Commun. Math. Phys. 367, 33–87 (2019)
Digital Object Identifier (DOI) https://doi.org/10.1007/s00220-019-03408-5

Communications in
Mathematical Physics

Sandpiles on the Square Lattice

Robert D. Hough[1], **Daniel C. Jerison**[2,3], **Lionel Levine**[2]

[1] Department of Mathematics, Stony Brook University, Stony Brook, NY 11794, USA. E-mail: robert.hough@stonybrook.edu
[2] Department of Mathematics, Cornell University, Malott Hall, Ithaca, NY 14853, USA. E-mail: levine@math.cornell.edu
[3] Department of Mathematics, Tel Aviv University, Tel Aviv, Israel. E-mail: dcjerison@gmail.com; jerison@mail.tau.ac.il

Received: 20 July 2017 / Accepted: 15 December 2017
Published online: 16 March 2019 – © Springer-Verlag GmbH Germany, part of Springer Nature 2019

Abstract: We give a non-trivial upper bound for the critical density when stabilizing i.i.d. distributed sandpiles on the lattice $\mathbb{Z}^2$. We also determine the asymptotic spectral gap, asymptotic mixing time, and prove a cutoff phenomenon for the recurrent state abelian sandpile model on the torus $(\mathbb{Z}/m\mathbb{Z})^2$. The techniques use analysis of the space of functions on $\mathbb{Z}^2$ which are harmonic modulo 1. In the course of our arguments, we characterize the harmonic modulo 1 functions in $\ell^p(\mathbb{Z}^2)$ as linear combinations of certain discrete derivatives of Green's functions, extending a result of Schmidt and Verbitskiy (Commun Math Phys 292(3):721–759, 2009. arXiv:0901.3124 [math.DS]).

1. Introduction

1.1. Stabilization of i.i.d. sandpiles. A *sandpile* on the integer lattice $\mathbb{Z}^2$ is a function $\sigma : \mathbb{Z}^2 \to \mathbb{Z}_{\geq 0}$, where $\sigma(x)$ represents the number of grains of sand at the site x. The sandpile σ is *stable* if each $\sigma(x) \leq 3$. If some $\sigma(x) \geq 4$, then we may *topple* the sandpile at x by passing one grain of sand from x to each of its four nearest neighbors. We say that σ *stabilizes* if it is possible to reach a stable configuration from σ by toppling each vertex finitely many times. If the heights $(\sigma(x))_{x\in\mathbb{Z}^2}$ are i.i.d. random variables, we refer to σ as an i.i.d. sandpile.

Meester and Quant [32] asked which i.i.d. sandpiles stabilize almost surely. It was proved by Fey and Redig [19] that such a sandpile σ must satisfy $\mathbf{E}[\sigma(x)] \leq 3$. This condition is not sufficient for stabilization: for every $p > 0$, the i.i.d. sandpile where each $\sigma(x) = 2$ with probability $1 - p$ and $\sigma(x) = 4$ with probability p almost surely fails to stabilize [16]. Thus, for each $2 < \rho \leq 3$, there are some i.i.d. sandpiles with $\mathbf{E}[\sigma(x)] = \rho$ that do stabilize almost surely (e.g. when each $\sigma(x) \in \{0, 1, 2, 3\}$) and others that fail

This material is based upon work supported by the National Science Foundation under Agreements No. DMS-1128155, http://www.nsf.gov/awardsearch/showAward?AWD_ID=1455272DMS-1455272, DMS-1712682, and DMS-1802336. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

to stabilize. This behavior contrasts with the closely related *divisible sandpile model*, in which stabilization of a nonconstant i.i.d. initial condition σ is determined entirely by the value of $\mathbf{E}[\sigma(x)]$ [30].

Our first main theorem shows that an i.i.d. sandpile with $\mathbf{E}[\sigma(x)]$ slightly less than 3 cannot stabilize almost surely unless $\sigma(x) \leq 3$ with high probability.

Theorem 1. *There are constants $c, d > 0$ such that any i.i.d. sandpile σ on $\mathbb{Z}^2$ that stabilizes almost surely satisfies*

$$\mathbf{E}[\sigma(x)] \leq 3 - \min\left(c, d\mathbf{E}[|X - X'|^{2/3}]\right) \tag{1}$$

where X, X' are independent and distributed as $\sigma(x)$.

If $3 - \mathbf{E}[\sigma(x)]$ is small, then the inequality $\mathbf{Prob}(X \neq X') \leq \mathbf{E}[|X - X'|^{2/3}]$ implies that the law of $\sigma(x)$ is concentrated at a single value, which must be at most 3. Some extra work would be required to extract explicit values for the constants c and d from our proof of Theorem 1; see the discussion following Lemma 15.

Theorem 1 answers a question posed by Fey et al. [18] by demonstrating that an i.i.d. Poisson sandpile with mean sufficiently close to 3 almost surely does not stabilize. An interesting question that remains open is whether there exists $\epsilon > 0$ such that the only i.i.d. stabilizing sandpiles with $\mathbf{E}[\sigma(x)] > 3 - \epsilon$ are those which are already stable.

1.2. Cutoff for sandpiles on the torus. We also consider sandpile dynamics on the discrete torus $\mathbb{T}_m = (\mathbb{Z}/m\mathbb{Z})^2$, given as follows. The point $(0, 0)$ is designated *sink* and is special. Each non-sink point on the torus has a sand allocation

$$\sigma : \mathbb{T}_m \backslash \{(0, 0)\} \to \mathbb{Z}_{\geq 0}. \tag{2}$$

As on the integer lattice, if at some time a non-sink vertex has allocation at least 4 it may topple, passing one grain of sand to each of its neighbors; if a grain of sand falls on the sink it is lost from the model. Those states $\mathscr{S}_m$ for which $\sigma \leq 3$ are stable. We consider the discrete time dynamics, where a single step consists of dropping a grain of sand on a uniformly randomly chosen vertex and then performing all legal topplings until the model reaches a stable state. The *abelian property* [11] ensures that this stable state does not depend on the order in which the topplings were performed.

Those stable states $\mathscr{R}_m$ which may be reached from the maximal state $\sigma \equiv 3$ are recurrent, whereas all other states are transient. Started from any stable state, the sandpile model forms a Markov chain with transition kernel P_m, which converges to the uniform measure $\mathbb{U}_{\mathscr{R}_m}$ on recurrent states.

Theorem 2. *Let $m \geq 2$. There is a constant $c_0 = 0.348661174(3)$ and $t_m^{\text{mix}} = c_0 m^2 \log m$ such that the following holds. For each fixed $\epsilon > 0$,*

$$\begin{aligned} \lim_{m\to\infty} \min_{\sigma\in\mathscr{S}_m} \left\| P_m^{\lceil(1-\epsilon)t_m^{\text{mix}}\rceil}\delta_\sigma - \mathbb{U}_{\mathscr{R}_m}\right\|_{\text{TV}(\mathscr{S}_m)} &= 1, \\ \lim_{m\to\infty} \max_{\sigma\in\mathscr{S}_m} \left\| P_m^{\lfloor(1+\epsilon)t_m^{\text{mix}}\rfloor}\delta_\sigma - \mathbb{U}_{\mathscr{R}_m}\right\|_{\text{TV}(\mathscr{S}_m)} &= 0. \end{aligned} \tag{3}$$

Informally, the convergence to uniformity of the sandpile model on the torus has total variation mixing time asymptotic to $c_0 m^2 \log m$ and the transition to uniformity satisfies a cutoff phenomenon. Implicit in the statement of Theorem 2 is that, with high probability, the time to reach a recurrent state started from a general state in the model is less than the mixing time. In Sect. 5 we give an easy proof using a coupon collector-type argument that this hitting time is almost surely $O(m^2\sqrt{\log m})$. Also, the asymptotic mixing time of order $m^2 \log m$ is at a later point than is sampled in some statistical physics studies regarding sandpiles, see [36].

We also determine asymptotically the absolute spectral gap of the torus sandpile Markov chain.

Theorem 3. *Let $m \geq 1$. There is a constant*

$$\gamma = 2.868114013(4)$$

such that the absolute spectral gap of the sandpile Markov chain restricted to its recurrent states satisfies

$$\mathrm{gap}_m = \frac{\gamma + o(1)}{m^2} \qquad \text{as } m \to \infty. \tag{4}$$

The constants in the preceding theorems are reciprocals: $c_0\gamma = 1$. An explicit formula for γ in terms of the Green's function on $\mathbb{Z}^2$ is given in "Appendix B".

1.3. Functions harmonic modulo 1. Functions which are *harmonic modulo* 1 play a central role in the proofs of Theorems 1–3. For $X = \mathbb{Z}^2$ or $X = \mathbb{T}_m$, we say that $f : X \to \mathbb{C}$ is harmonic modulo 1 if

$$(\Delta f)(i, j) := 4f(i, j) - f(i-1, j) - f(i+1, j) - f(i, j-1) - f(i, j+1) \tag{5}$$

is in $\mathbb{Z}$ for all $(i, j) \in X$. The operator Δ is the *graph Laplacian* on X.

Schmidt and Verbitskiy [35] characterized the set of all functions in $\ell^1(\mathbb{Z}^2)$ that are harmonic modulo 1. Their result can be stated using discrete derivatives of the *Green's function* on $\mathbb{Z}^2$. Let

$$\nu := \frac{1}{4}\left(\delta_{(-1,0)} + \delta_{(1,0)} + \delta_{(0,-1)} + \delta_{(0,1)}\right) \tag{6}$$

be the measure that drives simple random walk on $\mathbb{Z}^2$, and let ν^{*n} be its n-th convolution power, so that $\nu^{*n}(x)$ is the probability that a random walker started from the origin is at site x after n steps. The Green's function is defined by

$$G_{\mathbb{Z}^2}(x) := \frac{1}{4}\sum_{n=0}^{\infty}\left[\nu^{*n}(x) - \nu^{*n}(0, 0)\right]. \tag{7}$$

Evidently, $G_{\mathbb{Z}^2}(0, 0) = 0$. For nonzero $x = (x_1, x_2)$ with $\|x\|_2 = \sqrt{x_1^2 + x_2^2}$, it is known classically that $G_{\mathbb{Z}^2}(x) = -\frac{1}{2\pi}\log\|x\|_2 + O(1)$. As shown in [20], this is the start of an asymptotic expansion, whose first few terms we quote in Theorem 6.

It can easily be shown that $\Delta G_{\mathbb{Z}^2}(x) = \mathbf{e}_{(0,0)}(x) := \mathbf{1}\{x = (0, 0)\}$, so $G_{\mathbb{Z}^2}$ is harmonic modulo 1. By taking discrete derivatives, we can find harmonic modulo 1

functions that decay to zero with $\|x\|_2$. The discrete derivatives $D_1 f$, $D_2 f$ of any $f : \mathbb{Z}^2 \to \mathbb{C}$ are defined as

$$D_1 f(i, j) := f(i+1, j) - f(i, j), \quad D_2 f(i, j) := f(i, j+1) - f(i, j). \tag{8}$$

If f is harmonic modulo 1, then so is any finite linear combination with integer coefficients of translates of f, including $D_1 f$ and $D_2 f$.

From the asymptotic expansion, it follows that the k-th derivatives of $G_{\mathbb{Z}^2}$ decay like the inverse k-th power of the radius. That is, if $a + b = k$, then $D_1^a D_2^b G_{\mathbb{Z}^2}(x) = O\left(\|x\|_2^{-k}\right)$. When $k \geq 3$, this implies that $D_1^a D_2^b G_{\mathbb{Z}^2} \in \ell^1(\mathbb{Z}^2)$. Thus, the third derivatives of $G_{\mathbb{Z}^2}$, and all finite integer linear combinations of their translates, are harmonic modulo 1 functions in $\ell^1(\mathbb{Z}^2)$. (Note that the fourth and higher derivatives are linear combinations of translates of the third derivatives.)

For $1 \leq p < \infty$, let $\mathscr{H}^p(\mathbb{Z}^2)$ be the set of all functions in $\ell^p(\mathbb{Z}^2)$ that are harmonic modulo 1. Also, let $\langle\!\langle f_1, \ldots, f_n \rangle\!\rangle$ denote the set of all finite integer linear combinations of translates of the functions $f_1, \ldots, f_n$ on the domain $\mathbb{Z}^2$, so that for example $D_1^a D_2^b f \in \langle\!\langle f \rangle\!\rangle$ for any $a, b \geq 0$.

Theorem 4. *The sets $\mathscr{H}^p(\mathbb{Z}^2)$, for $1 \leq p < \infty$, admit the following characterization:*

$$\begin{aligned}
\mathscr{H}^1(\mathbb{Z}^2) &= \langle\!\langle D_1^3 G_{\mathbb{Z}^2}, D_1^2 D_2 G_{\mathbb{Z}^2}, D_1 D_2^2 G_{\mathbb{Z}^2}, D_2^3 G_{\mathbb{Z}^2}, \mathbf{e}_{(0,0)} \rangle\!\rangle \\
\mathscr{H}^p(\mathbb{Z}^2) &= \langle\!\langle D_1^2 G_{\mathbb{Z}^2}, D_1 D_2 G_{\mathbb{Z}^2}, D_2^2 G_{\mathbb{Z}^2} \rangle\!\rangle, \qquad 1 < p \leq 2 \\
\mathscr{H}^p(\mathbb{Z}^2) &= \langle\!\langle D_1 G_{\mathbb{Z}^2}, D_2 G_{\mathbb{Z}^2} \rangle\!\rangle, \qquad 2 < p < \infty.
\end{aligned} \tag{9}$$

The first equality in (9), which is the most delicate part to prove, is essentially a restatement of Theorem 2.4 in [35]. We provide a unified proof of all three parts of Theorem 4 in Sect. 3.

Since the function $\mathbf{e}_{(0,0)}$ is itself in $\mathscr{H}^p(\mathbb{Z}^2)$ for all p, it is implicit in the theorem statement that $\mathbf{e}_{(0,0)}$ is a linear combination of translates of second derivatives of $G_{\mathbb{Z}^2}$. This is true because $\Delta G_{\mathbb{Z}^2} = \mathbf{e}_{(0,0)}$, and the Laplacian Δ is a second-order discrete differential operator.

1.4. Discussion of method. This section outlines the methods used to prove Theorems 1–3.

Theorem 1 says that if σ is an i.i.d. sandpile on $\mathbb{Z}^2$ that stabilizes almost surely, then $3 - \mathbf{E}[\sigma(x)]$ is bounded below by a quantity that measures the typical difference between the heights at two locations $\sigma(x), \sigma(x')$. To prove the theorem, let $u(x)$ be the 'odometer' function that counts the number of times a vertex x topples in passing from σ to its stabilization σ^∞, so that $\sigma^\infty = \sigma - \Delta u$.

In Sect. 4 we observe that the modulo 1 harmonic functions are dual to toppling in the following sense: If $\xi \in \ell^1(\mathbb{Z}^2)$ is harmonic modulo 1, then

$$\langle \sigma, \xi \rangle \equiv \langle \sigma^\infty, \xi \rangle \mod 1, \qquad a.s. \tag{10}$$

where $\langle f, g \rangle = \sum_{x \in \mathbb{Z}^2} \overline{f(x)} g(x)$ is the usual pairing. This provides a collection of invariants which obstruct stabilization in the sandpile model.

To prove Theorem 1, we consider the characteristic functions

$$\chi(\sigma; \xi) = \mathbf{E}\left[e^{-2\pi i \langle \sigma, \xi \rangle}\right], \qquad \chi(\sigma^\infty; \xi) = \mathbf{E}\left[e^{-2\pi i \langle \sigma^\infty, \xi \rangle}\right] \tag{11}$$

which, by (10), are equal. If $\mathbf{E}[\sigma(x)] = \mathbf{E}[\sigma^\infty(x)]$ is close to 3 (which is the maximum possible value), then $\sigma^\infty(x)$ must equal 3 for most $x \in \mathbb{Z}^2$. Choosing ξ so that $\sum_{x\in\mathbb{Z}^2} \xi(x) = 0$, $\chi(\sigma^\infty;\xi)$ must be near $\chi(3;\xi) = 1$. On the other hand, since the starting values $\sigma(x)$ are i.i.d.,

$$\chi(\sigma;\xi) = \prod_{x\in\mathbb{Z}^2} \mathbf{E}\left[e^{-2\pi i\sigma(x)\xi(x)}\right]. \tag{12}$$

The modulus of each term $\mathbf{E}\left[e^{-2\pi i\sigma(x)\xi(x)}\right]$ decreases as the possible values of $\sigma(x)$ get more spread-out. In this way, the lower bound on $|\chi(\sigma;\xi)| = |\chi(\sigma^\infty;\xi)|$ translates into an upper bound on the amount that the starting values $\sigma(x)$ can vary.

We now turn to Theorem 3. The set $\mathscr{R}_m$ of recurrent sandpiles on the torus has a natural abelian group structure. This identifies $\mathscr{R}_m$ with the *sandpile group* $\mathscr{G}_m$, which is formally defined in Sect. 5. The sandpile Markov chain restricted to its recurrent states is a random walk on $\mathscr{G}_m$, meaning that its eigenvectors are given by the dual group $\hat{\mathscr{G}}_m$. We can express $\hat{\mathscr{G}}_m$ as the additive group of functions $\xi : \mathbb{T}_m \to \mathbb{R}/\mathbb{Z}$ such that $\xi(0,0) = 0$ and $\Delta\xi \equiv 0$ in $\mathbb{R}/\mathbb{Z}$. (The operation of ξ on sandpiles is $\sigma \mapsto \sum_{x\in\mathbb{T}_m\setminus\{(0,0)\}} \xi(x)\sigma(x)$.) In this way, an element $\xi \in \hat{\mathscr{G}}_m$ is naturally associated with the set of harmonic modulo 1 functions $\xi' : \mathbb{T}_m \to \mathbb{R}$ that reduce mod $\mathbb{Z}$ to ξ.

The eigenvalue of the Markov chain associated to ξ is the Fourier coefficient of the measure μ driving the random walk at frequency ξ:

$$\hat{\mu}(\xi) = \frac{1}{m^2} \sum_{x\in\mathbb{T}_m} e^{2\pi i\xi(x)}. \tag{13}$$

The mixing time is controlled by the frequencies for which $|\hat{\mu}(\xi)|$ is close to 1.

Given a frequency ξ, let $\xi' : \mathbb{T}_m \to \mathbb{R}$ be one of its harmonic modulo 1 representatives. The integer-valued function $v = \Delta\xi'$ will be referred to as a 'prevector' of ξ. To recover ξ' from v up to an additive constant, we convolve v with the Green's function $G_{\mathbb{T}_m}$ on the torus, which is defined by

$$G_{\mathbb{T}_m}(x) := \frac{1}{4} \sum_{n=0}^{\infty} \left(v^{*n}(x) - \frac{1}{m^2} \right) \tag{14}$$

and is the unique mean-zero function (i.e. $\sum_{x\in\mathbb{T}_m} G_{\mathbb{T}_m}(x) = 0$) satisfying

$$\Delta G_{\mathbb{T}_m}(x) = \mathbf{e}_{(0,0)}(x) - \frac{1}{m^2}. \tag{15}$$

It follows that $(G_{\mathbb{T}_m} * v)(x) = \xi'(x) - c$, where $c = \frac{1}{m^2}\sum_{y\in\mathbb{T}_m} \xi'(y)$.

Although we will not use this characterization, $G_{\mathbb{T}_m}$ can be considered as a mean-zero version of the Green's function for the simple random walk on $\mathbb{T}_m$ started from the origin and killed at a uniformly random point. To be precise, given $y \in \mathbb{T}_m$, let τ_y be the first time $t \geq 0$ that a simple random walker started from the origin reaches y, and define $g_y(x)$ to be the expected number of times $0 \leq t < \tau_y$ that the walker visits site x. If $g(x) = \frac{1}{m^2}\sum_{y\in\mathbb{T}_m} g_y(x)$, then $G_{\mathbb{T}_m}(x) = \frac{1}{4}\left[g(x) - \frac{1}{m^2}\sum_{x'\in\mathbb{T}_m} g(x')\right]$.

In Sect. 5.2, we specify for each frequency $\xi \in \hat{\mathscr{G}}_m$ a particular choice of ξ' such that the 'distinguished prevector' $v = \Delta\xi'$ satisfies

$$1 - |\hat{\mu}(\xi)| \asymp \frac{\|G_{\mathbb{T}_m} * v\|^2_{L^2(\mathbb{T}_m)}}{m^2}. \tag{16}$$

Each prevector v has mean zero because v is in the image of Δ. To find the absolute spectral gap of the Markov chain, which minimizes $1 - |\hat{\mu}(\xi)|$, we ask which mean-zero integer-valued vectors v make $\|G_{\mathbb{T}_m} * v\|^2_{L^2(\mathbb{T}_m)}$ as small as possible.

It is profitable to think of $G_{\mathbb{T}_m} * v$ as a linear combination of translates of discrete derivatives of $G_{\mathbb{T}_m}$. For example, if $v(a, b) = -1$, $v(a-1, b) = 1$, and $v(i, j) = 0$ at all other $(i, j) \in \mathbb{T}_m$, then

$$(G_{\mathbb{T}_m} * v)(x_1, x_2) = G_{\mathbb{T}_m}(x_1 + 1 - a, x_2 - b) - G_{\mathbb{T}_m}(x_1 - a, x_2 - b) \tag{17}$$

which is the translation by (a, b) of $D_1 G_{\mathbb{T}_m}$.

The Laplacian operator Δ acts locally. Its inverse, convolution with $G_{\mathbb{T}_m}$, is non-local but satisfies an approximate locality in that the discrete derivatives of $G_{\mathbb{T}_m}$, like those of $G_{\mathbb{Z}^2}$, decay to zero. Using these decay estimates, we show in Sect. 6.1 that $\|G_{\mathbb{T}_m} * v\|^2_{L^2(\mathbb{T}_m)}$ is minimized when $G_{\mathbb{T}_m} * v$ is an integer linear combination of the second derivatives $D_1^2 G_{\mathbb{T}_m}$, $D_1 D_2 G_{\mathbb{T}_m}$, $D_2^2 G_{\mathbb{T}_m}$ and their translates. These lead to gaps of order $1/m^2$ in (16).

It follows from (16), an upper bound on $\|\Delta\|_{L^2 \to L^2}$, and the inequality

$$\|v\|^2_{L^2(\mathbb{T}_m)} = \|\Delta(G_{\mathbb{T}_m} * v)\|^2_{L^2(\mathbb{T}_m)} \le \|\Delta\|^2_{L^2 \to L^2} \|G_{\mathbb{T}_m} * v\|^2_{L^2(\mathbb{T}_m)} \tag{18}$$

that if the L^2 norm of the prevector v is too high, then v cannot generate the spectral gap. Proposition 20 shows that if the support of v is too spread-out over $\mathbb{T}_m$, then by the approximate locality of convolution with $G_{\mathbb{T}_m}$, v can be separated into widely spaced clusters whose contributions to $1 - |\hat{\mu}(\xi)|$ are nearly additive. Just keeping one of the clusters and zeroing out the rest of v would produce a smaller gap. By this argument, the only prevectors with any chance of generating the spectral gap have bounded norm and bounded support, so the computation of the gap is reduced to a finite check.

To fill in the details of the proof, we require precise asymptotics for derivatives of $G_{\mathbb{T}_m}$. We obtain these using a local limit theorem, which is proved in "Appendix A". We also relate $G_{\mathbb{T}_m}$ as $m \to \infty$ to $G_{\mathbb{Z}^2}$, which translates the finite check for the spectral gap into a minimization problem involving functions in $\ell^2(\mathbb{Z}^2)$ that are harmonic modulo 1. The resulting search was performed using convex programming in the SciPy scientific computing package [27], and is described in "Appendix B". We find that for sufficiently large m, the gap is achieved for prevectors of the form $v(a, b) = v(a-1, b-1) = 1$, $v(a-1, b) = v(a, b-1) = -1$, $v(i, j) = 0$ elsewhere, which correspond to translates of $D_1 D_2 G_{\mathbb{T}_m}$.

For Theorem 2, we prove cutoff in both total variation and L^2 at time $\gamma^{-1} m^2 \log m$. The necessary ingredients are a total variation lower bound and an L^2 upper bound on mixing time.

First, we use the coupon-collector argument mentioned earlier to reduce to the case where the starting state σ is recurrent. Next we observe that, due to translation, there are m^2 different prevectors v whose corresponding frequencies ξ achieve $1 - |\hat{\mu}(\xi)| = \text{gap}_m$.

The L^2 distance from the uniform distribution on $\mathscr{R}_m$ of the chain started from σ after N steps satisfies

$$\left\| P_m^N \delta_\sigma - \mathbb{U}_{\mathscr{R}_m} \right\|^2_{L^2(d\mathbb{U}_{\mathscr{R}_m})} = \sum_{\xi \in \hat{\mathscr{G}}_m \setminus \{0\}} \left|\hat{\mu}(\xi)\right|^{2N} \geq m^2 (1 - \text{gap}_m)^{2N}. \tag{19}$$

Thus, the chain cannot mix in L^2 before time

$$N = \frac{1}{\text{gap}_m} \log m = \frac{1}{\gamma} m^2 \log m + o(m^2 \log m). \tag{20}$$

We strengthen this to a lower bound on total variation mixing time by a second moment method due originally to Diaconis [13,15] that builds a distinguishing statistic out of the top eigenvectors of the chain. See Lemma 27. To apply this lemma, we require an upper bound on $|\hat{\mu}(\xi_1 - \xi_2)|$ when the frequencies ξ_1, ξ_2 (both of which achieve the spectral gap) come from prevectors v_1, v_2 whose supports are separated. Since the contributions of v_1 and v_2 are nearly additive, we have $1 - |\hat{\mu}(\xi_1 - \xi_2)| \approx 1 - 2 \cdot \text{gap}_m$, which enables the argument to go through.

For the upper bound on the L^2 mixing time, we show that

$$\sum_{\xi \in \hat{\mathscr{G}}_m \setminus \{0\}} \left|\hat{\mu}(\xi)\right|^{2N} \tag{21}$$

tends to zero as $m \to \infty$ when $N = (1 + \epsilon)\gamma^{-1} m^2 \log m$. Our argument uses an agglomeration scheme in which we partition the support of each prevector v into widely spaced clusters. Lemma 25, the main step in the proof, shows that each small cluster contributes additively to the gap $1 - |\hat{\mu}(\xi)|$. The earlier additivity results in Sect. 6.1, most notably Proposition 20, hold only for prevectors with bounded L^1 norm, so the extension to the general case requires new arguments. We use techniques from the theory of exponential sums, including van der Corput's inequality. As a consequence of the clustering scheme, we can control the number of distinct frequencies ξ whose gap $1 - |\hat{\mu}(\xi)|$ might be small, giving the desired bound on (21).

The cutoff argument may be considered an extension of the classical analysis of mixing on the hypercube [14], and exploits the fact that the lattice which is quotiented to give the sandpile group is approximately cubic. See [22] for analysis of some random walks on the cycle where the L^1 and L^2 cutoff times differ by a constant.

1.5. Historical review. Sandpile dynamics on the square lattice were introduced by Bak et al. [3,4] as a model of self-organized criticality. Dhar [11] considered the case of an arbitrary finite underlying graph, proving many fundamental results. Subsequently, Dhar et al. [10] used harmonic modulo 1 functions (there called 'toppling invariants') to analyze the algebraic structure of the sandpile group for rectangular subsets of $\mathbb{Z}^2$.

