# OPINION PROPAGATION AND SANDPILES: A FEW MODELS OF AUTOMATA ON GRAPHS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Andrew Melchionna

December 2023

OPINION PROPAGATION AND SANDPILES: A FEW MODELS OF AUTOMATA ON GRAPHS

Andrew Melchionna, Ph.D.

Cornell University 2023

Each chapter of this thesis is a self-contained work, studying the long-time or many-particle limit of a different automaton.

**The sandpile identity element on an ellipse:** For certain elliptical subsets of the square lattice, the recurrent representative of the identity element of the sandpile group consists predominantly of a biperiodic pattern, along with some noise. It is shown that as the lattice spacing tends to 0, the fraction of the area taken up by the pattern in the identity element tends to 1.

**Stochastic sandpile on a cycle:** In the stochastic sandpile model on a graph, particles interact pairwise as follows: if two particles occupy the same vertex, they must each take an independent random walk step with some probability $0 < p < 1$ of not moving. These interactions continue until each site has no more than one particle on it. In this chapter, a formal coupling between the stochastic sandpile and activated random walk models is developed. This coupling is used to show that for the stochastic sandpile with n particles on the cycle graph $\mathbb{Z}_n$, the system stabilizes in $O(n^3)$ time for all initial particle configurations, provided that $p(n)$ tends to 1 sufficiently rapidly as $n \to \infty$.

**An urn model for opinion propagation on networks:** A coupled Polya's urn scheme for social dynamics on networks is considered. Agents hold continuum-valued opinions on a two-state issue and randomly converse with their neighbors on a graph, agreeing on one of the two states. The probability of agreeing on a given state is a simple function of both of agents' opinions, with higher importance given to agents who have participated in more conversations. Opinions are then updated based on the results of the conversation. I show that this system is governed by a discrete stochastic heat equation, and prove that a consensus of opinion is reached.

## BIOGRAPHICAL SKETCH

Andrew was born and raised in Massachusetts, and holds degrees in mathematics, physics, and music performance. He is excited to begin a career in quantitative finance after his time at Cornell.

## ACKNOWLEDGEMENTS

Thank you to Prof. Lionel Levine for your guidance and patience. You let me forge my own path, but were always there for me when I was stuck. I could not have asked for a better advisor.

Thank you to all of my professors, mentors and colleagues for teaching and inspiring me. While there are far too many people who deserve to be recognized, I'd like to thank some people who have humored me with their listening ear and insightful comments while I prepared for presentations and job interviews: Peter Vang Uttenthal and Kimoi Kemboi (from our lavish window-office), Kevin Miao and David Papale (even long after high school physics class!), Carrie Wang, Chris Wang, Irina Rasolonjanahary, Akane Wakai, Swee Hong Chan, Lila Greco, Ryan Webler, Peter Cody Fiduccia, Brenna Flatley, Profs. Tim Healey and Laurent Saloff-Coste, and many others.

Thank you to all of my friends for your support and companionship. In particular, thank you to my singing groups at Cornell- The Hangovers and Glee Club- for sharing music with me and for your lifelong friendship.

Finally, thank you to all of my family, especially my parents Donna and Emilio, my sister Alison and my brother Mark. Words can't really describe my gratitude, but maybe the equations below will (just kidding). Love you guys.

## TABLE OF CONTENTS

CHAPTER 1

# INTRODUCTION

This thesis includes a chapter for each of three independent works: *the sandpile identity element on an ellipse, the stochastic sandpile on a cycle,* and *an urn model for opinion propagation on networks*. In each work, a different model for interacting particles or agents on graphs is studied, with the goal of characterizing the long-time or many-particle limit of the system. A central theme of this thesis is to consider the question: "how do systems governed by simple local interactions exhibit interesting global behavior?"

Chapters 2 and 3 are devoted to *sandpile models* (see e.g. [10, 11, 15]), which feature interacting particles occupying the vertices of a graph and moving (deterministically or randomly) along the edges of the graph until they are sufficiently well-dispersed. First introduced by Bak, Tang and Wiesenfield in the 1980's [9], sandpiles serve as models of self-organized criticality [13], which is the tendency of certain physical systems to gravitate towards critical states without the tuning of any parameters. Such systems will naturally *stabilize* (in a sense made rigorous below), but will gravitate towards states that are on the boundary of instability, in the sense that microscopic changes (e.g. the addition or displacement of one particle) cause macroscopic changes to the system. An important signature of criticality found in these models is that for sandpiles on infinite lattices, avalanche sizes follow a power-law distribution [14], where an avalanche is the propagation of waves of high particle density through regions of low particle density.

In Chapter 2, the many-particle limit of the *abelian sandpile model* on the square, two-dimensional lattice is studied. The abelian sandpile model prescribes simple rules for diffusing regions of high particle concentration, and is called abelian because the final state of the system is independent of the order in which different regions of high density are diffused. In the many-particle limit, this system is governed by the so-called 'sandpile PDE' ([7]), which inherently features a fractal geometry catalogued by Levine, Pegden, and Smart ([4, 5]). Associated to the solutions of the sandpile PDE are certain periodic patterns in the sandpile, shown to be stable in [6]. In Chapter 2 of this thesis I use the machinery developed in [6] to assert that, under appropriate boundary conditions on the lattice, these periodic patterns can be isolated.

Chapter 3 features the *stochastic sandpile model*, in which particles perform random walks away from regions of high density. The model is studied on the cycle graph $\mathbb{Z}_n$, and in this chapter I prove that with appropriate parameter-tuning, the system stabilizes quickly, even when the graph is saturated with particles. This result follows from a coupling of the stochastic sandpile model with the closely-related *activated random walk* model, and a result due to Cairns, Ganguly and Levine pertaining to the fast stabilization of the activated random walk model ([17]).

In Chapter 4, a new model for social dynamics is presented. This framework is inspired by the Pólya's Urn problem, which can be stated simply as follows: begin with an urn that contains one red ball and one blue ball. Randomly draw one of the balls from the urn. Put this ball back, and add to the urn a new ball of the same color just drawn. Continue this process indefinitely. In this chapter I develop a model for opinion propagation in which agents living on the vertices of a graph converse with their neighbors, agreeing on one side of a two-sided issue (say, red or blue). The side on which they agree is a random function of both agents' histories: both players hold their own urn, with numbers of red and blue balls representing their propensity for the red and blue sides of the issue. When two players converse, they pool their urns, and draw a ball randomly. Each player's urn is then reinforced with balls of the color they've just agreed on. I show that the behavior of the system is governed by the stochastic heat equation, and prove consensus, i.e. that the fraction of balls in each player's urn that are red all converge to the same (random) constant.

CHAPTER 2

# THE SANDPILE IDENTITY ELEMENT ON AN ELLIPSE

## 2.1 Introduction

The abelian sandpile model is governed only by simple local interaction rules, yet demonstrates interesting and well-synchronized behavior on the large scale. The model is as follows. Begin with a function $\sigma : \mathbb{Z}^2 \to \mathbb{Z}_{\geq 0}$, representing the number of grains of sand on the individual vertices in $\mathbb{Z}^2$. $\sigma$ will be referred to as the *initial configuration.* If a site $x \in \mathbb{Z}^2$ has $\sigma(x) \geq 4$, the site is deemed unstable, and must be 'toppled', in the following manner. Remove 4 grains of sand from the unstable site, and donate them, one each, to the site's 4 nearest neighbors on $\mathbb{Z}^2$. Continue adjusting the sandpile in this manner, performing these toppling moves at unstable sites until every site is stable, i.e. has fewer than 4 grains of sand. Toppling moves occur at successive discrete time steps.

Consider the following sandpile process. Given a finite subset $E \subset \mathbb{Z}^2$, recall that the outer boundary $\partial E$ is equal to the set of vertices $x$ in $\mathbb{Z}^2 \backslash E$ such that $x$ is adjacent to $y$ for some $y$ in $E$. Initialize the sandpile with 0 grains of sand everywhere on $E$. Allow the outer boundary to be an infinite source of sand, in the following sense. For any fixed positive integer $n$, topple the outer boundary $n$ times, allowing the sites in $E$ that are adjacent to the boundary to accumulate sand from these toppling moves, and disregarding the sand that accumulates at any points not lying in $E$. Once the outer boundary has toppled $n$ times, a (possibly unstable) configuration has formed in $E$; in particular, the neighbors of the outer boundary that lie in $E$ have accumulated some sand, while the rest of the sites in $E$ still do not have any grains of sand on them (and any sand outside of $E$ is ignored). Now proceed to stabilize the sandpile on $E$ in the usual way, ignoring any grains of sand that leave $E$. After the $n$ topplings of the outer boundary, sites in $E^c$ (the complement of $E$) are never toppled again. One can think of $E$ as a tabletop; when a grain of sand falls off of $E$, it is lost forever. Note that the above discussion of stabilizing an unstable sandpile relies on the so-called 'abelian property' of the model, which states that any reasonable sequence of stabilizing toppling moves ultimately leads to the same unique stable configuration. This construction is made precise in Section 2.2.

It is a fact [1] that, given $E \subset \mathbb{Z}^2$, there exists a positive integer $N$ such that the stabilized sandpile resulting from the above process above will be identical for all values of $n \geq N$. This stable sandpile plays a crucial role, which is now illustrated. Consider the above process, in which one first topples the boundary $N$ times, and then stabilizes the interior of $E$. Call the resulting sandpile configuration $e$. If $e$ is now used as an initial configuration and the boundary is toppled $k$ (with $k$ an arbitrary positive integer) more times, and then stabilize the interior of $E$ again, then every site in $E$ will topple exactly $k$ times, and the resulting stable sandpile will again be the configuration $e$. More generally, let $r : E \to \mathbb{Z}_{\geq 0}$ be any initial sandpile configuration. Topple the boundary $k \geq 0$ times and perform the subsequent stabilization. If every site in $E$ topples $k$ times in the stabilization and the resulting stable sandpile is again $r$ (and this holds for all $k \geq 0$), then the configuration $r$ is called *recurrent*. The set of recurrent sandpiles on $E$ form a group under the operation of vertex-wise addition followed by stabilization. The stable sandpile $e$ is the identity element of this group, with the property that for any recurrent sandpile $r$, $\mathcal{S}(e + r) = r$, where $+$ denotes vertex-wise addition, and $\mathcal{S}$ denotes the stabilization. The identity element can be shown to satisfy a certain discrete boundary value problem, described in Section 2.3.

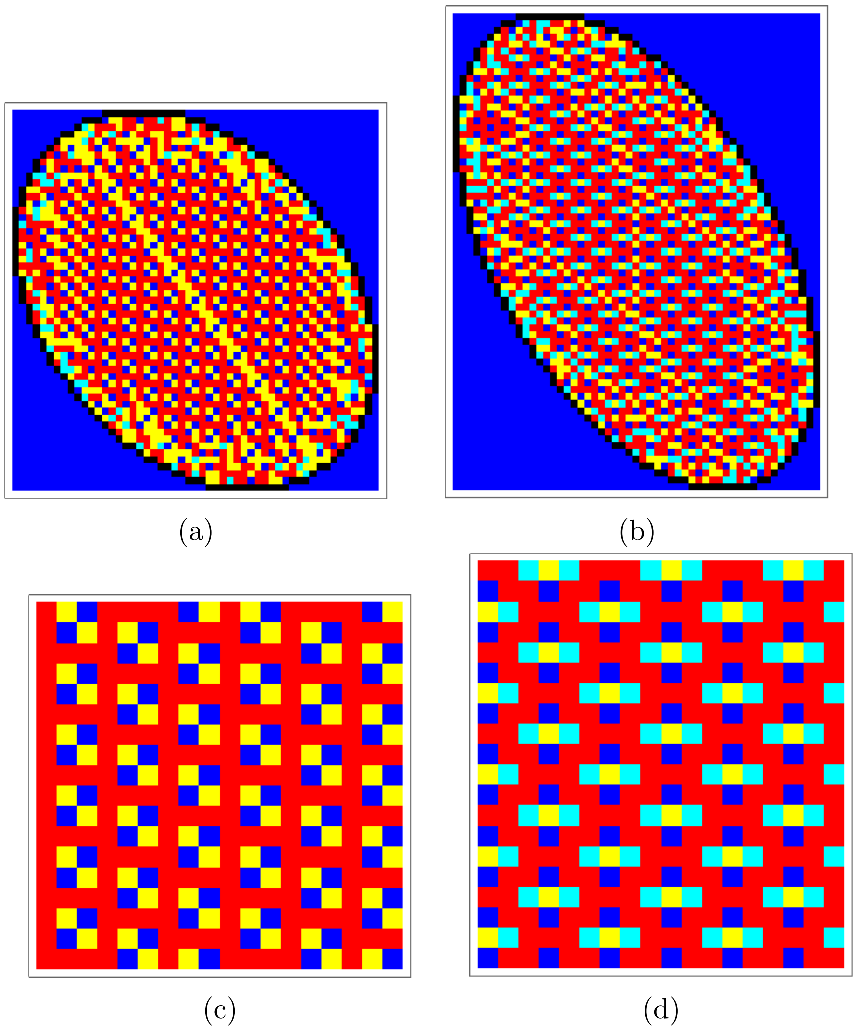



Figure 2.1: Subfigures a) and b) correspond to the identity elements for the ellipses $E_{A,k}$, with $A = \begin{pmatrix} \frac{10}{9} & \frac{1}{3} \\ \frac{1}{3} & 1 \end{pmatrix}$ and $\begin{pmatrix} \frac{4}{3} & \frac{1}{2} \\ \frac{1}{2} & \frac{3}{4} \end{pmatrix}$, respectively, with $k = 18^2$ in both. Subfigures c) and d) show the patterns $p_A$ corresponding to the ellipses in a) and b) respectively. A red vertex represents 3 grains of sand; yellow, 2; light blue, 1; navy, 0.

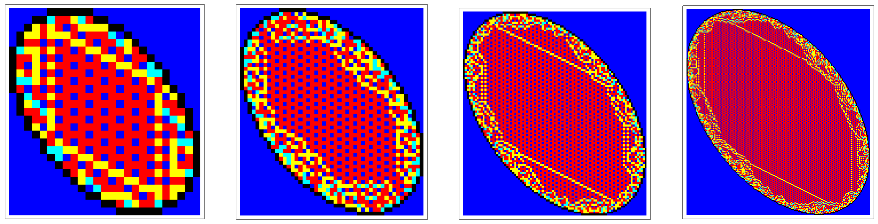
The goal of this chapter is to characterize the identity elements of various elliptical subsets of $\mathbb{Z}^2$. That is, for a certain set of $2 \times 2$ symmetric matrices $\Gamma^+ \subset \mathrm{Sym}(\mathbb{R}_{2\times 2})$ (defined in Section 2.2.3) this chapter considers the identity element on the graph

$$E_{A,k} = \{x \in \mathbb{Z}^2 : \frac{1}{2}x^T Ax < k\},$$

where $A \in \Gamma^+$ and $k > 0$ is a real number. In this chapter, I show that the identity element predominantly features a biperiodic pattern $p_A$ associated to the matrix $A$ (this association is established in Proposition 6). Figure 2.1 shows the identity elements for various $E_{A,k}$, along with the associated patterns $p_A$. Note that the identity elements also have some noise near the boundary of the ellipse and some one-dimensional defects in the interior of the ellipse. The main result of this chapter is that for $A \in \Gamma^+$, the fraction of the area inside the ellipse $E_{A,k}$ that conforms to this periodic pattern tends to 1 in the limit of $k \to \infty$ (see Figure 2.2). The proof adapts the estimates developed by Pegden and Smart in [6]. While [6] characterizes a sandpile on a square subset of $\mathbb{Z}^2$ in which many different periodic patterns appear, in this chapter I consider a sandpile on an elliptical geometry that isolates a single pattern. The main ingredient that is specific to the elliptical geometry is Lemma 2.4.2. Following [6], one can define a point $x \in E_{A,k}$ to be $r$-good if the sandpile identity element of $E_{A,k}$ matches some translation of the pattern $p_A$ in the set $B_r(x) \cap \mathbb{Z}^2$ (see Figure 2.7).

Below is a weak version of our Theorem 2.4.1, which is the main result of this chapter.

**Theorem.** *Fix $A \in \Gamma^+$. Fix $r$ to be a sufficiently large constant, dependent on $A$ (but not on $k$). Of the points that are distance at least $r$ from the boundary of the ellipse given by $A$ in $\mathbb{R}^2$, let $f(k, A)$ be the fraction of these points that are NOT*



(a) $k = 8^2$ (b) $k = 16^2$ (c) $k = 32^2$ (d) $k = 64^2$

Figure 2.2: The identity elements for the ellipses given by $E_{A,k}$, with $A = \begin{pmatrix} \frac{5}{4} & \frac{1}{2} \\ \frac{1}{2} & 1 \end{pmatrix}$ and various $k$. A red vertex represents 3 grains of sand; yellow, 2; light blue, 1; navy, 0.

*r-good in the identity element of $E_{A,k}$. Then*

$$\limsup_{k \to \infty} f(k, A) \cdot k^{1/4} \leq C_A,$$

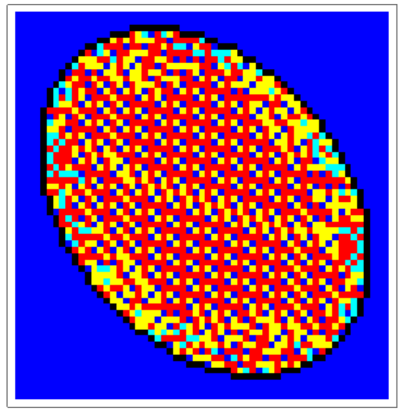
*where $C_A$ is a constant depending on the matrix $A$. In particular,*

$$\lim_{k \to \infty} f(k, A) = 0.$$

The remainder of the chapter is structured as follows: Section 2.2 lays out the preliminaries of the abelian sandpile model and the Apollonian structure of the growth rates attainable by odometer functions. In Section 2.3, the discrete boundary value problem that the stable sandpile solves is introduced. Section 2.4 features the proof of the main result and a discussion of an extension of the theorem to the case of an uncentered ellipse. That is, it is shown that this theorem still holds (but now with different constant $C_A$) for the identity element of $E^p_{A,k}$, with this graph given by

$$E^p_{A,k} = \{x \in \mathbb{Z}^2 | \frac{1}{2}(x - p)^T A(x - p) < k\}$$

for $p \in \mathbb{R}^2$. Figure 2.3 shows the identity element for the graph $E^p_{A,k}$, with $p = (.47, .5)$, and with the same $A$ and $k$ values as found in Figure 2.1a.



(a)

Figure 2.3: The identity element for the graph $E^p_{A,k}$, with $A = \begin{pmatrix} \frac{10}{9} & \frac{1}{3} \\ \frac{1}{3} & 1 \end{pmatrix}$, $k = 18^2$, and $p = (.47, .5)$.

## 2.2 Preliminaries

### 2.2.1 The sandpile on $\mathbb{Z}^2$

Given an initial configuration $\sigma_0 : \mathbb{Z}^2 \to \mathbb{Z}_{\geq 0}$, let the (finite or infinite) sequence $(x_1, x_2, ...)$ with $x_i \in \mathbb{Z}^2$ represent a sequence of toppling moves, with the vertex $x_i$ being toppled in the ith timestep. Demand that all vertices in $\mathbb{Z}^2$ are toppled finitely many times. Let the *odometer function* corresponding to the sequence $(x_1, x_2, ...)$ be a function $u : \mathbb{Z}^2 \to \mathbb{Z}_{\geq 0}$ that counts the number of times each site in $\mathbb{Z}^2$ appears in the sequence. Note that, since $u(x)$ is finite for all $x \in \mathbb{Z}^2$, one have that the toppling sequence gives a well-defined *final configuration*, given by $\sigma_0(x) + \Delta u(x)$, where $\Delta$ is the discrete Laplacian operator, acting on functions with domain $\mathbb{Z}^2$. It is given by

$$\Delta w(x) = \sum_{y \sim x} \big(w(y) - w(x)\big) = -4w(x) + \sum_{y \sim x} w(y).$$

The toppling sequence is called *legal* if a toppling move is only made when a vertex has 4 or more chips, and is called *stabilizing* if the resulting final configuration has fewer than 4 chips at every site on $\mathbb{Z}^2$. A configuration $\sigma$ is called *stabilizable* if it admits a stabilizing toppling sequence. A foundational result, justifying the the word 'abelian' in the name of the model, is that any two legal, stabilizing configurations give the same odometer function, and thus the same final sandpile:

**Proposition 1** (Abelian Property)**.** *[2] Let $\sigma$ be an initial configuration, and suppose that there exists a stabilizing sequence $(x_1, ..., x_n)$. Then there exists a legal stabilizing sequence, and any two legal stabilizing seqences are permutations of each other.*

With the abelian property in mind, it is then natural to define an odometer function $u : \mathbb{Z}^2 \to \mathbb{Z}_{\geq 0}$ corresponding to a stabilizable configuration $\sigma$, letting $u(x)$ represent the number of times a site x topples in a legal, stabilizing sequence. Given a stabilizable initial configuration $\sigma$, one can then write the final, stable sandpile configuration $s : \mathbb{Z}^2 \to \mathbb{Z}_{\geq 0}$ in the following way:

$$s(x) = \sigma(x) + \Delta u(x).$$

By the abelian property, the odometer function $u$ corresponding to $\sigma$ is well-defined. The odometer function can be shown to satisfy the following 'least action principle':

**Proposition 2** (Least Action Principle)**.** *[2] Let $u(x)$ be the odometer function for a stabilizable initial configuration $\sigma$. Then $u(x)$ satisfies*

$$u(x) = \inf\{w(x) \,|\, w : \mathbb{Z}^2 \to \mathbb{Z}_{\geq 0}, \, \sigma(y) + \Delta w(y) \leq 3 \quad \forall y \in \mathbb{Z}^2\}.$$

This proposition states that, during a stabilizing process, each vertex will topple as few times as necessary in order to stabilize the sandpile.

### 2.2.2 The sandpile on a general graph

Consider now a finite, connected, undirected multigraph $G = (V \cup \{q\}, E)$, where $q$ is a *sink vertex*. A sandpile configuration $s : V \to \mathbb{Z}$ is called *stable* if $s(v) < \deg(v)$ for all $v$ in $V$. Define the graph Laplacian $\Delta_G$ similarly as above: for any integer-valued function on the vertices $w : V \cup \{q\} \to \mathbb{Z}$, define the graph Laplacian $\Delta_G$ as

$$\Delta_G w(x) = \sum_{y \sim x} \big(w(y) - w(x)\big),$$

where $y \sim x \Leftrightarrow xy \in E$. By enumerating the vertices in $V \cup \{q\}$, one can think of a function on the graph as a vector (with components corresponding to the value of the function on each vertex), and the graph Laplacian $\Delta_G$ as a matrix acting on these vectors. This matrix can be written as

$$\big(\Delta_G\big)_{ij} = \begin{cases} -\deg i & i = j \\ M(i,j) & i \neq j \end{cases}$$

where $M(i,j)$ is the multiplicity of the edge connecting $i$ and $j$, with $M(i,j) = 0$ if $ij \notin E$.

With this construction in mind, define the reduced graph Laplacian $\tilde{\Delta}_G$ as the matrix obtained from $\Delta_G$ by deleting the row and column corresponding to the sink vertex $q$.

The set of sandpile configurations on $G$ enjoy a group structure, in the following sense. Consider the free abelian group $\mathbb{Z}^V$, corresponding to the set of all possible sandpile configurations (allowing for negative amounts of chips) on the nonsink vertices, with vertex-wise addition. Consider also the equivalence relation on $\mathbb{Z}^V$ given by $w \sim u$ if and only if there exists a function $v : V \to \mathbb{Z}$ such that $w = u + \tilde{\Delta}_G v$. $v(x)$ can be seen as the number of times that the vertex $x$ needs to be toppled in order to get from the configuration $u$ to the configuration $w$. Note that $v$ is allowed to take on negative values, corresponding to 'untopplings'. It can be shown that $\sim$ is an equivalence relation on $\mathbb{Z}^V$. It can also be seen that this equivalence relation respects vertex-wise addition of two sandpiles. Thus, the quotient $\mathbb{Z}^V/\sim$ with the operation of vertex-wise addition is a group, called the *sandpile group* of the graph $G$.

Next the notion of recurrency of a sandpile configuration is defined.

**Definition 1.** *Let $s : V \to \mathbb{Z}$ be a sandpile configuration. Given a nonempty subset $X \subset V$, and an element $x \in X$, define* ***indeg**$_X(x)$* *to be the number of nearest neighbors of $x$ that lie in $X$. A* ***forbidden subconfiguration*** *is a nonempty subset $X \subset V$ such that, for all $x \in X$, $s(x) < \mathit{indeg}_X(x)$. The sandpile $s$ is called* ***recurrent*** *if it possesses no forbidden subconfigurations.*

The name 'recurrent' comes from the study of Markov chains on the space of sandpile configurations. In this framework, these configurations are recurrent in the sense that $\mathbb{P}^s(\#\{n \in \mathbb{N} : s_n = s\} = \infty) = 1$, where $(s_n)_{n\in\mathbb{N}}$ is a Markov chain on the space of stable sandpiles on $G$ that evolves by dropping a grain of sand uniformly at random on a nonsink vertex and stabilizing. $\mathbb{P}^s$ is the measure given by starting the Markov chain at $s_0 = s$. The following standard fact linking recurrent configurations and the sandpile group is used:

**Proposition 3.** *[3] Every equivalence class in the sandpile group contains exactly one recurrent configuration.*

The goal of this chapter is to explore the recurrent representative of the identity element of the graph $E_{A,k}$, for certain matrices $A$ discussed in the next section.

### 2.2.3 Integer superharmonic matrices

Let $\mathrm{Sym}(\mathbb{R}_{2\times 2})$ be the set of symmetric $2 \times 2$ matrices with real entries. Let $q_A : \mathbb{Z}^2 \to \mathbb{Z}$ be defined by $q_A(x) = \frac{1}{2}x^T A x$, where in the previous equation, $x \in \mathbb{Z}^2$ is considered as a 2-component vector. Define the set of integer superharmonic matrices, $\Gamma$, as follows:

$$\Gamma = \{A \in \mathrm{Sym}_{2\times 2}(\mathbb{R}) \,|\, \exists o_A : \mathbb{Z}^2 \to \mathbb{Z},\, o_A(x) \geq q_A(x) + o(|x|^2),$$
$$\Delta o_A(x) \leq 3 \;\forall x \in \mathbb{Z}^2\}.$$

In words, a real, symmetric $2 \times 2$ matrix is integer superharmonic if there exists an integer-valued dominating (up to terms $o(|x|^2)$) function that satisfies $\Delta o_A(x) \leq 3$ for all $x$. Such a function will be called an 'integer superharmonic witness corresponding to A'. Note that the definition of a superharmonic function $o_A : \mathbb{Z}^2 \to \mathbb{Z}$ is usually that $\Delta o_A(x) \leq 0$ for all $x \in \mathbb{Z}^2$, while this chapter uses the convention that $o_A$ is required to satisfy $\Delta o_A(x) \leq 3$ for all $x \in \mathbb{Z}^2$. One can easily translate between these two conventions by simply adding a function such as $\frac{3}{2}x_1(x_1+1)$, which has Laplacian identically equal to 3.

When considering a subset $X \subset \mathbb{Z}^2$, the outer and inner boundaries $\partial X$ and $\overline{\partial} X$, respectively, will be defined as

**Definition 2.** *Given a subset $X \subset \mathbb{Z}^2$, define*

$$\partial X = \{y \in \mathbb{Z}^2 - X \,|\, y \sim x \text{ for some } x \in X\}.$$

**Definition 3.** *Given a subset $X \subset \mathbb{Z}^2$, define*

$$\overline{\partial} X = \{y \in X \,|\, y \sim x \text{ for some } x \in \mathbb{Z}^2 - X\}.$$

**Definition 4.** *Given a subset $X \subset \mathbb{Z}^2$, define the complement of $X$ as*

$$X^c = \{y \in \mathbb{Z}^2 \,|\, y \notin X\}.$$

Now define a subset $\Gamma^+ \subset \Gamma$ as follows:

**Definition 5.** $\Gamma^+ = \{A \in \Gamma \,|\, \exists \epsilon > 0 \;\; s.t. \;\; A - \epsilon I \leq B \in \Gamma \implies B \leq A\}$.

The matrices $A \in \Gamma^+$ are in some sense maximal, which is now discussed.

