The Geometry of Rounding *

Jason Vander Woude[1], Peter Dixon[†2], A. Pavan[3], Jamie Radcliffe[4], and N. V. Vinodchandran[5]

[1]*School of Computing and Department of Mathematics, University of Nebraska-Lincoln*
jasonvw@huskers.unl.edu
[2]*Ben-Gurion University of the Negev*
tooplark@gmail.com
[3]*Department of Computer Science, Iowa State University*
pavan@cs.iastate.edu
[4]*Department of Mathematics, University of Nebraska-Lincoln*
aradcliffe1@unl.edu
[5]*School of Computing, University of Nebraska-Lincoln*
vinod@cse.unl.edu

October 31, 2022

Abstract

Rounding has proven to be a fundamental tool in theoretical computer science. By observing that rounding and partitioning of $\mathbb{R}^d$ are equivalent, we introduce the following natural partition problem which we call the *secluded hypercube partition problem*: Given $k \in \mathbb{N}$ (ideally small) and $\epsilon > 0$ (ideally large), is there a partition of $\mathbb{R}^d$ with unit hypercubes such that for every point $\vec{p} \in \mathbb{R}^d$, its closed ϵ-neighborhood (in the ℓ_∞ norm) intersects at most k hypercubes?

We undertake a comprehensive study of this partition problem. We prove that for every $d \in \mathbb{N}$, there is an explicit (and efficiently computable) hypercube partition of $\mathbb{R}^d$ with $k = d + 1$ and $\epsilon = \frac{1}{2d}$. We complement this construction by proving that the value of $k = d + 1$ is the best possible (for any ϵ) for a broad class of "reasonable" partitions including hypercube partitions. We also investigate the optimality of the parameter ϵ and prove that any partition in this broad class that has $k = d + 1$, must have $\epsilon \leq \frac{1}{2\sqrt{d}}$. These bounds imply limitations of certain deterministic rounding schemes existing in the literature. Furthermore, this general bound is based on the currently known lower bounds for the dissection number of the cube, and improvements to this bound will yield improvements to our bounds.

While our work is motivated by the desire to understand rounding algorithms, one of our main conceptual contributions is the introduction of the *secluded hypercube partition problem*, which fits well with a long history of investigations by mathematicians on various hypercube partitions/tilings of Euclidean space.

Keywords. Rounding, Partition, Tiling, Packing, Tesselation, Cube, Hypercube, Reclusive, Sperner, Dissection, Triangulation

---

Contents

1 Introduction

Rounding has proven to be a fundamental tool in theoretical computer science. Generically, if $\mathcal{X} \subseteq \mathcal{Y}$ are metric spaces with the same metric, then a rounding algorithm $\mathcal{R}$ maps points in $\mathcal{X}$ to points in $\mathcal{Y}$ (typically so that x is close to $\mathcal{R}(x)$). Intuitively, the purpose of rounding is to limit the number of possible outcomes of the algorithm, which in turn can help reduce the complexity (e.g. space complexity, time complexity, or the number of random bits needed). Often $\mathcal{X} = \mathbb{R}^d$ with the metric induced by the l^∞ norm (we call this metric d_{max}).

In complexity theory, one of the earliest applications of rounding is in the seminal work of Saks and Zhou [1] in the context of space-bounded derandomization. They devised a randomized algorithm $\mathcal{R}$ rounding from $[0,1]^d$ to $[0,1]^d$ with the following property: for every $\vec{x} \in [0,1]^d$,

$$\Pr_{r \in \Sigma^\ell} \left[\forall \vec{y} \in B_\epsilon(\vec{x}) : \mathcal{R}(r, \vec{x}) = \mathcal{R}(r, \vec{y}) \right] \geq 1 - 1/\mathsf{poly}(d)$$

and for any $r \in \Sigma^\ell$, $d_{max}(\vec{x}, \mathcal{R}(r, \vec{x})) < \epsilon$.

Here, $B_\epsilon(\vec{x})$ is the open ϵ-ball around $\vec{x}$ with respect to the l^∞ norm and $\Sigma^\ell = \{0,1\}^\ell$ is the sample space of random bits. The property above expresses that for any ϵ-ball, with high probability over the choice of randomness, all points in the ball round to the same value. The Saks and Zhou rounding scheme is a randomized rounding scheme with $\epsilon = 1/\mathsf{poly}(d)$ using $\ell = O(\log d)$ bits of randomness. This randomized rounding is a critical step in the well known derandomization of probabilistic space-bounded algorithms, namely $\mathsf{BPSPACE}(S) \subseteq \mathsf{DSPACE}(S^{3/2})$.

A more recent example in which randomized rounding is employed is in the context of pseudodeterminism and multi-pseudodeterminism [2, 3, 4]. A pseudodetermistic algorithm is a randomized algorithm that on any particular input, returns a canonical solution with probability at least $2/3$ (which can be boosted to $1 - \delta$). A k-pseudodeterministic algorithm, a generalization of a pseudodeterministic algorithm, is a probabilistic algorithm such that for each input x, there is a set S_x of cardinality at most k, and on input x, the algorithm returns a value in S_x with probability at least $\frac{k+1}{k+2}$ (which again can be boosted to $1 - \delta$). In the context of designing multi-pseudodeterministic algorithms, Goldreich [3] designed a randomized algorithm $\mathcal{R}$ with the following property: for every $\vec{x} \in \mathbb{R}^d$, there exists $S_{\vec{x}}$ of cardinality at most $k = d+1$ such that

$$\Pr_{r \in \Sigma^\ell} [\forall \vec{y} \in B_\epsilon(\vec{x}) : \mathcal{R}(r, \vec{y}) \in S_{\vec{x}}] \geq \frac{k+1}{k+2}$$

and for any $r \in \Sigma^\ell$, $d_{max}(\vec{x}, \mathcal{R}(r, \vec{x})) < \epsilon$.

Goldreich's rounding has a different requirement than Saks and Zhou's: for any ϵ-ball, Goldreich does not require that with high probability all points in the ball round to the same value, but he does require that there is a small set such that with high probability all points in the ball round to points in that set. Also, in Saks and Zhou's scheme, the single value that the ball is rounded to can depend on random bits r, but Goldreich's scheme requires the set $S_{\vec{x}}$ to be independent of the randomness. Using this rounding scheme, Goldreich showed the existence of $(d+1)$-pseudodeterministic approximation algorithms for a class of functions whose range is $\mathbb{R}^d$. Grossman and Liu used a similar rounding scheme to design low *influential-bit* algorithms, a notion that generalizes pseudodeterministic algorithms [4]. Similar rounding schemes have found applications in a very recent work of Impagliazzo, Lei, Pitassi, and Sorrell [5] in the context of making statistical learning algorithms reproducible.

In [6], Hoza and Klivans designed a certain *deterministic* rounding scheme. Their work is motivated by the problem of reducing randomness for adaptive algorithms. Consider the problem of simulating m adaptive (adversarially chosen) invocations of a randomized estimation algorithm that returns a vector in $\mathbb{R}^d$. If the algorithm requires n bits of randomness, then the trivial way of simulating m invocations requires nm random bits. To reduce the randomness, they designed a deterministic rounding algorithm $\mathcal{R}$ that has the same properties as that of Goldreich. In other words, for every $\vec{x} \in \mathbb{R}^d$, the size of the

1

set $\{\mathcal{R}(\vec{y}) : \vec{y} \in B_\epsilon(\vec{x})\}$ has cardinality at most $d + 1$. Combining this rounding algorithm with the INW pseudorandom generator [7], Hoza and Klivans reduced the amount of randomness required for the above adaptive invocation problem to $n + O(m \log(d + 1))$. We give more elaborate details on all these rounding schemes and how they fit into our work in Appendix A.

The present work investigates *the geometry of rounding*. We equate the notion of deterministic rounding schemes to partitions of Euclidean space $\mathbb{R}^d$. This connection led us to the introduction of a very natural partition problem that we call the *secluded hypercube partition problem*. Using this geometric approach, we establish upper and lower bounds on the parameters of certain deterministic rounding schemes. Perhaps more importantly, the partition problem we investigate is very natural and should be of independent interest to a broader community. The introduction of the secluded hypercube partition problem is one of our main conceptual contributions.

Definition 1.1 (($k(d), \epsilon(d)$)-Deterministic Rounding). *A deterministic rounding scheme is a family of functions $\mathcal{F} = \{f_d\}_{d \in \mathbb{N}}$ where $f_d : \mathbb{R}^d \to \mathbb{R}^d$. We call $\mathcal{F}$ a $(k(d), \varepsilon(d))$-deterministic rounding scheme if (1) $\forall \vec{x} \in \mathbb{R}^d$, $d_{max}(\vec{x}, f_d(\vec{x})) \le 1$[1] (2) $\forall \vec{x} \in \mathbb{R}^d$, $|\{f_d(\vec{z}) : \vec{z} \in B_{\varepsilon(d)}(x)\}| \le k(d)$.*

Definition 1.2 ((k, ϵ)-Secluded Partition). *A partition $\mathcal{P}$ of $\mathbb{R}^d$ is called a (k, ϵ)-secluded partition if for every $\vec{p} \in \mathbb{R}^d$, $|\{X \in \mathcal{P} : X \cap \overline{B}_\epsilon(\vec{p}) \ne \emptyset\}| \le k$.*

Remark 1.3. *The use of the word "secluded" is meant as a synonym for "remote" or "isolated" and is meant to indicate that every point in space is only nearby (within ϵ) a few (at most k) members of the partition.*

We refer to k as the *degree* of the partition and ϵ as its *tolerance*. The values of k and ϵ will often be functions of the dimension d.

There is a natural equivalence between rounding schemes and partitions of Euclidean space in a very general sense. A rounding function $f : \mathbb{R}^d \to \mathbb{R}^d$ induces a partition $\mathcal{P}_f$ as follows: $\mathcal{P}_f = \{f^{-1}(y) : y \in \text{range}(f)\}$. Conversely, a partition $\mathcal{P}$ induces a deterministic rounding function $f_\mathcal{P}$ as follows: for each member $X \in \mathcal{P}$ let $\vec{p}_X \in X$ some fixed representative. Then the rounding function $f_\mathcal{P}$ maps any point $\vec{p} \in X$ to $\vec{p}_X$. This connection leads to the following observation.

Observation 1.4 (Equivalence of Rounding Schemes and Partitions). *A $(k(d), \epsilon(d))$-deterministic rounding scheme induces, for each $d \in \mathbb{N}$, a $(k(d), \epsilon(d))$-secluded partition of $\mathbb{R}^d$ in which each member has diameter[2] at most 2. Conversely, a sequence $\langle \mathcal{P}_d \rangle_{d=1}^\infty$ of partitions where $\mathcal{P}_d$ is $(k(d), \epsilon(d))$-secluded and contains only members of diameter at most 1 induces many[3] $(k(d), \epsilon(d))$-deterministic rounding schemes.*

The connection between rounding schemes and partitions of Euclidean space has been observed in earlier works. In particular, the works of Fiege et. al., Kindler et. al., and Braverman and Minzer investigated the foams problem in the context of parallel repetition [8, 9, 10, 11]. The foams problem asks to find a body that tiles $\mathbb{R}^d$ with $\mathbb{Z}^d$ with the smallest surface area. These works established connections between the foams problem and the parallel repetition theorem and gave a new foams construction. This new construction led to an optimal *noise-resistant* randomized rounding scheme (more precisely a distribution $\mathcal{D}$ over deterministic rounding schemes) with the following property for all $d \in \mathbb{N}$: $\forall \vec{x}, \vec{y} \in \mathbb{R}^d$ such that $d_{max}(\vec{x}, \vec{y}) \le \epsilon$, it holds that $\Pr_{f \sim \mathcal{D}}[f(\vec{x}) = f(\vec{y})] \ge 1 - O(\varepsilon)$. In [9, Definition 1.2], a randomized rounding scheme is defined as a distribution of functions for each dimension rather than a single function in each dimension. Not all rounding algorithms necessarily abide by this definition of a randomized rounding scheme, but all of those discussed in this paper do including [1, 3, 9, 10, 6, 5]. Thus, while our results

---

[1]The bound of 1 is not critical. We can use any constant and scale the parameters appropriately.

[2]The diameter is at most 2, not at most 1. For example, the rounding scheme which rounds a point by sending each coordinate to the nearest even integer (breaking ties by rounding up) is a $(2^d, 1)$-deterministic rounding scheme, and the partition it induces in each dimension consists of hypercubes of side length 2.

[3]For each member, any representative of that member can be chosen.

will be about deterministic rounding schemes, there is a connection between deterministic and randomized rounding schemes in this sense.

1.1 Our Contributions

In this work we conduct an extensive investigation of secluded partitions. Contributions of this work are two-fold. The first contribution is the formulation of a very natural partition problem known as the *secluded hypercube partition problem* and an explicit construction of such partition with degree $d+1$ and tolerance $1/2d$. The second contribution is establishing impossibility results on the degree and the tolerance parameters. In particular, we establish that any "reasonable" secluded partition must have degree at least $d+1$ and every reasonable partition with degree $d+1$ must have tolerance at most $1/2\sqrt{d}$.

We start our investigation by considering *unit hypercube partitions* of $\mathbb{R}^d$, which are very natural and extensively studied partitions of Euclidean space. A partition $\mathcal{P}$ of $\mathbb{R}^d$ is a unit hypercube partition if every $X \in \mathcal{P}$ is a d-dimensional unit hypercube. Note that the diameter of a unit hypercube is 1 in the d_{max} metric.

Question (Secluded Hypercube Partition Problem)**.** *Let $d \in \mathbb{N}$. For what values of $k \in \mathbb{N}$ and $\epsilon \in (0, \infty)$ does there exist a (k, ϵ)-secluded unit hypercube partition $\mathcal{P}$ of $\mathbb{R}^d$?*

One of our main conceptual contributions is the formulation of the above question which is a very natural geometric question and should be of broad interest. The question asks: given $d \in \mathbb{N}$, $\epsilon \in (0, \infty)$, and $k \in \mathbb{N}$, is there a partition of $\mathbb{R}^d$ with unit hypercubes so that for any point $\vec{p} \in \mathbb{R}^d$, its ϵ-neighborhood, in the d_{max} metric, intersects with at most k hypercubes? It is easy to see that the natural grid partition of $\mathbb{R}^d$ with unit hypercubes is a $(2^d, 1/2)$-secluded partition. Our first technical contribution is the design of a parameterized class of explicit $(d+1, \frac{1}{2d})$-secluded unit hypercube partitions for each $d \in \mathbb{N}$. (See Theorem 7.20 for a more elaborate statement of the following result and its proof.)

Theorem (Hypercube Partition Theorem)**.** *For all $d \in \mathbb{N}$, there exists $(d+1, \frac{1}{2d})$-secluded unit hypercube partitions of $\mathbb{R}^d$.*

Our next result shows that the above construction is optimal with respect to the degree parameter. In particular, we show that for any unit hypercube partition of $\mathbb{R}^d$ and any $\epsilon \in (0, \infty)$, the degree has to be at least $d + 1$. (See Theorem 9.8 for a slightly stronger version of the following result and its proof.)

Theorem (Degree Optimality for Unit Hypercube Partitions)**.** *For every (k, ϵ)-secluded unit hypercube partition $\mathcal{P}$ of $\mathbb{R}^d$, it must be that $k \geq d + 1$.*

This theorem raises the question of whether there exist non-hypercube partitions with degree smaller than $d+1$. There are trivial examples of such partitions since any partition with fewer than $d+1$ members trivially has degree smaller than $d + 1$ no matter what value of tolerance is used. Another trivial example is concentric "shells" in the d_{max} metric: $\mathcal{P} = \{\overline{B}_1, \ \overline{B}_2 \setminus \overline{B}_1, \ \overline{B}_3 \setminus \overline{B}_2, \ \overline{B}_4 \setminus \overline{B}_3, \ \ldots\}$ (where $\overline{B}_i$ denotes the ball of radius i centered at $\vec{0}$). For tolerance $\epsilon \leq \frac{1}{2}$, this is a $(2, \epsilon)$-secluded partition.

In the first of these examples, since there are a finite number of members, some member has infinite diameter and infinite measure. In the second example, though every member has finite measure and finite diameter, there is no bound on either of these quantities (i.e. for any D and any μ, the partition contains a member with diameter greater than D and measure greater than μ). If we consider partitions with either a bound on the diameter of the members or a bound on the measure of the members, then we can rule out the possibility of degree less than $d + 1$. We say a partition has bounded diameter (resp. bounded measure) if there exists some bound $b \in (0, \infty)$ such that every member of the partition has diameter (resp. measure) at most b. Such b will be called a diameter bound (resp. measure bound) for the partition. Our next contribution is a degree optimality theorem for bounded measure partitions. (See Theorem 9.3 for an equivalent statement of the following result and a proof.)

3

Theorem (Degree Optimality for Bounded Measure Partitions)**.** *For every (k, ϵ)-secluded partition $\mathcal{P}$ of $\mathbb{R}^d$ with bounded measure, it must be that $k \geq d + 1$.*

Since a partition with bounded diameter must also be a partition of bounded measure[4], we get the following corollary. (See Theorem 9.4 for a stronger version of this result and its proof.)

Corollary (Degree Optimality for Bounded Diameter Partitions)**.** *For every (k, ϵ)-secluded partition $\mathcal{P}$ of $\mathbb{R}^d$ with bounded diameter, it must be that $k \geq d + 1$.*

Now we turn from the degree parameter to the tolerance parameter. The Hypercube Partition Theorem above states that there is a (k, ϵ)-secluded partition with degree $k = d + 1$ (the minimum possible for bounded partitions) and tolerance $\epsilon = \frac{1}{2d}$. Is it possible to design (k, ϵ)-secluded partitions maintaining optimal $k = d + 1$, but with larger ϵ? We could just scale up the sizes of the hypercubes in the Hypercube Partition Theorem to obtain any ϵ we want, so this question only makes sense if we compare ϵ to the size of the largest member of the partition. (See Theorem 10.10 for a more detailed version of the following result and its proof.)

Theorem (Tolerance Upper Bound)**.** *For every $(d + 1, \epsilon)$-secluded partition $\mathcal{P}$ of $\mathbb{R}^d$ with diameter bound D, it must be that $\epsilon \leq \frac{D}{2\sqrt{d}}$.*

Our goal was to show an upper bound of $1/\Omega(d)$ on the tolerance parameter as that would establish that our construction is optimal in both degree and tolerance. However, we could only establish this for $d = 1$ and $d = 2$. Closing the gap between our upper and lower bounds on the tolerance parameter is the main open question that emerged from our investigation.

The above results and Observation 1.4 establish a limitation of deterministic rounding schemes as follows.

Corollary 1.5 (Limitations of Deterministic Rounding Schemes)**.** *For any $(k(d), \epsilon(d))$-deterministic rounding scheme, $k(d)$ must be at least $d + 1$, and for $(d + 1, \epsilon(d))$-deterministic rounding schemes, $\epsilon(d)$ must be at most[5] $\frac{1}{\sqrt{d}}$.*

Finally, we show an application of the hypercube partition theorem in the context of sample complexity of pseudodeterministic algorithms. Let $f_1, \ldots, f_d$ be functions from $\{0, 1\}^n$ to $[0, 1]$. Given blackbox access to the functions, the goal is to output a vector $\vec{\alpha} = \langle \alpha_1, \ldots, \alpha_d \rangle$ that is an (ϵ, δ)-approximation to the vector $\langle \mathbb{E}(f_i) \rangle_{i=1}^d$ with respect to d_{max} metric. This means that with probability at least $1 - \delta$, the vector $\vec{\alpha} \in \mathbb{R}^d$ returned by the algorithm is such that $d_{max}(\vec{\alpha}, \langle \mathbb{E}(f_i) \rangle_{i=1}^d) \leq \epsilon$ which is equivalent to saying that with probability at least $1 - \delta$, it holds for all i that $|\alpha_i - \mathbb{E}(f_i)| \leq \epsilon$. Goldreich proved that there is $(d + 1)$-pseudodeterministic algorithm for this task, and this algorithm has a sample complexity of $\widetilde{O}(d^4)$ [3]. Using our hypercube partitions, we show that the sample complexity can be improved to $\widetilde{O}(d^2)$ samples.

Proposition 11.7 (Application to Pseudodeterministic Algorithms)**.** *Let $f_1, \ldots, f_d : \{0, 1\}^n \to [0, 1]$ be functions and $\epsilon \in (0, \infty)$ and $\delta \in (0, \frac{1}{2})$. There is an algorithm that given sample access to each f_i can $(d + 1)$-pseudodeterministically (ϵ, δ)-approximate $\langle \mathbb{E}(f_i) \rangle_{i=1}^d$ relative to the d_{max} metric using $O\left(\frac{(d+1)^2}{\epsilon^2} \cdot \log\left(\frac{d}{\delta}\right)\right)$ samples.*

---

[4]Suppose $\mathcal{P}$ is a partition of $\mathbb{R}^d$ of bounded diameter. Then there is some D such every member $X \in \mathcal{P}$ has $\mathrm{diam}(X) \leq D$. Let $\mu = (2D)^d$. For any $X \in \mathcal{P}$, fix some $\vec{x} \in X$ so $\overline{B}_D(\vec{x}) \supset X$ and note that $\overline{B}_D(\vec{x})$ (in the d_{max} metric) has measure $(2D)^d = \mu$ which shows that each member of $\mathcal{P}$ has measure at most μ, so $\mathcal{P}$ has bounded measure.

[5]Observation 1.4 gives a diameter bound of $D = 2$, so $\frac{D}{2\sqrt{d}} = \frac{1}{\sqrt{d}}$

2 Main Results and Proof Outlines

In this section, we give formal statements and proof outlines for all our results. We start with some required notation. A complete set of notations that we use in this paper is given in Section 5.

For a point $\vec{p} \in \mathbb{R}^d$ and $\varepsilon \in (0, \infty)$, $\overline{B}_\epsilon(\vec{p})$ denotes the closed ball of radius ϵ around $\vec{p}$, with respect to the d_{max} metric (we will use this metric throughout this paper). In other words, $\overline{B}_\epsilon(\vec{p}) \overset{\text{def}}{=} \left\{ \vec{x} \in \mathbb{R}^d : \|\vec{x} - \vec{p}\|_\infty \leq \epsilon \right\} = \left\{ \vec{x} \in \mathbb{R}^d : \forall i \ |x_i - y_i| \leq \epsilon \right\}$.

When it is understood that we are considering a particular partition $\mathcal{P}$, then for any point $\vec{p} \in \mathbb{R}^d$, and any $\epsilon \in (0, \infty)$, we let $\mathcal{N}_\epsilon(\vec{p})$ denote the members of $\mathcal{P}$ intersecting the closed ϵ-neighborhood of $\vec{p}$:

$$\mathcal{N}_\epsilon(\vec{p}) \overset{\text{def}}{=} \left\{ X \in \mathcal{P} : X \cap \overline{B}_\epsilon(\vec{p}) \neq \emptyset \right\}.$$

In addition, we let $\mathcal{N}_{\overline{0}}(\vec{p})$ denote the members of $\mathcal{P}$ whose closure contains $\vec{p}$.

$$\mathcal{N}_{\overline{0}}(\vec{p}) \overset{\text{def}}{=} \left\{ X \in \mathcal{P} : \overline{X} \ni \vec{p} \right\}.$$

These two sets will be ubiquitous throughout the paper. We will restate the Secluded Hypercube Partition Problem using this notation as follows:

Question 2.1 (Secluded Hypercube Partition Problem)**.** *Let $d \in \mathbb{N}$. Given $k \in \mathbb{N}$ (degree) and $\epsilon \in (0, \infty)$ (tolerance) does there exists a partition $\mathcal{P}$ of $\mathbb{R}^d$ consisting of unit hypercubes such that for every point $\vec{p} \in \mathbb{R}^d$, $|\mathcal{N}_\epsilon(\vec{p})| \leq k$?*

This work is a comprehensive study of Question 2.1. It is easy to verify that for $k = 2^d$ and any $\epsilon \leq \frac{1}{2}$ such a partition exists. In particular, the standard grid partition demonstrated in Figure 1 has the desired property.

Can we do better? For $d = 2$ by shifting layers of the standard grid partition, as demonstrated in Figure 2 in ??, we can reduce the degree to 3 and have tolerance of $\frac{1}{4}$. While it is not obvious, for $d = 3$, this shifting method yields a partition with degree 4 and tolerance $\frac{1}{6}$ (again, see Figure 2 .

Figure 1: Simple unit hypercube partitions of $\mathbb{R}^d$ for $d = 1, 2, 3$. The shaded red regions are $\epsilon = \frac{1}{2}$ radius balls (in the d_{max} metric/$\|\cdot\|_\infty$ norm) showing that these neighborhoods don't intersect more than $k = 2^d$ members of the partition (it might look from the picture that centering the ball in one of the hypercubes would intersect more, but it does not because the hypercubes are half open). Furthermore, there are hypercubes that do intersect 2^d members as shown. (This figure was produced using Manim [12].)

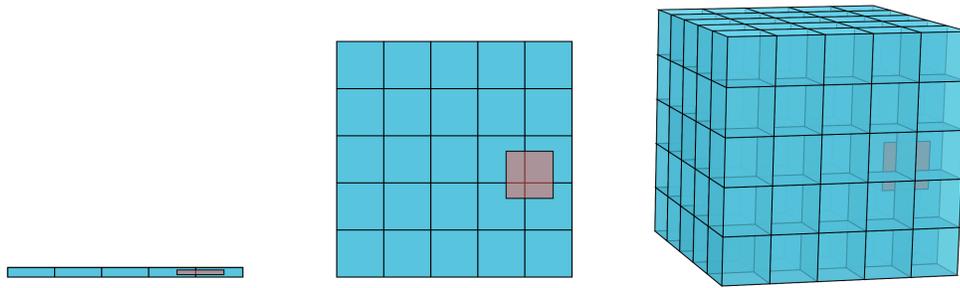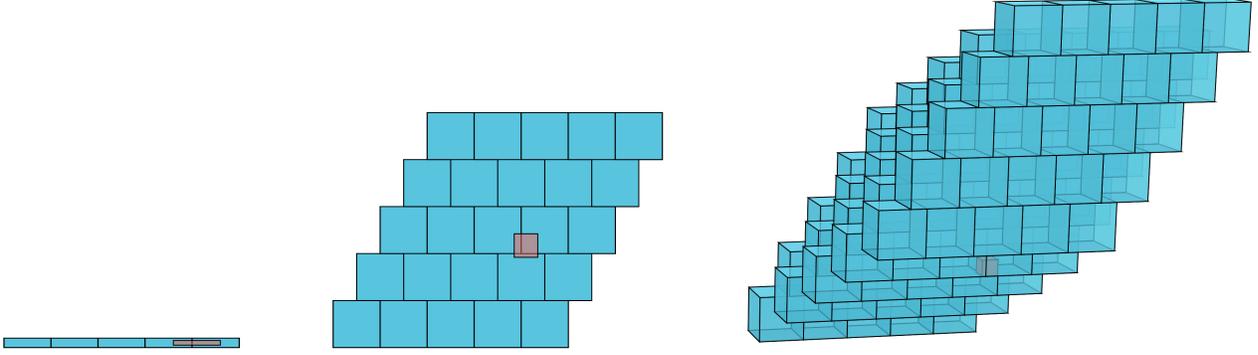Figure 2: This is a particular reclusive partition of $\mathbb{R}^d$ for $d = 1, 2, 3$. The shaded red regions are $\epsilon = \frac{1}{2d}$ radius balls (in the d_{max} metric/$\|\cdot\|_\infty$ norm) showing that these neighborhoods don't intersect more than $k = d + 1$ members of the partition. Observe that in the $d = 2$ case, the x-direction shift (horizontal) is $\frac{1}{2}$ unit; in the $d = 3$ case, there are shifts of $\frac{2}{3}$ as well as $\frac{1}{3}$ in various directions. (This figure was produced using Manim [12].)

Our first result shows that this intuition generalizes to all dimensions. We construct hypercube partitions with degree $d + 1$ and tolerance $\frac{1}{2d}$. We note that, while the intuition about the degree bound generalizes from lower dimensions, we do not believe it was obvious that $\epsilon = \frac{1}{2d}$ was achievable. We show at the end of Section 7 that a slight change to our partitions, while seemingly innocuous, results in significantly worse degree.

Theorem 7.20 (Hypercube Partition Theorem). *Let $d \in \mathbb{N}$. Then there exists a $(d+1, \frac{1}{2d})$-secluded unit hypercube partition $\mathcal{P}$ of $\mathbb{R}^d$. I.e., for any point $\vec{p} \in \mathbb{R}^d$, we have*

$$|\mathcal{N}_{\frac{1}{2d}}(\vec{p})| \leq d + 1.$$

Proof Outline. The proof of this theorem is constructive. We specify a particular basis for the vector space $\mathbb{R}^d$ and consider all linear combinations of these vectors in which all coefficients are integers (sometimes called integer linear combinations). These integer linear combinations are such that a unit hypercube can be placed at each position and it will form a partition of $\mathbb{R}^d$. We view the basis as the columns of a matrix in an appropriate order. We define a class of matrices which we call *reclusive matrices*—these are upper triangular matrices with 1's on the diagonal, and in each row the values that appear after the diagonal entry (inclusive) are decreasing. Given such a reclusive matrix A, we associate a lattice group L_A which is the set of vectors of the form $A\vec{v}$ where $\vec{v} \in \mathbb{Z}^d$. Now the members of the partition are the hypercube shifted by elements of the lattice group L_A. We prove that every reclusive matrix induces a $(d+1, \epsilon)$-secluded unit hypercube partition where the value of ϵ depends on a property of the matrix called the *reclusive distance* Δ.

For two hypercubes X, Y in this partition, we consider the vectors $\vec{n}$ and $\vec{m}$ (viewed as integer sequences) which define the positions of X and Y relative to the chosen basis. We show that X and Y are adjacent if and only if the vector/sequence $\vec{n} - \vec{m}$ is an alternating sequence of -1's and 1's padded by 0's (what we call *weak-alt-1 sequences*). We use this characterization to prove that for Δ, two members of the partition are either adjacent or their positions are at least Δ far apart. This means that for any ball of radius $\frac{\Delta}{2}$, all partition members that it intersects must be pairwise adjacent. I.e., they constitute a clique in the graph theory sense when the partition is viewed as an infinite graph with members being vertices and edges representing adjacency[6] of members. Because of this, it suffices to show that there are no cliques of size greater than $d + 1$ in this partition graph. We do this by giving an explicit graph coloring of the

---

[6]Members are adjacent if their closures intersect.

partition using $d+1$ colors. This will show that the partition defined is a $(d+1, \Delta/2)$-secluded partition. Finally, it is easy to construct (we give examples) reclusive matrices with $\Delta = 1/d$.

$\square$

Recall from the introduction that a deterministic rounding scheme yields a partition of $\mathbb{R}^d$. Thus using the rounding scheme of Hoza and Klivans [6], we can obtain a $(d+1, \frac{1}{6(d+1)})$-secluded partition of $\mathbb{R}^d$ whose diameter is bounded above by 1. However, the members of this partition are not unit hypercubes (See Figure 3b in Subsection A.4 for $d = 2$). For the sake of completeness, we provide a proof of the theorem in Subsection A.4.

Theorem 2.2 (Hoza-Klivans Partition). *Let $d \in \mathbb{N}$. Then there exists a $(d+1, \frac{1}{6(d+1)})$-secluded partition $\mathcal{P}$ of $\mathbb{R}^d$ with diameter at most 1.*

2.1 Lower Bound on Degree (k)

Our next set of results investigate the optimality of the degree parameter (k) of partitions. We first prove that the value of $d + 1$ in Theorem 7.20 is indeed optimal for hypercube partitions.

Theorem 9.8 (Optimality for Unit Hypercube Partitions). *Let $d \in \mathbb{N}$ and $\mathcal{P}$ a unit hypercube partition of $\mathbb{R}^d$. Then there exists $\vec{p} \in \mathbb{R}^d$ such that*

$$|\mathcal{N}_{\vec{0}}(\vec{p})| \geq d + 1$$

Furthermore, $\mathcal{P}$ contains a $(d+1)$-clique.

The above theorem states that there is a particular $\vec{p} \in \mathbb{R}^d$ so that for every choice of ϵ, $B_\epsilon(\vec{p})$ intersects at least $d+1$ members of the partition, and thus the value $d+1$ from Theorem 7.20 cannot be made any smaller. A $(d+1)$-clique in $\mathcal{P}$ means that there is a set of $d+1$ members of $\mathcal{P}$ that have pairwise intersecting closures (see Definition 6.1), and its existence follows trivially from the first claim of the theorem. This result is stronger than just claiming that such a partition $\mathcal{P}$ must have degree parameter at least $d+1$ (see the upcoming discussion of the three Optimality Theorems (or Subsection 9.1 and Subsection 9.4) for details.)

As discussed in the prior section, in generalizing from hypercubes to more general partitions, we have to impose some condition to avoid trivialities, and we previously mentioned partitions with bounded measure or bounded diameter. We also consider one other condition which imposes a type of local finiteness. Thus, we prove three different degree optimality theorems. Each has a slightly different requirement on the types of partitions under consideration (each bounding the size of the partition members in some way) and each has a slightly different sense in which the partitions are considered optimal.

In the statements below, m denotes the Lebesgue measure (intuitively the volume of a set). We present all three theorems before discussing the proof outlines. The First Optimality Theorem states that for partitions of $\mathbb{R}^d$ in which all members are bounded in measure, then there is no secluded partition with degree less than $d+1$, because no matter the value of ϵ, some point $\vec{p}$ can be found such that $|\mathcal{N}_\epsilon(\vec{p})| \geq d+1$.

Theorem 9.3 (First Optimality Theorem). *If $d \in \mathbb{N}$, and $\mathcal{P}$ is a partition of $\mathbb{R}^d$, and there exists $M \in (0, \infty)$ such that for all $X \in \mathcal{P}$, X is Lebesgue measurable, and $m(X) < M$, then for any $\epsilon \in (0, \infty)$, there exists $\vec{p} \in \mathbb{R}^d$ such that*

$$|\mathcal{N}_\epsilon(\vec{p})| \geq d + 1.$$

The Second Optimality Theorem states that under the stronger hypothesis that the members of the partition are bounded in diameter, then not only there is no secluded partition with degree less than $d+1$, but it is even false if we allow ϵ to be different for each point in the space. The ϵ function in the statement below should be viewed as some fixed ϵ for each point of $\mathbb{R}^d$.

Theorem 9.4 (Second Optimality Theorem). *If $d \in \mathbb{N}$, and $\mathcal{P}$ is a partition of $\mathbb{R}^d$, and there exists $D \in (0, \infty)$ such that for all $X \in \mathcal{P}$, it holds that $\mathrm{diam}(X) < D$, then for any $\epsilon : \mathbb{R}^d \to (0, \infty)$, there exists $\vec{p} \in \mathbb{R}^d$ such that*

$$|\mathcal{N}_{\epsilon(\vec{p})}(\vec{p})| \geq d + 1.$$

The conclusion of the Second Optimality Theorem is stronger than the conclusion of the First Optimality Theorem, but the hypothesis is also stronger (assuming all members are measurable); could it be, though, that the hypothesis of the First Optimality Theorem implies the conclusions of the Second Optimality Theorem and we just didn't find a proof of this? We show in Section 9 that this is not the case. There really is a "gap" between these two theorems.

The Third Optimality Theorem uses a still stronger hypothesis and requires finiteness somewhere in the space. With this strengthened hypothesis, the conclusion can again be strengthened to say that we don't even really care about the ϵ neighborhoods at all because there is some point at the closure of $d + 1$ members. The theorem hypothesis uses the notion of a *strict pairwise bound D* (Definition 9.1) which just means that for each member $X \in \mathcal{P}$, all points in X are distance *strictly less* than D apart[7] (with respect to d_{max}).

Theorem 9.5 (Third Optimality Theorem). *If $d \in \mathbb{N}$, and $\mathcal{P}$ is a partition of $\mathbb{R}^d$, and there exists $D \in (0, \infty)$ such that for all $X \in \mathcal{P}$, D is a strict pairwise bound for X (it is sufficient but not necessary that $\mathrm{diam}(X) < D$), and if there exists some $\vec{\alpha} \in \mathbb{R}^d$ such that $\vec{\alpha} + [0, D]^d$ intersects finitely many members of $\mathcal{P}$, then there exists $\vec{p} \in \mathbb{R}^d$ such that*

$$|\mathcal{N}_{\vec{0}}(\vec{p})| \geq d + 1.$$

Furthermore, $\mathcal{P}$ contains a $(d+1)$-clique.

Just as with the first two Optimality Theorems, we will show in Section 9 that the hypothesis of the Second Optimality Theorem does not imply the conclusion of the Third Optimality Theorem, so again there is a "gap". Now we turn to the proof outlines for the four theorems above.

