2013

Hypergraph Capacity with Applications to Matrix Multiplication

John Lee Thompson Peebles Jr.
Harvey Mudd College

Hypergraph Capacity with Applications to Matrix Multiplication

John Peebles

Nicholas Pippenger, Advisor

Michael Orrison, Reader

Department of Mathematics

May, 2013

Abstract

The capacity of a directed hypergraph is a particular numerical quantity associated with a hypergraph. It is of interest because of certain important connections to longstanding conjectures in theoretical computer science related to fast matrix multiplication and perfect hashing as well as various longstanding conjectures in extremal combinatorics.

We give an overview of the concept of the capacity of a hypergraph and survey a few basic results regarding this quantity. Furthermore, we discuss the Lovász number of an undirected graph, which is known to upper bound the capacity of the graph (and in practice appears to be the best such general purpose bound).

We then elaborate on some attempted generalizations/modifications of the Lovász number to undirected hypergraphs that we have tried. It is not currently known whether these attempted generalizations/modifications upper bound the capacity of arbitrary hypergraphs.

An important method for proving lower bounds on hypergraph capacity is to exhibit a large independent set in a strong power of the hypergraph. We examine methods for this and show a barrier to attempts to usefully generalize certain of these methods to hypergraphs.

We then look at cap sets: independent sets in powers of a certain hypergraph. We examine certain structural properties of them with the hope of finding ones that allow us to prove upper bounds on their size.

Finally, we consider two interesting generalizations of capacity and use one of them to formulate several conjectures about connections between cap sets and sunflower-free sets.

Contents

List of Figures

List of Conjectures and Questions

Acknowledgments

I would like to thank my thesis advisor—Prof. Pippenger—and second reader—Prof. Orrison.

Chapter 1

General Background on Graph and Hypergraph Capacities

1.1 Motivation from Coding Theory

In (Shannon, 1956), Shannon introduced the notion of the zero-error capacity of a channel. A sender wants to send some data to a receiver by sending several messages all of the same length one after the other. There is a particular finite fixed alphabet that the messages are composed from. However, some pairs of letters are **confusable** with each other. Two messages are **confusable** if for all i, the ith letters of the messages are either the same or confusable with one another.

Let's look at a concrete example. Suppose we have an alphabet $\{a, b, c\}$ of size 3 where

- "a" is confusable with "b",
- "b" is confusable with "c",
- but "c" is not confusable with "a".

Then the message "aba" is confusable with "acb" because it has the same first letter "a" and its remaining two letters "ba" are confusable with their corresponding entries "cb" in the latter string. However, the messages "aba" and "acc" are not confusable because they each have a different, non-confusable last letter, namely, "a" and "c."

Together, the alphabet and list of confusable letters from the alphabet define a **channel**.

Shannon asked the question of how efficiently one could transmit messages without any ambiguity given a particular channel. More specifically, Shannon was interested in the rate of information transmission as the length of the messages got large.

The **information conveyed** by a message from a particular set of messages is the log base 2 of the cardinality of the set. This simply measures how much information you can tell somebody by sending a single message from the set, as measured in bits.

The **rate of information transmission** using a set of messages (all with the same length) is the information conveyed by a message from the set divided by the length of the message. Thus, if we have a set of m non-confusable messages each of length n, the rate of information transmission we get by using this set is $\frac{\log_2 m}{n}$. This measures how much information we transmit per letter we send. It is easy to verify that this quantity is at most the log of the size of the alphabet.

For example, consider the set of non-confusable messages "aba" and "acc" from our earlier example. This set has two messages ($m = 2$) each of length three ($n = 3$) so the rate of information transmission corresponding to this set is $\frac{1}{3}\log_2 2 = 1/3$. This example also illustrates that the rate of information transmission need not be integral.

With this as motivation, Shannon defined the zero-error capacity of a channel as

$$\sup_n \frac{\log_2 M(n)}{n}$$

where $M(n)$ is the size of the largest set of messages of length n over the alphabet of the channel such that no two messages in the set are confusable.

This definition is still in use in the fields of coding theory and information theory. However, in extremal combinatorics—which is the present interest of this paper—a different but interconvertible definition is used instead that is more combinatorial in nature. This definition is

$$\sup_n \sqrt[n]{M(n)}.$$

It is easy to see that the former definition is simply the log base 2 of the latter definition, so in this sense, they are interconvertible. Furthermore, combinatorialists usually refer to the quantity in question as the Shannon capacity rather than the zero-error capacity. For brevity of exposition, we simply use the term "**capacity**."

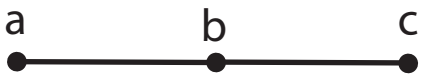

Figure 1.1 The graph corresponding to the example channel of Section 1.1.

Note that this alternative definition simply corresponds to looking at how the size of the largest set of non-confusable messages of length n grows as n gets large.

1.2 Combinatorial Definition of Capacity

Notice that the description of the channel in the previous section can be interpreted as an undirected graph. Specifically, a channel is just a set of letters (the alphabet) and a set of pairs of letters which are confusable. These can be interpreted as the vertex and edge sets respectively for a graph. In other words, the vertices of the graph correspond to letters, and there is an edge between two vertices iff their corresponding letters are confusable.

The example channel from the previous section of a three letter alphabet $\{a, b, c\}$ where

- "a" is confusable with "b",
- "b" is confusable with "c",
- but "c" is not confusable with "a"

corresponds to the graph shown in Figure 1.1.

An alternative definition of capacity uses this graph representation. In order to define it this way, though, we will need to define a few other things first.

Definition 1.1. *The **strong product** of two undirected graphs $G = (V_G, E_G)$ and $H = (V_H, G_H)$ is denoted $G \boxtimes H$. It is given by the graph with vertex set $V_G \times V_H$ and the following edge set. For any 2 vertices (g_1, h_1) and (g_2, h_2), there is an edge between them if and only if both of the following conditions hold*

1. *$g_1 = g_2$ or $(g_1, g_2) \in E_G$*
2. *$h_1 = h_2$ or $(h_1, h_2) \in E_H$.*

The following facts are well known and easy to prove.

Theorem 1.1. *The strong product of undirected graphs is commutative and associative up to isomorphism.*

Thus, the following is well-defined.

Definition 1.2. *For any graph G, let G^n denote $\underbrace{G \boxtimes \cdots \boxtimes G}_{n \text{ times}}$.*

We need one more definition before we can define capacity.

Definition 1.3. *An **independent set** in a graph is a set of vertices such that no edge has both its endpoints covered by vertices in the set. The **independence number** of a graph G, denoted $\alpha(G)$, is the size of the largest independent set in G.*

Then the capacity of an undirected graph can be defined as follows.

Definition 1.4. *The **capacity** of an undirected graph G denoted $\Theta(G)$ is defined as $\sup_n \sqrt[n]{\alpha(G^n)}$.*

It is not hard to show that $\alpha(G^n)$ is equal to the quantity $M(n)$ used at the end of the previous section. (Recall that $M(n)$ is the size of the largest set of messages, all of length n, such that no two messages in the set are confusable.) Thus, $\Theta(G) = \sup_n \sqrt[n]{\alpha(G^n)} = \sup_n \sqrt[n]{M(n)}$ so Definition 1.4 is equivalent to the modern definition in the preceding section.

1.3 Generalization to Directed Uniform Hypergraphs

The notion of capacity easily generalizes to directed uniform hypergraphs. Specifically, all we have to do is generalize the definitions of independence number and strong product, and then we can plug these in to the existing definition of capacity. Let's start by clarifying what we mean by a directed uniform hypergraph.

Definition 1.5. *Consider the pair (V, E) consisting of a set of vertices and and a set of k-tuples of vertices called **hyperedges** or simply edges where each hyperedge contains at least two edges. Such a pair is called a **k-uniform hypergraph**. A k-uniform hypergraph is called **directed** if the tuples are ordered and **undirected** if they are unordered. A **uniform hypergraph** is a hypergraph that is k-uniform for some k.*

In other words, a directed k-uniform hypergraph is similar to a directed graph, but now all edges go to exactly k vertices in some particular order (possibly with repeats) instead of two vertices. We maintain the requirement that an edge has to visit more than one vertex. Notice that an undirected 2-uniform hypergraph is simply an undirected graph.

Generalizing the definition of strong product, we have the following.

Definition 1.6. *The **strong product** of a pair of (un)directed k-uniform hypergraphs $G = (V_G, E_V)$ and $H = (V_H, E_H)$ is denoted $G \boxtimes H$ and defined as follows. It is the hypergraph with vertex set $V_G \times V_H$ and the following edge set. For any k vertices $(g_1, h_1), \ldots, (g_k, h_k)$, there is an edge between them if and only if both of the following conditions hold*

1. $g_1 = \cdots = g_k$ *or* $(g_1, \ldots, g_k) \in E_G$
2. $h_1 = \cdots = h_k$ *or* $(h_1, \ldots, h_k) \in E_H$.

For an undirected 2-uniform hypergraph (ie., an undirected graph) this is exactly Definition 1.4. Once again, we have

Theorem 1.2. *The strong product of (un)directed k-uniform hypergraphs is commutative and associative up to isomorphism.*

Thus, similarly to before, the following well-defined.

Definition 1.7. *For any (un)directed uniform hypergraph G, let the notation G^n denote $\underbrace{G \boxtimes \cdots \boxtimes G}_{n \text{ times}}$.*

Definition 1.8. *An **independent set** in an (un)directed hypergraph is a set of vertices such that no edge has all its endpoints covered by vertices in the set. The **independence number** of a directed or undirected hypergraph G denoted $\alpha(G)$ is the size of the largest independent set in G.*

Then the capacity of an undirected graph can be defined using essentially the same definition as before, but where the things referred to in the definition are the generalized quantities defined in this section.

Definition 1.9. *The **capacity** of an (un)directed uniform hypergraph G denoted $\Theta(G)$ is defined as* $\sup_n \sqrt[n]{\alpha(G^n)}$.

Similarly to the case of graphs, the capacity of a uniform directed hypergraph can be interpreted in terms of trying to find a large set of messages

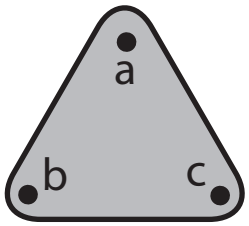

Figure 1.2 A 3-uniform undirected hypergraph G such that there is a bijection between independent sets in G^n and subsets of $\mathbb{Z}_3^n$ containing no three terms (not all equal) that sum to 0, where 0 is the tuple of n 0's.

where there can be no confusion, except that now, confusability is in terms of ordered k-tuples rather than unordered pairs.