The set $\Gamma$ shares a relationship to an Apollonian circle packing in the plane, which is now discussed (following the authors of [5]). Consider the set of lines $\{x = 2k\}$ for $k \in \mathbb{Z}$, along with the set of circles $C_k$ of radius 1, centered at $(2k+1, 0)$ for $k \in \mathbb{Z}$. For any three pairwise tangent *general circles* (that is, circles or lines), there are exactly two *Soddy general circles* that are tangent to all three (we consider two lines to be tangent if and only if they are adjacent, i.e. are given by $\{x = 2k\}$ and $\{x = 2k + 2\}$ for some $k$). The *Apollonian circle packing* generated by the lines $\cup_{k\in\mathbb{Z}}\{x = 2k\}$ and the circles $\cup_{k\in\mathbb{Z}}C_k$ is the minimal set of general circles that contains the generators and is closed under the addition of Soddy general circles for any pairwise-tangent triple in the packing (see Figure 2.4).

Let this Apollonian circle packing take place in the $\{z = 2\}$ plane in $\mathbb{R}^3$. Over each circle in the Apollonian circle packing, one can consider the cone protruding out of the plane (in the direction of positive-z) with slope 1, so that each circle is the base of a cone of height equal to the circle's radius. Define $\Theta \subset \mathbb{R}^3$ to be the downset of these cones, that is, the set of points $(x, y, z) \in \mathbb{R}^3$ such that $(x, y, z+\ell)$ is on a cone for some $\ell \in \mathbb{R}_{\geq 0}$ (see Figure 2.5).

The authors of [5] prove that $\Gamma = \Theta$, under the following coordinate identifica-

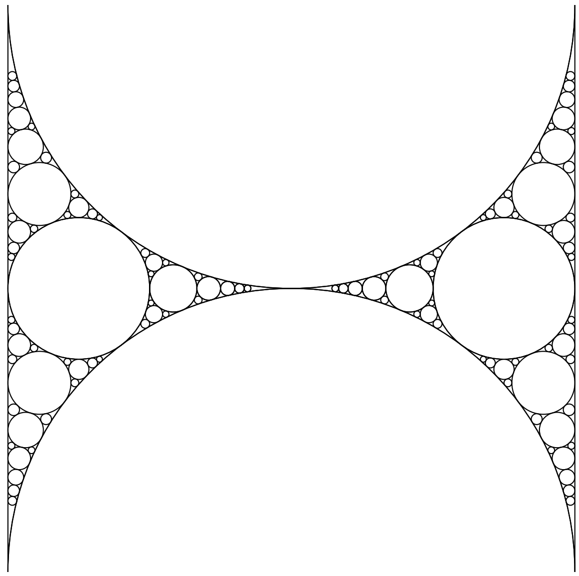


Figure 2.4: [5] A portion of the Apollonian circle packing between the lines $\{x = 0\}$ and $\{x = 2\}$

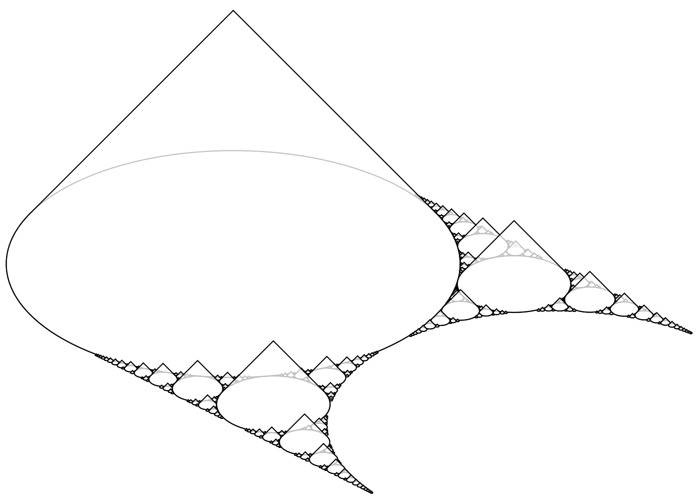


Figure 2.5: [5] A portion of the boundary of the set $\Theta$

tion on $\Gamma$:

$$\text{given } A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}, \text{ set } \begin{cases} x = a - c \\ y = 2b \\ z = a + c \end{cases}.$$

In other words,

**Proposition 4.** *[5] $A \in Sym(\mathbb{R})_{2\times 2}$ is integer superharmonic if and only if*

$(x(A), y(A), z(A)) \in \Theta$, *with the coordinates* $\big(x(A), y(A), z(A)\big)$ *defined above.*

The matrices $A \in \Gamma^+$ correspond exactly to the peaks of the cones in $\Theta$. A special property of matrices in $\Gamma^+$ is that they possess an integer superharmonic witness that is recurrent, in the following sense:

**Definition 6.** *A function* $v : \mathbb{Z}^2 \to \mathbb{Z}$ *is recurrent in* $X \subset \mathbb{Z}^2$ *if* $\Delta v \leq 3$ *in* $X$ *and*

$$\sup_Y (v - w) \leq \sup_{(X-Y)\cup\partial X} (v - w)$$

*whenever* $w : \mathbb{Z}^2 \to \mathbb{Z}$ *satisfies* $\Delta w \leq 3$ *in a finite* $Y \subset X$.

More precisely, the integer superharmonic witnesses for members of $\Gamma^+$ are recurrent on all of $\mathbb{Z}^2$. Below is a proposition relating the definition of a recurrent sandpile with that of a recurrent function.

**Proposition 5.** *A function* $v : \mathbb{Z}^2 \to \mathbb{Z}$ *is recurrent on a finite subset* $X \subset \mathbb{Z}^2$ *if and only if* $\Delta v|_X$ *is a recurrent sandpile on* $X$.

*Proof.* **v recurrent function** $\implies$ **$\Delta$v recurrent sandpile**

Assume that $\Delta v$ has an FSC $Y \subset X$, so that for all $x$ in $Y$, $\Delta v(x) < \text{indeg}_Y(x)$. Equivalently, $\Delta v(x) + \text{outdeg}_Y(x) < 4$, where $\text{outdeg}_Y(x) = 4 - \text{indeg}_Y(x)$. Since $\text{outdeg}_Y(x) = \Delta \mathbb{1}_{(X-Y)\cup\partial X}(x)$, one has that $w := v + \mathbb{1}_{(X-Y)\cup\partial X}$ satisfies $\Delta w \leq 3$ on $Y$ and

$$\sup_Y (v - w) = 0 > -1 = \sup_{(X-Y)\cup\partial X} (v - w),$$

violating the recurrency of $v$, and yielding a contradiction.

**$\Delta$v recurrent sandpile** $\implies$ **v recurrent function** Consider a function $w : X \cup \partial X \to \mathbb{Z}$ satisfying $\Delta w \leq 3$ in some finite $Y \subset X$. Define $y_0$ to be a point at which $v - w$ is maximized in Y. Consider $\tilde{w}(x) = w(x) - w(y_0) + v(y_0)$, so that $(v - \tilde{w})(y_0) = 0$, $\sup_Y(v - \tilde{w}) = 0$, and $\Delta\tilde{w} = \Delta w$. Define the set $Z \subset Y$ to be the maximal connected set containing $y_0$ on which $\tilde{w} - v = 0$. Define $\tilde{w}\backslash v(x) := \max\big((\tilde{w} - v)(x), 0\big)$. Clearly $\tilde{w}\backslash v$ is identically 0 on Z.

Now, if $(\tilde{w}\backslash v)(x) \geq 1$ for all $x \in \partial Z$, then one would have that

$$3 \geq \Delta\tilde{w}(z) \geq \Delta v(z) + \operatorname{outdeg}_Z(z) \qquad \forall z \in Z,$$

where the first inequality follows from our assumptions on $w$, and the second inequality follows from the fact that $\Delta(w - v)(z) \geq \operatorname{outdeg}(z)$, since $(\tilde{w}\backslash v)(x) \geq 1$ for all $x \in \partial Z$. But this says exactly that $Z$ is a forbidden subconfiguration of X for $\Delta v$, a contradiction our assumption that $\Delta v$ is a recurrent sandpile. Thus, there must be some $x \in \partial Z$ such that $(\tilde{w}\backslash v)(x) = 0$. By the definition of $Z$, $x \notin Y$, and thus $\sup_{(X-Y)\cup\partial X}(v - \tilde{w}) \geq 0 = \sup_Y(v - \tilde{w})$, giving that $v$ is a recurrent function on $X$.

□

In [5], the authors present an explicit recurrent integer superharmonic witness for all $A \in \Gamma^+$. A particularly interesting feature of such a witness is that the corresponding Laplacian is doubly periodic, and this period gives a hexagonal tiling of $\mathbb{Z}^2$, in the following way.

**Proposition 6.** *[5] For every matrix $A \in \Gamma^+$, there is a function $o_A : \mathbb{Z}^2 \to \mathbb{Z}$, a matrix $V \in \mathbb{Z}^{2\times 3}$, and a subset $T \subset \mathbb{Z}^2$ such that the following hold:*

*1) $o_A$ is recurrent and can be decomposed as $o_A(x) = q_A(x) + L(x) + \rho_A(x) + c$, where $q_A(x) = \frac{1}{2}x^T Ax$, $L(x) = b \cdot x$ for some $b \in \mathbb{R}^2$, $c \in \mathbb{R}$, and $\rho_A : \mathbb{Z}^2 \to \mathbb{Z}$ is a $V\mathbb{Z}^3$ periodic function*

*2) If $x \sim y \in \mathbb{Z}^2$, then there is $z \in \mathbb{Z}^3$ such that $x, y \in T + Vz$.*

*3) Let $z, w \in \mathbb{Z}^3$. $(T + Vz) \cap (T + Vw) \neq \emptyset$ if and only if $|z - w|_1 \leq 1$. Furthermore, $(T + Vz) \cap (T + Vw) \subset (\overline{\partial} T + Vz) \cap (\overline{\partial} T + Vw)$ for any $z \neq w$.*

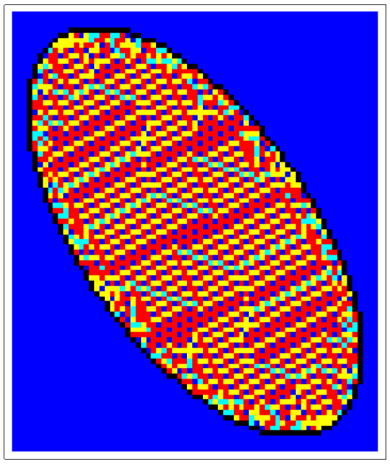
*4) $\cup_{z \in \mathbb{Z}^3} (\overline{\partial} T + Vz) \subset \{\Delta o_A = 3\}$*

*5) $1 \leq |V|^2 \leq C \det(V)$, where $|V|$ is the $\ell_2$ operator norm of the matrix, $\det(V)$ is the determinant of the $2 \times 2$ matrix formed by omitting the third column of $V$, and $C$ is a universal constant.*

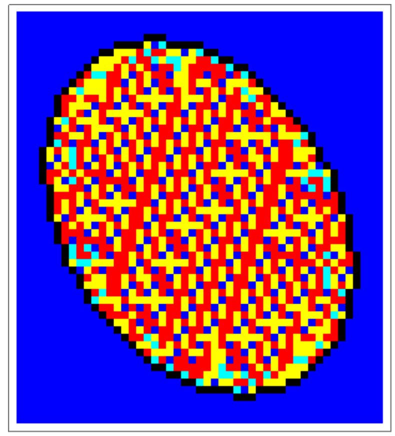
*6)* $V \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

The above proposition ensures that the tile T, along with its translation $T + Vz$ for $z \in \mathbb{Z}^3$, cover all of $\mathbb{Z}^2$, with overlap on the inner boundaries of adjacent tiles. The size of a tile is given by $(Tr(A) - 2)^{-1}$ [5]. It also gives that $\Delta o_A = \mathrm{Tr}(D^2 o_A)$ is periodic: for any $z \in \mathbb{Z}^3, \Delta o_A(x) = \Delta o_A(x + Vz)$. It also says that $\Delta o_A(x) = 3$ on the inner boundaries of all of the tiles. By fixing z in 3) and considering every w with $|z - w|_1 = 1$, one sees that every tile in $\mathbb{Z}^2$ has exactly 6 neighboring tiles. That $\det(V) \neq 0$ implies that $\dim \left( \mathrm{Span}_{\mathbb{R}} \{V_i\} \right) = 2$, where $\{V_i\}_{1 \leq i \leq 3}$ are the columns of V. Thus one can see that the vectors describing the periodicity of $\Delta o_A$ can be obtained by selecting any two columns of V.

The key property of matrices $A \in \Gamma^+$ that I explore in this chapter is that the identity element of $E_{A,k}$ will predominantly feature the biperiodic pattern $p_A = \Delta o_A$. In this sense, the identity elements corresponding to these matrices feature a high level of order; matrices not belonging to $\Gamma^+$ will generally produce more chaotic identity elements with no evident patterns. Figure 2.6 features the



(a) $E_{A_1,k}$



(b) $E_{A_2,k}$

Figure 2.6: Identity elements of $E_{A_i,k}$, $A_i \notin \Gamma^+$, $k = 18^2$ $A_1 \in \Gamma \backslash \partial\Gamma$, $A_2 \notin \Gamma$.

identity elements corresponding to various matrices that do not belong in $\Gamma^+$. $k = 18^2$ for all identity elements in Figure 2.6, and the matrices are as follows:

$$A_1 = \begin{bmatrix} 1 & \frac{49}{100} \\ \frac{49}{100} & \frac{2}{3} \end{bmatrix} \in \Gamma \backslash \partial\Gamma$$
$$A_2 = \begin{bmatrix} \frac{7}{4} & \frac{45}{99} \\ \frac{45}{99} & \frac{4}{3} \end{bmatrix} \notin \Gamma,$$

where $\partial\Gamma$ denotes the set of matrices represented by the Euclidean toplogical boundary of $\Theta \subset \mathbb{R}^3$.

## 2.3 The Sandpile Boundary Value Problem

Given a matrix $A \in \Gamma^+$ with $\det(A) > 0$ recall the following graph:

$$E_{A,k} = \{x \in \mathbb{Z}^2 : \frac{1}{2}xAx < k\}$$

Next, form the graph $E'_{A,k}$ by considering the outer boundary $\partial E_{A,k}$ of the above graph, and identifying all vertices of $\partial E_{A,k}$ as one sink vertex. The sink vertex has edges connecting to $E_{A,k}$ according to the adjacencies of the vertices of the outer boundary $\partial E_{A,k}$ with the inner boundary $\overline{\partial} E_{A,k}$ (counting multiplicities).

The goal of the remainder of this chapter is to characterize the recurrent representative of the sandpile identity element of $E'_{A,k}$ in the limit of large $k$. First a suitable definition of an odometer function for the recurrent identity element on $E'_{A,k}$, a graph with sink, is given. Our construction is first motivated by the following discrete boundary value problem:

Find the pointwise-minimal function $v : \mathbb{Z}^2 \to \mathbb{Z}$ satisfying

$$\begin{cases} \Delta v(x) \leq 3 & x \in E_{A,k} \\ v(x) \geq 0 & x \in E^c_{A,k}. \end{cases} \tag{2.1}$$

Where pointwise-minimality is interpreted in the following way: Let

$$\mathcal{O} := \{w : \mathbb{Z}^2 \to \mathbb{Z} \quad | \quad w \text{ satisfies } (2.1)\}.$$

Then $v(x) = \inf_{w \in \mathcal{O}}(w(x))$. Consider the following proposition:

**Proposition 7.** *Let $u, w : \mathbb{Z}^2 \to \mathbb{Z}$. If $\Delta u(x) \leq 3$ and $\Delta w(x) \leq 3$ for all $x \in \mathbb{Z}^2$, then $v(x) := \min(u(x), w(x))$ also satisfies $\Delta v(x) \leq 3$ for all $x \in \mathbb{Z}^2$.*

*Proof.* Let $x \in \mathbb{Z}^2$ be arbitrary. WLOG say that $v(x) = u(x)$. We then have

$$\Delta v(x) = -4v(x) + \sum_{y \sim x} v(y) = -4u(x) + \sum_{y \sim x} v(y) \leq -4u(x) + \sum_{y \sim x} u(y),$$

since $v(y) \leq u(y)$ for all $y$. The rightmost expression above is equal to $\Delta u(x)$, which is less than or equal to 3.

$\square$

Now it is shown that the solution, $v$, to this BVP is a recurrent function on $E_{A,k}$, giving that $\Delta v$ is a recurrent sandpile on $E'_{A,k}$.

**Proposition 8.** *The solution to the BVP (1) is a recurrent function on $E_{A,k}$.*

*Proof.* Given $Y \subset E_{A,k}$ and a function $w$ satisfying $\Delta w \leq 3$ on $Y$, define the integer $w_0 := \sup_{\partial Y}(v - w)$. Then define the function $\tilde{w} = w + w_0$, so that $\sup_{E_{A,k} \backslash Y \cup \partial E_{A,k}}(v - \tilde{w}) \geq 0$. Note that nonnegativity of the sup follows from the definitions of $w_0$ and $\tilde{w}$, and the fact that $\partial Y \subset E_{A,k} \backslash Y \cup \partial E_{A,k}$.

Now, it needs to be shown that $\sup_Y(v - \tilde{w}) \leq 0$. If it weren't, that is, if there were a $y_0 \in Y$ such that $\tilde{w}(y_0) < v(y_0)$, then the function

$$f(x) := \begin{cases} \min(\tilde{w}(x), v(x)) & x \in Y \\ v(x) & x \in Y^c \end{cases}$$

would satisfy the BVP, contradicting that $v$ is the pointwise-least solution (since $f$ is strictly less than $v$). That $f$ satisfies the BVP can be seen by noting that, on $E^c_{A,k} \subset Y^c$, $f(x) = v(x) \geq 0$. Further, by Proposition 7, one has that $\Delta f \leq 3$ on $E_{A,k}$, since on $Y \cup \partial Y$, $f(x) = \min(v(x), \tilde{w}(x))$ (note that $\tilde{w} \geq v$ on $\partial Y$). This gives that $\Delta f(x) \leq 3$ for all $x \in Y$. Next, for $x \in Y^c$, one has $\Delta f(x) \leq \Delta v(x)$, since $f = v$ at all $x \in Y^c$, and one has $f \leq v$ for points in $Y$ (which may neighbor points $x \in Y^c$). $\square$

Thus, by Proposition 5, $(\Delta v)|_{E_{A,k}}$ is a recurrent sandpile on $E_{A,k}$. Next it is shown that the sandpile $(\Delta v)|_{E_{A,k}}$ belongs in the same equivalence class as the all-zeroes configuration (thus proving that it is the recurrent representative of the identity element). It suffices to find a function $w : E_{A,k} \to \mathbb{Z}$ such that $\Delta v + \tilde{\Delta}_G w = 0$ on $E_{A,k}$. $-v$ is exactly this function. Note that $v|_{\partial E_{A,k}} \equiv 0$, since if there were a vertex $y \in \partial E_{A,k}$ such that $v(y) > 0$, then one would have $(1 - \mathbb{1}_{x=y})v \in \mathcal{O}$, since $v(x) \geq 0$ for all $x \in \partial E_{A,k}$ the condition $\Delta v(x) \leq 3$ for all $x \in E_{A,k}$ is preserved. This contradicts the fact that $v$ is the minimal function satisfying (1). Thus $v|_{\partial E_{A,k}} \equiv 0$, which gives that $\tilde{\Delta}_G v = (\Delta v)|_{E_{A,k}}$. This gives

$$\Delta v + \tilde{\Delta}_G(-v) = \Delta v + \Delta(-v) = 0$$

on $E_{A,k}$, by linearity of the Laplacian. Thus $\Delta v$ is the recurrent representative of the identity element.

In an effort to make the difference between the operators $\Delta$ and $\tilde{\Delta}_G$ clear, it is noted that the former operator may in general include topplings of the boundary $\Delta E_{A,k}$, while the latter operator does not represent toppling of the outer boundary; only topplings of the interior. The two operators are equivalent when $v|_{\partial E_{A,k}} \equiv 0$.

In what follows, $v$ is referred to as the odometer function.

## 2.4 Main results

Fix a matrix $A \in \Gamma^+$ such that $\det(A) > 0$ (so that the level set of $\frac{1}{2}x^T Ax$ is indeed an ellipse). Let $\lambda_1 < \lambda_2$ be the eigenvalues of the matrix A, with corresponding (unit-length) eigenvectors $v_1$, $v_2$.

In the following, this chapter only considers matrices lying in the fundamental domain of the Apollonian circle packing corresponding to $(x, y) \in [0, 2]^2$ (see the discussion around Proposition 4). Matrices lying in this fundamental domain that also have positive determinant can be shown to satisfy $0 < \lambda_1 \leq 1 < \lambda_2$. A complete description of these matrices can be found in [4]. We use below that these matrices satisfy $\mathrm{Tr}(A) > 2$, which follows immediately from the construction relating $\Gamma$ and $\Theta$ in the discussion before Proposition 4.

Since the matrix A is symmetric, $v_1$ and $v_2$ are orthogonal. Choose $v_1$ so that the angle $\theta$ that it makes with the x-axis is $\theta \in (-\pi, \pi]$, and choose $v_2$ so that $v_1 \times v_2$ points out of the page. Let $r_1$ and $r_2$ represent the semi-major and semi-minor axes of the ellipse $E_{A,k}$, respectively. They are given by $r_i = \sqrt{\frac{2k}{\lambda_i}}$.

Let $o_A$ represent a recurrent integer superharmonic witness for A, translated so that $o_A(0) = 0$. Let $v_{A,k}(x)$ be the solution to the boundary value problem in Section 2.3, translated so that $v_{A,k}(x)|_{\partial E_{A,k}} \equiv k$. In what follows, the subscripts A and k will sometimes be omitted to make the notation less cumbersome.

Experiments reveal that, except for some points near the boundary of $E_{A,k}$ and some 1-dimensional noise on the interior of the graph, the sandpile identity element of $E_{A,k}$ (i.e., the recurrent representative) almost perfectly matches the pattern given by $p_A = \Delta o_A = TrA + \Delta \rho_A$ (see Proposition 6). As k tends to infinity, the sandpile matches the pattern more and more closely; the noise takes up proportionally less area. The goal of this chapter (in particular, of Theorem 2.4.1 below) is to quantify the convergence of the pattern given by $\Delta v_{A,k}$ to the pattern given by $\Delta o_A$ in the limit of $k \to \infty$.

Some more notation is required before the main result is stated.

**Definition 7.** *Define the sets*

$$\tilde{E}_{A,k} := \{x \in \mathbb{R}^2 : \frac{1}{2}x^T Ax < k\} \subset \mathbb{R}^2$$

$$\tilde{E}_{A,k} \subset \tilde{F}_{L,A,k} := \{x \in \mathbb{R}^2 : d(x, \tilde{E}_{A,k}) \leq L\} \subset \mathbb{R}^2,$$

$$G_{L,A,k} := \{x \in E_{A,k} : d(x, \partial\tilde{E}_{A,k}) \geq L\} \subset E_{A,k} \subset \mathbb{Z}^2,$$

$$\tilde{G}_{L,A,k} := \{x \in \tilde{E}_{A,k} : d(x, \partial\tilde{E}_{A,k}) \geq L\} \subset \tilde{E}_{A,k} \subset \mathbb{R}^2,$$

where $L \in \mathbb{R}^+$, and $d(\cdot,\cdot)$ is the Euclidean distance. When referring to the set $\tilde{G}_{L,A,k}$ the subscripts $A$ and $k$ are often dropped when what is meant is clear from the context.

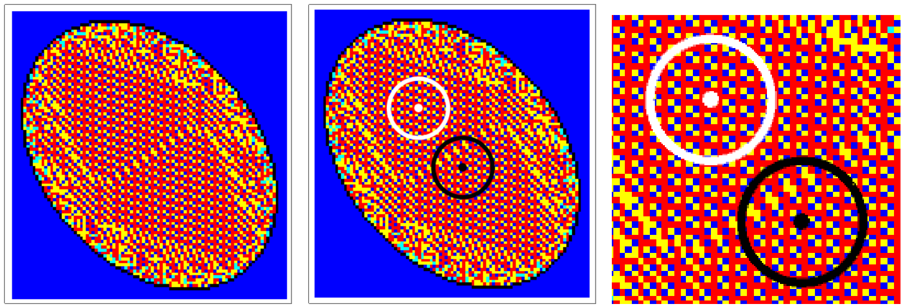
Note that $E_{A,k} = \tilde{E}_{A,k} \cap \mathbb{Z}^2$ and $G_{L,A,k} = \tilde{G}_{L,A,k} \cap \mathbb{Z}^2$. That is, if the name of a set has a tilde $\sim$ over it, then it is a subset of $\mathbb{R}^2$, and if it doesn't, then it is a subset of $\mathbb{Z}^2$.

Consider an open ball $B_r(x)$ (in the Euclidean metric) around every point $x \in E_{A,k}$. If on $B_r(x) \cap E_{A,k}$, the sandpile matches the $\Delta o_A$ pattern perfectly, this point is called r-good (see Figure 2.7).

**Definition 8.** *A point $x_0 \in E_{A,k}$ is r-good if there exist $y, z \in \mathbb{Z}^2$ and $w \in \mathbb{Z}$ such that $v(x) = o_A(x + y) + z \cdot x + w$ for all $x \in B_r(x_0) \cap E_{A,k}$*

In particular, by taking the Laplacian of both sides of the equation in the above definition, one sees that if a point is r-good, then the sandpile pattern given by $\Delta v$ exactly matches some translation of the pattern $\Delta o_A$ in $B_r(x)$.

**Theorem 2.4.1.** *Take $r(k) = o(k^{1/4})$ and $r(k) \geq 3|V|^3$ for all $k \in \mathbb{R}^+$. Let $f(k, A)$ be the fraction of points in $G_{r(k),A,k}$ that are **not** r-good. Then*



(a) The identity element of an ellipse

(b) $x_1$ (white point) and $x_2$ (black point), with balls of radius 10 centered at each

(c) Enlarged portion of figure b)

Figure 2.7: A demonstration of r-goodness. The white point $x_1$ is r-good, while the black point $x_2$ is not.

$$\limsup_{k\to\infty} f(k,A)\cdot\frac{k^{1/4}}{r(k)} \le C\cdot g(A),$$

*where $C$ is a universal constant, and $g(A)$ is a constant depending only on the matrix $A$, given by*

$$g(A)=\begin{cases}\sqrt{\lambda_1+\lambda_2}(\sqrt{\lambda_1}+\sqrt{\lambda_2})\left(\sqrt{\frac{1}{\lambda_1\lambda_2(1+2\lambda_1\lambda_2)}}+2\sqrt{\frac{\lambda_1\lambda_2}{1+2\lambda_1\lambda_2}}\right) & \lambda_1<\frac{1}{\sqrt{2}}\\ (\sqrt{\lambda_1}+\sqrt{\lambda_2})\left(\frac{1}{\sqrt{\lambda_2}}+\sqrt{2\lambda_2}\right) & \lambda_1\ge\frac{1}{\sqrt{2}}\end{cases}.$$

In particular, the above implies that the fraction of $r(k)$-good points tends to 1, provided that $r(k)$ is $o(k^{1/4})$. Note that the fraction of points in $G_{r,A,k}$ (rather than $E_{A,k}$) is considered in order to exclude points whose r-ball is not contained in the ellipse. The rate of $k^{1/4}$ appearing in the left hand side of the inequality in the theorem is a result of the $O(\sqrt{k})$ difference between the identity element's odometer function ($v$) and the integer superharmonic witness ($o$), detailed in Lemma 2.4.2 below.

The proof is adapted from Theorem 10 in [6]. Recurrence of $o$ and $v$ is used to give an upper bound for $|v-o|$ on $E_{A,k}$ (Lemma 2.4.2). Then, a 'touching map' whose range consists of good points (Lemma 2.4.3) is constructed. Finally, the area of the range of the touching map to is estimated give a lower bound on the number of r-good points.

**Lemma 2.4.2.** *There exists a recurrent integer superharmonic witness $o$ for $A$ such that*

$$\sup_{x \in E_{A,k}} |v(x) - o(x)| \leq h^2(A)\sqrt{k} + o(\sqrt{k}),$$

*where*

$$h(A)^2 := \begin{cases} \sqrt{\frac{\lambda_1+\lambda_2}{\lambda_1\lambda_2(1+2\lambda_1\lambda_2)}} + 2\sqrt{\frac{\lambda_1\lambda_2(\lambda_1+\lambda_2)}{1+2\lambda_1\lambda_2}} & \lambda_1 < \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{\lambda_2}} + \sqrt{2\lambda_2} & \lambda_1 \geq \frac{1}{\sqrt{2}} \end{cases}.$$

*Proof.* Using Proposition 6, write $o(x) = q(x) + L(x) + \rho(x)$, with $\rho(0) = 0$, and where the subscript $A$ has been dropped for convenience.

