Four Problems in Probability and Optimization

James Pfeiffer

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2013

Reading Committee:

Rekha Thomas, Chair

Monty McGovern

Christopher Hoffman

Program Authorized to Offer Degree:
Mathematics

University of Washington

Abstract

Four Problems in Probability and Optimization

James Pfeiffer

Chair of the Supervisory Committee:
Professor Rekha Thomas
Mathematics

This thesis studies bootstrap percolation, a problem in probability, as well as several topics in the application of sums of squares to combinatorial optimization.

In the chapter on percolation, we bound the critical probability for bootstrap percolation on the Hamming torus, as well as the critical probability for i-dimensional subgraphs to percolate. In the case $d = \theta = 3$ we exhibit a framework for deriving exact results within the scaling window using Poisson approximation.

In the chapters on combinatorial optimization, we consider the K_i-cover problem and the max cut problem. We show that a family of facets arising from K_i-p-holes is valid on the $i/2$ theta body. We also prove an integrality gap of $1/2$ for the triangle free problem, and show that at least $n/2$ steps are required for the triangle free problem's theta bodies to converge in the case $G = K_n$.

We introduce a criterion for an invariant polynomial to be a sum of squares on the hypercube. This gives a simple proof of Laurent's result that the theta body heirarchy requires at least $n/4$ steps to converge to the max cut polytope of K_n. It also allows us to give the first lower bounds on degrees of denominators in Hilbert's 17th problem.

In the last chapter, we consider the S_n-irreducible decomposition of the space of matchings on K_n as given by Barbasch and Vogan. We give an explicit map of the isomorphism in their result. We also generalize their approach to matchings on hypergraphs.

ACKNOWLEDGMENTS

I would like to thank Rekha Thomas for guiding me through my education in mathematical research, and for many interesting discussions over the years.

DEDICATION

To Katharine

TABLE OF CONTENTS

Chapter 1

INTRODUCTION

This thesis studies four problems in probability and optimization. In this introduction, we give an extended abstract of each chapter, and give some context and motivation. Chapter 2 studies bootstrap percolation, a probabilistic model of nucleation and growth. Chapters 3 and 4 apply sums of squares approximations to combinatorial optimization. Chapter 5 studies the representation theory of the set of matchings and is an offshoot of the work in Chapter 4.

1.1 Bootstrap Percolation on the Hamming Torus

Chapter 2 is taken from a paper coauthored with Janko Gravner, Christopher Hoffman, and Davis Sivakoff [23], and submitted to *Annals of Applied Probability.* It investigates the critical behavior of bootstrap percolation on the Hamming torus.

1.1.1 Problem Description

Let $G = (V, E)$ be a graph and $\theta \in \mathbb{N}$ a threshold. For any initial configuration $\omega_0 \in \{0,1\}^V$, define a sequence ω_i for $i \geq 1$ by the recursion

$$\omega_{j+1}(v) = \begin{cases} 1 & \text{if } \omega_j(v) = 1 \text{ or } \sum_{(v,w)\in E} \omega_j(w) \geq \theta \\ 0 & \text{else} \end{cases}$$

and let $\omega_\infty = \lim_i \omega_i$. This limit is defined pointwise, since $(\omega_i(v))_i$ is increasing for each v. Let $S_0 = \omega_0^{-1}(1)$.

We think of bootstrap percolation as process consisting of a growing set of *infected* or *active* vertices which starts with only the vertices in S_0 active. Once a vertex is active, it stays active, and new vertices become active once at least θ of their neighbors are active.

The main goal is to determine under what conditions the entire graph eventually becomes active, that is, when $\omega_\infty(v) = 1$ for all v. In this case, we say that percolation has occurred, or that S_0 *spans* G.

The initial set S_0 of active vertices can be deterministic or random. See [39] for some results on the minimal size of S_0 required for percolation in the deterministic case. In this paper, we let each vertex be active at time 0 independently with probability p. Thus, the bootstrap percolation process becomes a random variable. We are interested in the probability that percolation occurs. In Figure 1.1.1, we plot $\mathbb{P}(\text{percolation})$ vs p for the graph $[n]^2$ with Euclidean nearest neighbors and threshold $\theta = 2$. Here $[n] = \{1, \ldots, n\}$. As p increases, there is a relatively sharp transition from a low probability of percolation to a high probability. As we increase n, the transition becomes sharper.

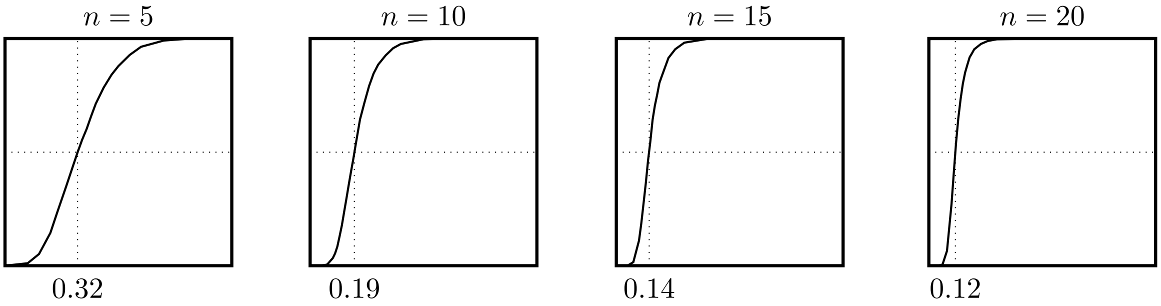


Figure 1.1: Plots of $\mathbb{P}(\text{percolation})$ vs p for nearest neighbor $[n]^2$, for $n = 5, 10, 15, 20$. Critical probabilities are indicated.

To pin down this transition, we define the *critical probability* p_c to be the p at which $\mathbb{P}(\text{percolation}) = \frac{1}{2}$. (Since the transition window narrows as n increases, we could choose any fixed $c \in (0, 1)$ in the definition without changing the theory.) We consider a family of graphs G_n and wish to find the dependence of p_c on n. For instance, in the example in Figure 1.1.1, it turns out that p_c scales as $\frac{1}{\log n}$.

For the graphs we consider, it turns out that the critical probability is essentially n^a for some $a < 0$ in the following sense. Let $p = n^b$. If $b < a$, then $\mathbb{P}(\text{percolation}) \to 0$ as $n \to \infty$, while if $b > a$, then $\mathbb{P}(\text{percolation}) \to 1$. In certain cases, we are able to look within the *scaling window* and determine $f(c) = \mathbb{P}(\text{percolation})$, when $p = cn^a$.

We study bootstrap percolation on the *Hamming torus* $G = ([n]^d, E)$, where x and y are adjacent if they differ in exactly one coordinate. In Figure 1.2 we see an example of an initial percolating set when $\theta = 2$ on the two-dimensional Hamming torus.

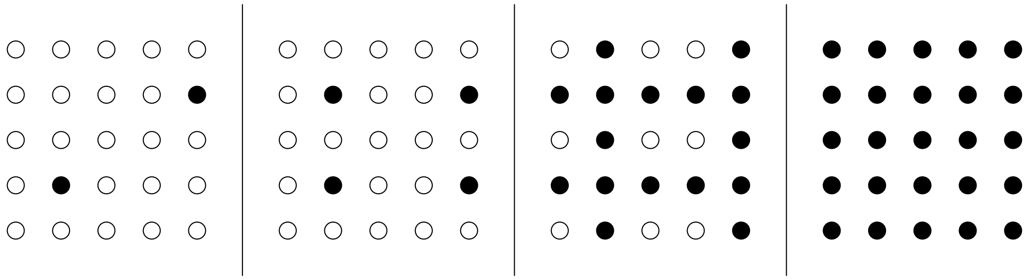


Figure 1.2: The bootstrap percolation process with threshold $\theta = 2$ on the Hamming torus with $n = 5$, $d = 2$. Nodes are adjacent whenever they share an x or y coordinate. This example takes four steps to percolate.

1.1.2 Background

Bootstrap percolation was introduced in 1979 by Chalupa, Leath and Reich [14] as a model of nucleation and metastability in physical processes such as crack formations, clustering, and magnetic spin alignment. For more applications and background see surveys by Adler and Levi [1] and Holroyd [29].

The first results in this area were by van Enter [48] and Schonmann [42], who proved that for the lattice $\mathbb{Z}^d$, the critical probability p_c is either 1 or 0 according to whether $\theta \leq d$ or $\theta > d$. For a large lattice cube $[n]^d \subset \mathbb{Z}^d$ (where each point is connected to the nearest $2d$ points), Aizenman and Lebowitz [2] proved that p_c behaves as $(\frac{1}{\log n})^{d-1}$ when $\theta = 2$. Later Cerf and Cirillo [12] and Cerf and Manzo [13] established the scaling $p_c \approx (\log_{\theta-1} n)^{-d+\theta-1}$ for $3 \leq \theta \leq d$. Here $\log_{\theta-1}$ denotes the $(\theta - 1)$'st iteration of the logarithm.

The Hamming torus is interesting because of computer simulations showing that the percolation on $[n]^d$ tends to proceed along straight lines. By using the Hamming metric for our graph distance, we can investigate this behavior.

1.1.3 Results

All our results are for the Hamming torus described in 1.1.1. In addition to finding the critical probability for the entire graph to percolate, we consider the event that an i-dimensional subgraph percolates. We find an expression for the critical probability of this event for $1 \le i \le d$.

- For $d = 2$, we find the critical probability exponent for all θ and describe the behavior within the scaling window as $n \to \infty$. For $d = \theta = 3$, we again find the critical probability and describe the limiting behavior within the scaling window. These results are powered by Stein's method [9], a method for dealing with almost independent events.

- For all d and all $i \ge 2$, we show that the critical probability for an i-dimensional subgraph to percolate is $p_c = n^{-1-\frac{2}{\theta}+\Theta(\theta^{-\frac{3}{2}})}$, independent of i. We prove this by constructing a necessary and a sufficient critical event for percolation, each of which is easier to analyze than the percolation event. We show that when $p \gg p_c$, the sufficient event happens with probability going to 1, and with $p \ll p_c$, the necessary event happens with probability going to 0.

- In particular, we show that the critical probability for the entire graph to percolate is $p_c = n^{-1-\frac{2}{\theta}+\Theta(\theta^{-\frac{3}{2}})}$.

1.1.4 Comments

Although our results are stated asymptotically in n and for large θ, the proofs also give results for large finite graphs. However, computing the constants involved is arduous.

It is interesting that the critical probabilities are all approximately $n^{-1-\frac{2}{\theta}}$ for $2 \leq i \leq d$. It turns out that for $i = 1$, the critical probability is $n^{-1-\frac{d}{\theta}}$, a qualitatively different behavior.

For $d = 3$, we get the same upper and lower bound for p_c when $\theta \leq 7$ and $\theta = 9, 11$, and therefore can determine the exact critical exponent. This raises the question of whether our proof technique can be improved to bring the bounds closer together for other values of θ.

1.2 Combinatorial Optimization and Sums of Squares

Chapters 3 and 4 study the application of sums of squares of polynomials to combinatorial optimization. In this section, we give a description of the general area of combinatorial optimization and of polynomial sums of squares in particular. We introduce many of the concepts which will be used later.

1.2.1 Combinatorial Optimization

Combinatorial optimization studies optimizing a function over a finite set. By interpolation, all such functions are polynomials in the decision variables. Applications are common in computer science, mathematics, and operations research. For example, the goal of the *traveling salesman problem* is to find a minimum-cost tour among a set of cities. The goal of the *shortest path problem* is to find a shortest-distance path between two nodes in a graph, and the goal of the *assignment problem* is to assign workers to jobs while minimizing the total wages paid.

In each of the above problems, we minimize a linear function over a finite set. For instance, in the traveling salesman problem (TSP), we model the problem as a complete graph $G = ([n], \binom{[n]}{2})$. For each $e \in E$, a cost c_e is given. The vertices represent cities, with specified costs to travel between them. We want to find the minimum of $f(T) = \sum_{e \in E} c_e x_e$ over all tours T. Here, a tour is represented by a 0/1 vector $x \in \mathbb{R}^{\binom{n}{2}}$, with $x_e = 1$ if and only if the edge e is part of T.

A common approach is to rewrite this as an *integer linear program* by replacing the constraints $x_e \in \{0, 1\}$ by $x_e \in \mathbb{Z}$, $0 \leq x_e \leq 1$, and encoding the fact that T is a tour using other linear constraints. But since we optimize a linear function, the optimum is unchanged if we instead optimize over the *convex hull* $\mathrm{TSP}(n)$ of all tours. The set $\mathrm{TSP}(n) \subset [0,1]^{\binom{n}{2}}$ is a polytope whose vertices are the vectors coming from tours. This is a natural object to study without referring to a specific set (c_e) of costs, since changing the costs is equivalent to optimizing in a different direction over $\mathrm{TSP}(n)$.

To explain this geometric approach in general, consider a problem where we optimize a linear function over some feasible collection $\mathcal{C}$ of subsets of a fixed set X. For instance, in the TSP we let $\mathcal{C}$ be the collection of all tours; here X is the set of edges. In the shortest-path problem we consider $\mathcal{C}$ to be the collection of all paths from a fixed $x \to y$; again X is the set of edges. A non-graph example is the *knapsack problem*, where $\mathcal{C}$ is the set of subsets of $X = \{1, \ldots, n\}$ satisfying a capacity constraint. In general, we will assume that $X = \{1, \ldots, n\}$ for some n; this can always be accomplished by enumerating X. For each $C \in \mathcal{C}$, define its *characteristic vector* $\chi_C \in \{0,1\}^n$, where for $1 \leq i \leq n$, $(\chi_C)_i = 1$ if and only if $i \in C$. Then define the polytope $P(\mathcal{C}) = \mathrm{conv}\{\chi_C : C \in \mathcal{C}\}$. If we can optimize any linear function in polynomial time over $P(\mathcal{C})$, we can optimize the same linear function over $\mathcal{C}$ in polynomial time.

Of course, many combinatorial optimization problems are NP-complete or harder. That is, we are unlikely to ever find efficient algorithms to solve them exactly. In some cases where the abstract mathematical problem does have an efficient algorithm, real-world constraints can make it intractable. For instance, when assigning medical students to residencies, the constraint that married pairs of students live in the same city makes the problem NP-complete. Therefore, much research has focused on developing methods to approximately solve these problems in polynomial time. Using various methods, we can construct a relaxation $P'(\mathcal{C})$ to $P(\mathcal{C})$ over which we can optimize efficiently, and try to show that the optimum objective values are close. In the next section, we describe one such method based on sums of squares of polynomials.

1.2.2 Sums of Squares and the Theta Body

Consider, for the moment, unconstrained polynomial optimization on $\mathbb{R}^n$. That is, we are given a polynomial $f \in \mathbb{R}[x_1, \ldots, x_n]$ and must find the global minimum value $f_{\min}$. It turns out that optimization is closely related to checking nonnegativity. If we can find the global minimum of a polynomial f, we can certainly check if f is nonnegative everywhere by checking if its minimum $f_{\min} \geq 0$. Conversely, if we can check whether polynomials are nonnegative, we can compute $f_{\min}$ by observing that $f_{\min} = \max\{r : f(x) - r \geq 0\}$. We can solve the latter problem by sampling different r (e.g. with binary search) and for each r checking whether $f - r$ is nonnegative.

Checking arbitrary polynomials for nonnegativity is NP-hard, so we seek an approximation. Note that if we can write $f(x)$ as a sum of squares (of polynomials), then it is certainly nonnegative. Therefore, we can define the relaxation $f_{\text{sos}} = \max\{r : f(x) - r \text{ is a sum of squares}\}$. We have that $f_{\min} \geq f_{\text{sos}}$, so this is a lower bound on the minimum. This idea is discussed at more length in [40] and [33].

So far, this only makes sense for polynomials on $\mathbb{R}^n$. We will define a notion of sums of squares that is applicable to combinatorial optimization problems. This definition originated with Lovász [35] and was used to produce an approximation to the *stable set problem*. It was generalized by Gouveia, Parrilo, and Thomas [21] to give an approximation to the convex hull of any algebraic set. This idea is also closely related to Lasserre's [33] and Parrilo's [37] [38] heirarchies for polynomial optimization.

As in Section 1.2.1, consider a collection $\mathcal{C}$ of subsets of $\{1, \ldots, n\}$. Let $\mathcal{V} = \{\chi_C : C \in \mathcal{C}\}$. Since $\mathcal{V}$ is a finite set, it is an *algebraic variety*. That is, it is the set of solutions of a finite list of polynomials. The *ideal* I of $\mathcal{V}$ is defined to be the set of all polynomials vanishing on $\mathcal{V}$. We can define an equivalence relation on polynomials, denoted $f \equiv g \mod I$, if and only if $f - g \in I$. With this definition, $f \equiv g \mod I$ if and only if $f(x) = g(x)$ for each $x \in \mathcal{V}$. For a polynomial f and an integer k, we say f is *k-sos mod I* if we can write $f \equiv \sum_i g_i^2 \mod I$ for some polynomials g_i with degree at most k. Now define the kth *theta body* of I, $\text{TH}_k(I)$, to be the intersection of the affine hyperplanes defined by all affine linear functions f which are k-sos mod I. This defines a relaxation of the convex hull of $\mathcal{V}$: $\operatorname{conv}(\mathcal{V}) \subseteq \mathrm{TH}_k(I)$ for each k, with the approximation improving as k increases. In our case, $\mathcal{V} \subseteq \{0,1\}^n$, so it turns out that $\mathrm{TH}_n(I) = \operatorname{conv}(\mathcal{V})$. We can optimize over $\mathrm{TH}_k(I)$ to fixed precision, for fixed k, in time polynomial in n. This is done using *semidefinite programming*, and provides an efficient relaxation to $P(\mathcal{C}) = \operatorname{conv}(\mathcal{V})$ for small k.

As an example of the theta body heirarchy, we describe the first theta body of the stable set problem. Historically, this was the first example where theta bodies were used, and where the general heirarchy was abstracted from. Given a graph $G = (V, E)$, a subset C of V is *stable* if for all $i, j \in C$, $ij \notin E$. Let $n = |V|$. As in Section 1.2.1, we will consider the set $\mathcal{V} \subset \{0,1\}^n$ of characteristic vectors of stable sets in G. Our goal is to describe the convex hull $\operatorname{conv}(\mathcal{V})$, denoted $\mathrm{STAB}(G)$. We can check that the ideal of $\mathcal{V}$ is

$$I = \langle x_i^2 - x_i \,\forall\, i \in [n]\,;\, x_i x_j \,\forall\, ij \in E\rangle.$$

To describe the first theta body $\mathrm{TH}_1(I)$, denoted $\mathrm{TH}(G)$ in this case, we must find all linear functions which are sums of squares of linear polynomials mod I. It turns out that $\mathrm{TH}(G)$ has a semidefinite description first discovered by Lovász [35]. We reproduce it here; see Chapter 9 of [25] for the derivation: $x \in \mathrm{TH}(G)$ if and only if there exists $M \in \mathbb{R}^{(n+1)\times(n+1)}$ such that $M \succeq 0$ and the following linear conditions hold:

1. For $1 \leq i \leq n$, $M_{0i} = M_{ii} = x_i$,
2. For each $ij \in E$, $M_{ij} = 0$,
3. $M_{00} = 1$.

The class of graphs for which $\mathrm{TH}(G) = \mathrm{STAB}(G)$ is known to be exactly the *perfect graphs.* A graph is G defined to be perfect if the chromatic number equals the clique number for each induced subgraph of G. Equivalently, G is perfect if G contains no induced odd cycle or the complement of an odd cycle; see Figure 1.3.

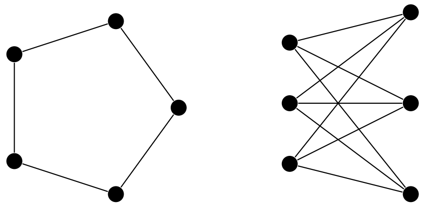


Figure 1.3: (Left) the odd cycle C_5. (Right) a perfect graph.

There is an older polyhedral relaxation to STAB(G) known as QSTAB(G). This is the polytope in $[0,1]^n$ cut out by the nonnegativities $x_i \geq 0$ and the *clique inequalities* $\sum_{i \in K} x_i \leq 1$ for each clique K in G. Since $f = 1 - \sum_{i \in K} x_i$ is idempotent mod I, that is, $f^2 \equiv f \mod I$, it is visibly a sum of squares. Therefore the clique inequalities are valid on TH(G), and STAB(G) $\subseteq$ TH(G) $\subseteq$ QSTAB(G).

1.3 A Semidefinite Approach to the K_i-Cover Problem

Chapter 3 is taken from a paper coauthored with João Gouveia [19] and submitted to *Operations Research Letters*. It is an application of theta bodies to the K_i-cover problem.

1.3.1 Problem Description

The *K_i-cover problem* is a generalization of the stable set problem discussed in Section 1.2. A graph $G = (V, E)$ is given. For any k, a k-clique in G, denoted K_i, is a collection of k nodes, each pair of which is connected by an edge in E. We define a covering relation where an i-clique H_1 is covered by an $(i-1)$-clique H_2 if $H_1 \supseteq H_2$, i.e., if H_2 is a subgraph of H_1. A K_i-cover in G is a collection C of $(i-1)$-cliques in G such that each i-clique in G is covered by some element of C. The K_i-cover problem is to find such a collection C of smallest size. In the graph G in Figure 1.4, a K_4-cover is given by $\{012, 034, 023\}$.

When $i = 2$, the K_i-cover problem is the *vertex cover problem*: to find the smallest set C of vertices such that each edge is covered by a vertex in C. To make the connection to the stable set problem, note that $C \subseteq V$ is a vertex cover if and only if $V \setminus C$ is a stable set.

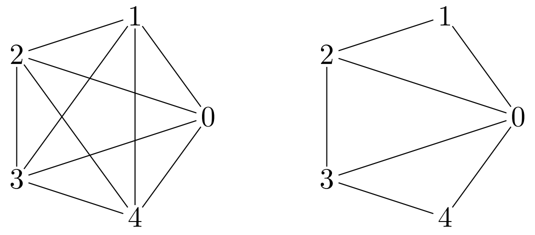


Figure 1.4: A graph with 5 copies of K_4: 0123,0124,0134,0234,1234; along with an example K_4-cover.

Therefore, the problems are essentially equivalent. In fact, the polytopes defined by these problems as, in section 1.1, are congruent via the involution $x_j \mapsto 1 - x_j$ in each coordinate. Similarly, a set C of K_{i-1}s is a K_i-cover if and only if its complement C^c is K_i-free.

We will need the description of the ideal I of the variety of K_i-free sets. We only state it here; for the derivation see Chapter 3. Let X be the collection of K_{i-1}s in G and Y the collection of K_is in G. Then we have

$$I = \langle x_H^2 - x_H \, \forall H \in X \, ; \, \prod_{H \subseteq H'} x_H \, \forall H' \in Y \rangle.$$

To understand this, recall that we have a variable x_H for each $(i-1)$-clique H in G. The functions that generate I come in two forms. The type $x_H^2 - x_H = 0$ enforces the constraint that $x_H \in \{0, 1\}$, i.e., that x is a characteristic vector. The type $\prod_{H \subseteq H'} x_H = 0$ ensures that x actually comes from a K_i-free set.

1.3.2 Background

The stable set and vertex cover problems have been studied in many contexts. The K_i-cover generalization was first studied in Conforti et al [15], wherein the associated polytope $P_i(G)$ was considered and several families of facets identified. As in Section 1.1, $P_i(G) \subseteq \mathbb{R}^N$ is the convex hull of characteristic vectors of K_i-covers, where N is the number of K_{i-1}s in G. Conforti et al provided polynomial-time *separation oracles* for many of these families of facets. A separation oracle for a family $\mathcal{F}$ of facets is a decision procedure which takes a point x and decides whether x satisfies each facet $F \in \mathcal{F}$, or whether x lies outside some facet $F \in \mathcal{F}$. Since a typical family $\mathcal{F}$ will contain exponentially many elements, it is not possible in general to enumerate the $F \in \mathcal{F}$ and check them one by one, making such an oracle a nontrivial result.

Conforti et al left open the existence of oracles for several families of facets, including the family associated with the *K_i-p-holes*. A graph H is a K_i-p-hole if it contains p copies of K_i arranged in a cycle, with neighboring K_i sharing a common K_{i-1}. This does not determine H up to isomorphism; see Figure 1.5 for three nonisomorpic K_3-9-holes. To understand the facet inequality associated to a K_i-p-hole, consider each graph H in Figure 1.5. To construct a K_3-cover of H of minimum size, we need to pick an edge from each K_3. Each edge is shared by at most two K_3s, so we need at least 5 edges. Therefore $\sum_{e \subseteq H} x_e \geq 5$ is valid on the polytope $P_3(G)$ for any graph G containing H as a subgraph. The inequality for general K_i-p-holes is derived from a similar argument and is given by $\sum_{e \subseteq H} x_e \geq \lceil \frac{p}{2} \rceil$. It defines a facet of $P_i(G)$ for $i \geq 3$ and odd p.

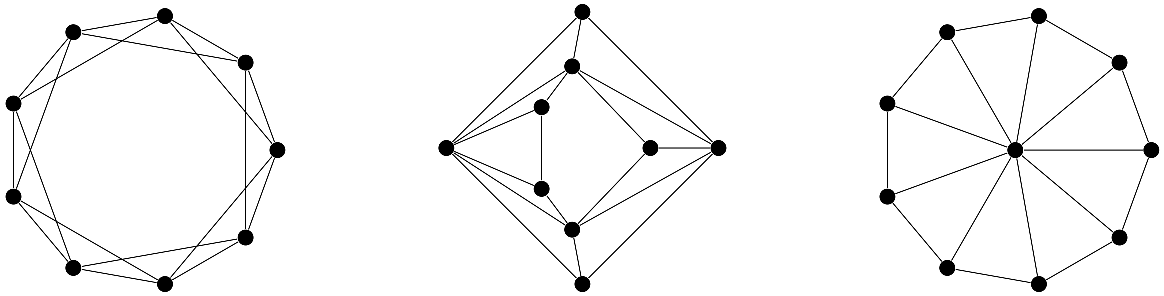


Figure 1.5: Three non-isomorphic K_3-9-holes.

1.3.3 Results

- Our main result is that the family of facets corresponding to K_i-p-holes is valid on the theta body $\mathrm{TH}_{\lceil i/2 \rceil}(I)$. Therefore, for fixed i, we have a polynomial-time algorithm to optimize over a relaxation at least as tight as the polyhedron described by this family. To show that the K_i-p-hole facets are valid on $\mathrm{TH}_{\lceil i/2 \rceil}(I)$, we exhibit a set of polynomials which are *idempotent* mod I, and whose sum is the facet-defining inequality.

We also prove two results about the triangle free problem, the case $i = 3$ of the K_i-free problem.

- For the triangle free problem on the complete graph K_n, we show that $P_3(K_n) \subsetneq \mathrm{TH}_k(I)$ for $k < n/2$. That is, it takes at least $n/2$ steps for the theta body heirarchy to converge to the triangle free polytope for K_n, and therefore this heirarchy does not give a polynomial-time algorithm for the triangle free problem. We show this by observing that the cut polytope and triangle free polytope of K_n share a facet, and apply a result of Laurent [34] that the theta heirarchy takes at least $n/2$ steps to reach this facet.

- We show that there is an *integrality gap* of 1/2 for the triangle cover problem's second theta relaxation. That is, if we optimize in the all-ones direction in the triangle cover problem, we have

  $$\min\{1 \cdot x : x \in \mathrm{TH}_2(I)\} \geq \frac{1}{2} \min\{1 \cdot x : x \in P_3(G)\}$$

  for all G. We prove this by applying a result of Krivelevich [32] on a fractional relaxation of the same problem, and proving that $\mathrm{TH}_2(I)$ is contained in this fractional relaxation.

1.3.4 Comments

In our main result, we consider the theta body $\mathrm{TH}_{\lceil i/2 \rceil}(I)$. The reason for choosing this level of the theta heirarchy is that the generators of the K_i-free ideal are polynomials of degree i. Therefore, they can't capture any inequalities on $\mathrm{TH}_k(I)$ when $2k < i$, so $\mathrm{TH}_{\lceil i/2 \rceil}(I)$ is the first theta body that it makes sense to consider.

We don't fully address the question raised in Conforti et al [15] of whether there is a polynomial time separation oracle for the family $\mathcal{F}$ of K_i-p-hole facets. We show that these facets are valid on $\mathrm{TH}_{\lceil i/2 \rceil}(I)$, and that we can check membership of $\mathrm{TH}_{\lceil i/2 \rceil}(I)$ in polynomial time. However, this does not give a separation oracle for $\mathcal{F}$. Indeed, let Q be the body defined as the intersection of all $F \in \mathcal{F}$. We have $P_i(G) \subseteq \mathrm{TH}_{\lceil i/2 \rceil}(I) \subseteq Q$, so in terms of an approximation to $P_i(G)$, the theta body is tighter, and we can optimize over it in polynomial time (for fixed i). However, this doesn't allow us to optimize over exactly Q in polynomial time. In fact, for $i = 2$, the stable set case, we have a similar phenomenon: $\mathrm{STAB}(G) \subseteq \mathrm{TH}(G) \subseteq \mathrm{QSTAB}(G)$; recall 1.2.2 for the definitions of these bodies. In this case it is NP-hard to optimize over either $\mathrm{STAB}(G)$ or $\mathrm{QSTAB}(G)$, while $\mathrm{TH}(G)$ can be optimized over in polynomial time.

Conforti et al [15] found additional facets of $P_i(G)$, associated to other subgraphs of G, for which no separation oracle is currently known. Numerical experiments indicated that they did not appear to be valid on $\mathrm{TH}_{\lceil i/2 \rceil}(I)$, but they may be valid on higher theta bodies.

1.4 Sums of Squares on the Unit Hypercube

Chapter 4 is taken from a paper co-authored with Greg Blekherman and João Gouveia [10], currently being written. It describes a criterion for certain polynomials to be sums of squares mod the ideal of the unit hypercube.

1.4.1 Problem Description

We begin by describing the *max cut problem.* Let $G = (V, E)$ be a graph, and $n = |V|$. A *cut* in G is a partition of V into two sets A and B; variously, a cut also refers to the set C of edges between A and B. The max cut problem is to find a cut containing the maximum number of edges. Alternatively, if a set of weights (w_{ij}) is given, it can refer to finding the cut C maximizing $\sum_{ij \in C} w_{ij}$. For example, in the complete graph K_n, max cuts with unit weights are attained by partitioning the vertices into two sets of size $\frac{n}{2}$ if n is even, or sizes $\frac{n-1}{2}$ and $\frac{n+1}{2}$ if n is odd.

