5-5-2023

Partitions of R^n with Maximal Seclusion and their Applications to Reproducible Computation

Jason Vander Woude
University of Nebraska-Lincoln, jasonvw@huskers.unl.edu

PARTITIONS OF $\mathbb{R}^n$ WITH MAXIMAL SECLUSION AND THEIR APPLICATIONS TO REPRODUCIBLE COMPUTATION

by

Jason Vander Woude

A DISSERTATION

Presented to the Faculty of

The Graduate College at the University of Nebraska

In Partial Fulfilment of Requirements

For the Degree of Doctor of Philosophy

Majors: Mathematics & Computer Science

Under the Supervision of Professors N. V. Vinodchandran & Jamie Radcliffe

Lincoln, Nebraska

May, 2023

PARTITIONS OF $\mathbb{R}^n$ WITH MAXIMAL SECLUSION
AND THEIR APPLICATIONS TO REPRODUCIBLE COMPUTATION

Jason Vander Woude, Ph.D.
University of Nebraska, 2023

Advisers: N. V. Vinodchandran & Jamie Radcliffe

We introduce and investigate a natural problem regarding unit cube tilings/partitions of Euclidean space and also consider broad generalizations of this problem. The problem fits well within a historical context of similar problems and also has applications to the study of reproducibility in randomized computation.

Given $k \in \mathbb{N}$ and $\varepsilon \in (0, \infty)$, we define a (k, ε)-secluded unit cube partition of $\mathbb{R}^d$ to be a unit cube partition of $\mathbb{R}^d$ such that for every point $\vec{p} \in \mathbb{R}^d$, the closed ℓ_∞ ε-ball around $\vec{p}$ intersects at most k cubes. The problem is to construct such partitions for each dimension d with the primary goal of minimizing k and the secondary goal of maximizing ε.

We prove that for every dimension $d \in \mathbb{N}$, there is an explicit and efficiently computable (k, ε)-secluded axis-aligned unit cube partition of $\mathbb{R}^d$ with $k = d + 1$ and $\varepsilon = \frac{1}{2d}$. We complement this construction by proving that for axis-aligned unit cube partitions, the value of $k = d + 1$ is the minimum possible, and when k is minimized at $k = d + 1$, the value $\varepsilon = \frac{1}{2d}$ is the maximum possible. This demonstrates that our constructions are the best possible.

We also consider the much broader class of partitions in which every member has at most unit volume and show that $k = d + 1$ is still the minimum possible. We also show that for any reasonable k (i.e. $k \leq 2^d$), it must be that $\varepsilon \leq \frac{\log_4(k)}{d}$. This demonstrates that when k is minimized at $k = d + 1$, our unit cube constructions

are optimal to within a logarithmic factor even for this broad class of partitions. In fact, they are even optimal in ε up to a logarithmic factor when k is allowed to be polynomial in d.

We extend the techniques used above to introduce and prove a variant of the KKM lemma, the Lebesgue covering theorem, and Sperner's lemma on the cube which says that for every $\varepsilon \in (0, \frac{1}{2}]$, and every proper coloring of $[0,1]^d$, there is a translate of the ℓ_∞ ε-ball which contains points of least $(1 + \frac{2}{3}\varepsilon)^d$ different colors.

ACKNOWLEDGMENTS

There are many people that I wish thank for the role that they have played in my education in no particular order.

I want to thank Nathan Tintle for giving me my first opportunity to do research during my junior year of undergrad and for providing funding for two academic years and three summers to work on computational and statistical genetics projects. Thank you for pushing me during that time and encouraging me to apply for the NSF GRFP. That you seriously recommended I apply for it gave me a lot of confidence in the quality of my research, and confidence to pursue graduate school. I am very appreciative that you took the time to meet every week despite being very busy with lots of other research, teaching, and administrative duties. Thank you also for the level of autonomy that you allowed me to have as that prepared me well for graduate research. You were an excellent mentor, and I am confident I would not have attended graduate school if I had not had the opportunities to experience the joys of research that you provided.

I want to thank Mike Janssen for his mathematical support in multiple ways during undergrad. It was your linear algebra course during my sophomore year that really got me interested in pure mathematics (in particular, your choice of starting with vector spaces instead of matrix computations). Thank you also for the opportunity to conduct pure mathematics research in the summer of 2017. Considering that you had another very qualified applicant (Thomas) for that research position along with the fact that I was also going to be doing research for Nathan Tintle that summer, I am extremely thankful to both you and Nathan Tintle and to whichever administrators were involved for working out the funding so that both Thomas and I could work on that mathematics project together. Thank you for your mentorship during that project and encouraging us to work toward publishing our work after the summer ended. I also have you and Tom Clark to thank for suggesting that I consider the joint math and computer science graduate program at UNL, otherwise I wouldn't be here.

I also want to thank too many people to name among the engineering faculty, computer science faculty, math faculty, and the Registrar's Office at Dordt University for putting up with me and my scheduling requests for four years of undergrad. At a small college, I wanted to double major in engineering and computer science, and after my sophomore year, I also wanted to take as many math courses as possible. Because of the fact that many classes only ran every two years and almost every class only had one section when run, this would have logistically been impossible if it weren't for many people in the aforementioned groups keeping me in the loop on sheduling, allowing me to work with professors to still take two classes one semester that overlapped one day a week, and letting me take the one course that was really difficult to schedule as an independent study. Without the efforts of all of these people, I would not have been on the same academic track in undergrad, and because engineering was my priority at that time, I don't think I would been able to pursue a graduate degree in mathematics or computer science, and I may not have explored graduate school at all.

I want to thank Jamie Radcliffe for taking me on as his student even when he was already advising four other students. I have really enjoyed working with you these past few years and taking courses from you. You have a lot of insight into many areas of mathematics and do an excellent job at effectively communicating that insight with your students. In the meetings where I didn't have a clue what you were trying to say, you were always good at rephrasing it in a different manner until it made sense to me. There were also many occasions in meetings where you pointed me in a direction that I would not have thought to look and ended up directing me down a very productive line of research.

I want to thank Vinod Variyam for taking me on as his student before I even started graduate school. I met with him during my visit to UNL and emailed him afterward expressing interest in eventually being his advisee, and to my surprise he accepted me right then. He wasted no time in getting me aquainted with various topics the field, and during my first semester, we started meeting weekly to work first through textbooks and eventually through papers. Thank you for funding me during the summer of 2019 and for the past five semesters. For me, the priveledge of focusing on nothing but research leads to huge gains which are certainly more than proportional to the increased amount of time available; without your funding, this dissertation would be only a shadow of what it is today. Thank you for your research guidance and your continual encouragement to pursue whichever research directions I found interesting. I spent the first semester of the research assistantship working on the construction results in this dissertation which was starting to diverge from your own research interests, but I found it very engaging, and I am appreciative that you allowed me to continue working on that problem instead of directing me to spend the assistantship time on questions of greater interest to you. I think we agree that this initially simple question and construction led to a much broader range of questions and many interesting results both in pure mathematics and in computer science, and I am glad that you allowed me to continue unconstrained down the paths where that initial question naturally led.

I want to thank Pavan Aduri, a collaborator of Vinod's, for weekly meetings and research guidance as well. I really valued your ability to easily translate between my more mathematical way of thinking and Vinod's more computer science theory way of thinking. I have no doubt that having you in the meetings led to faster and more effective communication of ideas between me and Vinod and led to much more productive research overall.

Last, but absolutely not least, I want to offer a very heartfelt thanks to both of my parents. From as early in school as I can remember, my mom worked diligently with me to ensure that I worked hard in school, and she was very involved in helping me with spelling words, daily and weekly assignments, and school projects—especially a select few that I remember being very difficult for me to complete. In high school, though my dad likely would have preferred that I was in his shop working with him after school so I would develop a good work ethic (like he did with his dad), during my junior and senior year of high school, my parents agreed together to let me spend the hours after school working on online college courses and also learning to program. I don't think they fully realize the impact that this had on my educational path—this allowed me to come into college with enough courses completed to be able to major in both engineering and computer science as well as take many additional math courses; without this, I would have only completed an engineering degree and almost certainly would not have attended a graduate program in math and computer science. In the same vein, having the chance to learn programming in high school was what allowed me to get my first research position in college with Nathan Tintle working on algorithms for computational genetics and to experience the joys of research. Thank you for always supporting my education and supporting my choices of what to study. Without your support, I wouldn't be here.

GRANT INFORMATION

This work was primarily supported by NSF grants 2130608 and 1934884. Additional support came from various awards through the University of Nebraska—Lincoln including the Dohrmann Fellowship (School of Computing), Outstanding First Year Student Award (Department of Mathematics), Walter Mientka Teaching Award (Department of Mathematics), David A. Klarner Fellowship (School of Computing), and the Outstanding Graduate Research and Creative Activities Award (Graduate College).

Table of Contents

List of Figures

List of Tables

List of Videos

List of Results

List of Conjectures

Notes to the Reader

Note. Throughout this dissertation, I will elect to use the noism writing style common in mathematics of using "we" and "our" instead of "I" and "my" (sometimes referred to "author's we" or "royal we"). While all of the work completed for this dissertation was done under advisement of N. V. Vinodchandran and Jamie Radcliffe along with collaborators A. Pavan and Peter Dixon, every formal result presented in this dissertation is one for which I was the primary researcher.

Note. There are some videos embedded in this dissertation. They will likely not be viewable if the PDF is opened in a web browser or some other less advanced PDF viewers. However, if the PDF is opened in Adobe Acrobat Reader (available on Windows and Mac) or Okular (available on Windows and Linux), then the videos should be viewable.

Note. As there is not really a single appropriate place to do so in the main body, we wish to credit the Manim Community Edition [The21] (as they request) which was the software used to create many of the images and videos included in this dissertation as well as the software used to create videos for multiple talks/presentations on this work.

Chapter 1

Introduction

The topic of this dissertation arises naturally in both a pure mathematics context and a theoretical computer science context, and these two perspectives will be illustrated next. The computational context provided the original motivation for this research, so we address that first.

1.1 Computational Motivation and Background

Randomized computation is a large field of study as some problems admit very simple and easy to analyze randomized algorithms which do not have known deterministic algorithms running in the same time or space bounds (e.g. the Boolean circuit acceptance probability problem[1]). Some problems have either a yes or no answer to each input (these are called formal languages), and in this case, due to randomness, any particular run of a randomized algorithm may return the wrong answer; the goal of randomized algorithms is to ensure that this does not happen

[1] In this problem one is given a representation of a Boolean circuit on n input bits and must estimate the fraction of all 2^n input strings on which the circuit "accepts" (evaluates to 1). Interpreted another way, the problem is to estimate the probability that a uniformly random input string is accepted by the circuit. In this view there is a very simple randomized algorithm which is to randomly sample input strings, evaluate the circuit on each, and output the proportion of the strings that were accepted. It is not known if this problem can be solved deterministically, and if it can it would imply that $\mathsf{BPP} = \mathsf{P}$.

often. Another type of problem that is studied (called search problems) have multiple valid solutions, and the goal is to design an algorithm that with high probability returns one of the valid solutions. An example is path-planning algorithms which attempt to find a shortest path between two vertices in a graph; the graph may have multiple shortest paths, and it would be acceptable and correct for the algorithm to return any one of these. As with all randomized algorithms, it is possible that a randomized search algorithm will provide an invalid solution due to randomness, but another issue specific to search problems also arises: the algorithm may produce different valid solutions each time the algorithm is run. For example, a randomized path planning algorithm could be run 10 times and may return a correct answer 9 of the 10 times and return 6 different valid answers over the 9 correct runs.

One approach to minimize the impact of this latter problem is to find a "pseudodeterministic" algorithm for the problem—that is, an algorithm for which there is a canonical solution for every input, and the algorithm, when given some input, returns the canonical solution for that input with high probability. This type of algorithm was first introduced in [GG11] and further studied in [GGR13]. Using the notation of the latter, a search problem is formally defined as a relation $R \subseteq X \times Y$, and for any input $x \in X$, we denote the valid solutions by $R(x) \stackrel{\text{def}}{=} \{y \in Y : (x, y) \in R\}$, and the domain of interest by $S_R \stackrel{\text{def}}{=} \{x \in X : R(x) \neq \emptyset\}$. In extremely formal contexts, $X = Y = \{0, 1\}^*$ is the set of finite strings of 0's and 1's. If A is a randomized algorithm, then the random variable defined by the response of A to input x is denoted by $A(x)$. Let $\perp \notin Y$, called "bottom", indicate some unique error-indication response that an algorithm may return. A randomized algorithm, A, is called a pseudodeterministic algorithm solving search problem R if the following holds: (1) for all inputs in the domain

$x \in S_R$, there exists a canonical solution $c_x \in R(x)$ such that $\mathbb{P}[A(x) = c_x] \geq 2/3$ and (2) for all $x \notin S_R$, $\mathbb{P}[A(x) = \perp] \geq 2/3$. In other words, for any input in the problem domain, with probability at least 2/3, the algorithm A returns the canonical solution, and for any input not in the problem domain, with probability at least 2/3, the algorithm indicates an error. By repeating algorithm A for $O(\log(1/\delta))$ times and then returning the most commonly occurring result, this probability of success can be increased from 2/3 to $1 - \delta$ for any $\delta \in (0, 1)$.

Unfortunately, pseudodeterminism is a very strong condition, and there are many natural search problems which do not have pseudodeterministic algorithms. For example, [GGR13, Theorem 4.2] exemplified simple problems where pseudodeterminism required exponential time algorithms. We give two example problems in Section 10.3 (Limitations on Learning) for which no pseudodeterministic algorithm exists, and [CMY23, Lemma 4] gave examples of similar problems.

As a partial remedy to this, Goldreich broadened the notion of pseudodeterministic algorithms to that of multi-pseudodeterministic algorithms by allowing more than one canonical solution [Gol19a]. Formally, a randomized algorithm A is called a k-pseudodeterministic algorithm solving search problem R if the following holds: (1) for all $x \in S_R$, there exists some $C_x \subseteq R(x)$ with $|C_x| \leq k$, called a canonical solution set, such that $\mathbb{P}[A(x) \in C_x] \geq \frac{k+1}{k+2}$ and (2) for all $x \notin S_R$, $\mathbb{P}[A(x) = \perp] \geq 2/3$. In other words, for any input in the problem domain, with probability at least $\frac{k+1}{k+2}$, the algorithm A returns a solution from the canonical set, and for any input not in the problem domain, with probability at least 2/3, the algorithm indicates an error[2]. When $k = 1$, this coincides with the definition of a pseudodeterministic algorithm

[2]The solution set may not be unique. For example, if an algorithm on each input x, returns y_1 with probability 1/2 and returns y_2 and y_3 each with probability 1/4, then this is a 2-pseudodeterministic algorithm, and for each x, C_x can either be $\{y_1, y_2\}$ or $\{y_1, y_3\}$.

above[3].

A special case of search problems is approximation problems: these are search problems $R \subseteq X \times Y$ where there is some notion of "closeness" on Y represented by a distance function[4] $\text{dist} : Y^2 \to [0, \infty]$ defining the distance between two points, and there is some underlying function $f : X \to Y$ which is being approximated. More formally, given some function $f : X \to Y$, and some distance function on Y, $\text{dist} : Y^2 \to [0, \infty]$, and some $\varepsilon \in [0, \infty)$, the ε-approximation problem for f is the search problem

$$R_f \stackrel{\text{def}}{=} \{(x, y) \in X \times Y : \text{dist}(f(x), y) \leq \varepsilon\} .$$

That is, it is the search problem where the domain of interest is all of X, and the valid solutions for any input, are all outputs that are at least "ε-close" to the actual function value.

One of the most natural contexts for approximation problems is when $Y = \mathbb{R}^d$ for some dimension $d \in \mathbb{N}$ and the notion of distance is the ℓ_∞ norm. The reason why this is so natural is that trying to approximate a single real value is a very natural and common problem, so having $Y = \mathbb{R}^d$ with the ℓ_∞ norm captures the idea of trying to simultaneously approximate d-many different real values where an approximation is considered ε-close if each of the d-many approximations is ε-close.

Goldreich showed in [Gol19a] that if a function $f : X \to \mathbb{R}^d$ is a function and for

[3] The reason the probability is set at $\frac{k+1}{k+2}$ is that this implies (by averaging) that there is some element of the canonical solution set with probability at least $\frac{k+1}{k(k+2)} > \frac{1}{k+1} > \frac{1}{k+2}$ and every element outside of the canonical solution set must have probability at most $\frac{1}{k+2}$ (because the rest of the probability is allocated to the canonical solution set). This is enough of a difference to be detectable in polynomial time.

[4] We specifically avoid the use of the term "metric" since there are other reasonable notions of closeness. In particular, one notion of closeness used often for approximation algorithms which is not equivalent to any metric is multiplicative distance on $(0, \infty)$ where $\text{dist}(a, b) \stackrel{\text{def}}{=} \frac{\max(a,b)}{\min(a,b)} - 1$ which expresses that a and b are distance ε apart if the larger is equal to a factor of $1 + \varepsilon$ of the smaller. The only property required of a distance function other than the codomain $[0, \infty]$ is that for any $y \in Y$, $\text{dist}(y, y) = 0$. For example, neither the triangle inequality nor symmetry is required.

every $x \in X$, the value $f(x)$ can be algorithmically approximated to tolerance ε_0 for every $\varepsilon_0 \in (0, \infty)$, then for any $\varepsilon_1 \in (0, \infty)$, $f(x)$ can be $(d+1)$-pseudodeterministically approximated to tolerance ε_1. This was a constructive result, and the algorithm consists of first approximating $f(x)$ to tolerance $\varepsilon_0 = \frac{\varepsilon_1}{10d^2}$ and then using a randomized rounding technique to convert this higher quality ε_0-approximation to a lower quality ε_1-approximation in a $(d+1)$-pseudodeterministic manner[5].

In $\mathbb{R}^1$, the strategy to do this is quite simple. Suppose $f : X \to \mathbb{R}$ is a function and for each $x \in X$, there is some method for approximating $f(x)$ to within any desired tolerance. Then to 2-pseudodeterministically ε_1-approximate a value $y = f(x) \in \mathbb{R}$, first obtain an $\varepsilon_0 = \frac{\varepsilon_1}{2}$ estimate $\hat{y}$ of y using an existing method. Second, round that estimate to the closest integer multiple of ε_1 (rounding up if $\hat{y}$ is equidistant from two integer multiples of ε_1). The reason this 2-pseudodeterministically ε_1-approximates y is explained next and demonstrated in Figure 1.1. Let $\hat{\hat{y}}$ denote the rounded value of $\hat{y}$. Then because $\hat{\hat{y}}$ is equal to the integer multiple of ε_1 which is closest to $\hat{y}$, we know that $\left|\hat{\hat{y}} - \hat{y}\right| \leq \frac{\varepsilon_1}{2}$ (i.e. half the distance between consecutive integer multiples of ε_1). Also, we know that $|\hat{y} - y| \leq \varepsilon_0 = \frac{\varepsilon_1}{2}$ because that was the specified approximation quality of the initial estimate. By the triangle inequality, we have $\left|\hat{\hat{y}} - y\right| \leq \varepsilon_1$, so the returned value, $\hat{\hat{y}}$, is indeed an ε_1-approximation to $y = f(x)$. The reason it is 2-pseudodeterministic is that regardless of what the value y is, we know that $\hat{y} \in [y - \varepsilon_0, y + \varepsilon_0]$ (by the quality of the initial estimate), and ε_0 was chosen to be

[5] This can either be seen implicitly in [Gol19a, Algorithm 3.1] by distinguishing the two implicit core pieces of the algorithm: (1) learning the initial approximation of function averages and (2) rounding that approximation. Though implicit, the rounding has nothing to do with the specific problem being solved in Algorithm 3.1.

Alternatively, this can be seen by combining Proposition 2.4 (which gives a way to 2-pseudodeterministically approximate $\mathbb{R}$-valued functions) with Theorem 3.3 (which taking $t = d$ and $m = 2$) shows how to apply the rounding technique to search problems beyond just approximation problems.

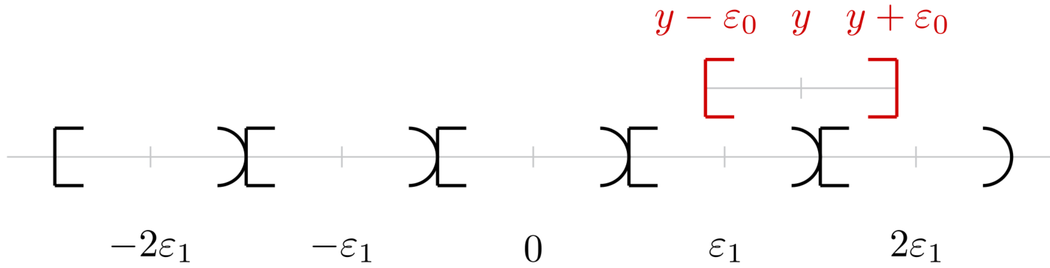


Figure 1.1: The partition of $\mathbb{R}^1$ induced by rounding every value to the nearest multiple of ε_1. For example, because points are rounded to the nearest multiple of ε_1 (rounding up for ties), the set of points that are rounded to 0 are the points in the half-open interval $[-\frac{\varepsilon_1}{2}, \frac{\varepsilon_1}{2})$. The value $\varepsilon_0 = \frac{\varepsilon_1}{2}$ is chosen to be the largest possible value such that for every $y \in \mathbb{R}$, the set of all possible ε_0-approximations of y (i.e. the closed interval $[y - \varepsilon_0, y + \varepsilon_0]$) intersects at most 2 members of the partition.

small enough that all points in this interval are rounded to at most 2 different values (see Figure 1.1). A nice view of this is that we can partition $\mathbb{R}^1$ by the equivalence classes of the rounding function—i.e. two points in $\mathbb{R}$ belong to the same member of the partition if they are both rounded to the value. Then we pick ε_0 as large as possible (so that we don't require more precise of an initial estimate than necessary) so that for every $y \in \mathbb{R}$, the set $[y - \varepsilon_0, y + \varepsilon_0]$ intersects at most 2 members of the partition (see Figure 1.1). This means, for a fixed $x \in X$, the value $y = f(x)$ is fixed, and the set of all possible ε_0-approximations of y (i.e. the closed interval $[y - \varepsilon_0, y + \varepsilon_0]$) is rounded to at most 2 different values.

Goldreich mentioned that the randomized rounding techniques employed in [Gol19a] for achieving $(d + 1)$-pseudodeterministic approximation of $f : X \to \mathbb{R}^d$ are reminiscent of the techniques of Grossman and Liu [GL19] used to estimate many possibilities using only a logarithmic amount of space. Randomized rounding is also used very famously in the work of Saks and Zhou [SZ99] in the context of derandomizing space-bounded computations. Randomized rounding was also used by Kindler, O'Donnell, Rao, and Wigderson to find an asymptotically optimal solution to the foams problem [KORW08, KROW12]. Because of the wide applicability of randomized rounding techniques, we wondered if we could derandomize the rounding function of Goldreich. The similarity of the techniques of [Gol19a] with those of [GL19] and [SZ99] indicates that our partitions may have multiple applications beyond multi-pseudodeterminism[6]. These applications might include the study of space bounded computation (as in [SZ99]), the reduction of the number of random bits needed to perform some computations (e.g. these save random bits compared to [Gol19a]), and it has been suggested to us that they may have applications in differential privacy[7].

Specifically, the question we asked was whether the randomized rounding technique of [Gol19a] could be replaced with a deterministic rounding technique of similar quality. Since it was the case in $\mathbb{R}^1$ that the partition consisted of half-open intervals of length ε_1, we wondered if we could construct partitions of $\mathbb{R}^d$ for each dimension $d \in \mathbb{N}$ consisting of translates of the half-open d-cube of side length ε_1 (i.e. translates of $[0, \varepsilon_1)^d$). In order to be able to achieve k-pseudodeterminism (for some $k \in \mathbb{N}$), this partition would have to have the property that there is some ε_0 such that for every point $\vec{y} \in \mathbb{R}^d$ it holds that $\prod_{i=1}^{d}[y_i - \varepsilon_0, y_i + \varepsilon_0]$ intersects at most k-many cubes in the partition. As before, this is because the set $\prod_{i=1}^{d}[y_i - \varepsilon_0, y_i + \varepsilon_0]$ represents all ε_0-approximations of $\vec{y}$ (in the ℓ_∞ norm), and the number of members of the partition it intersects corresponds to the number of

[6]While there is a resemblance of techniques, most of the papers and rounding techniques just mentioned have slightly different goals than each other regarding what properties the rounding function should have. Details about these papers and the properties that were desired can be found in Appendix F (Rounding Schemes in Prior Work). Nonetheless, we feel that our work will find applications beyond multi-pseudodeterminism.

[7]A reviewer of a conference submission of ours noted that the partition rounding functions can be seen as information hiding mechanisms because they hide the full information about the initial approximation $\hat{y}$ to y and reveal only which of k sets/locations $\hat{y}$ belonged. The reviewer indicated that this might be useful as a mechanism to re-use random bits.

different value these approximations are rounded to. We want to be able to, for any point $\vec{y} \in \mathbb{R}^d$, accept any ε_0-approximation to any point $\vec{y}$ and k-pseudodeterministically produce an ε_1-estimate[8].

In the context of multi-pseudodeterminism the primary goal is minimizing the value of k when designing k-pseudodeterministic algorithms. Thus, because Goldreich achieved a value of $k = d + 1$, our primary goal was to achieve this parameter in the partitions. Secondary to this goal would be to construct partitions where we could take ε_0 as large as possible because that means that our partition rounding function can tolerate the maximum possible initial error (for this type of rounding function) to still produce a final ε_1-approximation. We can always scale the partitions, so really we are not trying to maximize ε_0 but rather the ratio $\varepsilon \stackrel{\text{def}}{=} \frac{\varepsilon_0}{\varepsilon_1}$ which can be interpreted as saying, for a fixed final approximation quality ε_1, how large of an initial approximation quality ε_0 can be tolerated. Thus, for convenience, we consider fixing ε_1 to 1 so that we work only with unit cube partitions, and then ε reduces to ε_0. This line of thinking leads to the following questions.

[8] There is a slight subtlety that ε_1—the side lengths of the cubes in the partition—is not immediately the final approximation quality. The rounding process of using the partition will round each point in $\mathbb{R}^d$ to the center of the cube which contains it, so this rounds values by distance at most $\frac{\varepsilon_1}{2}$ in each coordinate, and we assume that the initial approximation is within ε_0 of the true value in each coordinate. Thus, the final rounded approximation will be at most $\frac{\varepsilon_1}{2} + \varepsilon_0$ away in each coordinate from the true value (i.e. it is at least an $(\frac{\varepsilon_1}{2} + \varepsilon_0)$-approximation to the true value in the ℓ_∞ norm). In $\mathbb{R}^1$ we could take $\varepsilon_0 = \frac{\varepsilon_1}{2}$ which guaranteed a final approximation quality of ε_1. In general dimensions, we will never end up taking $\varepsilon_0 > \frac{\varepsilon_1}{2}$, so we are in fact always guaranteed that the final approximation is at least an ε_1-approximation if we use cubes of side length ε_1; in general the final quality is guaranteed be slightly better than this, but not even by a factor of 2 more.

Question 1.1.1 (Motivating Computer Science Questions).

1. What is the minimum possible k for such partition constructions?
2. When k is minimized, what is the maximum possible ratio $\varepsilon \stackrel{\text{def}}{=} \frac{\varepsilon_0}{\varepsilon_1}$?
3. Can $k = d + 1$ be achieved so that we match the multi-pseudodeterminism parameter of Goldreich's randomized rounding technique?
4. If $k = d + 1$ is achievable, is $\varepsilon \in O(\frac{1}{d^2})$ simultaneously achievable so that we also match the approximation error tolerance of Goldreich's randomized rounding technique?
5. How is the trade-off between k and ε characterized? For example, is it possible to achieve much larger ε if we only impose that k is polynomial in the dimension d?
6. What can be said about deterministic rounding functions that are not based on cube partitions?

Following the mathematical motivation given next, we complete the introduction in Section 1.3 (Summary of Results) with a list of all main results of this dissertation which, among other things, includes nearly exact answers to every question above (i.e. we resolve k exactly, resolve ε exactly for cube partitions, and resolve both ε and the trade-off between k and ε up to a logarithmic factor in what is probably the most general setting one could wish for from a computational (or even pure mathematical) point of view).

While the motivation laid out so far is constructive motivation (i.e. asking if we can design such partition-based rounding functions), from the perspective of a computer scientist, the most important results of this dissertation are probably not the constructions but rather the very general impossibility results that we prove. The reason for this is that about a year after completing our partition constructions, it was brought to our attention by an anonymous reviewer that Hoza and Klivans [HK18] had previously designed similar deterministic rounding functions—though they did so implicitly for a specific problem and did not draw much attention to the fact that they had accomplished this more general feat.

Though the parameters achieved with the Hoza and Klivans rounding scheme asymptotically match ours, our parameters are slightly better[9]. An additional distinction is that our rounding function is based on unit cubes, so it has a very nice geometric interpretation whereas theirs was not defined based on a partition (though it does naturally induce one; see Figure F.1b). However, the definition of their rounding function is much simpler than ours. It is interesting that we have both independently come across very different methods for obtaining similarly parameterized rounding schemes—ours more beautiful from the geometric perspective and theirs with a simple elegance from the rounding perspective.

We discuss in Section 4.4 (New Partitions From Old) a way to use either our partition constructions or theirs to obtain new constructions that make trade-offs between the k and ε parameters, so in this way our family of constructions is much broader than theirs. We have also shown (Proposition 10.2.14 and Proposition 10.2.15) that up to a logarithmic factor, the rounding schemes from both us and from Hoza and Klivans are optimal in the ε parameter even under many generous allowances (e.g. this holds regardless of the norm used, regardless of how the deterministic algorithm is defined, regardless of efficient computability concerns, and regardless of whether k is minimized or allowed to be as large as any polynomial in the dimension).

[9] Their original proof is mixed in with many other details of their paper that are not relevant to us, so we include our reconstruction of their proof in Section F.4 (The Deterministic Rounding Scheme of Hoza and Klivans).

1.2 Mathematical Motivation and Background

The above question about the existence of partitions can also be taken as a purely mathematical pursuit independent of the context out of which it initially arose. That is, we ask the following natural geometric question: given $k \in \mathbb{N}$ and $\varepsilon \in (0, \infty)$, is there a partition of $\mathbb{R}^d$ consisting of axis-aligned unit cubes (i.e. sets of the form $\vec{\alpha} + [0,1)^d$) so that for any point $\vec{p} \in \mathbb{R}^d$, its ε-neighborhood (relative to the ℓ_∞ norm[10]) intersects at most k cubes? More formally we ask the following question where ${}^{\infty}\overline{B}_\varepsilon(\vec{p})$ denotes the closed ball of radius ε around $\vec{p}$ with respect to the ℓ_∞ norm.

Question 1.2.1 (Motivating Mathematical Question). Let $d \in \mathbb{N}$. For what values of $k \in \mathbb{N}$ and $\varepsilon \in (0, \infty)$ does there exist a partition $\mathcal{P}$ of $\mathbb{R}^d$ consisting of axis-aligned unit cubes such that for every point $\vec{p} \in \mathbb{R}^d$, we have that ${}^{\infty}\overline{B}_\varepsilon(\vec{p})$ intersects at most k members of the partition?

Ideally, we want ε to be large and k to be small, but as in the computational motivation, we regard k as a much more important parameter than ε. Thus, our first priority is to minimize k, and our secondary concern is to then maximize[11] ε. While we are interested in construction results which are restricted to using axis-aligned unit cubes, we are interested in more general impossibility results which allow for a much broader class of partitions.

While the problem we study is a natural geometric problem and we believe that it is of interest from a purely mathematical perspective, we have found no literature discussing Question 1.2.1. Nonetheless, there is a rich history of, and continued

[10] The ℓ_∞ norm is used because that is the natural norm in our motivating computational context. Furthermore, an ℓ_∞ ball is actually a cube, so the ball and the members of the partition are both the same geometric object which is convenient. Nonetheless, in this dissertation, we will give consideration to all norms in our impossibility results.

[11] It is not immediate that a maximum ε exists, so really we mean the supremum, but in informal contexts, we will refer to a maximum ε.

interest in, questions relating to partitions of $\mathbb{R}^d$, and in particular, partitions by unit cubes. Note that in essence, a partition by unit cubes is the same as a tiling of unit cubes—the only distinction is that in tilings, the boundaries are ignored and in a partition they are not (this is proved formally in Fact 3.6.8). Most of our results will deal with "partitions of unit cubes" and in this context "unit cube" refers to a translate of $[0,1)^d$. While we find partitions easier to work with, most of the results in the existing literature deal instead with "tilings of unit cubes" and in this context "unit cube" refers to a translate of $[0,1]^d$.

Below, we discuss a number of questions that have been investigated in the literature regarding tilings of unit cubes and other properties of unit cubes which will show up in this work. The purpose of these examples is fourfold: (1) to demonstrate broad interest in cubes and cube partitions, (2) to show that even though questions about cubes 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, so one should be careful not to assume that examples easily generalize.

Minkowski's lattice cube-tiling conjecture (1907) Minkowski's conjecture [Min07] states that in any lattice tiling of $\mathbb{R}^d$ by axis-aligned 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 an entire common edge, and in $\mathbb{R}^3$ there would be a pair of cubes sharing an entire common square side). A lattice tiling is one in which the centers of all of the cubes form a lattice. The conjecture was proven true in 1941 by Hajós [Haj42] by converting the problem to a purely algebraic one.

Keller's conjecture (1930) Keller's conjecture [Kel30] 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 axis-aligned 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—a total 113 years after Minkowski's original conjecture. In 1940, Perron [Per40a, Per40b] showed it was true for $d \leq 6$. Szabó [Sza86] recast the question in terms of periodic tilings in 1986 and then introduced the so-called Keller graphs along with Corrádi in 1990 [CS90]. In 1992, Lagarias and Shor [LS92] used the Keller graphs to show that the conjecture is false for all $d \geq 10$. This bound was refined by Mackey in 2002 [Mac02] 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 [DEL+11], and by Kisielewicz and Łysakowska in 2014 [KŁ14], and by Kisielewicz in 2017 [Kis17] and by Łysakowska in 2018 [Łys]. Finally, in 2020, Brakensiek, Heule, Mackey, and Narváez [BHMN20] 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 [Fur36] 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 cube is placed at each position, then every point of $\mathbb{R}^d$ either belongs to the boundary of some cube, or belongs to exactly k cubes). 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 [Haj42] that the conjecture was false for $d \geq 4$. In 1979, Robinson [Rob79] 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 [Fug74] conjectured that a set was spectral if and only if it could be used to tile $\mathbb{R}^d$. Though this was proven false by Tao in 2004 [Tao04], earlier work by Lagarias, Reeds, and Wang in 2000 [LRW00] 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 cubes in some partition of axis-aligned cubes. As a corollary, they used this result to show that extending an orthogonal set of functions to a basis is equivalent to extending a packing of cubes to a tiling. The ability to extend packings to tilings was also studied by Dutour, Itoh, and Poyarkov in 2018 [DIP18], though in a different context.

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 [Sal82b, Sal82a] 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. We utilized the dissection number of the d-cube for one of our upper bounds (see Section 7.2 (Upper Bound on ε via the Dissection Number of the Cube)) and, in particular, we used the lower bound of the dissection number given by Glazyrin in 2012 [Gla12].

The Lebesgue covering theorem (1911) In 1911, Brouwer was studying Euclidean d-space $\mathbb{R}^d$ and showed that $\mathbb{R}^d$ and $\mathbb{R}^{d'}$ are topologically different spaces if $d \neq d'$ [Bro11]. At the same time, Lebesgue was also studying $\mathbb{R}^d$ and showed that cubes in $\mathbb{R}^d$ could be assembled so that no more than $d+1$ meet at a common point and he conjectured that this was the minimum possible value [Leb11]; this conjecture was proven two years later by Brower [Bro13] when he improved on his prior work and was be able to define a dimension of a topological space and prove that the topological dimension of $\mathbb{R}^d$ is d[12]. This is summarized in the result known as the Lebesgue covering theorem (c.f. [HW48, Theorem IV 2]) which states that in any finite collection of closed sets which cover $[0,1]^d$, if no set intersects opposite faces of $[0,1]^d$, then there is some point belonging to at least $d+1$ of the sets in the collection. The work in this dissertation can be seen as a direct continuation of Lebesgue's line of research, but with the focus shifted beyond just the number of sets that must meet at a common point, and towards our parameter ε which generalizes from a point intersecting $d+1$ sets to an ℓ_∞ ball intersecting $d+1$ sets (and later other norms). We consider this question not just for $[0,1]^d$, but also for $\mathbb{R}^d$ where we generalize to allow covers (which we can reduce to partitions) with an upper bound on the outer measure of the members (see Chapter 8 (A Neighborhood Variant of the Lebesgue Covering Theorem, the Cubical KKM Lemma, and the Cubical Sperner's Lemma) as well as Chapter 6 (Optimality of k in General) and

[12]The papers discussed in this paragraph are not available in English, so the summary information about them is coming from the book [HW48].

Section 7.1 (Upper Bounds on ε via Brunn-Minkowski and Blichfeldt)).

Unit cubes more broadly Many of the above results are discussed in the 2005 survey paper "What Is Known About Unit Cubes" by Zong [Zon05] and in the followup 2006 book [ZBFK06]. 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.

1.3 Summary of Results

We now present a list of all main results in this dissertation with references provided to find the formal versions of each.

1.3.1 Main Results

Theorem 4.2.18 and Proposition 10.2.3: There exist efficiently computable axis-aligned unit cube partitions of $\mathbb{R}^d$ such that for $k = d + 1$ and $\varepsilon = \frac{1}{2d}$, every closed ℓ_∞ ε-ball intersects at most k cubes.

Theorem 5.1.1 and more generally Corollary 6.2.6: Regardless of the value of ε, there is no unit cube partition (axis-aligned or not) which can have a value of $k < d + 1$, so our constructions are optimal in this regard. As discussed earlier this result was known to Lebesgue (we give the proof from the Lebesgue covering theorem), but we offer two additional proofs of this fact—one of which is a very simple proof from first principles.

Corollary 9.8.5: When k is taken to be the minimum possible value $k = d + 1$, then for any axis-aligned unit cube partition, it must be that $\varepsilon \leq \frac{1}{2d}$, so our constructions are exactly optimal among all axis-aligned unit cube partitions.

Theorem 4.4.10: However, if we are willing to tolerate k which is not minimum but remains polynomial in d, then for any constant C, we can extend our constructions to (efficiently computable) axis-aligned unit cube partitions which have $\varepsilon = \frac{C}{d}$ and k polynomial in d.

Theorem 4.4.10: If we further weaken the restriction on k so that we only require it to be subexponential in d, we can get a vast improvement in ε. Specifically, for any $\varepsilon(d) \in o(1)$, we can extend our constructions to (efficiently computable) axis-aligned unit cube partitions which have the specified ε and have $k(d) \in \mathsf{weaksubexp}(d)$ (see footnote[13]).

Theorem 6.2.1: Beyond unit cube partitions, regardless of the value of ε, for every partition consisting of members with at most unit volume (which includes partitions with members of at most unit ℓ_∞ diameter[14]), it must be that $k \geq d+1$, so our constructions are optimal in this much more general regard.

Theorem 7.2.9 and Theorem 7.2.12: When k is taken to be the minimum value of $k = d+1$, then[15] for every partition consisting of members with at most unit ℓ_∞ diameter, it must be that $\varepsilon \leq \frac{2}{\sqrt{d}}$, and very importantly, this bound is given in terms of lower bounds on the dissection number of the cube. If the dissection number lower bounds can be improved enough (which is consistent with what is currently known) it would decrease our ε upper bound to $\varepsilon \leq \frac{C}{d}$ for a universal constant C. This would show that our constructions are optimal up to a constant factor even for the broad class of unit diameter bounded partitions

[13]There are two competing definitions in the research community for "subexponential" which we denote by $\mathsf{strongsubexp}(d)$ and $\mathsf{weaksubexp}(d)$. It holds that $\mathsf{strongsubexp}(d) \subseteq \mathsf{weaksubexp}(d)$ so that the stronger version is more restrictive and contains fewer functions. The stronger notion is defined as $\mathsf{strongsubexp}(d) = 2^{\mathsf{subpoly}(d)}$ which is more common in computational settings, and the weaker one is defined as $\mathsf{weaksubexp}(d) = 2^{o(d)} = \bigcap_{c\in(0,1)} o(c^d) = \bigcap_{c\in(0,1)} O(c^d)$ which we believe to be more mathematically natural considering it has multiple equivalent clean definitions (see Proposition D.0.3 for the proof of these qualities).

[14] This is a special property of the ℓ_∞ norm. See Fact 3.4.9 with $D = 1$.

[15]That k takes exactly the value $k = d+1$ is essential here.

(and we conjecture that this is the case (see Conjecture 7.3.2))

Theorem 7.1.9: In absence of improved dissection number lower bounds, we offer even better bounds on ε. For every partition consisting of members with at most unit volume (which includes partitions with members of at most unit ℓ_∞ diameter[14]), if k is taken to be the minimum value of $k = d + 1$, then it must be that $\varepsilon \leq \frac{\log_4(d+1)}{d}$ showing that our constructions are nearly optimal (up to a logarithmic factor) even for this extremely broad class of partitions.

Corollary 7.1.11: In fact, the above result is not specific to $k = d + 1$. For any $k \leq 2^d$, it must be that $\varepsilon \leq \frac{\log_4(k)}{d}$ showing in particular that even if k is allowed to be as large as polynomial in d, then ε is still at most $O\left(\frac{\log(d)}{d}\right)$ demonstrating that our constructions (either the original ones with $k = d + 1$ or the extended ones with polynomial k) are still optimal within a logarithmic factor.

Corollary 7.1.7 along with Theorem 7.1.9: Further still, the above result is not even specific to the ℓ_∞ norm. For any $k \leq 2^d$, and any norm $\|\cdot\|$ it must hold for any partition with at most unit diameter members (with respect to $\|\cdot\|$) that $\varepsilon \leq \frac{\log_4(k)}{d}$ (because for larger ε, there is an ε-ball with respect to $\|\cdot\|$ that intersects more than k members). There is also a corresponding version for any norm and partitions with at most unit volume members, but this version is less clean as there are other constants that show up relating the volume of the specific normed unit ball, so we will not state it here.

What the last few results indicate is that this bound on ε is currently better than the dissection based bound on ε in three ways (1) it holds for unit measure partitions which is a strict superset of the unit ℓ_∞ diameter partitions, (2) it holds for any $k \leq 2^d$ rather than just $k = d + 1$, and (3) it holds for any norm. Nonetheless, the dissection based ε bound remains relevant because the possibility is still open that that bound could some day give $\varepsilon \leq \frac{C}{d}$ which would be asymptotically better than the bound of $\varepsilon \leq \frac{\log_4(d+1)}{d}$ (again, in the much more specific context of $k = d+1$, the ℓ_∞ norm, and at most unit ℓ_∞ diameter members).

Theorem 8.0.7 and Theorem 8.0.8: Using a clever technique to deal with the boundary conditions, the methods used in the proof of Theorem 7.1.1 to get the above ε bound can be adapted from partitions of $\mathbb{R}^d$ to partitions of $[0,1]^d$ with members that don't included points on opposite faces. Appropriately rephrased, this gives a variant of the Lebesgue covering theorem discussed earlier (which can and should also be thought of as a variant of the KKM lemma) and we can discretize it to a variant of Sperner's lemma on the cube. Our variant states that in any coloring of $[0,1]^d$ in which no color is used on opposing faces, then there is a point where the open ε-ball in the ℓ_∞ norm contains points of at least $(1+\frac{2}{3}\varepsilon)^d$ different colors (for any ε in the sensible range $\varepsilon \in (0, \frac{1}{2}]$). If one thinks of ε as a function of the dimension d as we do in this work, then for $\varepsilon(d) \in O\left(\frac{1}{d}\right)$, this bound predicts nothing more than a constant number of colors, so asymptotically this is worse than the standard Sperner/KKM/Lebesgue result which predicts $d+1$ colors. However, if ε is large enough asymptotically that $\varepsilon(d) \in \omega\left(\frac{\log(d)}{d}\right)$, then the bound $(1+\frac{2}{3}\varepsilon(d))^d$ is super-polynomial in d giving vast improvement over the standard Sperner/KKM/Lebesgue bound of $d+1$. Finally, and obviously, if $\varepsilon(d) \in \Theta(1)$, then the bound $(1+\frac{2}{3}\varepsilon(d))^d$ is exponential in d. Considering the broad applicability of Sperner's lemma, we think that this will be a widely useful result in both mathematics and theoretical computer science.

Section 10.2: On the topic of computer science, because every deterministic function induces a natural partition of its domain (the one consisting of the fibers/preimages), the partition bounds above on k and ε imply restrictions on the types of deterministic rounding schemes/functions used in theoretical computer science which were described earlier. To complement this, our unit cube partition constructions can be used to define efficiently computable rounding functions with nearly matching bounds (again differing by the logarithmic factor discussed above).

Theorem 9.7.4: Our final main result diverges somewhat from our initial goal presented here. We consider, again, axis-aligned unit cube partitions, but we consider a natural generalization of the types of partitions in question by ignoring the ε parameter altogether. We consider instead the maximum number, k, of cubes that meet at a point which, as before, is still minimized at $k = d + 1$ (see Theorem 5.1.1). We can view the partition as an infinite graph where cubes are the vertices and two cubes/vertices are adjacent if they touch (i.e. their closures intersect). We show that the following are all equivalent:

1. The partition minimizes the number of cubes meeting at a point (in the graph theory sense, it minimizes the size of the largest clique[16]). This happens at $k = d + 1$.
2. The partition minimizes the maximum number of neighbors any cube has (in the graph theory sense, it minimizes the largest vertex degree). This

[16]This is true, but there is a subtlety that one may not immediately notice. Obviously if a collection of cubes meet at a single point, then this is a clique because all closures contain the point, so pairs of cubes have intersecting closures. However, it is probably not obvious that if a collection of cubes is a clique (i.e. for any pair of cubes, the closures intersect) that all of these cubes actually do meet at a single point. This fact follows from the special property of the ℓ_∞ norm that a set of diameter 1 can be contained in a ball of radius $\frac{1}{2}$ (see Fact 3.4.9). If the collection of cubes forms a clique, then each pair has intersecting closures which implies that the midpoints of the pair are ℓ_∞ distance at most 1 apart (because each closed cube is an ℓ_∞ ball of radius $\frac{1}{2}$ about its center). This means that if we consider the set of center positions of each cube, this set has ℓ_∞ diameter at most 1. Thus, there is some ball ${}^{\infty}\overline{B}_{\frac{1}{2}}(\vec{p})$ which contains the center points of each cube and so $\vec{p}$ is ℓ_∞ distance at most $\frac{1}{2}$ from the center of each cube in the clique. Since each cube closure is itself an ℓ_∞ ball of radius $\frac{1}{2}$ about its center point, this means each cube closure contains $\vec{p}$, so all cubes meet at the point $\vec{p}$.

happens at a value of $2^{d+1} - 2$ (see footnote[17]).

3. The partition consists only of cubes which are "neighborly" with each other as defined and studied in [Zak85, Zak87].

We also conjecture that we can add the following to the list of equivalencies:

4. **(Conjectured)** The partition has the property that in every collection of cubes which are all pairwise touching (a clique in the graph theory sense), the set of positions of these cubes are an affinely independent set of points (see footnote[18]).

We see this as a beautiful structural result about the types of cube partitions we have considered. In particular, we find it unsurprising but nice to know with certainty that the partitions which minimize the clique size are the same as those minimizing the vertex degree.

1.3.2 Additional Highlighted Results

This completes the list of the main results of this dissertation, but before proceeding, we do wish to highlight a small number of other results.

Definition 4.2.4 and Theorem 4.2.15: The initial constructions above have cubes positioned on a lattice, and the partition can be $(d+1)$-colored in the graph theory sense (two cubes that touch have different colors) which is the minimum possible number of colors for any unit cube partition (axis-aligned or not) because there always exists a $(d+1)$-clique.

[17] In fact, we show that in every axis-aligned unit cube partition, every cube touches at least $2^{d+1} - 2$ other cubes (i.e. every vertex has degree at least $2^{d+1} - 2$), so the minimizing partitions are in fact exactly those that ensure *every* vertex simultaneously has the smallest degree possible (i.e. every cube touches the minimum possible number of other cubes).

[18] This trivially implies the above because a set of affinely independent points in $\mathbb{R}^d$ has cardinality at most $d+1$, so the largest clique would have size $d+1$. It is the reverse implication that is conjectured.

Section 4.3: We initially defined a specific family of lattices/matrices for these constructions in such a way that we could analyze the ε parameter. Though we were just restrictive enough to allow our proofs to work, we suspected that we were far too restrictive in the sense that many other lattices/matrices should give rise to constructions with the same parameters. However, we observed surprisingly that there was something seemingly fundamental to the specific lattices/matrices we used. We believe further study of this phenomenon could be of mathematical interest, so we view this observation as an important one.

Chapter 6: We actually discuss three variations of the optimality of the parameter k in the general setting (the weakest being the one implied in the list of main results), and we give examples of simple partitions which demonstrate "gaps" between these notions as one strengthens or weakens the properties of the class of partitions under consideration. In particular, for partitions with a volume bound on the members, $k = d+1$ is optimal in the weakest sense. For partitions with a diameter bound on the members, $k = d+1$ is optimal in the middle sense. And for unit cube partitions, $k = d + 1$ is optimal in the strongest sense.

Proposition 7.2.10 and Proposition 7.2.10: We show that for $d = 1$ and $d = 2$ then $\varepsilon = \frac{1}{2d}$ is optimal when $k = d + 1$ for partitions with members of ℓ_∞ diameter at most 1.

Section 7.3: We give numerous conjectures on the true optimality of ε as our work leaves open the optimality in various contexts up to a logarithmic factor.

Chapter 9: The result we stated about minimizing degree being equivalent to minimizing clique size in axis-aligned unit cube partitions stems from properties that aren't really about partitions. Rather, the underlying properties are so local that they are really just about a single cube in the partition and all of its neighbors. For this reason much of the chapter is spent studying what we call "cube enclosures"—a single cube in space along with a set of other cubes that completely surround/enclose it. Though this particular structure has not, to our knowledge been studied before, there are other works that consider very local packings of cubes such as [DIP05, Poy07, SI10, LRW00, Zak85, Zak87] including in what circumstances the packings can be extended to tilings. One of our results is that in any such structure $(X, \mathcal{E})$ in $\mathbb{R}^d$ (where X is the central cube and the cubes in $\mathcal{E}$ surround X) the minimum size of $\mathcal{E}$ is $2^{d+1} - 2$ and we give exact conditions on how to achieve this minimum size—in particular, such minimum size structures fall out of our unit cube partition constructions in Chapter 4.

Section 10.3: We include a few impossibility results regarding the theory of learning which are not directly built upon the mathematical work here.

Chapter 2

Notation

We present here our notational conventions that will be used throughout this dissertation.

- Though we are of the opinion that the natural numbers include 0, in this work we shall use $\mathbb{N}$ to mean the strictly positive natural numbers.
- Throughout, we do our best to consistently use the following notational conventions:
  - The letters k and ε to refer to the two motivating quantities
  - Lower case letters such as x, y for real numbers and such as n, m for natural numbers (though in some contexts, m will mean Lebesgue measure)
  - Lower case vectors such as $\vec{x}, \vec{y}$ for points in $\mathbb{R}^d$ and such as $\vec{n}, \vec{m}$ for points in $\mathbb{Z}^d$
  - Upper case letters such as X, Y for subsets of $\mathbb{R}^d$
  - Upper case calligraphic letters such as $\mathcal{F}, \mathcal{P}, \mathcal{N}$ for families of subsets of $\mathbb{R}^d$
  - Subscripts for coordinates/vector components such as $x_1, \ldots, x_d$ to be the entries of $\vec{x}$ and X_i to be $\pi_i(X)$ when $X = \prod_{i=1}^d X_i$ is a product set
  - Parenthesized superscripts for collections of points or sets such as $\vec{x}^{(1)}, \ldots, \vec{x}^{(n)}$ to be a sequence of points or $X^{(1)}, \ldots, X^{(n)}$ be a sequence of

subsets of $\mathbb{R}^d$

- For $d \in \mathbb{N}$, the notation $[d]$ indicates the set of the first d natural numbers $[d] = \{i \in \mathbb{N} : i \leq d\} = \{1, 2, 3, \ldots, d\}$.
- When the space $\mathbb{R}^d$ is understood, then for $i \in [d]$ and a point $\vec{x} \in \mathbb{R}^d$, the notation $\pi_i(\vec{x}) \stackrel{\text{def}}{=} x_i$ means the ith coordinate projection of $\vec{x}$. We extend the notation to sets, so for $X \subseteq \mathbb{R}^d$, $\pi_i(X)$ means the ith coordinate projection of the set X: $\pi_i(X) = \{x_i : \vec{x} \in X\}$.
- When the space $\mathbb{R}^d$ is understood, and $X \subseteq \mathbb{R}^d$ is a product set $X = \prod_{i=1}^d X_i$ then for any $j \in [d]$ we write $\tau_j(X)$ to mean $\tau_j(X) = \prod_{i \in [d] \setminus \{j\}} X_i$ which is a product set in $\mathbb{R}^{[d] \setminus \{j\}} \cong \mathbb{R}^{d-1}$. We may also use the notation for a point so that $\tau_j(\vec{x}) \stackrel{\text{def}}{=} \langle x_i \rangle_{i \in [d] \setminus \{j\}}$.
- When the space $\mathbb{R}^d$ is understood, and $X \subseteq \mathbb{R}^d$, we use the following notation respectively for the interior, closure, and boundary of X: $\operatorname{int}(X)$, $\overline{X}$, $\operatorname{bd}(X)$.
- When the space $\mathbb{R}^d$ is understood, the notation $\vec{e}^{(i)}$ indicates the ith standard basis vector which has all 0 entries except for the ith entry which is 1.
- When the space $\mathbb{R}^d$ is understood, the notation $\vec{c}$ indicates the vector $\vec{c} \stackrel{\text{def}}{=} \langle c, c, \ldots, c \rangle \in \mathbb{R}^d$. We will frequently prefer to write $c \cdot \vec{1}$ to express this same vector (which looks notationally a little cleaner when c is written out in the text as a fraction).
- When the space $\mathbb{R}^d$ is understood, the notation $\operatorname{diam}_{\|\cdot\|}(X)$ to is used to denote the diameter of a set X with respect to the indicated norm $\|\cdot\|$. We will most often be interested in the ℓ_∞ norm, and we will use $\operatorname{diam}_\infty(X)$ instead.
- When the space $\mathbb{R}^d$ is understood, the notations ${}^{\|\cdot\|}B^\circ_\varepsilon(\vec{p})$ and ${}^{\|\cdot\|}\overline{B}_\varepsilon(\vec{p})$ will respectively denote the open and closed balls of radius ε about $\vec{p}$ with respect to the indicated norm $\|\cdot\|$. We will most often be interested in the ℓ_∞ norm, and we will denote such balls as ${}^{\infty}B^\circ_\varepsilon(\vec{p})$ and ${}^{\infty}\overline{B}_\varepsilon(\vec{p})$.

- When the space $\mathbb{R}^d$ is understood, then $H_\varepsilon(\vec{p})$ stands mnemonically for "half-open cube" and is defined as $H_\varepsilon(\vec{p}) \stackrel{\text{def}}{=} \vec{p} + [-\varepsilon, \varepsilon)^d$. Typically, it is useful to consider this as a set between ${}^{\infty}B^\circ_\varepsilon(\vec{p})$ and ${}^{\infty}\overline{B}_\varepsilon(\vec{p})$. That is, ${}^{\infty}B^\circ_\varepsilon(\vec{p}) \subseteq H_\varepsilon(\vec{p}) \subseteq {}^{\infty}\overline{B}_\varepsilon(\vec{p})$.
- When the space $\mathbb{R}^d$ is understood, and a family, $\mathcal{F}$, of subsets of $\mathbb{R}^d$ is also understood, we use the following notations:
  - ${}^{\|\cdot\|}\mathcal{N}^\circ_\varepsilon(\vec{p})$ (resp. ${}^{\|\cdot\|}\overline{\mathcal{N}}_\varepsilon(\vec{p})$) mnemonically stands for the "open ε neighborhood of $\vec{p}$" (resp. "closed ε neighborhood of $\vec{p}$") and consists of all sets in $\mathcal{F}$ which intersect the open (resp. closed) ε-ball at $\vec{p}$. Formally,

$$
\begin{aligned}
{}^{\|\cdot\|}\mathcal{N}^\circ_\varepsilon(\vec{p}) &\stackrel{\text{def}}{=} \{X \in \mathcal{F} : X \cap {}^{\|\cdot\|}B^\circ_\varepsilon(\vec{p}) \neq \emptyset\} \\
{}^{\|\cdot\|}\overline{\mathcal{N}}_\varepsilon(\vec{p}) &\stackrel{\text{def}}{=} \left\{X \in \mathcal{F} : X \cap {}^{\|\cdot\|}\overline{B}_\varepsilon(\vec{p}) \neq \emptyset\right\}
\end{aligned}
$$

    Because the ℓ_∞ norm will be used most frequently, we will use ${}^{\infty}\mathcal{N}^\circ_\varepsilon(\vec{p})$ (resp. ${}^{\infty}\overline{\mathcal{N}}_\varepsilon(\vec{p})$) to refer specifically to the ℓ_∞ norm.
  - $\mathcal{N}_{\overline{0}}(\vec{p})$ mnemonically stands for the "zero-closed neighborhood of $\vec{p}$" and consists of all sets in $\mathcal{F}$ whose closures contain $\vec{p}$. Formally,

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

    We realize that the $\overline{0}$ subscript is peculiar, but this is meant to be a reminder that this is defined relative to the closures of members and requires the closures to be distance 0 from $\vec{p}$.
- When the space $\mathbb{R}^d$ is understood, we will use m to denote the Lebesgue measure of a measurable set (i.e. the volume), and use m_{in} and m_{out} to denote the (induced) Lebesgue inner and outer measure of arbitrary sets when we do not

know if a set is measurable (i.e. it may not have a well-defined volume).

- The notation $v_{\|\cdot\|,d}$ denotes the Borel/Lebesgue measure (i.e. volume) of the unit ball in $\mathbb{R}^d$ with respect to a norm $\|\cdot\|$—that is $v_{\|\cdot\|,d} = m\left({}^{\|\cdot\|}B_1^\circ(\vec{0})\right) = m\left({}^{\|\cdot\|}\overline{B}_1(\vec{0})\right)$.
- When a partition $\mathcal{P}$ of a set is understood, and x is a point in the set, the notation $\operatorname{member}(x)$ is used to indicate the unique set in $\mathcal{P}$ which contains x. This notation is preferred over the more common notation of $[x]$ because we generally are not interested in the underlying equivalence relation and prefer to think conceptually of members of the partition. If there is ambiguity, we may write $\operatorname{member}_{\mathcal{P}}(x)$ to clarify the partition. In general we will refer to an arbitrary set in a partition as a "member of the partition."
- When the space $\mathbb{R}^d$ is understood, then for any bounded set $X \subseteq \mathbb{R}^d$, we use $\operatorname{corners}(X)$ to denote the set of 2^d corners of the smallest rectangle containing X. Formally,

$$\operatorname{corners}(X) \stackrel{\text{def}}{=} \prod_{i=1}^{d} \left\{\inf(\pi_i(X)),\ \ \sup(\pi_i(X))\right\}.$$

  Most commonly we use this notation when X is a rectangle (i.e. a product of intervals) and most commonly when X is an axis-aligned cube, but we don't require this.
- Some additional notation is introduced in Chapter 3 (Preliminaries) in context.
- Some additional notation is stated in Section 9.1 (Notation) which is used only in Chapter 9.

Chapter 3

Preliminaries

In this chapter we will state and prove a variety of fairly simple results. Nothing in this chapter is considered significant original research, but it nonetheless serves to provide the basic facts that we will later utilize.

3.1 The Notion of Seclusion

The following two definitions capture the main idea that we study. While we will for the most part be interested in studying partitions, we will sometimes wish to use the definitions for structures other than partitions (especially in Chapter 9) so we define them more generally.

Definition 3.1.1 ((k,ε)-Secluded). Let $d,k \in \mathbb{N}$ and $\varepsilon \in (0,\infty)$ and $\mathcal{F}$ a family of subsets of $\mathbb{R}^d$. We say that $\mathcal{F}$ is (k,ε)-*secluded* if for each $\vec{p} \in \mathbb{R}^d$ it holds that $\left|{}^{\infty}\overline{\mathcal{N}}_{\varepsilon}(\vec{p})\right| \leq k$.

We also define the notion of secluded when there is not a fixed ε that works for every point.

Definition 3.1.2 (k-Secluded). Let $d, k \in \mathbb{N}$ and $\mathcal{F}$ a family of subsets of $\mathbb{R}^d$. We say that $\mathcal{F}$ is *k-secluded* if for each $\vec{p} \in \mathbb{R}^d$, there exists $\varepsilon \in (0, \infty)$ such that $\left| {}^{\infty}\overline{\mathcal{N}}_{\varepsilon}(\vec{p}) \right| \leq k$.

Note that both of the definitions above are specific to the ℓ_∞ norm because that is the norm used in the neighborhoods. Nonetheless, we will sometimes be interested in other norms; when we are, we will simply discuss that $\left| {}^{\|\cdot\|}\overline{\mathcal{N}}_{\varepsilon}(\vec{p}) \right| \leq k$ for some ε and k and not use the term (k, ε)-secluded or k-secluded.

3.2 Packings

Definition 3.2.1 (Packing). Let $d \in \mathbb{N}$. A *packing* in $\mathbb{R}^d$ is a family $\mathcal{F}$ of subsets of $\mathbb{R}^d$ such sets in $\mathcal{F}$ have pairwise disjoint interiors.

3.3 Graph Theory Notions for Sets

Definition 3.3.1 (Adjacent). Let $d \in \mathbb{N}$ and X, Y be subsets of $\mathbb{R}^d$. X and Y are called *adjacent* if $\operatorname{int}(X) \cap \operatorname{int}(Y) = \emptyset$ and $\overline{X} \cap \overline{Y} \neq \emptyset$. We denote this as $X \overset{\text{adj}}{\sim} Y$. Using language from graph theory, if X and Y are adjacent, we will sometimes say that Y is a *neighbor* of X (and vice versa).

We emphasize that we define adjacency to require disjoint interiors. Also, we remark that we use the notation $X \overset{\text{adj}}{\sim} Y$ rather than simpler and more common notation $X \sim Y$ because we will later have need of a stronger notion of adjacency (where we call two sets cousins) and we denote this relationship as $X \overset{\text{cous}}{\sim} Y$, so we want to always be clear which notion we are referring to.

Using the definition above, any family of subsets of $\mathbb{R}^d$ naturally induces a graph.

Definition 3.3.2 (Set Family Graph/Partition Graph). Let $d \in \mathbb{N}$ and $\mathcal{F}$ be a family of subsets of $\mathbb{R}^d$. The *set family graph* of $\mathcal{F}$ is the graph G whose vertex set is the set $\mathcal{F}$ and whose edge set contains the edge $\{X, Y\}$ if and only if $X \overset{\text{adj}}{\sim} Y$ as sets. If $\mathcal{F}$ is a partition, we call this the *partition graph*.

We will usually not talk explicitly of this graph and will instead identify the partition and its graph structure at times. One structure we will be particularly interested in is cliques.

Definition 3.3.3 (Clique). Let $d \in \mathbb{N}$ and $\mathcal{C}$ a family of subsets of $\mathbb{R}^d$. Using the language from graph theory, $\mathcal{C}$ is called a *clique* if for all pairs of distinct $X, Y \in \mathcal{C}$ we have $X \overset{\text{adj}}{\sim} Y$. Equivalently, $\mathcal{C}$ is called a *clique* if its induced set family graph is a clique in the graph theoretic sense.

3.4 Unit Cubes

Definition 3.4.1 (Axis-Aligned Unit Cube). Let $d \in \mathbb{N}$ and $X \subseteq \mathbb{R}^d$. Then X is called an *axis-aligned unit cube* if there exists $\vec{x} \in \mathbb{R}^d$ such that $\vec{x} + (0,1)^d \subseteq X \subseteq \vec{x} + [0,1]^d$.

Remark 3.4.2. On occasion, we don't care if cubes are axis-aligned because we only care about the volume, so we try to state "axis-aligned" when we need that requirement, but there may be places where we forgot to be explicit. In such places, it should be possible to identify the need for axis-aligned cubes either because of the use of other results which assume axis-aligned cube or because we directly express the cubes in the form of a product set. △

The next fact follows immediately from the above definition.

Fact 3.4.3 (Unit Cube Interior and Closure)**.** *Let $d \in \mathbb{N}$ and X an axis-aligned unit cube in $\mathbb{R}^d$. Then there is some (unique) $\vec{x} \in \mathbb{R}^d$ such that* $\operatorname{int}(X) = \vec{x} + (0,1)^d$ *and* $\overline{X} = \vec{x} + [0,1]^d$. *Also, taking* $\vec{p} = \vec{x} + \frac{1}{2} \cdot \vec{1}$, *then* $\operatorname{int}(X) = \vec{p} + \left(-\frac{1}{2}, \frac{1}{2}\right)^d = {}^{\infty}B^{\circ}_{1/2}(\vec{p})$ *and* $\overline{X} = \vec{p} + \left[-\frac{1}{2}, \frac{1}{2}\right]^d = {}^{\infty}\overline{B}_{1/2}(\vec{p})$.

Definition 3.4.4 (Unit Cube Center and Representative). Let $d \in \mathbb{N}$ and X an axis-aligned unit cube in $\mathbb{R}^d$. Then the *center* of X is the unique point $\vec{p}$ (from Fact 3.4.3) such that $\operatorname{int}(X) = \vec{p} + \left(-\frac{1}{2}, \frac{1}{2}\right)^d = {}^{\infty}B^{\circ}_{1/2}(\vec{p})$ and $\overline{X} = \vec{p} + \left[-\frac{1}{2}, \frac{1}{2}\right]^d = {}^{\infty}\overline{B}_{1/2}(\vec{p})$, and this point $\vec{p}$ is denoted as $\operatorname{center}(X)$. The *representative corner* of X is the point $\langle \inf \pi_i(X) \rangle_{i=1}^d = \langle p_i - \frac{1}{2} \rangle_{i=1}^d = \vec{p} - \frac{1}{2} \cdot \vec{1}$ and is denoted as $\operatorname{rep}(X)$. We will also use the notation $\operatorname{center}_i(X) \stackrel{\text{def}}{=} \pi_i(\operatorname{center}(X))$ and $\operatorname{rep}_i(X) \stackrel{\text{def}}{=} \pi_i(\operatorname{rep}(X))$.

Remark 3.4.5. We will really only use $\operatorname{rep}(X)$ in the context of constructions of partitions by translates of $[0,1)^d$, and in such cases $\operatorname{rep}(X)$ is the unique corner of X is contained in X. The use of $\operatorname{rep}(X)$ at all is an old artifact of our work that remains in the results of Chapter 4 (Constructions), but it is generally much more convenient to work with $\operatorname{center}(X)$ because of the perspective that X sits between an open and closed ℓ_∞ ball of radius $\frac{1}{2}$ about its center. A convenient observation for converting between the two perspectives is that for two axis-aligned unit cubes X and Y, it holds that $\operatorname{center}(X) - \operatorname{center}(Y) = \operatorname{rep}(X) - \operatorname{rep}(Y)$ (as shown in the proof of the next result). △

The following is a simple well-known fact, but one that is very useful when working with axis-aligned unit cubes.

Fact 3.4.6 (Adjacency for Unit Cubes)**.** *Let $d \in \mathbb{N}$ and X, Y be axis-aligned unit cubes in $\mathbb{R}^d$. Then the following are equivalent:*

1. $X \overset{\text{adj}}{\sim} Y$
2. $\|\text{center}(X) - \text{center}(Y)\|_\infty = 1$
3. $\|\text{rep}(X) - \text{rep}(Y)\|_\infty = 1$

Proof. (1) $\iff$ (2) : By Definition 3.4.4 and Fact 3.4.3, we have that $\overline{X} = {}^\infty\overline{B}_{1/2}(\text{center}(X))$ and $\text{int}(X) = {}^\infty B^\circ_{1/2}(\text{center}(X))$ and similarly for Y. Thus $\text{int}(X) \cap \text{int}(Y) = \emptyset$ if and only if $\|\text{center}(X) - \text{center}(Y)\|_\infty \geq 1$ and $\overline{X} \cap \overline{Y} \neq \emptyset$ if and only if $\|\text{center}(X) - \text{center}(Y)\|_\infty \leq 1$. Thus, both happen simultaneously (i.e. the definition of adjacent) if and only if $\|\text{center}(X) - \text{center}(Y)\|_\infty = 1$.

(2) $\iff$ (3) : By Definition 3.4.4, we have $\text{rep}(X) = \text{center}(X) - \frac{1}{2} \cdot \vec{1}$ and $\text{rep}(Y) = \text{center}(Y) - \frac{1}{2} \cdot \vec{1}$, so in fact $\text{center}(X) - \text{center}(Y) = \text{rep}(X) - \text{rep}(Y)$, so in particular $\|\text{center}(X) - \text{center}(Y)\|_\infty = \|\text{rep}(X) - \text{rep}(Y)\|_\infty$. $\square$

If we add the context of a packing to Fact 3.4.6, we get one other natural equivalence added.

Fact 3.4.7 (Adjacency for Unit Cube Packings)**.** *Let $d \in \mathbb{N}$ and $\mathcal{F}$ a packing of axis-aligned unit cubes in $\mathbb{R}^d$ and $X, Y \in \mathcal{F}$ be distinct*[a]*. Then the following are equivalent:*

1. $X \overset{\text{adj}}{\sim} Y$
2. $\|\text{center}(X) - \text{center}(Y)\|_\infty = 1$
3. $\|\text{rep}(X) - \text{rep}(Y)\|_\infty = 1$
4. $\|\text{center}(X) - \text{center}(Y)\|_\infty \leq 1$
5. $\|\text{rep}(X) - \text{rep}(Y)\|_\infty \leq 1$

[a] We now need the hypothesis that $X \neq Y$.

Proof. By Fact 3.4.6 we have (1) $\iff$ (2) $\iff$ (3), and as in that proof we have $\|\text{center}(X) - \text{center}(Y)\|_\infty = \|\text{rep}(X) - \text{rep}(Y)\|_\infty$ which shows (4) $\iff$ (5), so it suffices for us to show (2) $\iff$ (4).

(2) $\iff$ (4) : The forward implication is trivial, so we only need to handle the reverse implication. By Definition 3.4.4 and Fact 3.4.3, we have that $\text{int}(X) = {}^{\infty}B^{\circ}_{1/2}(\text{center}(X))$ and $\text{int}(Y) = {}^{\infty}B^{\circ}_{1/2}(\text{center}(Y))$, and because $X \neq Y$ by hypothesis, we have by definition of $\mathcal{F}$ being a packing that $\text{int}(X) \cap \text{int}(Y) = \emptyset$ which implies that $\|\text{center}(X) - \text{center}(Y)\|_\infty \geq 1$. Thus, if we assume for the reverse implication that $\|\text{center}(X) - \text{center}(Y)\|_\infty \leq 1$, we obtain the equality $\|\text{center}(X) - \text{center}(Y)\|_\infty = 1$ as desired. □

For clarity, we state the following which one probably already anticipates.

Definition 3.4.8 (Axis-Aligned Unit Cube Packing/Clique/Partition). A packing/clique/partition $\mathcal{F}/\mathcal{C}/\mathcal{P}$ of $\mathbb{R}^d$ is called an *axis-aligned unit cube packing/clique/partition* if every member of $\mathcal{F}/\mathcal{C}/\mathcal{P}$ is an axis-aligned unit cube.

It is an elementary fact of metric spaces, that if a set X has finite diameter D, then there exists a point x in the metric space such that $X \subseteq \overline{B}_D(x)$. This ball has twice the diameter of X. When working specifically with the ℓ_∞ norm, though, this factor of 2 is not needed.

Fact 3.4.9 (ℓ_∞ Diameter Ball)**.** *Let $d \in \mathbb{N}$ and $X \subseteq \mathbb{R}^d$ be such that $D = \text{diam}_\infty(X)$ is finite. Then there exists $\vec{p} \in \mathbb{R}^d$ such that $X \subseteq \overline{B}_{D/2}(\vec{p})$. As a consequence, $m_{out}(X) \leq D^d$ where m_{out} denotes outer Lebesgue measure.*

Proof. For each coordinate $i \in [d]$, consider the set $X_i = \{\pi_i(\vec{x}) : \vec{x} \in X\} \subseteq \mathbb{R}$ of the ith coordinates of each point in X. The infimum and supremum are distance at most D apart, because otherwise there would be points $\vec{y}, \vec{z} \in X$ such that $|\pi_i(\vec{z}) - \pi_i(\vec{y})| >$

D which means $\|\vec{z}-\vec{y}\|_\infty > D$. Thus, taking $\vec{p} = \langle \frac{\inf(X_i)+\sup(X_i)}{2} \rangle_{i=1}^{d}$ we have

$$X \subseteq \prod_{i=1}^{d} X_i \subseteq \prod_{i=1}^{d} [\inf(X_i), \sup(X_i)] \subseteq \prod_{i=1}^{d} [p_i - \tfrac{D}{2}, p_i + \tfrac{D}{2}] = {}^{\infty}\overline{B}_{D/2}(\vec{p}).$$

The consequence is trivial; by definition, the outer measure of X is at most the measure of any superset. □

The usefulness of this fact about the ℓ_∞ norm is that in any axis-aligned unit cube clique, there exists at least one point that belongs to the closure of all the cubes. This is a special property not just of the ℓ_∞ norm, but also of the fact that the members are unit cubes. It is otherwise quite easy to construct cliques of sets where there is no point common to the closure of all members. See Figure 3.1 for example.

Definition 3.4.10 (Clique-Point). Let $d \in \mathbb{N}$, and $\mathcal{C}$ be an axis-aligned unit cube clique in $\mathbb{R}^d$. A point $\vec{p}$ is called a *clique-point* of $\mathcal{C}$ if for all $X \in \mathcal{C}$ we have that $\vec{p} \in \overline{X}$ (i.e. $\vec{p} \in \bigcap_{X \in \mathcal{C}} \overline{X}$).

The following result shows that in an axis-aligned unit cube clique, there is a point $\vec{p}$ which is at the closure of all of cubes in the clique, and so for any choice of $\varepsilon \in (0,\infty)$, the ball ${}^{\infty}\overline{B}_\varepsilon(\vec{p})$ will intersect every one of these cubes. The usefulness of this result is that it is possible to show that an axis-aligned unit cube partition is not (k,ε)-secluded by finding a clique of size $k+1$.

Lemma 3.4.11 (Clique-Points Exist for Axis-Aligned Unit Cube Cliques)**.** *Let $d \in \mathbb{N}$ and $\mathcal{C}$ be an axis-aligned unit cube clique in $\mathbb{R}^d$ (i.e. a clique containing only axis-aligned unit cubes). Then $\mathcal{C}$ has at least one clique-point.*

Proof. Let $C = \{\text{center}\, X : X \in \mathcal{C}\}$ denote the centers of all cubes in $\mathcal{C}$. By Fact 3.4.6, for any $X, Y \in \mathcal{C}$, $\|\text{center}(X) - \text{center}\, Y)\|_\infty = 1$, so $\text{diam}_\infty(C) = 1$. Then let $\vec{p}$ as in

Fact 3.4.9 so that $C \subseteq {}^{\infty}\overline{B}_{1/2}(\vec{p})$. By the definition of C, this means that for any $X \in \mathcal{C}$, $\|\text{center}(X) - \vec{p}\|_{\infty} \leq \frac{1}{2}$ and thus (by Fact 3.4.3), $\vec{p} \in {}^{\infty}\overline{B}_{1/2}(\text{center}(X)) = \overline{X}$. □

Observation 3.4.12. *Lemma 3.4.11 along with an upcoming result (Fact 3.6.5 (Locally Finite: Enlarged Neighborhood)) shows that an axis-aligned unit cube partition of $\mathbb{R}^d$ is k-secluded (Definition 3.1.2) if and only if the partition graph (Definition 3.3.2) has clique number at most k. This shows that we can study k-secluded axis-aligned unit cube partitions of $\mathbb{R}^d$ by just studying their graph structure, and this is something we do throughout Chapter 9 (Secluded Partitions Without ε).*

Proof. For one direction, every clique $\mathcal{C}$ has a clique-point $\vec{p}$ at the closure of every member of $\mathcal{C}$, so if the partition is k-secluded, then for some ε we have $\left|{}^{\infty}\overline{\mathcal{N}}_{\varepsilon}(\vec{p})\right| \leq k$ and because ${}^{\infty}\overline{\mathcal{N}}_{\varepsilon}(\vec{p}) \supseteq \mathcal{N}_{\overline{0}}(\vec{p}) \supseteq \mathcal{C}$, we have $|C| \leq k$ showing that the clique number of the partition graph is at most k.

Conversely, if the partition graph has clique number at most k, then in particular, for every point $\vec{p} \in \mathbb{R}^d$ it holds that $|\mathcal{N}_{\overline{0}}(\vec{p})| \leq k$ because this trivially constitutes a clique as all members of this family have closures intersecting at $\vec{p}$. By the upcoming Fact 3.6.5, for each point $\vec{p}$ it holds for all sufficiently small $\varepsilon > 0$ that $\mathcal{N}_{\overline{0}}(\vec{p}) = {}^{\infty}\overline{\mathcal{N}}_{\varepsilon}(\vec{p})$ which shows that there exists ε such that ${}^{\infty}\overline{\mathcal{N}}_{\varepsilon}(\vec{p}) \leq k$, so the partition is k-secluded by definition. □

The following fact will sometimes be applied to a closed ℓ_{∞} ball (which is a closed rectangle) along with a half-open unit cube, and sometimes it will be applied to a half-open unit cube and the closure of a different half-open unit cube.

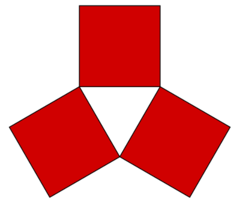


Figure 3.1: A cube clique with 3 members that has no point common to all members of the clique. By Lemma 3.4.11, this cannot occur with axis-aligned unit cube cliques.

Fact 3.4.13 (Containing Corners of Small Rectangles)**.** *Let* $B = \prod_{i=1}^{d} [a_i, b_i]$ *(even allowing* $a_i = b_i$*) and* $Y = \prod_{i=1}^{d} [x_i, y_i)$ *such that for all* $i \in [d]$ *it holds that* $y_i - x_i \geq b_i - a_i$ *(i.e.* Y *is at least as long as* B *in each coordinate). If* $Y \cap B \neq \emptyset$*, then* Y *contains a corner of* B*.*

Proof. We want to show that $C \stackrel{\text{def}}{=} Y \cap \text{corners}(B) \neq \emptyset$. Note that $C = \prod_{i=1}^{d} \{a_i, b_i\} \cap [x_i, y_i)$, so we will show for arbitrary $i \in [d]$ that $\{a_i, b_i\} \cap [x_i, y_i) \neq \emptyset$ by considering the four possible cases depending on how a_i and x_i relate and how y_i and b_i relate.

(1) $a_i \geq x_i$ **and** $b_i < y_i$**:**

Since $a_i \leq b_i$, this gives the ordering $x_i \leq a_i \leq b_i < y_i$, so $\{a_i, b_i\} \cap [x_i, y_i) = \{a_i, b_i\}$.

(2) $a_i \geq x_i$ **and** $b_i \geq y_i$**:**

It cannot be the case that $y_i \leq a_i$ because then the ordering would be $x_i \leq y_i \leq a_i \leq b_i$ which would contradict that $Y \cap B \neq \emptyset$. Thus $a_i < y_i$ and the ordering is $x_i \leq a_i < y_i \leq b_i$, so $\{a_i, b_i\} \cap [x_i, y_i) = \{a_i\}$.

(3) $a_i < x_i$ **and** $b_i < y_i$**:**

Similar to the last case, it cannot be the case that $b_i < x_i$ because then the ordering would be $a_i \leq b_i < x_i < y_i$ which would contradict that $Y \cap B \neq \emptyset$. Thus the ordering is $a_i < x_i \leq b_i < y_i$, so $\{a_i, b_i\} \cap [x_i, y_i) = \{b_i\}$.

(4) $a_i < x_i$ **and** $b_i \geq y_i$**:**

This case cannot occur, because it would give the ordering $a_i < x_i < y_i \leq b_i$ which would contradict that $y_i - x_i \geq b_i - a_i$.

In all cases, $C = \{a_i, b_i\} \cap [x_i, y_i) \neq \emptyset$. □

The following is virtually the same claim as above, but replacing the half-open set with a closed one. All cases of the proof remain the same except for case (2). It does not follow as a direct corollary because the intersection is larger.

Fact 3.4.14 (Containing Corners of Small Rectangles)**.** *Let* $B = \prod_{i=1}^{d}[a_i, b_i]$ *(even allowing* $a_i = b_i$*) and* $Y = \prod_{i=1}^{d}[x_i, y_i]$ *such that for all* $i \in [d]$ *it holds that* $y_i - x_i \geq b_i - a_i$ *(i.e.* Y *is at least as long as* B *in each coordinate). If* $Y \cap B \neq \emptyset$*, then* Y *contains a corner of* B*.*

Proof. We want to show that $C = Y \cap \text{corners}(B) \neq \emptyset$. Note that $C = \prod_{i=1}^{d} \{a_i, b_i\} \cap [x_i, y_i]$, so we will show for arbitrary $i \in [d]$ that $\{a_i, b_i\} \cap [x_i, y_i] \neq \emptyset$ by considering the four possible cases depending on how a_i and x_i relate and how y_i and b_i relate.

(1) $a_i \geq x_i$ **and** $b_i < y_i$**:**

Since $a_i \leq b_i$, this gives the ordering $x_i \leq a_i \leq b_i < y_i$, so $\{a_i, b_i\} \cap [x_i, y_i] = \{a_i, b_i\}$.

(2) $a_i \geq x_i$ **and** $b_i \geq y_i$**:**

It cannot be the case that $y_i < a_i$ because then the ordering would be $x_i \leq y_i < a_i \leq b_i$ which would contradict that $Y \cap B \neq \emptyset$. Thus $a_i \leq y_i$ and the ordering is $x_i \leq a_i \leq y_i \leq b_i$, so $\{a_i, b_i\} \cap [x_i, y_i]$ is either $\{a_i\}$ or $\{a_i, b_i\}$.

(3) $a_i < x_i$ **and** $b_i < y_i$**:**

Similar to the last case, it cannot be the case that $b_i < x_i$ because then the ordering would be $a_i \leq b_i < x_i < y_i$ which would contradict that $Y \cap B \neq \emptyset$. Thus the ordering is $a_i < x_i \leq b_i < y_i$, so $\{a_i, b_i\} \cap [x_i, y_i] = \{b_i\}$.

(4) $a_i < x_i$ and $b_i \geq y_i$:

This case cannot occur, because it would give the ordering $a_i < x_i < y_i \leq b_i$ which would contradict that $y_i - x_i \geq b_i - a_i$.

In all cases, $C = \{a_i, b_i\} \cap [x_i, y_i] \neq \emptyset$. □

The following corollary is the primary way we will utilize the prior two results.

Corollary 3.4.15 (Adjacent Cubes Share a Corner). *Let $d \in \mathbb{N}$ and $X, Y \subset \mathbb{R}^d$ be axis-aligned unit cubes such that $X \overset{\text{adj}}{\sim} Y$. Then there exists $\vec{p} \in \operatorname{corners}(X)$ such that $\vec{p} \in \overline{Y}$ (and vice versa by symmetry).*

Proof. Since $X \overset{\text{adj}}{\sim} Y$, by definition, $\overline{X} \cap \overline{Y} \neq \emptyset$. So by Fact 3.4.14, (taking B to be $\overline{X}$, and Y to be $\overline{Y}$), we have $\operatorname{corners}(\overline{X}) \cap \overline{Y} \neq \emptyset$ and since $\operatorname{corners}(\overline{X}) = \operatorname{corners}(X)$ we have $\operatorname{corners}(X) \cap \overline{Y} \neq \emptyset$, so there is some $\vec{p} \in \operatorname{corners}(X) \cap \overline{Y}$. □

3.5 Topology

The following is a well-known result in topology and real analysis.

Theorem 3.5.1 (Equivalence of Norms on $\mathbb{R}^d$). *Let $d \in \mathbb{N}$ and let $\|\cdot\|^a$ and $\|\cdot\|^b$ be two norms on $\mathbb{R}^d$. Then there exists constants $c_d, C_d \in (0, \infty)$ such that for all $\vec{x} \in \mathbb{R}^d$, it holds that $c_d\|\vec{x}\|^a \leq \|\vec{x}\|^b \leq C_d\|\vec{x}\|^a$.*

Remark 3.5.2. A consequence of the above result is that all norms on $\mathbb{R}^d$ generate the same topology on $\mathbb{R}^d$. In other words, the collection of open sets in $\mathbb{R}^d$ is the same no matter which norm we are using. This also means that the Borel (and Lebesgue) σ-algebra on $\mathbb{R}^d$ is the same no matter which norm is used, and thus balls with respect to any norm on $\mathbb{R}^d$ are measurable. △

3.6 Locally Finite Families

The following results will mostly be applied to sets of axis-aligned unit cubes, but on occasion they will be used for other sets. Also, for most of the paper they will only be applied to sets of cubes which are actually partitions of $\mathbb{R}^d$, but in Chapter 9 we will use these results for sets of cubes other than partitions.

The following is a standard definition from topology taken in the specific case where the space is $\mathbb{R}^d$ along with the standard topology (the one induced by any norm on $\mathbb{R}^d$).

Definition 3.6.1. A family $\mathcal{F}$ of subsets of $\mathbb{R}^d$ is called *locally finite if for every point* $\vec{p} \in \mathbb{R}^d$ *there exists* $\varepsilon \in (0,\infty)$ *such that* $\left|{}^{\infty}\overline{\mathcal{N}}_{\varepsilon}(\vec{p})\right| < \infty$ *(i.e.* ${}^{\infty}\overline{B}_{\varepsilon}(\vec{p})$ *intersects finitely many sets in* $\mathcal{F}$*).*

Remark 3.6.2. Note that by Theorem 3.5.1 (Equivalence of Norms on $\mathbb{R}^d$), the use of the ℓ_∞ norm above could be replaced with any other norm to recover an equivalent definition, and also the closed balls/neighborhoods could be replaced with open balls/neighborhoods to obtain an equivalent definition. △

The following says that if a packing in $\mathbb{R}^d$ has a non-zero lower bound on the measure of the interior[1] of every member and an upper bound on the diameter of every member, then it is a locally finite finite set. In fact, something stronger is true: not only does there exist an ε for each $\vec{p}$, but every ε works for every $\vec{p}$.

Fact 3.6.3 (Some Locally Finite Families)**.** *Let* $d \in \mathbb{N}$, *and* $\|\cdot\|$ *any norm on* $\mathbb{R}^d$, *and* $\mathcal{F}$ *be a packing in* $\mathbb{R}^d$ *such that there exists* $\mu, D \in (0,\infty)$ *so that for all* $X \in \mathcal{F}$, $\mu < m(\mathrm{int}(X))$ *and* $\mathrm{diam}_{\|\cdot\|}(X) \le D$. *Then* $\mathcal{F}$ *is locally finite. In fact, for any* $\vec{p} \in \mathbb{R}^d$ *and* $\varepsilon \in (0,\infty)$, *the cardinality of* ${}^{\|\cdot\|}\overline{\mathcal{N}}_{\varepsilon}(\vec{p})$ *is finite.*

[1]The interior is always Borel measurable as it is open.

Proof. Recall that ${}^{\|\cdot\|}\overline{\mathcal{N}}_\varepsilon(\vec{p}) = \left\{X \in \mathcal{F} : X \cap {}^{\|\cdot\|}\overline{B}_\varepsilon(\vec{p}) \neq \emptyset\right\}$. Consider the ball ${}^{\|\cdot\|}\overline{B}_{\varepsilon+D}(\vec{p})$ (which is a ball containing ${}^{\|\cdot\|}\overline{B}_\varepsilon(\vec{p})$ and all points within distance D of it). Let $X \in {}^{\|\cdot\|}\overline{\mathcal{N}}_\varepsilon(\vec{p})$ be arbitrary and note that because $\operatorname{diam}_\infty(X) \leq D$, we have $X \subseteq {}^{\|\cdot\|}\overline{B}_{\varepsilon+D}(\vec{p})$. Thus $\bigsqcup_{X \in {}^{\|\cdot\|}\overline{\mathcal{N}}_\varepsilon(\vec{p})} \operatorname{int}(X) \subseteq {}^{\|\cdot\|}\overline{B}_{\varepsilon+D}(\vec{p})$. By Theorem 3.5.1 (Equivalence of Norms on $\mathbb{R}^d$), $m({}^{\|\cdot\|}\overline{B}_{\varepsilon+D}(\vec{p}))$ is finite (see justification[2]). By a simple volume argument, we have that ${}^{\|\cdot\|}\overline{\mathcal{N}}_\varepsilon(\vec{p})$ has cardinality at most $\frac{m\left({}^{\|\cdot\|}\overline{B}_{\varepsilon+D}(\vec{p})\right)}{\mu} < \infty$.

□

The following fact is a standard fact about finite families of sets, though we have not before seen it stated for locally finite families of sets (though it is certainly known).

Fact 3.6.4 (Locally Finite: Closure of Union = Union of Closures)**.** *Let $d \in \mathbb{N}$ and $\mathcal{F}$ be a locally finite family of subsets of $\mathbb{R}^d$. Then for any $\mathcal{S} \subseteq \mathcal{F}$ we have $\bigcup_{X \in \mathcal{S}} \overline{X} = \overline{\bigcup_{X \in \mathcal{S}} X}$.*

Proof. The containment $\bigcup_{X \in \mathcal{S}} \overline{X} \subseteq \overline{\bigcup_{X \in \mathcal{S}} X}$ is true in general for any family of sets $\mathcal{S}$ in any topology, so we must only show the other containment; this will hold because we can essentially reduce the case where $\mathcal{S}$ has infinite cardinality to a case where it has finite cardinality.

Let $\vec{p} \in \overline{\bigcup_{X \in \mathcal{S}} X}$. By definition of closure, there is a sequence $\langle \vec{x}^{(n)} \rangle_{n=1}^\infty$ of points in $\bigcup_{X \in \mathcal{S}} X$ converging to $\vec{p}$. Because $\mathcal{F}$ is a locally finite family, it follows trivially that $\mathcal{S}$ is also a locally finite family. Thus, by the definition of locally finite, let $\varepsilon \in (0, \infty)$ be such that ${}^\infty\overline{\mathcal{N}}_\varepsilon(\vec{p}) = \left\{X \in \mathcal{S} : X \cap {}^\infty\overline{B}_\varepsilon(\vec{p}) \neq \emptyset\right\}$ has finite cardinality.

Let $N \in \mathbb{N}$ be such that for $n > N$, $\left\|\vec{x}^{(n)} - \vec{p}\right\|_\infty \leq \varepsilon$ so that for $n > N$ we have $\vec{x}^{(n)} \in {}^\infty\overline{B}_\varepsilon(\vec{p})$. Thus, for $n > N$, because $\vec{x}^{(n)} \in \bigcup_{X \in \mathcal{S}} X$, there is some set $X^{(n)} \in \mathcal{S}$

[2] By Theorem 3.5.1 (Equivalence of Norms on $\mathbb{R}^d$), there exists some constant C such that ${}^{\|\cdot\|}\overline{B}_{\varepsilon+D}(\vec{p}) \subseteq {}^\infty\overline{B}_C(\vec{p})$. The latter has measure $(2C)^d$ so the former also has finite measure.

with $\vec{x}^{(n)} \in X^{(n)}$ and because also $\vec{x}^{(n)} \in {}^{\infty}\overline{B}_\varepsilon(\vec{p})$, we have $X^{(n)} \cap {}^{\infty}\overline{B}_\varepsilon(\vec{p}) \neq \emptyset$ which demonstrates that $X^{(n)} \in {}^{\infty}\overline{\mathcal{N}}_\varepsilon(\vec{p})$. Thus, for $n > N$ we have $\vec{x}^{(n)} \in \bigcup_{X \in {}^{\infty}\overline{\mathcal{N}}_\varepsilon(\vec{p})} X$ (which is a finite union by choice of ε) which shows that $\vec{p} \in \overline{\bigcup_{X \in {}^{\infty}\overline{\mathcal{N}}_\varepsilon(\vec{p})} X}$. A standard topological fact is that the closure of a finite union is the same as the finite union of the closures, so we have

$$\vec{p} \in \overline{\bigcup_{X \in {}^{\infty}\overline{\mathcal{N}}_\varepsilon(\vec{p})} X} = \bigcup_{X \in {}^{\infty}\overline{\mathcal{N}}_\varepsilon(\vec{p})} \overline{X} \subseteq \bigcup_{X \in \mathcal{S}} \overline{X}$$

which proves the other containment. □

The next fact will allow a technique that we will frequently apply. In locally finite families, to consider neighborhood at a single point, we may instead consider an ε neighborhood for some small ε. This will be especially convenient in the context of partitions because we can consider the disjoint members rather than the closures of the members which will in general not be disjoint. Roughly, in the proof, we "zoom in" to a point $\vec{p}$ far enough to only "see" finitely many sets, and then "zoom in" further just long enough to longer "see" any set which does not contain $\vec{p}$ in the closure.

Fact 3.6.5 (Locally Finite: Enlarged Neighborhood)**.** *Let $d \in \mathbb{N}$ and $\|\cdot\|$ any norm on $\mathbb{R}^d$ and $\mathcal{F}$ a locally finite family of subsets of $\mathbb{R}^d$ and $\vec{p} \in \mathbb{R}^d$. Then for all sufficiently small $\varepsilon > 0$ we have ${}^{\|\cdot\|}\overline{\mathcal{N}}_\varepsilon(\vec{p}) = {}^{\|\cdot\|}\mathcal{N}^\circ_\varepsilon(\vec{p}) = \mathcal{N}_{\vec{0}}(\vec{p})$.*

Proof. Observe that for any $\varepsilon \in (0, \infty)$, we have $\mathcal{N}_{\vec{0}}(\vec{p}) \subseteq {}^{\|\cdot\|}\mathcal{N}^\circ_\varepsilon(\vec{p}) \subseteq {}^{\|\cdot\|}\overline{\mathcal{N}}_\varepsilon(\vec{p})$ (see justification[3]). Thus, we only need to show that ${}^{\|\cdot\|}\overline{\mathcal{N}}_\varepsilon(\vec{p}) \subseteq \mathcal{N}_{\vec{0}}(\vec{p})$.

[3]For the latter containment, trivially, ${}^{\|\cdot\|}\mathcal{N}^\circ_\varepsilon(\vec{p}) \subseteq {}^{\|\cdot\|}\overline{\mathcal{N}}_\varepsilon(\vec{p})$ because ${}^{\|\cdot\|}B^\circ_\varepsilon(\vec{p}) \subseteq {}^{\|\cdot\|}\overline{B}_\varepsilon(\vec{p})$. For the former containment, if $X \in \mathcal{N}_{\vec{0}}(\vec{p})$ then $\vec{p} \in \overline{X}$ by definition, so X contains points arbitrarily close to $\vec{p}$ (with respect to every norm including $\|\cdot\|$), so ${}^{\|\cdot\|}B^\circ_\varepsilon(\vec{p}) \cap X \neq \emptyset$ so $X \in {}^{\|\cdot\|}\mathcal{N}^\circ_\varepsilon(\vec{p})$ by definition.

By local finiteness, let $\delta \in (0,\infty)$ such that ${}^{\|\cdot\|}\overline{\mathcal{N}}_\delta(\vec{p})$ has finite cardinality. For a set $X \subseteq \mathbb{R}^d$, let $d_{\|\cdot\|}(\vec{p},X) \overset{\text{def}}{=} \inf_{\vec{x}\in X}\|\vec{x}-\vec{p}\|$ noting that $d_{\|\cdot\|}(\vec{p},X) = 0$ if and only if $\vec{p} \in \overline{X}$. Let

$$\begin{aligned} C &= \left\{ d_{\|\cdot\|}(\vec{p},X) : X \in {}^{\|\cdot\|}\overline{\mathcal{N}}_\delta(\vec{p}) \text{ and } d_{\|\cdot\|}(\vec{p},X) > 0 \right\} \\ &= \left\{ d_{\|\cdot\|}(\vec{p},X) : X \in {}^{\|\cdot\|}\overline{\mathcal{N}}_\delta(\vec{p}) \text{ and } \vec{p} \notin \overline{X} \right\}. \end{aligned}$$

Because ${}^{\|\cdot\|}\overline{\mathcal{N}}_\delta(\vec{p})$ naturally surjects onto C, it follows that C has finite cardinality and thus has some minimum; let $c = \min(C)$. Note that $c \le \delta$ because for every $X \in {}^{\|\cdot\|}\overline{\mathcal{N}}_\delta(\vec{p})$ we have by definition that X intersects ${}^{\|\cdot\|}\overline{B}_\delta(\vec{p})$ so $d_{\|\cdot\|}(\vec{p},X) \le \delta$.

Let $\varepsilon \in (0,c)$ be arbitrary. To see that ${}^{\|\cdot\|}\overline{\mathcal{N}}_\varepsilon(\vec{p}) \subseteq \mathcal{N}_{\overline{0}}(\vec{p})$ let $X \in {}^{\|\cdot\|}\overline{\mathcal{N}}_\varepsilon(\vec{p})$ be arbitrary (so ${}^{\|\cdot\|}\overline{B}_\varepsilon(\vec{p}) \cap X \neq \emptyset$) noting that this implies $d_{\|\cdot\|}(\vec{p},X) \le \varepsilon < c$. Since $\varepsilon < c \le \delta$ we also have $X \in {}^{\|\cdot\|}\overline{\mathcal{N}}_\delta(\vec{p})$ and because $\varepsilon \notin C$, this implies by definition of C that $d_{\|\cdot\|}(\vec{p},X) = 0$ so $\vec{p} \in \overline{X}$ and thus be definition of neighborhood, $X \in \mathcal{N}_{\overline{0}}(\vec{p})$. □

The main corollary we will use from above is that unit cube packings (which includes tilings and partitions as special cases) are locally finite families.

Corollary 3.6.6 (Unit Cube Packings are Locally Finite)**.** *Let $d \in \mathbb{N}$, and $\mathcal{F}$ be a packing of unit cubes in $\mathbb{R}^d$ (i.e. a set of unit cubes with pairwise disjoint interiors). Then $\mathcal{F}$ is locally finite.*

Proof. Each member of $\mathcal{F}$ has Lebesgue measure[4] 1 and diameter 1 so by Fact 3.6.3 $\mathcal{F}$ is locally finite. □

The following says that largely for the purposes we are interested in, we don't need to distinguish between closed axis-aligned unit cubes or half-open axis-aligned unit

[4]To address the measurability, because the cube X is contained between an open unit cube $\vec{x}+(0,1)^d$ and a closed unit cube $\vec{x}+[0,1]^d$ by definition, $X = (0,1)^d \sqcup S$ where $S \subseteq \vec{x}+([0,1]^d \setminus (0,1)^d)$, so S is a subset of a null set and is thus Lebesgue measurable, and thus so is X.

cubes—in particular in the case of tilings versus partitions (see Fact 3.6.8). Also, we won't need to distinguish between the two in the finite structures (cube enclosures) we work with in Chapter 9 (Secluded Partitions Without ε). While cube tilings tend to be more common in the literature (see Section 1.2 (Mathematical Motivation and Background)), we choose to work primarily with partitions in this work because they are generally simpler to work with. The main reason is that points in the union of the cubes belong to a unique cube, and this property allows for avoiding a lot of really annoying details about boundary cases. Another consequence is that an axis-aligned unit cube partition $\mathcal{P}$ of $\mathbb{R}^d$ induces an n-dimensional axis-aligned unit cube partition of on every axis aligned n-dimensional affine subspace of $\mathbb{R}^d$. For example, Szabó [Sza86] worked with tilings and had to go through some amount of work to deal with the fact that a tiling does not as easily induce a tiling in lower dimensional subspaces because points might belong to multiple cubes; working instead with partitions avoids this altogether.

Corollary 3.6.7 (Unit Cube Packings Closed or Half-Open)**.** *Let $d \in \mathbb{N}$, and $\mathcal{F}$ be a packing of axis-aligned unit cubes in $\mathbb{R}^d$. Letting $\mathcal{F}_{clos} = \left\{\overline{X} : X \in \mathcal{F}\right\}$ be the closed versions of all cubes and $\mathcal{F}_{half} = \left\{H_{1/2}(\text{center}(X)) : X \in \mathcal{F}\right\}$ be the half-open versions of all cubes, we have that $\mathcal{F}_{clos}$ and $\mathcal{F}_{half}$ are also unit cube packings, and*

$$\text{int}\left(\bigcup_{X\in\mathcal{F}} X\right) \subseteq \text{int}\left(\bigcup_{X\in\mathcal{F}_{clos}} X\right) = \text{int}\left(\bigcup_{X\in\mathcal{F}_{half}} X\right)$$

and all cubes in $\mathcal{F}_{half}$ are disjoint.

Proof. That $\mathcal{F}_{clos}$ and $\mathcal{F}_{half}$ are also unit cube packings follows from Fact 3.4.3 because the interior is the same regardless of the type of cube, so the interiors of the half-open cubes or the closed cubes remain disjoint.

The containment is trivial because $\bigcup_{X\in\mathcal{F}} X \subseteq \bigcup_{X\in\mathcal{F}_{clos}} X$, and the "$\supseteq$" direction of the equality is trivial by the same reasoning. Thus, we must show that $\operatorname{int}\left(\bigcup_{X\in\mathcal{F}_{clos}} X\right) \subseteq \operatorname{int}\left(\bigcup_{X\in\mathcal{F}_{half}} X\right)$.

Let $\vec{p} \in \operatorname{int}\left(\bigcup_{X\in\mathcal{F}_{clos}} X\right)$ be arbitrary. By definition of interior, there exists some $\varepsilon \in (0,\infty)$ such that ${}^{\infty}\overline{B}_{\varepsilon}(\vec{p}) \subseteq \operatorname{int}\left(\bigcup_{X\in\mathcal{F}_{clos}} X\right)$. Furthermore, by Corollary 3.6.6 and Fact 3.6.5 we may assume ε is small enough that ${}^{\|\cdot\|}\overline{\mathcal{N}}_{\varepsilon}(\vec{p}) = \mathcal{N}_{\vec{0}}(\vec{p})$.

Let $\vec{q} = \vec{p} + \varepsilon \cdot \vec{1}$. Then we have $\vec{q} \in {}^{\infty}\overline{B}_{\varepsilon}(\vec{p}) \subseteq \operatorname{int}\left(\bigcup_{X\in\mathcal{F}_{clos}} X\right) \subseteq \bigcup_{X\in\mathcal{F}_{clos}} X$, so there is some $Y \in \mathcal{F}_{clos}$ such that $\vec{q} \in Y$. Furthermore by definition of neighborhood, $\vec{q}$ witnesses that $Y \in {}^{\|\cdot\|}\overline{\mathcal{N}}_{\varepsilon}(\vec{p}) = \mathcal{N}_{\vec{0}}(\vec{p})$ which shows that $\vec{p} \in \overline{Y} = Y$ (because all cubes in $\mathcal{F}_{clos}$ are closed). Because Y is a closed cube (a product set) which contains $\vec{p}$ and $\vec{q}$, we have that $Y \supseteq \prod_{i=1}^{d}[\min(p_i,q_i),\max(p_i,q_i)] = \prod_{i=1}^{d}[p_i,p_i+\varepsilon] \supseteq \prod_{i=1}^{d}(p_i,p_i+\varepsilon)$. Since $\prod_{i=1}^{d}(p_i,p_i+\varepsilon)$ is an open subset of Y, it is in fact a subset of the interior of Y. Letting $\vec{y} = \operatorname{center}(Y)$, this gives the following:

$$\prod_{i=1}^{d}(p_i,p_i+\varepsilon) \subseteq \operatorname{int}(Y) \subseteq H_{1/2}(\vec{y}) = \prod_{i=1}^{d}[y_i - \tfrac{1}{2}, y_i + \tfrac{1}{2}).$$

Then, since $\prod_{i=1}^{d}[y_i - \frac{1}{2}, y_i + \frac{1}{2})$ is "left closed" it contains $\vec{p}$ by the following: the above containment shows for all $i \in [d]$ that $y_i - \frac{1}{2} \leq p_i < p_i + \varepsilon \leq y_i + \frac{1}{2}$ so $p_i \in [y_i - \frac{1}{2}, y_i + \frac{1}{2})$. Thus $\vec{p} \in \prod_{i=1}^{d}[y_i - \frac{1}{2}, y_i + \frac{1}{2}) = H_{1/2}(\vec{y})$.

Finally, the fact that all cubes in $\mathcal{F}_{half}$ are disjoint is because for two distinct cubes $X, Y \in \mathcal{F}$ we have by assumption that X and Y have disjoint interiors which occurs if an only if $\|\operatorname{center}(X) - \operatorname{center}(Y)\|_{\infty} \geq 1$ (because $\operatorname{int}(X) = {}^{\infty}B^{\circ}_{1/2}(\operatorname{center}(X))$ and $\operatorname{int}(Y) = {}^{\infty}B^{\circ}_{1/2}(\operatorname{center}(Y))$). Thus, letting $\vec{x} = \operatorname{center}(X)$ and $\vec{y} = \operatorname{center}(Y)$ we have some coordinate $i_0 \in [d]$ such that $|x_i - y_i| \geq 1$ and

$$H_{1/2}(\vec{x}) = \prod_{i=1}^{d}[x_i - \tfrac{1}{2}, x_i + \tfrac{1}{2})$$
$$H_{1/2}(\vec{y}) = \prod_{i=1}^{d}[y_i - \tfrac{1}{2}, y_i + \tfrac{1}{2}).$$

Thus $[x_{i_0} - \frac{1}{2}, x_{i_0} + \frac{1}{2})$ and $[y_{i_0} - \frac{1}{2}, y_{i_0})$ are disjoint intervals, so $H_{1/2}(\vec{x})$ and $H_{1/2}(\vec{y})$ are disjoint cubes. □

As a corollary, we get the unsurprising fact that tilings and partitions of $\mathbb{R}^d$ are equivalent in the sense of the sets of positions which define them. Though unsurprising, we do want to establish this result so that we know with certainty that we can work with partitions instead of or as a proxy to tilings.

Fact 3.6.8 (Equivalence of Axis-Aligned Partitions and Tilings)**.** *Let* $d \in \mathbb{N}$ *and* $\Lambda \subseteq \mathbb{R}^d$*. Then* $\mathcal{P} \overset{\text{def}}{=} \left\{H_{1/2}(\vec{\lambda}) : \lambda \in \Lambda\right\}$ *is a partition of* $\mathbb{R}^d$ *if and only if* $\mathcal{T} \overset{\text{def}}{=} \left\{{}^{\infty}\overline{B}_{1/2}(\vec{\lambda}) : \lambda \in \Lambda\right\}$ *is a tiling of* $\mathbb{R}^d$*.*

Proof. If $\mathcal{T}$ is a tiling (i.e. it covers $\mathbb{R}^d$ and all cubes have disjoint interiors), then by definition of a cover we have $\bigcup_{X \in \mathcal{T}} X = \mathbb{R}^d$ and so $\operatorname{int}\left(\bigcup_{X \in \mathcal{T}} X\right) = \mathbb{R}^d$, then by Corollary 3.6.7 (because all cubes are axis-aligned), $\operatorname{int}\left(\bigcup_{Y \in \mathcal{P}} Y\right) = \mathbb{R}^d$ (which implies $\bigcup_{Y \in \mathcal{P}} Y = \mathbb{R}^d$) and all cubes in $\mathcal{P}$ are disjoint. Thus, $\mathcal{P}$ is a partition.

Conversely, if $\mathcal{P}$ is a partition of $\mathbb{R}^d$, then clearly $\mathcal{T}$ is a cover of $\mathbb{R}^d$ (because each member of $\mathcal{P}$ is a subset of a member of $\mathcal{T}$). To see the disjointness of interiors, consider any distinct $X, Y \in \mathcal{T}$ and let $\vec{x} = \operatorname{center}(X) \in \Lambda$ and $\vec{y} = \operatorname{center}(Y) \in \Lambda$. Then $\operatorname{int}(X) = {}^{\infty}B^{\circ}_{1/2}(\vec{x}) \subseteq H_{1/2}(\vec{x}) \in \mathcal{P}$ and similarly $\operatorname{int}(Y) = {}^{\infty}B^{\circ}_{1/2}(\vec{y}) \subseteq H_{1/2}(\vec{y}) \in \mathcal{P}$. Because $H_{1/2}(\vec{x})$ and $H_{1/2}(\vec{y})$ are distinct members of the partition $\mathcal{P}$, they are disjoint by definition, so it follows by subsets that the interiors of X and Y are also disjoint. $\square$

3.7 Partition Neighborhoods

The following fact gives alternate expressions for ${}^{\|\cdot\|}\overline{\mathcal{N}}_\varepsilon(\vec{p})$ and ${}^{\|\cdot\|}\mathcal{N}^\circ_\varepsilon(\vec{p})$ in the case of a partition. It says nothing surprising—only that the set of members of partition $\mathcal{P}$ that intersect the ε ball is exactly the set obtained by considering the member associated to each point in the ε ball.

Fact 3.7.1 (Neighborhood Expression by Members)**.** *Let $d \in \mathbb{N}$ and $\varepsilon \in (0, \infty)$ and $\|\cdot\|$ any norm on $\mathbb{R}^d$ and $\mathcal{P}$ an axis-aligned unit cube partition of $\mathbb{R}^d$. Then*

$$ {}^{\|\cdot\|}\overline{\mathcal{N}}_\varepsilon(\vec{p}) = \left\{ \text{member}(\vec{x}) : \vec{x} \in {}^{\|\cdot\|}\overline{B}_\varepsilon(\vec{p}) \right\} $$

and

$$ {}^{\|\cdot\|}\mathcal{N}^\circ_\varepsilon(\vec{p}) = \left\{ \text{member}(\vec{x}) : \vec{x} \in {}^{\|\cdot\|}B^\circ_\varepsilon(\vec{p}) \right\}. $$

Proof. We prove only the former, as the latter is identical. If $X \in {}^{\|\cdot\|}\overline{\mathcal{N}}_\varepsilon(\vec{p})$, then $X \cap {}^{\|\cdot\|}\overline{B}_\varepsilon(\vec{p}) \neq \emptyset$ by definition, so let $\vec{x}^{(0)} \in X \cap {}^{\|\cdot\|}\overline{B}_\varepsilon(\vec{p})$ which means $\text{member}(\vec{x}^{(0)}) = X$, so $X \in \left\{ \text{member}(\vec{x}) : \vec{x} \in {}^{\|\cdot\|}\overline{B}_\varepsilon(\vec{p}) \right\}$. Conversely, if $X \in \left\{ \text{member}(\vec{x}) : \vec{x} \in {}^{\|\cdot\|}\overline{B}_\varepsilon(\vec{p}) \right\}$ then there exists $\vec{x}^{(0)} \in {}^{\|\cdot\|}\overline{B}_\varepsilon(\vec{p})$ such that $X = \text{member}(\vec{x}^{(0)})$, so in particular $\vec{x}^{(0)} \in X$ and thus $X \cap {}^{\|\cdot\|}\overline{B}_\varepsilon(\vec{p}) \neq \emptyset$ so $X \in {}^{\|\cdot\|}\overline{\mathcal{N}}_\varepsilon(\vec{p})$. $\square$

For the ℓ_∞ norm and small enough ε, we don't even need to consider the members of all points, just the members of the corner points of the ball (for the closed ball case at least).

Fact 3.7.2 (Neighborhood Expression by Corners)**.** *Let $d \in \mathbb{N}$ and $\mathcal{P}$ an axis-aligned unit cube partition of $\mathbb{R}^d$. Then for each $\vec{p} \in \mathbb{R}^d$, the following holds for all sufficiently small $\varepsilon > 0$:*

$$\begin{aligned}\mathcal{N}_{\vec{0}}(\vec{p}) &= \left\{\text{member}(\vec{x}) : \vec{x} \in \text{corners}\left({}^{\infty}\overline{B}_{\varepsilon}(\vec{p})\right)\right\} \\ &= \left\{\text{member}(\vec{x}) : \vec{x} \in \vec{p} + \{-\varepsilon, \varepsilon\}^d\right\}.\end{aligned}$$

Proof. Fix $\vec{p} \in \mathbb{R}^d$. All sets in the claim above are members of $\mathcal{P}$, so let $X \in \mathcal{P}$ be arbitrary. Let $\varepsilon \in (0, \frac{1}{2})$ be sufficiently small so that Fact 3.6.5 (Locally Finite: Enlarged Neighborhood) holds for $\vec{p}$. Then we have

$$\begin{aligned}X \in \mathcal{N}_{\vec{0}}(\vec{p}) &\iff X \in {}^{\infty}\overline{\mathcal{N}}_{\varepsilon}(\vec{p}) && \text{(By Fact 3.6.5)} \\ &\iff X \cap {}^{\infty}\overline{B}_{\varepsilon}(\vec{p}) \neq \emptyset && \text{(Def'n of neighborhood)} \\ &\iff X \cap \text{corners}\left({}^{\infty}\overline{B}_{\varepsilon}(\vec{p})\right) \neq \emptyset && \text{(See below)}\end{aligned}$$

The reverse direction of the above is trivial: if X intersects the corners, then it intersects the ball, because the closed ball contains all of its corners. The forward direction is because we may assume $\varepsilon \in (0, \frac{1}{2})$ (since we only claim the result for "sufficiently small ε"), so it follows from Fact 3.4.13 (Containing Corners of Small Rectangles). We continue.

$$\iff X \in \left\{\text{member}(\vec{x}) : \vec{x} \in \text{corners}\left({}^{\infty}\overline{B}_{\varepsilon}(\vec{p})\right)\right\} \qquad \text{(See below)}$$

For the reverse direction above, if X is the member of one of the corners, then it contains that corner, so it intersects the set of corners. For the forward direction, if X intersects the corners, then it contains one of the corners, so it is the member of that corner.