First note that, since o and v are both recurrent on $E_{A,k}$ (see Propositions 6 and 8), one has

$$\sup_{E_{A,k}} |o - v| \leq \sup_{\partial E_{A,k}} |o - v|.$$

Thus it suffices to show that $\sup_{\partial E_{A,k}} |o-v| \leq h^2(A)\sqrt{k} + o(k^{1/2})$. One has that

$$\sup_{x\in\partial E} |o - v|(x) \leq \sup_{x\in\partial E} (q_A(x) + |L(x)| - k) + \sup_{x\in\partial E} |\rho(x)|,$$

since $v|_{\partial E_{A,k}} = k$ and $q_A|_{\partial E_{A,k}} \geq k$.

Note that $\sup_{x\in\partial E} |\rho(x)| = o(\sqrt{k})$, since $\rho$ is a periodic function. Next consider the term $\sup_{x\in\partial E}(q_A(x) + |L(x)| - k)$, and note the following inclusion of sets: $\partial E_{A,k} \subset \tilde{F}_{1,A,k}$. This is apparent from the fact that for any $x \in \partial E_{A,k}$, $d(x, \tilde{E}_{A,k}) \leq 1$, since $x \sim y$ for some $y \in E_{A,k}$. 0 For each $x \in \partial E_{A,k}$ one can assign a point

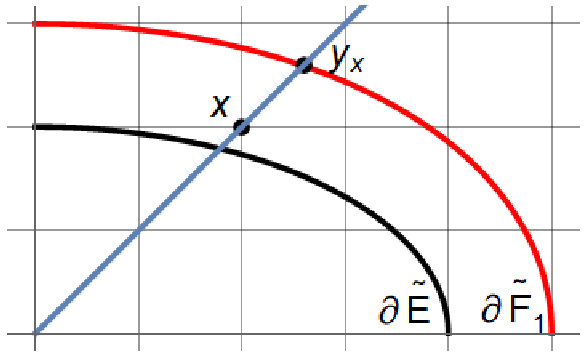



Figure 2.8: $x \in \partial E$ and the corresponding $y_x \in \partial \tilde{F}_1$. $\partial \tilde{E}$ is pictured in black, and $\partial \tilde{F}_1$ is pictured in red.

$y_x \in \partial \tilde{F}$ to it by letting $y_x$ be the point on $\partial \tilde{F}$ that also lies on the line passing through the origin and $x$. See Figure 2.8.

Write $L(x) = b \cdot x = b_1 x_1 + b_2 x_2$. Without loss of generality, take $|b_1| \leq \frac{1}{2}$ and $|b_2| \leq \frac{1}{2}$. If this *wasn't* the case, one could subtract some vector $b' \in \mathbb{Z}^2$, so that the modified superharmonic representative is still integer valued, and still has the appropriate growth at infinity. The Laplacian of the integer superharmonic witness is unaffected by the change in linear term. Thus, for all $x$, one has the inequality

$$|L(x)| \leq \|b\| \|x\| \leq \frac{1}{\sqrt{2}} \|x\| \leq \frac{1}{\sqrt{2}} \|y_x\|,$$

where the last inequality follows from the fact that x lies between the origin and $y_x$ on the line that determines $y_x$.

One can bound $Q(x)$ from above in a similar manner, noting that, as vectors in $\mathbb{R}^2$, $y_x = (1+\epsilon)x$ for some $\epsilon \geq 0$. Thus

$$Q(x) = \frac{1}{2} x^T A x \leq \frac{(1+\epsilon)^2}{2} x^T A x = Q(y_x).$$

Thus one has, for any $x \in \partial E_{A,k}$, that $Q(x) + |L(x)| \leq Q(y_x) + \frac{1}{\sqrt{2}} |y_x|$. Now $Q(y) + \frac{1}{\sqrt{2}} |y|$ needs to be maximized for $y \in \partial \tilde{F}$.

Now let $(x_1, x_2)$ parametrize $\mathbb{R}^2$, and switch to coordinates $(x_1', x_2')$ in which the matrix is diagonalized, so that the semi-major axis aligns with the $x_1'$ axis, while the semi-minor axis aligns with the $x_2'$ axis.

One then has that

$$E_{A,k} = \{x \in \mathbb{Z}^2 \,:\, \frac{1}{2}x'^T A' x' < k\},$$

where

$$A' = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}.$$

In what follows, work in the primed coordinate system, but drop the primes for ease of notation.

It now suffices to bound $Q(z) + \frac{1}{\sqrt{2}}|z|$ for $z \in \partial\tilde{F}$ with $z_2 \geq 0$, that is, to consider only the top half of $\partial\tilde{F}$, by symmetries of $Q(z) + \frac{1}{\sqrt{2}}|z|$.

First write the top half of the ellipse, $\partial\tilde{E}$, as a function of x. One then has that

$$x_2 = \sqrt{\frac{2k - \lambda_1 x_1^2}{\lambda_2}}.$$

Next note that, at any point on the top of $\partial\tilde{E}$, the outward facing unit normal vector is given by

$$n(x_1, x_2) = \frac{1}{\Delta}\begin{bmatrix} \lambda_1 x_1 \\ \lambda_2 x_2 \end{bmatrix},$$

where $\Delta := \sqrt{\lambda_1^2 x_1^2 + \lambda_2^2 x_2^2}$.

Now, all of the points $z \in \partial\tilde{F}^+$ can be written as $z = (x_1, x_2) + n(x_1, x_2)$. Thus, parametrize $\partial\tilde{F}^+$ by its $x_1$-coordinate in the following way:

$$\partial\tilde{F}^+ = \left\{ \Big(x_1\big(1 + \frac{\lambda_1}{\Delta}\big), \sqrt{\frac{2k - \lambda_1 x_1^2}{\lambda_2}}\big(1 + \frac{\lambda_2}{\Delta}\big)\Big) \quad | \quad -\sqrt{\frac{2k}{\lambda_1}} \leq x_1 \leq \sqrt{\frac{2k}{\lambda_1}} \right\}.$$

Using this parametrization, first consider the linear term, $\frac{1}{\sqrt{2}}|z|$, as a function of $x_1$. One has

$$\begin{aligned}\frac{1}{\sqrt{2}}|z| &= \frac{1}{\sqrt{2}}\Big(x_1^2\big(1+\frac{\lambda_1^2}{\Delta^2}+\frac{2\lambda_1}{\Delta}\big)+\frac{2k-\lambda_1 x_1^2}{\lambda_2}\big(1+\frac{\lambda_2^2}{\Delta^2}+\frac{2\lambda_2}{\Delta}\big)\Big)^{1/2}\\ &= \frac{1}{\sqrt{2}}\Big(1+\big(x_1^2+\frac{2k-\lambda_1 x_1^2}{\lambda_2}\big)+\frac{4k}{\Delta}\Big)^{1/2}.\end{aligned}$$

Next, write Q(z) in the same manner. We have

$$\begin{aligned}Q(z) &= \frac{1}{2}z^T A z = \frac{1}{2}\Big(\lambda_1 x_1^2\big(1+\frac{\lambda_1^2}{\Delta^2}+\frac{2\lambda_1}{\Delta}\big)+(2k-\lambda_1 x_1^2)\big(1+\frac{\lambda_2^2}{\Delta^2}+\frac{2\lambda_2}{\Delta}\big)\Big)\\ &= k+\Delta+\frac{\lambda_1^3 x_1^2+\lambda_2^3 x_2^2}{2\Delta^2}\end{aligned}$$

Now one has, for $z=(x_1,x_2)\in\partial\tilde{F}^+$, that

$$\begin{aligned}&\frac{1}{\sqrt{2}}|z|+Q(z)-k \qquad (2.2)\\ &\quad= \frac{1}{\sqrt{2}}\Big(1+\big(x_1^2+\frac{2k-\lambda_1 x_1^2}{\lambda_2}\big)+\frac{4k}{\Delta}\Big)^{1/2}+\Delta+\frac{\lambda_1^3 x_1^2+\lambda_2^3 x_2^2}{2\Delta^2}\\ &\quad\le \frac{1}{\sqrt{2}}\Big(1+\big(x_1^2+\frac{2k-\lambda_1 x_1^2}{\lambda_2}\big)+\frac{4\sqrt{k}}{C_2}\Big)^{1/2}+\Delta+\frac{C_1}{2C_2^2}\\ &\quad\le \frac{1}{\sqrt{2}}\Big(\big(x_1^2+\frac{2k-\lambda_1 x_1^2}{\lambda_2}\big)\Big)^{1/2}+\Delta+\frac{C_1}{2C_2^2}+\Big(1+\frac{4\sqrt{k}}{C_2}\Big)^{1/2}.\end{aligned}$$

The above inequalities use the fact that there exist constants $C_1$ and $C_2$ (which depend on $\lambda_1$ and $\lambda_2$ but not on $k$) such that, for all $z\in\partial\tilde{F}^+$, $\Delta\ge C_2\sqrt{k}$ and $\lambda_1^3 x_1^2+\lambda_2^3 x_2^2\le C_1 k$.

One needs to find the maximum value of the $x$-dependent terms, that is, of

$$M(x_1) := \sqrt{\frac{\frac{2k}{\lambda_2}+x_1^2(1-\frac{\lambda_1}{\lambda_2})}{2}}+\Delta$$

for $|x_1|\le\sqrt{2k/\lambda_1}$. Note that

$$M'(x_1) = \frac{1}{\sqrt{2\lambda_2}}\frac{(\lambda_2-\lambda_1)x_1}{\sqrt{2k+(\lambda_2-\lambda_1)x_1^2}} - \frac{\lambda_1(\lambda_2-\lambda_1)x_1}{\sqrt{2k\lambda_2-\lambda_1(\lambda_2-\lambda_1)x_1^2}},$$

which is nonsingular on $|x_1| \leq \sqrt{2k/\lambda_1}$. For $\lambda_1 < \frac{1}{\sqrt{2}}$, the solutions to $M'(x_1) = 0$ are the following:

$$x_1 = 0$$
$$\pm x_1 = \pm\nu := \pm\sqrt{2k\lambda_2}\frac{\sqrt{1-2\lambda_1^2}}{\sqrt{-\lambda_1^2+\lambda_1\lambda_2-2\lambda_1^3\lambda_2+2\lambda_1^2\lambda_2^2}}.$$

When $\lambda_1 \geq \frac{1}{\sqrt{2}}$, the only solution is

$$x_1 = 0.$$

Taking another derivative:

$$M''(x_1) = -\frac{\lambda_1^2(\lambda_2-\lambda_1)^2x_1^2}{(2\lambda_2k-\lambda_1(\lambda_2-\lambda_1)x_1^2)^{3/2}} - \frac{\lambda_1(\lambda_2-\lambda_1)}{(2\lambda_2k-\lambda_1(\lambda_2-\lambda_1)x_1^2)^{1/2}}$$
$$-\frac{\lambda_2^{3/2}(x_1-\frac{\lambda_1x_1}{\lambda_2})^2}{\sqrt{2}(2k+(\lambda_2-\lambda_1)x_1^2)^{3/2}} + \frac{\lambda_2-\lambda_1}{\sqrt{2\lambda_2}\sqrt{2k+(\lambda_2-\lambda_1)x_1^2}}.$$

Note that $M''$ is nonsingular on the domain.

**Case 1:** $\lambda_1 < \frac{1}{\sqrt{2}}$,

Note that $M''(\pm\nu) < 0$ (this can be seen by noting that $\lambda_2 > 1$). Thus at $\pm\nu$, $M'$ must switch sign from positive to negative, giving that $\pm\nu$ are absolute maxima on the domain in the case of $\lambda_1 < \frac{1}{\sqrt{2}}$.

Plugging $\pm\nu$ into $M$ and simplifying, one gets

$$M(\pm\nu) = \sqrt{\frac{\lambda_1+\lambda_2}{\lambda_1\lambda_2(1+2\lambda_1\lambda_2)}k} + 2\sqrt{\frac{\lambda_1\lambda_2(\lambda_1+\lambda_2)}{1+2\lambda_1\lambda_2}k}.$$

Dividing by $\sqrt{k}$, the result follows.

**Case 2:** $\lambda_1 \geq \frac{1}{\sqrt{2}}$

Using that $\lambda_1 \geq \frac{1}{\sqrt{2}}$ and $\lambda_2 > 1$, it is easily seen that $M(0) > M(\pm\sqrt{\frac{2k}{\lambda_1}})$.

Again plugging $x_1 = 0$ into $M$ and simplifying, one gets that

$$M(0) = \frac{\sqrt{k}}{\sqrt{\lambda_2}} + \sqrt{2\lambda_2 k}.$$

Combining the above two cases with inequality (2) and the material preceding, one sees that

$$\sup_{x \in E_{A,k}} |v(x) - o(x)| \leq h(A)^2\sqrt{k} + o(\sqrt{k}). \qquad \square$$

The next lemma, due to Pegden and Smart, states that if the quadratically-lowered integer superharmonic witness touches $v$ from below at 0 in $B_R$, then the point 0 is $C^{-1}R$-good for some universal constant $C$. The proof of this lemma, found in [6], makes use of the maximum principle used in the definition of recurrent functions.

**Lemma 2.4.3.** *[6, Lemma 14], There is a universal constant $C > 1$ such that, if*

*1) $R \geq C|V|^3$*

*2) $v : \mathbb{Z}^2 \to \mathbb{Z}$ is recurrent in $B_R$*

*3) $\psi_y(x) = o(x) - \frac{1}{2}\frac{|V|^2}{R^2}|x - y|^2 + k$ for some $k \in \mathbb{R}$*

*4) $\psi_y$ touches $v$ from below at 0 in $B_R$, that is, $\max_{B_R} \psi_y - v = 0$ and this maximum is achieved at 0*

*then 0 is $C^{-1}R$-good.*

In order to estimate the fraction of r-good points, the work below compares subsets of $\mathbb{Z}^2$ to their analogues in $\mathbb{R}^2$, and uses various geometric and measure-theoretic techniques in the continuum. This chapter proceeds by proving some properties of the sets $\tilde{G}$ and $\tilde{E}$ (which are subsets of $\mathbb{R}^2$).

**Lemma 2.4.4.** *$\tilde{G}_{L,A,k}$ is convex.*

*Proof.* Let $P, Q \in \tilde{G}_{L,A,k}$ be arbitrary. It needs to be shown that for all points $S \in \overline{PQ}$, $B_L(S) \subset \mathrm{cl}(\tilde{E}_{L,A,k})$ where $\overline{PQ}$ is the line segment connecting $P$ and $Q$, and cl denotes topological closure. Of course, $S \in \tilde{E}_{L,A,k}$, by convexity of $\tilde{E}_{L,A,k}$ and $\tilde{G}_{L,A,k} \subset \tilde{E}_{L,A,k}$.

Now, let $T$ be the midpoint of the line segment $\overline{PQ}$, and consider a coordinate parametrization $(x_1, x_2)$ of $\mathbb{R}^2$ where $T$ sits at the origin, and $P$ and $Q$ sit on the negative and positive $x_1$ axes, respectively. Consider also a topologically open rectangle $R \subset \mathbb{R}^2$ centered on the origin ($T$) with sides parallel to the coordinate axes, height $2L$ and width $|\overline{PQ}|$, so that the segment $\overline{PQ}$ bisects the rectangle horizontally. Note that each of the corners of $R$ are exactly a distance $L$ from one of the points $P$ or $Q$, so that the corners of $R$ are all contained in the closure of $\tilde{E}_{L,A,k}$. Then by convexity of $\mathrm{cl}(\tilde{E}_{L,A,k})$, $R \subset \mathrm{cl}(\tilde{E}_{L,A,k})$. This gives that for any point $x \in B_L(S) \cap R$, $x \in \mathrm{cl}(\tilde{E}_{L,A,k})$.

Now consider a point $x \in B_L(S) \cap R^c$, with coordinates $(x_1, x_2)$. Without loss of generality, let $x_1 \geq 0$. Let $S$ and $Q$ have coordinates $(s_1, s_2)$ and $(q_1, q_2)$ respectively. It now suffices to show that $x$ is within distance $L$ of $Q$, because $B_L(Q) \subset \mathrm{cl}(\tilde{E}_{L,A,k})$. One has:

$$\begin{aligned} d^2(x, Q) &= (x_1 - q_1)^2 + (x_2 - q_2)^2 = (x_1 - q_1)^2 + (x_2 - s_2)^2 \\ &\leq (x_1 - s_1)^2 + (x_2 - s_2)^2 < L^2. \end{aligned}$$ □

The following construction connects the relevant discrete sets with their continuum counterparts.

For any $x = (x_1, x_2) \in \mathbb{Z}^2$, define $x^{\square} \subset \mathbb{R}^2$ as the closed square of unit length centered on $x$, i.e.

$$x^{\square} = \left[x_1 - \frac{1}{2}, x_1 + \frac{1}{2}\right] \times \left[x_2 - \frac{1}{2}, x_2 + \frac{1}{2}\right] \subset \mathbb{R}^2.$$

Next, for any number $L$, define the sets

$$A = \{x \in \mathbb{Z}^2 \,:\, x^{\square} \subset \tilde{G}_L\}$$

$$B = \cap_{B_i \in \mathcal{B}} B_i,$$

where

$$\mathcal{B} = \{C \subset \mathbb{Z}^2 \,:\, \cup_{x \in C} x^{\square} \supset \tilde{G}_L\}.$$

Further, define

$$\tilde{A} = \cup_{x \in A} x^{\square}$$

$$\tilde{B} = \cup_{x \in B} x^{\square}$$

That is, $\tilde{A}$ is the largest union of unit squares (centered around lattice points) such that $\tilde{A} \subset \tilde{G}_L$, and $\tilde{B}$ is the smallest union of unit squares (centered around lattice points) so that $\tilde{B} \supset \tilde{G}_L$.

The following lemma is due to Levine:

**Lemma 2.4.5.** *$|\tilde{B}| - |\tilde{A}| \leq 16|\partial \tilde{E}_k|$, where $|\partial \tilde{E}_k|$ is the length of the circumference of the ellipse and $|\tilde{A}|$ and $|\tilde{B}|$ are the areas of $\tilde{A}$ and $\tilde{B}$ respectively.*

*Proof.* First, note that for all $x \in B - A$, $d(x, \partial\tilde{G}) \leq \frac{1}{\sqrt{2}}$, since $x^{\square} \not\subset \tilde{G}$. Thus $B_1(x)$ intersects $\partial\tilde{G}$ in an arc of length at least $2 - \sqrt{2}$, where $B_1(x)$ is the ball of radius 1 centered at $x$. Define $\alpha_x := B_1(x) \cap \partial\tilde{G}$ (see Figure 2.9). Next, note that for every point $z \in \partial\tilde{G}$, $z$ lies on at most 9 arcs $\alpha_x$. This follows by noting that for any point $z \in \mathbb{R}^2$, $B_1(z) \subset B_2(z')$, where $z'$ is a closest lattice point to z. $B_2(z')$ contains exactly 9 points, thus there are no more than 9 points with distance less than 1 from z. Then write, noting that $B - A$ is a finite set, and using $|B - A|$ to denote the number of points in $B - A$:

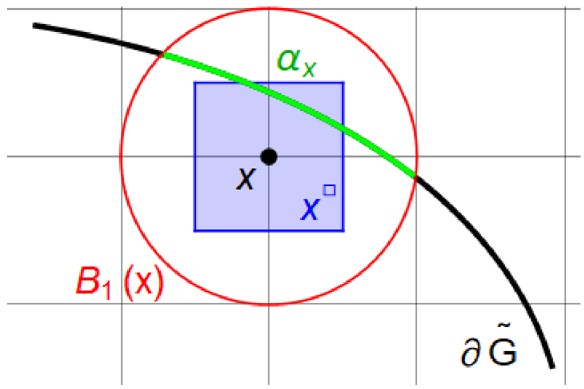



Figure 2.9: A schematic for Lemma 2.4.5. $x \in B - A$.

$$(2-\sqrt{2})|B-A| \le \sum_{x \in B-A} |\alpha_x| = \int_{\partial \tilde{G}} \sum_{x \in B-A} \mathbb{1}_{\alpha_x}(z)dz \le 9 \int_{\partial \tilde{G}} dz = 9|\partial \tilde{G}|.$$

Conclude the proof by noting that the closures of the sets $\tilde{G}_k \subset \tilde{E}_k$ are compact and convex (Lemma 2.4.4). Thus $|\partial \tilde{G}| \le |\partial \tilde{E}|$ (see, e.g., [8]). Since $\frac{9}{2-\sqrt{2}} < 16$, the proof is finished. □

The final lemma is a calculation of the area of the set $\tilde{G}$.

**Lemma 2.4.6.** *For constant $L$ and for sufficiently large $k$,*

$$|\tilde{E} - \tilde{G}_{Lr(k)k^{1/4}}| = Lr(k)k^{1/4}(C_E - \pi Lr(k)k^{1/4}),$$

*where $C_E$ is the circumference of the ellipse, and the vertical bars above refer to the Lebesgue measure of the continuum set.*

*Proof.* Consider, for every point $x \in \partial \tilde{E}$, a line segment $\ell_x$ of length $Lrk^{1/4}$ running inward normal to $\partial \tilde{E}$. Let $z_x$ refer to the end of $\ell_x$ that lies inside the ellipse. It is now proven that for sufficiently large $k$, the end of each line segment will coincide with a point on the boundary of $\tilde{G}_{Lrk^{1/4}}$, and that these segments will be non-intersecting (i.e. $\ell_x \cap \ell_y = \emptyset$ for $x \neq y$.) Certainly, $\partial \tilde{G}_{Lrk^{1/4}}$ will not coincide with interior points of any line segment, lest the conditions defining $\tilde{G}_{Lrk^{1/4}}$ be violated.

Then, $z_x$ will lie on $\partial\tilde{G}_{Lrk^{1/4}}$, provided that there does not exist a $y \in \partial\tilde{E}, y \neq x$ such that $d(y, z_x) < d(x, z_x)$. This will be satisfied if $Lrk^{1/4} < \inf_{p\in\partial\tilde{E}} r_p$, where $r_p$ is the radius of curvature of $\partial\tilde{E}$ at a point $p \in \partial\tilde{E}$. The smallest radius of curvature occurs at the points on the semi-major axis, where $r_p = \frac{\sqrt{2k\lambda_1}}{\lambda_2}$. Thus, require that $Lrk^{1/4} < \frac{\sqrt{2k\lambda_1}}{\lambda_2}$. This criterion also suffices for non-intersection of the line segments. Of course, this condition will be met in the limit of large $k$, where it is noted that $r = o(k^{1/4})$. Then use coordinates $(\theta, n)$, where $\theta$ gives points on the ellipse according to $x = \left(\sqrt{\frac{2k}{\lambda_1}}\cos(\theta), \sqrt{\frac{2k}{\lambda_2}}\sin(\theta)\right)$, and $n$ gives the position on the fiber $\ell_x$, with $n = 0$ coinciding with $\partial\tilde{E}$. Note that $\theta$ does not give the angle from the positive horizontal axis in general. In these coordinates, the Jacobian determinant becomes

$$J = \frac{-nr_1r_2}{r_2^2\cos^2(t) + r_1^2\sin^2(t)} + r_1r_2\sqrt{\frac{\cos^2(t)}{r_1^2} + \frac{\sin^2(t)}{r_2^2}}.$$

The area $|\tilde{E} - G_{Lrk^{1/4}}|$ then becomes the integral of the Jacobian over $0 \leq \theta < 2\pi$ and $0 \leq n \leq Lrk^{1/4}$, giving $Lrk^{1/4}(C_E - Lrk^{1/4}\pi)$, as desired. Note that the nonnegativity of the Jacobian (for sufficiently large $k$) on the region of integration follows from $r = o(k^{1/4})$.

□

*Proof of Theorem 2.4.1.* Consider the set $G_{7Ch(A)rk^{1/4},A,k} \subset \mathbb{Z}^2$, with $h(A)$ defined in 2.4.2 (and referred to henceforth as $h$), $C$ the constant from Lemma 2.4.3, and the set $G_{L,A,k}$ given in Definition 7. Consider the function

$$\phi_y(x) = o(x) - \frac{1}{2}\frac{|V|^2}{(2Cr)^2}|x - y|^2$$

for $y \in G_{7Chrk^{1/4}}$. Note that, for all sufficiently large $k$, $v(y) - \phi_y(y) = v(y) - o(y) \leq$

$h^2\sqrt{k} + o(\sqrt{k})$ (from Lemma 2.4.2). Furthermore, for $z \in E - B_{3hk^{1/4}\cdot 2Cr}(y)$,

$$\begin{aligned}(v - \phi_y)(z) &= (v - o)(z) + \frac{1}{2}\frac{|V|^2}{(2Cr)^2}|z - y|^2 \\ &\geq -h^2\sqrt{k} + \frac{9}{2}|V|^2 h^2\sqrt{k} + o(\sqrt{k}) \\ &\geq 3h^2\sqrt{k} + o(\sqrt{k}),\end{aligned}$$

where the above inequality uses that $|V|^2 \geq 1$ from Proposition 6. Combining the two previous inequalities gives that, for sufficiently large $k$, $(v - \phi_y)$ attains its minimum in E on the set $B_{6Chk^{1/4}r}(y)$, say at the point $x_y$. Thus $\phi_y$ can be translated to touch $v$ from below in $E$ at $x_y$. In particular, for $k$ sufficiently large (so that $6Chk^{1/4}r(k) \geq 2Cr(k)$), $\phi_y$ can be translated to touch $v$ from below in $B_{2Cr(k)}(x_y)$ at $x_y$. Next apply Lemma 2.4.3 (taking $R = 2Cr(k)$) to see that $x_y$ is $2r(k)$-good. Following [6], refer to the map $\theta : y \mapsto x_y$ as the 'touching map'.

Since, by Lemma 2.4.3, points in the range of the touching map are guaranteed to have a ball surrounding them that witness the matching of the sandpile identity element with the pattern given by $\Delta o_A$, one needs to estimate the number of points in the range of the touching map. If the touching map were injective, one would have that the number of good points is equal to $|G_{7Ch(A)rk^{1/4}}|$, the size of the domain of the touching map. However, this map is not injective in the standard sense; and one must instead make use of a weaker form of injectivity.

**Claim [6]:** For every $y \in G_{7Chrk^{1/4}}$, there are sets $y \in T_y \subset E$ and $x_y \in S_y \subset B_{|V|}(x_y)$ such that $|T_y| \leq |S_y|$ and $S_y \cap S_{\tilde{y}} \neq \emptyset$ implies $S_y = S_{\tilde{y}}$ and $T_y = T_{\tilde{y}}$.

Following [6], choose a $\mathcal{Y} \subset G_{7Chrk^{1/4}}$ maximal subject to $\{S_y | y \in \mathcal{Y}\}$ being disjoint. One then has

$$|\cup_{y\in\mathcal{Y}} S_y| \geq \sum_{y\in\mathcal{Y}} |T_y| \geq |\cup_{y\in\mathcal{Y}} T_y| \geq |G_{7Chrk^{1/4}}|,$$

where the above uses that $|S_y| \geq |T_y|$ for all y, the subadditivity of the measure, and the above claim for each of the inequalities, respectively. Lastly, note that for all $y$, every point in $S_y$ is r-good. This follows because every $S_y$ contains a 2r-good point, the diameter of $S_y$ is at most $2|V|$, and $r \geq 3|V|$. One thus have that the fraction of good points in $G_r$ is at least $\frac{|G_{7Chrk^{1/4}}|}{|G_r|}$.