Sandpiles are examples of *abelian networks*, which are systems of communicating automata satisfying a local commutativity condition [7,12]. By a theorem of Cairns [8], an abelian network on $\mathbb{Z}^2$ can emulate a Turing machine, as can a sandpile on $\mathbb{Z}^3$. In particular, for a periodic configuration of sand on $\mathbb{Z}^3$ plus a finite number of additional sand grains, the question of stabilization is algorithmically undecidable! It is not known whether the same question is undecidable on $\mathbb{Z}^2$. A related open problem, highlighted in [30], is the following: "Given a probability distribution μ on $\mathbb{Z}$ (say, supported on

$\{0, 1, 2, 3, 4\}$ with rational probabilities), is it algorithmically decidable whether the i.i.d. abelian sandpile on $\mathbb{Z}^2$ with marginal μ stabilizes almost surely?" Theorem 1 and its method of proof can be viewed as a slight advance on this problem.

The question of stabilization of i.i.d. sandpiles was posed by Meester and Quant [32] and by Fey and Redig [19]. A fundamental result is the *conservation of density* proved by Fey et al. [18], which in particular implies the earlier result of [19]: An i.i.d. stabilizing sandpile σ on $\mathbb{Z}^2$ must satisfy $\mathbf{E}[\sigma(x)] \leq 3$. To get strictly below 3 in the upper bound of Theorem 1, we use harmonic modulo 1 functions to construct additional conserved quantities; see Lemma 14.

Theorems 2 and 3 are concerned with the sandpile Markov chain on the discrete torus $\mathbb{T}_m$, whose stationary distribution is uniform on the (finite) set of recurrent states. These finite Markov chains are related to sandpiles on the infinite grid $\mathbb{Z}^2$ by theorems of [2,24]. Athreya and Járai [2] proved that the restriction of a uniform recurrent sandpile on the d-dimensional cube $[-m, m]^d \cap \mathbb{Z}^d$ to any fixed finite subset of $\mathbb{Z}^d$ converges in law as $m \to \infty$. Hence, there is a limiting measure μ on recurrent sandpiles on $\mathbb{Z}^d$. By equality of the free and wired uniform spanning forests, replacing the cube with the d-dimensional discrete torus results in the same limit μ. Járai and Redig [24] proved that in dimensions $d \geq 3$, a μ-distributed sandpile plus one additional chip stabilizes almost surely. They used this fact to construct an ergodic Markov process on recurrent sandpiles on $\mathbb{Z}^d$ having μ as its stationary distribution. In dimension 2, it is not known whether a μ-distributed sandpile plus one additional chip stabilizes almost surely. (Possibly Lemma 14 could help resolve this question.) Some further studies of sandpile dynamics on $\mathbb{Z}^d$ are [6,25,31].

The mixing of the sandpile Markov chain on finite graphs arises in relating sandpiles with different boundary conditions: the dependence of observables such as the 'density' (average amount of sand per vertex) on the boundary conditions is a symptom of slow mixing. In particular, the extra log factor in the mixing time t_m^{mix} of Theorem 2 could be viewed as the cause for the failure of the 'density conjecture' [17,29].

The proof of cutoff in Theorem 2 estimates a significant piece of the spectrum of the transition kernel of the sandpile walk on the torus. See [5,9] for further applications of spectral techniques related to sandpiles.

The eigenvectors and eigenvalues of the sandpile Markov chain on an arbitrary finite graph were characterized in [26] using 'multiplicative harmonic functions' (these are complex exponentials of the harmonic modulo 1 functions, as explained in Sect. 5.1). In [26] it was shown that the sandpile Markov chain on any connected graph with n vertices mixes in $O(n^3 \log n)$ steps, and that cutoff for the complete graph (both in total variation and in L^2) occurs at time $\frac{1}{4\pi^2} n^3 \log n$.

Regarding the discrete torus $\mathbb{T}_m$, it was proved in [26] that the sandpile chain on any graph with m^2 vertices and maximum degree 4 has spectral gap at least $1/(2m^2)$, and mixes in at most $\frac{5}{2} m^2 \log m$ steps. Theorems 2 and 3 improve these results by obtaining asymptotics for the mixing time and spectral gap, and by demonstrating cutoff. We expect that our techniques can also prove cutoff for the sandpile chain on the finite box $[-m, m]^2 \cap \mathbb{Z}^2$, with boundary vertices identified as the sink, at a constant multiple of $m^2 \log m$ steps.

Schmidt and Verbitskiy [35] characterize the set $\mathscr{H}^1(\mathbb{Z}^2)$ of harmonic modulo 1 functions in $\ell^1(\mathbb{Z}^2)$ in terms of third derivatives of the Green's function $G_{\mathbb{Z}^2}$. Our Theorem 4 provides a similar characterization of the sets $\mathscr{H}^p(\mathbb{Z}^2)$, for $1 \leq p < \infty$.

Organization. Section 2 fixes notation and provides background on discrete derivatives and Fourier transforms, the graph Laplacian and Green's function on $\mathbb{Z}^2$ and $\mathbb{T}_m$, and results from the theory of exponential sums. Section 3 proves Theorem 4, while Sect. 4 proves Theorem 1. Section 5 defines the sandpile group $\mathscr{G}_m$ and its dual $\hat{\mathscr{G}}_m$, describes the eigenvalues and eigenvectors of the sandpile Markov chain using $\hat{\mathscr{G}}_m$, and shows that we may assume the starting state is recurrent when proving Theorem 2. Section 6 proves Theorem 3 and provides the main technical estimates needed for Theorem 2. Finally, Sect. 7 proves Theorem 2.

"Appendix A" proves a local limit theorem for repeated convolutions of the simple random walk measure on $\mathbb{Z}^2$ that is used to obtain asymptotics for derivatives of the discrete Green's function on $\mathbb{T}_m$. "Appendix B" uses convex programming to find an exact formula for the leading constant γ in the spectral gap and mixing time of the sandpile chain.

2. Function Spaces and Conventions

The additive character on $\mathbb{R}/\mathbb{Z}$ is $e(x) = e^{2\pi i x}$. Its real part is denoted $c(x) = \cos 2\pi x$ and imaginary part $s(x) = \sin 2\pi x$. For real x, $\|x\|_{\mathbb{R}/\mathbb{Z}}$ denotes the distance of x to the nearest integer.

We use the notations $A \ll B$ and $A = O(B)$ to mean that there is a constant $0 < C < \infty$ such that $|A| < CB$, and $A \asymp B$ to mean $A \ll B \ll A$. A subscript such as $A \ll_R B$, $A = O_R(B)$ means that the constant C depends on R. The notation $A = o(B)$ means that A/B tends to zero.

Given a measurable space $(\mathscr{X}, \mathscr{B})$, the *total variation distance* between two probability measures μ and ν on $(\mathscr{X}, \mathscr{B})$ is

$$\|\mu - \nu\|_{\mathrm{TV}} = \sup_{A \in \mathscr{B}} |\mu(A) - \nu(A)|. \tag{22}$$

If μ is absolutely continuous with respect to ν, the total variation distance may be expressed as

$$\|\mu - \nu\|_{\mathrm{TV}} = \frac{1}{2} \int_{\mathscr{X}} \left| \frac{d\mu}{d\nu} - 1 \right| d\nu. \tag{23}$$

In this case an $L^2(d\nu)$ distance may be defined by

$$\|\mu - \nu\|^2_{L^2(d\nu)} = \int_{\mathscr{X}} \left(\frac{d\mu}{d\nu} - 1 \right)^2 d\nu, \tag{24}$$

and Cauchy-Schwarz gives $\|\mu - \nu\|_{\mathrm{TV}} \le \frac{1}{2} \|\mu - \nu\|_{L^2(d\nu)}$.

Consider $\mathbb{Z}^2$ and the discrete torus $\mathbb{T}_m$ to be metric spaces with the graph distance given by the ℓ^1 norm on $\mathbb{Z}^2$ and the quotient distance, for $x, y \in \mathbb{T}_m$,

$$\|x - y\|_1 = \min\{\|x' - y'\|_1 : x', y' \in \mathbb{Z}^2, [x'] = x, [y'] = y\} \tag{25}$$

where $[x']$, $[y']$ are the images of x', y' under the quotient map $\mathbb{Z}^2 \to \mathbb{T}_m$. The ball of radius $R > 0$ around a point x is

$$B_R(x) = \{y : \|y - x\|_1 \le R\}. \tag{26}$$

We also use $\|x\|_2$ to denote the ℓ^2 norm of $x = (x_1, x_2) \in \mathbb{Z}^2$. The argument of x, denoted $\arg(x)$, is the angle $0 \le \theta < 2\pi$ such that $(x_1, x_2) = (\|x\|_2 \cos\theta, \|x\|_2 \sin\theta)$.

Denote the usual function spaces

$$\ell^p\left(\mathbb{Z}^2\right) = \left\{ f : \mathbb{Z}^2 \to \mathbb{C}, \|f\|_p^p = \sum_{x\in\mathbb{Z}^2} |f(x)|^p < \infty \right\}, \quad 1 \le p < \infty \tag{27}$$

and

$$L^p\left(\mathbb{T}_m\right) = \left\{ f : \mathbb{T}_m \to \mathbb{C}, \|f\|_p^p = \sum_{x\in\mathbb{T}_m} |f(x)|^p \right\}, \quad 1 \le p < \infty. \tag{28}$$

The latter functions may be considered as functions on $\mathbb{Z}^2$ which are $m\mathbb{Z}^2$-periodic. Let $\ell^\infty(\mathbb{Z}^2)$ and $L^\infty(\mathbb{T}_m)$ be the spaces of bounded functions on $\mathbb{Z}^2$ and $\mathbb{T}_m$, with $\|f\|_{\ell^\infty(\mathbb{Z}^2)} = \sup_{x\in\mathbb{Z}^2} |f(x)|$ and $\|f\|_{L^\infty(\mathbb{T}_m)} = \max_{x\in\mathbb{T}_m} |f(x)|$.

On the torus, the subspace of mean zero functions is indicated by

$$L_0^2(\mathbb{T}_m) = \left\{ f \in L^2(\mathbb{T}_m) : \sum_{x\in\mathbb{T}_m} f(x) = 0 \right\}. \tag{29}$$

The notation $\mathbb{Z}_0^{\mathbb{T}_m}$ is used for the integer-valued functions in $L_0^2(\mathbb{T}_m)$.

On either $\mathbb{Z}^2$ or the torus, the standard basis vectors are written

$$\mathbf{e}_{(i,j)}(k,\ell) = \mathbf{1}\{i = k\}\mathbf{1}\{j = \ell\}. \tag{30}$$

For functions other than the standard basis vectors, the notation $f_x = f(x)$ is used interchangeably.

For $X = \mathbb{Z}^2$ or $X = \mathbb{T}_m$ the support of a function on X is

$$\operatorname{supp} f = \{x \in X : f(x) \neq 0\}. \tag{31}$$

Given $(i, j) \in X$, the translation operator $T_{(i,j)}$ acts on functions by

$$T_{(i,j)} f(k,\ell) = f(k - i, \ell - j). \tag{32}$$

The convolution of functions $f \in \ell^1(\mathbb{Z}^2)$, $g \in \ell^\infty\left(\mathbb{Z}^2\right)$ or $f, g \in L^2(\mathbb{T}_m)$ is given by

$$(f * g)(i, j) = \sum_{(k,\ell)\in X} f(i - k, j - \ell) g(k, \ell) \tag{33}$$

where again X represents $\mathbb{Z}^2$ or $\mathbb{T}_m$.

The averaging operator with respect to the uniform probability measure on $\mathbb{T}_m$ is indicated by

$$\mathbf{E}_{x\in\mathbb{T}_m}[f] = \frac{1}{m^2} \sum_{x\in\mathbb{T}_m} f(x). \tag{34}$$

Given $x, y \in \mathbb{R}/\mathbb{Z}$ and $f \in \ell^1(\mathbb{Z}^2)$, the Fourier transform of f is

$$\hat{f}(x, y) = \sum_{(i,j)\in\mathbb{Z}^2} f(i, j) e(-(ix + jy)). \tag{35}$$

Given $x, y \in \mathbb{Z}/m\mathbb{Z}$ and $f \in L^2(\mathbb{T}_m)$ the Fourier transform of f is

$$\hat{f}(x, y) = \sum_{(i,j)\in\mathbb{T}_m} f(i, j)e\left(-\frac{ix + jy}{m}\right). \tag{36}$$

The Fourier transform has the familiar property of carrying convolution to pointwise multiplication. For $f \in \ell^2(\mathbb{Z}^2)$, Parseval's identity is

$$\|f\|_2^2 = \int_{(\mathbb{R}/\mathbb{Z})^2} \left|\hat{f}(x, y)\right|^2 dxdy. \tag{37}$$

For $f \in L^2(\mathbb{T}_m)$ the corresponding identity is

$$\|f\|_2^2 = \frac{1}{m^2} \sum_{x\in\mathbb{T}_m} \left|\hat{f}(x)\right|^2. \tag{38}$$

For a function f on $\mathbb{Z}^2$ or $\mathbb{T}_m$, the discrete derivatives $D_1 f(i, j)$, $D_2 f(i, j)$ are defined by (8). Discrete differentiation is expressed as a convolution operator by introducing

$$\delta_1(i, j) = \begin{cases} -1 & (i, j) = (0, 0) \\ 1 & (i, j) = (-1, 0) \\ 0 & \text{otherwise,} \end{cases}$$
$$\delta_2(i, j) = \begin{cases} -1 & (i, j) = (0, 0) \\ 1 & (i, j) = (0, -1) \\ 0 & \text{otherwise.} \end{cases} \tag{39}$$

For integers $a, b \geq 0$, one has

$$D_1^a D_2^b f = \delta_1^{*a} * \delta_2^{*b} * f. \tag{40}$$

For $X = \mathbb{Z}^2$ or $\mathbb{T}_m$ and functions $f_1, \ldots, f_n$ on X, recall that

$$\langle\!\langle f_1, \ldots, f_n \rangle\!\rangle = \text{span}_{\mathbb{Z}}\{T_x f_1, \ldots, T_x f_n : x \in X\}, \tag{41}$$

where $\text{span}_{\mathbb{Z}}$ refers to the finite integer span. It is convenient to introduce classes of integer-valued functions:

$$\begin{aligned} C^0(X) &= \langle\!\langle \mathbf{e}_{(0,0)} \rangle\!\rangle = \{f : X \to \mathbb{Z}, \|f\|_1 < \infty\}, \\ C^1(X) &= \langle\!\langle \delta_1, \delta_2 \rangle\!\rangle, \\ C^2(X) &= \langle\!\langle \delta_1^{*2}, \delta_1 * \delta_2, \delta_2^{*2} \rangle\!\rangle, \\ C^3(X) &= \langle\!\langle \delta_1^{*3}, \delta_1^{*2} * \delta_2, \delta_1 * \delta_2^{*2}, \delta_2^{*3} \rangle\!\rangle. \end{aligned} \tag{42}$$

One has the equivalent characterizations

$$C^1(X) = \left\{ f \in C^0(X) : \sum_{x\in X} f(x) = 0 \right\}, \tag{43}$$

$$C^2(X) = \left\{ f \in C^0(X) : \sum_{x\in X} f(x) = 0, \sum_{x\in X} f(x)x = 0 \right\} \tag{44}$$

and, for each $1 \leq k \leq 3$,

$$C^k(X) = \{\delta_1 * f + \delta_2 * g : f, g \in C^{k-1}(X)\}. \tag{45}$$

Note the special cases $C^0(\mathbb{T}_m) = \mathbb{Z}^{\mathbb{T}_m}$ and $C^1(\mathbb{T}_m) = \mathbb{Z}_0^{\mathbb{T}_m}$.

2.1. The graph Laplacian and Green's function. The graph Laplacian Δ on either $\mathbb{Z}^2$ or $\mathbb{T}_m$ is the second-order discrete differential operator defined by (5). On $\mathbb{Z}^2$ its Fourier transform is given by

$$\widehat{(\Delta f)}(x, y) = (4 - 2[c(x) + c(y)])\hat{f}(x, y), \quad x, y \in \mathbb{R}/\mathbb{Z}, \tag{46}$$

and on $\mathbb{T}_m$ the Fourier transform is

$$\widehat{(\Delta f)}(x, y) = (4 - 2[c(x/m) + c(y/m)])\hat{f}(x, y), \quad x, y \in \mathbb{Z}/m\mathbb{Z}. \tag{47}$$

Lemma 5. *The graph Laplacians satisfy the operator bound*

$$\begin{aligned} \|\Delta\|_{\ell^2(\mathbb{Z}^2)\to\ell^2(\mathbb{Z}^2)}, \|\Delta\|_{L^2(\mathbb{T}_m)\to L^2(\mathbb{T}_m)} &\leq 8, \\ \|\Delta\|_{\ell^\infty(\mathbb{Z}^2)\to\ell^\infty(\mathbb{Z}^2)}, \|\Delta\|_{L^\infty(\mathbb{T}_m)\to L^\infty(\mathbb{T}_m)} &\leq 8. \end{aligned} \tag{48}$$

Proof. The ℓ^∞ and L^∞ estimates are immediate. For $f \in \ell^2(\mathbb{Z}^2)$, by Parseval

$$\|\Delta f\|_2^2 = \int_{(\mathbb{R}/\mathbb{Z})^2} (4 - 2(c(x) + c(y)))^2 \left|\hat{f}(x, y)\right|^2 dxdy \leq 64\|f\|_2^2. \tag{49}$$

The bound on $L^2(\mathbb{T}_m)$ is similar. □

On either $X = \mathbb{Z}^2$ or $X = \mathbb{T}_m$, let ν be the probability measure given by (6), which drives simple random walk on X. The Green's function G is a distribution on $C^1(X)$ given by

$$G * f = \frac{1}{4}\sum_{n=0}^{\infty}(\nu^{*n} * f), \qquad f \in C^1(X). \tag{50}$$

Since $\Delta f = 4\left(\delta_{(0,0)} - \nu\right) * f$, the formal computation

$$\Delta^{-1} = \frac{1}{4}\left(\delta_{(0,0)} - \nu\right)^{-1} = \frac{1}{4}\sum_{n=0}^{\infty}\nu^{*n} = G \tag{51}$$

indicates that G is in some sense the inverse of Δ. Precise versions of this statement are given below.

On $\mathbb{Z}^2$, G may be realized as the function (7):

$$G_{\mathbb{Z}^2}(x) = \frac{1}{4}\sum_{n=0}^{\infty}\left[\nu^{*n}(x) - \nu^{*n}(0, 0)\right]. \tag{52}$$

This is a classical object of probability theory. We quote the asymptotics from [20].

Theorem 6 ([20], Remark 2). *Let $x = (x_1, x_2) \in \mathbb{Z}^2$. There are constants $a, b > 0$ such that*

$$G_{\mathbb{Z}^2}(x) = \begin{cases} 0 & x = (0,0) \\ -\frac{\log \|x\|_2}{2\pi} - a - b\frac{\frac{8x_1^2x_2^2}{\|x\|_2^4} - 1}{\|x\|_2^2} + O(\|x\|_2^{-4}) & x \neq (0,0). \end{cases} \tag{53}$$

It follows from (52) that

$$\Delta G_{\mathbb{Z}^2}(x) = \sum_{n=0}^{\infty} \left[\nu^{*n}(x) - \nu^{*(n+1)}(x) \right] = \mathbf{e}_{(0,0)}(x), \tag{54}$$

so $\Delta(G_{\mathbb{Z}^2} * f) = f$ for all $f \in C^0(\mathbb{Z}^2)$. The Fourier transform of $G_{\mathbb{Z}^2}$ is

$$\hat{G}_{\mathbb{Z}^2}(x, y) = \frac{1}{4 - 2\left(c(x) + c(y)\right)}. \tag{55}$$

When combined with (46), this shows that $G_{\mathbb{Z}^2} * \Delta f = f$ whenever $f \in \ell^2(\mathbb{Z}^2)$.

On $\mathbb{T}_m$ a realization of G as a function is obtained by (14):

$$G_{\mathbb{T}_m}(x) = \frac{1}{4} \sum_{n=0}^{\infty} \left(\nu^{*n}(x) - \frac{1}{m^2} \right). \tag{56}$$

This converges absolutely, as is most easily checked by passing to frequency space, where the zeroth Fourier coefficient vanishes, and the remaining Fourier coefficients are convergent geometric series. Summing (56) over all $x \in \mathbb{T}_m$ shows that $G_{\mathbb{T}_m}$ has mean zero. As well, it follows from (56) that

$$\Delta G_{\mathbb{T}_m}(x) = \sum_{n=0}^{\infty} \left[\nu^{*n}(x) - \nu^{*(n+1)}(x) \right] = \mathbf{e}_{(0,0)}(x) - \frac{1}{m^2}. \tag{57}$$

Therefore, $\Delta(G_{\mathbb{T}_m} * f) = f - \mathbf{E}_{x \in \mathbb{T}_m}[f]$ for any $f \in L^2(\mathbb{T}_m)$. In particular, if $f \in L_0^2(\mathbb{T}_m)$ then $\Delta(G_{\mathbb{T}_m} * f) = f$.

It is also true that $G_{\mathbb{T}_m} * \Delta f = f - \mathbf{E}_{x \in \mathbb{T}_m}[f]$ for all $f \in L^2(\mathbb{T}_m)$. To prove this, observe that since $\Delta f \in L_0^2(\mathbb{T}_m)$, $\Delta(G_{\mathbb{T}_m} * \Delta f) = \Delta f$. Only the constant functions are in the kernel of Δ, so $G_{\mathbb{T}_m} * \Delta f = f - c$ for some constant c. Since $G_{\mathbb{T}_m} * \Delta f$ has mean zero, $c = \mathbf{E}_{x \in \mathbb{T}_m}[f]$.

Both operators, Δ and convolution with $G_{\mathbb{T}_m}$, have image $L_0^2(\mathbb{T}_m)$. The two observations $\Delta(G_{\mathbb{T}_m} * f) = f - \mathbf{E}_{x \in \mathbb{T}_m}[f]$ and $G_{\mathbb{T}_m} * \Delta f = f - \mathbf{E}_{x \in \mathbb{T}_m}[f]$ imply that the composition in either order of the two operators results in orthogonal projection onto $L_0^2(\mathbb{T}_m)$. Restricted to $L_0^2(\mathbb{T}_m)$, the two operators are inverses. On $L^2(\mathbb{T}_m)$, $G_{\mathbb{T}_m}$ is the *Moore-Penrose pseudoinverse* of Δ.

We will require the following statements regarding discrete derivatives of $G_{\mathbb{T}_m}$. Recall that the notation $A \ll_{a,b} B$ means that there is a constant $0 < C < \infty$ depending on a, b such that $|A| \leq C B$.

Lemma 7. *For $a, b \in \mathbb{Z}_{\geq 0}$, $1 \leq a + b$, for $|i|, |j| \leq \frac{m}{2}$,*

$$D_1^a D_2^b G_{\mathbb{T}_m}(i, j) \ll_{a,b} \frac{1}{1 + \left(i^2 + j^2\right)^{\frac{a+b}{2}}}. \tag{58}$$

In the case $a + b = 1$ the following asymptotic evaluation holds.

Lemma 8. *Let $m \geq 2$ and $0 \leq i, j \leq \frac{m}{2}$. Set $R = i^2 + j^2$. There is a constant $c > 0$ such that, as $m \to \infty$, for $0 < R < \frac{m}{(\log m)^{\frac{1}{2}}}$,*

$$D_1 G_{\mathbb{T}_m}(i, j) = -\frac{ci}{i^2 + j^2} + O\left(\frac{1}{i^2 + j^2}\right),$$
$$D_2 G_{\mathbb{T}_m}(i, j) = -\frac{cj}{i^2 + j^2} + O\left(\frac{1}{i^2 + j^2}\right). \tag{59}$$

The proofs of Lemmas 7 and 8 are given in "Appendix A".

Lemma 9. *If $a+b \geq 2$ then $D_1^a D_2^b G_{\mathbb{Z}^2}$ is in $\ell^2(\mathbb{Z}^2)$ and for each fixed i, j, $D_1^a D_2^b G_{\mathbb{T}_m}(i, j) \to D_1^a D_2^b G_{\mathbb{Z}^2}(i, j)$ as $m \to \infty$.*

Proof. The Fourier transform of $D_1^a D_2^b G_{\mathbb{Z}^2}$ is given by, for $x, y \in \mathbb{R}/\mathbb{Z}$, not both 0,

$$\widehat{D_1^a D_2^b G_{\mathbb{Z}^2}}(x, y) = \frac{(e(x) - 1)^a \, (e(y) - 1)^b}{4 - 2(c(x) + c(y))}. \tag{60}$$

This function is bounded on $(\mathbb{R}/\mathbb{Z})^2$, which proves the first claim by Parseval.

The Fourier transform of $D_1^a D_2^b G_{\mathbb{T}_m}$ at frequency $(x, y) \in (\mathbb{Z}/m\mathbb{Z})^2$ is given by $\widehat{D_1^a D_2^b G_{\mathbb{Z}^2}}\left(\frac{x}{m}, \frac{y}{m}\right)$. Taking the group inverse Fourier transform,

$$D_1^a D_2^b G_{\mathbb{T}_m}(i, j) = \frac{1}{m^2} \sum_{x, y \in \mathbb{Z}/m\mathbb{Z}} \widehat{D_1^a D_2^b G_{\mathbb{Z}^2}}\left(\frac{x}{m}, \frac{y}{m}\right) e\left(\frac{ix + jy}{m}\right). \tag{61}$$

Treating this as a Riemann sum and letting $m \to \infty$ obtains the limit

$$D_1^a D_2^b G_{\mathbb{Z}^2}(i, j) = \int_{(\mathbb{R}/\mathbb{Z})^2} \widehat{D_1^a D_2^b G_{\mathbb{Z}^2}}\,(x, y)\, e\,(ix + jy)\, dxdy. \tag{62}$$

□

2.2. Exponential sums. This section collects the two results from the classical theory of exponential sums that are needed for the proof of Lemma 25, which is the key ingredient in the upper bound of Theorem 2. For further references, see [23,33,38].

The first result is van der Corput's inequality. We will only need the case $H = 1$. See [37] for a motivation and proof of this statement.

Theorem 10 (van der Corput's Lemma). *Let H be a positive integer. Then for any complex numbers $y_1, y_2, ..., y_N$,*

$$\left|\sum_{n=1}^{N} y_n\right|^2 \leq \frac{N + H}{H + 1} \sum_{n=1}^{N} |y_n|^2 + \frac{2(N + H)}{H + 1} \sum_{h=1}^{H} \left(1 - \frac{h}{H + 1}\right) \left|\sum_{n=1}^{N-h} y_{n+h} \overline{y_n}\right|. \tag{63}$$

The second result treats summation of a linear phase function and is fundamental.