Thus, by examining an arbitrary member X that could be in any of these sets, we have established the claim that

$$\mathcal{N}_{\vec{0}}(\vec{p}) = \left\{ \text{member}(\vec{x}) : \vec{x} \in \text{corners}\left({}^{\infty}\overline{B}_{\varepsilon}(\vec{p}) \right) \right\}.$$

Noting that $\text{corners}\left({}^{\infty}\overline{B}_{\varepsilon}(\vec{p}) \right) = \vec{p} + \{-\varepsilon, \varepsilon\}^{d}$ proves the stated claim. □

Chapter 4

Constructions

In this chapter, we will construct a specific family of axis-aligned unit cube partitions (which we call reclusive partitions) which are (k, ε)-secluded with $k = d + 1$ and $\varepsilon = \frac{1}{2d}$. The word reclusive in English 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, and we will use the term reclusive partition to talk about unit cube partitions which have the very specific linear algebra structure that we develop in this section.

Remark 4.0.1. After having completed the reclusive constructions that are discussed in Section 4.2 (Reclusive Partitions), we became aware of work by Hoza and Klivans [HK18] which implicitly constructed partitions with similar parameters—though they were not unit cube partitions. For each $d \in \mathbb{N}$, they defined a deterministic rounding function $f : \mathbb{R}^d \to \mathbb{R}^d$ (which naturally induces the partition $\mathcal{P} = \{f^{-1}(\vec{y}) : \vec{y} \in \text{range}(f)\}$) which scaled appropriately gives a $(d + 1, \frac{1}{6(d+1)})$-secluded partition of $\mathbb{R}^d$ with sets of ℓ_∞ diameter at most 1 (so it is on the same footing as unit cube partitions in this regard). Furthermore, every set in this partition is identical up to translation, and the partition is efficiently computable in the sense that it is possible to efficiently compute the member of the partition in which a point belongs. See Section F.4 (The Deterministic Rounding Scheme of Hoza and Klivans) for additional details including Figure F.1 which shows the partition for $\mathbb{R}^2$.

We were aware of techniques similar to this due to Goldreich [Gol19b], but the rounding function in that case was a randomized function (i.e. randomized algorithm), so it did not give rise to a partition. See Section F.3 (The Randomized Rounding Scheme of Goldreich) for additional details.

Nonetheless, Goldreich was able to obtain parameters of $d+1$ and $O(\frac{1}{d})$ in the context of randomized computation and this was what originally led us to believe that (k,ε)-secluded partitions with $k=d+1$ and $\varepsilon=O(\frac{1}{d})$ existed and we hoped we could construct them only with unit cubes.

For perspective, this means that from the theoretical computation point of view, our reclusive unit cube partitions are not a significant advancement as they offer nothing asymptotically better than the partitions of Hoza and Klivans. However, from a mathematical perspective, the fact that this can be done with unit cubes is quite nice. Furthermore, the constructions that we offer in Section 4.4 (New Partitions From Old) which use either our reclusive partitions or the partitions of Hoza and Klivans as building blocks *are* new from the computational point of view. $\triangle$

We will begin by giving an example of a reclusive partition which has worse parameters than this. The example serves to give a quick introduction to how we conceptually will view the larger class of all reclusive partitions. Following the example, we will define the reclusive partitions in full generality using motivations found in the example. After this, we will give a discussion of why we believe the reclusive partitions should be of mathematical interest, and how they capture (at least to some extent) a fundamental property of matrix/lattice based unit cube partitions. Then, we will discuss how to use the reclusive partitions, or in fact any partitions which are (k, ε)-secluded for some values of k and ε, and show how to build new partitions out of the old ones. This will give explicit constructions which make a trade-off between the k and ε parameters.

4.1 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 $\frac{1}{2^{d-1}}$ (when $k = d+1$) and as we have mentioned, some of the reclusive partitions will achieve ε as large as $\frac{1}{2d}$ (when $k = d + 1$). 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 cubes. After presenting the definition, we elaborate on how to interpret it geometrically.

Definition 4.1.1 ($\mathcal{P}_d$). For each $d \in \mathbb{N}$, define[a] $\vec{v}^{(d)} \stackrel{\text{def}}{=} \langle \frac{1}{2^{d-1}}, \ldots, \frac{1}{2^{d-1}}, 1 \rangle$, 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:

$$\mathcal{P}_1 \stackrel{\text{def}}{=} \{[0,1) + n \colon n \in \mathbb{Z}\} = \left\{[0,1) + n \cdot \vec{v}^{(1)} \colon n \in \mathbb{Z}\right\}.$$

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:

$$\mathcal{P}_d \stackrel{\text{def}}{=} \left\{B \times [0,1) + n \cdot \vec{v}^{(d)} \colon n \in \mathbb{Z},\ B \in \mathcal{P}_{d-1}\right\}$$

[a] By this definition, in the case that $d = 1$, $v^{(d)} = \langle 1 \rangle$.

The following discussion motivates why we are interested in this partition and how to understand it geometrically. The first partition ($\mathcal{P}_d = \mathcal{P}_1$) breaks up $\mathbb{R}^d = \mathbb{R}^1$ into unit intervals which are half open. This partition has the property that for any point $x \in \mathbb{R}^1$, if you consider all points with 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 4.1).

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 $n = 0$ 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 4.1), the point $(1, 1)$ would be at the closure 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 one in the prior dimension, 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 (though the proof will probably not lend insight to the rest of the paper).

Claim 4.1.2. *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 so that the last coordinate is never an issue, this holds if and only if $\langle x_i - n \cdot v_i^{(d)} \rangle_{i=1}^{d-1} \in B$. By the inductive hypothesis, there exists a unique $B \in \mathcal{P}_{d-1}$ such that this holds. Thus $\mathcal{P}_d$ partitions $\mathbb{R}^d$. □

Based on our discussion above, we hope we have provided the intuition that each member of any $\mathcal{P}_d$ is a unit cube 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 cube partitions.

4.1.1 Motivating Properties

While the inductive definition of the running example partitions $\mathcal{P}_d$ is useful, it is also useful to consider unit cube partitions from another perspective. One can note that in the partition $\mathcal{P}_d$, the representative corner of each unit cube is an integer linear combination of the vectors $\vec{v}^{(1)}, \ldots, \vec{v}^{(d)}$ as defined in Definition 4.1.1 (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 cube representatives as points within the lattice group will be useful (this algebraic view is the same one that Minkowski had in formulating his conjecture (see Section 1.2)). In particular, this gives motivation to look at certain regularly structured unit cube 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 4.1.1 into $\mathbb{R}^5$ (by padding with zeros), 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 cubes whose representatives were integer linear combinations of the columns of this matrix; in other words, $\mathcal{P}_5$ could have been defined as the set of unit cubes 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 cube 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 cubes. 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 cubes to prevent points in $\mathbb{R}^d$ from being "close" to too many cubes. 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 cube partition, but it would not have the property we ultimately desire of limiting the adjacencies (in fact, this is the partition in Figure 4.1).

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 (we will give no justification of this). Instead, we would like the shifts to change linearly and work with partitions that have matrices more like one of the ones below where any two entries in the same row differ by at least $\frac{1}{5}$:

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

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 these matrices have the same structural properties mentioned above.

4.2 Reclusive Partitions

We now define in full generality the type of matrices that we will be interested in based on the ideas just discussed.

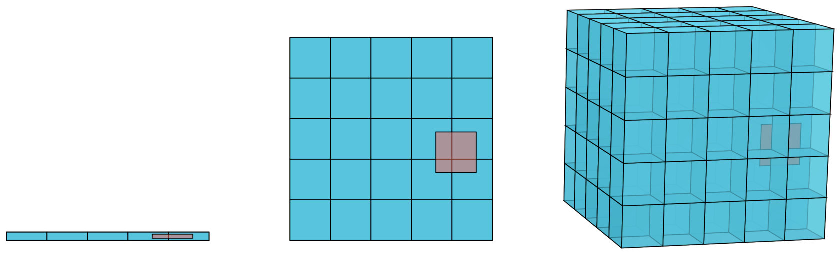


Figure 4.1: Simple unit cube partitions of $\mathbb{R}^d$ for $d = 1, 2, 3$. The shaded red regions are $\varepsilon = \frac{1}{2}$ radius balls (in the ℓ_∞ norm) showing that these neighborhoods don't intersect more than $k = 2^d$ members of the partition (it might look from the picture that centering the ball in one of the cubes would intersect more, but it does not because the cubes are half open). Furthermore, there are cubes that do intersect 2^d members as shown.

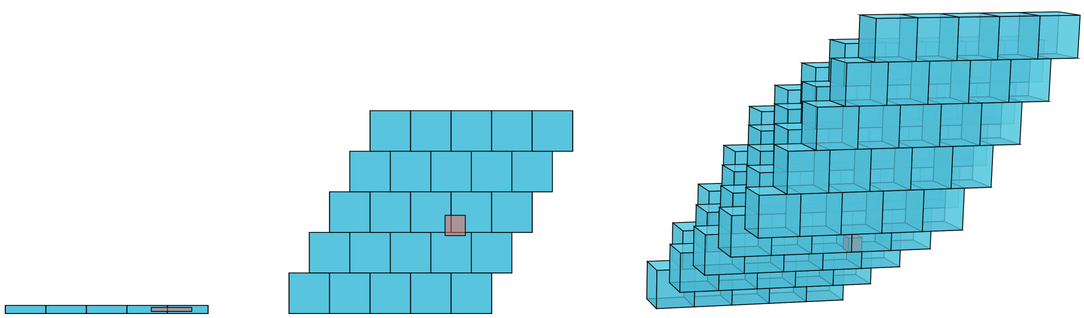


Figure 4.2: This is a particular reclusive partition of $\mathbb{R}^d$ for $d = 1, 2, 3$. The shaded red regions are $\varepsilon = \frac{1}{2d}$ radius balls (in the ℓ_∞ norm) showing that these neighborhoods don't intersect more than $k = d + 1$ members of the partition. Observe that in the $d = 2$ case, the x-direction shift (horizontal) is $\frac{1}{2}$ unit; in the $d = 3$ case, there are shifts of $\frac{2}{3}$ in the y-direction (vertical) and shifts of both $\frac{1}{3}$ and $\frac{2}{3}$ in the x-direction (horizontal). In dimensions $d \geq 5$, the "order" in which these shifts occur is vital as discussed in Section 4.3 (Fundamental Property of the Reclusive Definition).

Definition 4.2.1 (Reclusive Matrix). Informally, a square matrix A will be called a reclusive matrix if it (1) is upper triangular, (2) has only 1's on the main diagonal, and (3) 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 $i, j \in [n]$ with $i > j$, $a_{ij} = 0$ (upper triangular)
2. For all $i \in [n]$, $a_{ii} = 1$ (1's on diagonal)
3. For all $i, j, k \in [n]$ with $i \leq j < k$, $a_{ij} > a_{ik} > 0$ (decreasing non-zero)

Remark 4.2.2. 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 4.2.3 (Lattice Group). For any invertible $d \times d$ matrix A, define the set $L_A \stackrel{\text{def}}{=} \left\{A\vec{v} \colon \vec{v} \in \mathbb{Z}^d\right\}$.

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

Definition 4.2.4 (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 4.2.5. We could have also defined

$$\mathcal{P}_A \stackrel{\text{def}}{=} \left\{ H_{1/2}(\vec{a}) \colon \vec{a} \in L_A \right\} \\ = \left\{ \vec{a} + [-\tfrac{1}{2}, \tfrac{1}{2})^d \colon \vec{a} \in L_A \right\}$$

to use cubes centered at the points of the lattice which leads to no significant differences in results. We only use the definition we did, because it is more common to think of translates of $[0,1)^d$ rather than to think of translates of $[-\frac{1}{2}, \frac{1}{2})^d$. △

Remark 4.2.6. 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 because we reserve the $+$ notation for Minkowski sums in this dissertation as mentioned in Chapter 2 (Notation). △

Notice that every member of $\mathcal{P}_A$ is a unit cube shifted by an element of the lattice group so that for all $X \in \mathcal{P}_A$, $\operatorname{rep}(X) \in L_A$ and conversely, for every $\vec{a} \in L_A$ there is a unit cube X whose representative is $\vec{a}$.

The proof that this is a partition will not be presented as it is similar to the proof of Claim 4.1.2 and in fact a direct consequence of Proposition 10.2.3 (Efficient Computation of Reclusive Representatives) presented later 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 $\operatorname{rep}(X)$.

We will have use for another equivalent notion of adjacency in the reclusive partitions using algebraic properties of the positions, 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 4.2.7 (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 4.2.8. 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, but with additional details that will be necessary for us. 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 4.2.9 (Alternating Sequence Lemma). *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. *$\operatorname{sign}(s) = \operatorname{sign}(c_1)$ (resp. $\operatorname{sign}(s) = \operatorname{sign}(c_n)$)*

Proof. Note that (3) is implied by (1) for $n = 1$ and implied by (2) for $n \geq 2$ because $\langle a_i \rangle$ is a strictly decreasing (resp. strictly increasing) sequence of positive values. Thus (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

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

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

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

which proves (4). □

The reason we required $\langle a_i \rangle$ to be strictly monotonic as opposed to weakly monotonic was because otherwise $\operatorname{sign}(c_n)$ 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 equivalent algebraic definition of adjacency. The $A\vec{c}$ in the statement will end up being $\operatorname{rep}(X) - \operatorname{rep}(Y)$ for unit cubes X and Y, so this proposition will give an equivalent condition for $\|\operatorname{rep}(X) - \operatorname{rep}(Y)\|_\infty \leq 1$ (which by Fact 3.4.7 is equivalent to X and Y being adjacent).

Proposition 4.2.10 (Weak-Alt-1 Property for Reclusive Matrices)**.** *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 $i \in [d]$

$$\begin{aligned} x_i &= \sum_{j=1}^{d} a_{ij} c_j && \text{(Def'n of matrix multiplication)} \\ &= \sum_{j=i}^{d} a_{ij} c_j. && \text{(If } j < i \text{ then } a_{ij} = 0\text{)} \end{aligned}$$

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 4.2.9 (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

$$\begin{aligned} \|\vec{x}\|_\infty &= \max_{i \in [d]} |x_i| && \text{(Def'n)} \\ &\geq |x_k| \\ &= \left| \sum_{j=k}^{d} a_{kj} c_j \right| && \text{(By expression of } x_k\text{)} \\ &= \left| c_k + \sum_{j=k+1}^{d} a_{kj} c_j \right| && \text{(}A \text{ is reclusive, so } a_{kk} = 1\text{)} \end{aligned}$$

If $|c_d| > 1$ then $\langle c_j \rangle_{j=d}^{d}$ is not weak-alt-1, so $k = d$. Then 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 4.2.9 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 4.2.9 $|s| \leq a_{kj_1}$ (regardless of the value of n). This gives the following inequalities:

$$\begin{aligned}
\|\vec{x}\|_\infty &\geq |c_k + s| && \text{(Work above)} \\
&\geq |c_k| - |s| && \text{(Triangle inequality)} \\
&\geq 2 - |s| && \text{(Assumption on } |c_k|\text{)} \\
&\geq 2 - a_{kj_1} && \text{(Lemma 4.2.9)} \\
&= 1 + (1 - a_{kj_1}) && \\
&> 1 && (j_1 \geq k+1 \text{ so } a_{kj_1} < 1)
\end{aligned}$$

Case 2: If $|c_k| = 1$, then because $\langle c_j \rangle_{j=k}^d$ is not weak-alt-1, it implies that $\operatorname{sign}(c_k) =$

sign(c_{j_1}). We then get the following inequalities:

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

From here we have two cases depending on if $n = 1$ or $n \geq 2$. If $n = 1$, then by Lemma 4.2.9 $|s| = a_{kj_1}$, 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 4.2.9 $|s| \geq a_{kj_1} - a_{kj_2}$, so from the above we have

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

□

Upon inspection, one may note that we can improve upon the statement of Proposition 4.2.10. 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 $\{A\vec{v} \colon \vec{v} \in \mathbb{Z}^d\}$ is isomorphic to $\mathbb{Z}^d$ (as stated in the discussion of Definition 4.2.3), and this is a well-known property of lattices. Specifically, if $\vec{c}$ was not weak-alt-1, the above proof showed that one of the following four equations held:

$$\|A\vec{c}\|_\infty \geq 1 + 1 \qquad (|c_d| > 1)$$

$$\|A\vec{c}\|_\infty \geq 1 + (1 - a_{kj_1}) \qquad (|c_d| \leq 1 \text{ and } |c_k| \geq 2)$$

$$\|A\vec{c}\|_\infty \geq 1 + (a_{kj_1}) \qquad (|c_d| \leq 1 \text{ and } |c_k| = 1 \text{ and } n = 1)$$

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

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 consequence. In the definition below, the minimum of $(1 - a_{kj})$ should be understood as the minimum difference between entries past the diagonal of the matrix and the entries of the diagonal (which are all 1). The minimum of (a_{kj}) should be understood as the minimum difference between entries past the main diagonal and 0. And the minimum of $(a_{kj} - a_{kj'})$ should be understood as the minimum difference between two entries in the same row.

Definition 4.2.11 (Reclusive Distance)*.* If A is a $d \times d$ reclusive matrix, define Δ_A, the reclusive distance of A, as follows (taking $\min \emptyset = \infty$):

$$\delta_1 \overset{\text{def}}{=} 1$$

$$\delta_2 \overset{\text{def}}{=} \min_{k \in [d]} \min_{k < j \leq d} (1 - a_{kj})$$

$$\delta_3 \overset{\text{def}}{=} \min_{k \in [d]} \min_{k < j \leq d} (a_{kj})$$

$$\delta_4 \overset{\text{def}}{=} \min_{k \in [d]} \min_{k < j < j' \leq d} (a_{kj} - a_{kj'})$$

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

Observe that because A is reclusive, $\Delta_A > 0$ which follows immediate from the definition of a reclusive matrix because reclusive matrices are strictly decreasing with positive entries in each row starting at the main diagonal (Definition 4.2.1).

We can now state a strengthened version of Proposition 4.2.10.

Porism 4.2.12 (Weak-Alt-1 Property for Reclusive Matrices)**.** *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. Furthermore, 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 4.2.10. □

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

Lemma 4.2.13 (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 $X, Y \in \mathcal{P}_A$ be distinct cubes such that X and Y are not adjacent. Then $\|\mathrm{rep}(X) - \mathrm{rep}(Y)\|_\infty \geq 1 + \Delta_A$.*

Proof. By definition of $\mathcal{P}_A$, there exists $\vec{m}, \vec{n} \in \mathbb{Z}^d$ such that $A\vec{m} = \mathrm{rep}(X)$ and $A\vec{n} = \mathrm{rep}(Y)$ (in particular $\vec{m} = A^{-1}\,\mathrm{rep}(X)$ and $\vec{n} = A^{-1}\,\mathrm{rep}(Y)$). By Fact 3.4.7, because X and Y are not adjacent, $\|\mathrm{rep}(X) - \mathrm{rep}(Y)\|_\infty > 1$, so

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

By Porism 4.2.12, since $(\vec{m} - \vec{n}) \in \mathbb{Z}^d$ and $\|A(\vec{m} - \vec{n})\| > 1$ it must be that $\|A(\vec{m} - \vec{n})\| \geq 1 + \Delta_A$. □

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

Lemma 4.2.14 (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 $X, Y \in \mathcal{P}_A$ be distinct cubes such that X and Y are not adjacent. For all $\vec{x} \in X$ and $\vec{y} \in Y$, it holds that $\|\vec{x} - \vec{y}\|_\infty > \Delta_A$.*

Proof. We have $\vec{x} \in X = \operatorname{rep}(X) + [0,1)^d$ and $\vec{y} \in Y = \operatorname{rep}(Y) + [0,1)^d$, so let $\vec{\alpha}, \vec{\beta} \in [0,1)^d$ such that $\vec{x} = \operatorname{rep}(X) + \vec{\alpha}$ and $\vec{y} = \operatorname{rep}(Y) + \vec{\beta}$. As in the proof of the prior lemma, let $\vec{m} = A^{-1} \operatorname{rep}(X)$ and $\vec{n} = A^{-1} \operatorname{rep}(Y)$ so $\vec{m}, \vec{n} \in \mathbb{Z}^d$.

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:

$$\begin{aligned}
\|\vec{x} - \vec{y}\|_\infty &= \left\|(A\vec{m} + \vec{\alpha}) - (A\vec{n} + \vec{\beta})\right\|_\infty \\
&= \left\|(A(\vec{m} - \vec{n})) - (\vec{\beta} - \vec{\alpha})\right\|_\infty \\
&\geq \|(A(\vec{m} - \vec{n}))\|_\infty - \|\vec{\beta} - \vec{\alpha}\|_\infty && \text{(Triangle inequality reordered)} \\
&\geq (1 + \Delta_A) - \|(\vec{\beta} - \vec{\alpha})\|_\infty && \text{(Lemma 4.2.13)} \\
&> (1 + \Delta_A) - 1 && (\|\vec{\beta} - \vec{\alpha}\|_\infty < 1) \\
&= \Delta_A
\end{aligned}$$

Noting the strict inequality in the second to last line completes the proof. □

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 cubes can be, we want to turn our attention to how many cubes 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 cubes). A trivial result in graph theory is that 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 4.2.15 (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 cubes as follows:

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

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

$\vec{n} = A^{-1}\,\mathrm{rep}(Y)$. Then we have the following chain of equalities:

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

Observe that by Proposition 4.2.10, because $X \overset{\mathrm{adj}}{\sim} Y$ 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 $\mathrm{rep}(X) = \mathrm{rep}(Y)$ 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 4.2.9 (applied to the (non-empty) subsequence of non-zero terms), we have that the magnitude of the dot product is positive and at most the value of the last term in the increasing sequence (i.e. d). That is, we have

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

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

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 trivial corollary to Theorem 5.1.1 (Optimality of $k = d + 1$ for Cube Partitions), which appears in the next chapter, so the following proof is not strictly necessary; however, Theorem 5.1.1 uses more machinery, and it is fairly simple to find an explicit clique in the reclusive partitions.

Proposition 4.2.16 (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. Let $V = \left\{\vec{0}\right\} \cup \left\{\vec{e}^{(i)} : i \in [d]\right\}$. Note that for any two vectors $\vec{v}, \vec{w} \in V$, that $\vec{v} - \vec{w}$ is a weak-alt-1 vector, and if $\vec{v} \neq \vec{w}$, then the difference is a weak-alt-1 sequence with a non-zero term[1].

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

Consider distinct cubes $X, Y \in \mathcal{C}$, so $\operatorname{rep}(X), \operatorname{rep}(Y) \in R$, so $A^{-1}\operatorname{rep}(X), A^{-1}\operatorname{rep}(Y) \in V$. Let $\vec{m} = A^{-1}\operatorname{rep}(X)$ and $\vec{n} = A^{-1}\operatorname{rep}(Y)$. As described, since $\vec{m}, \vec{n} \in V$, we have that $\vec{m} - \vec{n}$ is a weak-alt-1 sequence. By Proposition 4.2.10 this implies that $\|A(\vec{m} - \vec{n})\|_\infty \leq 1$. Thus,

$$\|\operatorname{rep}(X) - \operatorname{rep}(Y)\|_\infty = \|A(\vec{m} - \vec{n})\|_\infty \leq 1$$

so by Fact 3.4.7, X and Y are adjacent. Since this holds for any two cubes in $\mathcal{C}$, it must be that $\mathcal{C}$ is a clique. □

The following theorem is the main result of this section showing that we have constructed (k, ε)-secluded partitions of $\mathbb{R}^d$ with $k = d+1$ and $\varepsilon = \frac{\Delta_A}{2}$. A simple corollary will be that we have constructed (k, ε)-secluded partitions of $\mathbb{R}^d$ with $k = d+1$ and $\varepsilon = \frac{1}{2d}$.

---

[1]If $\vec{v} = \vec{w}$, then $\vec{v} - \vec{w} = \vec{0}$ which is trivially weak-alt-1. Otherwise assume $\vec{v} \neq \vec{w}$. If one of the vectors is $\vec{0}$, then the difference is a standard basis vector or negated standard basis vector; thus, the difference 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 distinct $i, j \in [d]$. Then the difference is all zeros except for a single term which is $+1$ and a single term which is -1; any such vector is a weak-alt-1 sequence.

Theorem 4.2.17 (Reclusive 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. Let $\varepsilon = \frac{\Delta_A}{2}$. Then for any $\vec{p} \in \mathbb{R}^d$,*

$$\left| {}^{\infty}\overline{\mathcal{N}}_{\varepsilon}(\vec{p}) \right| \leq d + 1.$$

Proof. It suffices to prove that ${}^{\infty}\overline{\mathcal{N}}_{\varepsilon}(\vec{p})$ is a clique since by Theorem 4.2.15, any clique contains at most $d+1$ cubes because $\mathcal{P}_A$ is $(d+1)$-colorable. Let $X, Y \in {}^{\infty}\overline{\mathcal{N}}_{\varepsilon}(\vec{p})$ be distinct, so by definition of ${}^{\infty}\overline{\mathcal{N}}_{\varepsilon}(\vec{p})$, we have ${}^{\infty}\overline{B}_{\vec{p}}(\cap)X \neq \emptyset$ and ${}^{\infty}\overline{B}_{\vec{p}}(\cap)Y \neq \emptyset$. Thus, there exists $\vec{x} \in X$ with $\|\vec{x} - \vec{p}\|_{\infty} < \varepsilon$ and $\vec{y} \in Y$ with $\|\vec{y} - \vec{p}\|_{\infty} < \varepsilon$. By the triangle inequality $\|\vec{x} - \vec{y}\|_{\infty} \leq 2\varepsilon = \Delta_A$.

By the contrapositive of Lemma 4.2.14, since there exists $\vec{x} \in X$ and $\vec{y} \in Y$ such that $\|\vec{x} - \vec{y}\|_{\infty} \leq \Delta_A$, and X and Y are distinct, it must be that X and Y are adjacent. Thus, ${}^{\infty}\overline{\mathcal{N}}_{\varepsilon}(\vec{p})$ is a clique. □

To conclude this section, we prove the existence of $(d+1, \frac{1}{2d})$-secluded reclusive partitions by applying Theorem 4.2.17 to a specific matrix. To motivate the choice of matrix, consider again the definition of reclusive distance (Definition 4.2.11) and the discussion leading up to it. To make the most of Theorem 4.2.17, 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 $k \leq j < j'$ 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; and the third encourages "even spacing" of the terms in a given row. We balance these goals with the following $d \times d$ matrix (with a clarifying example for $d = 5$).

$$\begin{bmatrix} 1 & \frac{d-1}{d} & \frac{d-2}{d} & \frac{d-3}{d} & \cdots & \frac{2}{d} & \frac{1}{d} \\ 0 & 1 & \frac{d-2}{d-1} & \frac{d-3}{d-1} & \cdots & \frac{2}{d-1} & \frac{1}{d-1} \\ 0 & 0 & 1 & \frac{d-3}{d-2} & \cdots & \frac{2}{d-2} & \frac{1}{d-2} \\ 0 & 0 & 0 & 1 & \cdots & \frac{2}{d-3} & \frac{1}{d-3} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix} \qquad \text{Example, } d = 5\text{:} \begin{bmatrix} 1 & \frac{4}{5} & \frac{3}{5} & \frac{2}{5} & \frac{1}{5} \\ 0 & 1 & \frac{3}{4} & \frac{2}{4} & \frac{1}{4} \\ 0 & 0 & 1 & \frac{2}{3} & \frac{1}{3} \\ 0 & 0 & 0 & 1 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{4.1}$$

Formally, this is the upper triangular matrix in which each element a_{ij} on or above the diagonal is given by $a_{ij} = \left(\frac{d-j+1}{d-i+1}\right)$. It is easy to verify that the reclusive distance of this matrix is $\frac{1}{d}$ (Definition 4.2.11) because $\delta_1 = 1$, $\delta_2 = \frac{1}{d}$, $\delta_3 = \frac{1}{d}$, and $\delta_4 = \frac{1}{d}$, so the reclusive distance is $\min\{\delta_1, \delta_2, \delta_3, \delta_4\} = \frac{1}{d}$.

Theorem 4.2.18 (Existence of $(d+1, \frac{1}{2d})$-Secluded Unit Cube Partitions)**.** *Let $d \in \mathbb{N}$. There exists an axis-aligned unit cube partition of $\mathbb{R}^d$ which is $(d+1, \frac{1}{2d})$-secluded.*

Proof. Let A be the matrix of Equation 4.1 and $\mathcal{P}_A$ its reclusive partition and Δ_A its reclusive distance. By the discussion above, $\Delta_A = \frac{1}{d}$, so by Theorem 4.2.17, we have that for any $\vec{p} \in \mathbb{R}^d$,

$$\left| {}^{\infty}\mathcal{N}_{\frac{1}{2d}}(\vec{p}) \right| \leq d + 1$$

which means by definition that $\mathcal{P}_A$ is $(d+1, \frac{1}{2d})$-secluded (Definition 3.1.1). $\square$

Remark 4.2.19. There are a variety of matrices other than the one in Equation 4.1 which have reclusive distance $\frac{1}{d}$ and thus give rise to $(d+1, \frac{1}{2d})$-secluded partitions.

In particular, there is a lot of "wiggle room" in all but the first row of that matrix. We will see another example in the next section. △

In Chapter 5 (Optimality of k and ε for Unit Cube Partitions) and Section 9.8 (Optimal ε For Unit Cube Enclosures) we will see that there is no axis-aligned unit cube partitions with better seclusion parameters; that is, $d+1$ is the smallest possible value of k regardless of ε, and when k is minimized at $d+1$, then ε cannot be any smaller than $\frac{1}{2d}$.

Next, we give some discussion on why we defined reclusive matrices as we did.

4.3 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 chapter in order to move from ε with the denominator growing exponentially to ε with the denominator growing only linearly. The generalization followed by noting the properties in Subsection 4.1.1.

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 section, we were incorrect, and the definition of a reclusive matrix really seems to capture something fundamental to the properties we are interested in.

Consider, for example, the following two reclusive matrices A' and A'' (A'' was the one we used in the previous section).

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

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

Then consider the matrices B' and B''.

$$B' = \begin{bmatrix} 1 & \frac{1}{d} & \frac{2}{d} & \frac{3}{d} & \cdots & \frac{d-2}{d} & \frac{d-1}{d} \\ 0 & 1 & \frac{2}{d} & \frac{3}{d} & \cdots & \frac{d-2}{d} & \frac{d-1}{d} \\ 0 & 0 & 1 & \frac{3}{d} & \cdots & \frac{d-2}{d} & \frac{d-1}{d} \\ 0 & 0 & 0 & 1 & \cdots & \frac{d-2}{d} & \frac{d-1}{d} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & \frac{d-1}{d} \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix} \quad \text{Example, } d = 5\text{:} \begin{bmatrix} 1 & \frac{1}{5} & \frac{2}{5} & \frac{3}{5} & \frac{4}{5} \\ 0 & 1 & \frac{2}{5} & \frac{3}{5} & \frac{4}{5} \\ 0 & 0 & 1 & \frac{3}{5} & \frac{4}{5} \\ 0 & 0 & 0 & 1 & \frac{4}{5} \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$B'' = \begin{bmatrix} 1 & \frac{1}{d} & \frac{2}{d} & \frac{3}{d} & \cdots & \frac{d-2}{d} & \frac{d-1}{d} \\ 0 & 1 & \frac{1}{d-1} & \frac{2}{d-1} & \cdots & \frac{d-3}{d-1} & \frac{d-2}{d-1} \\ 0 & 0 & 1 & \frac{1}{d-2} & \cdots & \frac{d-4}{d-2} & \frac{d-3}{d-2} \\ 0 & 0 & 0 & 1 & \cdots & \frac{d-5}{d-3} & \frac{d-4}{d-3} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix} \quad \text{Example, } d = 5\text{:} \begin{bmatrix} 1 & \frac{1}{5} & \frac{2}{5} & \frac{3}{5} & \frac{4}{5} \\ 0 & 1 & \frac{1}{4} & \frac{2}{4} & \frac{3}{4} \\ 0 & 0 & 1 & \frac{1}{3} & \frac{2}{3} \\ 0 & 0 & 0 & 1 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Clearly B' and B'' are not reclusive matrices[2] 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[3]. 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

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

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

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 both coordinates. 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.

Because the reclusive distance of both A' and A'' is $\frac{1}{d}$, by Theorem 4.2.17, the partitions $\mathcal{P}_{A'}$ and $\mathcal{P}_{A''}$ are $(d+1, \frac{1}{2d})$-secluded, so a natural question would be whether the partitions $\mathcal{P}_{B'}$ and $\mathcal{P}_{B''}$ associated to B' and B'' are also $(d+1, \frac{1}{2d})$-secluded. Our intuition was that the answer to this question would be "yes", but this is not the case. For $d \geq 5$, not only are $\mathcal{P}_{B'}$ and $\mathcal{P}_{B''}$ not $(d+1, \frac{1}{2d})$-secluded, they are not $(d+1, \varepsilon)$-secluded for any $\varepsilon \in (0, \infty)$. In particular, for $d \geq 5$, both $\mathcal{P}_{B'}$ and $\mathcal{P}_{B''}$ contain cliques of size $d+2$ which implies[4] that there is a point $\vec{p} \in \mathbb{R}^d$ with $|\mathcal{N}_{\bar{0}}(\vec{p})| \geq d+2$, and so for any $\varepsilon \in (0, \infty)$ we have $\left|{}^{\infty}\overline{\mathcal{N}}_{\varepsilon}(\vec{p})\right| \geq |\mathcal{N}_{\bar{0}}(\vec{p})| \geq d+2$ showing that they are not $(d+1, \varepsilon)$-secluded[5].

We found this out by exhaustive computer search[6]. Nonetheless, our claim that for $d = 5$, both $\mathcal{P}_{B'}$ and $\mathcal{P}_{B''}$ contain cliques of size $d+2 = 7$ can easily be verified. Consider the following set of 7 vectors in $\mathbb{Z}^5$:

[4]The details of this implication are not important at the moment, but Lemma 3.4.11 implies that for a clique, there is a point $\vec{p}$ belonging to the closure of every cube in the clique, and thus $\mathcal{N}_{\bar{0}}(\vec{p})$ includes at least every member of the clique.

[5]For $d \leq 4$, the partitions $\mathcal{P}_{B'}$ and $\mathcal{P}_{B''}$ do not have cliques of size $d+2$, so because of the discrete structure of the lattice, they are $(d+1, \varepsilon)$-secluded for some value of $\varepsilon \in (0, \infty)$, but we have not checked if ε can be taken to be $\frac{1}{2d}$ or not.

[6]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 cubes 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 cubes with representative corners at these 7 locations form a clique, check that the ℓ_∞ distance between any pair is exactly 1 (and apply Fact 3.4.6).

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. We have not yet devoted any time to this conjecture, but we suspect there is probably a simple inductive-style proof of it.

Conjecture 4.3.1 (Maximum Clique Sizes in $\mathcal{P}_{B'}$ and $P_{B''}$)**.** *Based on our computations, we conjecture that the maximum clique size in the partitions $\mathcal{P}_{B'}$ and $\mathcal{P}_{B''}$ is greater than $d + 1$ for all $d \geq 5$, and we know the maximum clique sizes for dimensions given in Table 4.1.*

Dimension d	1	2	3	4	5	6	7	8	9	10	11	12
Max Clique in $\mathcal{P}_{B'}$	2	3	4	5	7	9	12	16	22	30	39	51
Max Clique in $\mathcal{P}_{B''}$	2	3	4	5	7	9	11	16	21	28	36	47

Table 4.1: Maximum clique sizes in non-reclusive partitions associated to matrices 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[7]. The above example shows that this equivalence does not hold if we relax the requirement that the matrix entries decrease in each row.

Another seemingly reasonable question is the following "The fact that there exists unit cube partitions for which the $\varepsilon = \frac{1}{2d}$ ℓ_∞-ball only intersects at most $d+1$ cubes is a simple statement, can you find a simpler proof of it than the one in the prior section?" The discussion above suggests that the answer in some sense is probably no—there is probably not a significantly simpler proof of this fact. The reason we believe this is that lattice based partitions are some of the most natural, and any proof of this result which explicitly constructs a partition from a lattice (equivalently from a matrix[8]) must somehow distinguish between matrices like A' or A'' and matrices like B' or B''. In other words these matrices are all very similar and yet some of them generate partitions with the desired property and other don't.

Next, we turn our attention to doing something very typical in mathematics: we will use the partitions we constructed in Section 4.2 (Reclusive Partitions) as building

[7] We have not explicitly stated this equivalence, but it comes from combining Proposition 4.2.10 with Lemma 4.2.13. Distinct cubes X and Y in a reclusive partition are adjacent if and only if $\|\text{rep}(X) - \text{rep}(Y)\|_\infty \leq 1$ (which is in turn equivalent to $\|\text{rep}(X) - \text{rep}(Y)\|_\infty = 1$ as the distance can't be strictly less than 1 without the cubes overlapping which is not possible in a partition).

[8] It is known that any lattice in $\mathbb{R}^d$ can be associated to a matrix.

blocks for other partitions.

4.4 New Partitions From Old

We have so far shown that (k, ε)-secluded unit cube partitions exist for $k = d+1$ and $\varepsilon = \frac{1}{2d}$, and as mentioned already, we will show in Chapter 5 (Optimality of k and ε for Unit Cube Partitions) that $k = d+1$ is the smallest value possible and that $\varepsilon = \frac{1}{2d}$ is the largest value when k is minimized at $k = d+1$. However, we wonder what the trade-off is between ε and k.

For example, taking $A^{(d)}$ to be the $d \times d$ matrix in Equation 4.1, the proof of Theorem 4.2.18 showed that $\mathcal{P}_{A^{(d)}}$ was $(d+1, \frac{1}{2d})$-secluded, but we could also ask how secluded this same partition $\mathcal{P}_{A^{(d)}}$ is for a larger ε parameter: for each $\varepsilon \in (0, \infty)$, what is the smallest $k \in \mathbb{N}$ such that the partition $\mathcal{P}_A$ is (k, ε)-secluded? Some such k exists by a trivial result that we state shortly (Fact 4.4.3). To be more concrete, suppose that for some constant $c > 1$ we want $\varepsilon = c\frac{1}{2d}$ instead of $\varepsilon = \frac{1}{2d}$, for what value of k can we say that this same $\mathcal{P}_{A^{(d)}}$ is $(k, c\frac{1}{2d})$-secluded? One might hope that the answer is something like $k = (d+1)^c$ so that we can consider linear increases in ε while only increasing k by a polynomial factor, or possibly hope that it is even better than this.

However, we have evidence suggesting that this is not the case, and it appears that for any $c > 1$, the parameter k immediately jumps from $d+1$ to an expression which is exponential in d which would demonstrate an extremely sharp threshold[9].

[9]We have sketched out a proof of the conjecture that follows, but we have not yet worked out the details to a sufficient degree to comfortably claim the conjecture as true.

Conjecture 4.4.1 (Bifurcation of Seclusion Parameters)**.** *For each $d \in \mathbb{N}$, let $A^{(d)}$ denote the $d \times d$ reclusive matrix as in Equation 4.1, and let $\mathcal{P}_{A^{(d)}}$ denote its reclusive partition of $\mathbb{R}^d$. Fix any constant $c > 1$. Then for each $d \in \mathbb{N}$, there exists $\vec{p}^{(d)} \in \mathbb{R}^d$ such that the sequence[a] $\left\langle \left| {}^{\infty}\overline{\mathcal{N}}_{c\frac{1}{2d}}(\vec{p}^{(d)}) \right| \right\rangle_{d=1}^{\infty}$ is asymptotically exponential in d (with hidden constants depending on c).*

[a]Where ${}^{\infty}\overline{\mathcal{N}}_{c\frac{1}{2d}}(\vec{p}^{(d)})$ is with respect to the partition $\mathcal{P}_{A^{(d)}}$.

In other words, we believe that in terms of an asymptotic perspective of d, taking ε to be a constant multiple larger than $\frac{1}{2d}$ results in some ℓ_∞ ε-ball intersecting an exponential (in d) number of cubes in $\mathcal{P}_{A^{(d)}}$ rather than intersecting only $d+1$ cubes in $\mathcal{P}_{A^{(d)}}$.

Based on this conjecture, we are motivated to construct *other* unit cube partitions of $\mathbb{R}^d$ which hopefully attain better parameters, and in fact we can do basically what we suggested above. For each $c > 1$, we can construct (for each $d \in \mathbb{N}$) an axis-aligned unit cube partition of $\mathbb{R}^d$ which is $(k, c\frac{1}{2d})$-secluded for $k \approx (d+1)^c$. The distinction is that the conjecture indicates we cannot use just a single partition in each dimension but instead have to construct a partition in each dimension which is specific to the choice of constant c.

The partitions that we construct are of a very natural form: we view $\mathbb{R}^d$ as $\prod_{i=1}^n \mathbb{R}^{d_i}$ where $\sum_{i=1}^n d_i = d$ and separately partition each $\mathbb{R}^{d_i}$ using a $(d_i + 1, \frac{1}{2d_i})$-secluded reclusive partition of Section 4.2. One could also partition each $\mathbb{R}^{d_i}$ with the partitions implicit in [HK18] which have similar seclusion parameters as the reclusive partitions, though they are not unit cube partitions (see Remark 4.0.1 from earlier). We then combine the partitions using a natural product construction to obtain the desired partition of $\mathbb{R}^d$.

To begin, we will define the construction very generically, and we will need two basic results. The following observation notes that if a partition is (k,ε)-secluded, then we can *increase* k to k' and *decrease* ε to ε' and the partition is trivially (k',ε')-secluded.

Observation 4.4.2 (Monotonicity in k and ε)**.** *Let $d \in \mathbb{N}$, $k, k' \in \mathbb{N}$ with $k' \geq k$, $\varepsilon, \varepsilon' \in (0,\infty)$ with $\varepsilon' \leq \varepsilon$, and $\mathcal{P}$ a (k,ε)-secluded partition of $\mathbb{R}^d$. Then $\mathcal{P}$ is also a (k',ε')-secluded partition of $\mathbb{R}^d$.*

Proof. Since $\mathcal{P}$ is (k,ε)-secluded, by definition every ε-ball intersects at most k members of $\mathcal{P}$, so trivially every (no larger) ε'-ball intersects at most $k' \geq k$ members of $\mathcal{P}$. □

We will frequently refer to the above observation just using the phrase "by monotonicity, $\mathcal{P}$ is (k',ε')-secluded"

Fact 4.4.3 (Trivial k for Unit Cube Partitions)**.** *Let $d \in \mathbb{N}$, $\varepsilon \in (0,\infty)$, and $\mathcal{P}$ be a unit cube partition of $\mathbb{R}^d$. Then $\mathcal{P}$ is (k,ε)-secluded for $k = \lfloor (2+2\varepsilon)^d \rfloor$.*

Proof. Consider any point $\vec{p} \in \mathbb{R}^d$. By definition, ${}^{\infty}\mathcal{N}_{\varepsilon}(\vec{p}) = \left\{ X \in \mathcal{P} \colon X \cap {}^{\infty}\overline{B}_{\varepsilon}(\vec{p}) \neq \emptyset \right\}$. Note that this is a subset of $\left\{ X \in \mathcal{P} \colon X \subseteq {}^{\infty}\overline{B}_{1+\varepsilon}(\vec{p}) \right\}$ because $\mathcal{P}$ is a unit cube partition so for each $X \in \mathcal{P}$, $\operatorname{diam}_{\infty}(X) = 1$. Because (1) each $X \in \mathcal{P}$ has measure 1, and (2) all of the X are disjoint (because $\mathcal{P}$ is a partition), and (3) the measure of ${}^{\infty}\overline{B}_{1+\varepsilon}(\vec{p})$ is $(2+2\varepsilon)^d$, it follows that the cardinality of this latter set is at most $\left\lfloor (2+2\varepsilon)^d \right\rfloor$. Thus, for every $\vec{p} \in \mathbb{R}^d$ we have that $\left| {}^{\infty}\mathcal{N}_{\varepsilon}(\vec{p}) \right| \leq \left\lfloor (2+2\varepsilon)^d \right\rfloor$ which shows that $\mathcal{P}$ is (k,ε)-secluded for $k = \left\lfloor (2+2\varepsilon)^d \right\rfloor$ as claimed. □

4.4.1 Construction

Definition 4.4.4 (Partition Product). Let $d_1, \ldots, d_n \in \mathbb{N}$ and $\mathcal{P}_1, \ldots, \mathcal{P}_n$ be partitions of $\mathbb{R}^{d_1}, \ldots, \mathbb{R}^{d_n}$ respectively. Letting $d = \sum_{i=1}^{n} d_i$ we define the product partition of $\mathbb{R}^d$ as

$$\prod_{i=1}^{n} \mathcal{P}_i \overset{\text{def}}{=} \left\{ \prod_{i=1}^{n} X^{(i)} \colon X^{(i)} \in \mathcal{P}_i \right\}$$

where $\prod_{i=1}^{n} X^{(i)}$ is viewed as a subset of $\mathbb{R}^d$.

Remark 4.4.5. We specifically stated that $\prod_{i=1}^{n} X^{(i)}$ is viewed as a subset of $\mathbb{R}^d$, because technically it is a subset of $\prod_{i=1}^{n} \mathbb{R}^{d_i}$, but this is naturally isomorphic to $\mathbb{R}^d = \mathbb{R}^{\sum_{i=1}^{n} d_i}$.

For example, technically, if $d_1 = d_2 = d_3 = 2$, then the elements of $\prod_{i=1}^{n} \mathbb{R}^{d_i}$ are of the form $\langle \langle x_1, x_2 \rangle, \langle x_3, x_4 \rangle, \langle x_5, x_6 \rangle \rangle$, but this is trivially isomorphic to $\mathbb{R}^6$ by instead considering the element as $\langle x_1, x_2, x_3, x_4, x_5, x_6 \rangle$. $\triangle$

Example 4.4.6. Let $d_1 = 1$, $d_2 = 2$, (so $n = 2$, and $d = 3$). Let $\mathcal{P}_1 = \{m + [0.5, 1.5) \colon m \in \mathbb{Z}\}$ (a partition of $\mathbb{R}^1$) and $\mathcal{P}_2 = \{\vec{m} + [0,1)^2 \colon \vec{m} \in \mathbb{Z}^2\}$ (a partition of $\mathbb{R}^2$). Then $\prod_{i=1}^{2} \mathcal{P}_i$ is the partition of $\mathbb{R}^3$ where each member is of the form

$$[m_1 + [0.5, 1.5)] \times [m_2 + [0,1)] \times [m_3 + [0,1)]$$

for $\vec{m} = \langle m_1, m_2, m_3 \rangle \in \mathbb{Z}^3$.

Observe that if the original partitions were unit cube partitions, then the product partition is also a unit cube partition. This is basically due to the fact unit cubes are balls with respect to ℓ_∞, and the ℓ_∞ norm behaves nicely in products:

Fact 4.4.7 (Unit Cube Preservation)**.** *If* $d_1, \ldots, d_n \in \mathbb{N}$ *and* $\mathcal{P}_1, \ldots, \mathcal{P}_n$ *are* unit cube *partitions of* $\mathbb{R}^{d_1}, \ldots, \mathbb{R}^{d_n}$ *respectively, then* $\prod_{i=1}^{n} \mathcal{P}_i$ *is also a unit cube partition.*

Proof. Each member of $\prod_{i=1}^{n} \mathcal{P}_i$ is of the form $\prod_{i=1}^{n} X^{(i)}$ where $X^{(i)} \in \mathcal{P}_i$. Since $\mathcal{P}_i$ is a unit cube partition, each $X^{(i)}$ is a product of translates of $[0, 1)$, and thus (up to the natural isomorphism) $\prod_{i=1}^{n} X^{(i)}$ is also a product of d-many translates of $[0, 1)$, where $d = \sum_{i=1}^{n} d_i$. Thus, the member is a unit cube. □

We can now present the main result of this section which is that if we take a product of partitions, and we have a guarantee for each $\mathcal{P}_i$ that it is (k_i, ε_i)-secluded, then we can guarantee that the product partition is (k, ε)-secluded where k is the product of the k_i's and ε is the minimum of the ε_i's.

Proposition 4.4.8 (Product Partition Seclusion Guarantees)**.** *Let* $n \in \mathbb{N}$*. For each index* $i \in [n]$*, let* $d_i, k_i \in \mathbb{N}$*,* $\varepsilon_i \in (0, \infty)$ *and* $\mathcal{P}_i$ *be a* (k_i, ε_i)*-secluded partition of* $\mathbb{R}^{d_i}$*. Then the product partition* $\mathcal{P} = \prod_{i=1}^{n} \mathcal{P}_i$ *is a* (k, ε)*-secluded partition of* $\mathbb{R}^d$ *where* $d = \sum_{i=1}^{n} d_i$*, and* $k = \prod_{i=1}^{n} k_i$*, and* $\varepsilon = \min_{i \in [n]} \varepsilon_i$*.*

Proof Sketch. The basic idea is that for any point $\vec{p} \in \mathbb{R}^d$, we consider how many members of $\mathcal{P}$ intersect ${}^{\infty}\overline{B}_{\varepsilon}(\vec{p})$. Conceptually[10], we think of $\vec{p}$ as a sequence $\langle \vec{p}^{(i)} \rangle_{i=1}^{n}$ of n points where the ith point $\vec{p}^{(i)}$ belongs to $\mathbb{R}^{d_i}$. Because we are working with the ℓ_∞ norm (that is the norm used by definition of secluded), the ε-ball around $\vec{p}$ is the product of the ε-balls around each $\vec{p}^{(i)}$ which is smaller than the product of ε_i-balls around each $\vec{p}^{(i)}$ because we chose ε as the minimum size. Thus, if the ε-ball around $\vec{p}$ intersects a member X of the partition $\mathcal{P}$, then viewing X as the product $\prod_{i=1}^{n} X^{(i)}$ where $X^{(i)}$ is a member of $\mathcal{P}_i$, it must be for each $i \in [n]$ that the ε-ball around

[10]In other words we identify the set $\mathbb{R}^d$ with $\mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \times \cdots \times \mathbb{R}^{d_{n-1}} \times \mathbb{R}^{d_n}$.

$\vec{p}^{(i)}$ intersects $X^{(i)}$ (and thus so does the ε_i-ball since $\varepsilon_i \geq \varepsilon$). This means (for each $i \in [n]$) that $X^{(i)}$ is one of at most k_i members of $\mathcal{P}_i$ because at most k_i members of $\mathcal{P}_i$ intersect the ε_i-ball around $\vec{p}^{(i)}$ (by definition of $\mathcal{P}_i$ being (k_i, ε_i)-secluded). Thus X is one of at most $\prod_{i=1}^{n} k_i = k$ members of $\mathcal{P}$. That is, there are at most k members of $\mathcal{P}$ that intersect the ε-ball around $\vec{p}$ which is the definition of $\mathcal{P}$ being (k, ε)-secluded. □

Utilizing the construction above, we will now take a $(d_i + 1, \frac{1}{2d_i})$-secluded reclusive partition for each $\mathbb{R}^{d_i}$ and take the product to obtain a new partition[11]. Since the dimension of each d_i is smaller than the dimension d, this allows us to get a larger value of ε_i for each partition, and thus a larger value of ε for the partition of $\mathbb{R}^d$ than if we had used the original partition. The price we pay for this is that the value of k also increases. The following result is nothing more than Proposition 4.4.8 where each partition in the product is specifically one of these reclusive partitions. We will shortly refine this result to specify parameters that we are ultimately interested in.

Lemma 4.4.9 (Secluded Partition Product Guarantees)**.** *Let $f : \mathbb{N} \to \mathbb{N}$ be any function[a]. For each $d \in \mathbb{N}$, there exists a $(k(d), \varepsilon(d))$-secluded axis-aligned unit cube partition of $\mathbb{R}^d$ where $k(d) = (f(d) + 1)^{\lceil \frac{d}{f(d)} \rceil}$ and $\varepsilon(d) = \frac{1}{2f(d)}$*

[a]It will only be useful though if $f(d) < d$ for each $d \in \mathbb{N}$.

Proof. Fix $d \in \mathbb{N}$. Let $d' = f(d)$ and $n = \left\lceil \frac{d}{f(d)} \right\rceil = \lceil \frac{d}{d'} \rceil$. Let $\mathcal{P}'$ be a $(d' + 1, \frac{1}{2d'})$-secluded unit cube partition of $\mathbb{R}^{d'}$ (which exist by Theorem 4.2.18).

By Proposition 4.4.8 and Fact 4.4.7, $\mathcal{P} = \prod_{i=1}^{n} \mathcal{P}'$ is a (k, ε)-secluded unit cube partition of $\mathbb{R}^{n \cdot d'}$ where $k = (d' + 1)^n$ and $\varepsilon = \frac{1}{2d'}$. Since $n \cdot d' = \lceil \frac{d}{d'} \rceil \cdot d' \geq d$, this trivially (by ignoring extra coordinates) gives a partition of $\mathbb{R}^d$ with these same

[11]As mentioned already, we could also use the partitions of [HK18] to get the same asymptotics, but would have to remove the "unit cube" portion of the following statement.

properties (alternatively, see footnote[12]). Recalling the definitions of $d' = f(d)$ and $n = \left\lceil \frac{d}{f(d)} \right\rceil$ gives the stated result. □

4.4.2 Parameter Analysis

Now we turn our attention to analyzing exactly how good the secluded partitions of Lemma 4.4.9 are. The main result of this subsection is the following.

Theorem 4.4.10. *Let $\varepsilon : \mathbb{N} \to (0, \infty)$. Then there exists $k : \mathbb{N} \to \mathbb{N}$ such that for every $d \in \mathbb{N}$ there exists a $(k(d), \varepsilon(d))$-secluded unit hypercube partition of $\mathbb{R}^d$, and k has the following properties:*

1. *If $\varepsilon(d) \in O(1)$, then $k(d) \in \mathsf{exp}(d)$*
2. *If $\varepsilon(d) \in o(1)$, then $k(d) \in \mathsf{weaksubexp}(d)$*
3. *If $\varepsilon(d) \in O(\frac{1}{d})$, then $k(d) \in \mathsf{poly}(d)$*

We find the most interesting result above to be that for any subconstant function $\varepsilon(d)$, we can achieve $k(d)$ which is subexponential, so if one's concern is just that $k(d)$ is subexponential, then one can take $\varepsilon(d)$ to be arbitrarily close to constant. We start by establishing a few results about the asymptotics of certain functions which will be used to prove this result.

Lemma 4.4.11. *Let $f : \mathbb{N} \to \mathbb{N}$ such that $f(d) \in \omega(1) \cap O(d)$. Then $(f(d) + 1)^{\left\lceil \frac{d}{f(d)} \right\rceil} \in \mathsf{weaksubexp}(d)$.*

Proof. Since $f(d) \in O(d)$, let N_1 and C such that for $d \geq N_1$, $f(d) \leq Cd$. Observe

[12] An alternate perspective is to let $d_1, \ldots, d_n$ be such that $\sum_{i=1}^{n} d_i = d$ and the first portion of the list $d_i = d'$, and the second portion of the list $d_i = d'' \stackrel{\text{def}}{=} d' - 1$. Then let $\mathcal{P}'$ a $(d'+1, \frac{1}{2d'})$-secluded partition of $\mathbb{R}^{d'}$ as before, and let $\mathcal{P}''$ a $(d''+1, \frac{1}{2d''})$-secluded partition of $\mathbb{R}^{d''}$. Since $d'' < d'$, $\mathcal{P}''$ is (by monotonicity) a $(d'+1, \frac{1}{2d'})$-secluded partition. Then take $\mathcal{P}_i = \mathcal{P}'$ when $d_i = d'$ and $\mathcal{P}_i = \mathcal{P}''$ when $d_i = d''$. Again, we get that $\mathcal{P}$ is (k, ε)-secluded for $k = (d'+1)^n$ and $\varepsilon = \frac{1}{2d'}$.

the following inequalities for $d \geq N_1$:

$$\begin{aligned}
(f(d)+1)^{\left\lceil \frac{d}{f(d)} \right\rceil} &\leq (f(d)+1)^{\frac{d}{f(d)}+1} && (f(d) \in \mathbb{N} \text{ so } f(d) \geq 0)\\
&= (f(d)+1)^{\frac{d+f(d)}{f(d)}} && \text{(Algebra)}\\
&\leq (f(d)+1)^{\frac{(C+1)d}{f(d)}} && (f(d) \leq Cd \text{ and } (f(d)+1) \geq 1)\\
&= \left(\left((f(d)+1)^{\frac{1}{f(d)}}\right)^{C+1}\right)^{d}
\end{aligned}$$

Since $f \in \omega(1)$, then $\lim_{d\to\infty} f(d) = \infty$, so by Fact D.0.1, $\lim_{d\to\infty}(f(d)+1)^{\frac{1}{f(d)}} = 1$ which implies by basic real analysis results that $\lim_{d\to\infty}\left((f(d)+1)^{\frac{1}{f(d)}}\right)^{C+1} = 1$. Finally, by Lemma D.0.5, $\left(\left((f(d)+1)^{\frac{1}{f(d)}}\right)^{C+1}\right)^{d}$ belongs to $2^{o(d)}$. □

Alternate Proof. To show that a function is in $2^{o(d)}$, we (by definition) show that its logarithm is in $o(d)$. Since $f \in O(d)$, let N_1 and C such that for $d \geq N_1$, $f(d) \leq Cd$. Observe the following inequalities for $d \geq N_1$:

$$\begin{aligned}
\lim_{d\to\infty} \frac{\log_2\left((f(d)+1)^{\left\lceil \frac{d}{f(d)} \right\rceil}\right)}{d} &= \lim_{d\to\infty} \left\lceil \frac{d}{f(d)} \right\rceil \frac{\log_2(f(d)+1)}{d}\\
&\leq \lim_{d\to\infty} \left(\frac{d}{f(d)}+1\right) \frac{\log_2(f(d)+1)}{d}\\
&= \lim_{d\to\infty} \frac{d+f(d)}{f(d)} \cdot \frac{\log_2(f(d)+1)}{d}\\
&\leq \lim_{d\to\infty} \frac{(C+1)d}{f(d)} \cdot \frac{\log_2(f(d)+1)}{d}\\
&= \lim_{d\to\infty} \frac{(C+1)}{f(d)} \cdot \log_2(f(d)+1)\\
&= (C+1) \lim_{d\to\infty} \frac{\log_2(f(d)+1)}{f(d)} && \left(\lim_{x\to\infty} \frac{\log_2(x+1)}{x} = 0\right)\\
&= 0
\end{aligned}$$

All limits above exist by the real analysis "squeeze theorem" because all terms of each sequence are non-negative, and the last limit exists. □

Remark 4.4.12. In the above result, the requirement that $f \in O(d)$ was relevant because taking for example $f(d) = 2^d$ which is $\omega(1)$ but not $O(d)$ would give for sufficiently large d,

$$(f(d)+1)^{\left\lceil \frac{d}{f(d)} \right\rceil} = (f(d)+1)^1 = 2^d + 1$$

which is not sub-exponential.

The requirement that $f \in \omega(1)$ was relevant because taking for example $f(d) = 1$ which is $O(d)$ but not $\omega(1)$ would give

$$(f(d)+1)^{\left\lceil \frac{d}{f(d)} \right\rceil} = (1+1)^{\left\lceil \frac{d}{1} \right\rceil} = 2^d$$

which is also not sub-exponential. △

Lemma 4.4.13. *Let $f : \mathbb{N} \to \mathbb{N}$ such that $f(d) \in \Omega(d) \cap O(d) = \Theta(d)$. Then $(f(d)+1)^{\left\lceil \frac{d}{f(d)} \right\rceil} \in \mathsf{poly}(d)$.*

Proof. Let $N \in \mathbb{N}$ and $c, C \in (0, \infty)$ such that for $d \geq N$, $cd \leq f(d) \leq Cd$. Then for $d \geq N$, we have

$$\begin{aligned}
(f(d)+1)^{\left\lceil \frac{d}{f(d)} \right\rceil} &\leq (Cd+1)^{\left\lceil \frac{d}{f(d)} \right\rceil} && \text{(Increase positive base; exponent non-negative)} \\
&\leq (Cd+1)^{\left\lceil \frac{d}{cd} \right\rceil} && \text{(Increase exponent; base } \geq 1\text{)} \\
&= (Cd+1)^{\left\lceil \frac{1}{c} \right\rceil} \\
&\in O\left(d^{\left\lceil \frac{1}{c} \right\rceil}\right) \\
&\subseteq \mathsf{poly}(d)
\end{aligned}$$

which completes the proof. □

Now we restate and prove the main result of this subsection.

Theorem 4.4.10. *Let $\varepsilon : \mathbb{N} \to (0, \infty)$. Then there exists $k : \mathbb{N} \to \mathbb{N}$ such that for every $d \in \mathbb{N}$ there exists a $(k(d), \varepsilon(d))$-secluded unit hypercube partition of $\mathbb{R}^d$, and k has the following properties:*

1. *If $\varepsilon(d) \in O(1)$, then $k(d) \in \mathsf{exp}(d)$*
2. *If $\varepsilon(d) \in o(1)$, then $k(d) \in \mathsf{weaksubexp}(d)$*
3. *If $\varepsilon(d) \in O(\frac{1}{d})$, then $k(d) \in \mathsf{poly}(d)$*

Proof. We begin with (1). By Fact 4.4.3, for each $d \in \mathbb{N}$, there exists a $(k(d), \varepsilon(d))$-secluded unit cube partition of $\mathbb{R}^d$ where $k(d) = \left\lfloor (2 + 2\varepsilon(d))^d \right\rfloor$. If $\varepsilon(d) \in O(1)$, then $k(d) \in \mathsf{exp}(d)$. This proves (1).

Now we do some setup for all of the remaining cases. Since

$$\frac{O(1)}{d} \subseteq \frac{\mathsf{poly}(d)}{d} \subseteq \frac{\mathsf{subpoly}(d)}{d} \subseteq \frac{o(d)}{d} = o(1)$$

in all remaining cases, there exists some $N_1 \in \mathbb{N}$ such that for $d \geq N_1$, $\varepsilon(d) \leq \frac{1}{2}$. Let $f : \mathbb{N} \to \mathbb{N}$ be defined by

$$f(d) = \begin{cases} 1 & d < N_1 \\ d & d \geq N_1 \text{ and } \varepsilon(d) \leq \frac{1}{2d} \\ \left\lfloor \frac{1}{2\varepsilon(d)} \right\rfloor & d \geq N_1 \text{ and } \varepsilon(d) > \frac{1}{2d} \end{cases}$$

noting that the codomain $\mathbb{N}$ of f is valid because f does not produce 0 or negative values[13]. Also observe that $f(d) \in O(d)$—in fact, for all $d \in \mathbb{N}$, $f(d) \leq d$ because if

[13]This is because if $d \geq N_1$, then $\varepsilon(d) \leq \frac{1}{2}$, so $\left\lfloor \frac{1}{2\varepsilon(d)} \right\rfloor \geq \left\lfloor \frac{1}{2(\frac{1}{2})} \right\rfloor = 1$

$\varepsilon(d) > \frac{1}{2d}$, then $\left\lfloor \frac{1}{2\varepsilon(d)} \right\rfloor \leq \left\lfloor \frac{1}{2\left(\frac{1}{2d}\right)} \right\rfloor = d$.

Now define the function $k : \mathbb{N} \to \mathbb{N}$ by

$$k(d) = \begin{cases} \left\lfloor (2+2\varepsilon(d))^d \right\rfloor & d < N_1 \\ (f(d)+1)^{\left\lceil \frac{d}{f(d)} \right\rceil} & d \geq N_1 \end{cases}$$

We claim that for each $d \in \mathbb{N}$, there exists a $(k(d), \varepsilon(d))$-secluded unit cube partition of $\mathbb{R}^d$. To see this, if $d < N_1$, then a $(k(d), \varepsilon(d))$-secluded unit cube partition of $\mathbb{R}^d$ exists by Fact 4.4.3. Also, for each $d \in \mathbb{N}$, there exists a $\left((f(d)+1)^{\left\lceil \frac{d}{f(d)} \right\rceil}, \frac{1}{2f(d)}\right)$-secluded unit cube partition $\mathcal{P}_d$ of $\mathbb{R}^d$, and for $d \geq N_1$, by definition of k, this is a $\left(k(d), \frac{1}{2f(d)}\right)$-secluded unit cube partition. Also, for $d \geq N_1$, we have $f(d) = \left\lfloor \frac{1}{2\varepsilon(d)} \right\rfloor$, so

$$\frac{1}{2f(d)} = \frac{1}{2\left(\left\lfloor \frac{1}{2\varepsilon(d)} \right\rfloor\right)} \geq \frac{1}{2\left(\frac{1}{2\varepsilon(d)}\right)} = \varepsilon(d),$$

so by monotonicity, $\mathcal{P}_d$ is a $(k(d), \varepsilon(d))$-secluded unit cube partition of $\mathbb{R}^d$. To complete the proof, we now must show that this function k has the desired asymptotic properties.

(2): If $\varepsilon(d) \in o(1)$, then we claim that $f(d) \in \omega(1)$. This is because for any $C \in (0, \infty)$, there exists $N_2 \in \mathbb{N}$ such that for $d \geq N_2$, $\varepsilon(d) \leq \frac{1}{2(C+1)}$. For $d \geq \max\{N_1, N_2, C\}$ we either have $\varepsilon(d) \leq \frac{1}{2d}$ in which case $f(d) = d \geq C$, or we have $\varepsilon(d) > \frac{1}{2d}$ in which case

$$f(d) = \left\lfloor \frac{1}{2\varepsilon(d)} \right\rfloor \geq \left\lfloor \frac{1}{2\left(\frac{1}{2(C+1)}\right)} \right\rfloor = \lfloor C+1 \rfloor \geq C$$

which shows that $f(d) \in \omega(1)$. So combining the fact that $f(d) \in \omega(1)$ with the prior established fact that $f(d) \in O(d)$, Lemma 4.4.11 shows that $(f(d)+1)^{\left\lceil \frac{d}{f(d)} \right\rceil} \in$

$\mathsf{weaksubexp}(d)$, and so also $k(d) \in \mathsf{weaksubexp}(d)$.

(3): If $\varepsilon(d) \in O\left(\frac{1}{d}\right)$, then we claim that $f(d) \in \Omega(d)$. This is because there exists $C \in [1, \infty)$ and $N_2 \in \mathbb{N}$ such that for $d \geq N_2$, $\varepsilon(d) \leq \frac{c}{2} \cdot \frac{1}{d}$. For $d \geq \max\{N_1, N_2, 2C\}$ we either have $\varepsilon(d) \leq \frac{1}{2d}$ in which case $f(d) = d \geq \frac{d}{2C}$, or we have $\varepsilon(d) > \frac{1}{2d}$ in which case

$$f(d) = \left\lfloor \frac{1}{2\varepsilon(d)} \right\rfloor \geq \left\lfloor \frac{1}{2\left(\frac{C}{2d}\right)} \right\rfloor = \left\lfloor \frac{d}{C} \right\rfloor \geq \frac{d}{C} - 1 = \frac{d-c}{c} \geq \frac{d - d/2}{C} = \frac{d}{2C}$$

which shows that $f(d) \in \Omega(d)$. So combining the fact that $f(d) \in \Omega(d)$ with the prior established fact that $f(d) \in O(d)$, Lemma 4.4.13 shows that $(f(d)+1)^{\lceil \frac{d}{f(d)} \rceil} \in \mathsf{poly}(d)$, and so also $k(d) \in \mathsf{poly}(d)$.

□

Remark 4.4.14. This result is a little unfortunate in the following sense. From the computational perspective that we have so far encountered, taking $\varepsilon(d) = \frac{c}{2d}$ for some constant c is typically not meaningfully better than taking $\varepsilon(d) = \frac{1}{2d}$ (as in our reclusive partitions) because both are $\varepsilon(d) \in O(\frac{1}{d})$. In our computational context, we have wanted to keep $k(d) \in \mathsf{poly}(d)$ as that is the typically accepted notion of efficient. However, this result makes no guarantees that we can have $k(d) \in \mathsf{poly}(d)$ unless we keep $\varepsilon(d) \in O(\frac{1}{d})$.

However, while we had hoped these new constructions would have better guarantees than this (i.e. we hoped for constructions where we could get $k(d) \in \mathsf{poly}(d)$ and $\varepsilon(d) \in \omega(\frac{1}{d})$), it turns out that not much better is possible. We will see in Section 7.1 (Upper Bounds on ε via Brunn-Minkowski and Blichfeldt) that if $k(d) \in \mathsf{poly}(d)$, then it must be that $\varepsilon(d) \in O(\frac{\log(d)}{d})$, so we cannot hope for more than a logarithmic factor improvement in $\varepsilon(d)$. Furthermore, the bound that $\varepsilon(d) \in O(\frac{\log(d)}{d})$ is not just a bound for unit cube partitions, but it is a bound for any partitions with members of (outer) measure at most 1 (which includes all partitions with members of ℓ_∞ diameter at most 1). $\triangle$

Chapter 5

Optimality of k and ε for Unit Cube Partitions

It turns out that our constructions in Chapter 4 with $k = d+1$ and $\varepsilon = \frac{1}{2d}$ are *exactly* optimal with regard to both parameters[1] k and ε for all axis-aligned unit cube partitions. In this chapter, we will do three things: (1) we will prove (in multiple ways) that $k = d+1$ is the minimum possible value of k for an axis-aligned unit cube partition, (2) we will sketch a very clean argument for why $\varepsilon = \frac{1}{2d}$ is *almost* optimal when $k = d+1$ which will motivate the generalizations which serve as our main proof technique in Chapter 7 (Near Optimality of ε in General), and (3) we will provide an outline of the proof that $\varepsilon = \frac{1}{2d}$ is in fact exactly optimal for axis-aligned unit cube partitions when $k = d+1$ (the full proof will appear later and more appropriately[2] in Section 9.8 (Optimal ε For Unit Cube Enclosures)).

[1] Note that there isn't really an optimal value of the ε parameter—there is only an optimal value of ε *with respect to* a specified value of the parameter k. When we say that both parameters are optimal, we really mean that $k = d+1$ is the minimum possible value of k, and that with respect to the parameter $k = d+1$, the value $\varepsilon = \frac{1}{2d}$ is the maximum.

[2] The reason we delay the proof that $\varepsilon = \frac{1}{2d}$ is optimal for axis-aligned unit cube partitions is that it is not fundamentally a result about unit cube partitions; rather it is a result about a single cube which is enclosed by a bunch of other cubes—a structure which we call a "cube enclosure" and study throughout Chapter 9 (Secluded Partitions Without ε).

5.1 Optimality of k for Unit Cube Partitions

The formal claim that k is optimal for axis-aligned unit cube partitions is stated as follows:

> **Theorem 5.1.1** (Optimality of $k = d + 1$ for Cube Partitions)**.** *Let $d \in \mathbb{N}$, and $\mathcal{P}$ an axis-aligned unit cube partition of $\mathbb{R}^d$. Then there exists $\vec{p} \in \mathbb{R}^d$ such that $|\mathcal{N}_{\vec{0}}(\vec{p})| \geq d + 1$.*

In other words, there is some point at the closure of at least $d+1$ cubes in an axis-aligned unit cube partition. As mentioned in Section 1.2 (Mathematical Motivation and Background), Lebesgue had conjectured in 1911 that this was true [Leb11] and it was known to be true two years later from the work of Brouwer [Bro13]. We will nonetheless offer three extremely short and distinct proofs of this result.

The first proof will utilize the Lebesgue covering theorem[3] to give the historical context of the knowledge. The second proof follows almost immediately from a result of Alon and Füredi. The third proof will demonstrate that this optimality result can be obtained from a relatively short proof from first principles utilizing an intuitive lemma about sums of powers of 2; this stands in stark contrast the to the fact that both the Lebesgue covering theorem (or Sperner's lemma) and the result of Alon and Füredi themselves utilize significant algebraic topology or combinatorics. However, this is not surprising because the proof of optimality utilizing the Lebesgue covering theorem will actually hold for a broader class of partitions than axis-aligned unit cube partitions, and the proof of optimality utilizing Alon and Füredi's theorem will

[3]The work of Brouwer [Bro13] was on the dimension of a topological space, and it is well-known that the invariance of dimension is related to Brouwer's fixed-point theorem which is in turn equivalent to the combinatorial Sperner's lemma. Lebesgue's covering theorem can (and for our purposes should) be viewed as equivalent to a certain variant of Sperner's lemma on the cube. We will discuss and make use of this variant of Sperner's lemma in Chapter 6 (Optimality of k in General).

only use a weak version of their full result. Thus, the significance of the Lebesgue covering theorem proof is both historical and broader than the result stated above; the significance of the proof using Alon and Füredi's theorem is to make a quick connection between our work and prior work on cubes; and the significance of the proof from first principles is that this optimality result does not require deep tools because of the simplicity that unit cubes can offer compared to other sets.

5.1.1 Proof via Lebesgue Covering Theorem

To begin, we state the Lebesgue covering theorem as taken from the textbook [HW48, Theorem IV 2].

Theorem 5.1.2 (Lebesgue Covering Theorem)**.** *Let $\mathcal{S}$ be a finite closed cover of $[0,1]^d$ in which no set contains points from opposite faces of $[0,1]^d$. Then there exists a point $\vec{p} \in [0,1]^d$ belonging to at least $d+1$ sets in $\mathcal{S}$.*

Obviously, the Lebesgue cover theorem is not specific to a unit cube, so we will apply it to the cube $[0,D]^d$ for some $D>1$.

Proof of Theorem 5.1.1 via the Lebesgue covering theorem. Pick any $D > 1$ and consider the cube $[0,D]^d$. Let $\mathcal{S}$ be the subset of members of $\mathcal{P}$ which intersect $[0,D]^d$: $\mathcal{S} = \left\{X \in \mathcal{P} : X \cap [0,D]^d \neq \emptyset\right\}$. By a simple volume argument[4], the cardinality of $\mathcal{S}$ is finite. Because $\mathcal{P}$ is a partition, the collection $\mathcal{S}$ is clearly a cover of $[0,D]^d$, and thus the set $\left\{\overline{X} : X \in \mathcal{S}\right\}$ is trivially a finite closed cover of $[0,D]^d$. Furthermore, no set contains points on opposite faces of $[0,D]^d$ because points on opposite faces are ℓ_∞ distance exactly D apart but members of $\mathcal{P}$ have ℓ_∞ diameter 1 and so the closures also have ℓ_∞ diameter 1.

[4]Because each member of $\mathcal{P}$ has ℓ_∞ diameter at most 1, each member which intersects $[0,D]^d$ is completely contained in $[-1,D+1]^d$, and since all members of $\mathcal{P}$ are disjoint and have Lebesgue measure (i.e. volume) 1, there can be at most $((D+1)-(-1))^d < \infty$ such members.

Thus, by the Lebesgue covering theorem, there is a point $\vec{p} \in [0, D]^d$ belonging to at least $d+1$ sets in $\{\overline{X} : X \in \mathcal{S}\} \subseteq \{\overline{X} : X \in \mathcal{P}\}$. That is, $\vec{p}$ belongs to the closure of at least $d+1$ members of $\mathcal{P}$. In our notation, $|\mathcal{N}_{\vec{0}}(\vec{p})| \geq d+1$. □

Remark 5.1.3. The observant reader will note that the above proof never used the fact that $\mathcal{P}$ consisted of axis-aligned unit cubes—only that every member of $\mathcal{P}$ had ℓ_∞ diameter strictly less than D (so that no member contained points on opposite faces of $[0, D]^d$) and that $[0, D]^d$ intersected finitely many members of $\mathcal{P}$ (for this, it sufficed that there was a common upper bound on the diameter of all members and that there was a common lower bound on the measures[5] of all members). These variants of the optimality result are stated respectively in Theorem 6.2.3 (Strongest Optimality Theorem) and Corollary 6.2.6 (Strongest Optimality Corollary). △

5.1.2 Proof via Alon and Füredi

The following theorem is due to Alon and Füredi [AF93] which states that covering all but one of the vertices of a d-dimensional cube with hyperplanes requires at least d hyperplanes. This result is a claim even about any hyperplanes including ones which are not orthogonal to an axis, however we will only need the (weaker) result as it applies to axis-orthogonal hyperplanes.

Theorem 5.1.4 ([AF93]). *Let $d, t \in \mathbb{N}$. Let $H_1, H_2, \ldots, H_t$ be affine hyperplanes which cover all vertices of $\{-1, 1\}^d \subseteq \mathbb{R}^d$ except for the vertex $\langle -1 \rangle_{i=1}^d$. Then $t \geq d$.*

Proof of Theorem 5.1.1 via Alon and Füredi. Without loss of generality, assume $X = [-1, 0)^d$ is one of the cubes in $\mathcal{P}$, and let $\vec{p} = \vec{0}$. Since $\mathcal{P}$ is an axis-aligned unit cube

[5] A common lower bound on the inner Lebesgue measures would have sufficed, so there is actually no need for the members to be measurable.

partition, by Corollary 3.6.6, $\mathcal{P}$ is locally finite, so by Fact 3.6.5 there exists $\varepsilon \in (0, \infty)$ such that ${}^{\infty}\overline{\mathcal{N}}_{\varepsilon}(\vec{p}) = \mathcal{N}_{\vec{0}}(\vec{p})$. Fixing such ε, to complete the proof, it suffices to show that $\left|{}^{\infty}\overline{\mathcal{N}}_{\varepsilon}(\vec{p})\right| \geq d+1$.

Consider the set $E = \{-\varepsilon, \varepsilon\}^d$. Observe that the only point in E that is also in X is $-\vec{\varepsilon} = \langle -\varepsilon \rangle_{i=1}^d$. Consider any member $Y \in \mathcal{P}$ such that $Y \neq X$ and Y contains some point of E. Clearly, because $\mathcal{P}$ is a partition, Y is disjoint from X, so $-\vec{\varepsilon} \notin Y$. We will find a hyperplane (in fact, an axis orthogonal one) which contains at least the points of E which Y contains and yet still does not contain $-\vec{\varepsilon}$; that is, we will find a hyperplane H_Y such that $H_Y \cap E \supseteq Y \cap E$ and $-\vec{\varepsilon} \notin H_Y$.

Since Y is a translate of $[0,1)^d$, for some $\vec{y} \in \mathbb{R}^d$, $Y = \prod_{i=1}^d [y_i, y_i + 1)$. Thus, $Y \cap E = \prod_{i=1}^d [y_i, y_i+1) \cap \{-\varepsilon, \varepsilon\}$, and it must be that there is some $i_0 \in [d]$ such that $[y_{i_0}, y_{i_0} + 1) \cap \{-\varepsilon, \varepsilon\} = \{\varepsilon\}$ (see justification[6]). Let H_Y be the hyperplane $H_Y = \left\{\vec{x} \in \mathbb{R}^d \colon x_{i_0} = \varepsilon\right\}$. Then

$$H_Y \cap E = \prod_{i=1}^{d} \begin{cases} \{\varepsilon\} & i = i_0 \\ \{-\varepsilon, \varepsilon\} & \text{otherwise} \end{cases}$$

which is trivially a superset of $Y \cap E$ and does not contain $-\vec{\varepsilon}$.

Thus, the set of cubes other than X which contain points of E generates a set (of no greater cardinality) of hyperplanes which cover all points of $E = \{-\varepsilon, \varepsilon\}^d$ except one. Since this requires at least d hyperplanes by Theorem 5.1.4 ([AF93]), there must be at least d cubes in $\mathcal{P}$ other than X which contain points of E, and thus at least $d+1$ cubes in total that intersect $E \subseteq [-\varepsilon, \varepsilon]^d = {}^{\infty}\overline{B}_{\varepsilon}(\vec{p})$, and thus (by definition) at least $d+1$ cubes that belong to ${}^{\infty}\overline{\mathcal{N}}_{\varepsilon}(\vec{p}) = \mathcal{N}_{\vec{0}}(\vec{p})$. □

[6] If it were the case for every $i \in [d]$ that $-\varepsilon \in [y_i, y_i+1) \cap \{-\varepsilon, \varepsilon\}$, then $-\vec{\varepsilon} \in Y \cap E$, but $-\vec{\varepsilon} \notin Y$ since $Y \neq X$. Thus, there is some coordinate i_0 such that $-\varepsilon \notin [y_{i_0}, y_{i_0}+1) \cap \{-\varepsilon, \varepsilon\}$, but it must be that $[y_{i_0}, y_{i_0}+1) \cap \{-\varepsilon, \varepsilon\} \neq \emptyset$ because otherwise $Y \cap E = \emptyset$ but we assumed Y contained some point of E. Thus, $[y_{i_0}, y_{i_0}+1) \cap \{-\varepsilon, \varepsilon\} = \{\varepsilon\}$.

5.1.3 Proof from First Principles

To prove Theorem 5.1.1 from first priciples, we need only the following lemma which says that if 2^d is written as a sum of at most $d+1$ powers of 2 and $2^0 = 1$ is one of the terms, then (up to ordering), the sum must be exactly these $d+1$ terms: $1+1+2+4+8+\ldots+2^{d-1}$. In other words, there is a unique way to write $2^d - 1$ as a sum of at most d powers of two. The intuition for why this is true is that removing one of the terms which is 1 means that d-many non-negative-powers of 2 sum to 2^d-1, so the binary expression for $2^d - 1$ shows that the sum could consist of one of each smaller power of 2. However, this is not a rigorous argument because though binary expressions are unique, they assume that each power of 2 is used at most once in the summation, and that is not an assumption that we are making on the terms here. We are mostly interested in the fact that $d+1$ terms are required, but we make a stronger statement below so that we have a stronger inductive hypothesis when proving the result.

The details of the proof somewhat hide the main idea, so we quickly sketch the proof in this paragraph. The proof uses induction and the base case will be a triviality. For the inductive case, since the sum is a power of 2, it is even, so there must be an even number of terms which are 1 (and by hypothesis there is at least one term which is 1). The terms which are 1 get added together in pairs to obtain a second sequence with strictly fewer terms (so the last index is at most $d-1$) and all of the terms are even. We divide all terms of this second sequence by 2 to obtain a third sequence whose sum is 2^{d-1} and apply the inductive hypothesis to conclude that the terms are $1, 1, 2, 4, 8, \ldots, 2^{d-2}$. Doubling these shows that the second sequence was $2, 2, 4, 8, 16, \ldots, 2^{d-1}$. We will be careful to show that the original sequence had exactly one pair of 1's that got paired up (this is the main subtlety of the proof) which will show that the original sequence was $1, 1, 2, 4, 8, 16, \ldots, 2^{d-1}$.

Lemma 5.1.5 (Summing Powers of 2)**.** *If $d \in \mathbb{N}$ and $k \in [d] \cup \{0\}$ and $\{a_i\}_{i=0}^{k}$ is a non-decreasing sequence of non-negative-powers of 2 with $a_0 = 1$ and $\sum_{i=0}^{k} a_i = 2^d$, then $k = d$ and for $i \in [k] = [d]$, $a_i = 2^{i-1}$.*

Proof. If $d = 1$, then $k \in [d] \cup \{0\} = \{0, 1\}$ implies that $k = 1$ (k cannot be 0 because $\sum_{i=0}^{0} a_i = a_0 = 1 \neq 2 = 2^d$). Since $a_0 = 1 = 2^{1-1}$ and $a_0 + a_1 = 2$ we have $a_1 = 1$ which proves the base case.

For the inductive case, let $d > 1$ and suppose the claim holds for $d - 1$; we consider $k \in [d] \cup \{0\}$ and a sequence $\{a_i\}_{i=0}^{k}$ (as before, it is trivial that $k \neq 0$). Since $\sum_{i=0}^{k} a_i = 2^d$ is even and all terms are positive integers (because they are non-negative-powers of 2), an even number of terms must be 1; since $a_0 = 1$ it must be a strictly positive even number of terms, and because the sequence is non-decreasing, it must be terms a_0 through a_{2n-1} for some positivie n (also, since $2n - 1 \leq k$ it must be that $n \leq (k+1)/2$ which is at most k since $k \geq 1$; i.e. $n \in [k]$).

Now consider a second sequence $\{b_i\}_{i=0}^{k-n}$ which "collapses" the terms of $\{a_i\}_{i=0}^{k}$ which are 1 into pairs, and leaves the remaining terms as is:

$$b_i = \begin{cases} a_{2i} + a_{2i+1} \ (= 2) & 0 \leq i < n \\ a_{i+n} & n \leq i \leq k - n. \end{cases}$$

(Note that the indexing is valid.)

Thus, every term of $\{b_i\}_{i=0}^{k-n}$ is even, and this sequence is non-decreasing because for $0 \leq i < n$, $b_i = 2$, and $b_n = a_{2n} \geq 2$, and for $i > n \geq 1$, $b_i = a_{i+n} \geq a_{i+n-1} = b_{i-1}$.

Now consider a third sequence $\{b_i/2\}_{i=0}^{k-n}$ which is also non-decreasing with every term a non-negative-power of 2 (because every term of $\{b_i\}_{i=0}^{k-n}$ is a positive power of 2). Since $\sum_{i=0}^{k-n} b_i = \sum_{i=0}^{k} a_i = 2^d$, then $\sum_{i=0}^{k-n} \frac{b_i}{2} = 2^{d-1}$, so we can apply the inductive hypothesis (because also $\frac{b_0}{2} = \frac{2}{2} = 1$ and $n \geq 1$ so $k - n \leq d - 1$). By inductive hypothesis, $k - n = d - 1$ and thus $n = 1$ (since $k \leq d$) and so $k = d$; also by inductive hypothesis, for $i \in [k-n] = [d-1]$, $\frac{b_i}{2} = 2^{i-1}$ so $b_i = 2^i$. Now we use this to complete the proof and show for $i \in [k] = [d]$ that $a_i = 2^{i-1}$.

Since we established that $n = 1$, we can simplify the expression of b_i to

$$b_i = \begin{cases} a_0 + a_1 & i = 0 \\ a_{i+1} & i \in [k-1] = [d-1]. \end{cases}$$

Then $a_1 = 1$ because $a_0 + a_1 = b_0 = 2$ and $a_0 = 1$. As just indicated, for $i \in [k-1] = [d-1]$, $b_i = a_{i+1}$, so reindexing, for $i \in [d] \setminus \{1\}$, $a_i = b_{i-1} = 2^{i-1}$. □

Now we can prove Theorem 5.1.1. The first paragraph is identical to the Alon and Füredi proof.

Proof of Theorem 5.1.1 from first principles. Without loss of generality, assume $X = [-1, 0)^d$ is one of the cubes in $\mathcal{P}$, and let $\vec{p} = \vec{0}$. Since $\mathcal{P}$ is an axis-aligned unit cube partition, by Corollary 3.6.6, $\mathcal{P}$ is locally finite, so by Fact 3.6.5 there exists $\varepsilon \in (0, \infty)$ such that ${}^{\infty}\overline{\mathcal{N}}_{\varepsilon}(\vec{p}) = \mathcal{N}_{\vec{0}}(\vec{p})$. Fixing such ε, to complete the proof, it suffices to show that $\left|{}^{\infty}\overline{\mathcal{N}}_{\varepsilon}(\vec{p})\right| \geq d + 1$.

Consider the set $E = \{-\varepsilon, \varepsilon\}^d$ of 2^d-many points and observe the following three facts:

1. Each point in E belongs to a unique member of $\mathcal{P}$.
2. Each member of $\mathcal{P}$ contains either 0 or a power of 2 many points of E (see justification[7]).

[7]Given $Y \in \mathcal{P}$, Y is a translate of $[0,1)^d$, so for some $\vec{y} \in \mathbb{R}^d$, $Y = \prod_{i=1}^{d}[y_i, y_i + 1)$. Also, $E = \prod_{i=1}^{d} \{-\varepsilon, \varepsilon\}$, so $Y \cap E = \prod_{i=1}^{d}[y_i, y_i + 1) \cap \{-\varepsilon, \varepsilon\}$. The cardinality of this set is $\prod_{i=1}^{d}|[y_i, y_i +$

3. Exactly one point in E belongs to X.

Thus, we consider the finite set of members of $\mathcal{P}$ which contain at least one point of E. If there are at least $d+1$ such members, we are done as $E \subseteq [-\varepsilon, \varepsilon]^d = {}^{\infty}\overline{B}_{\varepsilon}(\vec{p})$, so ${}^{\infty}\overline{B}_{\varepsilon}(\vec{p})$ intersects at least $d+1$ members. Otherwise, there are at most $d+1$ such members, and X is one of them, so there are at most d other members. Since X contains exactly one point of E, this is a sum of powers of 2, one of the terms is 1, and the sum is equal to 2^d (the cardinality of E). By Lemma 5.1.5, this sum consists of $d+1$ terms, so there are $d+1$ members of $\mathcal{P}$ which contain points of E, and thus $d+1$ members of $\mathcal{P}$ which intersect $[-\varepsilon, \varepsilon]^d = {}^{\infty}\overline{B}_{\varepsilon}(\vec{p})$, and thus (by definition) at least $d+1$ cubes that belong to ${}^{\infty}\mathcal{N}_{\varepsilon}(\vec{p}) = \mathcal{N}_{\vec{0}}(\vec{p})$. □

5.2 Near Optimality of ε for Unit Cube Partitions

We constructed axis-aligned unit cube partitions in Chapter 4 which were (k, ε)-secluded for $k = d+1$ and $\varepsilon = \frac{1}{2d}$, and we have now established in Theorem 5.1.1 that $k = d+1$ is in fact the minimum possible value regardless of ε because in any such partition there will always be at least $d+1$ cubes meeting at a single point. Having shown this, we now sketch what we believe is a very beautiful and simple proof that $\varepsilon = \frac{1}{2d}$ is nearly optimal. Specifically, we show that if $k = d+1$ (and in fact much more generally if $k \leq 2^d$) then it must be that $\varepsilon \leq \frac{\log_4(k)}{d}$. Thus, for $k = d+1$, $\varepsilon \leq \frac{\log_4(d+1)}{d}$ which differs from our construction by only a logarithmic factor (and also differs only by a logarithmic factor when k is a polynomial of d).

We will provide only a sketch of this result for two reasons: (1) the result itself will be superseded by Corollary 9.8.5 which states that $\varepsilon = \frac{1}{2d}$ is exactly optimal (a proof which we will also outline shortly in Section 5.3 (Optimality of ε for Unit

---

$1) \cap \{-\varepsilon, \varepsilon\}|$, and every term of this product is either 0, 1, or 2, so the cardinality is either 0 or a power of 2.

Cube Partitions)) and (2) the proof techniques of this result will be adapted to more powerful techniques in Chapter 7 (Near Optimality of ε in General). Nonetheless, the sketch we give is relatively simple[8], quite elegant, and inspires the more powerful techniques in Chapter 7.

We emphasize once more that while this entire section will be very informal, the stated results can be formally backed up, and most of them will be superseded by more formal results in Chapter 7.

In order to establish an upper bound on ε for a given value of k, we want to show (by the contrapositive of being (k,ε)-secluded) that for large enough ε, there is some point $\vec{p} \in \mathbb{R}^d$ for which ${}^{\infty}\overline{\mathcal{N}}_\varepsilon(\vec{p})$ has large cardinality meaning that ${}^{\infty}\overline{B}_\varepsilon(\vec{p})$ intersects a substantial number of cubes in the partition. Since cubes in the partition are half-open, it will suffice to show that the ball intersects many open cubes[9]. To show this, we will use the following key change of perspective:

Observation 5.2.1 (A Change of Perspective)**.** *An open axis-aligned unit cube $X = \vec{x} + (-\frac{1}{2}, \frac{1}{2})^d$ intersects the ball ${}^{\infty}\overline{B}_\varepsilon(\vec{p})$ if and only if the point $\vec{x}$ belongs to the ball ${}^{\infty}B^{\circ}_{\frac{1}{2}+\varepsilon}(\vec{p})$.*

Proof. Note that X is itself an open ℓ_∞ ball of radius $\frac{1}{2}$ centered at $\vec{x}$. Thus, X intersects ${}^{\infty}\overline{B}_\varepsilon(\vec{p})$ if and only if $\|\vec{x} - \vec{p}\|_\infty < \frac{1}{2} + \varepsilon$ which happens if and only if $\vec{x} \in {}^{\infty}B^{\circ}_{\frac{1}{2}+\varepsilon}(\vec{p})$. □

What this change in perspective allows us to do is to consider the set of all positions of cubes in an axis-aligned unit cube partition and consider how many of them we are guaranteed to be able to capture inside of an open ℓ_∞ ball of radius $\frac{1}{2} + \varepsilon$ (i.e. inside

[8] It is relatively simple if the intuitive/visual idea of the proof Blichfeldt's theorem is clear to the reader, however this is challenging to describe clearly in text. Hopefully the provided visuals help.

[9] Assuming the cubes are open will make the discussion in this section more straightforward and will not have any significant impact on the results here.

of an open cube of side length $\frac{1}{2} + \varepsilon$). The problem of determining how many points are contained inside of a specific region is a common task in discrete mathematics; often it is the case that the points form a lattice, and we have already seen that our reclusive constructions have positions forming a lattice, so it makes sense to temporarily restrict ourselves to considering axis-aligned unit cube partitions whose center points[10] form a lattice. Trying to solve this new problem led us to Blitchfeldt's theorem. The following can be generalized to lattices, but we will state it just for the standard integer lattice for the time being.

Theorem 5.2.2 (Blitchfeldt's Theorem)**.** *Let $d \in \mathbb{N}$ and $S \subseteq \mathbb{R}^d$ be a set with volume (Lebesgue measure) V. Then there exists a translation of S which contains at least $\lceil V \rceil$ points of the integer lattice (i.e. points with all integer coordinates).*

The proof is the real key to applying this result, and all of the insight can be demonstrated in two dimensions, so we sketch the proof in $\mathbb{R}^2$ below and visually demonstrate it in Video 5.2.1. This proof idea is present in Blichfeldt's original paper [Bli14] where it was credited to Professor Birkhoff.

Proof Sketch. Separate the 2-dimensional plane into unit squares with centers given by the integer lattice points. Use the squares to chop up the set S so that there is zero or one piece of S in each square region. Translate each piece of S to the square at the origin so that its position relative to the origin square is the same position it was originally in relative to the square it was in. At this point, all pieces of S are located in the same square.

[10] In Section 4.2 (Reclusive Partitions), we used the representative corners as the positions of the cubes, but for this discussion it will be more useful to consider the positions to be the center point of every cube so that we can view each cube as a superset of the open ℓ_∞ $\frac{1}{2}$-ball located at its center position.

Because S has a total area of V and has been chopped up to fit inside of a unit square (which has area 1). Consider how many pieces of S each point of this square is contained in. If each point of the square is contained in at most n pieces, then S has area at most n. Intuitively, if S is a shape cut out of paper and we cut up that shape into lots of pieces and lay them all down on a square table of area 1 and no matter where on the table we look we only see n layers of paper, then the total area of all the pieces in total is at most n, so the area of S to begin with was at most n.

Since the area of S is V, this means that $V \leq n$ and since n is an integer, we have $\lceil V \rceil \leq n$. This means there is some point $\vec{p}$ within the origin square (a point on the table) that is covered by at least $n \geq \lceil V \rceil$ pieces of S.

Now pierce the table at $\vec{p}$ so that every piece of S covering $\vec{p}$ is also pierced there. This puts a piercing/mark on at least $\lceil V \rceil$ pieces of S. Importantly, this piercing is in the same place relative to the origin on all pieces.

Now move every piece of S back to its original location so the set/shape S is reconstructed back to how it originally looked and there are at least $\lceil V \rceil$ marks/piercings on the shape S. Finally, translate the entire shape S by the vector $-\vec{p}$. This translation will move every mark/piercing on S to one of the integer lattice points because relative to the square they are in, all of the marks/piercings are in the same position, and we know that when all the pieces were in the origin square they were located at position $\vec{p}$, so translating them by $-\vec{p}$ would place them at the origin with respect to the origin square.

Thus, this translation of S contains at least $\lceil V \rceil$ marks/piercings, and each one is located at an integer lattice point which means this translation of S contains at least $\lceil V \rceil$ points of the integer lattice. □

Importantly, though the exact statement of Blichfeldt's theorem will vary depending on the source, this proof idea is not dependent on the points being from the integer lattice. This chopping into squares works no matter how the plane is chopped into squares. In other words, for any partitioning of the plane into squares, there is a translation of S that contains at least $\lceil V \rceil$ different centers of the squares. And this works not just in the plane, but in any partition of any Euclidean space. From this proof, we get the following result.

Video 5.2.1: Blichfeldt's theorem

Porism 5.2.3 (Blichfeld Porism)**.** *Let $d \in \mathbb{N}$ and $\varepsilon \in (0, \infty)$ and $\mathcal{P}$ be an axis-aligned unit cube partition of $\mathbb{R}^d$. There exists $\vec{p} \in \mathbb{R}^d$ such that $\vec{p} + (-\frac{1}{2} - \varepsilon, \frac{1}{2} + \varepsilon)^d = {}^{\infty}B^{\circ}_{\frac{1}{2}+\varepsilon}(p)$ contains the center point of at least $\lceil (1 + 2\varepsilon)^d \rceil$-many cubes in $\mathcal{P}$.*

Proof. The set $(-\frac{1}{2} - \varepsilon, \frac{1}{2} + \varepsilon)^d$ has volume $(1 + 2\varepsilon)^d$, so by the same proof technique as Blichfeldt's theorem above, there is some translation[11] $\vec{p} \in \mathbb{R}^d$ such that $\vec{p} + (-\frac{1}{2} -$

[11] In fact, following the proof sketch of Blichfeldt's theorem, $\vec{p} \in [-\frac{1}{2}, \frac{1}{2}]^d$, and if one is especially

$\varepsilon, \frac{1}{2} + \varepsilon)^d$ contains the centers of at least $\lceil (1+2\varepsilon)^d \rceil$-many cubes in $\mathcal{P}$. $\square$

Combining this with Observation 5.2.1 gives the following corollary.

Corollary 5.2.4 (ε-Ball Intersects Many Cubes)**.** *Let $d \in \mathbb{N}$ and $\varepsilon \in (0, \infty)$ and $\mathcal{P}$ be an axis-aligned unit cube partition of $\mathbb{R}^d$. There exists $\vec{p} \in \mathbb{R}^d$ such that $\left| {}^{\infty}\mathcal{N}_{\varepsilon}(\vec{p}) \right| \geq \lceil (1+2\varepsilon)^d \rceil$.*

Proof. From Porism 5.2.3 there exists $\vec{p} \in \mathbb{R}^d$ such that $\vec{p} + (-\frac{1}{2} - \varepsilon, \frac{1}{2} + \varepsilon)^d$ contains the centers of at least $\lceil (1+2\varepsilon)^d \rceil$-many cubes in $\mathcal{P}$, so by Observation 5.2.1, ${}^{\infty}\overline{B}_{\varepsilon}(\vec{p})$ intersects the open version of each such cube (and thus trivially intersects the half-open version of each such cube). Thus, ${}^{\infty}\overline{B}_{\varepsilon}(\vec{p})$ intersects at least $\lceil (1+2\varepsilon)^d \rceil$ cubes in $\mathcal{P}$, so by definition $\left| {}^{\infty}\mathcal{N}_{\varepsilon}(\vec{p}) \right| \geq \lceil (1+2\varepsilon)^d \rceil$. $\square$

With a bit of analysis, this becomes the upper bound on ε. We use the following fact.

Fact G.0.3. *For $\varepsilon \in [0, \frac{1}{2}]$ it holds that $\log_4(1+2\varepsilon) \geq \varepsilon$.*

One can note that equality holds for $\varepsilon = 0$ and $\varepsilon = \frac{1}{2}$. This inequality allows us to state a nice bound on ε in terms of k. Recall that we wish for (k, ε)-secluded partitions to have a small value of k, so we almost certainly don't want k to be larger than 2^d, so including the assumption that it is not in the result below is not a very strong hypothesis.

Theorem 5.2.5 (Near Optimality of $\varepsilon = \frac{1}{2d}$ for Unit Cube Partitions)**.** *Let $d, k \in \mathbb{N}$ with $k \leq 2^d$, $\varepsilon \in (0, \infty)$, and $\mathcal{P}$ a (k, ε)-secluded axis-aligned unit cube partition of $\mathbb{R}^d$. Then $\varepsilon \leq \frac{\log_4(k)}{d}$. In particular, if $k = d+1$, then $\varepsilon \leq \frac{\log_4(d+1)}{d}$.*

careful it can be guaranteed that $\vec{p} \in [-\frac{1}{2}, \frac{1}{2})^d$ because all of the squares in the proof sketch of Blichfeldt's theorem can be taken to be half-open so that every point belongs to a unique square, and the translation vector belongs to the origin square $[-\frac{1}{2}, \frac{1}{2})^d$.

Proof. Recall that by definition of a (k, ε)-secluded partition, for every point $\vec{x} \in \mathbb{R}^d$, it must be the case that ${}^{\infty}\overline{B}_{\varepsilon}(\vec{x})$ intersects at most k members of $\mathcal{P}$. By Corollary 5.2.4, there is a point $\vec{p}$ such that ${}^{\infty}B^{\circ}_{\varepsilon}(\vec{p})$ intersects at least $(1+2\varepsilon)^d$ members of $\mathcal{P}$, so the closed ball ${}^{\infty}\overline{B}_{\varepsilon}(\vec{p})$ trivially intersects at least $(1+2\varepsilon)^d$ members of $\mathcal{P}$. Thus, we have $k \geq (1+2\varepsilon)^d$.

Because $k \leq 2^d$ by hypothesis, this implies $\varepsilon \leq \frac{1}{2}$, so by Fact G.0.3 we have

$$\log_4(k) \geq \log_4\left((1+2\varepsilon)^d\right) = d\log_4(1+2\varepsilon) \geq d\varepsilon$$

so solving for ε, we have $\varepsilon \leq \frac{\log_4(k)}{d}$. Also, for each $d \in \mathbb{N}$ it holds that $d+1 \leq 2^d$ which shows the "in particular" part of the statement. □

We will later adapt the ideas in the proof of Blichfeldt's theorem to handle more general partitions in Section 7.1 (Upper Bounds on ε via Brunn-Minkowski and Blichfeldt). Essentially, we will reverse perspective of Observation 5.2.1—rather than thinking of the ball growing and the cube members shrinking to points, we think of the ball shrinking to a point and the members growing. Then we want to find a single point located inside many enlarged members. We will show that the members gain substantial volume when enlarged and use a volume-averaging/pigeonhole argument similar to that in the proof sketch of Blichfeldt's theorem (though much more formal) along with some other measure theory results to argue that such a point exists. We will get the same result that an ℓ_∞ ball of radius ε intersects at least $\lceil (1+2\varepsilon)^d \rceil$ members of the partition and use the approximation ideas in the proof of Theorem 5.2.5 to obtain the same bound that $\varepsilon \leq \frac{\log_4(k)}{d}$ when $k \leq 2^d$. However, this alternate change of perspective will allow us to consider a much more extensive class of partitions: any partition for which each member has (outer) Lebesgue measure at most 1 (which includes all partitions for which each member has ℓ_∞ diameter at most 1). These more general techniques will also allow us to obtain results for every norm—not just the ℓ_∞ norm.

5.3 Optimality of ε for Unit Cube Partitions

Before turning to a much more general setting of the partitions under consideration, we will outline the proof that for axis-aligned unit cube partitions, $\varepsilon = \frac{1}{2d}$ is the maximum possible value of ε when $k = d + 1$. We state the result (Corollary 9.8.5) here for completeness of the presentation of this chapter. Section 9.8 (Optimal ε For Unit Cube Enclosures) is devoted to the proof of the result about cube enclosures which implies the stated corollary for axis-aligned unit cube partitions. However, Section 9.8 also builds on other results from Chapter 9 (Secluded Partitions Without ε), so it should not be read in isolation. Fundamentally, the following result is not really a result about d-dimensional space as it is about fitting $d + 1$ points along a unit length line which requires two points to be distance at most $\frac{1}{d}$ apart.

Corollary 9.8.5 (Optimality of $\varepsilon = \frac{1}{2d}$ for Unit Cube Partitions)**.** *Let $d \in \mathbb{N}$ and $\varepsilon \in (0, \infty)$ and $\mathcal{P}$ a $(d+1, \varepsilon)$-secluded axis-aligned unit cube partition of $\mathbb{R}^d$. Then $\varepsilon \leq \frac{1}{2d}$.*

Proof Outline. We begin by choosing an arbitrary cube $X \in \mathcal{P}$ to focus on. Let $\vec{x} = \text{center}(X)$ for convenience. We then considering the corner $\vec{p} = \vec{x} - \frac{1}{2} \cdot \vec{1} = \langle x_i - \frac{1}{2} \rangle_{i=1}^d$ of X which is negative of the center in every direction (we could use any corner, but this one is the most convenient). We will work with all of the cubes of the partition which contain $\vec{p}$ in their closure (i.e. the set $\mathcal{N}_{\bar{0}}(\vec{p})$). As show earlier using the powers of 2 in the proof of Theorem 5.1.1 in Subsection 5.1.3 (Proof from First Principles), it must be that $|\mathcal{N}_{\overline{0}}(\vec{p})| \geq d+1$, and since $\mathcal{P}$ is $(d+1,\varepsilon)$-secluded it must be that equality holds because there cannot be $d+2$ cubes that meet at a single point.

Next, we utilize a result (Theorem 9.5.9 and Theorem 9.5.8) which extends the ideas in the proof of Theorem 5.1.1 to give additional structural information about the cubes in $\mathcal{N}_{\overline{0}}(\vec{p})$; specifically, there is a so-called Minkowski twin of X in this neighborhood—a cube Y with position $\text{center}(Y)$ identical to $\vec{x}$ in every coordinate except for a unique $i_0 \in [d]$ and at this coordinate $\text{center}_{i_0}(Y) = x_{i_0} - 1$. This coordinate i_0 will be important for the rest of the argument as we will consider the edge of X along coordinate i_0 which includes $\vec{p}$—that is, the edge between corners $\vec{p}$ and $\vec{q} \overset{\text{def}}{=} \vec{p} + \vec{e}^{(i_0)}$.

We then show that for any two cubes $Z, Z' \in \mathcal{N}_{\overline{0}}(\vec{p})$, it holds that $\text{center}_{i_0}(Z) \neq \text{center}_{i_0}(Z')$ so that all cubes are uniquely positioned along the edge in question. For each cube in $\mathcal{N}_{\overline{0}}(\vec{p})$, we consider the point $\vec{s}^{(Z)}$ on this edge at the "far end" of the cube. That is, for each cube Z, we consider the i_0th projection of Z (i.e. the interval $\pi_{i_0}(Z) = \text{center}_{i_0}(Z) + [-\frac{1}{2}, \frac{1}{2})$) and consider the point

$$\vec{s}^{(Z)} = \left\langle \begin{cases} p_i = q_i & i \neq i_0 \\ \text{center}_{i_0}(Z) + \frac{1}{2} & i = i_0 \end{cases} \right\rangle_{i=1}^{d}$$

which is at the rightmost extreme of the interval in the i_0th coordinate. Because $Z \in \mathcal{N}_{\overline{0}}(\vec{p})$, by definition $\vec{p} \in \overline{Z}$ which implies that $\text{center}_{i_0}(Z) \in [p_{i_0} - \frac{1}{2}, p_{i_0} + \frac{1}{2}]$ so that $s_{i_0}^{(Z)} = \text{center}_{i_0}(Z) + \frac{1}{2} \in [p_{i_0}, p_{i_0} + 1] = [p_{i_0}, q_{i_0}]$ showing that $\vec{s}^{(Z)}$ is a point which lies on the edge between $\vec{p}$ and $\vec{q}$.

We proceed by showing that for each $Z \in \mathcal{N}_{\overline{0}}(\vec{p})$, the neighborhood $\mathcal{N}_{\overline{0}}(\vec{s}^{(Z)})$ has cardinality $d+1$ and for distinct $Z, Z' \in \mathcal{N}_{\overline{0}}(\vec{p})$, the neighborhoods $\mathcal{N}_{\overline{0}}(\vec{s}^{(Z)})$ and $\mathcal{N}_{\overline{0}}(\vec{s}^{(Z')})$ are distinct.

It is at this point that we reach the crux of the argument. The set $S = \left\{ \vec{s}^{(Z)} : Z \in \mathcal{N}_{\overline{0}}(\vec{z}) \right\}$ consists of $|\mathcal{N}_{\overline{0}}(\vec{p})| = d + 1$ distinct points along the unit length edge/line from $\vec{p}$ to $\vec{q}$. Since there are $(d + 1)$-many points along a unit length line, there must be two points $\vec{a}, \vec{b} \in S$ which are distance at most $\frac{1}{d}$ apart (because $d + 1$ points in a unit interval will segment the interval into at least d different pieces, and it cannot be that all $\geq d$ pieces have length exceeding $\frac{1}{d}$). If we assume for contradiction that ε is larger than $\frac{1}{2d}$, we can consider the midpoint $\vec{c}$ of $\vec{a}$ and $\vec{b}$ and note that ${}^{\infty}\overline{B}_{\varepsilon}(\vec{c})$ contains both $\vec{a}$ and $\vec{b}$ in its interior. This implies that ${}^{\infty}\mathcal{N}_{\varepsilon}(\vec{c}) \supseteq \mathcal{N}_{\overline{0}}(\vec{a}) \cup \mathcal{N}_{\overline{0}}(\vec{b})$ because any cube containing $\vec{a}$ (resp. $\vec{b}$) in its closure will then intersect ${}^{\infty}\overline{B}_{\varepsilon}(\vec{c})$. Since $\mathcal{N}_{\overline{0}}(\vec{a})$ and $\mathcal{N}_{\overline{0}}(\vec{b})$ are distinct and each have cardinality $d+1$, their union has cardinality at least $d+2$ which implies that $\left|{}^{\infty}\mathcal{N}_{\varepsilon}(\vec{c})\right| \geq d+2$. This would contradict the hypothesis that $\mathcal{P}$ is $(d + 1, \varepsilon)$-secluded, and so we conclude that $\varepsilon \leq \frac{1}{2d}$. □

To conclude Chapter 5 (Optimality of k and ε for Unit Cube Partitions), we summarize the main results which show that we have fully resolved the question of the level of seclusion that can possibly be attained in axis-aligned unit cube partitions.

Summary Result 5.3.1. *Let $d, k \in \mathbb{N}$ and $\varepsilon \in (0, \infty)$ and $\mathcal{P}$ be a (k, ε)-secluded axis-aligned unit cube partition of $\mathbb{R}^d$. Then $k \geq d+1$ and if $k = d+1$, then it must be that $\varepsilon \leq \frac{1}{2d}$. Furthermore, these bounds are tight as there exists $(d+1, \frac{1}{2d})$-secluded axis-aligned unit cube partitions.*

Proof. This is a restatement of Theorem 5.1.1 (Optimality of $k = d + 1$ for Cube Partitions), Corollary 9.8.5 (Optimality of $\varepsilon = \frac{1}{2d}$ for Unit Cube Partitions), and Theorem 4.2.18 (Existence of $(d + 1, \frac{1}{2d})$-Secluded Unit Cube Partitions). □

Chapter 6

Optimality of k in General

We have now seen in Chapter 5 (Optimality of k and ε for Unit Cube Partitions) that our reclusive unit cube partition constructions in Chapter 4 (Constructions) are exactly optimal, and we also mentioned in Remark 5.1.3 that the Lebesgue Covering Theorem (Theorem 5.1.2) could be used to show that even in many partitions that don't have cube members, there must exist a point at the closure of $d+1$ partition members, and so $k = d+1$ is optimal with regard to these families of partitions as well. In this chapter we will make this more precise and consider very general families of partitions and show that $k = d+1$ (matching our unit cube constructions) remains optimal.

To begin with, we consider what it means for $k = d+1$ to be optimal for a fixed class[1] of partitions of $\mathbb{R}^d$. Naturally, this means that (1) there is a partition in the class which is $(d+1, \varepsilon)$-secluded for some ε (to witness that this parameter is attainable) and (2) that there is no partition which is (d, ε)-secluded for any choice of ε (to show that better partitions in the family do not exist).

Thus, once a positive $(d+1, \varepsilon)$-secluded example in the class is known, we must show for every partition $\mathcal{P}$ in the class and for every $\varepsilon \in (0, \infty)$ that $\mathcal{P}$ is not (d, ε)-

[1] We do not mean class in the formal set theoretic language, but rather mean a family of partitions with a certain shared property.

secluded. By definition of seclusion, this means showing for each $\mathcal{P}$ and ε that there exists a point $\vec{p} \in \mathbb{R}^d$ for which $\left|{}^{\infty}\overline{\mathcal{N}}_{\varepsilon}(\vec{p})\right| > d$ (where the neighborhood is with respect to $\mathcal{P}$).

However, when we proved the optimality of $k = d + 1$ for the class of axis-aligned unit cube partitions in Theorem 5.1.1, we actually proved something stronger. Not only was it the case that "$\forall \mathcal{P}\, \forall \varepsilon\, \exists \vec{p}\ \left|{}^{\infty}\overline{\mathcal{N}}_{\varepsilon}(\vec{p})\right| > d$", it was actually the case that there was a single point at the closure of greater than d (i.e. at least $d+1$) members which in particular implies[2] that "$\forall \mathcal{P}\, \exists \vec{p}\, \forall \varepsilon\ \left|{}^{\infty}\overline{\mathcal{N}}_{\varepsilon}(\vec{p})\right| > d$" which is a stronger qualified statement than the former (and still a weaker statement than saying there is a point at the closure of at least $d+1$ members). Thus, we have just introduced three (possibly distinct) notions of the optimality of the k parameter, and it turns out that all three arise naturally.

We summarize the three types below. The "standard optimality" is the natural definition of optimality of $k = d + 1$ based on the definition of seclusion. The "strongest optimality" is the notion of optimality that we found for axis-aligned unit cube partitions, and between them is the "stronger optimality" notion which uses the changed quantifier order.