Now aim to estimate this fraction by instead considering the areas of the continuum counterparts of the above sets; $|\tilde{G}_{7Chrk^{1/4}}|$ and $|\tilde{G}_r|$. Note that

$$|G_L| - |\tilde{G}_L| \leq |B| - |\tilde{A}| = |\tilde{B}| - |\tilde{A}| \leq 16|\partial \tilde{E}_k|$$

and

$$|\tilde{G}_L| - |G_L| \leq |\tilde{B}| - |A| = |\tilde{B}| - |\tilde{A}| \leq 16|\partial \tilde{E}_k|,$$

where $A, B, \tilde{A}, \tilde{B}$ are given in the construction before Lemma 2.4.5 for fixed $L$, and one can appeal directly to Lemma 2.4.5 for the last inequality in each of the above two lines. Thus $\big||\tilde{G}_L| - |G_L|\big| \leq 16|\partial \tilde{E}_k|$. This gives:

$$\begin{aligned}\frac{|G_{7Chrk^{1/4}}|}{|G_r|} &\geq \frac{1}{|G_r|}(|\tilde{G}_{7Chrk^{1/4}}| - \big||G_{7Chrk^{1/4}}| - |\tilde{G}_{7Chrk^{1/4}}|\big|) \\ &\geq \frac{1}{|G_r|}(|\tilde{G}_{7Chrk^{1/4}}| - 16|\partial \tilde{E}_k|) \geq \frac{|\tilde{G}_{7Chrk^{1/4}}| - 16|\partial \tilde{E}_k|}{|\tilde{G}_r| + \big||\tilde{G}_r| - |G_r|\big|}\end{aligned}$$

Now consider the fraction of points that are **not** r-good. By the above, this fraction is bounded from above by

$$\begin{aligned}&\frac{|\tilde{G}_r| + \big||\tilde{G}_r| - |G_r|\big| - |\tilde{G}_{7Chrk^{1/4}}| + 16|\partial \tilde{E}_k|}{|\tilde{G}_r| + \big||\tilde{G}_r| - |G_r|\big|} \leq \frac{|\tilde{G}_r| - |\tilde{G}_{7Chrk^{1/4}}| + 32|\partial \tilde{E}_k|}{|\tilde{G}_r|} \\ &= \frac{|\tilde{E}_k - \tilde{G}_{7Chrk^{1/4}}| - |\tilde{E}_k - \tilde{G}_r|}{|\tilde{E}_k| - |\tilde{E}_k - \tilde{G}_r|} + \frac{32|\partial \tilde{E}_k|}{|\tilde{E}_k| - |\tilde{E}_k - \tilde{G}_r|}\end{aligned}$$

Appealing now to Lemma 2.4.6, the fraction of r-bad points in $G_r$ is at most

$$\frac{7Chrk^{1/4}C_E - \pi(7Chr)^2k^{1/2} - rC_E + \pi r^2}{\pi\frac{2k}{\sqrt{\lambda_1\lambda_2}} - rC_E + \pi r^2} + \frac{32|\partial \tilde{E}_k|}{\pi\frac{2k}{\sqrt{\lambda_1\lambda_2}} - rC_E + \pi r^2}$$

Now, using the estimate $C_1(r_1+r_2) \leq C_E \leq C_2(r_1+r_2)$ (for some constants $C_1$ and $C_2$) and manipulating, one gets:

$$f(k,r) \leq \frac{(7ChrC_2 - \frac{C_1 r}{k^{1/4}})\sqrt{\frac{2}{\lambda_1\lambda_2}}(\sqrt{\lambda_1}+\sqrt{\lambda_2}) + \pi(\frac{r^2}{k^{3/4}} - \frac{49C^2h^2r^2}{k^{1/4}})}{\frac{2\pi}{\sqrt{\lambda_1\lambda_2}}k^{1/4} - \frac{r}{k^{1/4}}C_2\sqrt{\frac{2}{\lambda_1\lambda_2}}(\sqrt{\lambda_1}+\sqrt{\lambda_2}) + \frac{r^2}{k^{3/4}}\pi}$$
$$+ \frac{1}{k^{1/4}} \frac{32C_2\sqrt{\frac{2}{\lambda_1\lambda_2}}(\sqrt{\lambda_1}+\sqrt{\lambda_2})}{\frac{2\pi}{\sqrt{\lambda_1\lambda_2}}k^{1/4} - \frac{r}{k^{1/4}}C_2\sqrt{\frac{2}{\lambda_1\lambda_2}}(\sqrt{\lambda_1}+\sqrt{\lambda_2}) + \frac{r^2}{k^{3/4}}\pi}$$

Multiplying by $\frac{k^{1/4}}{r}$ and taking the limit as $k \to \infty$, obtain

$$\limsup_{k\to\infty} f(k,r) \cdot \frac{k^{1/4}}{r} \leq Ch(\sqrt{\lambda_1}+\sqrt{\lambda_2}),$$

where all constants have been absorbed into $C$. This proves the result. □

The discussion below aims to generalize the result slightly. Note that the graph $E_{A,k}$ is formed by taking the intersection of the square grid with an ellipse centered at the origin. Note that the main theorem can be extended to ellipses centered at an arbitrary point in $\mathbb{R}^2$, with 2.4.1 still holding (only with different, weaker constants $g(A)$, see Figure 2.3).

When replacing $(0,0)$ with an arbitrary centerpoint $p = (p_1, p_2) \in \mathbb{R}^2$, note that only Lemma 2.4.2 changes, with the rest of the proof of Theorem 2.4.1 being carried out as above. The dicussion below gives the adaptations that can be made to Lemma 2.4.2 in order to accommodate the recentered ellipse.

First note that the statement

$$\sup_{E^p} |o-v| \leq \sup_{\partial E^p} |o-v| \leq -k + \sup_{\partial E^p} \left(q(x) + |L(x)|\right) + \sup_{\partial E^p} p(x) \tag{2.3}$$

still holds, as both $v$ and $o$ are still recurrent on $E^p$, and $v \equiv k$ on $\partial E^p$ by definition. Without loss of generality, take $p \in [-\frac{1}{2}, \frac{1}{2}) \times [-\frac{1}{2}, \frac{1}{2})$, since one can always translate the witness $o_A$, and $v$ is the solution to a $\mathbb{Z}^2-$translation-invariant BVP.

Consider the second term in the rightmost expression above. As in Lemma 2.4.2, one can bound the linear function $L(x)$ by $\frac{1}{\sqrt{2}}|x|$. Next consider the quadratic term $q(x)$. Define $q'(x) = \frac{1}{2}(x-p)^T A(x-p)$. One then has

$$q(x) = q'(x) - \frac{1}{2}p^T Ap + p^T Ax \leq q'(x) + p^T Ax \leq q'(x) + \frac{1}{\sqrt{2}}\lambda_2|x|,$$

where the last inequality has used Cauchy-Schwarz and that $\lambda_2$ is the largest eigenvalue of $A$. One now has

$$\sup_{\partial E^p}\big(q(x) + |L(x)|\big) \leq \sup_{\partial E^p}\big(q'(x) + \frac{1}{\sqrt{2}}(1+\lambda_2)|x|\big) \leq \sup_{\partial E^p} q'(x) + \sup_{\partial E^p}\frac{1}{\sqrt{2}}(1+\lambda_2)|x|.$$

Heuristically speaking, the worst-case scenario for $q'(x)$ occurs when there is a point $y_x$ lying on $\partial E^p$ with unit distance from $x \in E^p$, with both points lying along the semi-minor axis. This gives

$$q'(y_x) = \frac{1}{2}(r_2+1)^2 v_2^T A v_2 = \frac{\lambda_2}{2}(\sqrt{\frac{2k}{\lambda_2}}+1)^2 = k + \sqrt{2k\lambda_2} + \frac{\lambda_2}{2}$$

Plugging back into equation (2.3), one sees that

$$\sup_{E^p}|o-v| \leq \sqrt{2k\lambda_2} + \sup_{\partial E^p}\frac{1}{\sqrt{2}}(1+\lambda_2)|x| + \sup_{\partial E^p} p(x) + \frac{\lambda_2}{2}$$

On $\partial E^p$, $|x|$ can be bounded by $\sqrt{\frac{2k}{\lambda_1}} + 1 + \frac{1}{\sqrt{2}}$. This is due to the fact that, for a boundary point $x \in \partial E^p$, $d(0,x) \leq d(0,p) + d(p,x)$. The terms $\sqrt{\frac{2k}{\lambda_1}} + 1$ come from $d(p,x)$, as it is the semimajor axis of the ellipse plus one, which accommodates for the farthest that a boundary point can be away from the interior of the ellipse. The $\frac{1}{\sqrt{2}}$ term bounds $d(p,0)$.

This gives that

$$\sup_{E^p}|o-v| \leq \big(\sqrt{2\lambda_2} + \frac{1+\lambda_2}{\sqrt{\lambda_1}}\big)\sqrt{k} + o(\sqrt{k}).$$

The proof of Theorem 2.4.1 now proceeds as above, with the above inequality taking the place of Lemma 2.4.2, and the constant $\sqrt{2\lambda_2} + \frac{1+\lambda_2}{\sqrt{\lambda_1}}$ taking the place of $h^2$.

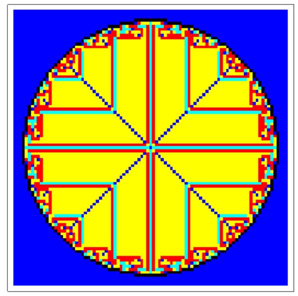


Figure 2.10: The identity element of $B_{1800}(0) \cap \mathbb{Z}^2$.

## 2.5 Open Questions

Related open questions involve identity elements for non-maximal integer superharmonic ellipses. For example, take two maximal matrices $A$ and $B$, corresponding to Apollonian circles that are tangent. Consider a straight line running between these two centers in the $\{z = 2\}$ plane, and then consider the inverse image of the projection operator from $\partial\Gamma$ onto the $\{z = 2\}$ plane. What do the identity elements of matrices lying on this curve in $\partial\Gamma$ look like? In particular, what does the identity element of the matrix at the point of tangency look like?

As one deviates from the peak of a cone (a maximal matrix) down the side of the cone in $\partial\Gamma$, experiments reveal that defects in the identity element appear in the form of stripes (Figure 2.6). Future work may include explorations of the rates at which stripes appear, and a characterization of the stripes appearing as one descends from a peak in $\partial\Gamma$ in arbitrary directions down the cone.

One may also attempt to characterize the identity element of a circular region of the square grid. The matrix $\mathbb{I}_2$ which gives the circle is not itself a maximal integer superharmonic matrix, but is the limit of such matrices. Experiments reveal that the identity element is a constant background of $\Delta v = 2$, with some line-shaped defects in the interior of the circle (see Figure 2.10). Note further that vertical and horizontal lines only feature sites with 1 or 3 grains of sand, while the diagonal lines have 0 grains of sand.

CHAPTER 3

# THE STOCHASTIC SANDPILE ON A CYCLE

## 3.1 Introduction

This chapter is concerned primarily with the so-called stochastic sandpile model (abbreviated SS) and its dynamics on the cycle graph, particularly with the number of stochastic updates needed until a stable state has been reached. A overview of the model is presented here, giving the details in Section 3.2.

Begin with $n$ indistinguishable particles on the vertices of $\mathbb{Z}_n$ in an arbitrary configuration. The locations of the particles are updated in discrete time steps as follows. If at any time a site has two or more particles on it, the site is deemed *unstable* and must be *toppled*. In toppling the unstable site, let two particles at the site each perform an independent, lazy, symmetric random walk on the graph. That is, for *lazy parameter* $0 < p(n) < 1$, a particle does not move with probability $p(n)$, moves one step clockwise with probabilty $\frac{1-p(n)}{2}$, and moves one step counter-clockwise with probability $\frac{1-p(n)}{2}$. Both particles move independently according to this distribution. This process continues, toppling one unstable site per timestep, until each site has exactly one particle, at which time the system is deemed *SS-stable*. Define the *odometer function for SS* to be the function $v : \mathbb{Z}_n \to \mathbb{N} \cup \{0\}$ that counts the number of times each site topples in the stabilization process. The stochastic sandpile model enjoys the so-called abelian property, which states that given an arbitrary initial particle configuration and quenched randomness, the odometer function is independent of the choice of unstable site that is toppled at each time step, making $v$ a well-defined random function (see Section 3.2.3).

The main result of this chapter will compare SS to a similar system: the activated random walk model (abbreviated ARW) (see [12]). ARW is briefly described here as a continuous-time process, and is later presented as a discrete-time construction in Section 3.2. Begin with $n$ particles on $\mathbb{Z}_n$. A particle in the ARW model can be in one of two states, "active" or "sleeping". Active particles perform continuous-time symmetric random walk along the graph, dictated by a Poisson clock with jump rate 1. An active particle also carries a second, independent Poisson clock with rate $\lambda$ whose ticks tell the particle to try to fall asleep. If the particle is alone on its vertex, it successfully changes its state to "sleeping". If it shares its vertex with other particles, it fails to sleep and remains in its "active" state. Sleeping particles can only exist on a vertex where they are the only particle, and remain asleep until another particle visits their vertex, at which time the particle instantaneously changes to an active state. This process continues until the configuration is *ARW-stable*, that is, until there is exactly one sleeping particle at every vertex (and no active particles). We define the *odometer function for ARW* to be the function $u : \mathbb{Z}_n \to \mathbb{N} \cup \{0\}$ that counts the number of times a Poisson clock (including the jump clocks and the sleep clocks) ticked at each site $x$ during the stabilization process.

Let $\eta$ represent a state of ARW (i.e. the number particles in each state are on each site of $\mathbb{Z}_n$), and let $|\eta|$ represent the corresponding SS state with the same number of particles on each site in $\eta$, ignorant of whether the particles are active or sleeping. Insofar as this chapter is concerned with ARW as it relates to the SS dynamics, define the random time $T_{-1}$ to be the first time that the ARW model has one particle (regardless of state) on each site in $\mathbb{Z}_n$, and define $\overline{u}$ to be the odometer function immediately after $T_{-1}$ (see Section 3.2 for a rigorous formulation).

**Theorem 3.1.1.** *Consider the ARW model with sleep rate $\lambda$ starting from arbitrary*

*$n$-particle state $\eta_0$, and let $u$ and $\overline{u}$ represent the odometer functions for the full stabilization process and stabilization process stopped at time $T_{-1}$. Similarly, let $v$ represent the odometer function for SS with lazy parameter $p = \frac{\lambda}{1+\lambda}$ and starting with state $|\eta_0|$. Then for all $x \in \mathbb{Z}_n$,*

$$\lceil \frac{\overline{u}(x)}{2} \rceil \overset{d}{=} v(x)$$

*and $\overline{u}(x) \leq u(x)$ almost surely for all $x \in \mathbb{Z}_n$.*

By combining Theorem 3.1.1 with Lemma 3.2.3, it follows that for $p(n)$ tending to 1 sufficiently rapidly, the stochastic sandpile on the $n$-cycle stabilizes in time $O(n^3)$ when starting from an arbitrary $n$-particle configuration.

**Corollary 3.1.1.1.** *Consider the stochastic sandpile model starting from an arbitrary configuration of $n$ particles on $\mathbb{Z}_n$, and let $0 < p(n) < 1$ be a function of $n$ satisfying*

$$\limsup_{n\to\infty} \left[\log(n) \cdot \big(1 - p(n)\big)\right] < 2.$$

*Then*

$$\lim_{n\to\infty} \mathbb{P}(T > An^3) = 0,$$

*where $T$ is the total time to stabilization, and $A$ is a constant independent of $n$.*

Lack of control over the effects of these avalanches remains an obstacle to proofs of fast stabilization of the stochastic sandpile model. For example, say the n-cycle contains a region of stability, that is, an interval along which each site contains exactly one particle. If a particle enters that interval and causes one of the sites to become unstable, large chain reactions can occur in which many of the sites in the interval must topple, leaving disastrously unpredictable particle configurations in its wake.

The ARW model couples nicely to the SS model, has dynamics on the cycle graph which are well-studied [16, 17], and can alleviate some of the aforementioned difficulties in analyzing SS. Aside from its own mathematical interest, the ARW model is useful to proofs of fast stabilization of the stochastic sandpile model for two main reasons. Firstly, the ARW model provides more paths to stability, in the sense that particles move independently rather than in pairs. This flexibility can be exploited by choosing the combinations of particle moves that are simplest and best-suited for a given analysis. The second important facet of ARW compared to SS is that its stabilization takes longer, since all particles must be asleep in addition to being alone on a vertex in order to be stable. This inequality between the stabilization times of the two models is useful to the goal of upper-bounding the stabilization time of SS. The proof of the Corollary 1 uses an analogous result proven in [17], which shows that ARW on the $n$-cycle undergoes a phase transition: for $\lambda(n)$ growing sufficiently rapidly with $n \to \infty$, stabilization occurs in $O(n^3)$ time, with stabilization time exponential in $n$ otherwise. The fast-phase result is stated below in Lemma 3.2.3.

An outstanding open problem, whose investigation led to this result, is to prove that SS on the $n$-cycle with *constant* (i.e. independent of $n$) lazy parameter $p$ stabilizes in polynomial time. Simulations suggest that the system does indeed stabilize in time $O(n^3)$. Figure 3.1 shows a log-log plot of average stabilization time vs. $n$. The result appears linear with slopes close to 3, giving evidence for $O(n^3)$ stabilization time of the constant-laziness SS model. The heuristic for fact that fast stabilization of the SS model with large lazy parameter appears easier to prove than that of the constant laziness SS model is as follows: despite there being more null topplings (topplings with no particles being displaced) in the lazy version, there are fewer moves resulting in both particles being displaced, which is

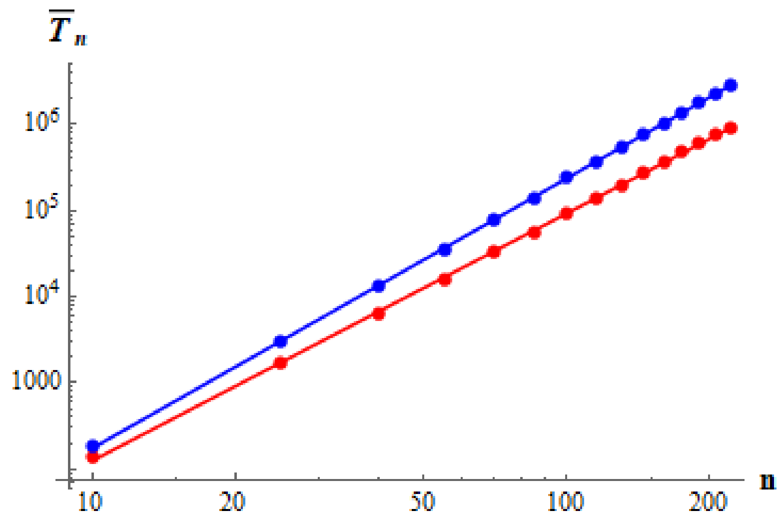



Figure 3.1: **SS simulations:** A log-log plot of average time to stabilization $\overline{T}_n$ over 200 trials for various values of $n$, with linear fits. The colors of the plots correspond to lazy parameters $p = \frac{1}{2}$ and $p = \frac{\log(n)}{1+\log(n)}$. The slope of the blue line is $\approx 3.14$ and the slope of the red line is $\approx 2.88$.

the mechanism driving the unpredictability of avalanches.

Some previous investigations have focused on phase transitions of SS of infinite lattices. In [15], the authors consider SS on $\mathbb{Z}$, with initial particle configuration given by i.i.d. Poisson random variables with mean $\mu$ at each site. They show the existence of a critical density $\mu_c$ such that for $\mu < \mu_c$ the system will eventually stabilize, but for $\mu > \mu_c$, the system does not stabilize. On finite graphs with sufficiently few particles, the system stabilizes in finite time almost surely, so the question of how long it the system will take to stabilize seems like a natural one. It is worth noting that for any more than $n$ particles on a finite graph with $n$ vertices, the system cannot stabilize, since there are no configurations with at most one particle on each site. Thus, this chapter investigates the stabilization properties of a system with the maximum number of particles.

The remainder of the chapter is structured as follows. In Section 3.2, both models and some of their important properties are rigorously introduced. In Section 3.3, some Markov chain machinery that allows us to couple the two models is developed, with the goal of expressing the SS Markov chain as a "quotient" of ARW Markov chain. Section 3.4 gives the coupling and the proof of Theorem 3.1.1.

## 3.2 Preliminaries

### 3.2.1 The Stochastic Sandpile Model on $\mathbb{Z}_n$

Label the sites in $\mathbb{Z}_n$ as $\{0, 1, 2, ..., n-1\}$, with indices increasing as we go counterclockwise. Consider the following definition:

**Definition 9.** *A **stochastic sandpile configuration** is a function $s : \mathbb{Z}_n \to \mathbb{N} \cup \{0\}$ satisfying*

$$\sum_{x \in \mathbb{Z}_n} s(x) = n, \tag{3.1}$$

*and $S$ refers to the set of all stochastic sandpile configurations.*

One can think of $s(x)$ as representing the number of particles on site $x \in \mathbb{Z}_n$, with equation (3.1) enforcing that there must always be exactly $n$ total particles on $\mathbb{Z}_n$.

For a given configuration $s$, a site $x$ is called *unstable for $s$* if $s(x) \geq 2$. If such a site exists, call $s$ unstable, otherwise, call $s$ stable.

The stochastic sandpile model begins with an initial particle configuration $s_0$. Stabilize the sandpile in discrete time steps as follows. Let $s_t$ represent the particle configuration at time $t \in \mathbb{N} \cup \{0\}$. Choose a site $x_t$ that is unstable for $s_t$ and *topple* this site by letting two particles perform independent lazy symmetric random walk

| $\tau^x$ | Transformation of $s$ | Probability |
|---|---|---|
| $\tau^x_{(0,0)}$ | None | $p^2$ |
| $\tau^x_{(1,0)}$ | $s(x) \to s(x) - 1$<br>$s(x-1) \to s(x-1) + 1$ | $p(1-p)$ |
| $\tau^x_{(0,1)}$ | $s(x) \to s(x) - 1$<br>$s(x+1) \to s(x+1) + 1$ | $p(1-p)$ |
| $\tau^x_{(1,1)}$ | $s(x) \to s(x) - 2$<br>$s(x+1) \to s(x+1) + 1$<br>$s(x-1) \to s(x-1) + 1$ | $\frac{(1-p)^2}{2}$ |
| $\tau^x_{(2,0)}$ | $s(x) \to s(x) - 2$<br>$s(x-1) \to s(x-1) + 2$ | $\frac{(1-p)^2}{4}$ |
| $\tau^x_{(0,2)}$ | $s(x) \to s(x) - 2$<br>$s(x+1) \to s(x+1) + 2$ | $\frac{(1-p)^2}{4}$ |

Table 3.1: All values of the SS toppling operator $\tau^x$ with corresponding effects on the configuration and probabilities of occurence.

steps with lazy parameter $0 < p(n) < 1$. That is, two particles at site $x_t$ each relocate independently according to the following distribution: do not move with probability $p(n)$, step clockwise with probability $\frac{1-p(n)}{2}$, and step counterclockwise with probability $\frac{1-p(n)}{2}$. Henceforth, we leave implicit the dependence of $p$ on $n$ for ease of notation. One can summarize the results of a stochastic sandpile toppling with the pair of nonnegative integers $(\rho_-, \rho_+)$, where $\rho_-$ and $\rho_+$ represent the number of particles that stepped clockwise and counterclockwise, respectively, in a toppling. Note that $(\rho_-, \rho_+)$ must satisfy the following two conditions:

$$\rho_-, \rho_+ \geq 0 \tag{3.2}$$

$$\rho_- + \rho_+ \leq 2. \tag{3.3}$$

Let $\tau^x$ be the (random) operator that represents a toppling at site $x$, with $\tau^x$ taking values $\tau^x_{(\rho_-, \rho_+)}$, and denote the resulting configuration by $s_{t+1} = \tau^{x_t} s_t$. Table 3.1 summarizes the possible values of $\tau^x$ and their effects on the configuration $s$.

### 3.2.2 The Activated Random Walk Model on $\mathbb{Z}_n$

In the activated random walk model, particles are in one of two states, "active" or "sleeping". Label the sites in $\mathbb{Z}_n$ in the same manner as above. Let $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$, and define the ordered set $\mathbb{N}_\mathfrak{s} = \mathbb{N}_0 \cup \{\mathfrak{s}\}$, with $0 < \mathfrak{s} < 1 < 2 < \ldots$. Definition 10 concerns an activated random walk particle configuration, which describes the number and type of particles at each site.

**Definition 10.** *An **activated random walk configuration** is a function* $\eta : \mathbb{Z}_n \to \mathbb{N}_\mathfrak{s} \cup \{0\}$ *with*

$$\sum_{x \in \mathbb{Z}_n} |\eta(x)| = n, \tag{3.4}$$

*where* $|\mathfrak{s}| := 1$. *Let* $H$ *refer to the set of all activated random walk configurations.*

$\eta(x) = \mathfrak{s}$ indicates that there is a single sleeping particle on site $x$, $\eta(x) \in \mathbb{N}$ indicates that there are $\eta(x)$ active particles on $x$, and $\eta(x) = 0$ indicates that there are no particles on $x$. Equation (3.2) enforces that there are $n$ total particles on $\mathbb{Z}_n$.

For a given configuration $\eta$, call a site $x$ *unstable* if $\eta(x) \geq 1$. A configuration with at least one unstable site is deemed unstable, otherwise it is stable. Note that $\eta$ is stable if and only if $\eta(x) = \mathfrak{s}$ for all $x$, that is, if each site contains one sleeping particle and no particles are active.

Below, the discrete-time formulation of the model is described. It has dynamics are equivalent with the continuous-time formulation given in the introduction. The activated random walk model will begin with an arbitrary initial $n$-particle configuration $\eta_0$, and is stabilized in discrete time steps as follows. Let $\eta_t$ represent the particle configuration at time $t \in \mathbb{N} \cup \{0\}$. Choose a site $x_t$ that is unstable for $\eta_t$ and *topple* this site by letting an active particle try to fall asleep with probability $\frac{\lambda}{1+\lambda}$, and otherwise taking a (non-lazy) symmetric random walk step with probability $\frac{1}{1+\lambda}$. Note that an ARW toppling involves the possible displacement of one particle, while an SS toppling involves the possible displacements of two particles.

The effects of each toppling move on the configuration $\eta$ are now described. Firstly, sleeping particles can only exist on a site by themselves. So, if one is toppling an active particle at a site $x$ where $\eta(x) \geq 2$ and that particle tries to go to sleep, it fails and remains active (remaining at site $x$). If it is alone (i.e. $\eta(x) = 1$), then it successfully goes to sleep (so that $\eta(x)$ becomes $\mathfrak{s}$). Secondly, if an active particle takes a step and arrives at a site containing a sleeping particle, the sleeping particle immediately wakes up, resulting in two active particles.

Let $\delta^x$ be the operator that represents a toppling at site $x$, with $\delta^x$ taking values $\delta^x_{-1}, \delta^x_{+1}, \delta^x_{\mathfrak{s}}$, which represent a particle taking a step clockwise, counterclockwise, and attempting to fall asleep, respectively. Denote the resulting configuration by $\eta_{t+1} = \delta^{x_t}\eta_t$.

Formalize the above by defining the following operations on $\mathbb{N}_{\mathfrak{s}}$ :

$$1 \cdot \mathfrak{s} = \mathfrak{s} \tag{3.5}$$

$$k \cdot \mathfrak{s} = k \qquad k \geq 2 \tag{3.6}$$

$$k + \mathfrak{s} = k + 1 \qquad k \geq 1 \tag{3.7}$$

$$k - \ell = k - \ell \qquad k \geq \ell \geq 0, \quad k, \ell \in \mathbb{N}_0 \tag{3.8}$$

and summarizing the distribution of the results of an ARW toppling at site $x$ in Table 3.2.

Update the configuration at time $t$ by choosing an unstable site $x_t$ and toppling it, and let the effects of the toppling move performed at $x_t$ be encoded in the (random) toppling operator $\delta_{x_t}$, so that $\eta_{t+1} = \delta_{x_t}\eta_t$.

### 3.2.3 Sitewise Representation for SS and ARW on $\mathbb{Z}_n$

I now introduce framework for the two models that prescribes all of the randomness from the start, making the dynamics easier to study (see [18] for an early example

| $\delta^x$ | Transformation of $\eta$ | Probability |
|---|---|---|
| $\delta^x_{\mathfrak{s}}$ | $\eta(x) \to \eta(x) \cdot \mathfrak{s}$ | $\frac{\lambda}{1+\lambda}$ |
| $\delta^x_{+1}$ | $\eta(x) \to \eta(x) - 1$<br>$\eta(x+1) \to \eta(x+1) + 1$ | $\frac{\lambda}{2(1+\lambda)}$ |
| $\delta^x_{-1}$ | $\eta(x) \to \eta(x) - 1$<br>$\eta(x-1) \to \eta(x-1) + 1$ | $\frac{1}{2(1+\lambda)}$ |

Table 3.2: All values of the ARW toppling operator $\delta^x$ with corresponding effects on the configuration and probabilities of occurence.

of this formulation). For each model, consider a *random field of instructions*

$$\mathcal{I}_{SS} = \{\tau_x^i \,:\, x \in \mathbb{Z}_n,\, i \in \mathbb{N}\}$$

$$\mathcal{I}_{ARW} = \{\delta_x^i \,:\, x \in \mathbb{Z}_n,\, i \in \mathbb{N}\}.$$

The set of instructions can be thought of as an infinite stack of instructions at each site in $\mathbb{Z}_n$, and each time a site $x$ is toppled, one uses the toppling operator at the top of the stack at $x$ and discard it. All of the randomness of the model is stored in the random set of instructions.