For example, consider the undirected 3-uniform hypergraph consisting of three vertices $0, 1, 2$ with a single edge visiting each vertex once, as shown in Figure 1.2. Independent sets the nth power of this hypergraph have a 1-to-1 correspondence with subsets of $\mathbb{F}_3^n$ containing no three terms that sum to 0, where 0 is the tuple of n 0's. (Such subsets are called cap sets.) A well-studied open problem in extremal combinatorics is to obtain better bounds on the size of cap sets, so determining the capacity of this hypergraph would essentially resolve this problem. Currently, the best known lower bound is 2.21 due to (Edel, 2004) whereas the best upper bound is the trivial value of 3 we get by Theorem 1.4. Furthermore, certain strengthenings and weakenings of conjectures regarding its capacity are connected to matrix multiplication (Alon et al., 2012).

There are several other hypergraphs with the property that finding the hypergraph's capacity resolves or may resolve (depending on the capacity found) an important open problem. Such hypergraphs include a two vertex hypergraph which corresponds to the Erdős-Szemerédi Conjecture of (Erdös and Szemerédi, 1978) and a larger hypergraph corresponding to the strong USP conjecture of (Cohn et al., 2005). Furthermore, it is known that if the former hypergraph has a capacity of 2 or the latter hypergraph has a capacity of $3/2^{2/3}$ then there exists an $O(n^{2+\epsilon})$ time algorithm for matrix multiplication for every $\epsilon > 0$ (Alon et al., 2012; Cohn et al., 2005).

Even in the case of graphs, there is not a great deal known about the capacities of specific graphs. It took many years before the capacity of the 5-cycle (pentagon) was determined in the celebrated paper of (Lovász, 1979) and the capacity of the 7-cycle is still open.

1.4 Relevant Basic Results about Hypergraphs

We mentioned in the previous section that the capacity of an undirected graph is generalized by the notion of the capacity of a directed hypergraph. This is true in the following sense.

Theorem 1.3. *Let $G = (V, E)$ be an undirected uniform hypergraph. Let G' be the directed hypergraph with vertex set equal to V and with all permutations of every edge in E. Then the capacities of G and G' are equal.*

The proof follows immediately from the definitions. As such, any result true of directed hypergraphs is true of undirected hypergraphs so we will no longer state explicitly that results about capacity apply undirected hypergraphs unless they do not also apply to directed hypergraphs.

For any hypergraph, there are some easy, but very weak bounds on the capacity.

Theorem 1.4. *Let $G = (V, E)$ be a directed uniform hypergraph. Then $1 \leq \Theta(G) \leq |V|$.*

Proof. For the lower bound, can take a single vertex from the first power of the graph and that forms an independent set. For the upper bound, note that G^n has $|V|^n$ vertices so no independent set in this graph can exceed size $|V|^n$. □

In theory, we can actually compute arbitrarily good lower bounds on capacity. The problems are that (1) we have no guaranteed way of knowing how good the bound we have computed is and (2) doing the computation is only feasible in very small powers of the graph because it takes so long. Specifically, we have the following.

Theorem 1.5. *Suppose G is a directed uniform hypergraph and there is an independent set of size s in G^n. Then $\Theta(G) \geq \sqrt[n]{s}$.*

Proof. This follows immediately from the definition of capacity. □

As it turns out, we can view capacity as a limit rather than a supremum. The first step in doing so is the following theorem which is useful in its own right.

Definition 1.10. *Let G, H be directed uniform hypergraphs. Let f be a function mapping any subset of hypergraphs to real numbers. The function f is **submultiplicative** if $f(G \boxtimes H) \leq f(G)f(H)$. The function f is **supermultiplicative** if $f(G \boxtimes H) \geq f(G)f(H)$. The function f is **multiplicative** $f(G \boxtimes H) = f(G)f(H)$.*

Theorem 1.6. *The independence number is supermultiplicative.*

Proof. Suppose we have directed hypergraphs G, H. Let S, T be independent sets in G and H, respectively. Observe that $S \times T$ is an independent set in $G \boxtimes H$. □

We also need a lemma from analysis.

Lemma 1.1 (Fekete's Lemma)**.** *Let a_n be a bounded nonnegative sequence. If $a_{n+m} \geq a_n + a_m$ for all n, m then the sequence $\{a_n/n\}$ converges to $\sup_n a_n/n$.*

Proof. Fix any $\epsilon > 0$. Let $c = \sup a_n/n$. Choose k such that $c - a_k/k \leq \epsilon/2$. (Note that this quantity is always nonnegative.) Let n be any large integer and consider any j such that $kn \leq j \leq k(n+1)$. Then we have that $a_j/j \geq a_{kn}/(k(n+1)) \geq a_k n/(k(n+1))$.

Note that $\lim_{n\to\infty} a_k n/(k(n+1)) = a_k/k$ so for all sufficiently large n, $a_k/k - a_k n/(k(n+1)) \leq \epsilon/2$. Thus, $c - a_j/j \leq \epsilon/2 + \epsilon/2 = \epsilon$. This implies that for all sufficiently large m, $c - a_m/m \leq \epsilon$ which is the desired result. □

Theorem 1.7. *For any directed uniform hypergraph G, the capacity of G is given by $\Theta(G) = \lim_{n\to\infty} \sqrt[n]{\alpha(G^n)}$.*

Proof. Let $a_n = \log \alpha(G^n)$. Since α is supermultiplicative, a_n satisfies the conditions for Fekete's lemma. Applying this and taking the limit of $e^{a_n/n}$ yields the desired result. □

Another very useful theorem is a generic method for proving that a particular function upper bounds capacity. Specifically, it gives two conditions which together imply a function upper bounds capacity.

Theorem 1.8. *Let f map (un)directed uniform hypergraphs to real numbers. Furthermore, let f satisfy the following two conditions.*

1. *$f(G) \geq \alpha(G)$ for all (un)directed uniform hypergraphs G.*
2. *f is submultiplicative.*

Then $f(G)$ upper bounds the capacity of G.

Proof. Consider any k. Then $\sqrt[k]{\alpha(G^k)} \leq \sqrt[k]{f(G^k)} \leq f(G)$. So, $f(G)$ upper bounds everything we are taking the supremum over in the definition of capacity, so it upper bounds the capacity itself. □

Chapter 2

Upper Bounding Hypergraph Capacity

2.1 The Lovász Number of a Graph

The best general purpose upper bound on the capacity of an undirected graph is the Lovász number. Lovász introduced this bound in (Lovász, 1979) which resolved the open problem of the capacity of the pentagon. There are several very different looking, but nonetheless equivalent formulations of the Lovász number proven in (Lovász, 1979). Of these, it is the author's opinion that the best candidate for generalizing to hypergraphs is the eigenvalue formulation which we give after developing a prerequisite definition.

Definition 2.1. *For any matrix A, let $\lambda_{max}(A)$ denote the largest eigenvalue of A.*

Definition 2.2. *Let G be an undirected graph with n vertices. Let $\mathcal{A}_G$ be the set of all $n \times n$ matrices A such that $a_{ij} = 1$ if $i = j$ or (i, j) is not an edge in G. The **Lovász number** of a graph is defined as $\vartheta(G) := \inf_{A \in \mathcal{A}_G} \lambda_{max}(A)$.*

As mentioned earlier, Lovász proved that this upper bounds the capacity.

Theorem 2.1. *For any undirected graph G, $\vartheta(G) \geq \Theta(G)$.*

We defer proof of this fact until the next section where we prove it more generally.

Besides its relationship to capacity, the Lovász number is of independent mathematical interest for other reasons. Perhaps the most beautiful of these is the so-called "sandwich theorem".

Theorem 2.2 (Sandwich Theorem)**.** *The Lovász number is no more than the independence number of the graph and no less than its chromatic number.*

This is a rather amazing fact, as while the Lovász number can be computed in polynomial time, neither the independence number nor the chromatic number can be!

We now develop a natural generalization of this definition of the Lovász number to hypergraphs. The following is known.

Theorem 2.3 (Min-Max Principle)**.** *For any matrix A, the largest eigenvalue of A is given by* $\max_{\|x\|_2=1} x^T Ax = \max_{\|x\|_2=1} \sum_{i,j} a_{ij} x_i x_j$.

Thus, we can rewrite the Lovász number as follows.

Theorem 2.4. $\vartheta(G) = \inf_{A \in \mathcal{A}_G} \max_{\|x\|_2=1} \sum_{i,j} a_{ij} x_i x_j$.

Notice that the sum starts by taking the matrix A, then takes every entry times the entries of x corresponding to its coordinates. Since matrices are 2-dimensional (they have height and width), there are only two items we are summing over. However, this sum has a natural generalization to tensors.

2.2 The Lovász Number of a Hypergraph

First, we define what a tensor is.

Definition 2.3. *A width n **tensor** of order k is a set of real numbers indexed by the family of k-tuples* $(\underbrace{[n], \ldots, [n]}_{k \text{ times}})$.

For example, a width 3 tensor of order 2 is a 3×3 matrix, while a width 2 tensor of order 3 is a $2 \times 2 \times 2$ array that is 3-dimensional.

Note that some authors choose to use the term multidimensional array, multidimensional matrix, or hypermatrix instead and reserve the term tensor for a formally different or more general object.

We generalize the notion of largest eigenvalue to tensors using the natural extension of the min-max principle.

Definition 2.4. *For any order k tensor A, the largest eigenvalue of A is defined as* $\lambda_{max}(A) := \left(\sup_{\|x\|_k=1} \sum_{i_1,\ldots,i_k} a_{i_1 \cdots i_k} (x_{i_1} \cdots x_{i_k})\right)^{1/(k-1)}$.

It may not be immediately clear why the $1/(k-1)$ in the exponent is natural. The reason is that the largest eigenvalue of the all 1's matrix is n and we would like the same thing to hold for tensors. Using this exponent gives us this result. Note that we also could have obtained the same result instead by using the $\ell_{k/(k-1)}$ norm instead of the ℓ_k norm. A potential direction for future research is to investigate this alternative.

We now generalize the Lovász number to any undirected hypergraph as follows.