Lemma 11. *Let $\alpha \in \mathbb{R}\backslash\mathbb{Z}$ and let $N \geq 1$. Then*

$$\left|\sum_{j=1}^{N} e(\alpha j)\right| \ll \min\left(N, \|\alpha\|_{\mathbb{R}/\mathbb{Z}}^{-1}\right). \tag{64}$$

Proof. Sum the geometric series. □

3. Classification of Functions Harmonic Modulo 1

This section proves Theorem 4. Let $1 \le p < \infty$, and recall that $\mathscr{H}^p(\mathbb{Z}^2)$ is the set of all harmonic modulo 1 functions in $\ell^p(\mathbb{Z}^2)$. If $f \in \mathscr{H}^p(\mathbb{Z}^2)$, then as $\|x\|_2 \to \infty$, $f(x) \to 0$ and therefore also $\Delta f(x) \to 0$. Since Δf is integer-valued, it must be identically zero outside a ball of finite radius. Thus,

$$\mathscr{H}^p\left(\mathbb{Z}^2\right) = \left\{ f \in \ell^p\left(\mathbb{Z}^2\right) : \Delta f \in C^0(\mathbb{Z}^2) \right\}, \qquad 1 \le p < \infty, \tag{65}$$

where $C^0(\mathbb{Z}^2)$ is the space of integer-valued functions on $\mathbb{Z}^2$ with finite support, as in (42).

From Theorem 6, we can derive the following formulas.

Lemma 12. *For nonzero $x = (x_1, x_2) \in \mathbb{Z}^2$, we have:*

$$\begin{aligned}
D_1 G_{\mathbb{Z}^2}(x_1, x_2) &= \frac{1}{2\pi} \cdot \frac{-x_1}{x_1^2 + x_2^2} + O\left(\|x\|_2^{-2}\right) \\
D_2 G_{\mathbb{Z}^2}(x_1, x_2) &= \frac{1}{2\pi} \cdot \frac{-x_2}{x_1^2 + x_2^2} + O\left(\|x\|_2^{-2}\right) \\
D_1^2 G_{\mathbb{Z}^2}(x_1, x_2) &= \frac{1}{2\pi} \cdot \frac{x_1^2 - x_2^2}{(x_1^2 + x_2^2)^2} + O\left(\|x\|_2^{-3}\right) \\
D_1 D_2 G_{\mathbb{Z}^2}(x_1, x_2) &= \frac{1}{2\pi} \cdot \frac{2x_1 x_2}{(x_1^2 + x_2^2)^2} + O\left(\|x\|_2^{-3}\right) \\
D_2^2 G_{\mathbb{Z}^2}(x_1, x_2) &= \frac{1}{2\pi} \cdot \frac{x_2^2 - x_1^2}{(x_1^2 + x_2^2)^2} + O\left(\|x\|_2^{-3}\right) \\
D_1^a D_2^b G_{\mathbb{Z}^2}(x_1, x_2) &= O\left(\|x\|_2^{-3}\right), \qquad a + b = 3.
\end{aligned} \tag{66}$$

Proof. For $(x_1, x_2) \notin \{(0, 0), (-1, 0)\}$, Theorem 6 gives

$$\begin{aligned}
&D_1 G_{\mathbb{Z}^2}(x_1, x_2) \\
&\quad = -\frac{1}{4\pi} \log\left(1 + \frac{2x_1 + 1}{x_1^2 + x_2^2}\right) + b\left[\frac{1}{(x_1 + 1)^2 + x_2^2} - \frac{1}{x_1^2 + x_2^2}\right] \\
&\quad\quad -8b\left[\frac{(x_1 + 1)^2 x_2^2}{[(x_1 + 1)^2 + x_2^2]^3} - \frac{x_1^2 x_2^2}{(x_1^2 + x_2^2)^3}\right] + O\left(\|x\|_2^{-4}\right).
\end{aligned} \tag{67}$$

Expand the log term into a Taylor series. The quantities in brackets are $O\left(\|x\|_2^{-3}\right)$, as follows from using a common denominator. Thus

$$D_1 G_{\mathbb{Z}^2}(x_1, x_2) = -\frac{1}{2\pi} \cdot \frac{x_1}{x_1^2 + x_2^2} + \frac{1}{4\pi} \cdot \frac{x_1^2 - x_2^2}{(x_1^2 + x_2^2)^2} + O\left(\|x\|_2^{-3}\right). \tag{68}$$

This and the analogous statement for $D_2 G_{\mathbb{Z}^2}$ prove the first two formulas in (66). The remainder of the lemma is proved similarly by taking further discrete derivatives; we omit the details. □

Proof of Theorem 4. Using the terminology introduced in Sect. 2, the desired statements are:

$$\begin{aligned}\mathscr{H}^1(\mathbb{Z}^2) &= \{G_{\mathbb{Z}^2} * v : v \in C^3(\mathbb{Z}^2)\} + C^0(\mathbb{Z}^2)\\ \mathscr{H}^p(\mathbb{Z}^2) &= \{G_{\mathbb{Z}^2} * v : v \in C^2(\mathbb{Z}^2)\}, \quad 1 < p \le 2\\ \mathscr{H}^p(\mathbb{Z}^2) &= \{G_{\mathbb{Z}^2} * v : v \in C^1(\mathbb{Z}^2)\}, \quad 2 < p < \infty. \end{aligned} \tag{69}$$

If $v \in C^k(\mathbb{Z}^2)$ for $1 \le k \le 3$, then $G_{\mathbb{Z}^2} * v$ is a finite integer linear combination of translates of k-th derivatives of $G_{\mathbb{Z}^2}$. It follows from Lemma 12 that $(G_{\mathbb{Z}^2} * v)(x) = O\left(\|x\|_2^{-k}\right)$, so $G_{\mathbb{Z}^2} * v \in \ell^p(\mathbb{Z}^2)$ as long as $p > 2/k$. Since $\Delta(G_{\mathbb{Z}^2} * v) = v$ is $\mathbb{Z}$-valued, we conclude that $G_{\mathbb{Z}^2} * v \in \mathscr{H}^p(\mathbb{Z}^2)$. Along with the observation that $C^0(\mathbb{Z}^2) \subset \mathscr{H}^1(\mathbb{Z}^2)$, this proves that for each line of (69), the set on the left side contains the set on the right side.

We prove the forward inclusions in (69) in reverse order, from the third line to the first line. Let $f \in \mathscr{H}^p(\mathbb{Z}^2)$ for some $1 \le p < \infty$, and let $v = \Delta f$. By (65), $v \in C^0(\mathbb{Z}^2)$, so there is $R > 0$ such that the support of v is contained in the ℓ^1-ball of radius R about the origin. In

$$(G_{\mathbb{Z}^2} * v)(x) = \sum_{y \in \mathbb{Z}^2} G_{\mathbb{Z}^2}(x-y)v(y) = \sum_{\|y\|_1 \le R} G_{\mathbb{Z}^2}(x-y)v(y), \tag{70}$$

write $G_{\mathbb{Z}^2}(x) - G_{\mathbb{Z}^2}(x-y)$ as a sum of at most R first derivatives of $G_{\mathbb{Z}^2}$, and use Lemma 12 to see that $G_{\mathbb{Z}^2}(x) - G_{\mathbb{Z}^2}(x-y) = O_R\left(\|x\|_2^{-1}\right)$. Thus, setting $B = \|v\|_1$ and $a = \sum_{y \in \mathbb{Z}^2} v(y)$,

$$(G_{\mathbb{Z}^2} * v)(x) = aG_{\mathbb{Z}^2}(x) + O_{B,R}\left(\|x\|_2^{-1}\right). \tag{71}$$

Set $h(x) = (G_{\mathbb{Z}^2} * v)(x) - f(x)$, so that $\Delta h \equiv 0$. If $a \neq 0$, then as $\|x\|_2 \to \infty$, we have $(G_{\mathbb{Z}^2} * v)(x) \to -\mathrm{sgn}(a) \cdot \infty$ while $f(x) \to 0$, meaning that $h(x) \to -\mathrm{sgn}(a) \cdot \infty$. Since $D_1 h(x), D_2 h(x) \to 0$ as $\|x\|_2 \to \infty$, it follows from the maximum principle that h is constant, a contradiction. Thus $a = 0$ and $v \in C^1(\mathbb{Z}^2)$. We now have $h(x) \to 0$ as $\|x\|_2 \to \infty$, so again by the maximum principle, $h \equiv 0$ and $f = G_{\mathbb{Z}^2} * v$. This proves the forward inclusion in the third line of (69).

Suppose that $p \le 2$. Since $v \in C^1(\mathbb{Z}^2)$, we can write $v = \delta_1 * v_1 + \delta_2 * v_2$ for some $v_1, v_2 \in C^0(\mathbb{Z}^2)$. Then

$$\begin{aligned}(G_{\mathbb{Z}^2} * v)(x) &= (D_1 G_{\mathbb{Z}^2} * v_1)(x) + (D_2 G_{\mathbb{Z}^2} * v_2)(x)\\ &= \sum_{\|y\|_1 \le R+1} D_1 G_{\mathbb{Z}^2}(x-y)v_1(y) + D_2 G_{\mathbb{Z}^2}(x-y)v_2(y)\\ &= b_1 D_1 G_{\mathbb{Z}^2}(x) + b_2 D_2 G_{\mathbb{Z}^2}(x) + O_{B,R}\left(\|x\|_2^{-2}\right), \end{aligned} \tag{72}$$

where each $b_i = \sum_{y \in \mathbb{Z}^2} v_i(y)$. In the last equality we wrote $D_i G_{\mathbb{Z}^2}(x) - D_i G_{\mathbb{Z}^2}(x-y)$ as a sum of $O(R)$ second derivatives of $G_{\mathbb{Z}^2}$ and used the bound from Lemma 12. Again using Lemma 12, we obtain for nonzero $x = (x_1, x_2)$ that

$$(G_{\mathbb{Z}^2} * v)(x) = \frac{1}{2\pi} \cdot \frac{-b_1 x_1 - b_2 x_2}{x_1^2 + x_2^2} + O_{B,R}\left(\|x\|_2^{-2}\right). \tag{73}$$

Suppose b_1 and b_2 are not both zero. Then, there are $0 \le \theta_1 < \theta_2 < 2\pi$ such that $(G_{\mathbb{Z}^2} * v)(x) \asymp \|x\|_2^{-1}$ for all $x \ne (0,0)$ with $\theta_1 \le \arg(x) \le \theta_2$. This contradicts the assumption that $f = G_{\mathbb{Z}^2} * v \in \ell^2(\mathbb{Z}^2)$. We conclude that $b_1 = b_2 = 0$, so $v_1, v_2 \in C^1(\mathbb{Z}^2)$ and therefore $v \in C^2(\mathbb{Z}^2)$ by (45).

Finally, suppose that $p = 1$. Since $v \in C^2(\mathbb{Z}^2)$, we can write $v = \delta_1^{*2} * w_1 + (\delta_1 * \delta_2) * w_2 + \delta_2^{*2} * w_3$ for some $w_1, w_2, w_3 \in C^0(\mathbb{Z}^2)$. Set $c_i = \sum_{y \in \mathbb{Z}^2} w_i(y)$. By the same reasoning as in the previous case,

$$\begin{aligned}
&(G_{\mathbb{Z}^2} * v)(x) \\
&\quad = c_1 D_1^2 G_{\mathbb{Z}^2}(x) + c_2 D_1 D_2 G_{\mathbb{Z}^2}(x) + c_3 D_2^2 G_{\mathbb{Z}^2}(x) + O_{B,R}\left(\|x\|_2^{-3}\right) \\
&\quad = \frac{1}{2\pi} \cdot \frac{(c_1 - c_3)(x_1^2 - x_2^2) + 2c_2 x_1 x_2}{(x_1^2 + x_2^2)^2} + O_{B,R}\left(\|x\|_2^{-3}\right)
\end{aligned} \tag{74}$$

for all nonzero $x = (x_1, x_2)$. This implies that $c_1 = c_3$ and $c_2 = 0$; if not, the first term would have asymptotic order $\|x\|_2^{-2}$ for $\arg(x)$ in some range $[\theta_1, \theta_2]$, contradicting that $f \in \ell^1(\mathbb{Z}^2)$.

Set $c = c_1 = c_3$, and let $v' = \Delta \mathbf{e}_{(0,0)} = -\delta_1^{*2} * \mathbf{e}_{(1,0)} - \delta_2^{*2} * \mathbf{e}_{(0,1)}$. Then

$$v + cv' = \delta_1^{*2} * (w_1 - c\,\mathbf{e}_{(1,0)}) + (\delta_1 * \delta_2) * w_2 + \delta_2^{*2} * (w_3 - c\,\mathbf{e}_{(0,1)}). \tag{75}$$

Since all three of $w_1 - c\,\mathbf{e}_{(1,0)}$, w_2, and $w_3 - c\,\mathbf{e}_{(0,1)}$ are in $C^1(\mathbb{Z}^2)$, we have $v + cv' \in C^3(\mathbb{Z}^2)$. As well, $G_{\mathbb{Z}^2} * cv' = c\,\mathbf{e}_{(0,0)} \in C^0(\mathbb{Z}^2)$. Hence

$$f = G_{\mathbb{Z}^2} * (v + cv' - cv') \in \{G_{\mathbb{Z}^2} * w : w \in C^3(\mathbb{Z}^2)\} + C^0(\mathbb{Z}^2), \tag{76}$$

which completes the proof. □

4. Stabilization on $\mathbb{Z}^2$

Consider a sandpile $\sigma : \mathbb{Z}^2 \to \mathbb{Z}_{\ge 0}$. The *parallel toppling* procedure attempts to stabilize σ by defining a sequence of sandpiles $\sigma = \sigma^0, \sigma^1, \sigma^2, \ldots$ where σ^{n+1} is obtained from σ^n by simultaneously toppling all vertices x with $\sigma^n(x) \ge 4$. Formally, set $v^n(x) = \mathbf{1}\{\sigma^n(x) \ge 4\}$ and define $\sigma^{n+1} = \sigma^n - \Delta(v^n)$. Define the sequence of *odometer functions* $u^1, u^2, \ldots$ by $u^n = v^0 + v^1 + \cdots + v^{n-1}$, so that $u^n(x)$ is the number of times vertex x has toppled in the first n topplings. In particular, $\|u^n\|_{\ell^\infty} \le n$ and $\sigma^n = \sigma - \Delta(u^n)$. It is shown in [18] that σ stabilizes if and only if $u^n \uparrow u^\infty$ for some $u^\infty : \mathbb{Z}^2 \to \mathbb{Z}_{\ge 0}$, in which case the stabilization is given by $\sigma^\infty = \sigma - \Delta u^\infty$.

Our proof uses the following 'conservation of density' result of [18].

Lemma 13 ([18], Lemma 2.10). *Let $(\sigma_x)_{x \in \mathbb{Z}^2}$ be i.i.d. and stabilize almost surely, with stabilization $(\sigma_x^\infty)_{x \in \mathbb{Z}^2}$. Then $\mathbf{E}[\sigma_0] = \mathbf{E}[\sigma_0^\infty]$.*

In particular, if the i.i.d. sandpile σ stabilizes almost surely, then $\mathbf{E}[\sigma_0] \le 3$.

We now show that if $\xi \in \mathscr{H}^1(\mathbb{Z}^2)$, the pairing $\langle \sigma, \xi \rangle = \sum_{x \in \mathbb{Z}^2} \sigma(x)\xi(x)$ remains invariant modulo 1 when the sandpile σ is stabilized.

Lemma 14. *Let $(\sigma_x)_{x \in \mathbb{Z}^2}$ be an i.i.d. sandpile which stabilizes almost surely, and let $\xi \in \mathscr{H}^1(\mathbb{Z}^2)$. Then*

$$\langle \sigma, \xi \rangle \equiv \langle \sigma^\infty, \xi \rangle \mod 1, \quad a.s. \tag{77}$$

Proof. Lemma 13 implies that $\mathbf{E}[\sigma_0] < \infty$. Since $\xi \in \ell^1(\mathbb{Z}^2)$,

$$\mathbf{E}[\langle \sigma, |\xi| \rangle] = \sum_{x \in \mathbb{Z}^2} |\xi_x| \mathbf{E}[\sigma_x] = \|\xi\|_1 \mathbf{E}[\sigma_0] < \infty \tag{78}$$

and so $\langle \sigma, \xi \rangle$ converges absolutely almost surely. Write $\sigma^n = \sigma - \Delta u^n$ and use self-adjointness of Δ to obtain

$$\langle \sigma^n, \xi \rangle = \langle \sigma - \Delta u^n, \xi \rangle = \langle \sigma, \xi \rangle - \langle u^n, \Delta \xi \rangle. \tag{79}$$

Since u^n is integer-valued, increasing and converges almost surely, while $\Delta\xi$ is integer-valued and has finite support, the increment $\langle u^n, \Delta\xi \rangle$ converges a.s. to $\langle u^\infty, \Delta\xi \rangle \in \mathbb{Z}$.

Note that the parallel toppling property implies that, for $n \geq 0$,

$$\sigma^{n+1}(x) \leq \max\left(\sigma^n(x), 7\right). \tag{80}$$

Thus, whenever $\langle \sigma, |\xi| \rangle$ is finite and σ stabilizes to σ^∞,

$$\lim_{n\to\infty} \langle \sigma^n, \xi \rangle = \lim_{n\to\infty} \sum_{x \in \mathbb{Z}^2} \sigma^n(x)\xi_x = \sum_{x \in \mathbb{Z}^2} \lim_{n\to\infty} \sigma^n(x)\xi_x = \langle \sigma^\infty, \xi \rangle \tag{81}$$

where the second equality is justified by dominated convergence:

$$|\sigma^n(x)\xi_x| \leq \max(\sigma(x), 7)|\xi_x|, \qquad \sum_{x \in \mathbb{Z}^2} \max(\sigma(x), 7)|\xi_x| < \infty. \tag{82}$$

Sending $n \to \infty$ in (79) completes the proof. □

For definiteness, our argument uses the particular function

$$\xi = G_{\mathbb{Z}^2} * \delta_1^{*3} = D_1^3 G_{\mathbb{Z}^2}, \tag{83}$$

which is in $\mathscr{H}^1(\mathbb{Z}^2)$ by Lemma 12. The next lemma estimates the tail of $\|\xi\|_2^2$.

Lemma 15. *Let $R \geq 1$ be a parameter. As $R \to \infty$,*

$$\sum_{x \in \mathbb{Z}^2 : 0 < |\xi_x| < \frac{1}{2R}} |\xi_x|^2 \gg R^{-\frac{4}{3}}. \tag{84}$$

Proof. Arguing as in Lemma 12, we see that there are $0 \leq \theta_1 < \theta_2 < 2\pi$ such that, for nonzero $x \in \mathbb{Z}^2$ satisfying $\theta_1 \leq \arg(x) \leq \theta_2$,

$$|\xi_x| \asymp \|x\|_2^{-3}. \tag{85}$$

Thus

$$\sum_{x \in \mathbb{Z}^2 : 0 < |\xi_x| < \frac{1}{2R}} |\xi_x|^2 \gg \int_{R^{\frac{1}{3}}}^\infty \frac{dr}{r^5} \gg R^{-\frac{4}{3}}. \tag{86}$$

□

As the proof below makes clear, an explicit constant in the lower bound (84) would lead to explicit values of c, d in the statement of Theorem 1. To obtain a fully quantitative version of Lemma 15, it would be enough to bound the error in (53) by finding an explicit $C > 0$ such that

$$\left| G_{\mathbb{Z}^2}(x) + \frac{\log \|x\|_2}{2\pi} + a + b \frac{\frac{8x_1^2 x_2^2}{\|x\|_2^4} - 1}{\|x\|_2^2} \right| \leq C \|x\|_2^{-4} \tag{87}$$

for all $(0,0) \neq x \in \mathbb{Z}^2$. A result in this direction [28, Section 4] is that

$$\left| G_{\mathbb{Z}^2}(x) + \frac{\log \|x\|_2}{2\pi} \right| \leq 0.01721 \|x\|_2^{-2}. \tag{88}$$

(Indeed, the constant 0.01721 is optimal and an exact formula for it is given.) It is likely that extending the techniques developed in [28] would lead to a bound of the form (87), and thence to an explicit numerical bound in Theorem 1.

Proof of Theorem 1. Consider the characteristic functions

$$\chi(\sigma; \xi) = \mathbf{E}\left[e^{-2\pi i \langle \sigma, \xi \rangle} \right], \qquad \chi(\sigma^\infty; \xi) = \mathbf{E}\left[e^{-2\pi i \langle \sigma^\infty, \xi \rangle} \right]. \tag{89}$$

Since $\langle \sigma, \xi \rangle \equiv \langle \sigma^\infty, \xi \rangle \bmod 1$ a.s., $\chi(\sigma; \xi) = \chi(\sigma^\infty; \xi)$.

Let $\mathbf{E}[\sigma_0] = \mathbf{E}[\sigma_0^\infty] = 3 - \epsilon$. Using $|1 - e^{2\pi i t}| \leq 2\pi |t|$ and $\sum_{x \in \mathbb{Z}^2} \xi_x = 0$,

$$\begin{aligned} |1 - \chi(\sigma^\infty; \xi)| &= \left| \mathbf{E}\left[1 - e^{-2\pi i \langle \sigma^\infty - 3, \xi \rangle} \right] \right| \\ &\leq \mathbf{E}[2\pi |\langle \sigma^\infty - 3, \xi \rangle|] \\ &\leq 2\pi \|\xi\|_1 \epsilon. \end{aligned} \tag{90}$$

Thus, $|\chi(\sigma^\infty; \xi)| \geq 1 - 2\pi \|\xi\|_1 \epsilon$.

Meanwhile, since $(\sigma_x)_{x \in \mathbb{Z}^2}$ is i.i.d.,

$$\chi(\sigma; \xi) = \prod_{x \in \mathbb{Z}^2} \mathbf{E}\left[e^{-2\pi i \xi_x \sigma_0} \right]. \tag{91}$$

Use the inequality $-\log t \geq \frac{1-t^2}{2}$ in $0 < t \leq 1$ to obtain

$$-\log |\chi(\sigma; \xi)| \geq \frac{1}{2} \sum_{x \in \mathbb{Z}^2} \left(1 - \left| \mathbf{E}\left[e^{-2\pi i \xi_x \sigma_0} \right] \right|^2 \right). \tag{92}$$

Let X, X' be independent and distributed as σ_0. One has

$$\left| \mathbf{E}\left[e^{-2\pi i \xi_x \sigma_0} \right] \right|^2 = \mathbf{E}\left[e^{-2\pi i \xi_x X} \right] \mathbf{E}\left[e^{2\pi i \xi_x X'} \right] = \mathbf{E}\left[e^{-2\pi i \xi_x (X - X')} \right]. \tag{93}$$

This quantity is equal to its real part $\mathbf{E}[c(\xi_x(X - X'))]$. (Recall $c(t) = \cos 2\pi t$.) Therefore, using $1 - c(t) \geq 8t^2$ for $|t| \leq \frac{1}{2}$,

$$-\log |\chi(\sigma; \xi)| \geq \frac{1}{2} \sum_{x \in \mathbb{Z}^2} \Big(1 - \mathbf{E}[c(\xi_x(X - X'))] \Big)$$

$$\geq 4\mathbf{E}\left[\sum_{0<|\xi_x(X-X')|<\frac{1}{2}} \xi_x^2 (X-X')^2\right]$$

$$= 4\sum_{k=1}^{\infty} \mathbf{E}\left[\mathbf{1}\{|X-X'|=k\} \sum_{0<|\xi_x|<\frac{1}{2k}} \xi_x^2 k^2\right]. \tag{94}$$

Lemma 15 now implies that

$$-\log|\chi(\sigma;\xi)| \gg \sum_{k=1}^{\infty} \mathbf{E}\left[\mathbf{1}\{|X-X'|=k\}k^{2/3}\right] = \mathbf{E}[|X-X'|^{2/3}], \tag{95}$$

and therefore

$$1-|\chi(\sigma;\xi)| \gg \min\left(1, \mathbf{E}[|X-X'|^{2/3}]\right). \tag{96}$$

The result follows on combining this with (90). □

5. The Sandpile Group

Recall the designations $\mathscr{R}_m \subset \mathscr{S}_m$ for the recurrent and stable states, respectively, of the sandpile model on $\mathbb{T}_m$ with sink at $(0,0)$. Any sandpile $\sigma : \mathbb{T}_m \backslash \{(0,0)\} \to \mathbb{Z}_{\geq 0}$ can be stabilized by repeatedly performing legal topplings until the resulting configuration is stable. By the abelian property [11], the final state does not depend on the order in which the topplings are performed, and is called the *stabilization* of σ.

If we view functions on $\mathbb{T}_m$ as $m^2 \times 1$ column vectors, then the Laplacian operator Δ on $\mathbb{T}_m$ can be considered as an $m^2 \times m^2$ matrix, so that for example $\Delta\mathbb{Z}^{\mathbb{T}_m}$ is the integer span of the columns of Δ. The null space of Δ is one-dimensional, and is spanned by the all-ones vector. The *reduced Laplacian* Δ' is obtained by omitting the row and column corresponding to the sink $(0,0)$, and is invertible.

The recurrent states $\mathscr{R}_m$ of the sandpile model are naturally identified with the abelian group

$$\mathscr{G}_m := \mathbb{Z}^{\mathbb{T}_m\backslash\{(0,0)\}}/\Delta'\mathbb{Z}^{\mathbb{T}_m\backslash\{(0,0)\}}, \tag{97}$$

which is the *sandpile group* of $\mathbb{T}_m$. Indeed, each equivalence class

$$\sigma + \Delta'\mathbb{Z}^{\mathbb{T}_m\backslash\{(0,0)\}} \subset \mathbb{Z}^{\mathbb{T}_m\backslash\{(0,0)\}}, \quad \sigma \in \mathbb{Z}^{\mathbb{T}_m\backslash\{(0,0)\}}, \tag{98}$$

contains exactly one recurrent sandpile [21]. Addition in $\mathscr{G}_m$ corresponds via this bijection to the operation on $\mathscr{R}_m$ of pointwise addition followed by stabilization.

The *sandpile Markov chain* has state space $\mathscr{S}_m$ and transition operator P_m. To take a single step from a sandpile σ, choose a site $x \in \mathbb{T}_m$ uniformly at random. If $x \neq (0,0)$, replace σ with the stabilization of $\sigma + \mathbf{e}_x$; if $x = (0,0)$, remain at σ. The recurrent states of the chain are precisely $\mathscr{R}_m$, and the chain restricted to $\mathscr{R}_m$ is a random walk on the group $\mathscr{G}_m$. See [26], which develops this construction in the setting of an arbitrary underlying graph, for further background.

Using (97), the matrix-tree theorem implies that $\mathscr{G}_m$ is in bijection with the spanning trees of $\mathbb{T}_m$. It is shown in [26] that $|\mathscr{G}_m| = \exp\left(\left(\frac{4\beta(2)}{\pi} + o(1)\right)m^2\right)$ where $\beta(2)$ is the Catalan constant,

$$\frac{4\beta(2)}{\pi} = 1.1662\ldots. \tag{99}$$

Thus the recurrent states make up an exponentially small fraction of the 4^{m^2-1} stable states.

The following proposition bounds the hitting time started from a deterministic stable state to reach a recurrent state.

Proposition 16. *There is a constant $C > 0$ such that, as $m \to \infty$, for any stable state $\sigma \in \mathscr{S}_m$, if $n > Cm^2\sqrt{\log m}$ then*

$$\mathbf{Prob}\left(P_m^n \delta_\sigma \in \mathscr{R}_m\right) = 1 - o(1).$$

Remark. Starting from $\sigma = 0$, at least order m^2 steps are necessary to reach a recurrent state, since only one chip is added at a time. We do not claim that the extra factor of $\sqrt{\log m}$ above is optimal. Because we will show that the mixing time of the sandpile chain has order $m^2 \log m$, the bound in Proposition 16 is sufficient for understanding the mixing behavior.