The following discussion applies to both SS and ARW. Let $\alpha = \{x_0, x_1, ...\}$ represent the chronologically-ordered sequence of sites that are toppled, with length $T \in \mathbb{N} \cup \{\infty\}$. Call $\alpha$ *legal* if $x_t$ is unstable for all $t$. Call $\alpha$ *stabilizing* if the sequence of toppling moves results in the stable configuration. Define the odometer function $v_\alpha : \mathbb{Z}_n \to \mathbb{N}_0 \cup \{\infty\}$ to be the function that counts the number of times each site has toppled in the sequence $\alpha$. The least action principle is stated below. It asserts that any legal sequence can be at most as long as a legal, stabilizing sequence.

**Lemma 3.2.1** (Least Action Principle, [12, 15]). *Let $\alpha$ and $\beta$ both be legal sequences for SS. If $\beta$ is stabilizing, then $v_\alpha(x) \leq v_\beta(x)$ for all $x$.*

*The same result holds for ARW odometers.*

Next, the abelian property is presented, which asserts that given a random set of instructions, the odometer function is independent of the choice of legal, stabilizing sequence.

**Lemma 3.2.2** (Abelian Property, [12, 15]). *Fix a stack of instructions for SS, and let $\alpha$ and $\beta$ both be legal and stabilizing sequences. Then $v_\alpha = v_\beta$.*

*The same result holds for ARW odometers.*

The abelian property asserts that the odometer function for a given stack of instructions is well-defined and independent of the choice of legal and stabilzing sequence. Accordingly, I henceforth drop the subscript and refer to the odometer functions (for a given stack of instructions) as $v$ and $u$ for SS and ARW, respectively. The stabilization time is defined to be number of toppling moves required for stabilization:

$$T_{SS} = \sum_{x \in \mathbb{Z}_n} v(x)$$
$$T_{ARW} = \sum_{x \in \mathbb{Z}_n} u(x),$$

which by the abelian property is independent of the choice of legal, stabilizing sequence. I emphasize that the above stabilization times are functions of the random stack of instructions.

### 3.2.4 Fast-Phase Stabilization for ARW

This section is concluded by stating the fast-phase stabilization result from [17]: for sufficiently high sleep rate, $ARW$ on the $n$-cycle stabilizes in $O(n^3)$ time.

**Lemma 3.2.3** (Fast-Phase ARW Stabilization [17])**.** *Let $T_{ARW}$ be the (random) stabilization time for ARW with sleep rate $\lambda(n)$ on $\mathbb{Z}_n$ with arbitrary initial particle configuration. If*

$$\liminf_{n\to\infty} \frac{\lambda(n)}{\log(n)} > \frac{1}{2},$$

*then*

$$\lim_{n\to\infty} \mathbb{P}(T_{ARW}(n) > Bn^3) = 0$$

*for some constant $B$ independent of $n$.*

## 3.3 Quotients of Markov Chains

I now develop a framework by which one can couple SS to ARW. The goal will be to compare ARW to SS by "projecting out" the particle states from the ARW Markov chain to obtain a Markov chain that is isomorphic to SS.

Let $H$ be a finite set, and consider the measurable space $(H, 2^H)$. Let $S = \cup_{i=-1}^{|S|-2} H_i$ be a partition of $H$ into at least two sets, with none of the $H_i$ empty. Let the map $\pi : H \to S$ project an element of $H$ to its cell in $S$. I will consider the Markov chain $\{\eta_t\}_{t\in\mathbb{N}_0}$ taking values in $H$ with transition probability $p^H : H \times 2^H \to \mathbb{R}$, and that starts in a state $\eta_0$ such that $\pi(\eta_0) = H_0$.

**Definition 11.** *Call the partition $S = \{H_{-1}, H_0, ...\}$* ***Markov-compatible with*** *$\{\eta_t\}$* ***up to*** *$H_{-1}$ if for all $H_i \in S\backslash H_{-1}$ and $H_j \in S$, $p^H(\eta, H_j)$ is independent of the choice of $\eta \in H_i$.*

**Definition 12.** *Let $\{\eta_t\}$ be a Markov chain on a finite set $H$ with initial state $\eta_0$, and let $S = \{H_{-1}, H_0, ...\}$ be a partition of $H$ that is Markov-compatible with $\{\eta_t\}$ up to $H_{-1}$. The **quotient transition probability** $p_S^H$ is defined as follows:*

$$p_S^H(H_i, H_j) = \begin{cases} p^H(\eta, H_j) \text{ for any } \eta \in H_i & H_i \neq H_{-1} \\ \delta_{H_i,H_j} & H_i = H_{-1}. \end{cases}$$

*Furthermore, let $\{s_t\}_{t\in\mathbb{N}_0}$ be a Markov chain taking values in $S$ with transition probability $p_S^H$, and that starts in state $s_0 = H_0 \ni \eta_0$. Refer to this as the **quotient Markov chain** of $\{\eta_t\}$.*

Note that the quotient transition probability is well-defined by the definition of Markov-compatibility, and is easily seen to be a bona-fide transition probability. Before stating our main lemma from this section, another definition is provided:

**Definition 13.** *Let $T_{-1}$ be the **hitting time for the set** $H_{-1}$:*

$$T_{-1} := \min\{t \geq 0 \,:\, \eta_t \in H_{-1}\},$$

*and let $\overline{\eta}_t := \eta_{\min(t,T_{-1})}$ be the corresponding **stopped Markov Chain**.*

Consider now the following claim:

**Lemma 3.3.1** (Quotient Markov Chains)**.** *$(s_0, s_1, ...)$ has the same law as $(\pi(\overline{\eta}_0), \pi(\overline{\eta}_1), ...)$.*

*Proof.* It suffices to show that:

$$p_S^H(H_i, H_j) = \mathbb{P}\big(\pi(\overline{\eta}_{t+1}) = H_j \,|\, \pi(\overline{\eta}_t) = H_i\big). \tag{3.9}$$

First assume that $H_i \neq H_{-1}$. Consider the RHS of (9). Since $\pi(\overline{\eta}_t) \neq H_{-1}$, we know that $t < T_{-1}$. Thus the RHS of (9) can be written as

$$\mathbb{P}\big(\pi(\eta_{t+1}) = H_j \,|\, \pi(\eta_t) = H_i\big) = \mathbb{P}\big(\eta_{t+1} \in H_j \,|\, \eta_t \in H_i\big) = p^H(\eta \in H_i, H_j)$$

where the last expression is well-defined over the choice of $\eta$ because $S$ is Markov-compatible up to $H_{-1}$. Using the definition of $p_S^H$ for $H_i \neq H_{-1}$, we are done with the case of $t < T_{-1}$.

Now let $H_i = H_{-1}$ in the RHS of (9). Since we are conditioning on $\pi(\overline{\eta}_t) = H_{-1}$, we have that $t \geq T_{-1}$. Thus we can write

$$\mathbb{P}\big(\pi(\overline{\eta}_{t+1}) = H_j \,|\, \pi(\overline{\eta}_t) = H_{-1}\big) = \mathbb{P}\big(\pi(\eta_{T_{-1}}) = H_j \,|\, \pi(\eta_{T_{-1}}) = H_{-1}\big) = \delta_{H_j, H_{-1}},$$

which is equal to $p_S^H(H_{-1}, H_j)$. This proves the lemma. □

## 3.4 SS as a quotient of ARW

This section is devoted to showing that the stochastic sandpile model is isomorphic to a quotient Markov chain of activated random walk. First define a (legal and stabilizing) toppling prescription for both models, so that each can be viewed as a well-defined Markov chain.

### 3.4.1 A Toppling Prescription for ARW and SS

I now would like to fix a legal, stabilizing toppling prescription for ARW, which I will write as a function of time: $j : \mathbb{N}_0 \to \mathbb{Z}_n$. Note that $j$ is also a function of the random field of instructions, though this dependence is left implicit for ease of notation. Define $j(t)$ for even and odd times separately. For all $k \in \mathbb{N}_0$:

$$j(2k) = \begin{cases} \min\{x : s_{2k}(x) \geq 2\} & \text{if } \max_{x \in \mathbb{Z}_n} s_{2k}(x) \geq 2 \\ \min\{x : x \text{ has an active particle}\} & \text{else} \end{cases} \tag{3.10}$$

$$j(2k+1) = \begin{cases} j(2k) & \text{if } \max_{x \in \mathbb{Z}_n} s_{2k}(x) \geq 2 \\ \min\{x : ix \text{ has an active particle}\} & \text{else.} \end{cases} \tag{3.11}$$

In words: at even times, topple the first site counterclockwise of the origin with 2 or more particles. If one is unable to find such a site, topple the first site counterclockwise of the origin with an active particle. Note that in either case, the toppling move is always legal. If the former case occurs, that is, if we've just toppled a site with 2 or more particles, topple this same site again at the next odd time. This toppling move is legal, since the particle that remains at the site (there is at least one) will still be active. If the latter case occured, that is, if we toppled a single active particle, we again topple the first active particle counterclockwise of the origin, if one exists. Stop toppling once we've stabilized, that is, once every particle is asleep. With the toppling prescription fixed, we can now think of ARW as a Markov chain. The state space $H$ of this Markov chain is the set of all ARW configurations, that is, the set of functions $\eta : \mathbb{Z}_n \to \mathbb{N}_{\mathfrak{s}}$ satisfying Definition 10.

The toppling prescription for SS will simply be to topple the first site counterclockwise of the origin, that is, to topple the site

$$j(t) = \min\{x : s_t(x) \geq 2\}. \tag{3.12}$$

### 3.4.2 Coupling

**Definition 14.** *The **SS Markov Chain** $\{s_t\}_{t\in\mathbb{N}_0}$ takes values in the set $S$ of all SS configurations. Let one Markov chain step correspond to one SS toppling, according to the toppling prescription given in (12).*

*The **ARW Markov Chain** $\{\eta_t\}_{t\in\mathbb{N}_0}$ takes values in the set $H$ of all ARW configurations. Let one Markov chain step correspond to two ARW topplings, according to the toppling prescription given in (10) and (11).*

For ease of coupling SS with ARW, let one step in the ARW Markov chain correspond to two ARW toppling moves. That is, if ARW stabilizes in $T_{ARW}$ toppling moves, this is thought of as $\lceil \frac{T_{ARW}}{2} \rceil$ Markov Chain steps. Each SS toppling move will constitute one SS Markov chain step.

Consider the equivalence relation on the set of ARW configurations defined as follows:

$$\text{For } \eta_1, \eta_2 \in H, \eta_1 \sim \eta_2 \iff |\eta_1(x)| = |\eta_2(x)| \quad \forall x. \tag{3.13}$$

It is clear that the partition given by $\sim$ can be identified with the set of all stochastic sandpile configurations $S$. Using notation from the previous section, define $H_{-1} = (1, 1, ..., 1)$. Consider the following lemma:

**Lemma 3.4.1.** *The partition $S$ is Markov-compatible with the ARW Markov chain up to $H_{-1}$.*

*Proof.* Let $H_i \in S\backslash H_{-1}$, $\eta_1, \eta_2 \in H_i$, and $H_j \in S$ be artbirary. We need to show that $p^H(\eta_1, H_j) = p^H(\eta_2, H_j)$. Let $s_i, s_j$ be the stochastic sandpile configurations (ignoring states of particles) representing $H_i$ and $H_j$, respectively. Since neither of $\eta_1, \eta_2$ are in $H_{-1}$, topple (in the ARW sense) the clockwise-most site with two or more particles twice. In particular, the site to be toppled is the same for $\eta_1$ as it is for $\eta_2$. Call this site $x$.

I now state a pair of conditions that are both necessary for a state $\eta \in H_i$ to have non-zero transition probability to $H_j$. Firstly, note that the numbers and states of particles at sites other than $x, x-1, x+1$ are unchanged by a toppling move, so that a set $H_j$ is accessible to $\eta \in H_i$ only if $s_j|_{\{x-1,x,x+1\}^c} = s_i|_{\{x-1,x,x+1\}^c}$. Second, make the following definitions:

$$\rho_- := s_j(x-1) - s_i(x-1) \tag{3.14}$$

$$\rho_+ := s_j(x+1) - s_i(x+1) \tag{3.15}$$

where $x-1$ and $x+1$ are to be interpreted modulo $n$. $H_j$ is accessible from $\eta \in H_i$ only if $(\rho_-, \rho_+)$ satisfy equations (3.2) and (3.3), and $s_j(x) - s_i(x) = -(\rho_- + \rho_+)$. This follows from particle conservation and the fact that at most two particles can be displaced per toppling.

Now assume that $H_i$ and $H_j$ are such that both of the above conditions are met (if this is not the case, then $p^H(\eta, H_j) = 0$ for all $\eta \in H_i$). Enumerate all possible sets $H_j$ to which $\eta \in H_i$ can transition with positive probability by the pair $(\rho_-, \rho_+)$. In what follows, I argue that the transition from $H_i$ to $H_j$ is independent of the state in $H_i$ that the system moves from. I discuss each possibility for $(\rho_-, \rho_+)$, case by case.

**Case 1:** $(\rho_-, \rho_+) = (0, 0)$: This case holds if and only if neither particle is displaced from $x$, that is, that both particles being toppled attempted to go to sleep. This happens with probability $\left(\frac{\lambda}{1+\lambda}\right)^2$, independently of the details of the state $\eta_1$ or $\eta_2$. Note that this case implies $H_j = H_i$.

**Case 2:** $(\rho_-, \rho_+) = (1, 0)$: There are two possible toppling outcomes corresponding to this case: a) the first particle steps clockwise, while the second one tries to fall asleep, and b) the first particle tries to fall asleep, and the second one steps clockwise. Each of these two possibilities occurs with probability $\frac{1/2}{1+\lambda}\frac{\lambda}{1+\lambda}$, giving a total transition probability of $\frac{\lambda}{(1+\lambda)^2}$.

**Case 3:** $(\rho_-, \rho_+) = (0, 1)$: Both particles step counterclockwise. This is similar to Case 2.

**Case 4:** $(\rho_-, \rho_+) = (1, 1)$: One particle steps clockwise, and one particle steps counterclockwise. This happens with probability $2\big(\frac{1/2}{1+\lambda}\big)^2$, where the factor of 2 represents that the left and right steps can occur in any order.

**Case 5:** $(\rho_-, \rho_+) = (2, 0)$: Both particles step clockwise. This occurs with probability $\big(\frac{1/2}{1+\lambda}\big)^2$.

**Case 6:** $(\rho_-, \rho_+) = (0, 2)$: Both particles step counterclockwise. This is similar to Case 5.

Since each transition probability is independent of the choice of ARW configuration in $H_i$, I have shown that $S$ is Markov-compatible with the ARW Markov chain up to $H_{-1}$. See Table 3.3 for the nonzero transition probabilities. □

**Lemma 3.4.2** (Coupling Lemma)**.** *Consider ARW with sleep rate $\lambda(n)$. For any initial configuration $\eta_0$, the stopped, projected Markov Chain $\{\pi(\overline{\eta}_t)\}_{t\in\mathbb{N}_0}$ on $S$ is isomorphic to the SS Markov chain $\{s_t\}_{t\in\mathbb{N}_0}$ on $S$, starting from $s_0 = \pi(\eta_0)$, with lazy parameter $p(n) = \frac{\lambda(n)}{1+\lambda(n)}$.*

*Proof.* By Lemmas 3.4.1 and 3.3.1, it suffices to show that the SS Markov chain evolves according to the transition probability $p_S^H$. This is easily checked by comparing corresponding rows in Tables 3.1 and 3.3. For any given pair $(\rho_-, \rho_+)$ which represents an accessible $H_j$ in the quotient chain, $p_S^H$ is obtained by summing the transition probabilities of all accessible states in $H_j$ in Table 3.3. Setting $p(n) = \frac{\lambda(n)}{1+\lambda(n)}$, this gives transitions identical to those found in Table 1. □

| Original State $\eta\vert_{\{x-1,x,x+1\}}$ | $\big(\eta(x-1), \eta(x), \eta(x+1)\big)$ | |
|---|---|---|
| $(\rho_-, \rho_+)$ | Transformed states $\eta'\vert_{\{x-1,x,x+1\}}$ | Transition Probabilities |
| (0,0) | $\big(\eta(x-1), \eta(x), \eta(x+1)\big)$ | $\left(\frac{\lambda}{1+\lambda}\right)^2$ |
| (1,0) | $\big(\eta(x-1)+1, \eta(x)-1, \eta(x+1)\big)$ | $\frac{\lambda}{1+\lambda}\frac{1/2}{1+\lambda}$ |
| | $\big(\eta(x-1)+1, (\eta(x)-1)\cdot\mathfrak{s}, \eta(x+1)\big)$ | $\frac{1/2}{1+\lambda}\frac{\lambda}{1+\lambda}$ |
| (0,1) | $\big(\eta(x-1), \eta(x)-1, \eta(x+1)+1\big)$ | $\frac{\lambda}{1+\lambda}\frac{1/2}{1+\lambda}$ |
| | $\big(\eta(x-1), (\eta(x)-1)\cdot\mathfrak{s}, \eta(x+1)+1\big)$ | $\frac{1/2}{1+\lambda}\frac{\lambda}{1+\lambda}$ |
| (1,1) | $\big(\eta(x-1)+1, \eta(x)-2, \eta(x+1)+1\big)$ | $2\left(\frac{1/2}{1+\lambda}\right)^2$ |
| (2,0) | $\big(\eta(x-1)+2, \eta(x)-2, \eta(x+1)\big)$ | $\left(\frac{1/2}{1+\lambda}\right)^2$ |
| (0,2) | $\big(\eta(x-1), \eta(x)-2, \eta(x+1)+2\big)$ | $\left(\frac{1/2}{1+\lambda}\right)^2$ |

Table 3.3: Possible transitions of the ARW Markov Chain. Each pair $(\rho_-, \rho_+)$ represents an $H_j$ accessible from $\eta$, with the middle column listing the accessible states $\eta' \in H_j$ from $\eta$.

*Proof of Theorem 3.1.1.* Theorem 3.1.1 is a simple consequence of Lemma 3.4.2. Consider the coupled isomorphic Markov chains. For a fixed field of instructions, the number of times each site has SS-toppled is exactly two times the number of times each site has ARW-toppled up until the very last step, thanks to our compatbile toppling prescriptions defined in (10), (11), and (12). The last step can correspond to either one or two ARW topplings, giving the formula $v(x) = \lceil \frac{\overline{u}}{2} \rceil$. That $\overline{u}(x) \leq u(x)$ almost surely for all $x$ is trivial. □

CHAPTER 4

# AN URN MODEL FOR OPINION PROPAGATION ON NETWORKS

## 4.1 Introduction

### 4.1.1 Statement of Problem and Result

Let $G = (\mathcal{V}, \mathcal{E})$ be a simple, connected graph, with each vertex $i \in \mathcal{V}$ representing an individual agent. In my model for opinion propagation, agents discuss an issue with their neighbors, each conversation resulting randomly in either an agreement on state $U$ or an agreement on state $V$. If two learners agree on state $U$ or $V$, both of the learners increase their propensity to prefer state $U$ or $V$, respectively, in the future. This is made precise in the following discussion.

For every vertex $i \in \mathcal{V}$ and timestep $t \in \mathbb{N} \cup \{0\}$, let the weights $(u_t^i, v_t^i) \geq 0$ represent the propensities of vertex $i$ for $U$ and $V$, respectively, at time $t$. For ease of notation, I write $(\vec{u}_t, \vec{v}_t)$, where $\vec{u}_t, \vec{v}_t \in \mathbb{R}^{\mathcal{V}}$ have components $(u_t^i)_{i\in\mathcal{V}}$ and $(v_t^i)_{i\in\mathcal{V}}$. For convenience, define the total weight of vertex $i$ and the fraction of that total weight stored in state $U$ to be

$$g_t^i := u_t^i + v_t^i$$
$$x_t^i := \frac{u_t^i}{g_t^i},$$

respectively. Consolidate notation with $\vec{g}_t$ and $\vec{x}_t$, similarly to the above. Enforce the initial conditions $(\vec{u}_0, \vec{v}_0)$ to be such that $u_0^i + v_0^i =: g_0^i > 0$ for all $i$, and define $\vec{\gamma}_t$ to be a vector with $\vec{\gamma}_t^i := \frac{1}{g_t^i}$ for later convenience.

The dynamics are as follows: at every timestep $t \geq 1$ choose a random edge $e = (i, j) \in \mathcal{E}$. Increment (only) each the two $g$ values:

$$g_t^i = g_{t-1}^i + 1$$

$$g_t^j = g_{t-1}^j + 1$$

with all other $g_t^k = g_{t-1}^k$ unchanged for $k \notin \{i, j\}$. Define:

$$p_t^e := \frac{u_t^i + u_t^j}{g_t^i + g_t^j} = \frac{x_t^i g_t^i + x_t^j g_t^j}{g_t^i + g_t^j},$$

as the pooled opinion of agents $i$ and $j$, and let $p_{t-1}^e$ give the probability of $i$ and $j$ agreeing on state $U$ at time $t$, given that edge $e$ is chosen at time $t$. If the chosen $i$ and $j$ agree on state $U$, increment each of their $u$ values:

$$u_t^i = u_{t-1}^i + 1$$

$$u_t^j = u_{t-1}^j + 1.$$

If they agree on opinion $V$, do not alter the $u$-values. All other $u_t^k = u_{t-1}^k$ for $k \notin \{i, j\}$ remain unchanged regardless of the outcome of the conversation along edge $e$.

I show that the dynamics of the system are governed by a discrete, stochastic version of the heat equation, with an "influence matrix" $L$ driving the propagation of opinions. The influence matrix acts like the graph Laplacian, but gives higher weight to vertices that have high degree, which have more conversations on average and therefore develop strong opinions more rapidly. Similarly to the graph Laplacian, the influence matrix has right-eigenvector $\vec{1}$ (the $|\mathcal{V}|$-dimensional vector with each component equal to 1); let $a_t$ be the coordinate of $\vec{x}_t$ corresponding to $\vec{1}$ with respect to a fixed, generalized eigenbasis of $L$ (discussed below). $a_t$ will be referred to as the *consensus coordinate*.

The goal of this chapter is to prove the following theorem, which states that a consensus of opinion is reached in the long-time limit.

**Theorem 4.1.1.** *There exists a random scalar* $0 \leq a_\infty \leq 1$ *such that*

$$\mathbb{E}[\|\vec{x}_t - a_\infty \vec{1}\|^2] \xrightarrow{t \to \infty} 0$$

### 4.1.2 Related Work

A similar class of frameworks for opinion propagation, called voter models, also feature randomly selected pairs of agents exchanging opinions. For example, in the Deffuant model, pairs of neighbors interact only when their opinions are within some threshold of one another, with consensus and/or polarization being driven by threshold size ([19]). Another example of a voter model is the Hegselmann-Krause model, in which an agent is randomly selected to have their opinion replaced by some determinstic function of their neighbors' opinions ([20]). The model presented in this chapter could perhaps be considered a stochastic voter model (stochastic in the sense that outcomes of conversations are random). A unique property of this model, however, is the pooled-experience nature of conversations, resulting in influences between agents that are random and dynamic, but that tend towards a graph-dependent object (the influence matrix). It should also be noted that this model features continuous opinions ($x_t^i \in [0, 1]$) with discrete actions (agents agree on either $U$ or $V$); different combinations of opinion and action spaces are featured throughout the literature.

This model can also be compared to the DeGroot model for learning, in which updates are made according to some constant 'trust matrix' $T$: $\vec{x}_{t+1} = T\vec{x}_t$ ([21]). The trust matrix can represent how much each agent trusts their neighbors as well as themself, giving a weighted average of their neighbors' beliefs and their own prior opinions. Other, similar models of opinion propagation have been studied, considering the effects of agents' self-confidence and network topology on long-term behavior ([22, 23, 24, 29]). Yet another related class of models for opinion propagation are 'probabilistic fuzzy models', which include agents' perceptions of some exogenous, albeit 'fuzzy' (the exact state is unclear) variables ([30]). I finally note that much of the literature on opinion propagation focuses on simulation-based studies, while rigorous proofs such as the one found in this paper are less common.

### 4.1.3 Outline of Chapter

The rest of the chapter proceeds as follows. In Section 4.2, I derive the fact that the behavior of $x_t$ is governed by a discrete-time stochastic heat equation, and give some important properties of the (stochastic) Laplacian operator driving the diffusion. In Section 4.3, I prove convergence of the consensus coordinate of $\vec{x}_t$, and in Section 4.4, I prove the decay of $\vec{x}_t - a_t\vec{1}$ (the *disagreement component*). In Section 4.5, I give a proof of Theorem 4.1.1, and in Section 4.6, I provide a conjecture that may lead to future work.

## 4.2 Stochastic Heat Equation

### 4.2.1 Preliminaries

At each timestep, an edge is randomly selected to host a 'conversation' between its two vertices. The following heuristic is equivalent and useful: let all edges have conversations, and uniformly at random select one edge to actually contribute to the dynamics.

Let $\Omega_t = \{\omega_t^e\}_{e\in\mathcal{E}} \in \{0,1\}^{\mathcal{E}}$ be the results of all conversations occuring at timestep $t$, with $\omega_t^e = 1$ if opinion $U$ is agreed on, and 0 otherwise. Similarly, let $\psi_t \in \mathcal{E}$ be the edge chosen at time $t$, and let $S_t^e = \mathbb{1}_{\{\psi_t = e\}}$. Define the filtrations

$$\mathcal{H}_1 \subset \mathcal{H}_2 \subset \mathcal{H}_3 \subset ...$$
$$\mathcal{G}_1 \subset \mathcal{G}_2 \subset \mathcal{G}_3 \subset ...$$
$$\mathcal{F}_1 \subset \mathcal{F}_2 \subset \mathcal{F}_3 \subset ...,$$

where $\mathcal{H}_t = \sigma(\Omega_t, \Omega_{t-1}, ..., \Omega_1)$ corresponds to the information received up to and immediately after discussions in the $t^{\text{th}}$ round, $\mathcal{G}_t = \sigma(\psi_t, \psi_{t-1}, ..., \psi_1)$ corresponds to the information received given all of the chosen edges up to and including time $t$, and let $\mathcal{F}_t = \sigma(\mathcal{H}_t, \mathcal{G}_t, ..., \mathcal{H}_1, \mathcal{G}_1)$. Note that $\Omega_t \in m(\mathcal{H}_t)$, but $\Omega_{t+1} \notin m(\mathcal{H}_t)$. Since $\psi_t$ does not care about the previous edges chosen or the results of any concurrent or previous conversations, let $\sigma(\psi_t)$ be independent of $\sigma(\mathcal{F}_{t-1} \cup \mathcal{H}_t)$. Furthermore, for $e \neq f$, let $\omega_t^e$ and $\omega_t^f$ be conditionally independent given $\mathcal{F}_{t-1}$ (they are not fully independent, since they are both affected by the history of conversations over the network). Define the full sample space and sigma-algebra to be

$$\Omega \times \Psi := \big(\{U,V\}^{\mathcal{E}}\big)^{\mathbb{N}} \times \mathcal{E}^{\mathbb{N}}$$
$$\mathcal{F} := \sigma\big(\cup_{t\in\mathbb{N}} \mathcal{F}_t\big).$$

Using the notation established above, one has the following update rule for $g$ and $u$:

$$g^i_{t+1} - g^i_t := \sum_{e\rightarrowtail i} S^e_{t+1}$$
$$u^i_{t+1} - u^i_t := \sum_{e\rightarrowtail i} S^e_{t+1}\omega^e_{t+1},$$

where $e \rightarrowtail i$ means that edge $e$ is incident to vertex $i$, and we are summing over all such edges.

Decompose $\omega^e_t$ into a $\mathcal{F}_{t-1}$-measurable random variable and a mean-0 $\sigma(\mathcal{F}_{t-1},\mathcal{H}_t)$-measurable fluctuation:

$$\omega^e_t = p^e_{t-1} + \tilde{w}^e_t,$$

so that

$$\tilde{w}^e_t = \begin{cases} 1 - p^e_{t-1} & \text{with probability } p^e_{t-1} \\ -p^e_{t-1} & \text{with probability } 1 - p^e_{t-1}. \end{cases}$$

Let $w^i_t := \sum_{e\rightarrowtail i} \tilde{w}^e_t S^e_t$. From here forward, the notation $\mathbb{E}_t[Z] := \mathbb{E}[Z|\mathcal{F}_t]$ will be used to represent conditional expectation with respect to the sigma-algebra $\mathcal{F}_t$.