Definition 2.5. *Let G be an undirected k-uniform hypergraph with n vertices. Let $\mathcal{A}_G$ be the set of all width n, order k tensors A such that $a_{ij} = 1$ if $i = j$ or (i, j) is not an edge in G. The **Lovász number** of a graph is defined as $\vartheta(G) := \inf_{A \in \mathcal{A}_G} \lambda_{max}(A)$.*

This is a true generalization of the Lovász number insofar as for undirected 2-uniform hypergraphs (aka. undirected graphs), it amounts to just the definition of the Lovász number. However, we are primarily interested in generalizing the Lovász number because of its usefulness in bounding the capacity of graphs. As such, the more important consideration is whether this quantity bounds the capacity of arbitrary undirected hypergraphs (not just 2-uniform ones).

It is easy to show that ϑ upper bounds the independence number of a graph.

Theorem 2.5. *For any undirected uniform hypergraph G, $\vartheta(G) \geq \alpha(G)$.*

Proof. Consider any tensor A in $\mathcal{A}_G$. Let S be an independent set in G, and let x be a vector such that $x_i = 1/\sqrt[k]{|S|}$ if $i \in S$ and $x_i = 0$ otherwise. Then $\|x\|_k = 1$.

So,

$$\lambda_{\max}(A) \geq \left(\sum_{i_1,\ldots,i_k} a_{i_1 \cdots i_k} (x_{i_1} \cdots x_{i_k}) \right)^{1/(k-1)} = \left(|S|^k / S \right)^{1/(k-1)} = |S|.$$

Since this holds for all A, $\vartheta \geq |S|$. □

This gives us one out of the two conditions needed to apply Theorem 1.8 to prove that ϑ is an upper bound on capacity for general undirected uniform hypergraphs. However, proving the remaining item (submultiplicativity) appears much harder. The above proof uses the same essential idea as the case in graphs. However, the proof of submultiplicativity for graphs

requires converting the Lovász number as it is defined here to a fairly different form where the proof is essentially transparent. Furthermore, this conversion process appears to make critical use of the multiple equivalent characterizations of eigenvalues of matrices. Unfortunately, these multiple characterizations of eigenvalues do not always remain equivalent when generalized to tensors. Furthermore, we are unable to think of a good natural (or even unnatural) generalization for the definition of the Lovász number where proving submultiplicativity is more direct and transparent.

2.3 Converting the Capacity of an Odd-Uniform Hypergraph to that of an Even One

We attempted other methods of generalizing the Lovász number other than the one described in this paper. Many of these methods either did not work or yielded only very trivial bounds for k-uniform hypergraphs with k odd. As such, it was useful to be able to convert k-uniform hypergraphs with k odd to k'-uniform hypergraphs with k' even. Although we currently have no use for such a transformation with the generalization of the Lovász number described in this paper, we give the transformation found anyway as it seems to be something somewhat useful to be aware of.

Specifically, we give something a bit stronger: that for any k-uniform directed hypergraph H and any positive integer j, there exists a jk-uniform hypergraph with the same capacity as H.

In order to prove this, we define the algorithm, then prove two lemmas about it.

Definition 2.6. *For any k-uniform directed hypergraph H, $w_j(H)$ is defined as the hypergraph with the same vertex set as H and the edge set consisting of (e, e) for every hyperedge E of H.*

In other words, we are "doubling up" the hyperedge. For example, the hyperedge (a, b) would become (a, b, a, b).

Lemma 2.1. *The independence number of $w_j(H)$ is the same as that of H.*

Proof. Observe that any independent set in H is an independent set in $w_j(H)$ and vice versa. □

We now prove that w_j is multiplicative under the strong product.

Lemma 2.2. *For any k-uniform directed hypergraphs $G = (V_G, E_G)$, $H = (V_H, E_H)$, $w_j(G \boxtimes H) = w_j(G) \boxtimes w_j(H)$.*

Proof. Let $X = w_j(G \boxtimes H)$ and $Y = w_j(G) \boxtimes w_j(H)$. Let $E_{w_j(G)}$ be the edges of $w_j(G)$ and $E_{w_j(H)}$ be the same for H. By definition, X and Y have the same vertex sets. We need only verify that they both have their edges in the same place. Let E_X be the edge set of X and E_Y be the edge set of Y.

The set E_Y set consists of hyperedges of exactly the form

$$e_Y = ((g_1, h_1), \ldots, (g_{jk}, h_{jk}))$$

such that at least one of the following conditions holds

1. $g_1 = \cdots = g_{jk}$ and $h_1 = \cdots = h_{jk}$
2. $g_1 = \cdots = g_{jk}$ and $(h_1, \ldots, h_{jk}) \in E_{w_j(H)}$
3. $(g_1, \ldots, g_{jk}) \in E_{w_j(G)}$ and $h_1 = \cdots = h_{jk}$
4. $(g_1, \ldots, g_{jk}) \in E_{w_j(G)}$ and $(h_1, \ldots, h_{jk}) \in E_{w_j(H)}$.

Notice that no matter which of these four conditions is true, we can write $e_Y = ((g_1, h_1), \ldots, (g_k, h_k))^j$. Thus, E_Y is the set of all hyperedges of the form

$$e_Y = ((g_1, h_1), \ldots, (g_k, h_k))^j$$

such that at least one of the following conditions holds

1. $g_1 = \cdots = g_k$ and $h_1 = \cdots = h_k$
2. $g_1 = \cdots = g_k$ and $(h_1, \ldots, h_k) \in E_H$
3. $(g_1, \ldots, g_k) \in E_G$ and $h_1 = \cdots = h_k$
4. $(g_1, \ldots, g_k) \in E_G$ and $(h_1, \ldots, h_k) \in E_H$.

which is exactly the definition of E_X. □

This gives us the following theorem nearly immediately.

Theorem 2.6. *For any k-uniform directed hypergraph H, the capacity of H is the same as that of $w_j(H)$.*

Proof. The capacity of H is the limit of the sequence $a_n = (\alpha(H^n))^{1/n}$. The capacity of $w_j(H)$ is the limit of the sequence

$$b_n = (\alpha(w_j(H)^n))^{1/n} = (\alpha(w_j(H^n)))^{1/n} = (\alpha(H^n))^{1/n} = a_n.$$

Thus, the capacities of H and $w_j(H)$ are the limit of the same sequence, so they are equal. □

This allows us to use any bound methods developed for k-uniform hypergraphs with k odd on k'-uniform hypergraphs with k even.

Chapter 3

Lower Bounding Hypergraph Capacity

Theorem 1.5 allows us to search for maximum independent sets in powers of a graph in order to prove lower bounds on its capacity. However, this search is extremely slow, with the naive method requiring time 2^{n^k} for the kth power of an n-vertex graph. As such, developing faster algorithms for the maximum independent set (MIS) problem could allow slightly larger powers of graphs to be considered, possibly improving known bounds.

3.1 Chromatic Number and State of the Art MIS Algorithms

The state of the art for MIS algorithms designed to be fast in practice are branch and bound algorithms that use colorings to bound the size of the largest independent set. Specifically, these algorithms leverage the following well-known theorem.

Definition 3.1. *A **coloring** of a graph is an assignment of colors to vertices such that no edge has both its endpoints covered by vertices of the same color. The size of a coloring is the number of distinct colors used. The **chromatic number** of a graph is the size of its smallest coloring.*

Theorem 3.1. *The independence number of a graph is no more than the chromatic number of the complement of the graph.*

Any valid coloring upper bounds the chromatic number. These branch and bound algorithms try to get a good bound on the independence num-

ber by trying to get a good bound on the chromatic number. Examples of such algorithms include (Konc and Janezic, 2007; Tomita et al., 2010).

The first step for generalizing one of these algorithms to work on hypergraphs is to define a coloring of a hypergraph. Unfortunately, while the bound does hold under the natural generalization, it may be very weak. Specifically, we generalize as follows.

Definition 3.2. *A **coloring** of a directed hypergraph is an assignment of colors to vertices such that no edge has all its endpoints covered by vertices of the same color. The size of a coloring is the number of distinct colors used. The **chromatic number** of a directed hypergraph is the size of its smallest coloring.*

Theorem 3.2. *The independence number of a directed hypergraph is no more than the chromatic number of the complement of the undirected graph.*

Proof. Consider any maximum independent set in a hypergraph then take its complement. This gives us a clique. For any two vertices in this clique, there must be an edge that goes back and forth between them and visits no other vertices, so they cannot be the same color. Thus, a distinct color must be used for every vertex in the clique. □

The problem with this bound is that the proof that it works also applies to the following bound which subsumes it.

Theorem 3.3. *Let G be a directed hypergraph. Suppose that we take its complement, then remove all edges going to more than two vertices. Then the chromatic number of the resulting graph upper bounds the independence number by G.*

For the cap sets hypergraph (see Figure 1.2), this means that this bound will never give anything better than the total number of vertices in the graph. We have not yet evaluated it on other hypergraphs but expect that it would also provide a rather poor bound on them as well.

3.2 Reducing Hypergraph MIS to Graph MIS

Given that the bounding heuristics used in practical graph MIS algorithms don't seem to generalize effectively to hypergraphs, a natural alternative route to pursue is to reduce the hypergraph MIS problem to a graph MIS problem. Of course, one could do this reduction by way of the Cook-Levin Theorem, but the resulting graph MIS instance would likely not be very

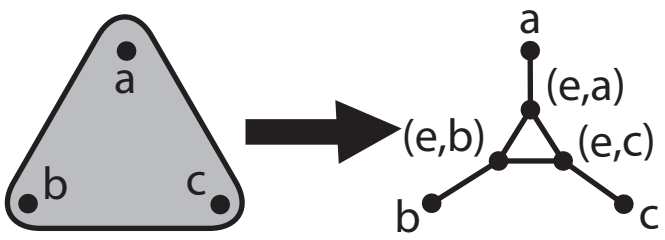

Figure 3.1 This shows how to replace any edge in a hypergraph with a gadget graph, while increasing the independent number of the resulting (hyper)graph by exactly 1. Note that if, say, a and c are the same vertex, the construction is the same but with the vertices a and c merged in the resulting graph.

nice. Since branch and bound algorithms for NP-hard problems rely critically on the problem instances being nice in order to find a solution quickly, a more natural reduction is desirable. We develop such a reduction here.

