Weizmann Institute of Science

Thesis for the degree
Master of Science

Submitted to the Scientific Council of the
Weizmann Institute of Science
Rehovot, Israel

A Different Perspective For Approximating Max Set Packing

Author:
Uri Sherman

Supervisor:
Prof. Uriel Feige

February 2014

Acknowledgments

First and foremost, I would like to express my deepest gratitude to my thesis advisor, Prof. Uriel Feige. His insights and ideas were inspiring to me, and provided me with the key elements that made this work possible.

I would also like to thank my fellow MSc candidates, For helpful discussions, and for making the experience more enjoyable altogether. For me, having good people around with which you can discuss ideas and other matters of research was an invaluable tool.

Abstract

Given an r-uniform hypergraph $\mathcal{H}(I, \mathcal{A})$, with $k, t \in \mathbb{N}$, consider the following problem: Does there exist a subset of hyperedges $\mathcal{S} \subseteq \mathcal{A}$ of size k, and an allocation function M such that $M(A) \subseteq A$ and $|M(A)| = t$ for all $A \in \mathcal{S}$, and furthermore $\{M(A)\}_{A \in \mathcal{S}}$ forms a packing. When $t = r$ this is exactly the famous set packing problem. In the optimization version of set packing, one asks to maximize k, while keeping $t = r$. We study a different optimization version of the problem, one where we ask to maximize t, and keep k as part of the input.

A hardness of approximation result of $1 - 1/e$ for the problem can be easily derived by a reduction from the well known max coverage problem. We show that a natural greedy algorithm obtains a 1/2-approximation guarantee, and no better for some instances. We formulate an LP relaxation and prove a tight integrality gap result of $1/2$ in the general case. For the special case in which the input hypergraph is linear, we prove the integrality gap tends to $1 - 1/e$ as r (the size of the hyperedges) tends to infinity.

Contents

1 Introduction

A hypergraph $\mathcal{H}$ is a pair $(I, \mathcal{A})$ where I is a set of elements, and $\mathcal{A}$ is a set of nonempty subsets of I called hyperedges. A hypergraph is said to be r-uniform if all hyperedges are of the same size r. Throughout this paper, we denote a hypergraph by $\mathcal{H}(I, \mathcal{A})$, reserving the notation $G(V, E)$ for graphs (which we assume the reader to be familiar with). In addition, we refer to members of the set I as items, while the term vertices is reserved for members of the vertex set V of a graph G. A *matching* in a hypergraph is a set of mutually disjoint hyperedges. A matching in G is a set of mutually disjoint edges.

Given any set of elements U and a family X of subsets of U, a *packing* is a subfamily $S \subseteq U$ of mutually disjoint subsets. A matching in a hypergraph $\mathcal{H}$ is just a special case of a packing where the subsets are hyperedges of $\mathcal{H}$.

Let $\mathcal{H}(I, \mathcal{A})$ be an r-uniform hypergraph, and let $\mathcal{S} \subseteq \mathcal{A}$ be some subset of hyperedges. An allocation of rank t for $\mathcal{S}$ is a mapping $M : \mathcal{S} \to P(I)$ (where $P(I)$ is the family of all subsets of I), such that for all $A \in \mathcal{S}$, $M(A) \subseteq A$, $|M(A)| = t$, and $\{M(A)\}_{A \in \mathcal{S}}$ forms a packing. Note that we do not say $\{M(A)\}_{A \in \mathcal{S}}$ is a matching in $\mathcal{H}$ since it might be that this set contains subsets that are not hyperedges in $\mathcal{H}$.

We say $\mathcal{S}$ is *a matching of rank* t if it admits an allocation of rank t but no allocation of rank $t + 1$. In sake of brevity, when the context is clear we simply say $\mathcal{S}$ is of rank t.

Given an r-uniform hypergraph $\mathcal{H}(I, \mathcal{A})$ and some integer $k \in \mathbb{N}$, the *Maximum Rank k-Matching* problem is that of selecting a subset $\mathcal{S} \subset \mathcal{A}$ of size k with maximum rank. We refer to this as the r-MRkM problem, and denote by $(I, \mathcal{A}, k)$ the input to the problem. We will see shortly that computing the rank of a given subset of hyperedges is a problem solvable in polynomial time, thus the heart of the r-MRkM problem is finding the right subset of hyperedges. An α-approximation algorithm (where $0 < \alpha \leq 1$) for the problem is an algorithm that produces a solution with rank at least α fraction of the optimal. We will mainly be focusing on a variant we call the *perfect* r-MRkM problem, in which we are assured there exists a solution of rank r, i.e. an actual matching of size k.

The work presented here deals with the algorithmic aspects of the r-MRkM problem. We believe this problem deserves attention as a somewhat natural combinatorial optimization problem stemming somewhere in the area between the well known **NP**-complete problems **Set Packing** and **Max Coverage**.

One way to think of the max rank k-matching problem is as a different maximization version of set packing. The maximum set packing problem, is that of selecting as many sets as possible while maintaining the property that they are all disjoint (that is, a packing). In the max rank k-matching problem, the size of the solution set is given as part of the input, while the quality of the packing is where we lose and what should be maximized.

In the max coverage problem, we are given a family of subsets and an integer $k \in \mathbb{N}$. The goal is to select k subsets with maximum coverage of elements given by their union. Obviously any solution of rank t to an r-MRkM instance is a solution with coverage $\geq kt$, and thus provides coverage of $\geq \frac{t}{r}$ fraction of the optimal. In a sense, max rank k-matching asks to maximize coverage while maintaining the property that each subset in the solution contributes the same amount of items. A $(1 - 1/e)$ hardness of approximation result for r-MRkM follows by a simple reduction from max coverage (see Appendix A.1).

To the best of our knowledge, the r-MRkM problem has not been studied previously (that is, not directly), but it can be viewed as a special case of a well studied problem called the **Max Min Allocation Problem** (*aka the Santa Claus problem*).

An instance of the max min allocation problem consists of a set of m agents, a set of n goods, and for each agent $i \in [m]$ and good $j \in [n]$, the valuation i gives j: $v_{ij} \in \mathbb{R}^+$. Given an allocation of goods to agents, that is, a partition $S_1, ..., S_m \subseteq [n]$ (meaning $\forall i, l \in [m]$, $S_i \cap S_l = \emptyset$ and $\bigcup_{i=1}^{m} S_i = [n]$), the utility of agent i is given by $u_i := \sum_{j \in S_i} v_{ij}$. The objective is finding an allocation for which $\min_{i \in [m]}\{u_i\}$ is maximized. The *restricted Santa Claus problem* is a special case of the same problem, in which each item $j \in [n]$ has an intrinsic value v_j, and each agent is either interested in it (and values it by this value), or not (values it with 0). More formally, $\forall j \in [n], i \in [m]$, $v_{ij} \in \{0, v_j\}$.

Given an r-MRkM instance $(I, \mathcal{A}, k)$, it can be viewed a special case of the restricted santa claus problem as follows. Identify hyperedges as agents and I as a tentative set of goods. Set $m := |\mathcal{A}|$, and add $m - k$ "dummy" goods for which all agents give valuation of ∞. As for the rest of the valuations, for each $i \in I$, $A \in \mathcal{A}$, the valuation that A gives i is 1 if $i \in A$ and 0 otherwise. Given a solution to the santa claus instance achieving minimal utility t, denote by $\mathcal{S}$ the subset of agents that did not receive any dummy items. Then $\mathcal{S}$ is of size k and thus is a solution with rank t to the r-MRkM instance.

1.1 Related Work

One line of research with direct relation to r-MRkM is that of the santa claus problem. Obviously, any positive results obtained for the santa claus problem would directly carry over to the r-MRkM problem, though unfortunately all such results are far from the quite trivial approximation guarantees we will present later on. As for the general santa claus setting, [CCK09] (following [BD05], [BS06] and [AS07]) give a $O(n^\epsilon)$-approximation algorithm that runs in polynomial time for any constant $\epsilon > 0$, and a $O(m^\epsilon)$-approximation algorithm that runs in quasi-polynomial time (n being the number of goods and m being the number of agents). These are the best positive results for the general case known to date.

Research focusing on the restricted santa claus version has produced stronger results. Asadpour, Feige and Saberi [AFS12] show that the $(1/4 - \epsilon, 1)$-gap problem can be solved in polynomial time for any $\epsilon > 0$. Subsequent work given in [PS12] also provides a quasi polynomial time algorithm that addresses the search problem and produces the actual solution. The first (and only to this date) constant factor polynomial time approximation algorithm for the search problem was established by Haeupler, Saha and Srinivasan [HSS11], relying on an earlier paper by Feige [Fei08].

The hardness results produced by this line of research (e.g. [BD05, BCG09]) show the restricted santa claus problem to be hard to approximate within any factor better than 1/2. These are all of the same nature and do not carry over to the r-MRkM problem.

The set packing problem, phrased in terms of hypergraphs, is that of finding a maximum matching in a given hypergraph. In the r-uniform version, all hyperedges are of the same size r. We note again that the objective in the max set packing problem is different than that of the r-MRkM problem, where the objective is to maximize the rank of the matching (rather than its size which is given as part of the input). A simple greedy algorithm for the max set packing problem can be easily shown to provide a $\frac{1}{r}$-approximation guarantee. Hurkens and Schrijver [HS89] present an improvement to this with a $\frac{2}{r}$-approximation algorithm, which is the best approximation factor known to date. Hazan, Safra and Schwartz [HSS06] prove the problem to be **NP**-hard to approximate within a factor of $O(\ln r/r)$. This improves upon an earlier hardness result by Trevisan [Tre01].

Another area of research in which our problem can be formulated, is that of Simplicial Complexes. A hypergraph $\mathcal{H}(I, \mathcal{A})$ is a simplicial complex if $A \in \mathcal{A} \Rightarrow B \in \mathcal{A}$ for all $B \subseteq A$. These structures have a strong geometric interpretation, and we refer to hyperedges of size j as simplices, or faces, of dimension $j-1$. A *pure* simplicial r-complex is one in which all simplices of maximal dimension (meaning they are not contained in any larger simplex) are of dimension $r-1$. Consider the following algorithmic problem: Given a pure simplicial r-complex $\mathcal{H}(I, \mathcal{A})$ that contains $k \in \mathbb{N}$ disjoint simplices of dimension $r-1$, find a set of k disjoint simplices of dimension t, where t is to be maximized. This problem is equivalent to the perfect r-MRkM problem, see Appendix A.2 for details.

1.2 Basic Observations

Proposition 1.1. *The perfect r-MRkM problem is* **NP***-hard to approximate within any factor better than* $1-1/e$*.*

The proof is by a simple reduction from the max coverage problem, and is given in Appendix A.1. Next, we note a few special cases in which the perfect r-MRkM problem can be solved efficiently. For the 2-uniform case, the problem reduces to finding a maximum matching in a simple graph.

Proposition 1.2. *The perfect* 2*-MRkM problem can be solved in polynomial time.*

Proof. In this case the hyperedges can be viewed as edges in a simple graph, and we can find a maximum matching using known polynomial time algorithms (e.g. [Edm65]). □

A *d-regular* hypergraph is one in which every item is contained in exactly d hyperedges. For 2-regular uniform hypergraphs we are able to find an actual matching if it covers all items:

Proposition 1.3. *There exists a polynomial time algorithm that takes as input a* 2*-regular* r*-uniform hypergraph* $\mathcal{H}(I, \mathcal{A})$*, and outputs a matching that covers all items, if one exists.*

Proof. Define an auxiliary graph G where the vertices are hyperedges of $\mathcal{H}$, and two vertices are connected with an edge if their corresponding hyperedges share an item in $\mathcal{H}$. Check if G is two-colorable. If so, then each color class forms a matching in $\mathcal{H}$, and both will be of size $\frac{|I|}{r} = \frac{|\mathcal{A}|}{2}$, hence either of them is a solution as desired. Conversely, if there exists a matching in $\mathcal{H}$ covering all items, it is easily verified that coloring its corresponding vertices in G with one color class, and the rest of the vertices (which actually must also be a solution) with another forms a two-coloring in G. □

For the 3-uniform case, we are able to find a matching of rank 2, again using as a subroutine an algorithm that finds a maximum matching in a simple graph.

Proposition 1.4. *Given a perfect* 3*-MRkM instance* $(I, \mathcal{A}, k)$*, we can find in polynomial time a solution of rank* 2*, while finding the perfect matching is* **NP***-hard.*

Proof. Let $(I, \mathcal{A}, k)$ be an instance of the 3-MRkM problem. Construct the auxiliary graph $G(V, E)$ where V is the set of items I, and E is the set of all subsets of size 2 of hyperedges in the original hypergraph $\mathcal{H}$. Now find a maximum matching in G. Since $\mathcal{H}$ is known to contain a matching of size k, it follows there must be a matching of size $\geq k$ in G, and thus the maximum matching algorithm will find it. Every two edges in G that originated from the same hyperedge in $\mathcal{H}$ intersect, thus the matching we have found in G can be uniquely identified with a matching of rank 2 in $\mathcal{H}$, and we are done.

