

Bounds on Rate of Convergence for the Shuffled Discrete Heat Equation in \mathbb{Z}^d

Luciano Viñas[†]

Sponsor: Atchar Sudhyadhom[‡]

Abstract

We explore the effects of interleaved shuffling on the rate of convergence for the discrete heat equation with Dirichlet boundary conditions. We derive a closed form for the expected value of the shuffled discrete heat equation and establish bounds on its rate of convergence. In particular for any connected region $D \subset \mathbb{Z}^d$ with volume $|D|$ and a non-negative initial state $h_0 \in \mathbb{R}^{|D|}$, there is an upper bound on the spectral radius associated with the shuffled discrete heat equation that grows on the order of $1 - \Omega(1/|D|^{1/d})$. An analogous lower bound for the standard discrete heat equation is also derived which grows on the order of $1 - \mathcal{O}(1/|D|^{2/d})$.

1 Introduction

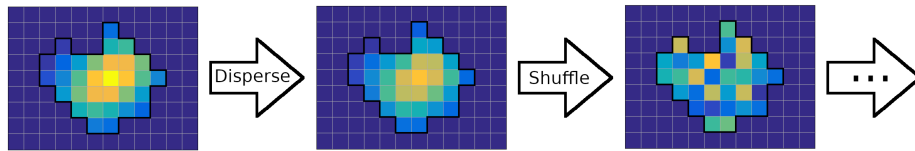


Figure 1: Shuffled Discrete Heat Equation on an Example Region in \mathbb{Z}^2

The heat equation is a partial differential equation well-known in science disciplines for its simplicity and aptness in modeling dispersion processes. Straightforward applications of the equation can be found in physics and biology, such as describing how temperature propagates down a rod or how the concentration of a mixture changes in space and time. In many real-world cases, dispersion processes operate within regions too complicated to solve analytically with the heat equation. It is instead more practical to obtain numerical approximations of the region using a discretized version of the heat equation. Interestingly enough, this discrete heat equation finds uses in computer science and applied math beyond numerical evaluation. Some noteworthy examples include its application in random walks and its use in exponential kernel families [6] [8].

In this paper we explore a supplementary topic on the rates of convergence for different discrete heat equation variants. The rate of convergence specifies how quickly a linearly convergent sequence converges to its limiting value. We induce

[†]University of California, Berkeley.

Email: lucianovinas@gmail.com

[‡]University of California, San Francisco.

Email: Atchar.Sudhyadhom@ucsf.edu

a sequence from the discrete heat equation by chronologically enumerating each discrete time step to a countable set. By characterizing the rate of convergence, we can determine how quickly each discrete heat variant reaches its equilibrium state.

Spectral theorem will provide a convenient eigenbasis to simplify the dispersion dynamics present in these discrete heat equation variants. By consequence of its formal definition, the rate of convergence will be exactly equal to the largest magnitude eigenvalue, the spectral radius, of the eigenbasis above. Knowing this, we refine our search to establishing bounds on the spectral radius of discrete heat equation variants. By providing these bounds we hope to aid discussion in classifying time complexities for more general discrete dispersion processes.

We begin this paper with a formal description of the discrete heat equation and its conditions for convergence. Next we define a shuffled variant of the discrete heat equation and solve for its expected value as a function of time. Finally we characterize the rate of convergence of each equation by bounding the spectral radius of pertinent linear operators.

Both the bound on the spectral radius and the closed form solution for the expectation of the shuffled heat equation are novel. Both spectral radii bounds share a similar function form which allows for straightforward comparison on the rate of convergence for the two discrete dispersion processes. The bound for the standard discrete heat equation can also be obtained by leveraging the eigenvalue bound for the negative Laplacian on bounded domains [11] [15].

2 Discrete Formulation of the Heat Equation

For a function $h : \mathbb{R}^d \times \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}$, the heat equation is expressed as

$$\frac{\partial}{\partial t} h(x, t) = \alpha \nabla^2 h(x, t). \quad (1)$$

It is common to refer to $x \in \mathbb{R}^d$ as the spatial position and $t \in \mathbb{R}_{\geq 0}$ as the temporal position of the function $h(\cdot, \cdot)$. The real-valued variable α is a constant describing how quickly $h(\cdot, \cdot)$ disperses in the medium of interest. The Laplacian in equation (1) acts on the spatial components of $h(\cdot, \cdot)$. That is to say, if our dispersive region occupies d dimensions such that $x = (x_1, \dots, x_d)$ then

$$\nabla^2 = \sum_{i=1}^d \frac{\partial^2}{\partial x_i^2}.$$

In the discrete case we add the condition that our spatio-temporal inputs must be integer-valued. This discrete input function is denoted by $h_t : \mathbb{Z}^d \rightarrow \mathbb{R}$ with discrete time indexes $t \in \mathbb{N}_0$. We have $h_t(\cdot)$ evolve according to the d -dimensional discrete heat equation defined by G. Lawler in [10]

$$\Delta h_t(x) = \alpha \mathcal{L} h_t(x) \quad \alpha \in (0, 1). \quad (2)$$

Here Δ is the forward difference operator in t and \mathcal{L} is the negative normalized discrete Laplacian with the following formal definition

$$\mathcal{L}h_t(x) = \frac{1}{2d} \sum_{\|y-x\|_1=1} [h_t(y) - h_t(x)]. \quad (3)$$

Borrowing further conventions from [10], for a set $D \subset \mathbb{Z}^d$ the boundary ∂D is defined as

$$\partial D = \{x \in \mathbb{Z}^d \setminus D \mid \inf_{y \in D} \|y - x\|_1 = 1\}. \quad (4)$$

We provide a visualization of equation (4) in Figure 2.

