^2 \left\| \partial_t \left( \frac{\partial_i D_j^k U}{\partial \nu} \right) \right\| \left( \partial_t^{i+1} D_j^k U \right) \leq c^2 \| \nabla \partial_t D_j^k U \| \left\| \partial_t^{i+1} D_j^k U \right\| \leq \frac{1}{2} c \left( \partial_t^{i+1} D_j^k U \right)^2 + \frac{1}{2} c^2 \| \nabla \partial_t D_j^k U \|^2.
\] (A28)

Therefore,
\[
\dot{E} \left( \partial_t^i D_j^k U \right) (\tau) \leq 0,
\] (A29)

and thus
\[
\tilde{E} \left( \partial_t^i D_j^k U \right) (\tau) \leq \tilde{E} \left( D_j^k \partial_t^i U \right) (0).
\] (A30)

Using (A21) or (A22)
\[
\tilde{E} \left( \partial_t^i D_j^k U \right) (\tau) \leq C \left( \| \nabla \phi \|^2_{H^{k+1} (B(\epsilon-\sigma) \cap B(\epsilon+\sigma))} + \| \psi \|^2_{H^{k+1} (B(\epsilon-\sigma) \cap B(\epsilon+\sigma))} \right).
\] (A31)

This holds for all \( j \in \{1, 2, 3, \ldots, 2^k \} \); therefore, yielding the result where the final constant will depend upon \( k \) and \( c \) (which depends on \( \bar{m} \)).

**Lemma A6.** Let \( \phi : \mathbb{R}^2 \to \mathbb{R} \). For \( \sigma > 0 \),
\[
\| \phi \|_{H^k (B(\epsilon-\sigma) \cap B(\epsilon+\sigma))} \leq \epsilon \| \phi \|_{H^k_\sigma}.
\] (A32)

**Proof.** Consider
\[
\left\| D^k \phi \right\|^2_{L^2 (B(\epsilon-\sigma) \cap B(\epsilon+\sigma))} = \int_{B(\epsilon-\sigma) \cap B(\epsilon+\sigma)} \left| D^k \phi (X) \right|^2 dX = \int_{B(\epsilon-\sigma) \cap B(\epsilon+\sigma)} \frac{|\epsilon^{\sigma} X|^{2\sigma-1}}{|\epsilon^{\sigma} X|^{2\sigma-1}} \left| D^k \phi (X) \right|^2 dX.
\] (A33)

Since \( \epsilon^{\sigma} |X| \geq 1 \) in the region of the integral, we have
\[
\left\| D^k \phi \right\|^2_{L^2 (B(\epsilon-\sigma) \cap B(\epsilon+\sigma))} \leq \epsilon^2 \int_{B(\epsilon-\sigma) \cap B(\epsilon+\sigma)} |X|^{2\sigma-1} \left| D^k \phi (X) \right|^2 dX.
\] (A34)

Thus
\[
\left\| D^k \phi \right\|^2_{L^2 (B(\epsilon-\sigma) \cap B(\epsilon+\sigma))} \leq \epsilon^2 \left\| \left( 1 + |\cdot| \right)^{\sigma-1} D^k \phi (\cdot) \right\|^2_{L^2}.
\] (A35)

Taking square roots, summing over \( j \), and comparing with the definition of \( H^k_\sigma \) in (97) yields the result.
A.3 | Residual estimates A

**Lemma A7.** Let \( f(x) : \mathbb{R}^2 \to \mathbb{R} \) be in \( H^2 \) and let \( x^{(j)} \in \mathbb{R}^2 \cap \prod_{i=1}^{2} [j_1, j_1+1] \) where \( j \in Q \subset \mathbb{Z}^2 \). Then

\[
\sum_{j \in Q} f(x^{(j)})^2 \leq \sum_{j \in Q} 4 \| f \|_{H^2}^2 \prod_{i=1}^{2} [j+i, j+1].
\] (A36)