Hardness follows by a reduction from the **NP**-hard **3-Dimensional Matching** problem (see [GJ90] for definition and hardness of 3DM). □

1.3 Our Results

We present a natural greedy algorithm for the uniform max rank k-matching problem, *Greedy-r-MRkM*, and prove a tight approximation guarantee (Section 3).

Theorem 1.5. *Algorithm Greedy-r-MRkM is a $\frac{1}{2}$-approximation algorithm for the r-MRkM problem, and the approximation guarantee is tight. That is, there exist instances for which Greedy-r-MRkM produces a solution with rank exactly $\frac{1}{2}$ of the optimal.*

In Section 4 we prove an existence result for bounded degree hypergraphs. The proof is constructive and is based on a polynomial time combinatorial algorithm. In section 5, amongst other results, we show the result stated in this theorem to be tight.

Theorem 1.6. *Let $\mathcal{H}(I, \mathcal{A})$ be an r-uniform hypergraph of bounded degree d (meaning all items are contained in at most d hyperedges). Then there exists $\mathcal{S} \subseteq \mathcal{A}$, $|\mathcal{S}| \geq \frac{|\mathcal{A}|}{d}$, with rank $\geq \max\left\{\left\lceil \frac{r}{2} \right\rceil, \left\lfloor \frac{d+1}{2d} r \right\rfloor\right\}$. Moreover, this set can be found in polynomial time.*

In Section 5, we consider a certain linear programming relaxation of the r-MRkM problem. We prove a tight integrality gap result.

Theorem 1.7. *The integrality gap of our LP relaxation is exactly $1/2$:*

1. *Let $(I, \mathcal{A}, k)$ be an r-MRkM instance for which our LP relaxation is feasible. Then there exists a set of hyperedges of size k with rank $\geq r/2$. Moreover, this set can be found in deterministic polynomial time.*

2. *There exist r-MRkM instances $(I, \mathcal{A}, k)$ such that our LP relaxation is feasible, yet any set of hyperedges of size k has rank $\leq r/2$.*

For r-MRkM instances in which the input hypergraph is *linear* (meaning every two hyperedges share at most one item), we devise a randomized rounding procedure for the LP relaxation establishing an approximation guarantee that approaches $(1 - 1/e)$ as r tends to infinity. We prove this in Section 6 and show this result to be nearly tight.

Theorem 1.8.

1. *Let $(I, \mathcal{A}, k)$ be a linear r-MRkM instance. If the LP relaxation is feasible, then there exists an integral solution of rank $\lfloor (1 - 1/e - \eta)r \rfloor$ where $\eta = O\left(\sqrt{\frac{\ln r}{r}}\right)$, and it can be found in expected polynomial time.*

2. *For any $\epsilon > 0$, there exist linear r-MRkM instances for which the LP relaxation is feasible, yet any integral solution has rank $< (1 - 1/e + \epsilon)r$.*

2 Preliminaries

In this section we record some simple results and give basic definitions which will be used throughout this work.

2.1 A Generalized Hall Condition

We prove a simple generalization of Hall's marriage theorem. For a graph $G(V, E)$, $v \in V$, denote by $N(v)$ the neighbor set of v. For $S \subseteq V$ denote by $N(S) = \cup_{v \in S} N(v)$ the neighbor set of all vertices in S.

Now, let $G(A \cup B, E)$ be a bipartite graph. Our goal is to allocate for each $v \in A$ some number $k_v \leq |N(v)|$ of his neighbors, so that no $b \in B$ is allocated more than once. More formally, we look for a subset of edges so that for all $v \in A$, $d(v) = k_v$, and for all $b \in B$, $d(b) \leq 1$.

Lemma 2.1. *A necessary and sufficient condition for the existence of an allocation as described above is that for any subset $S \subseteq A$, $|N(S)| \geq \sum_{v \in S} k_v$.*

Proof. The necessity of the condition is quite obvious - assume $M \subseteq E$ is an allocation with $d(v) \geq k_v$ for all $v \in A$, and let $N_M(v)$ denote the neighbor set of v under the allocation M. Since for any $v, u \in A$, $N_M(v) \cap N_M(u) = \emptyset$, it follows that for any $S \subseteq A$,

$$|N(S)| = \sum_{v \in S} |N_M(v)| \geq \sum_{v \in S} k_v$$

as desired.

Conversely, let us assume the condition holds, and we will show the existence of an allocation. Define an auxiliary graph $G'(A' \cup B', E')$ as follows - for each $v \in A$ we create k_v copies $Q(v) := \{v_j : 1 \leq j \leq k_v\}$, and set $A' := \cup_{v \in A} Q(v)$, $B' := B$, and for all $v_j \in A', t \in B'$,

$$(v_j, t) \in E' \iff (v, t) \in E$$

Let $S' \subseteq A'$, and denote $S := \{v \in A : Q(v) \cap S' \neq \emptyset\}$. Observe that $N(S') = N(S)$, and $S' \subseteq \bigcup_{v \in S} Q(v)$, hence

$$|N(S')| = |N(S)| \geq \sum_{v \in S} k_v = \sum_{v \in S} |Q(v)| \geq |S'|$$

This establishes that Hall's condition holds for G', and it follows there exists an A' saturating matching - denote it by $M' \subseteq E'$. Define

$$M := \bigcup_{v \in A} \bigcup_{v_j \in Q(v)} \{(v, w) : (v_j, w) \in M'\}$$

Let $w \in B$, assume there exist $v, u \in A$, $v \neq u$, but $(v, w) \in M$ and $(u, w) \in M$. Then there would exist $x, y \in A'$ such that $(x, w) \in M'$ and $(y, w) \in M'$ which is an obvious contradiction. Observe that for $v \in A$,

$$|N_M(v)| = |\bigcup_{v_j \in Q(v)} \{w : (v_j, w) \in M'\}| = |Q(v)| = k_v$$

It now follows that M is an allocation as desired. $\square$

Corollary 2.2. *Let $G(A \cup B, E)$ be a bipartite graph, with a weight function on the edges $w : E \to [0,1]$ satisfying for all $b \in B$:*

$$\sum_{e:b\in e} w(e) \leq 1$$

Denote for $v \in A$, $f(v) := \sum_{e:v\in e} w(e)$. Then there exists an allocation such that each $v \in A$ is allocated $\lfloor f(v) \rfloor$ of his neighbors.

Proof. We show the generalized hall condition holds. Let $S \subseteq A$, then

$$\sum_{v\in S} f(v) = \sum_{v\in S} \sum_{e:v\in e} w(e) = \sum_{b\in N(S)} \sum_{v\in S:(v,b)\in E} w(v,b) \leq \sum_{b\in N(S)} 1 = |N(S)|$$

$$\Longrightarrow |N(S)| \geq \sum_{v\in S} \lfloor f(v) \rfloor$$

$\square$

2.2 The Rank of Allocations and Matchings

Consider an r-uniform hypergraph $\mathcal{H}(I, \mathcal{A})$, and let $\mathcal{S} \subseteq \mathcal{A}$ be some subset of hyperedges. We will denote by $I(\mathcal{S}) := \cup_{A\in\mathcal{S}} A$ the set of items which are contained in some hyperedge from $\mathcal{S}$.

An *allocation* $M : \mathcal{S} \longrightarrow P(I)$ (where $P(I)$ is the family of all subsets of I) assigns each hyperedge some subset of its items.

Definition 2.3. *$M : \mathcal{S} \longrightarrow P(I)$ is an allocation if*

1. $\forall A \in \mathcal{S}, M(A) \subseteq A$
2. $\forall A, B \in \mathcal{S}, A \neq B, M(A) \cap M(B) = \emptyset$

Note that this is analogous to an allocation of the form discussed in Lemma 2.1 in the bipartite graph $G(\mathcal{S} \cup I, E)$ where $(A, i) \in E$ if $i \in A$. Let us denote by $M(\mathcal{S}, I)$ the set of all allocations of I to $\mathcal{S}$.

Definition 2.4. *The rank of an allocation $M : \mathcal{S} \longrightarrow P(I)$ is*

$$\phi(M) := \min_{A\in\mathcal{S}} \{|M(A)|\}$$

We say an allocation $M^* \in M(\mathcal{S}, I)$ is *fair* or *optimal* if $\phi(M^*) \geq \phi(M)$ for all $M \in M(\mathcal{S}, I)$. The rank of a subset of hyperedges $\mathcal{S} \subseteq \mathcal{A}$ is given now by the rank of a fair allocation for $\mathcal{S}$:

Definition 2.5. *The rank of $\mathcal{S} \subseteq \mathcal{A}$ is defined by*

$$\phi(\mathcal{S}) := \max_{M\in M(S,I)} \{\phi(M)\}$$

A *matching* $\mathcal{S} \subseteq \mathcal{A}$ is a subset of mutually disjoint hyperedges. A *matching of rank* t is a subset of hyperedges $\mathcal{S} \subseteq \mathcal{A}$ such that $\phi(\mathcal{S}) = t$. Note that an actual matching in an r-uniform hypergraph is a matching of rank r. Loosely speaking, a matching of rank t consists of mutually disjoint sub-hyperedges of size t. As an immediate corollary of Lemma 2.1 we have that

$$\phi(\mathcal{S}) \geq t \iff |I(\mathcal{T})| \geq |\mathcal{T}|t \quad \forall \mathcal{T} \subseteq \mathcal{S}$$

Observe that the rank of a matching is computable in polynomial time:

Proposition 2.6. *Let $\mathcal{H}(I, \mathcal{A})$ be an r-uniform hypergraph and let $\mathcal{S} \subseteq \mathcal{A}$. Then $\phi(\mathcal{S})$ is computable in polynomial time, as well as $M \in M(\mathcal{S}, I)$ for which $\phi(M) = \phi(\mathcal{S})$.*

Proof. Consider an algorithm mimicking the proof of Lemma 2.1 as a subroutine: Given a target value $t \in \mathbb{N}$, we can find an allocation of rank t if one exists. Identify the hypergraph $\mathcal{H}(I, \mathcal{A})$ with the bipartite graph $G(V, E)$ where $V := \mathcal{A} \cup I$ and $(A, i) \in E$ whenever $i \in A$. Construct the auxiliary graph $G'(\mathcal{A}' \cup I', E')$ as defined in Lemma 2.1, with t copies for each $A \in \mathcal{A}$. Now look for an $\mathcal{A}'$ saturating matching using known algorithms (see for example [KT05] Section 7.5).

Performing binary search on possible values for $t \in \{0, .., r\}$ provides the claimed algorithm. □

We will also be working with *fractional* allocations. For each hyperedge, a fractional allocation assigns a vector indexed by the item set. Each vector component has values in the range $[0, 1]$. An integral allocation can be thought of as a fractional allocation with vector components having values in the set $\{0, 1\}$.

Definition 2.7. $F : \mathcal{S} \longrightarrow [0, 1]^I$ *is a fractional allocation if*

1. $\forall A \in \mathcal{A}$, $\forall i \notin A$, $F(A)_i = 0$
2. $\forall i \in I, \sum_{A \in \mathcal{A}} F(A)_i \leq 1$

Let $M^f(\mathcal{S}, I)$ denote the set of all fractional allocations of I to $\mathcal{S}$. For $A \in \mathcal{S}$, $F \in M^f(\mathcal{S}, I)$, define $|F(A)| := \sum_{i \in A} F(A)_i$. The definition for the rank of F follows now similarly to that of an integral allocation.

Observe that by Corollary 2.2, if $M \in M(\mathcal{S}, I)$ and $F \in M^f(\mathcal{S}, I)$ are both be fair (e.g. of maximal rank), then

$$\lfloor \phi(F) \rfloor \leq \phi(M) \leq \phi(F)$$

3 The Greedy Approach

In this section we examine the performance of a few greedy algorithms for the r-MRkM problem.

3.1 The Max Coverage Greedy Algorithm

We begin by examining the well known greedy algorithm for the max coverage problem. Recall that for a hypergraph $\mathcal{H}(I, \mathcal{A})$ and a subset of hyperedges $\mathcal{S} \subseteq \mathcal{A}$, we denote $I(\mathcal{S}) := \cup_{T \in \mathcal{S}} T$.

Greedy Max Coverage$(I, \mathcal{A}, k)$

1. Set $\mathcal{S} \leftarrow \emptyset$
2. For $j = 1..k$:
   (a) Choose the hyperedge $A \in \mathcal{A}$ that maximizes $|I(\mathcal{S}) \cup A|$
   (b) Set $\mathcal{S} \leftarrow \mathcal{S} \cup \{A\}$
3. Return $\mathcal{S}$

Although it is well known *Greedy Max Coverage* achieves a $(1 - 1/e)$ approximation ratio for the max coverage problem, it is quite easily seen it actually provides no constant approximation guarantee for the r-MRkM problem.