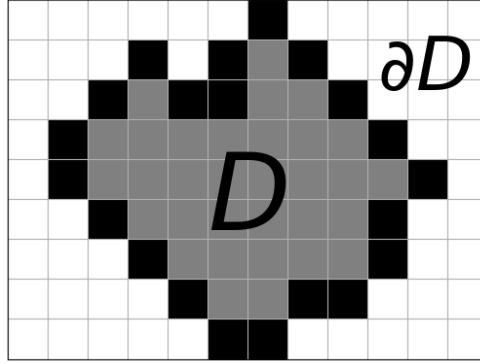


Figure 2: Example region (gray) in \mathbb{Z}^2 with boundary (black)

Our analysis is restricted to the case where D is finite and connected and the elements of ∂D satisfy the homogeneous Dirichlet boundary condition. We define an enumeration $\mathcal{I} : D \rightarrow \mathbb{N}$ such that function $h_t(\cdot)$ has a vector representation in $\mathbb{R}^{|D|}$. Using the same enumeration \mathcal{I} we introduce a matrix representation for \mathcal{L} where

$$\mathcal{L} = \frac{1}{2d} A_D - I. \quad (5)$$

Here A_D is the adjacency matrix of (D, E_D) , the natural induced graph structure of \mathbb{Z}^d , with $E_D = \{(u, v) \in D \times D \mid \|u - v\|_1 = 1\}$ and I is the identity matrix in $\mathbb{R}^{|D| \times |D|}$. Equations (3) and (5) are equivalent for homogeneous Dirichlet boundary conditions, but the representation in (5) will allow for more convenient handling of h_t . Revisiting the discrete formulation in (2), the following simplifications are produced

$$\begin{aligned} h_{t+1} - h_t &= \alpha \left(\frac{1}{2d} A_D - I \right) h_t \\ h_{t+1} &= \left(\frac{\alpha}{2d} A_D + (1 - \alpha) I \right) h_t = \left(\frac{\alpha}{2d} A_D + (1 - \alpha) I \right)^{t+1} h_0. \end{aligned} \quad (6)$$

Vector $h_0 \in \mathbb{R}^{|D|}$ is the non-negative initial state on D . The expression in (6) is especially nice as it identifies the time evolution operator for the discrete heat equation, $\frac{\alpha}{2d}A_D + (1-\alpha)I$, in terms of readily known quantities. Questions about the dynamics of (6) will conveniently simplify to analyzing the characteristics of the time evolution operator and its interaction with the initial state h_0 .

2.1 Convergence of Discrete Heat Equation

For the sake of compactness, we will have $B = \frac{\alpha}{2d}A_D + (1-\alpha)I$ and $n = |D|$. Matrix A_D is symmetric as it is the adjacency matrix of an undirected graph (D, E_D) . Consequently B is symmetric as well. By spectral theorem there exist eigenvectors $\{v_i\}_{i=1}^n$ with eigenvalues $\{\lambda_i\}_{i=1}^n$ which form an orthonormal eigenbasis for B . As $\{v_i\}_{i=1}^n$ is an orthonormal basis we can write h_0 as

$$h_0 = \sum_{i=1}^n \langle v_i, h_0 \rangle v_i.$$

Applying the above to (6) gives

$$h_{t+1} = \sum_{i=1}^n \langle v_i, h_0 \rangle \lambda_i^{t+1} v_i. \quad (7)$$

If the spectral radius of B , $\rho(B)$, is less than 1 then equation (7) will converge to $0 \in \mathbb{R}^n$ for all initial states h_0 . As we will see in the next subsections, the bound $\rho(B) < 1$ holds for all $\alpha \in (0, 1]$. We will not consider corner case $\alpha = 0$ as it is uninteresting with trivial convergence $h_t = h_0$ for all $t \in \mathbb{N}$.

2.1.1 Lemmas Bounding Spectral Radius $\rho(B)$

We start with a introduction of Perron–Frobenius Theorem and some standard lemmas from graph theory.

THEOREM 1 (Perron–Frobenius Theorem for Irreducible Matrices). *If A is a non-negative irreducible matrix, there exists a positive eigenvector of A with corresponding positive eigenvalue equal to the spectral radius. Furthermore if A is primitive, all other eigenvectors must have corresponding eigenvalue with magnitude less than the spectral radius.*

LEMMA 1. *If G is a connected graph, then its adjacency matrix A_G is irreducible.*

Proof. A matrix A is said to be irreducible if for every i, j index pair there exists $k \in \mathbb{N}$ such that $(A^k)_{ij} > 0$. When A is an adjacency matrix, $(A^k)_{ij}$ gives the number of walks of length k between vertex i and j . We are given in the problem statement that G is connected, so there always exists a walk of finite length between any pair of vertices. \square

LEMMA 2. *If H is a connected proper subgraph of graph G , then $\rho(H) < \rho(G)$.*

Proof. Refer to [12] for a more general version of this proof.

