

# On $k$ -Column Sparse Packing Programs

Nikhil Bansal\*    Nitish Korula†    Viswanath Nagarajan\*    Aravind Srinivasan‡

## Abstract

We consider the class of packing integer programs (PIPs) that are *column sparse*, i.e. there is a specified upper bound  $k$  on the number of constraints that each variable appears in. We give an  $ek+o(k)$ -approximation algorithm for  $k$ -column sparse PIPs, improving on recent results of  $k^2 \cdot 2^k$  [14] and  $O(k^2)$  [3, 5]. We also show that the integrality gap of our linear programming relaxation is at least  $2k - 1$ ; it is known that  $k$ -column sparse PIPs are  $\Omega(\frac{k}{\log k})$ -hard to approximate [8]. We also extend our result (at the loss of a small constant factor) to the more general case of maximizing a submodular objective over  $k$ -column sparse packing constraints.

## 1 Introduction

Packing integer programs (PIPs) are those of the form:

$$\max \{w^T x \mid Sx \leq c, x \in \{0, 1\}^n\}, \quad \text{where } w, c \in \mathbb{R}_+^n \text{ and } S \in \mathbb{R}_+^{m \times n}.$$

Above,  $n$  is the number of variables/columns and  $m$  is the number of rows/constraints. More generally, a PIP might have arbitrary upper-bounds on the variables: however, by replicating each such variable by several 0-1 variables, it suffices to consider those of the above form. In general PIPs are very hard to approximate, since a special case is the independent set problem, which is  $n^{1-o(1)}$ -hard to approximate [10]. In this paper, we consider  *$k$ -column sparse PIPs* (denoted  $k$ -CS-PIP), which is the special case of PIPs where the number of non-zero entries in each column of matrix  $S$  is upper bounded by a parameter  $k$ . A special case of  $k$ -CS-PIP is the  $k$ -set packing problem (when  $S$  is a 0-1 matrix and  $c$  is all ones), for which the best known approximation ratio is  $\frac{k+1}{2} + \epsilon$  [1] obtained using local search techniques; it is known to be  $\Omega(\frac{k}{\log k})$ -hard to approximate [8]. The projective plane instance of order  $k - 1$  also implies an integrality gap of  $k - 1 + 1/k$  for the natural LP relaxation of the  $k$ -set packing problem.

### 1.1 Results

We first study the natural LP relaxation for  $k$ -CS-PIP, and obtain an  $8k$ -approximation algorithm. Then we show that by adding some simple valid inequalities to this LP, one can obtain an improved  $(ek+o(k))$ -approximation algorithm. Our algorithm is based on randomized rounding with alterations. However, it needs to handle some subtle issues arising from conditioning. We also show that the integrality gap of this strengthened LP is at least  $2k - 1$ . Previously, a  $2^k \cdot k^2$  approximation for  $k$ -CS-PIP was given by Pritchard [14]; shortly after, the approximation ratio was improved to  $O(k^2)$  by Chekuri *et al.* [5] and

---

\*IBM T.J. Watson Research Center, Yorktown Heights, NY 10598. email: {nikhil,viswanath}@us.ibm.com.

†Dept. of Computer Science, University of Illinois, Urbana IL 61801. Partially supported by NSF grant CCF 07-28782. email: nkorula2@illinois.edu.

‡Dept. of Computer Science and Institute for Advanced Computer Studies, University of Maryland, College Park, MD 20742. Supported in part by NSF ITR Award CNS-0426683 and NSF Award CNS-0626636. email: srin@cs.umd.edu.

Chakrabarty-Pritchard [3]. Our  $O(k)$  approximation is nearly best possible since there is an  $\Omega(\frac{k}{\log k})$ -hardness of approximation [8].

We also consider the more general problem of maximizing a monotone submodular function over packing constraints that are  $k$ -column sparse. This problem is a common generalization of maximizing a submodular function over (a)  $k$ -dimensional knapsack [11], and (b)  $k$ -partition matroids [13]. Using the continuous greedy algorithm of Vondrák [17] in conjunction with our rounding algorithm, we obtain a  $(\frac{e^2 k}{e-1} + o(k))$ -approximation algorithm for the problem. However, the analysis of the approximation guarantee is more intricate: In particular, we need to generalize a result of Feige [7] showing that submodular functions are also *fractionally subadditive*. (See Section 5 for a statement of the new result, Theorem 5.3, and related context.) We believe that this theorem is also likely to be of independent interest in the study of submodular optimization.