Standard Optimality of $k = d + 1$:

For every $\mathcal{P}$ and every $\varepsilon \in (0, \infty)$ there exists $\vec{p}$ such that $\left|{}^{\infty}\overline{\mathcal{N}}_{\varepsilon}(\vec{p})\right| \geq d + 1$

Stronger Optimality of $k = d + 1$:

For every $\mathcal{P}$ there exists $\vec{p}$ such that for every $\varepsilon \in (0, \infty)$ we have $\left|{}^{\infty}\overline{\mathcal{N}}_{\varepsilon}(\vec{p})\right| \geq d+1$

Strongest Optimality of $k = d + 1$:

For every $\mathcal{P}$ there exists $\vec{p}$ such that $|\mathcal{N}_{\vec{0}}(\vec{p})| \geq d + 1$

Observe that "strongest optimality" implies[2] "stronger optimality" which in turn

[2] The implication is because if there is a point $\vec{p}$ such that $|\mathcal{N}_{\vec{0}}(\vec{p})| > d$, then this by definition means that $\vec{p}$ is at the closure of at least $d + 1$ members, so every open ε-ball will intersect those $d + 1$ members.

implies[3] "standard optimality". As alluded to already, the reason that these three notions of optimality will be considered is that they are the types of optimality that arise from some natural classes of partition.

With only a little consideration, it becomes obvious that $k = d + 1$ is not optimal for the class consisting of all partitions of $\mathbb{R}^d$ because for the partition of $\mathbb{R}^d$ which contains just one member (i.e. $\mathcal{P} = \left\{\mathbb{R}^d\right\}$) we have for any $\varepsilon \in (0, \infty)$ and $\vec{p} \in \mathbb{R}^d$ that $\left|{}^{\infty}\overline{\mathcal{N}}_{\varepsilon}(\vec{p})\right| = 1$ because $\mathcal{P}$ contains only one member. A similar argument works for any partition of $\mathbb{R}^d$ that consists of fewer than $d + 1$ members (and to a variety of other simple constructions). Thus, we need to impose some restriction on the class of partitions we are considering to have non-trivial results. A very natural restriction is to uniformly bound the size of the partition members in some way, with the two most obvious options being to upper bound the diameter of the members or to upper bound the (outer) measure[4] of the members.

Note that upper bounding the diameter of a set X (using a norm $\|\cdot\|$) is a strictly stronger condition than upper bounding the outer measure because if we insist that X has diameter at most D relative to $\|\cdot\|$, then X is contained in the ball ${}^{\|\cdot\|}\overline{B}_D(\vec{p})$ for some $\vec{p} \in \mathbb{R}^d$, so $m_{out}(X) \leq m\left({}^{\|\cdot\|}\overline{B}_D(\vec{0})\right)$ (we can take $\vec{p}$ to be any point in X if X is non-empty, otherwise we can take $\vec{p}$ to be any point). Thus, if we could show that $k = d + 1$ is optimal in the "strongest optimality" sense for the class of partitions which have a uniform upper bound on the outer measure of the members, then we would have no need to consider the class of partitions which have a uniform upper bound on the diameters of the members. However, we will find that this is not true (see Proposition 6.4.2 ("Stronger" and "Strongest" Optimality Gap)). What we will see, though, is that the following (informal) results hold.

[3] The implication is a general one for changing the quantifier order in this way.

[4] We use the outer measure so that we don't have to assume measurability of the partition members.

Informal Theorem (Standard Optimality Theorem (6.2.1)). *The value $k = d+1$ is optimal in the "standard optimality" sense for the class of partitions which have a uniform upper bound on the outer measure of the members.*

Informal Theorem (Stronger Optimality Theorem (6.2.2)). *The value $k = d+1$ is optimal in the "stronger optimality" sense for the class of partitions which have a uniform upper bound on the diameter of the members.*

Informal Theorem (Strongest Optimality Theorem (6.2.3)). *The value $k = d+1$ is optimal in the "strongest optimality" sense for the class of partitions which have a uniform upper bound on the diameter of the members and also a certain type of local finiteness.*

Informal Corollary (Strongest Optimality Corollary (6.2.6)). *The value $k = d+1$ is optimal in the "strongest optimality" sense for the class of partitions which have a uniform upper bound on the diameter of the members and also a non-zero lower bound on the inner measures of the members. (This includes unit cube partitions.)*

This will show that in the sense of optimality which we really care about ("standard optimality"), $k = d+1$ is the optimal value of k for (k,ε)-secluded partitions amongst the extremely broad class of partitions for which there is some upper bound on the outer measure of all members of the partition. Furthermore, we are able to strengthen this notion of optimality as we restrict the class to bounded diameter partitions, and to bounded diameter partitions with other additional conditions. We will also see that there are "gaps" between the three results above;

that is to say that the conclusions in each of the above three results are tight in some sense.

The Stronger Optimality Theorem and Strongest Optimality Theorem follow quite directly from an adaption of the Lebesgue Covering Theorem (Theorem 5.1.2), so they should probably be considered known results even if we are the first to state them in this particular fashion. However, the proof of the Standard Optimality Theorem requires a fair bit of work beyond the use of the Lebesgue Covering Theorem and constitutes an important contribution to the problem at hand. Again, because “standard optimality” is really what we care about, this result is very robust in its claim that we get this optimality of $k = d + 1$ for a very broad class of partitions which requires only the extremely weak assumption that there is a common upper bound on the outer measure of all members. Thus, the Standard Optimality Theorem should be considered the main result of this chapter.

One final (fairly minor) item that we have to address before proceeding is that on occasion, diameter is not quite what we want. For example, consider the sets $[0, 1)$ and $[0, 1]$. Both sets have diameter 1, but the former has the stronger property that all pairs of points have distance strictly less than the diameter. In other words, diameter[5] 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 6.0.1 (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 $\text{distance}(x, y) < D$, then we call D a strict pairwise bound of X with respect to this metric.

Unbounded sets do not have any strict pairwise bounds (though we could say ∞

[5]In general $\text{diameter}(X) \stackrel{\text{def}}{=} \sup \{\text{distance}(x, y) : x, y \in X\}$

is a strict pairwise bound), and bounded sets have infinitely many strict pairwise bounds (if D is a strict pairwise bound, then so is D' for any $D' > D$). This is unlike diameter where any bounded set has a unique diameter. Also, sets that do not attain their diameter have the property that the diameter is the smallest strict pairwise bound, and sets that do attain their diameter do not have a smallest strict pairwise bound. Also, we have the following implications for any constant D:

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

In general, the reverse implications do not hold which is exemplified by the sets $[0, D)$ and $[0, D]$. This demonstrates why we defined this notion: for a constant D it gives a condition on sets strictly weaker than requiring diameter less than D but strictly stronger than requiring diameter at most D.

6.1 Prerequisite Results: Lebesgue, Sperner, and KKM

Recall the statement of the Lebesgue Covering Theorem given earlier, and recall that a face F of the cube $[0,1]^d$ is a product set $F = \prod_{i=1}^{d} F_i$ where each F_i is one of three sets: $\{0\}$, $\{1\}$, or $[0,1]$, and two faces F, F' are said to be opposite each other if there is some coordinate $i_0 \in [d]$ such that $F_{i_0} = \{0\}$ and $F'_{i_0} = \{1\}$ (or vice versa).

Theorem 5.1.2 (Lebesgue Covering Theorem)**.** *Let $\mathcal{S}$ be a finite closed cover of $[0,1]^d$ in which no set contains points from opposite faces of $[0,1]^d$. Then there exists a point $\vec{p} \in [0,1]^d$ belonging to at least $d+1$ sets in $\mathcal{S}$.*

As stated, the Lebesgue Covering Theorem (Theorem 5.1.2) requires a cover of $[0,1]^d$ by closed sets in which no set contains points on opposite faces of $[0,1]^d$. Because the sets in this cover are both closed and can be assumed to be bounded by

$[0,1]^d$, they are compact which guarantees that they attain their diameter, and thus no set can have ℓ_∞ diameter 1 because it would then include points on opposite faces of $[0,1]^d$. Thus, each set in the cover has ℓ_∞ diameter strictly less than 1. Despite the fact that this is how the result is stated, this is not a necessary condition. We could instead state the conclusion of the Lebesgue Covering Theorem for any cover by (not necessarily closed) sets as long as no set contains points on opposite faces of the cube—i.e. each set in the cover has a strict pairwise bound of 1 in the ℓ_∞ norm. We will show that this modification can be seen as the limiting version of Sperner's lemma on the cube where no set (i.e. color) contains points on opposite faces of the cube (as shown in [LPS01] using the Bolzano–Weierstrass theorem). Alternatively, it can be viewed as a corollary of the KKM lemma on the cube which considers closed sets with certain coloring properties so that the sets can actually have ℓ_∞ diameter 1 (c.f. [Kom94, vdLTY99]). Finally, it can also be obtained directly from the statement of the Lebesgue Covering Theorem using our coloring extension argument which will appear later in the proof of Theorem 8.0.7 (Neighborhood KKM/Lebesgue Theorem).

Furthermore, we also want a version of the Lebesgue Covering Theorem which allows for covers with infinitely many members. The two results that we claim in this section are the following, and their proofs are really straightforward exercises using known results, and we provide said proofs in this section. If one accepts these two results as known, then they can skip the remainder of this section.

Theorem 6.1.1 (KKM/Lebesgue)**.** *Given a coloring of $[0,1]^d$ by finitely many colors in which no color includes points of opposite faces, there exists a point at the closure of at least $d+1$ different colors.*

If we ignore the finiteness condition, we still get the following.

Theorem 6.1.2 (Infinite KKM/Lebesgue)**.** *Given a coloring of* $[0,1]^d$ *in which no color includes points of opposite faces, there exists a point* $\vec{p}$ *such that every open set around* $\vec{p}$ *contains points of at least* $d+1$ *different colors.*

In either result, it is fine if the coloring assigns multiple colors to some or all points, because we can just pick one of the colors for each point and apply the result.

We will now lay out how these two theorems follow in particular from the work of De Loera, Peterson, and Su [LPS01] on polytope variants of Sperner's lemma. We define the notion of a Sperner/KKM coloring. This notion is used implicitly in [LPS01], though they focused on general polytopes, and we will need this notion only for the cube.

Definition 6.1.3 (Sperner/KKM Coloring). Let $d \in \mathbb{N}$ and $V = \{0,1\}^d$ denote the set of vertices of the cube $[0,1]^d$. We view V as both the vertices of the cube and a set of colors. Let $\chi : [0,1]^d \to V$ 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 $\vec{x}$ must be the color of one of the vertices defining the face F). Such a function χ will be called a Sperner/KKM coloring.

The above definition of a Sperner/KKM coloring is stated to be analogous to the definition of a Sperner coloring of a simplex. However, it can be equivalently defined in a way more analogous to the hypothesis of the Lebesgue Covering Theorem and to the hypotheses of the two theorems above.

Definition 6.1.4 (Equivalent Definition of Sperner/KKM Coloring). Let $d \in \mathbb{N}$ and $V = \{0,1\}^d$ denote the set of vertices of the cube $[0,1]^d$. Let $\chi : [0,1]^d \to V$ be any function so that (1) for each vertex $\vec{v}$ of the cube, $\chi(\vec{v}) = \vec{v}$ and (2) for a pair of opposite faces $F^{(0)}, F^{(1)}$ of the cube and any $\vec{x}^{(0)} \in F^{(0)}$ and $\vec{x}^{(1)} \in F^{(1)}$, it holds that $\chi(\vec{x}^{(0)}) \neq \vec{x}^{(1)}$ (i.e. no color from χ contains points on opposite faces of the cube).

Proof of Equivalence. (7.2.4) $\implies$ (6.1.4): A face F of $[0,1]^d$ is a d-fold product $F = \prod_{i=1}^d F_i$ where each F_i is either $\{0\}$ or $\{1\}$ or $[0,1]$. So for any vertex $\vec{v} \in \{0,1\}^d$ of the cube, the set $F = \prod_{i=1}^d \{v_i\} = \{\vec{v}\}$ is a face, so by hypothesis $\chi(\vec{v}) \in F$ showing that $\chi(\vec{v}) = \vec{v}$.

If $F^{(0)}, F^{(1)}$ are opposite faces of the cube, this means there is some coordinate i_0 such that (up swapping roles) $F^{(0)}_{i_0} = \{0\}$ and $F^{(1)}_{i_0} = \{1\}$. Then for every point $\vec{x}^{(0)} \in F^{(0)}$ we have by hypothesis that $\chi(x^{(0)}) \in F^{(0)}$ so the ith coordinate of the color $\chi(x^{(0)})$ is 0. Similarly, for every point $\vec{x}^{(1)} \in F^{(1)}$ we have by hypothesis that $\chi(x^{(1)}) \in F^{(1)}$ so the ith coordinate of the color $\chi(x^{(1)})$ is 1. Thus, no point in $F^{(0)}$ is given the same color as a point in $F^{(1)}$.

(6.1.4) $\implies$ (7.2.4): Consider any face F and point $\vec{x} \in F$, and we will show that $\chi(\vec{x}) \in F = \prod_{i=1}^d F_i$ which we do by showing for each $i_0 \in [d]$ that $\chi(\vec{x})_{i_0} \in F_{i_0}$. If $F_{i_0} = [0,1]$, this is trivial (because the codomain of χ is $V = \{0,1\}^d$ so $\chi(\vec{x})_{i_0}$ is either 0 or 1). So we only need to consider that F_{i_0} is $\{0\}$ or is $\{1\}$. We assume $F_{i_0} = \{0\}$ as the case $F_{i_0} = \{1\}$ is symmetric. Since we have $\vec{x} \in F$ so $x_{i_0} = 0$. Now consider any vertex $\vec{v}$ such that $v_{i_0} = 1$ and note that $\vec{v}$ and $\vec{x}$ belong to an opposite pair of faces[6] so $\chi(\vec{v}) \neq \chi(\vec{x})$ by hypothesis. And since we also have by hypothesis that $\chi(\vec{v}) = v$

[6] We have $\vec{v} \in \prod_{i=1}^d \begin{cases} [0,1] & i \neq i_0 \\ \{1\} & i = i_0 \end{cases}$ and $\vec{x} \in \prod_{i=1}^d \begin{cases} [0,1] & i \neq i_0 \\ \{0\} & i = i_0 \end{cases}$.

because $\vec{v}$ is a vertex, this gives $\chi(\vec{x}) \neq \vec{v}$. Since this is true for every vertex/color $\vec{v}$ with $v_{i_0} = 1$, we must have $\chi(\vec{x})_{i_0} = 0 \in \{0\} = F_{i_0}$ which completes the proof. □

We now state the primary result of interest from [LPS01] regarding Sperner/KKM colorings of the cube.

Theorem 6.1.5 ([LPS01] Corollary 3)**.** *In every Sperner/KKM coloring of* $[0,1]^d$ *there exists a point at the closure of at least* $d+1$ *different colors.*

The relevance of this is that any coloring which doesn't includes points on opposite faces of the cube can admit a Sperner/KKM coloring by combining certain colors together into one. That is to say that if we have an initial coloring function $\gamma : [0,1]^d \to C$ which assigns points of the cube to colors in C, then there is a mapping $f : C \to \{0,1\}^d$ of the colors in C to the standard set of colors (the vertices of the cube) so that the $\chi \stackrel{\text{def}}{=} f \circ \gamma$ is a Sperner/KKM coloring of $[0,1]^d$. Because no color $c \in C$ spans opposite faces of the cube, we can pick a vertex to associate to c so that even the union of all of the colors associated to that vertex don't span opposite faces of the cube.

Lemma 6.1.6 (Admitted Sperner/KKM Coloring)**.** *Given a coloring of* $[0,1]^d$ *(denoted by* $\gamma : [0,1]^d \to C$*) in which no color includes points of opposite faces, there is a function* $f : C \to \{0,1\}^d$ *such that* $\chi \stackrel{\text{def}}{=} f \circ \gamma$ *is a Sperner/KKM coloring of* $[0,1]^d$.

Proof. That no color contains points on opposite faces formally means that for each color $c \in C$ and coordinate $i \in [d]$, the set of points given colors c (i.e. $\gamma^{-1}(c)$) does not contain a point with ith coordinate 0 and a point with ith coordinate 1 (i.e. $\pi_i(\gamma^{-1}(c)) \not\supseteq \{0,1\}$).

For each $i \in [d]$, define $f_i : C \to \{0,1\}$ by

$$f_i(c) = \begin{cases} 0 & 0 \in \pi_i(\gamma^{-1}(c)) \\ 1 & 1 \in \pi_i(\gamma^{-1}(c)) \\ 0 & \text{otherwise} \end{cases}.$$

This is well-defined because the first two cases are mutually exclusive. Then define $f : C \to \{0,1\}^d$ by $f(c) = \langle f_i(c) \rangle_{i=1}^d$, and as in the statement, define the coloring $\chi : [0,1]^d \to \{0,1\}^d$ as the composition $f \circ \gamma$.

Consider any vertex $\vec{v} \in \{0,1\}^d$ and let $c = \gamma(\vec{v})$. Then equivalently $\vec{v} \in \gamma^{-1}(c)$ which implies by basic projection facts that $v_i \in \pi_i(\gamma^{-1}(c))$ which means $f_i(c) = v_i$ so $f(c) = \vec{v}$ so $\chi(\vec{v}) = f(\gamma(\vec{v})) = f(c) = \vec{v}$ which shows that χ has the first property of a Sperner/KKM coloring in Definition 6.1.4 (Equivalent Definition of Sperner/KKM Coloring).

For the second property, (that no color of χ contains points on opposite faces of the cube) as discussed above, we must show for each color $\vec{v} \in \{0,1\}^d$ and coordinate $i \in [d]$ that the set $\pi_i(\chi^{-1}(\vec{v}))$ either does not contain 0 or does not contain 1. In particular, we show that $\pi_i(\chi^{-1}(\vec{v}))$ does not contain the opposite value of v_i (i.e. doesn't contain 1 if $v_i = 0$ and doesn't contain 0 if $v_i = 1$).

Let $\vec{x} \in \chi^{-1}(\vec{v})$ be arbitrary and let $c = \gamma(\vec{x})$. This means $\vec{v} = \chi(\vec{x}) = f(\gamma(\vec{x})) = f(c)$ so $v_i = f_i(c)$. By definition of f_i this means that $\pi_i(\gamma^{-1}(c))$ does not contain the opposite value of v_i because if it did then $f_i(c)$ would be the opposite value of v_i. Since $x_i \in \pi_i(\gamma^{-1}(c))$ by definition of c, this means that x_i is not the opposite value of v_i. Since $\vec{x} \in \chi^{-1}(\vec{v})$ was arbitrary, this shows that $\pi_i(\chi^{-1}(\vec{v}))$ does not contain the opposite value of v_i as claimed.

Thus χ is a Sperner/KKM coloring by Definition 6.1.4 (Equivalent Definition of

Sperner/KKM Coloring).

□

With this result we can prove Theorem 6.1.1 and Theorem 6.1.2.

Proof of Theorem 6.1.1 and Theorem 6.1.2. Given a coloring $\gamma : [0,1]^d \to C$ for which no color contains points of opposite faces, let $f : C \to \{0,1\}^d$ and $\chi : [0,1]^d \to \{0,1\}^d$ with $\chi = f \circ \gamma$ as in Lemma 6.1.6. By Theorem 6.1.5, there is a point $\vec{p}$ at the closure of at least $d+1$ colors of χ. That is, the set $V' = \left\{\vec{v} \in V : \vec{p} \in \overline{\chi^{-1}(\vec{v})}\right\}$ has cardinality at least $d+1$. We will try to transfer this property to the coloring γ.

Note that for each $\vec{v} \in V$, we have

$$\chi^{-1}(\vec{v}) = (f \circ \gamma)^{-1}(\vec{v}) = \gamma^{-1}(f^{-1}(\vec{v})) = \bigcup_{c \in f^{-1}(\vec{v})} \gamma^{-1}(c). \tag{6.1}$$

Now, for each $\vec{v} \in V'$, because $\vec{p}$ is in the closure of $\chi^{-1}(\vec{v})$, any open set containing $\vec{p}$ intersects $\chi^{-1}(\vec{v}) = \bigcup_{c \in f^{-1}(\vec{v})} \gamma^{-1}(c)$ and thus intersects $\gamma^{-1}(c)$ for some $c \in f^{-1}(\vec{v})$. Let $g(\vec{v})$ denote one such color.

Because $f^{-1}(\vec{v})$ and $f^{-1}(\vec{v}\,')$ are trivially disjoint for $\vec{v} \neq \vec{v}\,'$, this means $g(\vec{v})$ and $g(\vec{v}\,')$ are distinct colors so $g : V' \to C$ is an injection which means there are at least $d+1$ colors in C that are intersected by any open set containing $\vec{p}$ which proves Theorem 6.1.2.

If also $|C|$ is finite, then for each $\vec{v} \in V$, $f^{-1}(\vec{v}) \subseteq C$ is finite, then we can use the fact the closure of a finite union is equal to the finite union of the closures to extend

Equation B.1 to

$$\begin{aligned}\vec{p} &\in \bigcap_{\vec{v}\in V'} \overline{\chi^{-1}(\vec{v})}\\ &= \bigcap_{\vec{v}\in V'} \overline{\bigcup_{c\in f^{-1}(\vec{v})} \gamma^{-1}(c)}\\ &= \bigcap_{\vec{v}\in V'} \bigcup_{c\in f^{-1}(\vec{v})} \overline{\gamma^{-1}(c)} && (f^{-1}(\vec{v}) \text{ is finite})\end{aligned}$$

and thus, for each $\vec{v} \in V'$, $\vec{p}$ belongs to the closure of $\gamma^{-1}(c)$ for some $c \in f^{-1}(\vec{v})$. By the same argument there are at least $d+1$ such colors in C. □

6.2 The Optimality Theorem Statements

We will now formally state all three optimality theorems (without proof) so that they are easily compared with each other and then we will offer some remarks and corollaries. For clear separation, we will give the proofs in the next section. We emphasize again that the Stronger Optimality Theorem and Strongest Optimality Theorem will follow very easily from the known results in Section 6.1 (Prerequisite Results: Lebesgue, Sperner, and KKM), and that it is the Standard Optimality Theorem that will require real work.

Theorem 6.2.1 (Standard Optimality Theorem)**.** *If $d \in \mathbb{N}$, and $\mathcal{P}$ is a partition of $\mathbb{R}^d$, and $\|\cdot\|$ is a norm on $\mathbb{R}^d$, and there exists $M \in (0,\infty)$ such that for all $X \in \mathcal{P}$, $m_{out}(X) < M$, then for any $\varepsilon \in (0,\infty)$, there exists $\vec{p} \in \mathbb{R}^d$ such that*

$$\left| {}^{\|\cdot\|}\overline{\mathcal{N}}_{\varepsilon}(\vec{p}) \right| \geq d+1.$$

Theorem 6.2.2 (Stronger Optimality Theorem)**.** *If $d \in \mathbb{N}$, and $\mathcal{P}$ is a partition of $\mathbb{R}^d$, and $\|\cdot\|$ is a norm on $\mathbb{R}^d$ and there exists $D \in (0,\infty)$ such that for all $X \in \mathcal{P}$, it holds that $\operatorname{diam}_{\|\cdot\|}(X) < D$, then there exists $\vec{p} \in \mathbb{R}^d$ such that for all $\varepsilon \in (0,\infty)$,*

$$\left|{}^{\|\cdot\|}\overline{\mathcal{N}}_{\varepsilon}(\vec{p})\right| \geq d+1.$$

Theorem 6.2.3 (Strongest 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*[a] *for X with respect to the ℓ_∞ norm, 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*

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

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

[a] Recall that it is sufficient but not necessary that $\operatorname{diam}(X) < D$.

Remark 6.2.4. We could have equivalently replaced the "${}^{\|\cdot\|}\overline{\mathcal{N}}_{\cdot}(\cdot)$","$\|\cdot\|$" with "${}^{\infty}\overline{\mathcal{N}}_{\cdot}(\cdot)$","$\|\cdot\|_\infty$" in the statements of Standard Optimality Theorem and Stronger Optimality Theorem above. Obviously, the stated versions apply in particular to the ℓ_∞ norm which gives one direction of the equivalence. The other direction is because for any $\varepsilon \in (0,\infty)$ there exists by Theorem 3.5.1 (Equivalence of Norms on $\mathbb{R}^d$) some $\varepsilon' \in (0,\infty)$ such that ${}^{\infty}\overline{B}_{\varepsilon'}(\vec{0}) \subseteq {}^{\|\cdot\|}\overline{B}_{\varepsilon}(\vec{0})$ and thus, for every $\vec{p} \in \mathbb{R}^d$, ${}^{\infty}\overline{B}_{\varepsilon'}(\vec{p}) \subseteq {}^{\|\cdot\|}\overline{B}_{\varepsilon}(\vec{p})$ and consequently for a partition $\mathcal{P}$ we have for every $\vec{p} \in \mathbb{R}^d$ that ${}^{\infty}\overline{\mathcal{N}}_{\varepsilon'}(\vec{p}) \subseteq {}^{\|\cdot\|}\overline{\mathcal{N}}_{\varepsilon}(\vec{p})$ (for Stronger Optimality Theorem this replacement of ε with ε' can be done for each point individually). Similarly, regarding the hypotheses of Stronger Optimality Theorem, if there is some value D for $\mathcal{P}$ such that for all

$X \in \mathcal{P}$, $\mathrm{diam}_\infty(X) < D$, then by Theorem 3.5.1 (Equivalence of Norms on $\mathbb{R}^d$), there is some value D' for $\mathcal{P}$ such that for all $X \in \mathcal{P}$, $\mathrm{diam}_{\|\cdot\|}(X) < D'$.

There are almost certainly variations of Strongest Optimality Theorem that could be stated for other norms, but this is quite a specific result as it is, and we will mostly be interested in Corollary 6.2.6 which follows from it and will be stated shortly. △

The conclusion of the Standard Optimality Theorem is that for partitions with a uniform upper bound on outer measure, $k = d + 1$ is optimal in the "standard optimality" sense (based on Remark 6.2.4).

The conclusion of the Stronger Optimality Theorem is that for partitions with a uniform upper bound on the diameter[7], $k = d + 1$ is optimal in the "stronger optimality" sense (based on Remark 6.2.4). Thus, we get a stronger sense in which $k = d+1$ is optimal than the Standard Optimality Theorem, but we also strengthened the hypotheses.

The conclusion of the Strongest Optimality Theorem is that for partitions with a certain local finiteness condition, $k = d + 1$ is optimal in the "strongest optimality" sense. Thus, we get still a stronger sense in which $k = d + 1$ is optimal than either the Standard Optimality Theorem or the Stronger Optimality Theorem, but we also strengthened the hypotheses yet again[8].

Remark 6.2.5. Observe that in the Standard Optimality Theorem (resp. Stronger 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

---

[7] It is important that the diameter uses a norm and not an arbitrary metric so that we can translate sets without affecting their diameter, but by Remark 6.2.4, it doesn't matter which norm is used.

[8] This is a stronger hypothesis than the Stronger Optimality Theorem (Theorem 6.2.2) because if D is a strict pairwise bound for each member with respect to $\|\cdot\|_\infty$, then each member has diameter at most D with respect to $\|\cdot\|_\infty$ and thus by either Theorem 3.5.1 (Equivalence of Norms on $\mathbb{R}^d$) or Remark 6.2.4, for any norm $\|\cdot\|$ there is some D' such that each member has diameter at most D' with respect to $\|\cdot\|$.

actual value of M (resp. D) is not used in the conclusion. However, in the Strongest Optimality Theorem the phrase "D is a strict pairwise bound for X" cannot be replaced by the phrase "$\operatorname{diam}_\infty(X) \leq D$". Consider for example if $\mathcal{P} = \left\{[0,1]^d\right\} \cup \left\{\{\vec{x}\} : \vec{x} \notin [0,1]^d\right\}$ (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 $X \in \mathcal{P}$, $\operatorname{diam}_\infty(X) \leq 1$ 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 $\vec{p} \in \mathbb{R}^d$, we have $\mathcal{N}_{\vec{0}}(\vec{p}) = \left\{X \in \mathcal{P} : \overline{X} \ni \vec{p}\right\} = \{X \in \mathcal{P} : X \ni \vec{p}\}$ which has cardinality $1 < d+1$ because $\mathcal{P}$ is a partition, so $\vec{p}$ belongs to only one member. Thus, the use of the strict pairwise bound in the statement will be necessary. It could of course be replaced by the strictly stronger hypothesis that "$\operatorname{diam}_\infty(X) < D$", but that would be a weaker result. △

The last hypothesis in the Strongest 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 many sets. This finiteness condition arises in the natural context below.

Corollary 6.2.6 (Strongest Optimality Corollary)**.** *Let $d \in \mathbb{N}$ and $\mathcal{P}$ a partition of $\mathbb{R}^d$ and $\|\cdot\|$ a norm on $\mathbb{R}^d$. If there exists $D \in (0,\infty)$ and $\mu \in (0,\infty)$ such that for all $X \in \mathcal{P}$, it holds that $\mu < m_{in}(X)$, and $\operatorname{diam}_{\|\cdot\|}(X) < D$, 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. It suffices to prove that the hypotheses of the corollary imply the hypothesis of the Strongest Optimality Theorem (Theorem 6.2.3). By Theorem 3.5.1 (Equivalence of Norms on $\mathbb{R}^d$) there exists some D' such that for every $X \in \mathcal{P}$, $\operatorname{diam}_\infty(X) < D'$. 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 is the ℓ_∞ ball containing B and all points within distance D' of B). Thus, for any $X \in \mathcal{B}$, because $\operatorname{diam}_\infty(X) < D'$, we have $X \subseteq B'$, and thus $\bigsqcup_{X \in \mathcal{B}} X \subseteq B'$. Thus, because each member in $\mathcal{B}$ has Lebesgue inner measure at least $\mu > 0$ we have the following volume argument (see Fact A.0.1 for the middle inequality):

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

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

6.3 The Optimality Theorem Proofs

We will now prove the three optimality theorems (with restatements included for convenience). As already mentioned multiple times, the Stronger Optimality Theorem and Strongest Optimality Theorem will follow very easily from the known results in Section 6.1 (Prerequisite Results: Lebesgue, Sperner, and KKM), and it is the Standard Optimality Theorem that will require effort.

Theorem 6.2.3 (Strongest 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[a] for X with respect to the ℓ_∞ norm, 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*

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

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

[a]Recall that it is sufficient but not necessary that $\operatorname{diam}(X) < D$.

Proof. Let $\mathcal{P}'$ be the finite set of members of $\mathcal{P}$ which intersect $\vec{\alpha} + [0,D]^d$. We define a coloring $\gamma : \vec{\alpha} + [0,D]^d \to \mathcal{P}'$ by mapping each point to the member in $\mathcal{P}'$ it is contained in. Points on opposing faces cannot be given the same color because no member includes points on opposing faces. By Theorem 6.1.1, there is a point $\vec{p}$ at the closure of at least $d+1$ colors. That is, there is a subset $\mathcal{P}'' \subseteq \mathcal{P}'$ of cardinality at least $d+1$ such that for each $X \in \mathcal{P}''$, $\vec{p} \in \overline{\gamma^{-1}(X)} \subseteq \overline{X}$ (see footnote[9]). Thus $|\mathcal{N}_{\overline{0}}(\vec{p})| \geq d+1$ and the set $\mathcal{N}_{\overline{0}}(\vec{p})$ is a clique because any pair of members are adjacent since they both contain $\vec{p}$ in their closure. □

Theorem 6.2.2 (Stronger Optimality Theorem)**.** *If $d \in \mathbb{N}$, and $\mathcal{P}$ is a partition of $\mathbb{R}^d$, and $\|\cdot\|$ is a norm on $\mathbb{R}^d$ and there exists $D \in (0,\infty)$ such that for all $X \in \mathcal{P}$, it holds that $\operatorname{diam}_{\|\cdot\|}(X) < D$, then there exists $\vec{p} \in \mathbb{R}^d$ such that for all $\varepsilon \in (0,\infty)$,*

$$\left| {}^{\|\cdot\|}\overline{\mathcal{N}}_{\varepsilon}(\vec{p}) \right| \geq d+1.$$

[9]Each $\vec{x} \in \gamma^{-1}(X)$ has $\gamma(\vec{x}) = X$ and thus by definition of γ, $\vec{x} \in X$. This is containment and not equality because $\gamma^{-1}(X)$ is by definition a subset of $\vec{\alpha} + [0,D]^d$ and $X \in \mathcal{P}' \subseteq \mathcal{P}$ might not be.

Proof. By Theorem 3.5.1 (Equivalence of Norms on $\mathbb{R}^d$) there exists some D' such that for all $X \in \mathcal{P}$, $\operatorname{diam}_\infty(X) < D'$. We define a coloring $\gamma : [0, D']^d \to \mathcal{P}$ by mapping each point to the member in $\mathcal{P}$ it is contained in. Points on opposing faces cannot be given the same color because no member of $\mathcal{P}$ includes points on opposing faces of $[0, D']^d$. By Theorem 6.1.2, there is a point $\vec{p}$ such that for every open set (in particular every open ε-ball with respect to $\|\cdot\|$) contains points of at least $d+1$ colors of γ. That is, for every $\varepsilon \in (0, \infty)$, there are at least $d+1$ colors $X^{(1)}, \ldots, X^{(d+1)} \in \operatorname{range}(\gamma) \subseteq \mathcal{P}$ (possibly depending on ε) such that ${}^{\infty}B^{\circ}_{\varepsilon}(\vec{p}) \cap \gamma^{-1}(X^{(j)}) \neq \emptyset$ for each $j \in [d+1]$. Because $\gamma^{-1}(X) \subseteq X$ (by definition of γ) we have that ${}^{\infty}B^{\circ}_{\varepsilon}(\vec{p}) \cap X^{(j)} \neq \emptyset$ for each $j \in [d+1]$ (and thus this trivially holds replacing the open ball with a closed one). That is, $\left|{}^{\infty}\overline{\mathcal{N}}_{\varepsilon}(\vec{p})\right| \geq d+1$ for each $\varepsilon \in (0, \infty)$. $\square$

Now we arrive at proving the main result of the chapter. Once we have done the initial setup within the proof, we will give an extensive outline of the main ideas of the proof.

Theorem 6.2.1 (Standard Optimality Theorem)**.** *If $d \in \mathbb{N}$, and $\mathcal{P}$ is a partition of $\mathbb{R}^d$, and $\|\cdot\|$ is a norm on $\mathbb{R}^d$, and there exists $M \in (0, \infty)$ such that for all $X \in \mathcal{P}$, $m_{out}(X) < M$, then for any $\varepsilon \in (0, \infty)$, there exists $\vec{p} \in \mathbb{R}^d$ such that*

$$\left|{}^{\|\cdot\|}\overline{\mathcal{N}}_{\varepsilon}(\vec{p})\right| \geq d+1.$$

Proof. Let $k = d+1$. If there exists $\vec{p} \in \mathbb{R}^d$ such that $|\mathcal{N}_{\varepsilon}(\vec{p})| \geq k$ then we are done. Otherwise, we may assume that for all $\vec{p} \in \mathbb{R}^d$ that $|\mathcal{N}_{\varepsilon}(\vec{p})| < k$. We remark that this is a strange way to start a proof, because we could have instead started the proof saying "Suppose for contradiction that for all $\vec{p} \in \mathbb{R}^d$ that $|\mathcal{N}_{\varepsilon}(\vec{p})| < d+1$" and then proceeded to a contradiction. However, morally, this is not a proof by contradiction, and the value k could be replaced by any other constant. All that is important for the proof to proceed is that there is some constant upper bound on the size of all of these neighborhoods which is why we deal immediately with the trivial case where the partition has a point with neighborhood of cardinality k or more.

Throughout the proof, for $\vec{x} \in \mathbb{R}^d$, let $H(\vec{x}) = \vec{x} + [-\frac{\varepsilon}{2}, \frac{\varepsilon}{2})^d$ (standing for half open hypercube) and note that $\overline{H(\vec{x})} = {}^{\infty}\overline{B}_{\frac{\varepsilon}{2}}(\vec{x})$. Let $V = \varepsilon \cdot \mathbb{Z}^d$ which should be viewed as the set of vertices defining a grid. Now observe that $\mathcal{G} = \{H(\vec{\alpha}) : \vec{\alpha} \in V\}$ is a partition of $\mathbb{R}^d$ (with $\mathcal{G}$ standing for grid). We view this grid as breaking up $\mathbb{R}^d$ into cells or pixels.

The main idea of the proof is easily lost in the details, so we first provide a high level overview.

Proof Outline.

1. Because of the assumption that $|\mathcal{N}_\varepsilon(\vec{p})| \leq k$, it will follow that for each $\vec{\alpha} \in V$, $H(\vec{\alpha})$ intersects at most k members of $\mathcal{P}$. That is, each cell of $\mathcal{G}$ intersects at most k members of $\mathcal{P}$. Thus, for each cell in $\mathcal{G}$, by averaging, there is some member of $\mathcal{P}$ (possibly multiple, but then we just pick one) which contains at least a $\frac{1}{k}$ fraction of the volume of the cell (i.e. a volume of $\frac{\varepsilon^d}{k}$), and we associate the cell with one such member of $\mathcal{P}$.
2. We interpret this process as "pixelating" the partition $\mathcal{P}$. That is, most information about $\mathcal{P}$ is disregarded except that in each cell/pixel we have an associated member of $\mathcal{P}$ which captures a significant portion of the volume of that cell/pixel. For each member $X \in \mathcal{P}$, we consider the union of all cells associated to X as a pixelated approximation to X. (Formally, this happens implicitly through the binary equivalence relation R_P defined later, though we never explicitly construct this pixelated partition.) We refer to these in this outline discussion as the primary pixelated member approximations.

3. Each primary pixelated member (which is a union of cells/pixels) will consist of at most $\frac{kM}{\varepsilon^d}$ cells/pixels because each cell/pixel associated to the member $X \in \mathcal{P}$ contains at least $\frac{\varepsilon^d}{k}$ (outer) volume of X, and since X has (outer) volume at most M, there are at most $\frac{kM}{\varepsilon^d}$ cells/pixels associated to X.
4. Because each primary pixelated member is a union of at most a constant number of pixels, each primary pixelated member has a finite diameter. However, the diameters of the primary pixelated members could be arbitrarily large, yet our goal will be to apply the Strongest Optimality Theorem (Theorem 6.2.3) to the pixelated partition. To deal with this, we next break up each primary pixelated member as a disjoint union of its connected components (formally, this can either be seen as the connected components of a certain graph or as the transitive closure of a certain relation $R_\cap$ defined later). The resulting partition consists of connected pixelated members, and we refer to these members as secondary pixelated approximations in this discussion. This partition is explicitly constructed as $\mathcal{A}$ in the proof.
5. We get an immediate diameter bound on the secondary pixelated members—each member has at most $\frac{kM}{\varepsilon^d}$ cells/pixels and is connected so cannot be more than $\frac{kM}{\varepsilon^d}$ cells/pixels in length in any direction.
6. This universal diameter bound on the secondary pixelated approximation along with the fact that a bounded set can trivially only intersect finitely many cell/pixels (and thus intersect only finitely many members of the secondary pixelated partition) allows us to apply the Strongest Optimality Theorem to find a point $\vec{p}$ at the closure of $d+1$ secondary pixelated members.
7. In the primary pixelated partition, the function mapping a pixelated member to the member of $\mathcal{P}$ it represents was an injection (but not necessarily a bijection). With the secondary pixelated partition, however, this function

might not be an injection anymore—there might be two different secondary pixelated members that are both associated to the same primary pixelated member and thus to the same member of $\mathcal{P}$ because the primary pixelated member was broken into multiple connected pieces. Thus, even though a point has been found at the closure of $d+1$ secondary pixelated members, it is not immediate that this point is at the closure of $d + 1$ primary pixelated members. Though not immediate, it is in fact true. We show so by demonstrating that no two of these $(d+1)$-many secondary pixelated members are part of the same primary pixelated member (because the secondary members meet at a single point, so if they were part of the same primary pixelated member to begin with, they would not have been broken into different connected components). Said another way, for any primary pixelated member which was broken up into multiple secondary pixelated members, the connected pieces it was broken up into are each separated by at least one cell/pixel because that is the only way to be disconnected. It follows that this point $\vec{p}$ is in fact at the closure of $d+1$-many primary pixelated members.

8. Finally, because this point is at the closure of $d+1$ primary pixelated members and each pixel has diameter ε, the ε-radius ball around this point will fully contain at least one cell/pixel of each of these primary pixelated members, and thus intersect each member of $\mathcal{P}$ that the primary pixelated members were associated to. Thus, this ε-ball at $\vec{p}$ intersects at least $d + 1$ members of $\mathcal{P}$ which is what we wanted to show.

■

We now continue formally with the proof. We want to argue that each cell ($H(\vec{\alpha})$ for each $\alpha \in V$) intersects some member of the partition with high outer volume/measure.

Claim A. *There exists some function $P : V \to \mathcal{P}$ such that for all $\vec{\alpha} \in V$, $m_{out}(P(\vec{\alpha}) \cap H(\vec{\alpha})) \geq \frac{\varepsilon^d}{k}$.*

Proof of Claim. For any $\vec{\alpha} \in V$ we have by assumption that $\left|{}^{\infty}\overline{\mathcal{N}}_{\varepsilon}(\vec{\alpha})\right| \leq k$, so (by definition) at most k members of $\mathcal{P}$ intersect ${}^{\infty}\overline{B}_{\varepsilon}(\vec{\alpha})$, and thus at most k members of $\mathcal{P}$ intersect the subset $H(\vec{\alpha})$. We index over these members to obtain the following:

$$\sum_{X \in \mathcal{P} : X \cap H(\vec{\alpha}) \neq \emptyset} m_{out}(X \cap H(\vec{\alpha})) \geq m_{out}\left(\bigsqcup_{X \in \mathcal{P} : X \cap H(\vec{\alpha}) \neq \emptyset} X \cap H(\vec{\alpha})\right)$$

(Countable/finite subaddativity of outer measures)

$$= m_{out}\left(H(\vec{\alpha})\right) \qquad \text{(Re-express union)}$$

$$= m\left(H(\vec{\alpha})\right) \qquad (H(\vec{\alpha}) \text{ is measurable})$$

$$= \varepsilon^d$$

Because the index set in the summation above has cardinality at most k, by the trivial averaging argument, there is some $X \in \mathcal{P}$ such that $m_{out}(X \cap H(\vec{\alpha})) \geq \frac{\varepsilon^d}{k}$. Thus, there exists some function $P : V \to \mathcal{P}$ such that for all $\vec{\alpha} \in V$, $m_{out}(P(\vec{\alpha}) \cap H(\vec{\alpha})) \geq \frac{\varepsilon^d}{k}$. ■

We view the function P above (which will be fixed for the remainder of the proof) as indicating, for each vertex/cell, a member of $\mathcal{P}$ that is sufficiently similar to the content of the cell. The name P is meant to indicate similarity to $\mathcal{P}$. This function will be the key to approximating $\mathcal{P}$ with cells.

Next, we provide a bound on how many vertices/cells can be mapped to a particular member of $\mathcal{P}$.

Claim B. *For each $X \in \mathcal{P}$ we have $|P^{-1}(X)| < \frac{kM}{\varepsilon^d}$ (i.e. there are fewer than $\frac{kM}{\varepsilon^d}$-many cells being associated/mapped to X).*

Proof of Claim. Let $X \in \mathcal{P}$ be arbitrary, and let $X' \supseteq X$ be such that X' is Lebesgue measurable and $m(X') < M$ (this is possible by definition of the Lebesgue outer measure and the fact that $m_{out}(X) < M$). Then noting that $\bigsqcup_{\vec{\alpha} \in P^{-1}(X)} (X' \cap H(\vec{\alpha}))$ is (trivially) a subset of X' and each term of the union is measurable[10], we have the following (where m_{in} denotes the Lebesgue inner measure):

$$
\begin{aligned}
M &> m(X') && \text{(By hypothesis)} \\
&= m_{in}(X') && (X' \text{ is measurable}) \\
&\geq m_{in}\left(\bigsqcup_{\vec{\alpha} \in P^{-1}(X)} X' \cap H(\vec{\alpha}) \right) && \text{(Subset; monotonicity of inner measures)} \\
&\geq \sum_{\vec{\alpha} \in P^{-1}(X)} m(X' \cap H(\vec{\alpha})) && \text{(Fact A.0.1; Emphasis: this is the correct direction)} \\
&\geq \sum_{\vec{\alpha} \in P^{-1}(X)} m_{out}(X \cap H(\vec{\alpha})) && (X \subseteq X') \\
&= \sum_{\vec{\alpha} \in P^{-1}(X)} m_{out}(P(\vec{\alpha}) \cap H(\vec{\alpha})) && (\text{For } \vec{\alpha} \in P^{-1}(X),\ P(\vec{\alpha}) = X) \\
&\geq \sum_{\vec{\alpha} \in P^{-1}(X)} \frac{\varepsilon^d}{k} && \text{(Def'n of } P\text{)} \\
&= \left| P^{-1}(X) \right| \frac{\varepsilon^d}{k}
\end{aligned}
$$

This shows that $|P^{-1}(X)| < \frac{kM}{\varepsilon^d}$. ■

Next, we will define three binary relations on the set V of vertices in order to arrive at a useful equivalence relation that will let us approximate $\mathcal{P}$ closely enough

[10] We cannot yet assume measurability of the entire union, though. For example, we don't yet know if $P^{-1}(X)$ is countable.

by using the cells (see Appendix E for details on binary relations.). We ultimately 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 label according to the function P and are also connected. With this goal, let

$$\begin{aligned} R_P &= \left\{(\vec{\alpha}, \vec{\beta}) \in V^2 : P(\vec{\alpha}) = P(\vec{\beta})\right\} \\ R_\varepsilon &= \left\{(\vec{\alpha}, \vec{\beta}) \in V^2 : \left\|\vec{\alpha} - \vec{\beta}\right\|_\infty \leq \varepsilon\right\} \\ R_\cap &= R_P \cap R_\varepsilon \\ R &= R_\cap^t && \text{(The transitive closure of } R_\cap\text{)} \end{aligned}$$

Note that both R_P and R_ε are reflexive and symmetric; R_P is also transitive, so R_P is an equivalence relation. It follows trivially that $R_\cap$ is reflexive and symmetric, and so by Fact E.0.2, R is an equivalence relation. Furthermore, because R_P is a transitive superset of $R_\cap$, it must be that the transitive closure of $R_\cap$ is a subset of R_P—that is $R \subseteq R_P$. A graph theory perspective on the relations above may be helpful[11].

We now show a diameter bound on the equivalence classes in R.

Claim C. *For any* $(\vec{\alpha}, \vec{\beta}) \in R$ *we have* $\left\|\vec{\alpha} - \vec{\beta}\right\|_\infty < \left(\frac{kM}{\varepsilon^d} - 1\right)\varepsilon$.

Proof of Claim. Let $(\vec{\alpha}, \vec{\beta}) \in R$ be arbitrary. By Fact E.0.7 because $\mathbb{R}$ is the transitive closure of $R_\cap$, there is a finite sequence $\langle \vec{x}^{(i)} \rangle_{i=0}^N$ (for some N) of distinct elements of

[11]One can view R_P and R_ε as specifying edges of two infinite graphs on vertex set V. Then $R_\cap$ specifies the edges in the intersection of these two graphs, and the equivalence classes in R are then the connected components of this infinite intersection graph. R_P ensures that each equivalence class in R consists of vertices representing the same member of $\mathcal{P}$, and R_ε ensures that each equivalence class in R is geometrically "connected" in the sense that the union of corresponding cells is a (topologically) connected set; the topology though is not really the key, from the pixelation view we just mean that an equivalence class consists of a collection of pixels which meet each other and are associated by P to the same member of $\mathcal{P}$. The connectedness will be key to ensuring that the diameters are bounded. Rigorous details are handled in the proof through the perspective of the binary relations rather than graphs, though these are quite fundamentally the same structure.

V with $\vec{x}^{(0)} = \vec{\alpha}$ and $\vec{x}^{(N)} = \vec{\beta}$ and for all $i \in [N]$, $(\vec{x}^{(i-1)}, \vec{x}^{(i)}) \in R_\cap$ (i.e. a chain of $R_\cap$-related vertices that witness why $(\vec{\alpha}, \vec{\beta}) \in R_\cap^t = R$). Because $R_\cap \subseteq R_P$ and R_P is an equivalence relation, all $\vec{x}^{(i)}$ belong to the same equivalence class of R_P, and because all $N+1$ of the $\vec{x}^{(i)}$ are distinct, it must be that $N+1$ is bounded above by the cardinality of some equivalence class in R_P. Each equivalence class in R_P is $P^{-1}(X)$ for some $X \in \mathcal{P}$, and thus each equivalence class in R_P has cardinality less than $\frac{kM}{\varepsilon^d}$ by Claim B. Thus, we have $N+1 < \frac{kM}{\varepsilon^d}$ which lets us obtain the stated bound:

$$
\begin{aligned}
\left\|\vec{\alpha} - \vec{\beta}\right\|_\infty &\leq \sum_{i=1}^{N} \left\|\vec{x}^{(i-1)} - \vec{x}^{(i)}\right\|_\infty && \text{(Triangle inequality)} \\
&\leq \sum_{i=1}^{N} \varepsilon && ((\vec{x}^{(i-1)}, \vec{x}^{(i)}) \in R_\cap \subseteq R_\varepsilon) \\
&\leq N\varepsilon \\
&< \left(\frac{kM}{\varepsilon^d} - 1\right)\varepsilon. && \text{(Constraint on } N+1 \text{ above)}
\end{aligned}
$$

This proves the claim. ■

We are now in a position to formally define the approximation partition. Using quotient notation, let V/R denote the family of equivalence classes of R (we emphasize 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 $\mathbb{R}^d$. In words, each $Y \in \mathcal{A}$ is a union of cells of the grid $\mathcal{G}$ which belong to the same equivalence class of R (if we identify vertices and cells of $\mathcal{G}$).

Thus, for any $Y \in \mathcal{A}$, there is an unique equivalence class $C \in V/R$ such that $Y = \bigsqcup_{\vec{\alpha} \in C} H(\vec{\alpha})$, and we denote this class as C_Y. We will extend the distance bound of Claim C to the members of $\mathcal{A}$.

Claim D. *For each $Y \in \mathcal{A}$, $\frac{kM}{\varepsilon^{d-1}}$ is a strict pairwise bound on Y in the ℓ_∞ norm.*

Proof of Claim. Let $Y \in \mathcal{A}$ be arbitrary and let $\vec{a}, \vec{b} \in Y$ be arbitrary. Because $Y = \bigsqcup_{\vec{v} \in C_Y} H(\vec{v})$ Then there must exist some $\vec{\alpha} \in C_Y$ with $\vec{a} \in H(\vec{\alpha})$ and some $\vec{\beta} \in C_Y$ (possibly the same as $\vec{\alpha}$) with $\vec{b} \in H(\vec{\beta})$. Utilizing the triangle inequality, along with Claim C and the fact that $H(\vec{\alpha}) \subseteq {}^{\infty}\overline{B}_{\frac{1}{2}}(\alpha)$ (and similarly with $\vec{\beta}$) we have the following:

$$\|\vec{a} - \vec{b}\|_\infty \leq \|\vec{a} - \vec{\alpha}\|_\infty + \|\vec{\alpha} - \vec{\beta}\|_\infty + \|\vec{\beta} - \vec{b}\|_\infty < \frac{\varepsilon}{2} + \left(\frac{kM}{\varepsilon^d} - 1\right)\varepsilon + \frac{\varepsilon}{2} = \frac{kM}{\varepsilon^{d-1}}.$$

This proves the claim. ■

Thus, $\mathcal{A}$ satisfies the hypothesis of the Strongest Optimality Theorem[12], so there is some point $\vec{p} \in \mathbb{R}^d$ (which is fixed for the remainder of the proof) which is at the closure of at least $d+1$ members of $\mathcal{A}$. Let ${}_{\mathcal{A}}\mathcal{N}_{\overline{0}}(\vec{p}) = \left\{Y \in \mathcal{A} : \overline{Y} \ni \vec{p}\right\}$ denote this set.[13]

We have now found a point $\vec{p}$ at the closure of $d+1$ different members of the approximation partition $\mathcal{A}$. We have two remaining things to show: (1) the ball ${}^{\infty}\overline{B}_\varepsilon(\vec{p})$ completely contains a cell in each of these approximation members and so ${}^{\infty}\overline{B}_\varepsilon(\vec{p})$ intersects the corresponding member of the original partition $\mathcal{P}$ and (2) each member of ${}_{\mathcal{A}}\mathcal{N}_{\overline{0}}(\vec{p})$ corresponds to a distinct member of the original partition $\mathcal{P}$ so

[12] For the first requirement, $D = \frac{kM}{\varepsilon^{d-1}}$ is a strict pairwise bound on each member of the partition $\mathcal{A}$. Then note that for any $\vec{\alpha} \in V$, the set $[0, D]^d$ intersects the cell $H(\vec{\alpha})$ only if $H(\vec{\alpha}) \subseteq [0-\varepsilon, D+\varepsilon]^d$, and because each cell has the same positive measure (specifically ε^d) and $[0-\varepsilon, D+\varepsilon]^d$ has finite measure, $[0-\varepsilon, D+\varepsilon]^d$ can contain only finitely many cells (because they are all disjoint), and thus $[0, D]^d$ intersects at most finitely many cells. It follows that $[0, D]^d$ intersects finitely many members of $\mathcal{A}$ because each member of $\mathcal{A}$ is a union of cells, so for each member of $\mathcal{A}$ which $[0, D]^d$ intersects, $[0, D]^d$ must intersect one of the cells which is a subset of that member. This demonstrates the second requirement for applying the Strongest Optimality Theorem.

[13] This is the usual definition of $\mathcal{N}_{\overline{0}}(\vec{p})$, but we have included a pre-subscript to distinguish that the members are coming from the partition $\mathcal{A}$ and not the partition $\mathcal{P}$.

that ${}^{\infty}\overline{B}_{\varepsilon}(\vec{p})$ intersects at least $|{}_{\mathcal{A}}\mathcal{N}_{\vec{0}}(\vec{p})| \geq d+1$ members of $\mathcal{P}$. We begin with the second of these two by defining a map and proving it is injective.

Observe that for any $Y \in {}_{\mathcal{A}}\mathcal{N}_{\vec{0}}(\vec{p})$ we have

$$\vec{p} \in \overline{Y} = \overline{\bigcup_{\vec{\alpha} \in C_Y} H(\vec{\alpha})} = \bigcup_{\vec{\alpha} \in C_Y} \overline{H(\vec{\alpha})}$$

where the second equality follows from Corollary 3.6.6 (Unit Cube Packings are Locally Finite) and Fact 3.6.4 (Locally Finite: Closure of Union = Union of Closures)[14]. Thus, there must be some vertex $\vec{\alpha}$ in the equivalence class C_Y such that $\vec{p} \in \overline{H(\vec{\alpha})}$, and we denote one such fixed vertex for each Y by $\vec{\alpha}_Y$ (we view this as a function).

Claim E. *The map ${}_{\mathcal{A}}\mathcal{N}_{\vec{0}}(\vec{p}) \to \mathcal{P}$ defined by $Y \mapsto P(\vec{\alpha}_Y)$ is an injection.*

Proof of Claim. Consider distinct members $Y, Y' \in {}_{\mathcal{A}}\mathcal{N}_{\vec{0}}(\vec{p})$, and the representing vertices $\vec{\alpha}_Y$ and $\vec{\alpha}_{Y'}$. Note that because Y and Y' are distinct we have $C_Y \neq C_{Y'}$ (otherwise $Y = \bigcup_{\vec{\alpha} \in C_Y} H(\vec{\alpha}) = \bigcup_{\vec{\alpha} \in C_{Y'}} H(\vec{\alpha}) = Y'$).

Thus, we have $\vec{\alpha}_Y \in C_Y \neq C_{Y'} \ni \vec{\alpha}_{Y'}$ which says that the vertices $\vec{\alpha}_Y$ and $\vec{\alpha}_{Y'}$ belong to distinct equivalence classes of the relation R; this can be expressed as $(\vec{\alpha}_Y, \vec{\alpha}_{Y'}) \notin R$. However, we do have $(\vec{\alpha}_Y, \vec{\alpha}_{Y'}) \in R_\varepsilon$ as follows: by definition of $\vec{\alpha}_Y$ and $\vec{\alpha}_{Y'}$, we have $\vec{p} \in \overline{H(\vec{\alpha}_Y)} = {}^{\infty}\overline{B}_{\frac{\varepsilon}{2}}(\vec{\alpha}_Y)$ and $\vec{p} \in \overline{H(\vec{\alpha}_{Y'})} = {}^{\infty}\overline{B}_{\frac{\varepsilon}{2}}(\vec{\alpha}_{Y'})$, so

$$\|\vec{\alpha}_Y - \vec{\alpha}_{Y'}\|_\infty \leq \|\vec{\alpha}_Y - \vec{p}\|_\infty + \|\vec{p} - \vec{\alpha}_{Y'}\|_\infty \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

which shows by definition of R_ε that $(\vec{\alpha}_Y, \vec{\alpha}_{Y'}) \in R_\varepsilon$. Thus, we have $(\vec{\alpha}_Y, \vec{\alpha}_{Y'}) \notin R \supseteq R_\cap = R_\varepsilon \cap R_P$ but $(\vec{\alpha}_Y, \vec{\alpha}_{Y'}) \in R_\varepsilon$ implying that $(\vec{\alpha}_Y, \vec{\alpha}_{Y'}) \notin R_P$, so by definition of

[14] Alternatively, the second equality could be justified because C_Y is an equivalence class of $R \subseteq R_P$ and thus has finite cardinality by Claim B.

of R_P, we have $P(\vec{\alpha}_Y) \neq P(\vec{\alpha}_{Y'})$. ■

As we did notationally with $\mathcal{A}$ earlier, let ${}_{\mathcal{P}}^{\infty}\mathcal{N}_{\varepsilon}(\vec{p})$ be the ε neighborhood for $\mathcal{P}$: ${}_{\mathcal{P}}^{\infty}\mathcal{N}_{\varepsilon}(\vec{p}) = \left\{X \in \mathcal{P} : {}^{\infty}\overline{B}_{\varepsilon}(\vec{p}) \cap X \neq \emptyset\right\}$. We must show that this set has cardinality at least $d+1$.

Claim F. *The range of the map* ${}_{\mathcal{A}}\mathcal{N}_{\overline{0}}(\vec{p}) \to \mathcal{P}$ *defined by* $Y \mapsto P(\vec{\alpha}_Y)$ *is a subset of* ${}_{\mathcal{P}}^{\infty}\mathcal{N}_{\varepsilon}(\vec{p})$*.*

Proof of Claim. Let $Y \in {}_{\mathcal{A}}\mathcal{N}_{\overline{0}}(\vec{p})$. Then by definition of $\vec{\alpha}_Y$, we have $\vec{p} \in \overline{H(\vec{\alpha}_Y)} = {}^{\infty}\overline{B}_{\frac{\varepsilon}{2}}(\vec{\alpha}_Y)$ showing that $\|\vec{\alpha}_Y - \vec{p}\|_{\infty} \leq \frac{\varepsilon}{2}$. Thus, it follows that ${}^{\infty}\overline{B}_{\frac{\varepsilon}{2}}(\vec{\alpha}_Y) \subseteq {}^{\infty}\overline{B}_{\varepsilon}(\vec{p})$.

Combining the containments, we have $H(\vec{\alpha}_Y) \subseteq {}^{\infty}\overline{B}_{\varepsilon}(\vec{p})$ (i.e. the ε-ball at $\vec{p}$ completely contains this cell). By the definition of P, we have that $P(\vec{\alpha}_Y)$ is a member of $\mathcal{P}$ such that $m_{out}(P(\vec{\alpha}_Y) \cap H(\vec{\alpha}_Y)) \geq \frac{\varepsilon^d}{k}$, so in particular, this intersection is not empty. Then replacing with the superset, we have that $P(\vec{\alpha}_Y) \cap {}^{\infty}\overline{B}_{\varepsilon}(\vec{p})$ is also not empty showing that $P(\vec{\alpha}_Y) \in {}_{\mathcal{P}}^{\infty}\mathcal{N}_{\varepsilon}(\vec{p})$. ■

Together, Claim E and Claim F show that the mapping $Y \mapsto P(\vec{\alpha}_Y)$ injects ${}_{\mathcal{A}}\mathcal{N}_{\overline{0}}(\vec{p})$ into ${}_{\mathcal{P}}^{\infty}\mathcal{N}_{\varepsilon}(\vec{p})$ showing that

$$d+1 \leq \left|{}_{\mathcal{A}}\mathcal{N}_{\overline{0}}(\vec{p})\right| \leq \left|{}_{\mathcal{P}}^{\infty}\mathcal{N}_{\varepsilon}(\vec{p})\right|$$

which proves the theorem. □

6.4 Optimality Theorem Gaps

The conclusions get stronger with each successive optimality theorem version, but each time the hypotheses were also made stronger. This begs the question of whether all three versions are really necessary; for example, could it be that the hypotheses of the Standard Optimality Theorem are mathematically sufficient to imply the conclusions of the Strongest 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—one example showing a gap between the Standard Optimality Theorem and the Stronger Optimality Theorem and the other example showing a gap between the Stronger Optimality Theorem and the Strongest Optimality Theorem.

Proposition 6.4.1 ("Standard" and "Stronger" Optimality Gap)**.** *The hypothesis of the Standard Optimality Theorem does not imply the conclusion of the Stronger 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 Standard Optimality Theorem and which does not satisfy the conclusion of the Stronger Optimality Theorem.

Recall that ${}^{\infty}\overline{B}_r(\vec{0}) = [-r, r]^d$ has diameter $2r$ and measure $(2r)^d$. Consider the sequence $\{r_n\}_{n=1}^{\infty}$ where $r_n = \frac{1}{2}n^{1/d}$ so that ${}^{\infty}\overline{B}_{r_n}(\vec{0})$ has measure n. Let $S_1 = {}^{\infty}\overline{B}_{r_1}(\vec{0})$ and inductively for $n > 1$, let $S_n = {}^{\infty}\overline{B}_{r_n}(\vec{0}) \setminus {}^{\infty}\overline{B}_{r_{n-1}}(\vec{0})$. 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 $n-(n-1) = 1$ which is 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 $\varepsilon(\vec{p})$ such that ${}^{\infty}\overline{B}_{\varepsilon(\vec{p})}(\vec{p})$ intersects at most 2 members of the partition. This shows that for $d \geq 2$, the conclusion of the Stronger Optimality Theorem does not hold[15]. For

[15]This also shows that the conclusion of the Strongest 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.

clarity, this construction works with $d = 1$ as well, but if $d = 1$, then $d + 1 = 2$, so for each $\vec{p} \in \mathbb{R}^d$, $\overline{B}_{\varepsilon(\vec{p})}(\vec{p})$ intersects at most $2 = d + 1$ members of the partition which does not contradict the conclusion of the Stronger Optimality Theorem.

In the case $d = 1$, let $\mathcal{P}$ be a partition of $\mathbb{R}$ such that for some M, it holds for all $X \in \mathcal{P}$ that $m_{out}(X) < M$, (so the hypothesis of the Standard Optimality Theorem is satisfied). Let $X \in \mathcal{P}$ be an arbitrary member. X is not empty because the partition does not contain empty sets, and $\mathbb{R} \setminus X$ is not empty because X has finite outer measure, so $X \neq \mathbb{R}$. Thus $\overline{X}$ and $\overline{\mathbb{R} \setminus X}$ are also both non-empty. Since $\mathbb{R}$ is connected, there is a point $\vec{p}$ common to both sets. Thus for any norm $\|\cdot\|$ it holds for all $\varepsilon \in (0, \infty)$ that ${}^{\|\cdot\|}\overline{B}_\varepsilon(\vec{p})$ intersects both X and $\mathbb{R} \setminus X$ so $\left|{}^{\|\cdot\|}\overline{\mathcal{N}}_\varepsilon(\vec{p})\right| \geq 2 = d + 1$. Thus, the conclusion of the Stronger Optimality Theorem follows from the hypothesis of the Standard Optimality Theorem in the case $d = 1$. □

Proposition 6.4.2 (“Stronger” and “Strongest” Optimality Gap)**.** *The hypothesis of the Stronger Optimality Theorem does not imply the conclusion of the Strongest Optimality Theorem for any $d \in \mathbb{N}$. (Also, the hypothesis of the Standard Optimality Theorem does not imply the conclusions of the Strongest Optimality Theorem for any $d \in \mathbb{N}$.)*

Proof. We will construct a partition which satisfies the hypotheses of both the Standard Optimality Theorem and the Stronger Optimality Theorem but which does not satisfy the conclusion of the Strongest Optimality Theorem.

Consider the partition of singletons $\mathcal{P} = \left\{\{\vec{x}\} : \vec{x} \in \mathbb{R}^d\right\}$. 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 $\mathcal{N}_{\bar{0}}(\vec{p}) = \left\{X \in \mathcal{P} : \overline{X} \ni \vec{p}\right\} = \{X \in \mathcal{P} : X \ni \vec{p}\} = \{\vec{p}\}$ which has size $1 < d + 1$. Thus, the conclusion of the Strongest Optimality Theorem does not hold. □

Chapter 7

Near Optimality of ε in General

At this stage, we have firmly established the optimality of the parameter value $k = d+1$ and will turn our attention to the ε parameter. Motivated by the results of the prior chapter, we will continue to focus on two families of partitions: those with a uniform upper bound on the outer measure of the members, and those with a uniform upper bound on the diameter of the members. Informally, in discussion we refer to these as measure bounded partitions and diameter bounded partitions.

Our concern then will be to consider holding k at the optimal value of $k = d+1$ and trying to determine for what values of $\varepsilon \in (0, \infty)$ there exist $(d+1, \varepsilon)$-secluded partitions among these families. Unlike with the k parameter, we now must be concerned with what the values of these upper bounds on measure or diameter are for the partitions we consider because otherwise for any $\varepsilon \in (0, \infty)$ we can trivially find a $(d+1, \varepsilon)$-secluded partition: just scale the $(d+1, \frac{1}{2d})$-secluded unit cube partitions we constructed in Chapter 4 (Constructions) so the cubes have side length $2d\varepsilon$.

In the first section of this chapter, Section 7.1 (Upper Bounds on ε via Brunn-Minkowski and Blichfeldt), we will demonstrate that our unit cube partitions are very nearly optimal even among unit bounded measure partitions (which subsumes the family of unit diameter bounded partitions in ℓ_∞); more specifically, for any dimension d, when $k = d + 1$, the best possible ε is at most $\frac{\log_4(d+1)}{d}$ which differs from our construction by only a logarithmic factor. Furthermore, the result we obtain is not specific to the value $k = d + 1$; for any $k \leq 2^d$—an extremely reasonable assumption—we must have $\varepsilon \leq \frac{\log_4(k)}{d}$. This implies that even if we don't restrict to the minimum value of $k = d + 1$ but instead allow for k to be polynomial in the dimension d, we get hardly any improvement in the possible values of ε because we still have $\varepsilon \in O\left(\frac{\log(d)}{d}\right)$ so even then we can only hope to gain a logarithmic factor. Furthermore, our result isn't even specific to the ℓ_∞ norm, and we get some bound on ε for any possibly normed ball one wants to consider.

In the second section of this chapter, Section 7.2 (Upper Bound on ε via the Dissection Number of the Cube), we present another bound on ε which is in terms of the dissection number of the cube. This was a result of ours that predated those discussed above and is much less powerful in three regards: (1) it only applies to the ℓ_∞ norm[1], (2) it is fundamentally tailored to the case when $k = d+1$ and (3) it only applies to diameter bounded partitions and not measure bounded partitions[2]. Furthermore, at the time of writing, the current known bounds on the dissection number are not good enough for this ε bound to improve on our bound above. Nonetheless, it is consistent with current knowledge that the bounds on the dissection number could improve enough that our bound on ε becomes $O(\frac{1}{d})$ when $k = d + 1$ for unit ℓ_∞ diameter bounded partitions. So, even though this is a much more limited result than those discussed above, it would be very interesting if we knew that our unit cube constructions have asymptotically optimal ε among unit ℓ_∞ diameter partitions when $k = d + 1$.

Lastly, in the third section of this chapter, Section 7.3 (Conjectures on Optimal

[1] It would not be difficult to modify the proof to get analogous versions for other norms, though.

[2] We considered using the same pixelation techniques as used in the proof of the Theorem 6.2.1, but they lead to basically useless bounds.

ε), we present some conjectures regarding the true optimal value of ε for the different classes of partitions we have discussed.

7.1 Upper Bounds on ε via Brunn-Minkowski and Blichfeldt

The main result we prove in this section is the following which guarantees that for a partition with a bound of M on the outer measure, it holds for every $\varepsilon \in (0,\infty)$ that there is a point $\vec{p}$ where the ball ${}^{\|\cdot\|}B^{\circ}_{\varepsilon}(\vec{p})$ intersects a substantial number of members of the partition. In other words, this expression gives a lower bound on our parameter k in terms of the parameter ε, and by solving the equations for ε we can obtain upper bounds on ε in terms of our parameter k (which has been our preferred perspective so far). Throughout this section, we will let $v_{\|\cdot\|,d}$ denote the measure of the unit ball in $\mathbb{R}^d$ with respect to a norm $\|\cdot\|$—that is $v_{\|\cdot\|,d} = m\left({}^{\|\cdot\|}B^{\circ}_{1}(\vec{0})\right)$. We now state the main result of this section (and delay the proof).

Theorem 7.1.1 (ε-Neighborhoods for Measure Bounded Partitions and Arbitrary Norm)**.** *Let $d \in \mathbb{N}$ and $M \in (0,\infty)$. Let $\mathbb{R}^d$ be equipped with any norm $\|\cdot\|$. Let $\mathcal{P}$ be a partition of $\mathbb{R}^d$ such that every member has outer Lebesgue measure at most M. Then for every $\varepsilon \in (0,\infty)$, there exists $\vec{p} \in \mathbb{R}^d$ such that ${}^{\|\cdot\|}B^{\circ}_{\varepsilon}(\vec{p})$ intersects at least $\left\lceil \left(1+\varepsilon\left(\frac{v_{\|\cdot\|,d}}{M}\right)^{1/d}\right)^d \right\rceil$ members of $\mathcal{P}$. That is*

$$\left|{}^{\|\cdot\|}\mathcal{N}^{\circ}_{\varepsilon}(\vec{p})\right| \geq \left\lceil \left(1+\varepsilon\left(\frac{v_{\|\cdot\|,d}}{M}\right)^{1/d}\right)^d \right\rceil.$$

The expression $\varepsilon\left(\frac{v_{\|\cdot\|,d}}{M}\right)^{1/d} = \frac{(\varepsilon^d \cdot v_{\|\cdot\|,d})^{1/d}}{M^{1/d}}$ should be viewed as a normalization factor. It is typical in measure theory contexts to see the dth roots of measures in d dimensions show up, and they often serve as a type of characteristic length scale and are sometimes more robust than actual distances. Thus, this expression should be viewed as the ratio of the characteristic length scale of the ε ball (which has measure $\varepsilon^d \cdot v_{\|\cdot\|,d}$) to the maximum characteristic length scale of the members of the partition (which have measure at most M).

We now want to examine a particular perspective on Theorem 7.1.1 which gives a much cleaner looking bound. In an asymptotic setting where one cares about a fixed norm $\|\cdot\|$ and how this bound changes as the dimension increases, it would be a bit peculiar to have interest in a fixed measure M that applies for all dimensions; this is because the measure of objects in one dimension is not really comparable to objects in another dimension. For this reason, it is probably not productive to think of the bound $\left(1+\varepsilon\left(\frac{v_{\|\cdot\|,d}}{M}\right)^{1/d}\right)^d$ as a reasonable asymptotic expression in d. What would be quite natural, though, is to have in each dimension a bound on the measures of the members relative to the measure of the unit $\|\cdot\|$-ball in that dimension. This is because the unit ball would serve as a reference point for the types of objects one might be interested in. In other words, it would be quite natural if there was some fixed radius r such that for each dimension $d \in \mathbb{N}$, the partition of $\mathbb{R}^d$ had members of measure no greater than the measure of the ball of radius r. In this natural setting, the $v_{\|\cdot\|,d}$ factor drops out and we have the immediate corollary below.

Corollary 7.1.2 (ε-Neighborhoods for Measure Bounded Partitions and Arbitrary Norm)**.** *Let $d \in \mathbb{N}$ and $r \in (0,\infty)$. Let $\mathbb{R}^d$ be equipped with any norm $\|\cdot\|$. Let $\mathcal{P}$ be a partition of $\mathbb{R}^d$ such that every member has outer Lebesgue measure at most $m\left({}^{\|\cdot\|}B_r^\circ(\vec{0})\right)$. Then for every $\varepsilon \in (0,\infty)$, there exists $\vec{p} \in \mathbb{R}^d$ such that ${}^{\|\cdot\|}B_\varepsilon^\circ(\vec{p})$ intersects at least $\left\lceil\left(1+\frac{\varepsilon}{r}\right)^d\right\rceil$ members of $\mathcal{P}$. That is*

$$\left|{}^{\|\cdot\|}\mathcal{N}_\varepsilon^\circ(\vec{p})\right| \geq \left\lceil\left(1+\frac{\varepsilon}{r}\right)^d\right\rceil.$$

Proof. Let $M = m\left({}^{\|\cdot\|}B_r^\circ(\vec{0})\right) = r^d \cdot v_{\|\cdot\|,d}$. Then by Theorem 7.1.1, there exists $\vec{p} \in \mathbb{R}^d$

such that

$$\left| {}^{\|\cdot\|}\mathcal{N}^{\circ}_{\|\cdot\|,\varepsilon}(\vec{p}) \right| \geq \left\lceil \left(1 + \varepsilon \left(\frac{v_{\|\cdot\|,d}}{M}\right)^{1/d}\right)^d \right\rceil \qquad \text{(Theorem 7.1.1)}$$

$$= \left\lceil \left(1 + \frac{\varepsilon}{r}\right)^d \right\rceil \qquad \text{(Definition of } M\text{)}$$

which completes the proof. □

This perspective also leads to a very clean result in terms of bounded diameter if we make use of the Isodiametric Inequality stated below as a way to convert diameter bounds into measure bounds. It states that no bounded set has greater measure than the ball of the same diameter. Thus, an upper bound on the diameters of members (in any norm) immediately gives an upper bound on the measures of the members.

Theorem A.2.1 (Isodiametric Inequality)**.** *Let $d \in \mathbb{N}$ and $\|\cdot\|$ be a norm on $\mathbb{R}^d$. Let $X \subseteq \mathbb{R}^d$ be a bounded set and $D = \operatorname{diam}_{\|\cdot\|}(X)$. Then the outer Lebesgue measure of X is at most the Lebesgue measure of the ball of diameter D. That is, (in three equivalent forms):*

$$\begin{aligned} m_{out}(A) &\leq m\left({}^{\|\cdot\|}B^{\circ}_{D/2}(\vec{0})\right) \\ &= \left(\tfrac{D}{2}\right)^d \cdot m\left({}^{\|\cdot\|}B^{\circ}_{1}(\vec{0})\right) \\ &= \left(\tfrac{D}{2}\right)^d \cdot v_{\|\cdot\|,d}. \end{aligned}$$

In some sources, this inequality is proven only for the ℓ_2 norm using Steiner symmetrization since it is geometrically quite intuitive, but there is a more general (and very short) proof for all norms using the Generalized Brunn-Minkowski Inequality (Theorem 7.1.4) which we will be introducing shortly. For convenience, we offer a standard proof of the Isodiametric Inequality (Theorem A.2.1) using the latter technique in the linked appendix.

Using this inequality, we get the following analog of Theorem 7.1.1 and Corollary 7.1.2 in terms of diameter bounded partitions.

Corollary 7.1.3 (ε-Neighborhoods for Diameter Bounded Partitions)**.** *Let $d \in \mathbb{N}$ and $D \in (0,\infty)$. Let $\mathbb{R}^d$ be equipped with any norm $\|\cdot\|$. Let $\mathcal{P}$ be a partition of $\mathbb{R}^d$ such that every member has diameter at most D (with respect to $\|\cdot\|$). Then for every $\varepsilon \in (0,\infty)$, there exists $\vec{p} \in \mathbb{R}^d$ such that ${}^{\|\cdot\|}B^\circ_\varepsilon(\vec{p})$ intersects at least $\left\lceil \left(1+\frac{2\varepsilon}{D}\right)^d \right\rceil$ members. That is*

$$|{}^{\|\cdot\|}\mathcal{N}^\circ_\varepsilon(\vec{p})| \geq \left\lceil \left(1+\frac{2\varepsilon}{D}\right)^d \right\rceil.$$

Proof. Let $r = \frac{D}{2}$. Then by the Isodiametric Inequality, for all $X \in \mathcal{P}$ we have $m_{out}(X) \leq m\left({}^{\|\cdot\|}B^\circ_r(\vec{0})\right)$ so by Corollary 7.1.2, there exists $\vec{p} \in \mathbb{R}^d$ such that

$$\begin{aligned} \left|{}^{\|\cdot\|}\mathcal{N}^\circ_{\|\cdot\|,\varepsilon}(\vec{p})\right| &\geq \left\lceil \left(1+\frac{\varepsilon}{r}\right)^d \right\rceil && \text{(Corollary 7.1.2)} \\ &= \left\lceil \left(1+\frac{2\varepsilon}{D}\right)^d \right\rceil && \text{(Definition of } r\text{)} \end{aligned}$$

□

The expression $\frac{2\varepsilon}{D}$ above can and should be viewed as a normalization factor which is the ratio of the diameter of the ε-ball to the maximum diameter of the members of the partition.

The rest of this section will consist of the following. In Subsection 7.1.1 (Proof of Main Result) we prove the main result of this chapter (i.e. Theorem 7.1.1). Then, in Subsection 7.1.2 (Upper Bounds on ε), we convert the three results above from lower bounds on k in terms of ε to the more desired format of upper bounds on ε in terms of k, and we also give a simple asymptotic corollary. We end in Subsection 7.1.3 (A Short Discussion of Specific Norms) with a very terse discussion of the bounds obtained from Theorem 7.1.1 for the ℓ_1, ℓ_2, and ℓ_∞ norms; this is by no means intended to be exhaustive and we emphasize again that we believe the perspective of Corollary 7.1.2 is more useful than that of Theorem 7.1.1 anyway, and in that perspective the specific norm is irrelevant to the value of the bound.

7.1.1 Proof of Main Result

We begin with five simple facts so that we don't have to justify them later. All have straightforward proofs provided in various appendices with hyperlinks provided by each for convenience.

The first fact (Fact G.0.1) will later allow us to pass a result through a limit because the answer will be an integer.

Fact G.0.1. *For any $\alpha \in \mathbb{R}$, there exists $\gamma \in \mathbb{R}$ such that $\gamma < \alpha$ and $\lceil \gamma \rceil = \lceil \alpha \rceil$.*

The second fact (Fact G.0.2) is a specific inequality that we will need to use. It can be interpreted as saying that for appropriate parameters, we can essentially factor our the "x" in $(x^{1/d} + \alpha)^d$ to get the no larger expression $x(1+\alpha)^d$.

Fact G.0.2. *For $d \in [1, \infty)$, $x \in [0, 1]$, and $\alpha \in [0, \infty)$, it holds that $(x^{1/d} + \alpha)^d \geq x(1+\alpha)^d$.*

The third fact (Fact C.0.1) says that the Minkowski sum of a set X and an open ball at the origin can be viewed as a union of open balls positioned at each point of X.

Fact C.0.1. *For any normed vector space, given a set X and $\varepsilon \in (0, \infty)$, then*

$$X + B^\circ_\varepsilon(\vec{0}) = \bigcup_{\vec{x} \in X} B^\circ_\varepsilon(\vec{x}).$$

The same can be said replacing open balls with closed balls.

The fourth fact (Fact C.0.2) says that we can decompose a ball into a Minkowski sum of two smaller balls.

Fact C.0.2. *For any normed vector space, and any $\alpha, \beta \in (0, \infty)$, it holds that $B^\circ_\alpha(\vec{0}) + B^\circ_\beta(\vec{0}) = B^\circ_{\alpha+\beta}(\vec{0})$.*

The final fact (Fact C.0.3), while also very simple, is the key change of perspective that allowed us to prove the main results of this section. It says that if we are checking if a member X in our partition intersects an ε-ball located at $\vec{p}$ (in order to see how many such members there are), we can instead enlarge X by taking its Minkowski sum with the origin-centered ε-ball, and check if this enlarged member contains the point $\vec{p}$.

Fact C.0.3. *For any normed vector space, for any set X and vector $\vec{p}$, the following are equivalent:*

1. $\overline{B}_\varepsilon(\vec{p}) \cap X \neq \emptyset$
2. $\vec{p} \in X + \overline{B}_\varepsilon(\vec{0})$

The same can be said replacing both closed balls with open balls.

Now we introduce the result which is the connection to the above mentioned key change of perspective. The result says to consider a subset $S \subseteq \mathbb{R}^d$ with finite measure and a collection $\mathcal{A}$ of (measurable) subsets of S. If we compute the sum of measures of all members in the collection $\mathcal{A}$ (i.e. intuitively the total volume that they take up), and compare this to the measure/volume of S, then whatever this ratio is, we can find a point in S covered by that many members of the collection $\mathcal{A}$. For example, in the simplest case that the total measure of members of $\mathcal{A}$ is larger than the measure of S, then there is no way for all of the members of $\mathcal{A}$ to be disjoint, so there has to be some point covered by two members.

In the more generic case, this result should be intuitively true by an averaging argument: if every point of S is covered only n times, then the total measure of members in $\mathcal{A}$ is at most $n \cdot m(S)$, so if the ratio of total measure in $\mathcal{A}$ to the measure of S is large, then n must also be large.

Proposition A.3.1 (Lower Bound Cover Number for $\mathbb{R}^d$)**.** *Let $d \in \mathbb{N}$ and $S \subset \mathbb{R}^d$ be measurable with finite measure. Let $\mathcal{A}$ be a family of measurable subsets of S and let $k = \left\lceil \frac{\sum_{A \in \mathcal{A}} m(A)}{m(S)} \right\rceil$. If $k < \infty$, then there exists $\vec{p} \in S$ such that $\vec{p}$ belongs to at least k members of $\mathcal{A}$. If $k = \infty$, then for any integer n, there exists $\vec{p} \in S$ such that $\vec{p}$ belongs to at least n members of $\mathcal{A}$.*

Unfortunately, though this result might be intuitive, proving it formally does require some effort. We first encountered this result as the main ingredient in the standard proof of Blitchfeldt's Theorem (Theorem 5.2.2) which was the source of motivation for our main technique as discussed in Section 5.2 (Near Optimality of ε for Unit Cube Partitions). However, many of the sources we found where proofs of Blitchfeldt's Theorem are presented did not prove the result above except in the special case where $k = 2$ (which is far simpler from the other cases from a measure theory perspective), so for convenience and completion we provide a proof of Proposition A.3.1 in the linked appendix[3].

[3]Our proof requires three stages which are proven as (1) Lemma A.3.2 (Exact Measure of Multiplicity), (2) Lemma A.3.3 (Upper Bound Measure of Multiplicity), and (3) Corollary A.3.4 (Lower Bound Cover Number). Then Proposition A.3.1 follows immediately as a corollary specific to the standard (Lebesgue or Borel) measure space on $\mathbb{R}^d$.

The next ingredient that we need is a way to measure how large the Minkowski sum in Fact C.0.3 is. In order to utilize the Lower Bound Cover Number for $\mathbb{R}^d$ we need a lower bound on the measures, and we can obtain one using the Generalized Brunn-Minkowski Inequality stated below.

Theorem 7.1.4 (Generalized Brunn-Minkowski Inequality)**.** *Let $d \in \mathbb{N}$ and $A, B \subseteq \mathbb{R}^d$ be non-empty and Lebesgue measurable such that $A + B$ is also Lebesgue measurable. Then*

$$m(A+B) \geq \left[m(A)^{\frac{1}{d}} + m(B)^{\frac{1}{d}}\right]^d .$$

This version of the statement can be obtained from [Gar02, Equation 11]; in that survey, Gardner states this theorem with a requirement that the sets be bounded, but in the following paragraph notes that this is not necessary and the requirement is only stated for convenience of the presentation in that survey.

In the theorem, the requirement that $A + B$ is Lebesgue measurable is not a triviality; Gardner discusses that there exist known Lebesgue measurable sets A and B such that the Minkowski sum $A+B$ is not Lebesgue measurable as shown in [Sie20]. The next result gives us a way to circumvent this issue in our application even if the members of our partition are not measurable by taking B to be an open set so that the sum $A + B$ is open (and thus measurable), and using the outer measure of A so that we don't need the assumption that A is measurable.

Lemma 7.1.5 (Brunn-Minkowsi with Balls)**.** *Let $d \in \mathbb{N}$ and let $\mathbb{R}^d$ be equipped with any norm $\|\cdot\|$. Let $Y \subseteq \mathbb{R}^d$, and $\varepsilon \in (0, \infty)$. Then $Y + {}^{\|\cdot\|}B^\circ_\varepsilon(\vec{0})$ is open (and thus Borel measurable), and $m(Y + {}^{\|\cdot\|}B^\circ_\varepsilon(\vec{0})) \geq \left(m_{out}(Y)^{\frac{1}{d}} + \varepsilon \cdot (v_{\|\cdot\|,d})^{\frac{1}{d}}\right)^d$.*

Proof. By Fact C.0.1, $Y + {}^{\|\cdot\|}B^\circ_\varepsilon(\vec{0}) = \bigcup_{\vec{y} \in Y} {}^{\|\cdot\|}B^\circ_\varepsilon(\vec{y})$ which is a union of open sets, so is itself open and thus Borel measurable. Now, for any $\varepsilon' \in (0, \varepsilon)$, observe that by Fact C.0.2, ${}^{\|\cdot\|}B^{\circ}_{\varepsilon}(\vec{0}) = {}^{\|\cdot\|}B^{\circ}_{\varepsilon-\varepsilon'}(\vec{0}) + {}^{\|\cdot\|}B^{\circ}_{\varepsilon'}(\vec{0})$ and thus, this sum is measurable because it is an open ball. Using this equality and the associativity of the Minkowski sum, we have

$$Y + {}^{\|\cdot\|}B^{\circ}_{\varepsilon}(\vec{0}) = Y + \left[{}^{\|\cdot\|}B^{\circ}_{\varepsilon-\varepsilon'}(\vec{0}) + {}^{\|\cdot\|}B^{\circ}_{\varepsilon'}(\vec{0})\right] = \left[Y + {}^{\|\cdot\|}B^{\circ}_{\varepsilon-\varepsilon'}(\vec{0})\right] + {}^{\|\cdot\|}B^{\circ}_{\varepsilon'}(\vec{0}).$$

Thus, we have the following inequalities:

$$\begin{aligned} m\left(Y + {}^{\|\cdot\|}B^{\circ}_{\varepsilon}(\vec{0})\right) &= m\left(\left[Y + {}^{\|\cdot\|}B^{\circ}_{\varepsilon-\varepsilon'}(\vec{0})\right] + {}^{\|\cdot\|}B^{\circ}_{\varepsilon'}(\vec{0})\right) && \text{(Above)} \\ &\geq \left(m\left(Y + {}^{\|\cdot\|}B^{\circ}_{\varepsilon-\varepsilon'}(\vec{0})\right)^{\frac{1}{d}} + m\left({}^{\|\cdot\|}B^{\circ}_{\varepsilon'}(\vec{0})\right)^{\frac{1}{d}}\right)^{d} \end{aligned}$$

The above comes from the Generalized Brunn-Minkowski Inequality (Theorem 7.1.4) noting that as demonstrated above, both sets $Y + {}^{\|\cdot\|}B^{\circ}_{\varepsilon-\varepsilon'}(\vec{0})$ and ${}^{\|\cdot\|}B^{\circ}_{\varepsilon'}(\vec{0})$ are open and thus measurable. We continue.

$$\geq \left(m_{out}\left(Y\right)^{\frac{1}{d}} + m\left({}^{\|\cdot\|}B^{\circ}_{\varepsilon'}(\vec{0})\right)^{\frac{1}{d}}\right)^{d}$$

The above inequality comes from the definition of the outer measure of Y being the infimum of the measures of all measurable supersets of Y. Since $Y \subseteq Y + {}^{\|\cdot\|}B^\circ_{\varepsilon'}(\vec{0})$, we get the inequality above. Continuing, we have the following:

$$= \left(m_{out}\left(Y\right)^{\frac{1}{d}} + m\left(\varepsilon' \cdot {}^{\|\cdot\|}B^\circ_1(\vec{0})\right)^{\frac{1}{d}} \right)^d$$

(Scaling of norm-based balls)

$$= \left(m_{out}\left(Y\right)^{\frac{1}{d}} + \left[(\varepsilon')^d \cdot m\left({}^{\|\cdot\|}B^\circ_1(\vec{0})\right) \right]^{\frac{1}{d}} \right)^d$$

(Scaling for Lebesgue measure)

$$= \left(m_{out}\left(Y\right)^{\frac{1}{d}} + \varepsilon' \cdot (v_{\|\cdot\|,d})^{\frac{1}{d}} \right)^d \quad \left(\text{Algebra and } v_{\|\cdot\|,d} \stackrel{\text{def}}{=} m\left({}^{\|\cdot\|}B^\circ_1(\vec{0})\right)\right)$$

Since the inequality above holds for all $\varepsilon' \in (0, \varepsilon)$, it must also hold in the limit (keeping d and Y fixed):

$$m\left(Y + {}^{\|\cdot\|}B^\circ_\varepsilon(\vec{0})\right) \geq \lim_{\varepsilon' \to \varepsilon} \left[\left(m_{out}\left(Y\right)^{\frac{1}{d}} + \varepsilon' \cdot (v_{\|\cdot\|,d})^{\frac{1}{d}} \right)^d \right] = \left(m_{out}\left(Y\right)^{\frac{1}{d}} + \varepsilon \cdot (v_{\|\cdot\|,d})^{\frac{1}{d}} \right)^d$$

which concludes the proof. □

At this point we are in a position to prove the main result of this section which we first restate for convenience. The idea of the proof is illustrated in Figure 7.1 for the ℓ_∞ norm and $d = 2$.

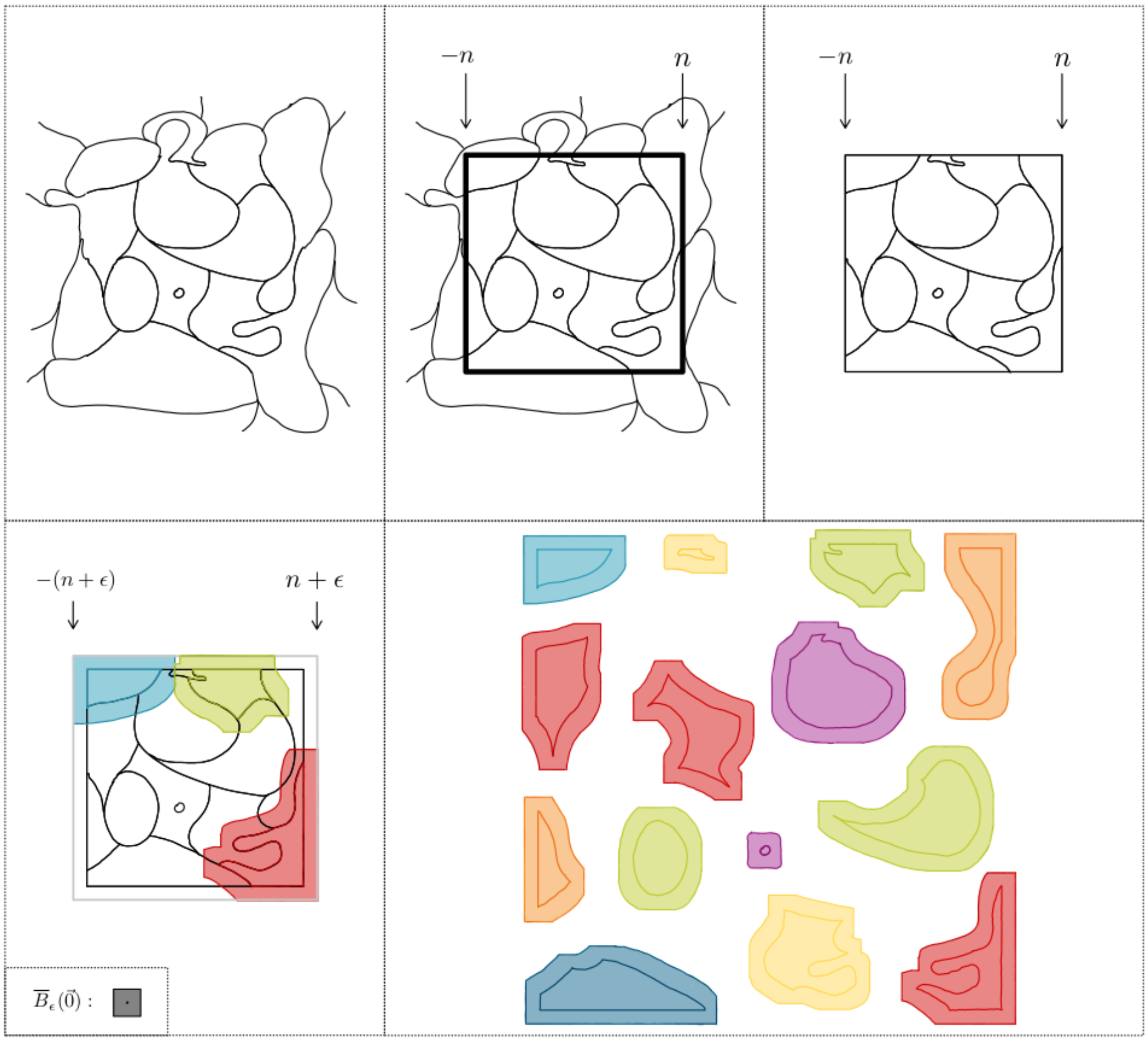


Figure 7.1: In the first pane, we have a partition of $\mathbb{R}^2$. In the second pane, we show that we will consider only members of the partition which intersect $[-n, n]^2$, and in the third pane we show the partition that $\mathcal{P}$ induces on $[-n, n]^2$. In the fourth pane, we consider enlarging each member by placing an ε-ball at each point of the member and show that these enlarged elements are still contained within $[-(n + \varepsilon), n + \varepsilon]^2$. In the fifth pane, we see all of the expanded members and observe that the sum of the areas of the enlarged members is "significantly" more that the area of $[-n, n]^2$. There is a "significant" amount of excess area, there is a point $\vec{p} \in [-n - \varepsilon, n + \varepsilon]^2$ belonging to "many" of the enlarged sets, and consequently $\vec{p} + [-\varepsilon, \varepsilon]^2$ intersects the corresponding original sets.

Theorem 7.1.1 (ε-Neighborhoods for Measure Bounded Partitions and Arbitrary Norm)**.** *Let $d \in \mathbb{N}$ and $M \in (0, \infty)$. Let $\mathbb{R}^d$ be equipped with any norm $\|\cdot\|$. Let $\mathcal{P}$ be a partition of $\mathbb{R}^d$ such that every member has outer Lebesgue measure at most M. Then for every $\varepsilon \in (0, \infty)$, there exists $\vec{p} \in \mathbb{R}^d$ such that ${}^{\|\cdot\|}B^\circ_\varepsilon(\vec{p})$ intersects at least $\left\lceil \left(1 + \varepsilon \left(\frac{v_{\|\cdot\|,d}}{M}\right)^{1/d}\right)^d \right\rceil$ members of $\mathcal{P}$. That is*

$$|{}^{\|\cdot\|}\mathcal{N}^\circ_\varepsilon(\vec{p})| \geq \left\lceil \left(1 + \varepsilon \left(\frac{v_{\|\cdot\|,d}}{M}\right)^{1/d}\right)^d \right\rceil.$$

Proof. Throughout the proof, all lengths will be with respect to $\|\cdot\|$, so we will eliminate the clutter by neglecting to use the $\|\cdot\|$ subscript/superscript/prescript anywhere in the proof. We will also be working in a single dimension d throughout the proof, so we write v instead of $v_{\|\cdot\|,d}$ throughout.

Consider the following for each $n \in \mathbb{N}$. Let $S_n = B^\circ_n(\vec{0})$ and $S'_n = B^\circ_{n+\varepsilon}(\vec{0}) = S_n + B^\circ_\varepsilon(\vec{0})$ and $\mathcal{S}$ be the partition of S_n induced[4] by $\mathcal{P}$. For each $Y \in \mathcal{S}_n$, let $C_Y = Y + B^\circ_\varepsilon(\vec{0})$. By Lemma 7.1.5, C_Y is measurable, and $m(C_Y) \geq \left(m_{out}(Y)^{\frac{1}{d}} + \varepsilon \cdot v^{\frac{1}{d}}\right)^d$.

[4] I.e. $\mathcal{S} = \{X \cap S_n : X \in \mathcal{P} \text{ and } X \cap S_n \neq \emptyset\}$. That is, $\mathcal{S}$ is the partition of S_n obtained by intersecting every member of $\mathcal{P}$ with the new domain S_n and keeping those which have non-empty intersection.

Also observe that $C_Y \subseteq S'_n$. Now consider the following inequalities:

$$\begin{aligned}
m(C_Y) &\geq \left(m_{out}\left(Y\right)^{\frac{1}{d}} + \varepsilon \cdot v^{\frac{1}{d}}\right)^d && \text{(Above)}\\
&= \left(M^{1/d}\left[\frac{m_{out}\left(Y\right)^{\frac{1}{d}}}{M^{1/d}} + \frac{\varepsilon \cdot v^{\frac{1}{d}}}{M^{1/d}}\right]\right)^d && \text{(Inroduce } M^{1/d}\text{)}\\
&= M\left(\left[\frac{m_{out}\left(Y\right)}{M}\right]^{\frac{1}{d}} + \varepsilon\left(\tfrac{v}{M}\right)^{\frac{1}{d}}\right)^d && \text{(Simlify)}\\
&\geq M \cdot \frac{m_{out}\left(Y\right)}{M} \cdot \left(1 + \varepsilon\left(\tfrac{v}{M}\right)^{\frac{1}{d}}\right)^d && \text{(Fact G.0.2 since } \tfrac{m_{out}(Y)}{M} \in [0,1]\text{)}\\
&= m_{out}\left(Y\right) \cdot \left(1 + \varepsilon\left(\tfrac{v}{M}\right)^{\frac{1}{d}}\right)^d && \text{(Simplify)}
\end{aligned}$$

Informally, the above shows that for each $Y \in \mathcal{S}_n$, the set $Y + B^\circ_\varepsilon(\vec{0})$ has substantially more (outer) measure than Y does—specifically a factor of $\left(1 + \varepsilon\left(\frac{v}{M}\right)^{\frac{1}{d}}\right)^d$. We will extend this to unsurprisingly show that this implies that $\left\{Y + B^\circ_\varepsilon(\vec{0})\right\}_{Y \in \mathcal{S}_n}$ also has this same factor more (outer) measure than $\mathcal{S}_n$ does, observing that $\mathcal{S}_n$ has total (outer) measure of about $m(S_n)$ since $\mathcal{S}_n$ is a partition of S_n (any discrepancy is due to non-measurable members in $\mathcal{S}_n$)

Formally, we claim that there exists a finite subfamily $\mathcal{F}_n \subseteq \mathcal{S}_n$ such that

$$\sum_{Y \in \mathcal{F}_n} m\left(Y + B^\circ_\varepsilon(\vec{0})\right) \geq \left(1 + \varepsilon\left(\tfrac{v}{M}\right)^{\frac{1}{d}}\right)^d \cdot m(S_n).$$

To see this, first consider the case that $\mathcal{S}_n$ has infinite cardinality. Let $\mathcal{F}_n \subset \mathcal{S}_n$ be any subfamily of finite cardinality at least $\left(1 + \varepsilon\left(\frac{v}{M}\right)^{\frac{1}{d}}\right)^d \cdot m(S_n) \cdot \frac{1}{\varepsilon^d v}$. This gives

$$\sum_{Y \in \mathcal{F}_n} m\left(Y + B^\circ_\varepsilon(\vec{0})\right) \geq \sum_{Y \in \mathcal{F}_n} m\left(B^\circ_\varepsilon(\vec{0})\right)$$

where this inequality is because $Y + B^\circ_\varepsilon(\vec{0})$ is a superset of some translation of $B^\circ_\varepsilon(\vec{0})$ since $Y \neq \emptyset$. Continuing, we use the standard fact that $m\left(B^\circ_\varepsilon(\vec{0})\right) = m\left(\varepsilon \cdot B^\circ_1(\vec{0})\right) = \varepsilon^d \cdot m\left(B^\circ_1(\vec{0})\right) = \varepsilon^d v$:

$$\begin{aligned}
&\geq \sum_{Y \in \mathcal{F}_n} \varepsilon^d v \\
&= \left[\left(1 + \varepsilon \left(\tfrac{v}{M}\right)^{\frac{1}{d}}\right)^d \cdot m(S_n) \cdot \frac{1}{\varepsilon^d v}\right] \cdot \varepsilon^d v && \text{(Cardinality of } \mathcal{F}_n\text{)} \\
&= \left(1 + \varepsilon \left(\tfrac{v}{M}\right)^{\frac{1}{d}}\right)^d \cdot m(S_n). && \text{(Simplify)}
\end{aligned}$$

Now consider the other (more interesting) case where $\mathcal{S}_n$ has finite cardinality[5]. Take $\mathcal{F}_n = \mathcal{S}_n$ so that

$$\begin{aligned}
\sum_{Y \in \mathcal{F}_n} m\left(Y + B^\circ_\varepsilon(\vec{0})\right) &= \sum_{Y \in \mathcal{S}_n} m\left(Y + B^\circ_\varepsilon(\vec{0})\right) && (\mathcal{F}_n = \mathcal{S}_n) \\
&\geq \sum_{Y \in \mathcal{S}_n} m_{out}(Y) \cdot \left(1 + \varepsilon \left(\tfrac{v}{M}\right)^{\frac{1}{d}}\right)^d && \text{(Above)} \\
&= \left(1 + \varepsilon \left(\tfrac{v}{M}\right)^{\frac{1}{d}}\right)^d \cdot \sum_{Y \in \mathcal{S}_n} m_{out}(Y) && \text{(Linearity of summation)} \\
&\geq \left(1 + \varepsilon \left(\tfrac{v}{M}\right)^{\frac{1}{d}}\right)^d \cdot m_{out}\left(\bigsqcup_{Y \in \mathcal{S}_n} Y\right)
\end{aligned}$$

where the above inequality is due to the countable subaddativity property of outer measures which states that a countable sum of outer measures of sets is at least as large as the outer measure of the union of the sets. The disjointness of the union is irrelevant. In the last step we get

$$= \left(1 + \varepsilon \left(\tfrac{v}{M}\right)^{\frac{1}{d}}\right)^d \cdot m(S_n) \qquad \left(\textstyle\bigsqcup_{Y \in \mathcal{F}_n} Y = S_n \text{ is measurable}\right)$$

[5]In fact this case also works if $\mathcal{S}_n$ is countable even though we have already dealt with that case.

Thus, regardless of whether $\mathcal{S}_n$ has infinite or finite cardinality, there exists a finite subfamily $\mathcal{F}_n \subseteq \mathcal{S}_n$ such that

$$\sum_{Y\in\mathcal{F}_n} m\left(Y + B^\circ_\varepsilon(\vec{0})\right) \geq \left(1+\varepsilon\left(\tfrac{v}{M}\right)^{\frac{1}{d}}\right)^d \cdot m(S_n).$$

Fix such a subfamily $\mathcal{F}_n$, and let $\mathcal{A}_n = \left\{Y + B^\circ_\varepsilon(\vec{0})\right\}_{Y\in\mathcal{F}_n}$ be a family indexed[6] by $\mathcal{F}_n$. Note that for each $A_Y \overset{\text{def}}{=} Y + B^\circ_\varepsilon(\vec{0}) \in \mathcal{A}_n$, since $Y \subseteq S_n = B^\circ_n(\vec{0})$, we have $A_Y \subseteq S_n + B^\circ_\varepsilon(\vec{0}) = S'_n$. Thus, by Corollary A.3.4[7], there is a point $\vec{p}^{(n)} \in S'_n$ which belongs to at least k_n-many sets in $\mathcal{A}_n$ where

$$\begin{aligned}
k_n &\overset{\text{def}}{=} \frac{\sum_{Y\in\mathcal{F}_n} m\left(Y + B^\circ_\varepsilon(\vec{0})\right)}{m(S'_n)} \\
&\geq \frac{\left(1+\varepsilon\left(\frac{v}{M}\right)^{\frac{1}{d}}\right)^d \cdot m\left(S_n\right)}{m(S'_n)} && \text{(Above)} \\
&= \left(1+\varepsilon\left(\tfrac{v}{M}\right)^{\frac{1}{d}}\right)^d \cdot \frac{m\left(B^\circ_n(\vec{0})\right)}{m\left(B^\circ_{n+\varepsilon}(\vec{0})\right)} && \text{(Def'n of } S_n \text{ and } S'_n\text{)} \\
&= \left(1+\varepsilon\left(\tfrac{v}{M}\right)^{\frac{1}{d}}\right)^d \cdot \frac{n^d\cdot v}{(n+\varepsilon)^d\cdot v} && \text{(Standard scaling fact)} \\
&= \left(1+\varepsilon\left(\tfrac{v}{M}\right)^{\frac{1}{d}}\right)^d \cdot \left(\frac{n}{n+\varepsilon}\right)^d. && \text{(Simplify)}
\end{aligned}$$

Since $\vec{p}^{(n)}$ belongs to at least k_n-many sets in $\mathcal{A}_n$, this means (by definition of $\mathcal{A}_n$) that there are at least k_n-many sets $Y \in \mathcal{F}_n \subseteq \mathcal{S}_n$ such that $\vec{p}^{(n)} \in A_Y = Y + B^\circ_\varepsilon(\vec{0})$. By Fact C.0.3, for each such Y, we have $Y \cap B^\circ_\varepsilon(\vec{p}^{(n)}) \neq \emptyset$. Also, for each such Y, there is (by definition of $\mathcal{S}_n$) a distinct $X_Y \in \mathcal{P}$ such that $Y \subseteq X_Y$ and thus $X_Y \cap B^\circ_\varepsilon(\vec{0}) \neq \emptyset$

[6] We require this to be an indexed family rather than just a set, because it is possible that there are distinct $Y, Y' \in \mathcal{S}_n$ such that $Y + B^\circ_\varepsilon(\vec{0}) = Y' + B^\circ_\varepsilon(\vec{0})$.

[7] We are taking X in Corollary A.3.4 to be S'_n in this proof, and taking μ to be m and taking $\mathcal{A}$ to be $\mathcal{A}_n$. We have that $\sum_{A\in\mathcal{A}_n} m(A) < \infty$ because $|\mathcal{A}_n| = |\mathcal{F}_n|$ is finite, and each $A \in \mathcal{A}_n$ is a subset of S'_n, so has finite measure.

showing that $X_Y \in \mathcal{N}^{\circ}_{\varepsilon}(\vec{p}^{(n)})$. Since there are at least k_n-many such Y, we have

$$\left|\mathcal{N}^{\circ}_{\varepsilon}(\vec{p}^{(n)})\right| \geq k_n \geq \left(1+\varepsilon\left(\tfrac{v}{M}\right)^{\frac{1}{d}}\right)^d \cdot \left(\frac{n}{n+\varepsilon}\right)^d.$$

For the last step of the proof, we perform a limiting process on n. By Fact G.0.1, let $\gamma \in \mathbb{R}$ such that $\gamma < \left(1+\varepsilon\left(\tfrac{v}{M}\right)^{\frac{1}{d}}\right)^d$ and $\lceil\gamma\rceil = \left\lceil\left(1+\varepsilon\left(\tfrac{v}{M}\right)^{\frac{1}{d}}\right)^d\right\rceil$. Then, because

$$\lim_{n\to\infty}\left(1+\varepsilon\left(\tfrac{v}{M}\right)^{\frac{1}{d}}\right)^d \cdot \left(\frac{n}{n+\varepsilon}\right)^d = \left(1+\varepsilon\left(\tfrac{v}{M}\right)^{\frac{1}{d}}\right)^d > \gamma,$$

we can take $N \in \mathbb{N}$ sufficiently large so that

$$\left(1+\varepsilon\left(\tfrac{v}{M}\right)^{\frac{1}{d}}\right)^d \cdot \left(\frac{N}{N+\varepsilon}\right)^d > \gamma.$$

Then considering the point $\vec{p}^{(N)}$, we have $\left|\mathcal{N}^{\circ}_{\varepsilon}(\vec{p}^{(N)})\right| \geq \gamma$, and since the cardinality is either infinite or an integer, we can take the ceiling on the right hand side, so by the choice of γ, we have

$$\left|\mathcal{N}^{\circ}_{\varepsilon}(\vec{p}^{(N)})\right| \geq \lceil\gamma\rceil = \left\lceil\left(1+\varepsilon\left(\tfrac{v}{M}\right)^{\frac{1}{d}}\right)^d\right\rceil$$

which completes the proof. □

Remark 7.1.6. In the proof above, we handled the case where infinitely members of $\mathcal{P}$ intersect some S_n, but this in fact wasn't necessary as this case is even more of a triviality than discussed in the proof. Suppose there is some bounded set $S \subseteq \mathbb{R}^d$ which intersects infinitely many members of $\mathcal{P}$. Then $\overline{S}$ is also bounded because the diameter is no larger and also intersects infinitely many members of $\mathcal{P}$. Consider the standard open cover $\{{}^{\|\cdot\|}B^{\circ}_{\varepsilon}(\vec{x})\}_{\vec{x}\in\overline{S}}$ of $\overline{S}$. Since $\overline{S}$ is closed and bounded, by the Heine-

Borel theorem there is a finite subcover $\mathcal{C} \subseteq \{^{\|\cdot\|}B^\circ_\varepsilon(\vec{x})\}_{\vec{x}\in\overline{S}}$ of $\overline{S}$. Since $\overline{S}$ intersects infinitely many sets in $\mathcal{P}$, the fact that $\mathcal{C}$ has finite cardinality implies one of the ε-balls in $\mathcal{C}$ must intersect infinitely many members of $\mathcal{P}$. Thus, for each $\varepsilon \in (0,\infty)$, there is in fact some $\vec{p} \in \mathbb{R}^d$ such that $^{\|\cdot\|}B^\circ_\varepsilon(\vec{p})$ intersects infinitely many members of $\mathcal{P}$. Thus, the only interesting partitions to deal with are those for which every bounded subset of $\mathbb{R}^d$ intersects only finitely many members. △

7.1.2 Upper Bounds on ε

For the most part, the content of this subsection is trivialities. We will write down statements which are immediate corollaries of our prior results, but we do so in order to make sure that the conversion of the results above into statements about limitations of (k,ε)-secluded partitions is very clear.

To begin with, the following two statements are trivial corollaries to Theorem 7.1.1 and Corollary 7.1.3 respectively.

Corollary 7.1.7. *Let $d \in \mathbb{N}$ and $\varepsilon \in (0,\infty)$ and $\|\cdot\|$ a norm on $\mathbb{R}^d$ and $\mathcal{P}$ a partition of $\mathbb{R}^d$ such that every member has diameter at most 1 (with respect to $\|\cdot\|$). Then there exists $\vec{p} \in \mathbb{R}^d$ such that $^{\|\cdot\|}B^\circ_\varepsilon(\vec{p})$ intersects at least $(1+2\varepsilon)^d$ members of $\mathcal{P}$.*

Proof. This is just the statement of Corollary 7.1.3 with $D = 1$. □

Corollary 7.1.8. *Let $d \in \mathbb{N}$ and $\varepsilon \in (0,\infty)$ and $\mathcal{P}$ a partition of $\mathbb{R}^d$ such that every member has Lebesgue outer measure at most 1. Then there exists $\vec{p} \in \mathbb{R}^d$ such that $^{\infty}B^\circ_\varepsilon(\vec{p})$ intersects at least $(1+2\varepsilon)^d$ members of $\mathcal{P}$.*

Proof. For the ℓ_∞ norm, $v_{\ell_\infty,d} = 2^d$, so by Corollary 7.1.3 with $M = 1$ and $v_{\|\cdot\|,d} = 2^d$ we get the result. □

Now recall the following fact presented earlier.

Fact G.0.3. *For $\varepsilon \in [0, \frac{1}{2}]$ it holds that $\log_4(1+2\varepsilon) \geq \varepsilon$.*

One can note that equality holds for $\varepsilon = 0$ and $\varepsilon = 1$. This inequality allows us to state the two optimality theorems that we really want. They are corollaries to the above, but we will label them as theorems for emphasis. Recall that we wish for (k, ε)-secluded partitions to have a small value of k, so we almost certainly don't want k to be larger than 2^d, so including the assumption that it is not in the result below is not a very strong hypothesis. The proof mirrors the proof of Theorem 5.2.5 (Near Optimality of $\varepsilon = \frac{1}{2d}$ for Unit Cube Partitions).

Theorem 7.1.9 (Near Optimality of ε)**.** *Let $d, k \in \mathbb{N}$ with $k \leq 2^d$, $\varepsilon \in (0, \infty)$, and $\mathcal{P}$ a (k, ε)-secluded partition of $\mathbb{R}^d$. If every member of $\mathcal{P}$ has Lebesgue outer measure at most 1, then $\varepsilon \leq \frac{\log_4(k)}{d}$. In particular, if $\mathcal{P}$ is $(d+1, \varepsilon)$-secluded, then $\varepsilon \leq \frac{\log_4(d+1)}{d}$.*

Proof. Recall that by definition of a (k, ε)-secluded partition, for every point $\vec{x} \in \mathbb{R}^d$, it must be the case that ${}^{\infty}\overline{B}_{\varepsilon}(\vec{x})$ intersects at most k members of $\mathcal{P}$ (note that we defined this specifically for the ℓ_∞ ball). By Corollary 7.1.8, there is a point $\vec{p}$ such that ${}^{\infty}B^{\circ}_{\varepsilon}(\vec{p})$ intersects at least $(1+2\varepsilon)^d$ members of $\mathcal{P}$, so the closed ball ${}^{\infty}\overline{B}_{\varepsilon}(\vec{p})$ trivially intersects at least $(1+2\varepsilon)^d$ members of $\mathcal{P}$. Thus, we have $k \geq (1+2\varepsilon)^d$.