Proof. We make two initial observations. First, any state satisfying $\sigma \geq 3$ can be toppled to a stable recurrent state. This is because such a state can evidently be reached from a recurrent state. Also, by performing a sequence of topplings, a single vertex with allocation h can be toppled to produce a disc of radius $\gg \sqrt{h}$ with height at least 3. This follows as a simple consequence of the analysis in [34], which studies the limiting shape of the configuration obtained by repeated toppling of a pile at a single vertex.

Let A be an integer, $A \ll \sqrt{\log m}$, and drop $n \sim \text{Poisson}(Am^2)$ grains of sand on the torus, while performing no topplings. Note that this is the same as independently dropping Poisson(A) grains of sand on each vertex. Also, $n < 2Am^2$ with probability $1 - o(1)$.

The probability that a non-sink vertex x has height at most a is

$$\mathbf{Prob}(h_x \leq a) = e^{-A} \sum_{j=0}^{a} \frac{A^j}{j!}. \tag{100}$$

For $a < \frac{A}{2}$ we obtain

$$\mathbf{Prob}(h_x \leq a) \asymp \frac{A^a}{a!} \exp(-A).$$

If $x_1, x_2, \ldots, x_s$ denote the points of a disc of area $s \gg a$, then, by independence,

$$\mathbf{Prob}\left(\bigwedge_{i=1}^{s} (h_{x_i} \leq a)\right) \leq \exp\left(-sA + sa \log \frac{A}{a} + s(a + O(1))\right). \tag{101}$$

Choose $s, a \asymp \sqrt{\log m}$ such that a point of height a in a disc of area s topples to cover the disc. Then choose A a sufficiently large constant times $\sqrt{\log m}$ so that the probability of (101) is $o\left(1/m^2\right)$. It follows that with probability $1 - o(1)$, the event (101) does not occur for any disc on the torus at distance $\gg \sqrt{\log m}$ from the sink. The sites closer to the sink have height ≥ 3 with probability $1 - o(1)$ by estimating using (100) and a union bound. □

The following proposition reduces the statements in Theorem 2 to estimates started from the fixed recurrent state $\sigma \equiv 3$.

Proposition 17. *For each constant $C > 0$, for $t = Cm^2 \log m$, as $m \to \infty$,*

$$\sup_{\sigma_0 \in \mathscr{S}_m} \left| \left\| P_m^t \delta_{\sigma_0} - \mathbb{U}_{\mathscr{R}_m} \right\|_{\mathrm{TV}} - \left\| P_m^t \delta_{\sigma \equiv 3} - \mathbb{U}_{\mathscr{R}_m} \right\|_{\mathrm{TV}} \right| = o(1). \tag{102}$$

Proof. Given $\sigma_0 \in \mathscr{S}_m$, let $\sigma_1 \in \mathscr{R}_m$ be the unique recurrent state in the equivalence class $\sigma_0 + \Delta' \mathbb{Z}^{\mathbb{T}_m \setminus \{(0,0)\}}$. By Proposition 16, $P_m^t \delta_{\sigma_0}(\mathscr{R}_m) = 1 - o(1)$, and thus $\left\| P_m^t \delta_{\sigma_0} - P_m^t \delta_{\sigma_1} \right\|_{\mathrm{TV}} = o(1)$. Since the chain restricted to $\mathscr{R}_m$ is transitive, it follows from the triangle inequality that

$$\begin{aligned} &\left| \left\| P_m^t \delta_{\sigma_0} - \mathbb{U}_{\mathscr{R}_m} \right\|_{\mathrm{TV}} - \left\| P_m^t \delta_{\sigma \equiv 3} - \mathbb{U}_{\mathscr{R}_m} \right\|_{\mathrm{TV}} \right| \\ &= \left| \left\| P_m^t \delta_{\sigma_0} - \mathbb{U}_{\mathscr{R}_m} \right\|_{\mathrm{TV}} - \left\| P_m^t \delta_{\sigma_1} - \mathbb{U}_{\mathscr{R}_m} \right\|_{\mathrm{TV}} \right| \\ &\le \left\| P_m^t \delta_{\sigma_0} - P_m^t \delta_{\sigma_1} \right\|_{\mathrm{TV}} = o(1). \end{aligned} \tag{103}$$

5.1. Random walk on the sandpile group. Going forward we assume that the sandpile Markov chain is started from the deterministic recurrent state $\sigma \equiv 3$ so that the dynamics is reduced to a random walk on the abelian group $\mathscr{G}_m$. In general, for any random walk on a finite abelian group $\mathscr{G}$ driven by the measure μ, the eigenfunctions of the transition kernel are given by the dual group, which is the additive group of characters $\hat{\mathscr{G}} = \{\xi : \mathscr{G} \to \mathbb{R}/\mathbb{Z}\}$. If $\xi \cdot g$ denotes the image of $g \in \mathscr{G}$ under $\xi \in \hat{\mathscr{G}}$, then the eigenfunction corresponding to ξ is $f_\xi(g) = e(\xi \cdot g)$. The corresponding eigenvalue is the Fourier coefficient of μ at frequency ξ, namely $\hat{\mu}(\xi) = \sum_{g \in \mathscr{G}} \mu(g) e(\xi \cdot g)$.

The sandpile chain on $\mathscr{G}_m$ is driven by the measure

$$\mu := \frac{1}{m^2} \left(\delta_0 + \sum_{x \in \mathbb{T}_m \setminus \{(0,0)\}} \delta_{\mathbf{e}_x} \right) \tag{104}$$

where, technically, $\mathbf{e}_x$ refers to the equivalence class $\mathbf{e}_x + \Delta' \mathbb{Z}^{\mathbb{T}_m \setminus \{(0,0)\}} \in \mathscr{G}_m$, and $0 \in \mathscr{G}_m$ is the identity. The dual group of $\mathscr{G}_m$ is

$$\hat{\mathscr{G}}_m = (\Delta')^{-1} \mathbb{Z}^{\mathbb{T}_m \setminus \{(0,0)\}} / \mathbb{Z}^{\mathbb{T}_m \setminus \{(0,0)\}}. \tag{105}$$

This can be seen by dualizing (97); a bare-hands proof is given in Section 3 of [26]. To define the meaning of $\xi \cdot g$ in this setting, we can view each frequency $\xi \in \hat{\mathscr{G}}_m$ as a function from $\mathbb{T}_m \setminus \{(0,0)\}$ to $\mathbb{R}/\mathbb{Z}$, and each group element $g \in \mathscr{G}_m$ as an equivalence class $\sigma + \Delta' \mathbb{Z}^{\mathbb{T}_m \setminus \{(0,0)\}}$, where $\sigma \in \mathbb{Z}^{\mathbb{T}_m \setminus \{(0,0)\}}$. Then, $\xi \cdot g = \sum_{x \in \mathbb{T}_m \setminus \{(0,0)\}} \xi_x \sigma_x \in \mathbb{R}/\mathbb{Z}$, whose value does not depend on the choice of the representative σ in the equivalence class. The eigenvalue corresponding to ξ is

$$\hat{\mu}(\xi) = \frac{1}{m^2} \left(1 + \sum_{x \in \mathbb{T}_m \setminus \{(0,0)\}} e(\xi_x) \right). \tag{106}$$

Given $\xi : \mathbb{T}_m \setminus \{(0,0)\} \to \mathbb{R}/\mathbb{Z}$, which may or may not be in $\hat{\mathscr{G}}_m$, set $v = \Delta' \xi$ (which is also $\mathbb{R}/\mathbb{Z}$-valued). Extend ξ to the domain $\mathbb{T}_m$ by setting $\xi(0,0) = 0$. Then $\Delta \xi(x) = v(x)$ for all $x \in \mathbb{T}_m \setminus \{(0,0)\}$, and since the columns of Δ all sum to zero, $\Delta \xi(0,0) = -\sum_{x \neq (0,0)} v(x)$. From (105), $\xi \in \hat{\mathscr{G}}_m$ if and only if $v \equiv 0$, which holds

if and only if $\Delta\xi \equiv 0$. This justifies the description of $\hat{\mathscr{G}}_m$ in Sect. 1.4 as the additive group of functions $\xi : \mathbb{T}_m \to \mathbb{R}/\mathbb{Z}$ such that $\xi(0,0) = 0$ and $\Delta\xi \equiv 0$ in $\mathbb{R}/\mathbb{Z}$. From this point forward, when we refer to a frequency $\xi \in \hat{\mathscr{G}}_m$, we mean a function that meets these conditions.

In [26], $\hat{\mathscr{G}}_m$ was identified with the group of 'multiplicative harmonic functions,' which in the present setting are the maps from $\mathbb{T}_m$ to $\mathbb{C}^*$ given by $x \mapsto e(\xi_x)$.

Abusing notation slightly, define for any $\mathbb{R}$-valued or $\mathbb{R}/\mathbb{Z}$-valued function ξ on $\mathbb{T}_m$,

$$\hat{\mu}(\xi) := \mathbf{E}_{x\in\mathbb{T}_m}\left[e(\xi_x)\right]. \tag{107}$$

When in fact $\xi \in \hat{\mathscr{G}}_m$, this definition agrees with (106).

5.2. Representations for frequencies. We use a concrete description of the frequencies in terms of the Green's function, which associates to the frequencies an approximate partial ordering. To describe this, given $\xi \in \hat{\mathscr{G}}_m$ recall that a 'prevector' for ξ is any integer-valued vector $\Delta\xi'$, where $\xi' : \mathbb{T}_m \to \mathbb{R}$ reduces mod $\mathbb{Z}$ to ξ. We choose a particular representative $\xi' : \mathbb{T}_m \to (-1, 1)$ by letting

$$C(\xi) = \frac{1}{2\pi}\arg\left(\hat{\mu}(\xi)\right) \in \left[-\tfrac{1}{2}, \tfrac{1}{2}\right) \tag{108}$$

and choosing each $\xi'_x \in \left(C(\xi) - \frac{1}{2}, C(\xi) + \frac{1}{2}\right]$. The 'distinguished prevector' of ξ is then given by

$$v = v(\xi) := \Delta\xi'. \tag{109}$$

Note that $v : \mathbb{T}_m \to \mathbb{Z}$ has mean zero and satisfies $\|v\|_{L^\infty} \le 3$.

Lemma 18. *For every $\xi \in \hat{\mathscr{G}}_m$, the distinguished prevector of ξ satisfies*

$$1 - \left|\hat{\mu}(\xi)\right| \gg \frac{\|v(\xi)\|_2^2}{m^2} \ge \frac{\|v(\xi)\|_1}{m^2}. \tag{110}$$

Proof. Choose ξ' as above, and define $\xi^* : \mathbb{T}_m \to \left(-\frac{1}{2}, \frac{1}{2}\right]$ by

$$\xi^*_x = \xi'_x - C(\xi), \tag{111}$$

so that $\Delta\xi^* = \Delta\xi' = v$ and

$$0 \le |\hat{\mu}(\xi)| = \frac{1}{m^2}\sum_{x\in\mathbb{T}_m} e\left(\xi^*_x\right) = \frac{1}{m^2}\sum_{x\in\mathbb{T}_m} c\left(\xi^*_x\right). \tag{112}$$

Approximating $1 - c(t) \gg t^2$ uniformly for $|t| \le \frac{1}{2}$ yields

$$1 - |\hat{\mu}(\xi)| \gg \frac{\|\xi^*\|_2^2}{m^2}. \tag{113}$$

Since Δ is bounded from $L^2(\mathbb{T}_m) \to L^2(\mathbb{T}_m)$,

$$\frac{\|v\|_2^2}{m^2} = \frac{\|\Delta\xi^*\|_2^2}{m^2} \ll \frac{\|\xi^*\|_2^2}{m^2} \ll 1 - |\hat{\mu}(\xi)| \tag{114}$$

as desired. Finally, $\|v\|_2^2 \ge \|v\|_1$ since v is integer-valued. □

To go in the reverse direction, for any $v \in \mathbb{Z}_0^{\mathbb{T}_m}$ define $\overline{\xi} = G_{\mathbb{T}_m} * v$, so that $\Delta\overline{\xi} = v$. Let $\xi''_x = \overline{\xi}_x - \overline{\xi}_{(0,0)}$, and set $\xi = \xi(v)$ to be the reduction mod $\mathbb{Z}$ of ξ''. Since $\xi''_{(0,0)} = 0$ and $\Delta\xi'' = v$, which is $\mathbb{Z}$-valued, it follows that $\xi \in \hat{\mathscr{G}}_m$.

If $\xi_0 \in \hat{\mathscr{G}}_m$ and $v = \Delta\xi'$ is any prevector of ξ_0, then $\xi(v) = \xi_0$; this is because $\Delta(\xi' - \xi'') \equiv 0$, so $\xi' - \xi'' \equiv c$ for some $c \in \mathbb{R}$, and in fact $c = \xi'_{(0,0)} - \xi''_{(0,0)} \in \mathbb{Z}$. Also, if $v_0 \in \mathbb{Z}_0^{\mathbb{T}_m}$ and v is any prevector of $\xi(v_0)$, then $v_0 - v \in \Delta\mathbb{Z}^{\mathbb{T}_m}$.

Lemma 19. *Given $\xi \in \hat{\mathscr{G}}_m$, let v be any prevector of ξ and let $\overline{\xi} = G_{\mathbb{T}_m} * v$. Then $|\hat{\mu}(\xi)| = |\hat{\mu}(\overline{\xi})|$. If v is the distinguished prevector of ξ, then in addition*

$$1 - |\hat{\mu}(\xi)| \asymp \frac{\|\overline{\xi}\|_2^2}{m^2}. \tag{115}$$

Equation (115) is equivalent to Theorem 3.8 in [26], and the argument below is the same as the proof given there.

Proof. Let $v = \Delta\xi'$, where $\xi' : \mathbb{T}_m \to \mathbb{R}$ reduces mod $\mathbb{Z}$ to ξ. Then $\overline{\xi} = G_{\mathbb{T}_m} * \Delta\xi' = \xi' - c$ where $c = \mathbf{E}_{x\in\mathbb{T}_m}[\xi']$, so

$$\hat{\mu}(\overline{\xi}) = \mathbf{E}_{x\in\mathbb{T}_m}[e(\xi'_x - c)] = e(-c)\hat{\mu}(\xi') = e(-c)\hat{\mu}(\xi) \tag{116}$$

and therefore $|\hat{\mu}(\overline{\xi})| = |\hat{\mu}(\xi)|$.

To prove the upper bound in (115),

$$\begin{aligned} 1 - |\hat{\mu}(\xi)| = 1 - |\hat{\mu}(\overline{\xi})| \le 1 - \operatorname{Re}\hat{\mu}(\overline{\xi}) &= \frac{1}{m^2}\sum_{x\in\mathbb{T}_m}\big[1 - c(\overline{\xi}_x)\big] \\ &\ll \frac{1}{m^2}\sum_{x\in\mathbb{T}_m}|\overline{\xi}_x|^2 = \frac{\|\overline{\xi}\|_2^2}{m^2}. \end{aligned} \tag{117}$$

For the lower bound, define ξ^* as in the proof of Lemma 18 and observe that $\overline{\xi} = G_{\mathbb{T}_m} * \Delta\xi^* = \xi^* - \mathbf{E}_{x\in\mathbb{T}_m}[\xi^*]$ is the orthogonal projection of ξ^* onto $L_0^2(\mathbb{T}_m)$. Thus $\|\overline{\xi}\|_2^2 \le \|\xi^*\|_2^2$, and the result follows from (113). □

6. Spectral Estimates

This section reduces the determination of the spectral gap to a finite check, and provides additive savings estimates for separated spectral components. Lemma 18 implies that each nonzero frequency $\xi \in \hat{\mathscr{G}}_m$ satisfies $1 - |\hat{\mu}(\xi)| \gg 1/m^2$, and if $1 - |\hat{\mu}(\xi)| \le c/m^2$, then the L^1 norm of the distinguished prevector $v(\xi)$ must be bounded by a constant depending only on c. Section 6.1 develops tools to deal with prevectors that have bounded L^1 norm, providing control over those frequencies that achieve the spectral gap or approach it to within a constant factor. This proves Theorem 3 and does most of the work for the lower bound in Theorem 2.

Section 6.2 extends the analysis to prevectors whose L^1 norm increases with m, but which are sparse enough that their supports can be partitioned into widely separated

clusters. This provides the main ingredient for the upper bound in Theorem 2. As we will show in Sect. 7, if ξ is a frequency for which $v(\xi)$ is not sparse, then the Lemma 18 lower bound on $1-|\hat{\mu}(\xi)|$ shows that the contribution of ξ is negligible when computing the mixing time.

To fix ideas, given $\xi \in \hat{\mathscr{G}}_m$ recall that $\hat{\mu}(\xi) = \mathbf{E}_{x\in\mathbb{T}_m}[e(\xi_x)]$. For any subset $S \subset \mathbb{T}_m$, it is evident that

$$\left|\sum_{x\in S} e(\xi_x)\right| \le |S|. \tag{118}$$

The 'savings from S' for the frequency ξ, denoted by $\mathrm{sav}(\xi; S)$, is the amount by which the left side falls short of this upper bound:

$$\mathrm{sav}(\xi; S) := |S| - \left|\sum_{x\in S} e(\xi_x)\right|. \tag{119}$$

By the triangle inequality, if $S_1, S_2 \subset \mathbb{T}_m$ are disjoint then

$$\mathrm{sav}(\xi; S_1) + \mathrm{sav}(\xi; S_2) \le \mathrm{sav}(\xi; S_1 \cup S_2). \tag{120}$$

The 'total savings' for ξ is defined by

$$\mathrm{sav}(\xi) := \mathrm{sav}(\xi; \mathbb{T}_m) = m^2 - \left|\sum_{x\in\mathbb{T}_m} e(\xi_x)\right| \tag{121}$$

and satisfies

$$1 - \left|\hat{\mu}(\xi)\right| = \frac{\mathrm{sav}(\xi)}{m^2}. \tag{122}$$

The notion of savings is well-suited for proving lower bounds on the gap $1-|\hat{\mu}(\xi)|$. Specifically, if $S_1, \ldots, S_k$ are disjoint subsets of $\mathbb{T}_m$ then

$$1 - |\hat{\mu}(\xi)| \ge \frac{1}{m^2}\sum_{i=1}^{k} \mathrm{sav}(\xi; S_i). \tag{123}$$

The spectral gap of the sandpile Markov chain is

$$\mathrm{gap}_m = \min_{0\neq\xi\in\hat{\mathscr{G}}_m} \frac{\mathrm{sav}(\xi)}{m^2}. \tag{124}$$

Observe that if v is the distinguished prevector of $\xi \in \hat{\mathscr{G}}_m$, then Lemma 18 gives $\mathrm{sav}(\xi) \gg \|v\|_1$. Also, given a set $S \subset \mathbb{T}_m$ and a function w on $\mathbb{T}_m$, write $w|_S$ for the function which is equal to w on S and 0 on S^c.

6.1. Determination of spectral gap up to finite check. Given constants $B, R > 0$, define the finite set

$$\mathscr{C}(B, R) := \{v \in C^2(B_R(0)) : \|v\|_1 \leq B\}. \tag{125}$$

Here $B_R(0)$ is the ℓ^1 ball of radius R about 0 in $\mathbb{Z}^2$, and $C^2(\cdot)$ is given by (44). Since $B_R(0)$ embeds into $\mathbb{T}_m$ for each $m > 2R$, we can view each $v \in \mathscr{C}(B, R)$ as an element either of $C^2(\mathbb{Z}^2)$ or of $C^2(\mathbb{T}_m)$ by setting $v \equiv 0$ outside $B_R(0)$.

For any $\mathbb{R}$- or $\mathbb{R}/\mathbb{Z}$-valued function ξ on $\mathbb{Z}^2$, define the functional

$$f(\xi) := \sum_{x \in \mathbb{Z}^2} (1 - c(\xi_x)). \tag{126}$$

We will see that this is the appropriate analogue to savings for functions on $\mathbb{Z}^2$. If $v \in C^2(\mathbb{Z}^2)$, then $f(G_{\mathbb{Z}^2} * v) < \infty$ by the bound $1 - c(t) \ll t^2$ combined with Lemma 9 or Lemma 12. For such v, $f(G_{\mathbb{Z}^2} * v) = 0$ if and only if $G_{\mathbb{Z}^2} * v$ is $\mathbb{Z}$-valued. Since $G_{\mathbb{Z}^2} * v \in \ell^2(\mathbb{Z}^2)$, if it is $\mathbb{Z}$-valued then it must be finitely supported, and in addition we have $\Delta(G_{\mathbb{Z}^2} * v) = v$. Thus, $f(G_{\mathbb{Z}^2} * v) = 0$ precisely for those v in the subset

$$\mathcal{I} := \{\Delta w : w \in C^0(\mathbb{Z}^2)\} \subset C^2(\mathbb{Z}^2). \tag{127}$$

If $v, v' \in C^2(\mathbb{Z}^2)$ and $v - v' \in \mathcal{I}$, then $f(G_{\mathbb{Z}^2} * v) = f(G_{\mathbb{Z}^2} * v')$.

Set

$$\gamma := \inf \left\{ f(G_{\mathbb{Z}^2} * v) : v \in C^2(\mathbb{Z}^2) \backslash \mathcal{I} \right\}. \tag{128}$$

The following are the main results of this section. Together with the computation in "Appendix B", they lead to a quick proof of Theorem 3.

Proposition 20. *We have $\gamma > 0$, and there exist constants $B_0, R_0 > 0$ such that:*

1. *For sufficiently large m, any $\xi \in \hat{\mathscr{G}}_m$ that achieves the spectral gap, $\mathrm{sav}(\xi) = m^2 \mathrm{gap}_m$, has a prevector v which is a translate of some $v' \in \mathscr{C}(B_0, R_0) \subset C^2(\mathbb{T}_m)$.*
2. *For any $v \in C^2(\mathbb{Z}^2)$ satisfying $f(G_{\mathbb{Z}^2} * v) < \frac{3}{2}\gamma$, there exists $v' \in \mathscr{C}(B_0, R_0) \subset C^2(\mathbb{Z}^2)$ such that a translate of v' differs from v by an element of $\mathcal{I}$. In particular, $f(G_{\mathbb{Z}^2} * v) = f(G_{\mathbb{Z}^2} * v')$.*

Proposition 21. *Fix $B, R_1 > 0$. For any $v \in \mathscr{C}(B, R_1)$ and $m > 2R_1$, let $\xi^{(m)} = \xi^{(m)}(v)$ be the frequency in $\hat{\mathscr{G}}_m$ corresponding to v, namely*

$$\xi_x^{(m)} = (G_{\mathbb{T}_m} * v)(x) - (G_{\mathbb{T}_m} * v)(0, 0) \quad \textit{(reduced mod } \mathbb{Z}\textit{)}, \tag{129}$$

and let $\xi = \xi(v) = G_{\mathbb{Z}^2} * v$. Then

$$\mathrm{sav}(\xi^{(m)}) \to f(\xi) \quad \textit{as } m \to \infty. \tag{130}$$

Part 2 of Proposition 20 implies that

$$\gamma = \min \left\{ f(G_{\mathbb{Z}^2} * v) : v \in \mathscr{C}(B_0, R_0) \backslash \mathcal{I} \right\}, \tag{131}$$

which reduces the computation of γ to a finite check. In "Appendix B" we verify that γ is obtained by $\xi = G_{\mathbb{Z}^2} * \delta_1 * \delta_2$ with numerical value

$$\gamma = 2.868114013(4). \tag{132}$$

Proof of Theorem 3. In this proof we use the notation $\xi(v) = G_{\mathbb{Z}^2} * v$. Take the constants B_0, R_0 from Proposition 20 and find $\gamma' > \gamma$ such that if $v \in \mathscr{C}(B_0, R_0)$ and $f(\xi(v)) > \gamma$, then $f(\xi(v)) \geq \gamma'$. Applying Proposition 21 with $B = B_0$ and $R_1 = R_0$, choose m large enough that

$$g(m) := \sup_{v \in \mathscr{C}(B_0, R_0)} |\operatorname{sav}(\xi^{(m)}(v)) - f(\xi(v))| < \frac{\gamma' - \gamma}{2}. \tag{133}$$

Let $v_0 \in \mathscr{C}(B_0, R_0)$ satisfy $f(\xi(v_0)) = \gamma$, and let $\xi_0^{(m)} = \xi^{(m)}(v_0) \in \hat{\mathscr{G}}_m$. Then

$$\operatorname{sav}(\xi_0^{(m)}) < \frac{\gamma + \gamma'}{2}. \tag{134}$$

Now suppose that $\xi^{(m)} \in \hat{\mathscr{G}}_m$ achieves the spectral gap. By translating, we may assume that $\xi^{(m)}$ has a prevector $v \in \mathscr{C}(B_0, R_0)$. We claim that $f(\xi(v)) = \gamma$: if not, then $f(\xi(v)) \geq \gamma'$ and

$$\operatorname{sav}(\xi^{(m)}) > \frac{\gamma + \gamma'}{2} > \operatorname{sav}(\xi_0^{(m)}), \tag{135}$$

a contradiction. Thus $f(\xi(v)) = \gamma$, and

$$|m^2 \operatorname{gap}_m - \gamma| = |\operatorname{sav}(\xi^{(m)}) - f(\xi(v))| \leq g(m), \tag{136}$$

with $g(m) \to 0$ as $m \to \infty$. Along with the formula (132) for γ, which is proved in "Appendix B", this concludes the proof. □

In the process of proving Propositions 20 and 21, we show two lemmas, Lemmas 22 and 23, regarding savings in the neighborhood of the support of v for prevectors $v \in \mathbb{Z}_0^{\mathbb{T}_m}$ that have bounded L^1 norm. Note that if $\xi \in \hat{\mathscr{G}}_m$ is the frequency corresponding to v and $\overline{\xi} = G_{\mathbb{T}_m} * v$, then for any $S \subset \mathbb{T}_m$, the proof of Lemma 19 implies that

$$\left| \sum_{x \in S} e(\xi_x) \right| = \left| \sum_{x \in S} e(\overline{\xi}_x) \right|, \tag{137}$$

so all savings computations can be done using $\overline{\xi}$. Indeed, if we extend the definitions (119), (121) from elements of $\hat{\mathscr{G}}_m$ to all $\mathbb{R}$- or $\mathbb{R}/\mathbb{Z}$-valued functions on $\mathbb{T}_m$, then $\operatorname{sav}(\xi; S) = \operatorname{sav}(\overline{\xi}; S)$ and $\operatorname{sav}(\xi) = \operatorname{sav}(\overline{\xi})$.

Lemma 22. *For all $A, B, R_1 > 0$ there exists an $R_2(A, B, R_1) > 2R_1$ such that if m is sufficiently large, then for any $x \in \mathbb{T}_m$ and any $v \in \mathbb{Z}^{\mathbb{T}_m}$ satisfying the following conditions:*

(1) $\|v\|_1 \leq B$
(2) $v|_{B_{R_1}(x)} \notin C^2(\mathbb{T}_m)$
(3) $d\left(x, \operatorname{supp} v|_{B_{R_1}(x)^c}\right) > 2R_2$

we have

$$\operatorname{sav}\left(G_{\mathbb{T}_m} * v; B_{R_2}(x)\right) \geq A. \tag{138}$$

Thus, if v has mean zero, then the corresponding frequency $\xi \in \hat{\mathscr{G}}_m$ satisfies $\operatorname{sav}\left(\xi; B_{R_2}(x)\right) \geq A$.