Proof Outline for Optimality for Unit Hypercube Partitions (Theorem 9.8). As a corollary to the Third Optimality Theorem, we get the same conclusion for partitions whose members have a uniform upper bound on diameter and a uniform lower bound on the measure (see Corollary 9.7), and since unit hypercube partitions have this property, the conclusion follows. □

Alternate Proof Sketch. It is also possible to directly prove this result without using the heavy machinery of the three Optimality Theorems by showing that if $\vec{p}$ is one of the corners of some hypercube X, then it is at the closure of at least $d + 1$ members. The argument goes by induction on d and considers the 2^d-many orthants locally around $\vec{p}$ showing that because X takes up exactly one orthant, the only ways to fill up the other $2^d - 1$ orthants requires at least d other hypercubes. □

We give the outlines of the three Optimality Theorems in order of increasing complexity (which happens to be reverse order).

Proof Outline for the Third Optimality Theorem. To prove this, let H denote the closed hypercube $\vec{\alpha} + [0, D]^d$ which intersects only finitely many elements of $\mathcal{P}$. We use the partition $\mathcal{P}$ to induce a (finite) partition $\mathcal{S}$ on H:

$$\mathcal{S} = \{X \cap H : X \in \mathcal{P}, \ X \cap H \neq \emptyset\}.$$

D remains a strict pairwise bound for all X in the induced partition $\mathcal{S}$. This means that no member of $\mathcal{S}$ intersects opposing facets of the hypercube H, and thus no two corners of H belong to the same member

---

[7]This is stronger than saying that the diameter of each member is $\leq D$, but weaker than saying that the diameter of each member is $< D$. Members can have diameter D, but can't have points that attain that diameter.

of $\mathcal{S}$. We show that this allows us to color all points of H using 2^d colors (one associated with each corner) in such a way to satisfy the coloring properties of a generalized Sperner's lemma/KMM lemma, and such that all points within any member of $\mathcal{S}$ are assigned the same color. This will let us conclude that there is some point $\vec{p}$ of the space belonging to the closure of $d+1$ colors, and because there are only finitely many members of $\mathcal{S}$, we use the fact that the closure of a finite union of members is equal to the finite union of the closures, and so $\vec{p}$ belongs not just to the closure of $d+1$ colors, but to the closure of $d+1$ members of $\mathcal{S}$. The existence of a $(d+1)$-clique is a trivial consequence because all of these $d+1$ members have closures that contain $\vec{p}$ and so each pair of these members is adjacent (see Definition 6.1). $\qquad\square$

The proof of the Second Optimality Theorem is quite similar to the previous proof.

Proof Outline for the Second Optimality Theorem. Let $H = [0, D]^d$ and let $\mathcal{S}$ denote the partition of H induced by $\mathcal{P}$, observing that as before, the fact that each $X \in \mathcal{P}$ has $\mathrm{diam}(X) < D$ implies no member of $\mathcal{P}$ (and thus $\mathcal{S}$) intersects opposing facets of H. Using the same Sperner/KMM technique as before we can find some point $\vec{p}$ in the closure of $d+1$ colors. However, because the closure of an infinite collection of members may be larger than the infinite union of the closures of those members, it is possible that $\vec{p}$ does not belong to the closure of $d+1$ members of $\mathcal{S}$. However, no matter what the value of $\epsilon(\vec{p})$ is, this neighborhood of $\vec{p}$ contains an open set around $\vec{p}$, and that open set must intersect some member of $\mathcal{S}$ corresponding to each of the $d+1$ colors. This gives the result. $\qquad\square$

The proof of the First Optimality Theorem is more involved.

Proof Outline for the First Optimality Theorem. Having only a bound on the measure of the members of $\mathcal{P}$, it is possible that they extend arbitrarily far (even infinitely far) within $\mathbb{R}^d$, so the goal will be to approximate $\mathcal{P}$ by a nicer partition that has bounded diameter elements and then utilize the same Sperner/KMM techniques as used for the Second and Third Optimality Theorems. In fact, the approximating partition $\mathcal{A}$ that we construct will satisfy the hypothesis of the Third Optimality Theorem, so there is a point $\vec{p}$ belonging to the closure of $d+1$ members of $\mathcal{A}$. We then argue that the partition $\mathcal{A}$ is similar enough to $\mathcal{P}$ that the set $B_\epsilon(\vec{p})$ intersects at least $d+1$ members of $\mathcal{P}$.

In more detail, the steps are as follows:
1. Partition $\mathbb{R}^d$ into a grid of ϵ-diameter hypercubes
2. Label each of these hypercubes with a member of $\mathcal{P}$ that it intersects a lot (i.e. the measure of the intersection is sufficiently large)
3. Define a graph (equivalently a symmetric binary relation) on the hypercubes so that they are adjacent in the graph if they are close together and also share the same label
4. Consider the transitive closure of the relation (i.e. connected components of the graph); show that each connected component has cardinality less than a fixed size, and so has bounded diameter
5. Consider the partition $\mathcal{A}$ where members correspond to equivalence classes and apply the Third Optimality Theorem to get a point $\vec{p}$ at the closure of $d+1$ members of $\mathcal{A}$
6. Show that each such element of $\mathcal{A}$ is a superset of a hypercube in the grid and that different elements of $\mathcal{A}$ which have $\vec{p}$ in their closure contain hypercubes with distinct labels
7. Show that each of the $d+1$ distinctly labeled hypercubes above are contained in $B_\epsilon(\vec{p})$
$\qquad\square$

2.2 Upper Bound on Tolerance (ϵ)

With the optimality of the parameter k firmly established, we wish to consider whether our value of $\epsilon = \frac{1}{2d}$ in Theorem 7.20 is maximal when we have k minimized at $k = d+1$. Ideally, we want to consider optimality of the ϵ parameter not just for unit hypercube partitions but for a broader class of partitions. A natural

class is partitions with the property that each member has a strict pairwise bound[8] (see Definition 9.1) of $D = 1$; all unit hypercube partitions belong to this class[9], and $D = 1$ is the smallest value of D we can use so that they do. For the sake of comparisons between some results in the paper we didn't want to fix D to be 1 in the statement of the results below, but taking any of the results below as a standalone result, the value of D is nothing more than a trivial scaling factor, so the reader can consider it to be fixed at $D = 1$. The next result gives a trivial bound on the tolerance parameter, but gives a tight bound for $d = 1$.

Proposition 10.1 (Trivial Tolerance (ϵ) Bound). *If $d \in \mathbb{N}$, and $\mathcal{P}$ is a partition of $\mathbb{R}^d$, and there exists $D \in (0, \infty)$ such that for all $X \in \mathcal{P}$, $\mathrm{diam}(X) \leq D$, and if $\epsilon \in (0, \infty)$ such that for all $\vec{p} \in \mathbb{R}^d$,*

$$|\mathcal{N}_\epsilon(\vec{p})| \leq 2^d$$

then $\epsilon \leq \frac{D}{2}$.

Proof Sketch. If $\epsilon > \frac{D}{2}$, we fix an arbitrary member X of $\mathcal{P}$, and shift the hypercube $[-\epsilon, \epsilon]^d$ so that X is completely contained in the interior. Then we observe that the 2^d-many corners of the shifted $[-\epsilon, \epsilon]^d$ are distance greater than D apart (in d_{max}), so they each belong to a distinct member of $\mathcal{P}$, and none of them belong to X. Taking $\vec{p}$ to be the center of the shifted $[-\epsilon, \epsilon]^d$, we have for contradiction that $|\mathcal{N}_\epsilon(\vec{p})| \geq 2^d + 1 > 2^d + 1$. $\qquad\square$

By considering the specific value $d = 1$, we get the following corollary which states that $\epsilon = \frac{D}{2d} = \frac{1}{2d}$ is indeed the optimal value for the class of partitions with strict pairwise bound of $D = 1$ for dimension $d = 1$.

Corollary 10.2 (Optimal Tolerance (ϵ) in $\mathbb{R}^1$). *If $d = 1$ and $\mathcal{P}$ is a partition of $\mathbb{R}^d = \mathbb{R}^1$, and there exists $D \in (0, \infty)$ such that for all $X \in \mathcal{P}$, D is a strict pairwise bound for X, and if $\epsilon \in (0, \infty)$ is such that for all $\vec{p} \in \mathbb{R}^d$,*

$$|\mathcal{N}_\epsilon(\vec{p})| \leq d + 1 = 2$$

then $\epsilon \leq \frac{D}{2} = \frac{D}{2d}$.

Proof. Since D is a strict pairwise bound for each member of $\mathcal{P}$, then every member of $\mathcal{P}$ has diameter at most D. Apply Proposition 10.1 with $d = 1$. $\qquad\square$

We can also show that the value of $\epsilon = \frac{D}{2d}$ is optimal for $d = 2$, but Proposition 10.1 evaluated with $d = 2$ does not give the correct value, so it requires a separate proof.

Proposition 10.8 (Optimal Tolerance (ϵ) in $\mathbb{R}^2$). *If $d = 2$ and $\mathcal{P}$ is a partition of $\mathbb{R}^d = \mathbb{R}^2$, and there exists $D \in (0, \infty)$ such that for all $X \in \mathcal{P}$, D is a strict pairwise bound for X, and if $\epsilon \in (0, \infty)$ is such that for all $\vec{p} \in \mathbb{R}^d$,*

$$|\mathcal{N}_\epsilon(\vec{p})| \leq d + 1 = 3$$

then $\epsilon \leq \frac{D}{4} = \frac{D}{2d}$.

For other dimensions, we do not know if $\epsilon = \frac{D}{2d}$ is optimal (though we conjecture that it is up to a constant factor; see Conjecture 10.11). The next result at least gives an upper bound that for any $d \in \mathbb{N}$, $\epsilon \leq \frac{D}{2\sqrt{d}}$[10].

---

[8] While it might seem that considering partitions with members of diameter at most $D = 1$ would be more natural than considering a strict pairwise bound of 1, working through the proofs of the results below one can begin to see that actually, the strict pairwise bound is the "right" thing to do as that is the condition that works best in the proofs. In fact, trying to use the diameter leads to slightly weaker results.

[9] This is at least true as we define unit hypercube partitions using half-open hypercubes (see Definition 6.4).

[10] There are two useful ways to interpret this upper bound. First, the value of $\epsilon = \frac{1}{2d}$ that we achieve is within a factor of $\frac{1}{\sqrt{d}}$ of the optimal value $\epsilon' \leq \frac{1}{2\sqrt{d}}$ which gives a bound on the quality in terms of the dimension. Second, the value of $\epsilon = \frac{1}{2d}$ that we achieve is correct to within a square: $\epsilon = \frac{1}{2d} > \left(\frac{1}{2\sqrt{d}}\right)^2 \geq (\epsilon')^2$ which gives a dimension invariant way to view it.

Theorem 10.10 (Universal Tolerance (ϵ) Bound). *If $d \in \mathbb{N}$, and $\mathcal{P}$ is a partition of $\mathbb{R}^d$, and there exists $D \in (0, \infty)$ such that for all $X \in \mathcal{P}$, it holds that D is a strict pairwise bound for X, and if there exists $\epsilon \in (0, \infty)$ such that for all $\vec{p} \in \mathbb{R}^d$,*

$$|\mathcal{N}_\epsilon(\vec{p})| \leq d + 1,$$

then $\epsilon \leq \frac{D}{2(\sqrt[d]{\mathrm{Sperner}(d)}-1)}$ (if $d > 1$). In particular, $\epsilon \leq \frac{D}{2\sqrt{d}}$ (for any d).

The Sperner(d) quantity in the theorem statement is something that will be discussed in depth prior to proving the theorem. It will be sufficient for now to note that Sperner(d) is bounded below by the minimum number of simplices needed in any dissection of the d-cube, and the "in particular" claim in the theorem uses a lower bound of $(d+1)^{\frac{d-1}{2}}$ for the dissection number. Any improvements that are made to the dissection number lower bound will translate directly to improvements of our ϵ upper bound.

Proof Methods of Theorem 10.10. We apply a stronger version of Sperner's lemma (or, equivalently, a stronger version of the KKM lemma) to find multiple points within a fixed volume that are near $d + 1$ different members of the partition. It is then argued that the ϵ-balls around these points must be disjoint, and the value of ϵ is bounded from above using a volume/measure argument. □

3   Related Work in Mathematics

We showed in Section 1 how our work is related to rounding schemes used in computational complexity theory, and in this section we connect our work to the rich history of and continued interest in partitions of $\mathbb{R}^d$, and in particular, partitions by unit hypercubes. Note that in essence, a partition by unit hypercubes is the same as a tiling or a packing of unit hypercubes. The only distinction is that in tilings and packings, the boundaries are ignored and in a partition they are not. Further, it is common in the literature to refer to $[0,1)^d$ and $[0,1]^d$ as a "cube", "d-cube", or "hypercube". Below, we discuss a number of questions that have been investigated in the literature regarding tilings and unit hypercubes. Most of these results deal with translated unit cubes (i.e. no rotation or the higher dimensional analogs of rotation).

The purpose of these examples is fourfold: (1) to demonstrate broad interest in hypercubes and hypercube partitions, (2) to show that even though questions about hypercubes may seem very simple, there remains active research in this area, (3) to preview a few questions related to some of our results, and (4) to demonstrate that there are many results which hold in $\mathbb{R}^1$, $\mathbb{R}^2$, and $\mathbb{R}^3$ but which may not hold in higher dimensions.

Minkowski lattice cube-tiling conjecture (1907)   Minkowski's conjecture [13] states that in any lattice tiling of $\mathbb{R}^d$ by translated unit cubes there exists a pair of cubes whose intersection is an entire $(d-1)$-dimensional face (e.g. in $\mathbb{R}^2$ there would be a pair of squares with a common edge, and in $\mathbb{R}^3$ there would be a pair of cubes sharing a common square side). A lattice tiling is one in which the centers of all of the cubes form a critical lattice (see Related Problems in [14] and [15] for definitions). It was proven true in 1941 by Hajós [16].

Keller's conjecture (1930)   Keller's conjecture [17] is a generalization of Minkowski's conjecture, which relaxes the assumption that the cubes form a lattice. Thus, it states that in any tiling of $\mathbb{R}^d$ by translated unit cubes, there exists a pair of cubes whose intersection is an entire $(d-1)$-dimensional face. The complete resolution of this conjecture took substantial effort only being completely resolved in 2020. In 1940 Perron [18, 19] showed it was true for $d \leq 6$. Szabó [20] recast the question in terms of periodic tilings in 1986 and then introduced the so-called Keller graphs along with Corrádi in 1990 [21]. In 1992, Lagarias and Shor [22] used the Keller graphs to show that the conjecture is false for all $d \geq 10$. This bound was refined by Mackey in 2002 [23] showing that the conjecture is false for $d \geq 8$. Progress on the only remaining case of $d = 7$ was made by Debroni, Eblen, Langston, Myrvold, Shor, and Weerapurage in 2011 [24], and by

Kisielewicz and Łysakowska in 2014 [25], and by Kisielewicz in 2017 [26] and by Łysakowska in 2018 [27]. Finally, in 2020, Brakensiek, Heule, Mackey, and Narváez [28] determined that the conjecture was true for $d = 7$ using automated satisfiability approaches which fully resolved the conjecture.

Furtwängler's conjecture (1936)   Furtwängler's conjecture [29] is another generalization of Minkowski's conjecture where instead of tilings, k-fold tilings are considered (a k-fold tiling is a collection of positions so that if a hypercube is placed at each position, then every point of $\mathbb{R}^d$ either belongs to the boundary of some hypercube, or belongs to exactly k hypercubes). Furtwängler's conjecture states that in any k-fold lattice tiling of $\mathbb{R}^d$, there exists a pair of cubes whose intersection is an entire $(d-1)$-dimensional face, and he proved it for $d \leq 3$. However, Hajós proved in 1942 [16] that the conjecture was false for $d \geq 4$. In 1979, Robinson [30] completely characterized the conjecture by proving for exactly which pairs (k, d) the conjecture held and which it did not.

Fuglede's set conjecture (1974) and functional analysis   A set $\Omega \subset \mathbb{R}^d$ is called a spectral set if it has positive measure and if there is a basis of certain exponential functions for the space $L^2(\Omega)$ of square integrable functions (the set generating the basis is denoted Λ). Fuglede [31] conjectured that a set was spectral if and only if it could be used to tile $\mathbb{R}^d$. Though this was proved false by Tao in 2004 [32], earlier work by Lagarias, Reeds, and Wang in 2000 [33] showed something similar for the special case of unit cubes. In particular, they showed that a set Λ will generate a basis for $L^2([0,1]^d)$ if and only if Λ is the set of center positions of hypercubes in some partition of translated hypercubes. They used this result to show that extending an orthogonal set of functions to a basis is equivalent to extending a packing of hypercubes to a tiling. The ability to extend packings to tilings was also studied by Dutour, Itoh, and Poyarkov in 2018 [34], though in a different context.

Countable partitions of $[0, 1]$ by closed sets   A problem solved by Sierpiński [35, 36] was that there is no (non-trivial) partition of the unit interval into countably many closed sets. This extends trivially to show that there is no (non-trivial) partition the unit hypercube $[0, 1]^d$ into countably many closed sets. This has some connection to our work. In particular, we state a conjecture which is equivalent to this problem when $d = 1$, and thus provides a natural interesting generalization of this problem to higher dimensions.

Coverings, dissections, and triangulations of the cube   A covering of the cube is a set of simplices (using the vertices of the cube) so that the union of the simplices is the entire cube. A dissection is a covering with the additional requirement that the only overlap occurs at the boundary of the simplices. A triangulation is a dissection with the additional requirement that the intersection of any two simplices is either empty or a face of each. It has long been known that there is a triangulation of the d-cube $[0, 1]^d$ using $d!$ simplices. In 1982, Sallee [37, 38] gave lower bounds for how many simplices are needed in a triangulation, and many others have tried to bound the minimal number of simplices needed for a covering, a dissection, and a triangulation since then. In this paper, we will utilize the dissection number for one of our upper bounds and will utilize the dissection number lower bound of Glazyrin from 2012 [39].

Unit cubes more broadly   Many of the above results are discussed in the 2005 survey paper "What Is Known About Unit Cubes" by Zong [40] and in his 2006 book [41]. In addition, many other properties of unit cubes are presented. Zong makes the case that despite the apparent simplicity of the d-cube, there is much that remains unknown about it.

4   Organization

The remainder of the paper provides a complete description and proofs of the established results in detail. In Section 5, we summarize notation that we will use. In Section 6 we present some basic definitions and results pertaining to partitions. In Section 7 we construct the a class of partitions which we call reclusive partitions and prove the Hypercube Partition Theorem. In Section 9 we prove three degree Optimality

Theorems to establish that $k = d+1$ is the minimum value of k for a broad range of reasonable partitions. In Section 10 we first prove the optimality of the tolerance parameter for dimensions 1 and 2 (Proposition 10.1 and Proposition 10.8). In this section, we further show the $1/2\sqrt{d}$ upper bound the tolerance for higher dimensions. In Section 11 we offer an application of secluded partitions to deterministic rounding and multipseudodeterminism. Section 12 contains concluding remarks and a discussion about future research directions.

5   Notation

We will use the following notation and conventions throughout this paper.

- We use $\mathbb{N}$ throughout to indicate the strictly positive integers.
- For any $d \in \mathbb{N}$, we let $[d]$ denote the set of the first d positive integers ($[d] = \{1, 2, \ldots, d\}$).
- If $d \in \mathbb{N}$ and $S \subseteq \mathbb{R}^d$ and $\vec{v} \in \mathbb{R}^d$, then we use the following notation to indicate shifting/translating the set S:

$$\vec{v} + S = S + \vec{v} \overset{\text{def}}{=} \{\vec{s} + \vec{v} \colon \vec{s} \in S\}.$$

- For $d \in \mathbb{N}$ we define the metric d_{max} on $\mathbb{R}^d$ which is the metric induced by the l^∞ norm[11]:

$$d_{max}(\vec{x}, \vec{y}) \overset{\text{def}}{=} \max_{i \in [d]} |x_i - y_i| = \|\vec{x} - \vec{y}\|_\infty.$$

  While we could use the $\|\cdot\|_\infty$ norm notation throughout the paper, we choose to use the d_{max} metric notation when possible.
- For any $X \subseteq \mathbb{R}^d$ we denote the closure of X with respect to the d_{max} metric by $\overline{X}$. Since the topology induced by the l^∞ norm on $\mathbb{R}^d$ is the same as the Euclidean topology (induced by the l^2 norm), and the Euclidean topology on $\mathbb{R}^d$ is the same as the product topology on $\mathbb{R}^d$, the closure is the same when taken with respect to any of these. In particular, if $X_1, \ldots, X_d \subseteq \mathbb{R}$ and $X = \prod_{i=1}^d X_i$, then $\overline{X} = \prod_{i=1}^d \overline{X_i}$ (because the product of closures is the same as the closure of the product).
- For any $\vec{x} \in \mathbb{R}^d$ and $\epsilon \in (0, \infty)$ we use $B_\epsilon^\circ(\vec{x}) \overset{\text{def}}{=} \{\vec{y} \in \mathbb{R}^d : d_{max}(\vec{x}, \vec{y}) < \epsilon\}$ to denote the open ball of radius ϵ around $\vec{x}$, and we use $\overline{B}_\epsilon(\vec{x}) \overset{\text{def}}{=} \{\vec{y} \in \mathbb{R}^d : d_{max}(\vec{x}, \vec{y}) \leq \epsilon\}$ to denote the closed ball of radius ϵ around $\vec{x}$, noting that $\overline{B}_\epsilon(\vec{x}) = \overline{B_\epsilon^\circ(\vec{x})}$. The standard notation for the open ball is $B_\epsilon(\vec{x})$, but we will not use the open ball very often in this paper and prefer to use the circle superscript to emphasize when we do. (Any occurrence in the paper of a ball without the overline or the circle superscript is a typo.)
- Though mentioned in the introduction, for ease of reference, we again define the ϵ-neighborhoods. When it is understood that we are considering a particular partition $\mathcal{P}$, then for any point $\vec{p} \in \mathbb{R}^d$, and any $\epsilon \in (0, \infty)$, we let

$$\mathcal{N}_\epsilon(\vec{p}) \overset{\text{def}}{=} \{X \in \mathcal{P} : X \cap \overline{B}_\epsilon(\vec{p}) \neq \emptyset\}.$$

$$\mathcal{N}_{\overline{0}}(\vec{p}) \overset{\text{def}}{=} \{X \in \mathcal{P} : \overline{X} \ni \vec{p}\}.$$

  An alternative expression of the former is as follows. This is justified in Lemma 5.1 at the end of this section.

$$\mathcal{N}_\epsilon(\vec{p}) = \{\text{member}(\vec{x}) : \vec{x} \in \overline{B}_\epsilon(\vec{p})\}.$$

- If the partition $\mathcal{P}$ of $\mathbb{R}^d$ is understood, we define a function member $: \mathbb{R}^d \to \mathcal{P}$ mapping each point $\vec{x} \in \mathbb{R}^d$ to the unique member containing it. Often this is denoted using the equivalence class notation of $[\vec{x}]$, but in this paper we shall prefer to use the member function notation for clarity since we will

---

[11]Technically this is a different metric for each dimension d, but it is not necessary to include this in the notation as the dimension will be clear from context.

in general not be thinking of the partition as being defined by an equivalence relation.

- When we say "countable" we mean finite or countably infinite.

Lemma 5.1.

$$\mathcal{N}_\epsilon(\vec{p}) = \left\{ \text{member}(\vec{x}) : \vec{x} \in \overline{B}_\epsilon(\vec{p}) \right\}.$$

Proof. If $X \in \mathcal{N}_\epsilon(\vec{p})$, then $X \cap \overline{B}_\epsilon(\vec{p}) \neq \emptyset$, so let $\vec{x}_0 \in X \cap \overline{B}_\epsilon(\vec{p})$ which means $\text{member}(\vec{x}_0) = X$, so $X \in \left\{ \text{member}(\vec{x}) : \vec{x} \in \overline{B}_\epsilon(\vec{p}) \right\}$. Conversely, if $X \in \left\{ \text{member}(\vec{x}) : \vec{x} \in \overline{B}_\epsilon(\vec{p}) \right\}$ then there exists $\vec{x}_0 \in \overline{B}_\epsilon(\vec{p})$ such that $X = \text{member}(\vec{x}_0)$, so in particular $\vec{x}_0 \in X$ and thus $X \cap \overline{B}_\epsilon(\vec{p}) \neq \emptyset$ so $X \in \mathcal{N}_\epsilon(\vec{p})$. $\qquad\square$

6 Definitions and Basic Partition Results

In this section, we will provide a variety of fairly basic claims about hypercube partitions that we will need later. We also introduce an example of a reclusive partition in order to present some geometric intuition before moving into the linear algebra perspective in the next section.

We first define a natural notion of adjacency for any partition of $\mathbb{R}^d$. By definition, no two members of a partition contain the same element, but if we consider the closures of the members, then this need not hold.

Definition 6.1 (Adjacent). *Let $d \in \mathbb{N}$, and $\mathcal{P}$ be a partition of $\mathbb{R}^d$, and $X, Y \in \mathcal{P}$. We say that X and Y are adjacent if the intersection of their closures is non-empty. That is, we say X and Y are adjacent if $\overline{X} \cap \overline{Y} \neq \emptyset$. We denote this relation as $X \overset{\text{adj}}{\sim} Y$.*

As defined, adjacency is a reflexive relation—for any partition $\mathcal{P}$, for all $X \in \mathcal{P}$, we have $X \overset{\text{adj}}{\sim} X$. Also note that despite the notation, adjacency is not an equivalence relation as it is not transitive in general.

Every partition of $\mathbb{R}^d$ naturally induces a graph in which adjacency in the graph aligns with adjacency in the partition.

Definition 6.2 (Partition Graph). *Let $d \in \mathbb{N}$ and $\mathcal{P}$ be a partition of $\mathbb{R}^d$. The partition graph of $\mathcal{P}$ is the graph G whose vertex set is the set $\mathcal{P}$, and whose edge set contains the edge $\{X, Y\}$ if and only if $X \overset{\text{adj}}{\sim} Y$ in the partition.*

Because adjacency is reflexive, every vertex in G has a self-loop. However, this will not be important for us. We will usually not talk explicitly of this graph and will instead identify the partition and its graph structure. We will frequently need to prove that two members of a partition are not adjacent, and we will use the following simple corollary.

Corollary 6.3 (Non-Adjacency Corollary). *Let $d \in \mathbb{N}$, and $\mathcal{P}$ be a partition of $\mathbb{R}^d$, and $X, Y \in \mathcal{P}$. If for all $\vec{x} \in \overline{X}$ and $\vec{y} \in \overline{Y}$ it holds that $d_{max}(\vec{x}, \vec{y}) > 0$ then X and Y are not adjacent.*

Proof. For proof by contrapositive, if X and Y are adjacent, then there exists $\vec{z} \in \overline{X} \cap \overline{Y}$, so let $\vec{x} = \vec{y} = \vec{z}$ and we have $d_{max}(\vec{x}, \vec{y}) = 0$. $\qquad\square$

We will be particularly interested in partitions of translated unit hypercubes as defined below.

Definition 6.4 (Unit Hypercubes and Unit Hypercube Partitions). *Let $d \in \mathbb{N}$. A set $X \subseteq \mathbb{R}^d$ is called a unit hypercube if there exists $\vec{a} \in \mathbb{R}^d$ such that X can be expressed as*

$$X = \vec{a} + [0, 1)^d.$$

If so, this expression is unique, and we call $\vec{a}$ the representative corner of X and denote it by $\text{rep}(X)$. A partition $\mathcal{P}$ of $\mathbb{R}^d$ is called a unit hypercube partition if each member of $\mathcal{P}$ is a unit hypercube.

It is worth noting that $\mathrm{rep}(X) \in X$.

The terminology of our definition may be somewhat deceptive because what we call a unit hypercube partition requires these hypercubes to be translations of $[0,1)^d$. For example, if $\mathbb{R}^2$ is partitioned by a unit grid and the partition is rotated, then it would not be considered a unit hypercube partition by our definition. As mentioned before, some other sources consider the closed unit hypercube $\vec{a} + [0,1]^d$ or rotations of this set. If we need a closed unit hypercube, we mention so explicitly.

The following simple lemma shows that shifting a unit hypercube changes the representative corner in the natural way.

Lemma 6.5 (Representative Corner Shift Lemma). *Let $d \in \mathbb{N}$, and $X \subset \mathbb{R}^d$ be a unit hypercube. Then for any vector $\vec{v} \in \mathbb{R}^d$,*
$$\mathrm{rep}(X + \vec{v}) = \mathrm{rep}(X) + \vec{v}.$$

Proof. This holds because if $X = \vec{a} + [0,1)^d$, then $X + \vec{v} = (\vec{a} + \vec{v}) + [0,1)^d$. $\qquad\square$

While the defined notion of adjacency may be the most natural, it will not be the easiest to work with, so the next two lemmas together give an equivalent characterization of adjacency of members of a unit hypercube partition. The first lemma states that in a unit hypercube partition it is impossible to have two distinct unit hypercubes X and Y with representative corners closer than a distance of 1 of each other, because otherwise the hypercubes would overlap.

Lemma 6.6 (Non-Overlapping Lemma). *Let $d \in \mathbb{N}$, and $\mathcal{P}$ be a unit hypercube partition of $\mathbb{R}^d$, and $X, Y \in \mathcal{P}$. Then either $X = Y$ or $d_{max}(\mathrm{rep}(X), \mathrm{rep}(Y)) \geq 1$.*

Proof. If $d_{max}(\mathrm{rep}(X), \mathrm{rep}(Y)) < 1$ then for all $i \in [d]$, $|\mathrm{rep}(X)_i - \mathrm{rep}(Y)_i| < 1$. Thus, for all $i \in [d]$, either

$$\mathrm{rep}(X)_i \in [\mathrm{rep}(Y)_i, \mathrm{rep}(Y)_i + 1) \qquad\qquad (\text{if } \mathrm{rep}(X)_i \geq \mathrm{rep}(Y)_i)$$

or

$$\mathrm{rep}(Y)_i \in [\mathrm{rep}(X)_i, \mathrm{rep}(X)_i + 1). \qquad\qquad (\text{if } \mathrm{rep}(X)_i \leq \mathrm{rep}(Y)_i)$$

In either case, this implies

$$[\mathrm{rep}(X)_i, \mathrm{rep}(X)_i + 1) \cap [\mathrm{rep}(Y)_i, \mathrm{rep}(Y)_i + 1) \neq \emptyset.$$

Note that by definition of $\mathrm{rep}(X)$ and $\mathrm{rep}(Y)$, we can express X and Y as

$$X = \prod_{i=1}^{d} [\mathrm{rep}(X)_i, \mathrm{rep}(X)_i + 1)$$

$$Y = \prod_{i=1}^{d} [\mathrm{rep}(Y)_i, \mathrm{rep}(Y)_i + 1).$$

This implies that $X \cap Y \neq \emptyset$, and because $\mathcal{P}$ is a partition, this implies $X = Y$. $\qquad\square$

The next lemma shows that in the case of of adjacent hypercubes, the above lower bound on the distance between the representative corners is tight. The intuition is that if X and Y are unit hypercubes, then because the side lengths of the hypercubes are all 1, it is the case that adjacency of X and Y is equivalent to the representatives of X and Y being within a distance of 1 (in the d_{max} metric, which measures the furthest coordinate).

Lemma 6.7 (Adjacency by Representatives Lemma). *Let $d \in \mathbb{N}$, and $\mathcal{P}$ be a unit hypercube partition of $\mathbb{R}^d$, and $X, Y \in \mathcal{P}$. Then $X \overset{\mathrm{adj}}{\sim} Y$ if and only if $d_{max}(\mathrm{rep}(X), \mathrm{rep}(Y)) \leq 1$.*

Proof. We begin with the reverse direction. If $d_{max}(\text{rep}(X), \text{rep}(Y)) \leq 1$ then for all $i \in [d]$, $|\text{rep}(X)_j - \text{rep}(Y)_j| \leq 1$. By similar reasoning as the previous lemma (but using the closed intervals instead of the half open intervals), we have for all $i \in [d]$,

$$[\text{rep}(X)_i, \text{rep}(X)_i + 1] \cap [\text{rep}(Y)_i, \text{rep}(Y)_i + 1] \neq \emptyset.$$

Since $X = \prod_{i=1}^{d}[\text{rep}(X)_i, \text{rep}(X)_i + 1)$ (and similarly for Y), we can express the closures as

$$\overline{X} = \prod_{i=1}^{d}[\text{rep}(X)_i, \text{rep}(X)_i + 1]$$

$$\overline{Y} = \prod_{i=1}^{d}[\text{rep}(Y)_i, \text{rep}(Y)_i + 1].$$

Thus $\overline{X}$ and $\overline{Y}$ intersect in each coordinate, so $\overline{X} \cap \overline{Y} \neq \emptyset$. Thus, by definition, $X \overset{\text{adj}}{\sim} Y$.

Now we prove the forward direction by contrapositive. Suppose $d_{max}(\text{rep}(X), \text{rep}(Y)) > 1$. This implies the existence of some $j \in [d]$ such that $|\text{rep}(X)_j - \text{rep}(Y)_j| > 1$. This means

$$[\text{rep}(X)_j, \text{rep}(X)_j + 1] \cap [\text{rep}(Y)_j, \text{rep}(Y)_j + 1] = \emptyset$$

thus $\overline{X} \cap \overline{Y} = \emptyset$, so X and Y are not adjacent. $\qquad\square$

The previous two lemmas give rise to a number of equivalent notions of adjacency in a unit hypercube partition which we now formalize.

Corollary 6.8 (Equivalent Definitions of Adjacency). *Let $d \in \mathbb{N}$, and $\mathcal{P}$ be a unit hypercube partition of $\mathbb{R}^d$, and $X, Y \in \mathcal{P}$. The following statements are all equivalent:*

1. *$X \overset{\text{adj}}{\sim} Y$*
2. *$X = Y$ or $d_{max}(\text{rep}(X), \text{rep}(Y)) = 1$*
3. *$d_{max}(\text{rep}(X), \text{rep}(Y)) = 0$ or $d_{max}(\text{rep}(X), \text{rep}(Y)) = 1$*
4. *$d_{max}(\text{rep}(X), \text{rep}(Y)) \leq 1$*

Proof. That (1) and (4) are equivalent is the statement of the previous lemma. Statements (2) and (3) are equivalent because $X = Y$ if and only if $\text{rep}(X) = \text{rep}(Y)$ if and only if $d_{max}(\text{rep}(X), \text{rep}(Y)) = 0$. Thus, it suffices to show that (3) and (4) are equivalent.

Trivially, (3) implies (4). To show that (4) implies (3), assume $d_{max}(\text{rep}(X), \text{rep}(Y)) \leq 1$; so by Lemma 6.6, either $X = Y$ implying $d_{max}(\text{rep}(X), \text{rep}(Y)) = 0$ implying (3), or $d_{max}(\text{rep}(X), \text{rep}(Y)) \geq 1$ implying $d_{max}(\text{rep}(X), \text{rep}(Y)) = 1$ (because $d_{max}(\text{rep}(X), \text{rep}(Y)) \leq 1$ by hypothesis) implying (3). $\qquad\square$

The following proposition shows that in a hypercube partition, if a collection of hypercubes is a clique in the partition graph (i.e. they are all pairwise adjacent), then that actually implies that there is a point $\vec{p}$ which is at the closure of all of them, and so for any choice of $\epsilon \in (0, \infty)$, $\overline{B}_\epsilon(\vec{p})$ will intersect every one of these hypercubes. This does not hold for more general partitions. The usefulness of this result is that it is possible to show that a unit hypercube partition is not $(k, \epsilon$-secluded by finding a clique of size $k + 1$.

Proposition 6.9 (Hypercube Cliques). *Let $d, n \in \mathbb{N}$, and $\mathcal{P}$ be a unit hypercube partition of $\mathbb{R}^d$, and $X_1, \ldots, X_n \in \mathcal{P}$ such that for all $i, j \in [n]$ $X_i \overset{\text{adj}}{\sim} X_j$. Then there is a point $\vec{p} \in \mathbb{R}^d$ such that for all $i \in [n]$, $\vec{p} \in \overline{X}_i$.*

Proof. Let $X_1, \ldots, X_n$ as stated. Each $\overline{X_i}$ can be expresses as $\overline{X_i} = \overline{B}_{\frac{1}{2}}(\vec{x}^{(i)})$ for some $\vec{x}^{(i)} \in \mathbb{R}^d$. Because each X_i and X_j are adjacent, by definition $\overline{X_i} \cap \overline{X_j} \neq \emptyset$, and because of each these sets is a metric ball of radius $\frac{1}{2}$, it follows that $d_{max}(\vec{x}^{(i)}, \vec{x}^{(j)}) \leq 1$. It is a property of the d_{max} metric that if a set of vectors are all pairwise within distance 1 of each other, then there is a point $\vec{p}$ which is within distance $\frac{1}{2}$ of all of them[12], so we can find such a $\vec{p}$ for the set $\vec{x}^{(i)}, \ldots, \vec{x}^{(n)}$. Since $\vec{p}$ is distance at most $\frac{1}{2}$ from each $\vec{x}^{(n)}$, that means $\vec{p} \in \overline{B}_{\frac{1}{2}}(\vec{x}^{(n)}) = \overline{X_i}$. $\qquad\square$

7 Reclusive Partitions

In this section we define what we call reclusive partitions. The word reclusive is a synonym of secluded, but we will have a different technical definition. We have used, and will continue to use the term secluded to discuss generic partitions. We will use the term reclusive partition to talk about unit hypercube partitions which have the very specific linear algebra structure that we develop in this section.