Applying Theorem 1 and Lemma 1, let x be the positive eigenvector with positive eigenvalue equal to the spectral radius of graph H . As H is a subgraph G and vector x is positive, it follows that $A_H x \leq A_G x$. Furthermore since the subgraph relation is proper there exists index $i \in \{1, \dots, n\}$ such that

$$(A_H x)_i < (A_G x)_i.$$

Using Rayleigh's quotient the following inequality is obtained

$$\rho(A_H) = \frac{\langle x, A_H x \rangle}{\langle x, x \rangle} < \frac{\langle x, A_G x \rangle}{\langle x, x \rangle} \leq \sup_{x' \neq 0} \left| \frac{\langle x', A_G x' \rangle}{\langle x', x' \rangle} \right| = \rho(A_G).$$

□

LEMMA 3. *Let G be a d -regular graph, the spectral radius of graph G has value d and can be attained by the ones vector $\mathbb{1}$.*

Proof. The 1-norm of a matrix calculates the maximum absolute column sum, which in the case of a d -regular graph is d . Consider a 1-norm normalized eigenvector x of adjacency matrix A_G . Using standard matrix norm inequalities we have

$$|\lambda| = \|A_G x\|_1 \leq \|A_G\|_1 \|x\|_1 = d.$$

Since the above holds for every eigenvector x ,

$$\rho(A_G) \leq d.$$

Each row of adjacency matrix A_G has d non-zero elements, so the spectral upper bound of d can be attained by the ones vector $\mathbb{1}$. □

2.1.2 Convergence Argument

Note that graph (D, E_D) is a proper subgraph of a $2d$ -regular graph. Utilizing Lemmas 2 and 3 we get that $\rho(A_D) < 2d$ and equivalently

$$\rho(B) = \frac{|\alpha|}{2d} \rho(A_D) + |1 - \alpha| \rho(I) < \frac{|\alpha|}{2d} 2d + |1 - \alpha| = 1.$$

The last equality holds for $\alpha \in (0, 1]$ which is our case of interest.

3 Shuffled Discrete Heat Equation

This next dispersion process introduces a random element into the standard discrete heat equation. After each dispersive step, we allow for a random independent permutation between the $h_t(\cdot)$ values of D (refer to Figure 1 for a visualization). This dispersion process is formalized as so

$$\begin{aligned} h_{t+1}(x) &= \sigma_t(\alpha \mathcal{L}h_t(x) + h_t(x)) & x \in D, \\ \mathcal{L}h_t(x) &= 0 & x \in \partial D. \end{aligned} \quad (8)$$

Here σ_t is a permutation map of the set $\{\alpha\mathcal{L}h_t(x)+h_t(x) \mid x \in D\}$. Permutation σ_t is drawn uniformly from the set of n -element permutations S_n .

Using the same enumeration \mathcal{I} as in (5), permutation map σ_t admits matrix representation M_t . To differentiate from (6) and emphasize $h_t(\cdot)$ is random quantity for $t \in \mathbb{N}$, we let $\mathbf{h}_t \in \mathbb{R}^n$ stand in as the random vector of values evolving according to the shuffled discrete heat equation defined above. Plugging in the new matrix and vector representations we obtain the following relation

$$\mathbf{h}_{t+1} = M_t \left(\frac{\alpha}{2d} A_D + (1 - \alpha) I \right) \mathbf{h}_t = \left(\prod_{i=0}^t M_i \left(\frac{\alpha}{2d} A_D + (1 - \alpha) I \right) \right) h_0. \quad (9)$$

Currently the time evolution operator for the shuffled process is dependent on time. In the next section, equation (9) will be simplified by considering the expectation of \mathbf{h}_{t+1} .

3.1 Expectation of the Shuffled Discrete Heat Equation

Each permutation M_i is identically distributed and independent so the expectation of \mathbf{h}_{t+1} becomes

$$\mathbb{E}[\mathbf{h}_{t+1}] = \left(\mathbb{E}[M_0] \left(\frac{\alpha}{2d} A_D + (1 - \alpha) I \right) \right)^{t+1} h_0. \quad (10)$$

For a given permutation matrix M , element $(M)_{ij}$ equals one if j is mapped to i and equals zero otherwise. The number of permutations mapping j to i for a set of n element permutations is $(n-1)!$. Paired with the fact that each of the $n!$ permutations is equiprobable the expectation of M_0 simplifies to

$$\begin{aligned} \mathbb{E}[(M_0)_{ij}] &= \sum_{M_0 \in P_n} (M_0)_{ij} \Pr(M_0) = \sum_{\sigma \in S_n: \sigma(j)=i} 1 \cdot \frac{1}{n!} = \frac{1}{n} \\ \mathbb{E}[M_0] &= \frac{1}{n} \mathbb{1} \mathbb{1}'. \end{aligned} \quad (11)$$

Here P_n is the space of $n \times n$ permutation matrices and S_n is the same set of n -element permutations introduced in (8).

Substituting $\mathbb{E}[M_0]$ into (10) yields

$$\begin{aligned} \mathbb{E}[\mathbf{h}_{t+1}] &= \frac{1}{n} \mathbb{1} \left(\frac{\alpha}{2dn} \mathbb{1}' A_D \mathbb{1} + (1 - \alpha) \right)^t \mathbb{1}' \left(\frac{\alpha}{2d} A_D + (1 - \alpha) I \right) h_0 \\ &= \left(\frac{\alpha}{2dn} \mathbb{1}' A_D \mathbb{1} + (1 - \alpha) \right)^t \frac{1}{n} \mathbb{1} \mathbb{1}' h_1. \end{aligned}$$

Vector $h_1 = \left(\frac{\alpha}{2d} A_D + (1 - \alpha) I \right) h_0$ is the evolution of h_0 after one time step according to the normal discrete heat equation. Let d_{tot} be the sum of all vertex degrees in (D, E_D) . Notice $\mathbb{1}' A_D \mathbb{1} = d_{tot}$ so

$$\mathbb{E}[\mathbf{h}_{t+1}] = \left(\frac{\alpha}{2dn} d_{tot} + (1 - \alpha) \right)^t \frac{1}{n} \mathbb{1} \mathbb{1}' h_1.$$