Proof. Given $v \in \mathbb{Z}^{\mathbb{T}_m}$, decompose $\overline{\xi} = G_{\mathbb{T}_m} * v$ into an internal and external component, $\overline{\xi} = \xi^i + \xi^e$, setting

$$\xi^i := G_{\mathbb{T}_m} * v|_{B_{R_1}(x)}, \qquad \xi^e := G_{\mathbb{T}_m} * v|_{B_{R_1}(x)^c}. \tag{139}$$

Treat R_2 as a parameter growing to infinity, and let R be a second parameter growing with R_2 such that $\frac{R_2}{R^3} \to \infty$ as $R_2 \to \infty$. In practice, these parameters are chosen large, but fixed, so that there is uniformity in all m sufficiently large. Since $|D_1 G_{\mathbb{T}_m}(y)|$ and $|D_2 G_{\mathbb{T}_m}(y)|$ have size $\ll 1/\|y\|_1$ as $\|y\|_1 \to \infty$, we have $\xi^e_{x+y} = \xi^e_x + O(\frac{BR}{R_2})$ for all $\|y\|_1 \le R$. Hence, by Taylor expansion,

$$\left| \sum_{\|y\|_1 \le R} e(\overline{\xi}_{x+y}) \right| = O\left(\frac{BR^3}{R_2}\right) + \left| \sum_{\|y\|_1 \le R} e(\xi^i_{x+y}) \right|. \tag{140}$$

Since the error tends to 0 as $R_2 \to \infty$, it suffices to prove that

$$\#\{y : \|y\|_1 \le R\} - \left| \sum_{\|y\|_1 \le R} e\left(\xi^i_{x+y}\right) \right| \to \infty \qquad \text{as } R \to \infty. \tag{141}$$

First suppose that $v|_{B_{R_1}(x)} \notin C^1(\mathbb{T}_m)$. For all $y = (y_1, y_2) \in \mathbb{T}_m$ with $|y_1|, |y_2| \le m/2$,

$$\xi^i_{x+y} = \sum_{\|z\|_1 \le R_1} G_{\mathbb{T}_m}(y - z) v(x + z). \tag{142}$$

Let $r = \sqrt{y_1^2 + y_2^2}$. Using the Lemma 7 bound on the first derivatives of $G_{\mathbb{T}_m}$ to approximate $G_{\mathbb{T}_m}(y - z)$ with $G_{\mathbb{T}_m}(y)$ yields

$$\xi^i_{x+y} = a G_{\mathbb{T}_m}(y) + O_{B,R_1}\left(r^{-1}\right), \tag{143}$$

where $a = \sum_{\|z\|_1 \le R_1} v(x + z) \neq 0$. The asymptotic for the first derivative of the Green's function in Lemma 8 now implies that $|\xi^i_{x+(j,0)} - \xi^i_x| \to \infty$ as $j \to \infty$, while $|\xi^i_{x+(j+1,0)} - \xi^i_{x+(j,0)}| \to 0$, so that $\{\xi^i_{x+(j,0)}\}_{j=0}^{\infty}$ is dense in $\mathbb{R}/\mathbb{Z}$, and hence

$$R - \left| \sum_{j=1}^{R} e(\xi^i_{x+(j,0)} - \xi^i_x) \right| \to \infty \qquad \text{as } R \to \infty, \tag{144}$$

which suffices for the claim.

Now suppose that $v|_{B_{R_1}(x)} \in C^1(\mathbb{T}_m) \setminus C^2(\mathbb{T}_m)$, so that it can be written as $\delta_1 * w_1 + \delta_2 * w_2$ where w_1, w_2 are $\mathbb{Z}$-valued, supported on $B_{R_1+1}(x)$, and not both in $C^1(\mathbb{T}_m)$ by (45). For all $y = (y_1, y_2)$,

$$\xi^i_{x+y} = \sum_{\|z\|_1 \le R_1 + 1} D_1 G_{\mathbb{T}_m}(y - z) w_1(x + z) + D_2 G_{\mathbb{T}_m}(y - z) w_2(x + z). \tag{145}$$

Use the Lemma 7 bound on the second derivatives of $G_{\mathbb{T}_m}$ to approximate $D_k G_{\mathbb{T}_m}(y - z)$ with $D_k G_{\mathbb{T}_m}(y)$ for $k = 1, 2$. The result is

$$\xi^i_{x+y} = a D_1 G_{\mathbb{T}_m}(y) + b D_2 G_{\mathbb{T}_m}(y) + O_{B,R_1}\left(r^{-2}\right) \tag{146}$$

for constants $a, b = O_{B,R_1}(1)$, not both zero. Lemma 8 now shows that for $1 \le r < m^{1/2}/(\log m)^{1/4}$,

$$\xi^i_{x+y} = \frac{-c(ay_1 + by_2)}{y_1^2 + y_2^2} + O_{B,R_1}\left(r^{-2}\right), \tag{147}$$

where $c > 0$ is a fixed constant. Thus $|\xi^i_{x+y}| \ll 1/r$, and there are $0 \le \theta_1 < \theta_2 < 2\pi$ such that if $\theta_1 \le \arg(y) \le \theta_2$, then $|\xi^i_{x+y}| \asymp 1/r$. It follows that

$$\sum_{\|y\|_1 \le R} (1 - c(\xi^i_{x+y})) \asymp \log R, \tag{148}$$

$$\left| \sum_{\|y\|_1 \le R} s(\xi^i_{x+y}) \right| \le \sum_{\|y\|_1 \le R} |s(\xi^i_{x+y})| \ll R. \tag{149}$$

To combine (148) and (149) we use that for all $a, b \in \mathbb{R}$ with $a > 0$,

$$\sqrt{a^2 + b^2} - \sqrt{a^2} = \int_{a^2}^{a^2+b^2} \frac{dt}{2\sqrt{t}} \le \frac{b^2}{2a}. \tag{150}$$

Letting a and b be the real and imaginary parts of $\sum_{\|y\|_1 \le R} e(\xi^i_{x+y})$, we conclude that

$$\#\{y : \|y\|_1 \le R\} - \left| \sum_{\|y\|_1 \le R} e\left(\xi^i_{x+y}\right) \right| \asymp \log R, \tag{151}$$

as required. □

Proof of Proposition 21. Given $v \in \mathscr{C}(B, R_1)$, set $\xi^* = G_{\mathbb{T}_m} * v$; we suppress the dependence on m for notational convenience. It will suffice to show that $\mathrm{sav}(\xi^*) \to f(\xi)$ as $m \to \infty$.

Write v as a sum of $O_{B,R_1}(1)$ translates of $\pm\delta_1^{*2}$, $\pm\delta_1 * \delta_2$, $\pm\delta_2^{*2}$. Since the second derivatives of the Green's function decay like the inverse square of the radius, an argument parallel to the one given in equations (145)–(146) shows that $|\xi^*_y| = O_{B,R_1}(1/r^2)$, where $r = \sqrt{y_1^2 + y_2^2}$ and $|y_1|, |y_2| \le m/2$. For all $R_1 < R < m/2$, Taylor expansion yields

$$\sum_{\|y\|_1 > R} (1 - c(\xi^*_y)) = O_{B,R_1}\left(R^{-2}\right),$$

$$\sum_{y \in \mathbb{T}_m} (1 - c(\xi^*_y)) = O_{B,R_1}(1),$$

$$\left| \sum_{y \in \mathbb{T}_m} s(\xi^*_y) \right| = O_{B,R_1}(1). \tag{152}$$

In the last estimate, we use that ξ^* is mean zero over $\mathbb{T}_m$ so that the contribution of the linear term in the Taylor expansion of $s(\xi^*_y)$ vanishes. Therefore, using (150) in the first equality,

$$\mathrm{sav}(\xi^*) = O_{B,R_1}\left(m^{-2}\right) + \sum_{y \in \mathbb{T}_m} (1 - c(\xi^*_y))$$

$$= O_{B,R_1}\left(R^{-2}\right) + \sum_{\|y\|_1 \le R} (1 - c(\xi_y^*)). \tag{153}$$

Sending $m \to \infty$ for fixed R, Lemma 9 shows that each $\xi_y^* \to \xi_y$. Thus

$$\lim_{m\to\infty} \operatorname{sav}(\xi^*) = O_{B,R_1}\left(R^{-2}\right) + \sum_{\|y\|_1 \le R} (1 - c(\xi_y)) \tag{154}$$

for each $R > R_1$. Sending $R \to \infty$ completes the proof. □

Lemma 23. *For all B, $R_1 > 0$ and $\alpha < 1$, there exists $R_2(\alpha, B, R_1) > 2R_1$ such that if m is sufficiently large, then for any $x \in \mathbb{T}_m$ and any $v \in \mathbb{Z}^{\mathbb{T}_m}$ satisfying the following conditions:*

(1) $\|v\|_1 \le B$
(2) $v|_{B_{R_1}(x)} \in C^2(\mathbb{T}_m)$
(3) $d\left(x, \operatorname{supp} v|_{B_{R_1}(x)^c}\right) > 2R_2$

we have

$$\operatorname{sav}\left(G_{\mathbb{T}_m} * v; B_{R_2}(x)\right) \ge \alpha \operatorname{sav}(\xi^*); \qquad \xi^* = G_{\mathbb{T}_m} * v|_{B_{R_1}(x)}. \tag{155}$$

Thus, if v has mean zero, then the corresponding frequency $\xi \in \hat{\mathscr{G}}_m$ satisfies $\operatorname{sav}\left(\xi; B_{R_2}(x)\right) \ge \alpha \operatorname{sav}(\xi^*)$.

Proof. First, we show that there is $\delta = \delta(B, R_1) > 0$ such that for sufficiently large m, either $\operatorname{sav}(\xi^*) = 0$ or $\operatorname{sav}(\xi^*) \ge \delta$. Translating v by $-x$ shows that $\operatorname{sav}(\xi^*) = \operatorname{sav}(G_{\mathbb{T}_m} * v')$ for some $v' \in \mathscr{C}(B, R_1)$. Let

$$\gamma' = \min\left\{f(G_{\mathbb{Z}^2} * v') : v' \in \mathscr{C}(B, R_1)\backslash\mathcal{I}\right\}, \tag{156}$$

so $\gamma' > 0$. By Proposition 21, if m is large enough then

$$|\operatorname{sav}(G_{\mathbb{T}_m} * v') - f(G_{\mathbb{Z}^2} * v')| < \gamma'/2 \quad \text{for all } v' \in \mathscr{C}(B, R_1). \tag{157}$$

Thus, if $v' \in \mathscr{C}(B, R_1)\backslash\mathcal{I}$, then $\operatorname{sav}(\xi^*) > \gamma'/2$.

If $v' \in \mathscr{C}(B, R_1) \cap \mathcal{I}$, we will show that $\operatorname{sav}(\xi^*) = 0$, allowing us to take $\delta = \gamma'/2$. Write $v' = \Delta w$ where $w \in C^0(\mathbb{Z}^2)$. Observe that $\operatorname{supp}(w)$ is a finite set, and any $(i, j) \in \mathbb{Z}^2\backslash \operatorname{supp}(w)$ that is adjacent to exactly one point in $\operatorname{supp}(w)$ must have $(\Delta w)(i, j) \ne 0$. Since $\operatorname{supp}(\Delta w) \subset B_{R_1}(0)$, it follows that $\operatorname{supp}(w) \subset B_{R_1-1}(0)$. Hence, we can consider v' and w as $\mathbb{Z}$-valued functions on $\mathbb{T}_m$ for $m > 2R_1$, and the equation $v' = \Delta w$ still holds in this context. Therefore, $G_{\mathbb{T}_m} * v' = w - c$ where c is the mean value of w on $\mathbb{T}_m$, and $\operatorname{sav}(G_{\mathbb{T}_m} * v') = 0$.

With δ in hand, we turn to the proof of (155). Set $\epsilon = \epsilon(\alpha, B, R_1) = (1 - \alpha)\delta > 0$. We will show that if m is sufficiently large,

$$\operatorname{sav}(G_{\mathbb{T}_m} * v; B_{R_2}(x)) > \operatorname{sav}(\xi^*) - \epsilon. \tag{158}$$

This implies (155), because if $\operatorname{sav}(\xi^*) = 0$ then (155) is trivial, while if $\operatorname{sav}(\xi^*) \ge \delta$ then $\operatorname{sav}(\xi^*) - \epsilon \ge \alpha \operatorname{sav}(\xi^*)$. By arguing as in Lemma 22 up to equation (140), it suffices to prove that if R is fixed but sufficiently large then

$$\operatorname{sav}(\xi^*; B_R(x)) > \operatorname{sav}(\xi^*) - \epsilon/2 \tag{159}$$

for all m sufficiently large. Writing $v|_{B_{R_1}(x)}$ as a sum of $O_{B,R_1}(1)$ translates of $\pm\delta_1^{*2}$, $\pm\delta_1 * \delta_2$, $\pm\delta_2^{*2}$, it follows as in the proof of Proposition 21 that for $y = (y_1, y_2) \in \mathbb{T}_m$ with $|y_1|, |y_2| \le m/2$ and $r = \sqrt{y_1^2 + y_2^2}$, $|\xi^*_{x+y}| = O_{B,R_1}(1/r^2)$. Taylor expansion gives

$$\sum_{\|y\|_1 \le R} (1 - c(\xi^*_{x+y})) = O_{B,R_1}(1),$$

$$\sum_{\|y\|_1 > R} (1 - c(\xi^*_{x+y})) = O_{B,R_1}\left(R^{-2}\right),$$

$$\left| \sum_{\|y\|_1 \le R} s(\xi^*_{x+y}) \right| \le \sum_{\|y\|_1 \le R} \left| s(\xi^*_{x+y}) \right| = O_{B,R_1}(\log R). \tag{160}$$

Thus,

$$\begin{aligned} \mathrm{sav}(\xi^*) &= m^2 - \left| \sum_{z \in \mathbb{T}_m} e(\xi^*_z) \right| \le \sum_{z \in \mathbb{T}_m} (1 - c(\xi^*_z)) \\ &= O_{B,R_1}\left(R^{-2}\right) + \sum_{\|y\|_1 \le R} (1 - c(\xi^*_{x+y})) \\ &= O_{B,R_1}\left(\frac{\log^2 R}{R^2}\right) + \#B_R(x) - \left| \sum_{\|y\|_1 \le R} e(\xi^*_{x+y}) \right|, \end{aligned} \tag{161}$$

using (150) in the last equality. Since this is $\mathrm{sav}(\xi^*; B_R(x))$ plus a quantity tending to zero with R, (159) is verified. □

Proof of Proposition 20. First we find a constant B_0 such that:

(I) For sufficiently large m, if $\xi^{(m)} \in \hat{\mathscr{G}}_m$ achieves the spectral gap, then its distinguished prevector $v^{(m)} = v(\xi^{(m)})$ must satisfy $\|v^{(m)}\|_1 \le B_0$.
(II) If $v \in C^2(\mathbb{Z}^2)$ satisfies $f(G_{\mathbb{Z}^2} * v) \le \frac{3}{2}\gamma + 1$, then v differs by an element of $\mathcal{I}$ from some $\tilde{v} \in C^2(\mathbb{Z}^2)$ with $\|\tilde{v}\|_1 \le B_0$.

To this end, fix any $v' \in C^2(\mathbb{Z}^2)\backslash\mathcal{I}$ and let $\gamma' = f(G_{\mathbb{Z}^2} * v') \ge \gamma$. Choose B', R' large enough that $v' \in \mathscr{C}(B', R')$, and let $m > 2R'$ so that $\mathscr{C}(B', R')$ embeds into $C^2(\mathbb{T}_m)$. Applying Proposition 21 shows that if $\xi'^{(m)}$ is the frequency in $\hat{\mathscr{G}}_m$ corresponding to v', then

$$\mathrm{sav}(\xi'^{(m)}) \to f(G_{\mathbb{Z}^2} * v') \quad \text{as } m \to \infty \tag{162}$$

and therefore $\mathrm{sav}(\xi'^{(m)}) < \gamma' + 1$ for sufficiently large m.

Suppose that $\xi^{(m)} \in \hat{\mathscr{G}}_m$ achieves the spectral gap, and let $v^{(m)}$ be the distinguished prevector of $\xi^{(m)}$. By Lemma 18,

$$\|v^{(m)}\|_1 \ll \mathrm{sav}(\xi^{(m)}) \le \mathrm{sav}(\xi'^{(m)}) < \gamma' + 1, \tag{163}$$

that is, $\|v^{(m)}\|_1$ is bounded by a universal constant. This verifies (I).

By choosing γ' arbitrarily close to γ, we could have obtained that for any $\epsilon > 0$, if m is sufficiently large then any $\xi^{(m)} \in \hat{\mathscr{G}}_m$ achieving the spectral gap must satisfy $\mathrm{sav}(\xi^{(m)}) < \gamma + \epsilon$. This will be used later.

For (II), given $v \in C^2(\mathbb{Z}^2)$, let $\xi = G_{\mathbb{Z}^2} * v$. Lemma 9 or Lemma 12 shows that $\xi \in \ell^2(\mathbb{Z}^2)$, so there are only finitely many $x \in \mathbb{Z}^2$ such that $|\xi_x| \geq \frac{1}{2}$. Reduce ξ to $\tilde{\xi} : \mathbb{Z}^2 \to \left[-\frac{1}{2}, \frac{1}{2}\right)$ by subtracting $w \in C^0(\mathbb{Z}^2)$, and let $\tilde{v} = \Delta\tilde{\xi} = v - \Delta w$, which differs from v by $\Delta w \in \mathcal{I}$ and is therefore in $C^2(\mathbb{Z}^2)$. Because $\tilde{v}$ is integer-valued and Δ is bounded from $\ell^2(\mathbb{Z}^2) \to \ell^2(\mathbb{Z}^2)$,

$$\|\tilde{v}\|_1 \leq \|\tilde{v}\|_2^2 = \|\Delta\tilde{\xi}\|_2^2 \ll \|\tilde{\xi}\|_2^2 = \sum_{x\in\mathbb{Z}^2} |\tilde{\xi}_x|^2 \ll \sum_{x\in\mathbb{Z}^2} (1 - c(\tilde{\xi}_x)), \tag{164}$$

where the last inequality uses $1 - c(t) \gg t^2$ for $|t| \leq \frac{1}{2}$. The right side is $f(\tilde{\xi}) = f(G_{\mathbb{Z}^2} * v)$, so an upper bound on $f(G_{\mathbb{Z}^2} * v)$ translates to an upper bound on $\|\tilde{v}\|_1$, confirming (II).

Fix B_0 to satisfy (I) and (II). For any $v \in \mathbb{Z}^{\mathbb{T}_m}$ with $\|v\|_1 \leq B_0$, we perform a clustering on $\mathrm{supp}(v)$, as follows.

(1) Initially all of $\mathrm{supp}\, v$ is uncovered and initialize a list $\mathscr{X}$ of centers of balls to be empty.
(2) Iterate until $\mathrm{supp}\, v$ is covered:
  (a) Choose $x \in \mathrm{supp}\, v$ which is uncovered and append x to $\mathscr{X}$.
  (b) Beginning from an initial guess $R_1(x) = 1$:
    (i) If $v|_{B_{R_1(x)}(x)} \notin C^2(\mathbb{T}_m)$, then choose $R_2(x)$ according to Lemma 22 with $A = \frac{3}{2}\gamma + 1$, $B = B_0$, and $R_1 = R_1(x)$. If $v|_{B_{R_1(x)}(x)} \in C^2(\mathbb{T}_m)$, then choose $R_2(x)$ according to Lemma 23 with $\alpha = \frac{7}{8}$, $B = B_0$, and $R_1 = R_1(x)$.
    (ii) If the condition of those lemmas holds,

$$d\left(x, \mathrm{supp}\, v|_{B_{R_1(x)}(x)^c}\right) > 2R_2(x),$$

then declare those $y \in \mathrm{supp}(v) \cap B_{R_1}(x)$ covered and continue to (c). Otherwise, replace $R_1(x) := 2R_2(x)$ and repeat step (i).
  (c) If all of $\mathrm{supp}\, v$ is covered, finish. If not, return to (a).
(3) Since $\|v\|_1 \leq B_0$, the process stops after boundedly many steps.

At the end of this process,

$$\mathrm{supp}\, v \subset \bigcup_{x\in\mathscr{X}} B_{R_1(x)}(x), \tag{165}$$

$$d\left(x, \mathrm{supp}\, v|_{B_{R_1(x)}(x)^c}\right) > 2R_2(x) \quad \text{for each } x \in \mathscr{X}. \tag{166}$$

We claim that a subset $\mathscr{X}' \subset \mathscr{X}$ can be chosen such that

$$\mathrm{supp}\, v \subset \bigsqcup_{x\in\mathscr{X}'} B_{R_1(x)}(x). \tag{167}$$

Note a particular consequence of (166) and (167) is that the balls

$$\{B_{R_2(x)}(x)\}_{x\in\mathscr{X}'} \tag{168}$$

are pairwise disjoint.

To verify (167), let x and x' be centers of balls of the process with x appearing prior to x' in the list. We will show that either $B_{R_1(x)}(x)$ and $B_{R_1(x')}(x')$ are disjoint, or $B_{R_1(x)}(x) \cap \operatorname{supp} v \subset B_{R_1(x')}(x')$. First, since $x' \notin B_{R_1(x)}(x)$, we have $d(x, x') > 2R_2(x) \geq 2R_1(x)$. Suppose y is in the intersection of $B_{R_1(x)}(x)$ and $B_{R_1(x')}(x')$, so that

$$d(x, x') \leq d(x, y) + d(y, x') \leq R_1(x) + R_1(x'). \tag{169}$$

Combining this with the lower bound on $d(x, x')$ gives $R_1(x) < R_1(x')$. Therefore, (169) implies that $d(x, x') < 2R_1(x') \leq 2R_2(x')$, whence $d(x, x') \leq R_1(x')$. For any $z \in B_{R_1(x)}(x) \cap \operatorname{supp} v$,

$$d(z, x') \leq d(z, x) + d(x, x') \leq R_1(x) + R_1(x') < 2R_2(x') \tag{170}$$

so that, in fact, $z \in B_{R_1(x')}(x')$ and $B_{R_1(x)}(x) \cap \operatorname{supp} v \subset B_{R_1(x')}(x')$. Hence, starting from $\mathscr{X}$ we obtain the desired list $\mathscr{X}'$ by discarding any x which satisfies $x \in B_{R_1(x')}(x')$ for some x' later in the list.

Let $\overline{R}_1(v) = \max\{R_1(x) : x \in \mathscr{X}'\}$. From the description of the clustering algorithm, there is a uniform in m upper bound on $\overline{R}_1(v)$ that depends only on B_0:

$$\overline{R}_1(v) \leq R_0 = R_0(B_0). \tag{171}$$

Fix

$$\gamma_0 = \min\left\{ f(G_{\mathbb{Z}^2} * v) : v \in \mathscr{C}(B_0, R_0) \backslash \mathcal{I} \right\}. \tag{172}$$

We will show that $\gamma = \gamma_0$, but a priori we only know that $\gamma_0 \geq \gamma$ and $\gamma_0 > 0$. Proposition 21 implies that if m is large enough,

$$|\operatorname{sav}(G_{\mathbb{T}_m} * v) - f(G_{\mathbb{Z}^2} * v)| < \gamma_0/8 \quad \text{for all } v \in \mathscr{C}(B_0, R_0). \tag{173}$$

We can now prove Part 1 of Proposition 20. Let $\xi \in \hat{\mathscr{G}}_m$ achieve the spectral gap, and take m large enough that $\operatorname{sav}(\xi) < \min(\frac{3}{2}\gamma + 1, \frac{3}{2}\gamma_0)$, noting that the upper bound is strictly greater than γ. Let $v \in \mathbb{Z}_0^{\mathbb{T}_m}$ be the distinguished prevector of ξ, with $\|v\|_1 \leq B_0$, and run the clustering algorithm on v. Set

$$\mathscr{X}'' = \left\{ x \in \mathscr{X}' : v|_{B_{R_1(x)}(x)} \notin \mathcal{I} \right\} \tag{174}$$

and define v' to equal v on each $B_{R_1(x)}(x)$ for $x \in \mathscr{X}''$, while $v' \equiv 0$ elsewhere. To get from v to v', we subtracted finitely many elements of $\mathcal{I}$. It follows that $v' = v - \Delta w$ for some $w \in \mathbb{Z}^{\mathbb{T}_m}$; the proof of Lemma 23 explains why this holds on $\mathbb{T}_m$ as well as on $\mathbb{Z}^2$. We conclude that v' is also a prevector of ξ.

Given $x \in \mathscr{X}''$, for notational convenience set $u^{(x)} = v|_{B_{R_1(x)}(x)}$. If some $u^{(x)} \notin C^2(\mathbb{T}_m)$, then by construction, $\operatorname{sav}(\xi) \geq \frac{3}{2}\gamma + 1$, a contradiction. Therefore, each $u^{(x)} \in C^2(\mathbb{T}_m)$. Since $R_1(x) \leq R_0$ and $u^{(x)} \notin \mathcal{I}$, we have $f(G_{\mathbb{Z}^2} * u^{(x)}) \geq \gamma_0$. It follows from (173) that $\operatorname{sav}(G_{\mathbb{T}_m} * u^{(x)}) > \frac{7}{8}\gamma_0$. Then, by step (i) of the clustering,

$$\operatorname{sav}(\xi; B_{R_2(x)}(x)) > \left(\frac{7}{8}\right)^2 \gamma_0 > \frac{3}{4}\gamma_0. \tag{175}$$

If $|\mathscr{X}''| \geq 2$, then $\operatorname{sav}(\xi) > \frac{3}{2}\gamma_0$, another contradiction. We conclude that $|\mathscr{X}''| = 1$, and moreover, for the unique $x \in \mathscr{X}''$, $v' = u^{(x)} \in C^2(\mathbb{T}_m)$. Translating v' by $-x$ yields an element of $\mathscr{C}(B_0, R_0)$. This proves Part 1 of Proposition 20.

Part 2 is proved along similar lines. Let $v \in C^2(\mathbb{Z}^2)$ satisfy

$$f(G_{\mathbb{Z}^2} * v) < \min(\tfrac{3}{2}\gamma + 1, \tfrac{3}{2}\gamma_0). \tag{176}$$

By property (II), we may assume that $\|v\|_1 \leq B_0$ (by subtracting an element of $\mathcal{I}$ if necessary). Since $\operatorname{supp} v$ is finite, it embeds into $\mathbb{T}_m$ for large enough m, and then v can be seen as an element of $C^2(\mathbb{T}_m)$. Let $\xi^{(m)} \in \hat{\mathscr{G}}_m$ be the frequency corresponding to v, so that $\operatorname{sav}(\xi^{(m)}) \to f(G_{\mathbb{Z}^2} * v)$ as $m \to \infty$, by Proposition 21.