7.1 An Example Reclusive Partition

We will shortly be working with partitions of $\mathbb{R}^d$ from a linear algebraic perspective because that allows us to state the results very generally. However, this makes the intuition of the geometry more difficult. As a partial remedy for this, we first introduce a very specific reclusive partition of $\mathbb{R}^d$ for each $d \in \mathbb{N}$—these are the partitions that we first studied, and they are mathematically very convenient to work with. We will not be interested in them as anything more than an example because the parameter value of ϵ that they achieve is only $\frac{1}{2^{d-1}}$ and as we have mentioned, some of the reclusive partitions will achieve as large as $\frac{1}{2d}$. Nonetheless, these partitions capture the essential geometric idea of the construction of the more general reclusive partitions.

The following defines for all $d \in \mathbb{N}$ a partition of $\mathbb{R}^d$ which consists solely of half-open/half-closed unit hypercubes. After presenting the definition, we elaborate on how to interpret it geometrically.

Definition 7.1 ($\mathcal{P}_d$)**.** *For each* $d \in \mathbb{N}$, *define*[13] $\vec{v_d} \overset{\text{def}}{=} \langle \frac{1}{2^{d-1}}, \ldots, \frac{1}{2^{d-1}}, 1 \rangle$, *the vector whose last entry is* 1 *and all other* $d-1$ *entries are* $\frac{1}{2^{d-1}}$.

Define $\mathcal{P}_1$, *which is a partition of* $\mathbb{R}^1$, *as follows:*

$$\mathcal{P}_1 \overset{\text{def}}{=} \{[0,1) + n \colon n \in \mathbb{Z}\} = \{[0,1) + n \cdot \vec{v_1} \colon n \in \mathbb{Z}\}.$$

Then define $\mathcal{P}_d$, *which is a partition of* $\mathbb{R}^d$, *inductively for all* $d \in \mathbb{N}$ *with* $d > 1$ *as follows:*

$$\mathcal{P}_d \overset{\text{def}}{=} \{B \times [0,1) + n \cdot \vec{v_d} \colon n \in \mathbb{Z}, \ B \in \mathcal{P}_{d-1}\}$$

The following discussion motivates why we are interested in this partition and how to understand it geometrically. The first partition ($d = 1$) breaks up $\mathbb{R}^d = \mathbb{R}^1$ into unit intervals which are half open. This partition has the property that for any fixed point $x \in \mathbb{R}^1$, if you consider all points with (Euclidean) distance less than or equal to $1/2^d = 1/2$ from x, all such points belong to at most $d + 1 = 2$ members of the partition $\mathcal{P}_1$ (see Figure 1).

Then consider how the second partition $\mathcal{P}_2$ is constructed by first examining only the members constructed when $n = 0$. In this case, each member is $B \times [0,1)$ for some $B \in \mathcal{P}_1$. We think of this as extruding each member of the previous partition one unit into the newest dimension. Restricted to $n = 0$, this would partition $\mathbb{R}^1 \times [0,1)$, so to capture all elements of $\mathbb{R}^2$, we need to make shifts not just for $n = 0$

---

[12]If the set of vectors is S, then the point $\vec{p}$ can defined coordinate-wise as $p_i = \frac{1}{2}(\sup_{\vec{x} \in S} x_i - \inf_{\vec{x} \in S} x_i)$.

[13]By this definition, in the case that $d = 1$, $v_d = \langle 1 \rangle$.

but for every integer. The last index of $\vec{v_d}$ is 1 to get integer shifts in the newest dimension so that for an arbitrary value of n we get a partition of $\mathbb{R}^1 \times [n, n+1)$.

Why is it that $\vec{v_d}$ is defined as it is? If we had taken $\vec{v_d} = \langle 0, \ldots, 0, 1 \rangle$ for example (so that $\vec{v_2} = \langle 0, 1 \rangle$), the definition above would still produce a partition. However, it would not have the desired property that for any point $\vec{x} \in \mathbb{R}^2$ the points within a distance $1/2^d = 1/4$ belong to at most $d+1 = 3$ members of $\mathcal{P}_2$. For example (see Figure 1), the point $(1, 1)$ would be arbitrarily close to points of the following four members: $[0, 1) \times [0, 1)$, $[0, 1) \times [1, 2)$, $[1, 2) \times [0, 1)$, $[1, 2) \times [1, 2)$.

To get this property, we shift the extrusions by a "little bit" in all of the other dimensions too in order to offset the "seams" or boundaries between members of the partition. With each new partition we build from an old one, the amount of shift in each dimension is reduced (the entries in the vector $\vec{v_d}$ decrease as d increases) so that shifts aren't "undone" by shifting too much and cycling back. The construction of $\mathcal{P}_3$ has a similar intuition, and beyond that, we find it difficult to visualize. For completeness, we next prove that the $\mathcal{P}_d$ are indeed partitions. The proof will probably not lend insight to the rest of the paper and should be freely skipped.

Claim 7.2. *For each $d \in \mathbb{N}$, $\mathcal{P}_d$ is a partition of $\mathbb{R}^d$.*

Proof. If $d = 1$ (for an inductive base case), let $x \in \mathbb{R}^1$ be arbitrary and let $n = \lfloor x \rfloor$ so $x \in [n, n+1)$ and $x \notin [m, m+1)$ for any $m \neq n$, so $\mathcal{P}_1$ partitions $\mathbb{R}^1$.

The inductive case follows similarly. Let $\vec{x} = \langle x_i \rangle_{i=1}^d \in \mathbb{R}^d$ be arbitrary. We want to prove the existence of unique $n \in \mathbb{Z}$ and $B \in \mathcal{P}_{d-1}$ such that $\vec{x} \in B \times [0, 1) + n \cdot \vec{v_d}$. Note that by necessity $n = \lfloor x_d \rfloor$ so that $x_d \in [0, 1) + n \cdot 1 = [n, n+1)$ (recall that the last coordinate of $\vec{v_d}$ is $v_{d_d} = 1$). Then we see that $\vec{x} \in B \times [0, 1) + n \cdot \vec{v_d}$ if and only if $\vec{x} - n \cdot \vec{v_d} \in B \times [0, 1)$, and since we have already established the value of n, this holds if and only if $\langle x_i - n \cdot v_{d_i} \rangle_{i=1}^{d-1} \in B$. By the inductive hypothesis, there exists a unique $B \in \mathcal{P}_{d-1}$ such that this holds. Thus $\mathcal{P}_d$ partitions $\mathbb{R}^d$. $\qquad\square$

Based on our discussion above, we hope we have provided the intuition that each member of any $\mathcal{P}_d$ is a unit hypercube with some amount of shift. The $\mathcal{P}_d$ example partitions will be useful to keep in mind as we work with more general unit hypercube partitions.

7.2 Motivating Properties

While the inductive definition of the running example partitions $\mathcal{P}_d$ is useful, it is also useful to consider unit hypercube partitions from another perspective. One can note that in the partition $\mathcal{P}_d$, the representative corner of each unit hypercube is an integer linear combination of the vectors $\vec{v_1}, \ldots, \vec{v_d}$ as defined in Definition 7.1 (padded with zeros in the trailing entries as necessary which correspond to the higher dimensions). The set of all integer linear combinations of a set of basis vectors for the vector space $\mathbb{R}^d$ is known as a lattice group (it is a group under vector addition). Viewing the hypercube representatives as points within the lattice group will be useful. In particular, this gives motivation to look at certain regularly structured unit hypercube partitions by examining a matrix associated with a set of basis vectors of $\mathbb{R}^d$. For example, consider the example partition $\mathcal{P}_d$ for $d = 5$. If we embed the vectors $\vec{v_1}, \ldots, \vec{v_5}$ from Definition 7.1 into $\mathbb{R}^5$ (by padding), and use those vectors as the columns of a matrix, then the matrix would be as follows (e.g. the first column is $\vec{v_1}$, the second column is $\vec{v_2}$, and so on with zeros padded at the end as necessary).

$$A = \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{4} & \frac{1}{8} & \frac{1}{16} \\ 0 & 1 & \frac{1}{4} & \frac{1}{8} & \frac{1}{16} \\ 0 & 0 & 1 & \frac{1}{8} & \frac{1}{16} \\ 0 & 0 & 0 & 1 & \frac{1}{16} \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

In fact, we could equivalently have defined $\mathcal{P}_5$ (and similarly for all $\mathcal{P}_d$ from the running example) to be the set of unit hypercubes whose representatives were integer linear combinations of the columns of this

18

matrix; in other words, $\mathcal{P}_5$ could have been defined as the set of unit hypercubes whose representatives are given by $A\vec{n}$ for some $\vec{n} \in \mathbb{Z}^d$.

In light of this, we shall define a more structured version of unit hypercube partitions by defining them in terms of a matrix. Observe the following four structural properties of the example matrix A above:

1. The matrix A for $\mathcal{P}_5$ above explicitly contains the structure of the partitions $\mathcal{P}_d$ for $d \leq 5$ in the sense that the submatrix consisting of the first 4 rows and first 4 columns is the matrix associated with the partition $\mathcal{P}_4$. Similarly the submatrix consisting of the first 3 rows and first 3 columns is the matrix associated with the partition $\mathcal{P}_3$, and so on.

2. The matrix is upper triangular. The reason is that in the inductive definition of $\mathcal{P}_d$, the vector $\vec{v_d}$ is in $\mathbb{R}^d$, so the lengths of these vectors grows by one with each iteration of the construction, and each iteration of the construction adds a row and column.

3. The diagonal of the matrix is all 1's. This is because in the inductive definition of $\mathcal{P}_d$, there has to be a unit shift in the current dimension to accommodate that the members of the partition are unit hypercubes. For example, in the definition of $\mathcal{P}_1$, the members are $[0, 1) + n\langle 1 \rangle$ for all $n \in \mathbb{Z}$. If the vector $\langle 1 \rangle$ was anything else, this would not be a partition. In the case of $\mathcal{P}_2$, each element is of the form $B \times [0, 1) + n\langle \frac{1}{2}, 1 \rangle$. Again, the last index of this vector must be 1 because we are extruding the elements of the previous partition by $[0, 1)$. This holds in each dimension of the construction.

4. The entries in each row are strictly decreasing. This is because in the inductive definition of $\mathcal{P}_d$, we wanted to offset the "seams" of the hypercubes to prevent points in $\mathbb{R}^n$ from being "close" to too many hypercubes. For example, if we had defined $\mathcal{P}_d$ by taking each $\vec{v_d}$ to be the vector of all 0's aside from the last entry being 1, then the associated matrix would be the identity matrix. This would indeed generate a unit hypercube partition, but it would not have the property we ultimately desire of limiting the adjacencies (in fact, this is the partition in Figure 1).

As mentioned earlier, while the $\mathcal{P}_d$ serve as nice examples, the exponentially decreasing nature of the shifts $(\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \ldots)$ leads to a need for exponentially decreasing ϵ parameter. Instead, we would like the shifts to change linearly and work with partitions that have matrices more like one of the ones below:

$$\begin{bmatrix} 1 & \frac{4}{5} & \frac{3}{5} & \frac{2}{5} & \frac{1}{5} \\ 0 & 1 & \frac{3}{5} & \frac{2}{5} & \frac{1}{5} \\ 0 & 0 & 1 & \frac{2}{5} & \frac{1}{5} \\ 0 & 0 & 0 & 1 & \frac{1}{5} \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} 1 & \frac{4}{5} & \frac{3}{5} & \frac{2}{5} & \frac{1}{5} \\ 0 & 1 & \frac{4}{5} & \frac{3}{5} & \frac{2}{5} \\ 0 & 0 & 1 & \frac{4}{5} & \frac{3}{5} \\ 0 & 0 & 0 & 1 & \frac{4}{5} \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

In fact, we will define the matrices of interest in a general enough way to include all of the matrices shown so far. Notice that both of these two matrices have the same structural properties mentioned above.

We now define the type of matrices that we will be interested in based on the ideas just discussed. These matrices will be used to define partitions which have the property that hypercubes will be adjacent to a small number of other hypercubes; in the language of graph theory, the members will have few neighbors, so we call them reclusive matrices and reclusive partitions. Stated again for emphasis, we will have a connection between minimizing the parameter k in the motivating question in the introduction, and minimizing the sizes of cliques in the partition graph.

7.3 Construction

Definition 7.3 (Reclusive Matrix). *Informally, a square matrix A will be called a reclusive matrix if it is upper triangular, has only 1's on the main diagonal, and is strictly decreasing with positive entries in each row starting at the main diagonal.*

More formally, a square $n \times n$ matrix $A = (a_{ij})$ will be called a reclusive matrix if all of the following hold:

1. *For all $i > j \in [n]$, $a_{ij} = 0$.* *(Upper Triangular)*

2. For all $i \in [n]$, $a_{ii} = 1$.   *(1's on Diagonal)*
3. For all $i \leq j < k \in [n]$, $a_{ij} > a_{ik} > 0$.   *(Decreasing and Non-zero after Diagonal)*

Remark 7.4. *Reclusive matrices are invertible because they are upper triangular so the determinant is equal to the product of the diagonal entries which are all 1.*

We view $d \times d$ reclusive matrices as a natural way to build partitions of $\mathbb{R}^d$ as just discussed. Before defining these partitions, we formalize the lattice group structure mentioned in the motivating discussion.

Definition 7.5 (Lattice Group). *For any invertible $d \times d$ matrix A, define the set $L_A \stackrel{\text{def}}{=} \{A\vec{v} \colon \vec{v} \in \mathbb{Z}^d\}$.*

The set L_A is a group under vector addition. Since A is an invertible linear map, it is actually an isomorphism between L_A and $\mathbb{Z}^d$.

Definition 7.6 (Reclusive Partition). *If A is a $d \times d$ reclusive matrix, then we associate to it a partition $\mathcal{P}_A$, called the reclusive partition for A, where*

$$\mathcal{P}_A \stackrel{\text{def}}{=} \left\{\vec{a} + [0,1)^d \colon \vec{a} \in L_A\right\}.$$

Remark 7.7. *Some sources use the notation $L_A + [0,1)^d$ to indicate the set $\mathcal{P}_A$, but we will elect to not do so here.*

Notice that every member of $\mathcal{P}_A$ is a unit hypercube shifted by an element of the lattice group so that for all $X \in \mathcal{P}_A$, $\text{rep}(X) \in L_A$ and conversely, for every $\vec{a} \in L_A$ there is a unit hypercube X whose representative is $\vec{a}$. The proof that this is a partition will not be presented as it is a direct consequence of Proposition 11.2 which shows that for any element $\vec{x} \in \mathbb{R}^d$ there is a unique $X \in \mathcal{P}$ such that $\vec{x} \in X$ by explicitly computing the representative $\text{rep}(X)$.

We will have use for yet another equivalent notion of adjacency, but to do so we must introduce two definitions for types of finite sequences (which we will apply to vectors). These defined sequences will play the essential role in proving that the reclusive partitions have the properties that we are looking for.

Definition 7.8 (Alt-1 and Weak-Alt-1 Sequences). *A finite sequence $\langle c_i \rangle_{i=1}^{n}$ is called alt-1 (alternating sequence of magnitude 1) if $\langle c_i \rangle_{i=1}^{n} = \langle (-1)^i \rangle_{i=1}^{n}$ or $\langle c_i \rangle_{i=1}^{n} = \langle (-1)^{i+1} \rangle_{i=1}^{n}$.*

A finite sequence $\langle c_i \rangle_{i=1}^{n}$ is called weak-alt-1 (weakly alternating sequence of magnitude 1) if all terms are -1, 0, or 1, and the subsequence $\langle c_{i_j} \rangle_{j=1}^{k}$ of non-zero terms is an alt-1 sequence.

Remark 7.9. *We consider the empty sequence to vacuously satisfy these definitions, so in particular, a finite sequence of all zeros is considered weak-alt-1.*

The next lemma is more or less an adaption of the alternating sequence convergence theorem from calculus (c.f. [42]). If we take a dot product of an alt-1 sequence and a strictly monotonic positive sequence, then we know what the sign of that dot product will be and can bound the magnitude.

Lemma 7.10. *Let $\langle c_i \rangle_{i=1}^{n}$ be an alt-1 sequence. Let $\langle a_i \rangle_{i=1}^{n}$ be a strictly decreasing (resp. strictly increasing) positive sequence. Letting $s = \sum_{i=1}^{n} c_i a_i$, the following hold:*
1. *If $n = 1$ then $|s| = a_1$ (resp. $|s| = a_n$)*
2. *If $n \geq 2$ then $a_1 - a_2 \leq |s| \leq a_1$ (resp. $a_n - a_{n-1} \leq |s| \leq a_n$)*
3. *$|s| > 0$*
4. *$\text{sign}(s) = \text{sign}(c_1)$ (resp. $\text{sign}(s) = \text{sign}(c_n)$)*

Proof. Note that (3) is implied by (1) and implied by (2) since $\langle a_i \rangle$ is strictly decreasing (resp. strictly increasing), so (3) need not be proven.

We prove the "increasing" version of the statement because the inductive indexing is cleaner. This immediately implies the stated "decreasing" version by reversing $\langle a_i \rangle$ and reversing $\langle c_i \rangle$.

We prove the "increasing" version by induction on n. If $n = 1$, the claim holds trivially. Otherwise $n > 1$ and assume for inductive hypothesis that the lemma holds for $n - 1$. Then let

$$s \stackrel{\text{def}}{=} \sum_{i=1}^{n} c_i a_i \qquad \text{and} \qquad s' \stackrel{\text{def}}{=} \sum_{i=1}^{n-1} c_i a_i \qquad \text{noting that} \qquad s = s' + c_n a_n.$$

Thus, we have

$$
\begin{aligned}
s &= s' + c_n a_n \\
&= \text{sign}(s')|s'| + c_n a_n & \text{(Decomposition of } s'\text{)} \\
&= \text{sign}(c_{n-1})|s'| + c_n a_n & \text{(Inductive hypothesis)} \\
&= -c_n|s'| + c_n a_n & (\text{sign}(c_{n-1}) = c_{n-1} = -c_n \text{ by alt-1 def'n}) \\
&= c_n(a_n - |s'|).
\end{aligned}
$$

Taking the magnitude we have

$$
\begin{aligned}
|s| &= |c_n||a_n - |s'|| \\
&= |a_n - |s'|| \\
&= a_n - |s'|. & (a_n \text{ and } |s'| \text{ both non-negative})
\end{aligned}
$$

Since $|s'|$ is non-negative, it follows that $|s| \leq a_n$. Further, by inductive hypothesis, $|s'| \leq a_{n-1}$ so again by the last line above, $|s| = a_n - |s'| \geq a_n - a_{n-1}$ which proves (2).

Lastly, noting again that $a_n > a_{n-1} \geq |s'|$ we have

$$
\begin{aligned}
\text{sign}(s) &= \text{sign}(c_n(a_n - |s'|)) \\
&= \text{sign}(c_n)\,\text{sign}(a_n - |s'|) \\
&= c_n \cdot 1 & (c_n = \text{sign}(c_n) \text{ and } a_n - |s'| > 0)
\end{aligned}
$$

which proves (4). $\qquad\qquad\square$

The reason we required $\langle a_i \rangle$ to be strictly monotonic was because otherwise $\text{sign}(c_n)$ could be 0. The above lemma extends very naturally to weak-alt-1 sequences which have at least one non-zero term by applying the lemma to the subsequence of non-zero terms and the corresponding entries of $\langle a_i \rangle$ and we shall sometimes use it as such.

The following proposition will be the key to establishing the (final) equivalent definition of adjacency. The $A\vec{c}$ in the statement will end up being $\text{rep}(X) - \text{rep}(Y)$ for unit hypercubes X and Y, so this proposition will give an equivalent condition for $\|\text{rep}(X) - \text{rep}(Y)\|_\infty \leq 1$ (which by Corollary 6.8 is equivalent to X and Y being adjacent).

Proposition 7.11. *Let A be a $d \times d$ reclusive matrix and $\vec{c} \in \mathbb{Z}^d$ (emphasis: $\vec{c}$ has integer coordinates). Then $\|A\vec{c}\|_\infty \leq 1$ if and only if $\vec{c} = \langle c_j \rangle_{j=1}^{d}$ is a weak-alt-1 sequence.*

Proof. Before proving either direction, let $\vec{x} = A\vec{c}$. Note that for any $i \in [d]$

$$x_i = \sum_{j=1}^{d} a_{ij}c_j \qquad \text{(Def'n of matrix multiplication)}$$

$$= \sum_{j=i}^{d} a_{ij}c_j. \qquad \text{(If } j < i \text{ then } a_{ij} = 0\text{)}$$

Note that $\langle a_{ij}\rangle_{j=i}^{d}$ is a strictly decreasing positive sequence.

We begin by proving the reverse direction. If $\langle c_j\rangle_{j=1}^{d}$ is a weak-alt-1 sequence, then for any $i \in [d]$, the subsequence $\langle c_j\rangle_{j=i}^{d}$ is also a weak-alt-1 sequence. Then by Lemma 7.10 (applied to the subsequence of non-zero terms) and the expression of $\vec{x}$ above, $|x_i| \leq a_{ii} = 1$. Since this holds for all $i \in [d]$, $\|\vec{x}\| \leq 1$.

For the reverse direction, assume that $\langle c_j\rangle_{j=1}^{d}$ is not a weak-alt-1 sequence, in which case we let $k \in [d]$ be the largest integer such that the subsequence $\langle c_j\rangle_{j=k}^{d}$ is not weak-alt-1. By the above expression of $\vec{x}$ we have

$$\|\vec{x}\|_\infty = \max_{i\in[d]}|x_i| \qquad \text{(Def'n)}$$

$$\geq |x_k|$$

$$= |\sum_{j=k}^{d} a_{kj}c_j| \qquad \text{(By expression of } x_k\text{)}$$

$$= |c_k + \sum_{j=k+1}^{d} a_{kj}c_j| \qquad \text{(} A \text{ is reclusive, so } a_{kk} = 1\text{)}$$

If $|c_d| > 1$ then $\langle c_j\rangle_{j=d}^{d}$ is not weak-alt-1, so $k = d$. Then the the above summation is empty, so this shows $\|\vec{x}\|_\infty \geq |c_d| > 1$ and we are done in this case. Otherwise we may assume $|c_d| \leq 1$; further, since $\vec{c} \in \mathbb{Z}^d$ this means c_d is -1, 0, or 1 and thus $\langle c_j\rangle_{j=d}^{d}$ is trivially weak-alt-1 which implies $k \neq d$. Then the sequence $\langle c_j\rangle_{j=k+1}^{d}$ is non-empty (because $k \neq d$) and is weak-alt-1 (by design of k) and contains at least one non-zero term (because otherwise it would be trivially weak-alt-1)—let n denote the number of non-zero terms. Let $\langle j_i\rangle_{i=1}^{n}$ be the sequence of non-zero terms of $\langle c_j\rangle_{j=k+1}^{d}$. Then $\langle c_{j_i}\rangle_{i=1}^{n}$ is an alt-1 sequence, and $\langle a_{kj_i}\rangle_{i=1}^{n}$ is a strictly decreasing sequence (since it is a subsequence of $\langle a_{kj}\rangle_{j=k+1}^{d}$ which is strictly decreasing because A is reclusive). Thus, Lemma 7.10 applies to $s = \sum_{j=k+1}^{d} a_{kj}c_j = \sum_{i=1}^{n} a_{kj_i}c_{j_i}$ (with different cases if $n = 1$ or $n \geq 2$). We complete the proof with cases on the magnitude of c_k (and handle the subcases of the value of n as needed).

Recall that $c_k \in \mathbb{Z}$ and note that $|c_k| \neq 0$ because $\langle c_j\rangle_{j=k+1}^{d}$ is weak-alt-1, so if $c_k = 0$, then $\langle c_j\rangle_{j=k}^{d}$ would be weak-alt-1, but it is not by choice of k. So we consider two cases: $|c_k| = 1$ and $|c_k| \geq 2$.

Case 1: If $|c_k| \geq 2$, then by Lemma 7.10 $|s| \leq a_{kj_1}$ (regardless of the value of n). This gives the following inequalities:

$$\|\vec{x}\|_\infty \geq |c_k + s| \qquad \text{(Work above)}$$

$$\geq |c_k| - |s| \qquad \text{(Triangle inequality)}$$

$$\geq 2 - |s| \qquad \text{(Assumption on } |c_k|\text{)}$$

$$\geq 2 - a_{kj_1} \qquad \text{(Lemma 7.10)}$$

$$= 1 + (1 - a_{kj_1})$$

$$> 1 \qquad \text{(} j_1 \geq k + 1 \text{ so } a_{kj_1} < 1\text{)}$$

22

Case 2: If $|c_k| = 1$, then because $\langle c_j \rangle_{j=k}^d$ is not weak-alt-1, it implies that $\text{sign}(c_k) = \text{sign}(c_{j_1})$. We then get the following inequalities:

$$
\begin{aligned}
\|\vec{x}\|_\infty &\geq |c_k + s| && \text{(Work above)} \\
&= |c_k| + |s| && \text{(Same sign)} \\
&= 1 + |s| && \text{(Assumption on } |c_k|)
\end{aligned}
$$

From here we have two cases depending on if $n = 1$ or $n \geq 2$. If $n = 1$, then by Lemma 7.10 $|s| = a_{kj_1}$, so from the above we have

$$
\begin{aligned}
\|\vec{x}\|_\infty &\geq 1 + a_{kj_1} \\
&> 1 && (j_1 \geq k + 1 \text{ so } a_{kj_1} > 0)
\end{aligned}
$$

If instead $n \geq 2$, then by Lemma 7.10 $|s| \geq a_{kj_1} - a_{kj_2}$, so from the above we have

$$
\begin{aligned}
\|\vec{x}\|_\infty &\geq 1 + (a_{kj_1} - a_{kj_2}) \\
&> 1 && (j_2 > j_1 \geq k + 1 \text{ so } a_{kj_1} - a_{kj_2} > 0)
\end{aligned}
$$

$\square$

Upon inspection, one may note that we can improve upon the statement of the prior proposition. In the proof above, if $\vec{c}$ was not a weak-alt-1 sequence, then not only was $\|A\vec{c}\|_\infty > 1$, but it could be bounded away from 1. This should not be surprising since $\{A\vec{v} : \vec{v} \in \mathbb{Z}^d\}$ is isomorphic to $\mathbb{Z}^d$ (as stated in the discussion of Definition 7.5). Specifically, if $\vec{c}$ was not weak-alt-1, the above proof showed that one of the following four equations held:

$$
\begin{aligned}
\|A\vec{c}\|_\infty &\geq 1 + 1 && (|c_d| > 1) \\
\|A\vec{c}\|_\infty &\geq 1 + (1 - a_{kj_1}) && (|c_d| \leq 1 \text{ and } |c_k| \geq 2) \\
\|A\vec{c}\|_\infty &\geq 1 + (a_{kj_1}) && (|c_d| \leq 1 \text{ and } |c_k| = 1 \text{ and } n = 1) \\
\|A\vec{c}\|_\infty &\geq 1 + (a_{kj_1} - a_{kj_2}) && (|c_d| \leq 1 \text{ and } |c_k| = 1 \text{ and } n \geq 2)
\end{aligned}
$$

If we ignore the specifics of how k, j_1, and j_2 were found, and minimize over all possibilities, it leads to the following definition and corollary.

Definition 7.12 (Reclusive Distance). *If A is a $d \times d$ reclusive matrix, define Δ_A, the reclusive distance of A, as follows (taking $\min \emptyset = \infty$)*

$$
\begin{aligned}
\delta_1 &\overset{\text{def}}{=} 1 \\
\delta_2 &\overset{\text{def}}{=} \min_{k \in [d]} \min_{k < j \leq d} (1 - a_{kj}) \\
\delta_3 &\overset{\text{def}}{=} \min_{k \in [d]} \min_{k < j \leq d} (a_{kj}) \\
\delta_4 &\overset{\text{def}}{=} \min_{k \in [d]} \min_{k < j < j' \leq d} (a_{kj} - a_{kj'})
\end{aligned}
$$

$$
\Delta_A \overset{\text{def}}{=} \min \{\delta_1, \delta_2, \delta_3, \delta_4\}
$$

Observe that because A is reclusive, $\Delta_A > 0$.

Corollary 7.13. *Let A be a $d \times d$ reclusive matrix and $\vec{c} \in \mathbb{Z}^d$ (emphasis: $\vec{c}$ has integer coordinates). Then $\|A\vec{c}\|_\infty \leq 1$ if and only if $\vec{c} = \langle c_j \rangle_{j=1}^d$ is a weak-alt-1 sequence. Further, if $\|A\vec{c}\|_\infty > 1$, then $\|A\vec{c}\|_\infty \geq 1 + \Delta_A$.*

Proof. The proof is implicit in the proof of Proposition 7.11. $\square$

A simple application of this corollary shows that if X and Y are non-adjacent hypercubes in a reclusive partition, then the distance between the representatives of the two hypercubes are separated by one plus the reclusive distance of the partition.

Lemma 7.14 (Adjacent or Far Lemma (Representatives))**.** *Let $d \in \mathbb{N}$, and A be a $d \times d$ reclusive matrix, and $\mathcal{P}_A$ its reclusive partition, and Δ_A its reclusive distance. Let $X, Y \in \mathcal{P}_A$ such that X and Y are not adjacent. Then $d_{max}(\text{rep}(X), \text{rep}\,Y) \geq 1 + \Delta_A$.*

Proof. By definition of $\mathcal{P}_A$, there exists $\vec{m}, \vec{n} \in \mathbb{Z}^d$ such that $A\vec{m} = \text{rep}(X)$ and $A\vec{n} = \text{rep}(Y)$ (in particular $\vec{m} = A^{-1}\text{rep}(X)$ and $\vec{n} = A^{-1}\text{rep}(Y)$). By Corollary 6.8, because X and Y are not adjacent, $d_{max}(\text{rep}(X), \text{rep}(Y)) > 1$, so

$$1 < d_{max}(\text{rep}(X), \text{rep}(Y))$$
$$= \|\text{rep}(X) - \text{rep}(Y)\|_\infty$$
$$= \|A\vec{m} - A\vec{n}\|_\infty$$
$$= \|A(\vec{m} - \vec{n})\|_\infty.$$

By Corollary 7.13, since $\|A(\vec{m} - \vec{n})\| > 1$ it must be that $\|A(\vec{m} - \vec{n})\| \geq 1 + \Delta_A$. $\square$

A similar result holds when considering general elements $\vec{x}, \vec{y}$ of non-adjacent hypercubes X and Y. We can get this result because the location of $\vec{x}$ relative to $\text{rep}(X)$ is similar to the location of $\vec{y}$ relative to $\text{rep}(Y)$.

Lemma 7.15 (Adjacent or Far Lemma (Points))**.** *Let $d \in \mathbb{N}$, and A be a $d \times d$ reclusive matrix, and $\mathcal{P}_A$ its reclusive partition, and Δ_A its reclusive distance. Let $X, Y \in \mathcal{P}_A$ such that X and Y are not adjacent. For all $\vec{x} \in X$ and $\vec{y} \in Y$, it holds that $d_{max}(\vec{x}, \vec{y}) > \Delta_A$.*

Proof. By the definition of unit hypercubes, $\vec{x} \in \text{rep}(X) + [0, 1)^d$, so let $\vec{\alpha} \in [0, 1)^d$ such that $\vec{x} = \text{rep}(X) + \vec{\alpha}$. Similarly, let $\vec{\beta} \in [0, 1)^d$ such that $\vec{y} = \text{rep}(Y) + \vec{\beta}$. As in the proof of the prior lemma, let $\vec{m} = A^{-1}\text{rep}(X)$ and $\vec{n} = A^{-1}\text{rep}(Y)$.

Since $\vec{\alpha}, \vec{\beta} \in [0, 1)^d$, it follows that $\vec{\beta} - \vec{\alpha} \in (-1, 1)^d$ so $\|\vec{\beta} - \vec{\alpha}\|_\infty < 1$. Then we have the following:

$$d_{max}(\vec{x}, \vec{y}) = \|\vec{x} - \vec{y}\|_\infty$$
$$= \|(A\vec{m} + \vec{\alpha}) - (A\vec{n} + \vec{\beta})\|_\infty$$
$$= \|(A(\vec{m} + \vec{n})) - (\vec{\beta} - \vec{\alpha})\|_\infty$$
$$\geq \|(A(\vec{m} + \vec{n}))\|_\infty - \|\vec{\beta} - \vec{\alpha}\|_\infty \qquad \text{(Triangle inequality)}$$
$$\geq (1 + \Delta_A) - \|(\vec{\beta} - \vec{\alpha})\|_\infty \qquad \text{(By previous lemma)}$$
$$> (1 + \Delta_A) - 1 \qquad (\|\vec{\beta} - \vec{\alpha}\|_\infty < 1)$$
$$= \Delta_A$$

Noting the strict inequality in the second to last line completes the proof. $\square$

This lemma will be important later when we need a fixed bound on the distances between non-adjacent partition members.

Now that we have given a bound on how close non-adjacent hypercubes can be, we want to turn our attention to how many hypercubes can be pairwise adjacent. In other words, we want to show a bound on the size of the largest clique in the partition graph of a reclusive partition. We actually do something stronger and give an explicit coloring of the graph (an explicit coloring of the hypercubes). If a graph can be properly colored with n colors, then the size of the largest clique in the graph is at most n.

Theorem 7.16 (Coloring Reclusive Partitions)**.** *Let $d \in \mathbb{N}$ and let A be a $d \times d$ reclusive matrix. The graph of the reclusive partition $\mathcal{P}_A$ can be properly $(d+1)$-colored.*

Proof. Let $\chi = [1, 2, \ldots, d]$ be a $1 \times d$ matrix. Define the coloring function on the hypercubes as follows:

$$\text{color} : \mathcal{P}_A \to \mathbb{Z}_{d+1}$$
$$\text{color}(X) = \chi \, A^{-1} \, \text{rep}(X) \mod (d+1).$$

Recall that if $X \in \mathcal{P}_A$, then $\text{rep}(X) = A\vec{m}$ for some $\vec{m} \in \mathbb{Z}^d$ (by definition of the reclusive partition $\mathcal{P}_A$), and $\vec{m}$ is unique because A is invertible, so the definition of color has an appropriate codomain because $(A^{-1}(A\vec{m})) = \vec{m} \in \mathbb{Z}^d$, and taken as a column vector it can be multiplied with χ to obtain an integer.

Let X and Y be distinct hypercubes in $\mathcal{P}_A$ such that $X \overset{\text{adj}}{\sim} Y$. We must show that $\text{color}(X) \neq \text{color}(Y)$. We do so by looking at the differences of the colors (and explicitly emphasize that it is being done mod $(d+1)$). Let $\vec{m} = A^{-1} \text{rep}(X)$ and $\vec{n} = A^{-1} \text{rep}(Y)$.

$$
\begin{aligned}
\text{color}(X) - \text{color}(Y) \mod (d+1) &= \big[(\chi\vec{m} \mod (d+1)) - (\chi\vec{n} \mod (d+1))\big] \mod (d+1) \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\text{(definition)} \\
&= \big[\chi\vec{m} - \chi\vec{n}\big] \mod (d+1) \qquad \text{(property of modular arithmetic)} \\
&= \big[\chi(\vec{m} - \vec{n})\big] \mod (d+1) \qquad\qquad\qquad \text{(linearity of } \chi\text{)}
\end{aligned}
$$

Observe that by Proposition 7.11, because $X \overset{\text{adj}}{\sim} Y$ it follows that $(\vec{m} - \vec{n})$ is a weak-alt-1 sequence/vector. Further, $(\vec{m} - \vec{n})$ has at least one non-zero term (if it did not, then $\vec{m} = \vec{n}$ so $\text{rep}(X) = \text{rep}(Y)$ but we assumed X and Y were distinct). Note also that the matrix product of χ with $(\vec{m} - \vec{n})$ is really a dot product of the increasing positive sequence $\langle 1, 2, \ldots, d \rangle$ with a weak-alt-1 sequence with a non-zero term. Thus, by Lemma 7.10 (applied to the (non-empty) subsequence of non-zero terms), we have

$$0 < |\chi(\vec{m} - \vec{n})| \leq d.$$

Thus, $\chi(\vec{m} - \vec{n}) \mod (d+1) \neq 0$ which proves that $\text{color}(X) \neq \text{color}(Y)$, so this is a proper coloring. $\qquad\square$

In fact, the coloring above is tight—the chromatic number (the smallest number of colors that can be used to color the graph) is $d+1$. To prove this, it suffices to show that $\mathcal{P}_A$ has a $(d+1)$-clique.