Because $k \leq 2^d$ by hypothesis, this implies $\varepsilon \leq \frac{1}{2}$, so by Fact G.0.3 we have

$$\log_4(k) \geq \log_4\left((1+2\varepsilon)^d\right) = d \log_4(1+2\varepsilon) \geq d\varepsilon$$

so solving for ε, we have $\varepsilon \leq \frac{\log_4(k)}{d}$. Also, for each $d \in \mathbb{N}$ it holds that $d + 1 \leq 2^d$ which shows the "in particular" part of the statement. □

For completeness and clarity, we will state the same result for ℓ_∞ unit diameter partitions. This either follows by using the proof outlined above and replacing the reference to Corollary 7.1.8 with a reference to Corollary 7.1.7, or it follows because a set with ℓ_∞ diameter at most 1 also has Lebesgue outer measure at most 1 (see Fact 3.4.9) so it is actually an immediate and direct corollary of Theorem 7.1.9.

Corollary 7.1.10 (Near Optimality of ε)**.** *Let $d, k \in \mathbb{N}$ with $k \leq 2^d$, $\varepsilon \in (0, \infty)$, and $\mathcal{P}$ a (k, ε)-secluded partition of $\mathbb{R}^d$. If every member of $\mathcal{P}$ has ℓ_∞ diameter at most 1, then $\varepsilon \leq \frac{\log_4(k)}{d}$. In particular, if $\mathcal{P}$ is $(d+1, \varepsilon)$-secluded, then $\varepsilon \leq \frac{\log_4(d+1)}{d}$.*

Proof. See discussion in the paragraph preceding the statement. □

We finish this section by stating an asymptotic corollary when $k(d)$ is polynomial.

Corollary 7.1.11 (Asymptotics of k and ε)**.** *Let $k : \mathbb{N} \to \mathbb{N}$ and $\varepsilon : \mathbb{N} \to (0, \infty)$ and let $\langle \mathcal{P}_d \rangle_{d=1}^\infty$ be a sequence of $(k(d), \varepsilon(d))$-secluded partitions of $\mathbb{R}^d$ such that every member of each $\mathcal{P}_d$ has outer Lebesgue measure at most 1 (it suffices if each member has ℓ_∞ diameter at most 1). If $k(d) \in \mathsf{poly}(d)$ then $\varepsilon(d) \in O(\frac{\log_4(d)}{d})$ (where the hidden constant can be taken to be anything exceeding the polynomial degree of k).*

Proof. Since $k(d) \in \mathsf{poly}(d)$, then there are constants C, n such that for sufficiently large d, we have $k(d) \leq Cd^n$ which for sufficiently large d is less than 2^d so by Theorem 7.1.9, for sufficiently large d we have

$$\varepsilon(d) \leq \frac{\log_4(k(d))}{d} \leq \frac{n \log_4(Cd)}{d} \in O\left(\frac{\log(d)}{d}\right).$$

More specifically, for any $n' > n$ we have for large enough d that $(n'-n)\log_4(d) \geq n\log_4(C)$, so for large enough d we have

$$\varepsilon(d) \leq \frac{n\log_4(Cd)}{d} = \frac{n[\log_4(C)+\log_4(d)]}{d} \leq \frac{(n'-n)\log_4(d)+n\log_4(d)}{d} = \frac{n'\log_4(d)}{d}$$

showing that the hidden constant can be taken to be any n' larger than the degree n of k. $\square$

7.1.3 A Short Discussion of Specific Norms

While we won't do an elaborate analysis with any specific norms other than ℓ_∞, we will at least mention how Theorem 7.1.1 evaluates when we consider some of the most common norms (ℓ_1, ℓ_2, and ℓ_∞), though we again emphasize that we believe the perspective of Corollary 7.1.2 is more useful. The information is summarized in Table 7.1. The main observation that we want to make is that when using ℓ_p norms other than ℓ_∞, the factor being multiplied by $\frac{\varepsilon}{M^{1/d}}$ is no longer a constant function of the dimension d, but instead a decreasing function of the dimension. This should not be too surprising since for ℓ_p norms which are not ℓ_∞, the unit ℓ_p-ball is a subset of the unit ℓ_∞ ball, so the measure will be smaller and actually tend to 0 as d tends to ∞, so we should expect to not be able to intersect as many members of the partition with this ball.

The ℓ_∞ norm For the ℓ_∞ norm, we have $v_{\ell_\infty,d} = 2^d$ so the stated bound of $\left(1+\varepsilon\left(\frac{v_{\|\cdot\|,d}}{M}\right)^{1/d}\right)^d$ is $\left(1+\varepsilon\frac{2}{M^{1/d}}\right)^d$.

The ℓ_2 norm For the ℓ_2 norm, it is well-known that $v_{\ell_2,d} = \frac{\pi^{d/2}}{\Gamma\left(\frac{d}{2}+1\right)}$ where Γ denotes the gamma-function which generalizes the factorial (specifically, for natural numbers n, $\Gamma(n+1) = n!$). Using Stirling's approximation and the approximation $\Gamma\left(\frac{d}{2}+1\right) \approx$

$\left\lceil \frac{d}{2} \right\rceil !$, we get

$$\left(\Gamma\left(\frac{d}{2}+1\right)\right)^{1/d} \approx \left(\left\lceil \frac{d}{2} \right\rceil !\right)^{1/d}$$
$$\approx \left(\sqrt{2\pi\left\lceil \frac{d}{2} \right\rceil}\left(\frac{\left\lceil \frac{d}{2} \right\rceil}{e}\right)^{\left\lceil \frac{d}{2} \right\rceil}\right)^{1/d}$$
$$\approx \left(\left(\frac{\left\lceil \frac{d}{2} \right\rceil}{e}\right)^{\left\lceil \frac{d}{2} \right\rceil}\right)^{1/d}$$
$$\approx \sqrt{\frac{d}{2e}}$$

leading to the approximate bound $\left(1+\frac{\varepsilon\sqrt{\pi}}{\left(M\cdot\Gamma\left(\frac{d}{2}+1\right)\right)^{1/d}}\right)^d \approx \left(1+\varepsilon\frac{\sqrt{2\pi e}}{M^{1/d}\cdot\sqrt{d}}\right)^d$.

The ℓ_1 norm For the ℓ_1 norm, the volume can be be obtained from [Wan05] (where the measure of the ℓ_p ball in general is discussed), or one can recognize that the ℓ_1 unit ball is (disregarding boundaries) a disjoint union of 2^d-many copies of the standard simplex—one in each orthant, and each simplex has measure $\frac{1}{d!}$ giving the unit ball a total volume of $v_{\ell_1,d} = \frac{2^d}{d!}$. We can use Stirling's approximation to say that $(d!)^{1/d} \approx \left(\sqrt{2\pi d}\left(\frac{d}{e}\right)^d\right)^{1/d} \approx \frac{d}{e}$ to obtain the approximate bound $\left(1+\frac{2\varepsilon}{(M\cdot(d!))^{1/d}}\right)^d \approx \left(1+\varepsilon\frac{2e}{M^{1/d}\cdot d}\right)^d$.

Norm	Lower Bound	
$\|\cdot\|$	$\left(1+\varepsilon\left(\frac{v_{\|\cdot\|,d}}{M}\right)^{1/d}\right)^d$	
ℓ_∞	$\left(1+\varepsilon\frac{2}{M^{1/d}}\right)^d$	$=\left(1+\frac{\varepsilon}{M^{1/d}}\cdot 2\right)^d$
ℓ_2	$\left(1+\frac{\varepsilon\sqrt{\pi}}{\left(M\cdot\Gamma\left(\frac{d}{2}+1\right)\right)^{1/d}}\right)^d$	$\approx\left(1+\frac{\varepsilon}{M^{1/d}}\cdot\sqrt{\frac{2\pi e}{d}}\right)^d$
ℓ_1	$\left(1+\frac{2\varepsilon}{(M\cdot(d!))^{1/d}}\right)^d$	$\approx\left(1+\frac{\varepsilon}{M^{1/d}}\cdot\frac{2e}{d}\right)^d$

Table 7.1: Lower bounds evaluated for some common norms

7.2 Upper Bound on ε via the Dissection Number of the Cube

In this section, we prove an upper bound on ε when $k = d+1$ for the class of partitions with members of ℓ_∞ diameter at most 1 (in this way, they are on the same footing as axis-aligned unit cube partitions). Though the bounds of Section 7.1 (Upper Bounds on ε via Brunn-Minkowski and Blichfeldt) are much better than the one given in this section and can be applied in far more general circumstances, the bound in this section remains relevant for the following reason. The bound in this section is given in terms of a property of $[0,1]^d$ which we call the "Sperner number of the d-cube," which we can lower bound by the previously studied property called the "dissection number of the d-cube". Using known lower bounds on the dissection number, we can obtain bound of $\varepsilon \leq \frac{2}{\sqrt{d}}$ for each dimension d when $k = d+1$. If the dissection number lower bounds can be improved enough (which is consistent with what is currently known) it would decrease this bound on ε to $\varepsilon \leq \frac{C}{d}$ for some universal constant C when $k = d+1$. This would show that our reclusive partition constructions (Theorem 4.2.18) are optimal up to a constant factor even for the broad class of unit ℓ_∞ diameter bounded partitions (which we conjecture to be true regardless of whether sufficient improvements to the dissection number lower bounds are possible (see Conjecture 7.3.2).

The method for obtaining the bound in this section is similar to the methods used in Chapter 6 (Optimality of k in General)—namely the use of Sperner's lemma and KKM lemma type results. However, unlike in Chapter 6, we will not use the standard or even cubical variants of these lemmas, but will use a quite strong polytopal variant due to De Loera, Peterson, and Su [LPS01] which guarantees the existence of not just one fully colored simplex as in Sperner's lemma, but guarantees the existence of many of them and guarantees a certain structural property that the collection has.

Combining this with the work of Glazyrin [Gla12] we obtain an improvement over the results stated in [LPS01] when we focus on the cube specifically. Using a simple volume argument (much simpler than the volume arguments in Section 7.1) we obtain a lower bound on ε (when $k = d+1$) in terms of the dissection number of the cube.

This section will be laid out as follows. First, in Subsection 7.2.1, we provide the necessary background on Sperner colorings, simplicial decompositions, cube covers, cube dissections, and cube triangulations to be able to present the result of [LPS01] and demonstrate that the statement can be improved for the particular case of the cube by using [Gla12]. Then, in Subsection 7.2.2 we use this result with a volume argument to get an upper bound on ε in terms of the Sperner number of the d-cube or the dissection number of the d-cube. Then, in Subsection 7.2.3 we provide the numeric bounds obtained by using the dissection number lower bound in [Gla12] and also show that in $\mathbb{R}^1$ and $\mathbb{R}^2$, the reclusive partitions (Theorem 4.2.18) achieve the optimal value of ε for unit ℓ_∞ diameter bounded partitions when $k = d+1$. Finally, in Subsection 7.2.4 we discuss how the numeric bound could be improved based on what is currently known about the dissection number of the cube.

7.2.1 Background on Various Simplicial Decompositions

We will need the definition of a simplicial subdivision.

Definition 7.2.1 (Simplicial Subdivision). A simplicial subdivision (sometimes called a triangulation) of $[0,1]^d$ is a finite set of d-simplices $\Sigma = \left\{\sigma^{(1)}, \ldots, \sigma^{(m)}\right\}$ (for some $m \in \mathbb{N}$) such that $[0,1]^d = \bigcup_{j=1}^m \sigma^{(j)}$, and for distinct $j, j' \in [m]$, $\sigma^{(j)} \cap \sigma^{(j')}$ is either empty, or a face of both $\sigma^{(j)}$ and of $\sigma^{(j')}$. The vertices of a simplicial subdivision are by definition the vertices of all of its d-simplices; that is, $V(\Sigma) \stackrel{\text{def}}{=} \bigcup_{j=1}^m V(\sigma^{(j)})$.

Next, we need the definition of a Sperner coloring of a simplicial subdivision. We will ultimately be interested in continuous colorings of $[0,1]^d$, but we need the discretized colorings first in order to properly utilize the results of [LPS01]. In the definition below, a face F of the cube $[0,1]^d$ is a product set $F = \prod_{i=1}^{d} F_i$ where each F_i is one of three sets: $\{0\}$, $\{1\}$, or $[0,1]$.

Definition 7.2.2 (Sperner Coloring). Let $d, m \in \mathbb{N}$ and $V([0,1]^d) \stackrel{\text{def}}{=} \{0,1\}^d$ denote the set of vertices of $[0,1]^d$ (which is identified with the colors that will be used). Let $\Sigma = \left\{\sigma^{(1)}, \ldots, \sigma^{(m)}\right\}$ be a simplicial subdivision of $[0,1]^d$. Let $\chi : V(\Sigma) \to V([0,1]^d)$ be a 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 defining the face F). Then the pair (Σ, χ) is called a *Sperner coloring* of $[0,1]^d$.

Furthermore, for any $\sigma \in \Sigma$, we define $\operatorname{colorset}(\sigma) \stackrel{\text{def}}{=} \{\chi(\vec{x}) : \vec{x} \in V(\sigma)\}$ which is the set of colors appearing on the vertices of the simplex σ.

The above definitions are generalized to convex polytopes in the natural way (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 [LPS01] that for any convex polytope P in $\mathbb{R}^d$ with n vertices, any Sperner coloring (Σ, χ) of P will have at least $(n-d)$-many "fully colored simplices". That is, if $\Sigma = \left\{\sigma^{(1)}, \ldots, \sigma^{(m)}\right\}$, then there must be least $(n-d)$-many simplices $\sigma^{(j_1)}, \ldots, \sigma^{(j_{n-d})} \in \Sigma$ for which $\left|\operatorname{colorset}(\sigma^{(j_i)})\right| = d+1$. Recall that each $\sigma \in \Sigma$ only has $d+1$ vertices, so this means that each vertex of $\sigma^{(j_i)}$ is assigned a different color—i.e. $\sigma^{(j_i)}$ is a "fully colored simplex". This is a generalization of Sperner's lemma to arbitrary convex polytopes, and Sperner's lemma is recovered when the polytope P is itself a simplex because then the number of vertices n is equal to $d+1$ (one more than the dimension), so this guarantees $(n-d) = ((d+1)-d) = 1$ fully colored simplex in the coloring. Additionally, [LPS01] proved that for distinct $\sigma^{(j_i)}$ and $\sigma^{(j_{i'})}$ that $\text{colorset}(\sigma^{(j_i)}) \neq \text{colorset}(\sigma^{(j_{i'})})$; In other words, not only are there $(n-d)$-many fully colored simplices, but there are at least $(n-d)$-many fully colored simplices which each represent a distinct set of $(d+1)$-many colors. From another perspective, [LPS01] says 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 the value $(n-d)$ is tight when P can be an arbitrary convex polytope, but that improvements may be made when restricting to specific types of convex polytopes (e.g. cubes). In fact, their bound can be improved significantly when restricting to cubes by combining the results of [LPS01] with other results about simplicial decompositions of cubes. Next, we give a definitional name to the best possible parameter for each dimension d. In other words, though the value $(n-d)$ is the best possible guarantee for the class of all convex n-vertex polytopes in $\mathbb{R}^d$, this is not the best possible guarantee for the $n = 2^d$-vertex cube $[0,1]^d$ in $\mathbb{R}^d$, but since we don't know what that best possible guarantee is, we will give it a name so that we can still work with it.

While we could get away without defining this value explicitly and instead just use bounds on this value (since that is all we will use anyway), we do want to make this quantity explicit since it relates nicely to some other areas of research regarding the d-dimension cube.

Definition 7.2.3 (Sperner Number of the Cube). Let $d \in \mathbb{N}$ 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 $[0,1]^d$, there exists at least s-many $(d+1)$-cardinality sets $J^{(1)}, \ldots, J^{(s)} \subseteq V([0,1]^d)$ such for each $J^{(i)}$, Σ contains a simplex $\sigma^{(J_i)}$ such that $\text{colorset}(\sigma^{(J_i)}) = J^{(i)}$.

We call the minimum value of S_d the *Sperner number of the d-cube* or more concisely the *dth Sperner number*, and we denote it by $\text{Sperner}(d) \stackrel{\text{def}}{=} \min(S_d)$.

Because $[0,1]^d$ is a convex polytope in d dimensions with $n = 2^d$ vertices, it follows immediately from [LPS01] that $\text{Sperner}(d) \geq 2^d - d$. However, they implicitly showed something much stronger than this. By [LPS01, Thm. 1] and the comment following the proof of [LPS01, 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 [Gla12], a "dissection" of $[0,1]^d$).

Said another way, for any Sperner coloring (Σ, χ) of $[0,1]^d$, there is a collection $J^{(1)}, \ldots, J^{(t)}$ of subsets of $V([0,1]^d)$ each of cardinality $d+1$ such that collection of simplices[8] $\text{conv}\, J^{(1)}, \ldots, \text{conv}(J^{(t)})$ forms a "face-to-face simplicial cover" of $[0,1]^d$ (what [Gla12] calls a "dissection" of $[0,1]^d$) and such that for each $J^{(j)}$, there is some simplex $\sigma^{(J^{(j)})} \in \Sigma$ with $\text{colorset}(\sigma^{(J^{(j)})}) = J^{(j)}$.

Thus, $\text{Sperner}(d)$ is at least as large as the number of simplices needed in a dissection of $[0,1]^d$ because by [LPS01] there is always guaranteed to be enough sets of $(d+1)$-many colors (all of which have associated fully colored simplices) to generate a dissection of $[0,1]^d$. In other words, the dissection number of the

[8] $J^{(j)}$ is a set of $d+1$ vertices of $[0,1]^d$, so the naturally associated simplex is the convex hull of these $d+1$ vertices: $\text{conv}(J^{(j)})$.

d-dimensional cube gives a lower bound on the Sperner number.

We can also get a trivial upper bound on Sperner(d) by noting that, by definition, it can be no larger than the size of the minimum cardinality simplicial subdivision of $[0,1]^d$ because we can consider any Sperner coloring (Σ, χ) in which Σ is a minimum cardinality simplicial subdivision, so we cannot guarantee more than $|\Sigma|$ fully colored simplices because there are only a total of $|\Sigma|$ simplices in the subdivision. Similarly, Sperner(d) can be no larger than the size of the minimum cardinality triangulation[9] of $[0,1]^d$—this is because any triangulation is a valid simplicial subdivision, so the same argument applies.

Summarizing the above paragraphs using the notation in [Gla12], 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 $[0,1]^d$ (i.e. no extra vertices are allowed). For example, Below, Brehm, De Loera, and Richter-Gebert showed in [BBDLa00] that the use of extra vertices can drastically reduce the necessary size of a simplicial subdivision.

$$\mathrm{cover}(d) \leq \mathrm{dis}(d) \leq \mathrm{Sperner}(d) \leq \mathrm{triang}(d) \tag{7.1}$$

Glazyrin showed in [Gla12] that $(d+1)^{\frac{d-1}{2}} \leq \mathrm{dis}(d)$ and Orden and Santos showed in [OS03] that $\mathrm{triang}(d) \in O(0.816^d d!)$ which are the best known asymptotic bounds to date, and they provide upper and lower[10] 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 [BBDLa00]), for cubes it is

[9] We use the term triangulation as in [Gla12], which is different from how it is used in [LPS01]. 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 vertices of the cube).

[10] For $d < 5$, the lower bound of $\mathrm{Sperner}(d) \geq 2^d - d$ can be used instead of $\mathrm{Sperner}(d) \geq (d+1)^{\frac{d-1}{2}}$, however, noting that 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.

still an open question if dis(d) equals triang(d) or not. We provide the first few values of Sperner(d) which are exactly the values of d where triang(d) is know to equal dis(d) (see [Gla12, Table 1]).

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

Table 7.2: Known values of Sperner(d).

We are not particularly interested in colorings of simplicial subdivisions, though. Rather, we are interested in colorings of the entire cube which have a similar property to Sperner colorings of simplicial subdivisions. We will call such colorings Sperner/KKM colorings since the relate not only to Sperner's lemma, but also to the KKM lemma.

Definition 7.2.4 (Sperner/KKM Coloring). Let $d \in \mathbb{N}$ and $V([0,1]^d) \stackrel{\text{def}}{=} \{0,1\}^d$. Let $\chi : [0,1]^d \to V([0,1]^d)$ 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 defining the face F). Such a function χ will be called a *Sperner/KKM coloring*.

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: consider a sequence of simplicial subdivisions with increasingly small simplices and use the finiteness of the number of possible sets of $(d+1)$ colors that could be used along with compactness and the Balzona-Weirstrass theorem to obtain limit points.

The lemma says that we can find at least $\text{Sperner}(d)$ different sets of colors (each set having cardinality $d+1$), and for each one of these sets, we can find a point in the closure of all $d+1$ of those colors.

Note that for a color/vertex v, the set of points assigned color v by a coloring function χ is the set $\chi^{-1}(v)$.

Lemma 7.2.5 (Mutli-Point Cubical KKM Lemma). *Let $d \in \mathbb{N}$ and $V([0,1]^d) \stackrel{\text{def}}{=} \{0,1\}^d$. In any Sperner/KKM coloring $\chi : [0,1]^d \to V([0,1]^d)$, there are at least $\text{Sperner}(d)$-many distinct $(d+1)$-cardinality sets $J^{(1)}, \ldots, J^{(\text{Sperner}(d))} \subseteq V([0,1]^d)$ such that for each $j \in [\text{Sperner}(d)]$, there is a point common to the closures of all colors in $J^{(j)}$ (formally, $\bigcap_{\vec{v} \in J^{(j)}} \overline{\chi^{-1}(\vec{v})} \neq \emptyset$).*

Proof Outline. The proof is identical to the proof of [LPS01, 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 $\text{Sperner}(d)$ different colorsets. □

A consequence of this is a similarly stated result for partitions of the cube.

Corollary 7.2.6 (Mutli-Point Cubical KKM Lemma for Partitions)**.** *Let $d \in \mathbb{N}$ and $\mathcal{S}$ a partition of $[0,1]^d$ such that for every $Y \in \mathcal{S}$, 1 is a strict pairwise bound*[a] *for Y with respect to the ℓ_∞ norm, and let $\delta \in (0,\infty)$. Then there are at least* $\text{Sperner}(d)$*-many distinct $(d+1)$-cardinality sets $\mathcal{F}^{(1)},\ldots,\mathcal{F}^{(\text{Sperner}(d))} \subseteq \mathcal{S}$ and corresponding points $\vec{s}^{(1)},\ldots,\vec{s}^{(\text{Sperner}(d))} \in [0,1]^d$ such that for each $j \in [\text{Sperner}(d)]$, the ball ${}^{\infty}\overline{B}_\delta(\vec{s}^{(j)})$ intersects all members in $\mathcal{F}^{(j)}$. That is,*

$$\mathcal{F}^{(j)} \subseteq {}^{\infty}\overline{\mathcal{N}}_\delta(\vec{s}^{(j)}).$$

Furthermore, if $\mathcal{S}$ has finite cardinality, then the $\mathcal{F}^{(j)}$'s and $\vec{s}^{(j)}$'s do not depend on δ, and in fact $\vec{s}^{(j)}$ is at the closure of every member in $\mathcal{F}^{(j)}$. That is,

$$\mathcal{F}^{(j)} \subseteq \mathcal{N}_{\overline{0}}(\vec{s}^{(j)}).$$

[a] See Definition 6.0.1. It is sufficient, but not necessary, that $\text{diam}_\infty Y < 1$.

Remark 7.2.7. Corollary 7.2.6 does *not* claim that the points $\vec{s}^{(1)},\ldots,\vec{s}^{(\text{Sperner}(d))} \in [0,1]^d$ are distinct, only that the $\mathcal{F}^{(1)},\ldots,\mathcal{F}^{(\text{Sperner}(d))} \subseteq \mathcal{S}$ are distinct. For example, it might be that $\vec{s}^{(1)} = \vec{s}^{(2)} = \vec{s}^{(3)}$ which would mean that every ball around this point intersects the members in $\mathcal{F}^{(1)}$ and $\mathcal{F}^{(2)}$ and $\mathcal{F}^{(3)}$. △

We first give a sketch of the proof and then the complete proof.

Proof Sketch. The basic idea is fairly simple (though this gets quickly lost in the notation of the full proof). Define a Sperner/KKM coloring function $\chi : [0,1]^d \to \{0,1\}^d$ which behaves nicely with the partition $\mathcal{S}$ (i.e. for any member $Y \in \mathcal{S}$ and any two points $\vec{y},\vec{y}' \in Y$ it holds that $\chi(\vec{y}) = \chi(\vec{y}')$). This essentially collapses all of the (possibly infinitely many) members of $\mathcal{S}$ into 2^d color classes where each color class is a union of members of $\mathcal{S}$. Then apply Lemma 7.2.5 (Mutli-Point Cubical

KKM Lemma) to obtain Sperner(d)-many sets of colors $J^{(1)}, \ldots, J^{(\text{Sperner}(d))}$ (each of cardinality $d+1$) and corresponding points $\vec{s}^{(1)}, \ldots, \vec{s}^{(\text{Sperner}(d))}$ in each intersection. Because of how the coloring χ will be defined, any open set containing $\vec{s}^{(j)}$ will intersect all $d+1$ colors in $J^{(j)}$ and thus it intersects at least $d+1$ members of $\mathcal{S}$. Furthermore, because χ behaves nicely with $\mathcal{S}$, the Sperner(d)-many sets of $d+1$ members are each distinct sets of $d+1$ members. The reason that we may not obtain quite the same closure properties is that $\mathcal{S}$ might contain infinitely many members, and so this point, though at the closure of $d+1$ colors, may not be at the closure of any member of $\mathcal{S}$. □

Proof. Throughout the proof, let $V = \{0,1\}^d$ which denotes the vertices of $[0,1]^d$ as well as the set of colors that will be used for a Sperner/KKM coloring that we will define. For each $i \in [d]$, define $f_i : \mathcal{S} \to \{0,1\}$ by

$$f_i(Y) = \begin{cases} 0 & 0 \in \pi_i(Y) \\ 1 & 1 \in \pi_i(Y) \\ 0 & \text{otherwise} \end{cases} .$$

This is well-defined because the first two cases are mutually exclusive[11]. Then define $f : \mathcal{S} \to \{0,1\}^d$ by $f(Y) = \langle f_i(Y) \rangle_{i=1}^d$. Finally, define a coloring function $\chi : [0,1]^d \to \{0,1\}^d$ by $\chi(\vec{y}) = f(\text{member}_{\mathcal{S}}(\vec{y}))$.

Claim A. *The function χ is a Sperner/KKM coloring of $[0,1]^d$.*

Proof of Claim. Let F be any face of $[0,1]^d$ (so 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

[11]This is because 1 is a strict pairwise bound for Y with respect to the ℓ_∞ norm which by definition means that Y does not contain a pair of points which are ℓ_∞ distance 1 or more apart, so it cannot be that there are points $\vec{y}, \vec{y}\,' \in Y$ with $y_i = 0$ and $y'_i = 1$—i.e. it is not the case that $0 \in \pi_i(Y)$ and $1 \in \pi_i(Y)$.

$\vec{v} = \chi(\vec{x}) = f(\text{member}_{\mathcal{S}}(\vec{x}))$; we must show that $\vec{v} \in F$ which we do by showing that $v_i \in F_i$ for each $i \in [d]$ (i.e. showing $f_i(\text{member}_{\mathcal{S}}(\vec{x})) \in F_i$.)

There are three cases. If $F_i = \{0\}$, then $x_i = 0$ (because $\vec{x} \in F$); also $\vec{x} \in \text{member}_{\mathcal{S}}(\vec{x})$, so $0 = x_i = \pi_i(\vec{x}) \in \pi_i(\text{member}_{\mathcal{S}}(\vec{x}))$, in which case $v_i = f_i(\text{member}_{\mathcal{S}}(\vec{x})) = 0$ by definition of f_i, so $v_i \in F_i$. The case $F_i = \{1\}$ is analogous. The remaining case is that $F_i = [0,1]$, but in this case $v_i \in F_i$ trivially since v_i is either 0 or 1. ■

By Claim A and Lemma 7.2.5 (Mutli-Point Cubical KKM Lemma), there are at least Sperner(d)-many distinct $(d+1)$-cardinality sets $J^{(1)}, \ldots, J^{(\text{Sperner}(d))} \subseteq \{0,1\}^d$ and corresponding points $\vec{s}^{(1)}, \ldots, \vec{s}^{(\text{Sperner}(d))} \in [0,1]^d$ such that for each $j \in [\text{Sperner}(d)]$ we have $\vec{s}^{(j)} \in \bigcap_{\vec{v} \in J^{(j)}} \overline{\chi^{-1}(v)}$.

So far, we have used the notation $\text{member}_{\mathcal{S}}(\cdot)$ primarily as notation only, but really this is a function

$$\text{member}_{\mathcal{S}} : [0,1]^d \to \mathcal{S}$$

so we can consider preimages with respect to this membership function. In particular, for any member $Y \in \mathcal{S}$ we have $\text{member}_{\mathcal{S}}^{-1}(Y) = Y$.

Note that for each $\vec{v} \in V$, we have

$$
\begin{aligned}
.\chi^{-1}(\vec{v}) &= (f \circ \text{member}_{\mathcal{S}})^{-1}(\vec{v}) && \text{(Def'n of } \chi\text{)} \\
&= \text{member}_{\mathcal{S}}^{-1}(f^{-1}(\vec{v})) && \text{(Preimages compose nicely)} \\
&= \bigcup_{Y \in f^{-1}(\vec{v})} \text{member}_{\mathcal{S}}^{-1}(Y) && \text{(Def'n of a preimage)} \\
&= \bigcup_{Y \in f^{-1}(\vec{v})} Y && \text{(Comment above)}
\end{aligned}
$$

Now, for each $j \in [\text{Sperner}(d)]$, and $\vec{v} \in J^{(j)}$, because $\vec{s}^{(j)}$ belongs to the closure of $\chi^{-1}(\vec{v}) = \bigcup_{Y \in f^{-1}(\vec{v})} Y$, the open ball ${}^{\infty}B^{\circ}_{\delta}(\vec{s}^{(j)})$ intersects $\bigcup_{Y \in f^{-1}(\vec{v})} Y$ and thus intersects some $Y \in f^{-1}(\vec{v}) \subseteq \mathcal{S}$. For each $j \in [\text{Sperner}(d)]$ and $\vec{v} \in J^{(j)}$, let $Y^{(j,\vec{v})}$ denote some such member which is fixed for the remainder of the proof. Then, for each $j \in [\text{Sperner}(d)]$, let $\mathcal{F}^{(j)} = \left\{Y^{(j,\vec{v})} : \vec{v} \in J^{(j)}\right\}$. Thus, for each $j \in [\text{Sperner}(d)]$ we have for every $Y \in \mathcal{F}^{(j)}$ that ${}^{\infty}B^{\circ}_{\delta}(\vec{s}^{(j)})$ intersects Y (i.e. $\mathcal{F}^{(j)} \subseteq {}^{\infty}\overline{\mathcal{N}}_{\delta}(\vec{s}^{(j)})$.)

We still have to show that all the $\mathcal{F}^{(j)}$'s are distinct and have cardinality $d+1$, and to do so we make use of the following three claims which demonstrate that for each $j \in [\text{Sperner}(d)]$, the function f restricted to $\mathcal{F}^{(j)}$ is a bijection between $\mathcal{F}^{(j)}$ and $J^{(j)}$.

First, we claim that when restricting f to $\mathcal{F}^{(j)}$, the set $J^{(j)}$ is a valid codomain.

Claim B. *Fix $j \in [\text{Sperner}(d)]$. For all $Y \in \mathcal{F}^{(j)}$ it holds that $f(Y) \in J^{(j)}$.*

Proof. If $Y \in \mathcal{F}^{(j)}$, then by definition of $\mathcal{F}^{(j)}$, $Y = Y^{(j,\vec{v})}$ for some $\vec{v} \in J^{(j)}$, and by definition of $Y^{(j,\vec{v})}$, we have $Y^{(j,\vec{v})} \in f^{-1}(\vec{v})$ which means $f\left(Y^{(j,\vec{v})}\right) = \vec{v} \in J^{(j)}$. □

Next, we claim that this restriction of f is an injection.

Claim C. *Fix $j \in [\text{Sperner}(d)]$. For distinct $Y, Y' \in \mathcal{F}^{(j)}$ it holds that $f(Y) \neq f(Y')$.*

Proof of Claim. By definition of $\mathcal{F}^{(j)}$, there must exists $\vec{v}, \vec{v}\,' \in J^{(j)}$ such that $Y = Y^{(j,\vec{v})}$ and $Y' = Y^{(j,\vec{v}\,')}$, and $\vec{v}, \vec{v}\,'$ must be distinct otherwise trivially $Y = Y'$. By definition of $Y^{(j,\vec{v})}$, we have $Y^{(j,\vec{v})} \in f^{-1}(\vec{v})$ which means $f\left(Y^{(j,\vec{v})}\right) = \vec{v}$ (and similarly for $\vec{v}\,'$). Thus, $f(Y) = \vec{v} \neq \vec{v}\,' = f(Y')$. ■

Finally, we claim that this restriction of f is also a surjection.

Claim D. *Fix $j \in [\text{Sperner}(d)]$. For each $\vec{v} \in J^{(j)}$, there exists some $Y \in \mathcal{F}^{(j)}$ such that $f(Y) = \vec{v}$.*

Proof. Take $Y = Y^{(j,\vec{v})}$. Then by definition of $\mathcal{F}^{(j)}$, because $\vec{v} \in J^{(j)}$ we have $Y^{(j,\vec{v})} \in \mathcal{F}^{(j)}$. By definition of $Y^{(j,\vec{v})}$, we have $Y^{(j,\vec{v})} \in f^{-1}(\vec{v})$ which means $f\left(Y^{(j,\vec{v})}\right) = \vec{v}$. □

Thus, because for each $j \in [\text{Sperner}(d)]$ the function f is a bijection between $\mathcal{F}^{(j)}$ and $J^{(j)}$ (i.e. the combination of the three claims), it follows that $\left|\mathcal{F}^{(j)}\right| = \left|J^{(j)}\right| = d + 1$. These claims also demonstrates that for distinct $j, j' \in [\text{Sperner}(d)]$ we have $\mathcal{F}^{(j)} \neq \mathcal{F}^{(j')}$ because $\left\{f(Y) : Y \in \mathcal{F}^{(j)}\right\} = J^{(j)}$ (by Claim B and Claim D) and similarly $\left\{f(Y) : Y \in \mathcal{F}^{(j')}\right\} = J^{(j')}$, and if $\mathcal{F}^{(j)}$ and $\mathcal{F}^{(j')}$ where the same set, then these images under f would also be the same set, but they are not because $J^{(j)} \neq J^{(j')}$.

This nearly proves the first part of the result. Currently, the $\mathcal{F}^{(j)}$'s are subsets of $\mathcal{S}$ and not subsets of $\mathcal{P}$, but we for each $\mathcal{F}^{(j)}$ and each $Y \in \mathcal{F}^{(j)}$, we can replace Y with the unique member $X \in \mathcal{P}$ which generated Y in the induced partition (so $X \supseteq Y$). This does not change the cardinality of $\mathcal{F}^{(j)}$ or the distinctness of $\mathcal{F}^{(j)}$ from $\mathcal{F}^{(j')}$ (because there isn't a single $X \in \mathcal{P}$ which generates two distinct $Y, Y' \in \mathcal{S}$) and doesn't change the intersection properties.

We now consider the case where $\mathcal{S}$ has finite cardinality. Let the $J^{(j)}$'s and $\vec{s}^{(j)}$'s as before, but we will redefine the $\mathcal{F}^{(j)}$'s more carefully. Because $|\mathcal{S}|$ is finite, we also have for each color/vertex $\vec{v} \in V$ that the set $f^{-1}(\vec{v}) \subseteq \mathcal{S}$ has finite cardinality. This will allow us to use the topological fact that the finite union of closures is equal to the closure of the finite union. For each $j \in [\text{Sperner}(d)]$, we have

$$
\begin{aligned}
\vec{s}^{(j)} &\in \bigcap_{\vec{v} \in J^{(j)}} \overline{\chi^{-1}(v)} && (J^{(j)} \text{ is from Mutli-Point Cubical KKM Lemma}) \\
&= \bigcap_{\vec{v} \in J^{(j)}} \overline{\bigcup_{Y \in f^{-1}(\vec{v})} Y} && (\text{Prior chain of equalities}) \\
&= \bigcap_{\vec{v} \in J^{(j)}} \bigcup_{Y \in f^{-1}(\vec{v})} \overline{Y} && (|f^{-1}(\vec{v})| < \infty; \text{ topological fact})
\end{aligned}
$$

which shows that for each $\vec{v} \in J^{(j)}$ there is some $Y \in f^{-1}(\vec{v})$ such that $\vec{s}^{(j)} \in \overline{Y}$. For each $j \in [\text{Sperner}(d)]$ and $\vec{v} \in J^{(j)}$, let $Y^{(j,\vec{v})}$ denote one such member of $\mathcal{S}$. Then, defining the $\mathcal{F}^{(j)}$'s as before using these more specific $Y^{(j,\vec{v})}$, we have $\vec{s}^{j} \in \bigcap_{Y \in \mathcal{F}^{(j)}} \overline{Y}$ (i.e. $\mathcal{F}^{(j)} \subseteq \mathcal{N}_{\bar{0}}(\vec{s}^{(j)})$) and by exactly the arguments above the $\mathcal{F}^{(j)}$'s are distinct and each have cardinality $d+1$ and we can again update them to be subsets of $\mathcal{P}$ with the same properties. □

7.2.2 Upper Bound on ε

Using the result above, we can prove the following upper bound on ε in terms of Sperner(d), or dis(d).

Remark 7.2.8. In the statement below, because Sperner(1) = dis(1) = 1, this gives the expression $\varepsilon \leq \frac{1}{0}$ when $d = 1$, and we interpret this under the typical convention as the (trivial) upper bound of $\varepsilon \leq \infty$. △

Theorem 7.2.9 (Dissection Upper Bound on ε for Unit Diameter Partitions)**.** *Let $d \in \mathbb{N}$ and $\varepsilon \in (0,\infty)$ and $\mathcal{P}$ a $(d+1,\varepsilon)$-secluded partition of $\mathbb{R}^d$ such that for all $X \in \mathcal{P}$, $\operatorname{diam}_\infty(X) \leq 1$. Then $\varepsilon \leq \frac{1}{2\left(\text{Sperner}(d)^{1/d}-1\right)}$. In particular, $\varepsilon \leq \frac{1}{2\left(\text{dis}(d)^{1/d}-1\right)}$.*

Proof. We will use a limiting argument, so let $D \in (1,\infty)$ and $\delta \in (0,\varepsilon)$. Consider the cube $[0,D]^d$. Let $\mathcal{S}$ be the partition of $[0,D]^d$ induced by $\mathcal{P}$ (i.e. $\mathcal{S} = \left\{X \cap [0,D]^d : X \in \mathcal{P} \text{ and } X \cap [0,D]^d \neq \emptyset\right\}$). Since members of $\mathcal{P}$ have ℓ_∞ diameter at most 1, members of $\mathcal{S}$ also have ℓ_∞ diameter at most $1 < D$. Thus, by Corollary 7.2.6 (applied to the scaled cube $[0,D]^d$ instead of $[0,1]^d$), there at least Sperner(d)-many distinct $(d+1)$-cardinality sets $\mathcal{F}^{(1)},\ldots,\mathcal{F}^{(\text{Sperner}(d))} \subseteq \mathcal{S}$ and corresponding points $\vec{s}^{(1)},\ldots,\vec{s}^{(\text{Sperner}(d))} \in [0,D]^d$ such that for each

$j \in [\mathrm{Sperner}(d)]$, the ball ${}^{\infty}\overline{B}_{\delta}(\vec{s}^{(j)})$ intersects all members in $\mathcal{F}^{(j)}$ (i.e. $\mathcal{F}^{(j)} \subseteq {}^{\infty}\overline{\mathcal{N}}_{\delta}(\vec{s}^{(j)})$).

Claim A. *For distinct* $j, j' \in [\mathrm{Sperner}(d)]$*. It must be that* $\left\|\vec{s}^{(j)} - \vec{s}^{(j')}\right\|_{\infty} \geq 2(\varepsilon - \delta)$.

Proof of Claim. Let $\alpha = \left\|\vec{s}^{(j)} - \vec{s}^{(j')}\right\|_{\infty}$ denote this distance. Let $\vec{c}$ be the midpoint between $\vec{s}^{(j)}$ and $\vec{s}^{(j')}$ and note that $\vec{s}^{(j)}, \vec{s}^{(j')} \in {}^{\infty}\overline{B}_{\alpha/2}(\vec{c})$. It follows by the triangle inequality that ${}^{\infty}\overline{B}_{\delta}(\vec{s}^{(j)}) \subseteq {}^{\infty}\overline{B}_{\alpha/2+\delta}(\vec{c})$ and similarly ${}^{\infty}\overline{B}_{\delta}(\vec{s}^{(j')}) \subseteq {}^{\infty}\overline{B}_{\alpha/2+\delta}(\vec{c})$. This implies that ${}^{\infty}\overline{B}_{\alpha/2+\delta}(\vec{c})$ intersects every member of $\mathcal{S}$ belonging to $\mathcal{F}^{(j)}$ and also every member belonging to $\mathcal{F}^{(j')}$ which includes at least $d+2$ members because $\mathcal{F}^{(j)}$ and $\mathcal{F}^{(j')}$ are two distinct sets of $d+1$ members each.

As $\mathcal{P}$ is $(d+1, \varepsilon)$-secluded, it follows that $\mathcal{S}$ is also $(d+1, \varepsilon)$-secluded, and this implies that $\varepsilon \leq \frac{\alpha}{2} + \delta$. Algebraically solving for α proves the claim. ∎

Now we consider the collection of open $(\varepsilon - \delta)$-balls at each $\vec{s}^{(j)}$:

$$\left\{ {}^{\infty}B^{\circ}_{\varepsilon-\delta}(\vec{s}^{(j)}) : j \in [\mathrm{Sperner}(d)] \right\}.$$

Each ball in this collection has a volume (Lebesgue measure) of $(2(\varepsilon - \delta))^d$, and by Claim A, these balls (because they are open) are all disjoint. Furthermore, each ball is a subset of $(-\varepsilon, D+\varepsilon)^d$ because $\vec{s}^{(j)} \in [0, D]^d$ and the radius of the ball is $(\varepsilon - \delta) < \varepsilon$. This gives the following volume/measure comparison argument:

$$\mathrm{Sperner}(d) \cdot (2(\varepsilon - \delta))^d = m \left(\bigsqcup_{j=1}^{\mathrm{Sperner}(d)} {}^{\infty}B^{\circ}_{\varepsilon-\delta}(\vec{s}^{(j)}) \right) \leq m \left((-\varepsilon, D+\varepsilon)^d \right) = (D + 2\varepsilon)^d.$$

Because the leftmost and rightmost expressions in this inequality are continuous functions of δ, and because the inequality holds for arbitrary $\delta \in (0, \infty)$, it holds also for $\delta = 0$. Fixing $\delta = 0$ the leftmost and rightmost expressions in this inequality are also continuous functions of D, and because the inequality holds for arbitrary $D \in (1, \infty)$, it holds also for $D = 1$. Thus, we have

$$\text{Sperner}(d) \cdot (2\varepsilon)^d \leq (1 + 2\varepsilon)^d.$$

Taking dth roots of both sides and manipulating the equations, we have[12]

$$\varepsilon \leq \frac{1}{2\left(\text{Sperner}(d)^{1/d} - 1\right)}.$$

In particular, we can increase (or maintain) this upper bound by replacing Sperner(d) with the smaller (or equal) quantity dis(d) (see footnote[13]) giving

$$\varepsilon \leq \frac{1}{2\left(\text{dis}(d)^{1/d} - 1\right)}.$$

This proves the result. □

7.2.3 Numeric Bound and Optimality in $\mathbb{R}^1$ and $\mathbb{R}^2$

We now present the main numeric bound in terms of Theorem 7.2.9 (Dissection Upper Bound on ε for Unit Diameter Partitions). For this, we will make use of Lemma D.0.2 proven in Appendix D (Asymptotics) which says that the ugly expression $\frac{1}{2\left((d+1)^{\frac{d-1}{2d}}-1\right)}$ is basically equal to $\frac{1}{2\sqrt{d}}$ for most dimensions d.

[12] We allow division by 0 (i.e. when Sperner(d) $= 1$) as the interpretation is consistent: the equation above says nothing about ε when Sperner(d) $= 1$, and division by 0 in the equation below gives the vacuous statement that $\varepsilon \leq \infty$.

[13] A minor, but important remark is that dis(d) ≥ 1, so decreasing the denominator in this manner does not change the sign of stated bound.

Lemma D.0.2. *The function* $f(x) = \frac{1}{2\left((x+1)^{\frac{x-1}{2x}}-1\right)}$ *is asymptotically equivalent to the function* $g(x) = \frac{1}{2\sqrt{x}}$ *(i.e.* $\lim_{x\to\infty} \frac{f(x)}{g(x)} = 1$*). Furthermore, for* $x \geq 3$*,* $f(x) \leq 4g(x) = \frac{2}{\sqrt{x}}$.

Because the inequality above only applies for $d \geq 3$, to get the numeric bound we want, we will additionally use the fact that we also have (exact) bounds in $\mathbb{R}^1$ and $\mathbb{R}^2$. These are stated next and will be proven after the general numeric upper bound on ε.

Proposition 7.2.10 (Optimal $\varepsilon = \frac{1}{2d}$ for Unit Diameter Partitions in $\mathbb{R}^1$)**.** *Let* $d = 1$ *and* $\varepsilon \in (0,\infty)$ *and* $\mathcal{P}$ *a* $(d+1,\varepsilon)$*-secluded partition of* $\mathbb{R}^d = \mathbb{R}^1$ *such that for all* $X \in \mathcal{P}$*,* $\operatorname{diam}_\infty(X) \leq 1$*. Then* $\varepsilon \leq \frac{1}{2d} = \frac{1}{2}$.

Proposition 7.2.11 (Optimal $\varepsilon = \frac{1}{2d}$ for Unit Diameter Partitions in $\mathbb{R}^2$)**.** *Let* $d = 2$ *and* $\varepsilon \in (0,\infty)$ *and* $\mathcal{P}$ *a* $(d+1,\varepsilon)$*-secluded partition of* $\mathbb{R}^d = \mathbb{R}^2$ *such that for all* $X \in \mathcal{P}$*,* $\operatorname{diam}_\infty(X) \leq 1$*. Then* $\varepsilon \leq \frac{1}{2d} = \frac{1}{4}$.

Now we can state and prove the main numeric upper bound on ε for this section (though we once again emphasize that Theorem 7.2.9 (Dissection Upper Bound on ε for Unit Diameter Partitions) is a more important result than this numeric bound because the bounds of Section 7.1 (Upper Bounds on ε via Brunn-Minkowski and Blichfeldt) are better than this numeric bound, so it is the fact that we have a bound in terms of the dissection number that is interesting).

Theorem 7.2.12 (Numeric Dissection Upper Bound on ε for Unit Diameter Partitions)**.** *Let* $d \in \mathbb{N}$ *and* $\varepsilon \in (0,\infty)$ *and* $\mathcal{P}$ *a* $(d+1,\varepsilon)$*-secluded partition of* $\mathbb{R}^d$ *such that for all* $X \in \mathcal{P}$*,* $\operatorname{diam}_\infty(X) \leq 1$*. Then* $\varepsilon \leq \frac{2}{\sqrt{d}}$.

Proof. For $d = 1$ this follows from Proposition 7.2.10 and for $d = 2$ it follows from Proposition 7.2.11 For $d \geq 3$, apply the bound of Theorem 7.2.9 together with the dissection number lower bound of $(d+1)^{\frac{d-1}{2}} \leq \mathrm{dis}(d)$ due to [Gla12] and the cleaner value of Lemma D.0.2 to obtain

$$\varepsilon \leq \frac{1}{2\left(\mathrm{dis}(d)^{1/d} - 1\right)} \leq \frac{1}{2\left((d+1)^{\frac{d-1}{2d}} - 1\right)} \leq \frac{2}{\sqrt{d}}.$$

□

Remark 7.2.13. By Lemma D.0.2, the bound of Theorem 7.2.12 can be asymptotically improved by a factor of 4. △

Now we prove Proposition 7.2.10 and Proposition 7.2.11. The former is an immediate consequence of the following simple bound.

Proposition 7.2.14 (Trivial ε Bound)**.** *Let $d \in \mathbb{N}$ and $\varepsilon \in (0,\infty)$ and $\mathcal{P}$ a $(2^d,\varepsilon)$-secluded partition of $\mathbb{R}^d$ such that for all $X \in \mathcal{P}$, $\mathrm{diam}_\infty(X) \leq 1$. Then $\varepsilon \leq \frac{1}{2}$.*

Proof. We show that for any $\varepsilon > \frac{1}{2}$, there is some point $\vec{p} \in \mathbb{R}^d$ such that ${}^{\infty}\overline{B}_\varepsilon(\vec{p})$ intersects at least $2^d + 1$ members of $\mathcal{P}$ which shows that $\mathcal{P}$ is not $(2^d,\varepsilon)$-secluded for any $\varepsilon > \frac{1}{2}$.

Let $\varepsilon > \frac{1}{2}$ and fix any $X \in \mathcal{P}$. For each $i \in [d]$, let $a_i = \inf(\pi_i(X))$ and let $b_i = \sup(\pi_i(X))$ noting that $b_i - a_i \leq 1$. Thus,

$$X \subseteq \prod_{i=1}^{d}[a_i,b_i] \subseteq \prod_{i=1}^{d}[a_i,a_i+1] = \vec{a} + [0,1]^d.$$

Let $\vec{p}$ be the center of $\vec{a}+[0,1]^d$ (i.e. $\vec{p} = \vec{a} + \frac{1}{2}\cdot\vec{1} = \langle a_i + \frac{1}{2}\rangle_{i=1}^{d}$), so $X \subseteq {}^{\infty}\overline{B}_{1/2}(\vec{p}) \subsetneq {}^{\infty}\overline{B}_\varepsilon(\vec{p})$. For any distinct $\vec{\alpha},\vec{\beta} \in \{-\varepsilon,\varepsilon\}^d$, we have $\vec{\alpha}+\vec{p}, \vec{\beta}+\vec{p} \in {}^{\infty}\overline{B}_\varepsilon(\vec{p})$, and $\vec{\alpha}+$

$\vec{p}, \vec{\beta} + \vec{p} \notin X$, and also

$$\|(\vec{p} + \vec{\alpha}) - (\vec{p} + \vec{\beta})\|_\infty = \|\vec{\alpha} - \vec{\beta}\|_\infty = 2\varepsilon > 1$$

which implies $\vec{p}+\vec{\alpha}$ and $\vec{p}+\vec{\beta}$ belong to different members of $\mathcal{P}$. Thus ${}^\infty\overline{B}_\varepsilon(\vec{p})$ intersects X and at least 2^d other members of $\mathcal{P}$ for a total of at least $2^d + 1$ members. □

Just to have it stated somewhere, the trivial bound above is tight.

Fact 7.2.15 (Existence of $(2^d, \frac{1}{2})$-Secluded Unit Cube Partition)**.** *For each $d \in \mathbb{N}$, there exists a $(2^d, \frac{1}{2})$-secluded axis-aligned unit cube partition of $\mathbb{R}^d$.*

Justification. The "standard grid partition" $\mathcal{P} = \left\{\vec{n} + [0,1)^d : \vec{n} \in \mathbb{Z}^d\right\}$ has this property. In fact, any partition $\mathcal{P}$ of $\mathbb{R}^d$ by translates of $[0,1)^d$ has this property because (by application of Fact 3.4.13) the cube ${}^\infty\overline{B}_{1/2}(\vec{p})$ intersects exactly one member of $\mathcal{P}$ per corner. □

As a simple corollary to the above for $\mathbb{R}^1$ we have that when $k = d+1$, then $\varepsilon = \frac{1}{2d}$ is the minimum possible (which is what we achieved with the reclusive partitions).

Proposition 7.2.10 (Optimal $\varepsilon = \frac{1}{2d}$ for Unit Diameter Partitions in $\mathbb{R}^1$)**.** *Let $d = 1$ and $\varepsilon \in (0, \infty)$ and $\mathcal{P}$ a $(d+1, \varepsilon)$-secluded partition of $\mathbb{R}^d = \mathbb{R}^1$ such that for all $X \in \mathcal{P}$, $\operatorname{diam}_\infty(X) \leq 1$. Then $\varepsilon \leq \frac{1}{2d} = \frac{1}{2}$.*

Proof. Because $2^d = d+1$ when $d = 1$, we can apply Proposition 7.2.14 with $d = 1$. □

We can also show that when $k = d + 1$, the value of $\varepsilon = \frac{1}{2d}$ is optimal also for $\mathbb{R}^2$, but Theorem 7.2.9 (Dissection Upper Bound on ε for Unit Diameter Partitions) evaluated with $d = 2$ (noting $\operatorname{Sperner}(2) = 2$ as in Table 7.2) only gives a bound of $\varepsilon \leq \frac{1}{2(\sqrt{2}-1)} \approx 1.207$ which is not very close to the bound of $\frac{1}{4}$. The proof uses the same ideas as those of Theorem 7.2.9, but is more precise by very carefully dealing with parameters and focusing on a cube centered around one of the $\vec{s}^{(j)}$. The proof is illustrated in Figure 7.2.

Proposition 7.2.11 (Optimal $\varepsilon = \frac{1}{2d}$ for Unit Diameter Partitions in $\mathbb{R}^2$)**.** *Let $d = 2$ and $\varepsilon \in (0, \infty)$ and $\mathcal{P}$ a $(d+1, \varepsilon)$-secluded partition of $\mathbb{R}^d = \mathbb{R}^2$ such that for all $X \in \mathcal{P}$, $\operatorname{diam}_\infty(X) \leq 1$. Then $\varepsilon \leq \frac{1}{2d} = \frac{1}{4}$.*

Proof. As in the proof of Theorem 7.2.9, we will use a limiting argument, so let $D \in (1, \infty)$ and $\delta \in (0, \varepsilon]$. By Theorem 6.2.2 (Stronger Optimality Theorem), because every member of $\mathcal{P}$ has diameter less than D, there is a point $\vec{p} \in \mathbb{R}^d$ such that ${}^{\infty}\overline{B}_\delta(\vec{p})$ intersects at least 3 members of $\mathcal{P}$. Because $\mathcal{P}$ is $(3, \varepsilon)$-secluded, by monotonicity (Observation 4.4.2) it is $(3, \delta)$-secluded, so ${}^{\infty}\overline{B}_\delta(\vec{p})$ intersects at most 3 members of $\mathcal{P}$, and so we get equality. That is, $\left|{}^{\infty}\overline{\mathcal{N}}_\delta(\vec{p})\right| = 3$. Let this $\vec{p}$ be fixed for the remainder of the proof.

Now consider the cube $S = {}^{\infty}\overline{B}_{D/2}(\vec{p}) = \vec{p} + [\frac{D}{2}, \frac{D}{2}]^d$ and the partition $\mathcal{S}$ of S induced by $\mathcal{P}$ (i.e. $\mathcal{S} = \{X \cap S : X \in \mathcal{P} \text{ and } X \cap S \neq \emptyset\}$). Because Sperner(2) $= 2$ (see Table 7.2), by Theorem 7.2.9, there exists two distinct sets $\mathcal{F}^{(1)}, \mathcal{F}^{(2)} \subseteq \mathcal{P}$ each of cardinality 3 and associated points $\vec{s}^{(1)}, \vec{s}^{(2)} \in S = \vec{p} + [\frac{D}{2}, \frac{D}{2}]^d$ such that ${}^{\infty}\overline{B}_\delta(\vec{s}^{(1)})$ intersects every member in $\mathcal{F}^{(1)}$ and ${}^{\infty}\overline{B}_\delta(\vec{s}^{(2)})$ intersects every member in $\mathcal{F}^{(2)}$ (in alternate notation, $\mathcal{F}^{(1)} \subseteq {}^{\infty}\overline{\mathcal{N}}_\delta(\vec{s}^{(1)})$ and $\mathcal{F}^{(2)} \subseteq {}^{\infty}\overline{\mathcal{N}}_\delta(\vec{s}^{(2)})$).

Because $\mathcal{F}^{(1)}$ and $\mathcal{F}^{(2)}$ are distinct subsets of $\mathcal{P}$ of cardinality 3, and ${}^{\infty}\overline{\mathcal{N}}_\delta(\vec{p})$ is also a subset of $\mathcal{P}$ of cardinality 3, we either have $\mathcal{F}^{(1)} \neq {}^{\infty}\overline{\mathcal{N}}_\delta(\vec{p})$ or $\mathcal{F}^{(2)} \neq {}^{\infty}\overline{\mathcal{N}}_\delta(\vec{p})$, and we assume without loss of generality that it is the former.

Now, because $\vec{s}^{(1)} \in S = \vec{p} + [\frac{D}{2}, \frac{D}{2}]^d = {}^{\infty}\overline{B}_{D/2}(\vec{p})$ we can let $\vec{c}$ be the midpoint between $\vec{p}$ and $\vec{s}^{(1)}$ and note that $\vec{p}, \vec{s}^{(1)} \in {}^{\infty}\overline{B}_{D/4}(\vec{c})$. It follows by the triangle inequality that ${}^{\infty}\overline{B}_\delta(\vec{p}) \subseteq {}^{\infty}\overline{B}_{D/4+\delta}(\vec{c})$ and similarly ${}^{\infty}\overline{B}_\delta(\vec{s}^{(1)}) \subseteq {}^{\infty}\overline{B}_{D/4+\delta}(\vec{c})$. This

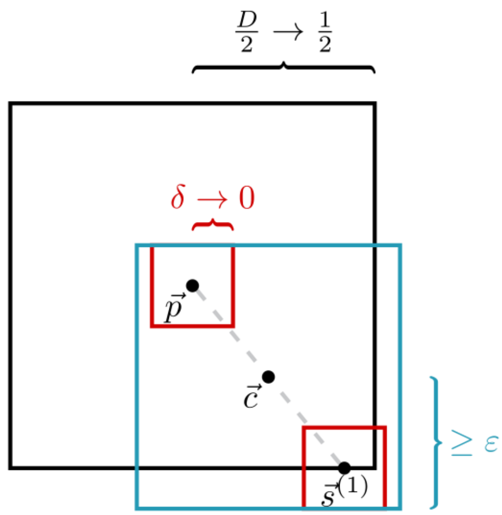


Figure 7.2: Illustration of the proof of Proposition 7.2.11 that $\varepsilon = \frac{1}{4}$ is optimal in $\mathbb{R}^2$. For all $D > 1$ and $\delta > 0$ the diagram shows that $\frac{D}{2} + 2\delta \geq 2\varepsilon$ so by a limiting argument we have $\frac{1}{4} \geq \varepsilon$.

implies that ${}^{\infty}\overline{B}_{D/4+\delta}(\vec{c})$ intersects every member of $\mathcal{P}$ belonging to ${}^{\infty}\overline{\mathcal{N}}_{\delta}(\vec{p})$ and also every member belonging to $\mathcal{F}^{(1)}$ which includes at least 4 members because these are two distinct sets of 3 members each.

As $\mathcal{P}$ is $(3, \varepsilon)$-secluded, this implies that $\varepsilon \leq \frac{D}{4} + \delta$. Since this is true for all $D \in (1, \infty)$ and $\delta \in (0, \varepsilon]$ it holds that $\varepsilon \leq \frac{1}{4}$. □

7.2.4 Possible Improvements to the Dissection Bound

The current upper and lower bounds on the dissection number due respectively to Glazyrin [Gla12] and Orden and Santos [OS03] are the following:

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

Thus, it is consistent with the current bounds that $\operatorname{dis}(d) \in \Omega\left(\frac{d!}{c^d}\right)$ for some constant $c \in (0, \infty)$. Using Stirling's approximation, we can reformat this and note that it is also consistent with these bounds that $\mathrm{dis}(d) \in \Omega\left(\left(\frac{d}{c'}\right)^d\right)$ for some constant $c' \in (0,\infty)$. To see this, note that by Stirling's approximation with $c' = ce$ we have

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

Thus, if we assume $\mathrm{dis}(d) \in \Omega\left(\left(\frac{d}{c'}\right)^d\right)$, we have $\frac{1}{2(\mathrm{dis}(d)^{1/d}-1)} \in O(\frac{1}{d})$. Thus, there is some constant $C \in (0,\infty)$ such that for all $d \in \mathbb{N}$, $\frac{1}{2\left(\sqrt[d]{\mathrm{dis}(d)}-1\right)} \leq \frac{C}{2d}$ (i.e. this inequality does not just hold asymptotically; see footnote[14]).

This demonstrates that if it is proven that $\mathrm{dis}(d) \in \Omega\left(\left(\frac{d}{c'}\right)^d\right)$, which is consistent with known bounds, then our reclusive partition constructions (Theorem 4.2.18) are optimal in ε up to a constant factor even for the broad class of unit ℓ_∞ diameter bounded partitions when $k = d+1$ (which we conjecture to be true regardless of whether sufficient improvements to the dissection number lower bounds are possible (see Conjecture 7.3.2).

7.3 Conjectures on Optimal ε

We complete this chapter by stating a number of conjectures of various strengths regarding the optimality of the ε parameter.

Firstly, we conjecture that our reclusive constructions attain the optimal value of ε for all unit ℓ_∞ diameter bounded partitions when $k = d+1$.

[14]In general, for any function $f(d) \in O(\frac{1}{d})$ this is true because $\frac{1}{d}$ is positive on $\mathbb{N}$. Specifically, if $f(d) \in O(\frac{1}{d})$, then there are some constants c''' and N such that for $d \geq N$, $f(d) \leq \frac{c'''}{d}$. Take $M = \max\{f(d) \colon d \in [N]\}$, so for $d \in [N]$, $f(d) \leq M \leq M \cdot \frac{N}{d}$. Thus for all $d \in \mathbb{N}$, $f(d) \leq \max\left\{\frac{c'''}{d}, \frac{MN}{d}\right\}$ so we can take $C = \max\{c''', MN\}$ and have $f(d) \leq \frac{C}{d}$ for all $d \in \mathbb{N}$.

Conjecture 7.3.1 (Exact Optimality of $\varepsilon = \frac{1}{2d}$ for Unit Diameter Bounded Partitions)**.** *Let $d \in \mathbb{N}$ and $\varepsilon \in (0, \infty)$ and $\mathcal{P}$ a $(d+1, \varepsilon)$-secluded partition of $\mathbb{R}^d$ such that for all $X \in \mathcal{P}$, $\operatorname{diam}_\infty(X) \leq 1$. Then $\varepsilon \leq \frac{1}{2d}$.*

Some motivation for believing this conjecture is that we at least know that $\frac{1}{2d}$ is exactly optimal for axis-aligned unit cube partitions (Corollary 9.8.5). Less convincing, though somewhat relevant is that we have also seen that $\frac{1}{2d}$ is exactly optimal in $\mathbb{R}^1$ and $\mathbb{R}^2$ for unit ℓ_∞ bounded diameter partitions when $k = d + 1$ (Proposition 7.2.10 and Proposition 7.2.11).

If this is not true, then we still conjecture a weaker result that this is optimal up to some universal constant factor.

Conjecture 7.3.2 (Optimality of $\varepsilon = \frac{1}{2d}$ for Unit Diameter Bounded Partitions Up To Constants)**.** *Let $d \in \mathbb{N}$ and $\varepsilon \in (0, \infty)$ and $\mathcal{P}$ a $(d+1, \varepsilon)$-secluded partition of $\mathbb{R}^d$ such that for all $X \in \mathcal{P}$, $\operatorname{diam}_\infty(X) \leq 1$. Then for some universal constant C independent of d it holds that $\varepsilon \leq \frac{C}{2d}$.*

As mentioned in Subsection 7.2.4 (Possible Improvements to the Dissection Bound), this conjecture would follow for example from sufficient improvements to the lower bounds on the dissection number of the cube, though it could be true even if the true value of the dissection number is too small to prove this result.

In fact, we believe this latter conjecture is true even when k can be polynomial in d.

Conjecture 7.3.3 (Optimality of $\varepsilon = \frac{1}{2d}$ for Unit Diameter Bounded Partitions Up To Constants for Polynomial k)**.** *Let $d \in \mathbb{N}$ and $\varepsilon \in (0, \infty)$ and $t \in [1, \infty)$ and $k \leq d^t$ and $\mathcal{P}$ a (k, ε)-secluded partition of $\mathbb{R}^d$ such that for all $X \in \mathcal{P}$, $\operatorname{diam}_\infty(X) \leq 1$. Then for some universal constant C_t dependent on t but independent of d, it holds that $\varepsilon \leq \frac{C_t}{2d}$.*

Some reason for believing this polynomial variant is that we have seen in our partition product constructions in Section 4.4 that our ideas for constructing other axis-aligned unit cube partitions of $\mathbb{R}^d$ beyond the reclusive constructions of Section 4.2 only yielded constant factor improvements of the ε parameter at the cost of polynomial increases to k. We also showed in Theorem 7.1.9 (Near Optimality of ε) or Corollary 7.1.10 (Near Optimality of ε) or Corollary 7.1.11 that our bound of $\varepsilon(d) \in O(\frac{\log(d)}{d})$ when $k = d + 1$ also held for any polynomial $k(d)$ which suggests that allowing k to increase from $d + 1$ to some other polynomial in d doesn't gain much. It is certainly possible that Conjecture 7.3.1 and/or Conjecture 7.3.2 holds and Conjecture 7.3.3 does not—this would imply that polynomial $k(d)$ allows for an $\omega(1)$ factor increase in $\varepsilon(d)$ compared to what can be accomplished with $k = d + 1$. However, we suspect that is not the case.

We also think that the latter two of these conjectures do not require the diameter condition, but only require the weaker outer measure condition[15].

Conjecture 7.3.4 (Optimality of $\varepsilon = \frac{1}{2d}$ for Unit Measure Bounded Partitions Up To Constants)**.** *Conjecture 7.3.2 and Conjecture 7.3.3 still hold if the condition "for all $X \in \mathcal{P}$, $\operatorname{diam}_\infty(X) \leq 1$" is replaced with the weaker condition "for all $X \in \mathcal{P}$, $m_{out}(X) \leq 1$."*

[15]Requiring outer measure at most 1 is weaker than requiring ℓ_∞ diameter at most 1 because the latter implies the former as shown in Fact 3.4.9 with $D = 1$.

We suspect this, because our best bounds (i.e. Theorem 7.1.9 (Near Optimality of ε) and the related results) have not needed the diameter condition, so at the moment we don't really have any stronger results for the stronger diameter condition, and we have been able to prove everything just using the outer measure condition.

However, one thing that we think we know[16] is that Conjecture 7.3.1 does *not* hold if we replace the unit diameter condition with a unit outer measure condition. This is based on the following which we state as a conjecture, but will outline our justification.

Conjecture 7.3.5 (Non-Uniform Ball in Standard Reclusive Partition). *Let $d \in \mathbb{N}$ and A the reclusive matrix of Equation 4.1 and $\mathcal{P}_A$ its reclusive partition of $\mathbb{R}^d$. By (the proof of) Theorem 4.2.18, we know that every translate of $[0, \frac{1}{d}]^d$ intersects at most $d+1$ members of $\mathcal{P}_A$, but we think something stronger is true: every translate of $\prod_{i=d}^{1}[0, \frac{1}{i}]$ intersects at most $d+1$ members of $\mathcal{P}_A$.*

Justification. The $\frac{1}{d}$ width of the cube $[0, \frac{1}{d}]^d$ is because of the top row of the matrix A, and so in the first coordinate, we suspect we do have to have the width of a rectangle be at most $\frac{1}{d}$. However, in the other coordinates (corresponding to other rows of the matrix A), the entries of the matrix are more spaced out, and so the cubes are more spaced out along these coordinates. We suspect that we can prove this by following closely through the proof of Proposition 4.2.10 and tracking distance calculations corresponding to each coordinate rather than blindly minimizing/maximizing over all coordinates. □

It this is true, it would imply that moving from a unit diameter bound to a unit measure bound allows for increasing ε asymptotically by a factor of $e \approx 2.718$.

[16] We haven't gone through all of the details, but we have sketched these ideas.

Conjecture 7.3.6 (Inexact Optimality of $\varepsilon = \frac{1}{2d}$ for Unit Measure Bounded Partitions)**.** *For each $d \in \mathbb{N}$, there is a $(d+1, \varepsilon(d))$-secluded partition of $\mathbb{R}^d$ consisting of rectangles with Lebesgue measure 1 where $\varepsilon(d) = \frac{1}{2}(\frac{1}{d!})^{1/d} \approx \frac{e}{d}$.*

Proof Sketch Assuming Conjecture 7.3.5. Letting $\mathcal{P}_A$ as in Conjecture 7.3.5, consider reversing the order of the coordinates so that we can consider the rectangle $\prod_{i=1}^{d}[0, \frac{1}{i}]$ rather than $\prod_{i=d}^{1}[0, \frac{1}{i}]$.

In each coordinate $i \in [d]$, stretch each member of $\mathcal{P}$ by a factor of i in coordinate i. Thus, each member of $\mathcal{P}$ is a translate of the rectangle $\prod_{i=1}^{d}[0, i)$ and we can now translate the closed cube $\prod_{i=1}^{d}[0, 1] = [0, 1]^d$ (the stretched version of $\prod_{i=1}^{d}[0, \frac{1}{i}]$) anywhere in space and intersect at most $d+1$ members of $\mathcal{P}$.

Each member of $\mathcal{P}$ has Lebesgue measure $d!$, so we now scale every member of $\mathcal{P}$ by a factor of $(\frac{1}{d!})^{1/d}$ (in every coordinate) so that the volume of each member of $\mathcal{P}$ is scaled by a factor of $\frac{1}{d!}$. Thus, at this point each member of $\mathcal{P}$ is a rectangle of Lebesgue measure 1, and we can translate the scaled cube $[0, (\frac{1}{d!})^{1/d}]^d$ anywhere in space and intersect at most $d+1$ members of $\mathcal{P}$.

Thus, we can take $\varepsilon = \frac{1}{2}(\frac{1}{d!})^{1/d}$ so that the cube $[0, (\frac{1}{d!})^{1/d}]^d$ is just a translate of ${}^{\infty}\overline{B}_{\varepsilon}(\vec{0})$. By Stirling's approximation we have that

$$\varepsilon = \frac{1}{2}\left(\frac{1}{d!}\right)^{1/d} \approx \frac{1}{2} \cdot \frac{1}{(2\pi d)^{1/2d}} \cdot \frac{e}{d}$$

which converges rapidly to $\frac{e}{2d}$ (as $d \to \infty$) yielding a factor of improvement of nearly e compared to our reclusive constructions. □

Remark 7.3.7. Note that in this construction, each member of $\mathcal{P}$ is a translate of $\prod_{i=1}^{d}\left[0, \frac{i}{(d!)^{1/d}}\right]$ which has diameter $\frac{d}{(d!)^{1/d}} \approx d \cdot \frac{1}{(2\pi d)^{1/2d}} \cdot \frac{e}{d}$ which tends to e as $d \to \infty$. △

Chapter 8

A Neighborhood Variant of the Lebesgue Covering Theorem, the Cubical KKM Lemma, and the Cubical Sperner's Lemma

The Lebesgue covering theorem (c.f. [HW48, Theorem IV 2], stated earlier as Theorem 5.1.2) along with polytopal variants of Sperner's lemma on the cube (c.f. [LPS01]) and polytopal variants of the KKM lemma on the cube (c.f. [Kuh60, Wol77, Kom94, vdLTY99]) are all equivalent and guarantee that in well-behaved colorings/partitions/coverings of the cube, there is a point at the closure of at least $d+1$ colors. We will state the KKM lemma and the Lebesgue covering theorem below (in a slightly different style than Theorem 5.1.2) and then present an equivalent version that we will prefer to work with before continuing the discussion of the purpose of this chapter.

Below, we use $[0,1]^d$ as the standard unit d-cube and $V=\{0,1\}^d$ as the set of vertices of it. A face F of the cube $[0,1]^d$ is a product set $F=\prod_{i=1}^d F_i$ where each F_i is one of three sets: $\{0\}$, $\{1\}$, or $[0,1]$, and two faces F, F' are said to be opposite each other if there is some coordinate $i_0 \in [d]$ such that $F_{i_0}=\{0\}$ and $F'_{i_0}=\{1\}$ (or vice versa).

Definition 8.0.1 (Lebesgue Cover). A *Lebesgue cover* of $[0,1]^d$ is an indexed family $\mathcal{C} = \{C_n\}_{n\in[N]}$ of closed sets for some $N \in \mathbb{N}$ such that for each $n \in [N]$ it holds that C_n does not contain points on opposite faces of the cube.

Definition 8.0.2 (KKM Cover). A *KKM cover* of $[0,1]^d$ is an indexed family $\mathcal{C} = \{C_{\vec{v}}\}_{\vec{v}\in\{0,1\}^d}$ of closed sets such that for each face F of the cube $[0,1]^d$ it holds that $F \subseteq \bigcup_{\vec{v}\in F\cap\{0,1\}^d} C_{\vec{v}}$.

While the two notions of covers are different from each other[1], they are both sufficient to guarantee the same conclusion as given in the known Cubical KKM Lemma and Lebesgue Covering Theorem below.

Theorem 8.0.3 (Lebesgue Covering Theorem). *Given a Lebesgue cover of* $[0,1]^d$*, there exists a point* $\vec{p} \in [0,1]^d$ *belonging to at least* $d+1$ *sets in the cover (i.e.* $|\{n \in [N] : \vec{p} \in C_n\}| \geq d+1$*).*

Theorem 8.0.4 (Cubical KKM Lemma). *Given a KKM cover of* $[0,1]^d$*, there exists a point* $\vec{p} \in [0,1]^d$ *belonging to at least* $d+1$ *sets in the cover (i.e.* $\left|\left\{\vec{v} \in \{0,1\}^d : \vec{p} \in C_{\vec{v}}\right\}\right| \geq d+1$*).*

For our purposes, we don't want to work directly with a KKM cover because we need to deal with extensions along boundaries that are cumbersome to work with for a KKM cover due to the fact that sets can intersect opposite faces. We also don't want to work with Lebesgue covers because the closedness in the definition of a Lebesgue cover actually implies that every set in the cover (presumed to be a subset of $[0,1]^d$)

[1]For example, the indexed family $\mathcal{C} = \{C_{\vec{v}}\}_{\vec{v}\in\{0,1\}^d}$ where for each $\vec{v}$, $C_{\vec{v}} = [0,1]^d$ is a KKM cover, but not a Lebesgue cover. Conversely, an finite Lebesgue cover with cardinality exceeding $\left|\{0,1\}^d\right| = 2^d$ is not a KKM cover.

has ℓ_∞ diameter strictly less than 1, but we are okay with sets that have diameter 1 as long as they don't contain points on opposite faces (i.e. no points in the set attain distance 1 from each other, so 1 is a strict pairwise bound (Definition 6.0.1)). In other words, we don't want to require the sets to be closed. For these reasons, we will find it easier to work not with a cover, but rather with a coloring (i.e. a partition), though our results trivially and immediately generalize to covers as well. This view will also make the connection to Sperner's lemma appear nicer later. Thus, opposed to KKM covers and Lebesgue covers, we define a type of coloring (called a non-spanning coloring) and an analogous result to the Cubical KKM Lemma and the Lebesgue Covering Theorem. In fact, the following theorem is equivalent to both, and the proofs of these equivalencies are provided in Section B.1 (Equivalencies).

Definition 8.0.5 (Non-Spanning Coloring). For a set $\Lambda \subseteq [0,1]^d$ (possibly $\Lambda = [0,1]^d$), a *non-spanning coloring* of Λ is a function $\chi : \Lambda \to C$ for some set C such that χ does not map points on opposite faces to the same value[a]. If $|C| < \infty$, we call χ a *finite non-spanning coloring*.

[a]That is, for $\vec{x}, \vec{x}\,' \in \Lambda$ such that $\vec{x}$ belongs to a face of $[0,1]^d$ and $\vec{x}\,'$ belongs to an opposite face, then $\chi(\vec{x}) \neq \chi(\vec{x}\,')$.

Then the Cubical KKM Lemma and the Lebesgue Covering Theorem can be equivalently stated as follows.

Theorem 8.0.6 (KKM/Lebesgue Theorem). *Given a finite non-spanning coloring χ of $[0,1]^d$, there exists a point $\vec{p} \in [0,1]^d$ belonging to the closure of at least $d+1$ color sets (i.e. $\left|\left\{c \in C : \vec{p} \in \overline{\chi^{-1}(c)}\right\}\right| \geq d+1$).*

The main result of this chapter is that if we are not interested in a single point, but rather a small open ε ball, then for a non-spanning coloring we can find a point where the open ℓ_∞ ε-ball intersects a significant number of colors. This is why we refer to it as a "neighborhood" variant of the Cubical KKM Lemma and Lebesgue Covering Theorem. We no longer need the finiteness assumption because we are not working with closures[2].

Theorem 8.0.7 (Neighborhood KKM/Lebesgue Theorem)**.** *Given an non-spanning coloring of* $[0,1]^d$*, then for any* $\varepsilon \in (0,\infty)$ *there exists a point* $\vec{p} \in [0,1]^d$ *such that* ${}^{\infty}B^{\circ}_{\varepsilon}(\vec{p})$ *contains points of at least* $\left\lceil \left(1+\frac{\varepsilon}{1+\varepsilon}\right)^d \right\rceil$ *different colors. In particular, if* $\varepsilon \in (0,\frac{1}{2}]$ *then* ${}^{\infty}B^{\circ}_{\varepsilon}(\vec{p})$ *contains points of at least* $\left\lceil \left(1+\frac{2}{3}\varepsilon\right)^d \right\rceil$ *different colors.*

The reason restricting to $\varepsilon \le \frac{1}{2}$ is reasonable is that for large ε, the ball can be placed at the center of the unit cube and is then a superset of the cube, so it contains every point. No two vertices of the cube can have the same color in a non-spanning coloring because they belong to some pair of opposite faces. This means that every non-spanning coloring contains at least 2^d colors, and there exist non-spanning colorings with only 2^d colors (e.g. color each of the 2^d orthants a distinct color), so for $\varepsilon > \frac{1}{2}$, the bound is 2^d and this is tight.

If one takes an asymptotic perspective and thinks of ε as a function of d as we did in Section 7.1 (Upper Bounds on ε via Brunn-Minkowski and Blichfeldt), then one gets the same results. (1) when $\varepsilon \in O\left(\frac{1}{d}\right)$, this theorem gives an $O(1)$ bound[3] which is asymptotically much worse than the value $d+1$ given in the Lebesgue Covering Theorem. (2) if $\varepsilon \in \omega\left(\frac{\ln(d)}{d}\right)$ then our bound is super-polynomial in the dimension[4].

[2]The reason the Lebesgue Covering Theorem requires finiteness is because for a finite collection of sets, the union of closures equals the closure of unions, but this does not in general hold for infinite collections. Even without the finiteness condition in the Lebesgue Covering Theorem, it is still the case that there exists a point $\vec{p}$ such that every open set containing $\vec{p}$ intersects at least $d+1$ colors. See for example the proof of Lemma B.1.4 where this is proven as part of the larger proof.

[3]This is because $\lim_{d\to\infty}(1+\frac{c}{d})^d = e^c$.

[4]Let $k = (1+\frac{2}{3}\varepsilon)^d$. Because $\varepsilon \in (0,\frac{1}{2}]$ we have $\frac{2}{3}\varepsilon \in (0,\frac{1}{3}]$. Using the inequality $\ln(1+x) \ge \frac{x}{2}$ for small enough x (in particular for $x \in (0,\frac{1}{3}]$) we have $\ln(k) = d\ln(1+\frac{2}{3}\varepsilon) \ge \frac{1}{3}d\varepsilon$, so $\varepsilon \le \frac{3\ln(k)}{d}$.

And (3) if $\varepsilon \in \Theta(1)$, then the bound of the theorem is exponential in d.

If we discretize the problem but know that points are what we call ρ-proximate (Definition 8.2.1, meaning that all points of the cube are within distance ρ of a point in the set *on the same face*) then we get the same result but have to account for this spacing so we get a term like "$(\varepsilon - \rho)$" instead of "ε".

Theorem 8.0.8 (Neighborhood Sperner's Lemma)**.** *Let $\rho \in [0, \infty)$ and $\Lambda \subseteq [0,1]^d$ be ρ-proximate. Let $\rho' = \min(\rho, \frac{1}{2})$. Given a non-spanning coloring of Λ, then for any $\varepsilon \in (0, \infty)$ there exists a point $\vec{p} \in [0,1]^d$ such that ${}^{\infty}B^{\circ}_{\varepsilon}(\vec{p})$ contains points of at least $\left\lceil \left(1 + \frac{(\varepsilon-\rho')}{1+(\varepsilon-\rho')}\right)^d \right\rceil$ different colors. In particular, if $\varepsilon \in (0, \frac{1}{2}]$ then ${}^{\infty}B^{\circ}_{\varepsilon}(\vec{p})$ contains points of at least $\left\lceil \left(1 + \frac{2}{3}(\varepsilon - \rho')\right)^d \right\rceil$ different colors.*

Remark 8.0.9. The Neighborhood Sperner's Lemma and the Neighborhood KKM/Lebesgue Theorem are in fact naturally equivalent. Our proof of the Neighborhood Sperner's Lemma will be as a corollary of the Neighborhood KKM/Lebesgue Theorem showing one direction of the equivalence. For the other direction, we recover the statement of the Neighborhood KKM/Lebesgue Theorem from the Neighborhood Sperner's Lemma in the special case $\Lambda = [0,1]^d$ and $\rho = 0$ (because $[0,1]^d$ is 0-proximate). △

Unsurprisingly, we can restate the Neighborhood KKM/Lebesgue Theorem in terms of Lebesgue covers or in terms of KKM covers as follows. Both are in fact equivalent to the Neighborhood KKM/Lebesgue Theorem; we don't prove the equivalence, but the main ideas can be found in Section B.1 where we prove the equivalence of the Cubical KKM Lemma, Lebesgue Covering Theorem, and KKM/Lebesgue Theorem.

Thus, if the bound k is polynomial in d, then $\varepsilon \in O(\frac{\ln(d)}{d})$, so if $\varepsilon \in \omega\left(\frac{\ln(d)}{d}\right)$ the k is not polynomial in d.

Corollary 8.0.10 (Neighborhood Lebesgue Theorem)**.** *Given a Lebesgue cover of* $[0,1]^d$*, then for any* $\varepsilon \in (0,\infty)$ *there exists a point* $\vec{p} \in [0,1]^d$ *such that* ${}^{\infty}B^{\circ}_{\varepsilon}(\vec{p})$ *intersects at least* $\left\lceil \left(1+\frac{\varepsilon}{1+\varepsilon}\right)^d \right\rceil$ *sets in the cover. In particular, if* $\varepsilon \in (0,\frac{1}{2}]$ *then* ${}^{\infty}B^{\circ}_{\varepsilon}(\vec{p})$ *intersects at least* $\left\lceil \left(1+\frac{2}{3}\varepsilon\right)^d \right\rceil$ *sets in the cover.*

Proof. Let $N \in \mathbb{N}$ and $\mathcal{C} = \{C_n\}_{n\in[N]}$ be a Lebesgue cover of $[0,1]^d$. Because this is a cover, every point of $[0,1]^d$ belongs to some set, so define χ as follows:

$$\chi : [0,1]^d \to [N]$$
$$\chi(\vec{x}) = \min\left\{n \in [N] : \vec{x} \in C_n\right\}.$$

This is trivially a finite non-spanning coloring of $[0,1]^d$ because the codomain of χ is finite and for $\vec{x}^{(0)}$ and $\vec{x}^{(1)}$ on opposite faces, there is no $n \in [N]$ for which both $\vec{x}^{(0)} \in C_n$ and $\vec{x}^{(1)} \in C_n$ and thus $\left\{n \in [N] : \vec{x}^{(0)} \in C_n\right\}$ and $\left\{n \in [N] : \vec{x}^{(1)} \in C_n\right\}$ are disjoint, so $\chi(\vec{x}^{(0)}) \neq \chi(\vec{x}^{(1)})$.

By the Neighborhood KKM/Lebesgue Theorem, there exists $\vec{p} \in [0,1]^d$ such that

$$\left|\left\{n \in [N] : {}^{\infty}B^{\circ}_{\varepsilon}(\vec{p}) \cap \chi^{-1}(n)\right\}\right| \geq \left(1+\tfrac{\varepsilon}{1+\varepsilon}\right)^d.$$

Fix such a $\vec{p}$ for the remainder of the proof. For each $n \in [N]$, observe that $\chi^{-1}(n) \subseteq C_n$ because for any $\vec{x} \in \chi^{-1}(n)$ we have $\chi(\vec{x}) = n$, so by definition of χ we have $\vec{x} \in C_n$. The following subset containment then follows immediately:

$$\left\{n \in [N] : {}^{\infty}B^{\circ}_{\varepsilon}(\vec{p}) \cap \chi^{-1}(n) \neq \emptyset\right\} \quad \subseteq \quad \left\{n \in [N] : {}^{\infty}B^{\circ}_{\varepsilon}(\vec{p}) \cap C_n) \neq \emptyset\right\}.$$

and since the former has cardinality at least $(1+\frac{\varepsilon}{1+\varepsilon})^d$, so does the latter which proves the result. ∎

Corollary 8.0.11 (Neighborhood KKM Theorem)**.** *Given a KKM cover of $[0,1]^d$, then for any $\varepsilon \in (0,\infty)$ there exists a point $\vec{p} \in [0,1]^d$ such that ${}^{\infty}B^{\circ}_{\varepsilon}(\vec{p})$ intersects at least $\left\lceil \left(1+\frac{\varepsilon}{1+\varepsilon}\right)^d \right\rceil$ sets in the cover. In particular, if $\varepsilon \in (0,\frac{1}{2}]$ then ${}^{\infty}B^{\circ}_{\varepsilon}(\vec{p})$ intersects at least $\left\lceil \left(1+\frac{2}{3}\varepsilon\right)^d \right\rceil$ sets in the cover.*

Proof. Let $\mathcal{C} = \{C_{\vec{v}}\}_{\vec{v}\in\{0,1\}^d}$ be a KKM cover of $[0,1]^d$. For each $\vec{x} \in [0,1]^d$, let $F_{\vec{x}}$ denote the smallest face of the cube containing $\vec{x}$ (i.e. $F_{\vec{x}}$ is the intersection of all faces containing $\vec{x}$). By the defining property of a KKM cover, we have $F_{\vec{x}} \subseteq \bigcup_{\vec{v}\in F_{\vec{x}}\cap\{0,1\}^d} C_{\vec{v}}$, so in particular there exists some $\vec{v} \in F_{\vec{x}} \cap \{0,1\}^d$ with $\vec{x} \in C_{\vec{v}}$. Define the function χ as follows where $\min_{\text{lex}}$ denotes the minimum element in a subset of $\{0,1\}^d$ under the lexicographic ordering:

$$\chi : [0,1]^d \to \{0,1\}^d$$
$$\chi(\vec{x}) = \min_{\text{lex}} \left\{ \vec{v} \in \{0,1\}^d \cap F_{\vec{x}} : \vec{x} \in C_{\vec{v}} \right\}$$

We have already demonstrated that the set in the definition is not empty, so χ is well-defined.

We claim that χ is a finite non-spanning coloring of $[0,1]^d$. The finiteness is trivial because the codomain of χ is finite, so we need only show it is a non-spanning coloring. Suppose $F^{(0)}$ and $F^{(1)}$ are opposite faces of the cube (i.e. there is some coordinate $j \in [d]$ such that $\pi_j(F^{(0)}) = \{0\}$ and $\pi_j(F^{(1)}) = \{1\}$) and let $\vec{x}^{(0)} \in F^{(0)}$ and $\vec{x}^{(1)} \in F^{(1)}$. Because $\pi_j(F^{(0)}) \cap \pi_j(F^{(1)}) = \emptyset$, it follows that $F^{(0)} \cap F^{(1)} = \emptyset$, so $F^{(0)}$ and $F^{(1)}$ are disjoint sets.

Because $\vec{x}^{(0)} \in F^{(0)}$ and $F_{\vec{x}^{(0)}}$ is by definition the intersection of all faces containing $\vec{x}$, we have $F_{\vec{x}^{(0)}} \subseteq F^{(0)}$ (and similarly replacing "0" with "1") so that also $F^{(0)}$ and $F^{(1)}$ are disjoint. By definition of χ we have $\chi(\vec{x}^{(0)}) \in F^{(0)}$ and $\chi(\vec{x}^{(1)}) \in F^{(1)}$ showing that $\chi(\vec{x}^{(0)}) \neq \chi(\vec{x}^{(1)})$, so χ is a non-spanning coloring.

By the Neighborhood KKM/Lebesgue Theorem, there exists $\vec{p} \in [0,1]^d$ such that

$$\left|\left\{\vec{v} \in \{0,1\}^d : {}^{\infty}B^{\circ}_{\varepsilon}(\vec{p}) \cap \chi^{-1}(\vec{v}) \neq \emptyset\right\}\right| \geq \left(1 + \tfrac{\varepsilon}{1+\varepsilon}\right)^d.$$

Fix such a $\vec{p}$ for the remainder of the proof. For each $\vec{v} \in \{0,1\}^d$, observe that $\chi^{-1}(\vec{v}) \subseteq C_{\vec{v}}$ because for any $\vec{x} \in \chi^{-1}(\vec{v})$ we have $\chi(\vec{x}) = \vec{v}$, so by definition of χ we have $\vec{x} \in C_{\vec{v}}$. The following subset containment then follows immediately:

$$\left\{\vec{v} \in \{0,1\}^d : {}^{\infty}B^{\circ}_{\varepsilon}(\vec{p}) \cap \chi^{-1}(\vec{v}) \neq \emptyset\right\} \quad \subseteq \quad \left\{\vec{v} \in \{0,1\}^d : {}^{\infty}B^{\circ}_{\varepsilon}(\vec{p}) \cap C_{\vec{v}}) \neq \emptyset\right\}.$$

and since the former has cardinality at least $(1+\frac{\varepsilon}{1+\varepsilon})^d$, so does the latter which proves the result. $\square$

The remainder of this chapter is laid out as follows. In Section 8.1 we prove the Neighborhood KKM/Lebesgue Theorem and then in Section 8.2 we prove the Neighborhood Sperner's Lemma). We finish in Section 8.3 with a discussion of how we hope the bounds are improved, and we mention some limitations on what is possible.

8.1 Proof of the Neighborhood KKM/Lebesgue Theorem

The proof of the Neighborhood KKM/Lebesgue Theorem will be very similar to the proof of Theorem 7.1.1 (ε-Neighborhoods for Measure Bounded Partitions and Arbitrary Norm), so we begin by restating a number of results used in its proof.

Fact G.0.1. *For any $\alpha \in \mathbb{R}$, there exists $\gamma \in \mathbb{R}$ such that $\gamma < \alpha$ and $\lceil \gamma \rceil = \lceil \alpha \rceil$.*

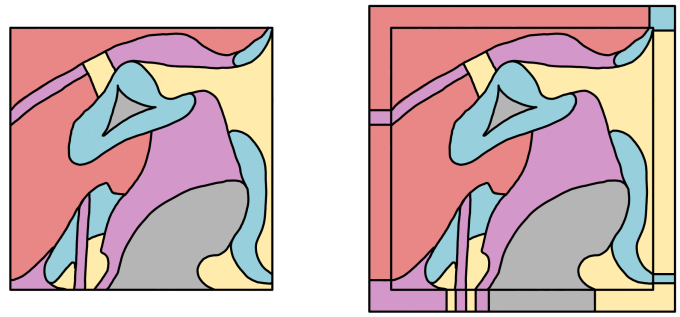
Fact G.0.2. *For $d \in [1, \infty)$, $x \in [0, 1]$, and $\alpha \in [0, \infty)$, it holds that $(x^{1/d} + \alpha)^d \geq x(1 + \alpha)^d$.*

Proposition A.3.1 (Lower Bound Cover Number for $\mathbb{R}^d$)**.** *Let $d \in \mathbb{N}$ and $S \subset \mathbb{R}^d$ be measurable with finite measure. Let $\mathcal{A}$ be a family of measurable subsets of S and let $k = \left\lceil \frac{\sum_{A \in \mathcal{A}} m(A)}{m(S)} \right\rceil$. If $k < \infty$, then there exists $\vec{p} \in S$ such that $\vec{p}$ belongs to at least k members of $\mathcal{A}$. If $k = \infty$, then for any integer n, there exists $\vec{p} \in S$ such that $\vec{p}$ belongs to at least n members of $\mathcal{A}$.*

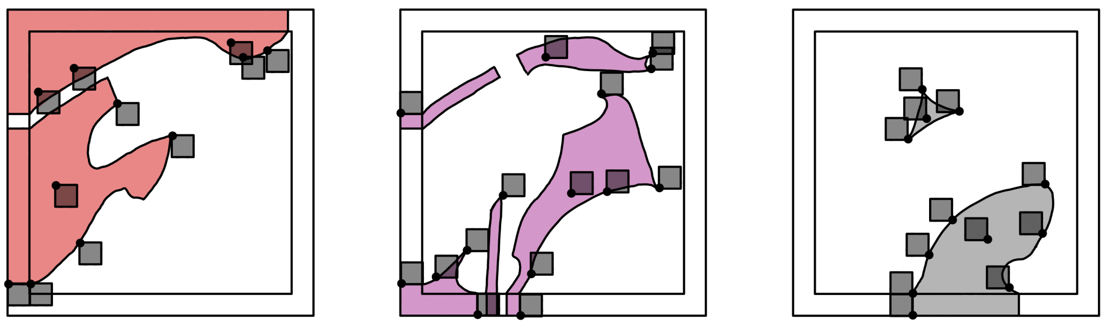
And the following is an immediate corollary of Lemma 7.1.5 (Brunn-Minkowsi with Balls) restricted to the ℓ_∞ norm.

Corollary 8.1.1 (Brunn-Minkowski with ℓ_∞ Balls)**.** *Let $d \in \mathbb{N}$ and $Y \subseteq \mathbb{R}^d$ and $\varepsilon \in (0, \infty)$. Then $Y + {}^{\infty}B^{\circ}_{\varepsilon}(\vec{0})$ is open (and thus Borel measurable), and $m(Y + {}^{\infty}B^{\circ}_{\varepsilon}(\vec{0})) \geq \left(m_{out}(Y)^{\frac{1}{d}} + 2\varepsilon \right)^d$.*

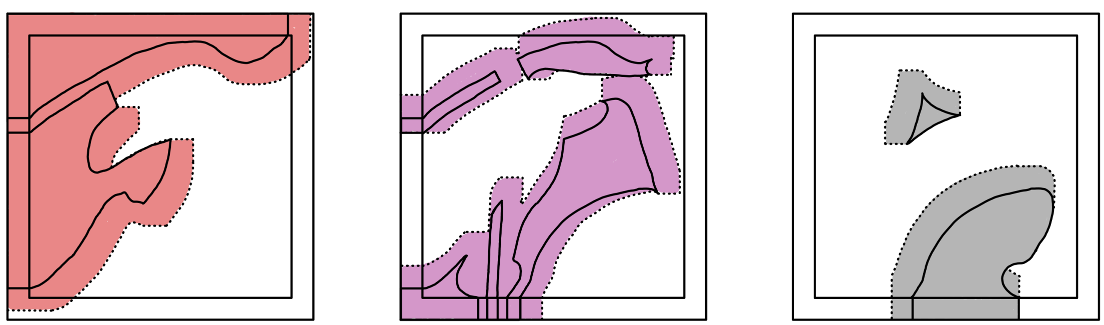
Now we prove the Neighborhood KKM/Lebesgue Theorem (restated for convenience). The proof is illustrated in Figure 8.1. The primary difference between this proof and the proof of Theorem 7.1.1 (ε-Neighborhoods for Measure Bounded Partitions and Arbitrary Norm) is that we have boundaries to be concerned with. This means we have to first create an extension of the initial coloring, and then we also cannot use a limiting argument, so obtain a worse bound of $(1 + \frac{2}{3}\varepsilon)^d$ than the bound of $(1 + 2\varepsilon)^d$ in Corollary 7.1.3 (ε-Neighborhoods for Diameter Bounded Partitions).



(a) Initial coloring χ (b) Extended coloring γ



(c) Red points (Y_{red}) (e) Purple points (Y_{purple}) (g) Gray points (Y_{gray})



(d) Ball added ($Y_{\text{red}} + B_{\vec{v}}$) (f) Ball added ($Y_{\text{purple}}+B_{\vec{v}}$) (h) Ball added ($Y_{\text{gray}} + B_{\vec{v}}$)

Figure 8.1: (a) shows a non-spanning coloring χ of the unit cube $[-\frac{1}{2}, \frac{1}{2}]^d$ for $d = 2$ (i.e. no color includes points on opposite edges). (b) shows the natural extension γ of that coloring to $[-\frac{1}{2} - \varepsilon, \frac{1}{2} + \varepsilon]^d$. The extension is obtained by mapping each point $\vec{y} \in [-\frac{1}{2} - \varepsilon, \frac{1}{2} + \varepsilon]^d$ to the point $\vec{x} \in [-\frac{1}{2}, \frac{1}{2}]^d$ for which each coordinate value is restricted to be within $[-\frac{1}{2}, \frac{1}{2}]$, and then $\vec{y}$ is given whatever color $\vec{x}$ had. (c), (e), and (g) show three of the five colors and demonstrate that there at least one quadrant of the ε-ball that can be Minkowski summed with the color so that the sum remains a subset of the extended cube. For red it is the lower right quadrant, for purple it is the upper right, and for gray it could be the upper left (shown) or the upper right. (d), (f), and (h) show the resulting Minkowski sum for each color. Utilizing the Brunn-Minkowski inequality, this set will have substantially greater area—by a factor of at least $(1 + \frac{\varepsilon}{1+\varepsilon})^d$.

Theorem 8.0.7 (Neighborhood KKM/Lebesgue Theorem)**.** *Given an non-spanning coloring of $[0,1]^d$, then for any $\varepsilon \in (0,\infty)$ there exists a point $\vec{p} \in [0,1]^d$ such that ${}^{\infty}B^{\circ}_{\varepsilon}(\vec{p})$ contains points of at least $\left\lceil \left(1+\frac{\varepsilon}{1+\varepsilon}\right)^d \right\rceil$ different colors. In particular, if $\varepsilon \in (0,\frac{1}{2}]$ then ${}^{\infty}B^{\circ}_{\varepsilon}(\vec{p})$ contains points of at least $\left\lceil \left(1+\frac{2}{3}\varepsilon\right)^d \right\rceil$ different colors.*

Proof. For convenience, we will assume that the cube is $[-\frac{1}{2},\frac{1}{2}]^d$ rather than $[0,1]^d$. Let C be a set (of colors) and $\chi\colon [-\frac{1}{2},\frac{1}{2}]^d \to C$ be a non-spanning coloring of the unit cube $[-\frac{1}{2},\frac{1}{2}]^d$. Let $C' = \text{range}(\chi)$ so that we know every color in C' appears for some point in the cube.

We first deal with the case where C' has infinite cardinality[5]. If C' has infinite cardinality, then because we can cover the cube with finitely many ε-balls, one of these balls must contain points of infinitely many colors, so the result is true. Thus, we assume from now on that C' has finite cardinality.

For each color $c \in C'$ we will let X_c denote the set of points assigned color c by χ—that is, $X_c = \chi^{-1}(c)$. Note that the hypothesis that no color includes points of opposite faces formally means that that for every color $c \in C'$, the set X_c has the property that for each coordinate $i \in [d]$, the projection $\pi_i(X_c) = \{x_i : \vec{x} \in X_c\}$ does not contain both $-\frac{1}{2}$ and $\frac{1}{2}$.

The first step in the proof is to extend the coloring χ to the larger cube $[-\frac{1}{2}-\varepsilon,\frac{1}{2}+\varepsilon]^d$ in a natural way. Consider the following function f which truncates points

[5]If one accepts the axiom of choice, then we don't need to deal with this as a special case, but by doing so, we can avoid requiring the axiom of choice in the proof.

in the larger interval to be in the smaller interval:

$$f\colon [-\tfrac{1}{2}-\varepsilon, \tfrac{1}{2}+\varepsilon] \to [-\tfrac{1}{2}, \tfrac{1}{2}]$$

$$f(y) \stackrel{\text{def}}{=} \begin{cases} -\frac{1}{2} & y \leq -\frac{1}{2} \\ y & y \in (-\frac{1}{2}, \frac{1}{2}) \\ \frac{1}{2} & y \geq \frac{1}{2} \end{cases}$$

Let $\vec{f}\colon [-\frac{1}{2}-\varepsilon, \frac{1}{2}+\varepsilon]^d \to [-\frac{1}{2}, \frac{1}{2}]^d$ be the function which is f in each coordinate: $\vec{f}(\vec{y}) \stackrel{\text{def}}{=} \langle f(y_i) \rangle_{i=1}^d$.

Now extend the coloring χ to the coloring $\gamma\colon [-\frac{1}{2}-\varepsilon, \frac{1}{2}+\varepsilon]^d \to C'$ defined by

$$\gamma(\vec{x}) \stackrel{\text{def}}{=} \chi\left(\vec{f}(\vec{x})\right).$$

For each color $c \in C'$, let $Y_c = \gamma^{-1}(c)$ denote the points assigned color c by γ and note that $X_c \subseteq Y_c$. Consistent with this notation, we will typically refer to a point in the unit cube as $\vec{x}$ and a point in the extended cube as $\vec{y}$.

We make the following claim which implies that for each color $c \in C'$, the set Y_c of points of that color in the extended coloring are contained in a set bounded away from one side of the extended cube $[-\frac{1}{2}-\varepsilon, \frac{1}{2}+\varepsilon]^d$ in each coordinate.

Claim A. *For each color $c \in C'$ there exists an orientation $\vec{v} \in \{-1,1\}^d$ such that $Y_c \subseteq \prod_{i=1}^d v_i \cdot (-\frac{1}{2}, \frac{1}{2}+\varepsilon]$.*

Proof of Claim. Fix an arbitrary coordinate $i \in [d]$. Note that for every $\vec{y} \in Y_c$ we have $\vec{f}(\vec{y}) \in X_c$ which is to say that the $\vec{y}$ has the same color in the extended coloring as $f(\vec{y})$ does in the original coloring (see justification[6]).

[6] For every $\vec{y} \in Y_c$ we have (by definition of Y_c) that $\gamma(\vec{y}) = c$ and (by definition of γ) that $\gamma(\vec{y}) = \chi(\vec{f}(\vec{y}))$ showing that $\chi(f(\vec{y})) = c$ and thus (by definition of X_c) that $f(\vec{y}) \in X_c$.

Note that if there is some $\vec{y} \in Y_c$ with $y_i \leq -\frac{1}{2}$, then $f(y_i) = -\frac{1}{2}$ so the fact that $X_c \ni \vec{f}(\vec{y})$ implies that $\pi_i(X_c) \ni f(y_i) = -\frac{1}{2}$. Similarly, if there is some $\vec{y} \in Y_c$ with $y_i \geq \frac{1}{2}$, then $\pi_i(X_c) \ni \frac{1}{2}$. Recall that by hypothesis, $\pi_i(X_c)$ does not contain both $-\frac{1}{2}$ and $\frac{1}{2}$ which means it is either the case that for all $\vec{y} \in Y_c$ we have $y_i > -\frac{1}{2}$ (so $\pi_i(Y_c) \subseteq (-\frac{1}{2}, \frac{1}{2} + \varepsilon])$ or it is the case that for all $\vec{y} \in Y_c$ we have $y_i < \frac{1}{2}$ (so $\pi_i(Y_c) \subseteq [-\frac{1}{2} - \varepsilon, \frac{1}{2})$).

Thus we can choose $v_i \in \{-1, 1\}$ such that $\pi_i(Y_c) \subseteq v_i \cdot (-\frac{1}{2}, \frac{1}{2} + \varepsilon]$. Since this is true for each coordinate $i \in [d]$ we can select $\vec{v} \in \{-1, 1\}^d$ such that

$$Y_c \subseteq \prod_{i=1}^{d} \pi_i(Y_c) \subseteq \prod_{i=1}^{d} v_i \cdot (-\tfrac{1}{2}, \tfrac{1}{2} + \varepsilon]$$

as claimed. ■

For an orientation $\vec{v} \in \{-1, 1\}^d$, let $B_{\vec{c}}$ denote the set $B_{\vec{v}} \stackrel{\text{def}}{=} \prod_{i=1}^{d} -v_i \cdot (0, \varepsilon)$ which should be interpreted as a on open orthant of the ℓ_∞ ε-ball centered at the origin—specifically the orthant opposite the orientation $\vec{v}$. Building on Claim A, we get the following:

Claim B. *For each color $c \in C'$, there exists an orientation $\vec{v} \in \{-1, 1\}^d$ such that $Y_c + B_{\vec{v}} \subseteq [-\frac{1}{2} - \varepsilon, \frac{1}{2} + \varepsilon]^d$.*

Proof of Claim. Let $\vec{v}$ be an orientation given in Claim A for color c. We get the following chain of containments:

$$\begin{aligned}
Y_c + B_{\vec{v}} &= Y_c + \left(\prod_{i=1}^{d} -v_i \cdot (0, \varepsilon)\right) && \text{(Def'n of } B_{\vec{v}}\text{)} \\
&\subseteq \left(\prod_{i=1}^{d} v_i \cdot (-\tfrac{1}{2}, \tfrac{1}{2} + \varepsilon]\right) + \left(\prod_{i=1}^{d} -v_i \cdot (0, \varepsilon)\right) && \text{(Claim A)} \\
&= \left(\prod_{i=1}^{d} v_i \cdot (-\tfrac{1}{2}, \tfrac{1}{2} + \varepsilon]\right) + \left(\prod_{i=1}^{d} v_i \cdot (-\varepsilon, 0)\right) && \text{(Factor a negative)} \\
&= \prod_{i=1}^{d} v_i \cdot (-\tfrac{1}{2} - \varepsilon, \tfrac{1}{2} + \varepsilon) && \text{(Minkowski sum of rectangles)} \\
&\subseteq [-\tfrac{1}{2} - \varepsilon, \tfrac{1}{2} + \varepsilon]^d. && (v_i \in \{-1, 1\})
\end{aligned}$$

This proves the claim. ■

We also claim that $Y_c + B_{\vec{v}}$ has substantial measure.

Claim C. *For each color $c \in C'$ and any orientation $\vec{v} \in \{-1,1\}^d$, the set $Y_c + B_{\vec{v}}$ is Borel measurable and $m(Y_c + B_{\vec{v}}) \geq m_{out}(Y_c) \cdot \left(1 + \frac{\varepsilon}{1+\varepsilon}\right)^d$.*

Proof of Claim. Let $M = (1+\varepsilon)^d$ which is the measure of $\prod_{i=1}^{d} v_i \cdot (-\frac{1}{2}, \frac{1}{2} + \varepsilon]$, and because by Claim A, Y_c is a subset of some such set, we have $m_{out}(Y_c) \leq M$.

We have that $Y_c + B_{\vec{v}}$ is Borel measurable and that $m\left(Y_c + B_{\vec{v}}\right) \geq \left(m_{out}(Y_c)^{\frac{1}{d}} + \varepsilon\right)^d$ by Corollary 8.1.1 (because $B_{\vec{v}}$ is some translation of ${}^{\infty}B^{\circ}_{\frac{\varepsilon}{2}}(\vec{0})$ and translations are irrelevant to the measure concerns of

Corollary 8.1.1). Thus, we have the following chain of inequalities:

$$\begin{aligned}
m(Y_c + B_{\vec{v}}) &\geq \left(m_{out}(Y_c)^{1/d} + \varepsilon\right)^d && \text{(Above)} \\
&= M \cdot \left(\frac{m_{out}(Y_c)^{1/d}}{M^{1/d}} + \frac{\varepsilon}{M^{1/d}}\right)^d && \text{(Factor out } M\text{)} \\
&\geq M \cdot \left(\frac{m_{out}(Y_c)}{M}\right) \cdot \left(1 + \frac{\varepsilon}{M^{1/d}}\right)^d && \text{(Fact G.0.2)} \\
&= m_{out}(Y_c) \cdot \left(1 + \frac{\varepsilon}{1+\varepsilon}\right)^d && \text{(Simplify and use } M = (1+\varepsilon)^d\text{)}
\end{aligned}$$

■

Now, consider the indexed family $\mathcal{A} = \left\{Y_c + B_{\vec{v}(c)}\right\}_{c \in C'}$ (where $\vec{v}(c)$ is an orientation for c as in Claim A and Claim B) noting that this family has finite cardinality because C' has finite cardinality. Considering the sum of measures of sets in $\mathcal{A}$, we have the following:

$$\begin{aligned}
\sum_{A \in \mathcal{A}} m(A) &= \sum_{c \in C'} m\left(Y_c + B_{\vec{v}(c)}\right) && \text{(Def'n of } \mathcal{A}\text{; measurability was shown above)} \\
&\geq \left(1 + \frac{\varepsilon}{1+\varepsilon}\right)^d \cdot \sum_{c \in C'} m_{out}(Y_c) && \text{(Claim C and linearity of summation)} \\
&\geq \left(1 + \frac{\varepsilon}{1+\varepsilon}\right)^d \cdot m_{out}\left(\bigcup_{c \in C'} Y_c\right) \\
& && \text{(Countable/finite subaddativity of outer measures)} \\
&= \left(1 + \frac{\varepsilon}{1+\varepsilon}\right)^d \cdot m_{out}\left([-\tfrac{1}{2} - \varepsilon, \tfrac{1}{2} + \varepsilon]^d\right) \\
& && \text{(The } Y_c\text{'s partition } [-\tfrac{1}{2} - \varepsilon, \tfrac{1}{2} + \varepsilon]^d\text{)} \\
&= \left(1 + \frac{\varepsilon}{1+\varepsilon}\right)^d \cdot (1 + 2\varepsilon)^d && \text{(Evaluate outer measure)}
\end{aligned}$$

By Claim B, each member of $\mathcal{A}$ is a subset of $[-\frac{1}{2}-\varepsilon, \frac{1}{2}+\varepsilon]^d$, so by Proposition A.3.1, there exists a point $\vec{p} \in [-\frac{1}{2}-\varepsilon, \frac{1}{2}+\varepsilon]^d$ that belongs to at least

$$\left\lceil \frac{\left(1+\frac{\varepsilon}{1+\varepsilon}\right)^d \cdot (1+2\varepsilon)^d}{(1+2\varepsilon)^d} \right\rceil = \left\lceil \left(1+\frac{\varepsilon}{1+\varepsilon}\right)^d \right\rceil$$

sets in $\mathcal{A}$. That is, $\vec{p}$ belongs to $Y_c + B_{\vec{v}(c)}$ for at least $\left\lceil \left(1+\frac{\varepsilon}{1+\varepsilon}\right)^d \right\rceil$ colors $c \in C'$. For each such color c, it follows that $\vec{p} + (-\varepsilon, \varepsilon)^d$ intersects Y_c (see justification[7]). Note that $\vec{p}+(-\varepsilon,\varepsilon)^d = {}^\infty B^\circ_\varepsilon(\vec{p})$ showing that ${}^\infty B^\circ_\varepsilon(\vec{p})$ contains points of at least $\left\lceil \left(1+\frac{\varepsilon}{1+\varepsilon}\right)^d \right\rceil$ colors (according to the coloring of γ since we are discussing sets Y_c).

What we really want, though, is a point in the unit cube that has this property rather than a point in the extended cube, and we want it with respect to the original coloring χ rather than the extended coloring γ. We will show that the point $\vec{f}(\vec{p})$ suffices.

Claim D. *If $c \in C'$ is a color for which ${}^\infty B^\circ_\varepsilon(\vec{p}) \cap Y_c \neq \emptyset$, then also ${}^\infty B^\circ_\varepsilon(\vec{f}(\vec{p})) \cap X_c \neq \emptyset$.*

Proof of Claim. Let $\vec{y} \in {}^\infty B^\circ_\varepsilon(\vec{p}) \cap Y_{\vec{c}}$. Then because $\vec{y} \in {}^\infty B^\circ_\varepsilon(\vec{p})$, we have $\|\vec{y}-\vec{p}\|_\infty < \varepsilon$, so for each coordinate $i \in [d]$ we have $|y_i - p_i| < \varepsilon$. It is easy to analyze the 9 cases (or 3 by symmetries) arising in the definition of f to see that this implies $|f(y_i) - f(p_i)| < \varepsilon$ as well (i.e. f maps pairs of values in its domain so that they are no farther apart), thus $\left\|\vec{f}(\vec{y}) - \vec{f}(\vec{p})\right\|_\infty < \varepsilon$ and thus $\vec{f}(\vec{y}) \in {}^\infty B^\circ_\varepsilon(\vec{f}(\vec{p}))$.

Also, as justified in a prior footnote[6], for any $\vec{y} \in Y_c$ we have $\vec{f}(\vec{y}) \in X_c$ so that $\vec{f}(\vec{y}) \in {}^\infty B^\circ_\varepsilon(\vec{f}(\vec{p})) \cap X_c$ which shows that the intersection is non-empty. ■

Thus, because ${}^\infty B^\circ_\varepsilon(\vec{p})$ intersects Y_c for at least $\left\lceil \left(1+\frac{\varepsilon}{1+\varepsilon}\right)^d \right\rceil$ choices of color $c \in C'$, by Claim D $\vec{f}(\vec{p})$ is a point in the unit cube for which ${}^\infty B^\circ_\varepsilon(\vec{f}(\vec{p}))$ intersects X_c for at

[7] If $\vec{p} \in Y_c + B_{\vec{v}(c)} \subseteq Y_c + (-\varepsilon,\varepsilon)^d$, then by definition of Minkowski sum there exists $\vec{y} \in Y_c$ and $\vec{b} \in (-\varepsilon,\varepsilon)^d$ such that $\vec{p} = \vec{y} + \vec{b}$ so $Y_c \ni \vec{y} = \vec{p} - \vec{b}$ and also $\vec{p} - \vec{b} \in \vec{p} + (-\varepsilon,\varepsilon)^d$ demonstrating that these two sets contain a common point.

least $\left\lceil \left(1+\frac{\varepsilon}{1+\varepsilon}\right)^d \right\rceil$ different colors $c \in C'$. That is, this ball contains points from at least this many of the original color sets.