As in Section 1.2, we can associate a cut C with its charactistic vector $\chi_C \in \{0,1\}^E$. Let $\mathcal{C}$ be the set of all characteristic vectors of cuts. We let the *max cut polytope* be $P(G) = \operatorname{conv}(\mathcal{C})$. Then the max cut problem becomes $\max\{\sum_{ij} w_{ij}x_{ij} : x \in P(G)\}$. This is called the *edge model* of the polyhedral formulation of the max cut problem. In the case $G = K_n$, the max cuts have size $\lfloor \frac{n^2}{4} \rfloor$, so the function $f(x) = \lfloor \frac{n^2}{4} \rfloor - \sum x_{ij} \geq 0$ on $\mathcal{C}$. In fact, for odd n it defines a facet of the max cut polytope $P(K_n)$. To determine how well the kth theta body captures this inequality, we can ask for the smallest λ such that $\lambda - \sum x_{ij}$ is k-sos mod $I(\mathcal{C})$. As k increases, λ will converge to the true minimum $\lfloor \frac{n^2}{4} \rfloor$.

We can define another model of cuts, the *vertex model.* Let C be a cut arising from a partition $[A,B]$ of vertices. Associate to C the characteristic vector $\chi_A \in \mathcal{H} = \{0,1\}^n$, the unit hypercube. This preserves all the information contained in C: given a vector $y \in \mathcal{H}$, we can recover $A = \{i \in [n] : y_i = 1\}$ and $B = [n] \setminus A$. Note that the cut C also arises from the partition $[B,A]$, so each cut is represented by two vectors in $\mathcal{H}$. We can translate between the vertex and edge models. Given a point $y \in \mathcal{H}$, define $x = \tau(y) \in \{0,1\}^{\binom{n}{2}}$ by $x_{ij} = (y_i - y_j)^2$. This transformation converts cuts represented by the vertex model to cuts represented by the edge model. If $y_i = y_j$, then the edge ij is not a cut edge so $x_{ij} = 0$. If $y_i \neq y_j$, then ij is a cut edge, and indeed $x_{ij} = (y_i - y_j)^2 = 1$.

Recall that $f(x) = \lfloor \frac{n^2}{4} \rfloor - \sum x_{ij}$ defines a facet of the max cut polytope $P(K_n)$ in the edge model when n is odd. It turns out that $g(y) = (\lfloor \frac{n}{2} \rfloor - \sum y_i)(\lceil \frac{n}{2} \rceil - \sum y_i)$ defines this same function on the vertex model under the correspondence given by τ. To see this, one may apply τ and reason algebraically. Alternatively, this can be seen by considering a cut $C = [A,B]$ represented by x and y, and evaluating $f(x) = g(y)$ by hand. Results about the sum of squares rank of $g(y)$ can be converted into results about $f(x)$. If $f(x)$ is k-sos mod $I(\mathcal{C})$ for some k, the sum of squares can be written in the variables x_{ij}. Writing the same sum of squares expression, substituting $(y_i - y_j)^2$ for x_{ij} everywhere, shows that $g(y)$ is also k-sos mod $I(\mathcal{H})$. Thus the max cut problem is an example of optimizing a polynomial over the hypercube $\mathcal{H}$.

In Chapter 4 we consider the general problem of optimizing a polynomial function over the hypercube $\mathcal{H}$. We consider functions on $\mathcal{H}$ satisfying certain properties. First, since the ideal $I = I(\mathcal{H}) = \langle x_i^2 - x_i;\, 1 \le i \le n\rangle$, each function $f \in \mathbb{R}[\mathcal{H}] = \mathbb{R}[x]/I$ can be represented using squarefree monomials. This gives a notion of *degree* mod I that coincides with the notion of degree we have used in the definition of k-sos. We define a notion of *proper divisibility* mod I in which f is properly divisible by g mod I if there exists h such that $f = gh$ mod I and $\deg(f) = \deg(g) + \deg(h)$. Here $f, g, h \in \mathbb{R}[\mathcal{H}]$. We also consider the symmetric group S_n acting on $\mathbb{R}[\mathcal{H}]$ by permuting variables. A function $f \in \mathbb{R}[\mathcal{H}]$ is *invariant* if it is fixed by S_n. An invariant function has the same value on each *level* $\{x \in \mathcal{H} : \sum x_i = t\}$ of the hypercube, and is determined by these values. We say that f *vanishes to degree* k *on level* t if f is properly divisible by $(\sum x_i - t)^k$ but not by $(\sum x_i - t)^{k+1}$. For example, the function $g(y) = (\lfloor \frac{n}{2} \rfloor - \sum y_i)(\lceil \frac{n}{2} \rceil - \sum y_i)$ given above vanishes to degree 2 on level $\frac{n}{2}$ when n is even. When n is odd, $g(y)$ vanishes to degree 1 on levels $\frac{n-1}{2}$ and $\frac{n+1}{2}$.

We also define a generalization of k-sos mod I from Section 1.2. This notion is denoted (d_1, d_2)-sos mod I, or being d_2*-sos with* d_1*-sos multipliers.* A function $f \in \mathbb{R}[\mathcal{H}]$ is said to be (d_1, d_2)-sos mod I if there exist $g_1, g_2 \in \mathbb{R}[\mathcal{H}]$, with $g_1 > 0$ on $\mathcal{H}$, such that g_i is d_i-sos mod I, and $fg_1 \equiv g_2$. Since any f which is (d_1, d_2)-sos mod I is also nonnegative on $\mathcal{H}$, this leads to a generalization of the theta body heirarchy which can still be optimized over in polynomial time.

1.4.2 Background

A previous semidefinite approximation to the max cut problem was given by Goemans and Williamson [18]. They used semidefinite programming to construct a randomized algorithm that produces solutions with objective value at least 0.878 times the true optimum. However, since the max cut problem is NP-hard, an exact semidefinite algorithm should not be efficient. Indeed, Laurent [34] showed that the Lasserre relaxations of the max cut polytope $P(K_n)$, equivalent to the theta body heirarchy in this case, do not capture the facet $\lfloor \frac{n^2}{4} \rfloor - \sum x_{ij} \ge 0$ until at least the $n/4$th step of the heirarchy.

The idea of sums of squares multipliers came from work in real algebraic geometry, where the same idea is considered for polynomials on $\mathbb{R}^n$. It is a celebrated result of Hilbert [27] that for homogeneous polynomials f in n variables of even degree d, the implication

$$f \geq 0 \textit{ everywhere if and only if } f \textit{ is a sum of squares}$$

holds exactly for $d = 2$,$n = 2$, and the special case $(n, d) = (3, 4)$. In all other cases, being a sum of squares implies nonnegativity, but not conversely. Artin [3] generalized this result by showing that every nonnegative polynomial can be written as a sum of squares using sum of squares multipliers as discussed above. Upper bounds are known for the degrees of these multipliers, but no lower bounds are known so far.

1.4.3 Results

Our main results are as follows.

- Let f be an invariant function on $\mathcal{H}$ which vanishes to odd degree on level t, for some $t \leq n/2$ with $\deg(f) \leq t$. Then f is not k-sos mod I for $k \leq t$. To show this, we consider the vector space $\mathbb{R}[\mathcal{H}]$ as an S_n*-module* and decompose it into *irreducible representations*. Most of these irreducible representations cannot contribute to a sum of squares which vanishes on level t. The ones that can contribute each introduce a factor of $(t - \sum x_i)$ of even degree, leading to a parity contradiction.

- As a corollary, we show that such an f is not (d_1, d_2)-sos for $d_1 \leq \min\{\frac{n-\deg f}{2}, t\}$ and $d_2 \leq t$.

- We prove an upper bound on d_1 such that f is (d_1, d_2) mod I in terms of the Hilbert function of $\mathbb{R}[x]/I$, that is, the dimensions of the degree-graded components of $\mathbb{R}[x]/I$.

- We reprove the result of Laurent, mentioned in the Background and used in Chapter 3, that the theta body heirarchy does not converge to the max cut polytope $P(K_n)$ until at least the $n/4$th step. To do this, we use the translation between the vertex and edge models of max cuts on K_n. The function $g(y)$ is invariant and vanishes on level $t = \lfloor n/2 \rfloor$. Therefore it cannot be a sum of squares of degree $\leq n/2$.

- We give an explicit example of a degree-4 polynomial on $\mathbb{R}^n$ which is nonnegative, but requires sum of squares multipliers of degree at least $n/2$ to be written as a sum of squares. This polynomial is actually $g(y)$, extended from $\mathcal{H}$ to $\mathbb{R}^n$ by adding a positive scalar multiple of the sum of the ideal generators $\sum(y_i^2 - y_i)^2$. That is, we use $h(y) = g(y) + \lambda \sum(y_i^2 - y_i)^2$ for a small $\lambda > 0$. This gives a new constructive example of a polynomial that requires degree-$O(n)$ sum of squares multipliers.

1.4.4 Comments

For a fixed f, the question of whether f is (d_1, d_2)-sos can be solved by semidefinite programming (either mod I or globally). However, we lose the ability to optimize the constant coefficient of f such that f is (d_1, d_2)-sos, as introducing sos multipliers changes the problem to quadratic optimization, which is NP-hard. This doesn't affect complexity considerations, as in general we have a bound on the quantity we want to optimize (e.g. it is known a priori to lie in some interval $[0, C]$). Then, we can use binary search to home in on the value to any fixed precision. However, it is still an interesting geometric question whether the set Σ_{d_1,d_2} of (d_2, d_2)-sos functions is the feasible region of a semidefinite program.

The representation-theoretic decomposition of $\mathbb{R}[\mathcal{H}]$ in this chapter can be performed in a wider setting. For example, we performed calculations by hand indicating that the *matching polytope* of K_9 does not equal its second theta body. Carrying these out for general n and more complicated varieties will require more sophisticated representation theory. We give a partial example of this in Chapter 5.

1.5 The Representation Theory of Matchings

Chapter 5 is taken from a paper co-authored with Monty McGovern [36], currently being written. It describes the decomposition of the space of matchings on K_n into irreducible representations of S_n.

1.5.1 Problem Description

A *perfect matching* in a graph $G = (V, E)$ is a subset $M \subseteq E$ such that each $v \in V$ meets some $e \in M$, and where each pair $e_1, e_2 \in M$ is disjoint. Let $\mathcal{M}_n$ be the set of all perfect matchings in the complete graph K_{2n}. The symmetric group S_{2n} acts naturally on $\mathcal{M}_n$ by permuting the labels of vertices. We wish to decompose $\mathbb{C}[\mathcal{M}_n]$ into *irreducible representations* of S_{2n}.

We quickly recall the necessary concepts from representation theory. A *representation* of a group G is a vector space V on which G acts by linear transformations: that is, a group homomorphism $\rho : G \to GL(V)$. A *subrepresentation* of V is a vector subspace $W \subseteq V$ for which $GW = W$. A representation V is *irreducible* if the only subrepresentations of V are 0 and V itself. For the symmetric group S_n, the irreducible representations are called the *Specht modules* S^λ and are in bijection with the partitions λ of n. For their construction, see the textbooks [41] and [17].

Suppose V is a representation of a group G, and H is a subgroup of G. We can make V into a representation of H by simply forgetting the elements of $G \setminus H$; this operation is known as *restriction* and is denoted $\mathrm{Res}^G_H(V)$. A similar operation is *induction*. Here, we are given a group G, a subgroup H, and a representation V of H. We can define the *induced representation* $\mathrm{Ind}^G_H(V) = \mathbb{C}G \otimes_{\mathbb{C}H} V$. The theorem we need here is the *Branching rule* [41], that $\mathrm{Res}^{S_n}_{S_{n-1}}(S^\lambda) = \sum_{\lambda^-} S^{\lambda^-}$ and $\mathrm{Ind}^{S_n}_{S_{n-1}}(S^\lambda) = \sum_{\lambda^+} S^{\lambda^+}$. The sums λ^- and λ^+ range over all partitions obtained from λ by removing, respectively adding, a single cell.

1.5.2 Background

In [8] Barbasch and Vogan showed that $\mathbb{C}[\mathcal{M}_n] \cong \oplus_\lambda S^\lambda$, where the sum is over all partitions $\lambda \vdash 2n$ consisting of even parts. Their key lemma uses the restriction and induction operations to link the $n-1$ and n cases of the theorem:

$$\mathrm{Ind}^{S_{2n-1}}_{S_{2n-2}}(\mathbb{C}[\mathcal{M}_{n-1}]) \cong \mathrm{Res}^{S_{2n}}_{S_{2n-1}}(\mathbb{C}[\mathcal{M}_n]).$$

From this lemma, it is straightforward to use induction on n to prove the desired decomposition of $\mathbb{C}[\mathcal{M}_n]$.

From the perspective of representation theory, this settles the matter, because it describes $\mathbb{C}[\mathcal{M}_n]$ up to isomorphism. However, to use this in determining sums of squares mod I, we need to know the explicit decomposition of $\mathbb{C}[\mathcal{M}_n]$ into irreducible representations. That is, for each valid $\lambda \vdash 2n$, we require an embedding $S^\lambda \to \mathbb{C}[\mathcal{M}_n]$. From Barbasch and Vogan, we know such a map exists, but not necessarily how to find it.

A natural idea is to "increase 2 to m" in the definition of matching. There are two natural ways to do this: perfect matchings on hypergraphs, and permutations of cycle type (m^n). We describe hypergraphs first.

An *m-uniform hypergraph* is a collection $G = (V, E)$ of vertices and edges, where an edge is a collection of m vertices. The case $m = 2$ gives the usual definition of a graph. Denote the complete m-uniform hypergraph by $K_n^{(m)} = ([n], \binom{[n]}{m})$. A perfect matching in $K_{mn}^{(m)}$ is a collection of n disjoint edges whose union is $[mn]$. Equivalently, it is a set partition of $[mn]$ into n sets of size m. Let $\mathcal{M}_{m,n}$ be the set of all matchings in $K_n^{(m)}$.

The second generalization is based on permutations of a fixed *cycle type*. Let $\mathcal{Z}_{m,n} \subset S_{mn}$ be the set of all permutations consisting of n disjoint m-cycles, such as $(12)(34)(56)(78)$ for $n = 4, m = 2$. Let S_{mn} act on $\mathcal{Z}_{m,n}$ by conjugation: $\sigma \cdot \tau := \sigma\tau\sigma^{-1}$ for $\sigma \in S_{mn}$, $\tau \in \mathcal{Z}_{m,n}$. It turns out that $\mathbb{C}[\mathcal{M}_n] \cong \mathbb{C}[\mathcal{Z}_{2,n}]$; for example, the permutation $(12)(34)(56)(78) \in \mathcal{Z}_{2,4}$ is mapped to the matching also written $(12)(34)(56)(78)$. This suggests an alternate generalization of $\mathbb{C}[\mathcal{M}_n]$. Fix m and let $\mathcal{Z}_{m,n}$ be the set of all permutations in S_{mn} of cycle type m^n. For example, when $(m, n) = (4, 2)$, an example is $(1234)(5678)$.

1.5.3 Results

- We find an explicit isomorphism between $\mathbb{C}[\mathcal{M}_n]$ and $\oplus_\lambda S^\lambda$. To construct the isomorphism, we give a map from each S^λ to $\mathbb{C}[\mathcal{M}_n]$ by imposing a matching structure on the elements of S^λ. We check that this map is injective; by the Barbasch-Vogan result, it is an isomorphism.

- We also extend the induction lemma in Barbasch and Vogan's proof to both of our generalized cases:

  $$\mathrm{Ind}_{S_{m(n-1)}\times S_{m-1}}^{S_{mn-1}}(\mathbb{C}[\mathcal{M}_{m,n-1}]\otimes 1)\cong \mathrm{Res}_{S_{mn-1}}^{S_{mn}}(\mathbb{C}[\mathcal{M}_{m,n}])$$
  $$\mathrm{Ind}_{S_{m(n-1)}}^{S_{mn-1}}(\mathbb{C}[\mathcal{Z}_{m,n-1}]\otimes 1)\cong \mathrm{Res}_{S_{mn-1}}^{S_{mn}}(\mathbb{C}[\mathcal{Z}_{m,n}])$$

  We use this lemma to explicitly compute the decompositions of $\mathbb{C}[\mathcal{M}_{m,n}]$ and $\mathbb{C}[\mathcal{Z}_{m,n}]$ for small values of m and n.

- In particular, we compute enough of the decomposition of $\mathbb{C}[\mathcal{M}_{3,n}]$ to show that it is not multiplicity-free for $n \geq 5$. Combined with the growing complexity of the decompositions for $n \leq 6$, this would suggest that no nice pattern analogous to the $m = 2$ case exists.

1.5.4 Comments

We extended the induction lemma to our generalized cases. By analogy with the case $m = 2$, one might guess that $\mathbb{C}[\mathcal{M}_{m,n}]$ or $\mathbb{C}[\mathcal{Z}_{m,n}]$ would be isomorphic to $\oplus_\lambda S^\lambda$, where the sum is over all $\lambda \vdash mn$ with all parts divisible by m. In fact, even for $(m, n) = (3, 2)$ we get $\mathbb{C}[\mathcal{M}_{3,2}] \cong S^{(6)} + S^{(4,2)}$, instead of $S^{(6)} + S^{(3,3)}$ as conjectured. The decomposition of $\mathbb{C}[\mathcal{Z}_{3,2}]$ is even further from the conjectured $(6), (3, 3)$, instead consisting of the partitions $(6), (2, 1, 1, 1, 1), (2, 2, 2), (4, 2), (3, 1, 1, 1)$, and $(4, 1, 1)$, each with multiplicity 1. So far we have not been able to find the pattern behind the decomposition, and are not hopeful that one exists. However, if a pattern is identified, it should be straightforward to use the generalized induction lemma to prove it.

1.6 A Note on Notation

The following chapters were published as separate papers, and in some cases refer to different aspects of the same or related problems. As such, the chapters use different notation to fit the topic at hand, so the same object may have different notations in different chapters.

Chapter 2

BOOTSTRAP PERCOLATION ON THE HAMMING TORUS

2.1 Introduction

Bootstrap percolation is a simple growth model, introduced to understand nucleation and metastability in physical processes such as crack formations, clustering, and alignment of magnetic spins. It was introduced in 1979 by Chalupa, Leath and Reich [14]. For more applications and background see surveys by Adler and Levi [1] and Holroyd [29].

Given a graph $G = (V, E)$, *bootstrap percolation* with *threshold* θ is the following discrete-time growth process: given an initial configuration $\omega \in \{0,1\}^V$, an increasing sequence of configurations $\omega = \omega_0, \omega_1, \ldots$ is defined by

$$\omega_{j+1}(v) = \begin{cases} 1 & \text{if } \omega_j(v) = 1 \text{ or } \sum_{w \sim v} \omega_j(w) \geq \theta \\ 0 & \text{else} \end{cases}$$

and ω_∞ is the pointwise limit of ω_j as $j \to \infty$. The initial configuration ω is random; $\{\omega(v) : v \in V\}$ is a collection of i.i.d. Bernoulli random variables with parameter p. A natural quantity to study is $\mathbb{P}_p(\omega_\infty \equiv 1)$. Indeed, first results in this area were by van Enter [48] and Schonmann [42], who proved that for the lattice $\mathbb{Z}^d$ this probability is either 1 or 0 according to whether $\theta \leq d$ or $\theta > d$. Following the seminal work of Aizenman and Lebowitz [2], it became clear that this process is even more interesting on large *finite* graphs. For a family of graphs depending on a single parameter n, with the number of vertices going to infinity as n increases, we assume that $p = p(n)$, and study the dependence on n of the critical probability p_c defined by

$$\mathbb{P}_{p_c}(\omega_\infty \equiv 1) = 1/2.$$

We mention only a few prominent results on how p_c scales with n. Let $[n] = \{1, \ldots, n\}$. For a large lattice cube $[n]^d \subset \mathbb{Z}^d$ (where each point is connected to the nearest $2d$ points)

, Aizenman and Lebowitz [2] proved that p_c behaves as $(\frac{1}{\log n})^{d-1}$ when $\theta = 2$, and later Cerf and Cirillo [12] and Cerf and Manzo [13] established the scaling $(\log_{\theta-1} n)^{-d+\theta-1}$ for $3 \leq \theta \leq d$; here $\log_{\theta-1}$ denotes the $(\theta-1)$'st iteration of the logarithm. For the hypercube $\{0,1\}^n$, Balogh and Bollobás [4] proved that the scaling for p_c is $n^{-2}4^{-\sqrt{n}}$ when $\theta = 2$; by contrast, for the very large threshold $\theta = \lceil n/2 \rceil$, the *majority bootstrap percolation* studied by Balogh, Bollobás, and Morris [7], p_c is close to $1/2$.

Such scaling results do not tell the whole story. They suggest the existence of an *order parameter*, a function of p and n whose size determines whether $\mathbb{P}_p(\omega_\infty \equiv 1)$ is small or close to 1; e.g., on a lattice square $[n]^2$, such a function is $p \log n$. This leads to two natural questions: Does the probability exhibit a sharp jump from 0 to 1 as the order parameter increases? Does the location of the (purported) sharp jump converge as n increases? (There are good reasons to expect the answer to the first question to be positive in surprising generality [16].)

In a major breakthrough, Holroyd [28] established a positive answer to both questions in the lattice square case, and proved that $p_c \sim \frac{\pi^2}{18 \log n}$. This celebrated theorem contradicted conjectures based on simulations, which is due to the fact that $p_c \log n$ converges to its limit very slowly, as about $1/\sqrt{\log n}$ [24]. For lattice cubes $[n]^d$, $d \geq 3$ and $2 \leq \theta \leq d$, the sharp transition was established by Balogh, Bollobás, Duminil-Copin, and Morris [6, 5].

Besides varying the dimension of the lattice or the threshold, one can also vary the neighborhood of a point. For example, Holroyd, Liggett and Romik [30] consider the lattice square $[n]^2$, with the "cross" neighborhood of a point that consists of $k-1$ points in each of the 4 axis directions, and $\theta = k$. In this case, $p_c \sim \frac{\pi^2}{3k(k+1)\log n}$.

In this paper we consider bootstrap percolation on the *Hamming torus* (or Hamming graph) with vertex set $V = [n]^d$, where $v \in V$ and $w \in V$ are adjacent iff $v - w$ has exactly one nonzero coordinate. In $d = 2$, this graph could be interpreted as taking the Holroyd-Liggett-Romik neighborhood [30] with $k = \infty$. For any d, the neighborhood of a point v is the union of all d lines through v parallel to the axes. We emphasize, however, that the threshold θ remains fixed as n increases (although some of our results assume that θ is large). Other models of percolation, including bond percolation [11, 47] and site percolation [43], have been considered on the Hamming torus, and were shown to exhibit interesting behavior due to the large neighborhood sizes relative to nearest-neighbor lattices and hypercubes. For the same reason, we expect qualitatively different transition phenomena in bootstrap percolation on the Hamming torus from those described above. First, the critical probability is much smaller. In fact, our results suggest that p_c is of the order $n^{-\alpha}$, for some critical exponent $\alpha > 1$. We are able to determine α exactly in a few cases, and give estimates otherwise. Moreover, we expect that varying the order parameter $n^{\alpha}p$ does *not* lead to a sharp jump of $\mathbb{P}_p(\omega_\infty \equiv 1)$ from 0 to 1; instead, this probability gradually approaches 0 (resp. 1) as the order parameter approaches 0 (resp. ∞). When $d = 2$ this is easy to demonstrate for arbitrary θ, but when $d \geq 3$ the combinatorics are quite difficult even when α is known exactly. Nevertheless, we succeeded in analyzing the case $d = \theta = 3$, which has $\alpha = 2$: we give an explicit formula for the limit of $\mathbb{P}_p(\omega_\infty \equiv 1)$ when $pn^2 = a \in (0, \infty)$. See [31], Theorem 3.2, for an analogous result for bootstrap percolation on Erdös-Rènyi random graphs.

Moreover, in dimensions $d \geq 3$ we find two distinct critical exponents. When p is much smaller than $n^{-1-d/\theta}$, the model does not accomplish much; with high probability it does not even fill a single line. When p is much larger than $n^{-1-d/\theta}$, but smaller than $n^{-1-2/\theta-c'/\theta^{3/2}}$, for large enough θ, with high probability some lines become open, but no two dimensional subgraphs do, and thus $\mathbb{P}_p(\omega_\infty \equiv 1) \to 0$. When $p > n^{-1-2/\theta-c''/\theta^{3/2}}$, and θ is large enough, $\mathbb{P}_p(\omega_\infty \equiv 1) \to 1$. Here, $0 < c'' < c'$ are constants depending on d.

It remains an open question for $\theta > 2$ whether the critical exponents for the appearance of open subspaces with dimension i are distinct for each $2 \leq i \leq d$. However, in subsequent work, Slivken has proven that for $\theta = 2$, there are distinct critical exponents for the appearance of open subspaces with dimension $2i$ for $1 \leq i < \sqrt{d}$ [44].

2.2 Statement of Results

Let $\mathcal{F}$ be a family of subsets of $[n]^d$. Then

$$\mathbb{P}_p(\exists F \in \mathcal{F} : \ \omega_\infty|_F \equiv 1)$$

is a nondecreasing function in p. For $\mathcal{F}_i$ the collection of i-dimensional subgraphs of G, there exists a threshold function $p_c(i,d)$ such that

$$\mathbb{P}_{p_c(i,d)}(\exists F \in \mathcal{F}_i : \ \omega_\infty|_F \equiv 1) = 0.5.$$

If $\omega_j(v) = 1$ we say v is open at step j, and a set $S \subseteq V$ is open if each $v \in S$ is open, i.e., $\omega_j|_S \equiv 1$.

For $i = 0$ we have an additional critical probability $p_c^*(0,d)$. We would like to define it to be the threshold function for $\omega_\infty \not\equiv \omega_0$; unfortunately, this is not an increasing event. Instead, we define the event

$$\text{Above Threshold} = \left\{ \exists v : \sum_{w \sim v} \omega_0(w) \geq \theta \right\}$$

and $p_c^*(0,d)$ to be the p for which $P_p(\text{Above Threshold}) = 0.5$.

We write $f(n) \sim g(n)$ if $\frac{f(n)}{g(n)} \to 1$. We conjecture that for every $\theta, i, d \in \mathbb{N}$ with $i \leq d$, there exists $a_c = a_c(\theta, i, d)$ and $\alpha_c = \alpha_c(\theta, i, d)$ such that

$$p_c(i,d) \sim a_c n^{-\alpha_c}.$$

Moreover there exists a nondecreasing function $G = G(\theta, i, d) : \mathbb{R}^+ \to [0,1]$ such that $G(x) \to 0$ as $x \to 0$, $G(x) \to 1$ as $x \to \infty$, and if $p = an^{-\alpha_c}$ then

$$\mathbb{P}_p(\exists F \in \mathcal{F}_i : \ \omega_\infty|_F \equiv 1) \sim G(a).$$

We are able to prove that this is the case for $d = 2$.

Theorem 2.2.1. *Let $d = 2$, $k = \lceil \theta/2 \rceil > 1$ and $p = an^{-1-\frac{1}{k}}$. Then*

$$\mathbb{P}(\omega_\infty \equiv 1) \to \begin{cases} 1 - e^{-2a^k/k!} & \textit{if } \theta \textit{ is odd,} \\ (1 - e^{-a^k/k!})^2 & \textit{if } \theta \textit{ is even.} \end{cases}$$

Thus

$$p_c(2,2) = p_c(1,2) = p_c^*(0,2) = n^{-1-\frac{2}{\theta}+o(\theta^{-3/2})}.$$

Furthermore,

$$\mathbb{P}\big(\{\omega_\infty \not\equiv \omega_0\} \setminus \{\omega_\infty \equiv 1\}\big) = o(1).$$

As d increases the problem becomes more intricate. For $d = 3$, we are able to identify the limit under critical scaling when $\theta = 3$.

Theorem 2.2.2. *Let $d = 3$, $\theta = 3$ and $p = an^{-2}$ with $a > 0$. Then as $n \to \infty$*

$$\mathbb{P}_p(\omega_\infty \equiv 1) \to 1 - e^{-a^3-(3/2)a^2(1-e^{-2a})}\left[\frac{3}{2}a^2\left(\left(e^{-a}+ae^{-3a}\right)^2 - e^{-2a}\right)e^{-a^2e^{-2a}} + e^{a^3e^{-3a}}\right]. \tag{2.2.1}$$

Other three dimensional results include determining the critical exponents (β_c) for $d = 3$ and low thresholds, but not the exact constants α_c; see Section 5 for details.

Many of our results state that

$$p_c(i,d) = n^{-1-\frac{c_1(i,d)}{\theta}-\Theta(\theta^{-3/2})},$$

where $c_1 = c_1(i,d)$ is a constant. This shorthand notation means that, for a large n, we can get a lower bound and and upper boud for $p_c(i,d)$ of the stated form, with constants in the correction term $\Theta(\theta^{-3/2})$ depending on i and d.

For general $d \geq 3$, we calculate $p_c^*(0,d)$ and $p_c(1,d)$ for all $d \geq 2$ quite precisely.

Theorem 2.2.3. *Let $p = f(n)n^{-1-\frac{d}{\theta}}$ and $d, \theta \geq 3$. If $f(n) \to 0$ then*

$$\mathbb{P}(\textit{Above Threshold}) \to 0$$

and if $f(n) \to \infty$ then

$$\mathbb{P}(\exists \textit{ a line } \ell \textit{ such that } \omega_\infty|_\ell \equiv 1) \to 1.$$

Furthermore, we get good bounds on $p_c(2,d)$, the threshold for existence of two dimensional subspaces in the final configuration.

Theorem 2.2.4. *Fix d and fix θ sufficiently large depending on d. For n sufficiently large*

$$n^{-1-\frac{2}{\theta}-\frac{4d^2+3}{\theta^{3/2}}} \leq p_c(2,d) \leq n^{-1-\frac{2}{\theta}-\frac{\sqrt{8(d-2.1)}}{\theta^{3/2}}}.$$

(We have not attempted to optimize the constants $\sqrt{8(d-2.1)}$ and $4d^2+3$ in the above theorem.) The key arguments in the proof of Theorem 2.2.4 are Lemmas 2.8.1 and 2.5.1.