The following result actually follows as a corollary to the Optimality Theorem for Unit Hypercube Partitions (Theorem 7.20), so the following proof is not strictly necessary; however, Theorem 7.20 uses far more machinery, and it is fairly simple to find an explicit clique in the partition.

Proposition 7.17 (Chromatic Number)**.** *Let $d \in \mathbb{N}$, and A be a $d \times d$ reclusive matrix, and $\mathcal{P}_A$ its reclusive partition. There exists a clique in $\mathcal{P}_A$ of size $d + 1$.*

Proof. For each $i \in [d]$, let $\vec{e}_i \in \mathbb{Z}^d$ denote the ith standard basis vector (i.e. all zeros except for the ith term which is 1). Let $V = \left\{\vec{0}\right\} \cup \{\vec{e}_i : i \in [d]\}$. Note that for any two vectors $\vec{v}, \vec{w} \in V$, that $\vec{v} - \vec{w}$ is a weak-alt-1 vector. This is obvious if one of the vectors is $\vec{0}$ because then the difference is a standard basis vector which is all zeros except for a single term which is 1, so it is weak-alt-1. Otherwise, $\vec{v} = \vec{e}_i$ and $\vec{w} = \vec{e}_j$ for some $i, j \in [d]$. If $i = j$, then the difference is $\vec{0}$ which is trivially weak-alt-1, and if $i \neq j$, then

the difference is all zeros except for a term which is $+1$ and a term which is -1; such a vector must be a weak-alt-1 sequence.

Let $R = \{A\vec{v} \colon \vec{v} \in V\}$ denote a set of representatives, and let $C = \{X \in \mathcal{P}_A \colon \mathrm{rep}(X) \in R\}$ be the set of unit hypercubes in $\mathcal{P}_A$ whose representatives are in R. Then C (and V) have cardinality $d+1$, and we claim that C is a clique.

Consider hypercubes $X, Y \in C$, so $\mathrm{rep}(X), \mathrm{rep}(Y) \in R$, so $A^{-1}\mathrm{rep}(X), A^{-1}\mathrm{rep}(Y) \in V$. Let $\vec{m} = A^{-1}\mathrm{rep}(X)$ and $\vec{n} = A^{-1}\mathrm{rep}(Y)$. As described, since $\vec{m}, \vec{n} \in V$, we have that $\vec{m} - \vec{n}$ is a weak-alt-1 sequence. By Proposition 7.11 this implies that $\|A(\vec{m} - \vec{n})\|_\infty \leq 1$. Thus,

$$
\begin{aligned}
d_{max}(\mathrm{rep}(X), \mathrm{rep}(Y)) &= \|\mathrm{rep}(X) - \mathrm{rep}(Y)\|_\infty \\
&= \|A(\vec{m} - \vec{n})\|_\infty \\
&\leq 1
\end{aligned}
$$

and by Corollary 6.8, X and Y are adjacent. Since this holds for any two hypercubes in C, it must be that C is a $(d+1)$-clique. $\qquad\square$

The following theorem is the main result of this section showing that we have constructed partitions satisfying the motivating question with $k = d+1$ and $\epsilon = \frac{\Delta_A}{2}$. Then the Hypercube Parition Theorem (Theorem 7.20) mentioned in Section 2 will be a simple corollary of this by demonstrating the existence of a reclusive partition with $\Delta_A = \frac{1}{d}$.

Theorem 7.18 (Partition Theorem). *Let $d \in \mathbb{N}$, and A be a $d \times d$ reclusive matrix, and $\mathcal{P}_A$ its reclusive partition, and Δ_A its reclusive distance. Then for any $\vec{p} \in \mathbb{R}^d$,*

$$
|\mathcal{N}_{\Delta_A/2}(\vec{p})| \leq d+1.
$$

Proof. It suffices to prove that $\mathcal{N}_{\Delta_A/2}(\vec{p})$ is a clique since by Theorem 7.16, any clique contains at most $d+1$ hypercubes. Let $X, Y \in \mathcal{N}_{\Delta_A/2}(\vec{p})$ be arbitrary. By definition of $\mathcal{N}_{\Delta_A/2}(\vec{p})$, there exists $\vec{x} \in X$ such that $d_{max}(\vec{p}, \vec{x}) \leq \frac{\Delta_A}{2}$ and similarly there exists $\vec{y} \in Y$ such that $d_{max}(\vec{p}, \vec{y}) \leq \frac{\Delta_A}{2}$. By the triangle inequality of metrics,

$$
d_{max}(\vec{x}, \vec{y}) \leq d_{max}(\vec{p}, \vec{x}) + d_{max}(\vec{p}, \vec{y}) \leq \frac{\Delta_A}{2} + \frac{\Delta_A}{2} = \Delta_A
$$

By the contrapositive of the Adjacent or Far Lemma (Lemma 7.15), since there exists $\vec{x} \in X$ and $\vec{y} \in Y$ such that $d_{max}(\vec{x}, \vec{y}) \leq \Delta_A$, it must be that X and Y are adjacent. $\qquad\square$

In order to state the Hypercube Parition Theorem (Theorem 7.20) as promised, we first need a simple lemma. To motivate the following choice, consider again the definition of reclusive distance (Definition 7.12) and the discussion leading up to it. To make the most of the Partition Theorem, we want to have a large reclusive distance, and that is accomplished by keeping three key quantities in a reclusive matrix large—for each row k and arbitrary entries $j < j'$ within that row, we want the following to be large: $(1 - a_{kj})$, (a_{kj}), and $(a_{kj} - a_{kj'})$. The first discourages using matrix entries greater than 1 (which is partially why reclusive matrices were defined to not allow that) and encourages using small entries; the second encourages using large entries; the third seems to encourage "even spacing" of the terms in a given row.

Based on these considerations, we make all entries in the matrix multiples of a common denominator.

Lemma 7.19. *Let $d \in \mathbb{N}$, and A be a $d \times d$ reclusive matrix, and Δ_A its reclusive distance. If for all $i, j \in [d]$ it holds that a_{ij} is a multiple of $\frac{1}{k}$ (including 0), then $\Delta_A \geq \frac{1}{k}$. Further, such reclusive matrices exist for $k \geq d$.*

Proof. The first claim follows from the definitions of reclusive matrix and reclusive distance. For the second claim, let $k \geq d$ and consider the $d \times d$ reclusive matrix A which has 1's on the main diagonal and 0's in the lower triangle (as required for all reclusive matrices) and for any other entry, $a_{ij} = \frac{d-i+1}{k}$.

$$A = \begin{bmatrix} 1 & \frac{d-1}{k} & \frac{d-2}{k} & \frac{d-3}{k} & \cdots & \frac{2}{k} & \frac{1}{k} \\ 0 & 1 & \frac{d-2}{k} & \frac{d-3}{k} & \cdots & \frac{2}{k} & \frac{1}{k} \\ 0 & 0 & 1 & \frac{d-3}{k} & \cdots & \frac{2}{k} & \frac{1}{k} \\ 0 & 0 & 0 & 1 & \cdots & \frac{2}{k} & \frac{1}{k} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & \frac{1}{k} \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix}$$

It is easily verified that this is in fact a reclusive matrix. $\qquad\square$

We shall be particularly interested in reclusive $d \times d$ matrices A as in the lemma with $k = d$ in which case the lemma says that $\Delta_A \geq \frac{1}{d}$, and by examination of the first row $\Delta_A \leq \frac{1}{d}$ giving equality.

Theorem 7.20 (Hypercube Partition Theorem). *Let $d \in \mathbb{N}$. Then there exists a $(d+1, \frac{1}{2d})$-secluded unit hypercube partition $\mathcal{P}$ of $\mathbb{R}^d$. I.e., for any point $\vec{p} \in \mathbb{R}^d$, we have*

$$|\mathcal{N}_{\frac{1}{2d}}(\vec{p})| \leq d + 1.$$

Proof. As just discussed, there exists a reclusive partition $\mathcal{P}_A$ with with reclusive distance $\Delta_A = \frac{1}{d}$, so the conclusion follows from Theorem 7.18. $\qquad\square$

8 Fundamental Property of the Reclusive Definition

It would be a fair question for one to ask, "Why bother developing this general notion of reclusive partitions if you end up only using one specific example?" The initial reason for why we defined reclusive partitions is that we found ourselves looking to generalize the example of exponentially decaying shifts that we presented at the beginning of the section in order to move from ϵ with the denominator growing exponentially to ϵ with the denominator growing linearly. The generalization followed by noting the properties in Subsection 7.2.

Initially, we expected this to be a much stronger generalization than we needed and expected that the order in which the shifts were applied would not really matter. However, as we will demonstrate in this subsection, we were incorrect, and the definition of a reclusive matrix really seems to capture something that seems fundamental to the properties we are interested in—particularly the property of considering partitions with minimal possible clique sizes.

Consider, for example, the following two reclusive matrices A' and A'' (the shorthand notation denotes the value only of entries in the strict upper diagonal since the diagonal must be 1's and the subdiagonal must be 0's).

$$A' = \left(\frac{d-j+1}{d}\right) = \begin{bmatrix} 1 & \frac{d-1}{d} & \frac{d-2}{d} & \frac{d-3}{d} & \cdots & \frac{2}{d} & \frac{1}{d} \\ 0 & 1 & \frac{d-2}{d} & \frac{d-3}{d} & \cdots & \frac{2}{d} & \frac{1}{d} \\ 0 & 0 & 1 & \frac{d-3}{d} & \cdots & \frac{2}{d} & \frac{1}{d} \\ 0 & 0 & 0 & 1 & \cdots & \frac{2}{d} & \frac{1}{d} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & \frac{1}{d} \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix} \quad \text{Example, } d = 5: \begin{bmatrix} 1 & \frac{4}{5} & \frac{3}{5} & \frac{2}{5} & \frac{1}{5} \\ 0 & 1 & \frac{3}{5} & \frac{2}{5} & \frac{1}{5} \\ 0 & 0 & 1 & \frac{2}{5} & \frac{1}{5} \\ 0 & 0 & 0 & 1 & \frac{1}{5} \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

27

$$A'' = \left(\frac{d-j+1}{d-i+1}\right) = \begin{bmatrix} 1 & \frac{d-1}{d} & \frac{d-2}{d} & \frac{d-3}{d} & \cdots & \frac{2}{d} & \frac{1}{d} \\ 0 & 1 & \frac{d-2}{d-1} & \frac{d-3}{d-1} & \cdots & \frac{2}{d-1} & \frac{1}{d-1} \\ 0 & 0 & 1 & \frac{d-3}{d-2} & \cdots & \frac{2}{d-2} & \frac{1}{d-2} \\ 0 & 0 & 0 & 1 & \cdots & \frac{2}{d-3} & \frac{1}{d-3} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix} \qquad \text{Example, } d = 5: \begin{bmatrix} 1 & \frac{4}{5} & \frac{3}{5} & \frac{2}{5} & \frac{1}{5} \\ 0 & 1 & \frac{3}{4} & \frac{2}{4} & \frac{1}{4} \\ 0 & 0 & 1 & \frac{2}{3} & \frac{1}{3} \\ 0 & 0 & 0 & 1 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Then consider the matrices B' and B''.

$$B' = \left(\frac{j-1}{d}\right) = \begin{bmatrix} 1 & \frac{1}{d} & \frac{2}{d} & \frac{3}{d} & \cdots & \frac{d-2}{d} & \frac{d-1}{d} \\ 0 & 1 & \frac{2}{d} & \frac{3}{d} & \cdots & \frac{d-2}{d} & \frac{d-1}{d} \\ 0 & 0 & 1 & \frac{3}{d} & \cdots & \frac{d-2}{d} & \frac{d-1}{d} \\ 0 & 0 & 0 & 1 & \cdots & \frac{d-2}{d} & \frac{d-1}{d} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & \frac{d-1}{d} \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix} \qquad \text{Example, } d = 5: \begin{bmatrix} 1 & \frac{1}{5} & \frac{2}{5} & \frac{3}{5} & \frac{4}{5} \\ 0 & 1 & \frac{2}{5} & \frac{3}{5} & \frac{4}{5} \\ 0 & 0 & 1 & \frac{3}{5} & \frac{4}{5} \\ 0 & 0 & 0 & 1 & \frac{4}{5} \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$B'' = \left(\frac{i-j}{d-i+1}\right) = \begin{bmatrix} 1 & \frac{1}{d} & \frac{2}{d} & \frac{3}{d} & \cdots & \frac{d-2}{d} & \frac{d-1}{d} \\ 0 & 1 & \frac{1}{d-1} & \frac{2}{d-1} & \cdots & \frac{d-3}{d-1} & \frac{d-2}{d-1} \\ 0 & 0 & 1 & \frac{1}{d-2} & \cdots & \frac{d-4}{d-2} & \frac{d-3}{d-2} \\ 0 & 0 & 0 & 1 & \cdots & \frac{d-5}{d-3} & \frac{d-4}{d-3} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix} \qquad \text{Example, } d = 5: \begin{bmatrix} 1 & \frac{1}{5} & \frac{2}{5} & \frac{3}{5} & \frac{4}{5} \\ 0 & 1 & \frac{1}{4} & \frac{2}{4} & \frac{3}{4} \\ 0 & 0 & 1 & \frac{1}{3} & \frac{2}{3} \\ 0 & 0 & 0 & 1 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Clearly B' and B'' are not reclusive matrices[14] since the entries in the rows are not decreasing. Nonetheless, B' and B'' define partitions of $\mathbb{R}^d$ in the same fashion as a reclusive matrix[15]. For example, the partition associated with B' in dimension 5 is the one constructed by partitioning $\mathbb{R}^1$ into unit intervals, then extruding those intervals into unit squares in $\mathbb{R}^2$, copying the extrusion to multiple layers, and shifting each layer by $1/5$ of a unit more to the right than the previous layer; then extruding this partition of $\mathbb{R}^2$ into $\mathbb{R}^3$ and shifting each layer by $2/5$ of a unit more than the previous layer. In this sense, this partition is constructed in a *very* similar way to the partition for A': layers are still offset by multiples of $1/5$, and the only difference is that smaller shifts happen first.

By Theorem 7.16, the partitions $\mathcal{P}_{A'}$ and $\mathcal{P}_{A''}$ do not contain any cliques of size $d+2$, so a natural question is whether the partitions associated with B' and B'' also have this property since their constructions are so similar. Our intuition was that the answer to this question would "yes", but this is not the case. For $d \leq 4$, the partitions for both B' and B'' do not have cliques of size $d+2$, but for $d = 5$ they do.

We found this by exhaustive computer search[16], but our claim that for $d = 5$, B' and B'' each have a $d + 2 = 7$ clique can easily be verified. Consider the following set of 7 vectors in $\mathbb{Z}^5$:

---

[14] Well, they are for $d = 1$ and $d = 2$, but not for any $d \geq 3$.

[15] The partition is the set of all unit hypercubes with representatives in the set $\{B'\vec{v}: \vec{v} \in \mathbb{Z}^d\}$ (resp. $\{B''\vec{v}: \vec{v} \in \mathbb{Z}^d\}$).

[16] Because of the repetitive structure of the partition (since the underlying structure is a lattice), it suffices to check a sufficiently large but finite subset of the partition for cliques to determine the size of the largest clique in the whole partition.

$$\begin{bmatrix} 0 \\ 1 \\ -1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ -1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ -1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

Multiplying B' with each vector results in the representative corner of the hypercubes associated with each vector (one can similarly compute them for B''):

$$\begin{bmatrix} 3/5 \\ 7/5 \\ -1/5 \\ 4/5 \\ 1 \end{bmatrix}, \begin{bmatrix} 2/5 \\ 2/5 \\ -1/5 \\ 4/5 \\ 1 \end{bmatrix}, \begin{bmatrix} 1/5 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -2/5 \\ 7/5 \\ -1/5 \\ 4/5 \\ 1 \end{bmatrix}, \begin{bmatrix} -1/5 \\ 4/5 \\ 4/5 \\ 4/5 \\ 1 \end{bmatrix}, \begin{bmatrix} -2/5 \\ 3/5 \\ 3/5 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 3/5 \\ 3/5 \\ 3/5 \\ 1 \\ 0 \end{bmatrix}$$

To see that the hypercubes with representative corners at these 7 locations form a clique, check that the d_{max} distance between any pair is exactly 1 (and apply Corollary 6.8). By Proposition 6.9, and the discussion preceeding it, this shows that this partition has points where any ball of any radius centered at that point intersect all 7 of these hypercubes.

It could be that there is some sufficiently large dimension d such that B' and B'' have cliques of size at most $d + 1$, but we conjecture that this is not the case.

Conjecture 8.1 (Maximum Clique Sizes in Non-Reclusive Partitions B' and B''). *Based on our computations, we conjecture that the maximum clique size in the partitions associated with B' and B'' is greater than $d + 1$ for all $d \geq 5$, and we know the maximum clique sizes for dimensions given in Table 1.*

d	1	2	3	4	5	6	7	8	9	10	11	12
B'	2	3	4	5	7	9	12	16	22	30	39	51
B"	2	3	4	5	7	9	11	16	21	28	36	47

Table 1: Maximum clique sizes in non-reclusive partitions B' and B'' for various dimensions.

The above example demonstrates that our definition of reclusive partitions is not an arbitrary one and captures a certain structure of lattice based partitions that is sufficient to ensure that no large cliques exist. The key is that because our definition of reclusive partitions demands that the terms be decreasing in each row, we get an equivalent definition of adjacency in terms of weak-alt-1 sequences. The above example shows that this equivalence does not hold if we relax the decreasing requirement.

9 Optimality of the Degree Parameter (k)

Given our results that we can find a uniformly-sized neighborhood around every point that intersects at most $d + 1$ members of the partition, it is natural to ask if this is optimal—can this value be improved from $d + 1$ to d or even smaller? In other words, is there a partition of $\mathbb{R}^d$ and a neighborhood around each point in $\mathbb{R}^d$ such that each of these neighborhoods intersects at most d members of the partition? The answer is basically no; our partitions are optimal, and in this section we will discuss the technicalities of what we mean by "basically" and then prove the optimality formally.

It is obvious that this question above is only worth asking if some restrictions are added, because otherwise the partition of $\mathbb{R}^d$ which contains just one member ($\mathcal{P} = \{\mathbb{R}^d\}$) along with any neighborhoods trivially has the property that each neighborhood intersects only one member of the partition (because there is only one). The context above suggests that we are interested in partitions similar to tilings and

packings, and so we will restrict to the context in which the size of the partition members is uniformly bounded in some way.

Two natural ways to bound the size of the members are by bounding the diameter and by bounding the measure. Note that bounding the diameter is a strictly stronger condition (assuming all members are measurable), because if we insist that each member of the partition has diameter at most D in the d_{max} metric, then the measure of each member is at most $(2D)^d$ (fix some point in the member, and all other points are within D of it, so the member is contained in a ball of radius D in the d_{max} metric, and this ball has measure $(2D)^d$ because this ball is a d-dimensional hypercube with side length $2D$).

Thus, if we can answer this question by just restricting the measure, that would be ideal since it is a weaker hypothesis. Unfortunately, we will see that we cannot get the strongest version of the conclusion that we want just by bounding the measure, and so we will also consider diameter bounds. It will occur frequently in our theorems, though, that to get nicely generalized statements, diameter is not quite what we want. For example, consider the sets $[0, 1)$ and $[0, 1]$. Both sets have diameter 1, but the the former has the stronger property that all pairs of points have distance strictly less than the diameter. In other words, diameter[17] is defined as a supremum of distances, and the former set does not attain the supremum while the latter does. We want to distinguish between these cases, so we give the following definition.

Definition 9.1 (Strict Pairwise Bound)**.** *If X is a subset of a metric space and $D \in (0, \infty)$ is a constant such that for all $x, y \in X$ it holds that $\mathrm{distance}(x, y) < D$, then we call D a strict pairwise bound of X.*

Remark 9.2. *Unbounded sets do not have any strict pairwise bounds, and bounded sets have infinitely many strict pairwise bounds (if D is a strict pairwise bound, then so is $D + \epsilon$ for any $\epsilon \in [0, \infty)$). Also, we have the following implications:*

$$\mathrm{diam}(X) < D \quad \implies \quad D \text{ is a strict pairwise bound of } X \quad \implies \quad \mathrm{diam}(X) \leq D$$

In general, the reverse implications do not hold which is exemplified by the sets $[0, D)$ and $[0, D]$.

Because of the distinction just mentioned between a strict bound on diameter, a strict pairwise bound, and a nonstrict bound on diameter, we don't want to restrict to unit diameter partition members in this section as we did in the last section, and instead will state many results for a general diameter (or strict pairwise bound) $D \in (0, \infty)$. It is worth noting that all of the results of the previous section used unit hypercubes, but if we allow for D diameter hypercubes (i.e., translations of $[0, D)^d$) then the results also scale so that we can still attain the parameter $k = d + 1$ and $\epsilon = \frac{D}{2d}$.

In this section, we will prove the three different Optimality Theorems presented in Section 2. We begin with a more in depth discussion (and restatement) of these theorems and some of the corollaries that will follow. We will then introduce the necessary tools (primarily a variant of Sperner's lemma) for proving the Optimality Theorems, and prove them. Lastly, we will prove that there are "gaps" between each of these theorems which justifies the need for multiple variants.

9.1 The Optimality Theorems

In the statements below, let m denote the Lebesgue measure of a set. For those unfamiliar with measure theory, sets that are "not too bizarre" are called Lebesgue measurable, and for any Lebesgue measurable set $X \subseteq \mathbb{R}^d$, $m(X)$ is a generalization of the volume of X. Sets that are not Lebesgue measurable (ones that are "too bizarre") don't have a well-defined volume/measure.

We will state each of the three theorems and then give the interpretation of each.

---

[17]In general $\mathrm{diam}(X) \overset{\mathrm{def}}{=} \sup \{\mathrm{distance}(x, y) : x, y \in X\}$

Theorem 9.3 (First Optimality Theorem). *If $d \in \mathbb{N}$, and $\mathcal{P}$ is a partition of $\mathbb{R}^d$, and there exists $M \in (0, \infty)$ such that for all $X \in \mathcal{P}$, X is Lebesgue measurable, and $m(X) < M$, then for any $\epsilon \in (0, \infty)$, there exists $\vec{p} \in \mathbb{R}^d$ such that*

$$|\mathcal{N}_\epsilon(\vec{p})| \geq d + 1.$$

The conclusion of the First Optimality Theorem is that for partitions satisfying the hypothesis, there is no way to put uniformly sized neighborhoods (each a ball of radius ϵ) centered at every point of $\mathbb{R}^d$ each intersecting only d members of the partition. For any chosen ϵ, there will be a point in the space where its neighborhood intersects at least $d + 1$ partition members.

Theorem 9.4 (Second Optimality Theorem). *If $d \in \mathbb{N}$, and $\mathcal{P}$ is a partition of $\mathbb{R}^d$, and there exists $D \in (0, \infty)$ such that for all $X \in \mathcal{P}$, it holds that $\mathrm{diam}(X) < D$, then for any $\epsilon : \mathbb{R}^d \to (0, \infty)$, there exists $\vec{p} \in \mathbb{R}^d$ such that*

$$|\mathcal{N}_{\epsilon(\vec{p})}(\vec{p})| \geq d + 1.$$

The ϵ function above should be viewed as some fixed ϵ for each point of $\mathbb{R}^d$. The conclusion of the Second Optimality Theorem is that for partitions satisfying the hypothesis, there is no way to put neighborhoods (each a ball of some radius) centered at every point of $\mathbb{R}^d$ each intersecting only d members of the partition even if we allow the radius of the neighborhood to depend on which point it is centered at. There will be a point in the space where its neighborhood intersects at least $d + 1$ partition members. This is a stronger conclusion than the First Optimality Theorem, but the hypothesis is also stronger (assuming all members are measurable) as discussed at the beginning of this section.

Theorem 9.5 (Third Optimality Theorem). *If $d \in \mathbb{N}$, and $\mathcal{P}$ is a partition of $\mathbb{R}^d$, and there exists $D \in (0, \infty)$ such that for all $X \in \mathcal{P}$, D is a strict pairwise bound for X (it is sufficient but not necessary that $\mathrm{diam}(X) < D$), and if there exists some $\vec{\alpha} \in \mathbb{R}^d$ such that $\vec{\alpha} + [0, D]^d$ intersects finitely many members of $\mathcal{P}$, then there exists $\vec{p} \in \mathbb{R}^d$ such that*

$$|\mathcal{N}_{\vec{0}}(\vec{p})| \geq d + 1.$$

Furthermore, $\mathcal{P}$ contains a $(d + 1)$-clique.

The conclusion of the Third Optimality Theorem is that for partitions satisfying the hypothesis, we don't even need to consider neighborhoods, because there is a single point in the closure of at least $d + 1$ members of the partition, and thus for this specific point, any sized neighborhood with intersect $d + 1$ members of the partition. This further implies that the partition has a $(d + 1)$-clique (the intersection of closures of these $d + 1$ members is non-empty because it contains the point $\vec{p}$). Again, this conclusion is stronger than the conclusions of the Second Optimality Theorem, but the hypothesis is also stronger.

Remark 9.6. *Observe that in the First (resp. Second) Optimality Theorem, the strict inequality on the measure (resp. diameter) could be replaced with a non-strict inequality, and the statement would be equivalent since the actual value of M (resp. D) is not used in the conclusion. However, this is not the case with the Third Optimality Theorem. Consider for example if $\mathcal{P} = \{[0, 1]^d\} \cup \{\{\vec{x}\} : \vec{x} \notin [0, 1]^d\}$ (i.e. the partition where one member is the hypercube $[0, 1]^d$, and every other member is a singleton set). Then, it holds for $D = 1$ that for all $X \in \mathcal{P}$, $\mathrm{diam}(X) \leq 1$ and taking $\vec{\alpha} = \vec{0}$, the set $\vec{\alpha} + [0, 1]^d$ only intersects one member of the partition (and in fact intersects the closure of only one member of the partition since all members are already closed sets). However, because all members are closed, for any $\vec{p} \in \mathbb{R}^d$, the set $\mathcal{N}_{\vec{0}}(\vec{p}) = \{X \in \mathcal{P} : \overline{X} \ni \vec{p}\}$ has cardinality $1 < d + 1$ because it contains only the set $\mathrm{member}(\vec{p})$. Thus, the use of the strict inequality of the statement will be necessary.*

The hypothesis added in the Third Optimality Theorem is that there is a local finiteness somewhere in the partition. Essentially this limits the resolution so that we can examine some portion of the partition and not deal with infinitely small sets. This finiteness condition arises in the natural context below.

Corollary 9.7. *Let $d \in \mathbb{N}$ and $\mathcal{P}$ a partition of $\mathbb{R}^d$. If there exists $D \in (0, \infty)$ and $\mu \in (0, \infty)$ such that for all $X \in \mathcal{P}$, it holds that X is Lebesgue measurable, and $\mu < m(X)$, and $\mathrm{diam}(X) < D$, then there exists $\vec{p} \in \mathbb{R}^d$ such that*

$$|\mathcal{N}_\epsilon(\vec{p})| \geq d + 1.$$

Furthermore, $\mathcal{P}$ contains a $(d+1)$-clique.

Proof. It suffices to prove that the hypotheses of the corollary imply the hypothesis of the Third Optimality Theorem.

Let $\vec{\alpha} \in \mathbb{R}^d$ be arbitrary. Let $B = \vec{\alpha} + [0, D]^d$, and $\mathcal{B} = \{X \in \mathcal{P} : X \cap B \neq \emptyset\}$ (the members intersecting B). Let $B' = \vec{\alpha} + [0 - D, D + D]^d$ (which should be viewed as ball containing B and all points within distance D of B in the d_{max} metric). Thus, for any $X \in \mathcal{B}$, because $\mathrm{diam}(X) < D$, we have $X \subseteq B'$, and thus $\bigsqcup_{X \in \mathcal{B}} X \subseteq B'$ (the square cup indicating a disjoint union).

Consider the following volume argument (see the Appendix for details on the second inequality):

$$(3D)^d = m(B') \geq m(\bigsqcup_{X \in \mathcal{B}} X) \geq \sum_{X \in \mathcal{B}} m(X) \geq |\mathcal{B}|\mu.$$

Dividing both sides by μ gives $|\mathcal{B}| \leq \frac{(3D)^d}{\mu}$ which is finite and thus demonstrates that the hypotheses of the Third Optimality Theorem hold and completes the proof. $\quad\square$

The following is also a simple corollary, but we state it as a theorem to emphasize that in the case of unit hypercube partitions, we get the strongest of the results above.

Theorem 9.8 (Optimality for Unit Hypercube Partitions). *Let $d \in \mathbb{N}$ and $\mathcal{P}$ a unit hypercube partition of $\mathbb{R}^d$. Then there exists $\vec{p} \in \mathbb{R}^d$ such that*

$$|\mathcal{N}_{\vec{0}}(\vec{p})| \geq d + 1$$

Furthermore, $\mathcal{P}$ contains a $(d+1)$-clique.

Proof. All members of $\mathcal{P}$ are measurable and have measure and diameter 1, so apply Corollary 9.7. $\quad\square$

As we have mentioned before, the reason we believe the bounds on measure and diameter of the partition members to be very reasonable is that we are interested in partitions which are something like tilings. In a tiling, there is some finite collection of members used with translation and rotation (and other orthogonal maps). Because all norms on $\mathbb{R}^d$ are equivalent. If each member of a finite collection has finite diameter in some norm, then they also have finite diameter in the $\|\cdot\|_\infty$ norm, and because the collection of members is finite, there is some value $D \in (0, \infty)$ so that all of these members have diameter less than D in the $\|\cdot\|_\infty$ norm. Also in a tiling, the members usually have non-zero measure, so again by finiteness there is some $\mu \in (0, \infty)$ which is a lower bound on the measure of each member (and measure is invariant under rotation and translation and other orthogonal maps). So by this reasoning we have the following simple corollary which is a slight generalization of the optimality theorem above for unit hypercubes.

Corollary 9.9 (Tiling Corollary). *Let $d \in \mathbb{N}$, and $\mathcal{F}$ be a finite collection of Lebesgue measureable subsets of $\mathbb{R}^d$ each with finite diameter and finite measure. Let $\mathcal{P}$ be a partition of $\mathbb{R}^d$ consisting only of translated orthogonal transformations of members of $\mathcal{F}$. Then there exists $\vec{p} \in \mathbb{R}^d$ such that*

$$|\mathcal{N}_{\vec{0}}(\vec{p})| \geq d + 1$$

and $\mathcal{P}$ also contains a $(d+1)$-clique.

Proof. The proof follows from the discussion above and Corollary 9.7. $\quad\square$

We will also state one more optimality theorem in light of the structure of Lemma 7.15 (Adjacent or Far Lemma (Points)).

Definition 9.10. *If $d \in \mathbb{N}$ and $\mathcal{P}$ is a partition of $\mathbb{R}^d$, then we say that $\mathcal{P}$ has the adjacent-or-far property if there exists $\epsilon \in (0, \infty)$ such that for all $X, Y \in \mathcal{P}$ it is either the case that $X \overset{\text{adj}}{\sim} Y$ or it is the case that for any $\vec{x} \in X$ and $\vec{y} \in Y$ that $d_{max}(\vec{x}, \vec{y}) > \epsilon$. Any such ϵ is called an adjacent-or-far constant for $\mathcal{P}$.*

The following result strengthens the both the hypothesis and conclusion of the Second Optimality Theorem but the conclusion is slightly weaker than that of the Third Optimality Theorem.

Proposition 9.11 (Adjacent or Far Optimality Theorem)**.** *If $d \in \mathbb{N}$, and $\mathcal{P}$ is partition of $\mathbb{R}^d$ with the adjacent-or-far property, and there exists $D \in (0, \infty)$ such that for all $X \in \mathcal{P}$, $\mathrm{diam}(X) < D$, then $\mathcal{P}$ contains a $(d+1)$-clique*

Proof Sketch. Use the same proof technique as the Second Optimality Theorem to restrict attention to $[0, D]^d$, induce a partition on it from $\mathcal{P}$ (which is non-spanning), and consider an admitted Sperner/KKM coloring to find a point $\vec{p}$ at the closure of $d + 1$ colors. Taking ϵ to be an adjacent-or-far constant for $\mathcal{P}$, then $\overline{B}_{\epsilon/2}(\vec{p})$ intersects at least $d + 1$ colors and thus at least $d + 1$ members of $\mathcal{P}$. By the triangle inequality with the point $\vec{p}$, for any pair of these members X, Y, there is a point $\vec{x} \in X$ and $\vec{y} \in Y$ such that $d_{max}(\vec{x}, \vec{y}) \leq \epsilon$, so by the definition of the adjacent or far property, $X \overset{\text{adj}}{\sim} Y$. Thus these $d+1$ members constitute a $(d + 1)$-clique. $\square$

9.2 Tools for Proving the Optimality Theorems

In the broader literature, when something like a Sperner coloring of a simplex is made continuous on the simplex (i.e. instead of coloring the vertices of some simplicial subdivision of the simplex, the entire simplex is colored) it is essentially a KKM covering. The Knaster-Kuratowski-Mazurkiewicz (KKM) lemma states that if a $(d+1)$-dimensional simplex is covered by a family $\{C_1, \ldots, C_{d+1}\}$ of $d+1$ "reasonable" closed sets, then there is a point belonging to the intersection of all these closed sets $\bigcap_{i \in [d]} C_i$. The "reasonableness" is stated in terms analogous to the conditions on a Sperner coloring of a $(d + 1)$-simplex. It is well known that the KKM lemma follows easily from Sperner's lemma using a sequence of subdivisions of finer and finer resolution and using the Bolzano–Weierstrass theorem (c.f. [43]), and it is known that the KKM lemma can also be used to prove Sperner's lemma (c.f. [44]). Thus, if we assign to every point within the simplex one of $d + 1$ colors in a "reasonable" way and take C_i to be the *closure* of the points of the ith color, then the KKM lemma guarantees the existence of some point in the intersection of all C_i (which is a point belonging to the closure of all of the colors).

A number of variations on Sperner's lemma are known, and as with the standard case, they can be used to easily prove variations of the KMM lemma. Of particular interest to us will be the hypercube versions of Sperner's lemma or the KKM lemma in which each vertex/corner of the cube is given a unique color. The hypercube version of Sperner's lemma was first proved by Kuhn [43], and generalized even further to convex polytopes by De Loera, Peterson, and Su [45]. Kuhn did not explicitly show the KKM/continuous variant for the cube, but it is easily proved using essentially the same arguments as in the case of the simplex, and De Loera *et. al.* did explicitly mention connections to the KKM lemma. There were also convex polotope variants of the KKM lemma proven directly by van der Laan *et. al.* [46] without proving them via Sperner's lemma.

The next proposition follows almost immediately from the results alluded to in the prior paragraph (in particular, [45, Corollary 3], [43, § 2-3], or [46].

Definition 9.12 (Sperner/KMM Coloring)**.** *Let $d \in \mathbb{N}$ and $V = \{0, 1\}^d$ denote a set of colors (which is exactly the set of vertices of $[0, 1]^d$ so that colors and vertices are identified). Let $\chi : [0, 1]^d \to V$ be a coloring function such that for any face F of $[0, 1]^d$, for any $\vec{x} \in F$, it holds that $\chi(\vec{x}) \in F$ (informally, the color of $\vec{x}$ is one of the vertices in the face F). Such a function χ will be called a Sperner/KMM coloring.*

Proposition 9.13 (Cubical Sperner/KMM lemma)**.** *Let $d \in \mathbb{N}$ and $V = \{0,1\}^d$ and $\chi : [0,1]^d \to V$ be a Sperner/KMM coloring. Then there exists a subset $J \subset V$ with $|J| = d+1$ and a point $y \in [0,1]^d$ such that for all $j \in J$, $y \in \overline{\chi^{-1}(j)}$ (informally, $\vec{y}$ is in the closure of at least $d+1$ different colors).*

This claim really does follow almost immediately from [45, Corollary 3], but we framed our proposition with notation and wording that is sufficiently different that we will offer a proof. We emphasize, though, that this proof amounts to nothing more than carefully working through this change of notation.