Note that

$$\begin{aligned}
\mathbb{E}_{t-1}[w_t^i] &= \sum_{e \rightarrowtail i} \mathbb{E}[\tilde{w}_t^e S_t^e | \mathcal{F}_{t-1}] \\
&= \sum_{e \rightarrowtail i} \mathbb{E}\Big[\mathbb{E}[\tilde{w}_t^e S_t^e | \sigma\big(\mathcal{F}_{t-1} \cup \mathcal{H}_t\big)] | \mathcal{F}_{t-1}\Big] \\
&= \sum_{e \rightarrowtail i} \mathbb{E}\Big[\tilde{w}_t^e \mathbb{E}[S_t^e | \sigma\big(\mathcal{F}_{t-1} \cup \mathcal{H}_t\big)] | \mathcal{F}_{t-1}\Big] \\
&= \frac{1}{|\mathcal{E}|} \sum_{e \rightarrowtail i} \mathbb{E}\Big[\tilde{w}_t^e | \mathcal{F}_{t-1}\Big] \\
&= 0,
\end{aligned}$$

where in the second, third, and fourth equalities I've used the tower property, 'taken out what was known', and used that $S_t^e \perp \sigma(\mathcal{F}_{t-1} \cup \mathcal{H}_t)$, respectively.

For later convenience, here is a consolidated list of definitions of important quantities, and the earliest sigma-algebra $\mathcal{F}$ with respect to which they are measurable:

**Definition 15.**

- ***Total weight***: $g_t^i \in m\mathcal{F}_t$, $\gamma_t^i = \frac{1}{g_t^i}$
- ***Weight on opinion*** $U$: $u_t^i \in m\mathcal{F}_t$
- ***Proportion of weight on*** $U$: $x_t^i = \frac{u_t^i}{g_t^i} \in m\mathcal{F}_t$
- ***Initial conditions***: $u_0^i, g_0^i > 0$
- ***Mutual weight on*** $U$: $p_t^e = \frac{u_t^i + u_t^j}{g_t^i + g_t^j} \in m\mathcal{F}_t$, *where* $e = (i, j)$
- ***Mean-0 fluctuation of conversation result***: $\tilde{w}_t^e \in m\mathcal{F}_t$
- ***Result of conversation***: $\omega_t^e = p_{t-1}^e + \tilde{w}_t^e \in m\mathcal{F}_t$
- ***Edge to play***: $\psi_t \in m\mathcal{F}_t$, $S_t^e = \mathbb{1}_{\{\psi_t = e\}} \in m\mathcal{F}_t$

I now define a Hadamard (elementwise) product between a vector and a matrix. Unless otherwise noted, the symbol $\|\cdot\|$ will refer to the Euclidean norm for vectors, and the operator norm between Euclidean vector spaces for matrices. This convention is carried through the end of the chapter.

**Definition 16.** *The* ***left-Hadamard product*** *between an $m$-dimensional row vector $b$ and $(m \times n)$ matrix $A$ is a $(m \times n)$ matrix with entries given as follows:*

$$(b \circ_L A)^{ij} := b^i A^{ij}.$$

*Similarly, the* ***right-Hadamard product*** *between an $n$-dimensional column vector and $(m \times n)$ matrix $A$ is an $(m \times n)$ matrix with entries as follows:*

$$(A \circ_R b)^{ij} = A^{ij} b^j.$$

Subscripts $L$ and $R$ will be omitted when it is clear from the context what is meant. It can readily be shown that the Euclidean norms are sub-multiplicative with respect to the right-Hadamard product. For an $m \times n$ matrix $A$ and an $n$-column vector $b$,

$$\|A \circ b\| \leq \|A\| \|b\|.$$

and similarly for left-products. Another important property of Hadamard multiplication is its associativity with matrix multiplication. For an $(m \times n)$ matrix $A_1$, an $n \times p$ matrix $A_2$, and an $n-$column vector $b$,

$$A_1(b^T \circ_L A_2) = (A_1 \circ_R b) A_2.$$

### 4.2.2 Deriving the Stochastic Heat Equation

Fix an arbitrary vertex $i \in V$. Consider the quantity $u^i_{t+1} - u^i_t$, which represents the increase in the propensity of vertex $i$ to play move $u$ after timestep $t+1$.

$$u^i_{t+1} - u^i_t = \sum_{e \rightarrowtail i} S^e_{t+1} \omega^e_{t+1} = \sum_{e \rightarrowtail i} S^e_{t+1} \Big( p^e_t + \tilde{w}^e_{t+1} \Big)$$

Use the equation above to write down the change in $x^i$ between timesteps $t$ and $t+1$:

$$\begin{aligned} x^i_{t+1} - x^i_t &= \frac{u^i_{t+1}}{g^i_{t+1}} - \frac{u^i_t}{g^i_t} = \frac{1}{g^i_{t+1}} \Big( u^i_{t+1} - u^i_t - \frac{g^i_{t+1} - g^i_t}{g^i_t} u^i_t \Big) \\ &= \frac{1}{g^i_{t+1}} \Big( \sum_{e \rightarrowtail i} S^e_{t+1} \Big( p^e_t + \tilde{w}^e_{t+1} \Big) - (g^i_{t+1} - g^i_t) x^i_t \Big) \\ &= \frac{w^i_{t+1}}{g^i_{t+1}} + \frac{1}{g^i_{t+1}} \Big( \Big[ \sum_{j \sim i} S^{ij}_{t+1} \frac{x^i_t g^i_t + g^j_t x^j_t}{g^i_t + g^j_t} \Big] - (g^i_{t+1} - g^i_t) x^i_t \Big) \\ &= \frac{1}{g^i_{t+1}} \big( w^i_{t+1} + \big( L_t \vec{x}_t \big)^i \big), \end{aligned}$$

where $L_t$ is defined as follows:

**Definition 17.** *The **diffusion matrix** $L_t \in m\mathcal{F}_{t+1}$ is a $|\mathcal{V}| \times |\mathcal{V}|$ matrix with entries:*

$$L^{ij}_t = \begin{cases} S^{ij}_{t+1} \frac{g^j_t}{g^i_t + g^j_t} & i \neq j, i \sim j \\ -\sum_{j \sim i} S^{ij}_{t+1} \frac{g^j_t}{g^i_t + g^j_t} & i = j \\ 0 & \textit{else.} \end{cases}$$

*Also define $\Lambda_t \in m\mathcal{F}_{t+1}$ to be a $|\mathcal{V}| \times |\mathcal{V}|$ matrix as follows:*

$$\Lambda_t = I + \gamma^T_{t+1} \circ L_t.$$

Note that $L_t$ will have exactly four non-zero entries, and takes the following form:

$$\begin{bmatrix} 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ \vdots & \ddots & \cdots & \cdots & \cdots & \cdots & \vdots \\ \vdots & \cdots & -a & \cdots & a & \cdots & \vdots \\ \vdots & \cdots & \cdots & \ddots & \cdots & \cdots & \vdots \\ \vdots & \cdots & b & \cdots & -b & \cdots & \vdots \\ \vdots & \cdots & \cdots & \cdots & \cdots & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \end{bmatrix}$$

for some $a, b > 0$.

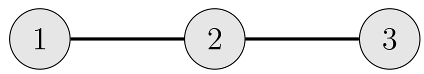
Although $L_t$ is sparse, its expectation given the previous timestep, $\mathbb{E}_t[L_t]$, is worthy of mention. It represents the aggregate effects after many rounds of conversations:

$$\mathbb{E}_t[L_t]^{ij} = \begin{cases} \frac{1}{\mathcal{E}} \frac{g_t^j}{g_t^i + g_t^j} & i \neq j, i \sim j \\ -\frac{1}{\mathcal{E}} \sum_{j \sim i} \frac{g_t^j}{g_t^i + g_t^j} & i = j \\ 0 & \text{else.} \end{cases}$$

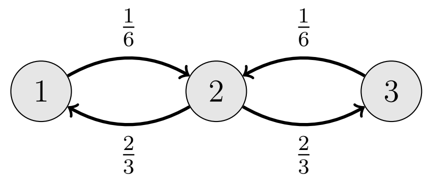
Also note that each $g_t^i - g_0^i$ is a binomial random variable with mean equal to $t\frac{d_i}{E}$, where $d_i$ is the degree of vertex $i$. One can therefore expect the leading order terms of $\mathbb{E}_t[L_t]$ to look like $\frac{1}{\mathcal{E}}$ times the following *influence matrix*, a graph dependent constant, defined below.

**Definition 18.** *The* ***influence matrix*** $L$ *is a* $|\mathcal{V}| \times |\mathcal{V}|$ *matrix with entries:*

$$L^{ij} = \begin{cases} \frac{d_j/d_i}{d_i + d_j} & i \neq j, i \sim j \\ -\sum_{j \sim i} \frac{d_j/d_i}{d_i + d_j} & i = j \\ 0 & \textit{else.} \end{cases}$$




(a) $G$




(b) $\mathcal{I}(G)$

Figure 4.1: A graph $G$ and the directed, weighted graph $\mathcal{I}(G)$ associated to $G$

*Also define $A_t$ to be a $|\mathcal{V}| \times |\mathcal{V}|$ matrix as follows:*

$$A_t = I + \frac{1}{t}L$$

The influence matrix corresponds to the graph Laplacian matrix for the weighted, directed graph $\mathcal{I}(G) := \big(\mathcal{V}, \mathcal{I}(\mathcal{E})\big)$, where $(i,j) \in \mathcal{E} \iff (i,j),(j,i) \in \mathcal{I}(\mathcal{E})$, and the edge weight from $j$ to $i$ is defined to be $L^{ij}$ (see Figure 4.1). Note that edge weights from $j$ to $i$ are high when $d_j$ is large relative to $d_i$. One can therefore think of $j$ as having more 'influence' than $i$ in this case.

With these definitions in place, I present the Stochastic Heat Equation (abbreviated SHE), derived above:

**Proposition 9** (Stochastic Heat Equation)**.** *The behavior of $\vec{x}_t$ is governed by the following Stochastic Heat Equation (SHE):*

$$\partial_t x_t = \gamma_t^T \circ (L_{t-1}\vec{x}_{t-1} + W_t),$$

*with solution*

$$\vec{x}_t = \sum_{j=0}^{t} [\Pi_{k=j}^{t-1}\Lambda_k](\gamma_j^T \circ W_j)$$

*where $\partial_t \vec{x}_t := \vec{x}_t - \vec{x}_{t-1}$ and $\vec{W}_0 := \vec{u}_0$.*

Throughout the chapter, define Π- products of matrices to have older matrices to the right, for example:

$$\Pi_{k=1}^{t}\Lambda_k = \Lambda_k\Lambda_{k-1}\cdots\Lambda_2\Lambda_1.$$

As intuition may suggest, the steady-state solution to the above heat equation is consensus: all $x_t^i$ will converge to the same (random) constant. At the heart of this idea is the Perron-Frobenius Theorem, which says that the eigenvector $\vec{1}$ of $I + L$ which represents consensus has strictly dominant eigenvalue 1. Below, the Perron-Frobenius Theorem for nonnegative matrices (Lemma 4.2.1) is stated, along with another necessary technical ingredient (Lemma 4.2.2).

**Lemma 4.2.1.** *[31, 32] Let $M$ be a square, nonnegative, irreducible, primitive matrix (i.e., there exists $k > 0$ such that $M^k > 0$ elementwise) with spectral radius $\rho$. Then the following hold:*

- *$\rho$ is an algebraically simple eigenvalue of $M$, and the corresponding normalized eigenvector $\vec{v}$ is unique and positive*
- *Any nonnegative eigenvecor of $M$ is a multiple of $\vec{v}$*
- *All other eigenvalues of $M$ have absolute value strictly smaller than $\rho$*

**Lemma 4.2.2.** *[33] Let $M$ be an $n \times n$ matrix, and define $\Gamma(M)$ to be a digraph with vertex set $\mathcal{V} = \{1, ..., n\}$ and directed edge set $\mathcal{E} = \{(i, j) : M_{ij} \neq 0\}$. If $\Gamma(M)$ is strongly connected, and every vertex $i$ of $\Gamma(M)$ has a self-loop, then $M$ is primitive.*

Having stated the above two ingredients, I now apply Perron-Frobenius to our system in the lemma below.

**Lemma 4.2.3.** *For all $k \geq 1$, 1 is a simple eigenvalue of $A_k := I + \frac{1}{k}L$. Furthermore, there exists $0 < \lambda < 1$ such that for all $k \geq 1$, and for all eigenvalues $\lambda^{(k)} \neq 1$ of $A_k$:*

$$|\lambda^{(k)}| \leq 1 - \frac{\lambda}{k}.$$

*Proof.* First notice that, for each row $i$ and for all times $t$, $\sum_j L^{ij} = \sum_j (\vec{\gamma}_{t+1}^T \circ L_t)^{ij} = 0$. From this it immediately follows that 0 is an eigenvalue of both matrices with corresponding right-eigenvector

$$\vec{1} := \begin{bmatrix} 1 \\ . \\ . \\ . \\ 1 \end{bmatrix},$$

and thus that $\vec{1}$ is a right-eigenvector of $A_k$ with eigenvalue 1. Next, label the eigenvalues $\mu_i$ of $L$ such that $\mu_1 = 0$. Notice that for any $k \geq 1$, the eigenvalues $\lambda_i^{(k)}$ of $A_k$ are given by $\lambda_i^{(k)} = 1 + \frac{\mu_i}{k}$, numbered such that $\lambda_1^{(k)} = 1$ for all $k$. It remains to show that 1 is a simple eigenvalue of $A_k$ and the bound given above.

One can invoke the Perron-Frobenius Theorem for irreducible non-negative matrices on $A_k := I + \frac{1}{k}L$. $A_k$ is nonnegative since it is clear that all off-diagonal elements are nonnegative, and for all $i$,

$$L^{ii} = -\sum_{j \sim i} \frac{\frac{\deg(j)}{\deg(i)}}{\deg(i) + \deg(j)} > -\sum_{j \sim i} \frac{1}{\deg(i)} = -1.$$

This gives that $A_k^{ii} > 0$ and thus that $A_k$ is nonnegative.

In order to show that $A_k$ is irreducible, consider its associated weighted digraph $\Gamma(A_k)$, which has vertex set $V$, a complete edge set $V \times V$, and weights $W : V \times V \rightarrow$

$\mathbb{R}_{\geq 0}$. By the definition of $L$, we have that for all $i \sim j$ in the original graph, there are edges with non-zero weights flowing from $i$ to $j$ and from $j$ to $i$. Since the original graph $V$ is connected, this implies that the weighted digraph associated to $A_k$ is strongly connected, giving that $A_k$ is irreducible. Also note that since the diagonal elements of $A_k$ are all strictly positive, each vertex in $\Gamma(A_k)$ has a self-loop, and thus $A_k$ is primitive by Lemma 4.2.2. Thus $A_k$ satisfies the assumptions of the Lemma 4.2.1. Since the eigenvector $\vec{1}$ has components that are all positive, Perron-Frobenius gives that associated eigenvalue 1 of $A_k$ is simple, that the spectral radius of $A_k$ is 1, and that all other eigenvalues of $A_k$ have modulus strictly less than 1.

Let $\lambda$ represent the spectral gap of $A_1$ (unless the spectral gap is 1, in which case I arbitrarily set $\lambda = \frac{1}{2}$):

$$\lambda := \begin{cases} 1 - \max_{i>1} |\lambda_i^{(1)}| & \text{if } \max_{i>1} |\lambda_i^{(1)}| \neq 0 \\ \frac{1}{2} & \text{else} \end{cases}$$

This gives that, for all $k \geq 1$ and $i > 1$,

$$|\lambda_i^{(k)}| = |1 + \frac{\mu_i}{k}| = \frac{1}{k}|\mu_i + k| \leq \frac{1}{k}|\lambda_i^{(1)}| + \frac{k-1}{k} \leq \frac{k-\lambda}{k} = 1 - \frac{\lambda}{k}.$$

□

From here forward, let $\lambda$ represent the number guaranteed by the above lemma. The next lemma shows that $L$ is similar to a symmetric matrix and hence is diagonalizable, which simplifies the long-time analysis involving products of $L$.

**Lemma 4.2.4.** *$L$ is diagonalizable, and can be written $L = PDP^{-1}$, where the first column of $P$ is $\vec{1}$, and $D_{11} = 0$.*

*Proof.* Let $E$ be the diagonal matrix with diagonal elements equal to the degree of each vertex:

$$E^{ij} = \begin{cases} d_i & i = j \\ 0 & i \neq j \end{cases}.$$

Note that $E$ has strictly positive entries on the diagonal and is therefore invertible with

$$(E^{-1})^{ij} = \begin{cases} \frac{1}{d_i} & i = j \\ 0 & i \neq j \end{cases}.$$

Then note that $ELE^{-1}$ is symmetric, because

$$\begin{aligned} \left(ELE^{-1}\right)^{ij} &= \sum_{k,\ell} E^{ik} L^{k\ell} (E^{-1})^{\ell j} \\ &= \frac{d_i}{d_j} L^{ij}. \end{aligned}$$

Now, by definition of $L$: if $i \nsim j$, then $(ELE^{-1})^{ij} = (ELE^{-1})^{ji} = 0$, and if $i \sim j$, then $(ELE^{-1})^{ij} = \frac{1}{d_i + d_j} = (ELE^{-1})^{ji}$. Thus $(ELE^{-1})$ is symmetric and therefore diagonalizable. Since $L$ is similar to a diagonalizable matrix, it is itself diagonalizable. □

From here forward, fix $P$ and $D$ as given in Lemma 4.2.4. The above two lemmas make a powerful combination, in the following sense. Note that the solution to the SHE (Proposition 9) involves a product of the $\Lambda$ matrices: $\Pi_{k=j}^{t-1} \Lambda_k$. In the discussion below, I show that this large product can be approximated by the following product of constant matrices: $\Pi_{k=j}^{t-1} A_k$, which in turn is similar to a product of diagonal matrices: $\Pi_{k=j}^{t-1} D_k$, where $D_k = P^{-1} A_k P$. Now, while the first entry of each of the $D_k$ is 1 (corresponding to consensus), the other entries are bounded by $1 - \frac{\lambda}{k}$ (due to Lemma 4.2.3). The last ingredient of this section is an application of the theory of gamma functions, due to Gautschi, which shows that while these eigenvalues approach 1 from below as $k \to \infty$, the approach is slow enough for their product to approach 0.

**Lemma 4.2.5.** *[34] For* $0 < s < 1$*:*

$$x^{1-s} < \frac{\Gamma(x+1)}{\Gamma(x+s)} < (x+1)^{1-s}$$

**Lemma 4.2.6.** *For all* $1 \leq j \leq t$ *and for* $0 < \lambda < 1$,

$$\Big(\frac{j-1}{t+1}\Big)^{\lambda} \leq \Pi_{k=j}^{t}(1-\frac{\lambda}{k}) \leq \Big(\frac{j}{t}\Big)^{\lambda}$$

*Proof.* Write

$$\Pi_{k=j}^{t}(1-\frac{\lambda}{k}) = \frac{\Pi_{k=j}^{t}(k-\lambda)}{\Pi_{k=j}^{t}(k)} = \frac{\Gamma(j)}{\Gamma(j-\lambda)}\frac{\Gamma(t+1-\lambda)}{\Gamma(t+1)},$$

and apply Gautschi's inequality (Lemma 4.2.5). □

## 4.3 Convergence of the Consensus Coordinate

Let $\vec{p}$ be the first row of $P^{-1}$, i.e. the left-eigenvector of $L$ with eigenvalue 0, and let $a_t := p \cdot x_t$ be the coordinate corresponding to $\vec{p}$ in the eigenbasis expansion of $L$ (where the eigenbasis is given by the columns of $P$). The goal of this section is to show the following lemma.

**Lemma 4.3.1.** *There exists a random constant* $a_\infty$ *such that* $a_t \to a_\infty$ *in* $\mathcal{L}^2$.

Decompose $a_t$ as follows:

$$\begin{aligned}
a_t &= a_0 + \sum_{j=0}^{t-1}(a_{j+1} - a_j) \\
&= a_0 + \vec{p} \cdot \sum_{j=0}^{t-1}(\vec{x}_{j+1} - \vec{x}_j) \\
&= a_0 + \vec{p} \cdot \sum_{j=0}^{t-1} \vec{\gamma}_{j+1}^T \circ L_j \vec{x}_j + \gamma_{j+1}^T \circ \vec{w}_{j+1} \\
&= a_0 + \vec{p} \cdot \sum_{j=0}^{t-1}(\vec{\gamma}_{j+1}^T \circ L_j - \frac{1}{j+1}L)\vec{x}_j + \frac{1}{j+1}L\vec{x}_j + \vec{\gamma}_{j+1}^T \circ \vec{w}_{j+1} \\
&= a_0 + \vec{p} \cdot \sum_{j=0}^{t-1}(\vec{\gamma}_{j+1}^T \circ L_j - \frac{1}{j+1}L)\vec{x}_j + \vec{\gamma}_{j+1}^T \circ \vec{w}_{j+1}
\end{aligned}$$

where I've used the SHE update in the third line, and the fact that $\vec{p}$ is a left 0-eigenvector of $L$ in the fifth.

Now, there are two main differences between the dampened diffusion matrix $\vec{\gamma}_{j+1}^T \circ L_j$ and the dampened influence matrix $\frac{1}{j+1}L$. The first is that the diffusion matrix only involves a random edge, while the influence matrix considers all edges. The second is that the $g_t$ are random functions of the $\psi$ variables, while $L$ is a constant. Separate out these two differences by adding and subtracting $\mathbb{E}_j[\vec{\gamma}_{j+1}^T \circ L_j]$:

$$a_t = a_0 + m_t + s_t,$$

where

$$\begin{aligned}
m_t &:= \vec{p} \cdot \sum_{j=0}^{t-1}(\gamma_{j+1}^T \circ L_j - \mathbb{E}_j[\gamma_{j+1}^T \circ L_j])\vec{x}_j + \vec{\gamma}_{j+1}^T \circ \vec{w}_{j+1} \\
s_t &:= \vec{p} \cdot \sum_{j=0}^{t-1} \Delta_j \vec{x}_j \\
\Delta_j &:= \mathbb{E}_j[\vec{\gamma}_{j+1}^T \circ L_j] - \frac{1}{j+1}L
\end{aligned}$$

I consider each of $s_t$ (which stands for 'small') and $m_t$ (which stands for martingale) separately; in order to show Lemma 4.3.1, it suffices to show that each of $s_t$ and $m_t$ converge in $\mathcal{L}^2$. While $s_t$ is nonzero due to the randomness of $g_t$, I show that each term is small in expectation and therefore that the sum is convergent, while $m_t$ is shown to be a martingale, on which I will invoke the martingale convergence theorem.

Before proceeding, I state a useful lemma that allows us to rigorously pass from sums to integrals:

**Lemma 4.3.2.** *Let $f(k)$ be nonnegative on $[t_1, t_2]$, non-decreasing on $[t_1, x]$ and non-increasing on $[x, t_2]$ for some $x \in [t_1, t_2]$. Then*

$$\sum_{k=t_1}^{t_2} f(k) \leq \int_{t_1}^{t_2} f(k)dk + 2f(x)$$

*Proof.* In the below, take sums with lower endpoint strictly greater than upper endpoint to be 0. We have:

$$\begin{aligned}\sum_{k=t_1}^{t_2} f(k) &\leq 2f(x) + \sum_{k=t_1}^{\lfloor x \rfloor - 1} f(k) + \sum_{k=\lceil x \rceil + 1}^{t_2} f(k) \\ &\leq \int_{t_1}^{\lfloor x \rfloor} f(k) + \int_{\lceil x \rceil}^{t_2} f(k) + 2f(x) \\ &\leq \int_{t_1}^{t_2} f(k)dk + 2f(x).\end{aligned}$$

□

### 4.3.1 $s_t$ : Fluctuations of $\vec{g}_t$

Recall that

$$\mathbb{E}_t[\vec{\gamma}_{t+1}^T \circ L_t]^{ij} = \begin{cases} \frac{1}{E(g_t^i+1)} \frac{g_t^j}{g_t^i+g_t^j} & i \sim j \\ -\sum_{k\sim i} \frac{1}{E(g_t^i+1)} \frac{g_t^k}{g_t^i+g_t^k} & i = j \\ 0 & i \nsim j. \end{cases}$$

Now, the random variable $g_t^i$ is equal to $g_0^i$ plus a binomial random variable resulting from $t$ trials with probability $\frac{d_i}{E}$ of success for each trial. Thus one can expect each $g_t^i$ to grow like $\frac{d_i}{E}t$, with standard deviation proportional to $\sqrt{t}$. This gives the heuristic that $\mathbb{E}[\|\Delta_j\|] = O(t^{-\frac{3}{2}})$. This idea is supported by the following concentration inequality for the binomial random variable, which can be used to show that the probability of $g_t^i - g_0^i$ deviating from its mean by $t^{\frac{1}{2}+\epsilon}$ is exponentially small in $t$.

**Lemma 4.3.3.** *[35] Let $B \sim Bin(n,p)$ be a binomial random variable, and let $a > 0$. Then*

$$\mathbb{P}(|B - np| > a) < 2\exp(-\frac{2a^2}{n}).$$

The above statement serves as the main tool for showing that $\Delta_j$ is indeed small. In particular, I prove the following lemma, which will be used to show that $s_t$ converges in $\mathcal{L}^2$. From here forward, I use the notation $f(t) \lesssim g(t)$ to mean that there exists a constant $c$, independent of $t$, such that $f(t) \leq cg(t)$.

**Lemma 4.3.4.** *For sufficiently large $s, t$:*

$$\mathbb{E}[\|\Delta_s\|\|\Delta_t\|] \lesssim \frac{1}{(st)^{\frac{5}{4}}} + \frac{s}{t^{\frac{5}{4}}}\exp(-\frac{2}{|\mathcal{E}|^2}s^{1/2}) + st\exp(-\frac{2}{|\mathcal{E}|^2}t^{1/2})$$

*Proof.* Define

$$C_t := \{\big||\mathcal{E}|\frac{g_t^i - g_0^i}{t} - d_i\big| \leq \frac{1}{t^{\frac{1}{4}}} \ : \forall i \in \mathcal{V}\}$$
$$\delta_t^i := |\mathcal{E}|\frac{g_t^i - g_0^i}{t} + |\mathcal{E}|\frac{g_0^i}{t} - d_i,$$

and note that $|\mathcal{E}|\frac{g_0^i}{t} - d_i \leq \delta_t^i \leq |\mathcal{E}|(g_0^i + 1) - d_i$ almost surely. Note also that using $a = \frac{t^{3/4}}{|\mathcal{E}|}$ in Lemma 4.3.3 produces

$$\mathbb{P}(C_t^c) < 2|\mathcal{V}| \exp(-\frac{2}{|\mathcal{E}|^2} t^{\frac{1}{2}}).$$

Now, for fixed $i \sim j$, we have that:

$$\begin{aligned}
(t+1)\Delta_t^{ij} &= \frac{t+1}{|\mathcal{E}|(g_t^i + 1)} \frac{g_t^j}{g_t^i + g_t^j} - \frac{\frac{d_j}{d_i}}{d_i + d_j} \\
&= \frac{1 + \frac{1}{t}}{(\delta_t^i + d_i + \frac{|\mathcal{E}|}{t})} \frac{\delta_t^j + d_j}{\delta_t^j + d_j + \delta_t^i + d_i} - \frac{d_j}{d_i(d_i + d_j)} \\
&= \frac{(1 + \frac{1}{t})(d_i + d_j)(\delta_t^j + d_j)d_i - d_j(\delta_t^i + d_i + \frac{|\mathcal{E}|}{t})(\delta_t^j + d_j + \delta_t^i + d_i)}{d_i(d_i + d_j)(\delta_t^i + d_i + \frac{|\mathcal{E}|}{t})(\delta_t^j + d_j + \delta_t^i + d_i)}
\end{aligned}$$

Almost surely:

$$(t+1)|\Delta_t^{ij}| \leq \frac{c_1|\delta_t^i| + c_2|\delta_t^j| + c_3\frac{1}{t}}{(d_i + \delta_t^i)(d_j + \delta_t^j)}$$

for some $t$-independent constants $c_1, c_2, c_3$. Further, on $C_t$, $|\delta_t^i| \lesssim \frac{1}{t^{1/4}}$ for all $i$. So, on $C_t$ for sufficiently large $t$,

$$|\Delta_t^{ij}| \lesssim \frac{1}{t^{5/4}}.$$

It's also easy to see that, almost surely (in particular, on $C_t^c$),

$$|\Delta_t^{ij}| \lesssim t$$

Further, since $\Delta_t$ is row-stochastic for all $t$, we can drop the requirement that $i \sim j$ for the above two inequalities on $|\Delta_t^{ij}|$ (perhaps at the cost of a larger constant).