## 1.2 Related Work

Pritchard [14] gave a  $2^k k^2$ -approximation algorithm for  $k$ -CS-PIP; this was the first result with approximation ratio depending only on  $k$ . Pritchard’s algorithm was based on solving an iterated LP relaxation, and then applying a certain randomized selection procedure. Independently, [5] and [3] showed that this final step could be derandomized, yielding an improved bound of  $O(k^2)$ . Prior to these results, an  $11.54k$ -approximation algorithm was known [6] for a special case of  $k$ -CS-PIP, the so-called column-restricted PIPs, where (i) in each column all non-zero entries are equal, and (ii) the maximum entry in  $S$  is at most the minimum entry in  $c$  (this is also known as the no bottle-neck assumption); later, it was observed in [4] that even without the second of these conditions, one can obtain an  $8k$  approximation<sup>1</sup>.

As mentioned previously, a well-studied special case of  $k$ -CS-PIP is the  $k$ -set packing problem, where the constraint matrix  $S$  is 0-1 and capacity  $c$  is all ones. The best-known approximation ratio for this problem is  $\frac{k+1}{2} + \epsilon$  [1] obtained using local search techniques. An improved bound of  $\frac{k}{2} + \epsilon$  is also known [9] when the weight vector  $w = \mathbf{1}$ . It is also known that the natural LP relaxation for this problem has a  $k - 1 + 1/k$  integrality gap, and in particular holds for the projective plane instance of order  $k - 1$ . Hazan et al [8] showed that  $k$ -set packing is  $\Omega(\frac{k}{\log k})$ -hard to approximate.

Shepherd and Vetta [15] studied the *demand matching* problem, which is  $k$ -CS-PIP with  $k = 2$  and the restriction that in each column the non-zero entries are equal. (That is, an instance of the demand matching problem is a column-restricted 2-column sparse PIP.) They gave an LP-based 3.264-approximation algorithm, and also showed that the natural LP relaxation for this problem has integrality gap at least 3.

## 1.3 Problem Definition

Let the items (i.e. columns) be indexed by  $i \in [n]$  and the constraints (i.e. rows) be indexed by  $j \in [m]$ . We consider the following packing integer program ( $k$ -CS-PIP).

$$\begin{aligned} \max \quad & \sum_{i=1}^n w_i \cdot x_i \\ & \sum_i s_{ij} \cdot x_i \leq c_j, \quad \forall j \in [m] \\ & x_i \in \{0, 1\}, \quad \forall i \in [n] \end{aligned}$$

where we have the restriction that for each column  $i$ , there are at most  $k$  non-zero entries in  $\{s_{ij}\}_{j=1}^m$ . By scaling the constraint matrix, we assume that  $c_j = 1$  for all  $j \in [m]$ . We also assume that  $s_{ij} \leq 1$  for each  $i, j$ ; otherwise, we can just fix  $x_i = 0$ .

<sup>1</sup>The problem setting and the results in [6] and [4] however are much more general. In particular, they show that the integrality gap of general column-restricted packing programs is within a constant factor of the corresponding 0-1 packing programs

We say that item  $i$  *participates* in constraint  $j$  if  $s_{ij} > 0$ ; for each constraint  $j$ , let  $P(j)$  denote the set of items participating in this constraint. For each  $i \in [n]$ , let  $N(i) := \{j \in [m] \mid s_{ij} > 0\}$  be the set of constraints that  $i$  participates in. In a  $k$ -column sparse PIP, we have  $|N(i)| \leq k$  for each  $i \in [n]$ .

**Outline.** All our algorithms are based on solving linear programs. To obtain an integral solution we round each variable independently; this gives a set of items of good expected weight, but some constraints may be violated. We then perform an *alteration* step, deleting some previously selected items; to obtain good approximation ratios, we bound the loss in this alteration step.

In Section 2, we study the natural LP relaxation for  $k$ -CS-PIP and present the  $8k$ -approximation algorithm. Then we consider a stronger LP relaxation and obtain an  $(ek + o(k))$ -approximation algorithm in Section 3. We also present the integrality gap for this strengthened LP in Subsection 3.1. In section 4, we show significantly better ratios if the capacities of all constraints are large relative to the sizes. Finally we consider column sparse packing problems over a submodular objective, and obtain an  $O(k)$ -approximation algorithm in Section 5.

## 2 The Algorithm for $k$ -CS-PIP

### 2.1 Preliminary Algorithms

Before presenting our final algorithm, we describe a (seemingly correct) algorithm that does not quite work, and a second algorithm that provides a weaker approximation ratio. Understanding where these algorithms fail and succeed will give useful insights into the design for the actual algorithm.