As mentioned in Section 2.1.2, graph (D, E_D) is a proper subgraph of a $2d$ regular graph. Note that d_{tot} is equal to $2dn$ minus the number of edges between D and ∂D , which we will denote as $E(D, \partial D)$. Applying this to the above yields

$$\mathbb{E}[\mathbf{h}_{t+1}] = \left(1 - \alpha \frac{E(D, \partial D)}{2dn}\right)^t \frac{1}{n} \mathbb{1} \mathbb{1}' h_1 \quad (12)$$

3.2 Convergence in Expectation

Convergence for the shuffled heat equation hinges on whether the multiplicative constant $1 - \alpha \frac{E(D, \partial D)}{2dn}$ is less than 1. The sum of degrees in any graph is non-negative, so $d_{tot} = 2dn - E(D, \partial D)$ implies $E(D, \partial D) \leq 2dn$. Consequently, any non-negative initial vector h_0 will converge in expectation to $0 \in \mathbb{R}^n$ for $\alpha \in (0, 1]$ and $\frac{1}{n} \mathbb{1} \mathbb{1}' h_0$ for $\alpha = 0$.

3.3 Convergence in Variance

LEMMA 4. *Let $\{X_k\}$ be a non-negative bounded sequence of random values with $\lim_{k \rightarrow \infty} \mathbb{E}[X_k] = 0$, then $\lim_{k \rightarrow \infty} \text{Var}(X_k) = 0$.*

Proof.

$$\text{Var}(X_k) = \mathbb{E}[X_k^2] - \mathbb{E}[X_k]^2 = \int_0^M x^2 f_{X_k}(x) dx - \mathbb{E}[X_k]^2,$$

where f_{X_k} is the probability density function of X_k and M is a bound such that $X_k \leq M$ for all $k \in \mathbb{N}$. Next note

$$f_{X_k}(x) \leq \int_x^\infty f_{X_k}(y) dy = \mathbb{P}(X_k \geq x).$$

Markov's inequality on $\mathbb{P}(X_k \geq x)$ is defined for $x > 0$ but with a little legwork we can show

$$\begin{aligned} \int_0^M x^2 \mathbb{P}(X_k \geq x) dx &= \lim_{\delta \rightarrow 0^+} \int_\delta^M x^2 \mathbb{P}(X_k \geq x) dx \\ &\leq \lim_{\delta \rightarrow 0^+} \int_\delta^M x \mathbb{E}[X_k] dx && \text{(Markov's Ineq.)} \\ &= \lim_{\delta \rightarrow 0^+} \mathbb{E}[X_k] (M^2 - \delta^2)/2 = \mathbb{E}[X_k] (M^2/2). \end{aligned}$$

Applying the above to $\text{Var}(X_k)$,

$$\text{Var}(X_k) \leq \int_0^M x^2 \mathbb{P}(X_k \geq x) dx - \mathbb{E}[X_k]^2 \leq \mathbb{E}[X_k] (M^2/2) - \mathbb{E}[X_k]^2.$$

For $\epsilon > 0$ find $K \in \mathbb{N}$ such that $k \geq K$ implies $\mathbb{E}[X_k] < \epsilon$. With the same $k \geq K$,

$$|\text{Var}(X_k) - 0| \leq |(M^2/2)\epsilon - \epsilon^2| < \max((M^2/2)\epsilon, \epsilon^2).$$

As ϵ can be made arbitrarily small we have $\lim_{k \rightarrow \infty} \text{Var}(X_k) = 0$. \square

Equation (9) is the composition of non-negative terms, so $\{(\mathbf{h}_t)_j\}_{t=1}^\infty$ is a non-negative sequence for $j \in \{1, \dots, n\}$. Let h_0^* be the largest component of h_0 ,

$$\begin{aligned} \mathbf{h}_1 &= M_0 \left(\frac{\alpha}{2d} A_D + (1 - \alpha) I \right) h_0 \leq M_0 \left(\frac{\alpha}{2d} A_D + (1 - \alpha) I \right) \mathbb{1} h_0^* \\ &\leq M_0 \mathbb{1} h_0^* = \mathbb{1} h_0^*. \end{aligned}$$

Iteratively applying $M_i(\frac{\alpha}{2d} A_D + (1 - \alpha) I)$ with the same inequality argument gives $(\mathbf{h}_t)_j \leq h_0^*$ for $j \in \{1, \dots, n\}$. Each $\{(\mathbf{h}_t)_j\}_{t=1}^\infty$ is a non-negative bounded sequence of random values and by Lemma 4 we have $\lim_{t \rightarrow \infty} \text{Var}((\mathbf{h}_t)_j) = 0$.

4 Spectral Radius Analysis

So far we have shown that the standard heat equation and the expectation of the shuffled variant converge to $0 \in \mathbb{R}^n$ for $\alpha \in (0, 1]$. Noting this, we turn our attention to defining the rate of convergence for different discrete time processes.

DEFINITION 1 (Linear Convergence). *Let $\{x_k\}$ be a sequence in \mathbb{R}^n with limit x^* . Sequence $\{x_k\}$ is said to converge linearly to x^* if there exists a rate of convergence $\mu \in (0, 1)$ such that*

$$\lim_{k \rightarrow \infty} \frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|} = \mu. \quad (13)$$

It is straightforward to see that the rate of convergence for (12) is equal to $1 - \alpha \frac{E(D, \partial D)}{2dn}$. The previous example was simple to solve because we had a closed form to work with. For other cases when we do not have a closed form it can be difficult to derive the rate of convergence. Luckily enough finding the rate of convergence for sequences with elements of the form (7) simplifies to computing the spectral radius of the operator generating the sequence. To see this notice that for large t

$$\begin{aligned} \|\mathbf{h}_{t+1}\| &= \left\| \sum_{i=1}^n \langle v_i, h_0 \rangle \lambda_i^{t+1} v_i \right\| \approx \|\langle v_*, h_0 \rangle \lambda_*^{t+1} v_*\| \\ &= \rho(B)^{t+1} \|\langle v_*, h_0 \rangle v_*\|, \end{aligned} \quad (14)$$

where v_* is the eigenvector associated with largest magnitude eigenvalue λ_* . Plugging the above into (13) gives equality between the rate of convergence and the spectral radius.