Run the clustering algorithm on v, noting that independent of the value of m, the algorithm will follow exactly the same steps and produce identical clusters. Define $\mathscr{X}''$, v', and the notation $u^{(x)}$ as in the proof of Part 1. If $u^{(x)} \notin C^2(\mathbb{T}_m)$ for some $x \in \mathscr{X}''$, then $\operatorname{sav}(\xi^{(m)}) \geq \frac{3}{2}\gamma + 1$. Since the property "$u^{(x)} \notin C^2(\mathbb{T}_m)$" is independent of m, this is a uniform lower bound on all $\operatorname{sav}(\xi^{(m)})$ and so $f(G_{\mathbb{Z}^2} * v) \geq \frac{3}{2}\gamma + 1$. Likewise, if each $u^{(x)} \in C^2(\mathbb{T}_m)$ but $|\mathscr{X}''| \geq 2$, then $\operatorname{sav}(\xi^{(m)}) > \frac{3}{2}\gamma_0$ for all m and so $f(G_{\mathbb{Z}^2} * v) \geq \frac{3}{2}\gamma_0$. Both possibilities contradict (176).

We conclude that any $v \in C^2(\mathbb{Z}^2)$ satisfying (176) differs by an element of $\mathcal{I}$ from some $v' \in C^2(\mathbb{Z}^2)$ that has a translate in $\mathscr{C}(B_0, R_0)$. It follows that γ is equal to the right side of (172), that is, $\gamma = \gamma_0$. In particular, $\gamma > 0$. Finally, the right side of (176) simplifies to $\frac{3}{2}\gamma$, proving Part 2 of Proposition 20. □

The following lemma provides the additive savings needed to prove the lower bound in Theorem 2.

Lemma 24. *Let $k \geq 1$ be fixed, and let $v_1, \ldots, v_k \in C^2(\mathbb{T}_m)$ be bounded functions of bounded support which are R-separated, in the sense that their supports have pairwise ℓ^1 distance at least R. Set $v = \sum_{i=1}^k v_i$. Then as $R \to \infty$,*

$$1 - \left|\hat{\mu}(\xi(v))\right| = O\left(\frac{\log(1+R)}{R^2 m^2}\right) + \sum_{i=1}^{k} \left(1 - \left|\hat{\mu}(\xi(v_i))\right|\right). \tag{177}$$

The implicit constant depends upon k and the bounds for the functions and their supports.

Proof. Set $\overline{\xi} = G_{\mathbb{T}_m} * v$ and $\overline{\xi}_i = G_{\mathbb{T}_m} * v_i$, so that $|\hat{\mu}(\xi(v))| = |\hat{\mu}(\overline{\xi})|$ and $|\hat{\mu}(\xi(v_i))| = |\hat{\mu}(\overline{\xi}_i)|$. Fix a point x_i in the support of each v_i, so that the balls $B_{R'}(x_i)$ are disjoint where $R' = \lfloor (R-1)/2 \rfloor$. As in the proof of Proposition 21, if $y = (y_1, y_2)$ with $|y_1|, |y_2| \leq m/2$ and $r = \sqrt{y_1^2 + y_2^2}$,

$$|\overline{\xi}_i(x_i + y)| = O(1/r^2). \tag{178}$$

We obtain the analogue to (153),

$$1 - |\hat{\mu}(\overline{\xi}_i)| = O\left(\frac{1}{R^2 m^2}\right) + \frac{1}{m^2} \sum_{\|y\|_1 \leq R'} \left(1 - c\left(\overline{\xi}_i(x_i + y)\right)\right). \tag{179}$$

If $\|y\|_1 \leq R'$, then $\overline{\xi}(x_i + y) = \overline{\xi}_i(x_i + y) + O(R^{-2})$, so that

$$c\left(\overline{\xi}(x_i + y)\right) = c\left(\overline{\xi}_i(x_i + y)\right) + O\left(\frac{\left|s(\overline{\xi}_i(x_i + y))\right|}{R^2}\right) + O\left(\frac{1}{R^4}\right). \tag{180}$$

As in (160),

$$\sum_{\|y\|_1 \le R'} \left| s\left(\overline{\xi}_i(x_i + y)\right)\right| = O(\log(1+R)). \tag{181}$$

Combining (179), (180), and (181) yields

$$1 - |\hat{\mu}(\overline{\xi}_i)| = O\left(\frac{\log(1+R)}{R^2 m^2}\right) + \frac{1}{m^2} \sum_{\|y\|_1 \le R'} \left(1 - c\left(\overline{\xi}(x_i + y)\right)\right). \tag{182}$$

Take the sum of (182) over $i = 1, 2, \ldots, k$. For $z \notin \bigcup_{i=1}^k B_{R'}(x_i)$, let r_i be the ℓ^2 distance from z to x_i, so that $|\overline{\xi}(z)| = O(1/r_1^2 + \cdots + 1/r_k^2)$. Use the inequality

$$\left(\frac{1}{r_1^2} + \cdots + \frac{1}{r_k^2}\right)^2 \le k\left(\frac{1}{r_1^4} + \cdots + \frac{1}{r_k^4}\right) \tag{183}$$

to conclude that

$$\sum_{z \notin \bigcup_{i=1}^k B_{R'}(x_i)} \left(1 - c\left(\overline{\xi}(z)\right)\right) = O\left(\frac{1}{R^2}\right). \tag{184}$$

In combination with (182), this yields

$$\sum_{i=1}^k \left(1 - \left|\hat{\mu}(\overline{\xi}_i)\right|\right) = O\left(\frac{\log(1+R)}{R^2 m^2}\right) + \frac{1}{m^2} \sum_{z \in \mathbb{T}_m} \left(1 - c\left(\overline{\xi}(z)\right)\right), \tag{185}$$

or equivalently,

$$1 - \mathrm{Re}\left(\hat{\mu}(\overline{\xi})\right) = O\left(\frac{\log(1+R)}{R^2 m^2}\right) + \sum_{i=1}^k \left(1 - \left|\hat{\mu}(\overline{\xi}_i)\right|\right). \tag{186}$$

We will finish the proof by applying (150). Proposition 21 implies that each $1 - \left|\hat{\mu}(\overline{\xi}_i)\right| = O(1/m^2)$, so $\mathrm{Re}\left(\hat{\mu}(\overline{\xi})\right) \gg 1$. Meanwhile,

$$\mathrm{Im}\left(\hat{\mu}(\overline{\xi})\right) = \frac{1}{m^2} \sum_{z \in \mathbb{T}_m} s\left(\overline{\xi}(z)\right), \qquad s(t) = 2\pi t + O\left(|t|^3\right). \tag{187}$$

In the Taylor expansion, the linear term vanishes since $\overline{\xi}$ has mean zero on $\mathbb{T}_m$. Also,

$$\left|\overline{\xi}(z)\right|^3 \le \left(\sum_{i=1}^k \left|\overline{\xi}_i(z)\right|\right)^3 \le k^2 \sum_{i=1}^k \left|\overline{\xi}_i(z)\right|^3, \tag{188}$$

and $\sum_{z \in \mathbb{T}_m} \left|\overline{\xi}_i(z)\right|^3 = O(1)$ by (178). Hence $\mathrm{Im}\left(\hat{\mu}(\overline{\xi})\right) = O(1/m^2)$, and we obtain the desired result from (186) using (150). □

6.2. Estimation of moderate size phases. In this section we give estimates for the savings of frequencies ξ whose distinguished prevectors $v = v(\xi)$ have $\|v\|_1$ growing with m. In particular, we prove an approximate additive savings estimate for separated parts of v, which is what is needed to prove the upper bound of Theorem 2.

Let $R > 1$ be a large fixed parameter. Given any $v \in \mathbb{Z}^{\mathbb{T}_m}$, for each $x \in \operatorname{supp} v$ let

$$\operatorname{nbd}(x) := B_R(x) = \{y \in \mathbb{T}_m : \|y - x\|_1 \le R\}. \tag{189}$$

Perform a simple agglomeration scheme, in which any two points $x, y \in \operatorname{supp} v$ whose neighborhoods overlap are joined in a common R-cluster. In other words, x and y belong to a common cluster if and only if there is a sequence of points $\{z_i\}_{i=0}^n \subset \operatorname{supp} v$ such that $x = z_0$, $y = z_n$ and, for $0 \le i < n$, $\|z_i - z_{i+1}\|_1 \le 2R$. Write $\mathscr{C}$ for the collection of clusters formed in this way. Given $C \in \mathscr{C}$, write

$$\operatorname{nbd}(C) := \bigcup_{x \in C} \operatorname{nbd}(x) \tag{190}$$

for the neighborhood of C, so that $\operatorname{supp} v \subset \bigsqcup_{C \in \mathscr{C}} \operatorname{nbd}(C)$.

Let $\mathscr{P} \subset \mathscr{C}$ be the collection of all clusters C such that $v|_C = \Delta w$ for some $w \in \mathbb{Z}^{\mathbb{T}_m}$, and let S be the union of all clusters $C \in \mathscr{C} \setminus \mathscr{P}$. The '$R$-reduction' of v is defined to be $\tilde{v} = v|_S$, which differs from v by a sum of terms of the form Δw, and whose L^1 and L^∞ norms are bounded by $\|v\|_1$, $\|v\|_{L^\infty}$ respectively. We say that v is 'R-reduced' if $\tilde{v} = v$. For any frequency $\xi \in \hat{\mathscr{G}}_m$, the '$R$-reduced prevector' of ξ is the R-reduction of the distinguished prevector $v(\xi)$, which is indeed a prevector of ξ.

The following is the main result of this section. It is similar to Lemmas 22 and 23, but does not require the prevector v to have bounded L^1 norm.

Lemma 25. *Let $B \ge 1$ be a fixed parameter. There is a function $\eta(B, R)$ tending to 0 as $R \to \infty$ such that for all m sufficiently large, if $v \in \mathbb{Z}^{\mathbb{T}_m}$ satisfies the following conditions:*

(1) *v is R-reduced*
(2) *$\|v\|_{L^\infty} \le 3$*
(3) *v has an R-cluster C for which $\| v|_C \|_1 \le B$*

then

$$\operatorname{sav}(G_{\mathbb{T}_m} * v; \operatorname{nbd}(C)) \ge m^2 \operatorname{gap}_m - \eta(B, R). \tag{191}$$

Thus, if v has mean zero, then the corresponding frequency $\xi \in \hat{\mathscr{G}}_m$ satisfies $\operatorname{sav}(\xi; \operatorname{nbd}(C)) \ge m^2 \operatorname{gap}_m - \eta(B, R)$.

The sufficiently large value of m above which (191) holds is allowed to depend on both B and R.

The upper bound on $\|v\|_{L^\infty}$ could be replaced by any fixed constant; we chose 3 because the distinguished prevector of every $\xi \in \hat{\mathscr{G}}_m$ satisfies $\|v(\xi)\|_{L^\infty} \le 3$, so the R-reduced prevector has the same bound.

Proof. Suppose that $v \in \mathbb{Z}^{\mathbb{T}_m}$ satisfies the conditions of the lemma. We decompose the phase function $\overline{\xi} = G_{\mathbb{T}_m} * v$ into an internal and external component, $\overline{\xi} = \xi^i + \xi^e$, where

$$\xi^i := G_{\mathbb{T}_m} * v|_C, \qquad \xi^e := G_{\mathbb{T}_m} * v|_{C^c}. \tag{192}$$

Our first observation is that the third derivatives of ξ^e are uniformly bounded over all $x \in \mathrm{nbd}(C)$:

$$\left| D_1^a D_2^b \xi_x^e \right| \ll \frac{1}{R}, \quad \text{for } x \in \mathrm{nbd}(C) \text{ and } a, b \geq 0, a + b = 3. \tag{193}$$

To see this, note that if $x \in \mathrm{nbd}(C)$, then every $y \in \mathrm{supp}(v) \backslash C$ satisfies $\|x - y\|_1 > R$. Therefore,

$$\begin{aligned} \left| D_1^a D_2^b \xi_x^e \right| &= \left| \sum_{y \in C^c} v(y) D_1^a D_2^b G_{\mathbb{T}_m}(x - y) \right| \\ &\leq \|v\|_{L^\infty} \sum_{y \in B_R(x)^c} |D_1^a D_2^b G_{\mathbb{T}_m}(x - y)|. \end{aligned} \tag{194}$$

The bound (193) then follows from the asymptotic of Lemma 7.

The rest of the proof is divided into three cases. Heuristically, ξ^e could be roughly constant over $\mathrm{nbd}(C)$, vary linearly over $\mathrm{nbd}(C)$, or vary quadratically over $\mathrm{nbd}(C)$. The bound on the third derivatives of ξ^e ensures that these are the only possibilities. If ξ^e is roughly constant, then we can prove (191) using the arguments developed in Sect. 6.1.

If ξ^e varies linearly over $\mathrm{nbd}(C)$, then we can find a region of $\mathrm{nbd}(C)$ far enough away from C that the internal phase is nearly constant, so $\overline{\xi} = \xi^i + \xi^e$ varies linearly. We then cite the geometric series bound of Lemma 11 to show that for any $A > 0$, if R is large enough then $\mathrm{sav}(\overline{\xi}; \mathrm{nbd}(C)) \geq A$. This is much stronger than the desired bound (191): as long as $A > \gamma = \lim_{m \to \infty} m^2 \mathrm{gap}_m$, we do not even need to subtract $\eta(B, R)$.

Finally, if ξ^e varies quadratically over $\mathrm{nbd}(C)$, we use van der Corput's inequality to reduce to the linear case.

The proof will use three auxiliary parameters R_1, R_2, R_3 which tend to infinity with R and satisfy $R_1 < R_2 < R_3 < R$. We require that

$$R_1 \to \infty, \quad \frac{R_2}{R_1^4} \to \infty, \quad \frac{R_3}{R_1 R_2^2} \to \infty, \quad \frac{R}{R_1^2 R_3^2} \gg 1, \quad \text{as } R \to \infty. \tag{195}$$

These properties are all satisfied if, for example, $R_1 = R^{1/26}$, $R_2 = R^{5/26}$, $R_3 = R^{12/26}$.

For the first case, suppose that for all $x \in C$ and $\|y - x\|_1 \leq R_1$, we have $\left\| \xi_y^e - \xi_x^e \right\|_{\mathbb{R}/\mathbb{Z}} < 1/R_1^3$. Perform the clustering algorithm from the proof of Proposition 20 on $v|_C$, using the same parameters (e.g. $\alpha = 7/8$). This partitions C into sub-clusters indexed by a set $\mathscr{X}'$: for each $x \in \mathscr{X}'$ there are radii $2\tilde{R}_1(x) < \tilde{R}_2(x)$ such that

$$C \subset \bigsqcup_{x \in \mathscr{X}'} B_{\tilde{R}_1(x)}(x), \tag{196}$$

while the balls $\{B_{\tilde{R}_2(x)}(x)\}_{x \in \mathscr{X}'}$ are disjoint, and each sub-cluster meets the conditions of either Lemma 22 or Lemma 23, as appropriate. As in (171), the radii $\tilde{R}_2(x)$ are uniformly bounded by some R_0 depending only on B. By taking R large enough with respect to B, we may assume that R_0 is arbitrarily small relative to R_1.

Let $x \in \mathscr{X}'$ and $R' \le R_1$. We use the assumption that $\left\|\xi^e_y - \xi^e_x\right\|_{\mathbb{R}/\mathbb{Z}} < 1/R_1^3$ for all $y \in B_{R_1}(x)$ to compute, by Taylor expansion,

$$\left|\sum_{y\in B_{R'}(x)} e(\xi^i_y + \xi^e_y)\right| = \left|\sum_{y\in B_{R'}(x)} e(\xi^i_y)\right| + O\left(\frac{1}{R_1}\right). \tag{197}$$

In other words, for all $R' \le R_1$,

$$\operatorname{sav}(\bar{\xi}; B_{R'}(x)) = \operatorname{sav}(\xi^i; B_{R'}(x)) + O\left(R_1^{-1}\right). \tag{198}$$

Define $\mathscr{X}'' \subset \mathscr{X}'$ as in the proof of Proposition 20. Since v is R-reduced, $\mathscr{X}''$ is nonempty. For $x \in \mathscr{X}''$, let $u^{(x)}$ be the restriction of v to $B_{\tilde{R}_1(x)}(x)$. By step (i) of the clustering,

$$\operatorname{sav}(\xi^i; B_{\tilde{R}_2(x)}(x)) \ge \begin{cases} \frac{3}{2}\gamma + 1, & u^{(x)} \notin C^2(\mathbb{T}_m), \\ \frac{7}{8}\operatorname{sav}(G_{\mathbb{T}_m} * u^{(x)}), & u^{(x)} \in C^2(\mathbb{T}_m). \end{cases} \tag{199}$$

Note that $\frac{3}{2}\gamma + 1 > m^2\mathrm{gap}_m$ for large enough m, and $\operatorname{sav}(G_{\mathbb{T}_m} * u^{(x)}) \ge m^2\mathrm{gap}_m$ by definition of $\mathscr{X}''$. Thus, the combination of (198) with (199) verifies the desired bound (191) except when $|\mathscr{X}''| = 1$ and $u^{(x)} \in C^2(\mathbb{T}_m)$. In that remaining situation, we observe from (161) that

$$\begin{aligned} \operatorname{sav}(\xi^i; B_{R_1}(x)) &= \operatorname{sav}(G_{\mathbb{T}_m} * u^{(x)}; B_{R_1}(x)) \\ &= \operatorname{sav}(G_{\mathbb{T}_m} * u^{(x)}) - O_B\left(\frac{\log^2 R_1}{R_1^2}\right) \\ &\ge m^2\mathrm{gap}_m - O_B\left(\frac{\log^2 R_1}{R_1^2}\right), \end{aligned} \tag{200}$$

which along with (198) completes the proof.

In the second and third cases, we assume that there exist $x \in C$ and $y \in B_{R_1}(x)$ such that $d := \|\xi^e_y - \xi^e_x\|_{\mathbb{R}/\mathbb{Z}} \ge 1/R_1^3$. Set $w = y - x$, so $\|w\|_1 \le R_1$. For the second case, suppose that for all integers $1 \le n \le \frac{R_2}{\|w\|_1}$,

$$\left\|\xi^e_{x+nw} - \xi^e_x - n\left(\xi^e_y - \xi^e_x\right)\right\|_{\mathbb{R}/\mathbb{Z}} < \frac{1}{R_1}. \tag{201}$$

Effectively, the external phase varies linearly along the discrete line $\{x + nw : n \in \mathbb{Z},\ 0 \le n \le \frac{R_2}{\|w\|_1}\}$.

We now find a segment along the line that is far away from C. Set

$$\ell = \left\lfloor \frac{R_2}{3^B\|w\|_1} \right\rfloor, \tag{202}$$

and consider the ℓ^1-balls of radius $3^k\ell\|w\|_1$ centered at $x + 2\cdot 3^k\ell w$, for $0 \le k \le B-1$. The interiors of these balls are disjoint, and $x \in C$ is not in any of the interiors. By the pigeonhole principle, the interior of at least one ball contains no elements of C. Choose k corresponding to one such ball, and set $U = 2\cdot 3^k\ell$, so that $x + Uw$ is the center.

Set $V = \lfloor \sqrt{\ell d^{-1}} \rfloor$. By (195), $\frac{\ell}{V} \geq \sqrt{\ell d} \to \infty$ with R, and certainly $V \to \infty$ with R. Any point y along a shortest path from $x + Uw$ to $x + nw$, with $U < n \leq U + V$, satisfies $d(y, C) \gg U\|w\|_1$. Since the first derivatives of $G_{\mathbb{T}_m}$ decay like the inverse of the radius, it follows that $\xi^i_{x+nw} - \xi^i_{x+Uw} = O\left(\frac{V}{U}\right) = o_R(1)$.

Consider the exponential sum

$$\begin{aligned}\sum_{n=U}^{U+V} e\left(\overline{\xi}_{x+nw}\right) &= \sum_{n=U}^{U+V} e\left(\xi^e_{x+nw} + \xi^i_{x+nw}\right) \\ &= \sum_{n=U}^{U+V} e\left(\xi^e_x + \xi^i_{x+Uw} + n\left(\xi^e_y - \xi^e_x\right) + O\left(\frac{1}{R_1}\right) + O\left(\frac{V}{U}\right)\right). \end{aligned} \tag{203}$$

Taylor expanding the error in the exponential, then summing the geometric series, we obtain

$$\sum_{n=U}^{U+V} e\left(\overline{\xi}_{x+nw}\right) = O\left(\frac{V}{R_1} + \frac{V^2}{U} + \frac{1}{d}\right) = o_R(V). \tag{204}$$

Hence, this segment of the line provides savings of $\gg V$ for $\overline{\xi}$. That is, $\mathrm{sav}(\overline{\xi}; \mathrm{nbd}(C))$ is bounded below by a constant that may be made arbitrarily high by taking R large enough.

For the third case, suppose that (201) fails for some $n = n_1 \leq \frac{R_2}{\|w\|_1}$. Set $W = \lfloor \frac{R_3}{\|w\|_1} \rfloor$, and apply van der Corput's inequality with $H = 1$ to estimate

$$\left|\sum_{n=1}^{W} e\left(\overline{\xi}_{x+nw}\right)\right|^2 \leq \frac{W(W+1)}{2} + \frac{W+1}{2}\left|\sum_{n=1}^{W-1} e\left(\overline{\xi}_{x+(n+1)w} - \overline{\xi}_{x+nw}\right)\right|. \tag{205}$$

Set $z_n = \xi^e_{x+(n+1)w} - \xi^e_{x+nw}$. By the definition of n_1,

$$\frac{1}{R_1} \leq \left\|\sum_{i=0}^{n_1-1}(z_i) - n_1 z_0\right\|_{\mathbb{R}/\mathbb{Z}} = \left\|\sum_{i=0}^{n_1-2}(n_1 - 1 - i)(z_{i+1} - z_i)\right\|_{\mathbb{R}/\mathbb{Z}} \tag{206}$$

and therefore there is $n_0 \leq n_1 - 2$ for which

$$\delta := \|z_{n_0+1} - z_{n_0}\|_{\mathbb{R}/\mathbb{Z}} \geq \frac{2}{R_1 n_1^2} \geq \frac{2\|w\|_1^2}{R_1 R_2^2}. \tag{207}$$

For all $1 \leq n, p \leq W$, the quantity $z_n - z_p - (n-p)(z_{n_0+1} - z_{n_0})$ is a sum of $O\left(W^2\|w\|_1^3\right)$ terms of the form $D_1^a D_2^b \xi^e_{x'}$ where $a + b = 3$ and $x' \in \mathrm{nbd}(C)$. (Each z_i is a sum of $O(\|w\|_1)$ first derivatives of ξ^e, so $z_{i+1} - z_i$ is a sum of $O\left(\|w\|_1^2\right)$ second derivatives of ξ^e, and then $(z_{i+1} - z_i) - (z_{n_0+1} - z_{n_0})$ is a sum of $O\left(\|w\|_1^2 \cdot W\|w\|_1\right)$ third derivatives of ξ^e. Finally, sum over all i between p and n.) Thus, (193) gives

$$\left\|z_n - z_p - (n-p)(z_{n_0+1} - z_{n_0})\right\|_{\mathbb{R}/\mathbb{Z}} = O\left(\frac{W^2\|w\|_1^3}{R}\right). \tag{208}$$

By the definition of W and (195), this quantity is $O(1/R_1)$.

We now repeat the argument of the previous case, using

$$\ell' = \left\lfloor \frac{R_3}{3^B \|w\|_1} \right\rfloor \tag{209}$$

to define $U' = 2 \cdot 3^k \ell'$ for an appropriately chosen $0 \le k \le B-1$, and $V' = \lfloor \sqrt{\ell' \delta^{-1}} \rfloor$. By (195), $\frac{\ell'}{V'} \ge \sqrt{\ell' \delta} \to \infty$ with R, and $\delta^{-1} = o_R(V')$. Arguing as before, we obtain

$$\sum_{n=U'}^{U'+V'} e\left(\overline{\xi}_{x+(n+1)w} - \overline{\xi}_{x+nw}\right) = O\left(\frac{V'}{R_1} + \frac{(V')^2}{U'} + \frac{1}{\delta}\right) = o_R(V'). \tag{210}$$

Thus, we have saved an arbitrary constant in the sum

$$\sum_{n=1}^{W-1} e\left(\overline{\xi}_{x+(n+1)w} - \overline{\xi}_{x+nw}\right), \tag{211}$$

and hence also in $\sum_{n=1}^{W} e\left(\overline{\xi}_{x+nw}\right)$, by (205). □

7. Proof of Theorem 2

In the process of proving Theorem 2, we also prove the following mixing result in L^2.

Theorem 26. *Let $m \ge 2$, let $c_0 = \gamma^{-1}$ be the constant of Theorem 2, and as there, set $t_m^{\mathrm{mix}} = c_0 m^2 \log m$. For each fixed $\epsilon > 0$,*

$$\begin{aligned} \lim_{m\to\infty} \min_{\sigma \in \mathscr{R}_m} \left\| P_m^{\lceil (1-\epsilon) t_m^{\mathrm{mix}} \rceil} \delta_\sigma - \mathbb{U}_{\mathscr{R}_m} \right\|_{L^2(d\mathbb{U}_{\mathscr{R}_m})} &= \infty \\ \lim_{m\to\infty} \max_{\sigma \in \mathscr{R}_m} \left\| P_m^{\lfloor (1+\epsilon) t_m^{\mathrm{mix}} \rfloor} \delta_\sigma - \mathbb{U}_{\mathscr{R}_m} \right\|_{L^2(d\mathbb{U}_{\mathscr{R}_m})} &= 0. \end{aligned} \tag{212}$$

Note that, since we restrict to recurrent states, Parseval gives the following characterization of the $L^2(d\mathbb{U}_{\mathscr{R}_m})$ norm,

$$\left\| P_m^N \delta_\sigma - \mathbb{U}_{\mathscr{R}_m} \right\|^2_{L^2(d\mathbb{U}_{\mathscr{R}_m})} = \sum_{\xi \in \hat{\mathscr{G}}_m \setminus \{0\}} \left|\hat{\mu}(\xi)\right|^{2N}. \tag{213}$$

7.1. Proof of the lower bound. Our proof of the lower bound in Theorem 2 uses the following second moment lemma, a variant of the method used by Diaconis and Shahshahani [15] to show cutoff in the Bernoulli–Laplace diffusion model (see also [13]).

Given any probability measure μ on a finite abelian group $\mathscr{G}$, recall from Sect. 5.1 the definitions of the dual group $\hat{\mathscr{G}}$ and the Fourier coefficients $\hat{\mu}(\xi)$, for $\xi \in \hat{\mathscr{G}}$.