The final step in the proof of the theorem is to clean up the expression with an inequality. Note that C' must contain of at least 2^d colors because each of the 2^d corners of the unit cube must be assigned a unique color since any pair of corners belong to an opposite pair of faces on the cube. For this reason it is trivial that for $\varepsilon > \frac{1}{2}$ there is a point $\vec{p}$ such that ${}^{\infty}B^{\circ}_{\varepsilon}(\vec{p})$ intersects at least 2^d colors: just let $\vec{p}$ be the midpoint of the unit cube. Thus, the only interesting case is $\varepsilon \in (0, \frac{1}{2}]$, and for such ε we have $1+\varepsilon \leq \frac{3}{2}$ and thus $\frac{\varepsilon}{1+\varepsilon} \geq \frac{2}{3}\varepsilon$ showing that $\left(1+\frac{\varepsilon}{1+\varepsilon}\right)^d \geq (1+\frac{2}{3}\varepsilon)^d$ which completes the proof.

□

8.2 Proof of the Neighborhood Sperner's Lemma

In this section, we present the discretized version of the Neighborhood KKM/Lebesgue Theorem which will be more analogous to Sperner's lemma on the cube. However, unlike typical Sperner's lemma variants, it will not be enough to have an arbitrary triangulation. Because we are working with an ε-ball, we need to have some information about how close together discrete points are within the unit cube. We define a notion of this below which matches the type we need to make our main result discrete.

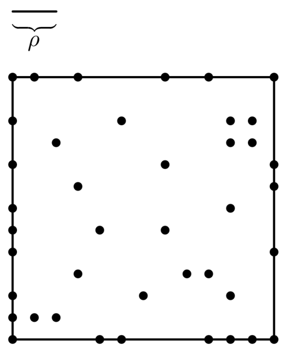
Definition 8.2.1 (ρ-Proximate Set). Let $\Lambda \subseteq [0,1]^d$. For $\rho \in [0,\infty)$, Λ is called ρ-proximate if for every face F of $[0,1]^d$ and for every $\vec{x} \in F$, there exists $\vec{y} \in F \cap \Lambda$ such that $\|\vec{x}-\vec{y}\|_{\infty} \leq \rho$.

Remark 8.2.2. We could have rephrased the last part of the definition as "... for every face F of $[0,1]^d$ and for every $\vec{x} \in F$, we have $F \cap \Lambda \cap {}^{\infty}\overline{B}_{\rho}(\vec{x}) \neq \emptyset$". Or we could have rephrased it as "... for every face F of $[0,1]^d$, we have $F \subseteq (F \cap \Lambda) + {}^{\infty}\overline{B}_\rho(\vec{0})$." △

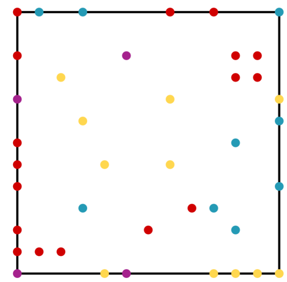
It should be unsurprising that the definition of ρ-proximate includes the property that every point of the cube is close to a point of Λ because we will want to show that some point $\vec{p}$ is ε-close to many colors, so we need to know that the color classes (and thus the points) aren't all mutually far apart (a set with only this requirement is called a ρ-net in some contexts). The condition that such points must belong to the same face may be less obvious, but probably not surprising considering the nature of Sperner's lemma and the KKM lemma; the reason we need such a property is demonstrated by the set $\Lambda = (0,1)^d$. If we didn't require that the point $\vec{y}$ in Definition 8.2.1 is in both F and Λ, then the set $\Lambda = (0,1)^d$ would be ρ-proximate for each $\rho \in (0,\infty)$, and yet it would be a valid non-spanning coloring to assign every point of Λ the same color, and thus we could not guarantee that some ε-ball intersects more than 1 color.

Remark 8.2.3. A consequence of the definition of ρ-proximate is that Λ contains all vertices of the cube (i.e. $\Lambda \supseteq \{0,1\}^d$). This is because for any vertex $\vec{v} \in \{0,1\}^d$, the singleton set $F = \prod_{i=1}^{d} \{v_i\} = \{\vec{v}\}$ is a face of the cube, so by definition of ρ-proximate, for each $\vec{x} \in F$, there is some associated $\vec{y} \in F \cap \Lambda$, so in particular $F \cap \Lambda \neq \emptyset$ and since F contains only the point $\vec{v}$, we must have $\vec{v} \in \Lambda$. △

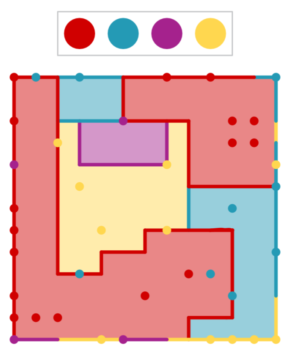
Now we restate and prove the Neighborhood Sperner's Lemma. Basically, we assign each point of $[0,1]^d$ to a color of a point in Λ nearby in a careful way and show that the resulting coloring still doesn't use the same color on opposite faces. We then use the Neighborhood KKM/Lebesgue Theorem to find the point $\vec{p}$ and pass back the result because Λ approximates $[0,1]^d$. Part of the proof is illustrated in Figure 8.2.



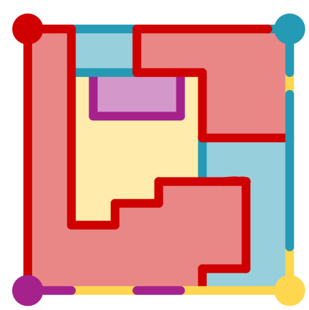
(a) A ρ-proximate set Λ with $\rho = \frac{1}{6}$



(b) Coloring $\chi : \Lambda \to \{R, B, P, Y\}$



(c) Sperner/KKM $\gamma : [0,1]^2 \to \{0,1\}^2$



(d) γ emphasized

Figure 8.2: (a) shows a set $\Lambda \subseteq [0,1]^2$ which is ρ-proximate for $\rho = \frac{1}{6}$ (i.e. every vertex is present, every point of each edge is within distance $\frac{1}{6}$ of another point on the same edge, and every point in the interior is within distance $\frac{1}{6}$ of some point). The distance ρ is shown visually in the upper left. (b) shows a non-spanning coloring of Λ (i.e. a function $\chi : \Lambda \to C = \{R_{ed}, B_{lue}, P_{urple}, Y_{ellow}\}$ in which no color is used on opposite faces). (c) shows how this coloring is used to produce a coloring of $[0,1]^2$ (i.e. a function $\gamma : [0,1]^d \to \{0,1\}^2$): an order is put on the set of colors C (in this case $R_{ed}, B_{lue}, P_{urple}, Y_{ellow}$ as shown at the top of (c)) and each point $\vec{x}$ of the cube is mapped to the first color in the ordering which belongs to the smallest face containing $\vec{x}$ and is within distance ρ. For example, if $\vec{x}$ is on an edge, then the only points considered are other points in Λ on the same edge which are distance at most ρ away. (d) clarifies the coloring γ by emphasizing the colors on the vertices, the edges, and the boundaries between colors.

Theorem 8.0.8 (Neighborhood Sperner's Lemma)**.** *Let $\rho \in [0,\infty)$ and $\Lambda \subseteq [0,1]^d$ be ρ-proximate. Let $\rho' = \min(\rho, \frac{1}{2})$. Given a non-spanning coloring of Λ, then for any $\varepsilon \in (0,\infty)$ there exists a point $\vec{p} \in [0,1]^d$ such that ${}^{\infty}B^{\circ}_{\varepsilon}(\vec{p})$ contains points of at least $\left\lceil \left(1 + \frac{(\varepsilon-\rho')}{1+(\varepsilon-\rho')}\right)^d \right\rceil$ different colors. In particular, if $\varepsilon \in (0, \frac{1}{2}]$ then ${}^{\infty}B^{\circ}_{\varepsilon}(\vec{p})$ contains points of at least $\left\lceil \left(1 + \frac{2}{3}(\varepsilon - \rho')\right)^d \right\rceil$ different colors.*

Proof. Let C be a set and $\chi : \Lambda \to C$ denote the coloring of Λ, and let $C' = \text{range}(\chi)$.

We first deal with the case where C' has infinite cardinality[8]. If C' has infinite cardinality, then because we can cover the cube (and thus Λ) with finitely many ε-balls, one of these balls must contain points of infinitely many colors, so the result is true. Thus, we assume from now on that C' has finite cardinality.

The next step in the proof is the following observation.

Claim A. *If $\rho \geq \frac{1}{2}$ then the fact that Λ is ρ-proximate implies that Λ is $\frac{1}{2}$-proximate.*

Proof of Claim. For any face F of the cube, we have $F = \prod_{i=1}^{d} F_i$ where each F_i is $\{0\}$, $\{1\}$, or $[0,1]$. This means that any $\vec{x} \in F$, we can find a vertex $\vec{v}$ of the cube (a point where each coordinate v_i is 0 or 1) which also belongs to F such that in each coordinate we have $|x_i - v_i| \leq \frac{1}{2}$. Since $\vec{v}$ is a vertex, by Remark 8.2.3, we have $\vec{v} \in \Lambda$. Thus, $\vec{v} \in F \cap \Lambda$ and $\|\vec{x} - \vec{v}\|_{\infty} \leq \frac{1}{2}$. ■

Thus, we may continue knowing that Λ is not only ρ-proximate but in fact ρ'-proximate where $\rho' = \min(\rho, \frac{1}{2})$. Next we comment on what happens if the term $(\varepsilon - \rho')$ is non-positive. Essentially, we have to take show that the expressions $(1 + \frac{(\varepsilon-\rho')}{1+(\varepsilon-\rho')})$ and $(1 + \frac{2}{3}(\varepsilon - \rho'))$ are never large negative values to make sure the bound we are giving does not take on large positive values when d is even.

[8] If one accepts the axiom of choice, then we don't need to deal with this as a special case, but by doing so, we can avoid requiring the axiom of choice in the proof.

Claim B. *The stated result holds when $\varepsilon \leq \rho'$.*

Proof of Claim. Note that when $\varepsilon \leq \rho'$, then because $\rho' \in [0, \frac{1}{2}]$ this implies $\varepsilon \in (0, \frac{1}{2}]$, so $(\varepsilon - \rho') \in (-\frac{1}{2}, 0]$. On the interval $x \in (-\frac{1}{2}, 0]$, the expression $1 + \frac{1}{1+x}$ is in $(0, 1]$, so we have $1 + \frac{(\varepsilon-\rho')}{1+(\varepsilon-\rho')} \in (0, 1]$ and thus $(1 + \frac{(\varepsilon-\rho')}{1+(\varepsilon-\rho')})^d \in (0, 1]$, so $\left\lceil (1 + \frac{(\varepsilon-\rho')}{1+(\varepsilon-\rho')})^d \right\rceil = 1$. Similarly, because $(\varepsilon - \rho') \in (-\frac{1}{2}, 0]$, we have $\frac{2}{3}(\varepsilon - \rho') \in (-\frac{1}{3}, 0]$, so $1 + \frac{2}{3}(\varepsilon - \rho') \in (\frac{2}{3}, 1]$, so $(1 + \frac{2}{3}(\varepsilon - \rho'))^d \in (0, 1]$, and again $\lceil (1 + \frac{2}{3}(\varepsilon - \rho'))^d \rceil = 1$. Because Λ is non-empty (by Remark 8.2.3) it is trivial to find a point where the ε ball contains points of at least 1 color showing that the result is true when $\varepsilon \leq \rho'$. ∎

Thus, we assume from now on that $\varepsilon > \rho'$ which implies $(\varepsilon - \rho') \in (0, \infty)$ (the hypothesis we need on the ball radius to apply the Neighborhood KKM/Lebesgue Theorem).

Next, for each $\vec{x} \in [0, 1]^d$, let $F^{(\vec{x})}$ denote the smallest face containing $\vec{x}$ (i.e. the intersection of all faces containing $\vec{x}$) and let $C^{(\vec{x})}$ denote the set of colors present in the face $F^{(\vec{x})}$ and within ρ' of $\vec{x}$ (formally, $C^{(\vec{x})} = \left\{ \chi(\vec{y}) : \vec{y} \in F^{(\vec{x})} \cap \Lambda \cap {}^{\infty}\overline{B}_{\rho'}(\vec{x}) \right\}$) noting that $C^{(\vec{x})}$ is non-empty by Remark 8.2.2.

Let $\gamma : [0, 1]^d \to C' \subseteq C$ map each $\vec{x}$ to some[9] color in $C^{(\vec{x})}$. We claim that γ is a non-spanning coloring so that we will be able to apply the Neighborhood KKM/Lebesgue Theorem to γ.

Claim C. *The coloring γ is does not assign the same color to points on opposite faces.*

Proof of Claim. We show that points on opposite faces are assigned different colors by γ. Let $F^{(0)}$ and $F^{(1)}$ be opposite faces of the cube—that is, there is some coordinate j such that $\pi_j(F^{(0)}) = \{0\}$ and $\pi_j(F^{(1)}) = \{1\}$.

[9] Because C' has finite cardinality, this does not require the axiom of choice.

Let $\vec{x}^{(0)} \in F^{(0)}$ be arbitrary. Because $F^{(\vec{x}^{(0)})}$ is the intersection of all faces containing $\vec{x}$, we have $\vec{x}^{(0)} \in F^{(\vec{x}^{(0)})} \subseteq F^{(0)}$. By definition of γ, there is some $\vec{y}^{(0)} \in F^{(\vec{x}^{(0)})} \cap \Lambda \subseteq F^{(0)} \cap \Lambda$ such that $\gamma(\vec{x}^{(0)}) = \chi(\vec{y}^{(0)})$. Similarly, there is some $\vec{y}^{(1)} \in F^{(1)} \cap \Lambda$ such that $\gamma(\vec{x}^{(1)}) = \chi(\vec{y}^{(1)})$.

By hypothesis of the coloring χ, because $\vec{y}^{(0)}$ and $\vec{y}^{(1)}$ belong to opposite faces of the cube (i.e. $F^{(0)}$ and $F^{(1)}$), we have $\chi(\vec{y}^{(0)}) \neq \chi(\vec{y}^{(0)})$ showing that $\gamma(\vec{x}^{(0)}) \neq \gamma(\vec{x}^{(1)})$. ■

The following claim says that for any point $\vec{p}$, all of the colors (of γ) appearing in the $(\varepsilon - \rho')$ ball at $\vec{p}$ also appear (as colors of χ) in Λ within the ε ball at $\vec{p}$. The connection back to Λ below is because for any $c \in C'$, we have $\chi^{-1}(c) \subseteq \Lambda$.

Claim D. *The following subset containment holds for each point $\vec{p} \in [0,1]^d$ (see comment[10]):*

$$\left\{c \in \text{range}(\gamma) : \gamma^{-1}(c) \cap {}^{\infty}B^{\circ}_{\varepsilon-\rho'}(\vec{p}) \neq \emptyset\right\} \subseteq \left\{c \in \text{range}(\chi) = C' : \chi^{-1}(c) \cap {}^{\infty}B^{\circ}_{\varepsilon}(\vec{p}) \neq \emptyset\right\}$$

Proof of Claim. If c belongs to the left set, then $\gamma^{-1}(c) \cap {}^{\infty}B^{\circ}_{\varepsilon-\rho'}(\vec{p}) \neq \emptyset$, so let $\vec{x} \in \gamma^{-1}(c) \cap {}^{\infty}B^{\circ}_{\varepsilon-\rho'}(\vec{p})$. Then $\gamma(\vec{x}) = c$ and using the defining property of γ that $\gamma(\vec{x}) \in C^{(\vec{x})}$, we have the following:

$$c = \gamma(\vec{x}) \in C^{(\vec{x})} = \left\{\chi(\vec{y}) : \vec{y} \in F^{(\vec{x})} \cap \Lambda \cap {}^{\infty}\overline{B}_{\rho'}(\vec{x})\right\}.$$

This means that there is some $\vec{y} \in \Lambda \cap {}^{\infty}\overline{B}_{\rho'}(\vec{x})$ such that $\chi(\vec{y}) = c$ (i.e. $\vec{y} \in \chi^{-1}(c)$). Also, for this $\vec{y}$, because $\|\vec{y} - \vec{x}\|_\infty \leq \rho'$ and $\|\vec{x} - \vec{p}\|_\infty < \varepsilon - \rho'$ we have by the triangle

[10] In the former set we express $c \in \text{range}(\gamma)$ rather than $c \in C'$, because it is possible that $\text{range}(\gamma) \subsetneq C'$. The coloring γ was defined as one of many choices, and in the natural choices we would have $\gamma(\vec{y}) = \chi(\vec{y})$ for each $\vec{y} \in \Lambda$, but we did not require this, so γ is not necessarily an extension of χ, and it is possible that it does not surject onto C'.

inequality that $\vec{y} \in {}^{\infty}B^{\circ}_{\varepsilon}(\vec{p})$. Thus, $\vec{y}$ demonstrates that $\chi^{-1}(c) \cap {}^{\infty}B^{\circ}_{\varepsilon}(\vec{p})$ is non-empty which shows that c belongs to the right set. ■

Now we can finish off the proof. By the Neighborhood KKM/Lebesgue Theorem (using the fact that $(\varepsilon - \rho') \in (0, \infty)$ along with Claim C), there exists $\vec{p} \in [0,1]^d$ such that ${}^{\infty}B^{\circ}_{\varepsilon-\rho'}(\vec{p})$ intersects at least $\left\lceil (1 + \frac{(\varepsilon-\rho')}{1+(\varepsilon-\rho')})^d \right\rceil$ colors (formally, $\left\{c \in \mathrm{range}(\gamma) : \gamma^{-1}(c) \cap {}^{\infty}B^{\circ}_{\varepsilon-\rho'}(\vec{p}) \neq \emptyset\right\}$ has cardinality at least $\left\lceil (1 + \frac{(\varepsilon-\rho')}{1+(\varepsilon-\rho')})^d \right\rceil$). Thus, by Claim D the set $\left\{c \in \mathrm{range}(\chi) = C' : \chi^{-1}(c) \cap {}^{\infty}B^{\circ}_{\varepsilon}(\vec{p}) \neq \emptyset\right\}$ also has cardinality at least $\left\lceil (1 + \frac{(\varepsilon-\rho')}{1+(\varepsilon-\rho')})^d \right\rceil$ which is what we set out to prove. (Informally, this latter set is the colors c for which there is a point in $\Lambda \cap {}^{\infty}B^{\circ}_{\varepsilon}(\vec{p})$ that is mapped to c by the original coloring χ.)

Finally, if $\varepsilon \in (0, \frac{1}{2}]$, then because we have at this point that $\varepsilon > \rho'$, we in fact have $\varepsilon - \rho' \in (0, \frac{1}{2}]$. Thus, by the same inequalities used in the proof of the Neighborhood KKM/Lebesgue Theorem, we have $\left\lceil \left(1 + \frac{(\varepsilon-\rho')}{1+(\varepsilon-\rho')}\right)^d \right\rceil \geq \left\lceil (1 + \frac{2}{3}(\varepsilon - \rho'))^d \right\rceil$ which completes the proof. □

We briefly remark that we have said this result is a neighborhood analog of Sperner's lemma on the cube, but in fact there are different ways to generalize Sperner's lemma to the cube. For example, [LPS01] considers subdividing the cube into simplices and using 2^d colors (because this is natural in the broader polytope setting in which they work) while [Kuh60] considers subdividing the cube into smaller cubes still with 2^d colors and [Wol77] considers a subdivision into cubes with either 2^d colors or with $d+1$ colors. Our generalization does not work for the cubical variants of Sperner's lemma using only $d+1$ colors (because we then can't possibly hope to intersect more than $d+1$ colors with an ε ball), but it does work for either simplicial subdivisions or cube subdivisions. Cube subdivisions are, however, more natural in our context because they align more nicely with the ℓ_∞

norm. In particular, note that the grid $\{0,\rho,2\rho,\ldots,(n-1)\rho,n\rho,1\}^d$ which naturally occurs in many contexts is ρ-proximate (even if ρ is not the reciprocal of an integer).

8.3 Discussion

In this section, we will give some discussion of the bound that we achieved in the Neighborhood KKM/Lebesgue Theorem (and equivalently in the Neighborhood Sperner's Lemma) including some limitations on improving that bound and some desires for future improvements of our result.

Consider replacing the expression "$\lceil(1+\frac{2}{3}\varepsilon)^d\rceil$" in the statement of the Neighborhood KKM/Lebesgue Theorem with a generic function "$k^\circ(d,\varepsilon)$" of the dimension and ball radius so we can have a discussion that is not specific to our bound but will apply to any improved bounds as well[11]:

Statement 8.3.1 (Neighborhood KKM/Lebesgue Generic Statement)**.** *Given a non-spanning coloring of* $[0,1]^d$*, then for any* $\varepsilon\in(0,\frac{1}{2}]$ *there exists a point* $\vec{p}\in[0,1]^d$ *such that* ${}^\infty B^\circ_\varepsilon(\vec{p})$ *contains points of at least* $k^\circ(d,\varepsilon)$ *different colors.*

We can assume for each $d\in\mathbb{N}$, that k° is non-decreasing in ε without any loss of generality[12] which is useful conceptually. For any (non-decreasing) k° for which the

[11] We denote the function by k° because one might also consider the same statement with respect to closed balls, and we might denote such a function with $\bar{k}$. Though we will not do so here, one could also consider intersections of balls not with the color sets but with the closures of the color sets which may or may not be more convenient with future techniques.

[12] We define $k^{\circ\prime}$ as $k^{\circ\prime}(d,\varepsilon)=\sup_{\varepsilon'\in(0,\varepsilon]}\lceil k^\circ(d,\varepsilon')\rceil$ and claim the statement also holds for $k^{\circ\prime}$ if it holds for k°. This is because for any $\varepsilon\in(0,\frac{1}{2}]$, there exists a point $\vec{p}\in[0,1]^d$ such that ${}^\infty B^\circ_\varepsilon(\vec{p})$ intersects at least $k^\circ(d,\varepsilon)$ color classes, and since the number of color classes intersected is an integer, ${}^\infty B^\circ_\varepsilon(\vec{p})$ intersects at least $\lceil k^\circ(d,\varepsilon)\rceil$ color classes. Furthermore, because $k^\circ(d,\varepsilon)\leq 2^d$ (see the footnote in the statement of Neighborhood KKM/Lebesgue Theorem (Theorem 8.0.7)), for any $\varepsilon\in(0,\frac{1}{2}]$ we have that the set $\{\lceil k^\circ(d,\varepsilon'):\varepsilon'\in(0,\varepsilon]\rceil\}$ is a bounded set of integers, so it contains its supremum and thus there exists some $\varepsilon'_0\in(0,\varepsilon]$ associated to ε with $\lceil k^\circ(d,\varepsilon'_0)\rceil=\sup_{\varepsilon'\in(0,\varepsilon]}\lceil k^\circ(d,\varepsilon')\rceil$.

Neighborhood KKM/Lebesgue Generic Statement holds, we can recover an analogous statement of the KKM/Lebesgue Theorem (Theorem 8.0.6) where the expression "$d+1$" is replaced with "$\lim_{\varepsilon\to 0} k^\circ(d,\varepsilon)$" using the compactness of the cube and the finiteness of the coloring (see Lemma B.2.1). Because the Lebesgue Covering Theorem and Cubical KKM Lemma are equivalent to our KKM/Lebesgue Theorem (see Section B.1) we also recover statements of those as well.

For this reason, it would be especially nice if for each dimension d we had $\lim_{\varepsilon\to 0} k^\circ(d,\varepsilon) = d+1$ as that would mean that our result was a strict generalization of the Lebesgue Covering Theorem and Cubical KKM Lemma. Unfortunately, our current bound has $\lim_{\varepsilon\to\infty}\left\lceil(1+\frac{2}{3}\varepsilon)^d\right\rceil = 1$ so our Neighborhood KKM/Lebesgue Theorem says literally nothing for very small ε.

Furthermore, let $K^\circ : \mathbb{N}\times(0,\infty)\to\mathbb{N}$ be the best possible function for which the Neighborhood KKM/Lebesgue Generic Statement holds. That is, $K^\circ(d,\varepsilon)$ is the pointwise maximum[13] of all functions k° for which the statement holds. Alternatively, we can express K° explicitly as

$$K^\circ(d,\varepsilon) \stackrel{\text{def}}{=} \max\left\{\kappa\in[2^d] : \begin{array}{l}\text{for every non-spanning coloring of}\\ [0,1]^d\text{, there exists a point } \vec{p}\in\\ [0,1]^d \text{ such that } {}^\infty B^\circ_\varepsilon(\vec{p}) \text{ intersects at}\\ \text{least } \kappa \text{ colors}\end{array}\right\}$$

(and we can define the function $\overline{K}$ similarly, but with respect to closed balls). It is trivially the case for all $d\in\mathbb{N}$ and $\varepsilon\in(0,\infty)$ that $\overline{K}(d,\varepsilon)\geq K^\circ(d,\varepsilon)$ because the

Thus, for each $\varepsilon\in(0,\frac{1}{2}]$, taking ε_0' as above, there exists some $\vec{q}\in[0,1]^d$ such that ${}^\infty B^\circ_{\varepsilon_0'}(\vec{q})$ intersects at least $k^\circ(d,\varepsilon_0')$ color classes, so as a superset we have that ${}^\infty B^\circ_\varepsilon(\vec{q})$ also intersects at least $\lceil k^\circ(d,\varepsilon_0')\rceil = \sup_{\varepsilon'\in(0,\varepsilon]}\lceil k^\circ(d,\varepsilon')\rceil = k^{\circ\prime}(d,\varepsilon)$ color classes. This shows that if the Neighborhood KKM/Lebesgue Generic Statement is true of k°, then it is also true of the monotonic non-decreasing $k^{\circ\prime}$.

[13] The maximum exists because for any $d\in\mathbb{N}$ and $\varepsilon\in(0,\infty)$ we have that $k^\circ(d,\varepsilon)\leq 2^d$ because there are non-spanning colorings of the cube satisfying the hypotheses with only 2^d colors.

closed ε-ball is a superset of the open ε-ball, and we also know that for each fixed $d \in \mathbb{N}$, both function K° and $\overline{K}$ are non-decreasing in ε because we could assume all k° and $\bar{k}$ had this property[12]. What else can we say about K° and $\overline{K}$? The information from the discussion that follows summarized in Table 8.1.

Immediate lower bound: The first obvious thing we can say is that for all $d \in \mathbb{N}$ and $\varepsilon \in (0, \infty)$, we know that $K^\circ(d, \varepsilon) \geq d + 1$ (ditto for $\overline{K}$) which follows straightforwardly from the KKM/Lebesgue Theorem taking care with the infinite cardinality[14].

Trivial tight bound for large ε: The second obvious thing we can say is that for any $d \in \mathbb{N}$ and $\varepsilon > \frac{1}{2}$, we have $K^\circ(d, \varepsilon) = 2^d$. This is because the open ℓ_∞ ε-ball placed at the center of the unit d-cube is a superset of the cube itself, so it intersects all 2^d corners—no two of which have the same color—so $K^\circ(d, \varepsilon) \geq 2^d$. And (by definition) $K^\circ(d, \varepsilon) \leq 2^d$ because there are non-spanning colorings with only 2^d colors. Similarly, for any $d \in \mathbb{N}$ and $\varepsilon \geq \frac{1}{2}$ (note the non-strict inequality this time), we have $\overline{K}(d, \varepsilon) = 2^d$.

Perspective of this paper: The third thing that we can say is that for all $d \in \mathbb{N}$ and $\varepsilon \in (0, \infty)$ we have $\overline{K}(d, \varepsilon) \geq K^\circ(d, \varepsilon) \geq \left\lceil (1 + \frac{\varepsilon}{1+\varepsilon})^d \right\rceil$ by the Neighborhood KKM/Lebesgue Theorem. In other words, the purpose of this paper is to put some lower bound on K° (and consequently on $\overline{K}$).

[14] One can "collapse" the possibly infinitely many colors into just 2^d colors to obtain a new finite non-spanning coloring, and by the Neighborhood KKM/Lebesgue Theorem, there is a point at the closure of at least $d+1$ of the collapsed colors, and any open set around this point intersects at least $d+1$ of the original color sets. See Lemma B.1.4 for example, where this result shows up as part of the larger proof.

Non-trivial tight bound for small ε: The fourth thing we can say is something that we don't believe is at all obvious; it turns out that we know the value of K° (and $\overline{K}$) exactly for all dimensions for a specific small regime of ε near 0. Specifically, for any $d \in \mathbb{N}$ and $\varepsilon \in (0, \frac{1}{2d}]$, we know that $d+1 \geq \overline{K}(d,\varepsilon) \geq K^\circ(d,\varepsilon)$; along with the "first obvious thing" we said, this gives equality with $d+1$. The reason for this is the following. We demonstrated in Theorem 4.2.18 (Existence of $(d+1, \frac{1}{2d})$-Secluded Unit Cube Partitions) that in each dimension d, there is a partition $\mathcal{P}_d$ of $\mathbb{R}^d$ consisting of translates of the half-open cube $[0,1)^d$ with the property that for every point $\vec{p} \in \mathbb{R}^d$, the closed ball ${}^{\infty}\overline{B}_{\frac{1}{2d}}(\vec{p})$ intersects at most $d+1$ cubes in the partition. This immediately gives a non-spanning coloring of $[0,1]^d$ with the same property: define the coloring $\chi : [0,1]^d \to \mathcal{P}_d$ by mapping each point in $[0,1]^d$ to the unique member of $\mathcal{P}_d$ which it belongs to. This is a non-spanning coloring because no member of $\mathcal{P}_d$ contains points distance 1 apart, and so no points distance 1 apart are given the same color; in particular, no points on opposite faces are given the same color. In fact, this coloring uses exactly 2^d colors[15] and consists of color sets which are rectangles[16]. This demonstrates the existence of (finite) non-spanning colorings for which no ε-ball intersects more than the $d+1$ colors even when taking ε as large as $\varepsilon = \frac{1}{2d}$. Once more for clarity: for any $d \in \mathbb{N}$ and $\varepsilon \in (0, \frac{1}{2d}]$ we have that $\overline{K}(d,\varepsilon) = K^\circ(d,\varepsilon) = d+1$.

[15]The range of χ can be shown to have cardinality exactly 2^d. This is because for any $X \in \mathcal{P}_d$ which intersects $[0,1]^d$, it follows by simple analysis that because X is a translate of $[0,1)^d$, it must be that X contains one of the corners of $[0,1]^d$. And since no member of $\mathcal{P}_d$ contains two corners of $[0,1]^d$ (because any two corners of $[0,1]^d$ are ℓ_∞ distance 1 apart, and no two points in a translate of $[0,1)^d$ are distance 1 apart), the subset of members of $\mathcal{P}_d$ which intersect $[0,1]^d$ (i.e. the range of χ) are in bijection with the 2^d corners of $[0,1]^d$.

[16]For each color, the set of points in $[0,1]^d$ of that color is $[0,1]^d$ intersected with some translate of $[0,1)^d$, and this intersection of a product of intervals is itself a product of intervals (i.e. a d-dimensional rectangle).

Non-trivial upper bound: We can generalize the result above by utilizing our unit cube partition products (Lemma 4.4.9 (Secluded Partition Product Guarantees)) rather than our initial constructions. Specifically, for each $d, n \in \mathbb{N}$ there is a partition $\mathcal{P}_{d,n}$ of $\mathbb{R}^d$ by translates of $[0,1)^d$ which is $\left(\frac{1}{2n}, (n+1)^{\lceil \frac{d}{n} \rceil}\right)$-secluded (i.e. for every $\vec{p} \in \mathbb{R}^d$, the closed ball ${}^{\infty}\overline{B}_{\frac{1}{2n}}(\vec{p})$ intersects at most $(n+1)^{\lceil \frac{d}{n} \rceil}$-many members of the partition)[17]. As before, this immediately gives rise to a non-spanning coloring of $[0,1]^d$ such that for every $\vec{p} \in [0,1]^d$, we have that ${}^{\infty}\overline{B}_{\frac{1}{2n}}(\vec{p})$ contains points of at most $(n+1)^{\lceil \frac{d}{n} \rceil}$ colors. Thus, (along with monotonicity) we have for all $d, n \in \mathbb{N}$ and $\varepsilon \in (0, \frac{1}{2n}]$ that $K^{\circ}(d, \varepsilon) \leq \overline{K}(d, \varepsilon) \leq (n+1)^{\lceil \frac{d}{n} \rceil}$. Taking $n = 1$, we recover the trivial upper bound of 2^d when $\varepsilon \in (0, \frac{1}{2}]$ (which is tight for $\overline{K}$ at $\frac{1}{2}$), and taking $n = d$ we recover the upper bound of $d+1$ when $\varepsilon \in (0, \frac{1}{2d}]$ (which we have already said is tight on this whole interval). Since we can freely choose n, this gives an upper bound on $\overline{K}$ and K° for every choice of d and ε, and based on the fact that it is tight at the extremes, we wonder if it is nearly tight everywhere.

We quickly note that we really only need to consider $n \in [d]$ and not $n \in \mathbb{N}$ because for $n > d$, we have $(n+1)^{\lceil \frac{d}{n} \rceil} > d+1$ and we already know this bound for $\varepsilon \leq \frac{1}{2n} < \frac{1}{2d}$, so we get no new information. We also emphasize that it is important for n to be an integer in the above bounds as is necessary in proving Lemma 4.4.9 (Secluded Partition Product Guarantees).

Difference between K° and $\overline{K}$: The final thing we will say about these functions is that $\overline{K}$ and K° are not in general equal, and in fact they differ noticeably when $\varepsilon = \frac{1}{2}$. We have already stated that $\overline{K}(d, \frac{1}{2}) = 2^d$, and we will now argue that $K^{\circ}(d, \frac{1}{2}) \leq 2^{d-1} + 1$. Consider first $d = 2$ along with Figure 8.3; it is obviously impossible for any positioning of the *open* ℓ_∞ $\frac{1}{2}$-ball to intersect all $2^d = 4$ colors

[17] Replace "$f(d)$" with "n" in Lemma 4.4.9 (Secluded Partition Product Guarantees).

because it can't include both of the colors which are single points as those points are ℓ_∞ distance 1 apart. This idea extends into higher dimensions by using a non-spanning coloring with 2^d colors which are identified with the vertices of the cube. Each of the 2^{d-1} colors in $\{0,1\}^d$ with even hamming weight are used only on the corresponding vertex (e.g. in Figure 8.3 these are $\langle 0,0\rangle$ (purple) and $\langle 1,1\rangle$ (blue)). Every other point on the cube is given a color with odd hamming weight; this is possible because aside from vertices (which are given their own color), every other point of the cube belongs to a face which is a superset of an edge of the cube, and every edge contains both an even and odd hamming weight vertex (by the standard definition of the d-dimensional hypercube graph) so there is at least 1 odd hamming weight color which can be used. This results in a KKM-style cover (but not with closed sets) which is in fact a non-spanning coloring because each point can be given a color of one of the vertices on the smallest face to which it belongs, and thus points on opposite faces are not given the same color. Such colorings have 2^{d-1} colors which are single points, so even the open ℓ_∞ $\frac{1}{2}$-ball cannot hit more than one of them because all of the vertices are distance 1 apart. Thus, the $\varepsilon = \frac{1}{2}$-ball can hit at most $2^{d-1}+1$ colors (possibly all of the odd hamming weight colors and one even hamming weight color), so $K^\circ(d,\frac{1}{2}) \leq 2^{d-1}+1$. And similarly, for $\varepsilon < \frac{1}{2}$ we have $\overline{K}(d,\varepsilon) \leq K^\circ(d,\frac{1}{2}) \leq 2^{d-1}+1$.

In light of this, we believe it is interesting to ask for each dimension d what the ranges of the functions $\overline{K}$ and K° are (as functions of ε).

Question 8.3.2. For each $d \in \mathbb{N}$, what is $\text{range}(\overline{K}(d,\cdot))$ and $\text{range}(K^\circ(d,\cdot))$? In particular, what are the cardinalities of these ranges?

We know that the range is trivially a subset of the integers between $d+1$ and 2^d as already discussed, but now we see that it is in fact a proper subset because $K^\circ(d,\cdot)$

and $\overline{K}(d,\cdot)$ are monotonic, so the hamming coloring discussion above shows that neither range includes any values strictly between $2^{d-1}+1$ and 2^d. We wonder if the functions $K^\circ(d,\cdot)$ and $\overline{K}(d,\cdot)$ are constant along these d intervals:

$$\left(0, \tfrac{1}{2d}\right), \quad \left(\tfrac{1}{2d}, \tfrac{1}{2(d-1)}\right), \quad \left(\tfrac{1}{2(d-1)}, \tfrac{1}{2(d-2)}\right), \quad \ldots, \quad \left(\tfrac{1}{8}, \tfrac{1}{6}\right), \quad \left(\tfrac{1}{6}, \tfrac{1}{4}\right), \quad \left(\tfrac{1}{4}, \tfrac{1}{2}\right).$$

If so, this might align nicely with the upper bounds that we obtained using Lemma 4.4.9 (Secluded Partition Product Guarantees) which gave a separate upper bound on each such interval.

We do at least know that $K^\circ(d,\cdot)$ is broken into intervals because it is left continuous (justified as follows). For any non-spanning coloring of $[0,1]^d$ and $\varepsilon \in (0,\infty)$, there is a point $\vec{p} \in [0,1]^d$ such that ${}^\infty B^\circ_\varepsilon(\vec{p})$ contains points of at least $K^\circ(d,\varepsilon)$ different colors. Pick one point of each color that is included (which is a finite number of points because $K^\circ(d,\varepsilon) \leq 2^d$). Because each of these finite number of points is contained in the *open* ball ${}^\infty B^\circ_\varepsilon(\vec{p})$, then for each $\varepsilon' < \varepsilon$ sufficiently large, we have that ${}^\infty B^\circ_{\varepsilon'}(\vec{p})$ contains all of these points, and so includes as many colors as ${}^\infty B^\circ_\varepsilon(\vec{p})$ does. Thus, $\lim_{\varepsilon' \uparrow \varepsilon} K^\circ(d,\varepsilon') \geq K^\circ(d,\varepsilon)$ and we get the other inequality by the monotonicity of K° showing that K° is right continuous:

$$\lim_{\varepsilon' \uparrow \varepsilon} K^\circ(d,\varepsilon') = K^\circ(d,\varepsilon).$$

Either by using a similar argument, or by noting that for all $\varepsilon' < \varepsilon$ we have $K^\circ(d,\varepsilon') \leq \overline{K}(d,\varepsilon') \leq K^\circ(d,\varepsilon)$ we have by a squeeze theorem that also

$$\lim_{\varepsilon' \uparrow \varepsilon} \overline{K}(d,\varepsilon') = K^\circ(d,\varepsilon).$$

Furthermore, on the interior of any interval $(\varepsilon_a, \varepsilon_b]$ for which $K^\circ(d,\cdot)$ is constant, we also know that $\overline{K}(d,\cdot)$ is constant because for $\varepsilon \in (\varepsilon_a, \varepsilon_b)$ we have ${}^{\infty}B^{\circ}_{\varepsilon_a}(\vec{0}) \subseteq {}^{\infty}\overline{B}_{\varepsilon}(\vec{0}) \subseteq {}^{\infty}B^{\circ}_{\varepsilon_b}(\vec{0})$ so $K^{\circ}(d,\varepsilon_a) \leq \overline{K}(d,\varepsilon) \leq K^{\circ}(d,\varepsilon_b)$.

ε	$K^{\circ}(d,\varepsilon)$	$\overline{K}(d,\varepsilon)$	Reason
$\in (0,\infty)$	$\leq 2^d$	$\leq 2^d$	Trivial
$\in (\frac{1}{2},\infty)$	$= 2^d$	$= 2^d$	Trivial
$= \frac{1}{2}$	$\leq 2^{d-1}+1$	$= 2^d$	Hamming coloring discussion
$\in (0,\infty)$	$\geq d+1$	$\geq d+1$	KKM/Lebesgue
$\in (0,\infty)$	$\geq (1+\frac{\varepsilon}{1+\varepsilon})^d$	$\geq (1+\frac{\varepsilon}{1+\varepsilon})^d$	Neighborhood KKM/Lebesgue
$\in (0,\frac{1}{2}]$	$\geq (1+\frac{2}{3}\varepsilon)^d$	$\geq (1+\frac{2}{3}\varepsilon)^d$	Neighborhood KKM/Lebesgue
$\in (0,\frac{1}{2d}]$	$= d+1$	$= d+1$	Theorem 4.2.18 & KKM/Lebesgue
$\in (0,\frac{1}{2n}], n \in \mathbb{N}$	$\leq (n+1)^{\lceil \frac{d}{n} \rceil}$	$\leq (n+1)^{\lceil \frac{d}{n} \rceil}$	Lemma 4.4.9

ε	$K^{\circ}(1,\varepsilon)$	$\overline{K}(1,\varepsilon)$	Reason
$\in (0,\infty)$	$= 2$	$= 2$	Combination of above bounds with $d=1$

ε	$K^{\circ}(2,\varepsilon)$	$\overline{K}(2,\varepsilon)$	Reason
$\in (0,\frac{1}{2})$	$= 3$	$= 3$	Combination of above bounds with $d=2$
$= \frac{1}{2}$	$= 3$	$= 4$	Combination of above bounds with $d=2$
$\in (\frac{1}{2},\infty)$	$= 4$	$= 4$	Combination of above bounds with $d=2$

Table 8.1: Known information about the ideal functions K° and $\overline{K}$.

Bounding the constant in the Neighborhood KKM/Lebesgue Theorem: A consequence of this last item is that any improvement to the Neighborhood KKM/Lebesgue Theorem using a bound of the form $(1+c\varepsilon)^d$ for some constant c must have the property that $(1+c\frac{1}{2})^d \leq 2^{d-1}+1$ for all d. Solving for c gives $c \leq 2\left(\left(2^{d-1}+1\right)^{1/d}-1\right)$. Graphing shows that $d=3$ is the integer where this

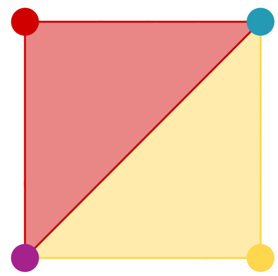


Figure 8.3: An example of a coloring in which each even hamming weight vertex has a color only used at that vertex. It is impossible for an $\varepsilon = \frac{1}{2}$ open ℓ_∞ ball to contain more than one of the even hamming weight colors because all vertices are distance 1 apart.

takes the smallest value and shows that $c \leq 1.42$. This is not an asymptotic claim; we are only saying that if we want to replace the constant $\frac{2}{3}$ in the Neighborhood KKM/Lebesgue Theorem with some other constant c, it must be that $c \leq 1.42$. In particular, we cannot hope to obtain the bound of $(1 + 2\varepsilon)^d$ (a bound which we were able to achieve in Corollary 7.1.8 for colorings (i.e. partitions) of $\mathbb{R}^d$ in which all color sets (i.e. partition members) had at most unit outer measure or at most unit diameter).

An explicit difference between $\mathbb{R}^d$ and $[0,1]^d$: The fact that we can't obtain the same bound on $[0,1]^d$ as we can in $\mathbb{R}^d$ is exemplified explicitly and exactly for $d = 2$ by our results. Table 8.1 shows that in $[0,1]^2$ it can be guaranteed that some translate of ${}^{\infty}\overline{B}_\varepsilon(\vec{0})$ intersects 4 colors in a non-spanning coloring if and only if $\varepsilon \geq \frac{1}{2}$. On the other hand, the combination of Theorem 4.2.18 (Existence of $(d+1, \frac{1}{2d})$-Secluded Unit Cube Partitions) and Proposition 7.2.11 (Optimal $\varepsilon = \frac{1}{2d}$ for Unit Diameter Partitions in $\mathbb{R}^2$) show that in $\mathbb{R}^2$ it can be guaranteed that some translate of ${}^{\infty}\overline{B}_\varepsilon(\vec{0})$ intersects 4 colors in a unit ℓ_∞ diameter bounded partition if and only if $\varepsilon > \frac{1}{4}$.

Conclusion: All of this demonstrates that the bound of $(1+\frac{\varepsilon}{1+\varepsilon})^d$ (or $(1+\frac{2}{3}\varepsilon)^d$) in our Neighborhood KKM/Lebesgue Theorem is not in general tight. For ε tending to 0, it predicts 1 even though the exact bound is $d+1$ for all $\varepsilon \in (0, \frac{1}{d}]$ (though we can get this bound by applying the KKM/Lebesgue Theorem), and it predicts only slightly more than $(1+\frac{1}{3})^d$ for ε slightly bigger than $\frac{1}{2}$ when the tight bound is 2^d. For these reasons, we suspect that our bound is not tight for any value of ε, and we hope that future work is able to establish better lower bounds on $K^\circ(d,\varepsilon)$ and $\overline{K}(d,\varepsilon)$. In addition, we hope that future work can either improve our upper bounds on these functions or prove that they are nearly tight.

We believe it would be particularly satisfying if all of this can be done with a single uniform technique; for example, the current lower bounds on K° and $\overline{K}$ for small ε follow from the Cubical KKM Lemma or the Lebesgue Covering Theorem (summarized in the KKM/Lebesgue Theorem) which are generally proved by very different techniques than those we used to prove our Neighborhood KKM/Lebesgue Theorem which provides the lower bounds on K° and $\overline{K}$ large ε.

Chapter 9

Secluded Partitions Without ε

In this chapter, we continue the study of secluded partitions, but we extend the definition to a new case. So far we have considered (k, ε)-secluded partitions which are parameterized by the two parameters k and ε (Definition 3.1.1). In this section we consider k-secluded partitions (Definition 3.1.2 restated below): for every point $\vec{p}$, there exists some ε such that ${}^{\infty}\overline{B}_{\varepsilon}(\vec{p})$ intersects at most k members of the partition. The distinction is necessary because there exist partitions which for some fixed $k_0 \in \mathbb{N}$ are k_0-secluded but not (k_0, ε)-secluded for any value of $\varepsilon \in (0, \infty)$. See Figure 9.1 for example.

Definition 3.1.2 (k-Secluded). Let $d, k \in \mathbb{N}$ and $\mathcal{F}$ a family of subsets of $\mathbb{R}^d$. We say that $\mathcal{F}$ is *k-secluded* if for each $\vec{p} \in \mathbb{R}^d$, there exists $\varepsilon \in (0, \infty)$ such that $\left| {}^{\infty}\overline{\mathcal{N}}_{\varepsilon}(\vec{p}) \right| \leq k$.

Remark 9.0.1. By Theorem 3.5.1 (Equivalence of Norms on $\mathbb{R}^d$), the ℓ_∞ norm in Definition 3.1.2 could be replaced with any other norm to recover the same definition.

△

We also have an equivalent definition in the cases we are interested in.

Lemma 9.0.2 (Equivalent Definitions of k-Secluded For Unit Cubes)**.** *Let $d \in \mathbb{N}$ and $\mathcal{F}$ be a packing of axis-aligned unit cubes in $\mathbb{R}^d$. For any $k \in \mathbb{N}$, the following are equivalent:*

1. *$\mathcal{F}$ is k-secluded.*
2. *For all $\vec{p} \in \mathbb{R}^d$ there exists $\varepsilon \in (0,\infty)$ such that $|{}^{\infty}\overline{\mathcal{N}}_{\varepsilon}(\vec{p})| \leq k$.*
3. *For all $\vec{p} \in \mathbb{R}^d$ we have $\mathcal{N}_{\overline{0}}(\vec{p}) \leq k$.*
4. *For all $\mathcal{C} \subseteq \mathcal{E} \sqcup \{X\}$ such that $\mathcal{C}$ is a cube clique, we have $|\mathcal{C}| \leq k$.*

Proof.

(1) $\iff$ (2): This is by definition.

(2) $\implies$ (3): This is trivial because for any $\varepsilon \in (0,\infty)$ we have $\mathcal{N}_{\overline{0}}(\vec{p}) \subseteq {}^{\infty}\overline{\mathcal{N}}_{\varepsilon}(\vec{p})$.

(3) $\implies$ (4): For a clique $\mathcal{C}$ of axis-aligned unit cubes there exists a clique-point $\vec{p}$ by Lemma 3.4.11 (Clique-Points Exist for Axis-Aligned Unit Cube Cliques) so that each member in $\mathcal{C}$ contains $\vec{p}$ in its closure and thus $\mathcal{C} \subseteq \mathcal{N}_{\overline{0}}(\vec{p})$, so $|\mathcal{C}| \leq |\mathcal{N}_{\overline{0}}(\vec{p})| \leq k$.

(4) $\implies$ (2): For any $\vec{p} \in \mathbb{R}^d$, consider the clique $\mathcal{N}_{\overline{0}}(\vec{p})$ (which by hypothesis has cardinality at most k) and by Corollary 3.6.6 (Unit Cube Packings are Locally Finite) and Fact 3.6.5 (Locally Finite: Enlarged Neighborhood) there exists $\varepsilon \in (0,\infty)$ such that $\mathcal{N}_{\overline{0}}(\vec{p}) = {}^{\infty}\overline{\mathcal{N}}_{\varepsilon}(\vec{p})$ so we have $\left|{}^{\infty}\overline{\mathcal{N}}_{\varepsilon}(\vec{p})\right| = |\mathcal{N}_{\overline{0}}(\vec{p})| \leq k$. $\square$

We discussed in Chapter 1 (Introduction), and will see in more detail in Chapter 10 (Computer Science Applications), that the (k,ε)-secluded partitions have applications in theoretical computer science, and this seems to primarily be the case because of the ε parameter. In absence of that, we suspect that the k-secluded partitions will not have the same level of applicability to computer science (though we obviously can never know for sure), but the k-secluded partitions can nonetheless be studied with a purely mathematical motivation. We will see in this chapter that the $(d+1)$-secluded partitions of axis-aligned unit cubes are quite beautiful because they have a

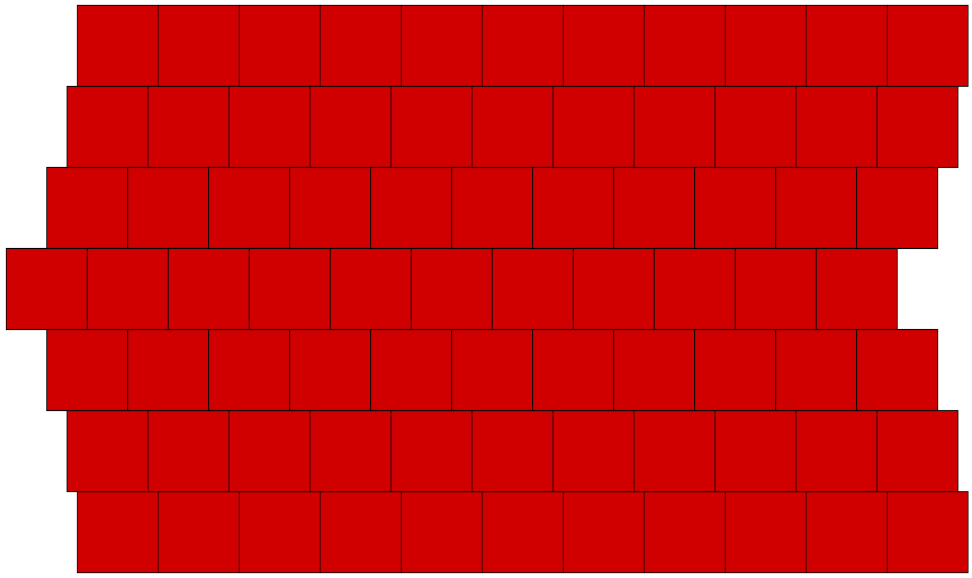


Figure 9.1: A partition of $\mathbb{R}^2$ which is 3-secluded, but is not $(3,\varepsilon)$-secluded for any $\varepsilon \in (0,\infty)$. The row above the middle has a shift of $\frac{1}{2}$, then the next row has a shift of $\frac{1}{4}$, then the next has a shift of $\frac{1}{8}$, and the pattern continues (beyond what is shown) for a shift of $\frac{1}{2^i}$. The same happens below the middle row. To see that it is not $(3,\varepsilon)$-secluded, fix any $\varepsilon \in (0,\infty)$. There is a point with large enough y-coordinate such that the x-shift from one layer to the next is too small so there is a point $\vec{p}$ where $^{\infty}\overline{B}_{\varepsilon}(\vec{p})$ intersects 4 members.

tremendous amount of structure and turn out to have numerous equivalent definitions which are not immediately obvious and take a lot of exploration to identify. We will also see that by moving away from considering the ε-parameter, we can place more focus on the partition graph (Definition 3.3.2). The partition graph contains no distance information, so it was less useful when studying (k,ε)-secluded partitions, but in the case of k-secluded axis-aligned unit cube partitions, the partition graph contains basically all of the structural information about the partition that we desire.

The chapter is laid out as follows. In Section 9.1 (Notation) we introduce a few additional notational conventions that we will use throughout the chapter. In Section 9.2 (Cliques and Cube Enclosures) we introduce the definition of a cube enclosure—the main structure other than partitions that we will want to study in this chapter—and prove some of the basic properties. In Section 9.3 (Approximation Structure of Adjacency) we present one lemma which demonstrates in gory detail the exact structural ways in which two cubes can be adjacent; the use of this result will be ubiquitous throughout the remainder of the chapter. In Section 9.4 (The Notion of Cousins) we utilize the prior lemma to get a higher level view of some different types of adjacency that occur in $(d+1)$-secluded families of cubes. We generalize the notion of a Minkowski twin pair in a way that lets us track the structure of $(d+1)$-secluded families of cubes in an effective way. In Section 9.5 (Cliques and Cube Enclosures Revisited) we make use of the notion of cousins to prove much more complex properties about cliques and cube enclosures. It is in this section that we see that $(d+1)$-secluded cube enclosures are cube enclosures with the minimum possible cardinality, and we prove the first of three main results in this chapter which gives many equivalent properties to a cube enclosure having minimum cardinality (Theorem 9.5.8 (Minimum Cube Enclosure Equivalencies)). In Section 9.6 (Characterizing the Difference Between Minimal and Secluded) we explore why the reverse implication is not true (minimum cardinality cube enclosures are in general not $(d+1)$-secluded) and discuss the work of Zaks which is critical for obtaining the partition equivalence result of the next section. In Section 9.7 (Partition Implications) we pass the results about cube enclosures over to partitions and prove the second main result of the chapter which shows three quite different equivalent ways to define a $(d+1)$-secluded cube partition (Theorem 9.7.4 ($(d+1)$-Secluded Iff Minimum Degree Iff Neighborly)). This result implies that not only are the $(d+1)$-secluded cube partitions exactly those that minimize the size of the largest clique in the partition graph, but equivalently they are exactly those that minimize the largest vertex degree in the partition graph.

Finally, in Section 9.8 (Optimal ε For Unit Cube Enclosures) we prove (independently of the results in Section 9.6 and Section 9.7) the third main result of this chapter: $\varepsilon = \frac{1}{2d}$ is the maximum possible value for a (k,ε)-secluded unit cube partition when k is taken to be minimized at $k = d+1$ (Theorem 9.8.4 ($\varepsilon = \frac{1}{2d}$ Optimal for Unit Cube Enclosures)).

9.1 Notation

- When a family, $\mathcal{F}$, of subsets of $\mathbb{R}^d$ is understood, and $X \in \mathcal{F}$, we will use the notation $\mathcal{N}(X)$ to mnemonically stands for "neighbors of X" and mean the set of all members of $\mathcal{F}$ which are adjacent to X. Formally, $\mathcal{N}(X) \overset{\text{def}}{=} \left\{ Y \in \mathcal{F} : X \overset{\text{adj}}{\sim} Y \right\}$. When $\mathcal{F}$ is a partition (which is how we will typically use it), this implies by the definition of adjacency (Definition 3.3.1) that $\mathcal{N}(X) = \left\{ Y \in \mathcal{F} : \overline{X} \cap \overline{Y} \neq \emptyset \right\}$.
- For a point $\vec{c} \in \mathbb{R}^d$ and value $\varepsilon \in (0,\infty)$, we use the notation $A_\varepsilon(\vec{c})$ to mean the corners of the ε radius cube at $\vec{c}$ (formally, $A_\varepsilon(\vec{c}) \overset{\text{def}}{=} \text{corners}\left({}^{\infty}\overline{B}_\varepsilon(\vec{c}) \right) = \vec{c} + \{-\varepsilon, \varepsilon\}^d$).
- For an axis-aligned unit cube X and parameter $\varepsilon \in (0,\infty)$ we will use the notation $A_\varepsilon(X)$ which mnemonically stands for "approximate corners" to indicate the set of corners of all ε radius cubes centered at each corner of X (see Figure 9.3); formally, using a Minkowski sum and letting $\vec{x} = \text{center}(X)$

we define in many equivalent forms

$$\begin{aligned}
A_\varepsilon(X) &\stackrel{\text{def}}{=} \operatorname{corners}(X) + \operatorname{corners}\left({}^{\infty}\overline{B}_\varepsilon(\vec{0})\right) \\
&= \bigsqcup_{\vec{c}\in\operatorname{corners}(X)} \operatorname{corners}\left({}^{\infty}\overline{B}_\varepsilon(\vec{c})\right) \\
&= \bigsqcup_{\vec{c}\in\operatorname{corners}(X)} A_\varepsilon(\vec{c}) \\
&= \bigsqcup_{\vec{c}\in\operatorname{corners}(X)} \left(\vec{c} + \{-\varepsilon, \varepsilon\}^d\right) \\
&= \prod_{i=1}^{d} \left\{x_i - \tfrac{1}{2} - \varepsilon, \quad x_i - \tfrac{1}{2} + \varepsilon, \quad x_i + \tfrac{1}{2} - \varepsilon, \quad x_i + \tfrac{1}{2} + \varepsilon\right\} \\
&= \prod_{i=1}^{d} \left\{x_i \pm \tfrac{1}{2} \pm \varepsilon\right\}
\end{aligned}$$

9.2 Cliques and Cube Enclosures

While partitions have been our primary interest to this point, it turns out that much of what we wish to study in regards to k-secluded axis-aligned unit cube partitions can be accomplished by studying two local structures in the partition. The first local structure will be to consider a single point $\vec{p} \in \mathbb{R}^d$ and consider $\mathcal{N}_{\overline{0}}(\vec{p})$. Observe that by definition, $\bigcap_{X\in\mathcal{N}_0(\vec{p})} \overline{X} \ni \vec{p}$, so all pairs of cubes in this set are adjacent (because they are also members of a partition so have disjoint interiors), and thus $\mathcal{N}_{\overline{0}}(\vec{p})$ is a clique in the partition (and also in the partition graph (Definition 3.3.2 (Set Family Graph/Partition Graph))). The second local structure will be to fix a cube X in the partition and consider $\mathcal{N}(X)$ (which can also be interpreted as the set of all vertices adjacent to vertex X in the partition graph). We will prove shortly (Fact 9.2.6) the unsurprising fact that X is completely enclosed by its set of neighbors (formally, $\overline{X} \subseteq \operatorname{int}\left(\overline{X} \cup \bigcup_{Y\in\mathcal{N}(X)} \overline{Y}\right)$). We will refer to such a collection as a cube enclosure (upcoming Definition 9.2.3).

While these two structures naturally fall out of partitions, they can be defined and studied independently outside of the context of a partition, and that is what we will do. The advantages are twofold: (1) we only focus on the essential structure and remove noise from the mathematics, and (2) there are properties that we will find about cube enclosures in partitions which are not properties of cube enclosures themselves but are actually emergent properties when $\mathbb{R}^d$ is partitioned by cubes. Thus, studying the structures in their own right allows us to distinguish the intrinsic properties from the emergent ones. The following are the general definitions of axis-aligned cube cliques and cube enclosures.

Definition 9.2.1 (Axis-Aligned Cube Clique). Let $d \in \mathbb{N}$. An *axis-aligned cube clique* in $\mathbb{R}^d$ is a set $\mathcal{C}$ of axis-aligned unit cubes such that for any distinct $X, Y \in \mathcal{C}$, it holds that $X \overset{\text{adj}}{\sim} Y$.

Remark 9.2.2. This is just Definition 3.3.3 (Clique) restricted to axis-aligned unit cubes. △

Definition 9.2.3 (Axis-Aligned Cube Enclosure). Let $d \in \mathbb{N}$. An *axis-aligned cube enclosure* in $\mathbb{R}^d$ is a pair $(X, \mathcal{E})$ where X is an axis-aligned unit cube in $\mathbb{R}^d$ and $\mathcal{E}$ is a family of axis-aligned unit cubes in $\mathbb{R}^d$ such that the following hold:

1. For all $Y \in \mathcal{E}$, $X \overset{\text{adj}}{\sim} Y$
2. For all distinct $Y, Y' \in \mathcal{E}$, $\operatorname{int}(Y) \cap \operatorname{int}(Y') = \emptyset$
3. $\overline{X} \subseteq \operatorname{int}\left(X \cup \bigcup_{Y \in \mathcal{E}} Y\right)$

The use of $\mathcal{E}$ is a mnemonic for "enclosure." Condition (1) requires that all elements of the enclosure $\mathcal{E}$ are adjacent to X and condition (2) requires that no cubes in the enclosure overlap with each other. Condition (3) ensures that the cubes in the enclosure $\mathcal{E}$ actually fully enclose X so that every point of $\overline{X}$ (i.e. every point in X or its boundary) is contained either in X or in one of the cubes in the enclosure $\mathcal{E}$.

Because formally a cube enclosure is a tuple, we clarify how we will use the terms "(k, ε)-secluded" and "k-secluded" for cube enclosures.

Definition 9.2.4 ((k, ε)-Secluded and k-Secluded Cube Enclosures). We say that an axis-aligned unit cube enclosure is (k, ε)-*secluded* (resp. k-*secluded*) if the set family $\mathcal{F} = \mathcal{E} \cup \{X\}$ is (k, ε)-secluded (resp. k-secluded), and use the notations $\mathcal{N}_{\overline{0}}(\vec{p})$, ${}^{\infty}\overline{\mathcal{N}}_{\varepsilon}(\vec{p})$, and ${}^{\infty}\mathcal{N}^{\circ}_{\varepsilon}(\vec{p})$ to be taken with respect to $\mathcal{F}$.

Fact 9.2.5 (Packing Neighborhoods Give Cube Cliques)**.** *Let $d \in \mathbb{N}$ and $\mathcal{F}$ a packing of axis-aligned unit cubes in $\mathbb{R}^d$. Then for any point $\vec{p} \in \mathbb{R}^d$, the set $\mathcal{N}_{\overline{0}}(\vec{p})$ is an axis-aligned unit cube clique.*

Proof. By definition of $\mathcal{N}_{\overline{0}}(\vec{p})$, for every distinct $X, Y \in \mathcal{N}_{\overline{0}}(\vec{p})$, we have $\vec{p} \in \overline{X}$ and $\vec{p} \in \overline{Y}$ so the closures intersect, and because $\mathcal{F}$ is a packing, X and Y have disjoint interiors, so $X \overset{\text{adj}}{\sim} Y$. □

Fact 9.2.6 (Partition Neighbors Give Cube Enclosures)**.** *Let $d \in \mathbb{N}$ and $\mathcal{P}$ a partition of $\mathbb{R}^d$ by axis-aligned unit cubes. Then for any cube $X \in \mathcal{P}$, the pair $(X, \mathcal{N}(X))$ is an axis-aligned unit cube enclosure.*

Proof. Condition (1) of the definition of cube enclosure is trivial because $\mathcal{N}(X) \overset{\text{def}}{=} \left\{Y \in \mathcal{P} : X \overset{\text{adj}}{\sim} Y\right\}$ (Section 9.1 (Notation)). Condition (2) is also trivial because $\mathcal{P}$ is a partition and $\mathcal{N}(X) \subseteq \mathcal{P}$. For condition (3), we will consider an arbitrary point $\vec{x} \in \overline{X}$ and show that there is some $\varepsilon \in (0, \infty)$ such that ${}^{\infty}B^{\circ}_{\varepsilon}(\vec{x}) \subseteq \left(X \cup \bigcup_{Y \in \mathcal{N}(X)} Y\right)$ to show that $\vec{x}$ belongs to the interior of said set.

Let $\vec{x} \in \overline{X}$ be arbitrary. Note that $\mathcal{N}_{\overline{0}}(\vec{x}) \subseteq \mathcal{N}(X) \sqcup \{X\}$ because if $Y \in \mathcal{N}_{\overline{0}}(\vec{x})$, then by definition $\vec{x} \in \overline{Y}$ so $\overline{X} \cap \overline{Y} \ni \vec{x}$ and is not empty; either $Y = X$ (and so trivially $Y \in \mathcal{N}(X) \sqcup \{X\}$) or $Y \neq X$ in which case because $\mathcal{P}$ is a partition $X \cap Y = \emptyset$, so $X \overset{\text{adj}}{\sim} Y$ and thus $Y \in \mathcal{N}(X)$ (and so again $Y \in \mathcal{N}(X) \sqcup \{X\}$).

By Corollary 3.6.6 (Unit Cube Packings are Locally Finite) and Fact 3.6.5 (Locally Finite: Enlarged Neighborhood), there exists $\varepsilon \in (0, \infty)$ such that ${}^{\infty}\mathcal{N}_{\varepsilon}^{\circ}(\vec{x}) = \mathcal{N}_{\overline{0}}(\vec{x})$ and by Fact 3.7.1 (Neighborhood Expression by Members) ${}^{\infty}\mathcal{N}_{\varepsilon}^{\circ}(\vec{x}) = \{\text{member}(\vec{x}) : \vec{x} \in {}^{\|\cdot\|}B_{\varepsilon}^{\circ}(\vec{x})\}$.

It follows that ${}^{\infty}B_{\varepsilon}^{\circ}(\vec{x}) \subseteq \left(X \cup \bigcup_{Y \in \mathcal{N}(X)} Y\right)$ because for $\vec{z} \in {}^{\infty}B_{\varepsilon}^{\circ}(\vec{x})$ we then have

$$\text{member}(\vec{z}) \in {}^{\infty}\mathcal{N}_{\varepsilon}^{\circ}(\vec{x}) = \mathcal{N}_{\overline{0}}(\vec{x}) \subseteq \mathcal{N}(X) \sqcup \{X\}$$

so

$$\vec{y} \in \bigcup_{Y \in \mathcal{N}(X) \sqcup \{X\}} Y = \left(X \cup \bigcup_{Y \in \mathcal{N}(X)} Y\right)$$

which completes the proof. □

Now we mention some initial properties about cube cliques and cube enclosures in Subsection 9.2.1 and Subsection 9.2.2 respectively before moving into much more elaborate study.

9.2.1 Properties of Axis-Aligned Unit Cube Cliques

We have already discussed in Chapter 3 (Preliminaries) an important property of axis-aligned unit cube cliques—there always exists a point common to the closure of every cube in the clique. Recall the following three (restated) definitions and results.

Definition 3.3.3 (Clique). Let $d \in \mathbb{N}$ and $\mathcal{C}$ a family of subsets of $\mathbb{R}^d$. Using the language from graph theory, $\mathcal{C}$ is called a *clique* if for all pairs of distinct $X, Y \in \mathcal{C}$ we have $X \overset{\text{adj}}{\sim} Y$. Equivalently, $\mathcal{C}$ is called a *clique* if its induced set family graph is a clique in the graph theoretic sense.

Definition 3.4.10 (Clique-Point). Let $d \in \mathbb{N}$, and $\mathcal{C}$ be an axis-aligned unit cube clique in $\mathbb{R}^d$. A point $\vec{p}$ is called a *clique-point* of $\mathcal{C}$ if for all $X \in \mathcal{C}$ we have that $\vec{p} \in \overline{X}$ (i.e. $\vec{p} \in \bigcap_{X \in \mathcal{C}} \overline{X}$).

Lemma 3.4.11 (Clique-Points Exist for Axis-Aligned Unit Cube Cliques)**.** *Let $d \in \mathbb{N}$ and $\mathcal{C}$ be an axis-aligned unit cube clique in $\mathbb{R}^d$ (i.e. a clique containing only axis-aligned unit cubes). Then $\mathcal{C}$ has at least one clique-point.*

Another result about axis-aligned unit cube cliques which follows from [DIP05] is that any cube clique (even if it was not defined within the context of a partition) can be *extended* to a partition, and even more so to a periodic partition.

Corollary 9.2.7 (Cube Cliques Extend to Periodic Partitions)**.** *Let $d \in \mathbb{N}$ and $\mathcal{C}$ be an axis-aligned unit cube clique in $\mathbb{R}^d$ where each cube is a translate of $[0,1)^d$. Then there exists a $2\mathbb{Z}^d$-periodic[a] axis-aligned unit cube partition $\mathcal{P}$ of $\mathbb{R}^d$ such that $\mathcal{C} \subseteq \mathcal{P}$.*

[a] This means that if $X \in \mathcal{P}$, then $2\vec{n} + X \in \mathcal{P}$ for all $\vec{n} \in \mathbb{Z}^d$.

Proof. By Fact 3.6.8 (Equivalence of Axis-Aligned Partitions and Tilings), we can assume each cube in $\mathcal{C}$ is a translate of the closed unit cube $[0,1]^d$ and show that there exists a $2\mathbb{Z}^d$-periodic axis-aligned unit cube tiling $\mathcal{T}$ such that $\mathcal{C} \subseteq \mathcal{T}$.

By Lemma 3.4.11 (Clique-Points Exist for Axis-Aligned Unit Cube Cliques), there exists $\vec{p} \in \bigcap_{X \in \mathcal{C}} \overline{X}$ which implies that each cube $X \in \mathcal{C}$ must be contained in

$\vec{p} + [-1, 1]^d$ (see justification[1]). By the definitions of clique, adjacent, and packing (Definition 3.3.3, Definition 3.3.1, Definition 3.2.1) $\mathcal{C}$ is a packing which is contained in $\vec{p} + [-1, 1]^d$. By [DIP05, Theorem 4] and the remark that follows, (and scaling their result[2]) we have that $\mathcal{C}$ can be extended to a $2\mathbb{Z}^d$-periodic axis-aligned unit cube tiling $\mathcal{T}$ which definitionally means that $\mathcal{C} \subseteq \mathcal{T}$. □

A much more trivial result (stated as a fact) about cube cliques is an upper bound of 2^d on their cardinality and this is tight as witnessed by using 2^d axis-aligned unit cubes positioned at all points of $\{0, 1\}^d$.

Fact 9.2.8 (Maximum Size Axis-Aligned Cube Clique)**.** *Let $d \in \mathbb{N}$ and $\mathcal{C}$ be an axis-aligned unit cube clique in $\mathbb{R}^d$. Then $|\mathcal{C}| \leq 2^d$.*

Proof. As in the proof of Corollary 9.2.7 (Cube Cliques Extend to Periodic Partitions), there exists $\vec{p} \in \mathbb{R}^d$ such that $\mathcal{C}$ is a packing contained in $\vec{p} + [-1, 1]^d$. The interior of each cube has Lebesgue measure 1, and $\vec{p} + [-1, 1]^d$ has Lebesgue measure 2^d, so $\mathcal{C}$ can contain at most 2^d cubes. □

9.2.2 Properties of Axis-Aligned Unit Cube Enclosures

As with cube cliques, cube enclosures are packings.

Fact 9.2.9 (Cube Enclosures Are Cube Packings)**.** *Let $d \in \mathbb{N}$ and $(X, \mathcal{E})$ a cube enclosure in $\mathbb{R}^d$. Then $\mathcal{E} \sqcup \{X\}$ is a cube packing.*

Proof. By definition of cube enclosure, each $Y \in \mathcal{E}$ is adjacent to X, and by definition of adjacency (Definition 3.3.1), that means X has disjoint interior from each $Y \in \mathcal{E}$.

[1] For any $X \in \mathcal{C}$, we have $\vec{p} \in \overline{X} = {}^{\infty}\overline{B}_{1/2}(\text{center}(X))$ which means the center of X is distance at most $1/2$ from $\vec{p}$ and all point of $\overline{X}$ are distance at most $1/2$ from the center. Thus, by the triangle inequality, all points of $\overline{X}$ are ℓ_∞ distance at most 1 from $\vec{p}$.

[2] They consider a packing of $[-2, 2]^d$ by translates of $[-1, 1]^d$ and we consider packings of $[-1, 1]^d$ by translates of [1]

Also, for every pair $Y, Y' \in \mathcal{E}$ we have by the definition of cube enclosure that Y and Y' have disjoint interiors. Thus, $\mathcal{E} \sqcup \{X\}$ is a packing. □

The following result says that not only is $\overline{X} = {}^{\infty}\overline{B}_{\frac{1}{2}}(\text{center}(X))$ contained in the interior as in the definition of a unit cube enclosure, but actually a strictly larger closed cube ${}^{\infty}\overline{B}_{\frac{1}{2}+\varepsilon}(\text{center}(X))$ is for some $\varepsilon \in (0, \infty)$.

Lemma 9.2.10 (Enclosure Enlargment)**.** *Let $d \in \mathbb{N}$ and $(X, \mathcal{E})$ be a cube enclosure in $\mathbb{R}^d$. Then there exists $\varepsilon \in (0, \infty)$ such that ${}^{\infty}\overline{B}_{\frac{1}{2}+\varepsilon}(\text{center}(X)) \subseteq \text{int}\left(X \cup \bigcup_{Y \in \mathcal{E}} Y\right)$.*

Proof. Suppose otherwise. Then for any $n \in \mathbb{N}$, there exists some $\vec{z}^{(n)} \in {}^{\infty}\overline{B}_{\frac{1}{2}+\frac{1}{n}}(\text{center}(X))$ but not in $\text{int}\left(X \cup \bigcup_{Y \in \mathcal{E}} Y\right)$. This gives an infinite sequence of points all of which are contained in ${}^{\infty}\overline{B}_{\frac{1}{2}+1}(\text{center}(X))$ which is a compact set, so by the Bolzano-Weirstrass theorem, there exists a subsequence converging to some value $\vec{z}$. Because each $\vec{z}^{(n)}$ also belongs to $\mathbb{R}^d \setminus \text{int}\left(X \cup \bigcup_{Y \in \mathcal{E}} Y\right)$ (which is a closed set) so must $\vec{z}$ (so $\vec{z} \notin \text{int}\left(X \cup \bigcup_{Y \in \mathcal{E}} Y\right)$).

Finally, for each $n \in \mathbb{N}$ all points in the tail of the sequence belong to ${}^{\infty}\overline{B}_{\frac{1}{2}+\frac{1}{n}}(\text{center}(X))$, and thus so does $\vec{z}$. Because this holds for all $n \in \mathbb{N}$ we have $\vec{z} \in {}^{\infty}\overline{B}_{\frac{1}{2}}(\text{center}(X)) = \overline{X}$. This shows that $\overline{X} \not\subseteq \text{int}\left(X \cup \bigcup_{Y \in \mathcal{E}} Y\right)$ which contradicts the definition of a cube enclosure. □

The next result is a corollary of Fact 9.2.9 (Cube Enclosures Are Cube Packings) and very similar in spirit. It says for the most part, we have no need to distinguish what type of cubes are used in a cube enclosure.

Corollary 9.2.11 (Cube Enclosures Closed or Half-Open)**.** *Let $d \in \mathbb{N}$, and $(X, \mathcal{E})$ be an axis-aligned unit cube enclosure in $\mathbb{R}^d$. Let $X_{clos} = \overline{X}$ and $\mathcal{E}_{clos} = \left\{\overline{Y} : Y \in \mathcal{E}\right\}$ be the closed versions of all cubes, and let $X_{half} = H_{1/2}(\text{center}(X))$ and $\mathcal{E}_{half} = \left\{H_{1/2}(\text{center}(Y)) : Y \in \mathcal{E}\right\}$ be the half-open versions of all cubes. Then both $(X_{clos}, \mathcal{E}_{clos})$ and $(X_{half}, \mathcal{E}_{half})$ are axis-aligned unit cube enclosures and*

$$\operatorname{int}\left(\bigcup_{Y \in \mathcal{E} \sqcup \{X\}} Y\right) \subseteq \operatorname{int}\left(\bigcup_{Y \in \mathcal{E}_{clos} \sqcup \{X_{clos}\}} Y\right) = \operatorname{int}\left(\bigcup_{Y \in \mathcal{E}_{half} \sqcup \left\{X_{half}\right\}} Y\right).$$

Furthermore, all cubes in $\mathcal{E}_{half} \sqcup \{X_{half}\}$ are disjoint.

Proof. Note that the first two conditions (1) and (2) in the definition of a cube enclosure (Definition 9.2.3) do not depend on the type of cube (i.e. open, closed, half-open, or something between), so all we have to show is that condition (3) of the definition holds for $(X_{clos}, \mathcal{E}_{clos})$ and $(X_{half}, \mathcal{E}_{half})$; in other words, it suffices to show the containment chain in the statement. The containment chain in the statement along with the "furthermore" claim follow from Corollary 3.6.7 because $(X, \mathcal{E})$ is a packing of axis-aligned unit cubes by Fact 9.2.9 (Cube Enclosures Are Cube Packings).

□

Remark 9.2.12 (Closed or Half-Open Assumption). It follows from Corollary 9.2.11 that any proof in which we want to argue about the cardinality of $\mathcal{E}$ in an axis-aligned unit cube enclosure $(X, \mathcal{E})$ in $\mathbb{R}^d$, we may assume without loss of generality that X and every cube in $\mathcal{E}$ is a translate of $[0, 1)^d$ and they are all disjoint because we can replace it with the cube enclosure $(X_{half}, \mathcal{E}_{half})$ which has $|\mathcal{E}_{half}| = |\mathcal{E}|$ and consists of disjoint cubes. This is useful because then $\mathcal{E}_{half} \sqcup \{X_{half}\}$ partitions the set $\bigcup_{Y \in \mathcal{E}_{half} \sqcup \{X_{half}\}} Y$ so we get the benefits of working with a partition. Similarly, we may assume all cube are half-open when proving any property that only depends on positions of the cubes (e.g. the cousins relationship which we later define).

By the same argument, we can assume X and every cube in $\mathcal{E}$ is closed which will sometimes be more convenient. △

We conjecture that, as with cube cliques (Corollary 9.2.7), every cube enclosure can be extended to a partition.

Conjecture 9.2.13 (Cube Enclosures Extend to Partitions)**.** *Let $d \in \mathbb{N}$ and $(X, \mathcal{E})$ be an axis-aligned unit cube enclosure in $\mathbb{R}^d$ where X and each cube in $\mathcal{E}$ is a translate of $[0,1)^d$. Then there exists an axis-aligned unit cube partition $\mathcal{P}$ of $\mathbb{R}^d$ such that $X \in \mathcal{P}$ and $\mathcal{E} \subseteq \mathcal{P}$.*

Remark 9.2.14. We cannot use the same techniques as the proof of Corollary 9.2.7 (Cube Cliques Extend to Periodic Partitions) because no axis-aligned unit cube enclosure is contained in a translate of $[-1,1]^d$ (see justification[3]). Furthermore, it was shown in the appendix of [LRW00] that there are non-trivial unit cube packings in $\mathbb{R}^d$ for all $d \geq 3$ which cannot be extended to tilings/partitions, so it is not obvious that cube enclosures are extendable. △

[3]Let $(X, \mathcal{E})$ be a cube enclosure in $\mathbb{R}^d$ and $\vec{x} = \text{center}(X)$. We want to show that $\bigcup_{Y \in \mathcal{E} \sqcup \{X\}} Y$ is not a subset of any translate of $[-1,1]^d$. In fact, for every coordinate $i \in [d]$ it holds that $\pi_i \left(\bigcup_{Y \in \mathcal{E} \sqcup \{X\}} Y \right)$ is an interval of length 3. Informally, for coordinate $i \in [d]$, consider any line passing through X which is aligned with the ith axis (i.e. $\vec{e}^{(i)}$). The intersection of X with this line is a unit interval (possibly closed or open or neither depending on X, but this is irrelevant). Because of property (3) of the definition of a cube enclosure, there must also be cubes containing points arbitrarily close to the left and right endpoints of this interval. Since the intersection of each of these cubes with the line is also a unit interval, we have three consecutive unit intervals. This shows that $\pi_i \left(\bigcup_{Y \in \mathcal{E} \sqcup \{X\}} Y \right)$ must also have length at least 3, and it can be shown that the length will be exactly 3. Thus, the smallest cube which contains all cubes in the cube enclosure is a cube of side length 3.

As with unit cube cliques, we have a trivial result (stated as a fact) about cube enclosures which is that there is an upper bound of 3^d on their cardinality and this is tight as witnessed by using 3^d axis-aligned unit cubes positioned at all points of $\{-1,0,1\}^d$ where the cube at the origin is X and the others constitute $\mathcal{E}$. We will later also prove an lower bound on the size of a cube enclosure of $2^{d+1}-2$ (Theorem 9.5.5).

Fact 9.2.15 (Maximum Size Axis-Aligned Cube Enclosure)**.** *Let $d \in \mathbb{N}$ and $(X, \mathcal{E})$ be an axis-aligned unit cube enclosure in $\mathbb{R}^d$. Then $|\mathcal{E}| \le 3^d - 1$.*

Proof. For each $Y \in \mathcal{E}$ we have by definition of a cube enclosure that $X \overset{\text{adj}}{\sim}$, so by Fact 3.4.6 (Adjacency for Unit Cubes) we have $\|\text{center}(X) - \text{center}(Y)\|_\infty = 1$ and because $Y \subseteq {}^\infty\overline{B}_{1/2}(\text{center}(Y))$ (by Definition 3.4.4 (Unit Cube Center and Representative)) we have by the triangle inequality that $Y \subseteq {}^\infty\overline{B}_{3/2}(\text{center}(X))$. With some of the same reasoning $X \subseteq {}^\infty\overline{B}_{1/2}(\text{center}(X)) \subseteq {}^\infty\overline{B}_{3/2}(\text{center}(X))$.

Thus $\left(\bigcup_{Y \in \mathcal{E} \sqcup \{X\}} Y\right) \subseteq {}^\infty\overline{B}_{3/2}(\text{center}(X))$, and because this has Lebesgue measure 3^d and the interiors of each cube in $\mathcal{E} \sqcup \{X\}$ are disjoint and have Lebesgue measure 1, we have $|\mathcal{E} \sqcup \{X\}| \le 3^d$. □

9.3 Approximation Structure of Adjacency

In this section we present only one result, but it will be critical for the rest of this chapter. Very roughly, we want to focus on some axis-aligned unit cube X along with some adjacent axis-aligned unit cube Y and consider in gory analytic detail how the adjacency relationship between them is structured. In particular, we will want to do two things. First, we will want to consider which corners of the cube X belong to the closure of the cube Y. Second, we will want to approximate each corner of X with 2^d points distance ε away, and consider which of those points belong to Y. These ideas are demonstrated in Figure 9.3 (which will be referenced later in more detail). Neither the statement nor the proof of Lemma 9.3.1 (ε-Approximation Structure of Adjacency) (the main result of this section) are difficult, but they are long and tedious. Though it is unfortunately extremely verbose, it says absolutely nothing surprising. The purpose is to give an explicit expression for four different set intersections that we will be interested in throughout the rest of the chapter. Each intersection is expressed as a product set, and the value of each coordinate of these product sets depends on one of 5 cases (3 cases with symmetries). All of these cases are shown in Figure 9.2. The statement of the result is broken up over two pages.

Lemma 9.3.1 (ε-Approximation Structure of Adjacency)**.** *Let $d \in \mathbb{N}$ and X, Y be axis-aligned unit cubes in $\mathbb{R}^d$ such that $X \overset{\text{adj}}{\sim} Y$. Let $\vec{x} = \text{center}(X)$ and $\vec{y} = \text{center}(Y)$. Then*

$$\text{corners}(X) \cap \overline{Y} = \prod_{i=1}^{d} \begin{cases} \left\{x_i + \frac{1}{2}\right\} = \left\{y_i - \frac{1}{2}\right\} & |x_i - y_i| = 1 \textit{ and } x_i < y_i \\ \left\{x_i - \frac{1}{2}\right\} = \left\{y_i + \frac{1}{2}\right\} & |x_i - y_i| = 1 \textit{ and } x_i > y_i \\ \left\{x_i - \frac{1}{2}, \quad x_i + \frac{1}{2}\right\} & |x_i - y_i| = 0 \\ \left\{x_i + \frac{1}{2}\right\} & |x_i - y_i| \in (0,1) \textit{ and } x_i < y_i \\ \left\{x_i - \frac{1}{2}\right\} & |x_i - y_i| \in (0,1) \textit{ and } x_i > y_i \end{cases} \tag{9.1}$$

Also, for all sufficiently small $\varepsilon > 0$*,*

$$\begin{aligned}
&A_\varepsilon(X) \cap \operatorname{int}(Y) \\
=&A_\varepsilon(X) \cap Y \\
=&A_\varepsilon(X) \cap \overline{Y} \\
=&\prod_{i=1}^{d} \begin{cases} \left\{x_i + \frac{1}{2} + \varepsilon\right\} & |x_i - y_i| = 1 \text{ and } x_i < y_i \\ \left\{x_i - \frac{1}{2} - \varepsilon\right\} & |x_i - y_i| = 1 \text{ and } x_i > y_i \\ \left\{x_i - \frac{1}{2} + \varepsilon, \quad x_i + \frac{1}{2} - \varepsilon\right\} & |x_i - y_i| = 0 \\ \left\{x_i + \frac{1}{2} - \varepsilon, \quad x_i + \frac{1}{2} + \varepsilon\right\} & |x_i - y_i| \in (0,1) \text{ and } x_i < y_i \\ \left\{x_i - \frac{1}{2} - \varepsilon, \quad x_i - \frac{1}{2} + \varepsilon\right\} & |x_i - y_i| \in (0,1) \text{ and } x_i > y_i \end{cases}
\end{aligned} \tag{9.2}$$

Also, for all $\varepsilon \in (0,1)$*,*

$$\begin{aligned}
&A_\varepsilon(X) \cap \operatorname{int}(X) \\
=&A_\varepsilon(X) \cap X \\
=&A_\varepsilon(X) \cap \overline{X} \\
=&\prod_{i=1}^{d} \left\{x_i - \tfrac{1}{2} + \varepsilon, x_i + \tfrac{1}{2} - \varepsilon\right\}
\end{aligned} \tag{9.3}$$

Also, for sufficiently small $\varepsilon > 0$*, for any* $\vec{p} \in \operatorname{corners}(X) \cap \overline{Y}$*,*

$$\begin{aligned}
&A_\varepsilon(\vec{p}) \cap \operatorname{int}(Y) \\
=&A_\varepsilon(\vec{p}) \cap Y \\
=&A_\varepsilon(\vec{p}) \cap \overline{Y} \\
=&\prod_{i=1}^{d} \begin{cases} \{p_i + \varepsilon \cdot \operatorname{sign}(p_i - x_i)\} & |x_i - y_i| = 1 \\ \{p_i - \varepsilon \cdot \operatorname{sign}(p_i - x_i)\} & |x_i - y_i| = 0 \\ \{p_i - \varepsilon, \quad p_i + \varepsilon\} & |x_i - y_i| \in (0,1) \end{cases}
\end{aligned} \tag{9.4}$$

Proof. We will prove that $A_\varepsilon(X) \cap \operatorname{int}(Y) = A_\varepsilon(X) \cap \overline{Y}$ and has the form specified above. Since $\operatorname{int}(Y) \subseteq Y \subseteq \overline{Y}$, this suffices to show that $A_\varepsilon(X) \cap Y$ does as well. Ditto for the other claims.

We begin by proving the third claim as it is trivial. For $\varepsilon \in (0,1)$,

$$A_\varepsilon(X) \cap \operatorname{int}(X) = \prod_{i=1}^{d} \left\{x_i \pm \tfrac{1}{2} \pm \varepsilon\right\} \cap \left(x_i - \tfrac{1}{2}, x_i + \tfrac{1}{2}\right) = \prod_{i=1}^{d} \left\{x_i - \tfrac{1}{2} + \varepsilon, x_i + \tfrac{1}{2} - \varepsilon\right\}$$

and

$$A_\varepsilon(X) \cap \overline{X} = \prod_{i=1}^{d} \left\{x_i \pm \tfrac{1}{2} \pm \varepsilon\right\} \cap \left[x_i - \tfrac{1}{2}, x_i + \tfrac{1}{2}\right] = \prod_{i=1}^{d} \left\{x_i - \tfrac{1}{2} + \varepsilon, x_i + \tfrac{1}{2} - \varepsilon\right\}$$

Now we prove the remainder of the claims. Let $\vec{p} \in \operatorname{corners}(X) \cap \overline{Y}$ be arbitrary. We proceed by letting $i \in [d]$ be arbitrary and examining each case noting that by Fact 3.4.6, $|x_i - y_i| \in [0,1]$. Within this proof, let (the admittedly ugly notation) $a \mp\!\!\!\!\pm\, b$ denote the three numbers $a - b$, a, $a + b$ (similar to how $\pm$ is used).

1. If $|x_i - y_i| = 1$, and $x_i < y_i$ then it must be that $y_i = x_i + 1$, so the following holds for any $\varepsilon \in (0, \frac{1}{2})$:

$$\begin{aligned}
\pi_i(A_\varepsilon(X)) \cap \pi_i(\operatorname{int}(Y)) &= \left\{x_i \pm \tfrac{1}{2} \pm \varepsilon\right\} \cap (y_i - \tfrac{1}{2}, y_i + \tfrac{1}{2}) \\
&= \left\{x_i \pm \tfrac{1}{2} \pm \varepsilon\right\} \cap (x_i + \tfrac{1}{2}, x_i + \tfrac{3}{2}) \\
&= \left\{x_i + \tfrac{1}{2} + \varepsilon\right\}
\end{aligned}$$

and similarly with the closure

$$\begin{aligned}\pi_i(A_\varepsilon(X)) \cap \pi_i(\overline{Y}) &= \left\{x_i \pm \tfrac{1}{2} \pm \varepsilon\right\} \cap [y_i - \tfrac{1}{2}, y_i + \tfrac{1}{2}] \\ &= \left\{x_i \pm \tfrac{1}{2} \pm \varepsilon\right\} \cap [x_i + \tfrac{1}{2}, x_i + \tfrac{3}{2}] \\ &= \left\{x_i + \tfrac{1}{2} + \varepsilon\right\}\end{aligned}$$

and

$$\begin{aligned}\pi_i(\mathrm{corners}(X)) \cap \pi_i(\overline{Y}) &= \left\{x_i \pm \tfrac{1}{2}\right\} \cap [y_i - \tfrac{1}{2}, y_i + \tfrac{1}{2}] \\ &= \left\{x_i \pm \tfrac{1}{2}\right\} \cap [x_i + \tfrac{1}{2}, x_i + \tfrac{3}{2}] \\ &= \left\{x_i + \tfrac{1}{2}\right\}.\end{aligned}$$

Since $\vec{p} \in \mathrm{corners}(X) \cap \overline{Y}$, this last line also shows that $p_i = x_i + \frac{1}{2}$ (so $\mathrm{sign}(p_i - x_i) = +1$), thus

$$\begin{aligned}\pi_i(A_\varepsilon(\vec{p}) \cap \mathrm{int}(Y)) &= \{p_i \pm \varepsilon\} \cap (y_i - \tfrac{1}{2}, y_i + \tfrac{1}{2}) \\ &= \left\{x_i + \tfrac{1}{2} \pm \varepsilon\right\} \cap (x_i + \tfrac{1}{2}, x_i + \tfrac{3}{2}) \\ &= \left\{x_i + \tfrac{1}{2} + \varepsilon\right\} \\ &= \{p_i + \varepsilon\} = \{p_i + \varepsilon \cdot \mathrm{sign}(p_i - x_i)\}\end{aligned}$$

and similarly with the closure

$$\begin{aligned}\pi_i(A_\varepsilon(\vec{p}) \cap \overline{Y}) &= \{p_i \pm \varepsilon\} \cap [y_i - \tfrac{1}{2}, y_i + \tfrac{1}{2}] \\ &= \left\{x_i + \tfrac{1}{2} \pm \varepsilon\right\} \cap [x_i + \tfrac{1}{2}, x_i + \tfrac{3}{2}] \\ &= \left\{x_i + \tfrac{1}{2} + \varepsilon\right\} \\ &= \{p_i + \varepsilon\} \\ &= \{p_i + \varepsilon \cdot \mathrm{sign}(p_i - x_i)\}\end{aligned}$$

2. If $|x_i - y_i| = 1$, and $x_i > y_i$ then it must be that $y_i = x_i - 1$, so the following holds for any $\varepsilon \in (0, \frac{1}{2})$:

$$\begin{aligned}\pi_i(A_\varepsilon(X)) \cap \pi_i(\mathrm{int}(Y)) &= \left\{x_i \pm \tfrac{1}{2} \pm \varepsilon\right\} \cap (y_i - \tfrac{1}{2}, y_i + \tfrac{1}{2}) \\ &= \left\{x_i \pm \tfrac{1}{2} \pm \varepsilon\right\} (x_i - \tfrac{3}{2}, x_i - \tfrac{1}{2}) \\ &= \left\{x_i - \tfrac{1}{2} - \varepsilon\right\}\end{aligned}$$

and similarly with the closure

$$\begin{aligned}\pi_i(A_\varepsilon(X)) \cap \pi_i(\overline{Y}) &= \left\{x_i \pm \tfrac{1}{2} \pm \varepsilon\right\} \cap [y_i - \tfrac{1}{2}, y_i + \tfrac{1}{2}] \\ &= \left\{x_i \pm \tfrac{1}{2} \pm \varepsilon\right\} \cap [x_i - \tfrac{3}{2}, x_i - \tfrac{1}{2}] \\ &= \left\{x_i - \tfrac{1}{2} - \varepsilon\right\}\end{aligned}$$

and

$$\begin{aligned}\pi_i(\text{corners}(X)) \cap \pi_i(\overline{Y}) &= \left\{x_i \pm \tfrac{1}{2}\right\} \cap [y_i - \tfrac{1}{2}, y_i + \tfrac{1}{2}] \\ &= \left\{x_i \pm \tfrac{1}{2}\right\} \cap [x_i - \tfrac{3}{2}, x_i - \tfrac{1}{2}] \\ &= \left\{x_i - \tfrac{1}{2}\right\}.\end{aligned}$$

Since $\vec{p} \in \text{corners}(X) \cap \overline{Y}$, this last line also shows that $p_i = x_i - \frac{1}{2}$ (so $\text{sign}(p_i - x_i) = -1$), thus

$$\begin{aligned}\pi_i(A_\varepsilon(\vec{p}) \cap \text{int}(Y)) &= \{p_i \pm \varepsilon\} \cap (y_i - \tfrac{1}{2}, y_i + \tfrac{1}{2}) \\ &= \left\{x_i - \tfrac{1}{2} \pm \varepsilon\right\} \cap (x_i - \tfrac{3}{2}, x_i - \tfrac{1}{2}) \\ &= \left\{x_i - \tfrac{1}{2} - \varepsilon\right\} \\ &= \{p_i - \varepsilon\} \\ &= \{p_i + \varepsilon \cdot \text{sign}(p_i - x_i)\}\end{aligned}$$

and similarly with the closure

$$\begin{aligned}\pi_i(A_\varepsilon(\vec{p}) \cap \overline{Y}) &= \{p_i \pm \varepsilon\} \cap [y_i - \tfrac{1}{2}, y_i + \tfrac{1}{2}] \\ &= \left\{x_i - \tfrac{1}{2} \pm \varepsilon\right\} \cap [x_i - \tfrac{3}{2}, x_i - \tfrac{1}{2}] \\ &= \left\{x_i - \tfrac{1}{2} - \varepsilon\right\} \\ &= \{p_i - \varepsilon\} \\ &= \{p_i + \varepsilon \cdot \text{sign}(p_i - x_i)\}.\end{aligned}$$

3. If $|x_i - y_i| = 0$, then the following holds for any $\varepsilon \in (0, \frac{1}{2})$:

$$\begin{aligned}\pi_i(A_\varepsilon(X)) \cap \pi_i(\mathrm{int}(Y)) &= \left\{x_i \pm \tfrac{1}{2} \pm \varepsilon\right\} \cap (y_i - \tfrac{1}{2}, y_i + \tfrac{1}{2}) \\ &= \left\{x_i \pm \tfrac{1}{2} \pm \varepsilon\right\} (x_i - \tfrac{1}{2}, x_i + \tfrac{1}{2}) \\ &= \left\{x_i - \tfrac{1}{2} + \varepsilon, x_i + \tfrac{1}{2} - \varepsilon\right\}\end{aligned}$$

and

$$\begin{aligned}\pi_i(A_\varepsilon(X)) \cap \pi_i(\bar{Y}) &= \left\{x_i \pm \tfrac{1}{2} \pm \varepsilon\right\} \cap [y_i - \tfrac{1}{2}, y_i + \tfrac{1}{2}] \\ &= \left\{x_i \pm \tfrac{1}{2} \pm \varepsilon\right\} \cap [x_i - \tfrac{1}{2}, x_i + \tfrac{1}{2}] \\ &= \left\{x_i - \tfrac{1}{2} + \varepsilon, x_i + \tfrac{1}{2} - \varepsilon\right\}\end{aligned}$$

and

$$\begin{aligned}\pi_i(\mathrm{corners}(X)) \cap \pi_i(\bar{Y}) &= \left\{x_i \pm \tfrac{1}{2}\right\} \cap [y_i - \tfrac{1}{2}, y_i + \tfrac{1}{2}] \\ &= \left\{x_i \pm \tfrac{1}{2}\right\} \cap [x_i - \tfrac{1}{2}, x_i + \tfrac{1}{2}] \\ &= \left\{x_i - \tfrac{1}{2}, x_i + \tfrac{1}{2}\right\}.\end{aligned}$$

Regarding $A_\varepsilon(\vec{p})$ we have two subcases (again the following is true for any $\varepsilon \in (0, \frac{1}{2})$):

a) If $p_i = x_i + \frac{1}{2}$ (so $\operatorname{sign}(p_i - x_i) = +1$), then

$$\begin{aligned}
\pi_i(A_\varepsilon(\vec{p}) \cap \operatorname{int}(Y)) &= \{p_i \pm \varepsilon\} \cap (y_i - \tfrac{1}{2}, y_i + \tfrac{1}{2}) \\
&= \left\{x_i + \tfrac{1}{2} \pm \varepsilon\right\} \cap (x_i - \tfrac{1}{2}, x_i + \tfrac{1}{2}) \\
&= \left\{x_i + \tfrac{1}{2} - \varepsilon\right\} \\
&= \{p_i - \varepsilon\} \\
&= \{p_i - \varepsilon \cdot \operatorname{sign}(p_i - x_i)\}
\end{aligned}$$

and similarly with the closure

$$\begin{aligned}
\pi_i(A_\varepsilon(\vec{p}) \cap \overline{Y}) &= \{p_i \pm \varepsilon\} \cap [y_i - \tfrac{1}{2}, y_i + \tfrac{1}{2}] \\
&= \left\{x_i + \tfrac{1}{2} \pm \varepsilon\right\} \cap [x_i - \tfrac{1}{2}, x_i + \tfrac{1}{2}] \\
&= \left\{x_i + \tfrac{1}{2} - \varepsilon\right\} \\
&= \{p_i - \varepsilon\} \\
&= \{p_i - \varepsilon \cdot \operatorname{sign}(p_i - x_i)\}
\end{aligned}$$

b) Otherwise $p_i = x_i - \frac{1}{2}$ (so $\operatorname{sign}(p_i - x_i) = -1$), then

$$\begin{aligned}
\pi_i(A_\varepsilon(\vec{p}) \cap \operatorname{int}(Y)) &= \{p_i \pm \varepsilon\} \cap (y_i - \tfrac{1}{2}, y_i + \tfrac{1}{2}) \\
&= \left\{x_i - \tfrac{1}{2} \pm \varepsilon\right\} \cap (x_i - \tfrac{1}{2}, x_i + \tfrac{1}{2}) \\
&= \left\{x_i - \tfrac{1}{2} + \varepsilon\right\} \\
&= \{p_i + \varepsilon\} \\
&= \{p_i - \varepsilon \cdot \operatorname{sign}(p_i - x_i)\}
\end{aligned}$$

and similarly with the closure

$$\begin{aligned}\pi_i(A_\varepsilon(\vec{p}) \cap \overline{Y}) &= \{p_i \pm \varepsilon\} \cap [y_i - \tfrac{1}{2}, y_i + \tfrac{1}{2}] \\ &= \left\{x_i - \tfrac{1}{2} \pm \varepsilon\right\} \cap [x_i - \tfrac{1}{2}, x_i + \tfrac{1}{2}] \\ &= \left\{x_i - \tfrac{1}{2} + \varepsilon\right\} \\ &= \{p_i + \varepsilon\} \\ &= \{p_i - \varepsilon \cdot \operatorname{sign}(p_i - x_i)\}\end{aligned}$$

4. If $|x_i - y_i| \in (0,1)$ and $x_i < y_i$, then $x_i - \frac{1}{2} < y_i - \frac{1}{2}$, so for sufficiently small $\varepsilon > 0$, $x_i - \frac{1}{2} \mp \varepsilon < y_i - \frac{1}{2}$ which shows that $x_i - \frac{1}{2} \mp \varepsilon \notin [y_i - \frac{1}{2}, y_i + \frac{1}{2}] \supseteq (y_i - \frac{1}{2}, y_i + \frac{1}{2})$. Further, that $|x_i - y_i| \in (0,1)$ and $x_i < y_i$ implies $y_i - 1 < x_i < y_i$, so $y_i - \frac{1}{2} < x_i + \frac{1}{2} < y_i + \frac{1}{2}$, and so for sufficiently small $\varepsilon > 0$, $y_i - \frac{1}{2} < x_i + \frac{1}{2} \mp \varepsilon < y_i + \frac{1}{2}$ and thus $x_i + \frac{1}{2} \mp \varepsilon \in (y_i - \frac{1}{2}, y_i + \frac{1}{2}) \subseteq [y_i - \frac{1}{2}, y_i + \frac{1}{2}]$. Together these establish that

$$\begin{aligned}\pi_i(A_\varepsilon(X)) \cap \pi_i(\operatorname{int}(Y)) &= \left\{x_i \pm \tfrac{1}{2} \pm \varepsilon\right\} \cap (y_i - \tfrac{1}{2}, y_i + \tfrac{1}{2}) \\ &= \left\{x_i + \tfrac{1}{2} \pm \varepsilon\right\}\end{aligned}$$

and

$$\begin{aligned}\pi_i(A_\varepsilon(X)) \cap \pi_i(\overline{Y}) &= \left\{x_i \pm \tfrac{1}{2} \pm \varepsilon\right\} \cap [y_i - \tfrac{1}{2}, y_i + \tfrac{1}{2}] \\ &= \left\{x_i + \tfrac{1}{2} \pm \varepsilon\right\}\end{aligned}$$

and

$$\pi_i(\text{corners}(X)) \cap \pi_i(\overline{Y}) = \left\{x_i \pm \tfrac{1}{2}\right\} \cap [y_i - \tfrac{1}{2}, y_i + \tfrac{1}{2}]$$
$$= \left\{x_i + \tfrac{1}{2}\right\}.$$

Since $\vec{p} \in \text{corners}(X) \cap \overline{Y}$, this last line also shows that $p_i = x_i + \frac{1}{2}$, thus $\{p_i \pm \varepsilon\} = \left\{x_i + \frac{1}{2} \pm \varepsilon\right\}$, and three lines above shows that this is a subset of $\pi_i(\text{int}(Y))$ and two lines above shows it is a subset of $\pi_i(\overline{Y})$ which shows that

$$\pi_i(A_\varepsilon(\vec{p}) \cap \text{int}(Y)) = \{p_i \pm \varepsilon\} \cap (y_i - \tfrac{1}{2}, y_i + \tfrac{1}{2})$$
$$= \{p_i \pm \varepsilon\}$$

and similarly with the closure

$$\pi_i(A_\varepsilon(\vec{p}) \cap \overline{Y}) = \{p_i \pm \varepsilon\} \cap [y_i - \tfrac{1}{2}, y_i + \tfrac{1}{2}]$$
$$= \{p_i \pm \varepsilon\}.$$

5. If $|x_i - y_i| \in (0, 1)$ and $x_i > y_i$, then $y_i + \frac{1}{2} < x_i + \frac{1}{2}$, so for sufficiently small $\varepsilon > 0$, $y_i + \frac{1}{2} < x_i + \frac{1}{2} \mp \varepsilon$ which shows that $x_i + \frac{1}{2} \mp \varepsilon \notin [y_i - \frac{1}{2}, y_i + \frac{1}{2}] \supset (y_i - \frac{1}{2}, y_i + \frac{1}{2})$. Further, that $|x_i - y_i| \in (0, 1)$ and $x_i > y_i$ implies $y_i < x_i < y_i + 1$, so $y_i - \frac{1}{2} < x_i - \frac{1}{2} < y_i + \frac{1}{2}$, and so for sufficiently small $\varepsilon > 0$, $y_i - \frac{1}{2} < x_i - \frac{1}{2} \mp \varepsilon < y_i + \frac{1}{2}$ and thus $x_i - \frac{1}{2} \mp \varepsilon \in (y_i - \frac{1}{2}, y_i + \frac{1}{2}) \subseteq [y_i - \frac{1}{2}, y_i + \frac{1}{2}]$. Together these establish that

$$\pi_i(A_\varepsilon(X)) \cap \pi_i(\text{int}(Y)) = \left\{x_i \pm \tfrac{1}{2} \pm \varepsilon\right\} \cap (y_i - \tfrac{1}{2}, y_i + \tfrac{1}{2})$$
$$= \left\{x_i - \tfrac{1}{2} \pm \varepsilon\right\}$$

and

$$\pi_i(A_\varepsilon(X)) \cap \pi_i(\overline{Y}) = \left\{x_i \pm \tfrac{1}{2} \pm \varepsilon\right\} [y_i - \tfrac{1}{2}, y_i + \tfrac{1}{2}]$$
$$= \left\{x_i - \tfrac{1}{2} \pm \varepsilon\right\}$$

and

$$\pi_i(\text{corners}(X)) \cap \pi_i(\overline{Y}) = \left\{x_i \pm \tfrac{1}{2}\right\} [y_i - \tfrac{1}{2}, y_i + \tfrac{1}{2}]$$
$$= \left\{x_i - \tfrac{1}{2}\right\}.$$

Since $\vec{p} \in \text{corners}(X) \cap \overline{Y}$, this last line also shows that $p_i = x_i - \frac{1}{2}$, thus $\{p_i \pm \varepsilon\} = \left\{x_i - \frac{1}{2} \pm \varepsilon\right\}$, and three lines above shows that this is a subset of $\pi_i(\text{int}(Y))$ and two lines above shows it is a subset of $\pi_i(\overline{Y})$ which shows that

$$\pi_i(A_\varepsilon(\vec{p}) \cap \text{int}(Y)) = \{p_i \pm \varepsilon\} \cap (y_i - \tfrac{1}{2}, y_i + \tfrac{1}{2})$$
$$= \{p_i \pm \varepsilon\}$$

and similarly with the closure

$$\pi_i(A_\varepsilon(\vec{p}) \cap \overline{Y}) = \{p_i \pm \varepsilon\} \cap [y_i - \tfrac{1}{2}, y_i + \tfrac{1}{2}]$$
$$= \{p_i \pm \varepsilon\}.$$

A final comment is that so far we have shown that many individual conditions hold for sufficiently small ε, but since we have only a finite number of dimensions/coordinates and a finite number of cases for each coordinate, we can in fact say that for sufficiently small ε, everything holds. □

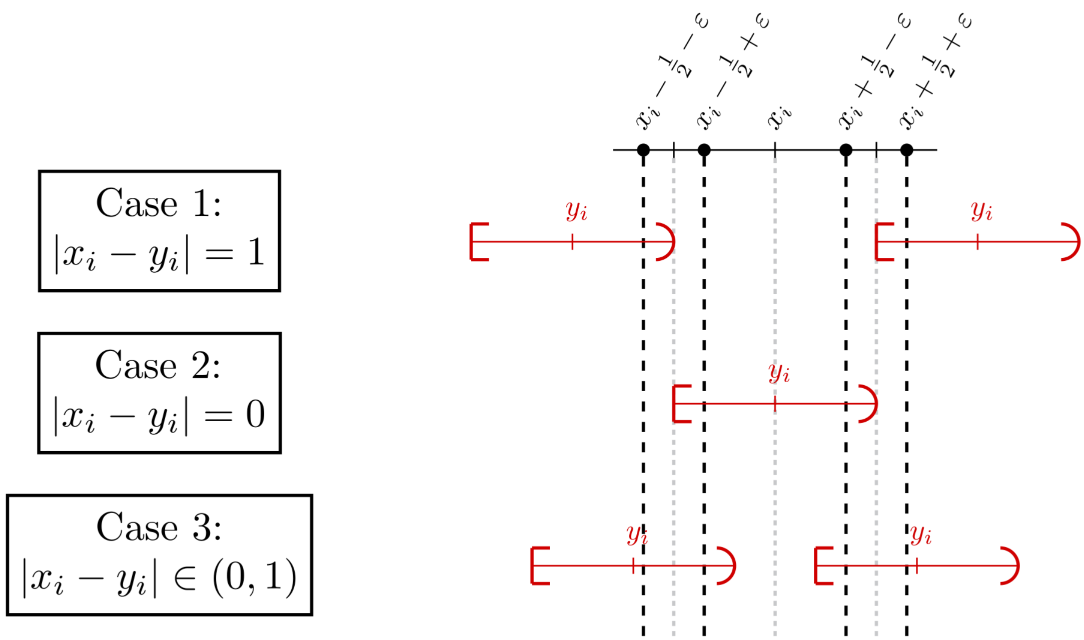


Figure 9.2: The cases dealt with in the proof of Lemma 9.3.1 (ε-Approximation Structure of Adjacency). In case 1, the interval $[y_i - \frac{1}{2}, y_i + \frac{1}{2})$ contains exactly one of the four points $\left\{x_i \pm \frac{1}{2} \pm \varepsilon\right\}$ for all sufficiently small $\varepsilon > 0$. In cases 2 and 3, for all sufficiently small $\varepsilon > 0$ the interval contains exactly 2 of the four points.

9.4 The Notion of Cousins

We now give additional specificity to exactly how unit cubes can be adjacent to each other in light of Lemma 9.3.1 (ε-Approximation Structure of Adjacency). To do so, we define the notion of two cubes being cousins. We use this language to extend the term "twin pair" used commonly in reference to Minkowski's conjecture and Keller's conjecture. A pair of axis-aligned unit cubes X and Y are said to be a Minkowski twin pair if there is exactly one coordinate i_0 where $|\text{center}_{i_0}(X) - \text{center}_{i_0}(Y)| = 1$ and for all other coordinates $i \neq i_0$ it holds that $|\text{center}_i(X) - \text{center}_i(Y)| = 0$. In other words, the positions differ by 1 in some coordinate so that the two cubes are adjacent (as required by Fact 3.4.6 (Adjacency for Unit Cubes)) but in as many coordinates as possible, the positions of X and Y are the same. We generalize by allowing the position differences in some coordinates to be non-zero as long as (1) they remain at most 1 so that the cubes are still adjacent and (2) they are not equal to 1.

Definition 9.4.1 (nth-Cousins). Let $d, n \in \mathbb{N}$ and X, Y be axis-aligned unit cubes in $\mathbb{R}^d$. We say X and Y are *nth-cousins* and denote it as $X \overset{\text{cous}}{\sim}_n Y$ if all of the following hold:

1. $X \overset{\text{adj}}{\sim} Y$
2. There is a unique coordinate $i_0 \in [d]$ such that $|\text{center}_{i_0}(X) - \text{center}_{i_0}(Y)| = 1$
3. There are n indices $i \in [d]$ such that $\text{center}_i(X) \neq \text{center}_i(Y)$

We say that X and Y are *cousins* and denote it as $X \overset{\text{cous}}{\sim} Y$ if there exists some $n \in [d]$ such that X and Y are nth-cousins.

We will refer to the unique coordinate $i_0 \in [d]$ such that $|\text{center}_{i_0}(X) - \text{center}_{i_0}(Y)| = 1$ as the *cousin-coordinate of X and Y*.

If one prefers, by Fact 3.4.6 (Adjacency for Unit Cubes) we could also use the following equivalent definition.

Definition (nth-Cousins). Let $d, n \in \mathbb{N}$ and X, Y be axis-aligned unit cubes in $\mathbb{R}^d$. We say X and Y are *nth-cousins* if all of the following hold:

1. For each $i \in [d]$ it holds that $|\text{center}_i(X) - \text{center}_i(Y)| \in [0, 1]$
2. There is a unique coordinate $i_0 \in [d]$ such that $|\text{center}_{i_0}(X) - \text{center}_{i_0}(Y)| = 1$
3. There are n indices $i \in [d]$ such that $|\text{center}_i(X) - \text{center}_i(Y)| \neq 0$

A few observations are in order. First, the unique coordinate i_0 where $|\text{center}_{i_0}(X) - \text{center}_{i_0}(Y)| = 1$ is counted in the set of indices for which

$\text{center}_i(X) \neq \text{center}_i(Y)$.

Second, X and Y are cousins if and only if the first two conditions of the above definition are satisfied (that is, by ignoring the n parameter, we are ignoring the third condition of the definition). This is because if the first two conditions are satisfied then $X \neq Y$ (by definition of adjacent), so there is at least one index $i \in [d]$ where $\text{center}_i(X) \neq \text{center}_i(Y)$, so the number n of such coordinates has $1 \leq n \leq d$—i.e. X and Y are nth cousins for some $n \in [d]$).

Third, observe that a pair of cubes X and Y are 1st-cousins if and only if there is a unique coordinate i_0 where $|\text{center}_{i_0}(X) - \text{center}_{i_0}(Y)| = 1$ and in all other coordinates $\text{center}_i(X) = \text{center}_i(Y)$—and this is the definition of X and Y being a Minkowski twin pair. Thus X and Y are 1st-cousins if and only if they are a Minkowski twin pair[4].

Fourth, observe that it would be sensible to extend the definition and call X a 0th cousin of itself. This is more of a nuisance than it is worth, but many of the claims we make about nth cousins extend trivially to claims about X taking $n = 0$.

Fifth, there is a notion in geometry in which two (typically closed) sets $X, Y \subseteq \mathbb{R}^d$ are called *neighborly* if X and Y have disjoint interiors and the intersection $\overline{X} \cap \overline{Y}$ is a $(d-1)$-dimensional set (c.f. [Zak85, Zak87] and the references therein). It turns out (see Fact 9.6.1 (Neighborly Iff Cousins)) that for axis-aligned unit cubes,

[4]From a language perspective, it would have been nice if our definition of 0th-cousin meant the same thing as Minkowski twin rather than 1st-cousin meaning the same thing as twin since a twin and first cousin are not the same thing in typical English usage of these words. We could have just defined nth-cousin it this way, but that would lead to lots of +1 terms showing up in properties and proofs.