Claim 3.1. *Let $r \in \mathbb{N}$, $L < r$. There exists an r-MRkM instance for which Greedy Max Coverage produces a solution with rank $\leq \frac{r}{L} + L$, while the optimal solution has rank r.*

Proof. Set $I := \{(i,j)\}_{1 \leq i,j \leq r}$. The set of "good" hyperedges which comprise a perfect solution is given by $A_l := \{(i,l)\}_{1 \leq i \leq r}$ where $1 \leq l \leq r$. The set of bad hyperedges which the algorithm will pick as the solution is given by

$$B_l := \{(l,i)\}_{1 \leq i \leq r} \quad \text{for } 1 \leq l \leq r - L + 1$$

and

$$B_l := \{(r-L+1,i)\}_{1 \leq i \leq r-L+1} \bigcup \{(l,i)\}_{r-L+2 \leq i \leq r} \quad \text{for } r - L + 2 \leq l \leq r$$

Figure 1 gives an illustration of our construction.

Assume the algorithm picks B_1 as the first hyperedge. Then the next $r-L-1$ hyperedges the algorithm will choose are $B_2, .., B_{r-L-1}$, since they are all disjoint (and we assume ties are broken in their favor). At this point, every good hyperedge A_l will provide additional coverage of $L-1$ items: $\{(i,l)\}_{r-L+2 \leq i \leq r}$. Every bad hyperedge B_l where $r - L + 2 \leq l \leq r$ will contribute its $L-1$ items $\{(l,i)\}_{r-L+2 \leq i \leq r}$. Again, since we assume that ties are broken in the favor of the bad hyperedges it follows the set produced by the algorithm is the set of bad hyperedges $B_1, ..., B_r$.

Observe that the last L bad hyperedges $B_{L+1}..B_r$ all share the same $r - L + 1$ items $\{(r-L+1,i)\}_{1 \leq i \leq r-L+1}$. Thus, in any allocation one of these hyperedges will be given no more than $\frac{r-L+1}{L} + L - 1 = \frac{r}{L} + \frac{1}{L} + L - 2 < \frac{r}{L} + L$ of its items as desired. $\square$

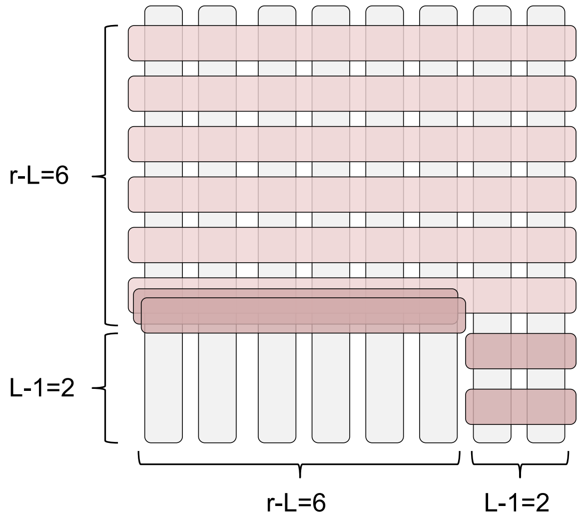

Figure 1: The construction for $r = 8, L = 3$

Set $L := \sqrt{r}$, and it follows that the solution produced by *Greedy Max Coverage* has rank $\leq 2\sqrt{r}$, meaning it is at best a $\frac{2}{\sqrt{r}}$-approximation algorithm for the r-MRkM problem.

3.2 A Max Rank Matching Greedy Algorithm

We present here a natural greedy algorithm for the r-MRkM problem, prove it provides a $\frac{1}{2}$ approximation guarantee, and show this guarantee is tight. That is, we present a family of instances for which the algorithm does not achieve an approximation better than $\frac{1}{2}$. This is the only result we prove for the non-perfect r-MRkM problem. Meaning, we have that the approximation guarantee holds for arbitrary instances which may have optimal solutions of rank $< r$.

Greedy-r-MRkM $(I, \mathcal{A}, k)$

1. Set $\mathcal{S} \leftarrow \emptyset$.
2. For $j = 1..k$:
   (a) For every $A \in \mathcal{A} \setminus \mathcal{S}$ compute $\phi(\mathcal{S} \cup \{A\})$
   (b) Choose $A \in \mathcal{A}$ achieving the maximum rank (with ties broken arbitrarily) and set $\mathcal{S} \leftarrow \mathcal{S} \cup \{A\}$.
3. Return $\mathcal{S}$

To prove this is a $\frac{1}{2}$-approximation algorithm for the uniform max rank k-matching problem, we begin with the following lemma.

Lemma 3.2. *Let $(I, \mathcal{A}, k)$ be an instance of the r-MRkM problem, in which there exists $\mathcal{S}^* \subseteq A$ with $\phi(\mathcal{S}^*) = t$. Then for any $\mathcal{S} \subseteq \mathcal{A}$, $|\mathcal{S}| < k$ with $\phi(\mathcal{S}) \geq \frac{t}{2}$, there exists $A \in \mathcal{A} \setminus \mathcal{S}$ such that $\phi(\mathcal{S} \cup \{A\}) \geq \frac{t}{2}$.*

Proof. If $\mathcal{S} \subseteq \mathcal{S}^*$ then the claim follows by adding any hyperedge $A \in \mathcal{S}^* \setminus \mathcal{S}$ (there must exist such a hyperedge since $|\mathcal{S}| < k = |\mathcal{S}^*|$). Otherwise, set $\mathcal{S}' := \mathcal{S} \setminus \mathcal{S}^* \neq \emptyset$, $\mathcal{S}_0^* := \mathcal{S}^* \setminus \mathcal{S}$, and $\mathcal{T} := \mathcal{S} \cap \mathcal{S}^*$. Observe

$$|\mathcal{S}'| = |\mathcal{S}| - |\mathcal{S}^* \cap \mathcal{S}| < |\mathcal{S}^*| - |\mathcal{S}^* \cap \mathcal{S}| = |S_0^*|$$

Fix $M^* \in M(\mathcal{S}^*, I)$ such that $|M^*(A)| = t$ for all $A \in \mathcal{S}^*$. What we will want now is to assert the existence of an allocation $M \in M(\mathcal{S}, I)$ that assigns a large portion of the items assigned by M^* to hyperedges in $\mathcal{T}$. Denote by $I(M) := \cup_{A \in \mathcal{S}} M(A)$ the set of all items assigned by M.

Claim 3.3. *There exists $M \in M(\mathcal{S}, I)$ assigning exactly $\lceil \frac{t}{2} \rceil$ items to each $A \in \mathcal{S}$, with the property that $|I(M) \cap M^*(A)| \geq \frac{t}{2}$ for all $A \in \mathcal{T}$.*

Proof. Let $M \in M(\mathcal{S}, I)$ be an allocation assigning exactly $\lceil \frac{t}{2} \rceil$ items to each $A \in \mathcal{S}$ (the existence of which is assured by our assumption that $\phi(\mathcal{S}) \geq \frac{t}{2}$). We give a procedure that transforms M into an allocation as desired. As long as there exists $A \in \mathcal{T}$ with $|I(M) \cap M^*(A)| < \frac{t}{2}$, release an allocated item $i \in M(A) \setminus M^*(A)$, and assign to A some other item $j \in M^*(A) \setminus I(M)$, i.e. $M(A) \leftarrow (M(A) \setminus \{i\}) \cup \{j\}$.

The existence of the item j is obvious, and there exists such an item i since A is allocated $\lceil \frac{t}{2} \rceil$ items, while M is using $< \frac{t}{2}$ items from $M^*(A)$. Observe that doing this strictly increases the amount of items A is allocated from the set $M^*(A)$, and does not affect the allocation of any hyperedge other than A. Hence, since we maintain an allocation of exactly $\lceil \frac{t}{2} \rceil$ to all hyperedges throughout, we will repeat the described step no more than $|\mathcal{T}| \lceil \frac{t}{2} \rceil$ times, meaning the procedure comes to an end after a finite amount of steps. □

Let $M \in M(\mathcal{S}, I)$ be such an allocation from the statement of the claim. Denote $I_{\mathcal{S}_0^*} := \cup_{T \in \mathcal{S}_0^*} M^*(T)$, and $I_{\mathcal{T}} := \cup_{T \in \mathcal{T}} M^*(T)$. Then $|I(M) \cap I_{\mathcal{T}}| \geq |\mathcal{T}| \frac{t}{2}$, hence

$$|I(M) \cap I_{\mathcal{S}_0^*}| \leq |\mathcal{S}| \frac{t}{2} - |\mathcal{T}| \frac{t}{2} = |\mathcal{S}'| \frac{t}{2}$$

This implies

$$|I_{\mathcal{S}_0^*} \setminus I(M)| \geq |\mathcal{S}_0^*| t - |\mathcal{S}'| \frac{t}{2} > |\mathcal{S}_0^*| \frac{t}{2}$$

Therefore, by an averaging argument it follows there must exist a hyperedge $A \in \mathcal{S}_0^*$ that contains $> \frac{t}{2}$ items that were not allocated at all by M, and we are done. □

Corollary 3.4. *Algorithm Greedy-r-MRkM is a $\frac{1}{2}$-approximation algorithm for the r-MRkM problem.*

Proof. Consider an r-MRkM instance $(I, \mathcal{A}, k)$. Let *opt* be the rank of the optimal solution, and let j denote the iteration index during the algorithm's execution. Correctness of approximation ratio follows by induction on $1 \leq j \leq k$. The base case is obvious, since a set consisting of a single hyperedge has rank r. When the algorithm is in iteration $2 \leq j \leq k$, by the inductive hypothesis $\phi(S) \geq \frac{opt}{2}$ at the beginning of the iteration. Since $|\mathcal{S}| = j - 1 < k$, by Lemma 3.2 it follows there must exist a hyperedge which we can add and maintain this rank lower bound. Obviously, a hyperedge maximizing the rank of the new set is such a hyperedge, and $\phi(\mathcal{S}) \geq \frac{opt}{2}$ at the end of the iteration as desired. □

We now turn to show the analysis we have given is tight.

Proposition 3.5. *Let $r \geq 2$ be even. There exists an instance of the r-MRkM problem for which Greedy-r-MRkM produces a set with rank $\leq \frac{r}{2}$, while the optimal solution has rank r.*

Proof. Set $X := \bigcup_{j=1}^{r} \{(j, l) : j \leq l \leq r\}$. This is the set of "problematic" items we will use.

$$\begin{array}{lllll}
(1,1) & & & & \\
(1,2) & (2,2) & & & \\
(1,3) & (2,3) & (3,3) & & \\
\vdots & \vdots & \vdots & \ddots & \\
(1,r) & (2,r) & (3,r) & \cdots & (r,r)
\end{array}$$

Let $(x_1, x_2, ..., x_R)$ be a lexicographic ordering of the items in X defined above (and $R = \frac{r(r+1)}{2}$).

The set of bad hyperedges we plan the algorithm to choose is given by

$$B_j := \{(i,l) \in X : i = j \text{ or } l = j\} \qquad 1 \leq j \leq r$$

Figure 2 illustrates the bad hyperedges in the $r = 4$ case.

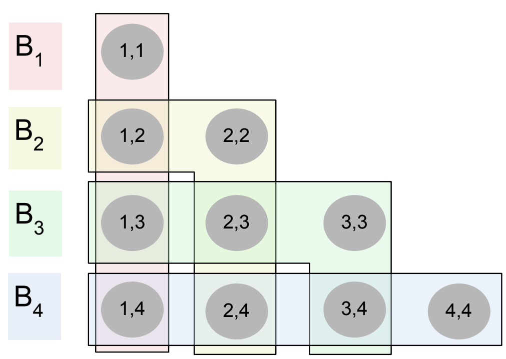

Figure 2: Bad hyperedges for $r = 4$

The set of good hyperedges comprising the perfect solution (i.e. of rank r) is

$$A_j := \{x_l \in X : l = j \mod r\} \cup P_j \qquad 1 \leq j \leq r$$

where P_j is a set of $r - |\{x_l \in X : l = j \mod r\}|$ items private to hyperedge A_j (private in the sense that they are not contained in any other hyperedge, good or bad). It is easily verified that $|\{x_l \in X : l = j \mod r\}| \leq r$ for all $1 \leq j \leq r$. Figure 3 illustrates the good hyperedges in the problematic items region.

Denote the set of bad hyperedges by $\mathcal{B}$, and the set of good hyperedges by $\mathcal{G}$.