**A strawman Algorithm:** Consider the following algorithm. Let  $x$  be some optimum solution to the LP relaxation. For each element  $i \in [n]$ , select it independently at random with probability  $x_i/(2k)$ . Let  $\mathcal{S}$  be the chosen set of items. For any constraint  $j \in [m]$ , if it is violated, then discard all items  $i \in \mathcal{S}$  that appear in  $j$ , i.e. for which  $s_{ij} > 0$ .

Since the probabilities are scaled down by  $2k$ , by Markov's inequality any constraint  $j$  is violated with probability at most  $1/(2k)$ . Hence, any constraint will discard its items with probability at most  $1/2k$ . By the  $k$ -sparse property, each element can be discarded by at most  $k$  constraints, and hence by union bound over those  $k$  constraints, it is discarded with probability at most  $k \cdot (1/2k) = 1/2$ . Since an element is chosen in  $\mathcal{S}$  with probability  $x_i/2k$ , this implies that it lies in the overall solution with probability at least  $x_i/(4k)$ , implying that the proposed algorithm is a  $4k$  approximation.

However, the above argument is not correct. Consider the following example: Suppose there is a single constraint

$$Mx_1 + x_2 + x_3 + x_4 + \dots + x_M \leq M$$

where  $M \gg k$  is a large integer. Clearly, setting  $x_i = 1/2$  for  $i = 1, \dots, M$  is a feasible solution. Now consider the execution of the above algorithm. Note that whenever item 1 is chosen in  $\mathcal{S}$ , it is very likely that some item other than 1 will also be chosen (since  $M \gg k$  and we pick each item with probability  $x_i/2k = 1/4k$ ). Thus the algorithm will almost always end up discarding item 1, violating the claim that it lies in the solution with probability at least  $x_1/4k = 1/8k$ .

The key point is that one must consider the probability of an item being discarded by some constraint, *conditional* on it being chosen in the set  $\mathcal{S}$ . This is not a problem if either all item sizes are small (i.e. say  $s_{ij} \leq c_j/2$ ), or all item sizes are large (say  $s_{ij} \approx c_j$ ). The algorithm we analyze below shows that the difficult case is when some constraints contain both large and small items, as in the example above.

**A correct, but weak Algorithm:** We now show that the strawman algorithm *can* be used if we restrict the item sizes in the constraint matrix.

**Definition 2.1.** Item  $i$  is big for constraint  $j$  if  $s_{ij} > 1/2$ . Otherwise, if  $s_{ij} > 0$ , item  $i$  is small for constraint  $j$ .

**Lemma 2.2.** If, for each constraint  $j$ , either all items are big for constraint  $j$  or small for the constraint, there is a randomized  $O(k)$ -approximation algorithm for  $k$ -CS-PIP.

*Proof.* The simple strawman algorithm described above gives a good solution for such instances. We select each item independently with probability  $x_i/4k$ ;  $\mathcal{S}$  is the set of items chosen in this step. If constraint  $j$  is violated, discard all items  $i$  such that  $s_{ij} > 0$ ;  $\mathcal{S}'$  is the set of finally selected items. Clearly,  $\mathcal{S}'$  is a feasible solution. To prove that it has good expected weight, it suffices to show that for each item  $i$ , the probability that it is in  $\mathcal{S}'$  conditioned on being selected for  $\mathcal{S}$  is at least  $1/2$ . In particular, we prove that the probability item  $i$  is discarded due to any constraint  $j$  it participates in is at most  $1/2k$ .

First, suppose  $i$  is small for  $j$ :  $E[\sum_{i' \in \mathcal{S}} s_{i'j}] \leq 1/4k$ , and so by Markov's inequality, the probability that  $\sum_{i' \in \mathcal{S}} s_{i'j} x_{i'} \geq 1/2$  is at most  $1/2k$ . Constraint  $j$  is violated only if the total size of other items in  $\mathcal{S}$  that participate in  $j$  is at most  $1 - s_{ij} \geq 1/2$ . Thus, even if item  $i$  is selected, the probability that the total size of other items is more than  $1/2 \leq 1 - s_{ij}$  is at most  $1/2k$ .