Proof. For each vertex/color $\vec{v} \in V$, let $C^{\vec{v}} = \overline{\chi^{-1}(\vec{v})}$ denote the closure of the points assigned the color $\vec{v}$. Let F be some face of $[0,1]^d$ and we will show that F is covered by $\bigcup C^{\vec{v}} : \vec{v} \in F \cap V$ (that is, F is covered by the collection $C^{\vec{v}}$ for each vertex $\vec{v} \in F$). To see this, let $\vec{x} \in F$ be arbitrary. Since the codomain of χ is V, we have trivially that $\chi(\vec{x}) \in V$. Let $\vec{v} = \chi(\vec{x})$ denote this value to emphasize that this is a vertex of $[0,1]^d$. Also, $\vec{v} = \chi(\vec{x}) \in F$ by hypothesis, and thus we have a particular $\vec{v} \in F \cap V$ such that

$$C^{\vec{v}} = \chi^{-1}(\vec{v}) = \chi^{-1}(\chi(\vec{x})) \ni \vec{x}$$

which shows F is covered as claimed.

This gives the hypothesis of [45, Corollary 3], so taking $\vec{p} \in [0,1]^d$ arbitrarily, there must exist $J \subset V$ with $|J| = d+1$ such that $\bigcap_{j \in J} C^j \neq \emptyset$ which proves our claim. $\qquad\square$

For the next few result, we will use standard projection maps $\pi_i : [0,1]^d \to [0,1]$ defined by $\pi_i(\vec{x}) \overset{\text{def}}{=} x_i$ which maps points to the ith coordinate value, and we will also apply this map to sets and mean that $\pi_i(X) = \{\pi_i(\vec{x}) : \vec{x} \in X\}$.

Definition 9.14 (Non-Spanning)**.** *Let $d \in \mathbb{R}^+$ and $\mathcal{S}$ be a partition of $[0,1]^d$. We say that $\mathcal{S}$ is a non-spanning partition if it holds for all $X \in \mathcal{S}$ and for all $i \in [d]$ that either $\pi_i(X) \not\ni 0$ or $\pi_i(X) \not\ni 1$ (or both).*

Informally, a non-spanning partition of a hypercube does not contain any members which span the hypercube so as to intersect an opposite pair of opposite facets $((d-1)$-dimensional faces). We will now show that such partitions admit a Sperner/KMM coloring which respects the structure of the partition (i.e. any two points in the same partition member are assigned the same color).

Lemma 9.15 (Coloring Admission)**.** *Let $d \in \mathbb{N}$, and $V = \{0,1\}^d$, and $\mathcal{S}$ a non-spanning partition of $[0,1]^d$. Then there exists a Sperner/KMM coloring $\chi : [0,1]^d \to V$ and a function $\chi_{\mathcal{S}} : \mathcal{S} \to V$ (which we will call a partition coloring) such that for all $\vec{x} \in \mathbb{R}^d$, $\chi(\vec{x}) = \chi_{\mathcal{S}}(\text{member}(\vec{x}))$.*

Proof. We begin by defining a coloring function $\chi_{\mathcal{S}}$ which colors members of the partition, and then define the coloring function for points which just uses the color of the containing member. Specifically, define $\chi_{\mathcal{S}} : \mathcal{S} \to V$ by

$$\chi_{\mathcal{S}}(X) \overset{\text{def}}{=} \langle c_i \rangle_{i=1}^d \quad \text{where } c_i = \begin{cases} 0 & 0 \in \pi_i(X) \\ 1 & 1 \in \pi_i(X) \\ 0 & \text{otherwise} \end{cases}$$

Observe that the cases are well-defined because $\mathcal{S}$ is non-spanning. Then define $\chi : [0,1]^d \to V$ by $\chi(\vec{x}) \overset{\text{def}}{=} \chi_{\mathcal{S}}(\text{member}(\vec{x}))$. We will now show that this has the required properties of a Sperner/KMM coloring.

Let F be any face of $[0,1]^d$. Thus, F can be expressed as $F = \prod_{i=1}^d F_i$ where each F_i is one of three sets: $\{0\}$, $\{1\}$, or $[0,1]$. Let $\vec{x} \in F$ be arbitrary and let $\vec{c} = \chi(\vec{x})$; we must show that $\vec{c} \in F$ which we do by showing that $c_i \in F_i$ for each $i \in [d]$. There are three cases. If $F_i = \{0\}$, then $x_i = 0$ (since $\vec{x} \in F$), so $0 = x_i = \pi_i(\vec{x}) \in \pi_i(\text{member}(\vec{x}))$, in which case $c_i = 0$ by definition. The case $F_i = \{1\}$ is analogous. The remaining case is that $F_i = [0,1]$, but in this case $c_i \in F_i$ trivially since c_i is either 0 or 1. $\qquad\square$

Remark 9.16. *Observe that in the previous few results, there is nothing special about the hypercube $[0,1]^d$; it was chosen only for convenience. It is trivial to extend the result to any rectangle $A = \prod A_i$ (where $A_i = [l_i, r_i]$ for some values l_i and r_i). This can either be done by rewriting the proofs (since these sets are convex polytopes) or by using the natural scaling map (which is continuous and so preserves the intersections of closures). We will freely use the results above for other hypercubes.*

9.3 Proofs of the Optimality Theorems

We will now prove all three optimality theorems. All three proofs will have the same flavor, though it gets somewhat lost in notation. Essentially, we consider some sufficiently sized hypercube in the space, use the partition $\mathcal{P}$ to induce a partition $\mathcal{S}$ on the hypercube, argue that the induced partition is non-spanning (thus admitting a Sperner/KMM coloring), use the coloring to find a point $\vec{p}$ in the closure of $d+1$ colors, and then argue that $\vec{p}$ has the desired property when we transition from color classes back to partition members. For convenience, we restate each theorem before its proof.

The proof of the Third Optimality Theorem requires the fewest technical details, so we begin with its proof.

Theorem 9.5 (Third Optimality Theorem). *If $d \in \mathbb{N}$, and $\mathcal{P}$ is a partition of $\mathbb{R}^d$, and there exists $D \in (0, \infty)$ such that for all $X \in \mathcal{P}$, D is a strict pairwise bound for X (it is sufficient but not necessary that $\operatorname{diam}(X) < D$), and if there exists some $\vec{\alpha} \in \mathbb{R}^d$ such that $\vec{\alpha} + [0, D]^d$ intersects finitely many members of $\mathcal{P}$, then there exists $\vec{p} \in \mathbb{R}^d$ such that*

$$|\mathcal{N}_{\vec{0}}(\vec{p})| \geq d + 1.$$

Furthermore, $\mathcal{P}$ contains a $(d+1)$-clique.

Proof. Let D, and $\vec{\alpha}$ as in the theorem statement. Without loss of generality, we may assume that $\vec{\alpha} = \vec{0}$, that $D = 1$, and that for all $X \in \mathcal{P}$, $\operatorname{diam}(X) < 1$ by applying a continuous transformation[18]. Observe that $\mathcal{P}$ induces a finite partition $\mathcal{S}$ on $[0,1]^d$ (specifically, $\mathcal{S} = \{X \cap [0,1]^d : X \in \mathcal{P} \text{ and } X \cap [0,1]^d\} \neq \emptyset$). Further, $\mathcal{S}$ is a non-spanning partition; this is because some $Y \in \mathcal{S}$ had the property for some $i \in [d]$ that $\pi_i(Y) \ni 0$ and $\pi_i(Y) \ni 1$, then let $\vec{y}, \vec{z} \in Y$ such that $y_i = \pi_i(\vec{y} = 0)$ and $z_i = \pi_i(\vec{z} = 1)$ so $\operatorname{diam}(Y) \geq d_{max}(\vec{z}, \vec{y}) \geq |z_i - y_i| = 1$ which would be a contradiction. Thus, by Lemma 9.15, letting $V = \{0,1\}^d$, there is a Sperner/KMM coloring $\chi : [0,1]^d \to V$ and partition coloring $\chi_{\mathcal{S}} : \mathcal{S} \to [0,1]^d$.

For each $j \in V$, let $\mathcal{S}_j = \{Y \in \mathcal{S} : \chi_{\mathcal{S}}(Y) = j\}$ (i.e. the members of the partition mapped to color j) noting that $\chi^{-1}(j) = \bigcup_{Y \in \mathcal{S}_j} Y$. By Proposition 9.13 there is some $J \subseteq V$ with $|J| \geq d + 1$ and some $\vec{p} \in [0,1]^d$ with $\vec{p} \in \bigcap_{j \in J} \overline{\chi^{-1}(j)}$. This gives the following:

$$\vec{p} \in \bigcap_{j \in J} \overline{\chi^{-1}(j)}$$

$$\in \bigcap_{j \in J} \overline{\bigcup_{Y \in \mathcal{S}_j} Y}$$

$$\in \bigcap_{j \in J} \bigcup_{Y \in \mathcal{S}_j} \overline{Y} \qquad \text{(Because $\mathcal{S}$ is a finite set)}$$

---

[18]In further detail, let $H = \vec{\alpha} + [0, D^+]^d$ and $\phi : \mathbb{R}^d \to \mathbb{R}^d$ the continuous bijection defined by $\phi(\vec{x}) = \frac{\vec{x} - \vec{\alpha}}{D}$. The map ϕ induces a partition $\mathcal{P}'$ of $\mathbb{R}^d$ (specifically $\mathcal{P}' = \{\phi(X) : X \in \mathcal{P}\}$). Importantly, the claim that there exists $\vec{p} \in \mathbb{R}^d$ such that $|\{X \in \mathcal{P} : \overline{X} \ni \vec{p}\}| \geq d + 1$ is equivalent to the claim that there exists $\vec{p'} \in \mathbb{R}^d$ such that $|\{Y \in \mathcal{P}' : \overline{Y} \ni \vec{p'}\}| \geq d + 1$; this is because ϕ is continuous, so for any $\vec{x} \in \mathbb{R}^d$ and any $X \in \mathcal{P}$, we have $\vec{x} \in \overline{X}$ iff $\phi(\vec{x}) \in \overline{\phi(X)}$. Thus we can work in $\mathcal{P}'$ instead. It is easily verified that $\phi(\vec{\alpha}) = \vec{0}$, and $\phi(H) = [0,1]^d$, and for any $X \in \mathcal{P}$ that $\operatorname{diam}(\phi(X)) \leq \frac{D}{D^+} < 1$

Thus, for each $j \in J$, $\vec{p}$ belongs $\overline{Y}$ for some $Y \in \mathcal{S}_j$; let Y_j denote one such element. Trivially, if $j \neq j'$, then $\mathcal{S}_j$ and $\mathcal{S}_{j'}$ contain no common members, so $Y_j \neq Y_{j'}$ and thus

$$\left|\left\{X \in \mathcal{P} : \overline{X} \ni \vec{p}\right\}\right| \geq \left|\left\{Y \in \mathcal{S} : \overline{Y} \ni \vec{p}\right\}\right| \geq \left|\{Y_j : j \in J\}\right| \geq d+1$$

$\square$

The finiteness condition above was critical. In essence, we found $\vec{p}$ to belong to the closures of $d+1$ color classes, and since each color class consisted of only finitely many members, in fact $\vec{p}$ had to belong to the closure of not just the union of the members, but the closure of a single member (or multiple members).

Theorem 9.4 (Second Optimality Theorem). *If $d \in \mathbb{N}$, and $\mathcal{P}$ is a partition of $\mathbb{R}^d$, and there exists $D \in (0, \infty)$ such that for all $X \in \mathcal{P}$, it holds that $\mathrm{diam}(X) < D$, then for any $\epsilon : \mathbb{R}^d \to (0, \infty)$, there exists $\vec{p} \in \mathbb{R}^d$ such that*
$$|\mathcal{N}_{\epsilon(\vec{p})}(\vec{p})| \geq d+1.$$

Proof. Let $D \in (0, \infty)$, $\epsilon : \mathbb{R}^d \to (0, \infty)$ as in the statement. As in the previous proof[19], we may assume without loss of generality that $D = 1$ so that all members of $\mathcal{P}$ have diameter less than 1, and we don't have to worry about scaling ϵ since it is arbitrary to begin with. As in the previous proof, $\mathcal{P}$ induces a non-spanning partition $\mathcal{S}$ on $[0,1]^d$ which admits Sperner/KMM coloring χ, so letting $V = \{0,1\}^d$, there is some $J \subset V$ with $|J| \geq d+1$ and some $\vec{p} \in \bigcap_{j \in J} \overline{\chi^{-1}(j)}$. Thus any open set containing $\vec{p}$ will intersect members of at least $d+1$ colors and so intersects at least $d+1$ members. Thus, in particular, $B^\circ_{\epsilon(\vec{p})}(\vec{p})$ intersects at least $d+1$ members of $\mathcal{S}$ and thus at least $d+1$ members of $\mathcal{P}$. $\square$

The proof of the First Optimality Theorem is the most complicated, and we recommend looking at the proof outline for the First Optimality Theorem given earlier to understand the overall structure of the proof.

Theorem 9.3 (First Optimality Theorem). *If $d \in \mathbb{N}$, and $\mathcal{P}$ is a partition of $\mathbb{R}^d$, and there exists $M \in (0, \infty)$ such that for all $X \in \mathcal{P}$, X is Lebesgue measurable, and $m(X) < M$, then for any $\epsilon \in (0, \infty)$, there exists $\vec{p} \in \mathbb{R}^d$ such that*
$$|\mathcal{N}_\epsilon(\vec{p})| \geq d+1.$$

Proof. The beginning of the proof will sound a bit strangely worded. This is because we want some bound throughout the proof of how many members of the partition an ϵ-ball can intersect with; the natural approach is thus proof by contradiction, but we can get a stronger result from the proof than stated in the theorem if we are careful not to arrive at a contradiction in the end.

Let $d \in \mathbb{N}$, and let $\mathcal{P}$ be a partition of $\mathbb{R}^d$ with Lebesgue measurable sets, and let $M \in (0, \infty)$ such that for all $X \in \mathcal{P}$, $m(X) < M$. Let $k \geq d+1$ be fixed throughout the proof. If there exists $\vec{p} \in \mathbb{R}^d$ such that $|\mathcal{N}_\epsilon(\vec{p})| > k$ then we are done. Otherwise, we may assume that for all $\vec{p} \in \mathbb{R}^d$ that $|\mathcal{N}_\epsilon(\vec{p})| \leq k$. (Note that since $k \geq d+1$, there *are* partitions such that for all $\vec{p} \in \mathbb{R}^d$ that $|\mathcal{N}_\epsilon(\vec{p})| \leq k$—the reclusive partitions for example—and thus we have *not* stated a contradiction assumption.)

Let $V = \epsilon \cdot \mathbb{Z}^d$ (representing vertices characterizing a scaled grid). For any $\vec{\alpha} \in V$, let $H(\vec{\alpha}) = \vec{\alpha} + [-\frac{\epsilon}{2}, \frac{\epsilon}{2})^d$ (standing for half open hypercube). Note that $\overline{H(\vec{\alpha})} = \overline{B}_{\epsilon/2}(\vec{\alpha})$. Now observe that $\{H(\vec{\alpha}) : \vec{\alpha} \in V\}$ is a partition of $\mathbb{R}^d$, and in this way we view the vertex set V as the centers of cells partitioning the space, and informally, we will identify the two for discussion.

Now we want to argue that each cell (each $H(\vec{\alpha})$) intersects some member of the partition with high volume/measure. For any $\vec{\alpha} \in V$ we have $H(\vec{\alpha}) \subseteq \overline{H(\vec{\alpha})} = \overline{B}_{\epsilon/2}(\vec{\alpha}) \subseteq \overline{B}_\epsilon(\vec{\alpha})$ and thus $\{X \in \mathcal{P} : X \cap H(\vec{\alpha}) \neq \emptyset\} \subseteq$

---

[19]We may use any $\vec{\alpha}$.

36

$\{X \in \mathcal{P} : X \cap \overline{B}_\epsilon(\vec{\alpha}) \neq \emptyset\}$ where the latter has cardinality at most k by assumption implying that the former has cardinality at most k. Then we have

$$\epsilon^d = m(H(\vec{\alpha})) = \sum_{X \in \mathcal{P}} m(X \cap H(\vec{\alpha})) = \sum_{X \in \mathcal{P}: X \cap H(\vec{\alpha}) \neq \emptyset} m(X \cap H(\vec{\alpha})).$$

Since the index set has size at most k, by a simple averaging argument, there is some $X \in \mathcal{P}$ such that $m(X \cap H(\vec{\alpha})) \geq \frac{\epsilon^d}{k}$. Thus, there exists some function $P : V \to \mathcal{P}$ such that for all $\vec{\alpha} \in V$, $m(P(\vec{\alpha}) \cap H(\vec{\alpha})) \geq \frac{\epsilon^d}{k}$. We view P as providing a labeling of each cell with a member of $\mathcal{P}$ that is sufficiently similar (hence the name P); this function will be the key to approximating $\mathcal{P}$ with cells.

Next, we provide a bound on how many cells can be mapped to a particular member of $\mathcal{P}$. For any $X \in \mathcal{P}$ we have

$$M > m(X) \qquad \text{(By hypothesis)}$$

$$= m\left(\bigsqcup_{\vec{\alpha} \in V} X \cap H(\vec{\alpha})\right) \qquad \text{(Decomposition of } \mathbb{R}^d\text{)}$$

$$\geq m\left(\bigsqcup_{\vec{\alpha} \in P^{-1}(X)} X \cap H(\vec{\alpha})\right) \qquad \text{(Subsets have equal or smaller measure)}$$

$$\geq \sum_{\vec{\alpha} \in P^{-1}(X)} m(X \cap H(\vec{\alpha})) \qquad \text{(Fact B.1)}$$

$$\geq \sum_{\vec{\alpha} \in P^{-1}(X)} \frac{\epsilon^d}{k} \qquad \text{(Def'n of } P \text{ and } X = P(\vec{\alpha}) \text{ for } \vec{\alpha} \in P^{-1}(X)\text{)}$$

$$= |P^{-1}(X)|\frac{\epsilon^d}{k}$$

which shows that $|P^{-1}(X)| < \frac{kM}{\epsilon^d}$.

Next, we will define three binary relations on the set of vertices in order to arrive at a useful equivalence relation that will let us approximate $\mathcal{P}$ closely enough by using the cells (see Appendix C for details on binary relations.). As with the first two Optimality Theorems, we want to control the diameter in the approximation and also utilize the labeling of P, so we will end up saying cells are equivalent (will be part of the same member of the approximating partition) if they have the same P label and are also close/connected. With this goal, let

$$R_P = \left\{(\vec{\alpha}, \vec{\beta}) \in V^2 : P(\vec{\alpha}) = P(\vec{\beta})\right\}$$

$$R_\epsilon = \left\{(\vec{\alpha}, \vec{\beta}) \in V^2 : d_{max}(\vec{\alpha}, \vec{\beta}) \leq \epsilon\right\}.$$

Note that both relations are reflexive and symmetric, and that R_P is also transitive (so R_P is an equivalence relation). Let R_ϵ^t denote the transitive closure of R_ϵ (so R_ϵ^t is an equivalence relation), and define $R = R_P \cap R_\epsilon^t$ which is also an equivalence relation. Observe that the equivalence classes of R_P are exactly the sets $P^{-1}(X)$ for $X \in \text{range}(P)$, and each equivalence class of R is a subset of an equivalence class of R_P. By these two observations and the previous paragraph, each equivalence class of R is finite and has cardinality (strictly) less than $\frac{kM}{\epsilon^d}$.

We can now show a type of bound on "diameter". Let $(\vec{\alpha}, \vec{\beta}) \in R$ be arbitrary. Let $R_0 = R_P \cap R_\epsilon$ and observe that R can be equivalently expressed as the transitive closure of R_0, so that there must be a sequence $\langle \vec{x}^{(j)} \rangle_{k=0}^N$ for some $N \in \mathbb{N}$ with $\vec{x}^{(0)} = \vec{\alpha}$ and $\vec{x}^{(N)} = \vec{\beta}$ and for all $j \in [N]$ that $(\vec{x}^{(j-1)}, \vec{x}^{(j)}) \in R_0 \subseteq R_\epsilon$.

Because of the bound on the cardinality of each equivalence class of R we may assume $N < \frac{kM}{\epsilon^d} - 1$ (the -1 is because of the zero-based indexing). Thus, we have

$$d_{max}(\vec{\alpha}, \vec{\beta}) \leq \sum_{j=1}^{N} d_{max}(\vec{x}^{(k-1)}, \vec{x}^{(j)}) \qquad \text{(Triangle inequality)}$$

$$\leq \sum_{j=1}^{N} \epsilon \qquad ((\vec{x}^{(k-1)}, \vec{x}^{(j)}) \in R_\epsilon)$$

$$\leq N\epsilon$$

$$< \left(\frac{kM}{\epsilon^d} - 1 \right) \epsilon \qquad \text{(Constraint on } N)$$

We are now in a position to define the approximation partition. As is common notation, let $V\!/_R$ denote the family of equivalence classes of R (note that for each $C \in V\!/_R$ we have $C \subseteq V$). Then define the approximation partition as $\mathcal{A} = \left\{ \bigsqcup_{\vec{\alpha} \in C} H(\vec{\alpha}) : C \in V\!/_R \right\}$ which is a partition of R^d. For any $X \in \mathcal{A}$, there is an unique equivalence class $C \in V\!/_R$ such that $X = \bigsqcup_{\vec{\alpha} \in C} H(\vec{\alpha})$, and we denote this class as C_X. We will extend the distance argument above the members of $\mathcal{P}$. Let $X \in \mathcal{A}$ be arbitrary and let $\vec{a}, \vec{b} \in X$ be arbitrary. Then there must exist some $\vec{\alpha} \in C_X$ with $\vec{a} \in H(\vec{\alpha})$ and similarly, there must exist some $\vec{\beta} \in C_X$ (possibly the same as $\vec{\alpha}$) with $\vec{b} \in H(\vec{\beta})$.

$$d_{max}(\vec{a}, \vec{b}) \leq d_{max}(\vec{a}, \vec{\alpha}) + d_{max}(\vec{\alpha}, \vec{\beta}) + d_{max}(\vec{\beta}, \vec{b}) < \frac{\epsilon}{2} + \left(\frac{dM}{\epsilon^d} - 1 \right) \epsilon + \frac{\epsilon}{2} = \frac{dM}{\epsilon^{d-1}}.$$

Thus, (noting the strict inequalities above), $\mathcal{A}$ satisfies the hypothesis of the Third Optimality Theorem (with $D = \frac{kM}{\epsilon^{d-1}}$ and $\vec{\alpha} = \vec{0}$), so there is some point $\vec{p} \in \mathbb{R}^d$ such that $|\{X \in \mathcal{A} : \overline{X} \ni \vec{p}\}| \geq d+1$. We let $\vec{p}$ denote such a point for the remainder of the proof.

Let $\mathcal{N} = \{X \in \mathcal{A} : \overline{X} \ni \vec{p}\}$ denote this set and observe that for any $X \in \mathcal{N}$ we have

$$\vec{p} \in \overline{X} = \overline{\bigcup_{\vec{\alpha} \in C_X} H(\vec{\alpha})} = \bigcup_{\vec{\alpha} \in C_X} \overline{H(\vec{\alpha})}$$

where the second equality is because the set C_X is finite. Thus, there must be some $\vec{\alpha} \in C_X$ such that $\vec{p} \in \overline{H(\vec{\alpha})}$; fix such an $\vec{\alpha}$ for X and denote it $\vec{\alpha}_X$. We will now show that for each $X, Y \in \mathcal{N}$, that $P(\vec{\alpha}_X) \neq P(\vec{\alpha}_Y)$ and that $\overline{B}_\epsilon(\vec{p}) \cap P(\vec{\alpha}_X) \neq \emptyset$. This will be enough to conclude the proof because this gives an injection from $\mathcal{N}$ to $\mathcal{N}_\epsilon(\vec{p}) = \{X \in \mathcal{P} : X \cap \overline{B}_\epsilon(\vec{p})\}$ proving that it has cardinality at least $d+1$.

Let $X, Y \in \mathcal{N}$ be arbitrary with $X \neq Y$. Then $C_X \neq C_Y$. We will show that $P(\vec{\alpha}_X) \neq P(\vec{\alpha}_Y)$ by showing that $(\vec{\alpha}_X, \vec{\alpha}_Y) \notin R_L$ which we do by showing that $(\vec{\alpha}_X, \vec{\alpha}_Y) \notin R$ and $(\vec{\alpha}_X, \vec{\alpha}_Y) \in R_\epsilon^t$. It is immediate that $(\vec{\alpha}_X, \vec{\alpha}_Y) \notin R$ because $\vec{\alpha}_X \in C_X$ and $\vec{\alpha}_Y \in C_Y$ which are different equivalence classes of R. Regarding R_ϵ^t, we have (by the definition of $\vec{\alpha}_X$ and $\vec{\alpha}_Y$) that

$$\overline{B}_{\epsilon/2}(\vec{\alpha}_X) = \overline{H(\vec{\alpha}_X)} \ni \vec{p} \in \overline{H(\vec{\alpha}_Y)} = \overline{B}_{\epsilon/2}(\vec{\alpha}_Y)$$

so by the triangle inequality we have $d_{max}(\vec{\alpha}_X, \vec{\alpha}_Y) \leq \epsilon$ which means $(\vec{\alpha}_X, \vec{\alpha}_Y) \in R_\epsilon \subseteq R_\epsilon^t$. Thus we have established that $P(\vec{\alpha}_X) \neq P(\vec{\alpha}_Y)$.

The last thing we need is to show that for any $X \in \mathcal{N}$ we have $\overline{B}_\epsilon(\vec{p})$ intersecting $P(\vec{\alpha}_X)$. Since $\vec{p} \in \overline{H(\vec{\alpha}_X)} = \overline{B}_{\epsilon/2}(\vec{\alpha}_X)$, it follows that $\overline{B}_{\epsilon/2}(\vec{\alpha}_X) \subseteq \overline{B}_\epsilon(\vec{p})$ (i.e. all points within $\epsilon/2$ of $\vec{\alpha}_X$ are within ϵ of $\vec{p}$ by the triangle inequality because $\vec{\alpha}_X$ is within $\epsilon/2$ of $\vec{p}$). Thus $H(\vec{\alpha}_X) \subseteq \overline{B}_\epsilon(\vec{p})$. We will intersect both

sides of this containment with $P(\vec{\alpha}_X)$ recalling that by the definition of P, $P(\vec{\alpha}_X)$ is a member of $\mathcal{P}$ such that $m(P(\vec{\alpha}_X) \cap H(\vec{\alpha}_X)) \geq \frac{\epsilon^d}{k}$, so in particular, this intersection is not empty. Thus

$$\emptyset \neq P(\vec{\alpha}_X) \cap H(\vec{\alpha}_X) \subseteq P(\vec{\alpha}_X) \cap \overline{B}_\epsilon(\vec{p}).$$

In words, the ϵ ball around $\vec{p}$ intersects the member $P(\vec{\alpha}_X)$ of $\mathcal{P}$. This completes the proof as we have shown that $|\{X \in \mathcal{P} : X \cap \overline{B}_\epsilon(\vec{p})\}| \geq d + 1$. $\qquad \square$

Remark 9.17. *In the above proof, we have actually shown something stronger than the statement of the First Optimality Theorem. Not only are we guaranteed the existence of a point $\vec{p}$ where the ϵ ball intersects at least $d + 1$ members of $\mathcal{P}$, but if we know some bound k for the partition, then we are guaranteed a point where the intersection with at least $d + 1$ members has substantial measure (at least $\frac{\epsilon^d}{k}$). This result also applies to the Second and Third Optimality Theorems (if the partitions are restricted to measurable members) because those theorems have otherwise stronger assumptions in the hypothesis.*

In this section we have shown that in our motivating question (Question 2.1), the value $k = d + 1$ is optimal not just for hypercube partitions, but also for any partition which has a uniform upper bound on the measures (or diameters) of the members (the First and Second Optimality Theorems). Further, if we strengthen this hypothesis to require that the diameters be uniformly upper bounded and require some finiteness, then we can actually conclude that there is a single point at the closure of $d + 1$ members of the partition, which also implies that the partition has a $(d + 1)$-clique (the Third Optimality Theorem). However, now that we know the optimal value of k, we can, in a sense, get this same strong conclusion even under the weaker assumptions if we insist that our partitions have the property of the motivating question for the optimal value of $k = d + 1$.

Proposition 9.18. *If $d \in \mathbb{N}$, and $\mathcal{P}$ is a partition of $\mathbb{R}^d$, and there exists $D \in (0, \infty)$ such that for all $X \in \mathcal{P}$, it holds that D is a strict pairwise bound for X, and if there exists $\epsilon \in (0, \infty)$ such that $\mathcal{P}$ has the property that for all $\vec{q} \in \mathbb{R}^d$,*

$$|\mathcal{N}_\epsilon(\vec{q})| \leq d + 1$$

then there exists $\vec{p} \in \mathbb{R}^d$ such that

$$|\mathcal{N}_\epsilon(\vec{p})| = d + 1.$$

Furthermore, $\mathcal{P}$ contains a $(d + 1)$-clique.

Proof. Using the same techniques as in the proof of the Second Optimality theorem, pick $\vec{\alpha} \in \mathbb{R}^d$ and consider the hypercube $H = \vec{\alpha} + [0, D]^d$ and the non-spanning partition $\mathcal{S}$ of it induced by $\mathcal{P}$, and an admitted Sperner/KKM coloring so that there is some $\vec{p} \in H$ at the closure of at least $d + 1$ colors. On the other hand, $\overline{B}_\epsilon(\vec{p})$ intersects at most $d + 1$ members by assumption, and thus it intersects exactly one member of each of these $d + 1$ colors. Further, for any $\delta \in (0, \epsilon]$, the ball $\overline{B}_\delta(\vec{p})$ will intersect exactly these same $d + 1$ members, so that $\vec{p}$ belongs to the closure of these $d + 1$ members (and it follows as before that these $d + 1$ members form a clique). $\qquad \square$

The significance of this result is that it makes a nice connection between the neighborhood property of the motivating question and the clique property of the partition graph.

9.4 Optimality Theorem Gaps

We now present two examples that demonstrate that there really is a "gap" between the three different Optimality Theorems. The conclusions get stronger with each successive version of the theorem, but each time the hypotheses were also made stronger (with the exception that the second and third theorems remove the requirement that the members be measurable). This begs the question of whether all three versions are necessary—for example, could it be that the hypothesis of the First Optimality Theorem are

mathematically sufficient to imply the conclusions of the Third Optimality Theorem and we just did not find a proof? The answer is no; the strengthening of the hypotheses really is necessary and we prove this by providing two (counter)examples.

Proposition 9.19 (First and Second Optimality Gap)**.** *The hypothesis of the First Optimality Theorem does not imply the conclusion of the Second Optimality Theorem for any $d \in \mathbb{N}$ with $d \geq 2$. (And it does for $d = 1$.)*

Proof. We will construct a partition which satisfies the hypothesis of the First Optimality Theorem and which does not satisfy the conclusion of the Second Optimality Theorem.

Recall that with respect to the d_{max} metric, the ball of radius r centered at the origin is $\overline{B}_r(0) = [-r, r]^d$. This ball (which is a hypercube) has diameter $2r$ and measure/volume $(2r)^d$.

Consider the sequence $\{r_n\}_{n=1}^{\infty}$ where $r_n = \frac{1}{2}n^{1/d}$. Then $\overline{B}_{r_n}(0)$ has measure n. Let $S_1 = \overline{B}_{r_1}(0)$ and inductively for $n > 1$, let $S_n = \overline{B}_{r_n} \setminus \overline{B}_{r_{n-1}}$. All of the S_n are disjoint by construction, and because $\{r_n\}_{n=1}^{\infty}$ increases without bound, they form a partition of $\mathbb{R}^d$. Further, the measure of S_1 is 1, and for $n > 1$, the measure of S_n is $n - (n-1) = 1$, the difference in measures of the two balls.

Thus, every member of this partition has a measure of 1, and because the S_n are concentric, each with non-zero "width", for every point $\vec{p} \in \mathbb{R}^d$, there exists some $\epsilon(\vec{p})$ such that $\overline{B}_{\epsilon(\vec{p})}(\vec{p})$ intersects at most 2 members of the partition. This shows that for $d \geq 2$, the conclusion of the Second Optimality Theorem does not hold[20]. For clarity, this construction works with $d = 1$ as well, but if $d = 1$, then $d + 1 = 2$, so for each $\vec{p} \in \mathbb{R}^d$, $\overline{B}_{\epsilon(\vec{p})}(\vec{p})$ intersects at most $2 = d + 1$ members of the partition which does not contradict the conclusion of the Second Optimality Theorem.

In the case $d = 1$, suppose for contradiction that there is some partition $\mathcal{P}$ of $\mathbb{R}$ and some M such that for all $X \in \mathcal{P}$, $m(X) < M$, (so the hypothesis of the First Optimality Theorem is satisfied) and suppose that there is some $\epsilon : \mathbb{R} \to (0, \infty)$ such that for all $\vec{p} \in \mathbb{R}^d$ that $|\mathcal{N}_{\epsilon(\vec{p})}(\vec{p})| \leq 1$ (so that the conclusion of the Second Optimality Theorem is not satisfied). Fix some $X \in \mathcal{P}$ and observe that for any $\vec{x} \in X$, $\overline{B}_{\epsilon(\vec{x})}(\vec{x})$ clearly intersects X, and so does not intersect any other member of P by assumption (since $|\mathcal{N}_{\epsilon(\vec{p})}(\vec{p})| \leq 1$, so $\overline{B}_{\epsilon(\vec{x})}(\vec{x})$ intersects at most one member of the partition). Thus $\overline{B}_{\epsilon(\vec{x})}(\vec{x}) \subseteq X$, and so $B^{\circ}_{\epsilon(\vec{x})}(\vec{x}) \subseteq X$. Since $\vec{x} \in X$ was arbitrary, this shows that X is an open set. Since $m(X)$ is finite, $X \neq \mathbb{R}$ and so X is not closed (because $\mathbb{R}$ is a connected set). Fix some $\vec{p} \in \overline{X} \setminus X$. Then $\overline{B}_{\epsilon(\vec{p})}(\vec{p})$ intersects two members of $\mathcal{P}$—namely X and member$(\vec{p})$—which is a contradiction. Thus, the conclusion of the Second Optimality Theorem follows from the hypothesis of the First Optimality Theorem in the case $d = 1$. $\square$

Proposition 9.20 (Second and Third Optimality Gap)**.** *The hypothesis of the Second Optimality Theorem does not imply the conclusion of the Third Optimality Theorem for any $d \in \mathbb{N}$. (Also, the hypothesis of the First Optimality Theorem does not imply the conclusions of the Third Optimality Theorem for any $d \in \mathbb{N}$.)*

Proof. We will construct a partition which satisfies the hypotheses of both the First and Second Optimality Theorems and which does not satisfy the conclusion of the Third Optimality Theorem.

Consider the partition of singletons $\mathcal{P} = \{\{\vec{x}\} : \vec{x} \in \mathbb{R}^d\}$. All sets are measurable and have diameter 0 (and thus measure 0), and because each member is already closed, for any point $\vec{p} \in \mathbb{R}^d$, we have $\mathcal{N}_{\overline{0}}(\vec{p}) = \{X \in \mathcal{P} : \overline{X} \ni \vec{p}\} = \{X \in \mathcal{P} : X \ni \vec{p}\} = \{\vec{p}\}$ which has size $1 < d + 1$. Thus, the conclusion of the Third Optimality Theorem does not hold. $\square$

One might notice that this example demonstrating the gap between the Second and Third Optimality Theorems seems rather contrived for two reasons: (1) all members of the partition have measure 0, and (2) the partition consists of uncountably many members. If we consider a partition in which no member has measure 0, then this actually implies that the partition has countably many members (see the Appendix

---

[20]This also shows that the conclusion of the Third Optimality Theorem does not hold for $d \geq 2$ either since it is stronger, but we will prove this for all $d \in \mathbb{N}$ in the next proposition.

for details), so insisting on countable partitions is a weaker condition. We believe that countable partitions are a very natural restriction and we briefly explore this and make some connections to partitions of closed sets.

As we have shown, the hypothesis of the Second Optimality Theorem does not imply the conclusion of the Third Optimality Theorem, but the example that we gave demonstrating this seemed contrived as discussed above, and we believe that the finiteness condition that we added to that hypothesis in the Third Optimality Theorem may stronger than necessary—we wonder if replacing the finite requirement with a countable requirement in the Third Optimality Theorem is mathematically sufficient to justify the same conclusion.