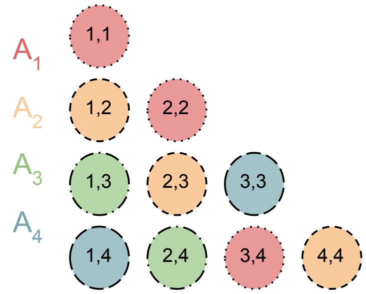

Figure 3: Good hyperedges in the problematic region, $r = 4$

Claim 3.6. *Let $\mathcal{S} \subseteq \mathcal{B}$, $|\mathcal{S}| = j$. Then $|I(\mathcal{S})| = j(r - \frac{j-1}{2})$ and $\phi(\mathcal{S}) = \left\lfloor r - \frac{j-1}{2} \right\rfloor$.*

Proof. Observe that every two hyperedges in $\mathcal{B}$ share exactly one item, and that this item is exclusive to them - no other hyperedge in $\mathcal{B}$ contains this item. So, every hyperedge in $\mathcal{S}$ has exactly $j-1$ items shared with other hyperedges in $\mathcal{S}$, thus $|I(\mathcal{S})| = jr - \binom{j}{2} = j(r - \frac{j-1}{2})$.

Define the following fractional allocation; for each hyperedge assign half of each one of its shared items, and all of his non-shared items. It is easily seen this is a legal fractional allocation that assigns exactly $r - \frac{j-1}{2}$ items to each hyperedge in $\mathcal{S}$. Moreover, this allocation assigns all of the items in $I(\mathcal{S})$ completely (there are no leftovers).

Obviously, any integral allocation can do only worse than this, thus $\phi(\mathcal{S}) \leq \left\lfloor r - \frac{j-1}{2} \right\rfloor$. Conversely, by Corollary 2.2 it follows that $\phi(S) \geq \left\lfloor r - \frac{j-1}{2} \right\rfloor$, and the claim follows. $\square$

Claim 3.7. *Let $1 \leq j \leq r$, $\mathcal{S} := \{B_1, .., B_j\}$. Then for any $A \in \mathcal{G}$, $|A \cap I(\mathcal{S})| \geq \left\lceil \frac{j}{2} \right\rceil$*

Proof. Set $J := |I(\mathcal{S})| = j\left(r - \frac{j-1}{2}\right)$, and observe that $I(\mathcal{S})$ consists of the first J items by the lexicographic order, meaning $I(\mathcal{S}) = \{x_1, .., x_J\}$.

Now let $1 \leq l \leq r$, then

$$|A_l \cap I(\mathcal{S})| = |\{x_l, x_{l+r}, ..., x_{l+cr} : l + cr \leq J\}| = \max_{l+(c-1)r \leq J} \{c\}$$

We want to show that for $c := \frac{j+1}{2}$, $l + (c-1)r \leq J$. Indeed

$$l + \left(\frac{j+1}{2} - 1\right) r \leq r + \frac{j-1}{2} r = \frac{j+1}{2} r$$

and

$$\begin{aligned} & \frac{j+1}{2} r \leq J \\ \iff \quad & \frac{j+1}{2} r \leq jr - \frac{j(j-1)}{2} \\ \iff \quad & \frac{j(j-1)}{2} \leq r\left(j - \frac{j+1}{2}\right) = \frac{j-1}{2} r \end{aligned}$$

where the last inequality holds since $j \le r$. Hence $|A_l \cap I(\mathcal{S})| \ge \frac{j+1}{2} \ge \left\lceil \frac{j}{2} \right\rceil$ as desired. $\square$

Assume the ties during execution are always broken in favor of the bad hyperedges by the order we have given them (thus the first hyperedge chosen will be B_1). We will now inductively prove that at the end of iteration j, $\mathcal{S} = \{B_1, ..., B_j\}$. The base case follows from our assumption. Let $2 \le j \le r$, and denote by $\mathcal{S}$ the solution set at the beginning of the iteration. By the inductive hypothesis we have that $\mathcal{S} = \{B_1, ..., B_{j-1}\}$. If j is odd, by Claim 3.6

$$\phi(\mathcal{S} \cup \{B_j\}) = \left\lfloor r - \frac{j-1}{2} \right\rfloor = r - \frac{j-1}{2}$$

By the same claim,

$$\phi(\mathcal{S}) = \left\lfloor r - \frac{j-2}{2} \right\rfloor = r - \frac{j-1}{2}$$

Since at the previous iteration B_{j-1} was chosen to be added to $\mathcal{S}$, it follows that for all $A \in \mathcal{G}$,

$$\phi(\mathcal{S} \cup \{A\}) \le \phi((\mathcal{S} \setminus \{B_{j-1})\} \cup \{A\}) \le \phi(\mathcal{S}) = r - \frac{j-1}{2}$$

Given our assumption on the way ties are broken, it follows that B_j will be chosen by the algorithm to be added to $\mathcal{S}$.

Now assume j is even. By Claim 3.6,

$$\phi(\mathcal{S} \cup \{B_j\}) = \left\lfloor r - \frac{j-1}{2} \right\rfloor = r - \frac{j}{2}$$

So, to complete the proof we want to show that $\phi(\mathcal{S} \cup \{A\}) \le r - \frac{j}{2}$ for all $A \in \mathcal{G}$.

Let $A \in \mathcal{G}$, and let $M^* \in M(\mathcal{S} \cup \{A\}, I)$ be a fair allocation. If $|M^*(A)| \le r - \frac{j}{2}$, we are done. Otherwise, since Claim 3.7 ensures us that $|A \cap I(\mathcal{S})| \ge \left\lceil \frac{j}{2} \right\rceil = \frac{j}{2}$, it must be that $M(G) \cap I(\mathcal{S}) \neq \emptyset$. But if this is the case, the allocation of I to $\mathcal{S}$ induced by M^* is an allocation of a proper subset of $I(\mathcal{S})$ to $\mathcal{S}$.

Observe that since $|I(\mathcal{S})| = (j-1)\left(r - \frac{j-2}{2}\right)$, any allocation of a proper subset of $I(\mathcal{S})$ to $\mathcal{S}$ would have to assign $< r - \frac{j-2}{2}$ to some hyperedge (because the average number of items per hyperedge is $< r - \frac{j-2}{2}$ in such a case). So, in any allocation of this sort there will be a hyperedge that is given no more than $r - \frac{j}{2}$ items, and we are done. $\square$

This concludes the proof of Theorem 1.5, restated here for completeness:

Theorem 1.5. *Algorithm Greedy-r-MRkM is a $\frac{1}{2}$-approximation algorithm for the r-MRkM problem, and the approximation guarantee is tight. That is, there exist instances for which Greedy-r-MRkM produces a solution with rank exactly $\frac{1}{2}$ of the optimal.*

Remark 3.8. *A plausible improvement for Greedy-r-MRkM would be to break ties in favor of the hyperedge providing maximal coverage, rather than arbitrarily. We have not been able to prove a better approximation guarantee, nor have we found an example showing the $\frac{1}{2}$-approximation given in the analysis is tight. The question of the improved algorithm's approximation guarantee is left as an open problem.*

3.3 A Linear Time Greedy Algorithm

We give here a linear time 1/2-approximation algorithm, for the perfect r-MRkM problem (in which we are assured there exists a solution which is an actual matching). The algorithm gradually constructs the allocation of items at each iteration, avoiding the computation of fair allocations altogether.

Lin-Greedy-r-MRkM $(I, \mathcal{A}, k)$

1. Give the hyperedges some arbitrary order $A_1, ..., A_{|\mathcal{A}|}$
2. Set $\mathcal{S} \leftarrow \emptyset$, $M \in M(\mathcal{S}, I)$.
3. For $j = 1..|\mathcal{A}|$:
   (a) Let $D := \cup_{T\in\mathcal{S}} M(T)$ (the subset of allocated items)
   (b) If $|A_j \setminus D| \geq \frac{r}{2}$, proceed to next step, otherwise continue to next j
   (c) Set $A'_j \subseteq A_j \setminus D$ to be an arbitrary choice of $\lceil \frac{r}{2} \rceil$ items from $A_j \setminus D$
   (d) $\mathcal{S} \leftarrow \mathcal{S} \cup \{A_j\}$, $M(A_j) := A'_j$
4. Return $\mathcal{S}$

It is easily seen *Lin-Greedy-r-MRkM* runs in linear time, since we have $|\mathcal{A}|$ iterations, perform $O(r)$ operations in each one, and the size of the input is $\geq |\mathcal{A}|r$.

Obviously, the set $\mathcal{S}$ produced by the algorithm is such that $\phi(\mathcal{S}) \geq \lceil \frac{r}{2} \rceil$. The following lemma establishes that $|\mathcal{S}| \geq k$.

Lemma 3.9. *Let $(I, \mathcal{A}, k)$ be a perfect r-MRkM instance, and let $\mathcal{S} \subseteq \mathcal{A}$, $M \in M(\mathcal{S}, I)$ be such that $|\mathcal{S}| < k$, and $|M(A)| = \lceil \frac{r}{2} \rceil$ for all $A \in \mathcal{S}$. Then there exists $B \in \mathcal{A} \setminus \mathcal{S}$, and $M' \in M(\mathcal{S} \cup \{B\}, I)$ such that $M'(A) = M(A)$ for all $A \in \mathcal{S}$, and $|M'(B)| = \lceil \frac{r}{2} \rceil$.*

The proof will be deferred for later and given in Section 5 in a more general setting (Lemma 5.3). Note that Lemma 3.2 does not imply the above lemma, since we here have the additional requirement that we do not change assignment of items that are already allocated.

Corollary 3.10. *Algorithm Lin-Greedy-r-MRkM is a 1/2 approximation algorithm for the perfect r-MRkM problem.*

Proof. Lemma 3.9 ensures that as long as the solution set S maintains $|\mathcal{S}| < k$, there must be a hyperedge that contains $\geq \lceil \frac{r}{2} \rceil$ items that have not been allocated yet by the algorithm. We only need to make sure this hyperedge is not one we have already examined in previous iterations. Indeed, assume the algorithm is at the end of iteration j, has just added A_j to $\mathcal{S}$, and that $|\mathcal{S}| < k$ still. It cannot be the good hyperedge of which existence is assured by the lemma is one that the algorithm has examined in previous iterations, since all of those are either already in $\mathcal{S}$, or do not have a sufficient amount of free items. Thus it follows that the desired hyperedge has not been examined yet and has index $> j$. $\square$

4 Bounded Degree Hypergraphs

In this section we examine the effect of bounding the degree of the input hypergraph. We say a hypergraph is of bounded degree d if every item is contained in $\leq d$ hyperedges. Following is the proof of Theorem 1.6, which we restate here for completeness.

Theorem 1.6. *Let $\mathcal{H}(I, \mathcal{A})$ be an r-uniform hypergraph of bounded degree d. Then there exists $\mathcal{S} \subseteq \mathcal{A}$, $|\mathcal{S}| \geq \frac{|\mathcal{A}|}{d}$, with rank $\geq \max\left\{\left\lceil \frac{r}{2} \right\rceil, \left\lfloor \frac{d+1}{2d} r \right\rfloor\right\}$. Moreover, this set can be found in polynomial time.*

Proof. Denote $n := |I|$, $m := |\mathcal{A}|$. Define the auxiliary graph $G(V := \mathcal{A}, E)$, where $(A, T)_i \in E$ if $i \in A \cap T$. Note this graph may contain parallel edges labeled by different items (but not loops).

Lemma 4.1. *There exists a set $\mathcal{S} \subseteq \mathcal{A}$ of size $\geq \frac{m}{d}$ with the property that each $A \in \mathcal{S}$ has $\leq \frac{d-1}{d} r$ edges with the other endpoint in $\mathcal{S}$, and this set can be found in polynomial time.*

Proof. The proof is given by the following algorithm:

1. Start with an arbitrary partition of $\mathcal{A}$ into d sets $\{\mathcal{S}_j\}_{j=1..d}$. Denote by $e(A)$ the number of inbound edges of A - edges with the other endpoint in the same set.

2. Pick a hyperedge for which there exists a set $\mathcal{S}_j$ such that if moved to it, $e(A)$ strictly decreases.

3. Repeat step two until no such hyperedge exists.

We claim the above algorithm repeats step two no more than $|E| \leq nd^2$ times (where $n = |I|$). Consider the amount of cut edges (edges with each endpoint in a different set) after each iteration. Assume hyperedge A was moved from one set to the other. Obviously, only edges incident in A have changed their status. Since we moved A, the amount of A's inbound edges strictly decreased, thus the amount of cut edges incident in A strictly increased. From this it follows that the total amount of cut edges strictly increased.