Now suppose that item  $i$  is large for constraint  $j$ . Let  $P(j)$  denote the set of items that participate in this constraint; note that for all  $i' \in P(j)$ ,  $s_{i'j} \geq 1/2$ . As  $\sum_{i' \in P(j)} s_{i'j} x_{i'j} \leq 1$ , we have  $\sum_{i' \in P(j)} x_{i'j} \leq 2$ . Because items are selected for  $\mathcal{S}$  with probabilities scaled down by a factor of  $4k$ , the probability that any item in  $P(j) - \{i\}$  is selected is again at most  $1/2k$ .  $\square$

Finally, we prove that one can achieve the hypothesis of Lemma 2.2 at a cost of  $2^k$  in the approximation ratio.

**Lemma 2.3.** There is a randomized  $O(k \cdot 2^k)$ -approximation algorithm for  $k$ -CS-PIP.

*Proof.* For each constraint  $j$  independently, with probability  $1/2$  discard all items large for constraint  $j$ , and with probability  $1/2$  discard all items small for constraint  $j$ . Let  $\mathcal{I} \subseteq [n]$  denote the set of items that were not discarded by any constraint they participate in; for each item the probability that  $i \in \mathcal{I}$  is at least  $1/2^k$ . Thus, the expected value of the optimal solution on the items in  $\mathcal{I}$  is at least  $1/2^k$  times the optimal value on the entire set of items.

Now, for each constraint  $j$ , either all remaining items in  $\mathcal{I}$  are large for  $j$ , or all items in  $\mathcal{I}$  are small for  $j$ . By Lemma 2.2, there is an  $O(k)$ -approximation algorithm for the problem restricted to items in  $\mathcal{I}$ , and hence an  $O(k \cdot 2^k)$  approximation for the original problem.  $\square$

To simultaneously handle large and small items, we use the more careful alteration process described below.

## 2.2 An $O(k)$ -approximation

The algorithm first solves the LP relaxation to obtain an optimal fractional solution  $x$ . Then we round to an integral solution as follows. With foresight, set  $\alpha = 4$ .

1. Sample each item  $i \in [n]$  independently with probability  $x_i/(\alpha k)$ .  
Let  $\mathcal{S}$  denote the set of chosen items. We call an item in  $\mathcal{S}$  an  $\mathcal{S}$ -item.
2. For each item  $i$ , mark  $i$  (for deletion) if, for any constraint  $j \in N(i)$ , either:
  - $\mathcal{S}$  contains some *other* item  $i'$  which is big for constraint  $j$  or
  - The sum of sizes of  $\mathcal{S}$ -items that are small for  $j$  exceeds 1. (i.e. the capacity).
3. Delete all marked items, and return  $\mathcal{S}'$ , the set of remaining items.

**Analysis:** We will show that this algorithm gives an  $8k$  approximation.

**Lemma 2.4.** *Solution  $\mathcal{S}'$  is feasible with probability one.*

*Proof.* Consider any fixed constraint  $j \in [m]$ .

1. Suppose there is some  $i' \in \mathcal{S}'$  that is big for  $j$ . Then the algorithm guarantees that  $i'$  will be the only item in  $\mathcal{S}'$  (either small or big) that participates in constraint  $j$ : Consider any other  $\mathcal{S}$ -item  $i$  participating in  $j$ ;  $i$  must have been deleted from  $\mathcal{S}$  because  $\mathcal{S}$  contains another item (namely  $i'$ ) that is big for constraint  $j$ . Thus,  $i'$  is the only item in  $\mathcal{S}'$  participating in constraint  $j$ , and so the constraint is trivially satisfied, as all  $s_{ij} \leq 1$ .
2. The other case is when all items in  $\mathcal{S}'$  are small for  $j$ . Let  $i \in \mathcal{S}'$  be some item that is small for  $j$  (if there are none such, then constraint  $j$  is trivially satisfied). Since  $i$  was not deleted from  $\mathcal{S}$ , it must be that the total size of  $\mathcal{S}$ -items that are small for  $j$  could not have exceeded 1. Now,  $\mathcal{S}' \subseteq \mathcal{S}$ , and so this condition is also true for items in  $\mathcal{S}'$ .

Thus every constraint is satisfied by solution  $\mathcal{S}'$  and we obtain the lemma.  $\square$

We now prove the main theorem.