The reduction is as follows. Start with a 3-uniform hypergraph H. (This reduction naturally generalizes to k-uniform hypergraphs, but we use 3-uniform hypergraphs here for simplicity and because we are principally interested in the capacities of such graphs.) We create a graph G from H by replacing each edge with a gadget graph. Specifically, for each vertex v in H, create a vertex v in G. For each edge $e = (a, b, c)$ in H, create 3 additional vertices (e, a), (e, b), and (e, c) in G and connect them to form a triangle. Also connect (e, x) with x for all x. See Figure 3.1 for an illustration.

We claim that $\alpha(G) = \alpha(H) + |E_H|$ where $|E_H|$ is the number of vertices in H.

We claim that the independence numbers of G and H are related as follows.

Theorem 3.4. *$\alpha(G) = \alpha(H) + |E_H|$ where $|E_H|$ is the number of vertices in H.*

Proof. First consider any maximum independent set in H. This set is also independent in G. Fix any hyperedge $(a, b, c) \in H$. Observe that a, b, c cannot all be in the set. Suppose without loss of generality that a is not in the set. Then there is nothing stopping us from adding (e, a) to the set. Doing this for all hyperedges from H, we obtain an independent set in G of size $\alpha(H) + |E_H|$.

Conversely, consider any maximum independent set in G. Without loss of generality, we will suppose that for any edge $e = (a, b, c)$ in H, at least one of (e, a), (e, b), (e, c) is in the independent set. To see that this is without loss of generality, note that if, say, c is in the independent set but none

of (e, a), (e, b), (e, c) are in the set, we can replace c with (e, c) in the independent set. Furthermore, if none of a, b, c are in the independent set, we can also add (e, c). Applying this repeatedly gives us an independent set S with the desired property.

Now remove all vertices of the form (e, x) (ie., all vertices except those corresponding to vertices in H) from S to form a new set S'. This new set has cardinality $\alpha(G) - |E_H|$. Furthermore, it is an independent set in H because for any a, b, c with an edge in H, at least one of (e, a), (e, b), (e, c) is in S which means that at least one of a, b, c is not in S nor S'. □

3.3 Implementing Edge Tests

When doing computation on powers of hypergraphs, it is sometimes useful to be able to test whether a set of vertices has an edge between them. Even for rather small powers of a hypergraph, it quickly becomes inefficient to precompute this information and store it in memory. As such, we developed a simple procedure for testing for an edge that uses only a few bitwise operations.

We illustrate this procedure for powers of the following hypergraph H. Let H have two vertices 1 and 0. Let H have one undirected edge $(0, 0, 1)$. (We will discuss this graph more in Chapter 5.)

We can think of the vertices of H^n as binary strings of length n. Then three vertices in H^n don't have an edge iff there is some index such that two of the vertices have a 1 and one of the vertices has a 0. For example, "010," "110," and "010" have an edge. However, if we replace the last vertex in that list with "000," then the vertices do not have an edge between them because the second coordinates of the first two vertices are 1 while that of the last is 0.

For our purposes, we never need to consider powers of H higher than $n = 32$ so we can simply use an int to represent any vertex in H^n. For example, if $n = 3$, the int $5 = 101_2$ represents the vertex "101." Then we have the following

Theorem 3.5. *If a, b, c are ints representing vectors in H^n for $n \leq 32$, the following C snippet evaluates to true if they have an edge between them and false otherwise:*

$$!((a\&b\&\sim c)|(a\&\sim b\&c)|(\sim a\&b\&c))$$

Note that & means bitwise AND, | means bitwise OR, ~ means bitwise NOT, and ! gives 0 if the expression it is applied to is nonzero and 1 otherwise. Thus, the code above says simply that there is at least one index such that two vertices have a 1 and one has a 0. One can use similar ideas for powers of other graphs and to enumerate the "neighborhood" of a given vertex.

Chapter 4

Cap Sets

From a hypergraph capacity perspective, a cap set of width n is simply an independent set in nth power of the graph shown in Figure 1.2. However, there is an equivalent linear algebraic definition.

Definition 4.1. *A **cap set** of width n is a subset S of $\mathbb{Z}_3^n$ such that no three elements in S (not all equal) sum to 0. The **asymptotic solidity** is defined as $\sigma = \lim_{n\to\infty}(p_n)^{1/n}$ where p_n denotes the size of the largest cap set of width n.*

The equivalence between the definitions follows from the fact that three vectors in $\mathbb{Z}_3^n$ sum to 0 iff for each i, their ith entries are either all equal or all distinct. Another way of stating the definition is that cap sets are subsets of $\mathbb{Z}_3^n$ with no nontrivial three-term arithmetic progressions. In this chapter, we will be concerned primarily with the linear algebraic view of cap sets.

The best known upper bound on the size of a cap set of width n is $O(3^n/n)$ due to Meshulam (Meshulam, 1995). Meshulam uses essentially the same proof as that of Roth's Theorem (Roth, 1953). Unfortunately, no progress has been made improving upon this result and at least one mathematician has speculated that the current upper bound may be at the limit of what these techniques can obtain (Tao, 2007). Additionally, the proof via the techniques of Roth's Theorem is non-elementary. For these reasons, it would be helpful to develop a proof of this bound by alternate means. In this chapter, we develop some ideas which could potentially lead to such a proof. However, we are presently unable to match the bound obtained by Meshulam.

We begin by noting the following Theorem which was known to at least Knuth in 2001 but was probably known to others previously (Knuth, 2001).

Theorem 4.1. *Let $S \subseteq \mathbb{Z}_3^n$ and τ be an invertible affine transformation from $\mathbb{Z}_3^n$ to itself. Then $\tau(S)$ is a cap set if and only if S is.*

Proof. We can write $\tau(x) = \phi(x) + b$ where ϕ is an invertible linear transformation and $b \in Z_3^n$. Suppose $a, b, c \in \mathbb{Z}_3^n$ and $a + b + c = 0$ for a, b, c not all equal. Then $\tau(a) + \tau(b) + \tau(c) = 0$ for $\tau(a), \tau(b), \tau(c)$ not all equal. The converse also holds. □

It will also be useful to prove a basic bound on the size of cap sets of larger width in terms of those of smaller width. (We don't have any reference for this result, but it seems easy enough that it is likely that it was known previously.)

Theorem 4.2. *The largest cap set of width $n-1$ is at least a third of the size of the largest cap set of width n. (Equivalently, the size of the largest cap set of width n is no more than 3 times the size of the largest cap set of width $n-1$.)*

Proof. Consider any cap set S of width n. We can partition the vectors in it into three disjoint subsets S_0, S_1, S_2 according to whether the first value of each vector is 0, 1, or 2. At least one of these sets in the partition must have cardinality $|S|/3$. Let $i \in \{0, 1, 2\}$ be such that $|S_i| \geq |S|/3$.

Consider the subset T of $\mathbb{Z}_3^{n-1}$ obtained by removing the i at the beginning of each vector in S_i. Formally, $T = \{[x_1, \ldots, x_{n-1}] \mid [i, x_1, \ldots, x_{n-1}] \in S_i\}$. Clearly, T is also a cap set and has cardinality $|S_i| \geq |S|/3$. Thus, T is a cap set of width $n-1$ that is at least a third the size of the largest cap set of width n. □

Finally, we give one last pair of definitions.

Definition 4.2. *A cap set is **maximal** if adding any vector in $\mathbb{Z}_3^n - S$ to S causes S to no longer be a cap set. We say that a cap set is **maximum** or that it has **maximum cardinality** if its cardinality is at least as large as every other cap set of the same width.*

4.1 A Simple Bound on the Size of Cap Sets.

Suppose we have a cap set containing a, b with $a \neq b$. Then $-(a+b)$ cannot also be in the cap set, as then we would have $a + b + -(a+b) = 0$. This implies the following bound on the size of cap sets.

Definition 4.3. *Let A, B be subsets of a vector space. Then $A \boxplus B$ denotes $\{a + b \mid a \in A,\ b \in B,\ a \neq b\}$. Let the notation $-A$ denotes $\{-a \mid a \in A\}$.*

Lemma 4.1. *If S is a cap set of width n then $|S \boxplus S| \leq 3^n - |S|$ with equality iff S is maximal.*

Proof. Let S be a cap set. We have already explained why the inequality holds. Now suppose it does not hold with equality. Then there exists some vector not in the set $-(S \boxplus S) \cup S$. Such a vector can be safely added to the cap set. Thus, the cap set is not maximal.

Inversely, suppose the equation holds with equality. Then we clearly cannot increase the size of the cap set without violating the equation, so the cap set is maximal. □

This inequality alone implies $\lceil 3^n/2 \rceil$, but the use of additional features of cap sets is required if one wishes to improve this bound beyond $3^n/3$. More specifically, there exist sets of this size (not cap sets) satisfying the inequality in the above lemma.

Observation 4.1. *Any subspace of $\mathbb{Z}_3^n$ of dimension $n-1$ satisfies the inequality in Lemma 4.1.*

It is natural, then, to ask what features of cap sets differentiate them from strict subspaces of $\mathbb{Z}_3^n$. We could then attempt to show that any large set with such a feature violates Lemma 4.1, so cap sets cannot be large. For this reason, we investigate several structural properties of cap sets that may differentiate them from large subspaces in the next section.

One feature of cap sets suggested by the second part of the lemma is that for maximal cap sets, the equation holds with equality. (Note that in the observation above, the equation does not hold with equality.) This raises the question of whether this feature alone can yield good bounds on the size of cap sets.

Question 4.1 (Using Maximal Cap Sets)**.** *Do there exist large subsets S of $\mathbb{Z}_3^n$ with $|S \boxplus S| = 3^n - |S|$?*

We do not know the answer, but we can modify the previous observation to obtain something very close to an affirmative answer.

Theorem 4.3. *There exists a subset S of $\mathbb{Z}_3^n$ with $|S \boxplus S| = 3^n - |S| + 2$ and $|S| = 3^{n-1} + 1$.*

Proof. Let T be the union of the $n-1$ dimensional subspace of all vectors starting with a 1. Let $S = T \cup \{0\}$. Clearly, $|S| = 3^{n-1} + 1$. Also, $S \boxplus S$ is composed of 0, T, and $-T$, which are disjoint. So, $|S \boxplus S| = 2 \cdot 3^{n-1} + 1 = 3^n - |S| + 2$. □

4.2 Structural Features of Cap Sets

4.2.1 Bases in Cap Sets

One feature of cap sets that differentiates them from strict subspaces of $\mathbb{Z}_3^n$ is that they contain a basis for $\mathbb{Z}_3^n$. As a warm up, we show that every large cap set needs to contain a basis for $\mathbb{Z}_3^n$.