Observe that a vertex $A \in V = \mathcal{A}$ has no more than $(d-1)r$ incident edges, thus there always must exist a set achieving $e(A) \leq \frac{(d-1)}{d} r$. So we have that when the algorithm halts, $e(A) \leq \frac{(d-1)}{d} r$ for all $A \in \mathcal{A}$. Since there are d sets, one of them must contain at least $\frac{m}{d}$ hyperedges, and the result follows. □

Lemma 4.2. *Let $\mathcal{S} \subseteq V = \mathcal{A}$ be a set such that for all $A \in \mathcal{S}$, A has $\leq \frac{d-1}{d} r$ inbound edges in G. Then*

$$\phi(\mathcal{S}) \geq \max\left\{\left\lceil \frac{r}{2} \right\rceil, \left\lfloor \frac{d+1}{2d} r \right\rfloor\right\}$$

Proof. We will show that the generalized hall condition (Lemma 2.1) holds. Let $\mathcal{T} \subseteq \mathcal{S}$ be some subset of hyperedges. We want to prove the following inequality:

$$|I(\mathcal{T})| \geq |\mathcal{T}| \max\left\{\left\lceil \frac{r}{2} \right\rceil, \left\lfloor \frac{d+1}{2d} r \right\rfloor\right\}$$

For $1 \le j \le d$ denote by S^j the set of items which are contained in exactly j hyperedges in $\mathcal{S}$. For $A \in \mathcal{S}$, $0 \le j \le d-1$, denote by $S_A^j := S^{j+1} \cap A$ the set of items in A shared with exactly j other hyperedges in $\mathcal{S}$. Then for all $0 \le j \le d-1$,

$$\sum_{A \in \mathcal{T}} |S_A^j| = \sum_{A \in \mathcal{T}} \sum_{i \in S_A^j} 1 = \sum_{i \in S^{j+1} \cap I(\mathcal{T})} \sum_{A \in \mathcal{T} : i \in S_A^j} 1 \le \sum_{i \in S^{j+1} \cap I(\mathcal{T})} (j+1) = |S^{j+1} \cap I(\mathcal{T})|(j+1)$$

and so

$$|I(\mathcal{T})| = \sum_{j=0}^{d-1} |I(\mathcal{T}) \cap S^{j+1}| \ge \sum_{j=0}^{d-1} \frac{1}{j+1} \sum_{A \in \mathcal{T}} |S_A^j| = \sum_{A \in \mathcal{T}} \sum_{j=0}^{d-1} \frac{1}{j+1} |S_A^j|$$

To complete the proof we will show that for any $A \in \mathcal{S}$

$$\sum_{j=0}^{d-1} \frac{1}{j+1} |S_A^j| \ge \max \left\{ \left\lceil \frac{r}{2} \right\rceil, \left\lfloor \frac{d+1}{2d} r \right\rfloor \right\}$$

Indeed,

$$\sum_{j=1}^{d-1} j |S_A^j| \le \left\lfloor \frac{d-1}{d} r \right\rfloor$$

(since A has $\le \frac{d-1}{d} r$ inbound edges), hence

$$\begin{aligned}
\sum_{j=0}^{d-1} \frac{1}{j+1} |S_A^j| &= |S_A^0| + \sum_{j=1}^{d-1} \frac{1}{j+1} |S_A^j| \\
&= (r - \sum_{j=1}^{d-1} |S_A^j|) + \sum_{j=1}^{d-1} \frac{1}{j+1} |S_A^j| \\
&= r - \sum_{j=1}^{d-1} \frac{j}{j+1} |S_A^j| \\
&\ge r - \frac{1}{2} \left\lfloor \frac{d-1}{d} r \right\rfloor
\end{aligned}$$

where the last inequality follows by the fact that

$$\sum_{j=1}^{d-1} \frac{j}{j+1} |S_A^j| \le \sum_{j=1}^{d-1} \frac{j}{2} |S_A^j| \le \frac{1}{2} \left\lfloor \frac{d-1}{d} r \right\rfloor$$

Now

$$r - \frac{1}{2} \left\lfloor \frac{d-1}{d} r \right\rfloor \ge r - \frac{1}{2}(r-1) = \frac{1}{2} r + \frac{1}{2} \ge \left\lceil \frac{1}{2} r \right\rceil$$

and

$$r - \frac{1}{2} \left\lfloor \frac{d-1}{d} r \right\rfloor \ge r - \frac{1}{2} \frac{d-1}{d} r = \frac{d+1}{2d} r \ge \left\lfloor \frac{d+1}{2d} r \right\rfloor$$

which completes the proof. $\square$

Combining the two lemmas, the theorem now follows. □

Obviously, the cases in which the theorem is interesting are those in which the average degree of the hypergraph is close to the maximum degree. We say a hypergraph is d-regular if every item is contained in exactly d hyperedges. Consider a perfect r-MRkM instance $(I, \mathcal{A}, k)$ in which the input hypergraph is d-regular. Since we know this instance contains an actual matching, it must be that $k \leq \frac{|I|}{r}$. Regularity and uniformity imply $\frac{|I|}{r} = \frac{|\mathcal{A}|}{d}$. Hence, for these instances of the problem, the algorithm given in the proof of Theorem 1.6 provides an approximation ratio strictly better than $1/2$ when $r \geq 2d$.

5 A Linear Programming Relaxation

Given an r-MRkM instance $(I, \mathcal{A}, k)$, consider the following linear program relaxation:

Find weights $w : \mathcal{A} \to [0, 1]$ so that
$\sum_{A \in \mathcal{A}} w(A) = k$
$\forall i \in I, \quad \sum_{A : i \in A} w(A) \leq 1$

We call a solution to this linear program a fractional solution for an r-MRkM instance. Note that this is in fact a fractional matching in the input hypergraph, in the sense that all items are covered at most once. If all the weights assigned by the LP were integers, the solution set would be an actual matching.

Obviously, if there exists an integral solution of rank r for an r-MRkM instance, then the corresponding linear program is feasible, and there would also exist a fractional solution. We consider the question of the integrality gap of this linear program. Assume $(I, \mathcal{A}, k)$ has a fractional solution. Then what can we expect the rank of the optimal integral solution to be?

More formally, denote by $MRkM_{LP}$ the family of all uniform MRkM instances for which our LP relaxation is feasible. Let $r(I, \mathcal{A}, k)$ denote the size of the hyperedges in the input hypergraph (and thus in a sense the rank of the fractional solution). Then we define the integrality gap of our LP as

$$\inf_{(I,\mathcal{A},k) \in MRkM_{LP}} \left\{ \frac{opt(I, \mathcal{A}, k)}{r(I, \mathcal{A}, k)} \right\}$$

Where $opt(I, \mathcal{A}, k)$ denotes the rank of the optimal integral solution to $(I, \mathcal{A}, k)$.

Proposition 5.1. *Let $d \in \mathbb{N}$, and set $k := d + 1$. There exists a d-uniform d-regular MRkM instance $(I, \mathcal{A}, k)$ with $|\mathcal{A}| = kd$ (and $|I| = kd$) for which our LP is feasible, yet any integral solution is such that all allocations must have a hyperedge receiving no more than $\lceil \frac{d}{2} \rceil$ items.*

Proof. The item set we will use is given by

$$I = \{i_{j,l} : 1 \leq j \leq d \text{ and } 0 \leq l \leq d\}$$

with hyperedges $\mathcal{A} := \{A_{j,l}\}$

$$A_{j,l} := \{i_{j,l'} : l \leq l' \leq l + d - 1 (\text{mod } d + 1)\}$$

for $1 \leq j \leq d$ and $0 \leq l \leq d$.

We refer to the set $\{i_{j,l} : 0 \leq l \leq d\}$ as cycle j, and note that each hyperedge is completely contained in one of the cycles. Figure 4 shows an illustration of the constructed graph for $d = 3$.

Consider the d-MRkM instance $(I, \mathcal{A}, k)$. Giving each hyperedge weight of $\frac{1}{d}$ can be easily verified as a fractional solution.

Now consider some subset of hyperedges of size k. Since there are $d = k-1$ cycles, by the pigeonhole principle there must exist at least two hyperedges in the same cycle. Since each cycle contains $d+1$ items, one of these two hyperedges will receive no more than $\lfloor \frac{d+1}{2} \rfloor = \lceil \frac{d}{2} \rceil$ items in any possible assignment. $\square$

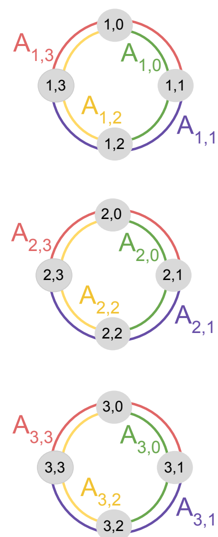

Figure 4: The cycles hypergraph construction for $d = 3$

Remark 5.2. *It follows the integrality gap is no better than* $1/2$*. By a very simple modification of the construction in which we consider multiple copies of each item, it can be seen that Theorem 1.6 is tight. That is, there exist* d*-regular* r*-uniform hypergraphs with* $r > 2d$*, such that every set of hyperedges of size* $\geq \frac{|\mathcal{A}|}{d}$ *has rank* $\leq \left\lfloor (\frac{1}{2} + \frac{1}{2d})r \right\rfloor$*.*

We will now establish that the integrality gap is no worse than $1/2$, thus achieving a tight result.

Lemma 5.3. *Let* $(I, \mathcal{A}, k)$ *be an* r*-MRkM instance that has a fractional solution. Let* $\mathcal{S} \subseteq \mathcal{A}$*,* $M \in M(\mathcal{S}, I)$ *be such that* $|\mathcal{S}| < k$*, and* $|M(A)| = \left\lceil \frac{r}{2} \right\rceil$ *for all* $A \in \mathcal{S}$*. Then there exists* $B \in \mathcal{A} \setminus \mathcal{S}$*, and* $M' \in M(\mathcal{S} \cup \{B\}, I)$ *such that* $M'(A) = M(A)$ *for all* $A \in \mathcal{S}$*, and* $|M'(B)| = \left\lceil \frac{r}{2} \right\rceil$*.*

Proof. Denote $x_i := \sum_{A: i \in A} w(A)$, and $X_V := \sum_{i \in V} x_i$ for any subset of items $V \in I$. We begin with establishing the following:

Claim 5.4. *Let* $U \subseteq I$ *be some subset of items. Then there exists a hyperedge* $A \in \mathcal{A}$ *such that*

$$|A \cap U| \geq \frac{X_U}{X_I} r$$

Proof. First observe

$$\sum_{A \in \mathcal{A}} |A \cap U| w(A) = \sum_{i \in U} x_i = X_U$$

Assume $|A \cap U| < \frac{X_U}{X_I} r$ for all $A \in \mathcal{A}$, then

$$X_U = \sum_{A \in \mathcal{A}} |A \cap U| w(A) < \sum_{A \in \mathcal{A}} \frac{X_U}{X_I} r w(A) = \frac{X_U}{X_I} \sum_{A \in \mathcal{A}} w(A) r = \frac{X_U}{X_I} X_I = X_U$$

Which is a contradiction. Therefore it must be that there exists $A \in \mathcal{A}$ s.t. $|A \cap U| \geq \frac{X_U}{X_I} r$. $\square$

Now, set $V := \cup_{A\in\mathcal{S}} M(A)$, and $U := I \setminus V$. Observe that $X_I = \sum_{i\in V}\sum_{A:i\in A} w(A) = kr$, $X_V \le |V|$, and $X_U = X_I - X_V = kr - X_V$.

We distinguish between the cases that r is even/odd. If r is even, $|V| < k\frac{r}{2}$, therefore $X_U > \frac{kr}{2}$ and by an application of the claim, there exists $B \in \mathcal{A}$ such that $|B\cap U| > \frac{kr/2}{kr} r = \frac{r}{2}$. Observe that $B \notin \mathcal{S}$, since $|A \cap U| = \frac{r}{2}$ for all $A \in \mathcal{S}$.

If r is odd, $|V| < k\frac{r+1}{2}$, therefore $X_U > kr - k\frac{r+1}{2} = k\frac{r-1}{2}$ and by an application of the claim, there exists $B \in \mathcal{A}$ such that $|B \cap U| > k\frac{r-1}{2}\frac{1}{kr} r = \frac{r-1}{2}$. Since $|B \cap U| \in \mathbb{N}$, it follows that $|B \cap U| \ge \lceil \frac{r}{2} \rceil$. Observe that $B \notin \mathcal{S}$, since $A \cap U = \frac{r-1}{2}$ for all $A \in \mathcal{S}$.

Therefore, in any case there exists $B \in \mathcal{A} \setminus \mathcal{S}$ such that it can be allocated $\lceil \frac{r}{2} \rceil$ items that are not already allocated by M to hyperedges in $\mathcal{S}$. $\square$

Remark 5.5. *Lemma 5.3 directly implies Lemma 3.9 which was stated earlier without a proof. Any r-MRkM instance that admits an integral solution with rank r has a fractional solution, as we have noted earlier.*

Corollary 5.6. *Let $(I, \mathcal{A}, k)$ be an r-MRkM instance for which our LP is feasible. Then there exists an integral solution with rank $\ge \frac{r}{2}$, and it can be found in polynomial time.*

Proof. The proof is implied directly by Lemma 5.3.

Consider the execution of Algorithm *Lin-Greedy-r-MRkM* on the instance $(I, \mathcal{A}, k)$. The proof given in Corollary 3.10 may be repeated in the same exact form with the only difference being that we rely on Lemma 5.3 (that only requires there existence of a fractional solution) rather than Lemma 3.9. $\square$

Corollary 5.6 and Proposition 5.1 together establish the proof for Theorem 1.7, restated here for completeness.