There is a slight sleight of hand in (14) where it is implicitly assumed that coefficient $\langle v_*, h_0 \rangle \neq 0$ and eigenvector v_* uniquely attains the spectral radius. In the next subsection we will confirm (12) and (6) satisfy these assumptions for $\alpha \in (0, 1)$.

4.1 Spectral Radius Eigenvector Uniqueness and Positivity

4.1.1 Shuffled Heat Equation

Starting with vector h_1 , the time evolution operator that which generates (12) is $(1 - \alpha \frac{E(D, \partial D)}{2dn}) \frac{1}{n} \mathbb{1} \mathbb{1}'$. The time evolution operator is symmetric and has positive eigenvector $(1/\sqrt{n}) \mathbb{1}$. The rest of the $n - 1$ eigenvectors lie in the subspace $\mathbb{1}^\perp$, the orthogonal complement to the ones vector, with eigenvalue 0. Since $(1/\sqrt{n}) \mathbb{1}$ is the eigenvector uniquely attaining the spectral radius and both $(1/\sqrt{n}) \mathbb{1}$ and h_1 are positive, we have equivalence between the rate of convergence and the spectral radius of the time evolution operator for the shuffled heat equation.

4.1.2 Discrete Heat Equation

Previously mentioned in Section 2.1.1, Theorem 1 states if the matrix is non-negative and primitive then there exists exactly one eigenvector attaining the spectral radius. Furthermore this eigenvector and it's corresponding eigenvalue must be positive.

LEMMA 5. *Given a connected graph G , the matrix $a_1 A_G + a_2 I$ is primitive for $a_1, a_2 > 0$.*

Proof. A matrix A is primitive if there exists a natural number k such that all entries $(A^k)_{ij}$ are positive. Lemma 1 implies A_G is irreducible, so for every entry $(A)_{ij}$ there exists some natural number m such that $(A^m)_{ij} > 0$. Let N be the smallest natural number such that for any i, j there exists $k \in \{1, \dots, N\}$ such that $(A^k)_{ij} > 0$. Next note that

$$(a_1 A_G + a_2 I)^N = \sum_{k=0}^N \binom{N}{k} a_1^k a_2^{N-k} A_G^k.$$

Since a_1, a_2 are positive it equivalently follows that for every i, j entry $((a_1 A_G + a_2 I)^N)_{ij}$ is positive. \square

The matrix $B = \frac{\alpha}{2d} A_D + (1 - \alpha) I$ with $\alpha \in (0, 1)$ is a specific case of Lemma 5, so there exists a unique positive eigenvector x attaining the spectral radius of B .

4.2 Spectral Radius Comparison

For an easier comparison both processes are arranged in a standard form with right vector h_1

$$h_{t+1} = \left(\frac{\alpha}{2d} A_D + (1 - \alpha) I \right)^t h_1,$$

$$\mathbb{E}[h_{t+1}] = \left(\frac{1}{\sqrt{n}} \mathbb{1}' \left(\frac{\alpha}{2d} A_D + (1 - \alpha) I \right) \frac{1}{\sqrt{n}} \mathbb{1} \right)^t \frac{1}{n} \mathbb{1} \mathbb{1}' h_1.$$

We introduce the shorthand matrix C to shorten the expression for $\mathbb{E}[\mathbf{h}_{t+1}]$ above

$$C = \left(\frac{1}{\sqrt{n}} \mathbb{1}' \left(\frac{\alpha}{2d} A_D + (1 - \alpha) I \right) \frac{1}{\sqrt{n}} \mathbb{1} \right) \frac{1}{n} \mathbb{1} \mathbb{1}',$$

$$\mathbb{E}[\mathbf{h}_{t+1}] = C^t \mathbf{h}_1.$$

Through some algebraic manipulations, an inequality can be produced between $\rho(C)$ and $\rho(B)$

$$\rho(C) = \frac{\langle \mathbb{1}, C \mathbb{1} \rangle}{\langle \mathbb{1}, \mathbb{1} \rangle} = \frac{\langle \frac{1}{\sqrt{n}} \mathbb{1}, B \frac{1}{\sqrt{n}} \mathbb{1} \rangle}{\langle \frac{1}{\sqrt{n}} \mathbb{1}, \frac{1}{\sqrt{n}} \mathbb{1} \rangle} \leq \sup_{x \neq 0} \frac{\langle x, Bx \rangle}{\langle x, x \rangle} = \rho(B).$$

The above inequality is actually strict as $\mathbb{1}$ cannot be an eigenvector of B . To prove this it is sufficient to show $\mathbb{1}$ is not an eigenvector of A_D . The matrix-vector product $A_D \mathbb{1}$ produces a vector of vertex degrees. For any vertex $u \in D$

$$\deg(u) = \begin{cases} < 2d, & \text{if } \text{dist}(u, \partial D) = 1. \\ 2d, & \text{otherwise.} \end{cases}$$

As there exists no scalar λ such that $A_D \mathbb{1} = \lambda \mathbb{1}$, the following strict inequality is obtained

$$\rho(C) < \rho(B). \tag{15}$$

Since $\rho(C^t) = \rho(C)^t$ and $\rho(B^t) = \rho(B)^t$, the above holds for C^t, B^t with $t \in \mathbb{N}$.