Theorem 4.4. *Every maximal cap set contains a basis for $\mathbb{Z}_3^n$.*

Proof. Let S be a cap set and suppose that it does not contain a basis for $\mathbb{Z}_3^n$. We claim it is not maximal. If S is empty, we are done. Otherwise, suppose it is nonempty. Note that S has rank at most $n-1$. As such, there exists an invertible linear transformation ϕ mapping S to the subspace of $\mathbb{Z}_3^n$ where the first entry in every vector is 0. Let $T = \{[1, x_1, \ldots, x_{n-1}] \mid [0, x_1, \ldots, x_{n-1}] \in \phi(S)$. In other words, we form T by taking every vector in $\phi(S)$ and changing the first entry from 0 to 1.Let $U = \phi(S) \cup T$. Then $\phi^{-1}(U)$ a strict superset of S (since S is nonempty, T must also be nonempty). It is straightforward to verify that U (and hence, $\phi^{-1}(U)$) is a cap set. □

Because of our earlier theorem about invertible affine transformations preserving whether a set is a cap set, we can assume without loss of generality that the basis we have found in this theorem is the standard basis. This has applications for using brute force to find large cap sets, as we only need to consider sets of vectors that contain the standard basis for $\mathbb{Z}_3^n$.

We can use a similar argument to show that every maximum cap set must be composed of a large number of disjoint bases for $\mathbb{Z}_3^n$.

Theorem 4.5. *Let S be a maximum cap set. Then there exists a subset R of S of cardinality at least $|S|/2 - 1$ such that R is the union of a pairwise-disjoint family of bases for $\mathbb{Z}_3^n$.*

Proof. Suppose that no such R exists. We form a new cap set V by iteratively removing a basis for $\mathbb{Z}_3^n$ from S until no such basis remains. Because a set R with the desired properties does not exist, the set V has cardinality at least $|S|/2 + 1$.

We then use the same trick as we used in the previous theorem of using an invertible linear transformation ϕ to project V into the subspace of $\mathbb{Z}_3^n$ where the first coordinate of every vector is 0. We then define T as we did previously by taking every vector in $\phi(V)$ and changing the first entry from 0 to 1. Finally, we note that $\phi(V) + T$ is a cap set of cardinality $2|V| \geq |S| + 2 > |S|$. □

The statement that R has cardinality at least $|S|/2 - 1$ was for convenience. The theorem actually holds if the cardinality is at least $|S|/2 - \epsilon$ for any $\epsilon > 0$.

Here is a conjecture which tries to use the fact that cap sets contain lots of bases to match the best known upper bound of $O(3^n/n)$

Conjecture 4.1 (Families of Bases)**.** *Let $S \subseteq \mathbb{Z}_3^n$ where $|S| \in \omega(3^n/n)$ and S is the union of $|S|/n$ disjoint bases for $\mathbb{Z}_3^n$. Then, $|S \boxplus S| > 3^n - |S|$.*[1]

This conjecture says that a large union of disjoin bases for $\mathbb{Z}_3^n$ must violate Lemma 4.1, so such a union cannot be a cap set. Since all cap sets must contain such a union, if this conjecture holds, there cannot exist large cap sets.

Theorem 4.6. *If Conjecture 4.1 holds then cap sets of width n have size $O(3^n/n)$.*

4.2.2 Closure of Cap Sets under Scalar Multiplication and Addition

Another way in which cap sets differ from subspaces is that they are not closed under linear combinations. This is a somewhat different scenario from that which is common in other areas of coding theory where optimal or near optimal constructions can be obtained by such sets. (Such sets are called **linear codes**.)

Here, we prove some results about the extent to which optimally large cap sets can be obtained using sets that are closed under these operations. First, we look at closure under scalar multiplication. Obviously, every cap set is closed under scalar multiplication by 1. We show that it is possible to construct maximum cap sets closed under scalar multiplication by 0 (ie., cap sets that contain the all zeros vector).

Theorem 4.7. *For every n, there exists a maximum cap set S of width n such that S contains the all zeros vector (ie., is closed under multiplication by 0).*

Proof. Let R be a maximum cap set of width n. Pick any vector $r \in R$ and set $S = R + (-r)$. Then S contains $r - r = 0$ and by Theorem 4.1, S is still a cap set. □

It is also possible to construct nearly maximum cap sets that are closed under multiplication by 2 (ie., cap sets closed under additive inverses).

[1] The usage of ω notation in this conjecture uses the "for infinitely many values of n" definition as opposed to the "for all sufficiently large n" definition. However, the usage of big-O notation still uses the "for all sufficiently large n" definition

Theorem 4.8. *For every n, there exists a cap set S of width n such that S is closed under additive inverses (multiplication by 2) and the cardinality of S is at least $2/3$ of the size of the largest cap set of width n.*

Proof. Fix any n. Let Q be any maximum cap set of width $n-1$. Let R be the cap set obtained by appending a 1 to the end of each element of Q. Let R' be obtained by negating each element in R. We claim that the set $S = R \cup R'$ has the desired properties.

First, we note that S is obviously closed under inverses by construction.

Second, we prove that S is within $2/3$ of the size of the largest cap set of width n. Let the size of the largest cap set of width n be denoted by p_n. By Theorem 4.2, $|R| = |R'| = |Q| \geq p_n/3$. Furthermore, R and R' are disjoint. Thus, $|S| \geq (2/3)p_n$.

Finally, we prove that S is actually a cap set. Recall that Q is a cap set. It is easy to see that this implies R and R' are cap sets.

Now consider any $a, b, c \in S$, not all equal. If a, b, c all come from R or they all come from R', $a + b + c \neq 0$. Otherwise, if two come from R and one comes from R' (or vice versa) the first coordinate of the vectors will sum to either $1 + 1 + 2 = 1 \pmod 3$ or $2 + 2 + 1 = 2 \pmod 3$ which is not 0. So, S is a cap set. □

This theorem can actually be seen as a first step in proving that large cap sets imply large sunflower free sets. See Chapter 5 and Theorem 5.5 for details.

Question 4.2. *Do there exist nearly optimal*

By essentially the same argument, the following is true. (Recall that σ is the asymptotic solidity, defined in Definition 4.1.)

Theorem 4.9. *For sufficiently large n, there exists a cap set S of width n such that S is closed under additive inverses (multiplication by 2) and the cardinality of S is at least $2/\sigma$ of the size of the largest cap set of width n.*

Note that this theorem means that if we could show a constant factor separation between the size of the largest cap set and the size of the largest cap set that is closed under additive inverses, we would immediately have lower bounds on the asymptotic solidity σ. In particular, showing that the constant $2/3$ in Theorem 4.8 is tight for sufficiently large n would imply that $\sigma = 3$.

We are unsure of whether this $2/3$ is tight, but we do conjecture that the $2/\sigma$ in Theorem 4.9 is tight.

Conjecture 4.2 (Additive Inverse Construction is Optimal)**.** *As $n \to \infty$, the size of the largest cap set of width n that is closed under additive inverses tends towards $2/\sigma$ of the size of the largest cap set of width n.*

Since there exist maximum cap sets that are closed under multiplication by 0, all cap sets are closed under multiplication by 1, and there exist nearly maximum cap sets closed under 2, one might conjecture that there exist nearly maximum cap sets that are closed under scalar multiplication. However, we show in a moment that it is impossible for nontrivial cap sets to be closed under both multiplication by 2 and 0 simultaneously.

Note the following observation which we will use a few times below.

Observation 4.2. *Any subset of $\mathbb{Z}_3^n$ that contains 0, an element $s \neq 0$, and $-s$ cannot be a cap set because $0, s, -s$ are not all equal and $0 + s + (-s) = 0$.*

Theorem 4.10. *The only cap set that is closed under both multiplication by 0 and multiplication by 2 is the cap set $\{0\}$ of cardinality 1.*

Proof. First, note that $\{0\}$ is a cap set closed under multiplication by 0 and 2. Now suppose by way of contradiction that $S \neq \{0\}$ is a cap set closed under multiplication by 2 and 0. Since S is closed under multiplication by 0, it contains 0. Since $S \neq \{0\}$ it contains some element $s \neq 0$. Then $-s \in S$ by closure under multiplication by 2. Thus, S is not a cap set. □

Since any linear combination of a subset of vectors in $\mathbb{Z}_3^n$ can be expressed as a sum of vectors from the set, this also implies that nontrivial cap sets cannot be closed under addition.

Corollary 4.1. *The only cap set that is closed under addition is $\{0\}$.*

These ideas also allow us to show that there exist maximum cap sets that contain no additive inverses.

Theorem 4.11. *For each n, there exists a maximum cap set S of width n such that $s \in S$ implies $-s \notin S$ for $s \neq 0$.*

Proof. By Theorem 4.7, there exists a maximum cap set S of width n that contains 0. By the observation above, if $s \in S$ and $s \neq 0$ then $-s \notin S$. □

In particular, this gives the following very simple bound on cap set size that is of course subsumed by the $O(3^n/n)$ bound via Roth's Theorem.

Corollary 4.2. *No cap set of width n can have size larger than $\lceil 3^n/2 \rceil$.*

4.2.3 Sums and Cap Sets

Since a cap set cannot be closed under addition, one naturally wonders about properties weaker than closure under addition that could hold for at least one large cap set. A natural such question is to ask is whether a large cap set S can be of the form $S = A + B$ for $A, B \subseteq \mathbb{Z}_3^n$. The answer is trivially yes if we set $A = S$ and $B = \{0\}$. Therefore, we ask the next natural question which is how large we can make A and B. (Note that $|S| \geq |A|, |B|$.)

We show that we can do so whenever $|A|$ is large and $|B|$ is constant, provided the asymptotic solidity is 3. Actually, we show something more general.

Theorem 4.12. *Let T be a subset of a group H, and set $p = |T|/|H|$. Then for any constant power of 2 denoted $c = 2^k$ for which we have $|H| \geq 2c/p^c$, there exists a cap set $S \subseteq T$ such that $S = A + B$ where $|A| \geq p^c|S|$ and $|B| \geq c$.*

For this, we need to first develop some probabilistic arguments.