We blame the fact that family trees don't start counting at 0: it would be reasonable that a person should be considered their own 0th-cousin (because the first common ancestor is 0 generations back), that siblings (and thus twins) should be considered 1st-cousins (because the first common ancestor is the parents which is 1 generation back), and that what are called cousins should really be called 2nd-cousins (because the first common ancestor would be the grandparents which is 2 generations back). If English used this system, then our definition would seem more sensible relative to the definition of a Minkowski twin pair. Oh the consequences one runs into using 1-based indexing instead of 0-based indexing!

this corresponds exactly with our notion of cousins. Nonetheless, we have reason to use our own terminology: sometimes the notion of neighborly is generalized and parameterized as λ-neighborly, but this generalization means something completely different than our parameterized version of nth-cousins, so we use the term "cousin" instead of "neighborly" to avoid any confusion. Further, we have already used the term "neighbor" a lot and feel that overloading it in this context would add confusion.

Next, we will give two additional equivalent definitions which are more substantial than above and make use of Lemma 9.3.1 (ε-Approximation Structure of Adjacency). The first will be an equivalent definition to cousins (i.e. of any type) and the second will be an equivalent definition of nth-cousins when we already know the cubes are cousins of some type.

Lemma 9.4.2 (Equivalent Definition of Cousins)**.** *Let $d \in \mathbb{N}$ and X, Y be axis-aligned unit cubes in $\mathbb{R}^d$ with $\operatorname{int}(X) \cap \operatorname{int}(Y) = \emptyset$. Then the following are equivalent:*

1. *X and Y are cousins*
2. *For all sufficiently small $\varepsilon > 0$, $|A_\varepsilon(X) \cap \operatorname{int}(Y)| = 2^{d-1}$*
3. *For all sufficiently small $\varepsilon > 0$, $|A_\varepsilon(X) \cap Y| = 2^{d-1}$*
4. *For all sufficiently small $\varepsilon > 0$, $\left|A_\varepsilon(X) \cap \overline{Y}\right| = 2^{d-1}$*

Proof. Let $\vec{x} = \operatorname{center}(X)$ and $\vec{y} = \operatorname{center}(Y)$. Recall that for $\varepsilon \in (0, \frac{1}{2})$ we have $A_\varepsilon(X) \overset{\text{def}}{=} \prod_{i=1}^{d} \left\{x_i \pm \frac{1}{2} \pm \varepsilon\right\}$, so

$$Y \cap A_\varepsilon(X) = \prod_{i=1}^{d} [y_i - \tfrac{1}{2}, y_i + \tfrac{1}{2}) \cap \left\{x_i \pm \tfrac{1}{2} \pm \varepsilon\right\}.$$

(1) $\implies$ (3): If X and Y are cousins, then by definition, $X \overset{\text{adj}}{\sim} Y$, and also there is a unique coordinate $i_0 \in [d]$ such that $|x_{i_0} - y_{i_0}| = 1$, and in all other coordinates $i \in$

$[d] \setminus \{i_0\}$ we have $|x_i - y_i| \in [0, 1)$, so by Lemma 9.3.1, for sufficiently small $\varepsilon > 0$, we have $|\pi_{i_0}(A_\varepsilon(X) \cap Y)| = 1$, and for all other $i \in [d] \setminus \{i_0\}$ we have $|\pi_i(A_\varepsilon(X) \cap Y)| = 2$, and so $|A_\varepsilon(X) \cap Y| = 2^{d-1}$.

(3) $\implies$ (1): If for all sufficiently small $\varepsilon > 0$, it holds that $|A_\varepsilon(X) \cap Y| = 2^{d-1}$, then we claim $\overline{X} \cap \overline{Y} \neq \emptyset$. To see this, suppose for contradiction that $\overline{X} \cap \overline{Y} = \emptyset$, then there would exist some coordinate $i_0 \in [d]$ such that $|x_{i_0} - y_{i_0}| > 1$. If such a coordinate existed, then let $\delta = |x_{i_0} - y_{i_0}| - 1$. Then for any $\varepsilon \in (0, \delta)$, we would have $A_\varepsilon(X) \cap Y$ empty because all points in $A_\varepsilon(X)$ are ℓ_∞ distance at most $\frac{1}{2} + \varepsilon < \frac{1}{2} + \delta$ from $\vec{x}$, and points in Y are ℓ_∞ distance at most $\frac{1}{2}$ from $\vec{y}$, but $\vec{x}$ and $\vec{y}$ are ℓ_∞ distance at least $1 + \delta$ apart. But $A_\varepsilon(X) \cap Y$ cannot be empty because by assumption it has cardinality 2^{d-1}. Thus $\overline{X} \cap \overline{Y} \neq \emptyset$ as we claimed. Since also $\operatorname{int}(Y) \cap \operatorname{int}(X) = \emptyset$ by hypothesis, we have $X \overset{\text{adj}}{\sim} Y$.

Then, by Lemma 9.3.1, there can be at most 1 coordinate $i \in [d]$ such that $|x_i - y_i| = 1$ (otherwise $|A_\varepsilon(X) \cap Y| < 2^{d-1}$), and by Fact 3.4.6 (Adjacency for Unit Cubes), there must be at least one such coordinate. Thus, there is exactly one coordinate $i_0 \in [d]$ such that $|x_{i_0} - y_{i_0}| = 1$, so X and Y are cousins by definition.

(2) $\iff$ (3) $\iff$ (4): These three statements are equivalent by Lemma 9.3.1, because for all sufficiently small ε, these three intersections are the same set.

□

Now we present an equivalent definition for nth-cousins assuming that the cubes are cousins of some sort.

Lemma 9.4.3 (Equivalent Definitions of nth-Cousins)**.** *Let $d \in \mathbb{N}$ and X, Y be axis-aligned unit cubes in $\mathbb{R}^d$ such that $X \overset{\text{cous}}{\sim} Y$. Then for any $n \in \mathbb{N}$, the following are equivalent:*

1. *X and Y are nth-cousins ($X \overset{\text{cous}}{\sim}_n Y$)*
2. $\left|\operatorname{corners}(X) \cap \overline{Y}\right| = 2^{d-n}$
3. *There exists $\vec{c} \in \operatorname{corners}(X)$ such that for all sufficiently small $\varepsilon > 0$ it holds that $|A_\varepsilon(\vec{c}) \cap \operatorname{int}(Y)| = 2^{n-1}$*
4. *There exists $\vec{c} \in \operatorname{corners}(X)$ such that for all sufficiently small $\varepsilon > 0$ it holds that $|A_\varepsilon(\vec{c}) \cap Y| = 2^{n-1}$*
5. *There exists $\vec{c} \in \operatorname{corners}(X)$ such that for all sufficiently small $\varepsilon > 0$ it holds that $\left|A_\varepsilon(\vec{c}) \cap \overline{Y}\right| = 2^{n-1}$*

Proof. Let $\vec{x} = \operatorname{center}(X)$ and $\vec{y} = \operatorname{center}(Y)$ throughout.

(1) $\iff$ (2): Since X and Y are cousins by assumption, $X \overset{\text{adj}}{\sim} Y$ (so Lemma 9.3.1 can be applied). Let m be the number of coordinates $i \in [d]$ such that $x_i \neq y_i$ (i.e. $|x_i - y_i| \neq 0$). By Lemma 9.3.1, $\left|\operatorname{corners}(X) \cap \overline{Y}\right| = \prod_{i=1}^{d} \left|\pi_i(\operatorname{corners}(X) \cap \overline{Y})\right| = 2^{d-m}$ which is a product of d-many integers—each of which is a 1 or 2. Further, the only way to express 2^{d-m} as a product of d-many 1's or 2's is to have m-many factors of 1 and $(d-m)$-many factors of 2. So in fact (by Lemma 9.3.1), $\left|\operatorname{corners}(X) \cap \overline{Y}\right| = 2^{d-m}$ if and only if $\vec{x}$ and $\vec{y}$ differ in exactly m-many coordinates. Thus $\left|\operatorname{corners}(X) \cap \overline{Y}\right| = 2^{d-n}$ if and only if $\vec{x}$ and $\vec{y}$ differ in n coordinates, and since X and Y are assumed to be cousins[5], they are nth-cousins if and only if $\vec{x}$ and $\vec{y}$ differ in n coordinates.

[5]The assumption that they are cousins is necessary, because without this assumption, it could be that $\vec{x}$ and $\vec{y}$ differ in n coordinates and so $\left|\operatorname{corners}(X) \cap \overline{Y}\right| = 2^{d-n}$, but they might differ by 1 in multiple coordinates, and so they would not be nth-cousins because they would not even be cousins.

(1) $\implies$ (3), (4), and (5): If $X \overset{\text{cous}}{\sim}_n Y$, then $X \overset{\text{adj}}{\sim} Y$, so by Corollary 3.4.15 (Adjacent Cubes Share a Corner) there exists $\vec{c} \in \text{corners}(X) \cap \overline{Y}$. Then by Lemma 9.3.1 ($\varepsilon$-Approximation Structure of Adjacency), for all sufficiently small $\varepsilon > 0$ we have

$$|A_\varepsilon(\vec{c}) \cap \text{int}(Y)| = |A_\varepsilon(\vec{c}) \cap Y| = \left|A_\varepsilon(\vec{c}) \cap \overline{Y}\right| = \prod_{i=1}^{d} \begin{cases} 1 & |x_i - y_i| = 1 \\ 1 & |x_i - y_i| = 0 \\ 2 & |x_i - y_i| \in (0,1) \end{cases} \tag{9.5}$$

and because $X \overset{\text{cous}}{\sim}_n Y$ there are exactly n coordinates where $x_i \neq y_i$ and exactly one of those where $|x_i - y_i| = 1$, which shows that the cardinality in Equation 9.5 is 2^{n-1}.

(3), (4), or (5) $\implies$ (1): Let $\vec{c}$ be some such corner of X that exists. Because it holds for sufficiently small $\varepsilon > 0$ that $|A_\varepsilon(\vec{c}) \cap \text{int}(Y)| = 2^{n-1} \neq 0$ (resp. Y, resp. $\overline{Y}$) this shows that there exist points of $\text{int}(Y)$ (resp. Y, resp. $\overline{Y}$) arbitrarily close to $\vec{c}$ and so $\vec{c} \in \overline{\text{int}(Y)}$ (resp. $\overline{Y}$, resp. $\overline{Y}$) which implies $\vec{c} \in \overline{Y}$ (see justification[6]). Thus, because $\vec{c} \in \text{corners}(X) \cap \overline{Y}$, by Lemma 9.3.1 ($\varepsilon$-Approximation Structure of Adjacency), Equation 9.5 holds for all sufficiently small $\varepsilon > 0$ and has cardinality 2^{n-1} by hypothesis. This implies that there are exactly $(n-1)$-many coordinates such that $|x_i - y_i| \in (0,1)$, and because $X \overset{\text{cous}}{\sim} Y$ by hypothesis, there is a unique coordinate such that $|x_i - y_i| = 1$ demonstrating that there are exactly n coordinates such that $x_i \neq y_i$ which shows that $X \overset{\text{cous}}{\sim}_n Y$. □

[6]Because up to translation we have $(0,1)^d \subseteq Y \subseteq [0,1]^d$ by Definition 3.4.1 (Axis-Aligned Unit Cube) we have $\overline{\text{int}(Y)} = \overline{Y}$.

9.5 Cliques and Cube Enclosures Revisited

In this section we now use the notion of cousins along with the approximation structure of adjacency studied in the last two sections to study more complex properties of cliques and cube enclosures. The main result of this section is Theorem 9.5.8 (Minimum Cube Enclosure Equivalencies) which gives numerous equivalent conditions for a cube enclosure to be of the minimum possible cardinality.

9.5.1 Cliques Revisited

There will be one type of clique that we are particularly interested in. By Lemma 3.4.11 (Clique-Points Exist for Axis-Aligned Unit Cube Cliques) every axis-aligned unit cube clique has a point where all of the cubes meet together, and we will be interested in the case where one such point belongs to the corner of one of the cubes in the clique. This will show up in many contexts including $(d+1)$-secluded partitions, $(d+1)$-secluded cube enclosures, and also cube enclosures of the minimum possible cardinality. Nonetheless, the main structural property we are interested in (Lemma 9.5.3 that follows) does not require the entire structure of a partition or cube enclosure, so we feel it worthwhile to introduce the following definition.

Definition 9.5.1 (Complete Corner Clique). Let $d \in \mathbb{N}$. A tuple $(\mathcal{C}, X, \vec{c})$ will be called an *axis-aligned unit cube complete cornered clique* in $\mathbb{R}^d$ if all of the following hold:

1. $\mathcal{C}$ is an axis-aligned unit cube clique in $\mathbb{R}^d$
2. $X \in \mathcal{C}$
3. $\vec{c} \in \text{corners}(X)$
4. $\vec{c} \in \bigcap_{Y \in \mathcal{C}} \overline{Y}$ (i.e. $\vec{c}$ is a clique-point of $\mathcal{C}$)
5. There exists $\varepsilon \in (0, \infty)$ such that ${}^{\infty}\overline{B}_{\varepsilon}(\vec{c}) \subseteq \text{int}\left(\bigcup_{Y \in \mathcal{C}} Y\right)$

The reason for the words "corner" and "clique" in the name are probably obvious, and we use the word "complete" to refer to the last property above which ensures that the clique $\mathcal{C}$ is complete in the sense that locally around $\vec{c}$ all points belong to some cube in the clique—there is no empty space. The relevance of this definition is exhibited in the following result.

Lemma 9.5.2 (Complete Corner Cliques in Cube Enclosures)**.** *Let $d \in \mathbb{N}$ and $(X, \mathcal{E})$ an axis-aligned unit cube enclosure in $\mathbb{R}^d$. Then for each $\vec{c} \in$ corners(X) we have that $(X, \mathcal{N}_{\overline{0}}(\vec{c}), \vec{c})$ is an axis-aligned unit cube complete cornered clique.*

Proof. The family $\mathcal{E} \sqcup \{X\}$ is an axis-aligned unit cube packing because all cubes have disjoint interiors by definition of a cube enclosure. By Fact 9.2.5 (Packing Neighborhoods Give Cube Cliques), $\mathcal{N}_{\overline{0}}(\vec{c})$ is an axis-aligned unit cube clique in $\mathbb{R}^d$. Also, $X \in \mathcal{N}_{\overline{0}}(\vec{c})$ and $\vec{c} \in \text{corners}(X)$ by choice of $\vec{c}$, and $\vec{c}$ is a clique-point for $\mathcal{N}_{\overline{0}}(\vec{c})$ because by definition every cube in $\mathcal{N}_{\overline{0}}(\vec{c})$ contains $\vec{c}$ in its closure.

For the final property, by the definition of cube enclosure, because $\vec{c} \in \overline{X}$ we have $\vec{c} \in \text{int}\left(X \cup \bigcup_{Y \in \mathcal{E}} Y\right)$ so for all sufficiently small $\varepsilon > 0$, we have ${}^{\infty}B^{\circ}_{\varepsilon}(\vec{c}) \subseteq X \cup \bigcup_{Y \in \mathcal{E}} Y$. Also, by Corollary 3.6.6 (Unit Cube Packings are Locally Finite) and Fact 3.6.5 (Locally Finite: Enlarged Neighborhood) for all sufficiently small $\varepsilon > 0$ we have

${}^{\|\cdot\|}\mathcal{N}^{\circ}_{2\varepsilon}(\vec{c}) = \mathcal{N}_{\vec{0}}(\vec{c})$ so ${}^{\infty}B^{\circ}_{2\varepsilon}(\vec{c})$ does not intersect any members of $\mathcal{E} \sqcup \{X\}$ that aren't in $\mathcal{N}_{\vec{0}}(\vec{c})$. Thus ${}^{\infty}B^{\circ}_{2\varepsilon}(\vec{c}) \subseteq \bigcup_{Y \in \mathcal{N}_{\vec{0}}(\vec{c})} Y$, and because this ball is open, by definition of interior we have ${}^{\infty}\overline{B}_{\varepsilon}(\vec{c}) \subseteq {}^{\infty}B^{\circ}_{2\varepsilon}(\vec{c}) \subseteq \operatorname{int}\left(\bigcup_{Y \in \mathcal{N}_{\vec{0}}(\vec{c})} Y\right)$. □

If one has the appropriate intuition, the following result is not particularly surprising (though arriving at the intuition does not necessarily come easily). Roughly, it says the following for an axis-aligned unit cube complete cornered clique $(\mathcal{C}, X, \vec{c})$. Consider the point $\vec{c}$ and place each of the d-many axis-orthogonal hyperplanes through $\vec{c}$ to split $\mathbb{R}^d$ into 2^d orthants about $\vec{c}$. Then X is entirely contained in one of these orthants. By same argument as the powers of 2 proof of Theorem 5.1.1 (Optimality of $k = d + 1$ for Cube Partitions) in Subsection 5.1.3 (Proof from First Principles), completely filling up space locally around $\vec{c}$ (i.e. the last condition in the definition of a complete cornered clique) requires at least d cubes other than X showing that $|\mathcal{C}| \geq d + 1$. Along with a hypothesis that $|\mathcal{C}| \leq d + 1$, this gives equality. Locally near $\vec{c}$, X is taking up a $\frac{1}{2^d}$ fraction of the volume. The only way to fill up the rest of the space locally around $\vec{c}$ requires one of the cubes to take up $\frac{1}{2}$ of the space (i.e. it is contained within 2^{d-1} "nicely grouped" orthants), another cube must take up $\frac{1}{2}$ of the remaining space for a total of $\frac{1}{4}$ of the space, and then another must take up $\frac{1}{4}$ of the remaining space for a total of $\frac{1}{8}$, and this continues with $\frac{1}{16}, \frac{1}{32}, \ldots, \frac{1}{2^d}$. There is an inductive structure all the way down where essentially each cube in this process removes one half-plane of the remaining space and thus gets associated to a unique coordinate which defines that half-plane. Furthermore it holds that the first cube in this listing (the one taking up half of the original space) is a dth-cousin of X, and then the next cube is a $(d-1)$th-cousin of X, and so on until the dth and final cube is a 1st-cousin of X. Thus, structurally, $\mathcal{C}$ consists of X and then for each $n \in [d]$ an nth-cousin of X and each cube also has a distinct coordinate $j \in [d]$ associated to it which is the cousin coordinate for itself and X (i.e. the unique coordinate in which its position and the position of X differ by exactly 1).

Lemma 9.5.3 (Structure of Small Complete Cornered Cliques)**.** *Let $d \in \mathbb{N}$ and $(\mathcal{C}, X, \vec{c})$ be an axis-aligned unit cube complete cornered clique in $\mathbb{R}^d$ such that $|\mathcal{C}| \leq d+1$. Then $|\mathcal{C}| = d+1$ and there exists bijections*

$$n : \mathcal{C} \setminus \{X\} \to [d]$$
$$j : \mathcal{C} \setminus \{X\} \to [d]$$

such that for each $Y \in \mathcal{C} \setminus \{X\}$, we have Y is an $n(Y)$th-cousin of X with cousin coordinate $j(Y)$ (i.e. $Y \overset{\text{cous}}{\sim}_{n(Y)} X$ and for $i \in [d]$, $|\text{center}_i(X) - \text{center}_i(Y)| = 1$ if and only if $i = j(Y)$).

Proof. By Corollary 3.6.7 (Unit Cube Packings Closed or Half-Open), we can replace every cube in $\mathcal{C}$ with its half-open version and $(\mathcal{C}, X, \vec{c})$ will still be a an axis-aligned unit cube complete cornered clique because the first four conditions in the definition are trivially still satisfied and this process results in an interior of the union which is a superset of what it was originally, so this fifth condition is still satisfied. Doing so does not affect the cardinality of $\mathcal{C}$ or the cousin structure of any of its members, so we proceed assuming all cubes in $\mathcal{C}$ are half-open.

Throughout the proof, let $\vec{x} = \text{center}(X)$. Now we prove the claim by induction on d.

Claim A. *The result holds for the base case of $d = 1$.*

Proof of Claim. For $d = 1$, note that $X = [a, a+1)$ for some $a \in \mathbb{R}^1$. Thus, it is either the case that $\vec{c} = a$ or $\vec{c} = a+1$. We handle only the former case as the latter is similar. If $\vec{c} = a$, then let ε as in the definition of a complete cornered clique and note that $a - \varepsilon$ must belong to some $Y \in \mathcal{C}$ and this Y must contain $\vec{c}$ in its closure, so $Y = [a-1, a)$. Then $\mathcal{C}$ cannot contain any other members because no subsets of $\mathbb{R}^1$ which are disjoint from X and Y contain $\vec{c}$ in their closure, so we have exactly established what $\mathcal{C}$ must be. Thus, $|C| = 2 = d+1$ and we can trivially define bijections $n, j : \mathcal{C} \setminus X = \{Y\} \to \{1\} = [d]$ (which both necessarily map the only element of the domain to the only element of the codomain). Then it in fact holds that for every member in $\mathcal{C} \setminus X$ (i.e. just for the member Y) that Y is a $n(Y) = 1$st-cousin of X (because $X \overset{\text{adj}}{\sim} Y$, their positions differ by 1 in exactly one coordinate, and their positions are different in 1 coordinate), and they have cousin coordinate $j(Y) = 1$ (because the unique coordinate in which the positions of X and Y differ is coordinate 1), so the result holds in this case. The case $\vec{c} = a + 1$ is similar and omitted. ■

Now we consider the inductive case for $d \in \mathbb{N} \setminus \{1\}$. We proceed with what is essentially the same argument as the powers of 2 proof of Theorem 5.1.1 (Optimality of $k = d+1$ for Cube Partitions) in Subsection 5.1.3 (Proof from First Principles) but we have to be careful to (1) deal with the fact that not every point in $\mathbb{R}^d$ belongs to some cube and (2) pay more attention to the structure of the cubes.

By definition of complete cornered clique, there exists some $\varepsilon > 0$ such that ${}^{\infty}\overline{B}_{\varepsilon}(\vec{c}) \subseteq \operatorname{int}\left(\bigcup_{Y \in \mathcal{C}} Y\right)$, so it trivially follows that this same containment holds also for all sufficiently small $\varepsilon > 0$. For each $\varepsilon \in (0, \infty)$, consider the set of points $A_{\varepsilon}(\vec{c}) = \vec{c} + \{-\varepsilon, \varepsilon\}^d$. Then for all sufficiently small $\varepsilon > 0$ we have

$$A_{\varepsilon}(\vec{c}) \subseteq {}^{\infty}\overline{B}_{\varepsilon}(\vec{c}) \subseteq \operatorname{int}\left(\bigsqcup_{Y \in \mathcal{C}} Y\right) \subseteq \bigsqcup_{Y \in \mathcal{C}} Y$$

(where the disjointness is because we have assumed all cubes are half-open). This implies that

$$A_\varepsilon(\vec{c}) = A_\varepsilon(\vec{c}) \cap \left(\bigsqcup_{Y \in \mathcal{C}} Y \right) = \bigsqcup_{Y \in \mathcal{C}} (A_\varepsilon(\vec{c}) \cap Y)$$

and shows that

$$2^d = |A_\varepsilon(\vec{c})| = \left| \bigsqcup_{Y \in \mathcal{C}} (A_\varepsilon(\vec{c}) \cap Y) \right| = \sum_{Y \in \mathcal{C}} |A_\varepsilon(\vec{c}) \cap Y|$$

and gives a second way to compute the cardinality.

By Lemma 9.3.1 (ε-Approximation Structure of Adjacency) (taking X and Y in the statement to both be X from here) X contains exactly 1 point in $A_\varepsilon(\vec{c})$. Also, for each $Y \in \mathcal{C}$, because $\vec{c} \in \text{corners}(X) \cap \overline{Y}$ (and so $X \overset{\text{adj}}{\sim} Y$) we have by Lemma 9.3.1 that for all sufficiently small $\varepsilon > 0$

$$|A_\varepsilon(\vec{c}) \cap Y| = \prod_{i=1}^{d} \begin{cases} 1 & |x_i - y_i| = 1 \\ 1 & |x_i - y_i| = 0 \\ 2 & |x_i - y_i| \in (0,1) \end{cases} \tag{9.6}$$

which is a power of 2 where the power is between 0 and $d-1$. We can thus define the function $n : \mathcal{C} \setminus \{X\} \to [d]$ by $n(Y) = \log_2(|A_\varepsilon(\vec{c}) \cap Y|) + 1$ (where ε is taken to be some sufficiently small value). This function is well-defined in two regards: (1) the exact choice of ε in the definition is irrelevant as long as it is sufficiently small because the cardinality in the equation above does not depend on the exact value and (2) the codomain is valid.

Claim B. *The function $n : \mathcal{C} \setminus \{X\} \to [d]$ is a bijection.*

Proof of Claim. We have by the above equations and definitions that

$$\begin{aligned} 2^d &= |A_\varepsilon(\vec{c}) \cap X| + \sum_{Y \in \mathcal{C} \setminus \{X\}} |A_\varepsilon(\vec{c}) \cap Y| \\ &= 1 + \sum_{Y \in \mathcal{C} \setminus \{X\}} |A_\varepsilon(\vec{c}) \cap Y| \\ &= 1 + \sum_{Y \in \mathcal{C} \setminus \{X\}} 2^{n(Y)-1} \end{aligned}$$

By hypothesis, $|\mathcal{C}| \leq d+1$, so the index set of the summation above has cardinality at most d. By Lemma 5.1.5 (Summing Powers of 2) we conclude that the index set has cardinality exactly d (so $\mathcal{C}$ has cardinality exactly $d+1$) and the required powers of 2 are such that n is a surjection to $[d]$ (and because n is a surjection between two sets of cardinality d, it is a bijection). ■

For each $i \in [d]$, let $Y^{(i)}$ denote the unique member in $\mathcal{C} \setminus \{X\}$ with $n(Y^{(i)}) = i$ (in bijection notation $Y^{(i)} = n^{-1}(i)$), noting for emphasis that $\left|A_\varepsilon(\vec{c}) \cap Y^{(i)}\right| = 2^{i-1}$, and also let $\vec{y}^{(i)} = \text{center}(Y^{(i)})$. Part of the claim of the result is that each $Y^{(i)}$ is an ith-cousin of X.

If we knew *a priori* that each $Y^{(i)}$ was a cousin of X, then by Lemma 9.4.3 (Equivalent Definitions of nth-Cousins) we could use the fact that $\left|A_\varepsilon(\vec{c}) \cap Y^{(i)}\right| = 2^{i-1}$ holds for all sufficiently small $\varepsilon > 0$ to immediately conclude that each $Y^{(i)}$ is an ith-cousin of X. However, we do not know at this point that $Y^{(i)}$ is a cousin of X, so we cannot use this approach.

Instead we show only that $Y^{(d)}$ is a dth-cousin of X and then use induction.

Claim C. *$Y^{(d)}$ is a dth-cousin of X.*

Proof of Claim. Because $\left|A_\varepsilon(\vec{c}) \cap Y^{(d)}\right| = 2^{d-1}$, then by Equation 9.6 there is exactly one coordinate $i \in [d]$ such that $\left|x_i - y_i^{(d)}\right| \notin (0,1)$, and because $X \overset{\text{adj}}{\sim} Y^{(d)}$, by Fact 3.4.6 (Adjacency for Unit Cubes) there is at least one $i \in [d]$ such that $\left|x_i - y_i^{(d)}\right| = 1$. Consequently, there are no coordinates $i \in [d]$ such that $\left|x_i - y_i^{(d)}\right| = 0$. Thus, X and $Y^{(d)}$ are adjacent, have positions differing by 1 in a unique coordinate, and have positions which differ in d coordinates, so by definition X and $Y^{(d)}$ are dth-cousins. ■

Having established that $Y^{(d)} \overset{\text{cous}}{\sim}_d X$, let $i_d \in [d]$ denote the cousin coordinate of the two (i.e. the unique coordinate where the positions differ by exactly 1). We note one more fact about this coordinate.

Claim D. $\left|y_{i_d}^{(d)} - c_{i_d}\right| = \frac{1}{2}$.

Proof of Claim. Because $\vec{c} \in \overline{X} = {}^{\infty}\overline{B}_{1/2}(\vec{x})$ and $\vec{c} \in \overline{Y^{(d)}} = {}^{\infty}\overline{B}_{1/2}(\vec{y}^{(d)})$ we must have $|x_{i_d} - c_{i_d}| \leq \frac{1}{2}$ and also $\left|y_{i_d}^{(d)} - c_{i_d}\right| \leq \frac{1}{2}$. By definition of i_d we also have $\left|x_{i_d} - y_{i_d}^{(d)}\right| = 1$ and thus it must be that c_{i_d} is the midpoint of x_{i_d} and $y_{i_d}^{(d)}$ (and thus $\left|y_{i_d}^{(d)} - c_{i_d}\right| = \frac{1}{2}$). ■

We now argue that $Y^{(d)}$ is the only cube in $\mathcal{C} \setminus \{X\}$ which "utilizes" the coordinate i_d—in this coordinate, every other cube has position identical to that of X.

Claim E. *For $Y \in \mathcal{C} \setminus \left\{Y^{(d)}\right\}$ (letting $\vec{y} = \text{center}(Y)$), it holds that $y_{i_d} = x_{i_d}$.*

Proof of Claim. This is trivial for $Y = X$. Because X and $Y^{(d)}$ are dth-cousins we have that $\left|x_{i_d} - y_{i_d}^{(d)}\right| = 1$ and for all other $i \in [d] \setminus \{i_d\}$ we have $\left|x_i - y_i^{(d)}\right| \in (0, 1)$. Together with Lemma 9.3.1 (ε-Approximation Structure of Adjacency), the following holds for all sufficiently small $\varepsilon > 0$ (as a reminder, the expression $\text{sign}(c_i - x_i)$ just indicates for each coordinate how the corner $\vec{c}$ is positioned relative to the center of

X and is used to determine the appropriate sign of ε based on this orientation):

$$A_\varepsilon(\vec{c}) \cap Y^{(d)} = \prod_{i=1}^{d} \begin{cases} \{c_i + \varepsilon \cdot \operatorname{sign}(c_i - x_i)\} & |x_i - y_i^{(d)}| = 1 \\ \{c_i - \varepsilon \cdot \operatorname{sign}(c_i - x_i)\} & |x_i - y_i^{(d)}| = 0 \\ \{c_i - \varepsilon, \quad c_i + \varepsilon\} & |x_i - y_i^{(d)}| \in (0,1) \end{cases}$$

$$= \prod_{i=1}^{d} \begin{cases} \{c_i + \varepsilon \cdot \operatorname{sign}(c_i - x_i)\} & i = i_d \\ \{c_i - \varepsilon, \quad c_i + \varepsilon\} & i \neq i_d \end{cases}$$

which shows that $Y^{(d)}$ contains every point of $A_\varepsilon(\vec{c})$ which has i_dth coordinate equal to $c_{i_d} + \varepsilon \cdot \operatorname{sign}(c_{i_d} - x_{i_d})$.

Also observe that for arbitrary $Y \in \mathcal{C}$ (letting $\vec{y} = \operatorname{center}(Y)$), by Lemma 9.3.1 ($\varepsilon$-Approximation Structure of Adjacency) the following set containment holds for all sufficiently small $\varepsilon > 0$:

$$A_\varepsilon(\vec{c}) \cap Y = \prod_{i=1}^{d} \begin{cases} \{c_i + \varepsilon \cdot \operatorname{sign}(c_i - x_i)\} & |x_i - y_i| = 1 \\ \{c_i - \varepsilon \cdot \operatorname{sign}(c_i - x_i)\} & |x_i - y_i| = 0 \\ \{c_i - \varepsilon, \quad c_i + \varepsilon\} & |x_i - y_i| \in (0,1) \end{cases}$$

$$\supseteq \prod_{i=1}^{d} \begin{cases} \{c_i + \varepsilon \cdot \operatorname{sign}(c_i - x_i)\} & |x_i - y_i| = 1 \\ \{c_i - \varepsilon \cdot \operatorname{sign}(c_i - x_i)\} & |x_i - y_i| = 0 \\ \{c_i + \varepsilon \cdot \operatorname{sign}(c_i - x_i)\} & |x_i - y_i| \in (0,1) \end{cases}$$

$$= \prod_{i=1}^{d} \begin{cases} \{c_i - \varepsilon \cdot \operatorname{sign}(c_i - x_i)\} & |x_i - y_i| = 0 \\ \{c_i + \varepsilon \cdot \operatorname{sign}(c_i - x_i)\} & |x_i - y_i| \in (0,1] \end{cases}$$

If $Y \in \mathcal{C}$ has the property that $y_{i_d} \neq x_{i_d}$, then (by Lemma 9.3.1 (ε-Approximation Structure of Adjacency)) $|y_{i_d} - x_{i_d}| \in (0,1]$, so by the equation above this implies that Y contains some point of $A_\varepsilon(\vec{c})$ which has i_dth coordinate equal to $c_{i_d} + \varepsilon \cdot \operatorname{sign}(c_{i_d} - x_{i_d})$. But we already showed that $Y^{(d)}$ contains all such points and distinct members in $\mathcal{C}$ are disjoint, so $Y^{(d)}$ is the only member with this property which shows that for all $Y \in \mathcal{C} \setminus \left\{Y^{(d)}\right\}$ it holds that $y_{i_d} = x_{i_d}$. ■

Because the coordinate i_d is only "meaningfully utilized" by $Y^{(d)}$, we consider dropping $Y^{(d)}$ from the clique $\mathcal{C}$ and projecting out the coordinate i_d so that we can apply the inductive hypothesis in $\mathbb{R}^{d-1}$.

Let $\hat{X} = \tau_{i_d}(X)$, and for $i \in [d] \setminus \{i_d\}$, let $\hat{Y}^{(i)} = \tau_{i_d}(Y^{(i)})$. Let $\hat{c} = \tau_{i_d}(\vec{c})$. Then let $\hat{\mathcal{C}} = \left\{\tau_{i_d}(Z) : Z \in \mathcal{C} \setminus \left\{Y^{(d)}\right\}\right\} = \left\{\hat{X}\right\} \sqcup \left\{\hat{Y}^{(i)} : i \in [d] \setminus \{i_d\}\right\}$. We interpret these as points, subsets, and families of subsets in $\mathbb{R}^d$, but we index $\mathbb{R}^{d-1}$ not with $[d-1]$ but with $[d] \setminus \{i_d\}$.

Claim F. *$(\hat{\mathcal{C}}, \hat{X}, \hat{c})$ is an axis-aligned unit cube complete corner clique in $\mathbb{R}^{d-1}$ (where we index $\mathbb{R}^{d-1}$ not with $[d-1]$ but with $[d] \setminus \{i_d\}$).*

Proof of Claim. We prove all of the conditions of Definition 9.5.1.

1. To see that $\hat{C}$ is an axis-aligned unit cube clique, first note that each member of $\hat{C}$ is an axis-aligned unit cube because projecting out one coordinate of an axis-aligned unit cube results in a lower-dimensional axis-aligned unit cube. To see that $\hat{C}$ is a clique, consider arbitrary distinct members $\hat{Z}, \hat{Z}' \in \hat{C}$ (and let $Z, Z' \in \mathcal{C} \setminus \left\{Y^{(d)}\right\}$ be the associated distinct members such that $\hat{Z} = \tau_{i_d}(Z)$ and $\hat{Z}' = \tau_{i_d}(Z')$). Because $Z \overset{\text{adj}}{\sim} Z'$, by Fact 3.4.6 (Adjacency for Unit Cubes) there exists some coordinate $i \in [d]$ such that $|\text{center}_i(Z) - \text{center}_i(Z')| = 1$ and because neither Z nor Z' is $Y^{(d)}$, we have that $i \neq i_d$ because by Claim E, $\text{center}_{i_d}(Z) = \text{center}_{i_d}(X) = \text{center}_{i_d}(Z')$. Thus, $i \in [d] \setminus \{i_d\}$, so $\left|\text{center}_i(\hat{Z}) - \text{center}_i(\hat{Z}')\right| = 1$ so by Fact 3.4.6 $\hat{Z} \overset{\text{adj}}{\sim} \hat{Z}'$.
2. Trivially, $\hat{X} \in \hat{\mathcal{C}}$ by definition of $\hat{C}$.
3. It follows easily from the definition of τ_{i_d} that $\hat{c} \in \text{corners}(\hat{X})$.
4. Similarly, because $\vec{c} \in \overline{Z}$ for all $Z \in \mathcal{C}$, it follows that $\hat{c} \in \tau_{i_d}(Z)$, so $\hat{c} \in \bigcap_{\hat{Z} \in \hat{\mathcal{C}}} \overline{\hat{Z}}$.
5. Note that it will suffice to show that there exists $\varepsilon \in (0, \infty)$ such that ${}^{\infty}\overline{B}_{\varepsilon}(\hat{c}) \subseteq \bigcup_{\hat{Z} \in \hat{\mathcal{C}}} \hat{Z}$ because we can then take $\varepsilon' < \varepsilon$ to have ${}^{\infty}\overline{B}_{\varepsilon'}(\hat{c}) \subseteq {}^{\infty}B^{\circ}_{\varepsilon}(\hat{c}) \subseteq {}^{\infty}\overline{B}_{\varepsilon}(\hat{c}) \subseteq \bigcup_{\hat{Z} \in \hat{\mathcal{C}}} \hat{Z}$, and because ${}^{\infty}B^{\circ}_{\varepsilon}(\hat{c})$ is an open subset of the union, it is a subset of the interior of the union.

   Because $(\mathcal{C}, X, \vec{c})$ is an axis-aligned complete cornered clique in $\mathbb{R}^d$, there exists $\varepsilon \in (0, \infty)$ such that ${}^{\infty}\overline{B}_{\varepsilon}(\vec{c}) \subseteq \bigcup_{Z \in \mathcal{C}} Z$. Fix such an ε, and we will show that

${}^{\infty}\overline{B}_{\varepsilon}(\hat{c}) \subseteq \bigcup_{\hat{Z}\in\hat{\mathcal{C}}} \hat{Z}$. Let

$$B \overset{\text{def}}{=} \prod_{i\in[d]} \begin{cases} [c_i-\varepsilon, c_i+\varepsilon] & i \in [d]\setminus\{i_d\} \\ \left\{c_i+\varepsilon\cdot\operatorname{sign}(c_i-y_i^{(d)})\right\} & i = i_d \end{cases}$$

and observe that

$$\begin{aligned} {}^{\infty}\overline{B}_{\varepsilon}(\hat{c}) &= \prod_{i\in[d]\setminus\{i_d\}} [c_i-\varepsilon, c_i+\varepsilon] \\ &= \tau_{i_d}\left(\prod_{i\in[d]} \begin{cases} [c_i-\varepsilon, c_i+\varepsilon] & i \in [d]\setminus\{i_d\} \\ \left\{c_i+\varepsilon\cdot\operatorname{sign}(c_i-y_i^{(d)})\right\} & i = i_d \end{cases}\right) \\ &= \tau_{i_d}(B) \end{aligned}$$

Now note two things: (1) B is trivially a subset of $\prod_{i\in[d]}[c_i-\varepsilon, c_i+\varepsilon] = {}^{\infty}\overline{B}_{\varepsilon}(\vec{c})$ which by hypothesis is a subset of $\bigcup_{Z\in\mathcal{C}} Z$ and (2) B does not intersect $Y^{(d)}$ (see justification[7]). Thus, together these show that $B \subseteq \bigcup_{Z\in\mathcal{C}\setminus\left\{Z^{(d)}\right\}} Z$. It follows

[7]By Fact 3.4.6 (Adjacency for Unit Cubes) it suffices to show that the i_dth coordinate of every point in this product is distance more than $\frac{1}{2}$ from the center of $Y^{(d)}$. By Claim D we have $\left|y_{i_d}^{(d)} - c_{i_d}\right| = \frac{1}{2}$. Thus, the difference between the i_dth coordinate of points in the product and the center of Y^d is $\left|(c_i+\varepsilon\cdot\operatorname{sign}(c_i-y_i^{(d)})) - y_{i_d}^{(d)}\right| = \left|(c_i-y_{i_d}^{(d)})+\varepsilon\cdot\operatorname{sign}(c_i-y_{i_d}^{(d)})\right| = \left|c_i-y_{i_d}^{(d)}\right| + \varepsilon = \frac{1}{2}+\varepsilon > \frac{1}{2}$.

that

$$\begin{aligned}
{}^{\infty}\overline{B}_{\varepsilon}(\hat{c}) &= \tau_{i_d}(B) && \text{(Above)} \\
&= \{\tau_{i_d}(\vec{z}) : \vec{z} \in B\} && \text{(B is a product set)} \\
&\subseteq \left\{ \tau_{i_d}(\vec{z}) : \vec{z} \in \bigcup_{Z \in \mathcal{C} \setminus \left\{ Z^{(d)} \right\}} Z \right\} && \text{(Subset above)} \\
&= \bigcup_{Z \in \mathcal{C} \setminus \left\{ Z^{(d)} \right\}} \{\tau_{i_d}(\vec{z}) : \vec{z} \in Z\} && \text{(General set equality)} \\
&= \bigcup_{Z \in \mathcal{C} \setminus \left\{ Z^{(d)} \right\}} \tau_{i_d}(Z) && \text{(Each Z is a product set)} \\
&= \bigcup_{\hat{Z} \in \hat{\mathcal{C}}} \hat{Z}. && \text{(Notation of $\hat{Z}$ and def'n of $\hat{Z}$)}
\end{aligned}$$

This proves the fifth propery of Definition 9.5.1 and completes the proof of the claim.

■

By Claim F and the inductive hypothesis, there exist bijections

$$\hat{n} : \hat{\mathcal{C}} \setminus \left\{ \hat{X} \right\} \to [d] \setminus \{d\}$$

$$\hat{j} : \hat{\mathcal{C}} \setminus \left\{ \hat{X} \right\} \to [d] \setminus \{i_d\}$$

such that for each $\hat{Z} \in \hat{\mathcal{C}} \setminus \left\{ \hat{X} \right\}$, we have $\hat{Z}$ is an $\hat{n}(\hat{Z})$th-cousin of $\hat{X}$ with cousin coordinate $\hat{j}(\hat{Z})$.

Now we define

$$n : \mathcal{C} \setminus \{X\} \to [d]$$
$$j : \mathcal{C} \setminus \{X\} \to [d]$$

by

$$n(Z) = \begin{cases} d & Z = Y^{(d)} \\ \hat{n}(\hat{Z}) & Z \neq Y^{(d)} \end{cases}$$

and similarly

$$j(Z) = \begin{cases} i_d & Z = Y^{(d)} \\ \hat{j}(\hat{Z}) & Z \neq Y^{(d)} \end{cases}$$

which are both bijections because $\hat{n}$ and $\hat{j}$ are bijections.

Now we want to show that for each $Z \in \mathcal{C} \setminus \{X\}$, we have Z is an $n(Z)$th-cousin of X with cousin coordinate $j(Z)$. We first note that this holds if $Z = Y^{(d)}$ because $n(Y^{(d)}) = d$ and Claim C shows that $Y^{(d)}$ is a dth-cousin of X and i_d (which equals $j(Y^{(d)})$) was by definition the cousin coordinate of X and $Y^{(d)}$.

For any other $Z \in \mathcal{C} \setminus \{X\}$ which is not $Z^{(d)}$, we have by inductive hypothesis that $\hat{Z}$ is an $\hat{n}(\hat{Z})$th-cousin of $\hat{X}$ with cousin coordinate $\hat{j}(\hat{Z})$ which means that center$(\hat{Z})$ and center$(\hat{X})$ (as vectors in $\mathbb{R}^{d-1}$ indexed by $[d] \setminus \{i_d\}$) differ by 1 only in coordinate $\hat{j}(\hat{Z})$ and are different in $\hat{n}(\hat{Z})$ total coordinates. Because center$(\hat{Z}) = \tau_{i_d}(\text{center}(Z))$ and similarly center$(\hat{X}) = \tau_{i_d}(\text{center}(X))$, and because by Claim E center(Z) and center(X) are identical in coordinate i_d, we conclude that center(Z) and center(X) differ from each other in exactly $\hat{n}(\hat{Z}) = n(Z)$-many coordinates and differ by 1 only in coordinate $\hat{j}(\hat{Z}) = j(Z)$. Thus, X and Z are $n(Z)$th-cousins with cousin coordinate $j(Z)$ which completes the proof. □

9.5.2 Cube Enclosures Revisited

One of our primary interests will be cube enclosures of the minimum possible cardinality. It is not hard to see that in an every axis-aligned unit cube enclosure $(X, \mathcal{E})$ in $\mathbb{R}^d$, it must be that $|\mathcal{E}| \geq 2^d - 1$. This because by Corollary 9.2.11 (Cube Enclosures Closed or Half-Open) we may assume that X and each cube in $\mathcal{E}$ is half open so that all cubes are disjoint. Then, because $\text{corners}(X) \subseteq \overline{X} \subseteq \bigsqcup_{Y \in \mathcal{E} \sqcup \{X\}} Y$ (the latter containment is by the definition of a cube enclosure), each of the 2^d corners of X belongs to a distinct[8] cube in $\mathcal{E} \sqcup \{X\}$, and since X (being half-open) contains only one of the corners, there must be $2^d - 1$ other corners of X covered each by a unique cube in $\mathcal{E}$. Thus $|\mathcal{E}| \geq 2^d - 1$.

However, we now use a more careful analysis to show that the $2^d - 1$ lower bound can be doubled to $2^{d+1} - 2$ (and we will later see that this is tight[9]). The idea of the proof is demonstrated by Figure 9.3 with additional intuition given in Figure 9.4. We will pick a sufficiently small $\varepsilon \in (0, \frac{1}{2})$ and employ the set $A_\varepsilon(X)$ and examine how many of the points of these approximate corners are contained in each member of $\mathcal{E}$. After giving the proof, we provide an alternate proof which demonstrates very literally how to double the bound of $2^d - 1$. We believe the latter proof is more slick, but the former gives us not only the tight bound of $2^{d+1} - 2$ but also a necessary and sufficient condition for when it is attained—a condition that will be quite useful to

[8]Each corner belongs to a distinct cube because the corners of X are ℓ_∞ distance 1 apart and no translate of the half open cube $[0, 1)^d$ can contain two points which are ℓ_∞ distance 1 apart.

[9]We know that a $(d+1)$-secluded axis-aligned unit cube partition $\mathcal{P}$ of $\mathbb{R}^d$ exists by Theorem 4.2.18 (Existence of $(d+1, \frac{1}{2d})$-Secluded Unit Cube Partitions), and by Fact 9.2.6 (Partition Neighbors Give Cube Enclosures) this proves that $(d+1)$-secluded axis-aligned unit cube enclosures $(X, \mathcal{E}) = (X, \mathcal{N}(()X))$ exists, and by Theorem 9.5.9 (Cube Enclosures: $(d+1)$-Secluded Implies Minimum), $|\mathcal{E}| = 2^{d+1} - 2$.

us.

First, we give a simple fact that will show up a few times.

Fact 9.5.4 (Powers of 2 Sum)**.** *Let $d \in \mathbb{N}$. Then $\sum_{n=1}^{d} 2^n = 2^{d+1} - 2$.*

Proof. This is true for $d = 1$. By induction, $\sum_{n=1}^{d} 2^n = 2^d + \sum_{n=1}^{d-1} 2^n = 2^d + \left(2^{(d-1)+1} - 2\right) = 2^{d+1} - 2$. □

Theorem 9.5.5 (Minimum Size Cube Enclosure: $2^{d+1} - 2$ Cousins)**.** *Let $d \in \mathbb{N}$ and $(X, \mathcal{E})$ an axis-aligned unit cube enclosure in $\mathbb{R}^d$. Then $|\mathcal{E}| \geq 2^{d+1} - 2$. Furthermore, equality is achieved if and only if every $Y \in \mathcal{E}$ is a cousin of X.*

Proof. By Remark 9.2.12 (Closed or Half-Open Assumption) we may assume that all cubes are half-open and thus disjoint. Let $\vec{x} = \text{center}(X)$. By Lemma 9.2.10, for all sufficiently small $\varepsilon > 0$, ${}^{\infty}\overline{B}_{\frac{1}{2}+\varepsilon}(\vec{x}) \subseteq \text{int}\left(X \sqcup \bigsqcup_{Y \in \mathcal{E}} Y\right)$, and so for sufficiently small $\varepsilon > 0$ we have

$$A_\varepsilon(X) \quad = \quad \prod_{i=1}^{d} \left\{x_i \pm \tfrac{1}{2} \pm \varepsilon\right\} \quad \subseteq \quad {}^{\infty}\overline{B}_{\frac{1}{2}+\varepsilon}(\vec{x}) \quad \subseteq \quad X \sqcup \bigsqcup_{Y \in \mathcal{E}} Y.$$

Thus, each point of $A_\varepsilon(X)$ belongs to X or some member of $\mathcal{E}$, and because all cubes are half-open and disjoint, each point belongs to exactly one member. Thus, we can count the cardinality of $A_\varepsilon(X)$ by summing the counts of its intersection with X and with each member of $\mathcal{E}$.

Observe that $|A_\varepsilon(X)| = 4^d$, and by Lemma 9.3.1,

$$|X \cap A_\varepsilon(X)| = \left| \prod_{i=1}^{d} \left\{x_i - \tfrac{1}{2} + \varepsilon, \;\; x_i + \tfrac{1}{2} - \varepsilon\right\} \right| = 2^d.$$

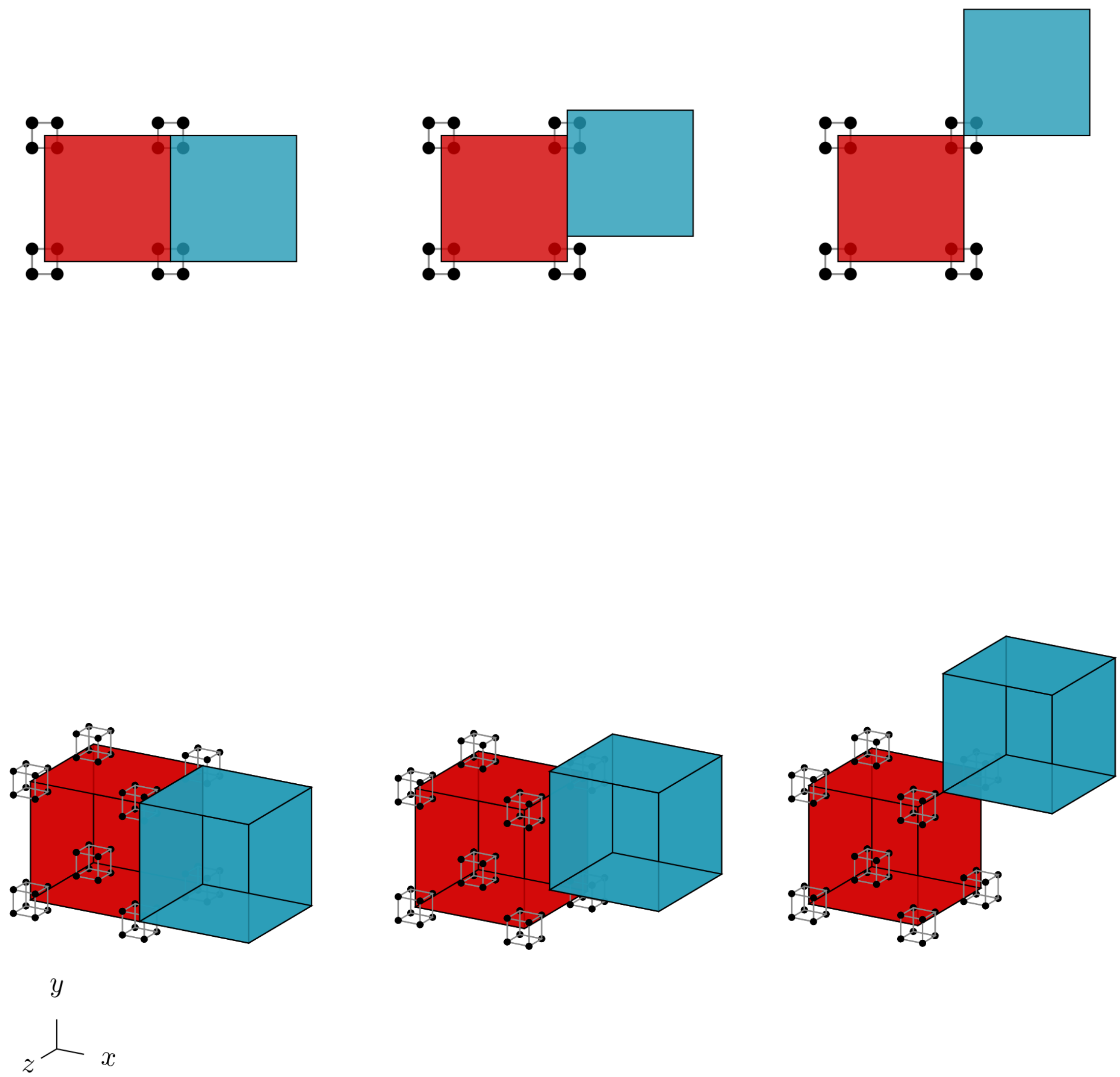


Figure 9.3: A cube X in $\mathbb{R}^2$ and $\mathbb{R}^3$ respectively along with the 4^d points in the set $A_\varepsilon(X)$ which approximate the positions of the corners of the cube. Exactly 2^d of these points belong to X, so there are $4^d - 2^d$ that are contained in cubes adjacent to X. Each cube Y that is adjacent to X must have a center that differs from X by 1 in at least one coordinate, and this means that along this coordinate Y contains only one layer of these points. In every other coordinate, Y contains at most 2 layers of these points because of how they are spaced (when ε is sufficiently small). In total Y contains at most 2^{d-1} of the points. Thus there have to be at least $\frac{4^d-2^d}{2^{d-1}} = 2^{d+1} - 2$ cubes adjacent to X to cover all $4^d - 2^d$ of these points not in X.

Now consider an arbitrary $Y \in \mathcal{E}$, and let $\vec{y} = \text{center}(Y)$. Because $Y \overset{\text{adj}}{\sim} X$, by Fact 3.4.6 (Adjacency for Unit Cubes) there is at least one coordinate $i_0 \in [d]$ such that $|y_i - x_i| = 1$. Consider the set $Y \cap A_\varepsilon(X)$. By Lemma 9.3.1, for coordinate i_0 it holds that $|\pi_{i_0}(Y \cap A_\varepsilon(X))| = 1$ and for every other $i \in [d] \setminus \{i_0\}$ it holds that $|\pi_i(Y \cap A_\varepsilon(X))| \leq 2$; thus $|Y \cap A_\varepsilon(X)| = \prod_{i=1}^{d} |\pi_i(Y \cap A_\varepsilon(X))| \leq 2^{d-1}$.

Then we have the following:

$$\begin{aligned}
4^d &= |A_\varepsilon(X)| \\
&= \sum_{Y \in \mathcal{E} \sqcup \{X\}} |Y \cap A_\varepsilon(X)| \\
&= |X \cap A_\varepsilon(X)| + \sum_{Y \in \mathcal{E}} |Y \cap A_\varepsilon(X)| \\
&= 2^d + \sum_{Y \in \mathcal{E}} |Y \cap A_\varepsilon(X)| \\
&\leq 2^d + |\mathcal{E}| \cdot 2^{d-1}.
\end{aligned}$$

This shows

$$|\mathcal{E}| \geq \frac{4^d - 2^d}{2^{d-1}} = 2^{d+1} - 2.$$

Furthermore, because we could run the above argument for any sufficiently small $\varepsilon > 0$, the equality is achieved if and only if it is the case that that for all sufficiently small $\varepsilon > 0$ it holds for all $Y \in \mathcal{E}$ that $|Y \cap A_\varepsilon| = 2^{d-1}$. By Lemma 9.4.2 this occurs if and only if every $Y \in \mathcal{E}$ is a cousin of X. □

The proof above provides insight into the structure that we are interested in by not only proving what the bound is, but also exactly what structural conditions must hold for equality with that bound to be attained. We now provide a second proof of the lower bound in Theorem 9.5.5 (Minimum Size Cube Enclosure: $2^{d+1} - 2$ Cousins),

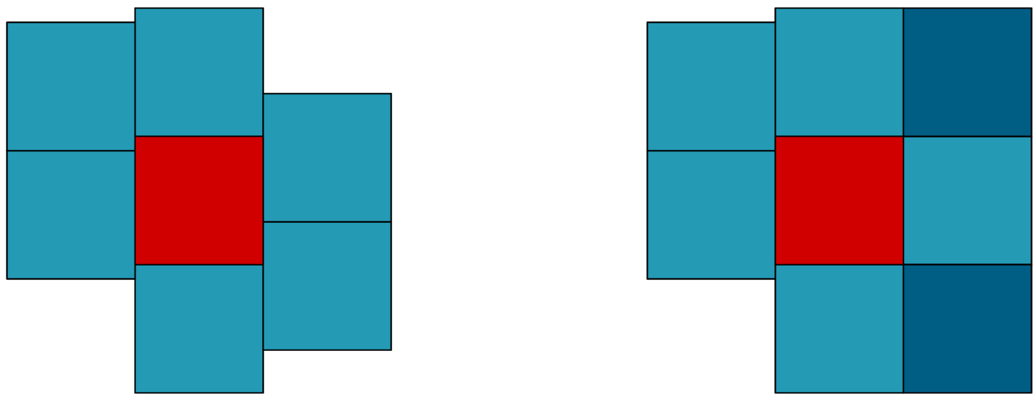


Figure 9.4: On the left is a minimum axis-aligned unit cube enclosure in $\mathbb{R}^2$ with 6 enclosing cubes. On the right is a non-minimum cube enclosure in $\mathbb{R}^2$ with 7 enclosing cubes. Observe that on the left, every cube is a cousin of the enclosed cube (there is a unique coordinate in which the center position differs by 1 from the center position of the enclosed cube) while on the right, the cubes in the top right and bottom right corner are not cousins of the enclosed cube because their positions differ by 1 from the enclosed cube position in both the x- and y-coordinates.

though it will not show when equality is attained. Nonetheless, it is arguably a slicker proof than the one above[10]. We mentioned earlier that it was possible to double the simple lower bound that for an axis-aligned unit cube enclosure $(X, \mathcal{E})$ it must be that $|E| \geq 2^d - 1$ to get to the (tight) lower bound of $2^{d+1} - 2$, and in the proof below, we quite literally double this bound. This second proof may offer additional insight into why the tight bound has the value that it does. The proof also highlights the benefits of studying tilings from the perspective of partitions—unique membership of points makes things much nicer to work with. Figure 9.5 demonstrates the idea of the proof: basically we apply the $2^d - 1$ bound twice using two "opposite" orientations

[10]For one reason, we believe it is a more clever proof, and for a second reason, it does not require the pages of ugly, tedious, brute force detail of Lemma 9.3.1.

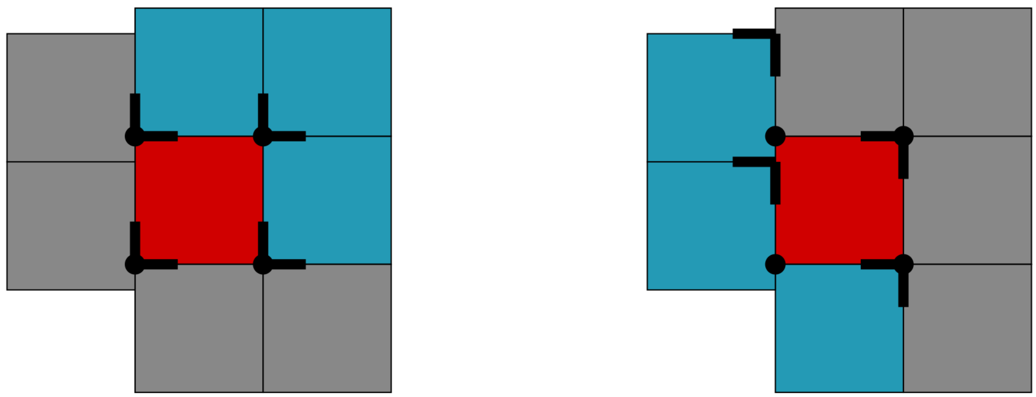


Figure 9.5: On the left is the cube enclosure $(X, \mathcal{E})$ which consists of translates of $[0,1)^d$. Each corner of the enclosed cube is marked, and the unique cube in the enclosure which contains the corner is also marked. On the right is a cube enclosure $(X', \mathcal{E}')$ which uses the same positions, but uses translates of $(0,1]^d$ instead, and again each corner of the enclosed cube is marked. In both cases, the unique cube in the enclosure which contains the corner is indicated. Note that the only cube which is marked in both the left and the right is the enclosed cube. Thus, there are $2^d - 1$ cubes indicated on the left, and a set of $2^d - 1$ cubes indicated on the right, and these sets are disjoint. Thus, there are at least $2(2^d - 1) = 2^{d+1} - 2$ cubes (not including the enclosed cube).

of all cubes in the enclosure and show that the cubes in $\mathcal{E}$ which contain corners of X in one orientation do not contain corners of X in the "opposite" orientation because of the half-openness, so we get two disjoint sets of $2^d - 1$ cubes in $\mathcal{E}$, so $|\mathcal{E}| \geq 2(2^d - 1) = 2^{d+1} - 2$. This alternate proof is given next.

Theorem (Lower Bound of Theorem 9.5.5)**.** *Let $d \in \mathbb{N}$ and $(X, \mathcal{E})$ an axis-aligned unit cube enclosure in $\mathbb{R}^d$. Then $|\mathcal{E}| \geq 2^{d+1} - 2$.*

Alternate Proof. Let $(X, \mathcal{E})$ be an axis-aligned unit cube enclosure in $\mathbb{R}^d$. By Remark 9.2.12 (Closed or Half-Open Assumption) we may assume that X and each cube in $\mathcal{E}$ is a translate of $[0,1)^d$. For each cube $Y \in \mathcal{E} \sqcup \{X\}$, let Y' denote the set $Y' = \text{center}(Y) + (-\frac{1}{2}, \frac{1}{2}]^d$ (which is the opposite half-open orientation). Let $\mathcal{E}' = \{Y' : Y \in \mathcal{E}\}$. By the same line of reasoning as Remark 9.2.12, $(X', \mathcal{E}')$ is also an axis-aligned unit cube enclosure. Now X' and every cube in $\mathcal{E}'$ is a translate of $(0,1]^d$ (which has the opposite orientation of X and all cubes in $\mathcal{E}$).

Let $\vec{x} = \text{center}(X) = \text{center}(X')$ and observe the following:

$$\text{corners}(X) = \prod_{i=1}^{d} \left\{x_i - \tfrac{1}{2},\ \ x_i + \tfrac{1}{2}\right\} = \text{corners}(X').$$

Being half-open, the only corner of X that is contained in X is $\langle x_i - \frac{1}{2}\rangle_{i=1}^d$, which means each of the other $2^d - 1$ corners of X belong to a cube in $\mathcal{E}$ (and belongs to a unique cube by disjointness). Also note that no cube in $\mathcal{E}$ can contain multiple corners of X because each pair of corners in X is ℓ_∞ distance 1 apart, and no half-open unit cube contains a pair of points ℓ_∞ distance 1 apart. Thus, letting $\mathcal{S} = \{Y \in \mathcal{E} : Y \cap \text{corners}(X) \neq \emptyset\}$ we have $|\mathcal{S}| \geq 2^d - 1$.

Similarly, the only corner of X' (recall these are the same corners as X) contained in X' is $\langle x_i + \frac{1}{2}\rangle_{i=1}^d$ so the other $2^d - 1$ corners belong to $2^d - 1$ distinct members of $\mathcal{E}'$. So let $\mathcal{S}' = \{Y \in \mathcal{E} : Y' \cap \text{corners}(X') \neq \emptyset\}$ which also has cardinality at least $2^d - 1$. We emphasize that $\mathcal{S}' \subseteq \mathcal{E}$ (not $\mathcal{S}' \subseteq \mathcal{E}'$).

We claim that $\mathcal{S}$ and $\mathcal{S}'$ are disjoint subsets of $\mathcal{E}$. To see this, fix an arbitrary $Y \in \mathcal{S}$ and and we will show that $Y \notin \mathcal{S}'$. Let $\vec{y} = \text{center}(Y) = \text{center}(Y')$. Since $Y \overset{\text{adj}}{\sim} X$ by definition of a cube enclosure, then by Fact 3.4.6 (Adjacency for Unit Cubes) there is at least one coordinate $i_0 \in [d]$ such that $|x_{i_0} - y_{i_0}| = 1$. It cannot be the case that $y_{i_0} < x_{i_0}$ because that would imply that $y_{i_0} = x_{i_0} - 1$ so

$$[y_{i_0} - \tfrac{1}{2}, y_{i_0} + \tfrac{1}{2}) = [x_{i_0} - \tfrac{3}{2}, x_{i_0} - \tfrac{1}{2})$$

which does not contain either $x_i - \frac{1}{2}$ or $x_i + \frac{1}{2}$ which would imply that Y does not contain any corners of X (contradicting that $Y \in \mathcal{S}$). Thus, it must be that $y_{i_0} > x_{i_0}$ and specifically $y_{i_0} = x_{i_0} + 1$. Now observe that the i_0th coordinate projection of Y' is

$$(y_{i_0} - \tfrac{1}{2}, y_{i_0} + \tfrac{1}{2}] = (x_{i_0} + \tfrac{1}{2}, x_{i_0} + \tfrac{3}{2}]$$

which does not contain either $x_i - \frac{1}{2}$ or $x_i + \frac{1}{2}$, and thus Y' does not contain a corner of X, so $Y \notin \mathcal{S}'$. This shows that $\mathcal{S}$ and $\mathcal{S}'$ are disjoint as $\mathcal{S}'$ shares no member with $\mathcal{S}$.

Since $\mathcal{S}$ and $\mathcal{S}'$ are disjoint subsets of $\mathcal{E}$ and each have cardinality at least $2^d - 1$, it follows that $|\mathcal{E}| \geq 2(2^d - 1) = 2^{d+1} - 2$. □

We briefly note that the above proof could have used any pair of "complementary" orientations of the half-open cube (there are 2^d orientations, so 2^{d-1} pairs). So one could have $\mathcal{S}_j \sqcup \mathcal{S}'_j$ for each $j \in [2^{d-1}]$, and one could further examine the structure of $\mathcal{E}$ by examining how these sets intersect.

We also want to remark that $\mathcal{S}_j \sqcup \mathcal{S}'_j$ is not in general the set of all cubes in $\mathcal{E}$ as can be seen for example in Figure 9.5 where the bottom right cube does not contain a corner of the enclosed cube in either half-open orientation. With this result, we define a minimum cube enclosure.

Definition 9.5.6 (Minimum Cube Enclosure). Let $d \in \mathbb{N}$ and $(X, \mathcal{E})$ an axis-aligned unit cube enclosure in $\mathbb{R}^d$. We say that $(X, \mathcal{E})$ is a *minimum cube enclosure* if $|\mathcal{E}| = 2^{d+1} - 2$.

Motivated by Theorem 9.5.5 (Minimum Size Cube Enclosure: $2^{d+1} - 2$ Cousins), we examine the structure of cube enclosures at the corners of the enclosed cube X when cubes in $\mathcal{E}$ are cousins. The next two results will eventually be applied to minimum cube enclosures which by Theorem 9.5.5 consist entirely of cousins of X.

Lemma 9.5.7 (Cousin Coordinates at a Corner)**.** *Let $d \in \mathbb{N}$ and $(X, \mathcal{E})$ an axis-aligned unit cube enclosure in $\mathbb{R}^d$ and $\vec{c} \in \text{corners}(X)$ and $Y, Y' \in \mathcal{N}_{\overline{0}}(\vec{c})$ distinct from X and each other. If $Y \overset{\text{cous}}{\sim} X$ and $Y' \overset{\text{cous}}{\sim} X$, then the cousin coordinate of X and Y is different from the cousin coordinate of X and Y'.*

Proof. Let $\vec{x} = \text{center}(X)$, $\vec{y} = \text{center}(Y)$, and $\vec{y}' = \text{center}(Y')$. Because Y and Y' are distinct elements of $\mathcal{E}$, they have disjoint interiors, which means there is some $i_0 \in [d]$ such that $\left|y_{i_0} - y'_{i_0}\right| = 1$. Because $\vec{c} \in \overline{Y}$ it must be that $|c_{i_0} - y_{i_0}| \leq \frac{1}{2}$ (and similarly $\left|c_{i_0} - y'_{i_0}\right| \leq \frac{1}{2}$). The only way for this to happen is if c_{i_0} is distance $\frac{1}{2}$ from both y_{i_0} and y'_{i_0} (i.e. c_{i_0} is the midpoint of y_{i_0} and y'_{i_0}). Without loss of generality, assume $y_{i_0} < y'_{i_0}$. Thus, $c_{i_0} = y_{i_0} + \frac{1}{2}$ and $c_{i_0} = y'_{i_0} - \frac{1}{2}$. Recall that because $\vec{c} \in \text{corners}(X) = \vec{x} + \left\{-\frac{1}{2}, \frac{1}{2}\right\}^d$ we either have $c_{i_0} = x_{i_0} - \frac{1}{2}$ or $c_{i_0} = x_{i_0} + \frac{1}{2}$.

If the former holds (i.e. $c_{i_0} = x_{i_0} - \frac{1}{2}$), then $y_{i_0} = c_{i_0} - \frac{1}{2} = x_{i_0} - 1$ and $y'_{i_0} = c_{i_0} + \frac{1}{2} = x_{i_0}$ which shows that i_0 is the cousin coordinate of X and Y, but it is not the cousin coordinate of X and Y', so in this case the two cousin coordinates are different.

If the latter case holds (i.e. $c_{i_0} = x_{i_0} + \frac{1}{2}$), then $y_{i_0} = c_{i_0} - \frac{1}{2} = x_{i_0}$ and $y'_{i_0} = c_{i_0} + \frac{1}{2} = x_{i_0} + 1$ which shows that i_0 is not the cousin coordinate of X and Y, but it is the cousin coordinate of X and Y', so also in this case the two cousin coordinates are different. $\square$

Theorem 9.5.8 (Minimum Cube Enclosure Equivalencies)**.** *Let $d \in \mathbb{N}$ and $(X, \mathcal{E})$ an axis-aligned unit cube enclosure in $\mathbb{R}^d$. Each of the following is equivalent.*

1. *$\mathcal{E}$ has the minimum possible cardinality*
2. *$|\mathcal{E}| = 2^{d+1} - 2$*
3. *Every cube in $\mathcal{E}$ is a cousin of X*
4. *For each $n \in [d]$, $\mathcal{E}$ contains at least 2^n-many nth-cousins of X*
5. *For each $n \in [d]$, $\mathcal{E}$ contains exactly 2^n-many nth-cousins of X and nothing else*
6. *For each $\vec{c} \in \text{corners}(X)$, $|\mathcal{N}_{\overline{0}}(\vec{c})| \leq d + 1$*
7. *For each $\vec{c} \in \text{corners}(X)$, $|\mathcal{N}_{\overline{0}}(\vec{c})| = d + 1$*
8. *For each $\vec{c} \in \text{corners}(X)$, $\mathcal{N}_{\overline{0}}(\vec{c})$ contains at least X and one nth-cousin of X for each $n \in [d]$*
9. *For each $\vec{c} \in \text{corners}(X)$, $\mathcal{N}_{\overline{0}}(\vec{c})$ contains exactly X and one nth-cousin of X for each $n \in [d]$ and contains nothing else*

Proof. We will prove enough implications to give the following two cyclic implications which proves everything is equivalent to (3):

- (3) $\implies$ (9) $\implies$ (7) $\implies$ (6) $\implies$ (3)
- (3) $\implies$ (9) $\implies$ (8) $\implies$ (4) $\implies$ (5) $\implies$ (2) $\implies$ (1) $\implies$ (3)

Claim A. (1) $\iff$ (2) $\iff$ (3)

Proof of Claim. This is the content of Theorem 9.5.5 (Minimum Size Cube Enclosure: $2^{d+1} - 2$ Cousins). ■

Claim B. (3) $\implies$ (9)

Proof of Claim. Fix an arbitrary corner $\vec{c} \in \text{corners}(X)$. Because each cube in $\mathcal{E}$ is a cousin of X by hypothesis, Lemma 9.5.7 (Cousin Coordinates at a Corner) gives an injection $\mathcal{N}_{\bar{0}}(\vec{c}) \setminus \{X\} \hookrightarrow [d]$ mapping each cube to its cousin-coordinate with X which demonstrates that $|\mathcal{N}_{\bar{0}}(\vec{c}) \setminus \{X\}| \leq d$. Thus, $|\mathcal{N}_{\bar{0}}(\vec{c})| \leq d+1$ (which happens to prove (6)). Then, with this cardinality bound, it follows by Lemma 9.5.2 (Complete Corner Cliques in Cube Enclosures) and Lemma 9.5.3 (Structure of Small Complete Cornered Cliques) that $\mathcal{N}_{\bar{0}}(\vec{c}) \setminus \{X\}$ contains exactly one nth-cousin of X for each $n \in [d]$ and contains nothing else (and also happens to show (7) that $|\mathcal{N}_{\bar{0}}(\vec{c})| = d+1$). ■

Claim C. (9) $\implies$ (7)

Proof of Claim. This is trivial. ■

Claim D. (7) $\implies$ (6)

Proof of Claim. This is trivial. ■

Claim E. (6) $\implies$ (3)

Proof of Claim. Consider arbitrary $Y \in \mathcal{E}$. By Corollary 3.4.15 (Adjacent Cubes Share a Corner) there exists $\vec{c} \in \overline{Y} \cap \text{corners}(X)$. Thus, by definition $Y \in \mathcal{N}_{\bar{0}}(\vec{c})$. By hypothesis, $|\mathcal{N}_{\bar{0}}(\vec{c})| \leq d+1$, so by Lemma 9.5.2 (Complete Corner Cliques in Cube Enclosures) and Lemma 9.5.3 (Structure of Small Complete Cornered Cliques) every member of $\mathcal{N}_{\bar{0}}(\vec{c}) \setminus \{X\}$ is a cousin of X, so in particular Y is a cousin of X. ■

Claim F. (9) $\implies$ (8)

Proof of Claim. This is trivial. ■

Claim G. (8) $\implies$ (4)

Proof of Claim. Consider arbitrary $n \in [d]$. Let $Y \in \mathcal{E}$ be an nth-cousin of X and note the following:

$$\{\vec{c} \in \text{corners}\, X : Y \in \mathcal{N}_{\overline{0}}(\vec{c})\} = \left\{\vec{c} \in \text{corners}\, X : \vec{c} \in \overline{Y}\right\} = \text{corners}(X) \cap \overline{Y}$$

By Lemma 9.4.3 (Equivalent Definitions of nth-Cousins), because $Y \overset{\text{cous}}{\sim}_n X$ the last set has cardinality 2^{d-n}, and so the first set does as well. Thus, each nth-cousin of X in $\mathcal{E}$ belongs to exactly 2^{d-n}-many corner neighborhoods and by hypothesis, each of the 2^d-many corner neighborhoods (which is a subset of $\mathcal{E}$) contains at least one nth-cousin of X. Thus, $\mathcal{E}$ must contain at least $2^d/2^{d-n} = 2^n$-many nth-cousins of X. ∎

Claim H. (4) $\iff$ (5)

Remark. The proof of this claim will in many ways mirror the main points in the proof of Theorem 9.5.5 (Minimum Size Cube Enclosure: $2^{d+1} - 2$ Cousins). △

Proof of Claim. Trivially (5) $\implies$ (4), so we need only show the reverse implication.

By Remark 9.2.12 (Closed or Half-Open Assumption) we assume all cubes are disjoint translates of $[0,1)^d$. Let $\vec{x} = \text{center}(X)$. By Lemma 9.2.10 (Enclosure Enlargment), for all sufficiently small $\varepsilon > 0$ it holds that ${}^{\infty}\overline{B}_{\frac{1}{2}+\varepsilon}(\subseteq) \operatorname{int}\left(X \cup \bigcup_{Y\in\mathcal{E}} Y\right)$. Also, for any $\varepsilon \in (0,\infty)$ we have $A_\varepsilon(X) \subseteq {}^{\infty}\overline{B}_{\frac{1}{2}+\varepsilon}(\vec{x})$. Combining this with the disjointness of all cubes, we have for sufficiently small $\varepsilon > 0$ that

$$A_\varepsilon(X) \subseteq X \sqcup \bigsqcup_{Y\in\mathcal{E}} Y.$$

Thus, we can decompose $A_\varepsilon(X)$ as

$$\begin{aligned} A_\varepsilon(X) &= A_\varepsilon(X) \cap \left(X \sqcup \bigsqcup_{Y \in \mathcal{E}} Y \right) \\ &= (X \cap A_\varepsilon(X)) \sqcup \bigsqcup_{Y \in \mathcal{E}} (Y \cap A_\varepsilon(X)). \end{aligned}$$

Let $\mathcal{E}_{cous} = \left\{ Y \in \mathcal{E} : Y \overset{\text{cous}}{\sim} X \right\}$ and $\mathcal{E}_{other} = \mathcal{E} \setminus \mathcal{E}$. The above allows us to write

$$\begin{aligned} 4^d &= |A_\varepsilon(X)| \\ &= |X \cap A_\varepsilon(X)| + \sum_{Y \in \mathcal{E}} |Y \cap A_\varepsilon(X)| \\ &= |X \cap A_\varepsilon(X)| + \sum_{Y \in \mathcal{E}_{cous}} |Y \cap A_\varepsilon(X)| + \sum_{Y \in \mathcal{E}_{other}} |Y \cap A_\varepsilon(X)| \end{aligned}$$

Now we note that by Lemma 9.4.2 (Equivalent Definition of Cousins) for each $Y \in \mathcal{E}_{cous}$ we have $|Y \cap A_\varepsilon(X)| = 2^{d-1}$ and by Lemma 9.3.1 (ε-Approximation Structure of Adjacency) $|X \cap A_\varepsilon(X)| = 2^d$, so we continue the expression above:

$$= 2^d + |\mathcal{E}_{cous}| \cdot 2^{d-1} + \sum_{Y \in \mathcal{E}_{other}} |Y \cap A_\varepsilon(X)|$$

Next, we observe that because each cube in $\mathcal{E}$ is adjacent to X, Lemma 9.3.1 (ε-Approximation Structure of Adjacency) shows that for each cube in $Y \in \mathcal{E}$, $Y \cap A_\varepsilon(X) \neq \emptyset$, so has cardinality at least 1. We continue the expression with an inequality:

$$\begin{aligned} &\geq 2^d + |\mathcal{E}_{cous}| \cdot 2^{d-1} + \sum_{Y \in \mathcal{E}_{other}} 1 \\ &= 2^d + |\mathcal{E}_{cous}| \cdot 2^{d-1} + |\mathcal{E}_{other}| \end{aligned}$$

We now note that the hypothesis implies that $\mathcal{E}$ contains at least $\sum_{n=1}^{d} 2^n = 2^{d+1} - 2$ cousins, so $|\mathcal{E}_{cous}| \geq 2^{d+1} - 2$ and continue the inequality:

$$\geq 2^d + \left(2^{d+1} - 2\right) 2^{d-1} + |\mathcal{E}_{other}|$$
$$= 4^d + |\mathcal{E}_{other}|$$

This implies that $|\mathcal{E}_{other}| = 0$ (so $\mathcal{E} = \mathcal{E}_{cous}$) and that all inequalities above are actually equalities, so in particular $|\mathcal{E}_{cous}| = 2^{d+1} - 2$. We conclude that $|\mathcal{E}| = |\mathcal{E}_{cous}| = 2^{d+1} - 2 = \sum_{n=1}^{d} 2^d$ and thus $\mathcal{E}$ contains nothing besides the 2^n-many nth-cousins of X that were hypothesized. ■

Claim I. (5) $\implies$ (2)

Proof of Claim. By hypothesis, $|\mathcal{E}| = \sum_{n=1}^{d} 2^d = 2^{d+1} - 2$. ■

Together, all of these implications complete the proof. □

The following lemma is now an immediate corollary but stated as a theorem for emphasis.

Theorem 9.5.9 (Cube Enclosures: $(d+1)$-Secluded Implies Minimum)**.** *Let $d \in \mathbb{N}$ and $(X, \mathcal{E})$ an axis-aligned unit cube enclosure in $\mathbb{R}^d$. If $(X, \mathcal{E})$ is $(d+1)$-secluded, then it is minimum cube enclosure. Furthermore, the reverse implication does not hold in general.*

Proof. By definition of a $(d+1)$-secluded cube enclosure, for every point $\vec{p} \in \mathbb{R}^d$ there exists $\varepsilon \in (0, \infty)$ such that $\left|{}^{\infty}\mathcal{N}_{\varepsilon}(\vec{p})\right| \leq d+1$ which implies by containment that $|\mathcal{N}_{\overline{0}}(\vec{p})| \leq d+1$. This holds in particular for $\vec{p} \in \operatorname{corners}(X)$, so by condition (6) of Theorem 9.5.8 (Minimum Cube Enclosure Equivalencies), $(X, \mathcal{E})$ is a minimal cube enclosure.

For the falsity of the reverse implication, Video 9.5.1 shows an example of a minimum cube enclosure $(X, \mathcal{E})$ in $\mathbb{R}^3$ which is not $(d+1) = 4$-secluded. It is a minimum cube enclosure by condition (2) of Theorem 9.5.8 (Minimum Cube Enclosure Equivalencies) because $\mathcal{E}$ contains $2^{d+1} - 2 = 2^4 - 2 = 14$ cubes. It is not 4-secluded because 4 cubes on the top layer along with the enclosed cube X demonstrate that there is a point where 5 cubes meet at a point. □

Video 9.5.1: Minimum but not $(d+1)$-secluded cube enclosure in $\mathbb{R}^3$

We now make a few observations.

Observation 9.5.10 (Minimum Cube Enclosures Have a Minkowski Twin Pair)**.** *Let $d \in \mathbb{N}$ and $(X, \mathcal{E})$ a minimum axis-aligned unit cube enclosure. Then there exists $Y \in \mathcal{E}$ such that X and Y are a Minkowski twin pair.*

Proof. By condition (4) of Theorem 9.5.8 (Minimum Cube Enclosure Equivalencies) $\mathcal{E}$ contains a 1st-cousin Y of X. As discussed after Definition 9.4.1 (nth-Cousins),

X and Y are 1st-cousins if and only if they are Minkowski twins (essentially by definition). $\square$

Observation 9.5.11 (Cube Enclosures Without a Twin Pair)**.** *Let $d \in \mathbb{N}$ with $d \geq 3$. Then there exists an axis-aligned unit cube enclosure $(X, \mathcal{E})$ such that no cube $Y \in \mathcal{E}$ is a Minkowski twin of X.*

Proof Sketch. The known tetrastix structure gives rise to an axis-aligned unit cube partition $\mathcal{P}$ of $\mathbb{R}^3$ for which some cubes X has no Minkowski twins. This generalizes to each $d > 3$ by using $\mathcal{P}$ as layers which are appropriately offset. Let $\mathcal{E} = \mathcal{N}(X)$ so that $(X, \mathcal{E})$ is an axis-aligned unit cube enclosure by Fact 9.2.6, and X is not a Minkowski twin of any of its neighbors. $\square$

Observation 9.5.12 (Cube Enclosures Without a Twin Pair)**.** *Let $d \in \mathbb{N}$ with $d \geq 8$. Then there exists an axis-aligned unit cube enclosure $(X, \mathcal{E})$ such that no two cubes in $\mathcal{E} \sqcup \{X\}$ are a Minkowski twin pair.*

Proof. Since Keller's conjecture is known to be false in dimensions $d \geq 8$, let $\mathcal{T}$ be an axis-aligned tiling of $\mathbb{R}^d$ which is a counterexample to Keller's conjecture. Pick an arbitrary cube $X \in \mathcal{T}$ and let $\mathcal{E} = \mathcal{N}(X)$. Then $(X, \mathcal{E})$ is an axis-aligned unit cube enclosure by Fact 9.2.6, and because $\mathcal{T}$ contains no Minkowski twin pairs, neither will the subset $\mathcal{E} \sqcup \{X\} \subseteq \mathcal{T}$. $\square$

9.6 Characterizing the Difference Between Minimal and Secluded

We saw in Theorem 9.5.9 (Cube Enclosures: $(d+1)$-Secluded Implies Minimum) that every $(d+1)$-secluded cube enclosure is a minimum cube enclosure, but the converse is not in general true. In this section we partially characterize the distinction between these two types of cube enclosures by giving a condition that can be added to the minimum cardinality hypothesis to imply $(d+1)$-secluded. In order to do so, we have to briefly examine a known result.

9.6.1 Zaks's Neighborly Families of Unit Cubes

In geometry, there is a notion of neighborly sets (c.f. [Zak85, Zak87] and the references therein) where (typically closed) sets $X, Y \subseteq R^d$ are called *neighborly* if they have disjoint interiors and $(d-1)$-dimensional intersection; a family $\mathcal{F}$ of subset of $\mathbb{R}^d$ is called a *neighborly family* if every pair of sets in $\mathcal{F}$ are neighborly. For closed axis-aligned unit cubes, it is quite easy to see that two cubes are neighborly if and only if they are cousins.

Fact 9.6.1 (Neighborly Iff Cousins)**.** *Let $d \in \mathbb{N}$ and X, Y be closed axis-aligned unit cubes in $\mathbb{R}^d$. Then X and Y are neighborly if and only if X and Y are cousins.*

Proof. Let $\vec{x} = \text{center}(X)$ and $\vec{y} = \text{center}(Y)$ so that $X = \prod_{i=1}^{d}[x_i - \frac{1}{2}, x_i + \frac{1}{2}]$ and $Y = \prod_{i=1}^{d}[y_i - \frac{1}{2}, y_i + \frac{1}{2}]$. The intersection $X \cap Y = \prod_{i=1}^{d}[x_i - \frac{1}{2}, x_i + \frac{1}{2}] \cap [y_i - \frac{1}{2}, y_i + \frac{1}{2}]$ is a rectangle, and the dimension is 0 if the intersection is empty, otherwise it is the number of coordinates which are not singleton sets. If the intersection is empty, then X and Y are neither cousins nor neighborly, so we need not deal with this case below. Thus we assume that the intersection is non-empty which implies that $\|\vec{x} - \vec{y}\|_\infty \leq 1$.

Then we have

$$\dim(X \cap Y) = \sum_{i=1}^{d} \begin{cases} 0 & [x_i - \frac{1}{2}, x_i + \frac{1}{2}] \cap [y_i - \frac{1}{2}, y_i + \frac{1}{2}] \text{ is a singleton} \\ 1 & [x_i - \frac{1}{2}, x_i + \frac{1}{2}] \cap [y_i - \frac{1}{2}, y_i + \frac{1}{2}] \text{ is not a singleton} \end{cases}$$

$$= \sum_{i=1}^{d} \begin{cases} 0 & |x_i - y_i| = 1 \\ 1 & |x_i - y_i| \in [0, 1) \end{cases}$$

This shows that the dimension of the intersection is $d - 1$ if and only if there is a unique coordinate $i \in [d]$ such that $|x_i - y_i| = 1$ which is exactly our definition of cousins. □

The main previously known result we will be interested is due to Zaks [Zak85].

Theorem 9.6.2 (Zaks's Theorem)**.** *Let $d \in \mathbb{N}$ and $\mathcal{C}$ be a neighborly family of axis-aligned unit cubes in $\mathbb{R}^d$. Then $|\mathcal{C}| \leq d + 1$.*

Using our language and Fact 9.6.1 (Neighborly Iff Cousins), the theorem can be restated as follows.

Theorem 9.6.3 (Zaks's Theorem)**.** *Let $d \in \mathbb{N}$ and $\mathcal{C}$ be family of axis-aligned unit cubes in $\mathbb{R}^d$ such that for distinct $X, Y \in \mathcal{C}$ we have $X \overset{\text{cous}}{\sim} Y$. Then $|\mathcal{C}| \leq d + 1$.*

Early in our work on this research we were trying to prove exactly this result to use as a lemma, and we suspected it would be a fairly simple result which we could prove using the same types of techniques that have been used in this chapter. However, we were unable to find such a proof, and we learned that there may be a reason for this: Zaks's theorem is in some ways a very deep result.

The way that Zaks proved the theorem was to prove that there was a bijection between neighborly families of axis-aligned unit cubes and complete bipartite decompositions of the complete graph; while this observation is not obvious, the bijection itself is a very simple one. In slightly more detail, a neighborly family of cardinality n in $\mathbb{R}^d$ corresponds to a decomposition of the complete graph on n vertices (K_n) into an edge disjoint union of d-many complete bipartite graphs. Using this bijection, Zaks utilized the Graham-Pollak theorem which says that any such decomposition of K_n requires at least $d \geq n - 1$. Rearranging shows $n \leq d + 1$ so that any neighborly family of axis-aligned unit cubes in $\mathbb{R}^d$ has cardinality at most $d + 1$.

Surprisingly, though the Graham-Pollak theorem is a simple graph theory result, there is no known graph theoretic or combinatorial proof of the theorem—all known proofs utilize (linear) algebraic methods (see [AZ10]). Thus, while Zaks's theorem is an extremely simple sounding statement, there is no known geometric proof of it (as such a proof would also constitute a non-algebraic proof of the Graham-Pollak theorem). Motivated by this, we spent quite a bit of time trying to find a geometric proof of Zaks's theorem by building on our research, but we were unsuccessful.

9.6.2 Difference Characterization

In the definition of an axis-aligned unit cube enclosure $(X, \mathcal{E})$, we require that X is adjacent to each cube in $\mathcal{E}$, and we have seen that if $|\mathcal{E}|$ is minimized it implies that all cubes in $\mathcal{E}$ are cousins of X. But this in some sense establishes a non-uniformity in the sense that we of X that every cube adjacent to X is actually a cousin of X, but we do not know the same thing about cubes in $\mathcal{E}$. We will call a cube enclosure *cousinly* if X is not special in this regard and if we know for every $Y \in \mathcal{E} \sqcup \{X\}$ that Y is a cousin of every cube in $\mathcal{E} \sqcup \{X\}$ which it is adjacent to.

Definition 9.6.4 (Cousinly Cube Enclosure). Let $d \in \mathbb{N}$ and $(X, \mathcal{E})$ an axis-aligned unit cube enclosure in $\mathbb{R}^d$. We say that $(X, \mathcal{E})$ is a *cousinly cube enclosure* if the following property holds: for all $Y, Y' \in \mathcal{E} \sqcup \{X\}$ if $Y \overset{\text{adj}}{\sim} Y'$ then $Y \overset{\text{cous}}{\sim} Y'$.

Compare this to a definition of minimum cube enclosure which is equivalent[11] to Definition 9.5.6 (Minimum Cube Enclosure) by Theorem 9.5.5 (Minimum Size Cube Enclosure: $2^{d+1} - 2$ Cousins) and stated in an analogous manner.

Definition 9.6.5 (Minimal Cube Enclosure (Alternate)). Let $d \in \mathbb{N}$ and $(X, \mathcal{E})$ an axis-aligned unit cube enclosure in $\mathbb{R}^d$. We say that $(X, \mathcal{E})$ is a *minimum cube enclosure* if the following property holds: for all $Y \in \mathcal{E}$ if $Y \overset{\text{adj}}{\sim} X$ then $Y \overset{\text{cous}}{\sim} X$.

This should make clear that a cousinly cube enclosure is a stronger condition than a minimum cube enclosure. In fact, it is strictly stronger because it implies $(d+1)$-secluded.

Theorem 9.6.6 (Cube Enclosures: Cousinly Implies $(d+1)$-Secluded)**.** *Let $d \in \mathbb{N}$ and $(X, \mathcal{E})$ an axis-aligned unit cube enclosure in $\mathbb{R}^d$. If $(X, \mathcal{E})$ is a cousinly cube enclosure then it is a $(d+1)$-secluded cube enclosure.*

Proof. Let $\vec{p} \in \mathbb{R}^d$ be arbitrary and consider $\mathcal{N}_{\bar{0}}(\vec{p})$. This is a clique (i.e. all pairs of cubes in $\mathcal{N}_{\bar{0}}(\vec{p})$ are adjacent), so because $(X, \mathcal{E})$ is cousinly, all pairs of cubes in $\mathcal{N}_{\bar{0}}(\vec{p})$ are cousins. By Theorem 9.6.3 (Zaks's Theorem) $|\mathcal{N}_{\bar{0}}(\vec{p})| \leq d+1$ and by Lemma 9.0.2 (Equivalent Definitions of k-Secluded For Unit Cubes) this shows that $(X, \mathcal{E})$ is $(d+1)$-secluded. □

[11] If $(X, \mathcal{E})$ is minimum by the standard definition, then by Theorem 9.5.5 every cube in $\mathcal{E}$ is a cousin of X, so the stated property holds. Conversely, if the stated property holds for $(X, \mathcal{E})$ then in particular because each $Y \in \mathcal{E}$ is adjacent to X by definition of a cube enclosure, we have by the property that each $Y \in \mathcal{E}$ is a cousin of X, so by Theorem 9.5.5, $(X, \mathcal{E})$ is a minimum cube enclosure by the standard definition.

We believe that cousinly is actually equivalent to $(d+1)$-secluded and have work to support this conjecture that is not yet formalized.

Conjecture 9.6.7 (Cube Enclosures: Cousinly Implies $(d+1)$-Secluded)**.** *Let $d \in \mathbb{N}$ and $(X, \mathcal{E})$ an axis-aligned unit cube enclosure in $\mathbb{R}^d$. Then $(X, \mathcal{E})$ is a cousinly cube enclosure if and only if it is a $(d+1)$-secluded cube enclosure.*

9.7 Partition Implications

Corollary 9.7.1 (Minimum Number of Neighbors in Partition)**.** *Let $d \in \mathbb{N}$ and $\mathcal{P}$ be an axis-aligned unit cube partition of $\mathbb{R}^d$. Then for every $X \in \mathcal{P}$, it must be that X has at least $2^{d+1} - 2$ neighbors (i.e. every vertex in the partition graph has degree at least $2^{d+1} - 2$).*

Proof. By Fact 9.2.6 (Partition Neighbors Give Cube Enclosures), $(X, \mathcal{N}(X))$ is an axis-aligned unit cube enclosure, so by Theorem 9.5.5 (Minimum Size Cube Enclosure: $2^{d+1} - 2$ Cousins), $|\mathcal{N}(X)| \geq 2^{d+1} - 2$. □

With this result, we define a minimum-degree unit cube partition similar to how we did with cube enclosures in Definition 9.5.6 (Minimum Cube Enclosure).

Definition 9.7.2 (Minimum Degree Unit Cube Partition). Let $d \in \mathbb{N}$ and $\mathcal{P}$ an axis-aligned unit cube partition of $\mathbb{R}^d$. We say that $\mathcal{P}$ is a *minimum degree partition* if for every $X \in \mathcal{P}$ we have $|\mathcal{N}(X)| = 2^{d+1} - 2$.

Remark 9.7.3. Probably we should use a more qualified name such as *minimum degree axis-aligned unit cube partition*, but that quickly gets verbose and we won't discuss any other types of partitions in this chapter. △

Now we present the analog of Theorem 9.5.9 (Cube Enclosures: $(d+1)$-Secluded Implies Minimum) for partitions, but in the case of partitions we have not only that $(d+1)$-secluded implies minimum but also that minimum also implies $(d+1)$-secluded.

Theorem 9.7.4 ($(d+1)$-Secluded Iff Minimum Degree Iff Neighborly)**.** *Let $d \in \mathbb{N}$ and $\mathcal{P}$ be an axis-aligned unit cube partition of $\mathbb{R}^d$. Then the following are equivalent:*

1. *$\mathcal{P}$ is $(d+1)$-secluded*
2. *$\mathcal{P}$ is a minimum degree partition*
3. *Every clique in $\mathcal{P}$ is a neighborly family*

Proof. (1) $\implies$ (2): Let $X \in \mathcal{P}$ be arbitrary and let $\mathcal{E} = \mathcal{N}(X)$. By Fact 9.2.6 (Partition Neighbors Give Cube Enclosures), $(X, \mathcal{E})$ is an axis-aligned unit cube enclosure in $\mathbb{R}^d$. Because $\mathcal{P}$ is $(d+1)$-secluded, it follows trivially that the cube enclosure $(X, \mathcal{E})$ is $(d+1)$-secluded because $\{X\} \sqcup \mathcal{E} \subseteq \mathcal{P}$. By Theorem 9.5.9 (Cube Enclosures: $(d+1)$-Secluded Implies Minimum), $(X, \mathcal{E})$ is a minimum cube enclosure, so by Theorem 9.5.5 (Minimum Size Cube Enclosure: $2^{d+1}-2$ Cousins) $|\mathcal{E}| = 2^{d+1} - 2$, so by definition (Definition 9.7.2 (Minimum Degree Unit Cube Partition)) $\mathcal{P}$ is a minimum degree partition.

(2) $\implies$ (3) Consider any clique $\mathcal{C} \subseteq \mathcal{P}$. Let $X, Y \in \mathcal{C}$ and let $\mathcal{E} = \mathcal{N}(X)$ (taken with respect to $\mathcal{P}$, not $\mathcal{C}$) and note that $\mathcal{E}$ contains Y and by hypothesis $|\mathcal{E}| = 2^{d+1}-2$. Then by Fact 9.2.6 (Partition Neighbors Give Cube Enclosures), $(X, \mathcal{E})$ is a minimum cube enclosure, so by Theorem 9.5.5 (Minimum Size Cube Enclosure: $2^{d+1}-2$ Cousins) every member of $\mathcal{E}$ is a cousin of X which means in particular Y is a cousin of X, so by Fact 9.6.1 (Neighborly Iff Cousins) X and Y are neighborly. Because X and Y were an arbitrary, ever pair of cubes in $\mathcal{C}$ are neighborly, so $\mathcal{C}$ is a neighborly family.

(3) $\implies$ (1) Let $\vec{p} \in \mathbb{R}^d$ be arbitrary and consider $\mathcal{N}_{\vec{0}}(\vec{p})$. This is a clique, so by hypothesis it is a neighborly family, so by Theorem 9.6.2 (Zaks's Theorem) $|\mathcal{N}_{\vec{0}}(\vec{p})| \leq d+1$ and by Lemma 9.0.2 (Equivalent Definitions of k-Secluded For Unit Cubes) this shows that $\mathcal{P}$ is $(d+1)$-secluded.

□

We can reinterpret this result as follows.

Corollary 9.7.5. *Let $d \in \mathbb{N}$ and $\mathcal{P}$ be an axis-aligned unit cube partition of $\mathbb{R}^d$. Then the maximum vertex degree of the partition graph of $\mathcal{P}$ is minimized if and only if the maximum clique size in the partition graph of $\mathcal{P}$ is minimized.*

We want to offer some discussion about why this equivalence emerges in the case of partitions but is not present in the case of cube enclosures. The reason is that in showing that minimum implied $(d+1)$-secluded we used the fact that we could show in the partition that all adjacent cubes are actually cousins; in a cube enclosure this is not necessarily the case[12]. We alluded to this in Section 9.6 (Characterizing the Difference Between Minimal and Secluded) where we discussed that minimal cube enclosure allow this non-uniformity because the enclosed cube is special in that context. However, in the partition context no cube is special because we can choose to create a cube enclosure around any cube in the partition that we want. This freedom allows us to take any two adjacent cubes in the partition and find a cube enclosure where one of them is the enclosed cube which guarantees by Theorem 9.5.5 (Minimum Size Cube Enclosure: $2^{d+1} - 2$ Cousins) that they must be cousins. In a cube enclosure we cannot do this because we don't have enough cubes available to find a second cube enclosure within an initial cube enclosure. This demonstrates that

[12] Video 9.5.1 shows a minimum cube enclosure in $\mathbb{R}^3$, but there are pairs of diagonal cubes in the top layer which are adjacent but not cousins because they have positions differing in 2 distinct coordinates.

this equivalence doesn't really emerge as a global property; it is still a fairly local property.

Remark 9.7.6. Though probably not surprising, it is noteworthy that the degree can be simultaneously minimized across all members in axis-aligned unit cube partitions; Corollary 9.7.1 (Minimum Number of Neighbors in Partition) indicates that every member in an axis-aligned unit cube partition must have degree at least $2^{d+1} - 2$, and by Theorem 9.7.4 (($d+1$)-Secluded Iff Minimum Degree Iff Neighborly) each member in our reclusive partition constructions (which are $(d+1, \frac{1}{2d})$-secluded axis-aligned unit cube partitions) achieve this degree simultaneously for every member. The result shows that all vertex degrees are minimized if and only if the size of the largest clique is minimized (by Lemma 9.0.2 (Equivalent Definitions of k-Secluded For Unit Cubes)).

In other words, if for a given partition $\mathcal{P}$, we let $\text{min-degree}(\mathcal{P}) = \min_{X \in \mathcal{P}} |\mathcal{N}(X)|$ and $\text{max-degree}(\mathcal{P}) = \max_{X \in \mathcal{P}} |\mathcal{N}(X)|$, and then considered the values

$$m = \min \left\{ \text{min-degree}(\mathcal{P}) : \mathcal{P} \text{ an axis-aligned unit cube partition of } \mathbb{R}^d \right\}$$
$$n = \min \left\{ \text{max-degree}(\mathcal{P}) : \mathcal{P} \text{ an axis-aligned unit cube partition of } \mathbb{R}^d \right\}$$

we would obviously have $m \leq n$, but it would not have been surprising if $m < n$ so that no partition can achieve the minimum degree across all members. This is especially true because the underlying structure we are examining is an infinite graph, and graphs are known to have disconnects between the local conditions and global conditions (for example, the clique number and chromatic number for a graph are in general not the same). △

We offer a few observations regarding Keller's conjecture (see Chapter 1). The first is that every cube in $(d+1)$-secluded axis-aligned unit cube partition contains a twin; this is less a result about Keller's conjecture and more a result indicating that the $(d+1)$-secluded axis-aligned unit cube partitions contain a lot of structure.

Observation 9.7.7 ($(d+1)$-Secluded Partitions Contain Twin Pairs)**.** *Let $d \in \mathbb{N}$ and $\mathcal{P}$ a $(d+1)$-secluded axis-aligned unit cube partition of $\mathbb{R}^d$. Then for every cube $X \in \mathcal{P}$ there is a Minkowski twin of X in $\mathcal{P}$.*

Proof. This follows applying Fact 9.2.6 (Partition Neighbors Give Cube Enclosures) then Theorem 9.5.9 (Cube Enclosures: $(d+1)$-Secluded Implies Minimum) then Observation 9.5.10 (Minimum Cube Enclosures Have a Minkowski Twin Pair). □

Observation 9.7.8 (Degree of Keller's Conjecture Counterexamples)**.** *Let $d \in \mathbb{N}$ and suppose $\mathcal{P}$ is an axis-aligned unit cube partition of $\mathbb{R}^d$ which is a counterexample of Keller's conjecture. Then every cube in $\mathcal{P}$ has at least $2^{d+1}-1$ neighbors.*

Proof. If even one cube $X \in \mathcal{P}$ had fewer neighbors, then it would have at most $2^{d+1}-2$ neighbors (and thus exactly $2^{d+1}-2$ neighbors by Corollary 9.7.1 (Minimum Number of Neighbors in Partition)), so the $(X, \mathcal{N}(()X))$ would be a minimum cube enclosure and thus by Observation 9.5.10 (Minimum Cube Enclosures Have a Minkowski Twin Pair) would contain a twin of X. □

We conjecture that there is another equivalence that can be added to Theorem 9.7.4 ($(d+1)$-Secluded Iff Minimum Degree Iff Neighborly).

Conjecture 9.7.9. *We can add the following item to Theorem 9.7.4 ($(d+1)$-Secluded Iff Minimum Degree Iff Neighborly):*

4. *For every clique $\mathcal{C} \subseteq \mathcal{P}$ the set $\{\text{center}(Y) : Y \in \mathcal{C}\}$ is affinely independent.*

It is trivial that this new item would imply the others because an affinely independent subset of $\mathbb{R}^d$ can have cardinality at most $d+1$, so this would imply that any clique has cardinality at most $d+1$ which by Lemma 9.0.2 (Equivalent Definitions of k-Secluded For Unit Cubes) is an equivalent condition to being $(d+1)$-secluded. One of the reasons we suspect that this new item follows from (3) is because as discussed in Subsection 9.6.1 (Zaks's Neighborly Families of Unit Cubes), Zaks's theorem is equivalent to the Graham-Pollak theorem which is typically proved with linear algebraic techniques of showing some set is linearly independent.

9.8 Optimal ε For Unit Cube Enclosures

In this section, we show that if an axis-aligned unit cube enclosure (and consequently an axis-aligned unit cube partition) is $(d+1, \varepsilon)$-secluded, then it must be that $\varepsilon \leq \frac{1}{2d}$. This matches the constructions given in Chapter 4. We begin with two lemmas which say nothing surprising given what we have already shown, but we present them anyway in order to be formal about our results.

The first lemma is just a generalization of the proof of Theorem 5.1.1 from first principles in Subsection 5.1.3. It really says nothing new other than that we didn't need an entire partition to run our argument before—a point at the corner of a cube is at the closure of at least $d+1$ cubes in total (if the point is completely surrounded by cubes in a cube packing).

Lemma 9.8.1 ($(d+1)$-Many Cubes At a Corner)**.** *Let $d \in \mathbb{N}$ and $\mathcal{F}$ a non-empty packing of axis-aligned unit cubes in $\mathbb{R}^d$. Let $X \in \mathcal{F}$ and $\vec{c} \in \operatorname{corners}(X)$. If it is the case that $\vec{c} \in \operatorname{int}\left(\bigcup_{Y \in \mathcal{F}} \overline{Y}\right)$, then $\vec{c}$ belongs to the closure of at least $d+1$ cubes in $\mathcal{F}$.*

Proof. We will assume without loss of generality[13] that $\vec{c}$ is the specific corner $\vec{c} = \langle \text{center}(X) - \frac{1}{2} \rangle_{i=1}^{d}$. We will assume without loss of generality[14] that all cubes in $\mathcal{F}$ are translates of $[0,1)^d$.

For each $\varepsilon \in (0, \infty)$, let

$$\begin{aligned} E_\varepsilon &= \vec{c} + \prod_{i=1}^{d} \{-\varepsilon, \varepsilon\} \\ &= \prod_{i=1}^{d} \{c_i - \varepsilon, c_i + \varepsilon\} \end{aligned}$$

For $\varepsilon > 0$ sufficiently small, by Fact 3.6.5 we have $\mathcal{N}_{\overline{0}}(\vec{c}) = {}^{\infty}\overline{\mathcal{N}}_{\varepsilon}(\vec{c})$ and since we have

$$\begin{aligned} {}^{\infty}\overline{\mathcal{N}}_{\varepsilon}(\vec{c}) &= \left\{ Y \in \mathcal{F} : Y \cap \left(\vec{c} + [-\varepsilon, \varepsilon]^d \right) \neq \emptyset \right\} \\ &\supseteq \{ Y \in \mathcal{F} : Y \cap E_\varepsilon \neq \emptyset \} \end{aligned}$$

we can get a lower bound on the number of cubes with $\vec{c}$ in their closure (i.e. $|\mathcal{N}_{\overline{0}}(\vec{c})|$) by counting how many cubes intersect the set E_ε. We denote this set as $\mathcal{M}_\varepsilon = \{ Y \in \mathcal{F} : Y \cap E_\varepsilon \neq \emptyset \}$.

Observe that because $\vec{c} \in \text{int}\left(\bigcup_{Y \in \mathcal{F}} \overline{Y} \right)$ by hypothesis, and the interior is an open set, there exists some open ball centered at $\vec{c}$ which is also a subset of $\text{int}\left(X \cup \bigcup_{Y \in \mathcal{E}} Y \right)$, and thus it holds for all $\varepsilon > 0$ sufficiently small that ${}^{\infty}\overline{B}_{\varepsilon}(\vec{c}) \subseteq \text{int}\left(\bigcup_{Y \in \mathcal{F}} \overline{Y} \right)$. By Corollary 3.6.7 (because we are assuming half-open members), we obtain a disjoint union: ${}^{\infty}\overline{B}_{\varepsilon}(\vec{c}) \subseteq \text{int}\left(\bigcup_{Y \in \mathcal{F}} \overline{Y} \right) = \text{int}\left(\bigsqcup_{Y \in \mathcal{F}} Y \right) \subseteq \bigsqcup_{Y \in \mathcal{F}} Y$. Thus, because $E_\varepsilon \subseteq {}^{\infty}\overline{B}_{\varepsilon}(\vec{c})$,

---

[13] If $\vec{c}$ is a different corner, then we can transform the family $\mathcal{F}$ into a new family $\mathcal{F}'$ be negating the coordinates where $c_i = \text{center}(X) + \frac{1}{2}$. Then $\vec{c}$ is transformed to $\langle \text{center}(X) - \frac{1}{2} \rangle_{i=1}^{d}$ and still belongs to the closure of the same number of cubes in $\mathcal{F}$.

[14] We can do so because this doesn't change what the closures are (Fact 3.4.3) and doesn't change whether the family is a packing (Corollary 3.6.7).

each of the 2^d points in E_ε belongs to a unique cube in $\mathcal{F}$ and in particular, to a unique cube in $\mathcal{M}_\varepsilon$.

Next, we show that X contains exactly one point of E_ε for any $\varepsilon \in (0,1)$. That is, $E_\varepsilon \cap X$ has cardinality 1. Let $\vec{x} = \text{center}(X)$.

$$\begin{aligned} E_\varepsilon \cap X &= \prod_{i=1}^{d} \{c_i - \varepsilon, c_i + \varepsilon\} \cap [x_i - \tfrac{1}{2}, x_i + \tfrac{1}{2}) \\ &= \prod_{i=1}^{d} \left\{x_i - \tfrac{1}{2} - \varepsilon, x_i - \tfrac{1}{2} + \varepsilon\right\} \cap [x_i - \tfrac{1}{2}, x_i + \tfrac{1}{2}) && \text{(Value of } \vec{c}\text{)} \\ &= \prod_{i=1}^{d} \left\{x_i - \tfrac{1}{2} + \varepsilon\right\} && \text{(if } \varepsilon \in (0,1)\text{)} \end{aligned}$$

which is a set of cardinality 1.

Now observe that each cube in $\mathcal{F}$ intersects either 0 or a power of 2 points in E_ε because for $Y \in \mathcal{F}$ we have (letting $\vec{y} = \text{center}(Y)$) that

$$E_\varepsilon \cap Y = \prod_{i=1}^{d} \begin{cases} \{c_i - \varepsilon, c_i + \varepsilon\} \cap [y_i - \frac{1}{2}, y_i + \frac{1}{2}) & i \neq i_0 \\ \{c_{i_0}\} \cap [y_i - \frac{1}{2}, y_i + \frac{1}{2}) & i \neq i_0 \end{cases}$$

and so

$$|E_\varepsilon \cap Y| = \prod_{i=1}^{d} \begin{cases} \left|\{c_i - \varepsilon, c_i + \varepsilon\} \cap [y_i - \frac{1}{2}, y_i + \frac{1}{2})\right| & i \neq i_0 \\ \left|\{c_{i_0}\} \cap [y_i - \frac{1}{2}, y_i + \frac{1}{2})\right| & i \neq i_0 \end{cases}$$

which is a product of 0's, 1's, or 2's, and thus the cardinality is either 0 or a power of 2. And since every cube in $\mathcal{M}_\varepsilon$ by definition has non-trivial intersection with E_ε, each cube in $\mathcal{M}_\varepsilon$ intersects a power of 2 points in E_ε.

Thus, (recalling that every point in E_ε belongs to exactly one cube in $\mathcal{M}_\varepsilon$ for sufficiently small $\varepsilon > 0$) and also noting that $X \in \mathcal{M}_\varepsilon$, we have

$$2^d = |E_\varepsilon| = \sum_{Y \in \mathcal{M}_\varepsilon} |Y \cap E_\varepsilon| = |X \cap E_\varepsilon| + \sum_{Y \in \mathcal{M}_\varepsilon \setminus \{X\}} |Y \cap E_\varepsilon| = 1 + \sum_{Y \in \mathcal{M}_\varepsilon \setminus \{X\}} |Y \cap E_\varepsilon|.$$

Since each term of this summation is a power of two, then by Lemma 5.1.5, there are at least d terms in the summation—i.e. $|\mathcal{M}_\varepsilon \setminus \{X\}| \geq d$ so $|\mathcal{M}_\varepsilon| \geq d + 1$ showing that $|\mathcal{N}_{\vec{0}}(\vec{c})| \geq |\mathcal{M}_\varepsilon| \geq d + 1$. □

The second lemma says that a point along an edge of a cube is at the closure of at least d cubes in total (if the point is completely surrounded by cubes in a cube packing). This is analogous to Lemma 9.8.1 which says that a point at the corner of a cube is at the closure of at least $d + 1$ cubes in total (if the point is completely surrounded by cubes in a cube enclosure). Morally, this result actually follows directly from Lemma 9.8.1 because we can assume half-open cubes and just project out the coordinate along which the edge exists so that the point is at the corner of a cube in a $(d-1)$-dimensional structure and apply Lemma 9.8.1. However, we haven't formally introduced this notion of projection (which is actually intersecting the structure with a hyperplane) and so we will prove it directly. The proof heavily mirrors the proof of Lemma 9.8.1.

Lemma 9.8.2 (d-Many Cubes Along an Edge)**.** *Let $d \in \mathbb{N}$ and $\mathcal{F}$ a non-empty packing of axis-aligned unit cubes in $\mathbb{R}^d$, and let $i_0 \in [d]$. Let $X \in \mathcal{F}$. Let $\vec{p}, \vec{q} \in \operatorname{corners}(X)$ such that $\vec{q} - \vec{p} = \vec{e}^{(i_0)}$ (i.e. $\vec{p}$ and $\vec{q}$ define an edge of X along direction i_0). Let $\vec{c} \in \mathbb{R}^d$ such that $c_{i_0} \in [p_{i_0}, q_{i_0}]$ and for $i \in [d] \setminus \{i_0\}$, $c_i = p_i = q_i$ (i.e. $\vec{c}$ is a point on the edge of X between $\vec{p}$ and $\vec{q}$). If it is the case that $\vec{c} \in \operatorname{int}\left(\bigcup_{Y \in \mathcal{F}} \overline{Y}\right)$, then $\vec{c}$ belongs to the closure of at least d cubes in $\mathcal{F}$.*

Proof. If $c_{i_0} = p_{i_0}$ (resp. $c_{i_0} = q_{i_0}$) then $\vec{c} = \vec{p}$ (resp. $\vec{c} = \vec{q}$) which is a corner of X, so by Lemma 9.8.1, $\vec{c}$ belongs to the closure of at least $d+1$ cubes in $\mathcal{F}$. Thus, we assume $c_{i_0} \in (p_{i_0}, q_{i_0})$.