Theorem 1.7. *The integrality gap of our LP relaxation is exactly* $1/2$*:*

1. *Let $(I, \mathcal{A}, k)$ be an r-MRkM instance for which our LP relaxation is feasible. Then there exists a set of hyperedges of size k with rank $\ge r/2$. Moreover, this set can be found in deterministic polynomial time.*

2. *There exist r-MRkM instances $(I, \mathcal{A}, k)$ such that our LP relaxation is feasible, yet any set of hyperedges of size k has rank $\le r/2$.*

6 Linear Hypergraphs

In this section we examine r-uniform linear hypergraphs. We show that as r grows, we are able to ensure a solution with rank arbitrarily close to $(1-1/e)r$. We do this by presenting a randomized polynomial time rounding procedure for our LP relaxation from Section 5.

Given an r-MRkM instance $(I, \mathcal{A}, k)$ equipped with a fractional solution $w : \mathcal{A} \longrightarrow [0,1]$, we would like to randomly select a subset of hyperedges for which we know a good fractional allocation of items to exist (recall Definition 2.7. A fractional *allocation* is not to be confused with a fractional *solution*).

The idea is to select a tentative subset of hyperedges according to a probability distribution defined by the fractional solution w with some added slackness. Thus, in expectation, we will choose more hyperedges than we need to. Next, we will compute a certain fractional allocation that is expected to be sufficiently good, and dismiss from the tentative set hyperedges that ended up with a low fractional assignment of items. Corollary 2.2 will then establish the existence of a good integral allocation.

6.1 Basic Concentration Bounds

We will repeatedly use Chernoff-Hoeffding bounds in this section. We state the exact form in which we use them here.

Lemma 6.1. *Let $X_1, ..., X_n$ be i.r.v. taking values in $[0,1]$, and $Y := \sum_{i=1..n} X_n$ Then for $0 < \delta < 1$*

$$\Pr\left[Y \le (1-\delta)\,\mathbb{E}\left[Y\right]\right] \le e^{-\frac{\delta^2\,\mathbb{E}[Y]}{2}}$$

and

$$\Pr\left[Y \ge (1+\delta)\,\mathbb{E}\left[Y\right]\right] \le e^{-\frac{\delta^2\,\mathbb{E}[Y]}{3}}$$

Lemma 6.2. *Let $X_1, ..., X_n$ be i.r.v. taking values in $\{0,1\}$, and $Y := \sum_{i=1..n} X_n$ Then*

$$\Pr\left[Y - \mathbb{E}\left[Y\right] \ge t\right] \le e^{-\frac{2t^2}{n}}$$

These bounds are well known. Proofs can be found, for example, in [AS92] Appendix A.1.

6.2 Randomized Selection and the Fair Contention Resolution Rule

Let $(I, \mathcal{A}, k)$ be an r-MRkM instance equipped with a fractional solution w. Let $0 < \delta \le \frac{1}{8}$ be a slackness parameter. Select randomly $\mathcal{S} \subseteq \mathcal{A}$ where each hyperedge $A \in \mathcal{A}$ is selected independently with probability $p_A := \min(1, w(A)(1+\delta))$. Assign now each item fractionally according to the Fair Contention Resolution Rule devised by Feige and Vondrak [FV10]. Following is the definition of the allocation.

Let $i \in I$. Denote by $\mathcal{T} \subseteq \mathcal{A}$ the subset of hyperedges that contain item i, and by $\mathcal{S}_i \subseteq \mathcal{S}$ the subset of selected hyperedges that contain i.

- If $\mathcal{S}_i = \emptyset$ do not allocate the item at all.
- If $|\mathcal{S}_i| = 1$ allocate the whole item to the single selected hyperedge that contains it.
- If $|\mathcal{S}_i| > 1$, for $A \in \mathcal{S}_i$ allocate the following fraction of i

$$\frac{1}{\sum_{B \in T} p_B} \left(\sum_{B \in \mathcal{S}_i \setminus \{A\}} \frac{p_B}{|\mathcal{S}_i| - 1} + \sum_{B \in \mathcal{T} \setminus \mathcal{S}_i} \frac{p_B}{|\mathcal{S}_i|} \right)$$

This concludes the allocation definition. Observe

$$\sum_{A \in \mathcal{S}_i} \left(\sum_{B \in \mathcal{S}_i \setminus \{A\}} \frac{p_B}{|\mathcal{S}_i| - 1} + \sum_{B \in \mathcal{T} \setminus \mathcal{S}_i} \frac{p_B}{|\mathcal{S}_i|} \right) = \sum_{B \in \mathcal{S}_i} (|\mathcal{S}_i| - 1) \frac{p_B}{|\mathcal{S}_i| - 1} + |\mathcal{S}_i| \sum_{B \in \mathcal{T} \setminus \mathcal{S}_i} \frac{p_B}{|\mathcal{S}_i|} = \sum_{B \in \mathcal{T}} p_B$$

Denote by $F \in M^f(\mathcal{S}, I)$ this fractional allocation given by the FCR rule (from the above it can easily be seen F maintains the requirements of Definition 2.7). The next lemma shows this allocation is expected to give a good fractional assignment.

Lemma 6.3. *For $A \in \mathcal{S}$, $i \in A$*

$$\mathbb{E}\left[F(A)_i\right] \geq 1 - 1/e - \delta$$

Proof. The original lemma established by Feige and Vondrak [FV10] shows that the fraction of i A is expected to collect is

$$\frac{1 - \prod_{T \in \mathcal{T}} (1 - p_T)}{\sum_{T \in \mathcal{T}} p_T}$$

An application of the arithmetic-geometric means inequality gives that

$$\prod_{T \in \mathcal{T}} (1 - p_T) \leq \left(\frac{\sum_{T \in \mathcal{T}} 1 - p_T}{|\mathcal{T}|} \right)^{|\mathcal{T}|} = \left(1 - \frac{\sum_{T \in \mathcal{T}} p_T}{|\mathcal{T}|} \right)^{|\mathcal{T}|}$$

Denote $C := \sum_{T \in \mathcal{T}} p_T$, then by a simple manipulation of the above

$$\frac{1 - \prod_{T \in \mathcal{T}} (1 - p_T)}{C} \geq \frac{1 - \left(1 - \frac{C}{|\mathcal{T}|}\right)^{|\mathcal{T}|}}{C}$$

An analysis of the RHS shows that it is decreasing as a function of C. Recall that $\sum_{T \in \mathcal{T}} w(T) \leq 1$ (since all hyperedges in $\mathcal{T}$ contain i), thus $C \leq 1 + \delta$. Meaning

$$\frac{1 - \left(1 - \frac{C}{|\mathcal{T}|}\right)^{|\mathcal{T}|}}{C} \geq \frac{1 - \left(1 - \frac{1+\delta}{|\mathcal{T}|}\right)^{|\mathcal{T}|}}{1 + \delta} \geq \frac{1}{1 + \delta} - \frac{e^{-1-\delta}}{1 + \delta}$$

Now

$$\begin{aligned}
& \frac{1}{1+\delta} - \frac{e^{-1-\delta}}{1+\delta} \geq 1 - \frac{1}{e} - \delta \\
\iff \quad & 1 - e^{-1-\delta} \geq 1 + \delta - \frac{1+\delta}{e} - \delta - \delta^2 \\
\iff \quad & \frac{1+\delta}{e} + \delta^2 \geq e^{-1-\delta}
\end{aligned}$$

which is true since $e^{-1-\delta} < \frac{1}{e}$. $\square$

The following lemma establishes a concentration result heavily relying on linearity of the hypergraph.

Lemma 6.4. *For $A \in \mathcal{S}$, assume $\mathbb{E}\left[F(A)_i\right] \geq \alpha$ for all $i \in A$, and denote $|F(A)| := \sum_{i \in A} F(A)_i$. Then*

$$\Pr\left[|F(A)| \leq (\alpha - \epsilon)r\right] \leq e^{-\frac{\epsilon^2 \alpha r}{2}}$$

Proof. Observe that the $\{F(A)_i\}_{i \in A}$ are *independent* r.v.'s - for $i, j \in A, i \neq j$, there do not exist hyperedges (other than A) that contain both i and j since that would contradict linearity of the hypergraph. Since $F(A)_i, F(A)_j$ depend only on the sets of other hyperedges containing in i, j respectively, and these sets are disjoint, it follows they are independent.

Thus we may use Chernoff's bound, obtaining

$$\Pr\left[|F(A)| \leq (\alpha - \epsilon)r\right] \overset{(\alpha \leq 1)}{\leq} \Pr\left[|F(A)| \leq \alpha(1-\epsilon)r\right] \leq e^{-\frac{\epsilon^2 \alpha r}{2}}$$

$\square$

6.3 The Rounding Procedure

Let $\eta > 0$ be our approximation parameter. We want to show a rounding procedure that produces an integral solution with rank $\geq \lfloor (1 - 1/e - \eta)r \rfloor$. The idea is to randomly select a subset of hyperedges with some appropriate amount of slackness as described in the previous section, then get rid of all hyperedges that weren't allocated enough items by the FCR allocation. By assuming r to be large enough (where r is the size of all hyperedges), one could expect the probability to get rid of a hyperedge to be dominated by our added slackness, which would leave us with enough hyperedges in expectation.

In essence, this is indeed what we'll be doing, though with some added complexity imposed by the fact that heavy hyperedges (heavy in the sense that the fractional solution gives them weight close to 1) cannot be added a sufficient amount of slackness.

Theorem 6.5. *Fix $\frac{1}{8} \geq \eta > 0$, and let $(I, \mathcal{A}, k)$ be a linear r-MRkM instance with $r = \Omega(\frac{1}{\eta^2} \ln \frac{1}{\eta})$. If our LP relaxation is feasible, then there exists a subset $\mathcal{S} \subseteq \mathcal{A}$ of size k with rank $\geq \lfloor (1 - 1/e - \eta)r \rfloor$.*

Proof. Let w denote the fractional solution to the r-MRkM instance, and set $\delta := \frac{\eta}{2}$. We distinguish between three types of hyperedges; *heavy* are those with weight $> \frac{1}{1+\delta}$, *light* are those with weight $< \frac{\delta}{1+\delta}$, and *medium* are all others. Note that heavy hyperedges may intersect only light hyperedges. To begin with, we describe the rounding procedure of the fractional solution.

1. Make a random selection of a tentative set $\mathcal{S}$ with slackness of δ: For every $A \in \mathcal{A}$, $\Pr[A \in \mathcal{S}] := p_A = \min\{w(A)(1+\delta), 1\}$.
2. Set $\mathcal{S}'$ to be the set of all heavy hyperedges in $\mathcal{S}$.
3. Compute the fractional allocation $F \in M^f(S, I)$ given by the FCR rule.
   (a) We call a hyperedge $A \in \mathcal{S}$ *poor* if $|F(A)| < (1 - \frac{1}{e} - 2\delta)r$.
4. Add to $\mathcal{S}'$ every medium $A \in \mathcal{S}$ that is not poor.
5. We say a heavy hyperedge $H \in \mathcal{S}$ is *stressed* if $|H \cap I(\mathcal{S} \setminus \{H\})| > \frac{r}{e}$. Add to $\mathcal{S}'$ every light $A \in \mathcal{S}$ that is not poor *and* does not intersect any stressed hyperedges.

This concludes the rounding procedure.

Lemma 6.6. *The expected size of $\mathcal{S}'$ given the random rounding procedure described above is $\geq k$.*

Proof. For any hyperedge $A \in \mathcal{A}$, denote by $q_A := \Pr[A \in \mathcal{S}']$ the overall probability for A to end up in $\mathcal{S}'$ as a result of our rounding procedure. Our strategy will be to show that $q_A \geq w(A)$ holds for all hyperedges $A \in \mathcal{A}$. Recall that $A \in \mathcal{A}$ are selected for $\mathcal{S}$ with probability $p_A = \min(1, w(A)(1+\delta))$.

If $A \in \mathcal{A}$ is heavy, then being selected for $\mathcal{S}$ makes it also selected for $\mathcal{S}'$ automatically (we never get rid of heavy hyperedges), thus $q_A = p_A \geq w(A)$.