**Theorem 2.5.** *For any item  $i \in [n]$ , the probability  $\Pr[i \in \mathcal{S}' \mid i \in \mathcal{S}] \geq 1 - \frac{2}{\alpha}$ . Equivalently, the probability that item  $i$  is deleted from  $\mathcal{S}$  conditional on it being chosen in  $\mathcal{S}$  is at most  $2/\alpha$ .*

*Proof.* For any item  $i$  and constraint  $j \in N(i)$ , let  $B_{ij}$  denote the event that  $i$  is marked for deletion from  $\mathcal{S}$  because there is some other  $\mathcal{S}$ -item that is big for constraint  $j$ . Let  $G_j$  denote the event that the total size of  $\mathcal{S}$ -items that are small for constraint  $j$  exceeds 1. For any item  $i \in [n]$  and constraint  $j \in [m]$ , we will show that:

$$\Pr[B_{ij} \mid i \in \mathcal{S}] + \Pr[G_j \mid i \in \mathcal{S}] \leq \frac{2}{\alpha k} \quad (1)$$

We prove (1) using the following intuition: The total extent to which the LP selects items that are *big* for any constraint cannot be more than 2 (each big item has size at least  $1/2$ ); therefore,  $B_{ij}$  is unlikely to occur. Ignoring for a moment the conditioning on  $i \in \mathcal{S}$ ,  $G_j$  is also unlikely, as we scaled down probabilities by a factor  $\alpha k$ . But items are selected for  $\mathcal{S}$  independently, so if  $i$  is big for constraint  $j$ , then its presence in  $\mathcal{S}$  does not affect the event  $G_j$  at all. If  $i$  is small for constraint  $j$ , then *even if*  $i \in \mathcal{S}$ , the total size of  $\mathcal{S}$ -items is unlikely to exceed 1.

We now prove (1) formally, using some care to save a factor of 2. Let  $B(j)$  denote the set of items that are big for constraint  $j$ , and  $Y_j := \sum_{\ell \in B(j)} x_\ell$ . By the LP constraint for  $j$ , it follows that  $Y_j \leq 2$  (since each  $\ell \in B(j)$  has size  $s_{\ell j} > \frac{1}{2}$ ). Now by a union bound,

$$\Pr[B_{ij} \mid i \in \mathcal{S}] \leq \frac{1}{\alpha k} \sum_{\ell \in B(j) \setminus \{i\}} x_\ell \leq \frac{Y_j}{\alpha k} \leq \frac{2}{\alpha k}. \quad (2)$$

Now, let  $G_{-i}(j)$  denote the set of items that are small for constraint  $j$ , *not counting* item  $i$ , even if it is small. Using the LP constraint  $j$ , we have:

$$\sum_{\ell \in G_{-i}(j)} s_{\ell j} \cdot x_\ell \leq 1 - \sum_{\ell \in B(j)} s_{\ell j} \cdot x_\ell \leq 1 - \frac{Y_j}{2}. \quad (3)$$

Since each item  $i'$  is chosen into  $\mathcal{S}$  with probability  $x_{i'}/(\alpha k)$ , inequality (3) implies that the expected total size of  $\mathcal{S}$ -items in  $G_{-i}(j)$  is at most  $\frac{1}{\alpha k} (1 - Y_j/2)$ . By Markov's inequality, the probability that the total size of these  $\mathcal{S}$ -items exceeds  $1/2$  is at most  $\frac{2}{\alpha k} (1 - Y_j/2)$ . Since items are chosen independently and  $i \notin G_{-i}(j)$ , we obtain this probability even conditioned on  $i \in \mathcal{S}$ .

If  $i$  is big for  $j$ , event  $G_j$  occurs only if the total size of  $\mathcal{S}$ -items in  $G_{-i}(j)$  exceeds 1. If  $i$  is small for  $j$ , event  $G_j$  occurs only if the total size of small  $\mathcal{S}$ -items participating in  $j$  exceeds 1; as  $s_{ij} \leq 1/2$ , the total size of  $\mathcal{S}$ -items in  $G_{-i}(j)$  must exceed  $1/2$ . Thus, whether  $i$  is big or small,

$$\Pr[G_j \mid i \in \mathcal{S}] \leq \frac{2}{\alpha k} \left(1 - \frac{Y_j}{2}\right) = \frac{2}{\alpha k} - \frac{Y_j}{\alpha k}.$$

Combined with inequality (2) we obtain (1):

$$\Pr[B_{ij} \mid i \in \mathcal{S}] + \Pr[G_j \mid i \in \mathcal{S}] \leq \frac{Y_j}{\alpha k} + \Pr[G_j \mid i \in \mathcal{S}] \leq \frac{Y_j}{\alpha k} + \frac{2}{\alpha k} - \frac{Y_j}{\alpha k} = \frac{2}{\alpha k}.$$

To see that (1) implies the theorem, for any item  $i$ , simply take the union bound over all  $j \in N(i)$ . Thus, the probability that  $i$  is deleted from  $\mathcal{S}$  conditional on it being chosen in  $\mathcal{S}$  is at most  $2/\alpha$ . Equivalently,  $\Pr[i \in \mathcal{S}' \mid i \in \mathcal{S}] \geq 1 - 2/\alpha$ .  $\square$

We are now ready to prove the final result.