We will assume without loss of generality that all cubes in $\mathcal{F}$ are translates of $[0,1)^d$ (we can do so because this doesn't change what the closures are (fact 3.4.3) and doesn't change whether the family is a packing (Corollary 3.6.7)).

For each $\varepsilon \in (0,\infty)$, let

$$E_\varepsilon = \vec{c} + \prod_{i=1}^{d} \begin{cases} \{-\varepsilon, \varepsilon\} & i \neq i_0 \\ \{0\} & i = i_0 \end{cases}$$
$$= \prod_{i=1}^{d} \begin{cases} \{c_i - \varepsilon, c_i + \varepsilon\} & i \neq i_0 \\ \{c_{i_0}\} & i = i_0 \end{cases}$$

For $\varepsilon > 0$ sufficiently small, by Fact 3.6.5 we have $\mathcal{N}_{\vec{0}}(\vec{c}) = {}^{\infty}\overline{\mathcal{N}}_\varepsilon(\vec{c})$ and since we have

$${}^{\infty}\overline{\mathcal{N}}_\varepsilon(\vec{c}) = \left\{Y \in \mathcal{F} : Y \cap \left(\vec{c} + [-\varepsilon, \varepsilon]^d\right) \neq \emptyset\right\}$$
$$\supseteq \{Y \in \mathcal{F} : Y \cap E_\varepsilon \neq \emptyset\}$$

we can get a lower bound on the number of cubes with $\vec{c}$ in their closure (i.e. $|\mathcal{N}_{\vec{0}}(\vec{c})|$) by counting how many cubes intersect the set E_ε. We denote this set as $\mathcal{M}_\varepsilon = \{Y \in \mathcal{F} : Y \cap E_\varepsilon \neq \emptyset\}$.

Observe that because $\vec{c} \in \operatorname{int}\left(\bigcup_{Y \in \mathcal{F}} \overline{Y}\right)$ by hypothesis, and the interior is an open set, there exists some open ball centered at $\vec{c}$ which is also a subset of $\operatorname{int}\left(X \cup \bigcup_{Y \in \mathcal{E}} Y\right)$, and thus it holds for all $\varepsilon > 0$ sufficiently small that ${}^{\infty}\overline{B}_\varepsilon(\vec{c}) \subseteq \operatorname{int}\left(\bigcup_{Y \in \mathcal{F}} \overline{Y}\right)$. By Corollary 3.6.7 (because we are assuming half-open members), we obtain a disjoint union: ${}^{\infty}\overline{B}_\varepsilon(\vec{c}) \subseteq \operatorname{int}\left(\bigcup_{Y \in \mathcal{F}} \overline{Y}\right) = \operatorname{int}\left(\bigsqcup_{Y \in \mathcal{F}} Y\right) \subseteq \bigsqcup_{Y \in \mathcal{F}} Y$. Thus, because $E_\varepsilon \subseteq {}^{\infty}\overline{B}_\varepsilon(\vec{c})$, each of the 2^{d-1} points in E_ε belongs to a unique cube in $\mathcal{F}$ and in particular, to a unique cube in $\mathcal{M}_\varepsilon$.

Next, we show that X contains exactly one point of E_ε for any $\varepsilon \in (0,1)$. Let $\vec{x} = \text{center}(X)$. Observe that because $\vec{p}, \vec{q} \in \text{corners}(X) = \vec{x} + \left\{-\frac{1}{2}, \frac{1}{2}\right\}^d = \prod_{i=1}^d \left\{x_i - \frac{1}{2}, x_i + \frac{1}{2}\right\}$ we have for all $i \in [d]$ that $p_i, q_i \in \left\{x_i - \frac{1}{2}, x_i + \frac{1}{2}\right\}$. Then, because $p_{i_0} < q_{i_0}$, we must have $p_{i_0} = x_{i_0} - \frac{1}{2}$ and $q_{i_0} = x_{i_0} + \frac{1}{2}$.

We also claim that for $i \in [d] \setminus \{i_0\}$ and $\varepsilon \in (0,1)$ we have $\{p_i - \varepsilon, p_i + \varepsilon\} \cap [x_i - \frac{1}{2}, x_i + \frac{1}{2}) = \{p_i - \varepsilon \cdot \text{sign}(p_i - x_i)\}$ which can be seen as follows in two cases: (1) $p_i = x_i - \frac{1}{2}$ and (2) $p_i = x_i + \frac{1}{2}$.

In case (1), we have $p_i = x_i - \frac{1}{2}$, so $\text{sign}(p_i - x_i) = -1$ and this gives

$$\begin{aligned}
\{p_i - \varepsilon, p_i + \varepsilon\} \cap [x_i - \tfrac{1}{2}, x_i + \tfrac{1}{2}) &= \left\{x_i - \tfrac{1}{2} - \varepsilon, x_i - \tfrac{1}{2} + \varepsilon\right\} \cap [x_i - \tfrac{1}{2}, x_i + \tfrac{1}{2}) \\
&= \left\{x_i - \tfrac{1}{2} + \varepsilon\right\} \qquad \text{(if } \varepsilon \in (0,1)\text{)} \\
&= \{p_i + \varepsilon\} \\
&= \{p_i - \varepsilon \cdot \text{sign}(p_i - x_i)\}
\end{aligned}$$

and in case (2), we have $p_i = x_i + \frac{1}{2}$, so $\text{sign}(p_i - x_i) = 1$ and this gives

$$\begin{aligned}
\{p_i - \varepsilon, p_i + \varepsilon\} \cap [x_i - \tfrac{1}{2}, x_i + \tfrac{1}{2}) &= \left\{x_i + \tfrac{1}{2} - \varepsilon, x_i + \tfrac{1}{2} + \varepsilon\right\} \cap [x_i - \tfrac{1}{2}, x_i + \tfrac{1}{2}) \\
&= \left\{x_i + \tfrac{1}{2} - \varepsilon\right\} \qquad \text{(if } \varepsilon \in (0,1)\text{)} \\
&= \{p_i - \varepsilon\} \\
&= \{p_i - \varepsilon \cdot \text{sign}(p_i - x_i)\}
\end{aligned}$$

so in either case $\{p_i - \varepsilon, p_i + \varepsilon\} \cap [x_i - \frac{1}{2}, x_i + \frac{1}{2}) = \{p_i - \varepsilon \cdot \text{sign}(p_i - x_i)\}$. This means that for $i \in [d] \setminus \{i_0\}$ because $c_i = p_i$ we have $\{c_i - \varepsilon, c_i + \varepsilon\} \cap [x_i - \frac{1}{2}, x_i + \frac{1}{2}) =$

$\{c_i - \varepsilon \cdot \text{sign}(p_i - x_i)\}$.

Now we can compute $E_\varepsilon \cap X$ and show in particular that it has cardinality 1.

$$E_\varepsilon \cap X = \prod_{i=1}^{d} \begin{cases} \{c_i - \varepsilon, c_i + \varepsilon\} \cap [x_i - \frac{1}{2}, x_i + \frac{1}{2}) & i \neq i_0 \\ \{c_{i_0}\} \cap [x_i - \frac{1}{2}, x_i + \frac{1}{2}) & i = i_0 \end{cases}$$
$$= \prod_{i=1}^{d} \begin{cases} \{c_i - \varepsilon \cdot \text{sign}(p_i - x_i)\} & i \neq i_0 \\ \{c_{i_0}\} & i = i_0 \end{cases}$$

which is a set of cardinality 1. Note that the second line used the hypothesis that $c_{i_0} \in (p_{i_0}, q_{i_0}) = (x_{i_0} - \frac{1}{2}, x_{i_0} + \frac{1}{2})$.

Now observe that each cube in $\mathcal{F}$ intersects either 0 or a power of 2 points in E_ε because for $Y \in \mathcal{F}$ we have (letting $\vec{y} = \text{center}(Y)$) that

$$E_\varepsilon \cap Y = \prod_{i=1}^{d} \begin{cases} \{c_i - \varepsilon, c_i + \varepsilon\} \cap [y_i - \frac{1}{2}, y_i + \frac{1}{2}) & i \neq i_0 \\ \{c_{i_0}\} \cap [y_i - \frac{1}{2}, y_i + \frac{1}{2}) & i \neq i_0 \end{cases}$$

and so

$$|E_\varepsilon \cap Y| = \prod_{i=1}^{d} \begin{cases} \left|\{c_i - \varepsilon, c_i + \varepsilon\} \cap [y_i - \frac{1}{2}, y_i + \frac{1}{2})\right| & i \neq i_0 \\ \left|\{c_{i_0}\} \cap [y_i - \frac{1}{2}, y_i + \frac{1}{2})\right| & i \neq i_0 \end{cases}$$

which is a product of 0's, 1's, or 2's, and thus the cardinality is either 0 or a power of 2. And since every cube in $\mathcal{M}_\varepsilon$ by definition has non-trivial intersection with E_ε, each cube in $\mathcal{M}_\varepsilon$ intersects a power of 2 points in E_ε.

Thus, (recalling that every point in E_ε belongs to exactly one cube in $\mathcal{M}_\varepsilon$ for sufficiently small $\varepsilon > 0$) and also noting that $X \in \mathcal{M}_\varepsilon$, we have

$$2^{d-1} = |E_\varepsilon| = \sum_{Y \in \mathcal{M}_\varepsilon} |Y \cap E_\varepsilon| = |X \cap E_\varepsilon| + \sum_{Y \in \mathcal{M}_\varepsilon \setminus \{X\}} |Y \cap E_\varepsilon| = 1 + \sum_{Y \in \mathcal{M}_\varepsilon \setminus \{X\}} |Y \cap E_\varepsilon|.$$

Since each term of this summation is a power of two, then by Lemma 5.1.5, there are at least $d-1$ terms in the summation—i.e. $|\mathcal{M}_\varepsilon \setminus \{X\}| \geq d-1$ so $|\mathcal{M}_\varepsilon| \geq d$ showing that $|\mathcal{N}_{\overline{0}}(\vec{c})| \geq |\mathcal{M}_\varepsilon| \geq d$. □

The final result we will need before proving that $\varepsilon = \frac{1}{2d}$ is optimal is the following. It says that we can underestimate the size of $\mathcal{N}_{\overline{0}}(\vec{p})$—the sets whose closures intersect $\vec{p}$—with the sets whose closures intersect $\vec{p} - \varepsilon \cdot \vec{e}^{(i_0)}$ or $\vec{p} + \varepsilon \cdot \vec{e}^{(i_0)}$ where i_0 is some coordinate in $[d]$ and $\varepsilon > 0$ is small enough. In particular, the following result will hold for unit cube packings by Corollary 3.6.6.

Fact 9.8.3 (Locally Finite: Two Points)**.** *Let $d \in \mathbb{N}$ and $\mathcal{F}$ a locally finite family of subsets of $\mathbb{R}^d$ and $\vec{p} \in \mathbb{R}^d$ and $i_0 \in [d]$. Then for all sufficiently small $\varepsilon > 0$ we have $\mathcal{N}_{\overline{0}}(\vec{p}) \supseteq \mathcal{N}_{\overline{0}}(\vec{p} - \varepsilon \cdot \vec{e}^{(i_0)}) \cup \mathcal{N}_{\overline{0}}(\vec{p} + \varepsilon \cdot \vec{e}^{(i_0)})$.*

Proof. Observe that for any set $X \subseteq \mathbb{R}^d$ and any $\varepsilon \in (0, \infty)$, we have the following implications[15] (note the closure in the first two expressions and not the third):

$$\overline{X} \ni \vec{p} \quad \Longrightarrow \quad \overline{X} \cap \left(\vec{p} + [-\varepsilon, \varepsilon]^d\right) \neq \emptyset \quad \Longrightarrow \quad X \cap \left(\vec{p} + [-2\varepsilon, 2\varepsilon]^d\right) \neq \emptyset. \tag{9.7}$$

[15]The first implication is trivial, and the second implication is justified as follows. If $\overline{X} \cap \left(\vec{p} + [-\varepsilon, \varepsilon]^d\right) \neq \emptyset$, then let $\vec{c} \in \overline{X} \cap \left(\vec{p} + [-\varepsilon, \varepsilon]^d\right)$. Then there a sequence $\left\{\vec{x}^{(n)}\right\}_{n=1}^\infty$ of points in X converging to $\vec{c}$ and thus there is some $N \in \mathbb{N}$ such that for $n \geq N$, $\left\|\vec{x}^{(n)} - \vec{c}\right\|_\infty \leq \varepsilon$, and since $\vec{c} \in \left(\vec{p} + [-\varepsilon, \varepsilon]^d\right) = {}^\infty\overline{B}_\varepsilon(\vec{p})$ we have $\left\|\vec{x}^{(N)} - \vec{p}\right\| \leq \left\|\vec{x}^{(N)} - \vec{c}\right\| + \|\vec{c} - \vec{p}\| \leq \varepsilon + \varepsilon = 2\varepsilon$ and thus $\vec{x}^{(N)} \in X \cap {}^\infty\overline{B}_{2\varepsilon}(\vec{p}) = X \cap \left(\vec{p} + [-2\varepsilon, 2\varepsilon]^d\right)$, so $X \cap \left(\vec{p} + [-2\varepsilon, 2\varepsilon]^d\right) \neq \emptyset$.

We also have the following implication by a trivial subset containment:

$$\overline{X} \cap \left\{\vec{p} \pm \varepsilon \cdot \vec{e}^{(i_o)}\right\} \neq \emptyset \qquad \Longrightarrow \qquad \overline{X} \cap \left(\vec{p} + [-\varepsilon, \varepsilon]^d\right) \neq \emptyset. \tag{9.8}$$

Then note that

$$\begin{aligned}
\mathcal{N}_{\vec{0}}(\vec{p}) &= \left\{X \in \mathcal{F} : \overline{X} \ni \vec{p}\right\} && \text{(Definition)} \\
&\subseteq \left\{X \in \mathcal{F} : \overline{X} \cap \left(\vec{p} + [-\varepsilon, \varepsilon]^d\right) \neq \emptyset\right\} && \text{(Equation 9.7)} \\
&\subseteq \left\{X \in \mathcal{F} : X \cap \left(\vec{p} + [-2\varepsilon, 2\varepsilon]^d\right) \neq \emptyset\right\} && \text{(Equation 9.7)} \\
&= {}^{\infty}\overline{\mathcal{N}}_{2\varepsilon}(\vec{p}) && \text{(Definition)}
\end{aligned}$$

and by Fact 3.6.5, for all sufficiently small $\varepsilon > 0$ we have that $\mathcal{N}_{\vec{0}}(\vec{p}) = {}^{\infty}\overline{\mathcal{N}}_{\varepsilon}(\vec{p})$, so all of the expressions above are equal.

Thus, for all sufficiently small $\varepsilon > 0$, we have

$$\begin{aligned}
\mathcal{N}_{\vec{0}}(\vec{p}) &= \left\{X \in \mathcal{F} : \overline{X} \cap \left(\vec{p} + [-\varepsilon, \varepsilon]^d\right) \neq \emptyset\right\} && \text{(Equalities above)} \\
&\supseteq \left\{X \in \mathcal{F} : \overline{X} \cap \left\{\vec{p} \pm \varepsilon \cdot \vec{e}^{(i_o)}\right\} \neq \emptyset\right\} && \text{(Equation 9.8)} \\
&= \mathcal{N}_{\vec{0}}(\vec{p} - \varepsilon \cdot \vec{e}^{(i_0)}) \cup \mathcal{N}_{\vec{0}}(\vec{p} + \varepsilon \cdot \vec{e}^{(i_0)})
\end{aligned}$$

which completes the proof. □

Now we state and prove the main result of this section. The details of the proof overwhelm the main idea, so we will outline the proof before providing it in full detail.

Theorem 9.8.4 ($\varepsilon = \frac{1}{2d}$ Optimal for Unit Cube Enclosures)**.** *Let $d \in \mathbb{N}$ and $\varepsilon \in (0, \infty)$ and $(X, \mathcal{E})$ a $(d+1, \varepsilon)$-secluded axis-aligned unit cube enclosure in $\mathbb{R}^d$. Then $\varepsilon \leq \frac{1}{2d}$.*

Proof Outline. We will consider the enclosed cube X along with the "bottom" corner $\vec{p} = \langle \text{center}_i(X) - \frac{1}{2} \rangle_{i=1}^{d}$. We consider the unique twin/1st-cousin of X contained in $\mathcal{N}_{\overline{0}}(\vec{p})$ and let i_0 denote the cousin-coordinate of them. We show that no two of the cubes in $\mathcal{N}_{\overline{0}}(\vec{p})$ have the same center position along direction/coordinate i_0. We consider the edge of X which starts at the corner $\vec{p}$ and finishes as the corner $\vec{q} = \vec{p} + \vec{e}^{(i_0)}$. We imagine starting at $\vec{p}$ and walking along this edge until we get to $\vec{q}$ and marking whenever we "leave" one of the cubes in $\mathcal{N}_{\overline{0}}(\vec{p})$—that is, whenever the i_0th coordinate of our position matches the supremum of $\pi_{i_0}(W)$ for some cube $W \in \mathcal{N}_{\overline{0}}(\vec{p})$. Because no cubes have the same position in the i_0th coordinate, we end up marking $d+1$ points along this edge of X—one for each of the $d+1$ cubes in $\mathcal{N}_{\overline{0}}(\vec{p})$—and we argue that for any two of these points $\vec{a}$ and $\vec{b}$, the neighborhoods $\mathcal{N}_{\overline{0}}(\vec{a})$ and $\mathcal{N}_{\overline{0}}(\vec{b})$ are distinct and each have cardinality $d+1$. Since there are $(d+1)$-many marked points along a unit length edge, there must be two points $\vec{a}$ and $\vec{b}$ which are distance at most $\frac{1}{d}$ apart (because $d+1$ marks in a unit interval segment the interval into at least d pieces). If we assume for contradiction that ε is larger than $\frac{1}{2d}$, we can consider the midpoint $\vec{c}$ of $\vec{a}$ and $\vec{b}$ and note that ${}^{\infty}\overline{B}_{\varepsilon}(\vec{c})$ contains both $\vec{a}$ and $\vec{b}$ in its interior which will imply that ${}^{\infty}\overline{\mathcal{N}}_{\varepsilon}(\vec{c}) \supseteq \mathcal{N}_{\overline{0}}(\vec{a}) \cup \mathcal{N}_{\overline{0}}(\vec{b})$. Since $\mathcal{N}_{\overline{0}}(\vec{a})$ and $\mathcal{N}_{\overline{0}}(\vec{b})$ are distinct and each have cardinality $d+1$, their union has cardinality at least $d+2$ which implies that $\left|{}^{\infty}\overline{\mathcal{N}}_{\varepsilon}(\vec{c})\right| \geq d+2$. This contradicts the hypothesis that $(X, \mathcal{E})$ is $(d+1, \varepsilon)$-secluded. Thus, we conclude that $\varepsilon \leq \frac{1}{2d}$. □

Proof. Let $\vec{x} = \text{center}(X)$ and assume for convenience and without loss of generality that $\vec{x} = \vec{0}$ and consider $\vec{p} = \langle -\frac{1}{2} \rangle_{i=1}^{d} \in \text{corners}(X)$. (The argument could be run with any corner, but it will be slightly more convenient to know that $p_i < x_i$ for all $i \in [d]$.) Since $(X, \mathcal{E})$ is $(d+1, \varepsilon)$-secluded, it is trivially $(d+1)$-secluded, so by Theorem 9.5.9 (Cube Enclosures: $(d+1)$-Secluded Implies Minimum) and Theorem 9.5.8 (Minimum

Cube Enclosure Equivalencies), $\mathcal{N}_{\vec{0}}(\vec{p})$ contains exactly X and one nth-cousin of X for each $n \in [d]$. Let $Y \in \mathcal{N}_{\vec{0}}(\vec{p})$ be the 1st-cousin/twin of X, and let $\vec{y} = \text{center}(Y)$, and let i_0 be the cousin-coordinate of X and Y (i.e. $|x_{i_0} - y_{i_0}| = 1$). Let $\vec{q} = \vec{p} + \vec{e}^{(i_0)}$ so that $\vec{q} \in \text{corners}(X)$, and $\vec{p}, \vec{q}$ defines an edge of X.

The first step is to show that no two cubes in $\mathcal{N}_{\vec{0}}(\vec{p})$ have centers with the same i_0 coordinate value. Observe that $y_{i_0} = -1$ because for every $Z \in \mathcal{N}_{\vec{0}}(\vec{p})$, $\vec{p} \in \overline{Z}$ by definition, so $\text{center}_{i_0}(Z) \in [p_{i_0} - \frac{1}{2}, p_{i_0} + \frac{1}{2}] = [-1, 0]$. Thus, because $|x_{i_0} - y_{i_0}| = 1$, and $x_{i_0} = 0$, we have $y_{i_0} = -1$.

Let $\mathcal{M} = \mathcal{N}_{\vec{0}}(\vec{p}) \setminus \{X, Y\}$ denote the remaining $d-1$ cubes. For each $Z \in \mathcal{M}$, we claim that not only is $\text{center}_{i_0}(Z) \in [0, 1]$ (as shown above), but in fact $\text{center}_{i_0}(Z) \in (0, 1)$. To see this fix some $Z \in \mathcal{F}$ and let $\vec{z} = \text{center}(Z)$. Because $Z \overset{\text{cous}}{\sim} X$, let $i_1 \in [d]$ be the cousin-coordinate of X and Z (i.e. $|z_{i_1} - x_{i_1}| = 1$) and since $z_{i_1} \in [-1, 0]$ and $x_{i_1} = 0$ we have $z_{i_1} = -1$.

For each $\delta \in (0, \infty)$, let $\vec{p}_\delta^{(-)} = \vec{p} - \delta \cdot \vec{e}^{(i_0)}$ and $\vec{p}_\delta^{(+)} = \vec{p} + \delta \cdot \vec{e}^{(i_0)}$. Note that for $\delta \in (0, 1)$, $\vec{p}_\delta^{(-)}$ belongs to the interior of an edge of Y along coordinate i_0 and $\vec{p}_\delta^{(+)}$ belongs to the interior of an edge of X along coordinate i_0 (see justification[16]).

Note that $\vec{p}_\delta^{(-)}$ and $\vec{p}_\delta^{(+)}$ are ℓ_∞ distance δ from $\vec{p}$, and since $\vec{p} \in \overline{X} = {}^\infty\overline{B}_{1/2}(\text{center}(X))$, by the triangle inequality we have that $\vec{p}_\delta^{(-)}$ and $\vec{p}_\delta^{(+)}$ are ℓ_∞ distance at most $\frac{1}{2} + \delta$ from $\text{center}(X)$. That is, $\vec{p}_\delta^{(-)}, \vec{p}_\delta^{(+)} \in {}^\infty\overline{B}_{\frac{1}{2}+\delta}(\text{center}(X))$.

By Lemma 9.2.10 (Enclosure Enlargment), for all sufficiently small $\delta > 0$, we have that ${}^\infty\overline{B}_{\frac{1}{2}+\delta}(\text{center}(X)) \subseteq \text{int}\left(X \cup \bigcup_{Z \in \mathcal{E}} Z\right)$ implying that for such δ, we have $\vec{p}_\delta^{(-)}, \vec{p}_\delta^{(+)} \in \text{int}\left(X \cup \bigcup_{Z \in \mathcal{E}} Z\right)$, so by Lemma 9.8.2 (recalling that $\vec{p}_\delta^{(-)}$ belongs to the interior of an edge of Y and $\vec{p}_\delta^{(+)}$ belongs to the interior of an edge of X) we have

[16] Regarding X, the point $\vec{p}_\delta^{(+)}$ is strictly between the corners $\vec{p}$ and $\vec{q}$ of X and $\vec{q} - \vec{p} = \vec{e}^{(i_0)}$ so these corners define an edge of X. Regarding Y, let $\vec{r} = \vec{p} - \vec{e}^{(i_0)}$ and note that $\vec{r} \in \text{corners}(Y)$ because $|r_{i_0} - y_{i_0}| = |(p_{i_0} - 1) - (-1)| = |p_{i_0}| = \frac{1}{2}$ and for $i \in [d] \setminus \{i_0\}$, $|r_i - y_i| = |p_i - x_i| = \frac{1}{2}$ showing that $\vec{r} \in \vec{y} + \left\{-\frac{1}{2}, \frac{1}{2}\right\}^d = \text{corners}(Y)$. Then the point $\vec{p}_\delta^{(-)}$ is strictly between the corners $\vec{p}$ and $\vec{r}$ of Y and $\vec{p} - \vec{r} = \vec{e}^{(i_0)}$ so these corners define an edge of Y.

$\left|\mathcal{N}_{\bar{0}}(\vec{p}_\delta^{(-)})\right| \geq d$ and $\left|\mathcal{N}_{\bar{0}}(\vec{p}_\delta^{(+)})\right| \geq d$. Also, for all sufficiently small $\delta > 0$, we have by Fact 9.8.3 that $\mathcal{N}_{\bar{0}}(\vec{p}) \supseteq \mathcal{N}_{\bar{0}}(\vec{p}_\delta^{(-)}) \cup \mathcal{N}_{\bar{0}}(\vec{p}_\delta^{(+)})$.

Now suppose that there is some $Z \in \mathcal{M} \subseteq \mathcal{N}_{\bar{0}}(\vec{p})$ with $\text{center}_{i_0}(Z) = -1$. Then $Z \in \mathcal{N}_{\bar{0}}(\vec{p})$, but $Z \notin \mathcal{N}_{\bar{0}}(\vec{p}_\delta^{(+)})$ (see justification[17]). By the same justification, $Y \in \mathcal{N}_{\bar{0}}(\vec{p})$ (by choice of Y), but $Y \notin \mathcal{N}_{\bar{0}}(\vec{p}_\delta^{(+)})$. Thus, $\mathcal{N}_{\bar{0}}(\vec{p}) \supseteq \{Y, Z\} \sqcup \mathcal{N}_{\bar{0}}(\vec{p}_\delta^{(+)})$ and so $|\mathcal{N}_{\bar{0}}(\vec{p})| \geq 2 + \left|\mathcal{N}_{\bar{0}}(\vec{p}_\delta^{(+)})\right| \geq d + 2$. But this would contradict that $(X, \mathcal{E})$ is $(d+1)$-secluded, so it must be that there is no $Z \in \mathcal{M}$ with $\text{center}_{i_0}(Z) = -1$.

Similarly, there can be no $Z \in \mathcal{M}$ with $\text{center}_{i_0}(Z) = 0$ because if there was, then we would have $X, Z \in \mathcal{N}_{\bar{0}}(\vec{p})$ but $X, Z \notin \mathcal{N}_{\bar{0}}(\vec{p}_\delta^{(-)})$, so $|\mathcal{N}_{\bar{0}}(\vec{p})| \geq 2 + \left|\mathcal{N}_{\bar{0}}(\vec{p}_\delta^{(1)})\right| \geq d+2$.

At this point, we know that $\text{center}_{i_0}(X) = 0$ and $\text{center}_{i_0}(Y) = -1$ and for $Z \in \mathcal{M} = \mathcal{N}_{\bar{0}}(\vec{p}) \setminus \{X, Y\}$ that $\text{center}_{i_0}(Z) \in (-1, 0)$. The next thing we want to show is that no two cubes in $\mathcal{M}$ have the same i_0th center coordinate, and we use the same technique as above. Assume that there are distinct $Z, Z' \in \mathcal{M}$ such that $\text{center}_{i_0}(Z) = \text{center}_{i_0}(Z')$. Let $\vec{z} = \text{center}_{i_0}(Z)$ and $\vec{z}' = \text{center}_{i_0}(Z')$, so $z_{i_0} = z'_{i_0} \in (-1, 0)$. Let

$$\vec{c} = \left\langle \begin{cases} p_i & i \neq i_0 \\ z_{i_0} + \frac{1}{2} & i = i + 0 \end{cases} \right\rangle_{i=1}^{d}$$
$$= \left\langle \begin{cases} -\frac{1}{2} & i \neq i_0 \\ z_{i_0} + \frac{1}{2} & i = i + 0 \end{cases} \right\rangle_{i=1}^{d}.$$

For each $\delta \in (0, \infty)$, let $\vec{c}_\delta^{(-)} = \vec{c} - \delta \cdot \vec{e}^{(i_0)}$ and $\vec{c}_\delta^{(+)} = \vec{c} + \delta \cdot \vec{e}^{(i_0)}$.

We claim that $\vec{c} \in \overline{Z}$ and $\vec{c} \in \overline{Z'}$ so that $Z, Z' \in \mathcal{N}_{\bar{0}}(\vec{c})$ (see justification[18]). We

[17] The i_0th coordinate of $\vec{p}_\delta^{(+)}$ is $p_{i_0} + \delta = -\frac{1}{2} + \delta$. This is more than distance $\frac{1}{2}$ from the i_0th coordinate of center(Z) which is -1. Thus $\vec{p}_\delta^{(+)}$ is not in the closure of $Z = {}^{\infty}\overline{B}_{1/2}(\text{center } Z)$.

[18] Because we have $\vec{p} \in \overline{Z'}$, we have for $i \in [d]$ that $|p_i - z'_i| \leq \frac{1}{2}$, so for $i \in [d] \setminus \{i_0\}$ we have

also claim that for all sufficiently small $\delta > 0$, both $\vec{c}_\delta^{(-)}$ and $\vec{c}_\delta^{(+)}$ belong to the interior of an edge of X (see justification[19]), so in particular they belong to $\overline{X}$. We also claim that for all $\delta \in (0, \infty)$, $\vec{c}_\delta^{(+)} \notin \overline{Z}$ and $\vec{c}_\delta^{(+)} \notin \overline{Z'}$ (see justification[20]) so that $Z \notin \mathcal{N}_{\overline{0}}(\vec{c}_\delta^{(+)})$ and $Z' \notin \mathcal{N}_{\overline{0}}(\vec{c}_\delta^{(+)})$.

By definition of a cube enclosure and the claim in the prior paragraph, we have for all sufficiently small $\delta > 0$ that $\vec{c}_\delta^{(+)} \in \overline{X} \subseteq \operatorname{int}\left(X \cup \bigcup_{W \in \mathcal{E}} W\right)$, so by Lemma 9.8.2 (recalling from the prior paragraph that $\vec{c}_\delta^{(+)}$ belong to the interior of an edge of X) we have $\left|\mathcal{N}_{\overline{0}}(\vec{c}_\delta^{(+)})\right| \geq d$. Also, for all sufficiently small $\delta > 0$, we have by Fact 9.8.3 that $\mathcal{N}_{\overline{0}}(\vec{c}) \supseteq \mathcal{N}_{\overline{0}}(\vec{c}_\delta^{(+)})$. By the prior paragraph, $Z, Z' \in \mathcal{N}_{\overline{0}}(\vec{c})$ but $Z \notin \mathcal{N}_{\overline{0}}(\vec{c}_\delta^{(+)})$ and $Z' \notin \mathcal{N}_{\overline{0}}(\vec{c}_\delta^{(+)})$ so we have $\mathcal{N}_{\overline{0}}(\vec{c}) \supseteq \{Z, Z'\} \sqcup \mathcal{N}_{\overline{0}}(\vec{c}_\delta^{(+)})$ and thus $|\mathcal{N}_{\overline{0}}(\vec{c})| \geq 2 + \left|\mathcal{N}_{\overline{0}}(\vec{c}_\delta^{(+)})\right| \geq d + 2$. This would contradict that $(X, \mathcal{E})$ is $(d+1)$-secluded, so it must be that there is no pair of distinct $Z, Z' \in \mathcal{M}$ with $\operatorname{center}_{i_0}(Z) = \operatorname{center}_{i_0}(Z')$.

At this point, we have established that for any two cubes in $\mathcal{N}_{\overline{0}}(\vec{p})$, the i_0th coordinates of the centers are distinct. For each $W \in \mathcal{N}_{\overline{0}}(\vec{p})$, let $\vec{s}^{(W)}$ denote the point

$$\vec{s}^{(W)} = \left\langle \begin{cases} p_i & i \neq i_0 \\ \operatorname{center}_{i_0}(W) + \frac{1}{2} & i = i_0 \end{cases} \right\rangle_{i=1}^{d}.$$

This should be viewed as a point along the edge of X from $\vec{p}$ to $\vec{q} = \vec{p} + \vec{e}^{(i_0)}$ which matches up with a face of W in the i_0th direction. Note that $\vec{s}^{(W)} \in \overline{W}$ (see justification[21]).

$p_i = c_i$, so $|c_i - z'_i| \leq \frac{1}{2}$. And for i_0 we have $z_{i_0} = z'_{i_0}$ so by definition of $\vec{c}$, $\left|c_{i_0} - z_{i'_0}\right| = \frac{1}{2}$. Thus $\vec{c} \in \vec{z}' + [-\frac{1}{2}, \frac{1}{2}]^d = \overline{Z'}$. The same argument shows $\vec{c} \in \overline{Z}$.

[19] The point $\vec{c}$ is strictly between the corners $\vec{p}$ and $\vec{q}$ of X (which we have already shown define an edge of X) because for $i \in [d] \setminus \{i_0\}$, $c_i = p_i = q_i$ and for i_0 we have $c_{i_0} = z_{i_0} + \frac{1}{2} \in (-1, 0) + \frac{1}{2} = (-\frac{1}{2}, \frac{1}{2}) = (p_{i_0}, q_{i_0})$. Thus, for sufficiently small δ, both $\vec{c}_\delta^{(-)}$ and $\vec{c}_\delta^{(+)}$ are strictly between $\vec{p}$ and $\vec{q}$.

[20] The i_0th coordinate of $\vec{c}_\delta^{(+)}$ is $c_{i_0} + \delta = z_{i_0} + \frac{1}{2} + \delta = z'_{i_0} + \frac{1}{2} + \delta$ which differs from z_{i_0} (resp. z'_{i_0}) by more than $\frac{1}{2}$, so $\vec{c}_\delta^{(+)} \notin \vec{z} + [-\frac{1}{2}, \frac{1}{2}]^d = \overline{Z}$ (resp. $\vec{c}_\delta^{(+)} \notin \vec{z}' + [-\frac{1}{2}, \frac{1}{2}]^d = \overline{Z'}$).

[21] This is because $W \in \mathcal{N}_{\overline{0}}(\vec{p})$, so $\vec{p} \in \overline{W} = \operatorname{center}(W) + [-\frac{1}{2}, \frac{1}{2}]^d$ by definition of

We claim that for each $W \in \mathcal{N}_{\bar{0}}(\vec{p})$, we have $\left|\mathcal{N}_{\bar{0}}(\vec{s}^{(W)})\right| = d+1$ which we prove in three cases. Case (1) is that $W = X$ in which case $\text{center}_{i_0}(W)+\frac{1}{2} = \text{center}_{i_0}(X)+\frac{1}{2} = \frac{1}{2}$, so that

$$\begin{aligned}\vec{s}^{(W)} &= \left\langle \begin{cases} p_i & i \neq i_0 \\ \frac{1}{2} & i = i_0 \end{cases} \right\rangle_{i=1}^{d} \\ &= \left\langle \begin{cases} q_i & i \neq i_0 \\ q_{i_0} & i = i_0 \end{cases} \right\rangle_{i=1}^{d} \\ &= \vec{q}.\end{aligned}$$

Thus, $\mathcal{N}_{\bar{0}}(\vec{s}^{(W)}) = \mathcal{N}_{\bar{0}}(\vec{q})$ and since $\vec{q} \in \text{corners}(X)$, we know by Theorem 9.5.9 (Cube Enclosures: $(d+1)$-Secluded Implies Minimum) and Theorem 9.5.8 (Minimum Cube Enclosure Equivalencies) that this neighborhood has cardinality $d+1$. Case (2) is that $W = Y$ in which case $\text{center}_{i_0}(W) + \frac{1}{2} = \text{center}_{i_0}(Y) + \frac{1}{2} = -\frac{1}{2}$, so that

$$\begin{aligned}\vec{s}^{(W)} &= \left\langle \begin{cases} p_i & i \neq i_0 \\ -\frac{1}{2} & i = i_0 \end{cases} \right\rangle_{i=1}^{d} \\ &= \left\langle \begin{cases} p_i & i \neq i_0 \\ p_{i_0} & i = i_0 \end{cases} \right\rangle_{i=1}^{d} \\ &= \vec{p}.\end{aligned}$$

---

neighborhood which implies that for all $i \in [d]$, we have $|\text{center}_i(W) - p_i| \leq \frac{1}{2}$. Thus, for $i \in [d] \setminus \{i_0\}$, $\left|\text{center}_i(W) - s_i^{(W)}\right| \leq \frac{1}{2}$. And trivially, $\left|\text{center}_{i_0}(W) - s_{i_0}^{(W)}\right| = \frac{1}{2}$. Thus, $s_{i_0}^{(W)} \in {}^{\infty}\overline{B}_{1/2}(\text{center}(W)) = \overline{W}$.

Thus, $\mathcal{N}_{\bar{0}}(\vec{s}^{(W)}) = \mathcal{N}_{\bar{0}}(\vec{p})$ and since $\vec{p} \in \text{corners}(X)$, we know by Theorem 9.5.9 (Cube Enclosures: $(d+1)$-Secluded Implies Minimum) and Theorem 9.5.8 (Minimum Cube Enclosure Equivalencies) that this neighborhood has cardinality $d+1$. Case (3) is that $W \in \mathcal{N}_{\bar{0}}(\vec{p}) \setminus \{X, Y\} = \mathcal{M}$ so by prior work in the proof, $\text{center}_{i_0}(W) \in (-1, 0)$ implying $s_{i_0}^{(W)} = \text{center}_{i_0}(W) + \frac{1}{2} \in (-\frac{1}{2}, \frac{1}{2})$. Thus, for all sufficiently small $\delta > 0$ we have $s_{i_0}^{(W)} \pm \delta \in (-\frac{1}{2}, \frac{1}{2})$ and thus both $\vec{s}^{(W)} - \delta \cdot \vec{e}^{(i_0)}$ and $\vec{s}^{(W)} + \delta \cdot \vec{e}^{(i_0)}$ belong to the interior of the edge of X between $\vec{p}$ and $\vec{q}$. For sufficiently small $\delta > 0$ we have by Fact 9.8.3 that $\mathcal{N}_{\bar{0}}(\vec{s}^{(W)}) \supseteq \mathcal{N}_{\bar{0}}(\vec{s}^{(W)} + \delta \cdot \vec{e}^{(i_0)})$. Since $\vec{s}^{(W)} \in \overline{W}$, then by definition $W \in \mathcal{N}_{\bar{0}}(\vec{s}^{(W)})$. Also, $W \notin \mathcal{N}_{\bar{0}}(\vec{s}^{(W)} + \delta \cdot \vec{e}^{(i_0)})$ because $\text{center}_{i_0}(W) = \vec{s}^{(W)} - \frac{1}{2}$ is distance greater than $\frac{1}{2}$ from $\vec{s}^{(W)} + \delta$ (which is the i_0th coordinate of $\vec{s}^{(W)} + \delta \cdot \vec{e}^{(i_0)}$). Thus, $\mathcal{N}_{\bar{0}}(\vec{s}^{(W)}) \supseteq \{W\} \sqcup \mathcal{N}_{\bar{0}}(\vec{s}^{(W)} + \delta \cdot \vec{e}^{(i_0)})$. By Lemma 9.8.2, $\left|\mathcal{N}_{\bar{0}}(\vec{s}^{(W)})\right| \geq d+1$. And since $(X, \mathcal{E})$ is $(d+1)$-secluded, we get equality.

We also claim that for distinct $W, W' \in \mathcal{N}_{\bar{0}}(\vec{p})$ we have $\mathcal{N}_{\bar{0}}(\vec{s}^{(W)}) \neq \mathcal{N}_{\bar{0}}(\vec{s}^{(W')})$. This is because $\text{center}_{i_0}(W) \neq \text{center}_{i_0}(W')$ so we assume without loss of generality that $\text{center}_{i_0}(W) < \text{center}_{i_0}(W')$. Then, because $\text{center}_{i_0}(W') = s_{i_0}^{(W')} - \frac{1}{2}$, the distance between $\text{center}_{i_0}(W)$ and $s_{i_0}^{(W')}$ is greater than $\frac{1}{2}$, so $\vec{s}^{(W')} \notin W$ showing that $W \notin \mathcal{N}_{\bar{0}}(\vec{s}^{(W')})$. However, $W \in \mathcal{N}_{\bar{0}}(\vec{s}^{(W)})$ because as shown previously, $\vec{s}^{(W)} \in \overline{W}$. Thus, W witnesses that $\mathcal{N}_{\bar{0}}(\vec{s}^{(W)}) \neq \mathcal{N}_{\bar{0}}(\vec{s}^{(W')})$.

Finally, note that for distinct $W, W' \in \mathcal{N}_{\bar{0}}(\vec{p})$ we have $\vec{s}^{(W)} \neq \vec{s}^{(W')}$, and in particular $s_{i_0}^{(W)} \neq s_{i_0}^{(W')}$ because we have already shown that $\text{center}_{i_0}(W) \neq \text{center}_{i_0}(W')$. Thus, the set $S = \left\{\vec{s}^{(W)} : W \in \mathcal{N}_{\bar{0}}(\vec{p})\right\}$ consists of $|\mathcal{N}_{\bar{0}}(\vec{p})|$-many points with distinct i_0th coordinates, and we know from Theorem 9.5.9 (Cube Enclosures: $(d+1)$-Secluded Implies Minimum) and Theorem 9.5.8 (Minimum Cube Enclosure Equivalencies) that $|\mathcal{N}_{\bar{0}}(\vec{p})| = d+1$. Also, note that for each $W \in \mathcal{N}_{\bar{0}}(\vec{p})$, we have already established that $\text{center}_{i_0}(W) \in [-1, 0]$, so it follows that $s_{i_0}^{(W)} \in [-\frac{1}{2}, \frac{1}{2}]$. Thus, S is a set of

$(d+1)$-many points, and each has a distinct i_0th coordinate in $[-\frac{1}{2}, \frac{1}{2}]$. Thus, there has to be a pair of these points with i_0th coordinates having distance at most $\frac{1}{d}$ apart. Let $\vec{a}, \vec{b} \in S$ denote such a pair of points, and let $\vec{c} = \frac{1}{2}(\vec{a} + \vec{b})$ be the midpoint. Since $|a_{i_0} - b_{i_0}| \leq \frac{1}{d}$ and for all other $i \in [d] \setminus \{i_0\}$, $a_i = b_i$, we have $\left\|\vec{a} - \vec{b}\right\|_\infty \leq \frac{1}{d}$, and since $\vec{c}$ is the midpoint we have $\|\vec{a} - \vec{c}\|_\infty \leq \frac{1}{2d}$ and $\left\|\vec{b} - \vec{c}\right\|_\infty \leq \frac{1}{2d}$. Suppose for contradiction that $\varepsilon > \frac{1}{2d}$. Then for some sufficiently small $\delta > 0$ we have ${}^{\infty}\overline{B}_\delta(\vec{a}) \subseteq {}^{\infty}\overline{B}_\varepsilon(\vec{c})$ as well as ${}^{\infty}\overline{B}_\delta(\vec{b}) \subseteq {}^{\infty}\overline{B}_\varepsilon(\vec{c})$ which means that

$$ {}^{\infty}\mathcal{N}_\varepsilon(\vec{c}) \supseteq {}^{\infty}\mathcal{N}_\delta(\vec{a}) \cup {}^{\infty}\mathcal{N}_\delta(\vec{b}) \supseteq \mathcal{N}_{\vec{0}}(\vec{a}) \cup \mathcal{N}_{\vec{0}}(\vec{b}). $$

By the prior two paragraphs, $\mathcal{N}_{\vec{0}}(\vec{a})$ and $\mathcal{N}_{\vec{0}}(\vec{b})$ are different sets of cardinality $d+1$, so their union must have cardinality at least $d+2$ implying that $\left|{}^{\infty}\mathcal{N}_\varepsilon(\vec{c})\right| \geq d+2$ which would contradict that $(X, \mathcal{E})$ is $(d+1, \varepsilon)$-secluded. Thus $\varepsilon \leq \frac{1}{2d}$. □

Corollary 9.8.5 (Optimality of $\varepsilon = \frac{1}{2d}$ for Unit Cube Partitions)**.** *Let $d \in \mathbb{N}$ and $\varepsilon \in (0, \infty)$ and $\mathcal{P}$ a $(d+1, \varepsilon)$-secluded axis-aligned unit cube partition of $\mathbb{R}^d$. Then $\varepsilon \leq \frac{1}{2d}$.*

Proof. Let $X \in \mathcal{P}$. By Fact 9.2.6, $(X, \mathcal{N}(X))$ is an axis-aligned unit cube enclosure, which is trivially $(d+1, \varepsilon)$-secluded because $\mathcal{P}$ is. By Theorem 9.8.4, $\varepsilon \leq \frac{1}{2d}$. □

Remark 9.8.6 ($\varepsilon = \frac{1}{2d}$ Optimal for Unit Cube Tilings). If one is interested in tilings by closed cubes instead of partitions, then the result still holds in spirit with minor technical differences. By close analysis of the proof of Theorem 9.8.4, if $\varepsilon = \frac{1}{2d}$, then there is some point $\vec{p} \in \mathbb{R}^d$ which that ${}^{\infty}\overline{B}_\varepsilon(\vec{p})$ intersects at least $d+2$ closed cubes[22]

[22]Using the points $\vec{a}$ and $\vec{b}$ from the proof and $\vec{c}$ their midpoint, ${}^{\infty}\overline{B}_{\frac{1}{2d}}(\vec{p})$ contains both $\vec{a}$ and $\vec{b}$, so it trivially intersects all closed cubes in $\mathcal{N}_{\vec{0}}(\vec{a})$ and $\mathcal{N}_{\vec{0}}(\vec{b})$ (because the cubes are closed so they contain $\vec{a}$ and $\vec{b}$ respectively).

so if a closed ball is used, then ε must be taken to be strictly less than $\frac{1}{2d}$. However, this suggests that for tilings of $\mathbb{R}^d$ by closed cubes it would make more sense to define the notions of (k, ε)-secluded and k-secluded using open ℓ_∞ balls rather than closed balls. $\triangle$

Chapter 10

Computer Science Applications

In this final chapter, we connect many of the mathematical results back to the original motivating computational context and also present a few results which don't directly relate to any of the mathematical research we have done (though they connect with our work in the sense that they use the Lebesgue covering theorem as the main ingredient). In Section 10.1 (Our Neighborhood Variants Lebesgue, KKM, and Sperner) we discuss how our variants of the Lebesgue covering theorem, the KKM lemma, and Sperner's lemma should prove useful in computer science. In Section 10.2 (Universal Deterministic Rounding Functions) we connect our main mathematical results throughout this dissertation back to the motivating question of constructing and proving limitations of deterministic rounding schemes which provide replicability/consistency in a universal black box manner. Lastly, in Section 10.3 (Limitations on Learning) we give two impossibility results in learning theory.

10.1 Our Neighborhood Variants Lebesgue, KKM, and Sperner

We want to begin by noting that we believe that one of our main contributions to the field of theoretical computer science is our variant of Sperner's lemma, the KKM lemma, and the Lebesgue covering theorem (Theorem 8.0.7 and Theorem 8.0.8). Sperner's lemma has found many applications in computer science, and we list a few as examples.

Sperner's lemma is also known to have applications in the context of fair division and economics, and so it finds applications in computational game theory [SH22, BCF+22, MS19, MZ19]. Polytopal variants of Sperner's lemma have also been used in the context of distributed and parallel computing to prove various impossibility results [AER21, Nis22, BRS11]. It has also found applications in communication complexity [GSP22] and computational complexity because determining the location of a full-colored simplex in a colored triangulation (which is guaranteed to exist by Sperner's lemma) is known to relate to the hardness of certain complexity classes [Gol15, Kin09], and we wonder if locating a point guaranteed by our variant could shed light on other complexity classes.

We utilized a cubical variant of Sperner's lemma to prove an impossibility result in a learning context in a prior paper [DPWV22], and we use the Lebesgue covering theorem later in this chapter to prove other learning impossibility results (Theorem 10.3.4 (No $(d+1)$-Pseudoterministic Algorithm for the d-Coin Bias Estimation Problem) and Theorem 10.3.10 (No $(d+1)$-Pseudoterministic Algorithm for the d-Threshold Estimation Problem)). A result called the Poincare-Miranda theorem—which is known to be equivalent to Sperner's lemma—was recently used in [CMY23] to prove other impossibility results about learning certain hypotheses classes. Finally, we use our variant to prove impossibility results of deterministic rounding functions when the functions are not $f : \mathbb{R}^d \to \mathbb{R}^d$ but rather $f : [0,1]^d \to [0,1]^d$ (Proposition 10.2.15 (Deterministic Rounding Function Limitations for the Cube)).

10.2 Universal Deterministic Rounding Functions

As discussed in Chapter 1 (Introduction), the original motivation for all of the mathematical work that we have done was a computer science motivation. For each dimension $d \in \mathbb{N}$, we wanted to design a "universal" deterministic rounding algorithm (i.e. a function) $f : \mathbb{R}^d \to \mathbb{R}^d$ that would not round values too much and that would round nearby values to a very small set. In this way, we can use this function as a black box on top of any randomized or deterministic approximation algorithm with codomain $\mathbb{R}^d$ and increase the amount of reproducibility of the approximation algorithm. In this section, we will explicitly lay out how our secluded partitions from Chapter 4 (Constructions) give rise to such rounding functions and also use our optimality and near optimality results of Chapter 6 (Optimality of k in General) and Chapter 7 (Near Optimality of ε in General) to prove impossibility results showing that one cannot hope for deterministic rounding functions of this type with parameters much better than ours.

Remark 10.2.1. While our impossibility results are stated for functions $f : \mathbb{R}^d \to \mathbb{R}^d$ and $f : [0,1]^d \to [0,1]^d$, the impossibility is not inherently one requiring the completeness or uncountability or non-computability of $\mathbb{R}^d$ or $[0,1]^d$, and one can adjust our results for countable or finite domains that are "dense enough" in $\mathbb{R}^d$ or $[0,1]^d$ as was done for example in obtaining the discrete result Theorem 8.0.8 (Neighborhood Sperner's Lemma) from the continuous result Theorem 8.0.7

(Neighborhood KKM/Lebesgue Theorem). △

10.2.1 Efficient Computation of Representatives

The first thing we want to establish is that all of the partitions we have constructed are "efficiently computable." We will use our partitions as the basis of rounding functions, and in order for them to computationally feasible, we have to be able to efficiently identify which member of the partition any given value $\vec{x} \in \mathbb{R}^d$ belongs to—that is, we can compute the representative corner (or the center point) of the unique cube in the partition which contains $\vec{x}$ (denoted as $\mathrm{member}(\vec{x})$). We begin by showing that this can be done efficiently for any reclusive partition (see Definition 4.2.1 (Reclusive Matrix) and Definition 4.2.4 (Reclusive Partition)). To prove this, we first need a basic fact.

Fact 10.2.2. *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 $a - b = \lfloor a \rfloor$. □

Proposition 10.2.3 (Efficient Computation of Reclusive 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 cube such that $\vec{x} \in X$. Then* $\mathrm{rep}(X)$ *can be efficiently*[a] *computed in terms of $\vec{x}$ and A.*

[a] By efficient, we mean that the computation can be done with $\mathsf{poly}(d)$ addition and multiplication operations.

Proof. Recall that by definition of $\mathcal{P}_A$, each cube $Y \in \mathcal{P}_A$ is of the form $Y = A\vec{n} + [0, 1)^d$ for some $\vec{n} \in \mathbb{Z}^d$. Thus, in particular $\mathrm{rep}(X) = A\vec{m}$ for a unique $\vec{m} \in \mathbb{Z}^d$, so it will suffice to compute $\vec{m}$ because rep(X) can then be computed via a single matrix multiplication of the $d \times d$ matrix A with the $d \times 1$ vector $\vec{m}$. 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 (i.e. we have vacuously computed m_i for all $i > d$ because there are no such indices i), so we proceed to the inductive case of computing m_k assuming we have computed m_i for all $i > k$.

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

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

Note that the summation above might be an empty summation which by convention is 0. We now reformulate in terms of m_k.

$$\begin{aligned}
m_k &= \left(x_k - \sum_{i=k+1}^{d} a_{ki} m_i \right) - \alpha_k && \text{(Solve for m_k)} \\
&= \left\lfloor x_k - \sum_{i=k+1}^{d} a_{ki} m_i \right\rfloor && \text{($\alpha_k \in [0,1)$, $m_k \in \mathbb{Z}$, and Fact 10.2.2)}
\end{aligned}$$

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 the 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 $A\vec{m}$. □

Next, we observe that because the representative corner of the unique member of a point can be efficiently computed for the reclusive partitions, it can also be efficiently computed for the partition product constructions; recall the definition of this construction restated below.

Definition 4.4.4 (Partition Product). Let $d_1, \ldots, d_n \in \mathbb{N}$ and $\mathcal{P}_1, \ldots, \mathcal{P}_n$ be partitions of $\mathbb{R}^{d_1}, \ldots, \mathbb{R}^{d_n}$ respectively. Letting $d = \sum_{i=1}^{n} d_i$ we define the product partition of $\mathbb{R}^d$ as

$$\prod_{i=1}^{n} \mathcal{P}_i \stackrel{\text{def}}{=} \left\{ \prod_{i=1}^{n} X^{(i)} : X^{(i)} \in \mathcal{P}_i \right\}$$

where $\prod_{i=1}^{n} X^{(i)}$ is viewed as a subset of $\mathbb{R}^d$.

Proposition 10.2.4 (Efficient Computation of Partition Product Representatives)**.** *Let $d_1, \ldots, d_n \in \mathbb{N}$ and $\mathcal{P}_1, \ldots, \mathcal{P}_n$ be reclusive partitions of $\mathbb{R}^{d_1}, \ldots, \mathbb{R}^{d_n}$ respectively with associated matrices $A_1, \ldots, A_n$ (where A_i is a $d_i \times d_i$ matrix), and let $d = \sum_{i=1}^{n} d_i$. Let $\mathcal{P} = \prod_{i=1}^{n} \mathcal{P}_i$ be the partition product which is a partition of $\mathbb{R}^d$. For any $\vec{x} \in \mathbb{R}^d$, let $X \in \mathcal{P}$ be the unique cube such that $\vec{x} \in X$. Then $\text{rep}(X)$ can be efficiently[a] computed in terms of $\vec{x}$ and $A_1, \ldots, A_n$.*

[a] By efficient, we mean that the computation can be done with $\mathsf{poly}(d)$ addition and multiplication operations.

Proof Sketch. Given $\vec{x} \in \mathbb{R}^d$, the member X that it is contained in can be found by determining which member of $\mathcal{P}_1$ the point $\langle x_i \rangle_{i=1}^{d_1}$ is in, and independently determining which member of $\mathcal{P}_2$ the point $\langle x_i \rangle_{i=d_1+1}^{d_1+d_2}$ is in, etc. The member of $\prod_{i=1}^{n} \mathcal{P}_i$ that contains $\vec{x}$ is just the product of members. By Proposition 10.2.3, this takes $O(\sum_{i=1}^{n} d_i^2) \subseteq O(d^2)$ additions and multiplications. $\square$

Remark 10.2.5 (Efficient Computation of Center Point). Typically, we would prefer to compute the center of a cube rather than the representative corner, because that is a better approximation for points in the cube in the worst case (i.e. the center is ℓ_∞ distance at most $\frac{1}{2}$ from all points in the cube, but the representative corner is ℓ_∞ distance arbitrarily close to 1 from some points in the cube). This is easy to do for the reclusive partitions or the partition products as in Proposition 10.2.3 or Proposition 10.2.4 because for any cube X, $\text{center}(X) = \text{rep}(X) + \frac{1}{2} \cdot \vec{1}$, so $\frac{1}{2}$ just has to be added to each coordinate of $\text{rep}(X)$. $\triangle$

10.2.2 Constructions

The significance of our partitions to rounding and reproducibility is that we can use our partitions to construct universal black box rounding functions. Specifically, a single function for each dimension which can accept any ε_0 approximation and be guaranteed to output an ε approximation, and is guaranteed to output one of at most $d+1$ distinct approximations.

Definition 10.2.6 (Unit Cube Partition Rounding Function). Let $d \in \mathbb{N}$ and $\mathcal{P}$ an axis-aligned unit cube partition of $\mathbb{R}^d$. The *partition rounding function of* $\mathcal{P}$ is the function

$$f_{\mathcal{P}} : \mathbb{R}^d \to \mathbb{R}^d$$
$$f_{\mathcal{P}}(\vec{x}) \stackrel{\text{def}}{=} \text{center}(\text{member}_{\mathcal{P}}(\vec{x}))$$

which maps each point to the center of the unique cube in $\mathcal{P}$ containing it.

Definition 10.2.7 (Unit Cube Partition Scaled Rounding Function). Let $d \in \mathbb{N}$ and $\mathcal{P}$ an axis-aligned unit cube partition of $\mathbb{R}^d$ and $c \in (0, \infty)$. The *c-scaled partition rounding function of* $\mathcal{P}$ is the function

$$f_{\mathcal{P},c} : \mathbb{R}^d \to \mathbb{R}^d$$
$$f_{\mathcal{P},c}(\vec{x}) \stackrel{\text{def}}{=} c f_{\mathcal{P}}(\tfrac{1}{c}\vec{x}).$$

The c-scaled partition rounding function of $\mathcal{P}$ essentially uses a scaled copy of $\mathcal{P}$ where cubes have side length c instead of 1.

Remark 10.2.8. If $\mathcal{P}$ is an efficiently computable partition (as is the case for the reclusive partitions (Proposition 10.2.3) and the partition products (Proposition 10.2.4)), then for any $c \in (0, \infty)$, the c-scaled partition rounding function $f_{\mathcal{P},c}$ is efficiently computable as in Remark 10.2.5. △

We next make a simple observation of properties of $f_{\mathcal{P}}$ for a secluded partition.

Observation 10.2.9. *If $\mathcal{P}$ is a (k, ε)-secluded axis-aligned unit cube partition of $\mathbb{R}^d$, then the partition rounding function, $f_{\mathcal{P}}$, of $\mathcal{P}$ has the following two properties:*

1. *For all $\vec{x} \in \mathbb{R}^d$, $\|f_{\mathcal{P}}(\vec{x}) - \vec{x}\|_\infty \leq \frac{1}{2}$*
2. *For each $\vec{p} \in \mathbb{R}^d$, the set $\left\{ f_{\mathcal{P}}(\vec{x}) : \vec{x} \in {}^{\infty}\overline{B}_\varepsilon(\vec{p}) \right\}$ has cardinality at most k*

Proof. The first property is because every point is mapped to the center of its containing axis-aligned unit cube, and the second is because ${}^{\infty}\overline{B}_\varepsilon(\vec{p})$ intersects at most k members/cubes of $\mathcal{P}$ by definition of $\mathcal{P}$ being (k, ε)-secluded. □

In our computational context we want something slightly different than this, though, because there is an added layer of approximation occurring; we somehow obtain an approximation $\hat{x}$ of an unknown true/target value $\vec{x}$ and we will round $\hat{x}$ and want the rounded value to remain close to $\vec{x}$. Mostly, this is nothing more than a triangle inequality being incorporated into the above. This leads to the following definition which defines a function f to be $(k, \varepsilon_0, \varepsilon_1)$-*consistent* if (1) after rounding an ε_0-approximation $\hat{x}$ of $\vec{x}$, the rounded value $f(\hat{x})$ remains at least an ε_1-approximation of $\vec{x}$ and (2) for any true/target value $\vec{x}$, the set of all ε_0-approximations of $\vec{x}$ are rounded to at most k distinct values.

Definition 10.2.10 ($(k, \varepsilon_0, \varepsilon_1)$-Consistent). Let $k \in \mathbb{N}$ and $\varepsilon_0, \varepsilon_1 \in (0, \infty)$ and M a metric space with $M' \subseteq M$ and $f : M' \to M$ a function. The function f is called $(k, \varepsilon_0, \varepsilon_1)$-*consistent* if the following two properties hold:

1. For any $\vec{x} \in M'$ and $\hat{x} \in \overline{B}_{\varepsilon_0}(\vec{x})$ it holds that $f(\hat{x}) \in \overline{B}_{\varepsilon_1}(\vec{x})$
2. For any $\vec{x} \in M'$, the set $\left\{ f(\hat{x}) : \hat{x} \in \overline{B}_{\varepsilon_0}(\vec{x}) \right\}$ has cardinality at most k.

Remark 10.2.11. All balls in the definition above are taken with respect to the metric on M, and where necessary the ball is restricted to the subspace M'. That is, for a point in M', the ball is taken relative to the subspace M' and not the entire space M. △

The following result indicates the level of consistency which we can obtain from the partition rounding function of a (k,ε)-secluded partition. The reason that we impose in the following statement that $\varepsilon \in (0,\frac{1}{2})$ is that for axis-aligned unit cube partitions, we generally don't desire the (k,ε)-secluded property for $\varepsilon > \frac{1}{2}$ because for any axis-aligned unit cube partition this would imply[1] $k \geq 2^d+1$ and we generally don't want k to be this large. We impose that $\varepsilon \in (0,\frac{1}{2}]$ to obtain an inequality, and see this as an extremely reasonable requirement.

Proposition 10.2.12 (Deterministic Rounding Function Guarantees)**.** *Let $d,k \in \mathbb{N}$ and $\varepsilon \in (0,\frac{1}{2}]$ and $\varepsilon_0 \in (0,\infty)$ and $\mathcal{P}$ a (k,ε)-secluded axis-aligned unit cube partition of $\mathbb{R}^d$. Let $\varepsilon_1 = \frac{\varepsilon_0}{\varepsilon}$. Then the ε_1-scaled partition rounding function, $f_{\mathcal{P},\varepsilon_1}$ of $\mathcal{P}$ is $(k,\varepsilon_0,\varepsilon_1)$-consistent.*

Proof. Let $f_{\mathcal{P}}$ denote the partition rounding function of $\mathcal{P}$, let $\vec{x} \in \mathbb{R}^d$ be arbitrary, and let $\hat{x} \in {}^{\infty}\overline{B}_{\varepsilon_0}(\vec{x})$. Note the following chain of inequalities.

$$\begin{aligned}
\left\| \tfrac{1}{\varepsilon_1} f_{\mathcal{P},\varepsilon_1}(\hat{x}) - \tfrac{1}{\varepsilon_1}\vec{x} \right\|_\infty &= \left\| f_{\mathcal{P}}(\tfrac{1}{\varepsilon_1}\hat{x}) - \tfrac{1}{\varepsilon_1}\vec{x} \right\|_\infty && \text{(Def'n of } f_{\mathcal{P},\varepsilon_1} \text{ from } f_{\mathcal{P}}) \\
&\leq \left\| f_{\mathcal{P}}(\tfrac{1}{\varepsilon_1}\hat{x}) - \tfrac{1}{\varepsilon_1}\hat{x} \right\|_\infty + \left\| \tfrac{1}{\varepsilon_1}\hat{x} - \tfrac{1}{\varepsilon_1}\vec{x} \right\|_\infty && \text{(Triangle ineq.)} \\
&= \left\| f_{\mathcal{P}}(\tfrac{1}{\varepsilon_1}\hat{x}) - \tfrac{1}{\varepsilon_1}\hat{x} \right\|_\infty + \tfrac{1}{\varepsilon_1}\left\| \hat{x} - \vec{x} \right\|_\infty && \text{(Norm scaling)} \\
&\leq \left\| f_{\mathcal{P}}(\tfrac{1}{\varepsilon_1}\hat{x}) - \tfrac{1}{\varepsilon_1}\hat{x} \right\|_\infty + \tfrac{\varepsilon_0}{\varepsilon_1} && (\hat{x} \in {}^{\infty}\overline{B}_{\varepsilon_0}(\vec{x})) \\
&\leq \tfrac{1}{2} + \tfrac{\varepsilon_0}{\varepsilon_1} && \text{(Observation 10.2.9 for point } \tfrac{1}{\varepsilon_1}\hat{x}) \\
&= \tfrac{1}{2} + \varepsilon && \text{(Def'n of } \varepsilon_1)
\end{aligned}$$

[1] Let $\vec{p}$ be the center of some cube X in the partition and consider ${}^{\infty}\overline{B}_{\varepsilon}(\vec{p})$. If $\varepsilon > \frac{1}{2}$, then ${}^{\infty}\overline{B}_{\varepsilon}(\vec{p})$ trivially intersects X, but none of the 2^d corners are contained in X and must each belong to some other cube in the partition, and no two of them belong to the same cube because they are too far apart.

$$\leq 1 \qquad (\varepsilon \in (0, \tfrac{1}{2}])$$

Scaling the first and last expression by ε_1 and applying the scaling property of norms proves property (1) of the definition of consistency:

$$\|f_{\mathcal{P},\varepsilon_1}(\hat{x}) - \vec{x}\|_\infty \leq \varepsilon_1.$$

To see that $f_{\mathcal{P},\varepsilon_1}$ has property (2) of the definition of consistency, note the following:

$$\begin{aligned}
\left\{f_{\mathcal{P},\varepsilon_1}(\hat{x}) : \hat{x} \in {}^{\infty}\overline{B}_{\varepsilon_0}(\vec{x})\right\} &= \left\{\varepsilon_1 f_{\mathcal{P}}(\tfrac{1}{\varepsilon_1}\hat{x}) : \hat{x} \in {}^{\infty}\overline{B}_{\varepsilon_0}(\vec{x})\right\} && \text{(Def'n of } f_{\mathcal{P},\varepsilon_1} \text{ from } f_{\mathcal{P}}) \\
&= \left\{\varepsilon_1 f_{\mathcal{P}}(\vec{a}) : \vec{a} \in {}^{\infty}\overline{B}_{\frac{\varepsilon_0}{\varepsilon_1}}(\vec{x})\right\} && \text{(Scaling norm balls)} \\
&= \left\{\varepsilon_1 f_{\mathcal{P}}(\vec{a}) : \vec{a} \in {}^{\infty}\overline{B}_{\varepsilon}(\vec{x})\right\} && \text{(Defn'n of } \varepsilon_1)
\end{aligned}$$

and since the scaling of $f_{\mathcal{P}}$ by $\varepsilon_1 > 0$ does not affect the cardinality of the set, we have by Observation 10.2.9 that

$$\left|\left\{f_{\mathcal{P},\varepsilon_1}(\hat{x}) : \hat{x} \in {}^{\infty}\overline{B}_{\varepsilon_0}(\vec{x})\right\}\right| = \left|\left\{f_{\mathcal{P}}(\vec{a}) : \vec{a} \in {}^{\infty}\overline{B}_{\varepsilon}(\vec{x})\right\}\right| \leq k$$

which completes the proof. □

Ideally, we would like to take ε_1 as small as possible so that to the extent possible we can limit how much worse the approximation might become after applying the rounding function $f_{\mathcal{P},\varepsilon_1}$. Initially, the if the approximation is at least ε_0 close, then after applying the rounding function $f_{\mathcal{P},\varepsilon_1}$ the approximation will be at least ε_1 close, and so we wish to minimize the ratio $\frac{\varepsilon_1}{\varepsilon_0}$ which implies maximizing ε. However, because our main concern in this work is having consistency/replicability we want to balance this with keeping k small—possibly the minimum possible, but ideally at least polynomial in d. This is why in our search for (k, ε)-secluded partitions, we wished to minimize k and then maximize ε conditional on that, and then later tried to maximize ε conditional on k being polynomial in d. So we will shortly present Proposition 10.2.12 with $\varepsilon = \frac{1}{2d}$ and $k = d+1$ (and one could likewise use the parameters from the partition products, but we don't gain more than a constant factor in ε with them if we keep k polynomial in d). This leads to the following immediate corollary which says that for the ℓ_∞ norm, we can achieve consistency parameter $d+1$ in $\mathbb{R}^d$ at the cost of the ratio of final approximation quality to initial approximation quality $(\varepsilon_1/\varepsilon_0)$ being linear in d.

Corollary 10.2.13 (Reclusive Deterministic Rounding Function Guarantees)**.** *Let $d, k \in \mathbb{N}$ and $\varepsilon_0 \in (0, \infty)$. There exists a $(d+1, \varepsilon_0, 2d\varepsilon_0)$-consistent function $f : \mathbb{R}^d \to \mathbb{R}^d$.*

Proof. By Theorem 4.2.18 (Existence of $(d+1, \frac{1}{2d})$-Secluded Unit Cube Partitions) there exists a $(d+1, \frac{1}{2d})$-secluded axis-aligned unit cube partition of $\mathbb{R}^d$, so apply Proposition 10.2.12 with $k = d+1$ and $\varepsilon = \frac{1}{2d} \in (0, \frac{1}{2}]$. □

Furthermore, our impossibility results of Chapter 6 (Optimality of k in General) and Chapter 7 (Near Optimality of ε in General) prove that we cannot do much better than the consistency of the deterministic rounding functions above: achieving $(k, \varepsilon_0, \varepsilon_1)$-consistency first of all requires $k \geq d+1$ and furthermore, and second, if $k \in \mathsf{poly}(d)$ then the loss of approximation quality $(\frac{\varepsilon_1}{\varepsilon_0})$ must be nearly linear in d—specifically, $\frac{\varepsilon_1}{\varepsilon_0} \in \Omega(\frac{d}{\log(d)})$. Not only is this true with regard to the ℓ_∞ norm (which is used in the constructive result above), but it is true of every norm.

Proposition 10.2.14 (Deterministic Rounding Function Limitations for $\mathbb{R}^d$). *Let $d, k \in \mathbb{N}$ and $\varepsilon_0, \varepsilon_1 \in (0, \infty)$ and $\|\cdot\|$ a norm on $\mathbb{R}^d$ and $f : \mathbb{R}^d \to \mathbb{R}^d$ a $(k, \varepsilon_0, \varepsilon_1)$-consistent function with respect to $\|\cdot\|$. Then $k \geq d + 1$ and $\frac{\varepsilon_1}{\varepsilon_0} \geq \frac{d}{2\log_4(k)}$. In particular, if k is at most d^t for some $t \in [1, \infty)$, then $\frac{\varepsilon_1}{\varepsilon_0} \geq \frac{d}{2t\log_4(d)} \in \Omega(\frac{d}{\log(d)})$.*

Proof. Let $\mathcal{P}$ be the partition of fibers/preimages of f: $\mathcal{P} = \{f^{-1}(\vec{y}) : \vec{y} \in \text{range}(f)\}$, and let $\varepsilon = \frac{\varepsilon_0}{\varepsilon_1}$ (the reciprocal of the quantity we are interested in).

Claim A. *Each member of $\mathcal{P}$ has diameter at most $2\varepsilon_1$ (with respect to $\|\cdot\|$).*

Proof of Claim. Let $X \in \mathcal{P}$ be arbitrary, so by definition there is some $\vec{y} \in \text{range}(f)$ such that $X = f^{-1}(\vec{y})$. Let $\vec{x}, \vec{x}\,' \in X$ be arbitrary. This implies that $f(\vec{x}) = \vec{y} = f(\vec{x}\,')$. Taking $\hat{x} = \vec{x} \in {}^{\|\cdot\|}\overline{B}_{\varepsilon_0}(\vec{x})$, we have by property (1) of the definition of consistency that $\|\vec{x} - \vec{y}\| = \|\hat{x} - f(\vec{x})\| \leq \varepsilon_1$ and similarly, $\|\vec{x}\,' - \vec{y}\| \leq \varepsilon_1$. By the triangle inequality, $\|\vec{x} - \vec{x}\,'\| \leq 2\varepsilon_1$ which demonstrates that $\text{diam}_{\|\cdot\|}(X) \leq 2\varepsilon_1$. ■

Claim B. *For each $\vec{x} \in \mathbb{R}^d$, ${}^{\|\cdot\|}\overline{B}_{\varepsilon_0}(\vec{x})$ intersects at most k members of $\mathcal{P}$ (i.e. $\left|{}^{\|\cdot\|}\overline{\mathcal{N}}_{\varepsilon_0}(\vec{x})\right| \leq k$) (i.e. $\mathcal{P}$ has the property analogous to being (k, ε_0)-secluded but with respect to $\|\cdot\|$ instead of the ℓ_∞ norm.)*

Proof of Claim. Let $\vec{x} \in \mathbb{R}^d$ be arbitrary, and consider the set of members of $\mathcal{P}$ which are intersected by ${}^{\|\cdot\|}\overline{B}_{\varepsilon_0}(\vec{x})$ (i.e. the set ${}^{\|\cdot\|}\overline{\mathcal{N}}_{\varepsilon_0}(\vec{x})$). We will establish that there is an injection from ${}^{\|\cdot\|}\overline{\mathcal{N}}_{\varepsilon_0}(\vec{x})$ into $\left\{f(\hat{x}) : \hat{x} \in {}^{\infty}\overline{B}_{\varepsilon_0}(\vec{x})\right\}$. Define the injection

$$g : {}^{\|\cdot\|}\overline{\mathcal{N}}_{\varepsilon_0}(\vec{x}) \hookrightarrow \left\{f(\hat{x}) : \hat{x} \in {}^{\infty}\overline{B}_{\varepsilon_0}(\vec{x})\right\}$$
$$g(f^{-1}(\vec{y})) = \vec{y}.$$

This is well defined function because ${}^{\|\cdot\|}\overline{\mathcal{N}}_{\varepsilon_0}(\vec{x}) \subseteq \mathcal{P}$, so by definition of $\mathcal{P}$, every element of this set has the form $f^{-1}(\vec{y})$ for some $\vec{y} \in \text{range}(f)$, and the codomain is valid because if $f^{-1}(\vec{y}) \in {}^{\|\cdot\|}\overline{\mathcal{N}}_{\varepsilon_0}(\vec{x})$, then by definition of the neighborhood $f^{-1}(\vec{y}) \cap {}^{\|\cdot\|}\overline{B}_{\varepsilon_0}(\vec{x}) \neq \emptyset$, so there is some $\hat{x} \in f^{-1}(\vec{y})$ such that also $\hat{x} \in {}^{\|\cdot\|}\overline{B}_{\varepsilon_0}(\vec{x})$ (i.e. there is some $\hat{x} \in {}^{\|\cdot\|}\overline{B}_{\varepsilon_0}(\vec{x})$ with $f(\hat{x}) = \vec{y}$, so $\vec{y}$ belongs to the specified codomain). Finally, the function g is an injection because for two different members $X, X' \in {}^{\|\cdot\|}\overline{\mathcal{N}}_{\varepsilon_0}(\vec{x})$ there exists unique $\vec{y}, \vec{y}\,' \in \operatorname{range}(f)$ such that $X = f^{-1}(\vec{y})$ and $X' = f^{-1}(\vec{y}\,')$ and $\vec{y}$ and $\vec{y}\,'$ must be distinct otherwise $X = X'$.

Because ${}^{\|\cdot\|}\overline{\mathcal{N}}_{\varepsilon_0}(\vec{x})$ can be injected into by $\left\{f(\hat{x}) : \hat{x} \in {}^{\infty}\overline{B}_{\varepsilon_0}(\vec{x})\right\}$ and by property (2) of the definition of consistency, this latter set has cardinality at most k, so does the former which demonstrates that ${}^{\|\cdot\|}\overline{B}_{\varepsilon_0}(\vec{x})$ intersects at most k members of $\mathcal{P}$. ■

Because there is a common diameter bound on every member of $\mathcal{P}$ (Claim A), by Theorem 6.2.2 (Stronger Optimality Theorem) there exists $\vec{p} \in \mathbb{R}^d$ such that $\left|{}^{\|\cdot\|}\overline{\mathcal{N}}_{\varepsilon_0}(\vec{p})\right| \geq d+1$, and thus by Claim B, $k \geq d+1$ which proves one part of the statement.

Similarly, combining the diameter bound of Claim A with the universal characterization of Claim B and the existential claim of Corollary 7.1.3 (ε-Neighborhoods for Diameter Bounded Partitions), there exists some point $\vec{p} \in \mathbb{R}^d$ such that

$$k \geq \left|{}^{\|\cdot\|}\overline{\mathcal{N}}_{\varepsilon_0}(\vec{p})\right| \geq \left(1 + \tfrac{2\varepsilon_0}{2\varepsilon_1}\right)^d = (1+\varepsilon)^d \tag{10.1}$$

We could stop here, but with a bit more work, we can get a more convenient form of the inequality. To do so, we need an upper bound on ε which is lower bound on ε_1 in terms of ε_0. In particular, it should be intuitively clear that $\varepsilon_1 \geq \varepsilon_0$ because it should be impossible to have a function $f : \mathbb{R}^d \to \mathbb{R}^d$ which increases the accuracy of approximations of all points, and indeed it is.

Claim C. *It must be that* $\varepsilon_1 \geq \varepsilon_0$

Proof of Claim. Let $\vec{v} \in \mathbb{R}^d$ be such that $\|\vec{v}\| = \varepsilon_0$ noting that also $\|-\vec{v}\| = \varepsilon_0$. Because $\vec{0} \in {}^{\|\cdot\|}\overline{B}_{\varepsilon_0}(\vec{v})$, by property (1) of the definition of consistency, we have $f(\vec{0}) \in {}^{\|\cdot\|}\overline{B}_{\varepsilon_1}(\vec{v})$. Similarly, because $\vec{0} \in {}^{\|\cdot\|}\overline{B}_{\varepsilon_0}(-\vec{v})$, we have $f(\vec{0}) \in {}^{\|\cdot\|}\overline{B}_{\varepsilon_1}(-\vec{v})$. Thus, $f(\vec{0}) \in {}^{\|\cdot\|}\overline{B}_{\varepsilon_1}(\vec{v}) \cap {}^{\|\cdot\|}\overline{B}_{\varepsilon_1}(-\vec{v})$ and that this intersection is non-empty implies that $\varepsilon_1 \geq \varepsilon_0$ which proves the claim[2]. ■

Now, because $\varepsilon_1 \geq \varepsilon_0$ we have that $\varepsilon = \frac{\varepsilon_0}{\varepsilon_1} \in (0, 1]$, so applying[3] Fact G.0.3, we can obtain the following from Equation 10.1.

$$\varepsilon \leq \frac{2\log_4(k)}{d}$$

Taking reciprocals and substituting $\varepsilon = \frac{\varepsilon_0}{\varepsilon_1}$ shows the desired result.

$$\frac{\varepsilon_1}{\varepsilon_0} \geq \frac{d}{2\log_4(k)}$$

The “in particular” part of the result then follows trivially. □

While it might seem intuitive that a similar result to Proposition 10.2.14 should hold for functions $f : [0,1]^d \to [0,1]^d$, it is not clear to us that such a result is actually true. We can obtain an analogous impossibility result for the ℓ_∞ norm, but we have no such result for other norms[4]. The analog is obtained by using our variant of the Lebesgue covering theorem and cubical KKM lemma. We will state the result as $f : [0,1]^d \to \mathbb{R}^d$ instead of $f : [0,1]^d \to [0,1]^d$ since this is more lenient.

[2] This reason this claim holds is similar to, but not quite the same as the Banach fixed point theorem, though we could probably use that theorem to prove this result if we went through enough setup.

[3] Replace “ε” with “$\varepsilon/2$”.

[4] One can use Theorem 3.5.1 (Equivalence of Norms on $\mathbb{R}^d$) to convert between the ℓ_∞ norm and any other chosen norm, but then the bound on $\varepsilon_1/\varepsilon_0$ is no longer the same (even asymptotically) and depends on the specifics of the chosen norm.

In the following statement we now require $\varepsilon_1 \in (0, \frac{1}{2})$ because otherwise we have a triviality; if $\varepsilon_1 \geq \frac{1}{2}$ then $f : [0,1]^d \to \mathbb{R}^d$ can map every point to $\frac{1}{2} \cdot \vec{1} = \langle \frac{1}{2} \rangle_{i=1}^d$ and for every $\varepsilon_0 \in (0, \infty)$ this function is trivially $(1, \varepsilon_0, \varepsilon_1)$-consistent, and so we can say absolutely nothing about the ratio $\frac{\varepsilon_1}{\varepsilon_0}$ because it could be anything in $(0, \infty)$. The proof of the following result will be very similar to that of Proposition 10.2.14.

Proposition 10.2.15 (Deterministic Rounding Function Limitations for the Cube)**.** *Let $d, k \in \mathbb{N}$ and $\varepsilon_0 \in (0, \infty)$ and $\varepsilon_1 \in (0, \frac{1}{2})$ and $f : [0,1]^d \to \mathbb{R}^d$ a $(k, \varepsilon_0, \varepsilon_1)$-consistent function with respect to the ℓ_∞ norm. Then $k \geq d+1$ and $\frac{\varepsilon_1}{\varepsilon_0} \geq \frac{d}{6\log_4(k)}$. In particular, if k is at most d^t for some $t \in [1, \infty)$, then $\frac{\varepsilon_1}{\varepsilon_0} \geq \frac{d}{6t\log_4(d)} \in \Omega(\frac{d}{\log(d)})$.*

Proof. Let $\mathcal{P}$ be the partition of $[0,1]^d$ consisting of the fibers/preimages of f: $\mathcal{P} = \{f^{-1}(\vec{y}) : \vec{y} \in \text{range}(f)\}$, and let $\varepsilon = \frac{\varepsilon_0}{\varepsilon_1}$ (the reciprocal of the quantity we are interested in). We state the following two claims without proof as the proofs are identical (up to trivialities) to the corresponding claims in the proof of Proposition 10.2.14.

Claim A. *Each member of $\mathcal{P}$ has diameter at most $2\varepsilon_1$ (with respect to ℓ_∞).*

Claim B. *For each $\vec{x} \in [0,1]^d$, ${}^\infty\overline{B}_{\varepsilon_0}(\vec{x})$ intersects at most k members of $\mathcal{P}$ (i.e. $\mathcal{P}$ is (k, ε_0)-secluded).*

Now we have to do something distinct from the proof of Proposition 10.2.14 which is to rescale the context in which we are working. Let $\varepsilon_1^+ \in (\varepsilon_1, \frac{1}{2})$ be arbitrary (we will eventually pass the result back from ε_1^+ to ε_1 in the limit). Now consider the subcube $[0, 2\varepsilon_1^+]^d \subseteq [0,1]^d$; this is the "correct" cube to think about because it makes ε_0 as large relative to the side length of the cube as possible while maintaining the following property.

Claim C. *No member of $\mathcal{P}$ contains points on opposite faces of the cube $[0, 2\varepsilon_1^+]^d$.*

Proof of Claim. Any pair of points on opposite faces of $[0, 2\varepsilon_1^+]^d$ are ℓ_∞ distance $2\varepsilon_1^+ > 2\varepsilon_1$ apart, but by Claim A each member of $\mathcal{P}$ has diameter at most $2\varepsilon_1$. ■

By Claim C and Theorem 6.1.2 (Infinite KKM/Lebesgue), there exists $\vec{p} \in \mathbb{R}^d$ such that the open set ${}^{\infty}B^{\circ}_{\varepsilon_0}(\vec{p})$ intersects at least $d+1$ members of $\mathcal{P}$ (which means ${}^{\infty}\overline{B}_{\varepsilon_0}(\vec{p})$ does to). Thus, by Claim B, $k \geq d+1$ which proves one part of the statement.

Now we can utilize the following claim.

Claim D. *There exists $\vec{p} \in [0, 2\varepsilon_1^+]^d$ such that ${}^{\infty}\overline{B}_{\varepsilon_0}(\vec{p})$ intersects at least $(1 + \frac{2}{3}\frac{\varepsilon_0}{2\varepsilon_1^+})^d$ members of $\mathcal{P}$.*

Proof. Using the hypothesis of Claim C, this follows from Theorem 8.0.7 (Neighborhood KKM/Lebesgue Theorem) by rescaling the cube $[0, 2\varepsilon_1^+]^d$, the ball ${}^{\infty}\overline{B}_{\varepsilon_0}(\vec{p})$, and every member of $\mathcal{P}$ by a factor of $\frac{1}{2\varepsilon_1^+}$ so that the scaled members of $\mathcal{P}$ cover the unit cube $[0, 1]^d$. □

Now, letting $\vec{p}$ as in Claim D and using Claim B we have the following:

$$k \geq \left| {}^{\|\cdot\|}\overline{\mathcal{N}}_{\varepsilon_0}(\vec{\vec{p}}) \right| \geq \left(1 + \frac{2}{3} \cdot \frac{\varepsilon_0}{2\varepsilon_1^+}\right)^d = \left(1 + \frac{1}{3} \cdot \frac{\varepsilon_0}{\varepsilon_1^+}\right)^d \tag{10.2}$$

Since this inequality is true for arbitrary $\varepsilon_1^+ \in (\varepsilon_1, \frac{1}{2})$, it also holds in the limit, so we have

$$k \geq \left(1 + \frac{1}{3} \cdot \frac{\varepsilon_0}{\varepsilon_1}\right)^d = \left(1 + \frac{1}{3}\varepsilon\right)^d \tag{10.3}$$

As before, to clean up the inequality we want an upper bound on ε which we again do by showing that $\varepsilon_1 \geq \varepsilon_0$, though we need to be slightly more careful this time.

Claim E. *It must be that $\varepsilon_1 \geq \varepsilon_0$.*

Proof of Claim. Let $\varepsilon_0' = \min(\varepsilon_0, \frac{1}{2})$, let $\vec{q} = \frac{1}{2} \cdot \vec{1} = \langle \frac{1}{2} \rangle_{i=1}^d$ which is the midpoint of $[0,1]^d = {}^{\infty}\overline{B}_{\frac{1}{2}}(\vec{q})$. Let $\vec{v} \in \mathbb{R}^d$ be such that $\|\vec{v}\|_\infty = \varepsilon_0' \leq \frac{1}{2}$ noting that also $\|-\vec{v}\|_\infty = \varepsilon_0' \leq \frac{1}{2}$. Let $\vec{w} = \vec{q} + \vec{v}$ and $\vec{w}\,' = \vec{q} - \vec{v}$ noting that $\|\vec{w} - \vec{q}\|_\infty = \varepsilon_0'$ and similarly $\|\vec{w}\,' - \vec{q}\|_\infty = \varepsilon_0'$.

Thus, because $\vec{q} \in {}^{\infty}\overline{B}_{\varepsilon_0'}(\vec{w}) \subseteq {}^{\infty}\overline{B}_{\varepsilon_0}(\vec{w})$, then by property (1) of the definition of consistency, we have $f(\vec{q}) \in {}^{\infty}\overline{B}_{\varepsilon_1}(\vec{w})$. Similarly, we have $f(\vec{q}) \in {}^{\infty}\overline{B}_{\varepsilon_1}(\vec{w}\,')$. Thus, $f(\vec{q}) \in {}^{\infty}\overline{B}_{\varepsilon_1}(\vec{w}) \cap {}^{\infty}\overline{B}_{\varepsilon_1}(\vec{w}\,')$, and that this intersection is non-empty implies that $\varepsilon_1 \geq \varepsilon_0'$ (because $\|\vec{w} - \vec{w}\,'\|_\infty = 2\varepsilon_0'$) and thus $\varepsilon_1 \geq \varepsilon_0' \geq \varepsilon_0$ which proves the claim. ■

Now, as before, because $\varepsilon_1 \geq \varepsilon_0$ we have that $\varepsilon = \frac{\varepsilon_0}{\varepsilon_1} \in (0,1]$, so applying[5] Fact G.0.3, we can obtain the following from Equation 10.1.

$$\varepsilon \leq \frac{6 \log_4(k)}{d}$$

Taking reciprocals and substituting $\varepsilon = \frac{\varepsilon_0}{\varepsilon_1}$ shows the desired result.

$$\frac{\varepsilon_1}{\varepsilon_0} \geq \frac{d}{6 \log_4(k)}$$

The "in particular" part of the result then follows trivially. □

We could also consider not just consistent functions (which are deterministic algorithms without computability concerns) but also randomized algorithms.

[5]This says $\log_4(1+x) \geq \frac{x}{2}$ for $x \in [0,1]$. We could tighten the approximation a little bit if desired because in our context we have $x = \frac{1}{3}\varepsilon \in [0, \frac{1}{3}]$.

Definition 10.2.16 ($(k, \varepsilon_0, \varepsilon_1, \delta)$-Consistent)*.* Let $k \in \mathbb{N}$ and $\varepsilon_0, \varepsilon_1 \in (0, \infty)$ and $\delta \in (0, \frac{1}{2})$ and M a metric space with $M' \subseteq M$ and A a randomized algorithm mapping inputs in M' to outputs in M. The randomized algorithm A is called $(k, \varepsilon_0, \varepsilon_1, \delta)$-*consistent* if the following two properties hold:

1. For any $\vec{x} \in M'$ and $\hat{x} \in \overline{B}_{\varepsilon_0}(\vec{x})$ it holds with probability at least $1 - \delta$ that $A(\hat{x}) \in \overline{B}_{\varepsilon_1}(\vec{x})$
2. For any $\vec{x} \in M'$, the set $\left\{ f(\hat{x}) : \hat{x} \in \overline{B}_{\varepsilon_0}(\vec{x}) \right\}$ has cardinality at most k.

We conjecture that the limitation results we have obtained also hold in the randomized setting as well and are currently working to show this.

Conjecture 10.2.17 (Randomized Rounding Limitations)**.** *Up to changes in constants and the inclusion of the δ parameter, the bounds of both Proposition 10.2.14 and Proposition 10.2.15 hold for $(k, \varepsilon_0, \varepsilon_1, \delta)$-consistent randomized algorithms.*

10.3 Limitations on Learning

In this section we will present two results on limitations of learning in two different models: (1) a statistical query model and (2) a PAC learning model. These two results will not use any of the mathematical results proven in this work, though they do use similar ideas and employ the cubical KKM lemma and/or Lebesgue covering theorem.

10.3.1 Biased Coin Estimation Problem in Statistical Query Model

We will demonstrate that there are certain problems (parameterized by a dimension d) for which no k-pseudodeterministic algorithm exists for $k < d + 1$. The d-Coin

Bias Estimation Problem defined below is one such problem.

Problem 10.3.1. The d-Coin Bias Estimation Problem is the following problem: design an algorithm A (possibly randomize) that given $\varepsilon \in (0, \infty)$, $\delta \in (0,1]$, and n independent tosses (for each coin) of an ordered collection of d-many biased coins with a bias vector $\vec{b} \in [0,1]^d$ outputs with probability at least $1-\delta$ an estimated bias vector $\vec{v} \in [0,1]^d$ such that $\|\vec{b} - \vec{v}\|_\infty \leq \varepsilon$.

Definition 10.3.2. We say an algorithm A for the d-Coin Bias Estimation Problem is k-pseudodeterministic if for any bias vector $\vec{b} \in [0,1]^d$, and parameters ε, δ, there is set $L_{\vec{b}} \subseteq {}^{\infty}\overline{B}_{\varepsilon}(\vec{b})$ and an integer n such that $\left|L_{\vec{b}}\right| \leq k$ and A on input ε and δ and n independent tosses (per coin) according to the bias vector $\vec{b}$, outputs with probability at least $1-\delta$ an estimate $\vec{v} \in L_{\vec{b}}$. The sample complexity of A is the value n.

Proposition 10.3.3. *There exists a $(d+1)$-pseudodeterministic algorithm for the d-Coin Bias Estimation Problem.*

Proof Sketch. Let $\varepsilon_0 = \frac{\varepsilon}{2d}$ so $\varepsilon = 2d\varepsilon_0$. The algorithm takes enough samples so that with probability at least $1-\delta$, the empirical bias $\vec{v}$ of the coins is within ε_0 of the true bias $\vec{b}$ (i.e. $\vec{v} \in {}^{\infty}\overline{B}_{\varepsilon_0}(\vec{b})$) and then round $\vec{v}$ using the $(d+1, \varepsilon_0, \varepsilon = 2d\varepsilon_0)$-consistent function $f : \mathbb{R}^d \to \mathbb{R}^d$ of Corollary 10.2.13. Then (if desired) truncate the final rounded version so that it is a valid bias vector in $[0,1]^d$; this truncation does not worsen the approximation quality. □

Theorem 10.3.4 (No $(d+1)$-Pseudoterministic Algorithm for the d-Coin Bias Estimation Problem)**.** *For $k < d+1$, there does not exist a k-pseudodeterministic algorithm for the d-*Coin Bias Estimation Problem.

Before proving the theorem, we need a lemma. The notation $\mathcal{D}_{A,\vec{b},n}$ means the distribution of the outputs of randomized algorithm A when it receives n independent flips of each of d coins according to the bias vector $\vec{b}$, and d_{TV} is the total variation distance.

Lemma 10.3.5. *For biases $\vec{a}, \vec{b} \in [0,1]^d$ we have $\mathrm{d}_{\mathrm{TV}}\left(\mathcal{D}_{A,\vec{a},n}, \mathcal{D}_{A,\vec{b},n}\right) \leq nd\|\vec{b} - \vec{a}\|_\infty$.*

Proof. We can view the model as algorithm A having access to a single draw from a distribution. The distribution giving one sample flip of each coin in a collection with bias $\vec{b}$ is the d-fold product of Bernoulli distributions $\prod_{i=1}^d \mathrm{Bern}(b_i)$ (which for notational brevity we denote as $\mathrm{Bern}(\vec{b})$, so the distribution which gives n independent flips of each coin is the n-fold product of this (using notation of [Can15] written as $\mathrm{Bern}(\vec{b})^{\otimes n}$).

Comparing the distributions of n independent flips of the d coins for bias $\vec{b}$ as compared to bias $\vec{a}$, we have for each $i \in [d]$ that

$$\mathrm{d}_{\mathrm{TV}}\left(\mathrm{Bern}(b_i), \mathrm{Bern}(a_i)\right) = |b_i - a_i|$$

so by C.1.2 and C.1.3 of [Can15] we have

$$\mathrm{d}_{\mathrm{TV}}\left(\mathrm{Bern}(\vec{b}), \mathrm{Bern}(\vec{a})\right) \leq \sum_{i=1}^{d} |b_i - a_i| \leq d\left\|\vec{b} - \vec{a}\right\|_\infty$$

and

$$\mathrm{d}_{\mathrm{TV}}\left(\mathrm{Bern}(\vec{b})^{\otimes n}, \mathrm{Bern}(\vec{a})^{\otimes n}\right) \leq nd\left\|\vec{b} - \vec{a}\right\|_\infty.$$

Because A is a randomized function of one draw of this distribution, by D.1.2 of

[Can15] we have that A cannot serve to increase the total variation distance, so

$$\mathrm{d_{TV}}\left(\mathcal{D}_{A,\vec{a},n},\mathcal{D}_{A,\vec{b},n}\right) \leq \mathrm{d_{TV}}\left(\mathrm{Bern}(\vec{b})^{\otimes n},\mathrm{Bern}(\vec{a})^{\otimes n}\right) \leq d\left\|\vec{b}-\vec{a}\right\|_\infty$$

which completes the proof. □

Proof of Theorem 10.3.4. Fix any $d \in \mathbb{N}$, and choose ε and δ as $\varepsilon < \frac{1}{2}$ and $\delta \leq \frac{1}{d+2}$.

Suppose for contradiction that such an algorithm does exist for some $k < d+1$. This means that for each possible bias $\vec{b} \in [0,1]^d$, there exists some set $L_{\vec{b}} \subseteq {}^{\infty}\overline{B}_{\varepsilon}(\vec{b})$ (not necessarily unique, but consider some fixed one) with $\left|L_{\vec{b}}\right| \leq k$ such that with probability at least least $(1-\delta) \geq (1-\frac{1}{d+2}) = \frac{d+1}{d+2} \geq \frac{k+1}{k+2}$, A returns an element of $L_{\vec{b}}$. By the trivial averaging argument (since $\left|L_{\vec{b}}\right| \leq k$), this means that there exists at least one element in $L_{\vec{b}}$ which is returned by A with probability at least $\frac{1}{k}\cdot\frac{k+1}{k+2}$. Let $f\colon [0,1]^d \to [0,1]^d$ be a function which maps each bias $\vec{b}$ to such an element of $L_{\vec{b}}$.

Since $\frac{1}{k}\cdot\frac{k+1}{k+2} > \frac{1}{k+1}$, let η be such that $0 < \eta < \frac{1}{k}\cdot\frac{k+1}{k+2} - \frac{1}{k+1}$.

The function f induces a partition $\mathcal{P}$ of $[0,1]^d$ where the members of $\mathcal{P}$ are the fibers of f (i.e. $\mathcal{P} = \{f^{-1}(\vec{y})\colon \vec{y} \in \mathrm{range}(f)\}$). By definition, for any member $X \in \mathcal{P}$ there exists some $\vec{y} \in \mathrm{range}(f)$ such that for all $\vec{b} \in X$, $f(\vec{b}) = \vec{y}$. By definition of k-pseudodeterministic ε-approximation, we have $f(\vec{b}) \in L_{\vec{b}} \subseteq {}^{\infty}\overline{B}_{\varepsilon}(\vec{b})$ showing that $\vec{y} \in {}^{\infty}\overline{B}_{\varepsilon}(\vec{b})$ and by symmetry $\vec{b} \in {}^{\infty}\overline{B}_{\varepsilon}(\vec{y})$. This shows that $X \subseteq {}^{\infty}\overline{B}_{\varepsilon}(\vec{y})$, so $\mathrm{diam}_{\infty}(X) \leq 2\varepsilon < 1$.

Let $r = \frac{\eta}{dn}$. Since every member of $\mathcal{P}$ has ℓ_∞ diameter less than 1, no member includes points of opposing faces, so by Theorem 6.1.2 (Infinite KKM/Lebesgue) there exists a point $\vec{p} \in [0,1]^d$ such that ${}^{\infty}\overline{B}_r(\vec{p})$ intersects at least $d+1 > k$ members of $\mathcal{P}$. Let $\vec{b}^{(1)},\ldots,\vec{b}^{(d+1)}$ be points belonging to distinct members of $\mathcal{P}$ that all belong to ${}^{\infty}\overline{B}_r(\vec{p})$. By definition of $\mathcal{P}$, this means for distinct $j,j' \in [d+1]$ that $f(\vec{b}^{(j)}) \neq f(\vec{b}^{(j')})$.

Now, for each $j \in [d+1]$, because $\left\|\vec{p}-\vec{b}^{(j)}\right\|_\infty \leq r$, by Lemma 10.3.5 we have

$d_{TV}(\mathcal{D}_{A,\vec{p},n}, \mathcal{D}_{A,\vec{b^{(j)}},n}) \leq ndr = \eta$. However, this gives rise to a contradiction because the probability that A with access to biased coins $\vec{b}^{(j)}$ returns $f(\vec{b}^{(j)})$ is at least $\frac{1}{k} \cdot \frac{k+1}{k+2}$ (by definition of f), and by the total variation distance, it must be that A with access to biased coins $\vec{p}$ returns $f(\vec{b}^{(j)})$ with probability at least $\frac{1}{k} \cdot \frac{k+1}{k+2} - \eta > \frac{1}{k+1}$. This is a contradiction because a distribution cannot have $d + 1 \geq k + 1$ disjoint events that each have probability strictly greater than $\frac{1}{k+1}$. □

10.3.2 Threshold Estimation Problem in PAC Model

In this section, we establish a similar type of impossibility result, but for the PAC (Probably Almost Correct) model. We begin by defining the PAC learning model.

Let $\mathcal{H}$ be a (hypothesis) class of Boolean functions over X, and $\mathcal{D}$ be a distribution over X. For a function $f \in \mathcal{H}$, let $\mathcal{D}_f$ denote the distribution over $X \times \{0, 1\}$ that is obtained by sampling an element $x \in X$ according $\mathcal{D}$ and outputting $\langle x, f(x)\rangle$. For a hypotheses h and h', the error with respect to a distribution $\mathcal{D}$ is denoted by $e_{\mathcal{D}}(h, h') \stackrel{\text{def}}{=} \Pr_{x\sim\mathcal{D}}[h(x) \neq h'(x)]$.

Definition 10.3.6. A hypothesis class (or concept class) $\mathcal{H}$ of binary functions on X is PAC learnable with sample complexity n if there is a learning algorithm A with the following property: for every $f \in \mathcal{H}$ and distribution $\mathcal{D}$ over X, for all $\varepsilon, \delta \in (0, 1)$, the algorithm A on inputs ε, δ and samples S drawn i.i.d. according to $\mathcal{D}_f$ where $|S| \leq n$ outputs with probability at least $(1 - \delta)$ a hypothesis h so that $e_{\mathcal{D}_f}(h, f) \leq \varepsilon$.

We will define a problem (parameterized by d) in the PAC learning model and then proceed to show that this problem is not k-pseudodeterministically learnable in the PAC model for $k < d + 1$.

Problem 10.3.7 (d-Threshold Estimation Problem). Fix some $d \in \mathbb{N}$. Let $X = [0,1]^d$. For each value $\vec{t} \in [0,1]^d$ (which happens to be the same as X), let $h_{\vec{t}} : X \to \{0,1\}$ be the function defined by

$$h_{\vec{t}}(\vec{x}) = \begin{cases} 1 & \text{for each } i \in [d] \text{, it holds that } x_i \leq t_i \\ 0 & \text{otherwise} \end{cases}.$$

This is the function which determines if each coordinate is less than or equal to the thresholds specified by $\vec{t}$. Let $\mathcal{H}$ be the hypothesis class consisting of all such threshold functions: $\mathcal{H} = \left\{ h_{\vec{t}} \mid \vec{t} \in [0,1]^d \right\}$.

In the rest of this section, we will use the notation in the definition of d-Threshold Estimation Problem. The proof that for $k < d+1$, there is no algorithm which learns d-Threshold Estimation Problem in the PAC model in a k-pseudodeterministic manner is similar to the proof of Theorem 10.3.4. The reason is that sampling d-many biased coins with biases $\vec{b}$ is similar to obtaining a point $\vec{x}$ uniformly at random from $[0,1]^d$ and evaluating the threshold function $h_{\vec{b}}$ on it—this corresponds to asking whether all of the coins were heads/1's. The two models differ though because in the sample model for the d-Coin Bias Estimation Problem, the algorithm sees for each coin whether it is heads or tails, but this information is not available in the PAC model for the d-Threshold Estimation Problem. Conversely, in the PAC model for the d-Threshold Estimation Problem, a random draw from $[0,1]^d$ is available to the algorithm, but in the sample model for the d-Coin Bias Estimation Problem the algorithm does not get this information.

Furthermore, there is the following additional complexity in the impossibility result for the d-Threshold Estimation Problem. In the d-Coin Bias

Estimation Problem, we said by definition that a collection of d coins parameterized by bias vector $\vec{a}$ was an ε-approximation to a collection of d coins parameterized by bias vector $\vec{b}$ if and only if $\|\vec{b}-\vec{a}\|_\infty \leq \varepsilon$, and we used this norm in applying Theorem 6.1.2 (Infinite KKM/Lebesgue). However, the notion of ε-approximation in the PAC model is quite different than this. It is possible to have a hypotheses $h_{\vec{a}}$ and $h_{\vec{b}}$ in the d-Threshold Estimation Problemsuch that $\|\vec{b}-\vec{a}\|_\infty > \varepsilon$ but with respect to some distribution $\mathcal{D}_X$ on the domain X we have $\mathrm{e}_{\mathcal{D}_X}(h_{\vec{a}}, h_{\vec{b}}) \leq \varepsilon$. For example, if $\mathcal{D}_X$ is the uniform distribution on $X = [0,1]^d$ and $\vec{a} = \vec{0}$ and $\vec{b}$ is the first standard basis vector $\vec{b} = \vec{e}^{(1)} = \langle 1, 0, \ldots, 0\rangle$, and $\varepsilon = \frac{1}{2}$, then $\|\vec{b}-\vec{a}\|_\infty = 1 > \varepsilon$, but $\mathrm{e}_{\mathcal{D}_X}(h_{\vec{a}}, h_{\vec{b}}) = 0 \leq \varepsilon$ because $h_{\vec{a}}(\vec{x}) \neq h_{\vec{b}}(\vec{x})$ if and only if all of the last $d-1$ coordinates of $\vec{x}$ are 0 and the first coordinate is > 0, but there is probability 0 of sampling such $\vec{x}$ from the uniform distribution on $X = [0,1]^d$.

For this reason, we can't just partition $[0,1]^d$ as we did with the proof of Theorem 10.3.4 and must do something more clever. It turns out that it is possible to find a subset $[\alpha,1]^d$ on which hypotheses parameterized by vectors on opposite faces of this cube $[\alpha,1]^d$ have high PAC error between them. A consequence by the triangle inequality of $\mathrm{e}_{\mathcal{D}_X}$ is that two such hypotheses cannot both be approximated by a common third hypothesis. That is what the following two lemmas state.

Lemma 10.3.8. *Let $d \in \mathbb{N}$ and $\alpha = \frac{d-1}{d} = 1 - \frac{1}{d}$. Then $(1-\alpha)\cdot\alpha^{d-1} > \frac{1}{4d}$.*

Proof. If $d = 1$, then $\alpha = 0$ so $(1-\alpha)\cdot\alpha^{d-1} = 1 \geq \frac{1}{4} = \frac{1}{4d}$ (see footnote[6]).

[6] This uses the interpretation that $0^0 = 1$ which is the correct interpretation in the context in which we will use the lemma.

If $d \geq 2$, then we utilize the fact that $(1-\frac{1}{d})^d \geq \frac{1}{4}$ in the following[7]:

$$\begin{aligned}(1-\alpha)\cdot\alpha^{d-1} &= (\tfrac{1}{d})(1-\tfrac{1}{d})^{d-1}\\ &= (\tfrac{1}{d})\frac{(1-\frac{1}{d})^d}{1-\frac{1}{d}}\\ &= \frac{(1-\frac{1}{d})^d}{d-1}\\ &\geq \frac{1}{4(d-1)}\\ &> \frac{1}{4d}.\end{aligned}$$

This completes the proof. As an aside, $\alpha = \frac{d-1}{d}$ is the value of α that maximizes the expression $(1-\alpha)\cdot\alpha^{d-1}$ which is why that value was chosen. □

Lemma 10.3.9. *Let $d \in \mathbb{N}$ and $\alpha = \frac{d-1}{d}$. Let $\vec{s}, \vec{t} \in [\alpha, 1]^d$ such that there exists a coordinate $i_0 \in [d]$ where $s_{i_0} = \alpha$ and $t_{i_0} = 1$ (i.e. $\vec{s}$ and $\vec{t}$ are on opposite faces of this cube). Let $\varepsilon \leq \frac{1}{8d}$. Then there is no point $\vec{r} \in X$ such that both $\mathrm{e}_{\mathrm{unif}}(h_{\vec{s}}, h_{\vec{r}}) \leq \varepsilon$ and $\mathrm{e}_{\mathrm{unif}}(h_{\vec{t}}, h_{\vec{r}}) \leq \varepsilon$ (i.e. there is no hypothesis which is an ε-approximation to both $h_{\vec{s}}$ and $h_{\vec{t}}$).*

Proof. Let $\vec{q} = \left\langle \begin{cases} s_i & i = i_0 \\ t_i & i \neq i_0 \end{cases} \right\rangle_{i=1}^{d}$ which will serve as a proxy to $\vec{s}$.

Claim A. *For each $\vec{x} \in X$, the following are equivalent:*

1. $h_{\vec{q}}(\vec{x}) \neq h_{\vec{t}}(\vec{x})$
2. $h_{\vec{q}}(\vec{x}) = 0$ *and* $h_{\vec{t}}(\vec{x}) = 1$
3. $x_{i_0} \in (q_{i_0}, t_{i_0}] = (\alpha, 1]$ *and for all* $i \in [d] \setminus \{i_0\}$, $x_i \in [0, t_i]$.

Furthermore, the above equivalent conditions imply the following:

[7] One can note that $(1-\frac{1}{d})^d = (1+\frac{(-1)}{d})^d$ which is the defining expression for e^{-1}, and even for $d=2$ this approximation is close enough to exceed $\frac{1}{4}$.

4. $h_{\vec{s}}(\vec{x}) \neq h_{\vec{t}}(\vec{x})$.

Proof of Claim. (2) $\implies$ (1): This is trivial.

(1) $\implies$ (2): Note that because $q_{i_0} = s_{i_0} = \alpha < 1 = t_{i_0}$, we have for all $i \in [d]$ that $q_i \leq t_i$. If $h_{\vec{t}}(\vec{x}) = 0$ then for some $i_1 \in [d]$ it must be that $x_{i_1} > t_{i_1}$, but since $t_{i_1} \geq q_{i_1}$ it would also be the case that $x_{i_1} > q_{i_1}$, so $h_{\vec{q}}(\vec{x}) = 0$ which gives the contradiction that $h_{\vec{q}}(\vec{x}) = h_{\vec{t}}(\vec{x})$. Thus $h_{\vec{t}}(\vec{x}) = 1$, and since $h_{\vec{q}}(\vec{x}) \neq h_{\vec{t}}(\vec{x})$ we have $h_{\vec{q}}(\vec{x}) = 0$.

(1) $\iff$ (3): We partition $[0,1]^d$ into three sets and examine these three cases.

Case 1: $x_{i_0} \in (q_{i_0}, t_{i_0}] = (\alpha, 1]$ and for all $i \in [d] \setminus \{i_0\}$, $x_i \in [0, t_i]$. In this case, $q_{i_0} < x_{i_0}$ so $h_{\vec{q}}(\vec{x}) = 0$ and for all $i \in [d]$ $x_i \leq t_i$, so $h_{\vec{t}}(\vec{x}) = 1$, so $h_{\vec{q}}(\vec{x}) \neq h_{\vec{t}}(\vec{x})$.

Case 2: $x_{i_0} \notin (q_{i_0}, t_{i_0}] = (\alpha, 1]$ and for all $i \in [d] \setminus \{i_0\}$, $x_i \in [0, t_i]$. In this case, because $x_{i_0} \in [0,1]$ and $x_{i_0} \notin (\alpha, 1]$ we have $x_{i_0} \leq \alpha = q_{i_0} \leq t_{i_0}$ and also for all other $i \in [d] \setminus \{i_0\}$, $x_i \leq t_i = q_i$ (by definition of $\vec{q}$). Thus $h_{\vec{q}}(\vec{x}) = 1 = h_{\vec{t}}(\vec{x})$.

Case 3: For some $i_1 \in [d] \setminus \{i_0\}$, $x_{i_1} \notin [0, t_{i_1}]$. In this case, because $x_{i_1} \in [0,1]$, we have $x_{i_1} > t_{i_1} = q_{i_1}$. Thus $h_{\vec{q}}(\vec{x}) = 0 = h_{\vec{t}}(\vec{x})$.

Thus, it is the case that $h_{\vec{q}}(\vec{x}) \neq h_{\vec{t}}(\vec{x})$ if and only if $x_{i_0} \in (q_{i_0}, t_{i_0}] = (\alpha, 1]$ and for all $i \in [d] \setminus \{i_0\}$, $x_i \in [0, t_i]$.

(1), (2), (3) $\implies$ (4): By (2), we have $x_{i_0} > q_{i_0}$, and since $q_{i_0} = s_{i_0}$ by definition of $\vec{q}$, it follows that $x_{i_0} > s_{i_0}$ which means $h_{\vec{s}}(\vec{x}) = 0$. By (3), $h_{\vec{t}}(\vec{x}) = 1$ which gives $h_{\vec{s}}(\vec{x}) \neq h_{\vec{t}}(\vec{x})$. ■

With this claim in hand, our next step will be two prove the following two inequalities:

$$2\varepsilon < \mathrm{e}_{\mathrm{unif}}(h_{\vec{q}}, h_{\vec{t}}) \leq \mathrm{e}_{\mathrm{unif}}(h_{\vec{s}}, h_{\vec{t}}).$$

For the second of these inequalities, note that by the (1) $\implies$ (4) part of claim above, since $h_{\vec{q}}(\vec{x}) \neq h_{\vec{t}}(\vec{x})$ implies $h_{\vec{s}}(\vec{x}) \neq h_{\vec{t}}(\vec{x})$ we have

$$\begin{aligned}
\mathrm{e}_{\mathrm{unif}}(h_{\vec{q}}, h_{\vec{t}}) &= \Pr_{\vec{x} \sim \mathrm{unif}(X)}[h_{\vec{q}}(\vec{x}) \neq h_{\vec{t}}(\vec{x})] \\
&\leq \Pr_{\vec{x} \sim \mathrm{unif}(X)}[h_{\vec{s}}(\vec{x}) \neq h_{\vec{t}}(\vec{x})] \\
&= \mathrm{e}_{\mathrm{unif}}(h_{\vec{s}}, h_{\vec{t}}).
\end{aligned}$$

Now, for the first of the inequalities above, we will use the (1) $\iff$ (3) portion of the claim, we will use our hypothesis that $\vec{t} \in [\alpha, 1]^d$ (which implies for each $i \in [d]$ that $[0, t_i] \subseteq [0, \alpha]$), we will use the hypothesis that $\varepsilon \leq \frac{1}{8d}$, and we will use Lemma 10.3.8. Utilizing these, we get the following:

$$\begin{aligned}
\mathrm{e}_{\mathrm{unif}}(h_{\vec{q}}, h_{\vec{t}}) &= \Pr_{\vec{x} \sim \mathrm{unif}(X)}[h_{\vec{q}}(\vec{x}) \neq h_{\vec{t}}(\vec{x})] \\
&= \Pr_{\vec{x} \sim \mathrm{unif}(X)}[x_{i_0} \in (\alpha, 1] \ \wedge \ \forall i \in [d] \setminus \{i_0\}, \, x_i \in [0, t_i]] \\
&= \Pr_{x_{i_0} \sim \mathrm{unif}([0,1])}[x_{i_0} \in (\alpha, 1]] \cdot \prod_{\substack{i=1 \\ i \neq i_0}}^{d} \Pr_{x \sim \mathrm{unif}([0,1])}[x \in [0, t_i]] \\
&\geq \Pr_{x_{i_0} \sim \mathrm{unif}([0,1])}[x_{i_0} \in (\alpha, 1]] \cdot \prod_{\substack{i=1 \\ i \neq i_0}}^{d} \Pr_{x \sim \mathrm{unif}([0,1])}[x \in [0, \alpha]] \\
&= (1 - \alpha) \cdot \alpha^{d-1} \\
&> \frac{1}{4d} \\
&\geq 2\varepsilon.
\end{aligned}$$

Thus, we get the desired two inequalities:

$$2\varepsilon < \mathrm{e}_{\mathrm{unif}}(h_{\vec{q}}, h_{\vec{t}}) \leq \mathrm{e}_{\mathrm{unif}}(h_{\vec{s}}, h_{\vec{t}}).$$

This nearly completes the proof. If there existed some point $\vec{r} \in X$ such that both $\mathrm{e}_{\mathrm{unif}}(h_{\vec{s}}, h_{\vec{r}}) \leq \varepsilon$ and $\mathrm{e}_{\mathrm{unif}}(h_{\vec{t}}, h_{\vec{r}}) \leq \varepsilon$, then it would follow from the triangle inequality of $\mathrm{e}_{\mathrm{unif}}$ that

$$\mathrm{e}_{\mathrm{unif}}(h_{\vec{s}}, h_{\vec{t}}) \leq \mathrm{e}_{\mathrm{unif}}(h_{\vec{s}}, h_{\vec{r}}) + \mathrm{e}_{\mathrm{unif}}(h_{\vec{t}}, h_{\vec{r}}) \leq 2\varepsilon$$

but this would contradict the above inequalities, so no such $\vec{r}$ exists. □

Theorem 10.3.10 (No $(d+1)$-Pseudoterministic Algorithm for the d-Threshold Estimation Problem)**.** *For $k < d+1$, there does not exist a k-pseudodeterministic algorithm for the d-*Threshold Estimation Problem *in the PAC model.*

The proof of this theorem will have similarities to the proof of Theorem 10.3.4.