Lemma 27. *Let $\mathscr{G}$ be a finite abelian group, let μ be a probability measure on $\mathscr{G}$ and let $N \geq 1$. Let $\mathscr{X} \subset \hat{\mathscr{G}} \backslash \{0\}$. Suppose that the following inequalities hold for some parameters $0 < \epsilon_1, \epsilon_2 < 1$,*

$$\sum_{\xi \in \mathscr{X}} \left|\hat{\mu}(\xi)\right|^N \geq \frac{|\mathscr{X}|^{\frac{1}{2}}}{\epsilon_1}$$

$$\sum_{\xi_1, \xi_2 \in \mathscr{X}} \left|\hat{\mu}(\xi_1 - \xi_2)\right|^N \leq (1 + \epsilon_2^2) \left(\sum_{\xi \in \mathscr{X}} \left|\hat{\mu}(\xi)\right|^N \right)^2. \tag{214}$$

Then

$$\left\| \mu^{*N} - \mathbb{U}_{\mathscr{G}} \right\|_{\mathrm{TV}(\mathscr{G})} \geq 1 - 4\epsilon_1^2 - 4\epsilon_2^2. \tag{215}$$

Proof. Define, for $\xi \in \mathscr{X}$, $w_\xi = \left(\frac{\overline{\hat{\mu}(\xi)}}{|\hat{\mu}(\xi)|} \right)^N$ or 1 if $\hat{\mu}(\xi) = 0$, and $f \in L^2(\mathscr{G})$ by

$$f(x) = \sum_{\xi \in \mathscr{X}} w_\xi e(\xi \cdot x). \tag{216}$$

Then

$$\begin{aligned} \mathbf{E}_{\mathbb{U}}[f] = 0, \qquad & \mathbf{E}_{\mathbb{U}}\left[|f|^2\right] = |\mathscr{X}|, \\ \mathbf{E}_{\mu^{*N}}[f] = \sum_{\xi \in \mathscr{X}} \left|\hat{\mu}(\xi)\right|^N, \qquad & \mathbf{E}_{\mu^{*N}}\left[|f|^2\right] = \sum_{\xi_1, \xi_2 \in \mathscr{X}} w_{\xi_1} \overline{w_{\xi_2}} \hat{\mu}(\xi_1 - \xi_2)^N. \end{aligned} \tag{217}$$

Define $A = \left\{ g \in \mathscr{G} : |f(g)| > \frac{1}{2} E_{\mu^{*N}}[f] \right\}$. By Chebyshev's inequality, $\mathbb{U}(A) \leq 4\epsilon_1^2$, while by the same inequality, $\mu^{*N}(A) \geq 1 - 4\epsilon_2^2$, from which the claim follows. □

Proof of Theorem 2, lower bound. In light of Proposition 20, choose $v \in \mathscr{C}(B_0, R_0)$ such that the frequency $\xi = \xi(v) \in \hat{\mathscr{G}}_m$ generates the spectral gap. Choose a large fixed constant $R > R_0$ and let $\{v_i\}_{i=1}^M$ be a collection of R-separated translates of v, with $M \asymp \frac{m^2}{R^2}$. The corresponding frequencies $\xi_i = \xi(v_i)$ all satisfy $|\hat{\mu}(\xi_i)| = |\hat{\mu}(\xi)| = 1 - \mathrm{gap}_m$.

Given $c > 0$, set $N = \left\lfloor (\log m - c) \mathrm{gap}_m^{-1} \right\rfloor$ and apply Lemma 27 with set of frequencies $\mathscr{X} = \{\xi_i\}_{i=1}^M$. Calculate

$$\left|\hat{\mu}(\xi_i)\right|^N = \frac{e^c}{m} \left[1 + O\left(\frac{\log m}{m^2} \right) \right]. \tag{218}$$

If c is sufficiently large, then the first condition of Lemma 27 is satisfied with $\epsilon_1 = O\left(Re^{-c}\right)$.

Write $d(v_i, v_j)$ for the ℓ^1 distance between the supports of v_i and v_j. If $d(v_i, v_j) \geq \rho$, then by Lemma 24,

$$1 - \left|\hat{\mu}(\xi_i - \xi_j)\right| = 2(1 - |\hat{\mu}(\xi)|) + O\left(\frac{\log(1 + \rho)}{\rho^2 m^2} \right) \tag{219}$$

and therefore we can compute

$$\left|\hat{\mu}(\xi_i - \xi_j)\right|^N = e^{2c} m^{-2+O(\log(1+\rho)/\rho^2)}. \tag{220}$$

Choose R large enough that when $\rho = R$ in (220), the power of m is less than -1. Then, by separating the cases $i = j$ and $i \neq j$,

$$\sum_{\substack{1\le i,j\le M \\ d(v_i,v_j)<\log m}} \left|\hat{\mu}(\xi_i - \xi_j)\right|^N = O\left(\frac{m^2}{R^2}\right) + e^{2c} O\left(\frac{m}{R^4}\right), \tag{221}$$

since the number of pairs (i, j) in the sum is $O(m^2 \log^2(m)/R^4)$. In addition, using $\rho = \log m$ in (220) and plugging in (218),

$$\begin{aligned}\sum_{\substack{1\le i,j\le M \\ d(v_i,v_j)\ge\log m}} \left|\hat{\mu}(\xi_i - \xi_j)\right|^N &= \sum_{\substack{1\le i,j\le M \\ d(v_i,v_j)\ge\log m}} e^{2c} m^{-2} \left(1 + O\left(\frac{\log\log m}{\log m}\right)\right) \\ &\le M^2 |\hat{\mu}(\xi)|^{2N} \left(1 + O\left(\frac{\log\log m}{\log m}\right)\right). \end{aligned} \tag{222}$$

Therefore, since $M^2|\hat{\mu}(\xi)|^{2N} \asymp e^{2c} m^2/R^4$,

$$\sum_{1\le i,j\le M} \left|\hat{\mu}(\xi_i - \xi_j)\right|^N \le M^2 |\hat{\mu}(\xi)|^{2N} \left(1 + O\left(\frac{\log\log m}{\log m} + \frac{R^2}{e^{2c}} + \frac{1}{m}\right)\right) \tag{223}$$

and the second condition of Lemma 27 is met with $\epsilon_2 = O\left(Re^{-c}\right)$. □

Proof of Theorem 26, lower bound. By Cauchy-Schwarz, the condition $\sum_{\xi\in\mathscr{X}} \left|\hat{\mu}(\xi)\right|^N \ge \frac{|\mathscr{X}|^{\frac{1}{2}}}{\epsilon_1}$ implies

$$\sum_{\xi\in\mathscr{X}} \left|\hat{\mu}(\xi)\right|^{2N} \ge \frac{1}{\epsilon_1^2}. \tag{224}$$

By the proof of the lower bound above, since ϵ is fixed, ϵ_1 may be taken arbitrarily small, which proves the L^2 lower bound. □

7.2. Proof of the upper bound. Recall that Proposition 17 reduces Theorem 2 to the case where the starting state is recurrent. We prove the upper bound of Theorem 26, which implies the upper bound of Theorem 2 by Cauchy-Schwarz. We consider mixing at step

$$N = \lfloor (1+\epsilon)\mathrm{gap}_m^{-1} \log m \rfloor \asymp m^2 \log m. \tag{225}$$

Let $R = R(\epsilon)$ be a parameter which is fixed as a function of m, to be determined at the end of the argument. Given frequency $\xi \in \hat{\mathscr{G}}_m$, let $v \in \mathbb{Z}_0^{\mathbb{T}_m}$ be its R-reduced prevector, and perform the clustering algorithm of Sect. 6.2 on v with parameter R.

Let $\mathscr{N}(V, K)$ denote the number of R-reduced prevectors v of L^1 mass V in K clusters.

Lemma 28. *The following upper bound holds:*

$$\mathscr{N}(V, K) \le \exp\Big(K \log(m^2) + O(V \log R)\Big). \tag{226}$$

Proof. We provide a recipe to generate all possible prevectors by adding mass one point at a time. Let Γ be a lattice path from $(0, 0)$ to (V, V), that moves either upward or rightward at each step, and that never passes above the main diagonal. Assume that Γ has exactly $K - 1$ intersection points with the main diagonal strictly between $(0, 0)$ and (V, V). Let the rightward edges go from $(i, k(i))$ to $(i + 1, k(i))$, for $0 \le i \le V - 1$. The sequence $\{k(i)\}_{i=0}^{V-1}$ is non-decreasing, with $0 \le k(i) \le i$, and there are K values of i for which $k(i) = i$ (including $i = 0$).

To generate a prevector v using the path Γ:

(1) Iterate from $i = 0$ to $V - 1$:
   (a) If $k(i) = i$, start a new cluster by adding one unit of mass to a point x_i that is separated from the set of previously placed points $\{x_j\}_{j<i}$ by a distance greater than $2R$.
   (b) If $k(i) < i$, add one unit of mass to a point x_i whose distance from the previously placed point $x_{k(i)}$ is at most $2R$. The possibility $x_i = x_{k(i)}$ is allowed.

(2) For each $x \in \operatorname{supp}(v) = \bigcup_{i=0}^{V-1}\{x_i\}$, let $w(x) = \#\{i : x_i = x\}$ be the total mass at x, and choose $v(x) \in \{-w(x), w(x)\}$.

Every prevector v with L^1 mass V in K clusters can be generated by this procedure. For each path Γ, since step (a) is taken K times, the number of possible prevectors is $O\left((m^2)^K \cdot (R^2)^V \cdot 2^V\right)$. The number of paths Γ is bounded by the V-th Catalan number, $\frac{1}{V+1}\binom{2V}{V} \le 2^{2V}$. Hence,

$$\mathscr{N}(V, K) = O\left(m^{2K} R^{2V} 8^V\right), \tag{227}$$

which has the desired form. □

Proof of Theorem 26, upper bound. In

$$\left\| P_m^N \delta_\sigma - \mathbb{U}_{\mathscr{R}_m} \right\|^2_{L^2(d\mathbb{U}_{\mathscr{R}_m})} = \sum_{0 \ne \xi \in \hat{\mathscr{G}}_m} \left|\hat{\mu}(\xi)\right|^{2N}, \tag{228}$$

write $\Xi(V, K)$ for the collection of nonzero frequencies $\xi \in \hat{\mathscr{G}}_m$ such that the R-reduced prevector of ξ has L^1 norm V in K R-clusters. Thus,

$$\left\| P_m^N \delta_\sigma - \mathbb{U}_{\mathscr{R}_m} \right\|^2_{L^2(d\mathbb{U}_{\mathscr{R}_m})} = \sum_{K \ge 1} \sum_{V \ge K} \sum_{\xi \in \Xi(V,K)} \left|\hat{\mu}(\xi)\right|^{2N}. \tag{229}$$

From the definition of R-reduction in Sect. 6.2, the bound of Lemma 18 applies also to R-reduced prevectors. Thus, there is a universal constant $c > 0$ such that every $\xi \in \Xi(V, K)$ satisfies

$$|\hat{\mu}(\xi)|^{2N} \le \exp(-cV \log m). \tag{230}$$

Let $A > 0$ be a fixed integer constant. Then,

$$\sum_{K \ge 1} \sum_{V \ge AK} \sum_{\xi \in \Xi(V,K)} \left|\hat{\mu}(\xi)\right|^{2N}$$

$$\leq \sum_{K\geq 1}\sum_{V\geq AK} \mathcal{N}(V,K)\exp\left(-cV\log m\right)$$
$$\leq \sum_{K\geq 1}\sum_{V\geq AK} \exp\Big(K\log(m^2) - V[c\log m - O(\log R)]\Big). \tag{231}$$

For sufficiently large m, the coefficient of V in the last expression is at least $\frac{c}{2}\log m$. Then, if $Ac > 4$, we sum the two geometric series:

$$\sum_{K\geq 1}\sum_{V\geq AK} \exp\Big(K\log(m^2) - V((c/2)\log m)\Big)$$
$$= \sum_{K\geq 1} \exp\left(2K\log m - AK(c/2)\log m\right)(1+o(1))$$
$$= m^{2-Ac/2}(1+o(1)), \tag{232}$$

where the $o(1)$ is as $m\to\infty$. Choose A so that $2 - Ac/2 \leq -1$.

To estimate the remaining sum over $K \leq V < AK$, let $\delta = \epsilon/3$ and set $B = A\delta^{-1}$. Choose $R = R(\epsilon)$ according to Lemma 25, so that the savings from each R-cluster of size at most B is at least $m^2\mathrm{gap}_m\,(1-\epsilon/2)$. If $\xi \in \Xi(V,K)$ with $V < AK$, then its R-reduced prevector has at least $(1-\delta)K$ clusters of size at most B. Hence,

$$1 - |\hat{\mu}(\xi)| \geq (1-\delta)K\cdot \mathrm{gap}_m\left(1-\frac{\epsilon}{2}\right) \geq \left(1-\frac{5\epsilon}{6}\right)\mathrm{gap}_m K, \tag{233}$$

and therefore

$$|\hat{\mu}(\xi)|^{2N} \leq \exp\left(\left[-2(1+\epsilon)\left(1-\frac{5\epsilon}{6}\right)\log m + O\left(m^{-2}\right)\right]K\right)$$
$$\leq \exp(-(2+\beta)(\log m)K) \tag{234}$$

for some constant $\beta = \beta(\epsilon) > 0$, as long as ϵ is sufficiently small.

We compute, for sufficiently large m,

$$\sum_{K\geq 1}\sum_{K\leq V<AK}\sum_{\xi\in\Xi(V,K)} |\hat{\mu}(\xi)|^{2N}$$
$$\leq \sum_{K\geq 1}\sum_{K\leq V<AK} \mathcal{N}(V,K)\exp(-(2+\beta)(\log m)K)$$
$$\leq \sum_{K\geq 1}\sum_{K\leq V<AK} \exp(-\beta(\log m)K + O(V\log R))$$
$$\leq \sum_{K\geq 1} \exp\Big(\left[-\beta\log m + O(A\log R)\right]K\Big)$$
$$\leq \sum_{K\geq 1} \exp(-(\beta/2)(\log m)K) = O\left(m^{-\beta/2}\right). \tag{235}$$

Thus, the entire sum (229) tends to zero like a small negative power of m, completing the proof. □

Appendix A: Local Limit Theorem

Let $\nu_{\mathbb{Z}^2}$ be the measure on $\mathbb{Z}^2$ given by

$$\nu_{\mathbb{Z}^2} := \frac{1}{4}\left(\delta_{(1,0)} + \delta_{(-1,0)} + \delta_{(0,1)} + \delta_{(0,-1)}\right), \tag{236}$$

while ν is the same measure on $\mathbb{T}_m$. Let

$$\nu_{\frac{1}{2}} = \frac{1}{2}(\delta_{(0,0)} + \nu), \qquad \nu_{\frac{1}{2},\mathbb{Z}^2} = \frac{1}{2}(\delta_{(0,0)} + \nu_{\mathbb{Z}^2}) \tag{237}$$

be the $\frac{1}{2}$-lazy versions. The Green's function may be expressed as a sum of convolutions of the lazy measures, which, unlike the non-lazy measures, do not exhibit periodicity.

Lemma 29. *The Green's functions on $\mathbb{Z}^2$ and $\mathbb{T}_m$ are given by*

$$G_{\mathbb{Z}^2}(x) = \frac{1}{8}\sum_{n=0}^{\infty}\left(\nu_{\frac{1}{2},\mathbb{Z}^2}^{*n}(x) - \nu_{\frac{1}{2},\mathbb{Z}^2}^{*n}(0,0)\right) \tag{238}$$

and

$$G_{\mathbb{T}_m}(x) = \frac{1}{8}\sum_{n=0}^{\infty}\left(\nu_{\frac{1}{2}}^{*n}(x) - \frac{1}{m^2}\right). \tag{239}$$

Proof. Recall

$$G_{\mathbb{Z}^2}(x) = \frac{1}{4}\sum_{n=0}^{\infty}\left(\nu_{\mathbb{Z}^2}^{*n}(x) - \nu_{\mathbb{Z}^2}^{*n}(0,0)\right)$$

$$G_{\mathbb{T}_m}(x) = \frac{1}{4}\sum_{n=0}^{\infty}\left(\nu^{*n}(x) - \frac{1}{m^2}\right).$$

By expanding by the binomial theorem,

$$\begin{aligned}\tilde{G}_{\mathbb{Z}^2}(x) &:= \frac{1}{8}\sum_{n=0}^{\infty}\left(\nu_{\frac{1}{2},\mathbb{Z}^2}^{*n}(x) - \nu_{\frac{1}{2},\mathbb{Z}^2}^{*n}(0,0)\right)\\ &= \frac{1}{8}\sum_{n=0}^{\infty}\frac{1}{2^n}\left(\sum_{k=0}^{n}\binom{n}{k}\left(\nu_{\mathbb{Z}^2}^{*k}(x) - \nu_{\mathbb{Z}^2}^{*k}(0,0)\right)\right).\end{aligned} \tag{240}$$

At even k, the largest value of $\nu_{\mathbb{Z}^2}^{*k}$ occurs at 0, so that by adjusting at most one term, the sum may be arranged to have a single sign. Exchange in the order of summation is thus justified by absolute convergence. This obtains

$$\tilde{G}_{\mathbb{Z}^2}(x) = \frac{1}{8}\sum_{k=0}^{\infty}\left(\nu_{\mathbb{Z}^2}^{*k}(x) - \nu_{\mathbb{Z}^2}^{*k}(0,0)\right)\sum_{n=k}^{\infty}\binom{n}{k}2^{-n}. \tag{241}$$

The inner sum has value 2 for every value of k, as follows from the generating function identity

$$\frac{1}{k!}\left(\frac{d}{dx}\right)^k\left(\frac{1}{1-x}\right) = \left(\frac{1}{1-x}\right)^k = \sum_{n=0}^{\infty}\binom{n+k}{k}x^n \tag{242}$$

which is valid in $|x| < 1$. Thus $\tilde{G}_{\mathbb{Z}^2} = G_{\mathbb{Z}^2}$.

The equality on $\mathbb{T}_m$ follows since both functions have the same discrete Fourier transform. □

Below we prove a local limit theorem for repeated convolutions of $\nu_{\frac{1}{2},\mathbb{Z}^2}$. Before doing so, we recall a Chernoff-type tail inequality.

Theorem 30 (Chernoff's Inequality). *Let X_i, $1 \le i \le n$, be mutually independent random variables with*

$$\textbf{Prob}\,(X_i = +1) = \textbf{Prob}\,(X_i = -1) = \frac{1}{2}. \tag{243}$$

Let $S_n = X_1 + \cdots + X_n$. For any $a > 0$,

$$\textbf{Prob}\,(|S_n| > a) \le 2e^{-\frac{a^2}{2n}}. \tag{244}$$

Proof. See [1], pp. 321–322. □

Chernoff's inequality implies bounds for the tail of $\nu^{*n}_{\frac{1}{2},\mathbb{Z}^2}$.

Lemma 31. *If $Y_1, Y_2, \ldots, Y_n$ are mutually independent random variables distributed according to $\nu_{\frac{1}{2},\mathbb{Z}^2}$ on $\mathbb{Z}^2$, and if $T_n = Y_1 + \cdots + Y_n$ then for any $a > 0$,*

$$\textbf{Prob}\,(\|T_n\|_2 > a) \le 4e^{-\frac{a^2}{4n}}. \tag{245}$$

Proof. If $\|T_n\|_2 > a$ then either the first or second coordinate is at least $\frac{a}{\sqrt{2}}$. The bound for either is at most $2e^{-\frac{a^2}{4n}}$, since this may be estimated by first making n choices, to move in the first direction, second direction, or hold. Conditioned on these choices, the first coordinate is the sum of $n_1 \le n$ variables of type X as in Theorem 30, and the conditional probability of their sum being larger than $\frac{a}{\sqrt{2}}$ is at most $2e^{-\frac{a^2}{4n}}$. □

Theorem 32 (Local Limit Theorem on $\mathbb{Z}^2$). *There are polynomials $\{P_k\}_{k=0}^{\infty}$ with P_k of degree at most k, such that for any $i, j \in \mathbb{Z}$ and any $a, b, N \ge 0$, we have*

$$\begin{aligned}
\delta_1^{*a} * \delta_2^{*b} * \nu^{*N}_{\frac{1}{2},\mathbb{Z}^2}(i,j) &= \exp\left(-\frac{2((i+\frac{a}{2})^2 + (j+\frac{b}{2})^2)}{N}\right) \\
&\times \left(\frac{P_a\left(\frac{i}{\sqrt{N}}\right) P_b\left(\frac{j}{\sqrt{N}}\right)}{N^{\frac{a+b+2}{2}}}\right. \\
&\left. + O\left(\frac{1}{N^{\frac{a+b}{2}+2}}\left[1 + \frac{|i|^{a+4}}{N^{\frac{a}{2}+2}} + \frac{|j|^{b+4}}{N^{\frac{b}{2}+2}} + \frac{|i|^a|j|^b(i^4+j^4)}{N^{\frac{a+b}{2}+2}}\right]\right)\right) \\
&+ O_\epsilon\left(\exp\left(-N^{\frac{1}{2}-\epsilon}\right)\right).
\end{aligned} \tag{246}$$

Moreover, P_k is an even function if k is even and an odd function if k is odd.

Proof. Apply Chernoff's inequality to reduce to $i^2 + j^2 \leq N^{\frac{3}{2}-\frac{\epsilon}{2}}$.

The quantity in question is given by

$$(2\sqrt{-1})^{a+b} \int_{(\mathbb{R}/\mathbb{Z})^2} s\left(\frac{x}{2}\right)^a s\left(\frac{y}{2}\right)^b \left(\frac{2+c(x)+c(y)}{4}\right)^N e\left(x\left(i+\frac{a}{2}\right)+y\left(j+\frac{b}{2}\right)\right) dxdy. \tag{247}$$

Truncate the integral at $\|x\|^2_{\mathbb{R}/\mathbb{Z}} + \|y\|^2_{\mathbb{R}/\mathbb{Z}} \leq N^{-\frac{1}{2}}$, since the remainder of the integral trivially satisfies the claimed bound. Now treat x, y as complex variables. Set

$$x_1 = \sqrt{\pi^2 N} x, \qquad y_1 = \sqrt{\pi^2 N} y. \tag{248}$$

Now replace

$$\begin{aligned} x_2 &:= x_1 - 2\sqrt{\frac{-1}{N}}\left(i+\frac{a}{2}\right) \\ y_2 &:= y_1 - 2\sqrt{\frac{-1}{N}}\left(j+\frac{b}{2}\right), \end{aligned} \tag{249}$$

so that

$$x = \frac{x_2 + 2\sqrt{\frac{-1}{N}}(i+\frac{a}{2})}{\pi\sqrt{N}}, \quad y = \frac{y_2 + 2\sqrt{\frac{-1}{N}}(j+\frac{b}{2})}{\pi\sqrt{N}}. \tag{250}$$

and shift the contour to $\mathrm{Im}(x_2) = \mathrm{Im}(y_2) = 0$. In doing so, an integral on $\mathrm{Re}(x_2)^2 + \mathrm{Re}(y_2)^2 = \pi^2\sqrt{N}$ is created, with $|\mathrm{Im}(x_2)| \leq \frac{2|i|+O(1)}{\sqrt{N}}$, $|\mathrm{Im}(y_2)| \leq \frac{2|j|+O(1)}{\sqrt{N}}$ on this integral. On this integral, for some $0 \leq t \leq 1$,

$$c(x) = 1 - 2\pi^2\left(\frac{\mathrm{Re}(x_2)}{\pi\sqrt{N}} + 2t\sqrt{-1}\frac{i+\frac{a}{2}}{N}\right)^2 + O\left(\|x\|^4\right) \tag{251}$$

and $\|x\|^4 = O(N^{-1})$, with a similar expression for $1-c(y)$. Since $i^2+j^2 = O\left(N^{\frac{3}{2}-\frac{\epsilon}{2}}\right)$, it follows that for some $C > 0$,

$$\mathrm{Re}\log\left(\frac{2+c(x)+c(y)}{4}\right) \leq -\frac{C}{\sqrt{N}}. \tag{252}$$

Throughout the integral $s\left(\frac{x}{2}\right)$ and $s\left(\frac{y}{2}\right)$ are bounded since the imaginary part is, and $e\left(x\left(i+\frac{a}{2}\right)+y\left(j+\frac{b}{2}\right)\right)$ is $O(1)$ since the real part of the exponential becomes negative. Thus this integral satisfies the claimed bound and may be discarded.

In the remaining integral with $\mathrm{Im}(x_2) = \mathrm{Im}(y_2) = 0$, expand

$$\begin{aligned} &\left(\frac{2+c(x)+c(y)}{4}\right)^N e\left(x\left(i+\frac{a}{2}\right)+y\left(j+\frac{b}{2}\right)\right) \\ &= \exp\left(-\frac{x_2^2+y_2^2}{2} - 2\frac{\left(i+\frac{a}{2}\right)^2+\left(j+\frac{b}{2}\right)^2}{N}\right) \end{aligned}$$

$$\times \left(1 + O\left(\frac{x_2^4 + y_2^4}{N} + \frac{1 + i^4 + j^4}{N^3}\right)\right). \tag{253}$$

Taylor expand $s\left(\frac{x}{2}\right)^a$, $s\left(\frac{y}{2}\right)^b$, to obtain

$$s\left(\frac{x}{2}\right)^a s\left(\frac{y}{2}\right)^b = (\pi x)^a (\pi y)^b \left(1 + O\left(\frac{x_2^2 + y_2^2}{N} + \frac{1 + i^2 + j^2}{N^2}\right)\right). \tag{254}$$

Now use that $x_2^2 + y_2^2 \ll 1 + x_2^4 + y_2^4$ and $\frac{i^2+j^2}{N^2} \ll \frac{1}{N} + \frac{i^4+j^4}{N^3}$ to obtain

$$\begin{aligned}
&\delta_1^{*a} * \delta_2^{*b} * \nu_{\frac{1}{2},\mathbb{Z}^2}^{*N}(i, j) + O\left(\exp(-N^{\frac{1}{2}-\epsilon})\right) \\
&\quad = \frac{(2\sqrt{-1})^{a+b}}{\pi^2 N} \exp\left(-\frac{2}{N}\left(\left(i + \frac{a}{2}\right)^2 + \left(j + \frac{b}{2}\right)^2\right)\right) \\
&\qquad \times \int_{\|x_2\|^2 + \|y_2\|^2 \le \pi^2 N^{\frac{1}{2}}} e^{-\frac{x_2^2+y_2^2}{2}} (\pi x)^a (\pi y)^b \\
&\qquad \times \left(1 + O\left(\frac{1 + x_2^4 + y_2^4}{N} + \frac{i^4 + j^4}{N^3}\right)\right) dx_2 dy_2.
\end{aligned} \tag{255}$$

Extend the integral to $\mathbb{R}^2$ with error $O\left(\exp\left(-N^{\frac{1}{2}-\epsilon}\right)\right)$. The integral

$$\pi^{a+b} \int_{\mathbb{R}^2} e^{-\frac{x_2^2+y_2^2}{2}} \left(\frac{x_2 + 2\sqrt{\frac{-1}{N}}(i + \frac{a}{2})}{\pi\sqrt{N}}\right)^a \left(\frac{y_2 + 2\sqrt{\frac{-1}{N}}(j + \frac{b}{2})}{\pi\sqrt{N}}\right)^b dx_2 dy_2$$

gives the claimed main term of the theorem. The claim regarding the parity of the polynomials P_k follows since in the main term, integration against odd powers of x or y vanishes by symmetry.