**Theorem 2.6.** *There is a randomized  $8k$ -approximation algorithm for  $k$ -CS-PIP.*

*Proof.* First observe that our algorithm always outputs a feasible solution (Lemma 2.4). To bound the objective value, recall that  $\Pr[i \in \mathcal{S}] = \frac{x_i}{\alpha k}$  for all  $i \in [n]$ . Hence Theorem 2.5 implies that

$$\Pr[i \in \mathcal{S}'] \geq \Pr[i \in \mathcal{S}] \cdot \Pr[i \in \mathcal{S}' \mid i \in \mathcal{S}] \geq \frac{x_i}{\alpha k} \cdot \left(1 - \frac{2}{\alpha}\right)$$

for all  $i \in [n]$ . Finally using linearity of expectation and  $\alpha = 4$ , we obtain the theorem.  $\square$

*Remark:* We note that the analysis above only uses Markov's inequality conditioned on a single item being chosen in set  $\mathcal{S}$ . Thus a pairwise independent distribution suffices to choose the set  $\mathcal{S}$ , and hence the algorithm can be easily derandomized.

### 3 Improved Approximation Ratio for $k$ -CS-PIP

In this section we give an improved  $(ek + o(k))$ -approximation algorithm for  $k$ -column sparse PIPs. This is tight up to a very small constant factor (less than 1.36). The improvement comes from *four* different steps. First, we use a slightly stronger LP relaxation. Second, we modify the alteration/deletion step to obtain  $\mathcal{S}'$  from  $\mathcal{S}$ ; the previous algorithm sometimes deleted items even when constraints were not violated, or deleted more items than necessary. Third, we are more careful in analyzing the probability that constraint  $j$  causes item  $i$  to be deleted from the set  $\mathcal{S}$ ; we take into account a certain discreteness of item sizes. And fourth, for each item  $i$ , we use the FKG inequality instead of the weaker union bound over all constraints in  $N(i)$ .

**Stronger LP relaxation:** Consider the following stronger LP relaxation. For each constraint  $j \in [m]$ , let  $B(j) = \{i \in [n] \mid s_{ij} > \frac{1}{2}\}$  denote the set of big items. Since no two items that are big for some constraint can be chosen in an integral solution, we add the following valid inequalities to the natural LP relaxation:

$$\sum_{i \in B(j)} x_i \leq 1, \quad \forall j \in [m]. \quad (4)$$

**Algorithm:** The algorithm obtains an optimal LP solution  $x$ , and rounds it to an integral solution  $\mathcal{S}'$  as follows (parameter  $\alpha$  will be set to 1 later).

1. Sample each item  $i \in [n]$  independently with probability  $x_i/(\alpha k)$ .  
Let  $\mathcal{S}$  denote the set of chosen items.
2. For any item  $i$  and constraint  $j \in N(i)$ , let  $E_{ij}$  denote the event that the items  $\{i' \in \mathcal{S} \mid s_{i'j} \geq s_{ij}\}$  have total size (in constraint  $j$ ) exceeding one. Mark  $i$  for deletion if  $E_{ij}$  occurs for any  $j \in N(i)$ .
3. Return set  $\mathcal{S}' \subseteq \mathcal{S}$  consisting of all items  $i \in \mathcal{S}$  not marked for deletion.

Note the different rule for deleting an item from  $\mathcal{S}$ . In particular, whether item  $i$  is deleted from constraint  $j$  only depends on items that are at least as large as  $i$  in  $j$ .

**Analysis:** It is clear that  $\mathcal{S}'$  is feasible with probability one. In the following we show that each item appears in  $\mathcal{S}'$  with good probability.

**Lemma 3.1.** *The probability  $\Pr[E_{ij} \mid i \in \mathcal{S}]$  is at most  $\frac{1}{\alpha k} \left(1 + \left(\frac{2}{\alpha k}\right)^{1/3}\right)$*

*Proof.* Let  $\ell := (4\alpha k)^{1/3}$ . We classify items in relation to constraints as:

- Item  $i \in [n]$  is *big* for constraint  $j \in [m]$  if  $s_{ij} > \frac{1}{2}$ .
- Item  $i \in [n]$  is *medium* for constraint  $j \in [m]$  if  $\frac{1}{\ell} \leq s_{ij} \leq \frac{1}{2}$ .
- Item  $i \in [n]$  is *tiny* for constraint  $j \in [m]$  if  $s_{ij} < \frac{1}{\ell}$ .