Proof of Theorem 10.3.10. Fix any $d \in \mathbb{N}$, and choose ε and δ as $\varepsilon \leq \frac{1}{4d}$ and $\delta \leq \frac{1}{d+2}$. We will use the constant $\alpha = \frac{d-1}{d}$ and consider the cube $[\alpha, 1]^d$. We will also consider only the uniform distribution over X.

Suppose for contradiction that such an algorithm A does exists for some $k < d+1$. This means that for each possible threshold $\vec{t} \in [0,1]^d$, there exists some set $L_{\vec{t}} \subseteq \mathcal{H}$ of hypotheses with three properties: (1) each element of $L_{\vec{t}}$ is an ε-approximation to $h_{\vec{t}}$, (2) $|L_{\vec{t}}| \leq k$, and (3) with probability at least $1-\delta$, A returns an element of $L_{\vec{t}}$.

By the trivial averaging argument, this means that there exists at least one element in $L_{\vec{t}}$ which is returned by A with probability at least $\frac{1}{k} \cdot (1-\delta) \geq \frac{1}{k} \cdot (1 - \frac{1}{d+2}) = \frac{1}{k} \cdot \frac{d+1}{d+2} \geq \frac{1}{k} \cdot \frac{k+1}{k+2}$. Let $f \colon [\alpha, 1]^d \to [0,1]^d$ be a function which maps each threshold $\vec{t} \in [\alpha, 1]^d$ to such an element of $L_{\vec{t}}$. This is slightly different from the proof of Theorem 10.3.4 because we are defining the function f on only a very specific subset of the possible thresholds. The reason for this was alluded to in the discussion following the statement of Theorem 10.3.10.

Since $\frac{1}{k} \cdot \frac{k+1}{k+2} > \frac{1}{k+1}$, let η be such that $0 < \eta < \frac{1}{k} \cdot \frac{k+1}{k+2} - \frac{1}{k+1}$.

The function f induces a partition $\mathcal{P}$ of $[\alpha, 1]^d$ where the members of $\mathcal{P}$ are the fibers of f (i.e. $\mathcal{P} = \{f^{-1}(\vec{y}) \colon \vec{y} \in \text{range}(f)\}$). For any member $W \in \mathcal{P}$ and any coordinate $i \in [d]$, it cannot be that the set $\{w_i \colon \vec{w} \in W\}$ contains both values α and 1—if it did, then there would be two points $\vec{s}, \vec{t} \in W$ such that $s_i = \alpha$ and $t_i = 1$, but because they both belong to W, there is some $\vec{y} \in [0,1]^d$ such that $f(\vec{s}) = \vec{y} = f(\vec{t})$, but by definition of the partition, $h_{\vec{y}}$ would have to be an ε-approximation (in the PAC model) of both $h_{\vec{s}}$ and $h_{\vec{t}}$, but by Theorem 10.3.10, this is not possible.

Because no member of $\mathcal{P}$ contains points on opposite faces of $[\alpha, 1]^d$, by Theorem 6.1.2 (Infinite KKM/Lebesgue) there is some point $\vec{p} \in [\alpha, 1]^d$ such that for every radius $r > 0$, it holds that ${}^{\infty}\overline{B}_r(\vec{p})$ intersects at least $d+1$ members of $\mathcal{P}$.

Similar to Lemma 10.3.5 and how it is used in the proof of Theorem 10.3.4, we can use the following two facts. First, the function γ_1 defined by $\gamma_1(\vec{s}, \vec{t}) = \mathrm{e}_{\mathrm{unif}}(h_{\vec{s}}, h_{\vec{t}})$ is continuous (with respect to the ℓ_∞ norm on the domain). Second, the function $\gamma_2(h_{\vec{s}}, h_{\vec{t}}) = \mathrm{d}_{\mathrm{TV}}(\mathcal{D}_{A,\vec{s},n}, \mathcal{D}_{A,\vec{t},n})$ is continuous (with respect to the $\mathrm{e}_{\mathrm{unif}}$ notion of distance on the domain). A consequence is that the composition $\gamma_{12}(\vec{s}, \vec{t}) = \mathrm{d}_{\mathrm{TV}}(\mathcal{D}_{A,\vec{s},n}, \mathcal{D}_{A,\vec{t},n})$ is continuous. Thus, we can find some radius $r > 0$ such that if $\left\|\vec{t} - \vec{s}\right\|_\infty \leq r$, then $\mathrm{d}_{\mathrm{TV}}(\mathcal{D}_{A,\vec{s},n}, \mathcal{D}_{A,\vec{t},n}) \leq \eta$.

Now we get the same type of contradiction as in the proof of Theorem 10.3.4: for the special point $\vec{p}$ we have that $\mathcal{D}_{A,\vec{p},n}$ is a distribution that has $d+1 \geq k+1$ disjoint events that each have probability greater than $\frac{1}{k+1}$. Thus, no k-pseudodeterministic algorithm exists.

□

Appendix A

Measure Theory

Throughout this appendix, we use the word "countable" to mean finite or countably infinite.

Fact A.0.1. *If μ is a measure or an inner measure and $\mathcal{A}$ is a (possibly uncountable) family of pairwise disjoint measurable sets, then*

$$\mu(\bigsqcup_{A\in\mathcal{A}} A) \geq \sum_{A\in\mathcal{A}} \mu(A).$$

Proof. By definition of the arbitrary summation (c.f. [Fol99, p. 11]) we have

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

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

$$\mu\left(\bigsqcup_{A\in\mathcal{A}} A\right) \geq \mu\left(\bigsqcup_{A\in\mathcal{F}} A\right) \geq \sum_{A\in\mathcal{F}} \mu(A)$$

where the second inequality holds either due to the superadditivity or inner measures or is actually an equality due to the countable additivity of measures. Thus $\mu(\bigsqcup_{A\in\mathcal{A}} A)$ is an upper bound for the set $\left\{\sum_{A\in\mathcal{F}} \mu(A) : \mathcal{F} \subseteq \mathcal{A},\ \mathcal{F} \text{ finite}\right\}$ and so greater than or equal to the supremum. □

Fact A.0.2 (Interchange of Countable Sums with Non-negative Terms)**.** *If I, J are countable sets, and $a_{i,j} \geq 0$ for all $(i,j) \in I \times J$, then*

$$\sum_{i\in I}\sum_{j\in J} a_{i,j} = \sum_{j\in J}\sum_{i\in I} a_{i,j}$$

Proof. This can be proved directly via basic analysis methods if I and J are assumed to be $\mathbb{N}$ and the definition of the infinite sum as a limit of finite sums is used. Alternatively, viewing the summation as an integral over a countable measure space, this can be viewed as a corollary to Tonelli's theorem. □

A.1 Countable Partitions

In this section, we prove that partitions of $\mathbb{R}^d$ which consist entirely of members with positive Lebesgue measure have only countably many members.

Fact A.1.1. *If μ is a measure and $\mathcal{A}$ is a (possibly uncountable) family of pairwise disjoint measurable sets and $\mu(\bigsqcup_{A\in\mathcal{A}} A) < \infty$, then the set $\{A \in \mathcal{A} : \mu(A) > 0\}$ is countable.*

Proof. Let $\mathcal{B} = \{A \in \mathcal{A} : \mu(A) > 0\}$ denote the set in question, and let $\mathcal{B}_n = \left\{A \in \mathcal{A} : \mu(A) > \frac{1}{n}\right\}$ so that $\mathcal{B} = \bigcup_{n=1}^{\infty} \mathcal{B}_n$. Clearly each $\mathcal{B}_n$ is finite since (by application of Fact A.0.1) we have

$$\infty > \mu\left(\bigsqcup_{A\in\mathcal{A}} A\right) \geq \mu\left(\bigsqcup_{A\in\mathcal{B}_n} A\right) \geq \sum_{A\in\mathcal{B}_n} \mu(A) \geq \sum_{A\in\mathcal{B}_n} \frac{1}{n} = \frac{|\mathcal{B}_n|}{n}.$$

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

Fact A.1.2. *If $\mathcal{P}$ is a partition of $\mathbb{R}^d$, and m is the Lebesgue measure on $\mathbb{R}^d$, 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 $\mathcal{A}_n = \left\{X \cap B_n(\vec{0}) : X \in \mathcal{P},\ m(X \cap B_n(\vec{0})) > 0\right\}$ 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 A_n has the same cardinality as $\mathcal{P}_n = \left\{X \in \mathcal{P} : m(S \cap B_n(\vec{0})) > 0\right\}$ (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. □

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.

A.2 Isodiametric Inequality

In this section we provide a proof of the known isodiametric inequality. We found the outline of this particular proof in [use14] who cites [Gru07]. In general, in a metric space, if a set X has diameter D, then it is possible to construct a closed ball of diameter $2D$ (radius D) centered somewhere in space so that it contains X (in particular it can be centered at any point in X). Fact 3.4.9 (ℓ_∞ Diameter Ball) showed that for the ℓ_∞ norm specifically, we could reduce the radius from D to $\frac{D}{2}$. The isodiametric inequality shows that even though we cannot reduce D to $\frac{D}{2}$ for general norms, if we are not actually interested in containment but just comparison of volumes, then we can replace D with $\frac{D}{2}$.

Theorem A.2.1 (Isodiametric Inequality)**.** *Let $d \in \mathbb{N}$ and $\|\cdot\|$ be a norm on $\mathbb{R}^d$. Let $X \subseteq \mathbb{R}^d$ be a bounded set and $D = \operatorname{diam}_{\|\cdot\|}(X)$. Then the outer Lebesgue measure of X is at most the Lebesgue measure of the ball of diameter D. That is, (in three equivalent forms):*

$$\begin{aligned} m_{out}(A) &\leq m\left({}^{\|\cdot\|}B^{\circ}_{D/2}(\vec{0})\right) \\ &= \left(\tfrac{D}{2}\right)^d \cdot m\left({}^{\|\cdot\|}B^{\circ}_{1}(\vec{0})\right) \\ &= \left(\tfrac{D}{2}\right)^d \cdot v_{\|\cdot\|,d}. \end{aligned}$$

Proof. Throughout the proof, all diameters and balls are considered with respect to $\|\cdot\|$. Let $D = \operatorname{diam}(A)$. Then $\operatorname{diam}(\overline{A}) = D$ as well. Consider the set $\overline{A} - \overline{A} \stackrel{\text{def}}{=} \left\{\vec{a}' - \vec{a}'' : \vec{a}', \vec{a}'' \in \overline{A}\right\}$ noting that $\overline{A} - \overline{A}$ is closed, centrally symmetric[1], and has diameter[2] at most $2D$. This implies[3] that $\overline{A} - \overline{A} \subseteq \overline{B}_D(\vec{0})$. Thus, $m(\overline{A} - \overline{A}) \leq m(\overline{B}_D(\vec{0}))$.

Also, letting $-\overline{A} \stackrel{\text{def}}{=} \left\{-\vec{a} \colon \vec{a} \in \overline{A}\right\}$ we have that $\overline{A} - \overline{A}$ is the Minkowski sum $\overline{A} - \overline{A} = \overline{A} + (-\overline{A})$. This allows us to use the Generalized Brunn-Minkowski Inequality

[1]That is, for any $\vec{a} \in \overline{A} - \overline{A}$, also $-\vec{a} \in \overline{A} - \overline{A}$, because $\vec{a} = \vec{a}' - \vec{a}''$ for some $\vec{a}', \vec{a}'' \in \overline{A}$, so also $\overline{A} - \overline{A} \ni \vec{a}'' - \vec{a}' = -\vec{a}$.

[2]Given any two vectors $\vec{a}, \vec{b} \in \overline{A} - \overline{A}$ we have $\vec{a} = \vec{a}' - \vec{a}''$ and $\vec{b} = \vec{b}' - \vec{b}''$ for some $\vec{a}', \vec{a}'', \vec{b}', \vec{b} \in \overline{A}$. So $\left\|\vec{a} - \vec{b}\right\| = \left\|(\vec{a}' - \vec{a}'') - (\vec{b}' - \vec{b}'')\right\| = \left\|(\vec{a}' - \vec{b}') + (\vec{b}'' - \vec{a}'')\right\| \leq \left\|(\vec{a}' - \vec{b}')\right\| + \left\|(\vec{b}'' - \vec{a}'')\right\| \leq 2D$, so the diameter of $\overline{A} - \overline{A}$ is at most $2D$.

[3]If there was some $\vec{a} \in \overline{A} - \overline{A}$ with $\|\vec{a}\| > D$, then by central symmetry, also $-\vec{a} \in \overline{A} - \overline{A}$, so $2D = \operatorname{diam} \overline{A} - \overline{A} \geq \|\vec{a} - (-\vec{a})\| = 2\|\vec{a}\| > 2D$ which would be a contradiction.

(Theorem 7.1.4) to obtain

$$\begin{aligned} m(\overline{A}-\overline{A}) &= m(\overline{A}+(-\overline{A})) \\ &\geq \left[m(\overline{A})^{\frac{1}{d}}+m(-\overline{A})^{\frac{1}{d}}\right]^d && \text{(Brunn-Minkowski)} \\ &= \left[2\cdot m(\overline{A})^{\frac{1}{d}}\right]^d \\ &= 2^d\cdot m(\overline{A}) \end{aligned}$$

Combining this with the inequality at the end of the last paragraph gives $m(\overline{A}) \leq \frac{1}{2^d}\cdot m(\overline{B}_D(\vec{0}))$.

We complete the proof noting a few simple inequalities. First, since $A\subseteq\overline{A}$, we have $m_{out}(A)\leq m(\overline{A})$. Second, by the scaling of Lebesgue measure, $m(\overline{B}_D(\vec{0})) = m(D\cdot\overline{B}_1(\vec{0})) = D^d\cdot m(\overline{B}_1(\vec{0}))$. Third, $m(B_1^\circ(\vec{0})) = m(\overline{B}_1(\vec{0}))$ because for any $\varepsilon\in(0,\infty)$ we have

$$m(B_1^\circ(\vec{0}))\leq m(\overline{B}_1(\vec{0}))\leq m(B_{1+\varepsilon}^\circ(\vec{0})) = (1+\varepsilon)^d\cdot m(B_1^\circ(\vec{0})).$$

Combining all of this gives the result:

$$m_{out}(A)\leq m(\overline{A})\leq\left(\frac{D}{2}\right)^d\cdot m(\overline{B}_1(\vec{0})) = \left(\frac{D}{2}\right)^d\cdot m(B_1^\circ(\vec{0})).$$

□

A.3 Lower Bound Cover Number for $\mathbb{R}^d$

This section is dedicated to proving the following result.

Proposition A.3.1 (Lower Bound Cover Number for $\mathbb{R}^d$)**.** *Let $d \in \mathbb{N}$ and $S \subset \mathbb{R}^d$ be measurable with finite measure. Let $\mathcal{A}$ be a family of measurable subsets of S and let $k = \left\lceil \frac{\sum_{A \in \mathcal{A}} m(A)}{m(S)} \right\rceil$. If $k < \infty$, then there exists $\vec{p} \in S$ such that $\vec{p}$ belongs to at least k members of $\mathcal{A}$. If $k = \infty$, then for any integer n, there exists $\vec{p} \in S$ such that $\vec{p}$ belongs to at least n members of $\mathcal{A}$.*

Lemma A.3.2 (Exact Measure of Multiplicity)**.** *Let $n \in \mathbb{N}$. Let X be a measurable set in some measure space (the measure being denoted by μ) and $\mathcal{A}$ a countable family of measurable subsets of X such that for each $x \in X$, x belongs to* exactly *n members of $\mathcal{A}$. Then*

$$\sum_{A \in \mathcal{A}} \mu(A) = n \cdot \mu(X).$$

Proof. We note that if $n = 0$, then the statement is trivially true because $\mathcal{A}$ is either empty or contains just the empty set, and in either case $\sum_{A \in \mathcal{A}} \mu(A) = 0 = 0 \cdot \mu(X)$ if we use the standard convention that the empty sum is 0.

For any $\mathcal{F} \subseteq \mathcal{A}$, let $G_{\mathcal{F}} = \bigcap_{A \in \mathcal{F}} A$ noting that this is a countable intersection of measurable sets, so it is measurable (mnemonically, the G represents an intersection as it does in the notation for G_δ sets).

Let $\binom{\mathcal{A}}{n}$ denote all subsets of $\mathcal{A}$ of size n noting that because $\mathcal{A}$ is countable, so is $\binom{\mathcal{A}}{n}$. Observe that for distinct $\mathcal{F}, \mathcal{F}' \in \binom{\mathcal{A}}{n}$, the sets $G_{\mathcal{F}}$ and $G_{\mathcal{F}'}$ are disjoint because

$$G_{\mathcal{F}} \cap G_{\mathcal{F}'} = \left(\bigcap_{A \in \mathcal{F}} A \right) \cap \left(\bigcap_{A \in \mathcal{F}'} A \right) = \bigcap_{A \in \mathcal{F} \cup \mathcal{F}'} A$$

and since $\mathcal{F}$ and $\mathcal{F}'$ are distinct and each contain n items, $|\mathcal{F} \cup \mathcal{F}'| \geq n + 1$, and by assumption no point in X belongs to $n+1$ members, so $\bigcap_{A \in \mathcal{F} \cup \mathcal{F}'} A = \emptyset$. Furthermore, for each $x \in X$, since x belongs to exactly n members $A_1, \ldots, A_n$ of $\mathcal{A}$, taking $\mathcal{F} = \{A_1, \ldots, A_n\}$ we have $x \in G_{\mathcal{F}}$ which shows that $\left\{G_{\mathcal{F}} \colon \mathcal{F} \in \binom{\mathcal{A}}{n}\right\}$ is a partition of X into countably many measurable sets (allowing that some $G_{\mathcal{F}}$ might be empty).

The last observation we need is that for any $\mathcal{F} \in \binom{\mathcal{A}}{n}$ and any $A \in \mathcal{A}$, it holds that if $A \in \mathcal{F}$, then $A \supseteq G_{\mathcal{F}}$ and if $A \notin \mathcal{F}$ then $A \cap G_{\mathcal{F}} = \emptyset$. To see this, note that for any $x \in G_{\mathcal{F}}$, x belongs to each of the n members of $\mathcal{F} \subseteq \mathcal{A}$, and since by assumption x belongs to exactly n members of $\mathcal{A}$, it does not belong to any other members of $\mathcal{A}$.

Now we have the following chain of equalities:

$$
\begin{aligned}
\sum_{A \in \mathcal{A}} \mu(A) &= \sum_{A \in \mathcal{A}} \mu(A \cap X) && (A \subseteq X \text{ so } A \cap X = A) \\
&= \sum_{A \in \mathcal{A}} \mu\left(A \cap \left[\bigsqcup_{\mathcal{F} \in \binom{\mathcal{A}}{n}} G_{\mathcal{F}}\right]\right) && (\text{Set equality; the } G_{\mathcal{F}} \text{ partition } X) \\
&= \sum_{A \in \mathcal{A}} \mu\left(\bigsqcup_{\mathcal{F} \in \binom{\mathcal{A}}{n}} [A \cap G_{\mathcal{F}}]\right) && (\text{Set equality}) \\
&= \sum_{A \in \mathcal{A}} \left[\sum_{\mathcal{F} \in \binom{\mathcal{A}}{n}} \mu\left(A \cap G_{\mathcal{F}}\right)\right] && (\text{Countable additivity of measures}) \\
&= \sum_{\mathcal{F} \in \binom{\mathcal{A}}{n}} \left[\sum_{A \in \mathcal{A}} \mu\left(A \cap G_{\mathcal{F}}\right)\right] && (\text{Interchange sums by Fact A.0.2}) \\
&= \sum_{\mathcal{F} \in \binom{\mathcal{A}}{n}} \left[\sum_{A \in \mathcal{A}} \begin{cases} \mu\left(A \cap G_{\mathcal{F}}\right) = \mu\left(G_{\mathcal{F}}\right) & A \in \mathcal{F} \\ \mu\left(A \cap G_{\mathcal{F}}\right) = \mu(\emptyset) = 0 & A \notin \mathcal{F} \end{cases}\right] && (\text{Previous paragraph}) \\
&= \sum_{\mathcal{F} \in \binom{\mathcal{A}}{n}} \left[\sum_{A \in \mathcal{F}} \mu\left(G_{\mathcal{F}}\right)\right] && (\text{Remove 0 terms from summation}) \\
&= \sum_{\mathcal{F} \in \binom{\mathcal{A}}{n}} \left[n \cdot \mu\left(G_{\mathcal{F}}\right)\right] && (|\mathcal{F}| = n)
\end{aligned}
$$

$$= n \sum_{\mathcal{F} \in \binom{\mathcal{A}}{n}} \mu\left(G_{\mathcal{F}}\right) \qquad \text{(Linearity of summation)}$$

$$= n \cdot \mu \left(\bigsqcup_{\mathcal{F} \in \binom{\mathcal{A}}{n}} G_{\mathcal{F}} \right) \qquad \text{(Countable additivity of measures)}$$

$$= n \cdot \mu\left(X\right) \qquad \text{(Set equality; the } G_{\mathcal{F}} \text{ partition } X\text{)}$$

This proves the result. ☐

Lemma A.3.3 (Upper Bound Measure of Multiplicity)**.** *Let $n \in \mathbb{N}$. Let X be a measurable set in some measure space (the measure being denoted by μ) and $\mathcal{A}$ a countable family of measurable subsets of X such that for each $x \in X$, x belongs to* at most *n members of $\mathcal{A}$. Then*

$$\sum_{A \in \mathcal{A}} \mu(A) \leq n \cdot \mu(X).$$

Proof. As in the last proof, for any $\mathcal{F} \subseteq \mathcal{A}$, let $G_{\mathcal{F}} = \bigcap_{A \in \mathcal{F}} A$ noting that this is a countable intersection of measurable sets, so it is measurable (mnemonically, the G represents an intersection as it does in the notation for G_δ sets). And for any $k \in [n] \cup \{0\}$, let $\binom{\mathcal{A}}{k}$ denote all subsets of $\mathcal{A}$ of size k noting that because $\mathcal{A}$ is countable, so is $\binom{\mathcal{A}}{k}$.

For each $k \in [n] \cup \{0\}$, let

$$S_k = \{x \in X \colon x \text{ belongs to } \textit{exactly } k \text{ members of } \mathcal{A}\}$$

$$S'_k = \{x \in X \colon x \text{ belongs to } \textit{at least } k \text{ members of } \mathcal{A}\}$$

We will show that S_k and S'_k are measurable.

To show that the S'_k are measurable, note that for any $k \in [n] \cup \{0\}$, S'_k can be expressed as $S'_k = \bigcup_{\mathcal{F} \in \binom{\mathcal{A}}{k}} G_{\mathcal{F}}$. This is because for any $x \in X$, if x belongs to at least k members of $\mathcal{A}$, then there is a subset $\mathcal{F} \subseteq \mathcal{A}$ with $|\mathcal{F}| = k$ such that $x \in \bigcap_{A \in \mathcal{F}} A = G_{\mathcal{F}}$. Conversely, if $x \in \bigcap_{A \in \mathcal{F}} A = G_{\mathcal{F}}$, then there is some $\mathcal{F} \in \binom{\mathcal{A}}{k}$ (i.e. some $\mathcal{F} \subseteq \mathcal{A}$ with $|\mathcal{F}| = k$) such that $x \in G_{\mathcal{F}} = \bigcap_{A \in \mathcal{F}} A$, so x belongs to at least k members of $\mathcal{A}$. Thus, since $\mathcal{A}$ is countable, so is $\binom{\mathcal{A}}{k}$ (for each k) implying that each S'_k is a countable union of measurable sets, so is itself measurable.

To show the measurability of each S_k, first consider $k = n$. Observe that $S_n = S'_n$ because by assumption each $x \in X$ belongs to at most n members of $\mathcal{A}$, so it belongs to exactly n members if and only if it belongs to at least n members. Thus, S_n is also measurable.

Now for $k \in [n-1] \cup \{0\}$ observe that $S_k = S'_k \setminus S'_{k+1}$ because some $x \in X$ belongs to exactly k members of $\mathcal{A}$ if and only if it belongs to at least k members and does not belong to at least $k+1$ members of $\mathcal{A}$. Thus, for $k \in [n-1] \cup \{0\}$, S_k is the set difference of two measurable sets, so is itself measurable.

Finally, note that $\{S_k \colon k \in [n] \cup \{0\}\}$ is a partition of X (allowing the possibility that some S_k are empty) because every point of x belongs to some number of members of $\mathcal{A}$, and that number is (by assumption) between 0 and n inclusive.

Now we have the following chain of inequalities:

$$
\begin{aligned}
\sum_{A \in \mathcal{A}} \mu(A) &= \sum_{A \in \mathcal{A}} \mu(A \cap X) && (A \subseteq X \text{ so } A \cap X = A) \\
&= \sum_{A \in \mathcal{A}} \mu\left(A \cap \left[\bigsqcup_{k \in [n] \cup \{0\}} S_k\right]\right) && (\text{Set equality; the } S_k \text{ partition } X) \\
&= \sum_{A \in \mathcal{A}} \mu\left(\bigsqcup_{k \in [n] \cup \{0\}} [A \cap S_k]\right) && (\text{Set equality})
\end{aligned}
$$

$$= \sum_{A \in \mathcal{A}} \left[\sum_{k \in [n] \cup \{0\}} \mu\left(A \cap S_k\right) \right] \qquad \text{(Countable additivity of measures)}$$

$$= \sum_{k \in [n] \cup \{0\}} \left[\sum_{A \in \mathcal{A}} \mu\left(A \cap S_k\right) \right] \qquad \text{(Interchange sums by Fact A.0.2)}$$

$$= \sum_{k \in [n] \cup \{0\}} \left[k \cdot \mu\left(S_k\right)\right] \qquad \text{(By Lemma A.3.2; see details below)}$$

$$= \sum_{k \in [n]} \left[k \cdot \mu\left(S_k\right)\right] \qquad (k = 0 \text{ term is } 0)$$

$$\leq \sum_{k \in [n]} \left[n \cdot \mu\left(S_k\right)\right] \qquad (k \leq n)$$

$$= n \sum_{k \in [n]} \left[\mu\left(S_k\right)\right] \qquad \text{(Linearity of summation)}$$

$$= n \cdot \mu \left(\bigsqcup_{k \in [n]} S_k \right) \qquad \text{(Countable additivity of measures)}$$

$$\leq n \cdot \mu\left(X\right)$$

(Set inequality; the S_k partition X, but S_0 is missing from the union)

After justifying the use of Lemma A.3.2, this completes the proof. For each $k \in [n] \cup \{0\}$, let $X_k = S_k$ and $\mathcal{A}_k = \{A \cap S_k \colon A \in \mathcal{A}\}$. Then observe that for each $x \in X_k = S_k$, by the definition of S_k, x belongs to exactly k members of $\mathcal{A}$, and thus (since it also belongs to S_k) belongs to exactly k members of $\mathcal{A}_k$. Applying Lemma A.3.2 once for each k with $X = X_k$ and $\mathcal{A} = \mathcal{A}_k$ shows that

$$\sum_{A \in \mathcal{A}} \mu(A \cap S_k) = \sum_{A' \in \mathcal{A}_k} \mu(A') = k \cdot \mu(X_k) = k \cdot \mu(S_k)$$

(the middle equality is where Lemma A.3.2 was applied). This is what we claimed in the long chain of equalities above and completes the proof. □

Corollary A.3.4 (Lower Bound Cover Number)**.** *Let X be a measurable set in some measure space (the measure being denoted by μ) such that $0 < \mu(X) < \infty$. Let $\mathcal{A}$ be a countable family of measurable subsets of X such that $\sum_{A\in\mathcal{A}} \mu(A) < \infty$. Then there exists $x \in X$ such that x belongs to at least $\left\lceil \frac{\sum_{A\in\mathcal{A}} \mu(A)}{\mu(X)} \right\rceil$-many members of $\mathcal{A}$.*

Proof. First observe that by hypothesis, $\left\lceil \frac{\sum_{A\in\mathcal{A}} \mu(A)}{\mu(X)} \right\rceil$ is finite. Suppose for contradiction that each $x \in X$ belongs to strictly less than $\left\lceil \frac{\sum_{A\in\mathcal{A}} \mu(A)}{\mu(X)} \right\rceil$-many members of $\mathcal{A}$. Let $n = \left\lceil \frac{\sum_{A\in\mathcal{A}} \mu(A)}{\mu(X)} \right\rceil - 1$ (noting that $n < \frac{\sum_{A\in\mathcal{A}} \mu(A)}{\mu(X)}$). Then each $x \in X$ belongs to at most n-many members of A, so we have

$$
\begin{aligned}
\sum_{A\in\mathcal{A}} \mu(A) &\le n \cdot \mu(X) && \text{(Lemma A.3.3)} \\
&< \frac{\sum_{A\in\mathcal{A}} \mu(A)}{\mu(X)} \mu(X) && \left(0 < \mu(X) < \infty \text{ and } n < \tfrac{\sum_{A\in\mathcal{A}} \mu(A)}{\mu(X)}\right) \\
&= \sum_{A\in\mathcal{A}} \mu(A)
\end{aligned}
$$

which is a contradiction. □

Remark A.3.5. In Corollary A.3.4 above, it was important that we required $\sum_{A\in\mathcal{A}} \mu(A)$ to be finite. If we allowed it to be infinite, then the claim would have been that there was some $x \in X$ belonging to infinitely many members of $\mathcal{A}$, but this is in general not true (see Example A.3.6 below). Nonetheless, it is true (and a straightforward corollary of the above) that if $\sum_{A\in\mathcal{A}} \mu(A) = \infty$, then for any $n \in \mathbb{N}_0$, there exists a point $x_n \in X$ that is contained in at least n-many sets of $\mathcal{A}$. The distinction is that this point might have to depend on the choice of n. △

Example A.3.6 (Harmonic Cover of Open Unit Interval). Let $X = (0,1)$ be equipped with the Borel or Lebesgue measure μ. Let $\mathcal{A} = \left\{(0, \frac{1}{i}) : i \in \mathbb{N}\right\}$. Then $\sum_{A \in \mathcal{A}} \mu(A) = \sum_{i \in \mathbb{N}} \frac{1}{i} = \infty$. For any $n \in \mathbb{N}$, we can consider the point $x_n = \frac{1}{n+1}$ which is contained in $(0, \frac{1}{i})$ for $i \in [n]$ and not for any other i, so it belongs to exactly n sets in $\mathcal{A}$.

However, no point in X belongs to infinitely many sets in $\mathcal{A}$. To see this, consider an arbitrary point $x \in X = (0,1)$. Then for sufficiently large $i \in \mathbb{N}$, $x \notin (0, \frac{1}{i})$ so x belongs to only finitely many members of $\mathcal{A}$.

The prior three results have been stated in typical measure theory notation, but in the body of the paper we present Corollary A.3.4 as follows for $\mathbb{R}^d$ specifically with notation matching what is used elsewhere in the paper.

Proposition A.3.1 (Lower Bound Cover Number for $\mathbb{R}^d$). *Let $d \in \mathbb{N}$ and $S \subset \mathbb{R}^d$ be measurable with finite measure. Let $\mathcal{A}$ be a family of measurable subsets of S and let $k = \left\lceil \frac{\sum_{A \in \mathcal{A}} m(A)}{m(S)} \right\rceil$. If $k < \infty$, then there exists $\vec{p} \in S$ such that $\vec{p}$ belongs to at least k members of $\mathcal{A}$. If $k = \infty$, then for any integer n, there exists $\vec{p} \in S$ such that $\vec{p}$ belongs to at least n members of $\mathcal{A}$.*

Proof. This follows trivially from Corollary A.3.4 and Remark A.3.5. □

Appendix B

KKM, Lebesgue, and Sperner Results

B.1 Equivalencies

Lemma B.1.1 (KKM/Lebesgue $\implies$ KKM)**.** *The KKM/Lebesgue Theorem (Theorem 8.0.6) implies the Cubical KKM Lemma (Theorem 8.0.4).*

Proof. Let $\mathcal{C} = \{C_{\vec{v}}\}_{\vec{v}\in\{0,1\}^d}$ be a KKM cover of $[0,1]^d$. For each $\vec{x} \in [0,1]^d$, let $F_{\vec{x}}$ denote the smallest face of the cube containing $\vec{x}$ (i.e. $F_{\vec{x}}$ is the intersection of all faces containing $\vec{x}$). By the defining property of a KKM cover, we have $F_{\vec{x}} \subseteq \bigcup_{\vec{v}\in F_{\vec{x}}\cap\{0,1\}^d} C_{\vec{v}}$, so in particular there exists some $\vec{v} \in F_{\vec{x}} \cap \{0,1\}^d$ with $\vec{x} \in C_{\vec{v}}$. Define the function χ as follows where $\min_{\text{lex}}$ denotes the minimum element in a subset of $\{0,1\}^d$ under the lexicographic ordering:

$$\chi : [0,1]^d \to \{0,1\}^d$$
$$\chi(\vec{x}) = \min_{\text{lex}} \left\{ \vec{v} \in \{0,1\}^d \cap F_{\vec{x}} : \vec{x} \in C_{\vec{v}} \right\}$$

We have already demonstrated that the set in the definition is not empty, so χ is well-defined.

We claim that χ is a finite non-spanning coloring of $[0,1]^d$. The finiteness is trivial because the codomain of χ is finite, so we need only show it is a non-spanning coloring. Suppose $F^{(0)}$ and $F^{(1)}$ are opposite faces of the cube (i.e. there is some coordinate $j \in [d]$ such that $\pi_j(F^{(0)}) = \{0\}$ and $\pi_j(F^{(1)}) = \{1\}$) and let $\vec{x}^{(0)} \in F^{(0)}$ and $\vec{x}^{(1)} \in F^{(1)}$. Because $\pi_j(F^{(0)}) \cap \pi_j(F^{(1)}) = \emptyset$, it follows that $F^{(0)} \cap F^{(1)} = \emptyset$, so $F^{(0)}$ and $F^{(1)}$ are disjoint sets.

Because $\vec{x}^{(0)} \in F^{(0)}$ and $F_{\vec{x}^{(0)}}$ is by definition the intersection of all faces containing $\vec{x}$, we have $F_{\vec{x}^{(0)}} \subseteq F^{(0)}$ (and similarly replacing "0" with "1") so that also $F^{(0)}$ and $F^{(1)}$ are disjoint. By definition of χ we have $\chi(\vec{x}^{(0)}) \in F^{(0)}$ and $\chi(\vec{x}^{(1)}) \in F^{(1)}$ showing that $\chi(\vec{x}^{(0)}) \neq \chi(\vec{x}^{(1)})$, so χ is a non-spanning coloring.

By the KKM/Lebesgue Theorem, there exists $\vec{p} \in [0,1]^d$ such that $\left|\left\{\vec{v} \in \{0,1\}^d : \vec{p} \in \overline{\chi^{-1}(\vec{v})}\right\}\right| \geq d+1$. Fix such a $\vec{p}$ for the remainder of the proof. For each $\vec{v} \in \{0,1\}^d$, observe that $\chi^{-1}(\vec{v}) \subseteq C_{\vec{v}}$ because for any $\vec{x} \in \chi^{-1}(\vec{v})$ we have $\chi(\vec{x}) = \vec{v}$, so by definition of χ we have $\vec{x} \in C_{\vec{v}}$. Because closures maintain subset containment and because $C_{\vec{v}}$ is a closed set by hypothesis of the Cubical KKM Lemma, we have $\overline{\chi^{-1}(\vec{v})} \subseteq \overline{C_{\vec{v}}} = C_{\vec{v}}$. It then follows immediately that

$$\left\{\vec{v} \in \{0,1\}^d : \vec{p} \in \chi^{-1}(\vec{v})\right\} \subseteq \left\{\vec{v} \in \{0,1\}^d : \vec{p} \in C_{\vec{v}}\right\}$$

and since the former has cardinality at least $d+1$, so does the latter which proves the Cubical KKM Lemma. □

Lemma B.1.2 (KKM/Lebesgue $\Longrightarrow$ Lebesgue)**.** *The KKM/Lebesgue Theorem (Theorem 8.0.6) implies the Lebesgue Covering Theorem (Theorem 8.0.3).*

Proof. Let $N \in \mathbb{N}$ and $\mathcal{C} = \{C_n\}_{n \in [N]}$ be a Lebesgue cover of $[0,1]^d$. Because this is a cover, every point of $[0,1]^d$ belongs to some set, so define χ as follows:

$$\chi : [0,1]^d \to [N]$$
$$\chi(\vec{x}) = \min \{n \in [N] : \vec{x} \in C_n\}.$$

This is trivially a finite non-spanning coloring of $[0,1]^d$ because the codomain of χ is finite and for $\vec{x}^{(0)}$ and $\vec{x}^{(1)}$ on opposite faces, there is no $n \in [N]$ for which both $\vec{x}^{(0)} \in C_n$ and $\vec{x}^{(1)} \in C_n$ and thus $\left\{n \in [N] : \vec{x}^{(0)} \in C_n\right\}$ and $\left\{n \in [N] : \vec{x}^{(1)} \in C_n\right\}$ are disjoint, so $\chi(\vec{x}^{(0)}) \neq \chi(\vec{x}^{(1)})$.

By the KKM/Lebesgue Theorem, there exists $\vec{p} \in [0,1]^d$ such that $\left|\left\{n \in [N] : \vec{p} \in \overline{\chi^{-1}(n)}\right\}\right| \geq d+1$. Fix such a $\vec{p}$ for the remainder of the proof. For each $n \in [N]$, observe that $\chi^{-1}(n) \subseteq C_n$ because for any $\vec{x} \in \chi^{-1}(n)$ we have $\chi(\vec{x}) = n$, so by definition of χ we have $\vec{x} \in C_n$. Because closures maintain subset containment and because C_n is a closed set by hypothesis of the Lebesgue Covering Theorem, we have $\overline{\chi^{-1}(n)} \subseteq \overline{C_n} = C_n$. It then follows immediately that

$$\left\{n \in [N] : \vec{p} \in \overline{\chi^{-1}(n)}\right\} \subseteq \{n \in [N] : \vec{p} \in C_n\}$$

and since the former has cardinality at least $d+1$, so does the latter which proves the Lebesgue Covering Theorem. □

Lemma B.1.3 (Lebesgue $\Longrightarrow$ KKM/Lebesgue)**.** *The Lebesgue Covering Theorem (Theorem 8.0.3) implies the KKM/Lebesgue Theorem (Theorem 8.0.6).*

This proof is probably the trickiest of the four. In order to use the hypothesis of the Lebesgue Covering Theorem (Theorem 8.0.3), we can't just close the sets in a non-spanning coloring because the closures might intersect opposite faces. Thus, we first have to extend the coloring, and we do so as we do in the proof of the main result of the paper (the Neighborhood KKM/Lebesgue Theorem (Theorem 8.0.7)).

Proof. For this proof, we assume that the cube is $[-\frac{1}{2}, \frac{1}{2}]^d$ instead of $[0,1]^d$. Let C be a finite set and $\chi : [-\frac{1}{2}, \frac{1}{2}]^d \to C$ a finite non-spanning coloring of $[-\frac{1}{2}, \frac{1}{2}]^d$. Let $\varepsilon \in (0, \infty)$ be any fixed value throughout the entirety of the proof. Let f and $\gamma = \chi \circ f$ as in the proof of the Neighborhood KKM/Lebesgue Theorem (Theorem 8.0.7) so that $\gamma : [-\frac{1}{2} - \varepsilon, \frac{1}{2} + \varepsilon]^d \to C$ is an extension of the coloring χ to the larger cube $\gamma : [-\frac{1}{2} - \varepsilon, \frac{1}{2} + \varepsilon]^d$ (see also Figure 8.1a and Figure 8.1b) with the property that for each color $c \in C$, there exists an orientation $\vec{v}^{(c)} \in \{-1, 1\}^d$ so that the set $Y_c \stackrel{\text{def}}{=} \gamma^{-1}(c)$ of points of color c (according to γ) is a subset of $\prod_{i=1}^d v_i^{(c)} \cdot (-\frac{1}{2}, \frac{1}{2} + \varepsilon]$ (see Claim A in the proof of the Neighborhood KKM/Lebesgue Theorem (Theorem 8.0.7)). Because closures maintain subset containment we have

$$\overline{\gamma^{-1}(c)} = \overline{Y_c} \subseteq \overline{\prod_{i=1}^{d} v_i^{(c)} \cdot (-\tfrac{1}{2}, \tfrac{1}{2} + \varepsilon]} = \prod_{i=1}^{d} v_i^{(c)} \cdot [-\tfrac{1}{2}, \tfrac{1}{2} + \varepsilon].$$

This demonstrates that for each color $c \in C$, the set $\gamma^{-1}(c)$ of points given color c does not include points on opposite faces of the cube $[-\frac{1}{2} - \varepsilon, \frac{1}{2} + \varepsilon]^d$ (because if it did, then there would be some coordinate $j \in [d]$ such that $\pi_j(\gamma^{-1}(c)) \supseteq \{-\frac{1}{2} - \varepsilon, \frac{1}{2} + \varepsilon\}$, but the containment above shows this is not the case). Thus, $\mathcal{C} = \{\gamma^{-1}(c)\}_{c \in \mathcal{C}}$ is a Lebesgue cover of the cube $[-\frac{1}{2} - \varepsilon, \frac{1}{2} + \varepsilon]^d$ (we could rescale it to $[0,1]^d$ and re-index the cover with $[N]$ for $N = |C|$ to be really formal, but we won't).

By the Lebesgue Covering Theorem, there exists $\vec{q} \in [-\frac{1}{2} - \varepsilon, \frac{1}{2} + \varepsilon]^d$ such that $\left|\left\{c \in \mathcal{C} : \vec{q} \in \overline{\gamma^{-1}(c)}\right\}\right| \geq d + 1$. Fix such a $\vec{q}$ for the remainder of the proof and let $\vec{p} = f(\vec{q})$ (recalling that $f : [-\frac{1}{2} - \varepsilon, \frac{1}{2} + \varepsilon]^d \to [-\frac{1}{2}, \frac{1}{2} =^d$ is as in the proof of the Neighborhood KKM/Lebesgue Theorem (Theorem 8.0.7)).

Claim A. *For each $c \in C$, if $\vec{q} \in \overline{\gamma^{-1}(c)}$, then $\vec{p} \in \overline{\chi^{-1}(c)}$.*

Proof of Claim. Let $c \in C$ be arbitrary. If $\vec{q} \in \overline{\gamma^{-1}(c)}$, then there is a sequence $\langle \vec{q}^{(j)} \rangle_{j=1}^{\infty}$ of points in $\gamma^{-1}(c)$ converging to $\vec{q}$. Because for each $j \in \mathbb{N}$ we have $\vec{q}^{(j)} \in \gamma^{-1}(c)$, we have $\gamma(\vec{q}^{(j)}) = c$. Then, by the definition of γ, we have

$$c = \gamma(\vec{q}^{(j)}) = \chi(f(\vec{q}^{(j)}))$$

showing that $f(\vec{q}^{(j)}) \in \chi^{-1}(c)$. Since f is a continuous function[1], then $\langle f(\vec{q}^{(j)}) \rangle_{j=1}^{\infty}$ converges to $f(\vec{q}) = \vec{p}$ demonstrating that $\vec{p} \in \overline{\chi^{-1}(c)}$. ■

It then follows immediately that

$$\left\{ c \in C : \vec{p} \in \overline{\gamma^{-1}(c)} \right\} \subseteq \left\{ c \in C : \vec{p} \in \overline{\chi^{-1}(c)} \right\}$$

and since the former has cardinality at least $d+1$ (established prior to the claim), so does the latter which proves the KKM/Lebesgue Theorem. □

Lemma B.1.4 (KKM $\implies$ KKM/Lebesgue)**.** *The Cubical KKM Lemma (Theorem 8.0.4) implies the KKM/Lebesgue Theorem (Theorem 8.0.6).*

In the proof we essentially condense a non-spanning coloring to have codomain of cardinality 2^d—one color associated to each vertex—and then close each color set to apply the Cubical KKM Lemma (Theorem 8.0.4).

Proof. Let C be a finite set and $\chi : [0,1]^d \to C$ a finite non-spanning coloring. Because χ does not map points on opposite faces to the same color, this means for each color

[1] It is argued implicitly in Claim D in the proof of the Neighborhood KKM/Lebesgue Theorem (Theorem 8.0.7) that f is Lipschitz with Lipschitz constant 1. Alternatively, this could be analyzed directly.

$c \in C$ and coordinate $i \in [d]$ that the set of points given color c (i.e. $\chi^{-1}(c)$) does not contain a point with ith coordinate 0 and also a point with ith coordinate 1 (i.e. $\pi_i(\chi^{-1}(c)) \not\supseteq \{0,1\}$).

For each $i \in [d]$, define $f_i : C \to \{0,1\}$ by

$$f_i(c) = \begin{cases} 0 & 0 \in \pi_i(\chi^{-1}(c)) \\ 1 & 1 \in \pi_i(\chi^{-1}(c)) \\ 0 & \text{otherwise} \end{cases}.$$

The function f_i is well-defined because the first two cases are mutually exclusive. Then define $f : C \to \{0,1\}^d$ by $f(c) = \langle f_i(c) \rangle_{i=1}^d$, and define the (coloring) function $\zeta : [0,1]^d \to \{0,1\}^d$ as the composition $f \circ \chi$.

For each $\vec{v} \in \{0,1\}^d$, let $C_{\vec{v}} = \overline{\zeta^{-1}(\vec{v})}$, and let $\mathcal{C} = \{C_{\vec{v}}\}_{\vec{v} \in \{0,1\}^d}$. We claim that $\mathcal{C}$ is a KKM cover of $[0,1]^d$ which we prove by the following claim.

Claim A. *For each face F of $[0,1]^d$, we have $F \subseteq \bigcup_{\vec{v} \in F \cap \{0,1\}^d} C_{\vec{v}}$.*

Proof of Claim. Let F be an arbitrary face of $[0,1]^d$; this means that $F = \prod_{i=1}^d F_i$ where each F_i is either $\{0\}$, $\{1\}$, or $[0,1]$. Let $\vec{x} \in F$ be arbitrary noting that this implies for each coordinate $i \in [d]$ that $x_i \in F_i$. Let $c = \chi(\vec{x})$ (so $\vec{x} \in \chi^{-1}(c)$).

We first show for each coordinate $i \in [d]$ that $f_i(c) \in F_i$ and do so in three cases.

1. If $x_i = 0$, then $0 = x_i \in \pi_i(\chi^{-1}(c))$, so by definition of f_i we have $f_i(c) = 0$ showing that $f_i(c) = 0 = x_i \in F_i$.
2. If $x_i = 1$, then $1 = x_i \in \pi_i(\chi^{-1}(c))$, so by definition of f_i we have $f_i(c) = 1$ showing that $f_i(c) = 1 = x_i \in F_i$.
3. Otherwise $x_i \in (0,1)$, so because $x_i \in F_i$ we cannot have $F_i = \{0\}$ or $F_i = \{1\}$ and so it must be that $F_i = [0,1]$. Thus, $f_i(c) \in \{0,1\} \subseteq F_i$.

Now let $\vec{v}^{(0)} = \zeta(\vec{x})$ (so $\vec{x} \in \zeta^{-1}(\vec{v}^{(0)})$) and observe the following:

$$\vec{v}^{(0)} = \zeta(\vec{x}) = f(\chi(\vec{x})) = f(c) = \langle f_i(c) \rangle_{i=1}^{d} \in \prod_{i=1}^{d} F_i = F.$$

Thus $\vec{v}^{(0)} \in F$, and also vacuously $\vec{v}^{(0)} = \zeta(\vec{x}) \in \{0,1\}^d$. This shows that

$$\vec{x} \in \zeta^{-1}(\vec{v}^{(0)}) \subseteq \overline{\zeta^{-1}(\vec{v}^{(0)})} = C_{\vec{v}^{(0)}} \subseteq \bigcup_{\vec{v} \in F \cap \{0,1\}^d} C_{\vec{v}}.$$

Since $\vec{x} \in F$ was arbitrary, we have $F \subseteq \bigcup_{\vec{v} \in F \cap \{0,1\}^d} C_{\vec{v}}$ as claimed. ■

Because $\mathcal{C}$ is a KKM cover, by the Cubical KKM Lemma, there exists $\vec{p} \in [0,1]^d$ such that the set $V' \stackrel{\text{def}}{=} \left\{ \vec{v} \in \{0,1\}^d : \vec{p} \in C_{\vec{v}} \right\}$ has cardinality at least $d+1$. Fix such a $\vec{p}$ for the remainder of the proof.

Note that for each $\vec{v} \in V$, we have

$$\zeta^{-1}(\vec{v}) = (f \circ \chi)^{-1}(\vec{v}) = \chi^{-1}(f^{-1}(\vec{v})) = \bigcup_{c \in f^{-1}(\vec{v})} \chi^{-1}(c). \tag{B.1}$$

Now, for each $\vec{v} \in V'$, because $\vec{p}$ is in the closure of $\zeta^{-1}(\vec{v})$, any open set containing $\vec{p}$ intersects $\zeta^{-1}(\vec{v}) = \bigcup_{c \in f^{-1}(\vec{v})} \chi^{-1}(c)$ and thus intersects $\chi^{-1}(c)$ for some $c \in f^{-1}(\vec{v})$. Let $g(\vec{v})$ denote one such color[2].

Because $f^{-1}(\vec{v})$ and $f^{-1}(\vec{v}\,')$ are trivially disjoint for $\vec{v} \neq \vec{v}\,'$, this means $g(\vec{v})$ and $g(\vec{v}\,')$ are distinct colors so $g : V' \to C$ is an injection which means there are at least $d+1$ colors in C that are intersected by any open set containing $\vec{p}$.

Because $|C|$ is finite, then for each $\vec{v} \in V$, $f^{-1}(\vec{v}) \subseteq C$ is finite, then we can use the fact the closure of a finite union is equal to the finite union of the closures to

[2]We don't need the full axiom of choice here because C has finite cardinality.

extend this to

$$\begin{aligned}\vec{p} &\in \bigcap_{\vec{v}\in V'} \overline{\zeta^{-1}(\vec{v})} \\ &= \bigcap_{\vec{v}\in V'} \overline{\bigcup_{c\in f^{-1}(\vec{v})} \chi^{-1}(c)} \\ &= \bigcap_{\vec{v}\in V'} \bigcup_{c\in f^{-1}(\vec{v})} \overline{\chi^{-1}(c)} && (f^{-1}(\vec{v}) \text{ is finite})\end{aligned}$$

and thus, for each $\vec{v} \in V'$, $\vec{p}$ belongs to the closure of $\chi^{-1}(c)$ for some $c \in f^{-1}(\vec{v})$. By the same argument there are at least $d + 1$ such colors in C. That is, $\left|\left\{c \in C : \vec{p} \in \overline{\chi^{-1}(c)}\right\}\right| \geq d+1$ which proves the KKM/Lebesgue Theorem. □

B.2 Recovering the Cubical KKM Lemma and Lebesgue Covering Theorem

The following lemma is a bit informal and we provide only a sketch.

Lemma B.2.1. *Let $k^\circ : \mathbb{N} \times (0, \frac{1}{2}] \to \mathbb{N}$. If the Neighborhood KKM/Lebesgue Generic Statement (Statement 8.3.1) holds for k°, then this implies that the KKM/Lebesgue Theorem (Theorem 8.0.6) holds where "$d + 1$" is replaced by "$\liminf_{\varepsilon\to 0} k(d, \varepsilon)$."*

Proof Sketch. By the comment after the Neighborhood KKM/Lebesgue Generic Statement (Statement 8.3.1), we may assume k° is non-decreasing in ε for each d, so $\liminf_{\varepsilon\to 0} k^\circ(d, \varepsilon)$ is actually $\lim_{\varepsilon\to 0} k^\circ(d, \varepsilon)$ (the range of the function is bounded between 0 and 2^d so this limit exists). Let $\kappa_d = \lim_{\varepsilon\to 0} k(d, \varepsilon)$.

This implies that for every ε, there is a point $\vec{p}$ where ${}^{\infty}B^\circ_\varepsilon(\vec{p})$ intersects at least κ_d colors. Consider a sequence $\langle \varepsilon^{(n)} \rangle_{n=1}^{\infty}$ such that $\lim_{n\to\infty} \varepsilon^{(n)} = 0$, and for each n, let $\vec{p}^{(n)} \in [0,1]^d$ be a point for $\varepsilon^{(n)}$ such that ${}^{\infty}B^{\circ}_{\varepsilon^{(n)}}(\vec{p^{(n)}})$ intersects points of at least κ_d colors.

By compactness, fix a subsequence for which the $\vec{p}^{(n)}$ converge, and let $\vec{p}$ be the point of convergence. Then for any $\delta \in (0,\infty)$, it holds that ${}^{\infty}B^{\circ}_{\delta}(\vec{p})$ intersects at least κ_d colors because there is some $N \in \mathbb{N}$ such that for all $n \geq N$ we have $\left\|\vec{p} - \vec{p}^{(n)}\right\|_{\infty} < \frac{\delta}{2}$ and $\varepsilon^{(n)} < \frac{\delta}{2}$ so ${}^{\infty}B^{\circ}_{\delta}(\vec{p}) \supseteq {}^{\infty}B^{\circ}_{\varepsilon^{(n)}}(\vec{p}^{(n)})$ which intersects at least κ_d colors.

To prove the KKM/Lebesgue Theorem with "κ_d" instead of "$d+1$," we may assume by hypothesis that only finitely many colors are used. This implies that there is some $N' \in \mathbb{N}$ so that for $n, m \geq N'$, the set of colors intersected by ${}^{\infty}B^{\circ}_{\varepsilon^{(n)}}(\vec{p})$ is the same as the set of colors intersected by ${}^{\infty}B^{\circ}_{\varepsilon^{(m)}}(\vec{p})$ (which is a set of at least κ_d colors); denote this set of colors as C'. This shows that $\vec{p}$ is in the closure of each color in C'. $\square$

Appendix C

Minkowski Sums

Fact C.0.1. *For any normed vector space, given a set X and $\varepsilon \in (0, \infty)$, then*

$$X + B^\circ_\varepsilon(\vec{0}) = \bigcup_{\vec{x} \in X} B^\circ_\varepsilon(\vec{x}).$$

The same can be said replacing open balls with closed balls.

Proof. We show this only for the open balls. Switching all open balls in the proof with closed ones gives the proof for closed balls.

($\subseteq$) A generic element of $X + B^\circ_\varepsilon(\vec{0})$ is $\vec{x} + \vec{b}$ for some $\vec{x} \in X$ and $\vec{b} \in B^\circ_\varepsilon(\vec{0})$ which means $\vec{x} + \vec{b} \in \vec{x} + B^\circ_\varepsilon(\vec{0}) = B^\circ_\varepsilon(\vec{x}) \subseteq \bigcup_{\vec{x} \in X} B^\circ_\varepsilon(\vec{x})$.

($\supseteq$) Given $\vec{y} \in \bigcup_{\vec{x} \in X} B^\circ_\varepsilon(\vec{x})$ there is some particular $\vec{x} \in X$ such that $y \in B^\circ_\varepsilon(\vec{x}) = \vec{x} + B^\circ_\varepsilon(\vec{0})$ which means $\vec{y} = \vec{x} + \vec{b}$ for some $\vec{b} \in B^\circ_\varepsilon(\vec{0})$, and since $\vec{x} \in X$, we have $\vec{y} \in X + B^\circ_\varepsilon(\vec{0})$. $\square$

Fact C.0.2. *For any normed vector space, and any $\alpha, \beta \in (0, \infty)$, it holds that $B^\circ_\alpha(\vec{0}) + B^\circ_\beta(\vec{0}) = B^\circ_{\alpha+\beta}(\vec{0})$.*

Proof. ($\subseteq$) A generic element of $B^\circ_\alpha(\vec{0}) + B^\circ_\beta(\vec{0})$ is $\vec{a} + \vec{b}$ for $\vec{a} \in B^\circ_\alpha(\vec{0})$ and $\vec{b} \in B^\circ_\beta(\vec{0})$. Then $\|\vec{a}\| < \alpha$ and $\left\|\vec{b}\right\| < \beta$, so $\left\|\vec{a} + \vec{b}\right\| < \alpha + \beta$ showing $\vec{a} + \vec{b} \in B^\circ_{\alpha+\beta}(\vec{0})$.

($\supseteq$) Let $\vec{x} \in B^\circ_{\alpha+\beta}(\vec{0})$ which implies $\|\vec{x}\| < \alpha + \beta$. If $\|\vec{x}\| < \alpha$, then $\vec{x} \in B^\circ_\alpha(\vec{0})$ and $\vec{0} \in B^\circ_\beta(\vec{0})$, so $\vec{x} = \vec{x} + \vec{0} \in B^\circ_\alpha(\vec{0}) + B^\circ_\beta(\vec{0})$. Similarly, if $\|\vec{x}\| < \beta$, then $\vec{x} = \vec{0} + \vec{x} \in B^\circ_\alpha(\vec{0}) + B^\circ_\beta(\vec{0})$. In either case we would be done, so we may now assume that $\|\vec{x}\| \geq \alpha, \beta$. Let $\varepsilon = \alpha + \beta - \|\vec{x}\| \in (0, \infty)$. Since $\|\vec{x}\| \geq \alpha$, we have $\varepsilon \leq \beta$, and because $\|\vec{x}\| \geq \beta$, we have $\varepsilon \leq \alpha$. This shows $\frac{\varepsilon}{2} < \alpha, \beta$. Let $\vec{a} = (\alpha - \frac{\varepsilon}{2})\frac{\vec{x}}{\|\vec{x}\|}$ and $\vec{b} = (\beta - \frac{\varepsilon}{2})\frac{\vec{x}}{\|\vec{x}\|}$ noting that $\|\vec{a}\| = \alpha - \frac{\varepsilon}{2} \in (0, \alpha)$ and $\left\|\vec{b}\right\| = \beta - \frac{\varepsilon}{2} \in (0, \beta)$. Also, note that $\vec{a} + \vec{b} = (\alpha - \frac{\varepsilon}{2} + \beta - \frac{\varepsilon}{2})\frac{\vec{x}}{\|\vec{x}\|} = \|\vec{x}\|\frac{\vec{x}}{\|\vec{x}\|} = \vec{x}$ which shows $\vec{x} \in B^\circ_\alpha(\vec{0}) + B^\circ_\beta(\vec{0})$. $\square$

Fact C.0.3. *For any normed vector space, for any set X and vector $\vec{p}$, the following are equivalent:*

1. $\overline{B}_\varepsilon(\vec{p}) \cap X \neq \emptyset$
2. $\vec{p} \in X + \overline{B}_\varepsilon(\vec{0})$

The same can be said replacing both closed balls with open balls.

Proof. We show this only for the closed balls. Switching all closed balls in the proof with open ones gives the proof for open balls.

($\Longrightarrow$) If $\overline{B}_\varepsilon(\vec{p}) \cap X \neq \emptyset$, then there exists $\vec{y} \in \overline{B}_\varepsilon(\vec{p}) \cap X$, and since $\vec{y} \in \overline{B}_\varepsilon(\vec{p})$ we have $\|\vec{y} - \vec{p}\| \leq \varepsilon$ so $\vec{p} \in \overline{B}_\varepsilon(\vec{y}) = \vec{y} + \overline{B}_\varepsilon(\vec{0})$, and since $\vec{y} \in X$, $\vec{y} + \overline{B}_\varepsilon(\vec{0}) \subseteq X + \overline{B}_\varepsilon(\vec{0})$ showing that $\vec{p} \in X + \overline{B}_\varepsilon(\vec{0})$.

($\Longleftarrow$) If $\vec{p} \in X + \overline{B}_\varepsilon(\vec{0})$ then there exists $\vec{x} \in X$ and $\vec{b} \in \overline{B}_\varepsilon(\vec{0})$ such that $\vec{p} = \vec{x} + \vec{b}$. Thus $\vec{p} - \vec{x} = \vec{b}$, so $\|\vec{p} - \vec{x}\| = \|\vec{b}\| \leq \varepsilon$, so $\vec{x} \in \overline{B}_\varepsilon(\vec{p})$. Since $\vec{x}$ belongs to both X and $\overline{B}_\varepsilon(\vec{p})$, their intersection is non-empty. $\square$

Appendix D

Asymptotics

Fact D.0.1. *The following limit holds:*

$$\lim_{x\to\infty}(x+1)^{\frac{1}{x}} = 1$$

Lemma D.0.2. *The function* $f(x) = \frac{1}{2\left((x+1)^{\frac{x-1}{2x}}-1\right)}$ *is asymptotically equivalent to the function* $g(x) = \frac{1}{2\sqrt{x}}$ *(i.e.* $\lim_{x\to\infty}\frac{f(x)}{g(x)} = 1$*). Furthermore, for* $x \geq 3$*,* $f(x) \leq 4g(x) = \frac{2}{\sqrt{x}}$*.*

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

$x > 1$.

$$\left(\frac{x}{x+1}\right)^{1/2} = \frac{x^{1/2}}{(x+1)^{1/2}}$$
$$\leq \frac{x^{1/2}}{\left((x+1)^{1/2}\right)^{\frac{x-1}{x}}}$$

This is because for $x > 1$ we have $(x+1)^{1/2} \geq 1$ and $\frac{x-1}{x} \in (0,1)$, so $\left((x+1)^{1/2}\right)^{\frac{x-1}{x}} \leq (x+1)^{1/2}$. Thus the denominator in the above expression decreased from the previous line, but remains positive, so the fraction increases. Continuing, we have the following.

$$\leq \frac{x^{1/2}}{\left((x+1)^{1/2}\right)^{\frac{x-1}{x}} - 1} \qquad = \frac{f(x)}{g(x)}$$

This follows because again the denominator has decreased but remained positive. To see that the denominator is still positive for $x > 1$, note that $(x+1)$ is strictly increasing on $(0, \infty)$ and so is $\frac{x-1}{x}$. Also, on $[0, \infty)$, we have $(x+1)^{1/2} \geq 1$, and thus $\left((x+1)^{1/2}\right)^{\frac{x-1}{x}}$ is strictly increasing on $(0, \infty)$ and takes value 1 at $x = 1$. Thus, for $x > 1$, we have $\left((x+1)^{1/2}\right)^{\frac{x-1}{x}} > \left((x+1)^{1/2}\right)^{\frac{x-1}{x}} - 1 > 0$ so the denominator remains positive. We continue.

$$\leq \frac{x^{1/2}}{\left(x^{1/2}\right)^{\frac{x-1}{x}} - 1}$$

We again have that the denominator has decreased and remains positive. For $x > 1$, we have $x^{1/2}, (x+1)^{1/2} > 1$ and $x^{1/2} < (x+1)^{1/2}$ and $\frac{x-1}{x} \in (0,1)$, so $1 < \left(x^{1/2}\right)^{\frac{x-1}{x}} < \left((x+1)^{1/2}\right)^{\frac{x-1}{x}}$ showing the denominator has decreased but remained positive. Then we have the following equality.

$$= \frac{1}{x^{-\frac{1}{2x}} - \frac{1}{\sqrt{x}}}$$

The limit as $x \to \infty$ of the first and last expressions is 1 (using the limit stated in the beginning of this proof), so by the squeeze theorem, $\lim_{x\to\infty} \frac{f(x)}{g(x)} = 1$.

Furthermore, if we continue from the last expression no longer for the case $x > 1$, but for the case $x \geq 3$, then we have the following.

$$\begin{aligned}
\frac{f(x)}{g(x)} &\leq \frac{1}{x^{-\frac{1}{2x}} - \frac{1}{\sqrt{x}}} \\
&= \frac{1}{\frac{1}{\sqrt{x^{1/x}}} - \frac{1}{\sqrt{x}}} && \text{(Re-express)} \\
&\leq \frac{1}{\frac{1}{\sqrt{e^{1/e}}} - \frac{1}{\sqrt{x}}}
\end{aligned}$$

This is because $x^{1/x}$ is maximized at $x = e$ (which can be confirmed using standard calculus techniques) so the main denominator has decreased, and as long as $x > e^{1/e} \approx 1.44466786101$, then the denominator remains positive, and this is true because we are considering $x \geq 3$. We continue.

$$\leq \frac{1}{\frac{1}{\sqrt{e^{1/e}}} - \frac{1}{\sqrt{3}}}$$

This is because for $x \geq 3$, $\frac{1}{\sqrt{x}} < \frac{1}{\sqrt{3}}$ so again the main denominator has decreased and as before remains positive. We now evaluate.

$$\approx 3.9271793385$$
$$< 4$$

This demonstrates the "furthermore" claim and completes the proof. □

Proposition D.0.3 (Subexponential Functions)**.** *The following function classes are equivalent:*

$$2^{o(n)} = \bigcap_{c\in(1,\infty)} o(c^n) = \bigcap_{c\in(1,\infty)} O(c^n)$$

Proof. We first show the second equality. For any constant $c \in (1, \infty)$, we have $o(c^n) \subseteq O(c^n) \subseteq o((2c)^n)$, and since the intersections are over all constants including $c' = 2c$, the equality holds[1].

We now prove the superset ($\supseteq$) containment of the first equality. Let $f \in \bigcap_{c\in(1,\infty)} o(c^n)$. To show that $f \in 2^{o(n)}$ we must show (by definition) that $\log_2 \circ f \in o(n)$. That is, (by definition) we must show that for any $\varepsilon \in (0, \infty)$ there exists N_ε such that for all $n \geq N_\varepsilon$ it holds that $\log_2(f(n)) \leq \varepsilon n$. So let $\varepsilon \in (0, \infty)$ be arbitrary. Let $\delta = 1$ and $c = 2^\varepsilon$ noting that $c > 1$ because $\varepsilon > 0$, so $f \in o(c^n)$. Then, by definition, there exists some N_δ such that for all $n \geq N_\delta$ it holds that $f(n) \leq \delta c^n$.

[1] To see that $O(c^n) \subseteq o((2c)^n)$, consider $f \in O(c^n)$. Thus, $\lim_{n\to\infty} \frac{f(n)}{c^n}$ is equal to some constant C, and thus we can show that $f \in o((2c)^n)$ by showing that $\lim_{n\to\infty} \frac{f(n)}{(2c)^n} = 0$ which is true because $\lim_{n\to\infty} \frac{f(n)}{(2c)^n} = \left[\lim_{n\to\infty} \frac{f(n)}{c^n}\right] \cdot \left[\lim_{n\to\infty} \frac{c^n}{(2c)^n}\right] = C \cdot 0 = 0$.

Let $N_\varepsilon = N_\delta$. This implies for all $n \geq N_\varepsilon = N_\delta$ that

$$\begin{aligned}
\log_2(f(n)) &\leq \log_2(\delta c^n) \\
&= \log_2(c^n) && (\delta = 1) \\
&= \log_2((2^\varepsilon)^n) && (c = 2^\varepsilon) \\
&= \varepsilon n
\end{aligned}$$

This shows that $f \in 2^{o(n)}$.

Lastly, we prove the subset ($\subseteq$) containment of the first equality. Let $f \in 2^{o(n)}$ be arbitrary. To show that $f \in \bigcap_{c \in (1,\infty)} o(c^n)$, let $c \in (1, \infty)$ be arbitrary and we will show that $f \in o(c^n)$ by the definition. Let $\varepsilon \in (0, \infty)$ be arbitrary, and let $\delta = \frac{\log_2(c)}{2}$ (noting that since $c > 1$, $\delta > 0$. Since $f \in 2^{o(n)}$, we have $\log_2 \circ f \in o(n)$, so by definition there exists some N_δ such that for all $n \geq N_\delta$ it holds that $\log_2(f(n)) \leq \delta n$. Let $N_\varepsilon = \max(N_\delta, \left\lceil \frac{-2\log_2(\varepsilon)}{\log_2(c)} \right\rceil)$. We have that for all $n \geq N_\varepsilon$ that

$$\begin{aligned}
\log_2(f(n)) &\leq \delta n && (n \geq N_\varepsilon \geq N_\delta) \\
&= \frac{n \log_2(c)}{2} \\
&\leq \log_2(\varepsilon) + n \log_2(c) && \left(n \geq N_\varepsilon \geq \left\lceil \frac{-2\log_2(\varepsilon)}{\log_2(c)} \right\rceil\right) \\
&= \log_2(\varepsilon c^n)
\end{aligned}$$

By monotonicity of $\log_2$, this implies for $n \geq N_\varepsilon$ that $f(n) \leq \varepsilon c^n$. Since ε was arbitrary, this shows $f \in o(c^n)$, and since c was also arbitrary, this shows $f \in \bigcap_{c \in (1,\infty)} o(c^n)$. □

Fact D.0.4. *If $f : \mathbb{N} \to \mathbb{R}$ has $\lim_{n \to \infty} f(n) = \ell$ for some $\ell \in (-1, 1)$, then $\lim_{n \to \infty} f(n)^n = 0$.*

Proof. We show this by the definition of limit, so let $\varepsilon \in (0, \infty)$ be arbitrary. Pick some $\gamma \in (|\ell|, 1)$ noting that $\ell \in (-\gamma, \gamma) \subsetneq (-1, 1)$. Since $\gamma \in (0, 1)$, $\lim_{n\to\infty} \gamma^n = 0$, so let N_1 be such that for $n \geq N_1$ it holds that $\gamma^n < \varepsilon$. And since $f(n)$ converges to $\ell \in (-\gamma, \gamma)$, let N_2 be such that for $n \geq N_2$ we have $|f(n)| < \gamma$. Thus, for $n \geq N \stackrel{\text{def}}{=} \max\{N_1, N_2\}$ we have

$$|f(n)^n| = |f(n)|^n \leq \gamma^n < \varepsilon$$

so $\lim_{n\to\infty} f(n)^n = 0$. □

Lemma D.0.5. *If $g : \mathbb{N} \to [0, \infty)$ is such that $\lim_{n\to\infty} g(n) = 1$, then for $h : \mathbb{N} \to \mathbb{R}$ defined by $h(n) = g(n)^n$ we have $h \in 2^{o(n)}$.*

Proof. Using Proposition D.0.3, we will show for all $c \in (1, \infty)$ that $h \in o(c^n)$. Let $c \in (1, \infty)$ be arbitrary for this purpose. Then

$$\frac{h(n)}{c^n} = \left(\frac{g(n)}{c}\right)^n$$

and since $g(n)$ converges to 1 and $c > 1$, $\frac{g(n)}{c}$ converges to $\frac{1}{c} \in (0, 1)$, so by Fact D.0.4, $\left(\frac{g(n)}{c}\right)^n$ converges to 0. Thus $\lim_{n\to\infty} \frac{h(n)}{c^n} = 0$, so $h \in o(c^n)$.

□

Lemma D.0.6 (Asymptotic Logs)**.** *If $f, g : \mathbb{N} \to (0, \infty)$ such that $\ln(f(n)) \in o(\ln(g(n))$ and either $\lim_{n\to\infty} g(n) = \infty$ or $\lim_{n\to\infty} g(n) < 1$, then $f(n) \in o(g(n))$.*

Proof. For any $C \in (0, \infty)$, there exists $N_C \in \mathbb{N}$ such that for $n \geq N_C$ we have $\ln(f(n)) \leq C \ln(g(n)) = (\ln(g(n)^C))$, so exponentiating, we have for $n \geq N_C$ that $f(n) \leq g(n)^C$.

For the first case, if $\lim_{n\to\infty} g(n) = \infty$ then take $C = \frac{1}{2}$ above. Let $\delta \in (0, \infty)$ be arbitrary, and let $N_\delta \in \mathbb{N}$ be such that for $n \geq N_\delta$, $g(n) \geq \frac{1}{\delta^2}$. Let $N = \max(N_C, N_\delta)$. Then for $n \geq N$ we have

$$\begin{aligned} f(n) \leq g(n)^{\frac{1}{2}} &= \frac{1}{\sqrt{g(n)}} \cdot g(n) && (C = \tfrac{1}{2} \text{ and } n \geq C) \\ &\leq \frac{\sqrt{\frac{1}{\delta^2}}}{g}(n) = \delta g(n) && (n \geq N_\delta) \end{aligned}$$

which shows $f(n) \in o(g(n))$.

For the second case, if $\lim_{n\to\infty} g(n) < 1$, then let ℓ denote this limit. Let $\delta \in (0, \infty)$ be arbitrary. Take C to be such that $\ell^{\frac{C-1}{2}} < \delta$. Since $\ell \in (0, 1)$, $\sqrt{(\ell)} > \ell$. Let $N_\delta \in \mathbb{N}$ be such that for $n \geq N_\delta$, $g(n) \leq \sqrt{(\ell)}$. Let $N = \max(N_C, N_\delta)$. Then for $n \geq N$ we have

$$\begin{aligned} f(n) \leq g(n)^C &= g(n)^{C-1} \cdot g(n) && (n \geq N_C) \\ &\leq \sqrt{\ell}^{C-1} \cdot g(n) && (n \geq N_\delta \text{ and a positive power is an increasing function}) \\ &\leq \delta g(n) && (\text{Choice of } C) \end{aligned}$$

which shows $f(n) \in o(g(n))$.

□

Remark D.0.7. If the two limit conditions on g above are removed, the conclusion may not hold. For example, let $C_f \in (0, 1)$ and $C_g \in [1, \infty)$ and suppose that $f, g : \mathbb{N} \to (0, \infty)$ are such that $\lim_{n\to\infty} f(n) = C_f$ and $\lim_{n\to\infty} g(n) = C_g$. Then by continuity, $\lim_{n\to\infty} \ln(f(n)) = \ln(C_f) < 0$ (since $C_f < 1$) and for any $\delta \in (0, \infty)$, $\lim_{n\to\infty} \delta \ln(g(n)) = \delta \ln(C_g) \geq 0$ (since $C_g \geq 1$). Thus, for sufficiently large $n \in \mathbb{N}$, $\ln(f(n)) \leq \delta \ln(g(n))$ showing that $\ln(f(n)) \in o(\ln(g(n))$. However, as both f and g

have finite limits, $f(n) \in \Theta_n(1)$ and $g(n) \in \Theta_n(1)$ so $f(n) \notin o(g(n))$. △

Appendix E

Binary Relations

Throughout this appendix, 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 properties 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 $\bigcap_{S\in\mathcal{S}} S^{-1} = \left(\bigcap_{S\in\mathcal{S}} S\right)^{-1}$.

Fact E.0.1. *If R is a binary relation, then $(R^{-1})^t = (R^t)^{-1}$.*

Proof. We have the following chain of equalities:

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

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

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

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

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

This completes the proof. □

Fact E.0.2. *If R is a reflexive and symmetric binary 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 implies 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. □

Fact E.0.3. *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 $y \in [x]_S$. □

Still letting R denote a binary relation on X, let $R^{t_0} = R$, and inductively for all $n \in \mathbb{N}$, let

$$R^{t_n} = R^{t_{n-1}} \cup \left\{ (x, y) \in X^2 : \exists z \in X \text{ with } (x, z) \in R^{t_{n-1}} \text{ and } (z, y) \in R^{t_{n-1}} \right\}.$$

Fact E.0.4. *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

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

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

Fact E.0.5. *Let $a, b \in X$. 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 inductive hypothesis, 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 inductive hypothesis, 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 $z_0 = x_0 = a$ and $z_{k=k'+k''} = y_{k''} = b$.

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 $\langle y_i \rangle_{i=k'}^{k'+k''=k}$ demonstrates that $(x_{k'}, x_k) \in R^{t_{n-1}}$ and thus $(a,b) = (x_0, x_k) \in R^{t_n}$. □

Fact E.0.6. *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 $(a,b) \in R^{t_n} \subseteq R^t$. □

Fact E.0.7. *For any $a, b \in X$, $(a,b) \in R^t$ if and only if there exists some finite sequence $\langle x_i \rangle_{i=0}^{N}$ of distinct terms with $x_0 = a$, and $x_N = b$, and for all $i \in [N]$ $(x_{i-1}, x_i) \in R$.*

Proof. Let $a, b \in X$. By Fact E.0.6, it suffices to prove that if there exists a finite sequence $\langle x_i \rangle_{i=0}^{N}$ with $x_0 = a$, and $x_N = b$, and for all $i \in [N]$ that $(x_{i-1}, x_i) \in R$, then such a sequence exists with distinct terms.

This is true, because given any sequence, a sequence with distinct terms can be obtained by "removing loops". Formally, let S be the set of lengths of such sequences:

$$S = \left\{ N \in \mathbb{N} : \text{exists } \langle x_i \rangle_{i=0}^{N} \text{ with } x_0 = a,\ x_N = b \text{ and for all } i \in [N], (x_{i-1}, x_i) \in R \right\}.$$

By Fact E.0.6, S is not empty, so let $n = \min(S)$, and $\langle x_i \rangle_{i=0}^{n}$ any sequence with $x_0 = a$, and $x_n = b$, and for all $i \in [n]$ that $(x_{i-1}, x_i) \in R$. Then this sequence must contain distinct terms because otherwise the following contradiction arises.

Suppose the terms are not distinct, and so there is some $i_0 \in [n-1] \cup \{0\}$ and $i_1 \in [n]$ with $i_0 < i_1$ such that $x_{i_0} = x_{i_1}$. Let $\langle z_j \rangle_{j=0}^{n-(i_1-i_0)}$ be defined by

$$z_j = \begin{cases} x_j & 0 \le j \le i_0 \\ x_{j+(i_1-i_0)} & i_0 < j \le n-(i_1-i_0) \end{cases}.$$

Then $\langle z_j \rangle_{j=0}^{n-(i_1-i_0)}$ is such a sequence[1] of length $n-(i_1-i_0) < n$ which contradicts that $n = \min(S)$.

□

[1] (1) $z_0 = x_0 = a$. (2) $z_{n-(i_1-i_0)} = x_n = b$ (if $i_1 = n$ then $n-(i_1-i_0) = i_0$ so $z_{n-(i_1-i_0)} = x_{n-(i_1-i_0)} = x_{i_1} = x_{i_0} = x_n = b$ and if $i_1 < n$ then $n-(i_1-i_0) > i_0$ so $z_{n-(i_1-i_0)} = x_{(n-(i_1-i_0))+(i_1-i_0)} = x_n = b$). (3) for $j \in [n-(i_1-i_0)]$, $(z_{j-1}, z_j) \in R$ which can be seen in three cases: (i) If $0 < j \le i_0$, then $0 \le j-1 \le i_0$ so $(z_{j-1}, z_j) = (x_{j-1}, x_j) \in R$. (ii) If $j = i_0+1$, then $j-1 = i_0$ so $(z_{j-1}, z_j) = (z_{i_0}, z_{i_0+1}) = (x_{i_0}, x_{i_1+1}) = (x_{i_1}, x_{i_1+1}) \in R$. (iii) Otherwise $i_0+1 < j$ so $i_0 < j-1$ so $(z_{j-1}, z_j) = (x_{i_1+j-1-i_0}, x_{i_1+j-i_0}) \in R$.

Appendix F

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.

F.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[1].

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

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 $\lfloor x\rfloor_\alpha \stackrel{\text{def}}{=} \alpha\lfloor x/\alpha\rfloor$. 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 $\lfloor x\rfloor_{\alpha,\beta} \stackrel{\text{def}}{=} \lfloor x - \beta\rfloor_\alpha + \beta = \alpha\lfloor(x-\beta)/\alpha)\rfloor$. The partition induced from this rounding scheme is $\mathcal{P}_{\lfloor\cdot\rfloor_{\alpha,\beta}} = \{[\alpha n + \beta, \alpha(n+1) + \beta) : n \in \mathbb{Z}\}$ [2]. 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

[2] 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').

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 $\lfloor x\rfloor_{\alpha,\beta,\gamma} \stackrel{\text{def}}{=} \lfloor x\rfloor_{\alpha,\beta} + \gamma = \alpha\lfloor(x-\beta)/\alpha)\rfloor + \gamma$. 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 functions $\left\{\lfloor\cdot\rfloor_{\alpha,\beta} : \beta \in [0, \alpha)\right\}$ with distribution corresponding to β being distributed uniformly over $[0, \alpha)$. This gives a randomized rounding scheme with the following property: For any $\varepsilon \in [0, \alpha)$, for any $x, y \in \mathbb{R}$ with $|x - y| \leq \varepsilon$, 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 $|x - y|/\alpha \leq \varepsilon/\alpha$.

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 $\lfloor \cdot \rceil_{\alpha,\vec{\beta},\vec{\gamma}} : \mathbb{R}^d \to \mathbb{R}^d$ coordinatewise in the expected way: $\lfloor \vec{x} \rceil_{\alpha,\vec{\beta},\vec{\gamma}} \stackrel{\text{def}}{=} \langle \lfloor x_i \rceil_{\alpha,\beta_i,\gamma_i} \rangle_{i=1}^d$. If $\vec{\gamma}$ is not specified, it will be assumed to be $\vec{0}$. The partition induced by this scheme is $\mathcal{P}_{\lfloor \cdot \rceil_{\alpha,\vec{\beta},\vec{\gamma}}} = \left\{ \vec{\gamma} + [\alpha n, \alpha(n+1))^d : n \in \mathbb{Z} \right\}$. In other words, the partition induced by rounding each coordinate is a grid of unit hypercubes with some shift applied to the grid.

F.2 The Randomized Rounding Scheme of Saks and Zhou

The rounding scheme used by Saks and Zhou is the basic rounding scheme just introduced[3]. We briefly state the parameters of their scheme.

Let $d \in \mathbb{N}$ and $t = O(\log d)$ and $D = O(\log d)$. Let $\varepsilon = d \cdot 2^{-t} \cdot 2^{-D}$. 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 $\vec{m} \in B_\varepsilon^\circ$ (w.r.t. the ℓ_∞ norm) that $\lfloor \vec{n} \rceil_{\alpha,\vec{\beta}} = \lfloor \vec{m} \rceil_{\alpha,\vec{\beta}}$. 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 $\lfloor n_i \rceil_{\alpha,\beta} \neq \lfloor m_i \rceil_{\alpha,\beta}$ 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.

[3]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.

This norm can be defined in either of these two well-known equivalent ways. Let M be an $d \times d$ matrix:

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

or

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

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 $\varepsilon' = \cdot 2^{-t} \cdot 2^{-D} = \varepsilon/d$ 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).

F.3 The Randomized Rounding Scheme of Goldreich

In [Gol19b, Algorithm 9], Goldreich uses the basic rounding scheme discussed above as well [4]. However, unlike Saks and Zhou, Goldreich's goals in using the partition are very relevant to our work in this paper. For an arbitrary $\varepsilon \in (0, \infty)$ let $\alpha = \varepsilon$. 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

[4] 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.

applies the function $\lfloor \cdot \rfloor_{\alpha,\vec{\beta}} : \mathbb{R}^d \to \mathbb{R}^d$.

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

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

where the ε-ball is with respect to the ℓ_∞ norm. 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 4.2, 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.

F.4 The Deterministic Rounding Scheme of Hoza and Klivans

In [HK18, Section 2], Hoza and Klivans have the same goal as Goldriech—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 $\varepsilon \in (0, \infty)$ and $d \in \mathbb{N}$. Let $\mathcal{I}$ be a partition of $\mathbb{R}$ by intervals of length $2\varepsilon(d+1)$ which are closed on the left and open on the right. Fix an arbitrary point $x \in \mathbb{R}$ and consider the interval $[x-\varepsilon, x+\varepsilon] = x + [-\varepsilon, \varepsilon]$. Because this interval has length 2ε (and is closed) and every interval in $\mathcal{I}$ has length $2\varepsilon(d+1)$ (and is half open), it follows that there is exactly one value $\Delta \in [d+1]$ such that the interval $[x+(2\varepsilon\Delta)-\varepsilon, x+(2\varepsilon\Delta)+\varepsilon] = x+2\varepsilon\Delta+[-\varepsilon, \varepsilon]$ 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) [5].

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

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

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

which is a closed interval of length $2\varepsilon\Delta$ and thus intersects exactly two intervals of $\mathcal{I}$, say $I_{left} = [\alpha - 2\varepsilon(d+1), \alpha)$ and $I_{right} = [\alpha, \alpha+2\varepsilon(d+1))$. The point α is either contained in the interior or right boundary of $[x+(2\varepsilon\Delta_0)-\varepsilon, x+(2\varepsilon\Delta_0)+\varepsilon]$ for some Δ_0 (if not, then $\bigcup_{\Delta\in[d+1]}[x+(2\varepsilon\Delta)-\varepsilon, x+(2\varepsilon\Delta)+\varepsilon]$ would not intersect I_{left}). This is the unique Δ_0 such that the interval intersects both I_{left} and I_{right}.

intervals of $\mathcal{I}$. That is,

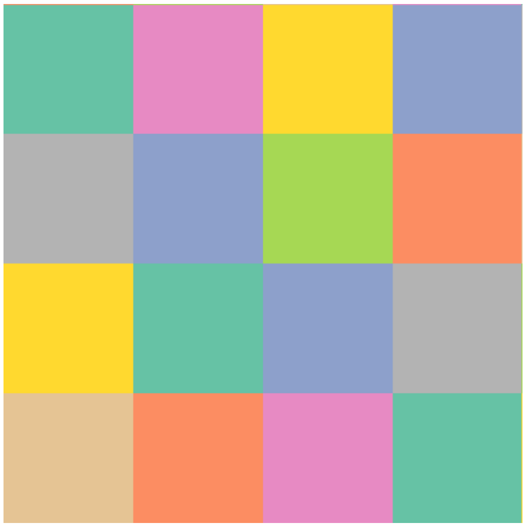
$$\mathcal{G} \stackrel{\text{def}}{=} \left\{ \prod_{i=1}^{d} I^{(i)} : I^{(1)}, \ldots, I^{(d)} \in \mathcal{I} \right\}.$$

Each member of $\mathcal{G}$ is a hypercube, and up to translation, the set of centers of these hypercubes is $2\varepsilon\mathbb{Z}^d$ (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 $\Lambda = \left\{ (2\varepsilon\Delta) \cdot \vec{1} : \Delta \in [d+1] \right\}$. We claim that for any point $\vec{x} \in \mathbb{R}^d$, there exists at least one $\vec{\lambda} \in \Lambda$ such that $\overline{B}_{\varepsilon}(\vec{x} + \vec{\lambda})$ 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 $\overline{B}_{\varepsilon}(\vec{x}+\vec{\lambda})$ (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 $[x_i + \lambda_i - \varepsilon, x_i + \lambda_i + \varepsilon] = [x_i + (2\varepsilon\Delta) - \varepsilon, x_i + (2\varepsilon\Delta) + \varepsilon]$ 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 $i \in [d]$.

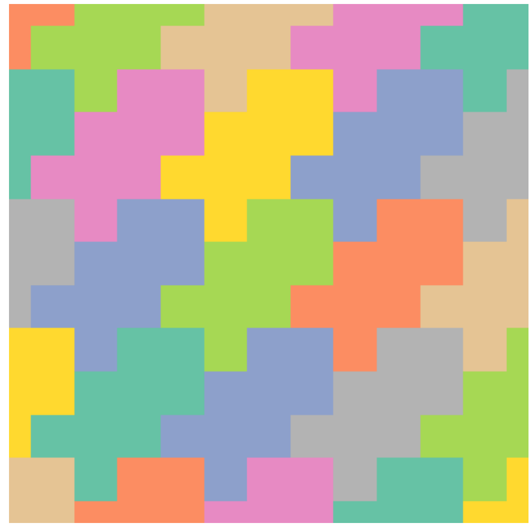
With these properties established, let $s : \mathbb{R}^d \to \Lambda$ be a function mapping each point $\vec{x}$ to one of the $\vec{\lambda} \in \Lambda$ that has the containment property above (e.g. take the smallest length $\vec{\lambda}$ that works). Also, define the representative function $\text{rep} : \mathcal{G} \to \mathbb{R}^d$ 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 F.1b.



(a) The grid partition $\mathcal{G}$ for $\mathbb{R}^2$.



(b) The Hoza-Klivans partition of $\mathbb{R}^2$.

Figure F.1: The partitions described in Section F.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.

Proposition F.4.1. *The partition induced by the rounding scheme described above has the following properties:*

- *Each member has diameter at most* $6\varepsilon(d+1)$ *(with respect to the* ℓ_∞ *norm)*
- *The partition is* $(d+1,\varepsilon)$*-secluded*

If this partition is scaled by a factor of $6\varepsilon(d+1)$, 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.

Corollary F.4.2 (Hoza-Klivans Partition)**.** *Let* $d \in \mathbb{N}$*. Then there exists a* $(d+1, \frac{1}{6(d+1)})$*-secluded partition* $\mathcal{P}$ *of* $\mathbb{R}^d$ *for which every member has* ℓ_∞ *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 interpreted as indicting $\mathcal{G}$.

Lemma F.4.3. *Let $d \in \mathbb{N}$ and $\varepsilon, D \in (0, \infty)$. Let $\mathcal{P}_0$ be a partition of $\mathbb{R}^d$ such that all members have diameter at most D. Let $\text{rep} : \mathcal{P}_0 \to \mathbb{R}^d$ 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 $\overline{B}_\varepsilon(\vec{x} + s(\vec{x}))$ is a subset of some member [a] 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 $f(\vec{x}) = \text{rep}(\text{member}_{\mathcal{P}_0}(\vec{x} + s(\vec{x})))$.*

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

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

Once this is proven, Proposition F.4.1 follows as an immediate corollary since in the initial partition $\mathcal{P}_0 = \mathcal{G}$, all members have diameter $D = 2\varepsilon(d+1)$, and the longest vector $\vec{\lambda} \in \Lambda$ has length $\ell = 2\varepsilon(d+1)$ (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 $(|\Lambda|, \varepsilon)$-secluded (i.e. that for any $\vec{p} \in \mathbb{R}^d$ it holds that $|\mathcal{N}_\varepsilon(\vec{p})| \leq |\Lambda|$) [6]. Let $\vec{p} \in \mathbb{R}^d$ be arbitrary. For any $\vec{x} \in \overline{B}_\varepsilon(\vec{p})$, it follows that

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

$\vec{p}+s(\vec{x}) \in \overline{B}_\varepsilon(\vec{x}+s(\vec{x}))$ [7]. Since $\overline{B}_\varepsilon(\vec{x}+s(\vec{x})) \subseteq \text{member}(\vec{x}+s(\vec{x}))$ (by our requirements on s), it then follows that $\vec{p}+s(\vec{x}) \in \text{member}(\vec{x}+s(\vec{x}))$ and so $\text{member}(\vec{p}+s(\vec{x})) = \text{member}(\vec{x}+s(\vec{x}))$. This allows us to show as follows that f takes on at most $|\Lambda|$ values on the set $\overline{B}_\varepsilon(\vec{p})$:

$$\begin{aligned}
\left\{f(\vec{x}) : \vec{x} \in \overline{B}_\varepsilon(\vec{p})\right\} &= \left\{\text{rep}(\text{member}_{\mathcal{P}_0}(\vec{x}+s(\vec{x}))) : \vec{x} \in \overline{B}_\varepsilon(\vec{p})\right\} && \text{(Def'n of } f\text{)} \\
&= \left\{\text{rep}(\text{member}_{\mathcal{P}_0}(\vec{p}+s(\vec{x}))) : \vec{x} \in \overline{B}_\varepsilon(\vec{p})\right\} \\
& && (\text{member}_{\mathcal{P}_0}(\vec{p}+s(\vec{x})) = \text{member}_{\mathcal{P}_0}(\vec{x}+s(\vec{x}))) \\
&\subseteq \left\{\text{rep}(\text{member}_{\mathcal{P}_0}(\vec{p}+\vec{\lambda})) : \vec{\lambda} \in \Lambda\right\} && (s(\vec{x}) \in \Lambda)
\end{aligned}$$

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

---

[7]

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

$$
\begin{aligned}
|\mathcal{N}_\varepsilon(\vec{p})| &= \left|\left\{\text{member}_{\mathcal{P}}(\vec{x}) : \vec{x} \in \overline{B}_\varepsilon(\vec{p})\right\}\right| && \text{(Fact 3.7.1)}\\
&= \left|\left\{f^{-1}(f(\vec{x})) : \vec{x} \in \overline{B}_\varepsilon(\vec{p})\right\}\right| && \text{(Def'n of the induced partition } \mathcal{P}\text{)}\\
&= \left|\left\{f^{-1}(\vec{w}) : \vec{w} \in \left\{f(\vec{x}) : \vec{x} \in \overline{B}_\varepsilon(\vec{p})\right\}\right\}\right| && \text{(Reformat)}\\
&\leq \left|\left\{f^{-1}(\vec{w}) : \vec{w} \in \left\{\text{rep}(\text{member}_{\mathcal{P}_0}(\vec{p}+\vec{\lambda})) : \vec{\lambda} \in \Lambda\right\}\right\}\right| && \text{(Prior paragraph)}\\
&\leq \left|\left\{\text{rep}(\text{member}_{\mathcal{P}_0}(\vec{p}+\vec{\lambda})) : \vec{\lambda} \in \Lambda\right\}\right| && (f^{-1} \text{ is a function})\\
&\leq |\Lambda| && \text{(Prior paragraph)}
\end{aligned}
$$

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 $\text{rep}(\text{member}_{\mathcal{P}_0}(\vec{x} + s(\vec{x})) = \text{rep}(\text{member}_{\mathcal{P}_0}(\vec{y} + s(\vec{y}))$. By definition of rep, the left hand side is contained in $\text{member}_{\mathcal{P}_0}(\vec{x}+s(\vec{x}))$ and the right hand side is contained in $\text{member}_{\mathcal{P}_0}(\vec{y} + s(\vec{y}))$, and since the left and right hand side are the same point, it must be that $\text{member}_{\mathcal{P}_0}(\vec{x} + s(\vec{x})) = \text{member}_{\mathcal{P}_0}(\vec{y} + s(\vec{y}))$, and because members of $\mathcal{P}_0$ have diameter at most D, it follows that $\|(\vec{x} + s(\vec{x})) - (\vec{y} + s(\vec{y}))\|_\infty \leq D$. Now observe that

$$
\begin{aligned}
\|\vec{x} - \vec{y}\|_\infty &\leq \|[\vec{x}] - [\vec{x} + s(\vec{x})]\|_\infty + \|[\vec{x} + s(\vec{x})] - [\vec{y} + s(\vec{y})]\|_\infty + \|[\vec{x}] - [\vec{x} + s(\vec{x})]\|_\infty\\
&\leq \|s(\vec{x})\|_\infty + D + \|s(\vec{y})\|_\infty\\
&\leq \ell + D + \ell && (s(\vec{x}), s(\vec{y}) \in \Lambda)
\end{aligned}
$$

so $\text{diam}(X) \leq 2\ell + D$. □

Appendix G

Miscellaneous Facts

Fact G.0.1. *For any $\alpha \in \mathbb{R}$, there exists $\gamma \in \mathbb{R}$ such that $\gamma < \alpha$ and $\lceil \gamma \rceil = \lceil \alpha \rceil$.*

Proof. Let $n = \lceil \alpha \rceil$. This implies that $\alpha > n - 1$ (otherwise $\alpha \leq n - 1$ so $\lceil \alpha \rceil \leq n - 1$). Thus $(n - 1, \alpha)$ is non-empty and we can take any $\gamma \in (n - 1, \alpha)$. Then $n - 1 < \lceil \gamma \rceil \leq \lceil \alpha \rceil = n$ showing $\lceil \gamma \rceil = n$ as well. □

Fact G.0.2. *For $d \in [1, \infty)$, $x \in [0, 1]$, and $\alpha \in [0, \infty)$, it holds that $(x^{1/d} + \alpha)^d \geq x(1 + \alpha)^d$.*

Proof. We will show that $(x^{1/d} + \alpha)^d - x(1 + \alpha) \geq 0$ for these parameters. Let For d, α as above, let $f : [0, 1] \to \mathbb{R}$ be defined by $f(x) = (x^{1/d} + \alpha)^d - x(1 + \alpha)$. Observe that $f(0) = \alpha^2 \geq 0$ and $f(1) = (1 + \alpha)^d - (1 + \alpha)^d = 0$. We will now prove that f is convex on the domain[1] $[0, 1]$. This will be sufficient to prove the claim because f is also non-negative at 0 and at 1.

[1] Actually we could have defined the domain of f to be $[0, \infty)$ and we show that f is convex on that domain. However, we only have need of the interval $[0, 1]$.

We show that f is convex on $[0,1]$ by considering its second derivative on $(0,1]$.

$$\begin{aligned}\frac{d}{dx}f(x) &= \frac{d}{dx}\left[(x^{1/d}+\alpha)^d - x(1+\alpha)^d\right] \\ &= d(x^{1/d}+\alpha)^{d-1}\cdot\frac{1}{d}x^{1/d-1} - (1+\alpha)^d \\ &= (x^{1/d}+\alpha)^{d-1}x^{1/d-1} - (1+\alpha)^d\end{aligned}$$

where we use the convention that $0^0 = 1$. Then

$$\begin{aligned}\frac{d^2}{dx^2}f(x) &= \frac{d}{dx}\left[(x^{1/d}+\alpha)^{d-1}x^{1/d-1} - (1+\alpha)^d\right] \\ &= (x^{1/d}+\alpha)^{d-1}\cdot\left(\frac{1}{d}-1\right)x^{1/d-2} + x^{1/d-1}\cdot(d-1)(x^{1/d}+\alpha)^{d-2}\frac{1}{d}x^{1/d-1} \\ &= (x^{1/d}+\alpha)(x^{1/d}+\alpha)^{d-2}\left(-\frac{d-1}{d}\right)\left(x^{1/d-2}\right) \\ &+ (x)\left(x^{1/d-2}\right)\left(\frac{d-1}{d}\right)\left(x^{1/d}+\alpha\right)^{d-2}\left(x^{1/d-1}\right) \\ &= \frac{(x^{1/d}+\alpha)^{d-2}(d-1)\left(x^{1/d-2}\right)}{d}\left[-(x^{1/d}+\alpha)+(x)\left(x^{1/d-1}\right)\right] \\ &= \frac{(x^{1/d}+\alpha)^{d-2}(d-1)\left(x^{1/d-2}\right)}{d}\left[-(x^{1/d}+\alpha)+\left(x^{1/d}\right)\right] \\ &= \frac{-\alpha(x^{1/d}+\alpha)^{d-2}(d-1)\left(x^{1/d-2}\right)}{d} \\ &= \frac{-\alpha(x^{1/d}+\alpha)^{d}(d-1)\left(x^{1/d}\right)}{d(x^{1/d}+\alpha)^2x^2}.\end{aligned}$$

Note that $\frac{d^2}{dx^2}f(x) \leq 0$ for $x \in (0,1]$ since $\alpha \geq 0$. This shows that f is convex on $(0,1]$ and by continuity on $[0,1]$ which completes the proof. □

Fact G.0.3. *For* $\varepsilon \in [0, \frac{1}{2}]$ *it holds that* $\log_4(1+2\varepsilon) \geq \varepsilon$.

Proof. The function $f(x) = \log_4(1+2x) - x$ is concave on its domain $(-\frac{1}{2}, \infty)$ because

$$\frac{d^2}{dx^2} f(x) = \frac{-4}{(1+2x)^2 \log(4)}$$

which is negative for all $x \in (-\frac{1}{2}, \infty)$. Also, $f(0) = 0$ and $f(\frac{1}{2}) = 0$ which shows by the other definition of concavity that $f(x) \geq 0$ for $x \in [0, \frac{1}{2}]$. Thus, for $\varepsilon \in [0, \frac{1}{2}]$ it holds that $\log_4(1 + 2\varepsilon) \geq \varepsilon$. $\square$

Bibliography

[AER21] Dan Alistarh, Faith Ellen, and Joel Rybicki. Wait-free approximate agreement on graphs, March 2021. arXiv:2103.08949 [cs].

[AF93] Noga Alon and Zoltán Füredi. Covering the cube by affine hyperplanes. *European Journal of Combinatorics*, 14(2):79–83, 1993.

[AZ10] Martin Aigner and Günter M. Ziegler. Lines in the plane and decompositions of graphs. In Martin Aigner and Günter M. Ziegler, editors, *Proofs from THE BOOK*, pages 63–67. Springer, Berlin, Heidelberg, 2010.

[BBDLa00] 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*, 24(1):35–48, January 2000.

[BCF$^+$22] Vittorio Bilò, Ioannis Caragiannis, Michele Flammini, Ayumi Igarashi, Gianpiero Monaco, Dominik Peters, Cosimo Vinci, and William S. Zwicker. Almost Envy-Free Allocations with Connected Bundles. *Games and Economic Behavior*, 131:197–221, January 2022. arXiv:1808.09406 [cs, econ].

[BHMN20] Joshua Brakensiek, Marijn Heule, John Mackey, and David Narváez. The Resolution of Keller's Conjecture. *arXiv:1910.03740 [cs, math]*, December 2020. arXiv: 1910.03740.

[Bli14] H. F. Blichfeldt. A new principle in the geometry of numbers, with some applications. *Transactions of the American Mathematical Society*, 15(3):227–235, 1914.

[Bro11] L. E. J. Brouwer. Beweis der Invarianz der Dimensionenzahl. *Mathematische Annalen*, 70(2):161–165, June 1911.

[Bro13] L. E. J. Brouwer. Über den natürlichen dimensionsbegriff. *Journal für die reine und angewandte Mathematik (Crelles Journal)*, pages 146 – 152, 1913.

[BRS11] Martin Biely, Peter Robinson, and Ulrich Schmid. Easy Impossibility Proofs for k-Set Agreement in Message Passing Systems, March 2011. arXiv:1103.3671 [cs].

[Can15] Clément L. Canonne. A survey on distribution testing: Your data is big. but is it blue? *Electron. Colloquium Comput. Complex.*, TR15-063, 2015.

[CMY23] Zachary Chase, Shay Moran, and Amir Yehudayoff. Replicability and stability in learning, April 2023. arXiv:2304.03757 [cs, math, stat].

[CS90] K. Corrádi and S. Szabó. A combinatorial approach for Keller's conjecture. *Periodica Mathematica Hungarica*, 1990.

[DEL+11] Jennifer Debroni, John D. Eblen, Michael A. Langston, Wendy Myrvold, Peter Shor, and Dinesh Weerapurage. A complete resolution of the Keller maximum clique problem. In *Proceedings of the 2011 Annual ACM-SIAM*

Symposium on Discrete Algorithms (SODA), Proceedings, pages 129–135. Society for Industrial and Applied Mathematics, January 2011.

[DIP05] Nikolai Dolbilin, Yoshiaki Itoh, and Alexei Poyarkov. On random tilings and packings of space by cubes. *The Proceedings of COE Workshop on Sphere Packings*, pages 70–79, March 2005. Kyushu University, Fukuoka.

[DIP18] Mathieu Dutour, Yoshiaki Itoh, and Alexei Poyarkov. Cube packings, second moment and holes. *arXiv:math/0509100*, September 2018. arXiv: math/0509100.

[DPWV22] Peter Dixon, A. Pavan, Jason Vander Woude, and N. V. Vinodchandran. Pseudodeterminism: Promises and lowerbounds. In *Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2022, page 1552–1565, New York, NY, USA, 2022. Association for Computing Machinery.

[Fol99] Gerald B. Folland. *Real Analysis: Modern Techniques and Their Applications*. John Wiley & Sons, April 1999. Google-Books-ID: N8jVDwAAQBAJ.

[Fug74] Bent Fuglede. Commuting self-adjoint partial differential operators and a group theoretic problem. *Journal of Functional Analysis*, 16(1):101–121, May 1974.

[Fur36] Ph. Furtwängler. Über Gitter konstanter Dichte. *Monatshefte für Mathematik und Physik*, 43(1):281–288, December 1936.

[Gar02] R. J. Gardner. The Brunn-Minkowski inequality. *Bulletin of the American Mathematical Society*, 39(03):355–406, April 2002.

[GG11] Eran Gat and Shafi Goldwasser. Probabilistic search algorithms with unique answers and their cryptographic applications. *Electron. Colloquium Comput. Complex.*, TR11-136, 2011.

[GGR13] Oded Goldreich, Shafi Goldwasser, and Dana Ron. On the possibilities and limitations of pseudodeterministic algorithms. In *Proceedings of the 4th conference on Innovations in Theoretical Computer Science - ITCS '13*, page 127, Berkeley, California, USA, 2013. ACM Press.

[GL19] Ofer Grossman and Yang 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*, pages 606–620. SIAM, 2019.

[Gla12] Alexey Glazyrin. Lower bounds for the simplexity of the n-cube. *Discrete Mathematics*, 312(24):3656–3662, December 2012.

[Gol15] Paul W. Goldberg. The Complexity of the Path-following Solutions of Two-dimensional Sperner/Brouwer Functions, June 2015. arXiv:1506.04882 [cs].

[Gol19a] Oded Goldreich. Multi-pseudodeterministic algorithms. *Electron. Colloquium Comput. Complex.*, TR19-012, 2019.

[Gol19b] Oded Goldreich. Multi-pseudodeterministic algorithms. *Electronic Colloquium on Computational Complexity (ECCC)*, 26:12, 2019.

[Gru07] Peter M. Gruber. *Convex and Discrete Geometry*. Springer Science & Business Media, May 2007. Google-Books-ID: bSZKAAAAQBAJ.

[GSP22] Anat Ganor, Karthik C. S., and Dömötör Pálvölgyi. On Communication Complexity of Fixed Point Computation, May 2022. arXiv:1909.10958 [cs].

[Haj42] Georg Hajós. Über einfache und mehrfache Bedeckung des n-dimensionalen Raumes mit einem Würfelgitter. *Mathematische Zeitschrift*, 47(1):427–467, December 1942.

[HK18] William M. Hoza and Adam R. Klivans. Preserving Randomness for Adaptive Algorithms. *arXiv:1611.00783 [cs]*, June 2018. arXiv: 1611.00783.

[HW48] Witold Hurewicz and Henry Wallman. *Dimension Theory*. Princeton University Press, 1948.

[Kel30] Ott-Heinrich Keller. Über die lückenlose Erfüllung des Raumes mit Würfeln. *Journal für die reine und angewandte Mathematik*, 1930(163):231–248, January 1930. Publisher: De Gruyter Section: Journal für die reine und angewandte Mathematik.

[Kin09] Shiva Kintali. Scarf is Ppad-Complete, May 2009. arXiv:0812.1601 [cs].

[Kis17] Andrzej P. Kisielewicz. Towards Resolving Keller's Cube Tiling Conjecture in Dimension Seven. *arXiv:1701.07155 [math]*, January 2017. arXiv: 1701.07155.

[KŁ14] Andrzej P. Kisielewicz and Magdalena Łysakowska. On Keller's conjecture in dimension seven. *arXiv preprint arXiv:1401.4689*, 2014.

[Kom94] Hidetoshi Komiya. A Simple Proof of K-K-M-S Theorem. *Economic Theory*, 4(3):463–466, 1994. Publisher: Springer.

[KORW08] Guy Kindler, Ryan O'Donnell, Anup Rao, and Avi Wigderson. Spherical Cubes and Rounding in High Dimensions. In *2008 49th Annual IEEE Symposium on Foundations of Computer Science*, pages 189–198, Philadelphia, PA, USA, October 2008. IEEE.

[KROW12] Guy Kindler, Anup Rao, Ryan O'Donnell, and Avi Wigderson. Spherical cubes: optimal foams from computational hardness amplification. *Communications of the ACM*, 55(10):90–97, October 2012.

[Kuh60] H. W. Kuhn. Some combinatorial lemmas in topology. *IBM Journal of Research and Development*, 4(5):518–524, 1960.

[Leb11] Henri Lebesgue. Sur la non-applicabilité de deux domaines appartenant respectivement à des espaces àn etn+p dimensions. *Mathematische Annalen*, 70(2):166–168, June 1911.

[LPS01] Jesus A. de Loera, Elisha Peterson, and Francis Edward Su. A Polytopal Generalization Of Sperner's Lemma. *J. Combin. Theory Ser. A*, 100:2002, 2001.

[LRW00] Jeffrey C. Lagarias, James A. Reeds, and Yang Wang. Orthonormal bases of exponentials for the n-cube. *Duke Mathematical Journal*, 103(1):25–37, May 2000. Publisher: Duke University Press.

[LS92] J. Lagarias and P. Shor. Keller's cube-tiling conjecture is false in high dimensions. *Bull. Amer. Math. Soc.*, 1992.

[Łys] Magdalena Łysakowska. Extended Keller graph and its properties: Quaestiones Mathematicae: Vol 42, No 4.

[Mac02] Mackey. A Cube Tiling of Dimension Eight with No Facesharing. *Discrete & Computational Geometry*, 28(2):275–279, August 2002.

[Min07] Hermann Minkowski. *Diophantische Approximationen*. Vieweg+Teubner Verlag, Wiesbaden, 1907.

[MS19] Frédéric Meunier and Francis Edward Su. Multilabeled versions of Sperner's and Fan's lemmas and applications, May 2019. arXiv:1801.02044 [cs, math].

[MZ19] Frédéric Meunier and Shira Zerbib. Envy-free cake division without assuming the players prefer nonempty pieces, January 2019. arXiv:1804.00449 [cs, math].

[Nis22] Susumu Nishimura. Proving Unsolvability of Set Agreement Task with Epistemic mu-Calculus, May 2022. arXiv:2205.06452 [cs].

[OS03] David Orden and Francisco Santos. Asymptotically Efficient Triangulations of thed-Cube. *Discrete & Computational Geometry*, 30(4):509–528, October 2003.

[Per40a] Oskar Perron. Über lückenlose Ausfüllung desn-dimensionalen Raumes durch kongruente Würfel. *Mathematische Zeitschrift*, 46(1):1–26, December 1940.

[Per40b] Oskar Perron. Über lückenlose Ausfüllung desn-dimensionalen Raumes durch kongruente Würfel. II. *Mathematische Zeitschrift*, 46(1):161–180, December 1940.

[Poy07] Alexey P. Poyarkov. Random packings by cubes. *Journal of Mathematical Sciences*, 146(1):5577–5583, October 2007.

[Rob79] Raphael M. Robinson. Multiple tilings ofn-dimensional space by unit cubes. *Mathematische Zeitschrift*, 166(3):225–264, October 1979.

[Sal82a] John F. Sallee. A note on minimal triangulations of an n-cube. *Discrete Applied Mathematics*, 4(3):211–215, 1982.

[Sal82b] John F. Sallee. A triangulation of the n-cube. *Discrete Mathematics*, 40(1):81–86, January 1982.

[SH22] Erel Segal-Halevi. Generalized Rental Harmony. *The American Mathematical Monthly*, 129(5):403–414, May 2022. arXiv:1912.13249 [cs, econ].

[SI10] M. D. Sikirić and Y. Itoh. Combinatorial cube packings in the cube and the torus. *Eur. J. Comb.*, 2010.

[Sie20] Wacław Sierpiński. Sur la question de la mesurabilité de la base de M. Hamel. *Fundamenta Mathematicae*, 1:105–111, 1920. Publisher: Instytut Matematyczny Polskiej Akademii Nauk.

[SZ99] M. Saks and S. Zhou. $\mathrm{BP_hSPACE(s)} \subseteq \mathrm{DSPACE(s^{3/2})}$. *J. Comput. Syst. Sci.*, 58(2):376–403, 1999.

[Sza86] S. Szabó. A reduction of Keller's conjecture. *Periodica Mathematica Hungarica*, 17(4):265–277, December 1986.

[Tao04] Terence Tao. Fuglede's conjecture is false in 5 and higher dimensions. *Mathematical Research Letters*, 11(2):251–258, March 2004. Publisher: International Press of Boston.

[The21] The Manim Community Developers. Manim – Mathematical Animation Framework, 11 2021.

[use14] user21467. Answer to "Isoperimetric inequality, isodiametric inequality, hyperplane conjecture... what are the inequalities of this kind known or conjectured?", April 2014. Mathematics Stack Exchange.

[vdLTY99] Gerard van der Laan, Dolf Talman, and Zaifu Yang. Intersection theorems on polytopes. *Mathematical Programming*, 84(1):25–38, January 1999.

[Wan05] Xianfu Wang. Volumes of Generalized Unit Balls. *Mathematics Magazine*, 78(5):390–395, 2005. Publisher: Mathematical Association of America.

[Wol77] Laurence A Wolsey. Cubical sperner lemmas as applications of generalized complementary pivoting. *Journal of Combinatorial Theory, Series A*, 23(1):78–87, July 1977.

[Zak85] Joseph Zaks. How Does a Complete Graph Split into Bipartite Graphs and How are Neighborly Cubes Arranged? *The American Mathematical Monthly*, 92(8):568–571, 1985. Publisher: Mathematical Association of America.

[Zak87] Joseph Zaks. Neighborly Families of Congruent Convex Polytopes. *The American Mathematical Monthly*, 94(2):151–155, 1987. Publisher: Mathematical Association of America.

[ZBFK06] C. Zong, B. Bollobas, W. Fulton, and A. Katok. *The Cube-A Window to Convex and Discrete Geometry.* Cambridge Tracts in Mathematics. Cambridge University Press, 2006.

[Zon05] Chuanming Zong. What is known about unit cubes. *Bulletin of The American Mathematical Society - BULL AMER MATH SOC*, 42:181–212, April 2005.