Conjecture 9.21 (Stronger Third Optimality Theorem)**.** *If $d \in \mathbb{N}$, and $\mathcal{P}$ is a partition of $\mathbb{R}^d$, and there exists $D \in (0, \infty)$ such that for all $X \in \mathcal{P}$, D is a strict pairwise bound of X, and if there exists some $\vec{\alpha} \in \mathbb{R}^d$ such that $\vec{\alpha} + [0, D]^d$ intersects countably many members of $\mathcal{P}$,*
then there exists $\vec{p} \in \mathbb{R}^d$ such that

$$|\mathcal{N}_{\vec{0}}(\vec{p})| \geq d + 1.$$

We can prove that the case of $d = 1$ in the above conjecture is true; not only that, we can prove that it is equivalent to a known result about the inability to partition the unit interval with countably many closed sets. In the statement below, non-trivial means that the partition contains at least two members.

Theorem 9.22 ([35])**.** *There is no non-trivial partition of $[0, 1]$ by countably many closed sets.*

A proof of the above theorem utilizing the Baire Category Theorem can be found in [36] and seems to be originally attributed to [35]. The natural generalization of this theorem to higher dimensions is not interesting because it is an immediate consequence of the theorem that there is no countable non-trivial partition of $[0, 1]^d$ for any $d \in \mathbb{N}$. If there was such a partition $\mathcal{S}$, then pick two points $\vec{x}, \vec{y} \in [0, 1]^d$ belonging to different members of $\mathcal{S}$ and consider the convex hull (i.e. the line segment between them) which is isomorphic to $[0, 1]$. Then $\mathcal{S}$ would induce a non-trivial countable partition of $[0, 1]$ by closed sets giving the contradiction. However, we will prove now that the theorem above is equivalent to the case $d = 1$ of our conjecture. Thus our conjecture (if true) serves as an interesting and fairly natural generalization of the above theorem.

Remark 9.23. *We will be talking about the topology on $\mathbb{R}$ as well as the subspace topology on $[0, 1]$, but because $[0, 1]$ is closed, every subset of $[0, 1]$ is closed in the subspace topology if and only if it is closed in the topology on $\mathbb{R}$, so we need not distinguish between these and can say sets are closed without ambiguity.*

Proof of equivalence. First assume the theorem and we will prove the conjecture for $d = 1$. Let $\mathcal{P}$ be a partition of $\mathbb{R}^1$ with D, α as in the conjecture hypothesis. We may assume that $\alpha = 0$ and $D = 1$ without loss of generality[21]. Then $\mathcal{P}$ induces a partition $\mathcal{S}$ on $[0, 1]$ (specifically $\mathcal{S} = \{X \cap [0, 1] : X \in \mathcal{P}, \ X \cap [0, 1] \neq \emptyset\}$), and this partition is non-trivial because if $\mathcal{S}$ contains just one member then the member is $[0, 1]$ which implies there is some member $X \in \mathcal{P}$ with $X \supseteq [0, 1]$ which contradicts that 1 is a strict diameter for all members of $\mathcal{P}$). $\mathcal{S}$ is countable as it has cardinality at most that of $\mathcal{P}$. By the theorem, $\mathcal{S}$ contains some set Y which is not closed, so let $\vec{p} \in \overline{Y} \setminus Y \neq \emptyset$. Let $M = \text{member}_\mathcal{S}(\vec{p})$. Since $\vec{p} \in M$ and $\vec{p} \notin Y$, we have $M \neq Y$, so $M, Y \in \{X \in \mathcal{S} : \overline{X} \ni \vec{p}\}$ which shows this set has cardinality at least two. It follows that $\{X \in \mathcal{P} : \overline{X} \ni \vec{p}\}$ has cardinality at least 2 as well[22] which proves the conjecture for $d = 1$.

For the other direction, assume that the theorem is false and we will show that our conjecture is false with $d = 1$. Let $\mathcal{S}$ be a non-trivial countable partition of $[0, 1]$ by closed sets. We will modify $\mathcal{S}$ to construct $\mathcal{S}'$ which will have the distance requirements we need. If $\text{member}_\mathcal{S}(0) \neq \text{member}_\mathcal{S}(1)$, then let $\mathcal{S}' = S$.

---

[21]See the footnote in the proof of the Third Optimality Theorem.
[22]Take Y' to be the member of $\mathcal{P}$ such that $Y' \cap [0, 1] = Y$, and similarly for M'. Then $M' \neq Y'$ because otherwise $Y = Y' \cap [0, 1] = M' \cap [0, 1] = M$. Also, $\overline{M'} \supseteq M' \supseteq M \ni \vec{p}$, and $\overline{Y'} \supseteq \overline{Y} \ni \vec{p}$ so that $Y', M' \in \{X \in \mathcal{P} : \overline{X} \ni \vec{p}\}$

Otherwise, let $M = \text{member}_{\mathcal{S}}(0) = \text{member}_{\mathcal{S}}(1)$; we will split M into disjoint closed sets. Since $\mathcal{S}$ is non-trivial, it contains at least two members, and thus $[0,1] \setminus M$ is non-empty, and since $M \ni 0, 1$ we have that $(0,1) \setminus M$ is non-empty and is an open set (in both topologies). Let a be an arbitrary element of this open set, and let $\epsilon > 0$ such that $(a - \epsilon, a + \epsilon) \subseteq (0,1) \setminus M$. In other words, $[0, a - \epsilon] \sqcup [a + \epsilon, 1] \subseteq M$. Let $M_0 = M \cap [0, a - \epsilon]$ and $M_1 = M \cap [0, a + \epsilon]$, and let $\mathcal{S}' = \mathcal{S} \setminus M \cup \{M_0, M_1\}$ (i.e. remove the member M from $\mathcal{S}$ and add back two members M_0 and M_1). Since M_0 and M_1 are closed and disjoint and $M_0 \sqcup M_1 = M$ we have that $\mathcal{S}'$ is also a non-trivial countable partition of $[0,1]$ by closed sets, and has the additional property that for any $Y \in \mathcal{S}'$ and $x, y \in Y$, $d_{max}(x, y) < 1$ (i.e. all members of $\mathcal{S}'$ have 1 as a strict diameter). Now consider the partition $\mathcal{P}$ of $\mathbb{R}^1$ given by $\mathcal{P} = \mathcal{S}' \cup \{\{x\} : x \in \mathbb{R}^1 \setminus [0,1]\}$ (i.e. the partition using the same members as $\mathcal{S}'$ and using singletons elsewhere). This satisfies the hypothesis of the conjecture with $\alpha = 0$ and $D = 1$, but all members of $\mathcal{P}$ are closed sets so $\{X \in \mathcal{P} : \overline{X} \ni \vec{p}\} = \{X \in \mathcal{P} : X \ni \vec{p}\} = \{\text{member}_{\mathcal{P}}(\vec{p})\}$ which has cardinality 1, thus the conclusion of the conjecture would not hold. $\square$

10 Upper Bounds on the Tolerance Parameter (ϵ)

In a particular reclusive partition, it was possible to center around every point in $\mathbb{R}^d$ a closed ball of radius $\frac{1}{2d}$ (in the d_{max} metric) which would only intersect $d+1$ members of the partition. In the previous section, we argued that this is optimal if there is a bound on the size of the partition members, (and that otherwise there are trivial counterexamples). We view this value as the primary concern, and now that we know $d+1$ is optimal, we consider a secondary concern which is the value of ϵ in the motivating question.

A simple argument shows that if all elements of $\mathcal{P}$ have diameter at most D, then it must be that $\epsilon \leq \frac{D}{2}$.

Proposition 10.1 (Trivial Tolerance (ϵ) Bound). *If $d \in \mathbb{N}$, and $\mathcal{P}$ is a partition of $\mathbb{R}^d$, and there exists $D \in (0, \infty)$ such that for all $X \in \mathcal{P}$, $\text{diam}(X) \leq D$, and if $\epsilon \in (0, \infty)$ such that for all $\vec{p} \in \mathbb{R}^d$,*

$$|\mathcal{N}_\epsilon(\vec{p})| \leq 2^d$$

then $\epsilon \leq \frac{D}{2}$.

Proof. We show that for any $\epsilon > \frac{D}{2}$, there is some point $\vec{p} \in \mathbb{R}^d$ such that $\overline{B}_\epsilon(\vec{p})$ intersects $2^d + 1$ members of $\mathcal{P}$. Let $\epsilon > \frac{D}{2}$ and fix any $X \in \mathcal{P}$. Then for each $i \in [d]$, let $a_i = \inf\{x_i : \vec{x} \in X\}$ and let $b_i = \sup\{x_i : \vec{x} \in X\}$ noting that $b_i - a_i \leq D$. Thus $X \subseteq \prod_{i=1}^d [a_i, b_i] \subseteq \prod_{i=1}^d [a_i, a_i + D] = \vec{a} + [0, D]^d$. Let $\vec{p} = \langle a_i + \frac{D}{2} \rangle_{i=1}^d$, so $X \subseteq \overline{B}_{\frac{D}{2}}(\vec{p}) \subset \overline{B}_\epsilon(\vec{p})$. For any distinct $\vec{\alpha}, \vec{\beta} \in \{-\epsilon, \epsilon\}^d$, we have $\vec{\alpha} + \vec{p}, \vec{\beta} + \vec{p} \in \overline{B}_\epsilon(\vec{p})$, and $\vec{\alpha} + \vec{p}, \vec{\beta} + \vec{p} \notin X$, and $d_{max}(\vec{p} + \vec{\alpha}, \vec{p} + \vec{\beta}) = d_{max}(\vec{\alpha}, \vec{\beta}) = 2\epsilon > D$ which implies $\vec{p} + \vec{\alpha}$ and $\vec{p} + \vec{\beta}$ belong to different members of $\mathcal{P}$. Thus $|\{X \in \mathcal{P} : X \cap \overline{B}_\epsilon(\vec{p}) \neq \emptyset\}| \geq 2^d + 1$. $\square$

By considering the specific value $d = 1$, we get the following corollary.

Corollary 10.2 (Optimal Tolerance (ϵ) in $\mathbb{R}^1$). *If $d = 1$ and $\mathcal{P}$ is a partition of $\mathbb{R}^d = \mathbb{R}^1$, and there exists $D \in (0, \infty)$ such that for all $X \in \mathcal{P}$, D is a strict pairwise bound for X, and if $\epsilon \in (0, \infty)$ is such that for all $\vec{p} \in \mathbb{R}^d$,*

$$|\mathcal{N}_\epsilon(\vec{p})| \leq d + 1 = 2$$

then $\epsilon \leq \frac{D}{2} = \frac{D}{2d}$.

Proof. Since D is a strict pairwise bound for each member of $\mathcal{P}$, then every member of $\mathcal{P}$ has diameter at most D. Apply Proposition 10.1 with $d = 1$. $\square$

Note that the reclusive partitions can be easily scaled from unit hypercubes to D-sidelength hypercubes, in which case a value of $k = d + 1$ and $\epsilon = \frac{D}{2d}$ can be achieved. The above proposition shows that for $\mathbb{R}^1$

(when $d = 1$), partitions with members with strict pairwise bound of D (or diameter at most D) must have $\epsilon \leq \frac{D}{2} = \frac{D}{2d}$, and thus (certain) reclusive partitions attain the maximal value of ϵ when k is minimized at $d + 1$.

The argument above did not use any interesting properties of the partitions though. All we did was to say that if the radius ϵ is larger than $\frac{D}{2}$, then it is possible to strictly contain an entire member within a hypercube and allow the corners far enough apart so that they must each belong to a different member. In order to get better bounds than this, we will need some additional properties. What we will show in this section is that the statements of the Optimality Theorems were actually quite a bit weaker than they could have been. In the first two optimality theorems, we picked an aribtrary D-sidelength hypercube and showed the existence of a point $\vec{p}$ with the desired properties in that cube. Obviously we could have picked any cube, and so there are infinitely many points $\vec{p}$ as in those two theorems. Not only this, the results of [47] show that because the d-dimensional hypercube is a d-dimensional polytope with 2^d vertices, we can actually find (basically[23]) $2^d - d$ such points in each D-sidelength hypercube. Further, their proof actually gives an even stronger result that we can use to get a general upper bound of $\epsilon \leq \dfrac{1}{2\left((d+1)^{\frac{d-1}{2d}} - 1\right)}$, and prove

that the value of $\epsilon = \frac{D}{2d}$ is optimal for $d = 2$.

Before moving on, we show that the (quite ugly) upper bound we just mentioned is bounded above by and also asymptotically equivalent to the much cleaner bound $\frac{1}{2\sqrt{d}}$.

Lemma 10.3. *The function* $f(x) = \dfrac{1}{2\left((x+1)^{\frac{x-1}{2x}} - 1\right)}$ *is asymptotically equivalent to the function* $g(x) = \frac{1}{2\sqrt{x}}$

(i.e. $\lim_{x\to\infty} \frac{g(x)}{f(x)} = 1$*). Furthermore, for* $x > 1$, $f(x) < g(x)$.

Proof. Note first that
$$\lim_{x\to\infty} x^{1/x} = \lim_{x\to\infty} e^{\ln(x)/x} = e^0 = 1$$

and so
$$\lim_{x\to\infty} x^{(-1)/(2x)} = \lim_{x\to\infty} \frac{1}{\sqrt{x^{1/x}}} = 1.$$

We will apply the squeeze theorem to the following inequalities which hold for $x > 0$.

$$1 < \left(\frac{x}{x+1}\right)^{1/2} = \frac{x^{1/2}}{(x+1)^{1/2}}$$
$$\leq \frac{x^{1/2}}{(x+1)^{(x-1)/(2x)} - 1} = \frac{g(x)}{f(x)}$$
$$\leq \frac{x^{1/2}}{x^{(x-1)/(2x)} - 1}$$
$$\leq \frac{1}{x^{(-1)/(2x)} - \frac{1}{\sqrt{x}}} \qquad \text{(divide top and bottom by } \sqrt{x}\text{)}$$

The limit as $x \to \infty$ of the first and last expressions is 1, so by the squeeze theorem, $\lim_{x\to\infty} \frac{g(x)}{f(x)} = 1$. Furthermore, since $1 < \frac{g(x)}{f(x)}$ (for $x > 1$) and $f(x)$ is positive for all $x > 1$, it follows that $f(x) < g(x)$ for $x > 0$. $\square$

---

[23]There is a caveat that the results of [47] are stated in terms of how many subsets of $d + 1$ colors have an intersection, so if say $d + 1 + n$ colors intersect, that accounts for $\binom{d+1+n}{d+1}$ of the $2^d - d$ points. When we actually apply these results in this section, the hypothesis that the parameter k from our motivating question is optimal at $d + 1$ will ensure we don't have to worry about this multiplicity.

Definition 10.4 (Simplicial Subdivision)**.** *A simplicial subdivision (sometimes called a triangulation) of a hypercube H is a finite set of d-simplices $\Sigma = \{\sigma_1, \ldots, \sigma_m\}$ such that $H = \bigcup_{i=1}^{m} \sigma_i$, and for every $i, j \in [m]$ with $i \neq j$, $\sigma_i \cap \sigma_j$ is either empty, or a face of both σ_i and of σ_j. The vertices of a simplicial subdivision are the vertices of all its d-simplices; that is, $V(\Sigma) \stackrel{\text{def}}{=} \bigcup_{i=1}^{m} V(\sigma_i)$.*

The following definition is comparable to Definition 9.12.

Definition 10.5 (Sperner Coloring)**.** *Let $d \in \mathbb{N}$, and $H = [0, 1]^d$, and $V(H) = \{0, 1\}^d$ denote a set of colors (which is exactly the set of vertices of $[0, 1]^d$ so that colors and vertices are identified). Let $\Sigma = \{\sigma_1, \ldots, \sigma_m\}$ be a simplicial subdivision of H. Let $\chi : V(\Sigma) \to V(H)$ be a coloring function such that for any face F of $[0, 1]^d$, for any $\vec{x} \in F \cap V(\Sigma)$, it holds that $\chi(\vec{x}) \in F$ (informally, the color of $\vec{x}$ is one of the vertices in the face F). Then Σ along with χ is called a Sperner coloring of H.*

Further, for any $\sigma \in \Sigma$, we define $\operatorname{colorset}(\sigma) = \{\chi(\vec{x}) : \vec{x} \in V(\sigma)\}$.

The above definitions are easily generalized to convex polytopes (but we will not need the more general definitions except for the discussion in this paragraph), and with the generalized definitions, De Loera, Peterson, and Su showed in [45] that for any convex polytope P in d dimensions with n vertices, any Sperner coloring (Σ, χ) of P will have at least $n-d$ "fully colored simplices". That is, Σ must contain $n-d$ simplices $\sigma_1, \ldots \sigma_{n-d}$ such that $|\operatorname{colorset}(\sigma_i)| = d + 1$ (recall that σ_i only has $d + 1$ vertices, so this means that each vertex of σ_i is assigned a different color). In fact, they show that if $i \neq j$, then $\operatorname{colorset}(\sigma_i) \neq \operatorname{colorset}(\sigma_j)$. From another perspective, they say that there are at least $(n - d)$ many $(d + 1)$-cardinality subsets of the colors, each being the colorset of some simplex in the subdivision.

They argue that this result $(n - d)$ is tight when P is an arbitrary convex polytope, but that improvements may be made when restricting to specific types of convex polytopes (e.g. hypercubes). In fact, their bound can be improved significantly when restricting to hypercubes by combining the results of [45] with other results about simplicial decompositions of hypercubes.

Next, we give a definitional name to the best possible parameter for each dimension d. We could get away without defining this value in our paper and just use bounds on this value (since that is all we will use anyway), but we want to make this quantity explicit since it relates nicely to some other areas of research regarding the d-dimension hypercube.

Definition 10.6 (Sperner Number)**.** *Let $d \in \mathbb{N}$, and $H = [0, 1]^d$, and let S_d be the set of all values $s \in \mathbb{N} \cup \{0\}$ such that the following statement holds:*

> *For any Sperner coloring (Σ, χ) of H, there exists at least s-many $(d + 1)$-cardinality sets $J_1, \ldots, J_s \subseteq V(H)$ such for each J_i, Σ contains a simplex σ_i such that $\operatorname{colorset}(\sigma_i) = J_i$.*

We call the minimum value of S_d the Sperner number of $[0, 1]^d$ or the dth Sperner number, and we denote it by $\operatorname{Sperner}(d) \stackrel{\text{def}}{=} \min(S_d)$.

Since $[0, 1]^d$ is a convex polytope in d dimensions with 2^d vertices, it follows immediately from [45] that any $\operatorname{Sperner}(d) \geq 2^d - d$. In fact, they implicitly showed something stronger than this. By [45, Thm. 1] and the comment following the proof of [45, Cor. 3] on pages 18-19 of the July 2001 version 8, the collection of colorsets of simplices in any simplicial subdivision induces a "face-to-face simplicial cover" of $[0, 1]^d$ using those colorsets to define simplices (in the terminology from [39], a "dissection" of $[0, 1]^d$). Thus, $\operatorname{Sperner}(d)$ is at least as large as the number of simplices needed in a dissection of $[0, 1]^d$. In other words, the dissection number of the d-cube gives a lower bound on the Sperner number.

We can also get a trivial upper bound on the dth Sperner number by noting that, by definition, it can be no larger than the size of the minimal cardinality simplicial subdivision of $[0, 1]^d$. Similarly, it can be no

larger than the size of the minimal cardinality triangulation[24] of $[0,1]^d$—this is because any triangulation is a valid simplicial subdivision.

Summarizing the above three paragraphs using the notation in [39], we have the following chain of inequalities for properties of $[0,1]^d$. We emphasize that all quantities below are with respect to using only vertices of the $[0,1]^d$ (i.e. no extra vertices are allowed). For example, Below *et. al.* show in [48] that the use of extra vertices can drastically reduce the necessary size of a triangulation.

$$\text{cover}(d) \leq \text{dis}(d) \leq \text{Sperner}(d) \leq \text{triang}(d) \tag{1}$$

Glazyrin showed in [39] that $(d+1)^{\frac{d-1}{2}} \leq \text{dis}(d)$ and Orden and Santos showed in [49] that $\text{triang}(d) \in O(0.816^d d!)$ which are the best known asymptotic bounds to date, and they provide upper and lower[25] bounds on the dth Sperner number.

While it is known for general polytopes that the dissection number and the triangulation number are not equal (see an example in [48]), for hypercubes it is still an open question if $\text{dis}(d)$ equals $\text{triang}(d)$ or not. We provide the first few values of $\text{Sperner}(d)$ which are exactly the values where $\text{triang}(d)$ is know to equal $\text{dis}(d)$ (see [39, Table 1]).

d	Sperner(d)
1	1
2	2
3	5
4	16

Table 2: Known values of Sperner(d).

The next lemma is really the result that we want regarding the quantity Sperner(d) because we have interest not in colorings of simplicial decompositions, but colorings of the entire hypercube as in Definition 9.12. Due to the compactness of $[0,1]^d$, we can transfer the defining property of Sperner(d) from Sperner colorings of simplicial subdivisions to Sperner/KKM colorings. The technique to do so is the same one that is used in many proofs that use variations of Sperner's lemma to prove variations of the KKM lemma.

Lemma 10.7. *Let $d \in \mathbb{N}$ and $H = [0,1]^d$ so $V(H) = \{0,1\}^d$. In any Sperner/KKM coloring $\chi : [0,1]^d \to V(H)$, there are $\text{Sperner}(d)$ distinct $(d+1)$-cardinality sets $J_1, \ldots, J_{\text{Sperner}(d)} \subseteq V(H)$ such that $\overline{\cap_{v \in V} \chi^{-1}(v)} \neq \emptyset$.*

Proof. The proof is identical to the proof of [45, Cor. 3] with the exception that we know each simplicial subdivision in the sequence (what they call triangulations) contains not just (in their notation) "$c(P)$" different colorsets, but in fact Sperner(d) different colorsets. □

With this lemma, we can now show that the value of $\epsilon = \frac{D}{2d}$ is optimal for $d = 2$.

Proposition 10.8 (Optimal Tolerance (ϵ) in $\mathbb{R}^2$). *If $d = 2$ and $\mathcal{P}$ is a partition of $\mathbb{R}^d = \mathbb{R}^2$, and there exists $D \in (0, \infty)$ such that for all $X \in \mathcal{P}$, D is a strict pairwise bound for X, and if $\epsilon \in (0, \infty)$ is such that for all $\vec{p} \in \mathbb{R}^d$,*

$$|\mathcal{N}_\epsilon(\vec{p})| \leq d + 1 = 3$$

---

[24]We use the term triangulation as in [39], which is different from how it is used in [45]. By triangulation, we mean a simplicial subdivision Σ such that $V(\Sigma) = V([0,1]^d)$ (i.e. the only vertices used in the subdivision are ones from the original hypercube).

[25]For $d < 5$, the lower bound of $\text{Sperner}(d) \geq 2^d - d$ can be used instead of $\text{Sperner}(d) \geq (d+1)^{\frac{d-1}{2}}$, however, noting that $\text{Sperner}(d)$ is an integer, this actually only gives an improved lower bound in the case $d = 3$, providing a bound of 5 instead of 4.

then $\epsilon \leq \frac{D}{4} = \frac{D}{2d}$.

Proof. We show that for any $\epsilon > \frac{D}{4}$, there is some point $\vec{p} \in \mathbb{R}^d$ such that $\overline{B}_\epsilon$ intersects at least $d + 2$ members of $\mathcal{P}$. Let $\epsilon > \frac{D}{4}$. By the Second Optimality Theorem there is some $\vec{p_0} \in \mathbb{R}^d$ such that $|\{X \in \mathcal{P} : X \cap \overline{B}_\epsilon(\vec{p_0}) \neq \emptyset\}| \geq d + 1 = 3$ (so along with the hypothesis we actually have cardinality exactly $d + 1 = 3$). Consider the hypercube $H = \overline{B}_{\frac{D}{2}}(\vec{p_0})$, and let $\mathcal{S}$ be the partition of H induced by $\mathcal{P}$, and let χ be an admitted Sperner/KKM coloring. By Proposition 9.18, since $\mathrm{Sperner}(2) = 2$, there exists distinct points $\vec{x}, \vec{y} \in H$ (one of them possibly equal to $\vec{p_0}$) each belonging to the closure of $d + 1$ members of $\mathcal{S}$ (and thus $\mathcal{P}$), denoted $\mathcal{N}_{\vec{x}}$ and $\mathcal{N}_{\vec{y}}$, and $\mathcal{N}_{\vec{x}} \neq \mathcal{N}_{\vec{y}}$. Thus, one of these points (wlog $\vec{x}$) is at the closure of a different set of $d + 1$ members of $\mathcal{S}$ than $\vec{p_0}$ is. Since $\vec{x} \in H$, $d_{max}(\vec{x}, \vec{p_0}) \leq \frac{D}{2}$. Consider the midpoint $\vec{c} = \frac{\vec{x} + \vec{p_0}}{2}$ observing that $d_{max}(\vec{x}, \vec{c}) \leq \frac{D}{4}$ and $d_{max}(\vec{p_0}, \vec{c}) \leq \frac{D}{4}$, and since $\epsilon > \frac{D}{4}$, $\overline{B}_\epsilon(\vec{c})$ contains an open set around $\vec{x}$ and an open set around $\vec{p_0}$, and thus intersects the $d + 1$ members containing $\vec{x}$ in their closure and also the $d + 1$ members containing $\vec{p_0}$ in their closure. Because these sets are not the same, $\overline{B}_\epsilon(\vec{c})$ intersects at least $d + 2$ members of $\mathcal{S}$ and thus $\mathcal{P}$. $\qquad\square$

Thus, we have shown that value of ϵ we are able to achieve with the reclusive partitions is optimal in $\mathbb{R}^1$ and $\mathbb{R}^2$, but these proofs were both special cases that resulted from the ϵ values in question being large enough relative to the diameter to easily argue about. We will now provide an upper bound on ϵ for all dimension $d \geq 2$, and while the technique will be more general than the technique for $\mathbb{R}^2$, it will use the same basic idea of locating points at the closures of $d + 1$ members, arguing about distances between these points, and limiting the number of occurences of such points based on ϵ.

Lemma 10.9. *If $d \in \mathbb{N}$, and $\mathcal{P}$ is a partition of $\mathbb{R}^d$, and there exists $D \in (0, \infty)$ such that for all $X \in \mathcal{P}$, it holds that D is a strict pairwise bound for X, and if there exists $\epsilon \in (0, \infty)$ such that for all $\vec{p} \in \mathbb{R}^d$,*

$$|\mathcal{N}_\epsilon(\vec{p})| \leq d + 1$$

then for any $\vec{\alpha} \in \mathbb{R}^d$, the hypercube $H = \vec{\alpha} + [0, D]^d$ contains at least $\mathrm{Sperner}(d)$ distinct points $\vec{s}_1, \ldots \vec{s}_{\mathrm{Sperner}(d)} \in H$ such that for each $i, j \in [\mathrm{Sperner}(d)]$,

$$|\mathcal{N}_{\overline{0}}(\vec{s}_i)| = d + 1,$$

and if $i \neq j$, then $\mathcal{N}_{\overline{0}}(\vec{s}_i) \neq \mathcal{N}_{\overline{0}}(\vec{s}_j)$.

Proof. As in other proofs, consider the set $H = \vec{\alpha} + [0, D]^d$, consider the partition $\mathcal{S}$ induced by $\mathcal{P}$ which is non-spanning and admits a Sperner/KKM coloring. So by Lemma 10.7, there are at least $\mathrm{Sperner}(d)$ distinct $(d + 1)$-cardinality sets $J_1, \ldots, J_{\mathrm{Sperner}(d)} \subseteq V(H)$ such that $\cap_{v \in V} \chi^{-1}(v) \neq \emptyset$ (i.e. for each set J_i of $d + 1$ colors, there is some point in the closure of all colors in J_i). For each i, let $\vec{s}_i \in \overline{\cap_{v \in V} \chi^{-1}(v)}$ be arbitrary (i.e. $\vec{s}_i$ is in the closure of all colors in J_i).

We now justify that if $i \neq j$, the $\vec{s}_i \neq \vec{s}_j$. By exactly the same argument as Proposition 9.18, we have $|\{X \in \mathcal{P} : \overline{X} \ni \vec{s}_i\}| = d + 1$. It follows immediately that $\{X \in \mathcal{P} : \overline{X} \ni \vec{s}_i\} \neq \{X \in \mathcal{P} : \overline{X} \ni \vec{s}_j\}$ because $J_i \neq J_j$ and the set $\{X \in \mathcal{P} : \overline{X} \ni \vec{s}_i\}$ contains exactly one member of each color in J_i (and similarly for J_j). $\qquad\square$

Theorem 10.10 (Universal Tolerance (ϵ) Bound)**.** *If $d \in \mathbb{N}$, and $\mathcal{P}$ is a partition of $\mathbb{R}^d$, and there exists $D \in (0, \infty)$ such that for all $X \in \mathcal{P}$, it holds that D is a strict pairwise bound for X, and if there exists $\epsilon \in (0, \infty)$ such that for all $\vec{p} \in \mathbb{R}^d$,*

$$|\mathcal{N}_\epsilon(\vec{p})| \leq d + 1,$$

then $\epsilon \leq \frac{D}{2(\sqrt[d]{\mathrm{Sperner}(d)} - 1)}$ (if $d > 1$). In particular, $\epsilon \leq \frac{D}{2\sqrt{d}}$ (for any d).

Proof. For any $\vec{\alpha} \in \mathbb{R}^d$, consider the hypercube $H = \vec{\alpha} + [0, D]^d$, and (as in many of our other proofs) an admitted Sperner/KKM coloring of an induced partition. Then let $\vec{s}_1, \ldots \vec{s}_{\mathrm{Sperner}(d)} \in H$ be points as in

46

Lemma 10.9. It follows that for any $i \neq j$ that $d_{max}(\vec{s}_i, \vec{s}_j) \leq \epsilon$ because if they were not, then the midpoint $\vec{c} = \frac{\vec{s}_i + \vec{s}_j}{2}$ would have the property that $\overline{B}_\epsilon(\vec{c})$ contains an open set around $\vec{s}_i$ and around $\vec{s}_i$, and thus $\overline{B}_\epsilon(\vec{c})$ would intersect strictly more than $d + 1$ members of $\mathcal{P}$ which would contradict the hypothesis.

Thus, the points $\vec{s}_i$ give rise to closed balls $\overline{B}_\epsilon(\vec{s}_i)$ which are pairwise disjoint. Further, because $\vec{s}_i \in H = \vec{\alpha} + [0, D]^d$, it follows that $\overline{B}_\epsilon(\vec{s}_i) \subseteq \vec{\alpha} + [-\epsilon, D + \epsilon]^d$. This gives the following volume/measure comparison argument:

$$\text{Sperner}(d) \cdot (2\epsilon)^d = m\left(\bigsqcup_{i=1}^{\text{Sperner}(d)} \overline{B}_\epsilon(\vec{s}_i)\right) \leq m(\vec{\alpha} + [-\epsilon, D + \epsilon]^d) = (D + 2\epsilon)^d.$$

Taking dth roots of both sides and manipulating the equations, we have the stated inequality for ϵ for $d \geq 2$ (with $d = 1$, there would be a division by 0 because $\text{Sperner}(1) = 1$).

To show the "in particular" statement, for $d \geq 2$, increase the bound by replacing $\text{Sperner}(d)$ with the lower bound $(d + 1)^{\frac{d-1}{2}} \leq \text{Sperner}(d)$ (Equation 1 and [39]) (note that for $d = 1$ this would give division by zero). Lastly then apply Lemma 10.3.

To show the "in particular" statement, for $d = 1$, note that $d = \sqrt{d}$, so apply Corollary 10.2. $\qquad\square$

Note that it is not an issue that the above bound does not hold for $d = 1$ because we already have exact bounds for $d = 1$ and $d = 2$ in Proposition 10.1 and Proposition 10.8. Nonetheless, we don't believe that the -1 term in the above expression is necessary, but we have not yet come up with an argument that removes it.

Since the first four values of $\text{Sperner}(d)$ are known, it is the first bound in the theorem above that gives the better bound. For convenience, we summarize the bounds on ϵ for $d \in \{1, 2, 3, 4\}$.

d	Sperner(d)	ϵ upper bound	Reason
1	1	$\frac{D}{2}$	Proposition 10.1
2	2	$\frac{D}{4}$	Proposition 10.8
3	5	$\frac{D}{2(\sqrt[3]{5}-1)} \approx \frac{D}{1.419952}$	Theorem 10.10
4	16	$\frac{D}{2(\sqrt[4]{16}-1)} = \frac{D}{2}$	Theorem 10.10

Table 3: Summary of upper bounds on ϵ.

While the above gives an upper bound whose denominator grows only as the square root of d, we conjecture that the true maximal value of ϵ has a tighter bound with the denominator growing linearly in d. If this conjecture is true, that means that reclusive partitions can achieve within a constant factor of the maximum possible value of ϵ.

Conjecture 10.11 (Linear Universal Tolerance (ϵ) Conjecture). *If $d \in \mathbb{N}$, and $\mathcal{P}$ is a partition of $\mathbb{R}^d$, and there exists $D \in (0, \infty)$ such that for all $X \in \mathcal{P}$, $\text{diam}(X) \leq D$, and if $\epsilon \in (0, \infty)$ such that for all $\vec{p} \in \mathbb{R}^d$,*

$$|\mathcal{N}_\epsilon(\vec{p})| \leq d + 1$$

then $\epsilon = \frac{D}{\Omega(d)}$.

In fact, if it holds that $\text{Sperner}(d) \geq \frac{d!}{c^d}$ for some constant $c \in (0, \infty)$, or even if $\text{Sperner}(d) \geq \left(\frac{d}{c'}\right)^d$, then the conjecture above holds. This is consistent with the current bounds mentioned earlier in this section

due to [39] and [49]:

$$(d+1)^{\frac{d-1}{2}} \leq \operatorname{dis}(d) \leq \operatorname{Sperner}(d) \leq \operatorname{triang}(d) \in O(0.816^d d!).$$

To see that the conjecture would hold in this case, note that we can use Stirling's approximation as a lower bound for the factorial case above:

$$\left(\frac{d}{ce}\right)^d \leq \frac{\sqrt{2\pi d}\left(\frac{d}{e}\right)^d}{c^d} \leq \frac{d!}{c^d} \leq \operatorname{Sperner}(d).$$

So the first case above reduces to the second case.

So assuming $\operatorname{Sperner}(d) \geq \left(\frac{d}{c'}\right)^d$ for some constant $c' \in (0, \infty)$, and using this lower bound in the previous upper bound for ϵ, we have

$$\epsilon \leq \frac{D}{2(\sqrt[d]{\operatorname{Sperner}(d)} - 1)} \leq \frac{D}{2(\frac{d}{c} - 1)}$$

showing that under these assumptions $\epsilon = \frac{D}{\Omega(d)}$.

11 Application of Reclusive Partitions to Deterministic Rounding

Throughout this section, let A denote a $d \times d$ reclusive matrix. Other notation from Section 7 will be used as well. Our first goal of this section will be to show that given $\vec{x} \in \mathbb{R}^d$ we can efficiently compute $\operatorname{rep}(X)$ for the unique $X \in \mathcal{P}_A$ such that $\vec{x} \in X$. In other words, we can quickly map points of $\mathbb{R}^d$ to the much sparser set of representative points $\{\operatorname{rep}(X): X \in \mathcal{P}_A\} = L_A$. First observe a simple fact.

Fact 11.1. *Let $a \in \mathbb{R}$ and $b \in [0, 1)$ such that $a - b \in \mathbb{Z}$. Then $a - b = \lfloor a \rfloor$.*

Proof. If $a - b > \lfloor a \rfloor$ then $a - b \geq \lfloor a \rfloor + 1$, so $a - \lfloor a \rfloor \geq 1 + b \geq 1$ which is a contradiction. If $a - b < \lfloor a \rfloor$ then $a - b \leq \lfloor a \rfloor - 1$, so $a - \lfloor a \rfloor \leq b - 1 < 0$ implying that $a < \lfloor a \rfloor$ which is a contradiction. Thus $a - b = \lfloor a \rfloor$. $\square$

Proposition 11.2 (Efficient Computation of Representatives). *Let $d \in \mathbb{N}$, and A be a $d \times d$ reclusive matrix, and $\mathcal{P}_A$ its reclusive partition. For any $\vec{x} \in \mathbb{R}^d$, let $X \in \mathcal{P}_A$ be the unique hypercube such that $\vec{x} \in X$. Then $\operatorname{rep}(X)$ can be efficiently[26] computed in terms of $\vec{x}$ and A.*

Remark 11.3. *For additional intuition of the following proof, see the proof of Claim 7.2 which is an inductive proof instead of an algebraic one.*

Proof. Recall that $\operatorname{rep}(X) = A\vec{m}$ for a unique $\vec{m} \in \mathbb{Z}^d$, so it will suffice to compute $\vec{m}$ because $\operatorname{rep}(X)$ can then be computed via a single matrix multiplication. We show by induction that if we have computed m_i for all $i > k$, then we can compute m_k. The main reason that we can do this is that A is a triangular matrix, so the technique for computing each m_i has the flavor of Gaussian elimination. The inductive base case is that we have not computed any m_i.