The higher the dimensions i and d, the more difficult it becomes to calculate $p_c(i,d)$. However, Theorems 2.2.3 and 2.2.4 are sufficient for us to get bounds on $p_c(i,d)$ for all $i, d \geq 2$.

Theorem 2.2.5. *For all $i \geq 2$ and d, and sufficiently large n,*

$$p_c(i,d) = n^{-1-\frac{2}{\theta}-\Theta(\theta^{-3/2})}.$$

Proof. It is easy to see that $p_c(i,d)$ is nondecreasing in i and decreasing in d. Also $p_c(d,d)$ is decreasing in d. To see this last inequality note that when $n \geq 3\theta$ and $d = j+1$

$$\mathbb{P}_{p_c(j,j)}(\exists \text{ at least } \theta\ i \text{ such that } \omega_\infty|_{(i,*,*,\cdots)} \equiv 1) > 1/2.$$

The event on the left hand side implies that $\omega_\infty \equiv 1$ and thus

$$p_c(j+1,j+1) \leq p_c(j,j)$$

and inductively

$$p_c(d,d) \leq p_c(3,3).$$

So

$$p_c(2,d) \leq p_c(i,d) \leq p_c(d,d) \leq p_c(3,3).$$

By Theorem 2.2.4

$$p_c(2,3) \leq n^{-1-\frac{2}{\theta}-\frac{\sqrt{7.2}+o(1)}{\theta^{3/2}}}.$$

By coupling it is easy to see that ω chosen when $p = 10\theta p_c(2,3)$ stochastically dominates the union of 10θ independent ω' chosen with $p = p_c(2,3)$. Then by the definition of $p_c(2,3)$

$$\mathbb{P}_{10\theta p_c(2,3)}(\exists \text{ at least } \theta\ i \text{ such that } \omega_\infty|_{(i,*,*)} \equiv 1) > 1/2.$$

The event on the left hand side implies $\omega|_\infty \equiv 1$ and thus

$$p_c(3,3) \leq 10\theta p_c(2,3). \tag{2.2.2}$$

And putting this all together for all $d \geq 3$ and $2 \leq i \leq d$,

$$n^{-1-\frac{2}{\theta}-\frac{4d^2+2+o(1)}{\theta^{3/2}}} \leq p_c(2,d) \leq p_c(i,d) \leq p_c(d,d) \leq p_c(3,3) \leq 10\theta p_c(2,3) \leq n^{-1-\frac{2}{\theta}-\frac{\sqrt{7.2}-o(1)}{\theta^{3/2}}},$$

which is the desired result. □

Remark 2.2.6. The above results are all asymptotic statements in n. One natural question is whether we can obtain non-asymptotic bounds on the critical parameters. Our arguments do in fact produce bounds on the critical probability for specific values of n. Keeping track of (or even stating) these bounds is quite challenging and we have made no attempt to optimize them. Different results kick in at different values of n, but all of them work if n is at least roughly $e^{\theta^{3/2}}$.

The rest of the paper is organized as follows. In Section 2.3 we prove the two-dimensional Theorem 2.2.1. In Section 2.4 we give a necessary condition for a plane to become open when $d = 3$ and in Section 2.5 we give a sufficient condition for this event for arbitrary d. Section 2.5 also features the resulting upper and lower bounds for critical exponents in three dimensions and the proof for the upper bound in Theorem 2.2.4. Section 2.6 features the proof of Theorem 2.2.2, although some details are deferred to Appendix 2.10. In Section 2.7 we study when a line is likely to become open and establish Theorem 2.2.3. In Section 2.8 we provide a lower bound on the value of p that makes it likely that a plane becomes open; this, together with results in Section 2.5, will conclude the proof of Theorem 2.2.4. We conclude with a short list of open questions in Section 2.9.

2.3 Precise two-dimensional results

In the two-dimensional case we can describe the limiting behavior exactly as $n \to \infty$. Let $k = \lceil \theta/2 \rceil$ and $p = an^{-1-\frac{1}{k}}$ for some constant a. Also assume $k > 1$; the cases $\theta = 1$ and

$\theta = 2$ are easy to work out separately. (For $\theta = 1$, $\omega_\infty \equiv 1$ if and only if $\omega_0 \not\equiv 0$; for $\theta = 2$, $\omega_\infty \equiv 1$ asymptotically if and only if ω_0 contains at least two non-collinear open points.)

Lemma 2.3.1. *Let $k = \lceil \theta/2 \rceil$ and $p = an^{-1-\frac{1}{k}}$. With probability going to 1, there are no lines with at least $k+1$ points initially open.*

Proof. The probability that a given line has $k+1$ points initially open is bounded above by $\binom{n}{k+1} p^{k+1} \leq n^{k+1} p^{k+1} \leq a^{k+1} n^{-1-\frac{1}{k}}$. Since there are $2n$ lines the union bound shows that, asymptotically almost surely, none of them have $k+1$ points. □

Lemma 2.3.2. *Fix an $\epsilon > 0$. Let $k = \lceil \theta/2 \rceil$ and $p = \epsilon n^{-1-\frac{1}{k}}$. Fix constants A, B and choose B fixed vertical (resp. horizontal) exceptional lines. With probability going to 1, there are at least A horizontal (resp. vertical) lines, which contain $k-1$ initially open points none of which are in the union of the exceptional lines.*

Proof. Each of the n horizontal lines satisfies the condition independently with probability at least

$$\binom{n-B}{k-1} p^{k-1} (1-p)^{n-k+1} = \Theta(n^{-1+\frac{1}{k}}).$$

The probability that there are at least A such lines therefore goes to 1. □

Let E_{horiz} be the event that some horizontal line contains at least k initially open points, E_{vert} the corresponding event for vertical lines, and $E_{\text{horiz}} \circ E_{\text{vert}}$ the event that the two occur disjointly.

Lemma 2.3.3. *Let $k = \lceil \theta/2 \rceil$ and $p = an^{-1-\frac{1}{k}}$. We have*

$$\mathbb{P}_p(E_{horiz} \cap E_{vert}) \setminus (E_{horiz} \circ E_{vert})) \to 0.$$

Furthermore,

$$\mathbb{P}_p(E_{horiz} \cap E_{vert}) \to (1 - e^{-a^k/k!})^2$$

and

$$\mathbb{P}_p(E_{horiz} \cup E_{vert}) \to 1 - (e^{-a^k/k!})^2.$$

Proof. The event $(E_{\text{horiz}} \cap E_{\text{vert}}) \setminus (E_{\text{horiz}} \circ E_{\text{vert}})$ happens only if some point v is open, and each of the two lines through v contains exactly $k-1$ additional open points. The probability that such a point exists is bounded by

$$n^2 p \left(n^{k-1} p^{k-1}\right)^2 = O(n^{-1+\frac{1}{k}}) \to 0.$$

This proves the first assertion.

As E_{horiz} and E_{vert} are increasing events, $\mathbb{P}_p(E_{\text{horiz}} \cap E_{\text{vert}}) \geq \mathbb{P}_p(E_{\text{horiz}})\mathbb{P}_p(E_{\text{vert}}) = \mathbb{P}_p(E_{\text{horiz}})^2$ by the FKG inequality. Conversely, the BK inequality gives $\mathbb{P}_p(E_{\text{horiz}})\mathbb{P}_p(E_{\text{vert}}) \geq \mathbb{P}_p(E_{\text{horiz}} \circ E_{\text{vert}})$. Thus $\mathbb{P}_p(E_{\text{horiz}} \cap E_{\text{vert}}) - \mathbb{P}_p(E_{\text{horiz}})^2 \to 0$. Moreover, the number of horizontal lines with at least k open points is Binomial and converges in distribution to a Poisson random variable with expectation $a^k/k!$. Thus $\mathbb{P}_p(E_{\text{horiz}}) \to 1 - e^{-a^k/k!}$, which easily ends the proof. ☐

Let G be the event that the entire graph becomes open; i.e., $G = \{\omega_\infty \equiv 1\}$.

Lemma 2.3.4. *Let $k = \lceil \theta/2 \rceil$ and $p = an^{-1-\frac{1}{k}}$. If θ is even, $\mathbb{P}_p(G) - P(E_{horiz} \cap E_{vert}) \to 0$, while if θ is odd, $\mathbb{P}_p(G) - \mathbb{P}_p(E_{horiz} \cup E_{vert}) \to 0$.*

Proof. If θ is odd the process adds no new open vertex unless there is some line with at least k vertices initially open. So $G \subseteq E_{\text{horiz}} \cup E_{\text{vert}}$. If θ is even, then by Lemma 2.3.1, $\mathbb{P}_p(G \setminus (E_{\text{horiz}} \cap E_{\text{vert}})) \to 0$.

Fix an $\epsilon > 0$ and let ω^*, ω', and ω'' be three independent configurations, the first with $p^* = (1-2\epsilon)n^{-1-\frac{1}{k}}$, and the other two are "sprinkled" with small $p' = \epsilon n^{-1-\frac{1}{k}}$. Observe that ω_0 (generated with p) stochastically dominates $\omega^* \cup \omega' \cup \omega''$.

Now suppose θ is odd and $E_{\text{horiz}} \cup E_{\text{vert}}$ occurs in ω^*. Then some line ℓ has k points open in ω^*. We now describe the events that occur with probability 1 as $n \to \infty$. By Lemma 2.3.2 there are θ lines $\{\ell'_i\}$ parallel to ℓ, each with $k-1$ points open in ω'. Moreover, again by Lemma 2.3.2, there are θ lines $\{\ell''_j\}$ perpendicular to ℓ, each with $k-1$ points, which are open in ω'' and avoid ℓ and all ℓ'_i.

Let G^* be the event that the initial configuration $\omega^* \cup \omega' \cup \omega''$ eventually causes every point to be open. We claim that if the events in the above paragraph all happen then G^*

happens. First, each point of intersection of ℓ''_j and ℓ becomes open as it sees $k-1$ open neighbors on ℓ''_j and k on ℓ. Then there are θ open points on ℓ, so ℓ becomes open. Now each point of intersection of ℓ''_j and ℓ'_i becomes open as it sees one open neighbor on ℓ, and $k-1$ additional open neighbors each on ℓ''_j and ℓ'_i. This results in θ open points on each ℓ''_i and ℓ'_i, so these 2θ lines all become open, and the entire graph becomes open in the next step.

It now follows that $\liminf \mathbb{P}_p(G) \geq \liminf \mathbb{P}_{p^*}(E_{\text{horiz}} \cup E_{\text{vert}})$, and the claim for odd θ follows by continuity (in a) of limits in Lemma 2.3.3.

Now suppose θ is even. If $E_{\text{horiz}} \cap E_{\text{vert}}$ occurs, then we may assume $E_{\text{horiz}} \circ E_{\text{vert}}$ occurs by Lemma 2.3.3. That is, there is a horizontal line ℓ_h and a vertical line ℓ_v, each with k points initially open, excluding their point of intersection. This point of intersection becomes open at the first time step.

As in the odd case, we may use sprinkling and Lemma 2.3.2 to produce θ horizontal lines ℓ'_i and θ vertical ℓ''_j, each with $k-1$ initially open points that avoid all other lines. Then every point of intersection between ℓ_h and ℓ''_j, and between ℓ_v and ℓ'_i, sees $\theta = (k+1)+(k-1)$ open sites, so it becomes open. Then ℓ_h and ℓ_v contain θ open sites, so they become open. Then every point of intersection of an ℓ'_i with an ℓ''_j sees $2 + 2(k-1) = \theta$ open sites, so becomes open. Now the entire graph becomes open in two additional steps. $\square$

Proof of Theorem 2.2.1. The claimed convergence follows from Lemmas 2.3.3 and 2.3.4. $\square$

2.4 Upper bound on critical exponent in three dimensions

It is easy to see that with $p = n^{-\alpha}$ for $\alpha > 1 + \frac{d}{\theta}$, with high probability, no points that are not initially open become open. (The expected number of vertices with at least θ open neighbors is at most $Cn^d(np)^\theta = O(n^{d+\theta-\alpha\theta}) = o(1)$.) In this section we will assume that $d = 3$ and $\theta \geq 3$ and establish a bound on α that ensures that no planes become open (and hence the entire Hamming torus does not become open) with high probability. A similar result is proved for general d in Section 2.8.

Lemma 2.4.1. *Let $d = 3$ and $\theta = 2k - 1 \geq 3$ be odd. Let $p = n^{-\alpha}$ for $\alpha > 1 + \frac{8}{3\theta - 1}$. Then $\mathbb{P}_p$(a plane becomes open) $\to 0$. The same holds for $\theta = 2k \geq 4$ when $\alpha > 1 + \frac{8}{3\theta - 2}$.*

Proof. We may assume $\theta \geq 4$, since the $\theta = 3$ bound of $\alpha > 2$ is equivalent to $\alpha > 1 + \frac{d}{\theta}$. We will prove the lemma for θ odd; the even case is similar. Define the following three conditions for a vertex v:

1. v is initially open,

2. v is on a line with at least k points initially open,

3. the neighborhood of v has at least θ points initially open.

We first prove

$$\mathbb{P}_p(\text{there exists a plane each of whose points satisfies one of (1)–(3)}) \to 0 \tag{2.4.1}$$

To prove (2.4.1), we fix a plane P, which we may assume to be the e_1, e_2-plane, and prove that the probability that all of its points satisfy one of (1)–(3) is exponentially small. Fix an $\epsilon \in (0, 1/3)$. Consider the lines perpendicular to P, horizontal lines in P, and vertical lines in P, that contain at least one initially open point. Let their respective numbers be S_1, S_2, and S_3, and note that each of these three numbers is Binomially distributed. The probability that a fixed line contains an initially open vertex is at most $np = o(1)$, so $\mathbb{P}_p(S_1 \geq \epsilon n^2)$, $\mathbb{P}_p(S_2 \geq \epsilon n)$, and $\mathbb{P}_p(S_3 \geq \epsilon n)$ are all exponentially small. With probability exponentially close to 1, the number of points in P included in one of the three types of lines is therefore at most $3\epsilon n^2$, which proves (2.4.1).

Let E_v be the event that the point v violates all three conditions (1)–(3), but that it becomes open and that no point violating these conditions becomes open earlier. It remains to show that

$$\mathbb{P}_p(E_v) = o(1/n^3). \tag{2.4.2}$$

We will denote by $\mathcal{N}(v)$ the neighborhood of a point v. If E_v occurs, then $\mathcal{N}(v)$ has m points initially open, for some $0 \leq m \leq \theta - 1$. Then $\mathcal{N}(v)$ contains $\theta - m$ other points

$w_1, \dots, w_{\theta-m}$, not initially open, which become open before v. Thus these w_i must satisfy (2) or (3). Because v violates (2), each w_i shares with v at most $k-1$ initially open neighbors. Therefore, whether w_i satisfies (2) or (3), $\mathcal{N}_i = \mathcal{N}(w_i) \setminus \mathcal{N}(v)$ must contain k initially open points.

Assume m and w_i are selected. Let N be the number of initially open points in $\mathcal{N}_i \cap \mathcal{N}_j$, for some $i \neq j$. (Note that the intersection of three or more $\mathcal{N}_i$ is empty.) Let H_b^m be the event that $\mathcal{N}(v)$ has m initially open points, $w_1, \dots, w_{\theta-m}$ exist such that $\mathcal{N}_i$ all contain k initially open points *and* that $N = b$. Then

$$P(H_0^m) \le C(np)^m n^{\theta-m} \left((np)^k\right)^{\theta-m}, \tag{2.4.3}$$

for some constant C. To estimate $P(H_b^m)$, observe that each increase of b by 1 contributes an additional factor of p and removes a factor $(np)^2$ from the right-hand side of (2.4.3). By monotonicty, we may assume $\alpha \le 2$ so $p \le (np)^2$ (recall $\theta \ge 5$ so $1 + 8/(3\theta - 1) < 2$); then $P(H_b^m) \le P(H_0^m)$ for all $b \ge 0$ and m. Furthermore, $n^k p^{k-1} = o(1)$ (since $k \ge 2$), thus the upper bound in (2.4.3) increases with m. It follows that $P(E_v)$ is bounded by the expression in (2.4.3) with $m = \theta - 1$, which gives

$$n^3 P(E_v) \le C n^{3k+2} p^{3k-2} \to 0,$$

proving (2.4.2). □

2.5 Internally spanned planes

In this section, we prove the upper bound in Theorem 2.2.4 regarding $p_c(2, d)$, the critical probability for the existence of two dimensional planes in the final configuration. We also introduce a dimension-reduction inequality that allows us to compute lower bounds on the spanning probabilities for arbitrary d and θ. To do so, we introduce the notation $\sigma_\theta(d, p)$ for the probability (as a function of n, which is suppressed in the notation) that the d-dimensional Hamming torus is spanned by bootstrap percolation with threshold θ and ω_0 chosen according to a product measure with probability p. Our first result is a lower bound on $\sigma_\theta(2, p)$, which will allow us to find lower bounds for all d later on.

Lemma 2.5.1. *Let $k = \lceil \theta/2 \rceil$ and $\liminf n^{\alpha} p = b > 0$ with $\alpha > 1 + 1/k$. Then there exists a constant $C > 0$ depending on θ and b such that for all sufficiently large n, $\sigma_{\theta}(2, p) \geq Cn^{-\beta}$ where*

$$\beta(\alpha) = \begin{cases} \alpha k^2 + a(a+1) - \alpha a(a-1) - (k+1)^2 & \theta \text{ odd} \\ \alpha k(k+1) + a(a+1) - \alpha a(a-1) - (k+1)(k+2) & \theta \text{ even}, \end{cases} \tag{2.5.1}$$

and $a = \lfloor \alpha/(\alpha - 1) \rfloor$.

Remark 2.5.2. If $\alpha = 1 + 1/k$ and $p = b/n^{\alpha}$ then $\sigma_{\theta}(2, p) \to c \in (0, 1)$ by Theorem 2.2.1, so $\beta(\alpha) = 0$ for $\alpha \leq 1 + 1/k$.

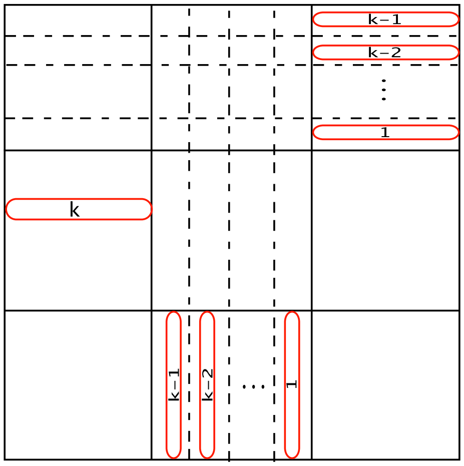


Figure 2.1: This configuration will span the two dimensional Hamming graph when $\theta = 2k - 1$ is odd. Each region bounded by solid lines is approximately $n/3 \times n/3$. The hashed lines are spaced $\frac{n}{3(k-2)}$ units apart, so each subregion has height and width on the order of n. A red oval represents the existence of at least one line (in the direction indicated) in that region with the specified number of open vertices.

Proof. Observe that the configuration in Figure 2.1 is sufficient for spanning for odd $\theta = 2k - 1$. In the figure, the two dimensional Hamming graph is first subdivided into nine regions that have dimensions $n/3 \times n/3$. The hashed lines further subdivide some of the regions, and are spaced $\frac{n}{3(k-2)}$ units apart, so each subregion has height and width on the order of n. Each red oval represents the existence of at least one line (in the direction indicated) in that region with the specified number of open vertices. To check that this configuration leads to spanning, observe that the horizontal line containing k open vertices is the first to span: after one step the vertex at the intersection of this line and the vertical line with $k-1$ open vertices becomes open, after two steps the vertex at the intersection of this line and the vertical line with $k-2$ open vertices becomes open, and so on until this line contains $2k-1$ open vertices and the entire line becomes open. As this line is made open, all of the vertical lines each gain one additional open vertex, so the vertical line with $k-1$ initially open vertices is next to span in the same fashion, followed by the horizontal line with $k-1$ open vertices and so on until all $2k-1$ lines with ovals span and cause the rest of the graph to become open. The reason for subdividing the graph into disjoint regions like we have is so that all of the events depicted are independent. Therefore, the spanning probability is bounded below as

$$\begin{aligned}\sigma_{2k-1}(2,p) &\geq \mathbb{P}_p\,(\text{configuration in Figure 2.1})\\ &= \left[1-\left(1-\frac{1}{k!}n^k p^k + o((np)^k)\right)^{n/3}\right] \times \prod_{\ell=1}^{k-1}\left[1-\left(1-\frac{1}{\ell!}(np)^\ell + o((np)^\ell)\right)^{n/3(k-2)}\right]^2.\end{aligned} \tag{2.5.2}$$

If $p \asymp n^{-\alpha}$ and $\alpha < 1+\frac{1}{k}$ then the lower bound in (2.5.2) tends to 1 as $n \to \infty$, in agreement with Theorem 2.2.1, so we assume $p \asymp n^{-\alpha}$ and $\alpha > 1+\frac{1}{k}$. In this case, the terms in the product in the last line of (2.5.2) for which $\ell \leq 1/(\alpha-1)$ either tend to 1 or (in the case of equality) are bounded away from 0 as $n \to \infty$. Therefore, by applying the bound $(1-x)^m \leq 1 - mx + m^2x^2$ for $x \in (0,1)$, we bound (2.5.2) from below by

$$C\left[n^{k+1}p^k - o(n^{k+1}p^k)\right]\prod_{\ell=a}^{k-1}\left[n^{\ell+1}p^\ell - o(n^{\ell+1}p^\ell)\right]^2 \tag{2.5.3}$$

where $a = \lfloor \alpha/(\alpha-1) \rfloor$ and the value of C here is not smaller than $(3 \cdot k!)^{-2k}$ for any $\alpha > 1 + 1/k$. We can take $p = (b/2)n^{-\alpha}$ by noting that $\sigma_\theta(2,p)$ is increasing in p, so the constant C appearing in the Lemma is not smaller than $(3 \cdot k!)^{-2k}(b/2)^{k(k+1)}$. Computing the exponent of the leading order term in (2.5.3) when $p = (b/2)n^{-\alpha}$ gives the formula for $\beta(\alpha)$ when θ is odd. A configuration similar to the one in Figure 2.1, but where there is one additional column with k initially open vertices, provides a sufficient condition for spanning when $\theta = 2k$. This leads to an expression like the one in (2.5.2), except with the first factor squared, and leads to the formula for $\beta(\alpha)$ when θ is even. $\square$

Our first application of Lemma 2.5.1 is to prove the upper bound in Theorem 2.2.4.

Theorem 2.5.3. *Fix* $d \geq 3$ *and fix* θ *large enough depending on* d *(*$\theta \geq 650(d-2.1)$ *is sufficient). For all sufficiently large* n

$$p_c(2,d) \leq n^{-1-\frac{2}{\theta}-\sqrt{8(d-2.1)}/\theta^{3/2}}$$

To prepare for the proof, we need a bound on the function $\beta(\alpha)$ in Lemma 2.5.1 that eliminates the use of the floor function. We isolate the reasoning by treating just the terms involving a.

Lemma 2.5.4. *If* $1 < \alpha \leq 2$ *and* $a = \lfloor \alpha/(\alpha-1) \rfloor$ *then*

$$a(a+1) - \alpha a(a-1) \leq \frac{1}{\alpha-1} + 1 + \frac{1}{2}(\alpha-1) \tag{2.5.4}$$

Proof. Let $\epsilon = \alpha - 1$ and suppose $\frac{1}{\epsilon} = m + u$ where $m \geq 1$ is an integer and $u \in [0,1)$. Then we can write (2.5.4) as

$$a(-\epsilon a + 2 + \epsilon) - \frac{1}{\epsilon} \leq 1 + \frac{1}{2}\epsilon,$$

so we must prove this inequality. Observe that

$$a = \left\lfloor \frac{1+\epsilon}{\epsilon} \right\rfloor = \lfloor m + u + 1 \rfloor = m + 1,$$

so we have

$$\begin{aligned} a(-\epsilon a + 2 + \epsilon) - \frac{1}{\epsilon} &= \frac{-(m+1)^2 + 2(m+u)(m+1) + m + 1 - (m+u)^2}{m+u} \\ &= 1 + \frac{u - u^2}{m+u} \leq 1 + \frac{1}{2}\epsilon. \end{aligned}$$

$\square$

Proof of Theorem 2.5.3. We can divide the d-dimensional Hamming torus into n^{d-2} disjoint 2-dimensional planes all parallel to the e_1, e_2-plane. Our goal is to show that at least one of these planes are internally spanned with high probability when $p = n^{-\alpha}$ with $\alpha = 1 + 2/\theta + \sqrt{8(d-2.1)}/\theta^{3/2}$. The number of these 2-planes that are internally spanned is binomially distributed, so we need only to show that the expected number of internally spanned planes tends to infinity. The expected number of internally spanned planes is

$$n^{d-2}\sigma_\theta(2, n^{-\alpha}) \geq Cn^{d-2-\beta(\alpha)}$$

by Lemma 2.5.1. By applying Lemma 2.5.4 we see that when $\theta = 2k-1$ is odd

$$\begin{aligned}
\beta(\alpha) &= \alpha k^2 - (k+1)^2 + a(a+1) - \alpha a(a-1) \\
&\leq \alpha k^2 - (k+1)^2 + \frac{1}{\alpha - 1} + 1 + \frac{1}{2}(\alpha - 1) \\
&= \left(1 + \frac{2}{\theta} + \frac{\sqrt{8(d-2.1)}}{\theta^{3/2}}\right)\left(\frac{\theta+1}{2}\right)^2 - \left(\frac{\theta+3}{2}\right)^2 + \frac{\theta}{2 + \sqrt{8(d-2.1)/\theta}} + 1 + \frac{1}{\theta} + \frac{\sqrt{8(d-2.1)}}{2\theta^{3/2}} \\
&\leq -\frac{\theta}{2} + \frac{3}{2\theta} + \frac{\sqrt{8(d-2.1)}}{4}(\theta^{1/2} + 2\theta^{-1/2} + \theta^{-3/2}) \\
&\qquad\qquad + \frac{\theta}{2}\left(1 - \frac{\sqrt{8(d-2.1)}}{2\theta^{1/2}} + \frac{8(d-2.1)}{4\theta}\right) + \frac{\sqrt{8(d-2.1)}}{2\theta^{3/2}} \\
&= d - 2.1 + \frac{3}{2\theta} + \frac{\sqrt{8(d-2.1)}}{4}(2\theta^{-1/2} + 3\theta^{-3/2}) \\
&< d - 2
\end{aligned}$$

where the last inequality holds for θ large relative to d, and in the fourth line we used the inequality $(1+x)^{-1} \leq 1 - x + x^2$ for $x > 0$. This implies that the expected number of internally spanned 2-dimensional planes tends to infinity with n, and completes the proof for odd θ. The proof for even θ is analogous. $\square$

The next theorem is a simple but powerful observation, which we refer to as the dimension reduction inequality.

Theorem 2.5.5. *For any $d \geq 2$, $\theta \geq 2$, and $1 \leq d' \leq d-1$*

$$\sigma_\theta(d, p) \geq \sigma_\theta(d - d', \sigma_\theta(d', p)). \tag{2.5.5}$$

Proof. We can subdivide the d-dimensional Hamming torus into $n^{d-d'}$ disjoint sub-Hamming tori of dimension d'. The probability of internally spanning a fixed sub-Hamming torus is $\sigma_\theta(d', p)$, and the initially open sets in the sub-Hamming tori are mutually independent. Therefore, we may identify each d'-dimensional sub-Hamming torus with a single vertex, which is open independently with probability $\sigma_\theta(d', p)$, and the result is a subset of a $(d - d')$-dimensional Hamming torus that spans with probability $\sigma_\theta(d - d', \sigma_\theta(d', p))$. If this procedure spans the $(d - d')$-dimensional Hamming torus, then the original configuration in the d-dimensional graph will span as well. $\square$

Since we can compute bounds for $\sigma_\theta(2, p)$ and $\sigma_\theta(1, p)$ for all θ and p, the dimension reduction inequality yields lower bounds on the critical exponents for all d and θ. In some cases, the lower bounds obtained this way match our upper bounds, so we can precisely compute the critical exponent. For instance, when $d = 3$ and $\theta = 4$ we see that the critical exponent is $\alpha_c = 1 + d/\theta = 7/4$. In this case, if $\alpha = (7 - \epsilon)/4$ with $0 < \epsilon < 1$ then Lemma 2.5.1 with $k = 2$ implies that $\sigma_4(2, n^{-\alpha}) \geq cn^{6-4\alpha} = cn^{-1+\epsilon}$. Then, since $\sigma_\theta(d, p)$ is increasing in p,

$$\sigma_4(3, n^{-\alpha}) \geq \sigma_4(1, \sigma_4(2, n^{-\alpha})) \geq \sigma_4(1, cn^{-1+\epsilon}) = P(\mathrm{Bin}(n, cn^{-1+\epsilon}) \geq 4) \to 1.$$

Theorem 2.7.6 implies that $1 + d/\theta$ is always an upper bound for the critical exponent, so in the case $d = 3$, $\theta = 4$ the critical exponent is $7/4$.