5 Bounds on the Rates of Convergence

In this section we will be looking at d -dimensional cubes in \mathbb{Z}^d which are subsets of D . Each element of D is considered a d -dimensional cube with side length 1, so as long as D is not empty a d -dimensional cube of side length $\ell \geq 1$ can always be found.

We will denote largest d -dimensional cube found in D by S_D (refer to Figure 3 for a visualization). It will be useful to relate the volume of S_D to the total volume of D through a proportionality constant γ . Formally this relation can be written as

$$\ell^d = \gamma n. \tag{16}$$

5.1 Upper Bound on $\rho(C)$

In Section 4.1.1 an equality between the spectral radius of the time evolution operator C and the rate of convergence for the shuffled heat equation was established

$$\rho(C) = 1 - \alpha \frac{E(D, \partial D)}{2dn}. \tag{17}$$

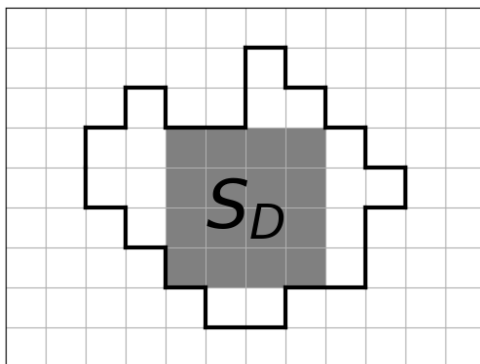


Figure 3: One choice of S_D (gray) with $\ell = 4$ for previous example region

In \mathbb{Z}^d the term $E(D, \partial D)$ can be interpreted as the surface area of the region D . Term $E(D, \partial D)$ can be bounded by noticing the surface area of D will always be greater than or equal to the surface area of S_D . The surface area of a d -dimensional cube of side length ℓ is $2d\ell^{d-1}$ so

$$E(D, \partial D) \geq 2d\ell^{d-1} = 2d\gamma^{\frac{d-1}{d}} n^{\frac{d-1}{d}}. \quad (18)$$

Applying the above relation to (17) yields the bound

$$\rho(C_1) \leq 1 - \alpha\gamma^{\frac{d-1}{d}} \frac{1}{n^{1/d}}. \quad (19)$$

This inequality is strict in cases where D is not a d -dimensional cube with side lengths equal to $n^{1/d}$.

A tighter bound for (18) can be obtained by utilizing the isoperimetric inequality for graphs in \mathbb{Z}^d found in [13] [1]. This isoperimetric inequality differs from (18) in that it does not require proportionality constant $\gamma^{\frac{d-1}{d}}$ to maintain the inequality between the surface area and volume of graph (D, E_D) . Surprisingly, this constant does not affect the asymptotic behavior of $\rho(C)$ for large n and will find a parallel use in bounding $\rho(B)$.

5.2 Lower Bound on $\rho(B)$

First we show a useful property of Cartesian graph products in the context of spectral radii.

LEMMA 6. *Let G_1, G_2 be graphs of size r, s respectively. If $A_{G_1}x_i = \lambda_i x_i$ and $A_{G_2}v_j = \mu_j v_j$, then the Kronecker product $x_i \otimes v_j$ is an eigenvector of the graph Cartesian product $G_1 \square G_2$ with eigenvalue $\lambda_i + \mu_j$.*

Proof. Referencing [7], the adjacency matrix of a graph formed through a Cartesian graph product can be represented as the Kronecker sum of the component

adjacency matrices

$$A_{G_1 \square G_2} = A_{G_1} \otimes I_r + A_{G_2} \otimes I_s.$$

Using the Mixed Product and Bilinear properties of Kronecker products

$$\begin{aligned} A_{G_1 \square G_2}(x_i \otimes v_j) &= (A_{G_1} \otimes I_r + A_{G_2} \otimes I_s)(x_i \otimes v_j) \\ &= (A_{G_1} \otimes I_r)(x_i \otimes v_j) + (A_{G_2} \otimes I_s)(x_i \otimes v_j) \quad (\text{Bilinearity}) \\ &= (A_{G_1} x_i) \otimes v_j + x_i \otimes (A_{G_2} v_j) \quad (\text{Mixed Product}) \\ &= (\lambda_i x_i) \otimes v_j + x_i \otimes (\mu_j v_j) \\ &= (\lambda_i + \mu_j)(x_i \otimes v_j) \quad (\text{Bilinearity}). \end{aligned}$$

□

A d -dimensional cube with side length ℓ is equivalent to the Cartesian product of d path graphs of length ℓ . The eigenvalues of path graphs are well-known and take the form

$$\lambda_k = 2 \cos \left(\frac{k\pi}{\ell+1} \right), \quad k \in \{1, \dots, \ell\}.$$

Graph (S_D, E_{S_D}) is a subgraph to (D, E_D) , which by Lemma 2 gives a lower bound for $\rho(A_D)$. The result from Lemma 6 holds for any finite number of Cartesian graph products so the largest positive eigenvalue of (S_D, E_{S_D}) equals

$$\lambda_{\max} = 2d \cos \left(\frac{\pi}{\ell+1} \right).$$

Applying this bound to $\rho(B)$,

$$\begin{aligned} \rho(B) &= \frac{\alpha}{2d} \rho(A_D) + (1 - \alpha) \rho(I) \geq \alpha \cos \left(\frac{\pi}{\ell+1} \right) + 1 - \alpha \\ &= 1 - \alpha \left(1 - \cos \left(\frac{\pi}{\ell+1} \right) \right). \end{aligned} \quad (20)$$