Now, for $s \neq t$,

$$\begin{aligned}\mathbb{E}[\|\Delta_s\|_{\max}\|\Delta_t\|_{\max}] &= \mathbb{E}[\max_{i,j,k,\ell}|\Delta_s^{ij}\Delta_t^{k\ell}||C_s\cap C_t]\mathbb{P}(C_s\cap C_t)\\ &+ \mathbb{E}[\max_{i,j,k,\ell}|\Delta_s^{ij}\Delta_t^{k\ell}||C_s^c\cap C_t]\mathbb{P}(C_s^c\cap C_t) + \mathbb{E}[\max_{i,j,k,\ell}|\Delta_s^{ij}\Delta_t^{k\ell}||C_t^c]\mathbb{P}(C_t^c)\\ &\lesssim \frac{1}{(st)^{\frac{5}{4}}} + \frac{s}{t^{\frac{5}{4}}}\exp(-\frac{2}{\mathcal{E}^2}s^{1/2}) + st\exp(-\frac{2}{\mathcal{E}^2}t^{1/2}),\end{aligned}$$

where $\|A\|_{\max} := \max_{i,j}|A^{ij}|$. Finally, note that $\|\Delta_t\| \lesssim \|\Delta_t\|_{\max}$. This gives the desired result for $s \neq t$.

When $s = t$, we have:

$$\begin{aligned}\mathbb{E}[\max_{i,j}|\Delta_t^{ij}|^2] &\leq \mathbb{E}[\max_{i,j}|\Delta_t^{ij}|^2||C_t]\mathbb{P}(C_t) + \mathbb{E}[\max_{i,j}|\Delta_t^{ij}|^2|C_t^c]\mathbb{P}(C_t^c)\\ &\lesssim \frac{1}{t^{\frac{5}{2}}} + t^2\exp(-\frac{2}{|\mathcal{E}|}t^{1/2}),\end{aligned}$$

and again I use that $\|\Delta_t\| \lesssim \|\Delta_t\|_{\max}$. $\square$

This allows us to prove the desired convergence of $s_t$.

**Lemma 4.3.5.** *$s_t$ converges in $\mathcal{L}^2$.*

*Proof.* It suffices to show Cauchy in $\mathcal{L}^2$, i.e. that for any $\epsilon > 0$, there exists $T$ such that for all $t_1, t_2 > T$, $\mathbb{E}[(s_{t_1} - s_{t_2})^2] \leq \epsilon$. Note that

$$(s_{t_2} - s_{t_1})^2 \lesssim \sum_{j,k=t_1}^{t_2-1}\|\Delta_j\|_2\|\Delta_k\|_2 \lesssim \sum_{j=t_1}^{t_2-1}\sum_{k=t_1}^{j}\|\Delta_j\|_2\|\Delta_k\|_2$$

Taking expectations and using the lemma,

$$\mathbb{E}[(s_{t_2} - s_{t_1})^2] \lesssim \sum_{j=t_1}^{t_2-1}\sum_{k=t_1}^{j}\frac{1}{(jk)^{\frac{5}{4}}} + \frac{k}{j^{\frac{5}{4}}}\exp(-\frac{2}{|\mathcal{E}|^2}k^{1/2}) + jk\exp(-\frac{2}{|\mathcal{E}|^2}j^{1/2}).$$

It's now clear, for example from Lemma 4.3.2, that the lemma follows. $\square$

### 4.3.2 $m_t$: Martingale Convergence

The goal of the subsection is to prove that $m_t$ converges. I begin by stating the $\mathcal{L}^2$ martingale convergence theorem without proof.

**Lemma 4.3.6.** *[36] Let $y_t$ be a martingale with $y_t \in \mathcal{L}^2$ for all $t$. Further assume that $\sup_t \|y_t\|_{\mathcal{L}^2} < \infty$. Then $y_t$ converges in $\mathcal{L}^2$.*

**Lemma 4.3.7.** *$m_t$ converges in $\mathcal{L}^2$.*

*Proof.* I first show that $m_t$ is a martingale. It is clearly an adpated process. Next, consider

$$\mathbb{E}_{t-1}[m_t - m_{t-1}] = \vec{p} \cdot \mathbb{E}_{t-1}\big[\vec{\gamma}_t^T \circ L_{t-1} - \mathbb{E}_{t-1}[\vec{\gamma}_t^T \circ L_{t-1}]\big]\vec{x}_{t-1} + \vec{p} \cdot \mathbb{E}_{t-1}[\vec{\gamma}_t^T \circ \vec{w}_t].$$

The first term is clearly 0. Next note that:

$$\begin{aligned}
\mathbb{E}_{t-1}[(\vec{\gamma}_t^T \circ \vec{w}_t)^i] &= \sum_{e \rightarrowtail i} \mathbb{E}[\frac{1}{g_t^i} \tilde{w}_t^e S_t^e | \mathcal{F}_{t-1}] \\
&= \sum_{e \rightarrowtail i} \mathbb{E}[\frac{1}{g_t^i} S_t^e \mathbb{E}[\tilde{W}_t^e | \sigma(\mathcal{F}_{t-1}, \sigma(\psi_t))] | \mathcal{F}_{t-1}] \\
&= 0.
\end{aligned}$$

Lastly, note that

$$m_t = a_t - a_0 - s_t,$$

so that

$$\|m_t\|_{\mathcal{L}^2} \leq \|a_t\|_{\mathcal{L}^2} + \|a_0\|_{\mathcal{L}^2} + \|s_t\|_{\mathcal{L}^2}.$$

$a_0$ is a constant, $a_t$ is a.s. bounded by virtue of $0 \leq x_t \leq 1$, and $\|s_t\|_{\mathcal{L}^2}$ is bounded since $s_t$ converges in $\mathcal{L}^2$. Thus $m_t$ is bounded in $\mathcal{L}^2$, proving the theorem. □

## 4.4 Decay of Disagreement

The goal of this section is to show that the component of $\vec{x}_t$ corresponding to any differing opinions converges to 0. Let $\vec{z}_t := \vec{x}_t - a_t\vec{1}$ represent this component of the opinion vector. I would like to show that

**Lemma 4.4.1.**

$$\mathbb{E}[\|\vec{z}_t\|^2] \to 0$$

### 4.4.1 Preliminary Discussion

I develop my approach to a proof as follows. With $Q := \text{diag}(0, 1, ..., 1)$, it's clear that $\vec{z}_t = PQP^{-1}\vec{x}_t$, where $P$ is the matrix of eigenvectors of $L$. Further, using the sum-product solution of the SHE from Proposition 9, we have that

$$\vec{z}_t = \sum_{j=0}^{t} PQP^{-1}[\Pi_{k=j}^{t-1}\Lambda_k](\vec{\gamma}_j^T \circ \vec{w}_j).$$

The intuition for why $\vec{z}_t$ is small is as follows: at each past timestep $j \leq t$, a random 'blip' $\vec{\gamma}_j^T \circ \vec{w}_j$ was introduced. In subsequent time steps $k \geq j$, this blip was smoothed by repeated application of the $\Lambda_k$ matrices. Now, as argued in the previous section (see Lemma 4.3.4), $\mathbb{E}[\Lambda_k] \approx A_k$. Using $PQP^{-1}$ to project out the Perron-Frobenius eigenvalue 1 of $A_k$ (corresponding to consensus), we get eigenvalues whose products decay sufficiently rapidly. So, sufficiently old blips are dampened by products of small eigenvalues with many factors, while newer blips will be small because the vector norm of $\vec{\gamma}_j$ is expected to decrease as $j$ increases.

An issue with the above heuristic, however, is that random draws of $\vec{\gamma}_{k+1}^T \circ L_k$ are not close to $\frac{1}{k}L$ (even though they approximately agree in expectation). This is circumvented by noting that the $\vec{\gamma}_{k+1}^T \circ L_k$ are Cesàro-summable with limit proportional to $L$: I expand the product $\Pi_{k=j}^{t-1}\Lambda_k = \Pi_{k=j}^{t-1}(I + \vec{\gamma}_{k+1}^T \circ L_k)$, show that the leading order terms (i.e. those linear in the dampened diffusion matrix) are proportional to $L$ due to a law of large numbers effect, and show that the lower order terms decay sufficiently rapidly because they have many factors of $\vec{\gamma}$.

More precisely: I group the $t-j$ factors in the product $\Pi_{k=j}^{t-1}\Lambda_k$ into subgroups of size $\tau := \lceil t^{1/4} \rceil$. This $\tau$ is large enough for the law of large numbers to kick in (allowing us to replace the group's average of the $\vec{\gamma}_{k+1}^T \circ L_k$ with a matrix proportional to $L$), but small enough so that there are enough factors of $L$ for the decay of the product of the non-dominant eigenvalues to be severe. Note that $j$ needs to be sufficiently small so that we have enough factors of $\Lambda$ to work with. With this in mind, I will separate the sum defining $z_t$ into $j \leq j_0$ and $j > j_0$ (for a value of $j_0$ to be specified later). The $j \leq j_0$ sum witnesses $PQP^{-1}[\Pi_{k=j}^{t-1}\Lambda_k]$ to have sufficiently small operator norm, while the $j > j_0$ sum is small because we expect $\vec{\gamma}_j$ to be small at such late values of $j$. This heuristic is illustrated in Figure 4.2.

For fixed $t$, define $r$ to be the remainder of $t$ divided by $\tau$, define $j_0 := r + \tau^2$, and let $H_{k,t}$ represent the aggregate effects of the $\Lambda$ factors from the $\tau$-window indexed by $k$. That is, for $1 \leq k \leq \frac{t-r}{\tau}$:

$$H_{k,t} = [\Pi_{j=r+(k-1)\tau+1}^{r+k\tau}\Lambda_j]$$

so that, for sufficiently large $t$ and $j \leq j_0$,

$$\Pi_{k=j}^{t}\Lambda_k = \left[\Pi_{k=\tau+1}^{\frac{t-r}{\tau}} H_{k,t}\right]\left[\Pi_{k=j}^{j_0}\Lambda_k\right]$$

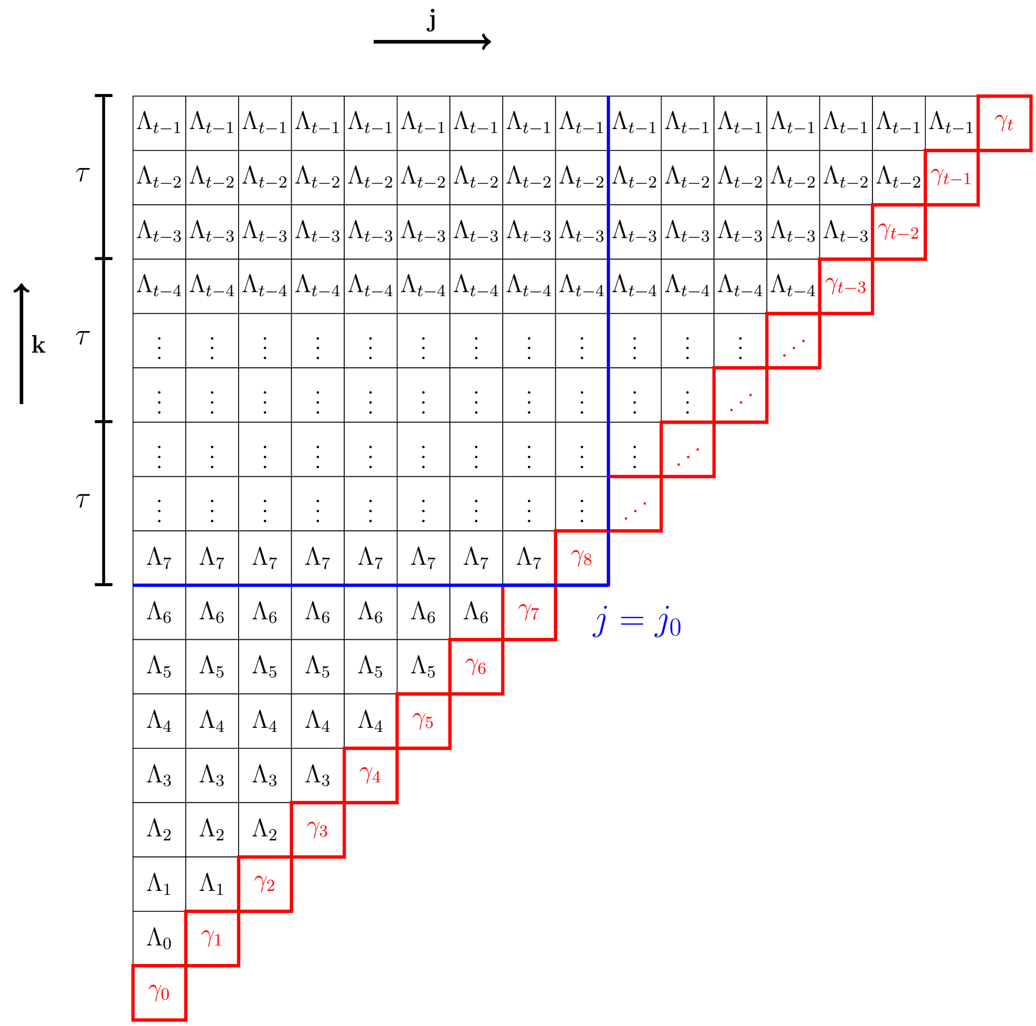


Figure 4.2: Schematic for the solution to the SHE: $\vec{x}_t = \sum_{j=0}^{t}[\Pi_{k=j}^{t-1}\Lambda_k](\vec{\gamma}_j^T \circ \vec{w}_j)$. $k$ parametrizes the factors in the product, $j$ parametrizes the terms in the sum. See the discussion above. $\gamma$ is written instead of $\gamma \circ w$ to save space.

### 4.4.2 Good and Bad Events

The above intuition only holds on 'good events' where the long-term randomness of the $\psi$ variables is close to expectation. In particular, this assumption is used when taking $\vec{\gamma}_j$ to be small for large $j$, and that the Cesàro mean of $\vec{\gamma}_{k+1}^T \circ L_k$ is roughly proportional to $L$. For the rest of the chapter we fix $\epsilon \ll \frac{1}{2}$, and for

$\tau + 1 \leq k \leq \frac{t-r}{\tau}$, define these good events as follows:

$$A_{k,t} = \left\{ |g_s^i - g_0^i - \frac{d_i}{E} k\tau| \leq (k\tau)^{\frac{1}{2}+\epsilon} \, : \, \forall i \in \mathcal{V}, \, \forall s \in \{(k-1)\tau + r + 1, ..., k\tau + r\} \right\}$$

$$B_{k,t} = \{ |( \sum_{s=(k-1)\tau+r+1}^{k\tau+r} S_s^e) - \frac{1}{E}\tau| \leq \tau^{\frac{1}{2}+\epsilon} \, : \, \forall e \in \mathcal{E} \}$$

$$E_t = \cap_{k=\tau+1}^{\frac{t-r}{\tau}} (A_{k,t} \cap B_{k,t}).$$

$A_{k,t}$ corresponds to the event that, for all $s$ in the $\tau$-window indexed by $k$, $g_s - g_0$ is close to the expectation of $g$ at the point $k\tau$ (which lies in the window). The event $B_{k,t}$ represents that, within the $\tau$ window indexed by $k$, the amount of conversations each edge hosts is close to its expectation. $E_t$ is the intersection of the $A$ and $B$ events for all windows $\tau + 1 \leq k \leq \frac{t-r}{\tau}$.

I first establish that the union of the bad events have exponentially small probability in $t$.

**Lemma 4.4.2.** *For $\epsilon < \frac{1}{2}$, there exist positive constants $c_1$ and $c_2$ such that, for sufficiently large $t$,*

$$\mathbb{P}(E_t^c) \leq c_1 \exp(-c_2 \tau^{2\epsilon}).$$

*Proof.*

$$\begin{aligned}
\mathbb{P}(E_t^c) &= \mathbb{P}\big((\cap_{k=k_{\min}}^{k_{\max}} (A_{k,t} \cap B_{k,t}))^c\big) \\
&= \mathbb{P}\big((\cup_{k=k_{\min}}^{k_{\max}} (A_{k,t}^c \cup B_{k,t}^c))\big) \\
&\leq \sum_{k=k_{\min}}^{k_{\max}} \big(\mathbb{P}(A_{k,t}^c) + \mathbb{P}(B_{k,t}^c)\big),
\end{aligned}$$

where a union bound has been invoked in the last line.

Now, for all $t \geq 0, k \geq 1$:

$$\mathbb{P}(B_{k,t}^c) \leq 2|\mathcal{E}| \exp(-2\tau^{2\epsilon}).$$

This follows directly from Lemma 4.3.3, with union bound.

Similarly, for $t$ sufficiently large and $k \geq \tau$:

$$\mathbb{P}(A_{k,t}^c) \leq 4|\mathcal{V}| \exp(-\frac{1}{4}(k\tau)^{2\epsilon}) \leq 4|\mathcal{V}| \exp(-\frac{1}{4}\tau^{4\epsilon}).$$

The proof of this claim is as follows: Note that

$$A_{k,t} = \{g_{s_0}^i - g_0^i - \frac{d_i}{\mathcal{E}}k\tau \geq -(k\tau)^{\frac{1}{2}+\epsilon} \ : \ \forall i \in \mathcal{V}\} \cap \{g_{s_1}^i - g_0^i - \frac{d_i}{\mathcal{E}}k\tau \leq (k\tau)^{\frac{1}{2}+\epsilon} \ : \ \forall i \in \mathcal{V}\}$$

where $s_0$ and $s_1$ represent the endpoints for a particular $\tau$-window:

$s_0 = (k-1)\tau + r + 1$ and $s_1 = k\tau + r$. Now, we have that

$$\begin{aligned}
&\mathbb{P}\big(\{g_{s_0}^i - g_0^i - \frac{d_i}{\mathcal{E}}k\tau \geq -(k\tau)^{\frac{1}{2}+\epsilon} \ : \ \forall i \in \mathcal{V}\}^c\big) \\
&\leq |\mathcal{V}|\mathbb{P}\big(\{g_{s_0}^i - g_0^i - \frac{d_i}{\mathcal{E}}k\tau \leq -(k\tau)^{\frac{1}{2}+\epsilon}\}\big) \\
&= |\mathcal{V}|\mathbb{P}\Big(\{g_{s_0}^i - g_0^i - \frac{d_i}{\mathcal{E}}s_0 \leq -(k\tau)^{\frac{1}{2}+\epsilon} + \frac{d_i}{\mathcal{E}}(\tau - r - 1)\}\big)\Big) \\
&\leq |\mathcal{V}|\mathbb{P}\Big(\{g_{s_0}^i - g_0^i - \frac{d_i}{\mathcal{E}}s_0 \leq -\frac{1}{2}(k\tau)^{\frac{1}{2}+\epsilon}\}\Big) \\
&\leq 2|\mathcal{V}| \exp(-\frac{1}{4}(k\tau)^{2\epsilon})
\end{aligned}$$

Similarly, for the second event,

$$\begin{aligned}
&\mathbb{P}\big(\{g_{s_1}^i - g_0^i - \frac{d_i}{E}k\tau \leq (k\tau)^{\frac{1}{2}+\epsilon} \ : \ \forall i \in \mathcal{V}\}^c\big) \\
&\leq |\mathcal{V}|\mathbb{P}\big(\{g_{s_1}^i - g_0^i - \frac{d_i}{\mathcal{E}}k\tau \geq (k\tau)^{\frac{1}{2}+\epsilon}\}\big) \\
&= |\mathcal{V}|\mathbb{P}\Big(\{g_{s_1}^i - g_0^i - \frac{d_i}{\mathcal{E}}s_1 \geq (k\tau)^{\frac{1}{2}+\epsilon} - \frac{d_i}{\mathcal{E}}r\}\big)\Big) \\
&\leq |\mathcal{V}|\mathbb{P}\Big(\{g_{s_1}^i - g_0^i - \frac{d_i}{\mathcal{E}}s_1 \geq \frac{1}{2}(k\tau)^{\frac{1}{2}+\epsilon}\}\big)\Big) \\
&\leq 2|\mathcal{V}| \exp(-\frac{1}{4}(k\tau)^{2\epsilon}).
\end{aligned}$$

This concludes the proof of the above claim. The proof is now finished by noting that

$$\big(\mathbb{P}(A_{k,t}^c) + \mathbb{P}(B_{k,t}^c)\big) \leq 4(|\mathcal{V}| + |\mathcal{E}|) \exp(-\frac{1}{4}\tau^{2\epsilon}),$$

so that, for sufficiently large $t$:

$$
\begin{aligned}
\mathbb{P}(E_t^c) &\leq \sum_{k=\tau+1}^{\frac{t-r}{\tau}} \big(\mathbb{P}(A_{k,t}^c) + \mathbb{P}(B_{k,t}^c)\big) \\
&\leq 4(|\mathcal{V}| + |\mathcal{E}|)\frac{t}{\tau}\exp(-\frac{1}{4}\tau^{2\epsilon}) \\
&\leq 4(|\mathcal{V}| + |\mathcal{E}|)\exp(-\frac{1}{8}\tau^{2\epsilon})
\end{aligned}
$$

□

### 4.4.3 Law of Large Numbers for Iterated Diffusion

I now show that, on good events, $H_{k,t}$ (representing the time-evolution over the $\tau$-window indexed by $k$) window is close to $A_k$. Begin by analyzing the leading-order terms in the product. The following lemma shows that, on good events, the dampened Laplacian matrices are Cesàro-summable, with average close to the influence matrix.

**Lemma 4.4.3.** *Fix $\epsilon \ll \frac{1}{2}$. There exists a constant $c$ such that, for $t$ sufficiently large, $k \geq \tau$ and on $A_{k,t} \cap B_{k,t}$:*

$$
\|\Big(\sum_{j=(k-1)\tau+z+1}^{k\tau+z} \vec{\gamma}_{j+1}^T \circ L_j\Big) - \frac{1}{k}L\| \leq \frac{c}{k\tau^{1/2-\epsilon}}
$$

*Proof.* Fix vertices $i \sim \ell$, and consider outcomes on $A_{k,t} \cap B_{k,t}$ only. In the below, the constant $c$ may change from line to line, but will never depend on $t$ or $k$.

$$
\begin{aligned}
\Big(\sum_{j=(k-1)\tau+z+1}^{k\tau+z} \gamma_{j+1}^{T} \circ L_j\Big)^{i\ell} &= \sum_{j=(k-1)\tau+r+1}^{k\tau+r} \frac{1}{g_{j+1}^i} S_{j+1}^{i\ell} \frac{g_j^\ell}{g_j^i + g_j^\ell} \\
&\leq \frac{1}{\frac{d_i}{|\mathcal{E}|}k\tau - (k\tau)^{\frac{1}{2}+\epsilon} + g_0^i} \frac{\frac{d_\ell}{|\mathcal{E}|}k\tau + (k\tau)^{\frac{1}{2}+\epsilon} + g_0^\ell}{\frac{d_\ell+d_i}{|\mathcal{E}|}k\tau - 2(k\tau)^{\frac{1}{2}+\epsilon} + g_0^i + g_0^\ell} \sum_{j=(k-1)\tau+r+1}^{k\tau+r} S_{j+1}^{i\ell} \\
&\leq \frac{1}{\frac{d_i}{|\mathcal{E}|}k\tau - (k\tau)^{\frac{1}{2}+\epsilon} + g_0^i} \frac{\frac{d_\ell}{|\mathcal{E}|}k\tau + (k\tau)^{\frac{1}{2}+\epsilon} + g_0^\ell}{\frac{d_\ell+d_i}{|\mathcal{E}|}k\tau - 2(k\tau)^{\frac{1}{2}+\epsilon} + g_0^i + g_0^\ell} \Big(\frac{\tau}{|\mathcal{E}|} + \tau^{\frac{1}{2}+\epsilon} + 1\Big) \\
&\leq \frac{|\mathcal{E}|}{k\tau}\Big(L^{i\ell} + \frac{c}{(k\tau)^{\frac{1}{2}-\epsilon}}\Big)\Big(\frac{\tau}{|\mathcal{E}|} + \tau^{\frac{1}{2}+\epsilon} + 1\Big) \\
&= \Big(L^{i\ell} + \frac{c}{(k\tau)^{\frac{1}{2}-\epsilon}}\Big)\Big(\frac{1}{k} + \frac{c}{k\tau^{\frac{1}{2}-\epsilon}}\Big) \\
&\leq \frac{1}{k}L^{i\ell} + \frac{c}{k\tau^{\frac{1}{2}-\epsilon}}.
\end{aligned}
$$

The opposite-direction inequality can be proven similarly, giving that

$$
\Big|\Big(\sum_{j=(k-1)\tau+r+1}^{k\tau+r} \vec{\gamma}_{j+1}^{T} \circ L_j\Big)^{i\ell} - \frac{1}{k}L^{i\ell}\Big| \leq \frac{c}{k\tau^{\frac{1}{2}-\epsilon}}.
$$

From this the lemma easily follows. □

I now use the above lemma to show that the $H$ product matrix over the $k^{\text{th}}$ $\tau$-window is indeed close to $A_k$ (sub-leading order terms included). Define the difference $\Theta_{k,t} := H_{k,t} - A_k$

**Lemma 4.4.4.** *Fix $\epsilon \ll \frac{1}{2}$. There exists a constant $c$ such that, for sufficiently large $t$, $k \geq \tau$ and on $A_{k,t} \cap B_{k,t}$:*

$$
\|\Theta_{k,t}\| \leq \frac{c}{k\tau^{1/2-\epsilon}}
$$

*Proof.* The matrix $H_{k,t} = \Pi_{j=(k-1)\tau+r+1}^{k\tau+r}(I + \vec{\gamma}_{j+1}^{T} \circ L_j)$ has $\tau$ factors in the product. It can be expanded as a sum:

$$
H_{k,t} = \sum_{n=0}^{\tau} h_{k,t,n}
$$

where $h_{k,t,n}$ collects the terms in the expansion with exactly $n$ factors of the dampened laplacian matrix $\vec{\gamma}^T \circ L$. Now, on good events, and for all $s_0 \le j \le s_1$ and for all $i \in \mathcal{V}$, we have that

$$g_s^i \gtrsim k\tau.$$

It's also to see that, almost surely, we have that $\|L_j\| \lesssim 1$. Then, using the submultiplicativity of the operator norm with respect to the Hadamard product, we have that, on good events,

$$\|\vec{\gamma}_{j+1}^T \circ L_j\| \le \frac{c}{k\tau}$$

for some constant $c$. So, collecting all terms with $n$ such matrices as factors in the binomial expansion (there are $\binom{\tau}{n}$ of them), we have that

$$\|h_{k,t,n}\| \le \binom{\tau}{n}\big(\frac{c}{k\tau}\big)^n \le \frac{1}{n!}\big(\frac{c}{k}\big)^n$$

So, using the previous lemma (which says that $\|h_{k,t,0} + h_{k,t,1} - A_k\| \le \frac{c'}{k\tau^{1/2-\epsilon}}$ for some constant $c'$:

$$\begin{aligned}
\|H_{k,t} - A_k\| &\le \frac{c'}{k\tau^{1/2-\epsilon}} + \sum_{n=2}^{\tau} \frac{1}{n!}\big(\frac{c}{k}\big)^n \\
&\le \frac{c'}{k\tau^{1/2-\epsilon}} + \big(\frac{c}{k}\big)^2 \sum_{n=0}^{\tau} \big(\frac{c}{k}\big)^n \\
&\le \frac{c'}{k\tau^{1/2-\epsilon}} + \frac{(\frac{c}{k})^2}{1-\frac{c}{k}} \\
&\le \frac{c'}{k\tau^{1/2-\epsilon}} + \frac{2c^2}{k\tau^{1/2-\epsilon}} \\
&= \frac{c' + 2c^2}{k\tau^{1/2-\epsilon}},
\end{aligned}$$

where I've used that $\tau$ is large, for example, enough to have $\frac{c}{\tau} \le \frac{1}{2}$, and that $k \ge \tau$. $\square$

### 4.4.4 Decay of Operator Norm

Now, recall that

$$\Pi_{k=j}^{t}\Lambda_k = \left[\Pi_{k=\tau+1}^{\frac{t-r}{\tau}} H_{k,t}\right]\left[\Pi_{k=j}^{j_0}\Lambda_k\right].$$

The late ($k > j_0$) $\Lambda_k$ matrices in the product are encapsulated in the $H$ matrices, while I 'chop off' the early ($k \leq j_0$) $\Lambda_k$ matrices. These matrices are removed because $t - j$ may not be divisible by $\tau$ (and thus that we cannot successfully partition all $\Lambda$ into groups of equal size). Note, however, that in the above decomposition, I chop off more than the remainder of $t - j$ divided by $\tau$; this is for later convenience.