For any constraint  $j \in [m]$ , let  $B(j), M(j), T(j)$  respectively denote the set of big, medium, tiny items for  $j$ . In the next three claims, we bound  $\Pr[E_{ij} \mid i \in \mathcal{S}]$  when item  $i$  is big, medium, and small respectively.

**Claim 3.2.** *For any  $i \in [n]$  and constraint  $j \in [m]$  such that item  $i$  is big for constraint  $j$ ,*

$$\Pr[E_{ij} \mid i \in \mathcal{S}] \leq \frac{1}{\alpha k}.$$

*Proof.* The event  $E_{ij}$  occurs if some item that is at least as large as  $i$  for constraint  $j$  is chosen in  $\mathcal{S}$ . Since  $i$  is big in constraint  $j$ ,  $E_{ij}$  occurs only if some big item other than  $i$  is chosen for  $\mathcal{S}$ . Now by the union bound, the probability that some item from  $B(j) \setminus \{i\}$  is chosen into  $\mathcal{S}$  is:

$$\Pr \left[ (B(j) \setminus \{i\}) \cap \mathcal{S} \neq \emptyset \mid i \in \mathcal{S} \right] \leq \sum_{i' \in B(j) \setminus \{i\}} \frac{x_{i'}}{\alpha k} \leq \frac{1}{\alpha k} \sum_{i' \in B(j)} x_{i'} \leq \frac{1}{\alpha k},$$

where the last inequality follows from the new LP constraint (4) on big items for  $j$ . □

**Claim 3.3.** *For any  $i \in [n], j \in [m]$  such that item  $i$  is medium for constraint  $j$ ,*

$$\Pr[E_{ij} \mid i \in \mathcal{S}] \leq \frac{1}{\alpha k} \left(1 + \frac{\ell^2}{2\alpha k}\right).$$

*Proof.* Here, if event  $E_{ij}$  occurs then it must be that either some big item is chosen or (otherwise) at least two medium items other than  $i$  are chosen, i.e.  $E_{ij}$  implies that either  $\mathcal{S} \cap B(j) \neq \emptyset$  or  $|\mathcal{S} \cap (M(j) \setminus \{i\})| \geq 2$ . This is because  $i$  together with any *one* other medium item is not enough to reach the capacity of constraint  $j$ . (Since  $i$  is medium, we do not consider tiny items for constraint  $j$  in determining whether  $i$  should be deleted.)

Just as in Claim 3.2, we have that the probability some big item for  $j$  is chosen is at most  $1/\alpha k$ , i.e.  $\Pr[\mathcal{S} \cap B(j) \neq \emptyset \mid i \in \mathcal{S}] \leq \frac{1}{\alpha k}$ .

Now consider the probability that  $|\mathcal{S} \cap (M(j) \setminus \{i\})| \geq 2$ , conditioned on  $i \in \mathcal{S}$ . We will show that this probability is much smaller than  $1/\alpha k$ . Since each item  $h \in M(j) \setminus \{i\}$  is chosen independently with probability  $\frac{x_h}{\alpha k}$  (even given  $i \in \mathcal{S}$ ):

$$\Pr \left[ |\mathcal{S} \cap (M(j) \setminus \{i\})| \geq 2 \mid i \in \mathcal{S} \right] \leq \frac{1}{2} \cdot \left( \sum_{h \in M(j)} \frac{x_h}{\alpha k} \right)^2 \leq \frac{\ell^2}{2\alpha^2 k^2}$$

where the last inequality follows from the fact that

$$1 \geq \sum_{h \in M(j)} s_{hj} \cdot x_h \geq \frac{1}{\ell} \sum_{h \in M(j)} x_h$$

(recall each item in  $M(j)$  has size at least  $\frac{1}{\ell}$ ). Combining these two cases, we have the desired upper bound on  $\Pr[E_{ij} \mid i \in \mathcal{S}]$ .  $\square$

**Claim 3.4.** For any  $i \in [n]$ ,  $j \in [m]$  such that item  $i$  is tiny for constraint  $j$ ,

$$\Pr[E_{ij} \mid i \in \mathcal{S}] \leq \frac{1}{\alpha k} \left( 1 + \frac{2}{\ell} \right).$$