Let $A \in \mathcal{A}$ be medium or light. If $A \in \mathcal{S}$, by Lemma 6.3, the fraction A is expected to collect of any of its items is $\geq (1 - 1/e - \delta)$. Thus, by applying Lemma 6.4 with $\epsilon := \delta$ and $\alpha := (1 - 1/e - \delta)$, we have that

$$\Pr[A \text{ is poor} \mid A \in \mathcal{S}] = \Pr\left[|F(A)| < (1 - \frac{1}{e} - 2\delta)r\right] \leq e^{-\frac{\delta^2 r}{4}}$$

Therefore, the probability for a medium hyperedge $A \in \mathcal{A}$ to not make it to $\mathcal{S}'$

$$1 - q_A = \Pr[(A \notin \mathcal{S}) \vee (A \in \mathcal{S} \text{ and is poor})] \leq 1 - w(A)(1+\delta) + w(A)(1+\delta)e^{-\frac{\delta^2 r}{4}}$$

As for light hyperedges, they have the added chance to not make it to $\mathcal{S}'$ on account of meeting a stressed hyperedge. For a heavy hyperedge $H \in \mathcal{S}$, denote the r.v. $X_H := |H \cap I(S \setminus \{H\})|$. Observe

$$\mathbb{E}[X_H] = \sum_{i \in H} \Pr[\exists A \in \mathcal{S} \setminus \{H\} \text{ s.t. } i \in A] \leq r\frac{\delta}{1+\delta}(1+\delta) = \delta r$$

and the probability of H being stressed can be bounded using Chernoff's inequality:

$$\Pr[H \text{ is stressed}] = \Pr\left[X_H > \frac{r}{e}\right] \leq \Pr\left[X_H - \delta r > \frac{r}{2e}\right] < e^{-\frac{2\frac{r^2}{4e^2}}{r}} < e^{-\frac{r}{20}}$$

Now let $A \in \mathcal{A}$ be light. Denote by E_A the event that A intersects a stressed hyperedge. By union bounding over A's r items, it follows that $\Pr[E_A] \leq re^{-\frac{r}{20}}$. Hence

$$\begin{aligned}1 - q_A &= \Pr[(A \notin \mathcal{S}) \vee (A \in \mathcal{S} \text{ and is poor}) \vee (A \in \mathcal{S} \text{ and } E_A)] \\ &\leq 1 - w(A)(1+\delta) + w(A)(1+\delta)e^{-\frac{\delta^2 r}{4}} + w(A)(1+\delta)re^{-\frac{r}{20}}\end{aligned}$$

What is left then is to require that

$$\begin{aligned} & 1 - w(A) \geq 1 - w(A)(1+\delta) + w(A)(1+\delta)e^{-\frac{\delta^2 r}{4}} + w(A)(1+\delta)re^{-\frac{r}{20}} \\ \iff \quad & 0 \geq -\delta + (1+\delta)e^{-\frac{\delta^2 r}{4}} + (1+\delta)re^{-\frac{r}{20}} \\ \iff \quad & \frac{\delta}{1+\delta} \geq e^{-\frac{\delta^2 r}{4}} + re^{-\frac{r}{20}}\end{aligned}$$

It is easily seen that if the last inequality holds, $q_A \geq w(A)$ for all medium and light hyperedges $A \in \mathcal{A}$. Indeed, assume $r \geq C\frac{1}{\delta^2}\ln\frac{1}{\delta} = \Omega(\frac{1}{\eta^2}\ln\frac{1}{\eta})$.

By using the following facts: $\delta < \frac{1}{10}$, $xe^{-\frac{x}{20}} \leq ye^{-\frac{y}{20}}$ for $50 \leq y \leq x$, and $\frac{1}{2^a} < a^{-1}$ for $a \geq 1$, we have

$$\begin{aligned} e^{-\frac{\delta^2 r}{4}} + re^{-\frac{r}{20}} &\leq e^{-\frac{\delta^2}{4}C\frac{1}{\delta^2}\ln\frac{1}{\delta}} + C\frac{1}{\delta^2}\ln\frac{1}{\delta}e^{-\frac{1}{20}C\frac{1}{\delta^2}\ln\frac{1}{\delta}} \\ &= e^{\frac{C}{4}\ln\delta} + C\frac{1}{\delta^2}\ln\frac{1}{\delta}e^{\frac{C}{20\delta^2}\ln\delta} \\ &= \delta^{\frac{C}{4}} + \delta^{\frac{C}{20\delta^2}}C\frac{1}{\delta^2}\ln\frac{1}{\delta} \\ &\leq \delta^{\frac{C}{4}} + \delta^{\frac{1}{20\delta^2}}\delta^{-3} \\ &\leq \delta^{\frac{C}{4}} + \delta^2 \leq \frac{\delta}{2}\end{aligned}$$

For $C = 8$. We have shown that for all $A \in \mathcal{A}$, $q_A \geq w(A)$, hence $\mathbb{E}[|\mathcal{S}'|] \geq \sum_{A\in\mathcal{A}} w(A) = k$. $\square$

To complete the proof, we will show there exists an integral allocation $M \in M(\mathcal{S}', I)$ such that $M(A) \geq \lfloor(1 - 1/e - \eta)r\rfloor$. This is true regardless of the random selection of $\mathcal{S}$, and is derived by the fact that $\mathcal{S}'$ has a good fractional allocation.

Lemma 6.7. *There exists an integral allocation of items to hyperedges in $\mathcal{S}'$ such that each hyperedge is assigned $\geq \lfloor(1 - 1/e - \eta)r\rfloor$ items.*

Proof. Let $M \in M^f(\mathcal{S}', I)$ be an allocation defined as follows; for all non-stressed hyperedges $A \in \mathcal{S}'$, set $M(A) := F(A)$. For stressed $H \in \mathcal{S}'$, set $M(H) := H$ (meaning $M(H)$ is set to the characteristic vector of H). We will verify M is a legal fractional allocation that has rank $\geq (1-1/e-\eta)r$.

To see that M is legal, observe that if $i \in I$ is such that no stressed hyperedge contains it,

$$\sum_{A\in\mathcal{S}'} M(A)_i \leq \sum_{A\in\mathcal{S}} F(A)_i \leq 1$$

If $i \in I$ is contained in some stressed hyperedge $H \in \mathcal{S}'$, it is not contained in any other hyperedge $B \in \mathcal{S}'$; since H is heavy it may only intersect light hyperedges, but since it is stressed, we dismissed all light hyperedges intersecting it and they never made it to $\mathcal{S}'$.

The fact that M has good rank is easily verified; All stressed hyperedges are allocated all their r items, and all other hyperedges $A \in \mathcal{S}'$ are allocated the same amount of items they were allocated by the FCR allocation F. Since we have left out of $\mathcal{S}'$ all non-stressed hyperedges that were assigned by F less than $(1-1/e-\eta)r$ items, we have it that the rank of M is $\geq (1-1/e-\eta)r$.

Now, an application of Corollary 2.2 gives that there exists an integral allocation assigning each hyperedges $\geq \lfloor(1-1/e-\eta)r\rfloor$ items, as desired. □

Summarizing, we have shown our rounding procedure produces the set $\mathcal{S}'$ which has the property of always having sufficiently good rank. In addition, the expected size of $\mathcal{S}'$ is $\geq k$, thus with positive probability $\mathcal{S}'$ is an integral solution as desired. □

Given a linear r-MRkM instance with $r = \Omega(\frac{1}{\eta^2}\ln\frac{1}{\eta})$ for which the LP is feasible, we can solve the LP and randomly construct an integral solution $\mathcal{S}'$ using the rounding procedure. As provided by the theorem, $\mathcal{S}'$ always has good rank, and thus we need only worry about its size for it to be an actual solution. To show that a solution is produced with sufficiently high probability, we use the fact that the r.v. $X := |S'|$ is a bounded $(0 \leq X \leq |\mathcal{A}|)$ integral r.v. with expectation $\geq k$ (by Lemma 6.6). Set $q := P(X \geq k)$, $m := |\mathcal{A}|$. Then $P(X < k) = 1 - q$, and

$$\begin{aligned}
k \leq E[X] &= \sum_{i=1..m} P(X=i)i \\
&\leq \sum_{i=1..k-1} P(X=i)(k-1) + \sum_{i=k..m} P(X=i)m \\
&\leq (1-q)(k-1) + qm \\
&= k-1+q(m-k+1)
\end{aligned}$$

$\implies q \geq \frac{1}{(m-k+1)}$. Obviously $k \geq 1$, thus we have that a solution is produced with probability $\geq \frac{1}{m}$. Repeating the rounding procedure a number of times polynomial in m gives us a solution with arbitrarily high probability. In addition, we can obviously check if $\mathcal{S}'$ is of size $\geq k$, thus we produce a solution in expected polynomial time.

Corollary 6.8. *Let $(I, \mathcal{A}, k)$ be a linear r-MRkM instance. Denote $\eta = \beta\sqrt{\frac{\ln r}{r}}$ where $\beta > 0$ is some universal constant that does not depend on r. If the LP relaxation is feasible, then there exists an integral solution of rank $\lfloor(1-1/e-\eta)r\rfloor$, and it can be found in expected polynomial time.*

Proof. Let $C \in \mathbb{N}$ be the constant from the condition on r in Theorem 6.5. Set $\beta := \frac{1}{\sqrt{C}}$. If r is such that $\eta = \beta\sqrt{\frac{\ln r}{r}} > \frac{1}{8}$, then $(1 - 1/e - \eta) \leq \frac{1}{2}$ and correctness follows by Corollary 5.6. Otherwise, $0 < \eta \leq \frac{1}{8}$, and

$$C\frac{1}{\eta^2}\ln\frac{1}{\eta} = \frac{C}{\beta^2}\frac{r}{\ln r}\ln\sqrt{\beta\frac{r}{\ln r}} \leq \frac{r}{\ln r}\ln\sqrt{r} \leq r$$

Thus the conditions of Theorem 6.5 hold and the result follows. ☐

This concludes the proof for the first part of Theorem 1.8.

6.4 An Integrality Gap Upper Bound

We give for any $\epsilon > 0$ a randomized construction of a linear r-MRkM instance for which our LP is feasible, yet any integral solution will have rank no greater than $(1 - 1/e + \epsilon)r$, thus showing our result from the previous section to be nearly tight. This will establish the proof for the second part of Theorem 1.8: For any $\epsilon > 0$, there exist r-MRkM instances for which the LP relaxation is feasible, yet any integral solution has rank $< (1 - 1/e + \epsilon)r$.

Proof. We first describe an (n, m, k, δ) random construction ($n, m, k \in \mathbb{N}$, $0 < \delta < 1$). This construction yields a $\frac{n}{k}(1-\delta)$-uniform MRkM instance $(I, \mathcal{A}, k')$ with $|I| = n$, $|\mathcal{A}| \leq m$, $k' \leq k$.

Let I be a set of items of size n. We start with a tentative set of hyperedges $\mathcal{A}'$ of size m by making m independent random selections of hyperedges. Every hyperedge is uniformly selected from the set of all subsets of items of size $r := \frac{n}{k}(1-\delta)$. It follows that for any $A \in \mathcal{A}'$, $i \in I$, $\Pr[i \in A] = \frac{1}{k}(1-\delta) =: q$. Now, we remove from $\mathcal{A}'$ any two hyperedges that intersect more than once, and define this new set as $\mathcal{A}$. This concludes the construction.

Fix $\epsilon_0, \eta > 0$ parameters to be taken with hindsight, and depend on ϵ the parameter from the statement of the theorem. We will want to choose (n, m, k, δ) such that with positive probability all of the following hold:

1. No item $i \in I$ is covered by more than $\frac{m}{k}$ hyperedges
2. For any $\mathcal{S} \subseteq \mathcal{A}$ of size k, $|I(\mathcal{S})| \leq (1 - 1/e - \epsilon_0)n$
3. $|\mathcal{A}| \geq (1 - \eta)m$

We will bound from above the probabilities of each bad event we want to avoid by an appropriate choice of paramaters. Let $\delta > 0$, $n \in \mathbb{N}$, and set $k := n^{4/5}$, $m := Ck\ln n$ where $C := \frac{60}{\delta^2}$.