Definition 4.4. *A **multiset** is a mathematical collection in which repeated elements are allowed. The **multiplicity** of an element in a multiset is the number of times it appears in the multiset. Let $|M|$ denote the number of elements in M, counting duplicates. We use the brackets $\langle, \rangle$ to denote multisets.*

For example, $M = \langle 1, 1, 2, 4, 5, 5, 5 \rangle$ is a multiset where 5 has multiplicity 3 and $|M| = 7$.

The following theorem allows us to estimate the number of unique elements in a multiset based on the probability that two elements drawn randomly from it are equal. Interestingly, this estimate does not depend explicitly on the size of the multiset.

Theorem 4.13. *Let M be a finite multiset with an associated equivalence relation. Suppose we draw two elements from M uniformly at random where the order in which the elements are drawn matters. Let p be the probability that these two elements are equivalent under the equivalence relation. Then the number of unique elements in the multiset (under the equivalence relation) is at least $1/p$.*

Proof. Let S be the set of elements in M (ie., with duplicates eliminated), and let m_s denote the multiplicity of s in M. Let m denote the mean of m_s over all $s \in S$. Equivalently, $m = |M|/|S|$ and $m = \mathbf{E}_{s \in S}[m_s]$ where $\mathbf{E}$ denotes the expected value.

If we choose two random elements of M (where order matters) and condition on the first element s from S that is selected, we have

$$p = \sum_{s \in S} \left(\frac{m_s}{|M|} \right)^2 = \frac{\mathbf{E}_{s \in S}[m_s^2]}{|S|m^2} \geq \frac{(\mathbf{E}_{s \in S}[m_s])^2}{|S|m^2} = \frac{1}{|S|}.$$

Thus, $|S| \geq 1/p$. □

Intuitively, what the above proof is doing is saying that $p = 1/|S|$ when the multiplicities of the elements are the same and that if they are different, this can only increase p.

Our next theorem says that if we have a lot of sets that don't overlap very much, their union must be large. This theorem is given in (Jukna, 2011) also using the probabilistic method.

Theorem 4.14. *Let $\mathcal{F}$ be a nonempty family of sets where the ith member of the family is denoted $\mathcal{F}_i$ and such that $|\mathcal{F}_i| = x$ for all i. Suppose that for a random pair of elements from distinct sets in the family, the probability that they are equal is p. Also suppose that there exists an $r \in (0, 1)$ such that $|\mathcal{F}| \geq \frac{1-\frac{1}{p}}{1-\frac{1}{r}}$. Then $|\bigcup_i \mathcal{F}_i| \geq r \cdot x/p$.*

Proof. Intuitively, we dump all the sets into a big multiset[2] and apply Theorem 4.13. More formally, form the set $S = \{(i, f) \mid f \in \mathcal{F}_i\}$. One can think of the first coordinate of each element in S as identifying which set in the family that element came from. For example, if we had the element $(2, \ldots)$, that means that the element came from the set $\mathcal{F}_2$. Let two elements in S be equivalent iff they share the same second coordinate. Clearly, the number of unique elements in S under this equivalence relation is equal to $|\bigcup_i \mathcal{F}_i|$.

We estimate the probability that a randomly chosen pair of elements from S (where order matters) are equal. Condition on the set in the family that each element in the pair came from; ie., on the values of the elements' first coordinates. If the sets they came from are the same, the probability q that the pair is equivalent is $1/x$. Otherwise, this probability is at most $px/x^2 = p/x$. By this, the fact that $(1 - p)$ is positive, and our assumption on $|\mathcal{F}|$, we have

$$q \leq \frac{1}{|\mathcal{F}|} \cdot \frac{1}{x} + \left(1 - \frac{1}{|\mathcal{F}|}\right) \frac{p}{x} = \frac{1}{|\mathcal{F}|} \cdot \frac{1}{x}(1 - p) + \frac{p}{x} \leq \frac{1}{r} \cdot \frac{p}{x}.$$

By Theorem 4.13, $|\bigcup_i \mathcal{F}_i| \geq 1/q \geq r \cdot x/p$. □

[2]Formally, it will be a set with an equivalence relation.

For our purposes, using $r = 1/2$ and considering the largest intersection of sets suffices.

Corollary 4.3. *Let $\mathcal{F}$ be a nonempty family of sets where the ith member of the family is denoted $\mathcal{F}_i$ and such that $|\mathcal{F}_i| = x$ for all i. Suppose that there exists a $p \in (0,1)$ such that $|\mathcal{F}| \geq \frac{1}{p} - 1$ and for all i, j, we have $|\mathcal{F}_i \cap \mathcal{F}_j| \leq px$. Then $|\bigcup_i \mathcal{F}_i| \geq \frac{1}{2} \cdot x/p$.*

By repeated application of this corollary, we obtain the following result due to Erdős (Jukna, 2011).

Corollary 4.4 (P. Erdős)**.** *Let $\mathcal{F}$ be a nonempty family of sets where the ith member of the family is denoted $\mathcal{F}_i$ and such that $|\mathcal{F}_i| = N/w$ for all i. Let c be a constant power of 2. If $|\mathcal{F}| \geq 2cw^c$, then there exist $\mathcal{F}_{i_1}, \ldots, \mathcal{F}_{i_k}$ where $|\mathcal{F}_{i_1} \cap \cdots \cap \mathcal{F}_{i_k}| \geq N/(2w^k)$*

Now, we can prove Theorem 4.12.

Proof of Theorem 4.12. Let $R_b = T + (-b)$. Note that $|R_b| = |T|$. Let $\mathcal{F}$ be the family of all R_b for $b \in H$. Then $|\mathcal{F}_i| = |R_i| = |T| = |H|/(1/p)$. Set $N = |H|$ and $w = 1/p$. Note that $|\mathcal{F}| = |H| \geq 2cw^c$ by assumption. Thus, There exists a subset B of H with size at least c such that for all $b \in B$, we have $|\bigcap_{b \in B} R_b| \geq p^k |H|/2$. Take $A = \bigcap_{b \in B} R_b$. This gives the desired result. □

However, we can also prove that if the asymptotic solidity is 3 then there cannot exist large cap sets that are the sum of large sets. Specifically, we prove the following.

4.2.4 Sum Graphs

In order to study this some of the ideas from the previous subsection further we developed a graph-theoretic interpretation of cap sets that is different from the hypergraph interpretation mentioned earlier.

Definition 4.5. *Let S be a subset of a group H. The **graph corresponding to S** denoted $G(S)$ is a graph defined according to its adjacency matrix as follows. The adjacency matrix M is a $|H| \times |H|$[3] matrix formed by indexing the rows and columns with the elements of H, then setting $m_{h_1,h_2} = 1$ if $h_1 + h_2 \in S$ and setting $m_{h_1,h_2} = 0$ otherwise.*

[3] $|\mathrm{H}|$ will always denote the number of elements in H

Note that this graph may contain self-loops. However, if we require that S be a cap set containing 0—which we can do without affecting its size by Theorem 4.7—then it only contains one self-loop and this self loop is on the vertex corresponding to 0.

We can immediately prove a number of things about these sorts of graphs for arbitrary groups.

Theorem 4.15. *Let S be a subset of a group H. Then the following hold.*

1. *$G(S)$ has $|H|$ vertices and $|S| \cdot |H|$ edges.*
2. *$G(S)$ is $|S|$-regular.*

Proof. (1) is a special case of (2). To see that $G(S)$ is $|S|$-regular, note that for any $h_1 \in H$ and $s \in S$, there is exactly one vector $h_2 \in H$ such that $h_1 + h_2 = s$. Thus, the vertex corresponding to h_1 has $|S|$ edges leaving it. Thus, every vertex has $|S|$ edges leaving it. □

There are also some properties that these graphs have by virtue of being taken from cap sets.

Theorem 4.16. *If S is a cap set then $G(S)$ is triangle-free. Equivalently, a cap set S cannot be written as $S = R + R$ for any R with $|R| > 1$.*

Proof. The edges in a triangle correspond to $-a$, $-b$, and $a + b$ for $a \neq b$. These elements cannot all be in S because they sum to 0. □

The following is equivalent to the fact that the Cartesian product of cap sets is a cap set.

Theorem 4.17. *Let S, T be a cap sets with widths n, k. Let $\otimes$ denote the tensor product of graphs which is given by taking the Kronecker product of their adjacency matrices. Then the $G(S) \otimes G(T)$ corresponds to a cap set of width nk and size $|S| \cdot |T|$.*

Proof. The tensor product of graphs is equivalent to taking the Cartesian product of the vertex sets and putting an edge between two vertices (a, b) and (c, d) iff (a, c) and (b, d) are edges in the original graphs, respectively. From this, it follows that exactly the pairs of elements from the original cap sets (one from each) are represented in the new graph. □

There are a number of interesting questions remaining to be asked regarding graphs corresponding to cap sets. For example,

Question 4.3 (Characterization of Cap Set Graphs)**.** *Is there a nice characterization of the graphs corresponding to cap sets that does not reference cap sets?*

A good place to start would be to determine whether the family of graphs corresponding to cap sets is closed under isomorphism.

Chapter 5

Generalizations of Capacity

5.1 Weighted Capacity

In this section, we introduce a weighted generalization of graph capacity. As we will see shortly, this is useful for getting some numerical evidence about whether the existence of large cap sets implies the existence of large sunflower free sets. (Sunflower free sets will be defined shortly.)

5.1.1 Weighted Generalization

Throughout this section, we will be dealing with hypergraphs that have weights on their vertices. We will use the notation $w(x)$ to denote the weight of vertex x and $w(S)$ to denote the weight of the set S. We first generalize the strong product.

Definition 5.1. *Let G, H be a pair of hypergraphs where each has real weights on its vertices. The **weighted strong product** of G and H is the hypergraph given by the strong product of G and H. Furthermore, it has weights on its vertices determined by setting the weight of vertex (g, h) in the product graph equal to the weight of g times the weight of h for $g \in G$, $h \in H$.*

The commutativity and associativity (up to isomorphism) of this product follow from the commutativity and associativity of the strong product and multiplication of reals.

We can also extend the definition of the independence number to hypergraphs with weighted vertices.