Next we expand $\cos(\pi/(\ell+1))$ about 0 with Taylor's Remainder Theorem,

$$\cos \left(\frac{\pi}{\ell+1} \right) = 1 - \left(\frac{\pi}{2!(\ell+1)} \right) + R_4 \left(\frac{\pi}{\ell+1} \right),$$

where $R_4(\pi/(\ell+1))$ is the remainder term for the expansion. Plugging in the expansion for $\cos(\pi/(\ell+1))$

$$\rho(B) \geq 1 - \alpha \frac{\pi^2}{2(\ell+1)^2} + \alpha R_4(\pi/(\ell+1)).$$

The remainder term is non-negative for $\ell \geq 1$ since

$$R_4(\pi/(\ell+1)) \geq \inf_{x \in [0, \pi/(\ell+1)]} \frac{\cos(x)}{4!} \left(\frac{\pi}{\ell+1} \right)^4 \geq 0.$$

Dropping the remainder term maintains

$$\rho(B) \geq 1 - \alpha \frac{\pi^2}{2(\ell + 1)^2}.$$

Using volume relation (16) and the fact $-1/(\ell + 1)^2 > -1/\ell^2$ for $\ell \geq 1$,

$$\rho(B) > 1 - \alpha \frac{\pi^2}{2\gamma^{2/d}} \frac{1}{n^{2/d}}. \quad (21)$$

5.3 Asymptotic Rates of Convergence for Shape-Preserved Growth

The bounds solved for in Section 5.1 and 5.2 are parametrized by two geometrical values, the volume of region D and the isoperimetric constant γ . Ideally we would like our bounds to grow solely with the volume n without having to worry about the behavior of γ . There is also the question of how will we maintain the shape of D while increasing volume n .

To combat these problems we introduce a method to increase the volume of a region in \mathbb{Z}^d that preserves the original shape of the region. This method of shape-preserved growth is a simple redefinition of the unit length in \mathbb{Z}^d . We define the new unit length of \mathbb{Z}^d such that the old unit length is an integral multiple of the new length. An example region after this redefinition is given in Figure 4. Take note that the grid mesh in Figure 4 is significantly more refined than the mesh present Figure 3.

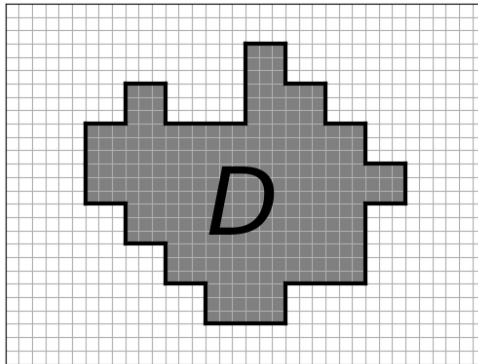


Figure 4: Unit length redefinition with $m = 3$ for previous example region

Let D' be the enlarged region produced by the shape-preserved growth method for integral multiple $m \in \mathbb{N}$. Similarly let n' be the volume of D' and ℓ' the side length of d -dimensional cube $S_{D'}$. Relating the quantities before and after a unit length redefinition gives

$$\begin{aligned} \ell' &= m\ell, \\ n' &= m^d n. \end{aligned}$$

A convenient consequence of this method is that γ is left unchanged after each length redefinition

$$(\ell')^d = m^d \ell^d = m^d \gamma n = \gamma n'.$$

Since γ is kept constant, the fractional n dependence in (19) and (21) correctly describes growth of the spectral radius for each process as D increases in volume. This leads to the following asymptotic rates of convergence for the shuffled and standard discrete heat equation

$$\rho(C) = 1 - \Omega(1/|D|^{1/d}), \quad (22)$$

$$\rho(B) = 1 - \mathcal{O}(1/|D|^{2/d}), \quad (23)$$

where the notation $n = |D|$ from Section 2.1 is reintroduced.

Equations (22) and (23) allow us to write down worst/best case functional forms for the rate of convergence of the shuffled/standard discrete heat equation without having to know the specific geometry of D . The similar forms of (22) and (23) makes it simple see that, for every region $D \subset \mathbb{Z}^d$, the asymptotic worst case rate of convergence for the shuffled heat equation will always be faster than the asymptotic best case rate of convergence for the standard discrete equation. This growth difference between the rates of convergence respects a previously solved inequality (15) in Section 4.2.

6 Application to Numerical Simulations

Suppose we are prototyping a three dimensional model specialized for dissipating heat. This model must satisfy certain geometric constraints and achieve an average temperature of s after some time T . The cooling constraint is tested through simulation with a finite difference method specified by the equation (6).