1. For $i \in I$, let $\mathcal{A}_i \subseteq \mathcal{A}'$ be the tentative subset of hyperedges containing i. Observe

   $$\mathbb{E}[|\mathcal{A}_i|] = m\frac{1}{k}(1-\delta)$$

   Thus by Chernoff's bound, relying on the fact that hyperedges were chosen independently

   $$\Pr\left[|\mathcal{A}_i| > \frac{m}{k}\right] \leq \Pr\left[|\mathcal{A}_i| > \frac{m}{k}(1-\delta)(1+\delta)\right] \leq e^{-\frac{\delta^2 m}{6k}}$$

and by union bounding over all items

$$\Pr\left[\exists i \in I \text{ s.t. } |\mathcal{A}_i| > \frac{m}{k}\right] \leq ne^{-\frac{\delta^2 m}{6k}} = ne^{-10\ln n} = \frac{1}{n^9} \xrightarrow[n\to\infty]{} 0$$

Since the final set of hyperedges $\mathcal{A}$ is a subset of $\mathcal{A}'$, it follows that

$$\Pr\left[\exists i \in I \text{ s.t. } i \text{ is covered by } \geq \frac{m}{k} \text{ hyperedges from } \mathcal{A}\right] \leq \Pr\left[\exists i \in I \text{ s.t. } |\mathcal{A}_i| > \frac{m}{k}\right]$$

2. Let $i \in I$, $\mathcal{S} \subseteq \mathcal{A}$, $|\mathcal{S}| = k$, then

$$\Pr\left[i \in I(\mathcal{S})\right] = q\sum_{j=0}^{k-1}(1-q)^j = 1 - q^k = 1 - (1 - \frac{1-\delta}{k})^k$$

Therefore for sufficiently small δ and large enough k

$$\mathbb{E}\left[|I(\mathcal{S})|\right] \leq (1 - 1/e + \epsilon_0/2)n$$

Observe that the random variables indicating for each $i \in I$ whether i was covered by some $A \in \mathcal{S}$ or not, are not independent. This is since the choices of items for the same hyperedge are not independent. Still, one can see they are *negatively* correlated; for $i, j \in I$, $A \in \mathcal{S}$, it holds that $\Pr\left[j \in A | i \in A\right] < \Pr\left[j \in A\right]$. Thus

$$\Pr\left[\bigwedge_{i\in I} i \in I(\mathcal{S})\right] < \prod_{i\in I}\Pr\left[i \in I(\mathcal{S})\right]$$

Therefore, we may apply Chernoff's bound (Panconesi and Srinivasan give a formal proof for this in [PS97]):

$$\Pr\left[|I(\mathcal{S})| > (1 - 1/e + \epsilon_0)\right] \leq e^{-\frac{\epsilon_0^2 n}{10}}$$

Union bounding over all subsets of $\mathcal{A}$ of size k, we have

$$\Pr\left[\exists \mathcal{S} \subseteq \mathcal{A}, |\mathcal{S}| = k \text{ s.t. } |I(\mathcal{S})| > (1 - 1/e + \epsilon_0)\right] \leq \binom{m}{k} e^{-\frac{\epsilon_0^2 n}{10}}$$

and

$$\binom{m}{k} e^{-\frac{\epsilon_0^2 n}{10}} \leq m^k e^{-\frac{\epsilon_0^2 n}{10}} \xrightarrow[n\to\infty]{} 0$$

To see this,

$$\ln\left(m^k e^{-\frac{\epsilon_0^2 n}{10}}\right) = k\ln m - \frac{\epsilon_0^2}{10}n = n^{4/5}\ln\left(Cn^{4/5}\ln n\right) - \frac{\epsilon_0^2}{10}n \xrightarrow[n\to\infty]{} -\infty$$

3. Denote by X_L the number of hyperedges we abandon due to the linearity requirement. Then

$$\mathbb{E}\left[X_L\right] \leq \binom{m}{2}\binom{r}{2}\left(\frac{r}{n}\right)^2 \leq m^2 r^2 \frac{r^2}{n^2} \leq m^2 \frac{n^2}{k^4}$$

Now, $m\frac{n^2}{k^4} = C \ln n \frac{n^2}{k^3} = \frac{C \ln n}{n^{2/5}} < \frac{\eta}{2}$ for sufficiently large n. Therefore $\mathbb{E}\left[X_L\right] \leq m^2 \frac{n^2}{k^4} \leq \frac{\eta}{2} m$, and by Markov's inequality

$$\Pr\left[X_L \geq \eta m\right] \leq \frac{\mathbb{E}\left[X_L\right]}{\eta m} \leq \frac{1}{2}$$

It now follows that taking $n \longrightarrow \infty$ and δ to be sufficiently small we have with positive probability a hypergraph with our desired properties (by union bounding over the three bad events). Let now $\mathcal{H}(I, \mathcal{A})$ be a hypergraph with these properties, and let $w : \mathcal{A} \to [0,1]$, $w(A) := \frac{k}{m}$ for all $A \in \mathcal{A}$. Then $\sum_{A:i \in A} w(A) \leq \frac{m}{k}\frac{k}{m} = 1$, and $\sum_{A \in \mathcal{A}} w(A) \geq m(1-\eta)\frac{k}{m} = (1-\eta)k$.

Hence we have that the r-MRkM instance $(I, \mathcal{A}, (1-\eta)k)$ has a fractional solution (assume w.l.o.g $(1-\eta)k \in \mathbb{N}$). Now, let $\mathcal{S} \subseteq \mathcal{A}$ be some subset of size $(1-\eta)k$. We have that $\mathcal{S}$ has coverage no larger than $(1 - 1/e - \epsilon_0)n$. Therefore

$$\frac{\phi\left(\mathcal{S}\right)}{r} \leq \frac{(1 - 1/e - \epsilon_0)n}{(1-\eta)k\frac{n}{k}(1-\delta)} = \frac{(1 - 1/e - \epsilon_0)}{(1-\eta)(1-\delta)}$$

So, for sufficiently small ϵ_0, η, δ it follows $(I, \mathcal{A}, (1-\eta)k)$ has no integral solutions with rank $> (1 - 1/e + \epsilon)r$, and we are done. □

7 Extensions

We consider the non-uniform MRkM problem. Let $\mathcal{H}(I, \mathcal{A})$ be a non-uniform hypergraph. In this case we would like to consider the notion of normalized rank of a matching $\mathcal{S} \subseteq \mathcal{A}$. Let $M \in M(\mathcal{S}, I)$, then the normalized rank of M is defined by

$$\overline{\phi}(M) := \min_{A \in \mathcal{S}} \left\{ \frac{|M(A)|}{|A|} \right\}$$

and the normalized rank of $\mathcal{S}$ is defined by

$$\overline{\phi}(\mathcal{S}) := \max_{M \in M(\mathcal{S}, I)} \left\{ \overline{\phi}(M) \right\}$$

The proof of Theorem 1.6 can be modified to apply for non-uniform hypergraphs. The result is immediate and requires only two minor changes. First, every use for r should be replaced with $|A|$ where A is the hyperedge being discussed in the context of the specific argument being made. Second, in Lemma 4.2 the generalized hall condition established should be of the following form:

$$|I(\mathcal{T})| \geq \sum_{A \in \mathcal{T}} \max \left\{ \left\lceil \frac{|A|}{2} \right\rceil, \left\lfloor \frac{d+1}{2d} |A| \right\rfloor \right\}$$

The proof is exactly the same as the one we have given for the uniform case and again, only requires the usage of $|A|$ instead of every appearance of r. This yields the following

Theorem 7.1. *Let $\mathcal{H}(I, \mathcal{A})$ be a hypergraph of bounded degree d. Then there exists $\mathcal{S} \subseteq \mathcal{A}$, $|\mathcal{S}| \geq \frac{|\mathcal{A}|}{d}$, and an allocation $M \in M(\mathcal{S}, I)$ such that $|M(A)| \geq \max \left\{ \left\lceil \frac{|A|}{2} \right\rceil, \left\lfloor \frac{d+1}{2d} |A| \right\rfloor \right\}$ for all $A \in \mathcal{A}$. Moreover, this set can be found in polynomial time.*

The proof for the first part of Theorem 1.8 does not actually use uniformity either. By repeating the proof replacing every use of r with the size of the hyperedge being discussed, we establish the following

Theorem 7.2. *Let $(I, \mathcal{A}, k)$ be a linear MRkM instance. Let $y \in \mathbb{N}$ denote the size of the smallest hyperedge in $\mathcal{H}$. If the LP relaxation is feasible, then there exists an integral solution $\mathcal{S} \subseteq \mathcal{A}$, and an allocation $M \in M(\mathcal{S}, I)$ such that*

$$|M(A)| \geq \lfloor (1 - 1/e - \eta)|A| \rfloor \qquad \forall A \in \mathcal{S}$$

where $\eta = O\left(\sqrt{\frac{\ln y}{y}}\right)$, and it can be found in expected polynomial time.

References

[AFS12] Arash Asadpour, Uriel Feige, and Amin Saberi, *Santa claus meets hypergraph matchings*, ACM Transactions on Algorithms **8** (2012), no. 3, 24.

[AS92] Noga Alon and Joel Spencer, *The probabilistic method*, John Wiley, 1992.

[AS07] Arash Asadpour and Amin Saberi, *An approximation algorithm for max-min fair allocation of indivisible goods*, STOC, 2007, pp. 114–121.

[BCG09] MohammadHossein Bateni, Moses Charikar, and Venkatesan Guruswami, *Maxmin allocation via degree lower-bounded arborescences*, Proceedings of the 41st annual ACM symposium on Theory of computing, STOC, 2009, pp. 543–552.

[BD05] Ivona Bezáková and Varsha Dani, *Allocating indivisible goods*, SIGecom Exchanges **5** (2005), no. 3, 11–18.

[BS06] Nikhil Bansal and Maxim Sviridenko, *The santa claus problem*, STOC, 2006, pp. 31–40.

[CCK09] Deeparnab Chakrabarty, Julia Chuzhoy, and Sanjeev Khanna, *On allocating goods to maximize fairness*, FOCS, 2009, pp. 107–116.

[Edm65] Jack Edmonds, *Paths, trees, and flowers*, Canad. J. Math. **17** (1965), 449–467.

[Fei98] Uriel Feige, *A threshold of ln n for approximating set cover*, J. ACM **45** (1998), no. 4, 634–652.

[Fei08] Uriel Feige, *On allocations that maximize fairness*, SODA, 2008, pp. 287–293.

[FV10] Uriel Feige and Jan Vondrák, *The submodular welfare problem with demand queries*, Theory of Computing **6** (2010), no. 1, 247–290.

[GJ90] Michael R. Garey and David S. Johnson, *Computers and intractability; a guide to the theory of np-completeness*, W. H. Freeman & Co., 1990.

[HS89] C. A. J. Hurkens and A. Schrijver, *On the size of systems of sets every t of which have an sdr, with an application to the worst-case ratio of heuristics for packing problems*, SIAM J. Discret. Math. **2** (1989), no. 1, 68–72.

[HSS06] Elad Hazan, Shmuel Safra, and Oded Schwartz, *On the complexity of approximating k-set packing*, Computational Complexity **15** (2006), no. 1, 20–39.

[HSS11] Bernhard Haeupler, Barna Saha, and Aravind Srinivasan, *New constructive aspects of the lovasz local lemma*, J. ACM **58** (2011), no. 6, 28:1–28:28.

[KT05] Jon Kleinberg and Eva Tardos, *Algorithm design*, Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 2005.

[PS97] Alessandro Panconesi and Aravind Srinivasan, *Randomized distributed edge coloring via an extension of the chernoff-hoeffding bounds*, SIAM J. Comput. **26** (1997), no. 2, 350–368.

[PS12] Lukas Polacek and Ola Svensson, *Quasi-polynomial local search for restricted max-min fair allocation*, ICALP (1), 2012, pp. 726–737.

[Tre01] Luca Trevisan, *Non-approximability results for optimization problems on bounded degree instances*, STOC, 2001, pp. 453–461.

A Appendix

A.1 Reduction from Max Coverage

The *max coverage* problem is defined as follows - as input we are given a collection of subsets $S_1, ...S_n$, and a number $k \in \mathbb{N}$. The objective is to select k of these sets so as to maximize coverage of elements, that is, the union of all these sets. Due to Feige [Fei98] we know this problem is **NP**-hard to approximate within any factor better than $1 - 1/e$:

Theorem A.1. *[Fei98] The max coverage problem is hard to approximate within any factor better than $1 - 1/e$, even in the specail case in which all sets are of the same size, and there exists a solution of disjoint sets covering all items.*

Assume now we have a polynomial time α-approximation algorithm Alg for the r-MRkM problem with $\alpha > 1 - 1/e$. Given an r-uniform max coverage instance that contains a perfect cover, we could run Alg (note that the input to both problems is of the exact same form) to find a family of k subsets which admits an allocation assigning αr elements to each subset. Thus the coverage of the selected subsets must be $\geq \alpha kr > (1 - 1/e)kr$, contradicting hardness of max-coverage.

So it follows that r-MRkM is hard to approximate within any factor better than $1 - 1/e$ as well.

A.2 Formulation in Terms of Simplicial Complexes

We repeat the problem definition for completeness: Given a pure simplicial r-complex $\mathcal{H}(I, \mathcal{A})$ that contains $k \in \mathbb{N}$ disjoint simplices of dimension $r - 1$, find a set of k disjoint simplices of dimension t, where t is to be maximized.

Observe that the output of this problem is slightly different than that of the Max Rank k-Matching problem, where we ask for the original hyperedges and not their subsets that form the packing. But, by Proposition 2.6 these subsets can be computed in polynomial time given the original hyperedges. Thus this difference is not significant.

Another difference, is that it might be that a solution to the simplicial complexes formulated problem does not originate from a set of k distinct maximal dimension simplices. If this is the case, it must be that the dimension of the simplices in the solution set is no larger than $\frac{r-1}{2}$. As we show in this work, the interesting range of values for the rank of a solution to the r-MRkM problem is between $\frac{r}{2}$ and r. In fact, we show that a solution with rank $\frac{r}{2}$ (i.e. of dimension $\frac{r-1}{2}$) can be found in polynomial time for the perfect r-MRkM problem (by Theorem 1.5 for example), thus this difference is not of great significance either.