As a second example of how to apply Lemma 2.5.1 and Theorem 2.5.5, consider the case $d = 6$, $\theta = 5$. Applying dimension reduction and Lemma 2.5.1 twice yields

$$\sigma_5(6, n^{-\alpha}) \geq \sigma_5(4, \sigma_5(2, n^{-\alpha})) \geq \sigma_5(4, Cn^{-\beta(\alpha)}) \geq \sigma_5(2, cn^{-\beta(\beta(\alpha))}).$$

The last term above tends to 1 as $n \to \infty$ if $\beta(\beta(\alpha)) < 4/3$ by Theorem 2.2.1, so finding the supremum over α satisfying this inequality gives a lower bound on the critical exponent in this case. With a little help from Matlab we can numerically compute this supremum, and generate lower bounds for other d and θ. See Figure 2.2 for plots of upper and lower bounds on α_c for $d \in \{2, 3, 4, 5, 6\}$ and $\theta \in \{2, \ldots, 20\}$. Table 2.1 lists all cases for which our upper

θ	2	3	4	5	6	7	8	9	10	11	12
Lower Bound	5/2	2	7/4	11/7	3/2	7/5	19/14	17/13	23/18	5/4	27/22
Upper Bound	5/2	2	7/4	11/7	3/2	7/5	15/11	17/13	9/7	5/4	21/17

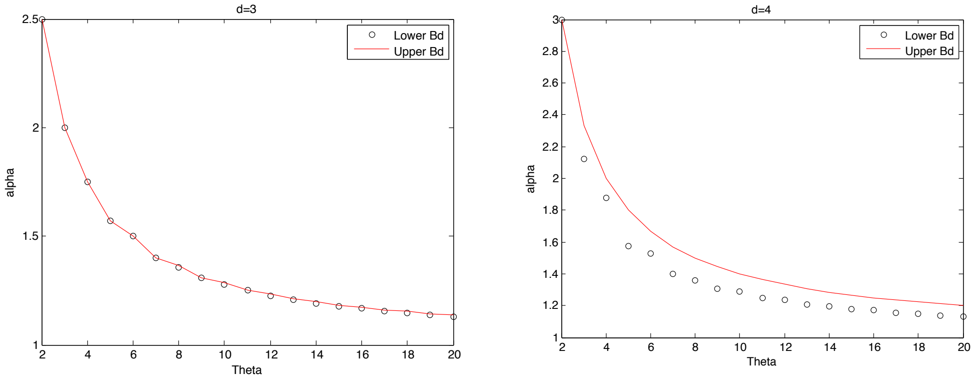
Table 2.1: Upper and lower bounds for the critical exponent when $d = 3$. If $p \asymp n^{-\alpha}$ and α is larger than the upper bound then spanning will not occur with high probability, while if α is smaller than the lower bound then spanning will occur with high probability.

and lower bounds match when $d = 3$, and a few cases for which they conspicuously do not ($\theta = 8, 10, 12$). The upper bounds in the table are the smaller of $1 + 3/\theta$ and the bounds from Theorem 2.4.1 – either $1 + 8/(3\theta - 1)$ or $1 + 8/(3\theta - 2)$, depending on whether θ is odd or even.

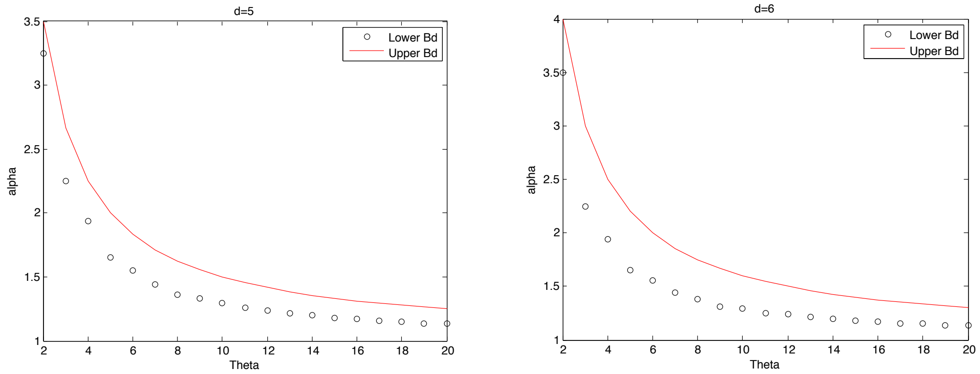
2.6 A precise three-dimensional result

In this section, we precisely compute the limiting spanning probability in the case $d = 3$ and $\theta = 3$. As computed in Section 2.5, the critical exponent in this case is $\alpha = 2$, so we consider the scaling $p = an^{-2}$ when $a > 0$ is a constant.

The resulting limit in Theorem 2.2.2 is a simplified expression for a probability involving Poisson random variables with means depending on a. Indeed, to compute the spanning probability, we identify the minimal ingredients that lead to spanning, and show that their frequencies of occurrence in ω_0 converge jointly to independent Poisson random variables by using the Chen-Stein method [9]. First, we identify two fundamental configurations, which we will define carefully later: points that see at least one open vertex in each direction (Figure 2.3(b)) and lines that contain at least two open vertices and at least one more open vertex in the same plane (Figure 2.3(a)). At least one of these configurations is necessary (in the limit) for spanning because lines that contain 3 or more open vertices do not appear when $p = an^{-2}$, as the expected number of such lines is $O(n^2(np)^3) = O(n^{-1})$. Note



(a) $d = 3$ (b) $d = 4$



(c) $d = 5$ (d) $d = 6$

Figure 2.2: Upper and lower bounds for the critical exponent when $p \asymp n^{-\alpha}$.

that in the definitions of our configurations we allow for there to be three or more open vertices in a line, even though this is unlikely to occur for large n. This is to maintain some monotonicity of the events, and simplifies the Poisson convergence proofs. Each fundamental configuration also has a corresponding "enhanced" configuration (Figures 2.4 and 2.6), which requires additional open vertices in certain planes. Each of these configurations has nonzero probability in the limit, and affects the limiting spanning probability.

We must now determine which combinations of these ingredients are asymptotically necessary and sufficient for spanning. This is summarized as follows:

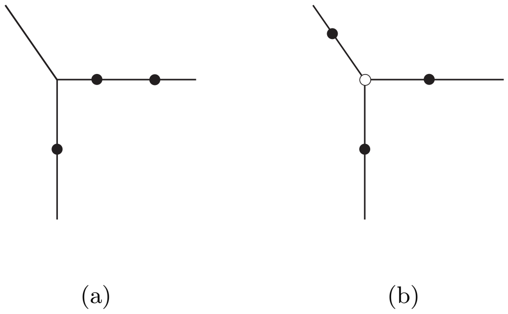

(a) (b)

Figure 2.3: Without one of these configurations appearing somewhere in the graph at time 0, nothing will become open at time 1 when $d = \theta = 3$. The open circle in (b) is to emphasize that this 'Basic' configuration is with respect to a focal vertex which will become open at time 1. The 'Line' configuration in (a) is indexed with respect to the line which contains two open points, and the single open vertex off of the horizontal line signifies that at least one vertex on one of the two planes containing the focal line must be open.

1. At least one 'basic' configuration like that in Figure 2.3(b), AND at least one 'line' configuration like that in Figure 2.3(a); OR
2. At least one 'enhanced basic' configuration like that in Figure 2.4; OR
3. At least one 'line' configuration, AND at least one askew (non-parallel, non-intersecting) line that contains at least two open vertices (see the configuration in Figure 2.5); OR
4. At least two 'line' configurations like the one in Figure 2.3(a); OR
5. At least one 'enhanced line' configuration like those in Figure 2.6.

We call ω_0 *good* if it contains at least one of the recipes (1)–(4) described above; a formal definition is given below. The event $\{\omega_0 \text{ is good}\}$ is asymptotically equivalent to the event $\{\omega_0 \text{ spans}\}$ in the sense of the following lemma.

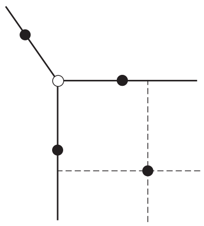


Figure 2.4: ‘Enhanced Basic’: First the two lines containing the open circle in the front plane will span, followed by the two dotted lines then the front plane. Once a plane is spanned, the rest of the graph is likely to be spanned (see the last paragraph in the proof of Lemma 2.6.1).

Lemma 2.6.1. *If $d = \theta = 3$ and $p = an^{-2}$, then as $n \to \infty$*

$$\mathbb{P}(\omega_0 \textit{ is good}) - \mathbb{P}(\omega_\infty \equiv 1) \to 0.$$

To formally define the event $\{\omega_0 \text{ is good}\}$, and for the proofs that follow, we need to introduce some notation.

Notation

Let e_1, e_2, e_3 denote the standard basis vectors in $\mathbb{R}^3$. For $v, w \in V$ let $d(v, w)$ be the number of nonzero coordinates of $v - w$. Let $\mathcal{N}(v) = \{w \in V : d(v, w) = 1\}$ denote the neighborhood of v, and for $A \subset V$ let $\mathcal{N}(A) = \cup_{v \in A} \mathcal{N}(v) \setminus A$.

The basic and enhanced basic configurations will be indexed by vertices, while the line and enhanced line configurations will be indexed by lines. So, we let

$$\mathscr{L} = \{\ell \subset V : |\ell| = n \text{ and } \forall v, w \in \ell,\ d(v, w) \leq 1\}$$

be the set of lines in V. Also, for $i = 1, 2, 3$, let

$$\mathscr{L}_i = \{\ell \in \mathscr{L} : \forall u, v \in \ell, \exists m = m(u, v) \in \mathbb{Z} \ s.t.\ u = v + me_i\}$$

denote the collection of lines in V parallel to the coordinate axis in the e_i direction. For the duration of this paper we will use ℓ to refer to a generic line.

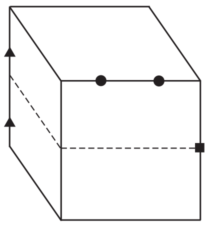


Figure 2.5: This configuration leads to the front plane being spanned, and the graph is likely to be spanned. There is a 'line' configuration with respect to the line that contains the two closed circles - the rectangle in the front plane completes the configuration and leads to the spanning of the top line in two steps. After the line with two circles spans, the line with two triangles is now in a 'line' configuration, and spans in two more steps. The vertex at the intersection of the dotted line and the line with the triangles is now open, and leads to the vertex at the intersection of the dotted lines becoming open, which leads to the spanning of the front plane in three more steps. Note that it is crucial for the lines with the circles and triangles to be askew – if these lines were parallel then the front plane would not span without additional help.

To apply the Chen-Stein method, we let `Basic`, `Line`, `Line`$\emptyset$, `EnhancedBasic`, `EnhancedLine`, and `NonEnhancedLine` be the random variables that count the number of occurrences of the corresponding configurations in ω_0, which we now define carefully. The relevant events are a bit difficult to describe, so we refer the reader to Figures 2.3-2.6 for guidance.

Define the *basic* event, for $v \in V$, to be

$$G_v^B = \{\exists\ w_1, w_2, w_3 \in \omega_0 \setminus \{v\} \text{ and } \exists\ m_1, m_2, m_3 \in \mathbb{Z}\ s.t.\ v = w_i + m_i e_i \text{ for } i = 1, 2, 3\}.$$

As Figure 2.3(b) indicates, the basic event occurs at v if v has at least one initially open

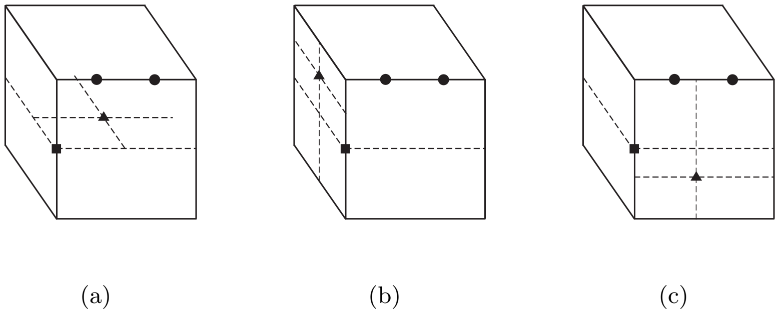


Figure 2.6: 'Enhanced Line': These configurations (and any rotations or shifts of them) are likely to span. The triangle vertex will cause a second line in the front plane to be spanned, thus the full front plane will be spanned if there is an additional open vertex anywhere in the graph that is not coplanar with this line or the line with two circles. Once a plane is spanned, the rest of the graph is likely to be spanned.

neighbor in each basis direction. Define the *enhanced basic* event, for $v \in V$, to be

$$\begin{aligned} G_v^{EB} = \{ &\exists\ w \in \omega_0\ s.t.\ d(v,w) = 2, \text{ and } \exists\ w_1, w_2, w_3 \in \omega_0 \setminus (\mathcal{N}(w) \cup \{v\}) \\ &\text{and } \exists\ m_1, m_2, m_3 \in \mathbb{Z}\ s.t.\ v = w_i + m_i e_i \text{ for } i = 1,2,3\}, \end{aligned}$$

As Figure 2.4 indicates, the enhanced basic event occurs at v if the basic event occurs at v and there is at least one open vertex in one of the planes containing v that is not a neighbor of v. Further, this additional open vertex should not be collinear with the sole open neighbor of v in any direction; if there were two open neighbors of v in a single direction, then we could allow the additional open vertex to be collinear with one of them, but this event is rare. Let I_v^B be the indicator random variable for the event G_v^B, so $\texttt{Basic} = \sum_v I_v^B$, and let I_v^{EB} be the indicator random variable for the event G_v^{EB}, so $\texttt{EnhancedBasic} = \sum_v I_v^{EB}$. In general, we will denote by $I_\dagger^*$ the indicator of the event $G_\dagger^*$.

For each line $\ell \in \mathscr{L}$ we define the *line* event

$$G_\ell^L = \{|\ell \cap \omega_0| = 2, |\mathcal{N}(\ell) \cap \omega_0 \setminus \mathcal{N}(\ell \cap \omega_0)| \geq 1\} \cup \{|\ell \cap \omega_0| \geq 3, |\mathcal{N}(\ell) \cap \omega_0| \geq 1\}.$$

As Figure 2.3(a) suggests, the line event occurs at ℓ if ℓ contains at least two initially open vertices, and there is at least one additional open vertex in the same plane as ℓ. This additional open vertex should not be in the neighborhood of the two open vertices in ℓ, though if there are three or more open vertices in ℓ then the location of the additional vertex does not matter. We now define $\texttt{Line} = \sum_{\ell \in \mathscr{L}} I_\ell^L$, and because we will also need to count the number of line events in a particular direction (for case (3) in the recipe for spanning), for $i = 1, 2, 3$ we let $\texttt{Line}_i = \sum_{\ell \in \mathscr{L}_i} I_\ell^L$. For each $\ell \in \mathscr{L}$ we define the *$\emptyset$-line* event

$$G_\ell^{\emptyset L} = \{|\ell \cap \omega_0| \geq 2\} \setminus G_\ell^L,$$

and let $I_\ell^{\emptyset L}$ be the corresponding indicator random variable so $\texttt{Line}\emptyset = \sum_{\ell \in \mathscr{L}} I_\ell^{\emptyset L}$ and for $i = 1, 2, 3$, $\texttt{Line}\emptyset_i = \sum_{\ell \in \mathscr{L}_i} I_\ell^{\emptyset L}$. The $\emptyset$-line event occurs at ℓ if ℓ contains at least two initially open vertices, and there are no other open vertices in the same plane as ℓ (except possibly those that are collinear with one of the two open vertices in ℓ).

For each line $\ell \in \mathscr{L}$ we define the *enhanced line* event

$$\begin{aligned} G_\ell^{EL} = \{&|\ell \cap \omega_0| = 2 \text{ and } \exists v \in \mathcal{N}(\ell) \cap \omega_0 \setminus \mathcal{N}(\ell \cap \omega_0) \textit{ s.t. } |\mathcal{N}(\mathcal{N}(v)) \cap \omega_0 \setminus \mathcal{N}(\ell \cap \mathcal{N}(v))| \geq 1\} \\ &\cup \{|\ell \cap \omega_0| \geq 3, \exists v \in \mathcal{N}(\ell) \cap \omega_0 \textit{ s.t. } |\mathcal{N}(\mathcal{N}(v)) \cap \omega_0 \setminus \mathcal{N}(\ell \cap \mathcal{N}(v))| \geq 1\} \end{aligned}$$

and let I_ℓ^{EL} be the corresponding indicator random variable so $\texttt{EnhancedLine} = \sum_{\ell \in \mathscr{L}} I_\ell^{EL}$ and for $i = 1, 2, 3$, $\texttt{EnhancedLine}_i = \sum_{\ell \in \mathscr{L}_i} I_\ell^{EL}$. For the enhanced line event to occur at ℓ, a line configuration must appear in ω_0 at ℓ and there must be at least one additional open vertex. This additional open vertex is coplanar with the open vertex in $\mathcal{N}(\ell)$ from the line configuration (there may be more than one), but is not counted if it is collinear with this vertex or on the other plane containing ℓ. Finally, define the non-enhanced line event

$$G_\ell^{NEL} = G_\ell^L \setminus G_\ell^{EL}$$

and its corresponding indicator I_ℓ^{NEL}, so that $I_\ell^{NEL} = I_\ell^L - I_\ell^{EL}$ for every $\ell \in \mathscr{L}$, $\texttt{NonEnhancedLine} = \texttt{Line} - \texttt{EnhancedLine}$ and for $i = 1, 2, 3$, $\texttt{NonEnhancedLine}_i = \texttt{Line}_i - \texttt{EnhancedLine}_i$.

Now we define the event that ω_0 is good by

$$\{\omega_0 \text{ is good}\} = \{\texttt{Basic} \geq 1, \texttt{Line} \geq 1\} \cup \{\texttt{EnhancedBasic} \geq 1\} \cup \bigcup_{i=1}^{3} \left\{ \texttt{Line}_i \geq 1, \sum_{j \neq i} \texttt{Line}\emptyset_j \geq 1 \right\}$$
$$\cup \{\texttt{Line} \geq 2\} \cup \{\texttt{EnhancedLine} \geq 1\}.$$

The third term above covers the scenario in Figure 2.5 when $\texttt{Line} \leq 1$, which is otherwise covered by the event $\{\texttt{Line} \geq 2\}$. Using inclusion-exclusion, exploiting obvious symmetries of the graph, and combining like terms:

$$\begin{aligned}
\mathbb{P}(\omega_0 \text{ is good}) = {} & \mathbb{P}(\texttt{Basic} \geq 1, \texttt{Line} = 1) + \mathbb{P}(\texttt{EnhancedBasic} \geq 1, \texttt{Line} = 0) \\
& + \mathbb{P}(\texttt{Line} \geq 2) + \mathbb{P}(\texttt{Basic} = 0, \texttt{EnhancedLine} = 1, \texttt{NonEnhancedLine} = 0) \\
& + 3\,\mathbb{P}(\texttt{Basic} = 0, \texttt{NonEnhancedLine}_1 = 1, \texttt{NonEnhancedLine}_2 + \texttt{NonEnhancedLine}_3 = 0, \\
& \qquad \texttt{EnhancedLine} = 0, \texttt{Line}\emptyset_2 + \texttt{Line}\emptyset_3 \geq 1)
\end{aligned} \tag{2.6.1}$$

Therefore, once we compute the probabilities in (2.6.1), Lemma 2.6.1 implies Theorem 2.2.2. Lemma 2.6.2 allows us to do just this, and is followed by the proof of Lemma 2.6.1. The proof of Lemma 2.6.2 uses the Chen-Stein method, and is outlined in Appendix 2.10.

Lemma 2.6.2. *If* $p = an^{-2}$*, then as* $n \to \infty$ *the table below gives the means of the random variables appearing in (2.6.1).*

Random Variable	*Mean*
$\texttt{Basic}$	a^3
$\texttt{EnhancedBasic}$	$a^3(1 - e^{-3a})$
$\texttt{Line}$	$\frac{3}{2}a^2(1 - e^{-2a})$
$\texttt{Line}\emptyset_i$	$\frac{1}{2}a^2 e^{-2a}$
$\texttt{NonEnhancedLine}_i$	$\frac{1}{2}a^2 \left[\left(e^{-a} + ae^{-3a}\right)^2 - e^{-2a} \right]$
$\texttt{EnhancedLine}$	$\frac{3}{2}a^2 \left[1 - \left(e^{-a} + ae^{-3a}\right)^2 \right]$

Furthermore, the two random variables $\texttt{EnhancedBasic}$ *and* $\texttt{Line}$ *converge jointly in distribution to independent Poisson random variables with the above means, as do the eight* *random variables* `Basic`*,* `EnhancedLine`*, and for* $i = 1, 2, 3$*,* $\texttt{NonEnhancedLine}_i$ *and* $\texttt{Line}\emptyset_i$*.*

Remark 2.6.3. Lemma 2.6.2 allows us to compute the limiting probability in (2.6.1) by treating all of the random variables that appear as independent Poisson random variables with the means given by the table. The means that appear in the limit are straightforward to compute. For example, to compute the expected number of basic events, the probability that a fixed vertex has at least one initially open neighbor in each direction is $\sim (np)^3 = a^3/n^3$, and there are n^3 vertices at which a basic configuration can be centered. To obtain the expected number of enhanced basic configurations, observe that a fixed vertex must first see a basic configuration, then independently at least one of the $3(n-2)^2$ coplanar but not collinear vertices must be present. This has probability $1-(1-p)^{3(n-2)^2} \sim 1-e^{-3a}$ of occurring.

Proof of Lemma 2.6.1. We will first show that spanning does not occur with high probability when ω_0 is not good. The expected number of lines that contain at least three initially open vertices is $\sim 3n^2\binom{n}{3}p^3 = O(n^{-1})$, so at least one line configuration or basic configuration is necessary for any vertices to become open after one step.

Any vertex that becomes open in the second step must be neighbors with at least one vertex that becomes open in the first step, i.e. with a vertex in $\omega_1 \setminus \omega_0$. If $\texttt{Line} = 0$ and $\texttt{EnhancedBasic} = 0$ then any two basic events located at vertices v and w cannot be coplanar unless $\mathcal{N}(v) \cap \mathcal{N}(w) \subset \omega_0$, otherwise a line or an enhanced basic configuration would exist. The probability that there exist two vertices, v and w, with $I_v^B I_w^B = 1$, $d(v,w) = 2$ and $\mathcal{N}(v) \cap \mathcal{N}(w) \subset \omega_0$ is at most $3n\binom{n^2}{2}(np)^2p^2 = O(n^{-1})$, so with high probability there are no coplanar basic events. Therefore, no pair of vertices in $\omega_1 \setminus \omega_0$ have a common neighbor, and no vertex in $\mathcal{N}(\omega_1 \setminus \omega_0) \setminus \omega_0$ has more than one neighbor in ω_0 (or else a line or enhanced basic configuration would have existed in ω_0). This implies that no vertices can become open in the second step, so spanning cannot occur with high probability when $\texttt{Line} = 0$ and $\texttt{EnhancedBasic} = 0$.

Similarly, if $\texttt{NonEnhancedLine}_1 = 1$, $\texttt{NonEnhancedLine}_2 + \texttt{NonEnhancedLine}_3 = 0$, $\texttt{Basic} =$

0,

`EnhancedLine` $= 0$ and `Line`$\emptyset_2 +$`Line`$\emptyset_3 = 0$ then spanning is unlikely to occur. The sole line configuration will span the focal line, ℓ, after two steps. There may be parallel lines that contain two occupied vertices, but they cannot be coplanar with ℓ or else the line configuration would be enhanced. These parallel lines will not span the cube as their neighborhoods do not intersect ℓ, so no other vertices will become open after two steps. Therefore, $\mathbb{P}(\{\omega_\infty \equiv 1\} \setminus \{\omega_0 \text{ is good}\}) \to 0$.

The probability of ω_0 containing a basic configuration and a line configuration that share a plane (that is, there exist v and ℓ so that $I_v^B I_\ell^L = 1$ and $v \in \mathcal{N}(\ell) \cup \ell$) is at most $Cn^3(n)(np)^3(np)^2 = O(n^{-1})$. Similarly, the probability of having two or more coplanar line configurations is $O(n^{-1})$. Conditional on the complements of these last two events, observe that a line configuration will cause a basic configuration to become an enhanced basic configuration in two steps. Likewise, a line configuration will cause a second line configuration to become an enhanced line configuration in two steps; and similarly a line configuration will with high probability cause an askew line with two initially open vertices to become a line configuration (and subsequently an enhanced line configuration).

Both the enhanced basic and enhanced line configurations lead to a plane becoming open. Once a plane is open, two non-neighboring, coplanar open vertices will cause another plane to become open, then one more open vertex elsewhere will cause the rest of the graph to become open. With probability exponentially close to 1, there are at least $n^{1/2}$ planes with at least two non-neighboring open vertices in ω_0. Therefore, $\mathbb{P}(\{\omega_0 \text{ is good}\} \setminus \{\omega_\infty \equiv 1\}) = O(n^{-1})$, and the two events are asymptotically equivalent. $\square$

2.7 Open one-dimensional subgraphs

In this section we obtain an upper bound on the threshold probability for lines, $p_c(1, d)$. The main idea is the following. Assume that the line ℓ contains $r \leq \theta$ initially open vertices, that it intersects one line with $\theta - r$ initially open sites (not on ℓ), and that it intersects θ other lines, each with $\theta - r - 1$ sites (not on ℓ) initially open. Then after one step, ℓ has $r + 1$

points open, and after two steps, θ points open. After three steps ℓ is completely open. See Fig. 2.7 for an illustration.

For a set $S \subset V$ and $x \in \mathbb{N}$, let $\operatorname{Initial}(S, \geq x)$ be the event that the set S has at least x points initially open, i.e.,

$$\operatorname{Initial}(S, \geq x) = \left\{ \sum_{v \in S} \omega_0(v) \geq x \right\}.$$

For a point $v \in V$, let $P_{1,2}(v)$ be the e_1, e_2-parallel plane through v:

$$P_{1,2}(v) = \{(a_1, a_2, v_3, v_4, \ldots, v_d) :\ a_1, a_2 \in [n]\}.$$

Let $\ell_2(v)$ be the e_2-parallel line through v:

$$\ell_2(v) = \{(v_1, a_2, v_3, v_4, \ldots, v_d) :\ a_2 \in [n]\}.$$

For any e_1-parallel line ℓ, define

$$\ell_l = \{w \in \ell, w_1 < n/3\}, \quad \ell_m = \{w \in \ell, n/3 \leq w_1 \leq 2n/3\}, \quad \text{and} \quad \ell_r = \{w \in \ell, w_1 > 2n/3\}$$

to be the left, middle and right thirds of ℓ. Define

$$\text{Cross Lines}_m(\ell) = \left\{ \sum_{v \in \ell_m} \mathbf{1}_{\operatorname{Initial}(\ell_2(v), \geq \theta - r)} \geq 1 \right\}$$

$$\text{Cross Lines}_r(\ell) = \left\{ \sum_{v \in \ell_r} \mathbf{1}_{\operatorname{Initial}(\ell_2(v), \geq \theta - r - 1)} \geq \theta \right\}$$

and

$$F_\ell = \operatorname{Initial}(\ell_l, \geq r) \bigcap \text{Cross Lines}_m(\ell) \bigcap \text{Cross Lines}_r(\ell).$$

Notice that the event F_ℓ depends only on the sites in $P_{1,2}(v)$ for any $v \in \ell$. Also note that

$$\text{Cross Lines}_m(\ell) = \text{Cross Lines}_m(\ell') \quad \text{and} \quad \text{Cross Lines}_r(\ell) = \text{Cross Lines}_r(\ell')$$

for any e_1-parallel lines $\ell \neq \ell'$ that lie in a common e_1, e_2-parallel plane. Finally, note that $\operatorname{Initial}(\ell_l, \geq r)$, $\text{Cross Lines}_m(\ell)$, and $\text{Cross Lines}_r(\ell)$ are independent, and $\operatorname{Initial}(\ell_l, \geq r)$ and $\operatorname{Initial}(\ell'_l, \geq r)$ are independent.

We exhibit the role of F_ℓ (see Fig. 2.7) in the following lemma.

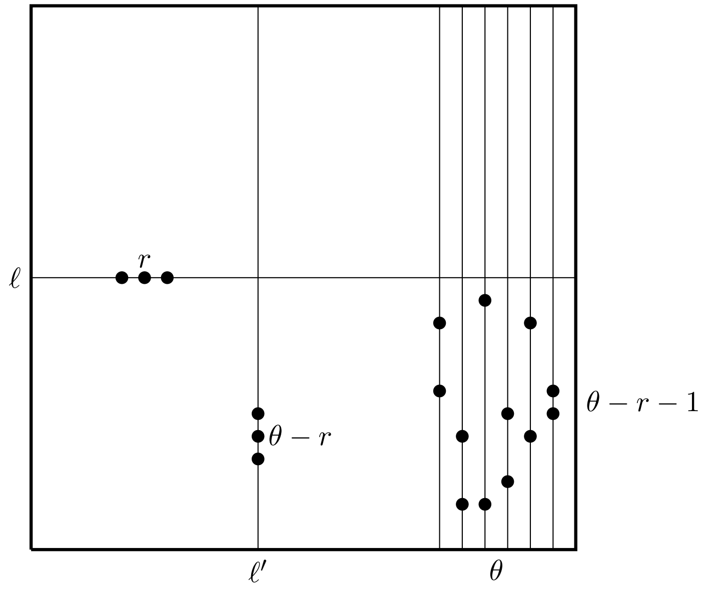


Figure 2.7: An instance of the event F_ℓ. Here $\theta = 6$, $r = 3$. After one step the intersection of lines ℓ and ℓ' becomes open so ℓ has $r + 1$ vertices open. At step 2 the θ intersections with ℓ and the other θ vertical lines become open. At step 3 all of ℓ becomes open.

Lemma 2.7.1. *If ℓ is a line parallel to the e_1 axis and F_ℓ occurs, then the entire line ℓ is open after three steps.*

Remark 2.7.2. Computation of $P(F_\ell)$ is facilitated by independence of the three events. A more natural definition would not restrict the orientations of the lines, or demand that the event happen in the left, middle or right sections thereof, and would increase the probability by a constant factor, independent of n.

We set $r = \left\lceil \frac{(d-1)\theta}{d} \right\rceil - 1$ and $p = n^{-1-\frac{d}{\theta}} f(n)$, where $f(n)$ is any function such that $f(n) \to \infty$. We will show that in this regime some line becomes open asymptotically almost surely. We will use the following elementary fact about the binomial distribution.

Lemma 2.7.3. *Assume that S is Binomial(n, p), with large n and $p = p(n)$, and that k does not depend on n. If $np = O(1)$, then $P(S \geq k) \geq c(np)^k$ for some constant c dependent on k. If $np \to \infty$, then $P(S \geq k) \to 1$.*

Lemma 2.7.4. *Fix $v \in V$ and $\theta, d \geq 3$. Let $p = n^{-1-\frac{d}{\theta}} f(n)$ where $f(n) \to \infty$. Then for any $c > 0$, the probability that there exists an e_1-parallel line ℓ in $P_{1,2}(v)$ such that F_ℓ occurs is at least cn^{2-d} for n sufficiently large.*

Proof. As the event in the statement is increasing, its probability is monotone in p. Thus we may assume that $f(n)$ grows to ∞ as slowly as we need in the proof.