By the definition of $\mathcal{P}_A$ and $\operatorname{rep}(X)$, we have that $\vec{x} \in \operatorname{rep}(X) + [0, 1)^d$, so let $\vec{\alpha} = \vec{x} - \operatorname{rep}(X)$ so

---

[26]By efficient, we mean that the computation can be done with $O(d)$ matrix multiplications with matrices of size $d \times d$.

$\vec{\alpha} \in [0,1)^d$. Now we consider just the kth coordinate.

$$
\begin{aligned}
x_k &= \alpha_k + \text{rep}(X)_k \\
&= \alpha_k + (A\vec{m})_k \\
&= \alpha_k + \sum_{i=1}^{d} a_{ki}m_i && \text{(Def'n of matrix multiplication)} \\
&= \alpha_k + \sum_{i=k}^{d} a_{ki}m_i && (A \text{ is reclusive, so } a_{ki} = 0 \text{ for } i < k) \\
&= \alpha_k + m_k + \sum_{i=k+1}^{d} a_{ki}m_i && (A \text{ is reclusive, so } a_{kk} = 1. \text{ Summation might be empty)}
\end{aligned}
$$

We now reformulate in terms of m_k.

$$
\begin{aligned}
m_k &= x_k - \sum_{i=k+1}^{d} a_{ki}m_i - \alpha_k && \text{(Solve for } m_k) \\
&= \left\lfloor x_k - \sum_{i=k+1}^{d} a_{ki}m_i \right\rfloor && \text{(Fact 11.1)}
\end{aligned}
$$

Thus, m_k can be computed as a floor in terms of A, $\vec{x}$, and the already known m_i for $i > k$. As mentioned, we can return vector $\text{rep}(X) = A\vec{m}$.

Altogether, this computation requires $O(d^2)$ additions and multiplications. We need $O(d)$ to compute each m_i and i takes d many values, and we need $O(d^2)$ operations to compute $A\vec{m}$. $\qquad \square$

Having shown that we can efficiently compute the representative corners, we turn to an application in pseudodeterministic computations. Consider a multi-valued function such as $f(1^n) =$ the set of n-bit primes. We know a probabilistic polynomial-time for this function. A deficiency of this algorithm is that two different runs of the algorithm may produce two different valid outputs. Is there a probabilistic algorithm that outputs a canonical prime number? I.e, most of the random choices of the algorithm will produce the same output. Motivated by this, Gat and Goldwasser [2] defined the notion of *pseudodeterministic algorithms*.

Definition 11.4. *Let f be a total multi-valued function[27]. We say that f admits a* polynomial-time, pseudodeterministic algorithm, *if there is a probabilistic polynomial-time algorithm A such that for every x, there exists $v \in f(x)$ such that $\Pr[A(x) = v] \geq 2/3$.*

Goldreich [3] generalized the notion to k-pseudodeterministic algorithms.

Definition 11.5. *Let f be total multi-valued function. We say that f admits a* polynomial-time, k-pseudodeterministic algorithm, *if there is a probabilistic polynomial-time algorithm A such that for every x, there exists a set $S_x \subseteq f(x)$, and $\Pr[A(x) \in S] \geq \frac{k+1}{k+2}$.*

Suppose there is some function $f : \{0,1\}^* \to \mathbb{R}^d$ (for some $d \in \mathbb{N}$) that admits a (ϵ, δ) additive approximation algorithm M. I.e,

$$
\Pr[|M(x) - f(x)|_\infty \leq \epsilon] \geq 1 - \delta
$$

---

[27]A multi-valued function maps inputs to non-empty sets of outputs.

A natural question that arises is whether we can make such algorithms pseudodeterministic. Using reclusive partitions, we show M can be we can arrive a $(d+1)$ modified to M' such that M' is a $(d+1)$-pseudodeterministic algorithm (with the same probability guarantee of $1-\delta$) at a small loss on approximation guarantee. The algorithm M' is describe below.

Algorithm 1 $(d+1)$-Pseudodeterministically approximating $f\colon \{0,1\}^* \to \mathbb{R}^d$

Let $d \in \mathbb{N}$, and A be a $d \times d$ reclusive matrix, $\mathcal{P}_A$ its reclusive partition, and Δ_A its reclusive distance. Let $f\colon \{0,1\}^* \to \mathbb{R}^d$ be any function, and let $\epsilon \in [0,\infty)$ and $\delta \in [0,1]$, and let M be an efficient (ϵ,δ)-approximation algorithm for f with respect to the d_{max} metric (i.e. for all $x \in \{0,1\}^*$, $\mathbb{P}[d_{max}(A(x), f(x)) \leq \epsilon] \geq 1-\delta$).

Let $\phi\colon \mathbb{R}^d \to \mathbb{R}^d$ be the bijection defined by $\phi(\vec{x}) \stackrel{\text{def}}{=} \dfrac{\Delta_A}{\epsilon} \cdot \vec{x}$.

Let M' be the following algorithm (which takes an input in $\{0,1\}^*$).

1: **procedure** M'(x)
2:     $\vec{\alpha} \leftarrow M(x)$ ▷ Run M just once on the input
3:     Determine the unique $X \in \mathcal{P}_A$ such that $\phi(\vec{\alpha}) \in X$ ▷ This does not happen explicitly
4:     $\vec{r} \leftarrow \text{rep}(X)$ ▷ Computed as in Proposition 11.2
5:     **return** $\phi^{-1}(\vec{r})$ ▷ ϕ^{-1} is just multiplication

---

Note that in Algorithm 1, we really do allow ϵ and δ to be as general as stated. However, in practice, because M must be an (ϵ,δ)-approximation algorithm, both ϵ and δ will be small. The domain is chosen as $\{0,1\}^*$ since this is the formal input of algorithms, but in practice the domain of the algorithm could be a different set.

Proposition 11.6 (Algorithm Guarantees). *Algorithm 1 is an efficient $(d+1)$-pseudodeterministic (ϵ',δ)-approximation algorithm for f where $\epsilon' = \epsilon(\frac{1}{\Delta_A} + 1)$.*

Proof. Let all notation be as in Algorithm 1. Note that for any $\vec{a}, \vec{b} \in \mathbb{R}^d$ we have $d_{max}(\vec{a}, \vec{b}) = \frac{\epsilon}{\Delta_A} d_{max}(\phi(\vec{a}), \phi(\vec{b}))$ via linearity of norms and the linearity of ϕ:

$$\begin{aligned}
d_{max}(\vec{a}, \vec{b}) &= \|\vec{a} - \vec{b}\|_\infty \\
&= \|\frac{\epsilon}{\Delta_A}\left(\frac{\Delta_A}{\epsilon}\vec{a} - \frac{\Delta_A}{\epsilon}\vec{b}\right)\|_\infty \\
&= \frac{\epsilon}{\Delta_A}\|\phi(\vec{a}) - \phi(\vec{b})\|_\infty \\
&= \frac{\epsilon}{\Delta_A}d_{max}(\phi(\vec{a}), \phi(\vec{b}))
\end{aligned}$$

By the defining property of M, for any $x \in \{0,1\}^*$, with probability at least $1-\delta$, we have $d_{max}(\vec{\alpha}, f(x)) \leq \epsilon$ which, by the prior comment, holds if and only if $d_{max}(\phi(\vec{\alpha}), \phi(f(x))) \leq \Delta_A$. In this case we can bound

the error of the approximation:

$$
\begin{aligned}
\text{error} &= d_{max}(\phi^{-1}(\vec{r},\ f(x)) \\
&= \frac{\epsilon}{\Delta_A} d_{max}(\phi(\phi^{-1}(\vec{r})),\ \phi(f(x))) && \text{(By the opening comment)} \\
&= \frac{\epsilon}{\Delta_A} d_{max}(\vec{r},\ \phi(f(x))) && (\phi \circ \phi^{-1} \text{ is identity map}) \\
&\leq \frac{\epsilon}{\Delta_A} [d_{max}(\vec{r},\ \phi(\vec{\alpha})) + d_{max}(\phi(\vec{\alpha}),\ \phi(f(x)))] && \text{(Triangle Inequality)} \\
&\leq \frac{\epsilon}{\Delta_A} [d_{max}(\vec{r},\ \phi(\vec{\alpha})) + \Delta_A] && \text{(By the second comment above)} \\
&\leq \frac{\epsilon}{\Delta_A} [1 + \Delta_A] && (\vec{r} \in X \text{ and } \phi(\vec{\alpha}) \in X, \text{ and } X \text{ is a unit hypercube}) \\
&= \epsilon \left(\frac{1}{\Delta_A} + 1 \right)
\end{aligned}
$$

In other words, on any input $x \in \{0,1\}^*$, with probability at least $(1 - \delta)$, the value $\vec{r}$ returned by M' will be within a distance $\epsilon' = \epsilon \left(\frac{1}{\Delta_A} + 1 \right)$ of the true function value $f(x)$. This proves that M' is an (ϵ', δ)-approximation of f.

That M' is $(d+1)$-pseudodeterministic follows from the Partition Theorem (Theorem 7.18). Using the notation $\vec{c}$, N, and $\mathcal{S}$ from that theorem, let $\vec{c} = \phi(f(x))$. As stated above, with probability at least $1 - \delta$ we have $d_{max}(\phi(\vec{\alpha}), \phi(f(x))) \leq \Delta_A$, in which case $\phi(\vec{\alpha}) \in N$, so the unique hypercube $X \in \mathcal{P}_d$ containing $\phi(\vec{\alpha})$ intersects N, and so $X \in \mathcal{S}$. Thus, with probability at least $1 - \delta$, $\phi^{-1}(\vec{r}) = \phi^{-1}(\text{rep}(X))$ is one of at most $d+1$ values. This proves that M' is $(d+1)$-pseudodeterministic[28].

The efficiency of M' follows from the efficiency of M as well as Proposition 11.2.

$\square$

We show another application in the context of sample complexity of pseudodeterministic algorithms. Let $f_1, \ldots, f_d$ be functions from $\{0,1\}^n$ to $[0,1]$. Consider algorithms that have a blackbox access to these functions. I.e, the algorithm can generate a query q ask for the values of $f_i(q)$ for some $1 \leq i \leq d$. The goal of the algorithm is to obtain a (ϵ, δ)-approximation to the vector $\langle \mathbb{E}(f_i) \rangle_{i=1}^d$. with respect to d_{max} metric. This means that with probability at least $1 - \delta$, the vector $\vec{\alpha} \in \mathbb{R}^d$ returned by the algorithm is such that $d_{max}(\vec{\alpha}, \langle \mathbb{E}(f_i) \rangle_{i=1}^d) \leq \epsilon$ which is equivalent to saying that with probability at least $1 - \delta$, it holds for all i that vector $|\alpha_i - \mathbb{E}(f_i)| \leq \epsilon$.

Goldreich proved that there is $(d+1)$-pseudodeterministic algorithm for this task, and this algorithm has a sample complexity of $\widetilde{O}(d^4)$. Using reclusive partitions, we show that the sample complexity can be improved to $\widetilde{O}(d^2)$ samples.

Proposition 11.7 (Application to Pseudodeterministic Algorithms)**.** *Let* $f_1, \ldots, f_d : \{0,1\}^n \to [0,1]$ *be functions and* $\epsilon \in (0, \infty)$ *and* $\delta \in (0, \frac{1}{2})$. *There is an algorithm that given sample access to each* f_i *can* $(d+1)$-*pseudodeterministically* (ϵ, δ)-*approximate* $\langle \mathbb{E}(f_i) \rangle_{i=1}^d$ *relative to the* d_{max} *metric using* $O \left(\frac{(d+1)^2}{\epsilon^2} \cdot \log \left(\frac{d}{\delta} \right) \right)$ *samples.*

Proof. First define an algorithm M as follows. M selects uniformly at random $m = O \left(\frac{(d+1)^2}{\epsilon^2} \cdot \log \left(\frac{d}{\delta} \right) \right)$ points $x_1, \ldots, x_m \in \{0,1\}^n$, queries each function f_i on all of those values and outputs the vector $\vec{v}$ where

---

[28] To meet the technical definition of k-pseudodeterminism of Goldreich, it should be required that $1 - \delta \geq \frac{k+1}{k+2}$, so in this case with $k = d+1$, it should be required that $1 - \delta \geq \frac{d+2}{d+3}$

$v_i = \frac{1}{m} \sum_{j=1}^{m} f(x_j)$ is the observed sample average. Then we have for each $i \in [d]$ that

$$\mathbb{P}\left[|v_i - \mathbb{E}f_i| > \frac{\epsilon}{d+1} \right] < \frac{\delta}{d}$$

(this follows because $O(\frac{1}{\epsilon'^2} \cdot \log(1/\delta'))$ samples is sufficient for (ϵ', δ')-approximating the average of a single function $f : \{0,1\}^n \to [0,1]$). Thus, by a union bound, we have

$$\mathbb{P}\left[\exists i \text{ s.t. } |v_i - \mathbb{E}f_i| > \frac{\epsilon}{d+1} \right] < \delta$$

and taking the complement we have the desired result that

$$\mathbb{P}\left[d_{max}\left(\vec{v}, \langle \mathbb{E}f_i \rangle_{i=1}^d \right) \leq \frac{\epsilon}{d+1} \right] > 1 - \delta.$$

Thus, M will $(\frac{\epsilon}{d+1}, \delta)$-approximate $\langle \mathbb{E}(f_i) \rangle_{i=1}^d$ with respect to the d_{max} metric.

To complete the proof of the claim, let A be a $d \times d$ reclusive matrix as in Lemma 7.19 so that the reclusive distance is $\Delta_A = \frac{1}{d}$. Apply Algorithm 1 with M and A to obtain an algorithm M' which $(d+1)$-pseudodeterministically (ϵ, δ)-approximates $\langle \mathbb{E}(f_i) \rangle_{i=1}^d$. $\qquad\square$

Both Goldreich's algorithm and our own begin by taking samples to obtain an (ϵ', δ)-approximation of the averages, and then apply a rounding technique to add pseudodeterminism. The difference is that Goldreich's rounding technique requires $\epsilon' = \frac{\epsilon}{d^2}$ whereas our algorithm only requires $\epsilon' = \frac{\epsilon}{d+1}$. Goldreich's requirement for a better initial approximation requires more samples than ours does. The reason for this could be that Goldreich's rounding technique is randomized, while ours is deterministic.

12 Conclusions and Future Work

One remaining open question from this work is to either prove Conjecture 10.11 (possibly by improving the bound in Theorem 10.10) or to offer a construction with better tolerance (ϵ). In particular, improvements of [45, Thm. 1] (along with the extra properties they proved that we discussed after our Definition 10.6) would decrease our upper bound on the tolerance by increasing the lower bound on Sperner(d) in Definition 10.6. In fact, improvement of [45, Thm. 1] just for the case that the polytope P is a hypercube would yield improvements on our upper bounds for generic partitions, because we only every applied their results to hypercubes (because balls in the d_{max} metric are hypercubes). As a simplification of this open question, one might ask if the conjectured bound on the tolerance is at least optimal for partitions of unit hypercubes.

A second open question that is of interest to us and is relevant to rounding schemes is the following: if we consider secluded partitions with degree $k = \mathsf{poly}(d)$ instead of exactly $k = d+1$, what improvements are possible for the tolerance ϵ? For example, in the results of Hoza and Klivans, a degree of $\mathsf{poly}(d)$ would have been fine. Viewed another way, one might consider the best possible degree as a function of the tolerance—given some tolerance ϵ, what is the smallest k such that there exists a (k, ϵ)-secluded partition with members of diameter at most 1. A simple result is that for a tolerance of $\epsilon > \frac{1}{2}$, $k \geq 2^d$ because for any point $\vec{p}$, the 2^d corners of $\overline{B}_\epsilon(\vec{p})$ are all d_{max} distance more than 1 apart, so because the partition has members of diameter at most 1, each corner must belong to different member of the partition, so $|\mathcal{N}_\epsilon(\vec{p})| \geq 2^d$. Thus, our results in this work show that for $\epsilon \leq \frac{1}{2d}$ one can achieve degree $k = d + 1 = \mathsf{poly}(d)$, and that is optimal, but for $\epsilon > \frac{1}{2}$, $k \geq 2^d = \mathsf{exp}(d)$, and we think it would be interesting to know the best value of k for $\epsilon \in (\frac{1}{2d}, \frac{1}{2})$ or at least know the largest ϵ such that $k = \mathsf{poly}(d)$.

A third area that is open is to better understand which lattice partitions are $(d+1, \epsilon)$-secluded (for some ϵ). We initially defined the notion of a reclusive matrix and reclusive partition in order to construct

some partition that was $(d + 1, \frac{1}{O(d)})$-secluded, and we expected that our definition imposed far more requirements on the matrix than were necessary (i.e. that there would be lots of partitions that did not meet our definition, but had similar structural properties). However, we discussed in Section 8 our conceptual understanding of the construction of these partitions by successive extrusions and shifts and expected that this would be robust to re-ordering the entries in a row of the matrix; however, we then showed by a specific example that doing so resulted in partitions that were not $(d + 1, \epsilon)$-secluded because there were points on the boundary of more than $d + 1$ members of the partition for dimensions greater than 5. This suggests to us that there is something deeper going on with these constructions than we initially realized, and it would be interesting to understand exactly what sets of basis vectors have the property that the integer linear combinations give the positions of hypercubes in a $(d + 1, \epsilon)$-secluded partition (for some ϵ).

A   Rounding Schemes in Prior Work

In this section, we will discuss in some detail how rounding is used in a number of publications and what properties of the rounding schemes are important in each of these papers. Not all of them benefit from our main construction and bounds, but we mention them nonetheless to highlight that there are a variety of perspectives one may reasonably take on what constitutes a good rounding scheme. Further, we think viewing each of these schemes as a partition (or distribution of partitions) highlights which publications have common goals in designing their rounding schemes. We begin by looking at a very simple rounding scheme, but though it is very simple, it shows up as a significant part of numerous publications. We have found that each of these publications independently walks through the construction, and we hope to demonstrate that each of these is doing the same thing under a different guise.

A.1   A Very Simple Rounding Scheme

Recalling that a deterministic rounding scheme for $\mathbb{R}$ is just a function $f : \mathbb{R} \to \mathbb{R}$, arguably, the most basic deterministic rounding scheme in $\mathbb{R}$ is the floor function, $\lfloor \cdot \rfloor$, which maps every real number to the largest integer that is not larger that it[29]. If one considers the partition induced by this deterministic rounding scheme, it is the partition of half-open unit intervals $\mathcal{P}_{\lfloor \cdot \rfloor} = \{[n, n+1) : n \in \mathbb{Z}\}$. There are three simple modifications one might wish to make to this rounding scheme.

First, one may want a "scaled" version. In the floor scheme, values might be rounded by as much as 1, but one might wish to have values rounded by at most α for some $\alpha \in (0, \infty)$. This can be accomplished by a modified floor function $\lfloor \cdot \rfloor_\alpha : \mathbb{R} \to \mathbb{R}$ defined by $\lfloor x \rfloor_\alpha \overset{\text{def}}{=} \alpha \lfloor x/\alpha \rfloor$. This function maps every real number to the largest integer multiple of α that is not larger than it. The partition induced by $\lfloor x \rfloor_\alpha$ is $\mathcal{P}_{\lfloor \cdot \rfloor_\alpha} = \{[\alpha n, \alpha(n+1)) : n \in \mathbb{Z}\}$.

The second modification that one might want is a "shift" of the floor scheme. For example, maybe it is desirable that 0.98 and 1.213 are rounded to the same value, and so one could (for example) choose the function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = \lfloor x - 0.3 \rfloor$ so that $f(0.98) = f(1.213) = 0$. More generally, one could pick any $\beta \in \mathbb{R}$ to shift by. This shift can be combined with a scaling $\alpha \in (0, \infty)$ to define the deterministic rounding scheme $\lfloor \cdot \rfloor_{\alpha, \beta}$ given by $\lfloor x \rfloor_{\alpha, \beta} \overset{\text{def}}{=} \lfloor x - \beta \rfloor_\alpha + \beta = \alpha \lfloor (x - \beta)/\alpha \rfloor$. The partition induced from this rounding scheme is $\mathcal{P}_{\lfloor \cdot \rfloor_{\alpha, \beta}} = \{[\alpha n + \beta, \alpha(n+1) + \beta) : n \in \mathbb{Z}\}$ [30]. Note, that by this definition, the difference between x and $\lfloor x \rfloor_{\alpha, \beta}$ will become relatively large as β is taken to be large, so typically β will only take values in $[0, \alpha)$.

The third modification that one might want to make to the floor scheme is to have a "different representative". In the floor function, each value in the interval $[n, n+1)$ is mapped/rounded to n, but it might make sense to map/round these values to some other point in the interval such as the midpoint (or it might even be desirable to map/round them to a point not in the interval). We can combine this with the scaling and shifting. Let α, β as before and $\gamma \in \mathbb{R}$ (it will be typical that γ is small, and to round to the midpoint we will let $\gamma = \alpha/2$). Define the deterministic rounding scheme $\lfloor \cdot \rfloor_{\alpha, \beta, \gamma}$ by $\lfloor x \rfloor_{\alpha, \beta, \gamma} \overset{\text{def}}{=} \lfloor x \rfloor_{\alpha, \beta} + \gamma = \alpha \lfloor (x - \beta)/\alpha \rfloor + \gamma$. The partition induced by this rounding scheme is the same as $\mathcal{P}_{\lfloor \cdot \rfloor_{\alpha, \beta}}$ because changing the value assigned to each member does not change the member.

The shift modification discussed above is typically most useful when applied in the context of a randomized rounding scheme (a distribution of functions) rather than a deterministic rounding scheme (a single function). The idea is that it is often desirable that for any fixed pair of points $x, y \in \mathbb{R}$ which are "sufficiently close", then it holds with "sufficiently high probability" (over the selection of function f from the distribution) that $f(x) = f(y)$. For example, fix some $\alpha \in (0, \infty)$ and consider the set of

---

[29] One could also consider the ceiling function, but floor tends to be used more often in practice as we shall see.

[30] To see this, observe that $\mathcal{P}_{\lfloor \cdot \rfloor_{\alpha, \beta}}$ is in fact a partition of $\mathbb{R}$ and that for any $n \in \mathbb{Z}$, if $x \in [\alpha n + \beta, \alpha(n+1) + \beta)$, then $(x - \beta)/\alpha \in [n, n+1)$ so $\lfloor (x - \beta)/\alpha \rfloor = n$ so $\lfloor x \rfloor_{\alpha, \beta} = \alpha n$. Thus all points in any member of $\mathcal{P}_{\lfloor \cdot \rfloor_{\alpha, \beta}}$ map to the same value, and points in two different members map to different values (i.e. n and n').

functions $\{\lfloor \cdot \rfloor_{\alpha,\beta} : \beta \in [0, \alpha)\}$ with distribution corresponding to β being distributed uniformly over $[0, \alpha)$. This gives a randomized rounding scheme with the following property: For any $\epsilon \in [0, \alpha)$, for any $x, y \in \mathbb{R}$ with $|x - y| \le \epsilon$, the probability that $\lfloor x \rfloor_{\alpha,\beta} = \lfloor y \rfloor_{\alpha,\beta}$ is greater than or equal to ϵ/α.

Intuitively this is because x and y end up in different members of the partition $\mathcal{P}_{\lfloor \cdot \rfloor_{\alpha,\beta}}$ if and only if one of the boundaries of that partition separate x and y which happens with probability $|x - y|/\alpha \le \epsilon/\alpha$.

We view the randomized rounding scheme above (for any distribution of $\beta \in \mathbb{R}$) as a distribution of partitions of $\mathbb{R}$ by half-open α-length intervals, and a value $x \in \mathbb{R}$ is randomly rounded by randomly obtaining a partition in the distribution, determining which member/interval of that partition contains x, and then returning the minimum value of that member/interval.

The ideas above easily generalizes to $\mathbb{R}^d$ for any $d \in \mathbb{N}$. One can view this generalization as being the above in each coordinate or (equivalently) as a vector version: for $\alpha \in (0, \infty)$ and a vector $\vec{\beta} = \langle \beta_i \rangle_{i=1}^d$, and a vector $\vec{\gamma} = \langle \gamma_i \rangle_{i=1}^d$ define $\lfloor \cdot \rfloor_{\alpha,\vec{\beta},\vec{\gamma}} : \mathbb{R}^d \to \mathbb{R}^d$ coordinatewise in the expected way: $\lfloor \vec{x} \rfloor_{\alpha,\vec{\beta},\vec{\gamma}} \stackrel{\text{def}}{=} \langle \lfloor x_i \rfloor_{\alpha,\beta_i,\gamma_i} \rangle_{i=1}^d$. If $\vec{\gamma}$ is not specified, it will be assumed to be $\vec{0}$. The partition induced by this scheme is $\mathcal{P}_{\lfloor \cdot \rfloor_{\alpha,\vec{\beta},\vec{\gamma}}} = \{\vec{\gamma} + [\alpha n, \alpha(n+1))^d : n \in \mathbb{Z}\}$. In other words, the partition induced by rounding each coordinate is a grid of unit hypercubes with some shift applied to the grid.

A.2 The Randomized Rounding Scheme of Saks and Zhou

The rounding scheme used by Saks and Zhou is the basic rounding scheme just introduced[31]. We briefly state the parameters of their scheme.

Let $d \in \mathbb{N}$ and $t = O(\log d)$ and $D = O(\log d)$. Let $\epsilon = d \cdot 2^{-t} \cdot 2^{-D}$. This will not be of much importance in this paper, but we want to highlight that this basic rounding scheme is used in multiple papers, so we briefly mention the parameters of the scheme for Saks and Zhou. Let $\alpha = 2^{-t}$ so $\alpha = 1/\mathsf{poly}(d)$. Let $S = \left\{ \frac{0}{2^D}, \frac{1}{2^D}, \ldots, \frac{2^D - 2}{2^D}, \frac{2^D - 1}{2^D} \right\}$ and let β be uniformly distributed over S and let $\vec{\beta}$ be the d^2 length vector in which every entry is β (i.e. $\vec{\beta} = \langle \beta \rangle_{i=1}^{d^2}$). Let $\vec{n} \in \mathbb{R}^{d^2}$. Then with probability at least $1 - \frac{O(d^3)}{2^D}$ (over the choice of β) it holds for all $\vec{m} \in B_\epsilon^\circ$ (w.r.t. the l^∞ norm/d_{max} metric) that $\lfloor \vec{n} \rfloor_{\alpha,\vec{\beta}} = \lfloor \vec{m} \rfloor_{\alpha,\vec{\beta}}$. In other words, with high probability, the entire ϵ-ball of vectors around $\vec{n}$ are rounded to the same value. The reason is that for any coordinate $i \in [d]$, there are at most $O(d)$ values of $\beta \in S$ such that $\lfloor n_i \rfloor_{\alpha,\beta} \ne \lfloor m_i \rfloor_{\alpha,\beta}$ so the result holds by a union bound over the d^2 coordinates.

The notion of distance that Saks and Zhou were interested in, though, is not the l^∞ norm, but the operator norm on matrices induced by the l^∞ norm on vectors. This norm can be defined in either of these two well-known equivalent ways. Let M be an $d \times d$ matrix:

$$\|M\|_{\infty-\mathrm{op}} \stackrel{\text{def}}{=} \sup \left\{ \|M\vec{x}\|_\infty : \vec{x} \in \mathbb{R}^d \text{ and } \|\vec{x}\|_\infty = 1 \right\}$$

or

$$\|M\|_{\infty-\mathrm{op}} \stackrel{\text{def}}{=} \max_{i \in [d]} \sum_{j=1}^d |M_{i,j}|.$$

If M is just viewed as the obvious vector $\vec{m}$ in $\mathbb{R}^{d^2}$, then it is easy to see using the second definition above that

$$\|\vec{m}\|_\infty \le \|M\|_{\infty-\mathrm{op}} \le d\|\vec{m}\|_\infty.$$

Thus, for any matrix N it holds with the above probability that for all matrices M within distance $\epsilon' = \cdot 2^{-t} \cdot 2^{-D} = \epsilon/d$ of N w.r.t. the operator norm M is rounded to the same value as N (they are rounded as they would be if they were viewed as d^2 length vectors).

---

[31] There is a small caveat that they consider only rounding matrices in $[0, 1]^{d \times d}$ and requiring them to be rounded to a value in $[0, 1]^{d \times d}$, but they just ensure everything is rounded down in each coordinate and then take 0 if the value was negative.

A.3   The Randomized Rounding Scheme of Goldreich

In [50, Algorithm 9], Goldreich uses the basic rounding scheme discussed above as well [32]. However, unlike Saks and Zhou, Goldreich's goals in using the partition are very relevant to our work in this paper. For an arbitrary $\epsilon \in (0, \infty)$ let $\alpha = \epsilon$. Goldreich selects β uniformly at random from the set $\left\{-(j - 0.5) \cdot \frac{\alpha}{10d^2}\right\}$ and takes $\vec{\beta}$ to be the vector of length d in which every coordinate is β (i.e. $\vec{\beta} = \langle\beta\rangle_{i=1}^d$) and then applies the function $\lfloor\cdot\rceil_{\alpha,\vec{\beta}} : \mathbb{R}^d \to \mathbb{R}^d$.

Goldreich shows that this randomized rounding scheme has the following property: For any point $\vec{x} \in \mathbb{R}^d$, there is a set $S_{\vec{x}}$ of cardinality at most $d + 1$ such that with high probability (at least $1 - \frac{1}{d+3}$ for $d > 12$) over the choice of β, it will hold that

$$\forall \vec{y} \in B_\epsilon(\vec{x}) : \lfloor\vec{y}\rceil_{\alpha,\vec{\beta}} \in S_{\vec{x}}$$

where the ϵ-ball is with respect to l^∞/d_{max}. In other words, Goldreich shows that for these parameters of the basic grid rounding scheme, for any ϵ-ball, there is a set of $d + 1$ members of the induced partition, and it will hold with high probability that that ball intersects no member other than these. We emphasize the order of quantifiers—for any ball there is a high probability that this occurs, but there is 0 probability that this occurs for all balls simultaneously because no matter which β is chosen, the induced partition is a grid, so the ϵ-ball at the corner of member will intersect 2^d different members.

Our work in this paper shows that this property that Goldreich desires can be achieved with a deterministic rounding scheme and that the randomness is not required. In other words, in Section 7, we construct a partition in each dimension $\mathbb{R}^d$ (which gives a deterministic rounding scheme) such that *every* ball of an appropriate radius ϵ intersects at most $d + 1$ members of the partition.

A.4   The Deterministic Rounding Scheme of Hoza and Klivans

In [6, Section 2], Hoza and Klivans have the same goal as Goldreich—ensuring that for any ϵ-ball, there are very few values that all points in that ball are rounded to. However, Hoza and Klivans do this with a deterministic rounding scheme, and the induced partition of this rounding scheme has quite good parameters regarding our motivating question. The analysis of the rounding scheme in their paper is somewhat obscured by other technical aspects that were relevant to other ideas they were discussing but are not necessary for the analysis of partition. For this reason, we will present their scheme here doing our best to preserve the notation that they used (so one can compare our presentation with their paper if desired) while also casting it in a way that is consistent with the perspective we take; we will then prove that the induced partition of $\mathbb{R}^d$ can be scaled to a $(d + 1, \frac{1}{6(d+1)})$-secluded partition with all members having diameter at most 1.

Let $\epsilon \in (0, \infty)$ and $d \in \mathbb{N}$. Let $\mathcal{I}$ be a partition of $\mathbb{R}$ by intervals of length $2\epsilon(d+1)$ which are closed on the left and open on the right. Fix an arbitrary point $x \in \mathbb{R}$ and consider the interval $[x - \epsilon, x + \epsilon] = x + [-\epsilon, \epsilon]$. Because this interval has length 2ϵ (and is closed) and every interval in $\mathcal{I}$ has length $2\epsilon(d + 1)$ (and is half open), it follows that there is exactly one value $\Delta \in [d+1]$ such that the interval $[x + (2\epsilon\Delta) - \epsilon, x + (2\epsilon\Delta) + \epsilon] = x + 2\epsilon\Delta + [-\epsilon, \epsilon]$ intersects two intervals in $\mathcal{I}$, and for every other $\Delta \in [d + 1]$, this interval is a subset of some interval of $\mathcal{I}$ (which interval that is may depend on Δ) [33].

---

[32]Goldreich uses t to denote the dimension that we refer to as d, uses ϵ to denote what we call α, and uses τ to denote what we call β. Further, Goldreich is is proving the property we are about to discuss in the context of learning the averages of t-many functions which is a detail showing up in the proof that is not needed for how we will state this property.

[33]The sketch of the reason for this is that

$$\bigcup_{\Delta \in [d+1]} [x + (2\epsilon\Delta) - \epsilon, x + (2\epsilon\Delta) + \epsilon] = [x + \epsilon, x + 2(d+1)\epsilon + \epsilon]$$

which is a closed interval of length $2\epsilon\Delta$ and thus intersects exactly two intervals of $\mathcal{I}$, say $I_{left} = [\alpha - 2\epsilon(d + 1), \alpha)$ and

Now consider the partition $\mathcal{G}$ of $\mathbb{R}^d$ where each member is a d-fold product of intervals of $\mathcal{I}$. That is,

$$\mathcal{G} \overset{\text{def}}{=} \left\{ \prod_{i=1}^{d} I^{(i)} : I^{(1)}, \ldots, I^{(d)} \in \mathcal{I} \right\}.$$

Each member of $\mathcal{G}$ is a hypercube, and up to translation, the set of centers of these hypercubes is $2\epsilon\mathbb{Z}^d$ (i.e. $\mathcal{G}$ should be interpreted as a grid of hypercubes). Let $\vec{1}$ denote the vector such that every entry is a 1, and define $\Lambda = \left\{ (2\epsilon\Delta) \cdot \vec{1} : \Delta \in [d+1] \right\}$. We claim that for any point $\vec{x} \in \mathbb{R}^d$, there exists at least one $\vec{\lambda} \in \Lambda$ such that $\overline{B}_\epsilon(\vec{x} + \vec{\lambda})$ is a subset of a member of $\mathcal{G}$ (intuitively, $\vec{x}$ can be shifted by one of these values, so that it is ϵ-far into the interior of some member). This is because $\overline{B}_\epsilon(\vec{x} + \vec{\lambda})$ (which is a hypercube) is a subset of a member of $\mathcal{G}$ (all of which are hypercubes) if and only if for all $i \in [d]$ it holds that $[x_i + \lambda_i - \epsilon, x_i + \lambda_i + \epsilon] = [x_i + (2\epsilon\Delta) - \epsilon, x_i + (2\epsilon\Delta) + \epsilon]$ is a subset of some member of $\mathcal{G}$. By what we showed, for each coordinate $i \in [d]$, there is exactly one $\Delta \in [d+1]$ (and thus one $\vec{\lambda} \in \Lambda$) such that this does not hold in coordinate i, and so there are at most d-many $\vec{\lambda}$'s for which this does not hold on some coordinate. Thus, there must be at least one $\vec{\lambda} \in \Lambda$ (i.e. at least one $\Delta \in [d+1]$) for which the containment holds for all coordinates $i \in [d]$.

With these properties established, let $s : \mathbb{R}^d \to \Lambda$ be a function mapping each point $\vec{x}$ to one of the $\vec{\lambda} \in \Lambda$ that has the containment property above (e.g. take the smallest length $\vec{\lambda}$ that works). Also, define the representative function rep $: \mathcal{G} \to \mathbb{R}^d$ so that $\text{rep}(X)$ is the midpoint of the hypercube X. Then, the deterministic rounding scheme of Hoza and Klivans is the function $f : \mathbb{R}^d \to \mathbb{R}^d$ defined by $f(\vec{x}) = \text{rep}(\text{member}_{\mathcal{G}}(\vec{x} + s(\vec{x})))$ (conceptually, $\vec{x}$ is rounded by first shifting $\vec{x}$ by some amount $\vec{\lambda}$ so that it is ϵ-far in the interior of some member of the partition $\mathcal{G}$, and then returning the center point of that member.

The partition induced by the Hoza-Klivans rounding scheme in $\mathbb{R}^2$ is shown in Figure 3b.

Proposition A.1. *The partition induced by the rounding scheme described above has the following properties:*

- *Each member has diameter (in the d_{max} metric) at most $6\epsilon(d+1)$*
- *The partition is $(d+1, \epsilon)$-secluded*

If this partition is scaled by a factor of $6\epsilon(d+1)$, then it trivially becomes a partition in which all members have diameter at most 1, and it is $(d+1, \frac{1}{6(d+1)})$-secluded. This result was stated in the paper and it restated here.