In order to use the finite difference method the original model in \mathbb{R}^3 is approximated in \mathbb{Z}^3 according to some user defined step size. Performing an exhaustive search of possible models is straightforward in \mathbb{Z}^3 but the computational complexity of such a search will depend on symmetries present in the geometric constraints. For complicated constraints this model generation step may take a long time, by optimizing the simulation step used in testing the cooling constraint we hope to save computational time and improve tractability. For the large T regime the average temperature can be approximated as

$$\frac{1}{|D|} \langle h_T, \mathbb{1} \rangle \approx \frac{\rho(B)^T}{|D|} \langle h_0, \mathbb{1} \rangle.$$

If T_{crit} is the critical time where the average of h_T equals s then,

$$T_{\text{crit}} = \log \left(\frac{s}{\langle h_0, \mathbb{1} \rangle} |D| \right) / \log(\rho(B)).$$

One issue with the above is that for large $|D|$ finding $\rho(B)$ exactly can be very computationally intensive. To combat this one can use (20) to produce a bound

in terms of ℓ (defined in Section 5),

$$T_{\text{crit}} \leq \log \left(\frac{s}{\langle h_0, \mathbb{1} \rangle |D|} \right) / \log \left(1 - \alpha \left(1 - \cos \left(\frac{\pi}{\ell + 1} \right) \right) \right). \quad (24)$$

Although ℓ was originally defined as the side length of the largest cube found inside $D \subset \mathbb{Z}^3$, bound (24) will also hold for ℓ associated to any smaller cubes found inside D . This second option would produce a looser bound on T_{crit} but can be used to avoid doing an extensive search of cubes within D . Depending on how one iterates through models in \mathbb{Z}^3 , the largest cube information may also be obtained in the model generation step.

If the bound on T_{crit} is less than T then region D will satisfy the cooling constraint for the large T regime. By comparing (24) to T we obtain a time efficient veto mechanism on whether a region $D \subset \mathbb{Z}^3$ satisfies the cooling constraint.

7 Future Work with the Shuffled Heat Equation

We have shown that shuffling elements during a heat equation dispersion process greatly speeds up the rate at which the process reaches equilibrium. In order to aid our physical intuition as to why this may be, we consider vector h_t to be the concentration of some solute dissolved throughout D . This mixture is connected to an infinite reservoir ∂D containing the same solvent that is within D . If this solute is left to sit in D its concentration will disperse according to the heat equation [14].

Suppose a strong mixing force was introduced to our domain D . Intuitively we would expect this mixing to aid the dispersed solute to reach the equilibrium state faster. In a discrete setting this strong physical mixing force shares similarities with the shuffling term introduced in Section 3. Equation (8) can be seen as discrete dispersion under the presence of a strong mixing force. In the presence of turbulent flow, a particle's motion can be modeled as a random walk [3]. Clearly a random walk is not the same as intermittent shuffling but one may view shuffling as a simplification to the random walk model after some modifications.

The main discrepancy between the random walk model and the shuffled model is that a random walk always propagates locally from some starting position while the shuffled model allows for nonlocal transport of elements across D . For sufficiently strong mixing where the variance of the random walk far surpasses the diameter of the domain D , it may be possible to excuse this nonlocal behavior but as a result we will be constrained to working with a very specific regime of mixing processes.

For future work it would be worthwhile to consider shuffling that is limited to displacing elements a maximum of r units away, where r is some predefined mixing length depending on the scale of the time step. One approach could be to select a random partition P of set D at each time step and require that each cell of P is connected and at most diameter r . Similar to the previous model we allow for uniform shuffling of elements within each cell of the partition.

Depending on the analysis carried out, one may define each random partition to be sampled with equal probability or allow for unequal sampling probabilities for different classes of possible partitions.

References

- [1] R. Ahlswede, S.L. Bezrukov. Edge Isoperimetric Theorems for Integer Point Arrays. *Applied Mathematics Letters*, 8(2):75–80, 1995.
- [2] N. Blair-Stahn. Random Permutation Matrices: An Investigation of the Number of Eigenvalues Lying in a Shrinking Interval. Report, University of Arizona, Tucson, AZ, 2000.
- [3] H. Chanson. 7 - Turbulent dispersion and mixing: 1. Vertical and transverse mixing. *Environmental Hydraulics of Open Channel Flows*, Butterworth-Heinemann, Oxford, 2004, pp. 81–98.
- [4] F. Escolano. Heat Kernels and Diffusion Processes, Matrix Computing subject 3168. Lecture notes, University of Alicante, Alicante, 2012.
- [5] W.G. Faris. Lectures on Partial Differential Equations. Lecture notes, University of Arizona, Tucson, AZ, 2004.
- [6] K. Hu. Partial Differential Equations and Random Walks, with Emphasis on the Heat Equation, Partial Differential Equations MAT394. Lecture notes, University of Toronto, Toronto, 2014.
- [7] A. Kaveh, H. Rahami. A unified method for eigendecomposition of graph products. *Commun. Numer. Meth. Engng.*, 21:377–388, 2005.
- [8] R.I. Kondor, J. Lafferty. Diffusion Kernels on Graphs and Other Discrete Structures. In *In Proceedings of the ICML*, pp. 315–322, 2002.
- [9] R.S. Laugesen. Spectral Theory of Partial Differential Equations. Lecture notes, University of Illinois, Urbana–Champaign, IL, 2012.
- [10] G.F. Lawler, *Random Walk and the Heat Equation*, The Student Mathematical Library 55, Chicago, IL, 2010.
- [11] M. Reed, B. Simon. *IV: Analysis of Operators*, volume 4. Elsevier, 1978, p. 271.
- [12] Seva. Spectral radius of a proper subgraph. MathOverflow, <https://mathoverflow.net/q/69855>, (updated July 16, 2011).
- [13] J. Tillich. Edge isoperimetric inequalities for product graphs. *Discrete Mathematics*, 213(1):291–320, 2000.
- [14] J.L. Wilson. Partial Differential Equations (PDEs), Quantitative Methods in Hydrology HYD510. Lecture notes, New Mexico Tech, Socorro, NM, 2010.
- [15] J. Xu, H. Zhang, L. Zikatanov. On the Weyl’s law for discretized elliptic operators. Preprint, <https://arxiv.org/pdf/1705.07803.pdf>, (updated October 31, 2019).