Definition 5.2. *For any hypergraph H with weighted vertices, the **independence number** of H denoted $\alpha(H)$ is the weight of the largest weight independent set in H.*

a. The hypergraph E. Independent sets in powers of this graph are **sunflower free sets**.

b. The hypergraph W. Independent sets in powers of this graph give lower bounds on the size of cap sets.

Figure 5.1 A weighted and unweighted hypergraph. Both hypergraphs have a single undirected edge which visits the right vertex twice and the left vertex once. We will use the label 1 to refer to the left vertex and 0 to refer to the right vertex in each of these graphs, respectively. (Note that these are purely descriptive labels, not weights.)

Having generalized the notion of independent sets and the strong product, the notion of capacity naturally extends to hypergraphs with weighted vertices. It is not hard to see that most of the basic facts about the capacity of a hypergraph still apply in the weighted setting. In particular, note the following.

Theorem 5.1. *For any hypergraph H with weighted vertices, $\alpha(H^n)^{1/n}$ is a lower bound on the capacity of H.*

This follows by essentially the same proof as in the unweighted case.

5.1.2 Applications to Cap Sets

In Figure 5.1, two hypergraphs are shown. For convenience, we name them E and W respectively. We claim that independent sets in powers of the hypergraph W shown in Figure 5.1b give lower bounds on the size of cap sets. This theorem implies that the capacity of W lower bounds the asymptotic solidity.

Theorem 5.2. *There exists a cap set of width n and size $\alpha(W^n)$.*

Proof. Any independent set S in W^n is already a cap set of width n. We need to show how to produce a new cap set T with $|T| = w(S)$. The basic idea is that we view each vector in S as acting like a "regular expression" that represents several vectors in T.

We construct T as follows. For each element s in S, we add a set R_s of vectors to T with $|R_s| = w(s)$. Let s be any element of S with weight $w = 2^k$. Then s has k 0's. Let R_s be the subset of $\mathbb{Z}_3^n$ of all vectors that have 1's in exactly the same places as s and some combination of 2's and 0's elsewhere. There are clearly $2^k = w$ such vectors, so $|R_s| = w$ as claimed. Take $T = \bigcup_{s \in S} R_s$. It is not hard to verify that T is a cap set. □

Note that a similar idea to that used in the proof of this theorem was known prior to this work: in papers proving lower bounds on the size of cap sets, it has been common for authors to specify large cap sets using what essentially amounts to a simplified form of regular expressions.

Notice that the trivial upper bound on the capacity of E is 2, and the trivial upper bound on the capacity of W is 3. A major open problem in extremal combinatorics is whether the capacity of E is equal to 2. (Most combinatorialists believe it is not.) We prove that E has capacity equal to its trivial bound iff W does.

Theorem 5.3. *E has capacity 2 if and only if W has capacity 3.*

In order to prove this we make use of a special case of Theorem 2.4 from (Alon et al., 2012).

Theorem 5.4 ((Alon et al., 2012))**.** *Let $\beta(E^n)$ denote the size of the largest independent set in E^n subject to the constraint that every vector in the set has exactly $\lfloor (2/3)n \rfloor$ 0's in it. Define the $(2/3)$-capacity of E as $\lim_{n\to\infty}(\beta(E^n))^{1/n}$. Then the $(2/3)$-capacity of E is $\frac{3}{2^{2/3}}$ if and only if E has capacity 2.*

Proof of Theorem 5.3. For the forward direction, note that Theorem 5.4 allows us to obtain sets of width n which are independent in W, which have $\lfloor (2/3)n \rfloor$ 0's in each vector in the set, and which also have $\left(\frac{3}{2^{2/3}} + o(1)\right)^n$ vectors in the set.[1] The middle fact implies that each vector in these sets has weight $2^{\lfloor (2/3)n \rfloor}$. Let a_n be an arbitrary such set. Then the capacity of W is lower-bounded by $\lim_{n\to\infty}(w(a_n))^{1/n} = \frac{3}{2^{2/3}} \cdot 2^{2/3} = 3$. Thus, the capacity of W is equal to 3.

For the reverse direction, suppose that W has capacity 3. Our strategy is to show that large independent sets in W^n have large subsets where each vector in the subset has weight about $2^{(2/3)n}$. First, we explain why this suffices. For each n, let r_n be the largest-weight independent set in W^n. Partition the vectors in r_n into subsets according to the fraction of entries in the vector that are 0. Set q_n equal to the largest-weight subset in the partition, and set δ_n equal to the fraction of entries in each vector in q_n that are 0. Clearly, $w(r_n) \geq w(q_n) \geq w(r_n)/(n+1)$. As such, $\lim_{n\to\infty}(w(q_n))^{1/n} = 3$.

Let b_n be the largest independent set in E^n subject to the constraint that there are exactly $\delta_n n$ 0's in each vector in the set. Since q_n satisfies this constraint, we have

[1] Note that when we use little-o notation in an equation like this, it simply means that there exists a function that is in the set defined by the notation and that satisfies the equation.

$$w(q_n)/2^{\delta_n n} = |q_n| \leq |b_n|. \tag{5.1}$$

Taking the nth root and the limit as $n \to \infty$ of both sides yields the inequality $\frac{3}{\lim_{n\to\infty} 2^{\delta_n}} \leq |b_n|$. Note that we are not assuming that the limit exists.

It suffices to show that $\delta_n \to (2/3)$ because then we can invoke Theorem 5.4 to finish the proof. In other words, we need to show that large independent sets in W^n have large numbers of vectors where $(2/3)$ their entries are 0's.

By Equation 5.1, the fact that $|q_n| \leq \binom{n}{\delta_n n}$, and the capacity of W, we have

$$(3 + o(1))^n = w(q_n) \leq \binom{n}{\delta_n n} 2^{\delta_n n} = (2 + o(1))^{(H(\delta_n)+\delta_n)n} \tag{5.2}$$

where $H(p) = -p \log_2 p - (1-p)\log_2(1-p)$ is the binary entropy of p. The last step in Equation 5.2 is justified by the fact that for large n, δ_n is bounded away from 0 and 1, which follows from the fact that if this were not true, we would have for sufficiently large n that $\binom{n}{\delta_n n} 2^{\delta_n n} \leq (2 + o(1))^n$ which is less than the left hand side of Equation 5.2 for sufficiently large n. Thus, $3 + o(1) \leq (2 + o(1))^{H(\delta_n)+\delta_n}$ which is equivalent to $3 + o(1) \leq 2^{H(\delta_n)+\delta_n}$. Note that the maximum of the right hand side is 3 which is uniquely attained when $\delta_n = 2/3$. The value $\delta_n = 2/3$ is also the δ_n coordinate of the unique critical point of the exponent of the right hand side. Because of this and the fact that 2^x is monotonically increasing, $2^{H(\delta_n)+\delta_n}$ decreases monotonically as δ_n gets further away from $2/3$.

For sufficiently large n, $3 + o(1)$ gets arbitrarily close to 3. This forces δ_n arbitrarily close to $2/3$. So, $\lim_{n\to\infty} \delta_n = 2/3$ as desired. This completes the proof. □

Note that it appears extremely difficult to get anything more specific than what we have stated in the theorem above. In particular, if one wishes to obtain a more explicit relationship between the capacity of E and W, one would not be able to make much use of Theorem 5.4 because—even if one works out the exact estimates it provides for capacities lower than the trivial bound—these estimates turn out to be very poor E.

As a corollary of this and Theorem 5.2, we have the following which is also proven in (Alon et al., 2012).

Corollary 5.1 ((Alon et al., 2012))**.** *If the capacity of E is 2 then the asymptotic solidity is 3.*

Sequence\Value of n	1	2	3	4	5	6	7
size of largest cap set of width n	2	4	9	20	45	112	?
$\alpha(E^n)$	1	2	3	4	7	11	16
$\alpha(W^n)$	2	4	8	20	40	96	224

Table 5.1 Comparison of the size of the largest cap sets of width n with $\alpha(E^n)$ and $\alpha(W^n)$ for small n. Note that the sequence of cap set sizes at the top comes from OEIS (Havermann, 2009).

Thus, the capacity of W gives us a number which "bridges the gap" between sunflower free sets and cap sets. In particular, it would be very interesting to know whether the capacity of W is equal to the asymptotic solidity. (Theorem 5.2 proves that the capacity of W is no more than the asymptotic solidity.) Note that if this were true, it would provide strong evidence that the asymptotic solidity is less than 3 because then, the asymptotic solidity equaling 3 would imply a statement that most mathematicians believe is likely false, namely, that the capacity of E is 2.

In Table 5.1, we provide numerical evidence that the capacity of W is equal to the asymptotic solidity. This evidence comes from the fact that the size of the largest cap set of width n is approximately equal to the weight of the largest-weight independent set in W^n. If this holds for all n, then the converse of Corollary 5.1 is true. We conjecture that this is the case.

Conjecture 5.1 (Equivalence of Cap Sets and Sunflower Free Sets)**.** *Let a_n be the size of the largest cap set of width n. Then* $\lim_{n\to\infty} a_n/\alpha(W^n) = 1$.

Another interesting pattern in the table is that for even n, $\alpha(W^{n+1})$ is given by twice the size of the largest cap set of width n. We conjecture this holds always.

Conjecture 5.2 (Doubling Conjecture)**.** *Let a_n denote the size of the largest cap set of width n. Then for all even n, $2a_n = \alpha(W^{n+1})$.*

Finally, we give a conjecture (also based on numerical evidence) for the value of $\alpha(E^n)$.

Conjecture 5.3 (Growth of Largest Sunflower Free Sets)**.**

$$\alpha(E^n) = \lceil 2^{(2/3)(n-1)} \rceil \approx (1.5874\ldots)^{n-1}.$$

This holds for $1 \leq n \leq 7$. It is slightly larger than the best known lower bound of $(1.551\ldots)^{n-2}$ due to (Deuber et al., 1997). The appearance of the

number $2^{2/3}$ in this conjecture is notable because $2^{2/3}$ is also the base of the exponential function describing the size of the best known construction of a related combinatorial object called a strong uniquely solvable puzzle (Cohn et al., 2005).

5.1.3 Proving Equivalence of Capacity and Solidity

Here, we give the start of an approach one might take to proving that the capacity of W is at least the asymptotic solidity. Specifically, the following theorem provides an approach.