To bound the error, estimate

$$\begin{aligned}
&\ll \frac{1}{N} \exp\left(-\frac{2}{N}\left(\left(i + \frac{a}{2}\right)^2 + \left(j + \frac{b}{2}\right)^2\right)\right) \\
&\quad \times \int_{\mathbb{R}^2} e^{-\frac{x_2^2+y_2^2}{2}} |x|^a |y|^b \left[\frac{1 + x_2^4 + y_2^4}{N} + \frac{i^4 + j^4}{N^3}\right] dx_2 dy_2.
\end{aligned}$$

Now bound $|x|^a \ll \left(\frac{|x_2|}{\sqrt{N}}\right)^a + \frac{1+|i|^a}{N^a}$ and $|y|^b \ll \left(\frac{|y_2|}{\sqrt{N}}\right)^b + \frac{1+|j|^b}{N^b}$. All of the powers of $|x|$ and $|y|$ integrate to a constant. This obtains the bound

$$\begin{aligned}
&\ll \frac{1}{N} \exp\left(-\frac{2}{N}\left(\left(i + \frac{a}{2}\right)^2 + \left(j + \frac{b}{2}\right)^2\right)\right) \\
&\quad \times \left(\frac{1}{N^{\frac{a}{2}}} + \frac{|i|^a}{N^a}\right)\left(\frac{1}{N^{\frac{b}{2}}} + \frac{|j|^b}{N^b}\right)\left[\frac{1}{N} + \frac{i^4 + j^4}{N^3}\right].
\end{aligned}$$

Each term in parentheses becomes equal when $|i|, |j| = N^{\frac{1}{2}}$. Hence the error is bounded by

$$\ll \frac{1}{N^{\frac{a+b}{2}+2}} \exp\left(-\frac{2}{N}\left(\left(i+\frac{a}{2}\right)^2+\left(j+\frac{b}{2}\right)^2\right)\right) \times\left[1+\frac{|i|^{a+4}}{N^{\frac{a}{2}+2}}+\frac{|j|^{b+4}}{N^{\frac{b}{2}+2}}+\frac{|i|^a|j|^b(i^4+j^4)}{N^{\frac{a+b}{2}+2}}\right].$$

□

Proof of Lemma 7. One has

$$D_1^a D_2^b G_{\mathbb{T}_m}(i,j) = \frac{1}{8}\sum_{n=0}^{\infty}\sum_{k,\ell\in\mathbb{Z}} \delta_1^{*a} * \delta_2^{*b} * \nu_{\frac{1}{2},\mathbb{Z}^2}^{*n}(i+km, j+\ell m). \tag{256}$$

Set $R = 2 + i^2 + j^2$. We must show $D_1^a D_2^b G_{\mathbb{T}_m}(i,j) \ll R^{-\frac{a+b}{2}}$. Write

$$D_1^a D_2^b G_{\mathbb{T}_m}(i,j) = \frac{1}{8}\sum_{k,\ell\in\mathbb{Z}} \delta_1^{*a} * \delta_2^{*b} * \sum_{0\le n<R} \nu_{\frac{1}{2},\mathbb{Z}^2}^{*n}(i+km, j+\ell m) \tag{257}$$

$$+\frac{1}{8}\delta_1^{*a} * \delta_2^{*b} * \sum_{R\le n} \nu_{\frac{1}{2}}^{*n}(i,j). \tag{258}$$

In bounding (257), for a sufficiently small constant $c > 0$, the part of the sum where $n \le c\frac{R}{\log R}$ satisfies the claimed bound by applying Chernoff's inequality, see Lemma 31. In those terms with $c\frac{R}{\log R} \le n < R$, those terms with $(|k|+|\ell|)^2 m^2 \ge CR\log R$ for a sufficiently large constant C may also be shown to satisfy the claimed bound by applying Chernoff's inequality. In the remaining terms, apply the local limit theorem to obtain a bound for (257) of

$$\ll \frac{1}{R^{\frac{a+b}{2}}} + \sum_{\frac{cR}{\log R}<n<R} \times \sum_{\substack{k,\ell\in\mathbb{Z}\\ (|k|+|\ell|)^2m^2\le CR\log R}} e^{-2\left(\frac{(i+\frac{a}{2}+km)^2+(j+\frac{b}{2}+\ell m)^2}{n}\right)} \frac{\left(1+\frac{|i+km|}{n^{\frac{1}{2}}}\right)^a\left(1+\frac{|j+\ell m|}{n^{\frac{1}{2}}}\right)^b}{n^{1+\frac{a+b}{2}}}.$$

For $n < R$, $n \ll m^2$ and hence the sum over k, ℓ is bounded by a constant times the term at $(k,\ell) = (0,0)$ by comparing with a geometric series. This obtains

$$(257) \ll \frac{1}{R^{\frac{a+b}{2}}} + \sum_{\frac{cR}{\log R}<n<R} e^{-2\left(\frac{(i+\frac{a}{2})^2+(j+\frac{b}{2})^2}{n}\right)} \frac{\left(1+\frac{|i|}{n^{\frac{1}{2}}}\right)^a\left(1+\frac{|j|}{n^{\frac{1}{2}}}\right)^b}{n^{1+\frac{a+b}{2}}}$$

$$\ll \frac{1}{R^{\frac{a+b}{2}}} + \int_{\frac{cR}{\log R} < x < \infty} e^{-2\left(\frac{(i+\frac{a}{2})^2+(j+\frac{b}{2})^2}{x}\right)} \frac{\left(1+\frac{|i|}{x^{\frac{1}{2}}}\right)^a \left(1+\frac{|j|}{x^{\frac{1}{2}}}\right)^b}{x^{1+\frac{a+b}{2}}} dx$$
$$\ll \frac{1}{R^{\frac{a+b}{2}}}.$$

Expand (258) in characters of $(\mathbb{Z}/m\mathbb{Z})^2$ to obtain

$$(258) \ll \frac{1}{m^2} \sum_{(0,0)\neq(\xi,\eta)\in(\mathbb{Z}/m\mathbb{Z})^2} \left|1-e\left(\frac{\xi}{m}\right)\right|^a \left|1-e\left(\frac{\eta}{m}\right)\right|^b \frac{\left|\frac{2+c\left(\frac{\xi}{m}\right)+c\left(\frac{\eta}{m}\right)}{4}\right|^R}{1-\left(\frac{c\left(\frac{\xi}{m}\right)+c\left(\frac{\eta}{m}\right)}{2}\right)}.$$

Use representatives ξ, η for $\mathbb{Z}/m\mathbb{Z}$ satisfying $|\xi|, |\eta| \leq \frac{m}{2}$, estimate $\left|1-e\left(\frac{\xi}{m}\right)\right|^a \left|1-e\left(\frac{\eta}{m}\right)\right|^b \ll \frac{|\xi|^a|\eta|^b}{m^{a+b}}$, $1-c\left(\frac{\xi}{m}\right) \asymp \frac{\xi^2}{m^2}$ and approximate the sum with an integral to obtain, for some $c > 0$,

$$(258) \ll \frac{1}{m^2} \sum_{\substack{|\xi|,|\eta|\leq\frac{m}{2} \\ (\xi,\eta)\neq(0,0)}} \left(\frac{|\xi|}{m}\right)^a \left(\frac{|\eta|}{m}\right)^b \frac{e^{-cR\frac{\xi^2+\eta^2}{m^2}}}{\frac{\xi^2}{m^2}+\frac{\eta^2}{m^2}}$$
$$\ll \int_{(x,y)\in[-\frac{1}{2},\frac{1}{2}]^2} \frac{|x|^a|y|^b}{x^2+y^2} e^{-cR(x^2+y^2)} dxdy$$
$$\ll \int_0^\infty r^{a+b-1} e^{-cRr^2} dr$$
$$\ll \frac{1}{R^{\frac{a+b}{2}}}.$$

□

Proof of Lemma 8. Without loss of generality, let $a = 1$, $b = 0$. Recall $R = i^2 + j^2$. Let, for a large constant C_1, $T = \frac{m^2}{C_1 \log m}$ and write

$$D_1 G_{\mathbb{T}_m}(i,j) = \frac{1}{8} \sum_{k,\ell\in\mathbb{Z}} \sum_{0\leq n<T} \delta_1 * \nu_{\frac{1}{2},\mathbb{Z}^2}^{*n}(i+km, j+\ell m)$$
$$+ \frac{1}{8} \sum_{T\leq n} \delta_1 * \nu_{\frac{1}{2}}^{*n}(i,j).$$

For $n < T$, by Chernoff's inequality, if C_1 is sufficiently large then the contribution of evaluating $\nu_{\frac{1}{2},\mathbb{Z}^2}^{*n}$ at terms with $(k,\ell) \neq (0,0)$ in the first sum is $O\left(\frac{1}{m^2}\right)$. Similarly,

those terms with $n \ll \frac{R}{\log R}$ may be bounded as an error term. Use the asymptotic from the local limit theorem to write the $(k, \ell) = (0, 0)$ term as

$$\sum_{0 \le n < T} \delta_1 * \nu^{*n}_{\frac{1}{2}, \mathbb{Z}^2}(i, j) = O\left(\frac{1}{i^2 + j^2}\right) + C_2 i \sum_{\frac{R}{\log R} \ll n < T} \frac{\exp\left(-\frac{2(i^2+j^2)}{n}\right)}{n^2}$$
$$= \frac{C_3 i}{i^2 + j^2} + O\left(\frac{|i| + 1}{T}\right) + O\left(\frac{1}{i^2 + j^2}\right).$$

This gives the main term. Bound the sum over large n as before, by taking Fourier transform. This obtains the bound, for some $C_4 > 0$,

$$\frac{1}{4m^2} \sum_{(0,0) \ne (\xi, \eta) \in (\mathbb{Z}/m\mathbb{Z})^2} \left|1 - e\left(\frac{\xi}{m}\right)\right| \frac{\left|\frac{2 + c\left(\frac{\xi}{m}\right) + c\left(\frac{\eta}{m}\right)}{4}\right|^T}{1 - \frac{1}{2}\left(c\left(\frac{\xi}{m}\right) + c\left(\frac{\eta}{m}\right)\right)}$$
$$\ll \frac{1}{m} \sum_{(0,0) \ne (\xi, \eta) \in (\mathbb{Z}/m\mathbb{Z})^2} \frac{|\xi|}{\xi^2 + \eta^2} \exp\left(-\frac{C_4 T}{m^2}(\xi^2 + \eta^2)\right)$$
$$\ll \frac{1}{\sqrt{T}}.$$

The claimed error holds, since $T \gg (i^2 + j^2)^2$. □

Appendix B: Determination of Spectral Gap

In this appendix, we compute the value of

$$\gamma = \inf\left\{f(G_{\mathbb{Z}^2} * v) : v \in C^2(\mathbb{Z}^2) \backslash \mathcal{I}\right\}. \tag{259}$$

Recall from Sect. 6.1 that

$$f(\xi) = \sum_{(i,j) \in \mathbb{Z}^2} \left(1 - c\left(\xi_{(i,j)}\right)\right), \tag{260}$$

while $\mathcal{I}$ is the set of those $v \in C^2(\mathbb{Z}^2)$ for which $G_{\mathbb{Z}^2} * v$ is $\mathbb{Z}$-valued (equivalently, those $v \in C^2(\mathbb{Z}^2)$ for which $f(G_{\mathbb{Z}^2} * v) = 0$).

Given any $v \in C^2(\mathbb{Z}^2)$, reduce $\xi = G_{\mathbb{Z}^2} * v$ to $\tilde{\xi} : \mathbb{Z}^2 \to \left[-\frac{1}{2}, \frac{1}{2}\right)$ by subtracting $w : \mathbb{Z}^2 \to \mathbb{Z}$. As we observed in the paragraph containing (164), $\tilde{v} = \Delta\tilde{\xi}$ is also in $C^2(\mathbb{Z}^2)$ and satisfies $f(G_{\mathbb{Z}^2} * \tilde{v}) = f(G_{\mathbb{Z}^2} * v)$. In addition, $\tilde{v} \in \mathcal{I}$ only if $\tilde{\xi} \equiv 0$ (so also $\tilde{v} \equiv 0$). Therefore we may write

$$\gamma = \inf\left\{f(\xi) : \xi \in \{G_{\mathbb{Z}^2} * v : 0 \not\equiv v \in C^2(\mathbb{Z}^2)\} \cap \left[-\frac{1}{2}, \frac{1}{2}\right)^{\mathbb{Z}^2}\right\}. \tag{261}$$

Write d for the ℓ^1 distance on $\mathbb{Z}^2$. Given a set $S \subset \mathbb{Z}^2$, let $N = N(S)$ be its distance-1 enlargement

$$N = \{(i, j) \in \mathbb{Z}^2 : d((i, j), S) \le 1\}. \tag{262}$$

A lower bound for $f(\xi)$ is obtained as the non-linear program $P(S, v)$,

$$\begin{aligned}
&\text{minimize: } \sum_{(i,j)\in N} \left(1 - \cos\left(2\pi x_{(i,j)}\right)\right) \\
&\text{subject to: } (x_{(i,j)})_{(i,j)\in N} \in \left[0, \frac{1}{2}\right]^N, \\
&\qquad \forall\, (k,\ell) \in S,\ 4x_{(k,\ell)} + \sum_{\|(i,j)-(k,\ell)\|_1=1} x_{(i,j)} \geq |v_{(k,\ell)}|.
\end{aligned}$$

Indeed, if $\xi = G_{\mathbb{Z}^2} * v \in \left[-\frac{1}{2}, \frac{1}{2}\right)^{\mathbb{Z}^2}$, then for any $S \subset \mathbb{Z}^2$, the function $x : N(S) \to \mathbb{R}$ given by $x_{(i,j)} = |\xi_{(i,j)}|$ satisfies the constraints and so

$$f(\xi) \geq \sum_{(i,j)\in N} (1 - \cos(2\pi x_{(i,j)})) \geq P(S, v). \tag{263}$$

Lemma 33. *The program $P(S, v)$ satisfies the following properties.*

1. *If $S, T \subset \mathbb{Z}^2$ satisfy $d(S, T) \geq 3$ then $P(S \cup T, v) = P(S, v) + P(T, v)$.*
2. *If $S \subset T \subset \mathbb{Z}^2$ then $P(S, v) \leq P(T, v)$.*
3. *Denote by $P'(S, v)$ the more constrained program in which*

$$(x_{(i,j)})_{(i,j)\in N} \in \left[0, \frac{1}{4}\right]^N,$$

with the same linear constraints. This program has a unique local minimum.

Proof. The first claim holds because the variables in the neighborhoods of S and T are disjoint and do not have a common constraint. The second claim holds since for $S \subset T$, the program $P(T, v)$ is more constrained. The last claim holds since the objective function is convex on $\left[0, \frac{1}{4}\right]^N$. □

We will also use the non-linear program $Q(S, v)$,

$$\begin{aligned}
&\text{minimize: } \sum_{(i,j)\in N} \left(1 - \cos\left(2\pi x_{(i,j)}\right)\right) \\
&\text{subject to: } (x_{(i,j)})_{(i,j)\in N} \in \left[-\frac{1}{2}, \frac{1}{2}\right]^N, \\
&\qquad \forall\, (k,\ell) \in S,\ 4x_{(k,\ell)} - \sum_{\|(i,j)-(k,\ell)\|_1=1} x_{(i,j)} = v_{(k,\ell)},
\end{aligned}$$

which is also a lower bound for $f(\xi)$.

We performed the following steps to confirm that the constant γ is obtained by $\xi^* = G_{\mathbb{Z}^2} * (\delta_1 * \delta_2)$.

(1) Calculate $f(\xi^*)$ by evaluating

$$\xi^*(m, n) = \frac{1}{4} \int_{(\mathbb{R}/\mathbb{Z})^2} \frac{(e(x) - 1)(e(y) - 1)}{1 - \frac{1}{2}(c(x) + c(y))} e(mx + ny)\, dxdy \tag{264}$$

for $|m|, |n| \leq M$. It is known that $\|\xi^*\|_2^2 = \frac{1}{2\pi}$, so we estimated

$$\sum_{|m|,|n|\leq M} \left(1 - \cos\left(2\pi\xi^*_{(m,n)}\right) - 2\pi^2\left(\xi^*_{(m,n)}\right)^2\right). \tag{265}$$

By the decay of the Green's function, this determines $f(\xi^*)$ to within precision M^{-6}. We thus obtained

$$f(\xi^*) = 2.868114013(4).$$

The precision was verified by estimating

$$\sum_{\max(|m|,|n|)>M} \left|1 - c(\xi^*_{(m,n)}) - 2\pi^2\left(\xi^*_{(m,n)}\right)^2\right| \leq \frac{2\pi^4}{3} \sum_{\max(|m|,|n|)>M} \left(\xi^*_{(m,n)}\right)^4 \tag{266}$$

and

$$\sum_{\max(|m|,|n|)>M} \left(\xi^*_{(m,n)}\right)^4 \leq 4\left(\sum_{\substack{0\leq m,n \\ \max(m,n)>M}} \left(\xi^*_{(m,n)}\right)^2\right)^2, \tag{267}$$

using the symmetries $|\xi^*_{(m,n)}| = |\xi^*_{(-1-m,n)}|$, $|\xi^*_{(m,n)}| = |\xi^*_{(m,-1-n)}|$.

(2) Suppose $\xi = G_{\mathbb{Z}^2} * v \in \left[-\frac{1}{2}, \frac{1}{2}\right)^{\mathbb{Z}^2}$ satisfies $f(\xi) < 2.869$. The condition $\Delta\xi = v$ implies that $|v| \leq 3$. We ruled out prevectors with some $|v_{(i,j)}| = 3$ by considering $P(S, v)$ with $S = \{(0, 0)\}$ and $|v_{(0,0)}| = 3$.

(3) We ruled out $v_{(0,0)} = 2$ by first considering $Q(S, v)$ with $S = \{(0, 0), (1, 0)\}$ and $v_{(0,0)} = 2$, $v_{(1,0)} \in \{-2, -1, 0, 1, 2\}$. The only choices giving $Q(S, v) < 2.869$ were $v_{(1,0)} = -1$ and $v_{(1,0)} = 0$. Then we considered $Q(S, v)$ for

$$S = \{(0, 0), (-1, 0), (1, 0), (0, -1), (0, 1)\} \tag{268}$$

with $v_{(0,0)} = 2$ and $v_{(i,j)} \in \{-1, 0\}$ for all $(i, j) \in S\backslash\{(0, 0)\}$. All possibilities led to $Q(S, v) \geq 2.869$. It follows that any v with $f(G_{\mathbb{Z}^2} * v) < 2.869$ satisfies $|v| \leq 1$.

(4) Consider a set $S \subset \mathbb{Z}^2$ to be connected if $N(S)$ is connected in the usual sense. Using the increasing property, we were able to enumerate all connected S containing $(0, 0)$, and such that $P(S, 1) < 2.869$ (all had $|S| \leq 6$). This can be done iteratively starting from $S = \{(0, 0)\}$ using the increasing property of $P(S, v)$ with set inclusion. The minima in $P(S, 1)$ were rapidly calculated in each case using the SLSQP algorithm in SciPy's minimize package. The minima can be verified, for instance, by noting that at most two variables can satisfy $x_{(i,j)} > \frac{1}{4}$ and by discretizing their values – the remaining variables are then confirmed if the point is a local minimum. (This approach also verifies the minima in steps (2) and (3).)

(5) Using the addition property of $P(S \cup T, v)$ for disconnected S and T, it was verified that there is not a configuration of $v \in C^2(\mathbb{Z}^2)$ with disconnected support that meets the condition $P(\operatorname{supp} v, 1) < 2.869$. Of the remaining connected components, those which admit a configuration in $C^2(\mathbb{Z}^2)$ were estimated using Fourier inversion, as in (264). The optimum was found to be ξ^*.

References

1. Alon, N., Spencer, J.H.: The Probabilistic Method. Wiley Series in Discrete Mathematics and Optimization, 4h edn. Wiley, Hoboken, NJ (2016)
2. Athreya, S.R., Járai, A.A.: Infinite volume limit for the stationary distribution of abelian sandpile models. Commun. Math. Phys., **249**(1), 197–213 (2004). https://doi.org/10.1007/s00220-004-1080-0
3. Bak, P., Tang, C., Wiesenfeld, K.: Self-organized criticality: an explanation of the 1/f noise. Phys. Rev. Lett., **59**, 381–384 (1987) https://doi.org/10.1103/PhysRevLett.59.381
4. Bak, P., Tang, C., Wiesenfeld, K.: Self-organized criticality. Phys. Rev. A (3) **38**(1), 364–374 (1988) https://doi.org/10.1103/PhysRevA.38.364
5. Baker, M., Shokrieh, F.: Chip-firing games, potential theory on graphs, and spanning trees. J. Combin. Theory Ser. A, **120**(1), 164–182 (2013). arXiv:1107.1313 [math.CO]
6. Bhupatiraju, S., Hanson, J., Járai, A.A.: Inequalities for critical exponents in d-dimensional sandpiles. Electron. J. Probab. 22 (2017) ArXiv e-prints, arXiv:1602.06475 [math.PR]
7. Bond, B., Levine, L.: Abelian networks I. Foundations and Examples. SIAM J. Discrete Math., **30**(2), 856–874 (2016). arXiv:1309.3445 [cs.FL]
8. Cairns, H.: Some halting problems for abelian sandpiles are undecidable in dimension three. ArXiv e-prints (2015). arXiv:1508.00161 [math.CO]
9. Chung, F., Ellis, R.B.: A chip-firing game and Dirichlet eigenvalues. Discrete Math., **257**(2-3), 341–355 (2002) https://doi.org/10.1016/S0012-365X(02)00434-X
10. Dhar, D., Ruelle, P., Sen, S., Verma, D.-N.: Algebraic aspects of abelian sandpile models. J. Phys. A, **28**(4), 805–831 (1995). arXiv:cond-mat/9408020
11. Dhar, D.: Self-organized critical state of sandpile automaton models. Phys. Rev. Lett, **64**(14), 1613–1616 (1990) https://doi.org/10.1103/PhysRevLett.64.1613
12. Dhar, D.: The abelian sandpile and related models. Physica A: Statistical Mechanics and Its Applications, **263**(1), 4–25 (1999). arXiv:cond-mat/9808047
13. Diaconis, P.: Group representations in probability and statistics, volume 11 of Institute of Mathematical Statistics Lecture Notes—Monograph Series. Institute of Mathematical Statistics, Hayward, CA (1988)
14. Diaconis, P., Graham, R.L., Morrison, J.A.: Asymptotic analysis of a random walk on a hypercube with many dimensions. Random Structures Algorithms, **1**(1), 51–72 (1990) https://doi.org/10.1002/rsa.3240010105
15. Diaconis, P., Shahshahani, M.: Time to reach stationarity in the Bernoulli-Laplace diffusion model. SIAM J. Math. Anal., **18**(1), 208–218 (1987) https://doi.org/10.1137/0518016
16. Fey, A., Levine, L., Peres, Y.: Growth rates and explosions in sandpiles. J. Stat. Phys., **138**(1-3), 143–159 (2010). arXiv:0901.3805 [math.CO]
17. Fey, A., Levine, L., Wilson, D.B.: Driving sandpiles to criticality and beyond. Phys. Rev. Lett., **104**, 145703 (2010). arXiv:0912.3206 [cond-mat.stat-mech]
18. Fey, A., Meester, R., Redig, F.: Stabilizability and percolation in the infinite volume sandpile model. Ann. Probab., **37**(2), 654–675 (2009). arXiv:0710.0939 [math.PR]
19. Fey-den Boer, A., Redig, F.: Organized versus self-organized criticality in the abelian sandpile model. Markov Process. Related Fields, **11**(3), 425–442 (2005). arXiv:math-ph/0510060
20. Fukai, Y., Uchiyama, K.: Potential kernel for two-dimensional random walk. Ann. Probab., **24**(4), 1979–1992 (1996) https://doi.org/10.1214/aop/1041903213
21. Holroyd, A.E., Levine, L., Mészáros, K., Peres, Y., Propp, J., Wilson, D.B.: Chip-firing and rotor-routing on directed graphs. In: In and out of equilibrium. 2, volume 60 of Progr. Probab. Birkhäuser, Basel, pp. 331–364 (2008). arXiv:0801.3306 [math.CO]
22. Hough, B.: Mixing and cut-off in cycle walks. Electron. J. Probab. 22 (2017) ArXiv e-prints, arXiv:1512.00571 [math.NT]
23. Iwaniec, H., Kowalski, E.: Analytic number theory, volume 53 of American Mathematical Society Colloquium Publications. American Mathematical Society, Providence, RI, (2004). https://doi.org/10.1090/coll/053
24. Járai, A.A., Redig, F.: Infinite volume limit of the abelian sandpile model in dimensions $d \geq 3$. Probab. Theory Related Fields, **141**(1-2), 181–212 (2008). arXiv:math/0408060
25. Járai Antal, A., Redig, F., Saada, E.: Approaching criticality via the zero dissipation limit in the abelian avalanche model. J. Stat. Phys., **159**(6), 1369–1407 (2015). arXiv:0906.3128 [math.PR]
26. Jerison, D.C., Levine, L., Pike, J.: Mixing time and eigenvalues of the abelian sandpile Markov chain. ArXiv e-prints, (November 2015). arXiv:1511.00666 [math.PR]
27. Jones, E., Oliphant, T., Peterson, P., et al.: SciPy: Open source scientific tools for Python, (2001). [Online; Accessed 13 Feb 2017]
28. Kozma, G., Schreiber, E.: An asymptotic expansion for the discrete harmonic potential. Electron. J. Probab., **9**(1), 1–17 (2004). arXiv:math/0212156

29. Levine, L.: Threshold state and a conjecture of Poghosyan, Poghosyan, Priezzhev and Ruelle. Commun. Math. Phys., **335**(2), 1003–1017 (2015). arXiv:1402.3283 [math.PR]
30. Levine, L., Murugan, M., Peres, Y., Ugurcan Baris, E.: The divisible sandpile at critical density. Ann. Henri Poincaré, **17**(7), 1677–1711 (2016). arXiv:1501.07258 [math.PR]
31. Maes, C., Redig, F., Saada, E.: The infinite volume limit of dissipative abelian sandpiles. Commun. Math. Phys., **244**(2), 395–417 (2004) https://doi.org/10.1007/s00220-003-1000-8
32. Meester, R., Quant, C.: Connections between 'self-organised' and 'classical' criticality. Markov Process. Related Fields, **11**(2), 355–370 (2005)
33. Montgomery, H.L.: Ten lectures on the interface between analytic number theory and harmonic analysis, volume 84 of CBMS Regional Conference Series in Mathematics. Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, (1994). https://doi.org/10.1090/cbms/084
34. Pegden, W., Smart Charles, K.: Convergence of the Abelian sandpile. Duke Math. J., **162**(4), 627–642 (2013). arXiv:1105.0111 [math.AP]
35. Schmidt, K., Verbitskiy, E.: Abelian sandpiles and the harmonic model. Commun. Math. Phys., **292**(3), 721–759 (2009). arXiv:0901.3124 [math.DS]
36. Sokolov, A., Melatos, A., Kieu, T., Webster, R.: Memory on multiple time-scales in an abelian sandpile. Physica A Stat. Mech. Appl. **428**, 295–301 (2015) https://doi.org/10.1016/j.physa.2015.02.001
37. Michael Steele, J.: An introduction to the art of mathematical inequalities. In: The Cauchy-Schwarz Master Class. MAA Problem Books Series. Mathematical Association of America, Washington, DC; Cambridge University Press, Cambridge (2004). https://doi.org/10.1016/10.1017/CBO9780511817106.
38. Titchmarsh, E.C.: The Theory of the Riemann Zeta-Function. The Clarendon Press, Oxford University Press, New York, 2nd edn, (1986). Edited and with a preface by D. R. Heath-Brown

Communicated by H. Duminil-Copin