Theorem 2.2 (Hoza-Klivans Partition). *Let $d \in \mathbb{N}$. Then there exists a $(d+1, \frac{1}{6(d+1)})$-secluded partition $\mathcal{P}$ of $\mathbb{R}^d$ with diameter at most 1.*

To prove this result, we will abstract this rounding scheme slightly and prove a corresponding version of the result so as to highlight the essential components of this deterministic rounding scheme if one wished to generalize it. In the statement of the following lemma, the only notation change is that $\mathcal{P}_0$ can be interpretted as indicting $\mathcal{G}$.

Lemma A.2. *Let $d \in \mathbb{N}$ and $\epsilon, D \in (0, \infty)$. Let $\mathcal{P}_0$ be a partition of $\mathbb{R}^d$ such that all members have diameter at most D. Let rep $: \mathcal{P}_0 \to \mathbb{R}^d$ be a function such that $\text{rep}(X) \in X$ (conceptually, this function defines a unique representative for each member of the partition). Let Λ be a finite set of vectors in $\mathbb{R}^d$ (conceptually a finite set of possible shifts). Let $s : \mathbb{R}^d \to \Lambda$ be a function such that $\overline{B}_\epsilon(\vec{x} + s(\vec{x}))$ is a subset*

$I_{right} = [\alpha, \alpha + 2\epsilon(d+1))$. The point α is either contained in the interior or right boundary of $[x + (2\epsilon\Delta_0) - \epsilon, x + (2\epsilon\Delta_0) + \epsilon]$ for some Δ_0 (if not, then $\bigcup_{\Delta \in [d+1]}[x + (2\epsilon\Delta) - \epsilon, x + (2\epsilon\Delta) + \epsilon]$ would not intersect I_{left}). This is the unique Δ_0 such that the interval intersects both I_{left} and I_{right}.
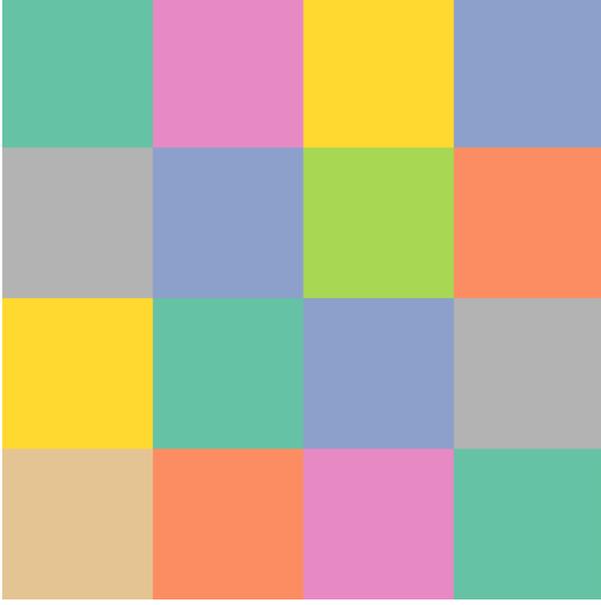
(a) The grid partition $\mathcal{G}$ for $\mathbb{R}^2$.
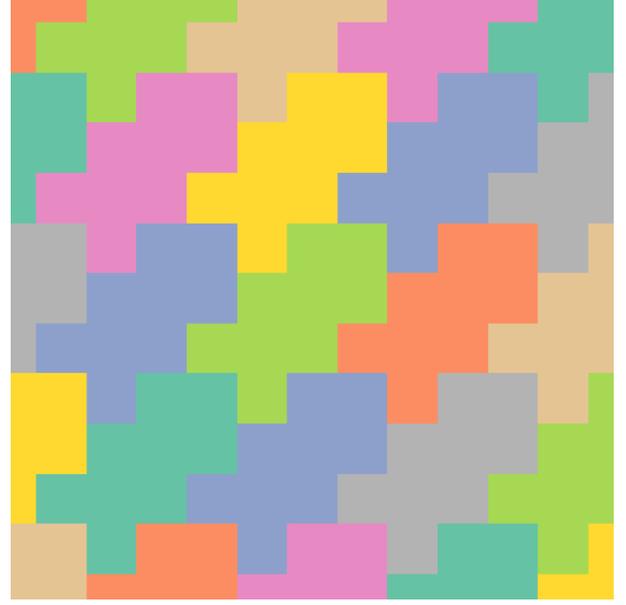
(b) The Hoza-Klivans partition of $\mathbb{R}^2$.

Figure 3: The partitions described in Section A.4. $\mathcal{G}$ is given arbitrary colors for reference (there are duplicated colors because otherwise colors were hard to distinguish). The color of a member of the Hoza-Klivans partition indicates that all points in the member are rounded to the center of the cube in the grid partition $\mathcal{G}$ of the same color. Note that the Hoza-Klivans partition has the same grid structure as $\mathcal{G}$ despite the members no longer being cubes. Also note, that at most 3 members of the Hoza-Klivans partition meet at a point.

of some member [34] *of $\mathcal{P}_0$. Let $\ell = \max_{\vec{\lambda} \in \Lambda} \|\lambda\|_\infty$ (the maximum length of a shift). Let $f : \mathbb{R}^d \to \mathbb{R}^d$ denote the function (i.e. deterministic rounding scheme) defined by $f(\vec{x}) = \mathrm{rep}(\mathrm{member}_{\mathcal{P}_0}(\vec{x} + s(\vec{x})))$.*

Then the partition induced by the deterministic rounding scheme f is $(|\Lambda|, \epsilon)$-secluded and has members of diameter at most $2\ell + D$.

Once this is proven, Proposition A.1 follows as an immediate corollary since in the initial partition $\mathcal{P}_0 = \mathcal{G}$, all members have diameter $D = 2\epsilon(d+1)$, and the longest vector $\vec{\lambda} \in \Lambda$ has length $\ell = 2\epsilon(d+1)$ (in the l^∞ norm), and $|\Lambda| = d + 1$.

Proof. Let $\mathcal{P}$ denote the partition induced by the deterministic rounding scheme f. We first show that $\mathcal{P}$ is $(|\Lambda|, \epsilon)$-secluded (i.e. that for any $\vec{p} \in \mathbb{R}^d$ it holds that $|\mathcal{N}_\epsilon(\vec{p})| \leq |\Lambda|$) [35]. Let $\vec{p} \in \mathbb{R}^d$ be arbitrary. For any $\vec{x} \in \overline{B}_\epsilon(\vec{p})$, since the d_{max} metric arises from a norm, it follows that $\vec{p} + s(\vec{x}) \in \overline{B}_\epsilon(\vec{x} + s(\vec{x}))$ [36]. Since $\overline{B}_\epsilon(\vec{x}+s(\vec{x})) \subseteq \mathrm{member}(\vec{x}+s(\vec{x}))$ (by our requirements on s), it then follows that $\vec{p}+s(\vec{x}) \in \mathrm{member}(\vec{x}+s(\vec{x}))$

---

[34] In particular, because $\vec{x} + s(\vec{x}) \in \overline{B}_\epsilon(\vec{x} + s(\vec{x}))$ the member of $\mathcal{P}_0$ containing this ball must also contain $\vec{x} + s(\vec{x})$, and so this member must be $\mathrm{member}_{\mathcal{P}_0}(\vec{x} + s(\vec{x}))$.

[35] The neighborhood notation $\mathcal{N}_\epsilon(\vec{p})$ throughout this proof is always relative to the partition $\mathcal{P}$ and never the partition $\mathcal{P}_0$.

[36]

$$
\begin{aligned}
\vec{x} \in \overline{B}_\epsilon(\vec{p}) &\iff d_{max}(\vec{x}, \vec{p}) \leq \epsilon \\
&\iff \|\vec{x} - \vec{p}\|_\infty \leq \epsilon \\
&\iff \|(\vec{x} + s(\vec{x})) - (\vec{p} + s(\vec{x}))\|_\infty \leq \epsilon \\
&\iff \vec{p} + s(\vec{x}) \in \overline{B}_\epsilon(\vec{x} + s(\vec{x}))
\end{aligned}
$$

and so $\mathrm{member}(\vec{p} + s(\vec{x})) = \mathrm{member}(\vec{x} + s(\vec{x}))$. This allows us to show as follows that f takes on at most $|\Lambda|$ values on the set $\overline{B}_\epsilon(\vec{p})$:

$$
\begin{aligned}
\big\{ f(\vec{x}) : \vec{x} \in \overline{B}_\epsilon(\vec{p}) \big\} &= \big\{ \mathrm{rep}(\mathrm{member}_{\mathcal{P}_0}(\vec{x} + s(\vec{x}))) : \vec{x} \in \overline{B}_\epsilon(\vec{p}) \big\} && \text{(Def'n of } f) \\
&= \big\{ \mathrm{rep}(\mathrm{member}_{\mathcal{P}_0}(\vec{p} + s(\vec{x}))) : \vec{x} \in \overline{B}_\epsilon(\vec{p}) \big\} \\
&\hspace{3em} (\mathrm{member}_{\mathcal{P}_0}(\vec{p} + s(\vec{x})) = \mathrm{member}_{\mathcal{P}_0}(\vec{x} + s(\vec{x}))) \\
&\subseteq \big\{ \mathrm{rep}(\mathrm{member}_{\mathcal{P}_0}(\vec{p} + \vec{\lambda})) : \vec{\lambda} \in \Lambda \big\} && (s(\vec{x}) \in \Lambda)
\end{aligned}
$$

The latter set clearly has cardinality at most $|\Lambda|$ because $\mathrm{rep}(\mathrm{member}(\vec{p} + \vec{\lambda}))$ is a mapping of the elements of Λ. This is morally why the induced partition $\mathcal{P}$ has the property $|\mathcal{N}_\epsilon(\vec{p})| \le |\Lambda|$; the following formalizes this, but the intuition of the above is somewhat lost in the notation.

$$
\begin{aligned}
|\mathcal{N}_\epsilon(\vec{p})| &= |\big\{ \mathrm{member}_{\mathcal{P}}(\vec{x}) : \vec{x} \in \overline{B}_\epsilon(\vec{p}) \big\}| && \text{(Lemma 5.1)} \\
&= |\big\{ f^{-1}(f(\vec{x})) : \vec{x} \in \overline{B}_\epsilon(\vec{p}) \big\}| && \text{(Def'n of the induced partition } \mathcal{P}) \\
&= |\big\{ f^{-1}(\vec{w}) : \vec{w} \in \{ f(\vec{x}) : \vec{x} \in \overline{B}_\epsilon(\vec{p}) \} \big\}| && \text{(Reformat)} \\
&\le |\big\{ f^{-1}(\vec{w}) : \vec{w} \in \big\{ \mathrm{rep}(\mathrm{member}_{\mathcal{P}_0}(\vec{p} + \vec{\lambda})) : \vec{\lambda} \in \Lambda \big\} \big\}| && \text{(Prior paragraph)} \\
&\le |\big\{ \mathrm{rep}(\mathrm{member}_{\mathcal{P}_0}(\vec{p} + \vec{\lambda})) : \vec{\lambda} \in \Lambda \big\}| && (f^{-1} \text{ is a function)} \\
&\le |\Lambda| && \text{(Prior paragraph)}
\end{aligned}
$$

We next show that every member of $\mathcal{P}$ has diameter at most $2\ell + D$. Let $X \in \mathcal{P}$ be arbitrary and let $\vec{x}, \vec{y} \in X$. This means that $f(\vec{x}) = f(\vec{y})$ so by definition of f, this means $\mathrm{rep}(\mathrm{member}_{\mathcal{P}_0}(\vec{x} + s(\vec{x})) = \mathrm{rep}(\mathrm{member}_{\mathcal{P}_0}(\vec{y} + s(\vec{y}))$. By definition of rep, the left hand side is contained in $\mathrm{member}_{\mathcal{P}_0}(\vec{x} + s(\vec{x}))$ and the right hand side is contained in $\mathrm{member}_{\mathcal{P}_0}(\vec{y} + s(\vec{y}))$, and since the left and right hand side are the same point, it must be that $\mathrm{member}_{\mathcal{P}_0}(\vec{x} + s(\vec{x})) = \mathrm{member}_{\mathcal{P}_0}(\vec{y} + s(\vec{y}))$, and because members of $\mathcal{P}_0$ have diameter at most D, it follows that $d_{max}(\vec{x} + s(\vec{x}), \vec{y} + s(\vec{y})) \le D$. Now observe that

$$
\begin{aligned}
d_{max}(\vec{x}, \vec{y}) &\le d_{max}(\vec{x}, \vec{x} + s(\vec{x})) + d_{max}(\vec{x} + s(\vec{x}), \vec{y} + s(\vec{y})) + d_{max}(\vec{x}, \vec{x} + s(\vec{x})) \\
&\le \|s(\vec{x})\|_\infty + D + \|s(\vec{y})\|_\infty \\
&\le \ell + D + \ell && (s(\vec{x}), s(\vec{y}) \in \Lambda)
\end{aligned}
$$

so $\mathrm{diam}(X) \le 2\ell + D$. $\qquad\square$

B Measure Theory

Throughout this section, by "countable" we mean finite or countably infinite.

Fact B.1. *If m is a measure and $\mathcal{A}$ is a (possibly uncountable) family of pairwise disjoint measurable sets, then*

$$
m(\bigsqcup_{A \in \mathcal{A}} A) \ge \sum_{A \in \mathcal{A}} m(A).
$$

Proof. By definition of the arbitrary summation (c.f. [51, p. 11]) we have

$$
\sum_{A \in \mathcal{A}} m(A) \overset{\text{def}}{=} \sup \left\{ \sum_{A \in \mathcal{F}} : \mathcal{F} \subseteq \mathcal{A}, \ \mathcal{F} \text{ finite} \right\}
$$

and for any $\mathcal{F} \subseteq \mathcal{A}$ we have

$$m(\bigsqcup_{A \in \mathcal{A}} A) \geq m(\bigsqcup_{A \in \mathcal{F}} A) = \sum_{A \in \mathcal{F}} m(A).$$

Thus $m(\bigsqcup_{A \in \mathcal{A}} A)$ is an upper bound for the set $\{\sum_{A \in \mathcal{F}} : \mathcal{F} \subseteq \mathcal{A}, \mathcal{F} \text{ finite}\}$ and thus greater than or equal to the supremum. $\qquad \square$

Fact B.2. *If m is a measure and $\mathcal{A}$ is a (possibly uncountable) family of pairwise disjoint measurable sets and $m(\bigsqcup_{A \in \mathcal{A}} A) < \infty$, then the set $\{A \in \mathcal{A} : m(A) > 0\}$ is countable.*

Proof. Let $\mathcal{B} = \{A \in \mathcal{A} : m(A) > 0\}$ denote the set in question, and let $\mathcal{B}_n = \{A \in \mathcal{A} : m(A) > \frac{1}{n}\}$ so that $\mathcal{B} = \bigcup_{n=1}^{\infty} \mathcal{B}_n$. Clearly each $\mathcal{B}_n$ is finite since

$$\infty > m(\bigsqcup_{A \in \mathcal{A}} A) \geq m(\bigsqcup_{A \in \mathcal{B}_n} A) \geq \sum_{A \in \mathcal{B}_n} m(A) \geq \sum_{A \in \mathcal{B}_n} \frac{1}{n}$$

and n is independent of the summation.

Thus $\mathcal{B}$ is a countable union of finite families, so $\mathcal{B}$ is countable. $\qquad \square$

Fact B.3. *If $\mathcal{P}$ is a partition of $\mathbb{R}^d$, and m is the Lebesgue measure on R^d, and for all $X \in \mathcal{P}$, X is measurable and $m(x) > 0$, then $\mathcal{P}$ is countable.*

Proof. We first show that for any $n \in \mathbb{N}$, the set $\mathcal{A}_n = \{X \cap B_n(\vec{0}) : X \in \mathcal{P}, m(X \cap B_n(\vec{0})) > 0\}$ is countable. Observe that $\mathcal{A}_n$ is pairwise disjoint and $\bigsqcup_{A \in \mathcal{A}} A \subseteq B_n(\vec{0})$ so $\infty > m(B_n(\vec{0})) \geq m(\bigsqcup_{A \in \mathcal{A}} A)$, so by the previous result, $\mathcal{A}_n$ is countable. Observe that $\mathcal{A}_n$ has the same cardinality as $\mathcal{P}_n = \{X \in \mathcal{P} : m(S \cap B_n(\vec{0})) > 0\}$ (it is easy to inject $\mathcal{P}_n$ into $\mathcal{A}_n$ via intersection with $B_n(\vec{0})$, and it is easy to inject $\mathcal{A}_n$ into $\mathcal{P}_n$ by mapping A to the unique member of $\mathcal{P}_n$ containing A). Thus $\mathcal{P}_n$ is countable.

Clearly $\mathcal{P} \subseteq \bigcup_{n=1}^{\infty} \mathcal{P}_n$, and we also get the other inclusion because for any $X \in \mathcal{P}$ there is some $n \in \mathbb{N}$ such that $m(X \cap B_n(\vec{0})) > 0$ (since $0 < m(X) = m(\bigcup_{n=1}^{\infty} (X \cap B_n(\vec{0}))) \leq \sum_{n=1}^{\infty} m(X \cap B_n(\vec{0}))$ so some term on the right must be positive). Thus $\mathcal{P}$ is a countable union of countable families, so $\mathcal{P}$ is countable. $\qquad \square$

Note that the above proof can be easily generalized from $\mathbb{R}^d$ to any (non-empty) σ-finite measure space by replacing the $B_n(\vec{0})$ with a σ-decomposition of the space.

C Binary Relations

Let R denote a binary relation on a set X, and let R^t denote the transitive closure and let R^{-1} denote the inverse relation $R^{-1} = \{(a, b) \in X^2 : (b, a) \in R\}$. The following are easily verified:

- R is symmetric if and only if $R = R^{-1}$.
- R is transitive if and only if R^{-1} is transitive.
- For another binary relation S, we have $R \subseteq S$ if and only if $R^{-1} \subseteq S^{-1}$.
- If $\mathcal{S}$ is a collection of relations, then $\bigcap_{S \in \mathcal{S}} S^{-1} = (\bigcap_{S \in \mathcal{S}} S)^{-1}$.

From these it follows that $(R^{-1})^t = (R^t)^{-1}$ as shown below.

$$(R^{-1})^t = \bigcap_{\substack{S \subset X^2 \\ S \text{ transitive} \\ S \supseteq R^{-1}}} S \qquad \text{(Common alternate definition of transitive closure)}$$

$$= \bigcap_{\substack{S \subset X^2 \\ S^{-1} \text{ transitive} \\ S^{-1} \supseteq R}} S \qquad \text{(Inverse preserves transitivity and subsets)}$$

$$= \bigcap_{\substack{T \subset X^2 \\ T \text{ transitive} \\ T \supseteq R}} T^{-1}$$

$$= \left(\bigcap_{\substack{T \subset X^2 \\ T \text{ transitive} \\ T \supseteq R}} T \right)^{-1} \qquad \text{(Inverse preserves intersections)}$$

$$= (R^t)^{-1} \qquad \text{(Common alternate definition of transitive closure)}$$

Fact C.1. *If R is a reflexive and symmetric relation on X, then R^t is an equivalence relation.*

Proof. Since R is reflexive, we have that for all $a \in X$, $(a, a) \in R \subseteq R^t$, so R^t is reflexive. Since R is symmetric, we have that $R = R^{-1}$, so $R^t = (R^{-1})^t = (R^t)^{-1}$ which impies that R^t is symmetric since it is equal to its inverse. That R^t is transitive follows from the definition of transitive closure. Thus R^t is an equivalence relation. $\qquad \square$

Fact C.2. *If R, S are equivalence relations on X, and $R \subseteq S$, then each equivalence class of R is a subset of some equivalence class of S.*

Proof. Let E_R denote an arbitrary equivalence class of R. Then E_R contains some $x \in X$, and we denote E_R using the standard notation $[x]_R$ which is the equivalence class containing x. We will show that $[x]_R$ is a subset of $[x]_S$. Let $y \in [x]_R$ be arbitrary. Then $(x, y) \in R \subseteq S$ which implies $(x, y) \in S$ and thus $y \in [x]_S$. $\qquad \square$

Still letting R denote a binary relation on X, let $R^{t_0} = R$, and inductively for all $n \in \mathbb{N}^+$, let $R^{t_n} = R^{t_{n-1}} \cup \{(x, y) \in X^2 : \exists z \in X \text{ with } (x, z) \in R^{t_{n-1}} \text{ and } (z, y) \in R^{t_{n-1}}\}$.

Fact C.3. *If R is a binary relation on X, then $R^t = \bigcup_{n=0}^{\infty} R^{t_n}$.*

Proof. To show that $R^t \subseteq \bigcup_{n=0}^{\infty} R^{t_n}$ it suffices to show that $\bigcup_{n=0}^{\infty} R^{t_n}$ is transitive. First note that for any $n \in \mathbb{N}_0$, $R^{t_n} \subseteq R^{t_{n+1}}$. Let $(a, b), (b, c) \in \bigcup_{n=0}^{\infty} R^{t_n}$; then there is some N such that $(a, b), (b, c) \in R^{t_N}$ which means that $(a, c) \in R^{t_{N+1}}$ and so $\bigcup_{n=0}^{\infty} R^{t_n}$ is transitive.

For the other containment, for an inductive base case note that $R^{t_0} = R \subseteq R^t$. Then for the inductive case, if $R^{t_n} \subseteq R^t$ for some n, then because R^t is transitive it follows that

$$\{(x, y) \in X^2 : \exists z \in X \text{ with } (x, z) \in R^{t_n} \text{ and } (z, y) \in R^{t_n}\} \subseteq R^t$$

and thus $R^{t_{n+1}} \subseteq R^t$. Thus $\bigcup_{n=0}^{\infty} R^{t_n} \subseteq R^t$. $\qquad \square$

Fact C.4. *Let $(a, b) \in X^2$, then $(a, b) \in R^{t_n}$ if and only if there exists $0 < k \leq 2^n$ and there exists a sequence $\langle x_i \rangle_{i=0}^{k}$ with $x_0 = a$ and $x_k = b$ and for all $i \in [k]$, $(x_{i-1}, x_i) \in R$.*

Proof. The case $n = 0$ is trivial and serves as an inductive base case. For induction, assume the statement for $n-1$. For the forward direction, if $(a, b) \in R^{t_n}$ then either $(a, b) \in R^{t_{n-1}}$ and the required sequence exists

by IH, or $(a,b) \in \{(x,y) \in X^2 : \exists z \in X \text{ with } (x,z) \in R^{t_{n-1}} \text{ and } (z,y) \in R^{t_{n-1}}\}$ and thus there exists $c \in X$ such that $(a,c), (c,b) \in R^{t_{n-1}}$ so by IH, there exists $0 < k', k'' \leq 2^{n-1}$ and sequences $\langle x_i \rangle_{i=0}^{k'}$ and $\langle y_0 \rangle_{i=1}^{k''}$ with $x_0 = a$, $x_{k'} = c = y_0$, and $y_{k''} = b$, and thus pasting the sequences together as $\langle z_i \rangle_{i=0}^{k'+k''}$ with $z_i = \begin{cases} x_i & 0 \leq i \leq k' \\ y_{i-k'} & k' \leq i \leq k' + k'' \end{cases}$ is a sequence with $0 < k = k' + k'' \leq 2^n$ and $z_i = a$ and $z_{k=k'+k''} = b$.

For the reverse direction, if a sequence $\langle x_i \rangle_{i=0}^k$ exists with $x_0 = a$, $x_k = b$, and $0 < k \leq 2^n$, then either $k = 1$ and we are done (because then $(a,b) \in R$) or $k > 1$ in which case we let $k' = \lceil k/2 \rceil$ and $k'' = \lfloor k/2 \rfloor$ so that $k' + k'' = k$ and $0 < k', k'' \leq 2^{n-1}$, so by inductive hypothesis, the sequence $\langle x_i \rangle_{i=0}^{k'}$ demonstrates that $(x_0, x_{k'}) \in R^{t_{n-1}}$ and the sequence $\langle y_i \rangle_{i=k'+0}^{k'+k''=k}$ demonstrates that $(x_{k'}, x_k) \in R^{t_{n-1}}$ and thus $(a,b) = (x_0, x_k) \in R^{t_n}$. $\qquad\square$

Fact C.5. *For any $a, b \in X$, $(a,b) \in R^t$ if and only if there exists some $N \in \mathbb{N}$ and some sequence $\langle x_i \rangle_{i=0}^N$ with $x_0 = a$, and $x_N = b$, and for all $i \in [N]$ $(x_{i-1}, x_i) \in R$.*

Proof. If $(a,b) \in R^t$, then $(a,b) \in R^{t_n}$ for some $n \in \mathbb{N} \cup \{0\}$, so by the prior fact there exists some $0 < N \leq 2^n$ for which a sequence as described exists. Conversely, if such a sequence $\langle x_i \rangle_{i=0}^N$ exists then $(a,b) \in R^{t_n} \subseteq R^t$. $\qquad\square$

References

[1] M. Saks and S. Zhou, "BP $_\mathrm{h}$space(s) subseteq dspace(s$^{3/2}$)," *J. Comput. Syst. Sci.*, vol. 58, no. 2, pp. 376–403, 1999.

[2] E. Gat and S. Goldwasser, "Probabilistic Search Algorithms with Unique Answers and Their Cryptographic Applications," Tech. Rep. 136, 2011. [Online]. Available: https://eccc.weizmann.ac.il/report/2011/136/

[3] O. Goldreich, "Multi-pseudodeterministic algorithms," *Electronic Colloquium on Computational Complexity (ECCC)*, vol. 26, p. 12, 2019.

[4] O. Grossman and Y. P. Liu, "Reproducibility and pseudo-determinism in log-space," in *Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2019, San Diego, California, USA, January 6-9, 2019.* SIAM, 2019, pp. 606–620.

[5] R. Impagliazzo, R. Lei, T. Pitassi, and J. Sorrell, "Reproducibility in Learning," *arXiv:2201.08430 [cs]*, Jan. 2022, arXiv: 2201.08430. [Online]. Available: http://arxiv.org/abs/2201.08430

[6] W. M. Hoza and A. R. Klivans, "Preserving Randomness for Adaptive Algorithms," *arXiv:1611.00783 [cs]*, Jun. 2018, arXiv: 1611.00783. [Online]. Available: http://arxiv.org/abs/1611.00783

[7] R. Impagliazzo, N. Nisan, and A. Wigderson, "Pseudorandomness for network algorithms," in *Proceedings of the twenty-sixth annual ACM symposium on Theory of computing*, 1994, pp. 356–364.

[8] U. Feige, G. Kindler, and R. O'Donnell, "Understanding parallel repetition requires understanding foams," in *22nd Annual IEEE Conference on Computational Complexity (CCC 2007), 13-16 June 2007, San Diego, California, USA.* IEEE Computer Society, 2007, pp. 179–192.

[9] G. Kindler, R. O'Donnell, A. Rao, and A. Wigderson, "Spherical Cubes and Rounding in High Dimensions," in *2008 49th Annual IEEE Symposium on Foundations of Computer Science.* Philadelphia, PA, USA: IEEE, Oct. 2008, pp. 189–198. [Online]. Available: http://ieeexplore.ieee.org/document/4690953/

[10] G. Kindler, A. Rao, R. O'Donnell, and A. Wigderson, "Spherical cubes: optimal foams from computational hardness amplification," *Communications of the ACM*, vol. 55, no. 10, pp. 90–97, Oct. 2012. [Online]. Available: https://dl.acm.org/doi/10.1145/2347736.2347757

[11] M. Braverman and D. Minzer, "Optimal tiling of the euclidean space using permutation-symmetric bodies," in *36th Computational Complexity Conference, CCC 2021, July 20-23, 2021, Toronto, Ontario, Canada (Virtual Conference)*, ser. LIPIcs, V. Kabanets, Ed., vol. 200. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2021, pp. 5:1–5:48.

[12] The Manim Community Developers, "Manim – Mathematical Animation Framework," 11 2021. [Online]. Available: https://www.manim.community/

[13] H. Minkowski, *Diophantische Approximationen*. Wiesbaden: Vieweg+Teubner Verlag, 1907. [Online]. Available: http://link.springer.com/10.1007/978-3-663-16055-7

[14] "Keller's conjecture," Aug. 2021, page Version ID: 1041468858. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Keller%27s_conjecture&oldid=1041468858

[15] "Lattice (group)," Jun. 2021, page Version ID: 1029466251. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Lattice_(group)&oldid=1029466251

[16] G. Hajós, "Über einfache und mehrfache Bedeckung desn-dimensionalen Raumes mit einem Würfelgitter," *Mathematische Zeitschrift*, vol. 47, no. 1, pp. 427–467, Dec. 1942. [Online]. Available: https://doi.org/10.1007/BF01180974

[17] O.-H. Keller, "Über die lückenlose Erfüllung des Raumes mit Würfeln." vol. 1930, no. 163, pp. 231–248, Jan. 1930, publisher: De Gruyter Section: Journal für die reine und angewandte Mathematik. [Online]. Available: https://www.degruyter.com/document/doi/10.1515/crll.1930.163.231/html

[18] O. Perron, "Über lückenlose Ausfüllung desn-dimensionalen Raumes durch kongruente Würfel," *Mathematische Zeitschrift*, vol. 46, no. 1, pp. 1–26, Dec. 1940. [Online]. Available: https://doi.org/10.1007/BF01181421

[19] ——, "Über lückenlose Ausfüllung desn-dimensionalen Raumes durch kongruente Würfel. II," *Mathematische Zeitschrift*, vol. 46, no. 1, pp. 161–180, Dec. 1940. [Online]. Available: https://doi.org/10.1007/BF01181436

[20] S. Szabó, "A reduction of Keller's conjecture," *Periodica Mathematica Hungarica*, vol. 17, no. 4, pp. 265–277, Dec. 1986. [Online]. Available: https://doi.org/10.1007/BF01848388

[21] K. Corrádi and S. Szabó, "A combinatorial approach for Keller's conjecture," 1990.

[22] J. Lagarias and P. Shor, "Keller's cube-tiling conjecture is false in high dimensions," 1992.

[23] Mackey, "A Cube Tiling of Dimension Eight with No Facesharing," *Discrete & Computational Geometry*, vol. 28, no. 2, pp. 275–279, Aug. 2002. [Online]. Available: https://doi.org/10.1007/s00454-002-2801-9

[24] J. Debroni, J. D. Eblen, M. A. Langston, W. Myrvold, P. Shor, and D. Weerapurage, "A complete resolution of the Keller maximum clique problem," in *Proceedings of the 2011 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, ser. Proceedings. Society for Industrial and Applied Mathematics, Jan. 2011, pp. 129–135. [Online]. Available: https://epubs.siam.org/doi/abs/10.1137/1.9781611973082.11

[25] A. P. Kisielewicz and M. Łysakowska, "On Keller's conjecture in dimension seven," *arXiv preprint arXiv:1401.4689*, 2014.

[26] A. P. Kisielewicz, "Towards Resolving Keller's Cube Tiling Conjecture in Dimension Seven," *arXiv:1701.07155 [math]*, Jan. 2017, arXiv: 1701.07155. [Online]. Available: http://arxiv.org/abs/1701.07155

[27] M. Łysakowska, "Extended Keller graph and its properties: Quaestiones Mathematicae: Vol 42, No 4." [Online]. Available: https://www.tandfonline.com/doi/abs/10.2989/16073606.2018.1462865?casa_token=hFqa5v9LiHUAAAAA:-M4aprzSdFNxtocgaEAxPogP69FlzQYvNhNQif5VPSeI2W8fjJaiPQLDkLmSJda1QDSB0LhQlfQYnQ

[28] J. Brakensiek, M. Heule, J. Mackey, and D. Narváez, "The Resolution of Keller's Conjecture," *arXiv:1910.03740 [cs, math]*, Dec. 2020, arXiv: 1910.03740. [Online]. Available: http://arxiv.org/abs/1910.03740

[29] P. Furtwängler, "Über Gitter konstanter Dichte," *Monatshefte für Mathematik und Physik*, vol. 43, no. 1, pp. 281–288, Dec. 1936. [Online]. Available: https://doi.org/10.1007/BF01707607

[30] R. M. Robinson, "Multiple tilings ofn-dimensional space by unit cubes," *Mathematische Zeitschrift*, vol. 166, no. 3, pp. 225–264, Oct. 1979. [Online]. Available: https://doi.org/10.1007/BF01214145

[31] B. Fuglede, "Commuting self-adjoint partial differential operators and a group theoretic problem," *Journal of Functional Analysis*, vol. 16, no. 1, pp. 101–121, May 1974. [Online]. Available: https://www.sciencedirect.com/science/article/pii/002212367490072X

[32] T. Tao, "Fuglede's conjecture is false in 5 and higher dimensions," *Mathematical Research Letters*, vol. 11, no. 2, pp. 251–258, Mar. 2004, publisher: International Press of Boston. [Online]. Available: https://www.intlpress.com/site/pub/pages/journals/items/mrl/content/vols/0011/0002/a008/abstract.php

[33] J. C. Lagarias, J. A. Reeds, and Y. Wang, "Orthonormal bases of exponentials for the n-cube," *Duke Mathematical Journal*, vol. 103, no. 1, pp. 25–37, May 2000, publisher: Duke University Press. [Online]. Available: https://projecteuclid.org/journals/duke-mathematical-journal/volume-103/issue-1/Orthonormal-bases-of-exponentials-for-the-n-cube/10.1215/S0012-7094-00-10312-2.full

[34] M. Dutour, Y. Itoh, and A. Poyarkov, "Cube packings, second moment and holes," *arXiv:math/0509100*, Sep. 2018, arXiv: math/0509100. [Online]. Available: http://arxiv.org/abs/math/0509100

[35] W. Sierpiński, "Un théorème sur les continus," *Tôhoku Mathematical Journal*, pp. 300–303, 1918.

[36] A. Blass, "reference request - Why are the integers with the cofinite topology not path-connected?" [Online]. Available: https://mathoverflow.net/questions/48970/why-are-the-integers-with-the-cofinite-topology-not-path-connected

[37] J. F. Sallee, "A triangulation of the n-cube," *Discrete Mathematics*, vol. 40, no. 1, pp. 81–86, Jan. 1982. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0012365X8290190X

[38] ——, "A Note on Minimal Triangulations of an n-cube."

[39] A. Glazyrin, "Lower bounds for the simplexity of the n-cube," *Discrete Mathematics*, vol. 312, no. 24, pp. 3656–3662, Dec. 2012. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0012365X12003974

[40] C. Zong, "What is known about unit cubes," *Bulletin of The American Mathematical Society - BULL AMER MATH SOC*, vol. 42, pp. 181–212, Apr. 2005.

[41] C. Zong, B. Bollobas, W. Fulton, and A. Katok, *The Cube-A Window to Convex and Discrete Geometry*, ser. Cambridge Tracts in Mathematics. Cambridge University Press, 2006. [Online]. Available: https://books.google.com/books?id=Ola6htFUQ1IC

[42] R. Johnsonbaugh and W. E. Pfaffenberger, *Foundations of mathematical analysis*, 1981.

[43] H. Kuhn, "Some Combinatorial Lemmas in Topology."

[44] H. Komiya, "A simple proof of K-K-M-S theorem," *Economic Theory*, vol. 4, no. 3, pp. 463–466, May 1994. [Online]. Available: https://doi.org/10.1007/BF01215383

[45] J. A. d. Loera, E. Peterson, and F. E. Su, "A Polytopal Generalization Of Sperner's Lemma," *J. Combin. Theory Ser. A*, vol. 100, p. 2002, 2001.

[46] G. van der Laan, D. Talman, and Z. Yang, "Intersection theorems on polytopes," *Mathematical Programming*, vol. 84, no. 1, pp. 25–38, Jan. 1999. [Online]. Available: http://link.springer.com/10.1007/s10107980024a

[47] J. A. De Loera, E. Peterson, and F. Edward Su, "A Polytopal Generalization of Sperner's Lemma," *Journal of Combinatorial Theory, Series A*, vol. 100, no. 1, pp. 1–26, Oct. 2002. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0097316502932747

[48] A. Below, U. Brehm, J. A. De Loera, and and J. Richter-Gebert, "Minimal Simplicial Dissections and Triangulations of Convex 3-Polytopes," *Discrete & Computational Geometry*, vol. 24, no. 1, pp. 35–48, Jan. 2000. [Online]. Available: https://doi.org/10.1007/s004540010058

[49] D. Orden and F. Santos, "Asymptotically Efficient Triangulations of thed-Cube," *Discrete & Computational Geometry*, vol. 30, no. 4, pp. 509–528, Oct. 2003. [Online]. Available: https://doi.org/10.1007/s00454-003-2845-5

[50] O. Goldreich, "Multi-pseudodeterministic algorithms," p. 13.

[51] G. B. Folland, *Real Analysis: Modern Techniques and Their Applications.* John Wiley & Sons, Apr. 1999, google-Books-ID: N8jVDwAAQBAJ.