Note that when $\theta, d \geq 3$ then $rd/\theta \geq 1$ as

$$rd/\theta \geq \left(\frac{(d-1)\theta}{d} - 1 \right) \frac{d}{\theta} = d - 1 - d/\theta.$$

The right hand side is strictly greater than 1 except if $d = \theta = 3$. We assume that at least one of d and θ is at least 4, and leave the exceptional case to the reader.

The three events that define F_ℓ depend on disjoint sets of sites, so they are independent and we compute their probabilities separately. Furthermore, for the set of lines ℓ we consider, the events Cross Lines$_m(\ell)$ and Cross Lines$_r(\ell)$ do not depend on ℓ, which will thus be dropped from the notation. For any ℓ, by Lemma 2.7.3

$$\begin{aligned}\mathbb{P}(\text{Initial}\,(\ell_l, \geq r)) &\geq c_1(np)^r \\ &\geq c_1(f(n)n^{-d/\theta})^r.\end{aligned}$$

As this is $o(1/n)$ we can use Lemma 2.7.3 again to get that

$$\mathbb{P}\big(\exists \ell \text{ such that Initial}\,(\ell_l, \geq r) \text{ occurs }\big) \geq c_2 n(f(n)n^{-d/\theta})^r.$$

To estimate the second probability, observe that

$$\mathbb{P}(\text{Initial}\,(\ell_2(v), \geq \theta - r)) \geq c_3(np)^{\theta - r},$$

which is $o(1/n)$, as $r < (d-1)\theta/d$. Thus

$$\begin{aligned}\mathbb{P}(\text{Cross Lines}_m) &\geq c_4 n(np)^{\theta - r} \\ &\geq c_4 n(f(n)n^{-d/\theta})^{\theta - r}.\end{aligned}$$

For the third probability,

$$\mathbb{P}(\text{Initial}\,(\ell_2(v), \geq \theta - r)) \geq c_5(np)^{\theta - r - 1},$$

and

$$n \cdot (np)^{\theta - r - 1} \geq f(n)^{\theta - r - 1} n^{1 - d + (r+1)\frac{d}{\theta}} \to \infty$$

as $n \to \infty$, so Lemma 2.7.3 implies that

$$\mathbb{P}(\text{Cross Lines}_r) \to 1,$$

and for large n the probability is bounded below by a constant $c_6 > 0$. Multiplying together the probabilities, we have that for any c and all sufficiently large n

$$\begin{aligned}&\mathbb{P}(\exists\ \ell \text{ in } P_{1,2}(v) \text{ such that that } F_\ell \text{ occurs}) \\ &\quad= \mathbb{P}(\exists \ell \text{ such that Initial}\,(\ell_l, \geq r))\mathbb{P}(\text{Cross Lines}_m)\mathbb{P}(\text{Cross Lines}_r) \\ &\quad\geq c_2 n(f(n)n^{-d/\theta})^r c_4 n(f(n)n^{-d/\theta})^{\theta - r} c_6 \\ &\quad= c_7 f(n)^\theta n^{2-d} \\ &\quad> cn^{2-d},\end{aligned}$$

ending the proof. ☐

Theorem 2.7.5. *Suppose* $p = n^{-1-\frac{d}{\theta}} f(n)$ *with* $f(n) \to \infty$. *Then* $\mathbb{P}(\cup_\ell F_\ell) \to 1$ *as* $n \to \infty$, *where the union is taken over all* e_1*-parallel lines. Thus, with probability going to 1, some line becomes open after three steps.*

Proof. We can choose n^{d-2} distinct vertices v_i such that $P_{1,2}(v_i)$ are disjoint. Then the events that there exist ℓ in $P_{1,2}(v_i)$ where F_ℓ occurs are independent. Moreover,

$$n^{d-2}\mathbb{P}(\exists \ell \text{ in } P_{1,2}(v_i) \text{ such that } F_\ell \text{ occurs}) \geq n^{d-2} c n^{2-d} = c$$

for any fixed c. Thus Thus $\mathbb{P}(\cup_\ell F_\ell) \to 1$ by Lemma 2.7.3. ☐

Theorem 2.7.6. *Let* $p = n^{-1-d/\theta} f(n)$. *If* $f(n) \to 0$, *then* $\mathbb{P}(\textit{Above Threshold}) \to 0$.

Proof. Using the union bound,

$$\begin{aligned}
\mathbb{P}(\text{Above Threshold}) &\leq \sum_{v \in V} \mathbb{P}\left(\sum_{w \sim v} \omega_0(w) \geq \theta\right) \\
&= n^d \mathbb{P}\left(\sum_{w \sim v} \omega_0(w) \geq \theta\right) \\
&\leq n^d \binom{n}{\theta} p^\theta \\
&\leq f(n)^\theta
\end{aligned}$$

which approaches 0 as $n \to \infty$. ☐

Proof of Theorem 2.2.3. Combining Theorems 2.7.5 and 2.7.6 proves the result. ☐

2.8 Open two-dimensional subgraphs

In previous sections we have encountered several possibilities for a vertex v to become open:

- v is initially open;

- the neighborhood of v has at least θ vertices initially open, causing v to become open by time 1; and

- a line containing v has at least $\theta(d-1)/d$ vertices initially open, with some additional open sites "nearby" (see Section 2.7).

Let Plane Active be the event that some plane eventually becomes open. In this section we show that if p is sufficiently small then with high probability all of the vertices that are eventually open satisfy a condition like one of the three above. By doing this we prove an upper bound on the probability of Plane Active and consequently a lower bound on the threshold probability $p_c(2,d)$.

Let A be some integer, $1 \leq A \leq \theta$, which we will specify later. Let E be the event that there exists a vertex v such that:

1. v is initially not open;
2. the neighborhood of v has at most A vertices initially open;
3. each line containing v has at most $A/2$ vertices initially open; and
4. v becomes open.

Our strategy to demonstrate that $\mathbb{P}(\text{Plane Active})$ is small for sufficiently small p is to show that $\mathbb{P}(E)$ and $\mathbb{P}(\text{Plane Active} \setminus E)$ are both small.

For each vertex v, let E_v be the event that v satisfies (1)–(4), and none among such vertices becomes open earlier. If the event E occurs then there must be a first time a vertex satisfying (1)–(4) exists, thus $E \subseteq \cup_v E_v$, and consequently, $\mathbb{P}(E) \leq n^d \mathbb{P}(E_v)$.

Lemma 2.8.1. *Suppose $p = o(n^{-1-\beta})$ with $\beta > (\frac{2d^2}{\theta - A} + 1)\frac{2}{A}$. Fix a line ℓ. The probability that ℓ contains at least $\frac{\theta - A}{2d}$ vertices v that have at least $A/2$ initially open points in $\mathcal{N}(v) \setminus \ell$ is*

$$o\left(n^{\frac{\theta-A}{2d}(1-\beta\frac{A}{2})}\right).$$

Proof. The reduced neighborhoods $\mathcal{N}(v) \setminus \ell$, $v \in \ell$, are pairwise disjoint, and in each the number of initially open vertices is a Binomial$((d-1)(n-1), p)$ random variable. The probability that such a random variable is at least $A/2$ is bounded by a constant times $(np)^{A/2} = o\left(n^{-\beta A/2}\right)$. These random variables are independent, thus the probability that at least $\frac{\theta-A}{2d}$ of them are at least $A/2$ is $o\left((n \cdot n^{-\beta A/2})^{\frac{\theta-A}{2d}}\right)$. $\square$

Lemma 2.8.2. *Assume p satisfies the same bound as in Lemma 2.8.1. Fix a line ℓ. The probability that ℓ has at least $\frac{\theta-A}{2d}$ vertices w, for which there exists a line $\ell' \neq \ell$ through w such that $\ell' \setminus \{w\}$ contains at least $A/2$ initially open points is*

$$o\left(n^{\frac{\theta-A}{2d}(1-\beta\frac{A}{2})}\right).$$

Proof. We need to bound the probability of at least $\frac{\theta-A}{2d}$ successes in $n(d-1)$ independent trials, each of which is a success with the probability that a given line has at least $A/2$ points initially open. Same estimates as in the proof of Lemma 2.8.1 apply. $\square$

Lemma 2.8.3. *Assume p satisfies the same bound as in Lemma 2.8.1. Then $\mathbb{P}(E) \to 0$ as $n \to \infty$.*

Proof. As we have already observed, $\mathbb{P}(E) \leq n^d \mathbb{P}(E_v)$. Now, if E_v occurs, by (2) at least $\theta - A$ vertices in the neighborhood of v must be initially closed but become open strictly before v; therefore, they violate at least one of (1)–(4). But since they are not open initially and become open, they must violate one of (2) or (3). By the pigeonhole principle, of the d lines through v, at least one must either contain $\frac{\theta-A}{2d}$ vertices which violate (2), or $\frac{\theta-A}{2d}$ vertices which violate (3).

By Lemmas 2.8.1 and 2.8.2, each of these happens with probability

$$o\left(n^{\frac{\theta-A}{2d}(1-\beta\frac{A}{2})}\right).$$

Rearranging using the inequality $\beta > (\frac{2d^2}{\theta-A}+1)\frac{2}{A}$, we see that $\mathbb{P}(E_v) = o(n^{-d})$, as claimed. $\square$

Lemma 2.8.4. *Let $p = n^{-1-\beta}$, with $\beta > 0$, and assume $A \geq 4$. Then $\mathbb{P}(\text{Plane Active} \setminus E) \to 0$ as $n \to \infty$.*

Proof. There are $\binom{d}{2}n^{d-2}$ planes, P, and Plane Active $= \cup_P \{P \text{ becomes open}\}$, so we have

$$\mathbb{P}(\text{Plane Active} \setminus E) \leq \binom{d}{2} n^{d-2} \mathbb{P}(\{P \text{ becomes open}\} \setminus E).$$

Now if P becomes open but E does not occur, then since each point in P becomes open, they must all violate one of (1), (2), or (3). By the pigeonhole principle, at least $n^2/3$ of these points must together violate a single condition. We will check that the probabilities of these three cases are $o(n^{-(d-2)})$. In fact, we will see that they are exponentially small by reducing each case to a large deviation probability involving a Binomial random variable with a small chance of success. We will use the fact that neighborhoods of two points in P do not intersect outside P.

- $\mathbb{P}(n^2/3$ vertices in P are initially open) is exponentially small in n^2, as $p = o(1)$.

- $\mathbb{P}(n^2/3$ vertices in P are each on a line with $A/2$ points initially open) is exponentially small in n.

  As every line covers at most n points in P, this event implies that there are at least $n/(3d)$ parallel lines, in some direction e_i, each with at least $A/2$ points initially open. The probability that a given line has at least $A/2$ points initially open is $O((np)^{\lfloor A/2 \rfloor}) = o(1)$, thus the probability that $n/(3d)$ lines in a given direction e_i satisfy this is exponentially small in n.

- $\mathbb{P}(n^2/3$ vertices in P each have at least A initially open vertices in their neighborhoods) is exponentially small in n.

  If a vertex w has at least A initially open vertices in its neighborhood then either one of the two lines through w in P contain at least $A/4$ initially open vertices or the $d-2$ lines through w not in P together contain at least $A/2$ initially open vertices. This implies that either (a) there are at least $n/12$ parallel lines in P with at least $A/4$ vertices initially open, or (b) there are at least $n^2/6$ vertices with at least $A/2$ vertices in their neighborhoods outside of P.

> The probability of (a) is exponentially small by the same argument as in the previous case. For a fixed w, the probability that $(d-2)(n-1)$ sites in $\mathcal{N}(w) \setminus P$ contain at least $A/2$ initially open sites is again $O(np) = o(1)$. Thus the probability of (b) is exponentially small in n^2.

Therefore, $\mathbb{P}(\text{Plane Active} \setminus E)$ goes to 0 exponentially fast. □

Proof of Theorem 2.2.4. To get the lower bound set $A = \lfloor \theta - \sqrt{\theta} \rfloor$. Then Lemmas 2.8.1–2.8.3 are (for large enough θ) satisfied with

$$\beta = \frac{2}{\theta} + \frac{4d^2 + 3}{\theta^{3/2}}.$$

The upper bound was proved in Theorem 2.5.3. □

2.9 Further questions and conjectures

We begin with a general form of threshold probabilities; we believe that the answer to the question below is positive.

Question 2.9.1. Do there exist positive constants $c_1 = c_1(i, d)$ and $c_{3/2} = c_{3/2}(i, d)$, so that, for all i and d, a lower bound and an upper bound for $p_c(i, d)$ are both of the form

$$n^{-1-\frac{c_1}{\theta}-\frac{c_{3/2}}{\theta^{3/2}}+o(\theta^{-3/2})},$$

for large enough n?

We next ask whether it is possible that generation of open planes does not likely lead to spanning of the entire graph when $d \geq 4$.

Question 2.9.2. Can one find d and $\theta > 2$ such that $\log_n(p_c(2, d)) - \log_n(p_c(d, d))$ is bounded away from 0 as $n \to \infty$, i.e. $p_c(2, d) \approx n^{-\zeta}$ and $p_c(d, d) \approx n^{-\xi}$ with $\zeta > \xi$? Does this hold for all θ and $d \geq 4$? Note that it does not hold for $d = 3$ by (2.2.2).

It would be desirable to have a general method to determine the critical exponent for any given (small) d and θ; here we merely recall the simplest unsolved instances.

Question 2.9.3. When $d = 3$, we know the critical exponents for $\theta = 2, 3, 4, 5, 6, 7, 9, 11$; what are the correct exponents for $\theta = 8, 10$ and $\theta \geq 12$?

2.10 Poisson convergence for $d = \theta = 3$

In this section we outline the proof of Lemma 2.6.2 regarding Poisson convergence of the random variables that count the configurations that lead to spanning when $d = \theta = 3$ and $p = an^{-2}$. Our approach is to apply the Chen-Stein method [9], and to do so we need to introduce some notation.

We want to show that a collection of random variables, which are sums of indicator random variables, converge to independent Poisson random variables in the limit. That is, suppose we have disjoint sets of indices, $\Gamma_1, \Gamma_2, \ldots, \Gamma_\ell$, let $\Gamma = \cup_{i=1}^{\ell} \Gamma_i$, and for each $\gamma \in \Gamma$ suppose I_γ is an indicator random variable. For $i = 1, \ldots, \ell$ let $W_i = \sum_{\gamma \in \Gamma_i} I_\gamma$ and suppose that $EW_i = \lambda_i$ and $EI_\gamma = p_\gamma$. In our application, the index sets are going to be V for the indicators of the basic and enhanced basic events, and $\mathscr{L}$ for the indicators of the line, $\emptyset$-line, enhanced line, and non-enhanced line events.

To apply the Chen-Stein method in many cases we need to construct a coupling for every fixed $\gamma \in \Gamma$ between I_η and $J_{\eta\gamma}$ so that

$$(J_{\eta\gamma})_{\eta \neq \gamma} \overset{d}{=} (I_\eta | I_\gamma = 1)_{\eta \neq \gamma}. \tag{2.10.1}$$

Many of the indicators that we have constructed are increasing functions of ω_0, which makes those sets of indicators positively related ([9], Section 2.1). However, the $\emptyset$-line and non-enhanced line indicators, $I_\ell^{\emptyset L}$ and I_ℓ^{NEL}, are not increasing functions of ω_0, so whenever these appear we are unable to use the simpler form of the Poisson convergence theorem. Instead, we will explicitly define the couplings below, and use Theorem 10.J of [9], which we state below as Lemma 2.10.1.

Suppose X and Y are two $\mathbb{Z}^m$-valued random variables with laws μ_X and μ_Y, and recall that the total variation distance between μ_X and μ_Y (or with an abuse of notation, between

X and Y or X and μ_Y) is

$$d_{TV}(X, Y) = d_{TV}(\mu_X, \mu_Y) := \sup_{A \subset \mathbb{Z}^m} |\mu_X(A) - \mu_Y(A)| = \frac{1}{2} \sum_{k \in \mathbb{Z}^m} |\mu_X(k) - \mu_Y(k)| .$$

Let P_λ denote the law of a Poisson(λ) random variable (taking values in $\mathbb{Z}_+$). The Chen-Stein method gives us the following bound on the total variation distance between the joint law of $(W_1, W_2, \ldots, W_m)$ and $\prod_{i=1}^m P_{\lambda_i}$.

Lemma 2.10.1. *([9], Theorem 10.J and Corollary 10.J.1) If W_i are defined as above with $\lambda_i = EW_i$ for $i = 1, \ldots, \ell$, with $EI_\gamma = p_\gamma$, then*

$$d_{TV}\left((W_1, \ldots, W_m), \prod_{i=1}^m P_{\lambda_i}\right) \le \sum_{\gamma \in \Gamma} p_\gamma^2 + \sum_{\substack{\gamma, \eta \in \Gamma \\ \gamma \neq \eta}} p_\gamma \mathbb{E} |J_{\eta\gamma} - I_\eta| . \tag{2.10.2}$$

If $\{I_\gamma\}_{\gamma \in \Gamma}$ are positively related then

$$d_{TV}\left((W_1, \ldots, W_m), \prod_{i=1}^m P_{\lambda_i}\right) \le \sum_{\gamma \in \Gamma} p_\gamma^2 + \sum_{\substack{\gamma, \eta \in \Gamma \\ \gamma \neq \eta}} \mathrm{Cov}(I_\gamma, I_\eta). \tag{2.10.3}$$

Remark 2.10.2. In all of our applications of Lemma 2.10.1, the first sum on the right hand side is easy to control, since it merely requires that p_γ are uniformly small. In the case of events indexed by $\mathscr{L}$ this sum is $O(n^{-2})$, since there are $O(n^2)$ summands and the probability of a line configuration is $O(n^2p^2) = O(n^{-2})$. Similarly, in the case of basic or enhanced basic events this sum is $O(n^{-3})$. The important part of the right hand side is the term $\mathbb{E} |J_{\eta\gamma} - I_\eta| = \mathbb{P}(J_{\eta\gamma} \neq I_\eta)$, which requires bounding the probability that our coupling destroys or creates the event indicated by I_η. In the case of positively related indicators, no explicit coupling is needed, and we must merely bound the covariances between the relevant indicators.

Construction of couplings

Observe that in equation (2.6.1), the last term involves random variables that are sums of indicators that are not positively related. So, for each of the indicators $I_v^B, I_\ell^{\emptyset L}, I_\ell^{EL}, I_\ell^{NEL}$

and every $v \in V$ and $\ell \in \mathscr{L}$, we must construct a suitable coupling between all of the remaining indicators and their conditioned versions as in (2.10.1). As in (2.10.1), we will use the letter J for coupled indicator random variables.

Once we show that these random variables appearing in the last term of (2.6.1) converge jointly to independent Poissons, we will be able to compute the limiting probabilities for all of the terms except the second, which involves the `EnhancedBasic` and `Line` random variables. We will treat this term separately using the simpler form of Lemma 2.10.1, since the enhanced basic and line indicators are positively related.

Our goal is to show that the second summation in (2.10.2) is $O(n^{-1})$ under the couplings that we construct. We will need to construct four couplings, one for each type of indicator, and for each coupling we have four comparisons (to each of the four types of indicators) that need to be made. Furthermore, for each comparison, there are several cases that need to be checked depending on the relative positions of the vertices and lines that index each event. There are many cases that need to be verified, but the arguments quickly become repetitive, thus we merely outline the proof and give complete details in two typical cases (see proofs of Lemmas 2.10.6 and 2.10.7).

We begin with the simplest case, the *basic coupling* for conditioning on $I_v^B = 1$ for a fixed $v \in V$. In this case, we merely need each of the three lines containing v to contain at least one open vertex. To achieve this, we extend the probability space by possibly resampling the vertices in each of the three lines until this condition is met. That is, if a line through v already contains an open vertex, nothing is resampled for that line, and the original configuration is kept, otherwise it is repeatedly replaced with an independent configuration until it does contain an open vertex. Also, it is important to note that none of the other vertices in the initial configuration, ω_0, are altered. Then, $J_{wv}^B, J_{\ell v}^{\emptyset L}, J_{\ell v}^{EL}, J_{\ell v}^{NEL}$ are the indicator random variables of the corresponding events after the local resampling is completed. Since v is fixed and the Hamming torus is transitive, we will drop the index v in the conditioning on $I_v^B = 1$.

Lemma 2.10.3. *Under the basic coupling, the following sums are all* $O(n^{-1})$

$$\sum_{v\in V}\sum_{\substack{w\in V\\ w\neq v}} EI_v^{\mathrm{B}}\ \mathbb{P}\big(I_w^{\mathrm{B}}\neq J_w^{\mathrm{B}}\big), \qquad \sum_{v\in V}\sum_{\ell\in\mathscr{L}} EI_v^{\mathrm{B}}\ \mathbb{P}\Big(I_\ell^{\emptyset\mathrm{L}}\neq J_\ell^{\emptyset\mathrm{L}}\Big),$$

$$\sum_{v\in V}\sum_{\ell\in\mathscr{L}} EI_v^{\mathrm{B}}\ \mathbb{P}\big(I_\ell^{\mathrm{NEL}}\neq J_\ell^{\mathrm{NEL}}\big), \qquad \sum_{v\in V}\sum_{\ell\in\mathscr{L}} EI_v^{\mathrm{B}}\ \mathbb{P}\big(I_\ell^{\mathrm{EL}}\neq J_\ell^{\mathrm{EL}}\big).$$

The next simplest coupling is the $\emptyset$*-line coupling* for the conditioning on $I_\ell^{\emptyset L}=1$ for a fixed $\ell\in\mathscr{L}$. For this coupling, we need the line ℓ to contain at least two initially open vertices, so we first resample the vertices in ℓ if necessary until this condition is met. Given the locations of the open vertices in ℓ, we need the two planes containing ℓ to have no open vertices that are not neighbors of the open vertices in ℓ. To achieve this, we simply remove any violating vertices from ω_0. In the next three lemmas, we use indicators J, with proper subscripts and superscripts, in an analogous fashion as in Lemma 2.10.3.

Lemma 2.10.4. *Under the* $\emptyset$*-line coupling the following sums are* $O(n^{-1})$

$$\sum_{\ell\in\mathscr{L}}\sum_{w\in V} EI_\ell^{\emptyset\mathrm{L}}\ \mathbb{P}\big(I_w^{\mathrm{B}}\neq J_w^{\mathrm{B}}\big), \qquad \sum_{\ell\in\mathscr{L}}\sum_{\substack{\ell'\in\mathscr{L}\\ \ell'\neq\ell}} EI_\ell^{\emptyset\mathrm{L}}\ \mathbb{P}\Big(I_{\ell'}^{\emptyset\mathrm{L}}\neq J_{\ell'}^{\emptyset\mathrm{L}}\Big),$$

$$\sum_{\ell\in\mathscr{L}}\sum_{\ell'\in\mathscr{L}} EI_\ell^{\emptyset\mathrm{L}}\ \mathbb{P}\big(I_{\ell'}^{\mathrm{NEL}}\neq J_{\ell'}^{\mathrm{NEL}}\big), \qquad \sum_{\ell\in\mathscr{L}}\sum_{\ell'\in\mathscr{L}} EI_\ell^{\emptyset\mathrm{L}}\ \mathbb{P}\big(I_{\ell'}^{\mathrm{EL}}\neq J_{\ell'}^{\mathrm{EL}}\big).$$

Next, we construct the *enhanced line coupling* for the conditioning on $I_\ell^{EL}=1$ for a fixed $\ell\in\mathscr{L}$. To achieve this, we will need the line ℓ to contain at least two open vertices, so we first resample the vertices in ℓ if necessary until this condition is met. Next, given the locations of the open vertices in ℓ, we need that at least one of the two planes containing ℓ has at least one open vertex that is not collinear with an open vertex in ℓ. Again, if necessary, we resample these two planes (excepting the vertices in ℓ) simultaneously until this condition is satisfied. At this point, if one of the two planes containing ℓ has at least two non-neighboring open vertices, then the coupling is completed. Otherwise, conditional on the location of the open vertex (or vertices) in $\mathcal{N}(\ell)$, we need there to be at least one open vertex in the same plane as this vertex (or vertices) but not in the same line. If one does not exist, then we resample the two (or four) planes containing the open vertex (or vertices) in $\mathcal{N}(\ell)$ but not containing ℓ until there is at least one open vertex in any of these planes (we do not resample the vertices in ℓ, $\mathcal{N}(\ell)$, or the neighborhood of the open vertices in $\mathcal{N}(\ell)$).

Lemma 2.10.5. *Under the enhanced line coupling the following sums are* $O(n^{-1})$

$$\sum_{\ell\in\mathscr{L}}\sum_{w\in V} EI_\ell^{\mathrm{EL}}\ \mathbb{P}\big(I_w^{\mathrm{B}} \neq J_w^{\mathrm{B}}\big), \qquad \sum_{\ell\in\mathscr{L}}\sum_{\ell'\in\mathscr{L}} EI_\ell^{\mathrm{EL}}\ \mathbb{P}\Big(I_{\ell'}^{\emptyset\mathrm{L}} \neq J_{\ell'}^{\emptyset\mathrm{L}}\Big),$$

$$\sum_{\ell\in\mathscr{L}}\sum_{\ell'\in\mathscr{L}} EI_\ell^{\mathrm{EL}}\ \mathbb{P}\big(I_{\ell'}^{\mathrm{NEL}} \neq J_{\ell'}^{\mathrm{NEL}}\big), \qquad \sum_{\ell\in\mathscr{L}}\sum_{\substack{\ell'\in\mathscr{L}\\ \ell'\neq\ell}} EI_\ell^{\mathrm{EL}}\ \mathbb{P}\big(I_{\ell'}^{\mathrm{EL}} \neq J_{\ell'}^{\mathrm{EL}}\big).$$

Finally, we construct the *non-enhanced line coupling* for the conditioning on $I_\ell^{NEL} = 1$ for a fixed $\ell \in \mathscr{L}$. To achieve this, we will need the line ℓ to contain at least two open vertices. So first we resample the vertices in ℓ if necessary until this condition is met. Next, given the locations of the open vertices in ℓ, we need: 1) that at least one of the two planes containing ℓ has at least one open vertex that is not collinear with an open vertex in ℓ, and 2) that neither plane containing ℓ has more than one non-collinear open vertex. Again, if necessary, we resample these two planes simultaneously until these conditions are met (here we do not resample ℓ). Now, conditional on the locations of the open points in $\mathcal{N}(\ell)$, we must guarantee that there are no other points outside of ℓ that are coplanar but not collinear with these points. For this part of the coupling, we simply remove any violating points from ω_0.

Lemma 2.10.6. *Under the non-enhanced line coupling the following sums are* $O(n^{-1})$

$$\sum_{\ell\in\mathscr{L}}\sum_{w\in V} EI_\ell^{\mathrm{NEL}}\ \mathbb{P}\big(I_w^{\mathrm{B}} \neq J_w^{\mathrm{B}}\big), \qquad \sum_{\ell\in\mathscr{L}}\sum_{\ell'\in\mathscr{L}} EI_\ell^{\mathrm{NEL}}\ \mathbb{P}\Big(I_{\ell'}^{\emptyset\mathrm{L}} \neq J_{\ell'}^{\emptyset\mathrm{L}}\Big),$$

$$\sum_{\ell\in\mathscr{L}}\sum_{\substack{\ell'\in\mathscr{L}\\ \ell'\neq\ell}} EI_\ell^{\mathrm{NEL}}\ \mathbb{P}\big(I_{\ell'}^{\mathrm{NEL}} \neq J_{\ell'}^{\mathrm{NEL}}\big), \qquad \sum_{\ell\in\mathscr{L}}\sum_{\ell'\in\mathscr{L}} EI_\ell^{\mathrm{NEL}}\ \mathbb{P}\big(I_{\ell'}^{\mathrm{EL}} \neq J_{\ell'}^{\mathrm{EL}}\big).$$

Proof. We now outline the proof by bounding the first summation above. There are three cases.

Case 1: $w \in \ell$. This term appears in the sum $O(n^3)$ times, and $EI_\ell^{NEL} = O(n^{-2})$, so we must show that $\mathbb{P}(I_w^B \neq J_w^B) = O(n^{-2})$. Now there are two subcases, *destruction* and *creation* respectively: $\mathbb{P}(I_w^B = 1, J_w^B = 0)$ and $\mathbb{P}(I_w^B = 0, J_w^B = 1)$. Clearly, $\mathbb{P}(I_w^B = 1, J_w^B = 0) \leq$

$\mathbb{P}(I_w^B = 1) = O(n^{-3})$. Next, in order for the creation event to occur, the resampling procedure must have generated at least one open vertex in both planes containing ℓ, and both of these points must lie in the neighborhood of w. The probability of this is $O(n^{-2})$, since we require an open vertex in each of two fixed lines.

Case 2: $w \in \mathcal{N}(\ell)$. This term appears in the sum $O(n^4)$ times, and $EI_\ell^{NEL} = O(n^{-2})$, so we must show that $\mathbb{P}(I_w^B \neq J_w^B) = O(n^{-3})$. Once again, there are two subcases as above. The creation event cannot occur in this case because an open vertex in $\mathcal{N}(\ell)$ that is collinear with w must not see any coplanar open vertices (off of ℓ), which includes a line in the neighborhood of w, so w can no longer see an open vertex in each direction. The probability of the destruction event can be trivially bounded by $O(n^{-3})$ as in Case 1.

Case 3: $w \notin \mathcal{N}(\ell) \cup \ell$. This term appears in the sum $O(n^5)$ times, and $EI_\ell^{NEL} = O(n^{-2})$, so we must show that $\mathbb{P}(I_w^B \neq J_w^B) = O(n^{-4})$. Once again, the creation event cannot occur for the same reason as cited in Case 2. The destruction event can only occur if one of the initially open points in the neighborhood of w is in one of the resampled planes. At most six planes are affected with probability $1 - O(n^{-1})$, and with the same probability none of the resampled planes contain a line in the neighborhood of w. The probability of the destruction event is at most $O(n^{-4})$, since w must first have three open neighbors initially (an event with probability $O(n^{-3})$), and at least one must coincide with one of the resampled planes (an event with probability $O(n^{-1})$). ☐

Positively related case

Since $\{I_v^{EB}\}_{v \in V}$ and $\{I_\ell^L\}_{\ell \in \mathscr{L}}$ are all increasing functions of ω_0, these collections of indicators are positively related so we may apply the simpler form of Lemma 2.10.1 by bounding the covariances.