Theorem 5.5. *Let σ be the asymptotic solidity. Suppose that there exist cap sets of size $(\sigma + o(1))^n$ with the property that if x in the cap set, then negating any one of x's coordinates produces a vector that is also in the cap set. Then σ is equal to the capacity of W.*

Proof. It is not hard to see that reversing the construction in Theorem 5.2 produces the desired result. □

Note that in Theorem 4.8, we prove a necessary condition for the existence of cap sets of the form described in the above theorem. Namely, we show large cap sets that are closed under taking the inverse of the entire vector. We conjecture that one can actually achieve cap sets of the stronger form described here.

Conjecture 5.4 (Existence of Special Cap Sets)**.** *Let σ be the asymptotic solidity. There exist cap sets of size $(\sigma + o(1))^n$ with the property that if x in the cap set, then negating any one of x's coordinates produces a vector that is also in the cap set.*

5.2 NAND and EQ Edges

Suppose we have an undirected graph where we label each edge as either NAND or EQ. We allow self-edges, but we do not allow two edges between the same pair of vertices.

The motivation for this sort of graph is the following generalization of the notion of an independent set.

Let G be a graph of the form described above. Let us associate boolean variables $x_1, \ldots, x_k$ with the vertices $v_1, \ldots, v_k$ of G. Then the edges can be interpreted as functions on these variables as follows. If there is an EQ edge between v_1 and v_2 this is the same as the function $x_1 \equiv x_2$. Similarly,

a NAND edge between them would be associated with the function $\neg(x_1 \wedge x_2)$. Then we can ask the natural question of how many variables we can set to true while all the edges (interpreted as functions) are satisfied. More formally, we give the following definition.

Definition 5.3. *Let G be a graph of the form described above. An **independent set** in G is defined as any set of vertices such that both the following two conditions hold.*

1. *There exist no two vertices in the set that have a NAND edge between them.*
2. *If a vertex is in the set, then every vertex that it has an EQ edge to is also in the set.*

Then the independence number of G denoted as $\alpha(G)$ is defined as the size of the largest independent set in G.

Then we can generalize the strong product as follows.

Definition 5.4. *Let G and H be graphs of the form described above. Then the **strong product** of G and H, denoted $G \boxtimes H$ is defined as follows.*

Its vertex set consists of the Cartesian product of the vertex sets of the original graphs. Suppose we have any pair of vertices (g_1, h_1) and (g_2, h_2) from the product graph.

There is an EQ edge between this pair of vertices if and only if both of the following two conditions hold.

1. *There is an EQ edge between g_1 and g_2 in G.*
2. *There is an EQ edge between h_1 and h_2 in H.*

There is a NAND edge between (g_1, h_1) and (g_2, h_2) if and only if all of the following conditions hold.

1. *There is an EQ or NAND edge between g_1 and g_2 in G.*
2. *There is an EQ or NAND edge between h_1 and h_2 in H.*
3. *At least one of the edges identified in the previous two conditions is a NAND edge.*

By expanding definitions, one can show that this product is associative. We can define a capacity for such graphs as follows

Definition 5.5. *Let G^n denote the strong product of G with itself a total of n times. Then the capacity of G is defined as $\lim_{n\to\infty} \alpha(G^n)$.*

This limit exists by a similar argument to that used in the case of the standard definition of capacity. In fact, this definition of capacity generalizes the standard definition of undirected graph capacity in a sense because we can convert any standard undirected graph to a graph of this form by making all edges NAND edges, then adding EQ self-edges to all vertices. The resulting graph will have the same capacity under this definition as the original graph has under the standard definition.

We now explain how to take a standard directed graph and convert it to a graph of the form considered in this section in a way that preserves a relationship between the capacities. Suppose we have a standard directed graph G. We convert it to a graph H of the form considered in this section as follows.

Definition 5.6. *Let G be a standard directed graph. define the function $s_1(G)$ as taking such a graph G and outputting a graph H of the form we are considering in this section as follows. For each vertex v in G, create two vertices v and v' in H with an EQ edge between v and v'. For each directed edge in G from u to v, create a NAND edge between u and v' in H. Define the function $s_2(G)$ as the same graph, except that each vertex has a self EQ edge.*

Notice that the graph constructed is bipartite. We now prove that these transformations give bounds on the capacity of the original graph. Specifically, we have the following.

Theorem 5.6. *For any standard directed graph G and $n \geq 1$,*

$$\alpha(s_2(G)^n) \leq 2^n \alpha(G^n)$$

Proof. We can partition the vertices in $s_2(G)^n$ into equivalence classes of size 2^n based on the vertex in G they correspond to under the following mapping. For any vertex $(v_1, \ldots, v_n)$ in $s_2(G)^n$, convert it to a vertex in G by replacing each v_i with a if $v_i = a'$ for some a. (In other words, go through the tuple and "unprime" each component.)

Notice that the EQ edges force us to take either every vertex in the a particular equivalence class or none of the vertices in that equivalence class.

Consider an independent set S in $s_2(G)^n$. We convert it to a set S' of size $|S|/2^n$ in G by applying the previously described mapping to each vertex.

We claim that S' is independent in G. To see this, suppose S' is not independent in G. Then there exist distinct vertices $(a_1, \ldots, a_n)$ and $(b_1, \ldots, b_n)$ in S' such that for each i, there is an edge from a_i to b_i in G. Consider the two vertices $(a_1, \ldots, a_n)$ and $(b'_1, \ldots, b'_n)$ in S. Then for all i, there is an edge between a_i and b'_i in $s_2(G)$. Thus, S is not independent which is a contradiction. □

Corollary 5.2. *The capacity of G is at least twice the capacity of $s_2(G)$.*

Theorem 5.7. *For any standard directed graph G and $n \geq 1$,*

$$\alpha(G^n) \leq 2\alpha(s_1(G)^n)$$

Proof. Consider any independent set S in G^n. For each vertex $(v_1, \ldots, v_n) \in S$, construct two distinct vertices $(v_1, \ldots, v_n)$ and $(v'_1, \ldots, v'_n)$ in $s_1(G)$. Let S' be the set of all such vertices in $s_1(G)$. Clearly, $|S'| = 2|S|$.

We claim S' is independent in G. To see this, suppose that S' is not independent in G. We satisfy all EQ edges by our construction, so there must be a NAND edge between two vertices in S'. By the way our graph product is defined, this means there exist two vertices $(a_1, \ldots, a_n)$ and $(b'_1, \ldots, b'_n)$ in S' such that for all i, either $a_i = b_i$ or a_i has a NAND edge to b'_i. This implies that there is an edge from $(a_1, \ldots, a_n)$ to $(b_1, \ldots, b_n)$ in G, contradicting the independence of S. □

Corollary 5.3. *The capacity of G is at most the capacity of $s_1(G)$.*

Bibliography

Alon, Noga, Amir Shpilka, and Christopher Umans. 2012. On Sunflowers and Matrix Multiplication. In *2012 IEEE 27th Conference on Computational Complexity*, 214–223. Porto, Portugal: IEEE. doi:10.1109/CCC.2012.26. URL http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=6243397.

Cohn, H., R. Kleinberg, B. Szegedy, and C. Umans. 2005. Group-theoretic Algorithms for Matrix Multiplication. In *FOCS '05 Proceedings of the 46th Annual IEEE Symposium on Foundations of Computer Science*, 379–388. IEEE. doi:10.1109/SFCS.2005.39. URL http://ieeexplore.ieee.org/xpl/downloadCitations http://ieeexplore.ieee.org/ielx5/10244/32664/01530730.pdf?tp=&arnumber=1530730&isnumber=32664.

Deuber, W.A, P Erdős, D.S Gunderson, A.V Kostochka, and A.G Meyer. 1997. Intersection Statements for Systems of Sets. *Journal of Combinatorial Theory, Series A* 79(1):118–132. doi:10.1006/jcta.1997.2778. URL http://dx.doi.org/10.1006/jcta.1997.2778.

Edel, Yves. 2004. Extensions of generalized product caps. *Designs, Codes and Cryptography* 31(1):5–14. URL http://www.springerlink.com/index/M55136X765161240.pdf.

Erdös, Paul, and Endre Szemerédi. 1978. Combinatorial properties of systems of sets. *Journal of Combinatorial Theory, Series A* 24(3):308–313. doi:10.1016/0097-3165(78)90060-2. URL http://dx.doi.org/10.1016/0097-3165(78)90060-2.

Havermann, Hans. 2009. A090245—OEIS. URL http://oeis.org/A090245.

Jukna, Stasys. 2011. *Extremal Combinatorics - With Applications in Computer Science*. Berlin: Springer Berlin / Heidelberg, 2nd ed. URL http://www.springer.com/computer/theoretical+computer+science/book/978-3-642-17363-9.

Knuth, Donald. 2001. SETSET-ALL. URL http://www-cs-faculty.stanford.edu/~uno/programs/setset-all.w.

Konc, Janez, and D Janezic. 2007. An improved branch and bound algorithm for the maximum clique problem. *proteins* 58:569–590. URL http://www.sicmm.org/konc/%C4%8CLANKI/MATCH58(3)569-590.pdf.

Lovász, L. 1979. On the Shannon capacity of a graph. *IEEE Transactions on Information Theory* 25(1):1–7. doi:10.1109/TIT.1979.1055985. URL http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1055985.

Meshulam, Roy. 1995. On subsets of finite abelian groups with no 3-term arithmetic progressions. *Journal of Combinatorial Theory, Series A* 71(1):168–172. doi:10.1016/0097-3165(95)90024-1. URL http://dx.doi.org/10.1016/0097-3165(95)90024-1.

Roth, K. F. 1953. On Certain Sets of Integers. *J London Math Soc* 28(1):104–109. URL http://jlms.oxfordjournals.org/content/s1-28/1/104.full.pdf.

Shannon, Claude. 1956. The zero error capacity of a noisy channel. *IEEE Transactions on Information Theory* 2(3):8–19. doi:10.1109/TIT.1956.1056798. URL http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=1056798.

Tao, Terence. 2007. Open question: best bounds for cap sets. URL http://terrytao.wordpress.com/2007/02/23/open-question-best-bounds-for-cap-sets/.

Tomita, Etsuji, Yoichi Sutani, and Takanori Higashi. 2010. A simple and faster branch-and-bound algorithm for finding a maximum clique. In *WALCOM: Algorithms and Computation*, 191–203. Springer Berlin / Heidelberg. URL http://www.springerlink.com/index/7778800225080533.pdf.

Index