The next lemma guarantees that these extra, 'loose' factors of $\Lambda$ have bounded norm. I present a straightforward proof that makes use of some simple matrix calculations. It can be noted, however, that this lemma can also be proven by noting that a discrete dynamical system driven by the $\Lambda$ matrices (with no random blips $W$) represent a version of the heat equation where the only randomness is in the edge selection, rather than in the outcome in the conversation, and long-term solutions must be bounded.

**Lemma 4.4.5.** *For all* $0 \leq j \leq t$,

$$\|\Pi_{k=j}^{t}\Lambda_k\| \leq \sqrt{|\mathcal{V}|}$$

*almost surely.*

*Proof.* I first aim to prove that $\Lambda_k$ is nonnegative. The offdiagonal elements are obviously nonnegative, so I focus only on the diagonal. Let $k$ be arbitrary. For any $i$,

$$\Lambda_k^{ii} = 1 - \frac{1}{g_{k+1}^{i}} \sum_{j\sim i} S_{k+1}^{ij} \frac{g_k^j}{g_k^i + g_k^j}.$$

Now, if $S^{ij}_{k+1} = 0$ for all $j \sim i$, then it's clear that $\Lambda^{ii}_k = 1$. Otherwise, let $S^{ij}_{k+1} = 1$ for some $j \sim i$. Note that this will be the only nonzero term in the sum. In this case, we are guaranteed that $g^i_{k+1} = g^i_k + 1 > 1$. So:

$$\Lambda^{ii}_k = 1 - \frac{1}{g^i_{k+1}} \frac{g^j_k}{g^i_k + g^j_k} > 1 - \frac{g^j_k}{g^i_k + g^j_k} > 0.$$

This gives that $\Lambda_k$ is nonnegative.

Fix $j, t$ as above, arbitrary. Note that

$$A_{j,t} := \Pi^t_{k=j} \Lambda_k$$

is row stochastic, as it is the product of row stochastic matrices. It's also nonnegative. So, let $\vec{v}$ be an aribtrary unit vector. Note that for all $j$, $|v^j| \leq 1$ so:

$$|(A_{j,t}v)_i| = |\sum_k (A_{j,t})^{ik} v^k| \leq \sum_j (A_{j,t})^{ik} = 1.$$

where I've used that $A$ is nonnegative. Thus for arbitrary unit vector, $\|A_{j,t}v\|^2 = \sum_k \left((A_{j,t}v)^k\right)^2 \leq \sum_k 1 = |\mathcal{V}|$. This proves that, for arbitrary $0 \leq j < t$, almost surely, $\|A_{j,t}\| \leq \sqrt{|\mathcal{V}|}$. □

Before tackling the main lemma of this section (Lemma 4.4.7), note the useful fact that for a square matrix $A$, the $\ell_2$ operator norm is equivalent to the max of the vector norms of the rows.

**Lemma 4.4.6.** *Let $A$ be an $n \times n$ matrix, and let $A^i$ represent the ith row of $A$. We then have the following two inequalities, for arbitrary $1 \leq i \leq n$:*

$$\|A^i\| \leq \|A\|$$

$$\|A\| \leq \sqrt{n} \max_j \|A^j\|$$

*Note that for $1 \times n$ matrices, the matrix norm coincides with the vector norm.*

*Proof.* Let $1 \le i \le n$ be arbitrary. Start with the the first inequality. If $A^i = \vec{0}$, we are done. Otherwise, define the vector $\vec{x}$ to have components $x^j := \frac{A^{ij}}{\|A^i\|}$. Then we have

$$\|A\| \ge \|A\vec{x}\| \ge \|A^i\|,$$

where the second inequality follows because the $i^{\text{th}}$ entry of $A\vec{x}$ is equal to $A^i \cdot x = \|A^i\|$.

Next I prove the second inequality. Let $\vec{x}$ be an arbitrary unit vector. We have

$$\|A\vec{x}\|^2 = \sum_j (A^j \cdot \vec{x})^2 \le \sum_j \|A^j\|^2 \le n \max_j \|A^j\|^2.$$

Taking the square root of both sides, we have the desired inequality. □

I now add the main ingredient in the proof of 4.4.7, which says that for sufficiently small $j$ and on good events, the product of diffusion matrices (with consensus projected out) decays with $t$.

**Lemma 4.4.7.** *There exists $\alpha > 0$ such that, on $E_t$, and for all $j \le j_0 := r + \tau^2$,*

$$\|QP^{-1}[\Pi_{k=j}^{t}\Lambda_k]P\| \lesssim \frac{1}{t^\alpha},$$

*Proof.* Define $k_{\min} := \tau + 1$, $k_{\max} = \frac{t-r}{\tau}$, $\Theta'_{k,t} := P^{-1}\Theta_{k,t}P$, and $D_k = P^{-1}A_kP$ (the diagonal matrix consisting of eigenvalues of $A_k$). Now,

$$\left[\Pi_{k=k_{\min}}^{k_{\max}} H_{k,t}\right]\left[\Pi_{k=j}^{j_0}\Lambda_k\right].$$

$$
\begin{aligned}
\|QP^{-1}[\Pi_{k=j}^{t}\Lambda_k]P\| &\leq \|QP^{-1}[\Pi_{k=j_0+1}^{t}\Lambda_k]P\|\|P^{-1}[\Pi_{k=j}^{j_0}\Lambda_k]P\| \\
&\lesssim \|QP^{-1}[\Pi_{k=j_0+1}^{t}\Lambda_k]P\| \\
&= \|QP^{-1}[\Pi_{k=k_{\min}}^{k_{\max}} H_{k,t}]P\| \\
&= \|QP^{-1}[\Pi_{k=k_{\min}}^{k_{\max}} (A_k + \Theta_{k,t})]P\| \\
&= \|Q[\Pi_{k=k_{\min}}^{k_{\max}} (D_k + \Theta'_{k,t})]\| \\
&\lesssim \max_{i\neq 1} \|[\Pi_{k=k_{\min}}^{k_{\max}} (D_k + \Theta'_{k,t})]^i\| \\
&= \max_{i\neq 1} \|R^i_{k_{\max},k_{\min},t}\|
\end{aligned}
$$

where Lemma 4.4.6 has been used in the second to last line and I've defined

$$
R_{k_1,k_2,t} := \Pi_{k=k_1}^{k_2}(D_k + \Theta'_{k,t}).
$$

One has, for a constant $c$, for $\epsilon < \frac{1}{2}$, and for all $i \neq 1$ (dropping primes on Theta for ease of notation),

$$
\begin{aligned}
R^i_{k_{\min},k_{\max},t} &= \Theta^i_{k_{\max},t} R_{k_{\min},k_{\max}-1,t} + |\lambda^{(k_{\max})}| R^i_{k_{\min},k_{\max}-1,t} \\
\|R^i_{k_{\min},k_{\max},t}\| &\leq \frac{c}{k_{\max}\tau^{1/2-\epsilon}} + (1 - \frac{\lambda}{k_{\max}})\|R^i_{k_{\min},k_{\max}-1,t}\|,
\end{aligned}
$$

where $\lambda^{(k)} \neq 1$ is an eigenvalue of $A_k$, and I've used Lemma 4.4.5, Lemma 4.4.4, and Lemma 4.2.3. By iterating the above, we obtain

$$
\begin{aligned}
\|R^i_{k_{\min},k_{\max},t}\| &\leq \Pi_{k=k_{\min}}^{k_{\max}}(1 - \frac{\lambda}{k}) + \frac{c}{\tau^{1/2-\epsilon}} \sum_{j=k_{\min}}^{k_{\max}} \frac{1}{j} \Pi_{k=j+1}^{k_{\max}}(1 - \frac{\lambda}{k}) \\
&\leq (\frac{k_{\min}}{k_{\max}})^{\lambda} + \frac{c}{\tau^{1/2-\epsilon}} \sum_{j=k_{\min}}^{k_{\max}} \frac{1}{j}(\frac{j+1}{k_{\max}})^{\lambda} \\
&\leq (\frac{k_{\min}}{k_{\max}})^{\lambda} + \frac{c}{k_{\max}^{\lambda}\tau^{1/2-\epsilon}} \sum_{j=k_{\min}}^{k_{\max}} j^{\lambda-1} \\
&\leq (\frac{k_{\min}}{k_{\max}})^{\lambda} + \frac{c}{\tau^{1/2-\epsilon}} \\
&\lesssim \frac{1}{t^{\lambda/2}} + \frac{1}{t^{1/8-\epsilon/4}},
\end{aligned}
$$

where in the second and fourth inequalities, I used Lemma 4.2.6 and Lemma 4.3.2, respectively, and the value of $c$ can change from line to line. Setting $\epsilon = \frac{1}{4}$ (for example) concludes the proof of the lemma. □

Our final ingredient is the summability of $\mathbb{E}[\|\vec{\gamma}_t\|^2]$.

**Lemma 4.4.8.** *The sum*

$$\sum_{t=0}^{\infty} \mathbb{E}[\|\vec{\gamma}_t\|^2]$$

*converges.*

*Proof.* For sufficiently large $t$:

$$\begin{aligned}
\mathbb{E}[\|\vec{\gamma}_t\|^2] &= \mathbb{E}[\|\vec{\gamma}_t\|^2 | A_{t,\frac{t-r}{\tau}}]\mathbb{P}(A_{t,\frac{t-r}{\tau}}) + \mathbb{E}[\|\vec{\gamma}_t\|^2 | A^c_{t\frac{t-r}{\tau}}]\mathbb{P}(A^c_{t,\frac{t-r}{\tau}}) \\
&\lesssim \mathbb{E}[\|\vec{\gamma}_t\|^2 | A_{t,\frac{t-r}{\tau}}] + \mathbb{P}(A^c_{t,\frac{t-r}{\tau}}) \\
&\lesssim \frac{1}{t^2} + \exp(-\frac{1}{4}\tau^{4\epsilon}) \\
&\lesssim \frac{1}{t^2}
\end{aligned}$$

□

### 4.4.5 Proof of Lemma 4.4.1

In the proof of Lemma 4.4.1, the following simple comparison between a nonnegative random variable's conditional and total expectation is useful:

**Lemma 4.4.9.** *Let $X$ be an almost-surely nonnegative random variable with $\mathbb{E}[X] < \infty$, and let $E$ be an event with $\mathbb{P}(E) \geq \frac{1}{2}$. Then*

$$\mathbb{E}(X|E) \leq 2\mathbb{E}[X]$$

*Proof.*

$$\mathbb{E}[X|E] = \frac{1}{\mathbb{P}[E]}(\mathbb{E}[X] - \mathbb{E}[X|E^c]\mathbb{P}(E^c)) \le 2\mathbb{E}[X].$$

□

*Proof of Lemma 4.4.1:* I aim to show that

$$\mathbb{E}[\|\vec{z}_{t+1}\|^2] \to 0,$$

where

$$\vec{z}_{t+1} = \sum_{j=0}^{t+1} PQP^{-1}\big[\Pi_{k=j}^{t}\Lambda_k\big]\vec{\gamma}_j^T \circ \vec{w}_j.$$

Expanding the square:

$$\begin{aligned}\|\vec{z}_{t+1}\|^2 = &\sum_{j=0}^{t+1} \|PQP^{-1}\big[\Pi_{k=j}^{t}\Lambda_k\big]\vec{\gamma}_j^T \circ \vec{w}_j\|^2 \\ &+ 2\sum_{0\le j_1<j_2\le t+1} \langle PQP^{-1}\big[\Pi_{k_1=j_1}^{t}\Lambda_{k_1}\big]\vec{\gamma}_{j_1}^T \circ \vec{w}_{j_1}, PQP^{-1}\big[\Pi_{k_2=j_2}^{t}\Lambda_{j_2}\big]\vec{\gamma}_{j_2}^T \circ \vec{w}_{j_2}\rangle\end{aligned}$$

Now take the expectation of the cross-terms. For $0 \le j_1 < j_2 \le t+1$:

$$\begin{aligned}&\mathbb{E}[\langle PQP^{-1}\big[\Pi_{k_1=j_1}^{t}\Lambda_{k_1}\big]\vec{\gamma}_{j_1}^T \circ \vec{w}_{j_1}, PQP^{-1}\big[\Pi_{k_2=j_2}^{t}\Lambda_{j_2}\big]\vec{\gamma}_{j_2}^T \circ \vec{w}_{j_2}\rangle] \\ &= \mathbb{E}\big[\mathbb{E}[\langle PQP^{-1}\big[\Pi_{k_1=j_1}^{t}\Lambda_{k_1}\big]\vec{\gamma}_{j_1}^T \circ \vec{w}_{j_1}, PQP^{-1}\big[\Pi_{k_2=j_2}^{t}\Lambda_{j_2}\big]\vec{\gamma}_{j_2}^T \circ \vec{w}_{j_2}\rangle|\sigma(\mathcal{F}_{j_2-1}, \mathcal{G}_{t+1})]\big] \\ &= \mathbb{E}\big[\big(PQP^{-1}\big[\Pi_{k_1=j_1}^{t}\Lambda_{k_1}\big]\vec{\gamma}_{j_1}^T \circ \vec{w}_{j_1}\big)^T PQP^{-1}\big[\Pi_{k_2=j_2}^{t}\Lambda_{j_2}\big](\vec{\gamma}_{j_2}^T \circ \mathbb{E}[\vec{w}_{j_2}|\sigma(\mathcal{F}_{j_2-1}, \mathcal{G}_{t+1})]\big] \\ &= 0,\end{aligned}$$

where I've used independence of $\sigma(\psi_t)$ and $\sigma(\mathcal{F}_{t-1}, \sigma(\Omega_t))$ as well as the fact that $\mathbb{E}_{t-1}[\vec{w}_t] = 0$.

Now deal with the expectation of the 'diagonal' elements:

$$\begin{aligned}
\mathbb{E}[\|\vec{z}_t\|^2] &= \sum_{j=0}^{t+1} \|PQP^{-1}\left[\Pi_{k=j}^{t}\Lambda_k\right]\vec{\gamma}_j^T \circ \vec{w}_j\|^2 \\
&\lesssim \sum_{j=0}^{t+1} \mathbb{E}[\|PQP^{-1}\left[\Pi_{k=j}^{t}\Lambda_k\right]\|^2\|\vec{\gamma}_j\|^2] \\
&= \sum_{j=0}^{r+\tau^2} \mathbb{E}\left[\|PQP^{-1}\left[\Pi_{k=j}^{t}\Lambda_k\right]\|^2\|\vec{\gamma}_j\|^2\right] + \sum_{j=r+\tau^2+1}^{t+1} \mathbb{E}\left[\|PQP^{-1}\left[\Pi_{k=j}^{t}\Lambda_k\right]\|^2\|\vec{\gamma}_j\|^2\right]
\end{aligned}$$

I show that each of the above two terms goes to zero. Let $\alpha > 0$ be the number guaranteed by Lemma 4.4.7. The first term:

$$\begin{aligned}
&\sum_{j=0}^{r+\tau^2} \mathbb{E}\left[\|PQP^{-1}\left[\Pi_{k=j}^{t}\Lambda_k\right]\|^2\|\vec{\gamma}_j\|^2\right] \\
&\lesssim \sum_{j=0}^{r+\tau^2} \mathbb{E}\left[\|PQP^{-1}\left[\Pi_{k=j}^{t}\Lambda_k\right]\|^2\|\vec{\gamma}_j\|^2|E_t\right] + \mathbb{E}\left[\|PQP^{-1}\left[\Pi_{k=j}^{t}\Lambda_k\right]\|^2\|\vec{\gamma}_j\|^2|E_t^c\right]\mathbb{P}(E_t^c) \\
&\lesssim \sum_{j=0}^{r+\tau^2} \frac{1}{t^\alpha}\mathbb{E}\left[\|\vec{\gamma}_j\|^2|E_t\right] + \mathbb{P}(E_t^c) \\
&\lesssim \sum_{j=0}^{r+\tau^2} \frac{1}{t^\alpha}\mathbb{E}\left[\|\vec{\gamma}_j\|^2\right] + \mathbb{P}(E_t^c) \\
&\lesssim \frac{1}{t^\alpha} + (r+\tau^2)\exp(-c\tau^{2\epsilon}) \\
&\to 0,
\end{aligned}$$

where in the second inequality I used Lemmas 4.4.7 and 4.4.5, in the third inequality I used Lemma 4.4.9, and in the fourth inequality I used Lemmas 4.4.8 and 4.4.2.

And in the second term of the expansion of the diagonal sum:

$$\sum_{j=r+\tau^2+1}^{t+1} \mathbb{E}\left[\|PQP^{-1}\left[\Pi_{k=j}^{t}\Lambda_k\right]\|^2\|\vec{\gamma}_j\|^2\right] \lesssim \sum_{j=r+\tau^2+1}^{t+1} \mathbb{E}\left[\|\vec{\gamma}_j\|^2\right],$$

where I've used Lemma 4.4.5. The right-hand side goes to 0 By Lemma 4.4.8, since the lower bound of the sum goes to $\infty$.

□

## 4.5 Proof of Theorem

*Proof of Theorem 4.1.1.* Let $a_\infty$ be the limit of $a_t = \vec{p} \cdot \vec{x}_t$, established in Lemma 4.3.1. Using the triangle inequality, we have:

$$\begin{aligned}\mathbb{E}[\|\vec{x}_t - a_\infty \vec{1}\|^2] &= \mathbb{E}[\|\vec{x}_t - a_t\vec{1} + a_t\vec{1} - a_\infty\vec{1}\|^2] \\ &\leq \mathbb{E}[\|a_t\vec{1} - a_\infty\vec{1}\|^2] + \mathbb{E}[\|\vec{x}_t - a_t\vec{1}\|^2] + 2\mathbb{E}[\|\vec{x}_t - a_t\vec{1}\|\|a_t\vec{1} - a_\infty\vec{1}\|].\end{aligned}$$

I now show that each term goes to 0. In the first term, we have that

$$\|a_t\vec{1} - a_\infty\vec{1}\|^2 = (a_t - a_\infty)^2\|\vec{1}\|^2,$$

the expectation of which goes to 0 by virtue of Lemma 4.3.1. Similarly, the second term goes to 0 due to Lemma 4.4.1.

To see that the last term goes to 0, note that $\|\vec{x}_t - a_t\vec{1}\|$ is almost surely bounded, so that

$$\mathbb{E}[\|\vec{x}_t - a_t\vec{1}\|\|a_t\vec{1} - a_\infty\vec{1}\|] \lesssim \mathbb{E}[\|a_t\vec{1} - a_\infty\vec{1}\|] = \|\vec{1}\|\mathbb{E}[|a_t - a_\infty|] \lesssim \|a_t - a_\infty\|_{\mathcal{L}^1}.$$

Finally, since $a_t \to a_\infty$ in $\mathcal{L}^2$, convergence also holds in $\mathcal{L}^1$, so that this term goes to 0 as well. □

## 4.6 Future Work

Future work might consider the rate of convergence, for example of the disagreement component $\vec{z}_t$ to 0. Simulations inspire the following conjecture:

**Conjecture 1.**

$$\mathbb{E}[\|\vec{z}_t\|^2] \lesssim \begin{cases} \frac{1}{t^{2\lambda}} & \lambda \leq \frac{1}{2} \\ \frac{1}{t} & \lambda > \frac{1}{2} \end{cases}$$

.

In the case of parallel updates (i.e. all edges converse with all of their neighbors simultaneously in each time step), the above conjecture can be proven readily using the techniques from Lemma 4.4.7. With the appropriate choice of $\tau(t)$, bounds on the decay rate can be proven for the present case (though these bounds seem loser than what simulation demonstrates). This discussion has been omitted because the bounds do not seem empirically tight, and the choice of $\tau(t) = \lceil t^{1/4} \rceil$ is convenient.

Figure 4.3 shows the decay of disagreement, averaged over 1000 runs for the interval graph $I_5 = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \{1, 2, 3, 4, 5\}$, and $(i, j) \in \mathcal{E}$ if and only if $|i - j| = 1$.

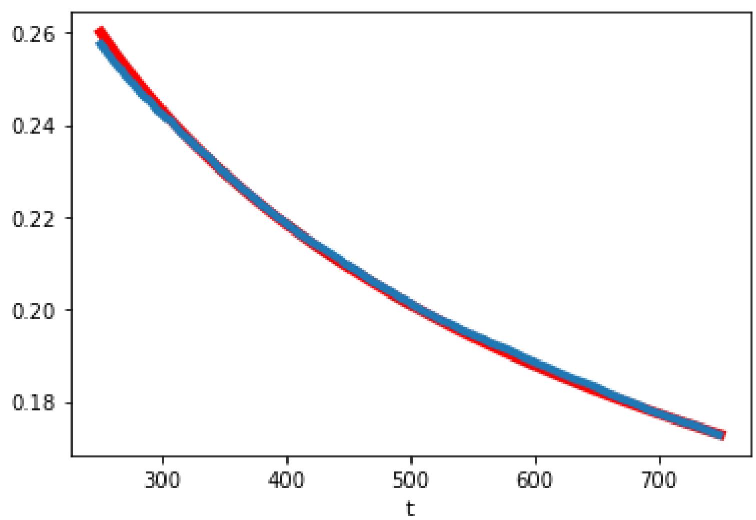



Figure 4.3: The blue curve represents the average of 1000 trajectories of $\|\vec{x}_t - a_t\vec{1}\|_2^2$, for $g_0^i \equiv 1$ and $\vec{x}_0 = (1, 1, 0, 0, 0)$ for the graph $I_5$.
The red curve is $\frac{2.02}{t^{2*0.185667}}$
(note: for $I_5$, $\lambda = \frac{13-\sqrt{73}}{24} \approx 0.185667$).

# BIBLIOGRAPHY


[1] Corry, Scott, and David Perkinson. (2018). *Divisors and Sandpiles: An Introduction to Chip-Firing.* American Mathematical Society.

[2] Fey, Anne, et al. “Growth Rates and Explosions in Sandpiles.” (2009). *Journal of Statistical Physics*, vol. 138, no. 1–3, pp. 143–159.

[3] Járai, Antal A. “Sandpile Models.” (2018). *Probability Surveys*, vol. 15, no. none.

[4] Levine, Lionel, et al. “Apollonian Structure in the Abelian Sandpile.” (2016). *Geometric and Functional Analysis*, vol. 26, no. 1, pp. 306–336.

[5] Levine, Lionel, et al. “The Apollonian Structure of Integer Superharmonic Matrices.” (2017). *Annals of Mathematics*, vol. 186, no. 1.

[6] Pegden, Wesley, and Charles K. Smart. “Stability of Patterns in the Abelian Sandpile.” (2020). *Annales Henri Poincaré*, vol. 21, no. 4, pp. 1383–1399.

[7] Pegden, Wesley, and Charles K. Smart. “Convergence of the Abelian Sandpile.” (2013). Duke Mathematical Journal, vol. 162, no. 4.

[8] Stefani, G. ”On the monotonicity of the perimeter of convex bodies.” preprint, arxiv:math/1612.00295

[9] Bak, P., Tang, C., Wiesenfeld, K. (1987). Self-organized criticality: An explanation of the 1/f noise. Physical Review Letters, 59(4), 381–384.

[10] Dhar, D. (1999). The abelian sandpile and related models. Physica A: Statistical Mechanics and Its Applications, 263(1-4), 4–25.

[11] Manna, S S. (1991). Two-state model of self-organized criticality. *J. Phys. A: Math. Gen.* 24 L363

[12] Rolla, Leonardo T. (2020). Activated random walks on $\mathbb{Z}^d$. *Probab. Surv.* 17, 478–544.

[13] Dhar, Deepak. (2006). Theoretical studies of self-organized criticality. *Phys. A* 369, no. 1, 29–70.

[14] Ktitarev, D. V., Lübeck, S., Grassberger, P., B. Priezzhev, V. (2000). Scaling of waves in the bak-tang-wiesenfeld sandpile model. Physical Review E, 61(1), 81–92.

[15] Rolla, Leonardo T.; Sidoravicius, Vladas. (2012). Absorbing-state phase transition for driven-dissipative stochastic dynamics on $\mathbb{Z}$. *Invent. Math.* 188, no. 1, 127–150.

[16] Basu, Riddhipratim; Ganguly, Shirshendu; Hoffman, Christopher; Richey, Jacob. (2019). Activated random walk on a cycle. *Ann. Inst. Henri Poincaré Probab. Stat.* 55, no. 3, 1258–1277

[17] Hannah Cairns, Shirshendu Ganguly, and Lionel Levine. (2021). Phase transition for Activated Random Walk on a cycle, In preparation.

[18] Diaconis, P.; Fulton, W. (1993). A growth model, a game, an algebra, Lagrange inversion, and characteristic classes. Commutative algebra and algebraic geometry, II (Italian) (Turin, 1990). *Rend. Sem. Mat. Univ. Politec. Torino* 49 (1991), no. 1, 95–119.

[19] Deffuant, Guillaume, et al. (2000). Mixing Beliefs among Interacting Agents. *Advances in Complex Systems*, vol. 03, no. 01n04, pp. 87–98.

[20] Hegselmann, Rainer, and Ulrich Krause. (2005). Opinion Dynamics Driven by Various Ways of Averaging. *Computational Economics*, vol. 25, no. 4, pp. 381–405.

[21] Degroot, Morris H. (1974). Reaching a Consensus. *Journal of the American Statistical Association*, vol. 69, no. 345, 1974, pp. 118–121.

[22] Banerjee, Abhijit, et al. (2021). Naïve Learning with Uninformed Agents. *American Economic Review*, vol. 111, no. 11, pp. 3540–3574.

[23] Ding, Zhaogang, et al. (2019). Consensus Reaching in Social Network DeGroot Model: The Roles of the Self-Confidence and Node Degree. *Information Sciences*, vol. 486, pp. 62–72.

[24] Gao, Yue, et al. (2020). The Dynamics of Two-State Public Opinion Propagation on Signed Networks. *Journal of Systems Science and Complexity*, vol. 34, no. 1, pp. 251–264.

[25] Li, Yun, and Jiakun Wang. (2021). Cross-Network Propagation Model of Public Opinion Information and Its Control in Coupled Double-Layer Online Social Networks. *Aslib Journal of Information Management*, vol. 74, no. 2, pp. 354–376.

[26] Martins, Andre C. (2008). Continuous Opinions and Discrete Actions in Opinion Dynamics Problems. *International Journal of Modern Physics C*, vol. 19, no. 04, pp. 617–624.

[27] Mohammadinejad, Amir, et al. (2018). Opiu: Opinion Propagation in Online Social Networks Using Influential Users Impact. *2018 IEEE International Conference on Communications (ICC).*

[28] Prasetya, Hafizh A., and Tsuyoshi Murata. (2020). A Model of Opinion and Propagation Structure Polarization in Social Media. *Computational Social Networks*, vol. 7, no. 1.

[29] Ureña, Raquel, et al. (2018). A New Influence Based Network for Opinion Propagation in Social Network Based Scenarios. *Procedia Computer Science*, vol. 139, pp. 329–337.

[30] Bashari, Masoud, and Mohammad-R. Akbarzadeh-T. (2023). Theoretical Development of a Probabilistic Fuzzy Model for Opinion Formation in Social Networks. *Fuzzy Sets and Systems*, vol. 454, pp. 125–148.

[31] Lemmens, Bas, and Roger D. Nussbaum. (2012). *Nonlinear Perron-Frobenius Theory*. Cambridge University Press.

[32] Cairns, Hannah. (2021). Perron's Theorem in an Hour *Taylor and Francis Online*.

[33] Hogben, Leslie. (2016). *Handbook of Linear Algebra. CRC Press/Taylor and Francis Group.*

[34] §5.6 Inequalities. Digital Library of Mathematical Functions, dlmf.nist.gov/5.6.

[35] Alon, Noga, and Joel H. Spencer. (2016). *The Probabilistic Method*. Wiley.

[36] Williams, David. (2020). *Probability with Martingales. Cambridge University Press.*