Lemma 2.10.7. *The collections of indicators* $\{I_v^{\mathrm{EB}}\}_{v \in V}$ *and* $\{I_\ell^{\mathrm{L}}\}_{\ell \in \mathscr{L}}$ *are positively related* *and the following sums are* $O(n^{-1})$

$$\sum_{v\in V}\sum_{\substack{w\in V\\ w\neq v}} \mathrm{Cov}(I_v^{\mathrm{EB}}, I_w^{\mathrm{EB}}), \qquad \sum_{v\in V}\sum_{\ell\in\mathscr{L}} \mathrm{Cov}(I_v^{\mathrm{EB}}, I_\ell^{\mathrm{L}}), \qquad \sum_{\ell\in\mathscr{L}}\sum_{\substack{\ell'\in\mathscr{L}\\ \ell'\neq\ell}} \mathrm{Cov}(I_\ell^{\mathrm{L}}, I_{\ell'}^{\mathrm{L}}).$$

Note that the bound on the last sum, which involves only indicators of line events, is implied by combining the results for the enhanced line and non-enhanced line couplings in Lemmas 2.10.5 and 2.10.6 by writing $I_\ell^L = I_\ell^{EL} + I_\ell^{NEL}$.

Proof. We will explain the proof of the bound on the first sum, as the second sum is evaluated in a similar fashion and the third is implied by previous lemmas. We break up the sum into three cases depending on the Hamming distance between v and w.

Case 1: $d(v,w) = 1$. There are $O(n^4)$ such terms in the sum, so we need to show that the covariance is $O(n^{-5})$. In this case it suffices to use the trivial bound $\mathrm{Cov}(I_v^{EB}, I_w^{EB}) \leq \mathbb{E}I_v^{EB}I_w^{EB} = \mathbb{P}(G_v^{EB} \cap G_w^{EB})$, which is the probability that an enhanced basic configuration appears at v and at w. For this event to occur, v must have one open neighbor in each direction, one of which is shared with w, so w needs only one open neighbor in each direction orthogonal to $w - v$. This is a total of at least five open points on five fixed lines, which has probability $O((np)^5) = O(n^{-5})$ as desired.

Case 2: $d(v,w) = 2$. There are $O(n^5)$ such terms in the sum, so we need to show that the covariance is $O(n^{-6})$. Again, it suffices to use the bound $\mathrm{Cov}(I_v^{EB}, I_w^{EB}) \leq \mathbb{E}I_v^{EB}I_w^{EB}$. In this case, the vertices v and w have exactly two common neighbors, so there are three cases: zero, one, or two of these common neighbors are initially open. If neither common neighbor is initially open, then v and w each independently need one open neighbor in each direction – a total of six open vertices in six fixed lines, which has probability $O(n^{-6})$. If one of the common neighbors is open, an event with probability $O(p) = O(n^{-2})$, then v and w each need an open neighbor in two other directions – a total of four open vertices in four fixed lines which has probability $O(n^{-4})$. This gives a probability of $O(n^{-6})$ to the case where one common neighbor is open. The event that both common neighbors are open has probability $p^2 = O(n^{-4})$, and v and w each require one more occupied neighbor in one direction, which has probability $O(n^{-2})$ for a total probability of $O(n^{-6})$.

Case 3: $d(v,w) = 3$. There are $O(n^6)$ such terms in the sum, so we need to show that the covariance is $O(n^{-7})$, and the trivial upper bound on the covariance will not suffice. Observe that the planes containing v and the planes containing w intersect only along 6 lines, and conditional on the event that none of the points on these lines are initially open, I_v^{EB} and I_w^{EB} are independent. Call this event E_{empty}, then since I_v^{EB} and I_w^{EB} are increasing functions of ω_0, the covariance is bounded by

$$\mathrm{Cov}(I_v^{EB}, I_w^{EB}) \leq \mathbb{P}(I_v^{EB} I_w^{EB} = 1, E_{\text{empty}}^c)$$

We now divide the event E_{empty}^c into subcases according to which vertices in the intersection are open. There are two types of vertices in the intersection – those which are neighbors to either v or w, and those which are only in the same plane as each vertex. There are exactly 6 vertices in the former category and $6(n-2)$ in the latter. The probability that j of the 6 vertices in $[\mathcal{N}(v) \cap \mathcal{N}(\mathcal{N}(w))] \cup [\mathcal{N}(w) \cap \mathcal{N}(\mathcal{N}(v))]$ are initially open is $O(p^j) = O(n^{-2j})$. Conditional on this, v and w collectively require an initially open vertex in each of the remaining $6-j$ lines in their neighborhoods, which has probability $O(n^{-6+j})$, giving a total probability of $O(n^{-6-j})$ to the event that there are j of these 6 vertices initially open and both enhanced basic events occur. Therefore, if $j \geq 1$ we are done, otherwise we must consider the case where $j = 0$ and then E_{empty}^c requires that at least one vertex among the $6(n-2)$ vertices in $\mathcal{N}(\mathcal{N}(v)) \cap \mathcal{N}(\mathcal{N}(w))$ are initially open. This event has probability $O(np) = O(n^{-1})$, and when $j = 0$, v and w still need one open vertex in each line of their neighborhoods, which has probability $O(n^{-6})$, giving a total probability of $O(n^{-7})$. $\square$

Proof of Lemma 2.6.2. The limiting means are straightforward to calculate, as outlined in remark 2.6.3. It is also not difficult to show that $d_{TV}(P_{\lambda_n}, P_\lambda) \leq |\lambda_n - \lambda|$ so if $\lambda_n \to \lambda$ then P_{λ_n} converges to P_λ. Therefore, applying Lemma 2.10.1 and using Lemmas 2.10.3-2.10.6 to bound the second summation in (2.10.2) implies that the random variables `Basic`, $\texttt{Line}\emptyset_i$, $\texttt{NonEnhancedLine}_i$, and `EnhancedLine` (where $i = 1, 2, 3$, so there are a total of 8 random variables) converge jointly to independent Poisson random variables with the appropriate limiting means. Similarly, applying Lemma 2.10.1 and using Lemma 2.10.7 to bound the second summation in (2.10.3) implies that the random variables `EnhancedBasic` and `Line` converge jointly to independent Poisson random variables with the appropriate limiting means. □

Chapter 3

A SEMIDEFINITE APPROACH TO THE K_i-COVER PROBLEM

3.1 Introduction

A common way to model a combinatorial optimization problem is as the optimization of a function over the set $S \subseteq \{0,1\}^n$ of characteristic vectors of the objects in question. When the objective function is linear, we may replace S by its convex hull $\operatorname{conv}(S)$. The problem can be solved efficiently if we can find a small description of this polytope. Since for NP hard problems we cannot expect this, we look instead for approximations to $\operatorname{conv}(S)$. One possibility is to use semidefinite approximations, as introduced by Lovász [35] with the construction of the *theta body* of the stable set polytope of a graph. Another famous example is the approximation algorithm for the max cut problem due to Goemans and Williamson [18]. In this paper we will use the semidefinite relaxations introduced by Gouveia, Parrilo and Thomas [21] to analyze the *K_i-cover problem.*

Recall that K_i denotes the complete graph, or clique, on i vertices. Given a graph G, let $\mathbf{K}_j(G)$ be the collection of cliques in G of size j (usually, the graph is clear from context, and we write $\mathbf{K}_j$). A (possibly empty) collection $C \subset \mathbf{K}_{i-1}$ is said to be a K_i-cover if for each $K \in \mathbf{K}_i$, there is some $H \in C$ with $H \subset K$. In this case we say that H covers K. The K_i-cover problem is, given a graph G and a set of weights on $\mathbf{K}_{i-1}$, to compute the minimum weight K_i-cover. The case $i = 2$ is more commonly known as the vertex cover problem, in which we seek a collection C of vertices such that each edge in G contains at least one vertex from C. However, note that the usage of "cover" is reversed here: the vertex cover problem is the K_2-cover problem, not the K_1-cover problem.

A closely related problem, and the setting in which we will prove our results, is the *K_i-free problem.* As before, we are given a graph and a collection of weights on $\mathbf{K}_{i-1}$. But now we seek the maximum weight collection $C \subseteq \mathbf{K}_{i-1}$ such that C is K_i-free. That is, for each $K \in \mathbf{K}_i$, there is some $H \in \mathbf{K}_{i-1}$, with $H \subset K$ and $H \notin C$. Again, the case $i = 2$ of this problem is well-known as the stable set problem: we seek a maximum weight *stable set* C, where C is stable if no two of its vertices are connected by an edge.

The vertex cover and stable set problems are related in the following sense: let $G = (V, E)$ be a graph. Then a subset C of vertices is a vertex cover if and only if $V \setminus C$ is a stable set. The same is true for the K_i-cover and K_i-free problems: a subset $C \subset \mathbf{K}_{i-1}$ is a K_i-cover if and only if $\mathbf{K}_{i-1} \setminus C$ is K_i-free. Therefore, for a given set of weights on $\mathbf{K}_{i-1}$, optimal solutions to the two problems are complementary, and so solving one solves the other.

In this paper, we consider the polytope associated with the K_i-free problem. Let $P_i(G) = \operatorname{conv}(\{\chi_S : S \subset \mathbf{K}_{i-1}(G) \text{ and } S \text{ is } K_i\text{-free}\})$, the convex hull of the incidence vectors of the K_i-free sets. Note that $P_i(G) \subseteq [0,1]^{\mathbf{K}_{i-1}(G)}$.

As the K_i-free problem is NP-complete (see [15]), we cannot expect a small description of $P_i(G)$ for general graphs G. However, for certain classes of facets of $P_i(G)$, Conforti, Corneil, and Mahjoub [15] show that we can solve the separation problem in polynomial time, allowing us to optimize efficiently over a relaxation of $P_i(G)$. We provide a strictly tighter relaxation of $P_i(G)$, improving their optimization result, but without proving the existence of polynomial separation oracles for any new family of facets.

The structure of this paper is: in Section 2, we outline the main algebraic machinery, *theta bodies*, a semidefinite relaxation hierarchy. In Section 3 we show that the K_i-p-hole facets are valid on $\lceil i/2 \rceil$ level of the theta body hierarchy. Finally, in Section 4 we focus on the triangle free problem. We show that in the case of $G = K_n$, the theta body relaxations cannot converge in less than $(n-2)/4$ steps. We also use a result of Krivelevich [32] to show an integrality gap of 2 for the second theta body.

Acknowledgements

The authors would like to thank the anonymous referees for their helpful comments. The authors were partially supported on this project as follows: JG by 'Centro de Matemática da

Universidade de Coimbra' and 'Fundação para a Ciência e a Tecnologia', through European program COMPETE/FEDER; and JP by NSF grant DMS-1115293.

3.2 Theta bodies

Theta bodies are semidefinite approximations to the convex hull of an algebraic variety. For background, see [22] and [21]. Here we state the necessary results for this paper without proofs.

Let $V \subseteq \mathbb{R}^n$ be a finite point set. One description of the convex hull of V is as the intersection of all affine half spaces containing V (recall that $f|_V$ is the restriction of f to V):

$$\operatorname{conv}(V) = \{x \in \mathbb{R}^n : f(x) \geq 0 \text{ for all linear } f \text{ such that } f|_V \geq 0\}.$$

Since it is computationally intractable to find whether $f|_V \geq 0$, we relax this condition. Let I be the vanishing ideal of V, i.e., the set of all polynomials vanishing on V. Recall that $f \equiv g \mod I$ means $f - g \in I$, and implies that $f(x) = g(x)$ for all $x \in V$. A function f is said to be a sum of squares of degree at most k mod I, or *k-sos mod I*, if there exist functions g_j, $j = 1, \ldots, m$ with degree at most k, such that $f \equiv \sum_{j=1}^{m} g_j^2 \mod I$. If f is k-sos mod I for any k, it is clear that $f|_V \geq 0$ since g_j^2 is visibly nonnegative on V. Therefore, we make the following definition of $\mathrm{TH}_k(I)$, the k-th theta body of I:

$$\mathrm{TH}_k(I) = \{x \in \mathbb{R}^n : f(x) \geq 0 \text{ for all linear } f \equiv k\text{-sos mod } I\}.$$

The reason why the theta bodies $\mathrm{TH}_k(I)$ provide a computationally tractable relaxation of $\operatorname{conv}(V)$ is that the membership problem for $\mathrm{TH}_k(I)$ can be expressed as a semidefinite program, using *moment matrices* that are reduced mod I.

For what follows, we will restrict ourselves to a special class of varieties, and suppose that our variety $V \subseteq \{0,1\}^n$ and is *lower-comprehensive*; i.e., if $x \leq y$ componentwise, and $y \in V$, then $x \in V$. Additionally, we will always assume that V contains the canonical basis of $\mathbb{R}^n$, $\{e_1, \cdots, e_n\}$, as otherwise we could restrict ourselves to a subspace. All combinatorial optimization problems of avoiding certain finite list of configurations, such as stable set,

K_i-free, etc., have lower-comprehensive varieties. The restriction to this class is not necessary, but makes the theta body exposition simpler. In particular, the ideal of a lower-comprehensive variety has the following simple description.

Lemma 3.2.1. *Let V be a lower-comprehensive subset of $\{0,1\}^n$. Then its vanishing ideal is given by*

$$I = \langle x_j^2 - x_j : j = 1, \ldots, n; x^S : S \notin V \rangle,$$

and a basis for $\mathbb{R}[V] = \mathbb{R}[x]/I$ is given by $B = \{x^S : S \in V\}$, where $x^S := \prod_{i \in S} x_i$ is a shorthand used throughout the paper.

Another important fact about $\mathrm{TH}_k(I)$ in this setting (when I is a real ideal) is that a linear inequality $f(x) \geq 0$ is valid on $\mathrm{TH}_k(I)$ if and only if f is actually k-sos modulo I. In Section 3, we will prove that certain facet-defining inequalities of $P_i(G)$ are also valid on its theta relaxations $\mathrm{TH}_k(I)$ by presenting a sum of squares representation modulo the ideal. For now, we observe that by considering degrees, we can get a bound on which theta bodies are trivial; that is, equal to the hypercube $[0,1]^n$.

Lemma 3.2.2. *Let $V \subseteq \{0,1\}^n$ be lower-comprehensive, and suppose that all elements $x \notin V$ have $\sum_j x_j \geq k$. Let I be the vanishing ideal of V. Then for $l < k/2$, $\mathrm{TH}_l(I) = [0,1]^n$.*

Proof. Let f be linear with $f \equiv \sum_j g_j^2 \mod I$ with each g_j of degree at most l. Then $f - \sum_j g_j^2 =: F \in I$, and F has degree at most $2l < k$. But the basis from Lemma 3.2.1 is a Groebner basis, and the only elements with degree less than k are $x_j^2 - x_j$, so $F \in I' := \langle x_j^2 - x_j ; j = 1, \ldots, n \rangle$. Thus $\mathrm{TH}_l(I) \supseteq \mathrm{TH}_l(I') = [0,1]^n$. □

Let V_k be the subset of V whose elements have at most k entries equal to one. For convenience, we will often identify the elements of V, characteristic vectors χ_S for $S \subseteq \{1, \ldots, n\}$, with their supports, via $S \leftrightarrow \chi_S$. Given $y \in \mathbb{R}^{V_{2k}}$ we denote the *reduced moment matrix* of y with respect to I to be the matrix $M_{V_k}(y) \in \mathbb{R}^{V_k \times V_k}$ defined by

$$[M_{V_k}(y)]_{X,Y} = \begin{cases} y_{X \cup Y} & \text{if } X \cup Y \in V, \\ \\ 0 & \text{otherwise.} \end{cases}$$

With these matrices we can finally give a semidefinite description of $\mathrm{TH}_k(I)$.

Proposition 3.2.3. *With I and V as before, $\mathrm{TH}_k(I)$ is the canonical projection onto $\mathbb{R}^n$ via the coordinates $(y_{e_1}, \cdots, y_{e_n})$ of the set*

$$\{y \in \mathbb{R}^{V_{2k}} \ : \ M_{V_k}(y) \succeq 0 \text{ and } y_0 = 1\}.$$

In particular, optimizing to arbitrary fixed precision over $\mathrm{TH}_k(I)$ can be done in time polynomial in n, for fixed k.

Now we can consider the specific case of the K_i-free problem. Here the variety $V \subseteq \mathbb{R}^{\mathbf{K}_{i-1}(G)}$ is the set of characteristic vectors of K_i-free subsets of $\mathbf{K}_{i-1}(G)$, V_k is the subset of V of elements of size at most k, and I is the vanishing ideal of V, described by Lemma 3.2.1. Since the K_is in G are the minimal elements not in V, by Lemma 3.2.1 we can write the ideal I as follows.

$$I = \langle x_j^2 - x_j : j \in \mathbf{K}_{i-1}(G); \prod_{j \subseteq K} x_j : K \in \mathbf{K}_i(G) \rangle.$$

For example, let G be a triangle, with edges A, B, C, and consider the triangle free problem on G. Then the ideal is

$$I = \langle x_A^2 - x_A, x_B^2 - x_B, x_C^2 - x_C, x_A x_B x_C \rangle,$$

and the variety V is as follows.

$$V = \{\emptyset, \{A\}, \{B\}, \{C\}, \{A,B\}, \{A,C\}, \{B,C\}\} \equiv \{0,1,2,3,4,5,6\}.$$

Note that here, we again use our identification of sets with their characteristic vectors. To avoid writing, e.g., $y_{\{A,C\}}$ or even $y_{\chi_{\{A,C\}}}$, we label the elements of V by numbers as above.

Then the moment matrix $M_{V_2}(y)$ is as follows:

$$M_{V_2}(y) = \begin{bmatrix} y_0 & y_1 & y_2 & y_3 & y_4 & y_5 & y_6 \\ y_1 & y_1 & y_4 & y_5 & y_4 & y_5 & 0 \\ y_2 & y_4 & y_2 & y_6 & y_4 & 0 & y_6 \\ y_3 & y_5 & y_6 & y_3 & 0 & y_5 & y_6 \\ y_4 & y_4 & y_4 & 0 & y_4 & 0 & 0 \\ y_5 & y_5 & 0 & y_5 & 0 & y_5 & 0 \\ y_6 & 0 & y_6 & y_6 & 0 & 0 & y_6 \end{bmatrix}$$

Projecting the set $\{y : y_0 = 1, M_{V_2}(y) \succeq 0\}$ onto (y_1, y_2, y_3) gives $\mathrm{TH}_2(I)$ for this graph.

3.3 Polynomial-time algorithm

A graph H is a K_i-p-hole if H is the union of $G_1, \dots, G_p$, each a copy of K_i, where G_j and G_l share a common K_{i-1} if and only if $j - l = \pm 1 \mod p$; see Figure 3.1. Theorem 3.5 in [15] establishes that for $i \geq 3$ and odd p, the inequality $\sum_{\mathbf{K}_{i-1}(H)} x_j \leq (\frac{p-1}{2})(2i-3) + i - 2$ defines a facet of $P_i(G)$ for each induced K_i-p-hole H of G. We will show that the facets corresponding to induced K_i-p-holes are valid on $\mathrm{TH}_{\lceil i/2 \rceil}(I)$, which can be optimized over in polynomial time for fixed i, and relate this complexity result with the ones in Conforti, Corneil and Mahjoub [15]. Note that in this section, the ideal I always refers to the K_i-free problem, and the associated graph G will be clear from context. Therefore, we will say k-sos for k-sos mod I.

The first lemma is an auxiliary result that certain functions are sums of squares. For an ideal I, a function f is said to be *idempotent* mod I if $f^2 \equiv f \mod I$. Since an idempotent is visibly a square, we can use it as a summand in our sum of squares. In practice, idempotents end up being very useful in constructing sums of squares.

Lemma 3.3.1. *Suppose* $A \subseteq B \subseteq \mathbf{K}_{i-1}(K_i)$. *Denote the variables corresponding to elements of* $\mathbf{K}_{i-1}(K_i)$ *by* $\{x_k : 1 \leq k \leq i\}$. *Then* $f(x) = |B \setminus A| - x^A + x^B - \sum_{k \in B \setminus A} x_k$ *is* $|B|$*-sos mod* I.

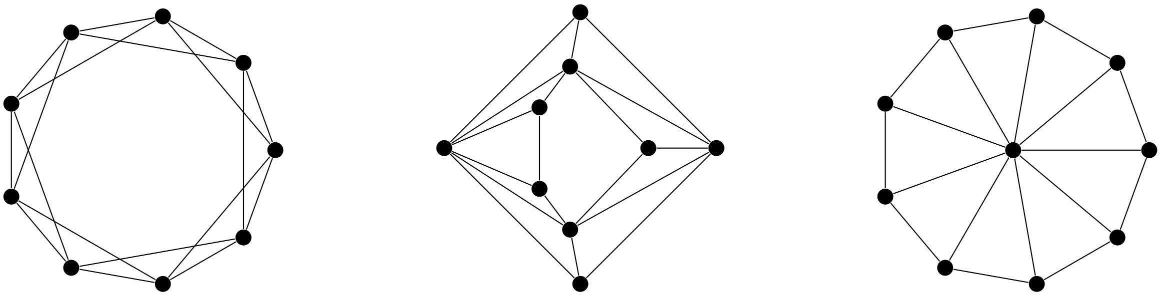


Figure 3.1: Three non-isomorphic K_3-9-holes.

Proof. Let $A = A_1 \subset A_2 \ldots \subset A_m = B$, where $|A_{k+1} \setminus A_k| = 1$, and without loss of generality, the variable corresponding to the element of $A_{k+1} \setminus A_k$ is x_k. Check that $g_k(x) = 1 - x_k - x^{A_k} + x^{A_{k+1}}$ is idempotent mod I. Adding them up we get that $f(x) = \sum_{k=1}^{m-1} g_k(x)$. Since each summand has degree at most $|B|$ the assertion holds. $\square$

The stable set polytope STAB(G) has a fractional relaxation FRAC(G), given by imposing nonnegativities $x_i \geq 0$, and inequalities $x_i + x_j \leq 1$ for each edge (i, j) of G. Similarly, we can define a fractional K_i-free polytope $\mathrm{FRAC}_i(G)$ by imposing nonnegativities, and the inequalities $\sum_{k \in \mathbf{K}_{i-1}(H)} x_k \leq i - 1$ for each $H \in \mathbf{K}_i(G)$. The following corollary shows that these inequalities are $\lceil i/2 \rceil$-sos, and therefore that the relaxation $\mathrm{TH}_{\lceil i/2 \rceil}(I) \subseteq \mathrm{FRAC}_i(G)$. This is parallel to the result that the Lovász theta body lies inside FRAC(G).

Corollary 3.3.2. *The inequality* $\sum_{k \in \mathbf{K}_{i-1}(H)} x_k \leq i - 1$ *is valid on* $\mathrm{TH}_{\lceil i/2 \rceil}(I)$ *for every* $H \in \mathbf{K}_i(G)$.

Proof. Let J be a subset of $\mathbf{K}_{i-1}(H)$ of size $\lceil i/2 \rceil$. Applying Lemma 3.3.1 with $A = \emptyset$ and $B = J$ we see that

$$f(x) = |J| - 1 + x^J - \sum_{l \in J} x_l$$

is $|J|$-sos. Similarly

$$g(x) = |J^c| - 1 + x^{J^c} - \sum_{l \in J^c} x_l$$

is $|J^c|$-sos (J^c is the complement of J in $\mathbf{K}_{i-1}(H)$). Finally observe that $h(x) = 1 - x^J - x^{J^c}$ is idempotent. Since these polynomials are all $\lceil i/2 \rceil$-sos, it remains to observe that their sum,

$$f(x) + g(x) + h(x) = i - 1 - \sum_{k \in \mathbf{K}_{i-1}(H)} x_k,$$

is also $\lceil i/2 \rceil$-sos. □

Now we are ready to prove that the K_i-p-hole inequalities are valid on $\mathrm{TH}_{\lceil i/2 \rceil}(I)$. Recall that if H is a K_i-p-hole, we write $G_1, \ldots, G_p$ for the K_is in H, with adjacent K_i sharing a common K_{i-1}.

Theorem 3.3.3. *[15, Theorem 3.4] If G has an induced K_i-p-hole H, then the inequality*

$$\frac{p-1}{2}(2i-3) + i - 2 - \sum_{i \in H} x_i \geq 0$$

defines a facet of $P_i(G)$ for $i \geq 3$.

For even i these inequalities are still valid, but not facets anymore. In the next result we give a sums of squares certificate to the validity of these inequalities.

Lemma 3.3.4. *The K_i-p-hole inequalities are $\lceil i/2 \rceil$-sos for p odd.*

Proof. Let $p = 2k+1$. For each $l = 1, \ldots, 2k+1$, there is exactly one K_{i-1} common to G_l and G_{l-1} (taking indices mod $2k+1$). Denote the corresponding variable by x_l. Now fix l. Let the variables $\{y_k\}$ correspond to the K_{i-1} contained in only one G_l. Then the variables corresponding to $\mathbf{K}_{i-1}(G_l)$ are $\{x_l, x_{l+1}, y_1, \ldots, y_{i-2}\}$. We will show that $p_l(x,y) = i - 2 - \sum y_k - x_l x_{l+1}$ is $\lceil i/2 \rceil$-sos.

Let $J_1 = \{1, \ldots, \lceil i/2 \rceil - 2\}$ and $J_2 = \{\lceil i/2 \rceil - 1, \ldots, i-2\}$. Applying Lemma 3.3.1, we see that the following two functions are $\lceil i/2 \rceil$-sos. First apply the lemma with $A = \{x_l, x_{l+1}\}$ and $B = \{y_j : j \in J_1\} \cup \{x_l, x_{l+1}\}$:

$$f(x,y) = |J_1| - x_l x_{l+1} + x_l x_{l+1} y^{J_1} - \sum_{j \in J_1} y_j.$$

Second, take $A = \emptyset$ and $B = J_2$:

$$g(x, y) = |J_2| - 1 + y^{J_2} - \sum_{j \in J_2} y_j.$$

Finally, observe that the following is idempotent mod I:

$$h(x, y) = 1 - x_l x_{l+1} y^{J_1} - y^{J_2}.$$

Adding these up we get that $p_l(x, y) = f(x, y) + g(x, y) + h(x, y)$ is $\lceil i/2 \rceil$-sos. Now with $p(x, y) = \sum_{l=1}^{2k+1} p_l(x, y)$, we have that p is $\lceil i/2 \rceil$-sos:

$$p(x, y) = (2k + 1)(i - 2) - \sum_{l=1}^{2k+1} \sum_{y_k \subseteq G_l} y_k - \sum_{l=1}^{2k+1} x_l x_{l+1},$$

where the sum $\sum y_k$ is over all K_{i-1} contained in a unique K_i. It remains to show that $k - \sum x_l + \sum x_l x_{l+1}$ is $\lceil i/2 \rceil$-sos. Observe that this is attained by adding the following two quantities, each of which is a sum of idempotents mod I:

$$\sum_{l=1}^{k} \left(1 - x_{2l-1} - x_{2l} - x_{2l+1} + x_{2l-1}x_{2l} + x_{2l-1}x_{2l+1} + x_{2l}x_{2l+1}\right),$$

$$\sum_{l=2}^{k} (x_{2l-1} - x_{2l-1}x_1 - x_{2l-1}x_{2l+1} + x_{2l+1}x_1).$$

□

In view of Lemma 3.3.4, we see that the K_i-p-hole inequalities are valid on $\mathrm{TH}_{\lceil i/2 \rceil}(I)$. But since these inequalities define facets of $P_i(G) \subseteq \mathrm{TH}_{\lceil i/2 \rceil}(I)$, we see that they also define facets of $\mathrm{TH}_{\lceil i/2 \rceil}(I)$.

In Section 3.3 of [15], Conforti, Corneil, and Mahjoub show that polynomial time separation oracles exist for the following facets of $P_i(G)$.

1. Nonnegativities $0 \leq x_k \leq 1$,

2. K_i (clique) inequalities $\sum_{k \subset K_i} x_k \leq i - 1$,

3. Odd wheels of order $i-1$.

Define the polytope $Q(G)$ as the intersection of these facets, and define $Q'(G)$ by replacing (3) by

(3′) K_i-p-hole facets,

a superclass of (3).

The separation oracle provided by Conforti, Corneil, and Mahjoub [15] allows us to use the ellipsoid method to optimize over $Q(G)$ in polynomial time. However, it is stated as an open problem in [15] whether there is also such a polynomial time oracle for the class (3′), which would allow us to optimize in polynomial time over $Q'(G)$.

The results in this section show that $\mathrm{TH}_{\lceil i/2 \rceil}(I)$, over which we can optimize in polynomial time (for fixed i and to fixed arbitrary precision), is a tighter relaxation than $Q'(G)$. Precisely, we have the following inclusions:

$$P_i(G) \subseteq \mathrm{TH}_{\lceil i/2 \rceil}(I) \subseteq Q'(G) \subseteq Q(G).$$

A big advantage of this result is how easy it is to use in practice. The polynomial time results derived from the separation oracle rely on the ellipsoid method, which is numerically unstable and poor in practice even for small instances. By contrast, optimizing over the theta body is a standard semidefinite program; hence it can be done using interior point methods and can be straightforwardly implemented using any off-the-shelf solver. The original question in [15] is however still open, since we have not provided any separation oracle for the class (3′).

There are other families of facets of $P_i(G)$ for which efficient separation oracles are given in [15]. We have not treated them here, as they would not yield any new polynomial time results.

3.4 Related Problems

Here we apply two results appearing in the literature to the triangle free problem.

3.4.1 A lower bound on theta convergence

In Section 3, we showed that the earliest possible theta body, $\mathrm{TH}_{\lceil i/2 \rceil}(I_G)$, satisfies several inequalities defining facets of $P_i(G)$. However, in general it can take many steps of the theta hierarchy before a given facet of $P_i(G)$ is valid on $\mathrm{TH}_k(G)$. This is the case even for the triangle free problem. In particular, we will show:

Theorem 3.4.1. *For* $k < \frac{n-2}{4}$, $P_3(K_n) \subsetneq \mathrm{TH}_k(I_{K_n})$.

To prove Theorem 4.1, we will apply a result of Laurent [34] on the *cut polytope.* Let $G = (N, E)$ be a graph. A *cut* in G arises from a partition of the node set N into two sets S_1 and S_2, whereupon the associated cut is the set of edges from S_1 to S_2. Let $C_G \subseteq \{0, 1\}^E$ be the collection of characteristic vectors of cuts in G. Then $\mathrm{CUT}(G) = \mathrm{conv}(C_G)$ is the cut polytope of G. Similarly, define $T_G \subseteq \{0, 1\}^E$ to be the set of characteristic vectors of triangle free sets in G, and as before, $P_3(G) = \mathrm{conv}(T_G)$. Note that a cut is by definition a bipartite graph; hence, it is triangle free. Therefore $C_G \subseteq T_G$ and $\mathrm{CUT}(G) \subseteq P_3(G)$. The theta body approach has also been applied to the cut polytope by Gouveia, Laurent, Parrilo, and Thomas [20]. The following lemma shows that inclusion among varieties extends to inclusion of theta bodies.