**Proof.** The proof is an extension of the proof for a 1D sum in Lemma 4.3 of Gaison et al.\textsuperscript{18}.

**Corollary A1.** In the same context as Lemma A7 with \( Q = \mathbb{Z}^2 \), we have

\[
\sum_{j \in \mathbb{Z}^2} f(x^{(j)})^2 \leq \varepsilon^{-2} 4 \| f \|_{H^2}^2.
\] (A37)

**Lemma A8.** In the same context as Lemma A7 with \( Q = D(0, \varepsilon^{-1} R + 1)_c \), we have

\[
\sum_{j \in Q} f(x^{(j)})^2 \leq \varepsilon^{-2} 4 \| f \|_{H^2(B(\varepsilon^{-1} R)_c)}^2.
\] (A38)

**Proof.** From Lemma A7,

\[
\sum_{j \in D(0, \varepsilon^{-1} R + 1)_c} f(x^{(j)})^2 \leq 4 \sum_{j \in D(0, \varepsilon^{-1} R + 1)_c} \| f(x) \|_{H^2(B(\varepsilon^{-1} R)_c)}^2.
\] (A39)

Note that

\[
\bigcup_{j \in D(0, \varepsilon^{-1} R + 1)_c} \prod_{i=1}^{2} [j, j + e_i] \subset B(\varepsilon^{-1} R)_c.
\] (A40)

Hence,

\[
\sum_{j \in D(0, \varepsilon^{-1} R + 1)_c} \| f(x) \|_{H^2(B(\varepsilon^{-1} R)_c)}^2 \leq \| f(x) \|_{H^2(B(\varepsilon^{-1} R)_c)}^2.
\] (A41)

Written in polar coordinates,

\[
\| f(x) \|_{H^2(B(\varepsilon^{-1} R)_c)}^2 = \int_0^{2\pi} \int_{\varepsilon^{-1} R}^\infty f(\theta, \varepsilon r)^2 r dr d\theta
\]

\[
= \varepsilon^{-2} \int_0^{2\pi} \int_{\varepsilon^{-1} R}^\infty (\phi, s)^2 s ds d\theta
\]

\[
= \varepsilon^{-2} \| f \|_{H^2(B(\varepsilon^{-1} R)_c)}^2.
\] (A42)

**Lemma A9.** Let \( \delta_{\varepsilon}^\pm \phi(X) := (\phi(X) - \phi(X - \varepsilon e_i)) \). Let \( w \) be as in assumptions (86) and (87). Then

\[
\| \delta_{\varepsilon}^\pm \phi \|_{H^k_w} \leq \varepsilon C \| \phi \|_{H^{k+1}_w}.
\] (A43)
Proof. The proof is analogous to a 1D version of the result found at the end of the proof of Lemma 3.3 in McGinnis and Wright.\footnote{15}

\textbf{Corollary A2.} Let $\Delta_{\epsilon_i} \phi(X) := \phi(X + \epsilon e_i) - 2\phi(X) + \phi(X - \epsilon e_i)$ and $w$ as in the previous lemma. Then

$$
\|\Delta_{\epsilon_i} \phi\|_{H^k_w} \leq \epsilon^2 C \|\phi\|_{H^{k+2}_w}.
$$

\textbf{(A44)}

\textbf{A.4 \quad Coarse graining A}\n
\textbf{Lemma A10.} Let $U : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be in $H^s$ with $s > 2$. Put $U_{\epsilon}(x) := U(\epsilon x)$. Then

$$
\lim_{\epsilon \to 0^+} \left\| \mathcal{L} S[U_{\epsilon}](\cdot/\epsilon) - U \right\|_{L^2} = 0,
$$

\textbf{(A45)}

where $\mathcal{L}$ and $S$ are defined by (153) and (151).

\textbf{Proof.} The proof for the 1D case is given in Lemma 5.2 in McGinnis and Wright.\footnote{15} The 2D version is analogous and also requires a nontrivial result regarding band-limited functions found in McNamee et al.\footnote{26}