Lemma 3.4.2. *Let* $X \subseteq Y$ *be two real varieties, with ideals* $I(X)$ *and* $I(Y)$. *Then for any* k, $\mathrm{TH}_k(I(X)) \subseteq \mathrm{TH}_k(I(Y))$.

Proof. If $X \subseteq Y$, then the reverse inclusion holds for their ideals: $I(Y) \subseteq I(X)$. Any function which is k-sos mod $I(Y)$ is then also k-sos mod $I(X)$. The result follows from the definition of $\mathrm{TH}_k(I)$. □

In particular, since $C_G \subseteq T_G$, we have $\mathrm{TH}_k(I(C_G)) \subseteq \mathrm{TH}_k(I(T_G))$ for all k. Note that $I(T_G) = I_G$ in our notation from the K_i-free problem.

For the complete graph K_n, when n is odd, the inequality

$$\sum_{e \in E} x_e \leq \frac{n^2 - 1}{4} \tag{3.4.1}$$

defines a facet of both $P_3(K_n)$ and $\mathrm{CUT}(K_n)$.

Theorem 3.4.3. *[34, Theorem 6] For $k < \frac{n-2}{4}$, $\mathrm{CUT}(K_n) \subsetneq \mathrm{TH}_k(I(C_{K_n}))$. In particular, equation (1) does not hold on $\mathrm{TH}_k(I(C_{K_n}))$.*

In [34] this result appears in terms of a different but related relaxation. A translation of that result to the theorem above can be found in Example 3.9 of [20]. We can now prove Theorem 4.1.

Proof. By Theorem 4.3, there is a point $x \in \mathrm{TH}_k(I(C_{K_n}))$ violating (1). But by Lemma 4.2, $x \in \mathrm{TH}_k(I(T_{K_n}))$. Since (1) is valid on $P_3(K_n)$, $x \notin P_3(K_n)$. $\square$

This implies that the theta body hierarchy does not polynomially capture the K_n inequalities, as the size of the reduced moment matrices associated with the $\lceil \frac{n-2}{4} \rceil$-th theta body is exponential in n. It is still an open question, for both $\mathrm{CUT}(K_n)$ and $P_3(K_n)$, what is the smallest k so that the k-th theta body is exact.

Recall that for $i = 2$, the K_i-free problem is simply the stable set problem. In that case it is well known that the clique inequalities are valid for the first theta body relaxation. A simple byproduct of Theorem 3.4.1 is that this fact fails to generalize even to $i = 3$, as it is impossible to capture all the clique inequalities with a constant rank of theta bodies in this case.

3.4.2 An integrality gap for triangle cover

Let G be a graph. A *triangle cover* is a collection of edges in G, containing at least one edge from every triangle in G. Let $\tau(G)$ be the minimum-size triangle cover in G (in the language of the introduction, the K_3-cover problem with unit weights). Let I be the ideal of the triangle cover problem. Define the following semidefinite relaxation:

$$\tau^{\dagger}(G) = \min \left\{ \sum_{e \in E} x_e : x \in \mathrm{TH}_2(I) \right\}.$$

Note that since C is a triangle cover if and only if $E \setminus C$ is a triangle free set, we can restate any statements about theta bodies for the triangle free problem using the change of variables $x \mapsto 1 - x$.

We can also define a natural LP relaxation for the triangle cover problem. Let

$$\tau^*(G) = \min \left\{ \sum_{e \in E} x_e : x \in [0,1]^E \text{ and for all triangles } \Delta, \sum_{e \in \Delta} x_e \geq 1 \right\}.$$

Krivelevich [32] proved that $\tau(G) \leq 2\tau^*(G)$. We can apply this to prove an integrality gap for $\tau^\dagger(G)$.

Theorem 3.4.4. *For any graph G, $\tau^\dagger(G) \geq \frac{\tau(G)}{2}$.*

Proof. By Corollary 3.2, $\tau^\dagger(G) \geq \tau^*(G)$, as the inequalities defining $\tau^*(G)$ are valid on the second theta body. Combining this with Krivelevich's inequality gives the result. ☐

Another way to interpret Krivelevich's result is in terms of a conjecture of Tuza. Define a *triangle packing* in a graph G to be a collection of triangles in G, no two of which share an edge. Let $v(G)$ be the maximum-size triangle packing in G. It is an easy exercise to check that $v(G) \leq \tau(G) \leq 3v(G)$. However, Tuza conjectured in [46] that the stronger inequality $\tau(G) \leq 2v(G)$ holds for all graphs G. The problem is currently open; see [26] for more information. $v(G)$ also has a natural LP relaxation.

$$v^*(G) = \max \left\{ \sum_{\Delta \in T} y_\Delta : y \in [0,1]^T \text{ and for all edges } e, \sum_{e \in \Delta} y_\Delta \leq 1 \right\}$$

By LP duality, $\tau^*(G) = v^*(G)$. Krivelevich [32] also proved that $v^*(G) \leq 2v(G)$. After applying the duality $\tau^*(G) = v^*(G)$, these become fractional versions of Tuza's conjecture: $\tau(G) \leq 2v^*(G)$ and $\tau^*(G) \leq 2v(G)$. A natural question to ask is whether, given the SDP relaxation $\tau^\dagger(G)$, whether the "semidefinite version" of Tuza's conjecture would hold: $\tau^\dagger(G) \leq 2v(G)$.

Chapter 4

SUMS OF SQUARES ON THE UNIT HYPERCUBE

4.1 Introduction

Our main results in this paper are upper and lower bounds on the degrees of sum of squares multipliers needed to write certain polynomials as sums of squares mod I, the ideal of the hypercube $C = \{0,1\}^n$. That is, given a polynomial f on C, we ask whether we can write $fg = h$, where both g and h are sums of squares of polynomials: $g = \sum_i g_i^2$ and $h = \sum_j h_j^2$ mod I. In Section 4.3 we prove a lower bound on these degrees for invariant polynomials f vanishing to odd degree on a slice of the hypercube. Correspondingly, in Section 4.4 we prove an upper bound on these degrees. In the case of quadratics on the hypercube these bounds are tight.

This work is motivated by the recent study of sum of squares relaxations and their application to combinatorial optimization. Testing whether $f - c$ is nonnegative can be repeated to find $\min(f)$. In practice, replacing the nonnegativity of $f - c$ with a sum of squares condition has been an effective relaxation in some problems [21] [40]. We can fix k and ask whether a nonnegative polynomial f is a sum of squares of degree at most k. For given f, k, this can be solved via semidefinite programming, leading to a hierarchy of approximations called *theta bodies* [21]. Proofs about theta bodies then involve proving that an f is, or is not, a sum of squares of degree k mod I, and our results give exactly this information. As a corollary to our lower bound, we reprove a result of Laurent [34] that the Lasserre rank of the cut polytope is at least $n/2$. Our upper bound in fact answers a conjecture of Laurent that this rank is exactly $n/2$.

A second motivation is Hilbert's 17th problem: whether every nonnegative polynomial on $\mathbb{R}^n$ can be written as a sum of squares of rational functions. Artin proved that this is the case, and there has been some work on considering the degrees of the functions involved, but no general lower bounds exist. As a corollary to our lower bound, in Section 4.5 we provide the first lower bounds on such degrees.

The layout of the paper is as follows: in section 4.2, we give the algebraic background necessary for the paper. In section 4.3, we prove the lower bound for sos multipliers on the hypercube C. In section 4.4, we prove the upper bound for multipliers on C. In section 4.5, we discuss applications to Hilbert's 17th problem, and to the max cut problem.

4.2 Algebraic background

Let $X \subseteq \mathbb{R}^n$ be an algebraic set and $I \subseteq \mathbb{R}[x] := \mathbb{R}[x_1, \ldots, x_n]$ its vanishing ideal. That is, $I = \{f \in \mathbb{R}[x] : f(x) = 0 \text{ on } X\}$. This lets us define an equivalence relation $f \equiv g \mod I$ if and only if $f - g \in I$; i.e. if f and g agree on X. The equivalence classes are the cosets $f + I$, and the space of equivalence classes is denoted $\mathbb{R}[X] := \mathbb{R}[x]/I$, which we can think of as polynomials on X. In the ring $\mathbb{R}[X]$, equality is defined by equivalence mod I, so for most of the paper we will write $f = g$ for $f \equiv g \mod I$. For $f \in \mathbb{R}[X]$, let $\deg(f) = \min\{\deg(g) : g \in f\}$. Our main example is the hypercube $C = \{0, 1\}^n$ with ideal $I = \langle x_i^2 - x_i : 1 \leq i \leq n\rangle$, so each $f \in \mathbb{R}[C]$ has a natural representative of smallest degree; namely, the representative using only squarefree monomials.

For $f \in \mathbb{R}[X]$ and $d \geq 0$, we say f is *d-sos mod I*, or simply d-sos if I is clear from context, if $f = \sum_{i=1}^k g_i^2$ for some $g_i \in \mathbb{R}[X]$ with $\deg(g_i) \leq d$. Note that this implies $f \geq 0$ on X.

We will also need to discuss the notion of divisibility mod I. It may happen that, due to cancellation mod I, the ring-theoretic definition of divisibility in $\mathbb{R}[X]$ doesn't correspond well to the usual notion of divisibility of polynomials. For instance, we may have $f = gh$ but $\deg(f) < \deg(g) + \deg(h)$; in the case of the hypercube, $x \cdot x = x$. To fix this, for $f, g \in \mathbb{R}[X]$, we say that *g properly divides f mod I* if there is $h \in \mathbb{R}[X]$ such that $f = gh$ and $\deg(f) = \deg(g) + \deg(h)$. Again, if I is clear we can drop "mod I." We will also say that *g properly divides f to order m* if g^m properly divides f, but g^{m+1} does not.

Recall our main example $C = \{0,1\}^n$ and $I = \langle x_i^2 - x_i,\ i = 1, \dots, n\rangle$. It will simplify the notation to use subsets of $[n]$ as exponents: $x^{\{1,4\}} = x_1 x_4$. Then the space of functions $\mathbb{R}[C]$ has a basis $\{x^m : m \subseteq [n]\}$ of squarefree monomials. Thus we can write any function $f \in \mathbb{R}[C]$ as $f = \sum_{m \subset [n]} c_m x^m$, and we have $\deg(f) = \max\{|m| : c_m \neq 0\}$. We define $\mathbb{R}[C]_d$ to be the collection of homogeneous degree-d functions, and $\mathbb{R}[C]_{\leq d} = \oplus_{i=0}^{d} \mathbb{R}[C]_i$ the collection of functions of degree at most d.

4.3 Lower Bound on Multipliers

In this section we prove our main results, which deal with S_n-invariant polynomials which vanish on a *level* $T = \{x \in C : \sum x_i = t\}$ of the hypercube. Such functions come up naturally in combinatorial optimization, where we are counting objects subject to some symmetric restrictions; see Section 4.5.1. We will show that such functions cannot be sums of squares of low degree. For the rest of this section, we abbreviate

$$l = t - \sum x_i,$$

the parameter t will be fixed.

To start, we decompose $\mathbb{R}[C]$ into *irreducible representations* of the symmetric group S_n. For background on representation theory in general and on the representation theory of S_n in particular, see the introduction by Sagan [41], whose notation we adopt here. Note that we treat $(n, 0)$ as an alias for the partition (n) to simplify our notation. We let S_n act on $\mathbb{R}[C]$ by permuting the variables directly: $(123)x_1 = x_2$.

We will define an isomorphism between tabloids and monomials. Suppose $k \leq n/2$. Define $\phi_k : M^{(n-k,k)} \to \mathbb{R}[C]$ by $\phi_k([m^c, m]) = x^m$, and extend by linearity. For example, $\phi_3([12345, 678]) = x_6 x_7 x_8$. The image of ϕ_k is exactly the subspace $\mathbb{R}[C]_k$ of homogeneous degree-k functions. We also have $\mathbb{R}[C]_k \cong \mathbb{R}[C]_{n-k}$ as S_n-modules, since we can take complements in the exponent: if $n = 6$, then $x_1 x_2 \in \mathbb{R}[C]_2 \leftrightarrow x_3 x_4 x_5 x_6 \in \mathbb{R}[C]_4$.

Proposition 4.3.1. *The S_n-module $\mathbb{R}[C]$ decomposes into $n + 1 - 2k$ copies of $S^{(n-k,k)}$, for $0 \leq k \leq \frac{n}{2}$.*

Proof. By Young's rule (Theorem 2.11.2 in [41]), $M^{(n-k,k)}$ contributes one copy of $S^{(n-i,i)}$ for each $0 \leq i \leq k$. By the above, if $k \leq n/2$, $\mathbb{R}[C]_{n-k} \cong \mathbb{R}[C]_k \cong M^{(n-k,k)}$. If n is odd, then

$$\begin{aligned}
\mathbb{R}[C] &= \bigoplus_{0 \leq k < n/2} (\mathbb{R}[C]_k \oplus \mathbb{R}[C]_{n-k}) \\
&\cong 2 \bigoplus_{0 \leq k < n/2} M^{(n-k,k)} \\
&\cong 2 \bigoplus_{0 \leq k < n/2} \left(\bigoplus_{i=0}^{k} S^{(n-i,i)} \right) \\
&\cong 2 \bigoplus_{i=0}^{\lfloor n/2 \rfloor} \left(\frac{n-1}{2} - i + 1 \right) S^{(n-i,i)},
\end{aligned}$$

which gives the result when n is odd. For even n just add the single copy of $\mathbb{R}[C]_{n/2} \cong M^{(n/2,n/2)}$. $\square$

Proposition 4.3.1 gave the decomposition of $\mathbb{R}[C]$ into irreducibles, up to isomorphism. But to analyze a specific $f \in \mathbb{R}[C]$ in terms of this decomposition, we need to know how the irreducible submodules lie inside $\mathbb{R}[C]$; that is, we need to know an isomorphism explicitly. We will choose a slightly idiosyncratic description which will be useful later. Fix $t \in \mathbb{R}$. Recalling that $S^{(n-k,k)} \subset M^{(n-k,k)}$, define $H_{k0} = \phi(S^{(n-k,k)}) \subseteq \mathbb{R}[C]_k$. Since ϕ is an S_n-module isomorphism, we have $H_{k0} \cong S^{(n-k,k)}$. Then for $i = 1, \ldots, n-2k$, define $H_{ki} = (t - \sum_j x_j)^i \cdot H_{k0}$. Note that no nonzero element of H_{k0} is properly divisible by ℓ.

Theorem 4.3.2. *$\mathbb{R}[C]$ has the following decomposition into irreducibles:*

$$\mathbb{R}[C] = \bigoplus_{k=0}^{\lfloor n/2 \rfloor} \left(\bigoplus_{i=0}^{n+1-2k} H_{ki} \right).$$

This decomposition respects degree: for any d,

$$\mathbb{R}[C]_{\leq d} = \bigoplus_{k+i \leq d} H_{ki}.$$

Proof. By Proposition 4.3.1, the proposed decomposition contains the correct number of each irreducible. Therefore, we need to show that the summands are linearly independent.

By Corollary 2.11 in [45], the map $U : \mathbb{R}[C]_k \to \mathbb{R}[C]_{\leq n-k}$ given by $U(f) = (t - \sum x_j)^{n-2k} f$ is a bijection. Therefore, the map $U' : \mathbb{R}[C]_k \to \mathbb{R}[C]_{\leq k+i}$ given by $f \mapsto (t - \sum x_j)^i f$ is injective for $i \leq n-2k$, since U' is a precomposition of U. Since $H_{ki} = U'(H_{k0})$, we have that $\deg(f) = k+i$ for each nonzero $f \in H_{ki}$; in particular, $H_{ki} \neq 0$. Since S_n acts trivially on $(t - \sum_j x_j)^i$, $H_{ki} \cong H_{k0}$. By irreducibility, we know that H_{ki} and $H_{k'i'}$ are linearly independent if $k \neq k'$. It remains to consider H_{ki} for varying i; but since each nonzero $f_i \in H_{ki}$ has degree exactly $k+i$, these are linearly independent as well.

The expression for $\mathbb{R}[C]_{\leq d}$ now follows from the linear independence of the H_{ki}. □

We now show that proper divisibility holds for functions of low degree vanishing on a level T.

Lemma 4.3.3. *Let $T = \{x \in C : \sum_i x_i = t\}$, for fixed $t \in \{0, \ldots, n\}$. Suppose $f \in \mathbb{R}[C]_{\leq d}$, and f vanishes on T. If $d \leq t \leq n-d$, then f is properly divisible by l.*

Proof. Let V be the S_n-submodule of $\mathbb{R}[C]_{\leq d}$ consisting of polynomials that are properly divisible by l and let

$$W = H_{00} \oplus \ldots \oplus H_{d0} \cong S^{(n)} \oplus \cdots \oplus S^{(n-d,d)}.$$

By Theorem 4.3.2 we have $\mathbb{R}[C]_{\leq d} = V \oplus W$. Let $U \subset W$ be the S_n-submodule of polynomials vanishing on T. It suffices to show that $U = 0$. Since the H_{i0} are nonisomorphic irreducible modules, it follows that

$$U = \bigoplus_{i \in I} H_{i0},$$

where I is a subset of $\{0, \ldots, d\}$. Now we claim that polynomials in H_{i0} do not identically vanish on T for all $0 \leq i \leq d$. Since H_{i0} is an irreducible S_n-module it suffices to exhibit a single polynomial $p \in H_{i0}$ not vanishing on T.

To see this, let q be the standard tableau of shape $(n-i, i)$ where the first row contains $\{1, \dots, n-i\}$ and the second row contains $\{n-i+1, \dots, n\}$. Let $\hat{x} \in C$ be given by

$$\hat{x} = e_{n-t+1} + \cdots + e_n.$$

Since $i \leq t \leq n-i$, the support of $\hat{x}$ contains the second row of q and does not contain any of the first i entries of the first row of q. Consider $p = \phi(e_q)$, $p \in H_{i0}$. It follows that $p(\hat{x}) = 1$, since only the monomial $\phi(q)$ is nonzero on $\hat{x}$ in $\phi(e_q)$ and $\phi(q)(\hat{x}) = 1$. See Figure 4.3 for an example. $\square$

$$q = \begin{array}{|c|c|c|c|c|c|c|} \hline 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ \hline 8 & 9 \\ \cline{1-2} \end{array}$$

$$\hat{x} = (0, 0, 0, 0, 0, 0, 1, 1, 1)$$

$$p = \phi(e_q) = x_8x_9 - x_1x_9 - x_8x_2 + x_1x_2$$

Figure 4.1: A standard tableau q with sorted rows, and an associated $\hat{x}$. Here $n = 9, i = 2, t = 3$. We have $p(\hat{x}) = 1$.

Now we can prove our main result.

Theorem 4.3.4. *Suppose $f \in \mathbb{R}[C]_{\leq t}$ with $t \leq n/2$ is an S_n-invariant polynomial, and f is properly divisible by $l = t - (x_1 + \cdots + x_n)$ to odd order. Then f is not (d_1, d_2)-sos for $d_1 \leq \min\left\{\frac{n - \deg f}{2}, t\right\}$, $d_2 \leq t$.*

Proof. Suppose that $f \sum g_i^2 = \sum h_j^2$ with $g_i \in \mathbb{R}[C]_{\leq d_1}$, $g_i \neq 0$ and $h_j \in \mathbb{R}[C]_{\leq d_2}$. Let $g = \sum g_i^2$ and $h = \sum h_j^2$. Without loss of generality we may assume that g and h are S_n-invariant polynomials, otherwise we may replace them by their S_n symmetrizations.

Since $d_2 \leq t$ by Lemma 4.3.3 we can write $h_j = l^{a_j} q_j$ with S_n-invariant polynomials q_j such that $\deg q_j = \deg h_j - \deg a_j$ and q_j is non-zero on any point of T. Therefore $h = l^{2a} q$

where $a = \min a_j$ and q is an S_n invariant polynomial, $\deg q = \deg h - 2a$, and q is strictly positive on T.

Similarly, since $d_1 \leq t$ we argue that $g = l^{2b} r$, where r is S_n-invariant polynomial positive on T, and $\deg r = \deg g - 2b$. Finally, $f = l^c p$ where c is odd and p is an S_n-invariant polynomial not identically 0 on T with $\deg p = \deg f - c$. Combining, we see that

$$l^{2b+c} pr - l^{2a} q = 0.$$

Let $\alpha = \min\{2a, 2b+c\}$. By factoring out l^α in the equation above we obtain

$$l^\alpha s = 0,$$

for an S_n-invariant polynomial $s \in \mathbb{R}[C]$ of degree strictly less than n since $d_1 \leq \min\left\{\frac{n-t}{2}, t\right\}$ and $d_2 \leq t$. Since q and r are strictly positive on T and p is not identically zero on T, it follows that s does not vanish on T. Thus s is a non-zero symmetric polynomial in $\mathbb{R}[C]$ vanishing on $C \setminus T$. Therefore $s = \beta \chi_T$ for some constant $\beta \neq 0$, where $\chi_T \in \mathbb{R}[C]$ is the polynomial vanishing on $C \setminus T$ and equal to 1 on T. However, it is not hard to check that $\deg \chi_T = n$ for any level T and therefore we arrive at a contradiction.

□

Corollary 4.3.5. *Fix $t \leq n/2$ and let $f \in \mathbb{R}[C]_{\leq t}$ be fixed by S_n. Suppose that f is properly divisible by $l = t - (x_1 + \cdots + x_n)$ to odd order. Then f is not d-sos for $d \leq t$.*

Proof. Apply Theorem 4.3.4 with $d_1 = 0$. □

Corollary 4.3.6. *Let $k = \lfloor \frac{n}{2} \rfloor$ and let $f \in \mathbb{R}[C]$ be given by*

$$f = (x_1 + \cdots + x_n - k)(x_1 + \cdots + x_n - k - 1).$$

Then f is nonnegative on C but f is not $(k-1, k)$-sos.

Proof. Apply Theorem 4.3.4. □

4.4 Upper Bound on Multipliers

Let $\Gamma = \{v_1, \ldots, v_m\}$ be a finite set of points in $\mathbb{R}^n$. We consider sums of squares in $\mathbb{R}[\Gamma]$.

Lemma 4.4.1. *Fix $d_1, d_2 \in \mathbb{N}$. The set $A \subseteq \mathbb{R}[\Gamma]_{\leq d}$ of polynomials which are (d_1, d_2)-sos is closed in $\mathbb{R}[\Gamma]_{\leq d}$ for all d.*

Proof. One can check that $\Sigma(\Gamma)_{\leq 2d}$ is a closed pointed cone. Suppose that $f_i \in \mathbb{R}[\Gamma]_{\leq d}$ are (d_1, d_2)-sos and $f_i \to f$. Then there exist g_i, h_i which are respectively d_1 and d_2-sos and $f_i g_i = h_i$. We may rescale g_i and assume that

$$\frac{1}{|\Gamma|} \sum_{x \in \Gamma} g_i(x) = 1.$$

The set of d_1-sos polynomials with average 1 on Γ is compact. Therefore a subsequence of $\{g_i\}$ converges to g, which is also d_1-sos. Then $f_i g_i$ converge to fg and since each $f_i g_i$ is d_2-sos it follows that fg is d_2-sos.

□

Let $\ell : R[\Gamma]_{\leq 2d} \to \mathbb{R}$ be a linear functional given as a combination of evaluations on Γ:

$$\ell(f) = \sum_{i=1}^{m} \mu_i f(v_i), \quad f \in R[\Gamma]_{2d},\ \mu_i \in \mathbb{R}.$$

Let $Q_\ell : R[\Gamma]_{\leq d} \to \mathbb{R}$ be the quadratic form associated to ℓ given by

$$Q_\ell(f) = \ell(f^2) = \sum_{i=1}^{m} \mu_i f^2(v_i).$$

We assume that the coefficients μ_i are non-zero and let m_+ and m_- be the number of positive and negative μ_i respectively.

Lemma 4.4.2. *Suppose that the quadratic form Q_ℓ is positive semidefinite. Then $m_+ \geq \dim R[\Gamma]_{\leq d}$.*

Proof. Let $\pi_\Gamma : R[\Gamma]_{\leq d} \to \mathbb{R}^m$ be the evaluation projection of forms in $R[\Gamma]_{\leq d}$ given by

$$\pi_\Gamma(f) = (f(v_1), \ldots, f(v_m)), \quad f \in R[\Gamma]_{\leq d}.$$

We observe that the map π_Γ has no kernel and therefore

$$\dim \pi_\Gamma(R[\Gamma]_{\leq d}) = \dim R[\Gamma]_{\leq d}.$$

Let $\bar{Q}_\ell$ be the quadratic form on $\mathbb{R}^m$ given by:

$$\sum_{i=1}^{m} \mu_i x_i^2.$$

By its definition, the form Q_ℓ is a composition of π_Γ and $\bar{Q}_\ell$:

$$Q_\ell = \bar{Q}_\ell \circ \pi_\Gamma.$$

The form $\bar{Q}_\ell$ has m_- negative eigenvalues, and thus $\bar{Q}_\ell$ is strictly negative on a subspace of dimension m_-. Recall that the form Q_ℓ is positive semidefinite, which implies that $\bar{Q}_\ell$ is positive semidefinite on the image of π_Γ. Thus the image of π_Γ has codimension at least m_- in $\mathbb{R}^m$. Since $m_+ + m_- = m$ the Lemma follows. □

Let $H_\Gamma(t)$ denote the Hilbert function of Γ:

$$H_\Gamma(t) = \dim \mathbb{R}[\Gamma]_{\leq t}.$$

Theorem 4.4.3. *Let $p \in \mathbb{R}[\Gamma]_{\leq 2s}$ be a polynomial of degree at most $2s$ nonnegative on Γ. Suppose that for some $k \in \mathbb{R}$ we have*

$$H_\Gamma(k+s) + H_\Gamma(k) > H_\Gamma(2k+2s).$$

Then p is $(k, k+s)$-sos on Γ, i.e. there exists $q \in \Sigma(\Gamma)_{\leq 2k}$ such that $pq \in \Sigma(\Gamma)_{\leq 2s+2k}$.

Proof. Suppose not. By Lemma 4.4.1, the set of all polynomials in $\mathbb{R}[\Gamma]_{\leq 2s}$ that is not $(k, k+s)$-sos is open. Thus we can find $p \in \mathbb{R}[\Gamma]_{\leq 2s}$ that is strictly positive on Γ but is not $(k, k+s)$-sos. Now consider the pointed, closed convex cones $p\Sigma(\Gamma)_{\leq 2k}$ and $\Sigma(\Gamma)_{\leq 2k+2s}$ in $\mathbb{R}(\Gamma)_{\leq 2k+2s}$. By our assumption

$$p\Sigma(\Gamma)_{\leq 2k} \cap \Sigma(\Gamma)_{\leq 2k+2s} = \{0\}.$$

Therefore there exists a linear functional $\ell : \mathbb{R}[\Gamma]_{\leq 2k+2s} \to \mathbb{R}$ strictly separating the two cones: $\ell(f) > 0$ for all nonzero $f \in \Sigma(\Gamma)_{\leq 2k+2s}$ and $\ell(f) < 0$ for all nonzero $f \in p\Sigma(\Gamma)_{\leq 2k}$.

Let $\Gamma' \subseteq \Gamma$ be a subset of Γ such that point evaluations on Γ' form a basis of $\mathbb{R}[\Gamma]^*_{\leq 2k+2s}$. In particular, $|\Gamma'| = \dim \mathbb{R}(\Gamma)_{\leq 2k+2s}$. Therefore the separating functional ℓ can be written as

$$\ell = \sum_{v_i \in \Gamma'} \mu_i \ell_{v_i}, \quad \mu_i \in \mathbb{R}.$$

Since ℓ strictly separates the two cones we may assume without loss of generality that all coefficients μ_i are non-zero. Let m_+ and m_- be the number of positive and negative μ_i respectively. Then by Lemma 4.4.2 we know that $m_+ \geq \dim \mathbb{R}[\Gamma]_{\leq k+s}$.

Now define $\ell' : \mathbb{R}[\Gamma]_{\leq 2k} \to \mathbb{R}$ by

$$\ell' = \sum_{v_i \in \Gamma'} \mu_i p(v_i) \ell_{v_i}.$$

It follows that $Q_{\ell'} : \mathbb{R}[\Gamma]_{\leq k} \to \mathbb{R}$ is a negative definite quadratic form. Therefore, by applying Lemma 4.4.2, we see that $m_- \geq \dim \mathbb{R}[\Gamma]_{\leq k}$, since $p(v_i) > 0$ for all $v_i \in \Gamma$. Combining, we see that

$$|\Gamma'| = m_+ + m_- = H_\Gamma(2k+2s) \geq H_\Gamma(k+s) + H_\Gamma(k),$$

which is a contradiction. □

Corollary 4.4.4. *Let $p \in \mathbb{R}[C]_{\leq 2}$ be a quadratic polynomial nonnegative on C and let $k = \lfloor \frac{n}{2} \rfloor$. Then p is $(k, k+1)$-sos.*

Proof. This follows immediately from Theorem 4.4.3 since $H_C(t) = \sum_{i=0}^{t} \binom{n}{i}$. □

This gives insight into a conjecture of Laurent [34] that the Lasserre rank of the max cut polytope on K_n is exactly $n/2$. It was already known to be at least $n/2$; see for instance Section 4.5.1 of this paper. By Corollary 4.4.4, rank $n/2$ is required when using multipliers. It remains to check whether allowing multipliers reduces the degree required in this case.

4.5 *Applications*

We give two applications of our results. Section 4.5.1 deals with the max cut problem on K_n, and is an application to combinatorial optimization. Section 4.5.2 deals with degree bounds in Hilbert's 17th problem.

4.5.1 The max cut problem

A *cut* in a graph arises from a partition of the vertices into two sets S_1, S_2; then the cut is the collection of all edges from S_1 to S_2. Note that switching S_1 and S_2 gives the same cut. We write $C = [S_1, S_2] = [S_2, S_1]$, and let $|S| =$ the number of edges from S_1 to S_2. A *max cut* is a cut maximizing $|S|$.

In the complete graph K_n, the max cuts come from any partition of $[n]$ into two sets of $n/2$ vertices, or $(n \pm 1)/2$ when n is odd. Given a cut $S = [S_1, S_2]$ in K_n, let $x = x(S) \in \{0,1\}^n$ be defined by $x_i = 1$ if $i \in S_1$; $x_i = 0$ if $i \in S_2$. Observe that for $i \neq j$, the quantity

$$(2x_i - 1)(2x_j - 1) = \begin{cases} 1 & \text{if } i, j \text{ are in the same half of } S, \\ -1 & \text{if } i, j \text{ are in different halves of } S. \end{cases}$$

We will calculate the function on $\{0,1\}^n$ that counts the number of edges in a cut. If x represents a cut with $|S|$ edges, then

$$\sum_{i<j} (2x_i - 1)(2x_j - 1) = \left(\binom{n}{2} - |S| \right) \cdot 1 + (|S|) \cdot -1 = \binom{n}{2} - 2|S|.$$

Therefore, $|S| = \frac{n^2-n}{4} - \frac{1}{2} \sum_{i<j} (2x_i - 1)(2x_j - 1)$. For n odd, the maximum size of a cut is $\frac{n-1}{2} \frac{n+1}{2} = \frac{n^2-1}{4}$. Put

$$q = \frac{n^2-1}{4} - \left(\frac{n^2-n}{4} - \frac{1}{2} \sum_{i<j} (2x_i - 1)(2x_j - 1) \right).$$

Then we have the following factorization and inequality.

Proposition 4.5.1. *Let n be odd and q as above. q factors mod I as $(\frac{n-1}{2} - \sum x_i)(\frac{n+1}{2} - \sum x_i)$. For $x \in \{0,1\}^n$, $q(x) \geq 0$, and $q(x) = 0$ only for the vectors representing max cuts in K_n.*

Proof. The factorization is routine algebra; recall that the ideal I of the cube allows simplification $x_i^2 = x_i$. The nonnegativity and equality statements follow from the preceding discussion. ☐

Note that the q defined above satisfies the conditions of Theorem 4.3.5, for $t = \frac{n-1}{2}$. Therefore, it is not d-sos mod I for $d \leq t$. This will allow us to reprove a result of Laurent. In [34], Theorem 4, it is shown that the Lasserre rank of the cut polytope of K_n, n odd, is at least $\frac{n+1}{2}$. The Lasserre relaxation is a hierarchy of semidefinite program approximations to a 0/1 polytope based on sums of squares, and we can reprove Laurent's result by showing that q is not d-sos for $d < \frac{n+1}{2}$.

Corollary 4.5.2. *Let n be odd and I the ideal of the hypercube C. Then q, as above, is not a sum of squares mod I of degree $\leq \frac{n-1}{2}$.*

Proof. This is just 4.3.6. ☐

4.5.2 Globally nonnegative function with large multipliers

We finish with an application to Hilbert's 17th problem about sums of rational squares on $\mathbb{R}^n$.

Theorem 4.5.3. *Let $k = \lfloor \frac{n}{2} \rfloor$. There exists a polynomial p of degree 4 nonnegative on $\mathbb{R}^n$ which is not $(k-2, k)$-sos.*

Proof. Let $k = \lfloor \frac{n}{2} \rfloor$ and let $f \in \mathbb{R}[x]$ be given by

$$f = (x_1 + \cdots + x_n - k)(x_1 + \cdots + x_n - k - 1).$$

By Corollary 4.3.6 we know that f is not $(k-1, k)$-sos in $\mathbb{R}[C]$. Using Lemma 4.4.1 with $\Gamma = C$ it follows that $f + \epsilon$ is not $(k-1, k)$-sos in $\mathbb{R}[C]$ for sufficiently small $\epsilon > 0$. Let $f' = f + \epsilon$ for a fixed such ϵ.

Let $r = \sum_{i=1}^{n} (x_i^2 - x_i)^2$. For sufficiently large $\lambda > 0$ the polynomial $q = f' + \lambda r$ is positive on $\mathbb{R}^n$. Suppose that q is $(k-2, k)$-sos in $\mathbb{R}[x]$: we have $qg = h$ with $(k-2)$-sos non-zero g, and k-sos h.

For $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{R}^n$ let C_α be the hypercube given by equations $(x_i - \alpha_i)(x_i - \alpha_i - 1) = 0$. By Lemma 4.4.1 it follows that q is not $(k-2, k)$-sos in $\mathbb{R}[C_\alpha]$ for all α sufficiently close to 0. However, there exist α arbitrarily close to 0 such that $g \not\equiv 0$ in $\mathbb{R}[C_\alpha]$. This is a contradiction since it follows that q is $(k-2, k)$-sos in $\mathbb{R}[C_\alpha]$ for such α. $\square$

Chapter 5

THE REPRESENTATION THEORY OF MATCHINGS

5.1 Introduction

Let $B_n \subseteq S_{2n}$ be the *hyperoctahedral group*; that is, the stabilizer of $\sigma = (12)(34)\dots(2n-1, 2n)$ in S_{2n}. Barbasch and Vogan [8] showed that the induced representation $\mathrm{Ind}_{B_n}^{S_{2n}}(1)$ decomposes into a sum of Specht modules $\oplus_\lambda S^\lambda$, one for each $\lambda \vdash 2n$ such that each λ_i is even. We define two subgroups $C_{m,n}$ and $D_{m,n}$ of S_{mn}, each of which is a natural generalization of B_n. Let $C_{m,n} = S_n \text{ wr } S_m$ and $D_{m,n} = S_n \text{ wr } C_m$. Here $C_m \subseteq S_m$ is the cyclic group generated by an m-cycle. When $m = 2$, $C_{m,n} = D_{m,n} = B_n$.

These generalizations arise naturally when using symmetry to reduce the dimension of semidefinite programs in combinatorial optimization. The S_{2n}-module $\mathrm{Ind}_{B_n}^{S_{2n}}(1)$ is naturally isomorphic to the vector space of perfect matchings on K_{2n}. Decomposing this vector space into irreducible representations corresponds to a block diagonalization of the semidefinite program underlying the *theta body* for these matchings, an approximation based on sums of squares.

Similarly, the S_{mn}-module $\mathrm{Ind}_{C_{m,n}}^{S_{mn}}(1)$ is naturally isomorphic to the vector space of perfect m-uniform hypermatchings on the m-uniform complete hypergraph $K_{mn}^{(m)}$. Likewise, $\mathrm{Ind}_{D_{m,n}}^{S_{mn}}(1)$ is naturally isomorphic to the vector space of decompositions of the vertex set $[mn]$ of the complete graph K_{mn} into n disjoint m-cycles. Decomposing these into irreducible representations would allow symmetry reduction of the corresponding combinatorial optimization problems.

We generalize Barbasch and Vogan's proof to recursively describe the decomposition of both $\mathrm{Ind}_{C_{m,n}}^{S_{mn}}(1)$ and $\mathrm{Ind}_{D_{m,n}}^{S_{mn}}(1)$. We do not believe that a simple pattern for the decomposition exists for $m > 2$ in either case. However, we are able to establish enough of the structure of $\mathrm{Ind}_{C_{3,n}}^{S_{3n}}(1)$ to show that, unlike the case $m = 2$, the irreducible representations are not multiplicity-free for $n \geq 5$.

The structure of this paper is as follows. In Section 2, we give a method for determining $\mathrm{Ind}_{C_{m,n}}^{S_{mn}}(1)$ from $\mathrm{Ind}_{C_{m,n-1}}^{S_{m(n-1)}}(1)$. In Section 3, we produce an explicit linear isomorphism corresponding to the $m = 2$ case. In Section 4, we prove that $\mathrm{Ind}_{C_{m,n}}^{S_{mn}}(1)$ is not multiplicity-free for $n \geq 5$.

5.2 Recursive construction

We generalize the induction step in the proof of Barbasch and Vogan to the cases of $C_{m,n}$ and $D_{m,n}$. First, we will recall the main ingredients in the case of B_n.

Lemma 5.2.1. *As homogeneous spaces,* $S_{2n}/B_n \cong S_{2n-1}/(B_n \cap S_{2n-1}) = S_{2n-1}/B_{n-1}$.

Proof. The second equality follows from $B_n \cap S_{2n-1} = B_{n-1}$. For the first, define a map $\phi : S_{2n}/B_n \to S_{2n-1}/B_{n-1}$ by $\phi(gB_n) = (gB_n) \cap S_{2n-1}$. When defining ϕ, choosing the coset representative $g \in S_{2n-1}$ shows that ϕ is well-defined. It's straightforward to check that the S_{2n-1} action commutes with ϕ. $\square$

The next step lets us determine $\mathrm{Ind}_{B_n}^{S_{2n}}(1)$ by considering its restriction to S_{2n-1}. Although in general a representation is not uniquely determined by its restriction to a subgroup, we will see that in this case there is enough extra information to determine the decomposition.

Lemma 5.2.2. *The following recursive rule holds:*

$$\mathrm{Res}_{S_{2n-1}}^{S_{2n}}\left(\mathrm{Ind}_{B_n}^{S_{2n}}(1)\right) = \mathrm{Ind}_{S_{2n-2}}^{S_{2n-1}}\left(\mathrm{Ind}_{B_{n-1}}^{S_{2n-2}}(1)\right).$$

Proof. By Lemma 5.2.1, $\mathrm{Res}_{S_{2n-1}}^{S_{2n}}(\mathrm{Ind}_{B_n}^{S_{2n}}(1)) = S_{2n-1}/B_{n-1}$. But this is just a restatement of the definition of $\mathrm{Ind}_{B_{n-1}}^{S_{2n-1}}(1) = \mathrm{Ind}_{S_{2n-2}}^{S_{2n-1}}(\mathrm{Ind}_{B_{n-1}}^{S_{2n-2}}(1))$. $\square$

We will use the original result of Barbasch and Vogan in Section 5.3, so we prove it here for completeness. Here we say a partition $\lambda \vdash 2n$ is *even* if each of its parts λ_i is even.

Theorem 5.2.3. *The decomposition of* $\operatorname{Ind}_{B_n}^{S_{2n}}(1)$ *into irreducibles is*

$$\operatorname{Ind}_{B_n}^{S_{2n}}(1) \cong \bigoplus_{\substack{\lambda \vdash 2n \\ \lambda \text{ is even}}} S^{\lambda}.$$

Proof. This is true for $n = 1$. We use induction. Assume $\operatorname{Ind}_{B_{n-1}}^{S_{2n-2}}(1)$ has the described decomposition. We use Lemma 5.2.1. By the branching rule, $\operatorname{Ind}_{S_{2n-2}}^{S_{2n-1}}(\operatorname{Ind}_{B_{n-1}}^{S_{2n-2}}(1))$ contains each $\mu \vdash 2n-1$ having exactly one odd part, and each such μ appears once. Suppose $\operatorname{Ind}_{B_n}^{S_{2n}}(1)$ contains λ with at least three rows and at least two odd parts. Then the restriction of λ contains a μ with at least two odd parts; thus these λ do not occur. To rule out $\lambda = (\lambda_1, \lambda_2)$ with λ_1 and λ_2 odd, note that $(2n)$ occurs in $\operatorname{Ind}_{B_n}^{S_{2n}}(1)$ by Frobenius reciprocity. Therefore $(2n-1, 1)$ can't occur in $\operatorname{Ind}_{B_n}^{S_{2n}}(1)$, as it would contribute a second copy of $(2n-1)$ to $\operatorname{Ind}_{S_{2n-2}}^{S_{2n-1}}(\operatorname{Ind}_{B_{n-1}}^{S_{2n-2}}(1))$. An induction on i shows that $(2n-i, i)$ occurs in $\operatorname{Ind}_{B_n}^{S_{2n}}(1)$ if and only if i is even.

Finally, consider a λ with at least three even odd rows. Each μ obtained by deleting a box from λ occurs in $\operatorname{Ind}_{S_{2n-2}}^{S_{2n-1}}(\operatorname{Ind}_{B_{n-1}}^{S_{2n-2}}(1))$ exactly once, and a single copy of λ in $\operatorname{Ind}_{B_n}^{S_{2n}}(1)$ is the only way remaining to account for these μ. □

We now generalize Lemmas 5.2.1 and 5.2.2 to the cases of $C_{m,n}$ and $D_{m,n}$.

Lemma 5.2.4. *The following two recursive rules hold:*

$$\begin{aligned}\operatorname{Res}_{S_{nm-1}}^{S_{nm}}\left(\operatorname{Ind}_{C_{m,n}}^{S_{mn}}(1)\right) &= \operatorname{Ind}_{S_{(n-1)m}\times S_{m-1}}^{S_{nm-1}}\left(\operatorname{Ind}_{C_{m,n-1}}^{S_{m(n-1)}}(1)\otimes 1\right),\\ \operatorname{Res}_{S_{nm-1}}^{S_{nm}}\left(\operatorname{Ind}_{D_{m,n}}^{S_{mn}}(1)\right) &= \operatorname{Ind}_{S_{(n-1)m}}^{S_{nm-1}}\left(\operatorname{Ind}_{D_{m,n-1}}^{S_{m(n-1)}}(1)\right).\end{aligned}$$

Proof. The proof is a straightforward generalization of Lemmas 5.2.1 and 5.2.2. Observe that $C_{m,n} \cap S_{mn-1} = C_{m,n-1} \times S_{m-1}$ and that $D_{m,n} \cap S_{mn-1} = D_{m,n-1}$. We then have that $S_{mn}/C_{m,n} \cong S_{mn-1}/(C_{m,n-1} \times S_{m-1})$ and $S_{mn}/D_{m,n} \cong S_{mn-1}/D_{m,n-1}$. The results follow. □

If we know the decomposition of $\operatorname{Ind}_{C_{m,n}}^{S_{mn}}(1)$ into irreducibles, we can use Lemma 5.2.4 and Pieri's rule to decompose $\operatorname{Ind}_{C_{m,n+1}}^{S_{m(n+1)}}(1)$ into irreducibles. The same is true for $\operatorname{Ind}_{D_{m,n}}^{S_{mn}}(1)$

and $\operatorname{Ind}_{D_{m,n+1}}^{S_{m(n+1)}}(1)$, except that we use the branching rule. See Table 5.1 for some results for $m = 3$ and small n.

5.3 *Explicit isomorphism for* $m = 2$

Recall that a *matching* in a graph is a set of disjoint edges; we say a matching is a k-matching if it consists of k edges. Take S to be the set of n-matchings in K_{2n}; these are also known as perfect matchings. If we let S_{2n} permute the vertices of K_{2n}, then S is an S_{2n}-set and $\mathbb{C}[S]$ an S_{2n}-module. Note that S_{2n} acts transitively on S.

Fix the matching $s = 12|34|\cdots|2n-1, 2n$. Then the stabilizer of s in S_{2n} is exactly B_n as defined in Section 5.1. Then $\operatorname{Ind}_{B_n}^{S_{2n}}(1) \cong \mathbb{C}[S]$ as S_{2n}-modules. We give an explicit decomposition of $\mathbb{C}[S]$ into irreducibles; i.e., we provide a concrete linear map from each summand $S^\lambda \to \mathbb{C}[S]$. Note that the decomposition is determined up to isomorphism by Theorem 5.2.3. Our contribution here is to give an effectively computable isomorphism.

Lemma 5.3.1. *Let S be the set of k-matchings in K_{2k}. Then $\mathbb{C}[S] \cong \bigoplus_\lambda S^\lambda$, where the direct sum is over all partitions λ of $2k$ consisting of even parts. The multiplicity of each S^λ is 1.*

Proof. Fix an even λ. We will define a map $f : M^\lambda \to \mathbb{C}[S]$. For a single-row tabloid R, let $f(R)$ be the sum of all matchings in R. For a tabloid T with rows R_i, let $f(T) = \prod_i f(R_i)$; we interpret the product of disjoint matchings as their union. For example:

$$f\left(\begin{array}{cccc}\hline 1 & 2 & 3 & 4 \\ \hline\end{array}\right) = 12|34 + 13|24 + 14|23$$

$$\begin{aligned} f\left(\begin{array}{cccc}\hline 1 & 2 & 3 & 4 \\ \hline 5 & 6 & & \\ \cline{1-2}\end{array}\right) &= f([1234])f([34]) \\ &= (12|34 + 13|24 + 14|23) \cdot (56) \\ &= 12|34|56 + 13|24|56 + 14|23|56 \end{aligned}$$

Extend by linearity to M^λ. This is a map of S_n-modules, so its restriction to S^λ is either 0 or an isomorphism.

Let t be the standard tableau with entries in increasing order. We will show $f(e_t) \neq 0$. $f(\{t\})$ contains the term $m = 12|34|\cdots|2k-1, 2k$. If $\pm\pi\{t\}$ is another term in e_t such that $f(\pi\{t\})$ also contains m, then $2i$ and $2i-1$ must be in the same row of $\pi\{t\}$ for all i. But using column group operations, this is only possible if we switch $2i$ with $2j$ and $2i-1$ with $2j-1$. Therefore π is a product of an even number of disjoint transpositions, and in particular, $\mathrm{sign}(\pi) = 1$. So $f(\{t\})$ appears with positive sign in $f(e_t)$, and therefore $f(e_t) \neq 0$.

The proof is completed by noting that, per Theorem 5.2.3, we have accounted for each irreducible representation that appears. □

5.4 Multiplicities occur for $m = 3, n \geq 5$

The recursion rules established in Lemma 5.2.4 can be used to compute the decompositions of $\mathrm{Ind}_{C_{m,n}}^{S_{mn}}(1)$ and $\mathrm{Ind}_{D_{m,n}}^{S_{mn}}(1)$ for small values of m and n; see Table 5.1 at the end of this section.

As discussed above, in all cases we computed, there was a unique solution to the recursion containing a copy of the trivial representation. However, unlike the case $m = 2$, there does not seem to be any simple pattern to the decomposition. In particular, the decompositions are not multiplicity-free after the first few values of n.

In this section, we consider $V_n := \mathrm{Ind}_{C_{3,n}}^{S_{3n}}(1)$, and determine enough of the structure of V_n to show that for $n \geq 5$, V_n is not multiplicity-free. We accomplish this by considering *partition patterns.* A partition pattern $\lambda = (*, \lambda_1, \ldots, \lambda_k)$ represents any partition of length $k+1$ whose second through last parts equal λ. We also abandon tuple notation and simply concatenate digits, as all our entries are at most 9. For instance, the partition pattern 42 represents the partition $(n-6, 4, 2)$ for any n. As a special case, we let 0 denote the pattern $\emptyset$, representing the partition (n) for any n.

For any partition pattern λ, let $\mathrm{mult}(\lambda, n)$ be the multiplicity of S^λ in V_n. Also let $\mathrm{mult}(\lambda, n^-)$ be the multiplicity of S^λ in $\mathrm{Res}_{S_{3n-1}}^{S_{3n}}(V_n) = \mathrm{Ind}_{S_{3n-3}}^{S_{3n-1}}(V_{n-1})$. It is also convenient to refer to V_n and $\mathrm{Res}_{S_{3n-1}}^{S_{3n}}(V_n)$ as *level* n and *level* n^-, respectively.

We will first determine the multiplicities of certain S^{λ} in V_n. Then, we will use this structure to show that V_n is not multiplicity-free for $n \geq 5$.

Lemma 5.4.1. *The following* λ *have multiplicity* 1 *in all levels* $n \geq 5$*:* $0, 2, 3, 4, 22, 5, 41, 32$. *The following* λ *do not appear in any level:* $1, 21, 31, 221, 311, 411$.

Proof. It is easy to check that this holds for $n = 5$; see Table 5.1. Assuming by induction that the given decomposition holds for $n - 1$, we get a partial list of multiplicities at level n^-:

λ	0	1	11	2	21	3	111	4	31	22	211	1111	5	41	32	311	221
$\mathrm{mult}(\lambda, n^-)$	1	1	0	2	1	2	0	3	2	2	0	0	3	3	3	0	1

It is then straightforward to check that the given decomposition for level n is the only way to recover these multiplicities at level n^-. □

Theorem 5.4.2. *All levels* V_n *for* $n \geq 5$ *have multiplicities.*

Proof. By Lemma 5.4.1, it follows that $\mathrm{mult}(51, n) + \mathrm{mult}(42, n) = 2$ for all n. By considering the relevant children at level $n-1$, we can see that $\mathrm{mult}(51, n^-) + \mathrm{mult}(42, n^-) = 9$. Therefore, one of $\mathrm{mult}(51, n^-)$, $\mathrm{mult}(42, n^-) \geq 5$. But since each of 51 and 42 has four parents at level n, we must have multiplicities at level n. □

Table 5.1: The decomposition of $\mathrm{Ind}_{C_{m,n}}^{S_{mn}}(1)$ into irreducible representations for $m = 3$ and small n. Note that $n = 5$ is the first to contain multiplicities.

n	m	$\mathrm{Ind}_{C_{m,n}}^{S_{mn}}(1)$
2	3	$[4,2]$, $[6]$
3	3	$[4,4,1]$, $[5,2,2]$, $[6,3]$, $[7,2]$, $[9]$
4	3	$[4,4,4]$, $[5,4,2,1]$, $[6,2,2,2]$, $[6,4,2]$, $[6,6]$, $[7,3,2]$, $[7,4,1]$, $[8,2,2]$, $[8,4]$, $[9,3]$, $[10,2]$, $[12]$
5	3	$[5,4,4,2]$, $[5,5,3,1,1]$, $[6,4,2,2,1]$, $[6,4,4,1]$, $[6,5,2,2]$, $[6,6,3]$, $[7,2,2,2,2]$, $[7,4,2,2]$, $[7,4,3,1]$, $[7,4,4]$, $[7,5,2,1]$, $[7,6,2]$, $[8,3,2,2]$, $[8,4,2,1]$, $[8,4,3]$, $[8,5,2]$, $[8,6,1]$, $[9,2,2,2]$, $[9,4,2]$, $[9,4,2]$, $[9,6]$, $[10,3,2]$, $[10,4,1]$, $[10,5]$, $[11,2,2]$, $[11,4]$, $[12,3]$, $[13,2]$, $[15]$

BIBLIOGRAPHY

[1] Joan Adler and Uri Levi. Bootstrap percolation: visualizations and applications. *Brazilian Journal of Physics*, 33(3), 2003.

[2] M. Aizenman and J. L. Lebowitz. Metastability effects in bootstrap percolation. *J. Phys. A*, 21(19):3801–3813, 1988.

[3] Emil Artin. Über die zerlegung definiter funktionen in quadrate. In *Abhandlungen aus dem mathematischen Seminar der Universität Hamburg*, volume 5, pages 100–115. Springer, 1927.

[4] József Balogh and Béla Bollobás. Bootstrap percolation on the hypercube. *Probability Theory Related Fields*, 134:624–648, 2006.

[5] József Balogh, Béla Bollobás, Hugo Duminil-Copin, and Robert Morris. The sharp threshold for bootstrap percolation in all dimensions. *Trans. Amer. Math. Soc.*, 364:2667–2701, 2012.

[6] József Balogh, Béla Bollobás, and Robert Morris. Bootstrap percolation in three dimensions. *Annals of Probability*, 37(4):1329–1380, 2009.

[7] József Balogh, Béla Bollobás, and Robert Morris. Majority bootstrap percolation on the hypercube. *Combinatorics, Probability and Computing*, 18(1-2):17–51, 2009.

[8] Dan Barbasch and David Vogan. Weyl group representations and nilpotent orbits. In *Representation theory of reductive groups*, volume 40 of *Progress in mathematics*, pages 21–33. Springer, 1983.

[9] A. D. Barbour, Lars Holst, and Svante Janson. *Poisson Approximation.* Oxford University Press, 1992.

[10] Greg Blekherman, João Gouveia, and James Pfeiffer. Sums of squares on the unit hypercube. In preparation.

[11] C. Borgs, J. T. Chayes, Remco van der Hofstad, G. Slade, and J. Spencer. Random subgraphs of finite graphs: Ii. the lace expansion and the triangle condition. *Annals of Probability*, 33:1886–1944, 2005.

[12] Raphaël Cerf and Emilio N. M. Cirillo. Finite size scaling in three-dimensional bootstrap percolation. *Ann. Probab.*, 27(4):1837–1850, 1999.

[13] Raphaël Cerf and F. Manzo. The threshold regime of finite volume bootstrap percolation. *Stochastic Process. Appl.*, 101(1):69–82, 2002.

[14] J Chalupa, P L Leath, and G R Reich. Bootstrap percolation on a Bethe lattice. *J. Phys. C*, 12(1):L31–L35, 1979.

[15] Michele Conforti, Derek Corneil, and Ali Mahjoub. K_i-covers. I. Complexity and polytopes. *Discrete Math.*, 58(2):121–142, 1986.

[16] Ehud Friedgut and Gil Kalai. Every monotone graph property has a sharp threshold. *Proc. Amer. Math. Soc.*, 10:2993—3002, 1996.

[17] William Fulton. *Young tableaux*, volume 35 of *London Mathematical Society Student Texts*. Cambridge University Press, Cambridge, 1997. With applications to representation theory and geometry.

[18] Michel Goemans and David Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *J. Assoc. Comput. Mach.*, 42(6):1115–1145, 1995.

[19] João Gouveia and James Pfeiffer. A semidefinite approach to the K_i cover problem. *arXiv preprint arXiv:1211.0039*, 2012.

[20] João Gouveia, Monique Laurent, Pablo A. Parrilo, and Rekha R. Thomas. A new semidefinite programming hierarchy for cycles in binary matroids and cuts in graphs. *Math. Program.*, 133(1-2, Ser. A):203–225, 2012.

[21] João Gouveia, Pablo A. Parrilo, and Rekha R. Thomas. Theta bodies for polynomial ideals. *SIAM J. Optim.*, 20(4):2097–2118, 2010.

[22] João Gouveia and Rekha R. Thomas. Spectrahedral approximations of convex hulls of algebraic sets. In *Semidefinite Optimization and Convex Algebraic Geometry*, volume 13 of *MOS-SIAM Series on Optimization*, pages 293–340. 2012.

[23] Janko Gravner, Christopher Hoffman, James Pfeiffer, and David Sivakoff. Bootstrap percolation on the hamming torus. *arXiv preprint arXiv:1202.5351*, 2012.

[24] Janko Gravner, Alexander E Holroyd, and Robert Morris. A sharper threshold for bootstrap percolation in two dimensions. *Probab. Theory Related Fields*, 153(1-2):1–23, 2012.

[25] Martin Grötschel, László Lovász, and Alexander Schrijver. *Geometric algorithms and combinatorial optimization*, volume 2 of *Algorithms and Combinatorics.* Springer-Verlag, Berlin, second edition, 1993.

[26] Penny Haxell, Alexandr Kostochka, and Stéphan Thomassé. A stability theorem on fractional covering of triangles by edges. *European J. Combin.*, 33(5):799–806, 2012.

[27] David Hilbert. Über die darstellung definiter funktionen durch quadrateber die darstellung definiter formen als summe von formenquadraten. *Mathematische Annalen*, 32(3):342–350, 1888.

[28] Alexander E Holroyd. Sharp metastability threshold for two-dimensional bootstrap percolation. *Probability Theory Related Fields*, 125:195–224, 2003.

[29] Alexander E Holroyd. Astonishing cellular automata. *Bulletin du Centre de Recherches Mathematiques*, 13(1):10–13, 2007.

[30] Alexander E. Holroyd, Thomas M. Liggett, and Dan Romik. Integrals, partitions, and cellular automata. *Trans. Amer. Math. Soc.*, 356(8):3349–3368, 2004.

[31] Svante Janson, Tomasz Łuczak, Tatyana Turova, and Thomas Vallier. Bootstrap percolation on the random graph $G_{n,p}$. *Ann. Appl. Probab.*, 22(5):1989–2047, 2012.

[32] Michael Krivelevich. On a conjecture of Tuza about packing and covering of triangles. *Discrete Math.*, 142(1-3):281–286, 1995.

[33] Jean B. Lasserre. Global optimization with polynomials and the problem of moments. *SIAM J. Optim.*, 11(3):796–817, 2000/01.

[34] Monique Laurent. Lower bound for the number of iterations in semidefinite hierarchies for the cut polytope. *Math. Oper. Res.*, 28(4):871–883, 2003.

[35] László Lovász. On the Shannon capacity of a graph. *IEEE Trans. Inform. Theory*, 25(1):1–7, 1979.

[36] Montgomery McGovern and James Pfeiffer. Representation theory of generalized matchings. In preparation.

[37] Pablo A Parrilo. *Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization.* PhD thesis, California Institute of Technology, 2000.

[38] Pablo A Parrilo. Semidefinite programming relaxations for semialgebraic problems. *Mathematical programming*, 96(2):293–320, 2003.

[39] Gábor Pete. *Disease Process and Bootstrap Percolation.* PhD thesis, Bolyai Institute, Hungary, 1998.

[40] Stephen Prajna, Antonis Papachristodoulou, Peter Seiler, and Pablo A. Parrilo. SOSTOOLS and its control applications. In *Positive polynomials in control*, volume 312 of *Lecture Notes in Control and Inform. Sci.*, pages 273–292. Springer, Berlin, 2005.

[41] Bruce E. Sagan. *The symmetric group*, volume 203 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 2001. Representations, combinatorial algorithms, and symmetric functions.

[42] Roberto H Schonmann. On the behavior of some cellular automata related to bootstrap percolation. *Annals of Probability*, 20(1):174–193, January 1992.

[43] David Sivakoff. Random site subgraphs of the d-dimensional Hamming torus. 2010. Preprint available at: arXiv:1001.1007.

[44] E Slivken. Bootstrap percolation on the hamming torus with threshold 2. *manuscript in preparation*, 2013.

[45] Richard P. Stanley. Variations on differential posets. In *Invariant theory and tableaux (Minneapolis, MN, 1988)*, volume 19 of *IMA Vol. Math. Appl.*, pages 145–165. Springer, New York, 1990.

[46] Zsolt Tuza. A conjecture on triangles of graphs. *Graphs and Combinatorics*, pages 373–380, 1990.

[47] Remco van der Hofstad and Malwina Luczak. Random subgraphs of the 2D Hamming graph: the supercritical phase. *Probability Theory and Related Fields*, 147(1-2):1–41, 2010.

[48] C. D. van Enter. Proof of Straley's argument for bootstrap percolation. *J. Statist. Phys.*, 48(3-4):943–945, 1987.