*Proof.* Since  $i$  is tiny, if event  $E_{ij}$  occurs then the total size (in constraint  $j$ ) of items  $\mathcal{S} \setminus \{i\}$  is greater than  $1 - \frac{1}{\ell}$ . So,

$$\Pr[E_{ij} \mid i \in \mathcal{S}] \leq \Pr \left[ \sum_{h \in \mathcal{S} \setminus \{i\}} s_{hj} > 1 - \frac{1}{\ell} \right] \leq \frac{1}{\alpha k} \cdot \frac{\ell}{\ell - 1} \leq \frac{1}{\alpha k} \left( 1 + \frac{2}{\ell} \right)$$

where the first inequality follows from the above observation and the fact that  $\mathcal{S} \setminus \{i\}$  is independent of the event  $i \in \mathcal{S}$ , the second is Markov's inequality, and the last uses  $\ell \geq 2$ .  $\square$

Thus, for any item  $i$  and constraint  $j \in N(i)$ ,  $\Pr[E_{ij} \mid i \in \mathcal{S}] \leq \frac{1}{\alpha k} \max\{(1 + \frac{2}{\ell}), (1 + \frac{\ell^2}{2\alpha k})\}$ . From the choice of  $\ell = (4\alpha k)^{1/3}$ , which makes the probability in Claims 3.3 and 3.4 equal, we obtain the lemma.  $\square$

We now prove the main result of this section

**Theorem 3.5.** For each  $i \in [n]$ , probability  $\Pr[i \in \mathcal{S}' \mid i \in \mathcal{S}] \geq (1 - \frac{1}{\alpha k} (1 + (\frac{2}{\alpha k})^{1/3}))^k$ .

*Proof.* For any item  $i$  and constraint  $j \in N(i)$ , the event that  $E_{ij}$  does not occur is a decreasing function (over the choice of the other items in the set  $\mathcal{S}$ ). Thus, by the FKG inequality, for item  $i$  the probability that no event  $E_{ij}$  occurs:

$$\Pr \left[ \bigwedge_{j \in N(i)} \neg E_{ij} \right] \geq \prod_{j \in N(i)} \Pr[\neg E_{ij}]$$

From Lemma 3.1,  $\Pr[\neg E_{ij}] \geq 1 - \frac{1}{\alpha k} (1 + (\frac{2}{\alpha k})^{1/3})$ . As each item is in at most  $k$  constraints, we obtain the theorem.  $\square$

Now, by setting  $\alpha = 1^2$  we have  $\Pr[i \in \mathcal{S}] = 1/k$ , and  $\Pr[i \in \mathcal{S}' \mid i \in \mathcal{S}] \geq \frac{1}{e+o(1)}$ , which immediately implies:

**Theorem 3.6.** There is a randomized  $(ek + o(k))$ -approximation algorithm for  $k$ -CS-PIP.

---

<sup>2</sup>Note that this is optimal only asymptotically; in the case of  $k = 2$ , for instance, it is better to choose  $\alpha \approx 2.8$ .

### 3.1 Integrality Gap

Recall that the LP relaxation for the  $k$ -set packing problem has an integrality gap of  $k - 1 + 1/k$ , as shown by the instance given by the projective plane of order  $k - 1$ . If we set each right hand side of the LP to  $2 - \epsilon$ , this directly implies an integrality gap arbitrarily close to  $2(k - 1 + 1/k)$  for the (weak) LP relaxation for  $k$ -CS-PIP. This is because the LP can set each  $x_i = (2 - \epsilon)/k$  hence obtaining a profit of  $(2 - \epsilon)(k - 1 + 1/k)$ , while the integral solution can only choose one item. However, for our stronger LP relaxation used in the previous section, this does not work and the projective plane instance only implies a gap of  $k - 1 + 1/k$  (note that each item is big in every constraint that it appears in).

However, using another instance of  $k$ -CS-PIP, we show that even the stronger LP relaxation has an integrality gap at least  $2k - 1$ . Consider the instance on  $n = m = 2k - 1$  items and constraints defined as follows. We view the indices  $[n] = \{0, 1, \dots, n - 1\}$  as integers modulo  $n$ . The weights  $w_i = 1$  for all  $i \in [n]$ . The sizes are:

$$s_{ij} := \begin{cases} 1 & \text{if } i = j \\ \epsilon & \text{if } j \in \{i + 1, \dots, i + k - 1 \pmod{n}\} \\ 0 & \text{otherwise} \end{cases}, \quad \forall i, j \in [n].$$

where  $\epsilon > 0$  is arbitrarily small, in particular  $\epsilon \ll \frac{1}{nk}$ .

Observe that setting  $x_i = 1 - k\epsilon$  for all  $i \in [n]$  is a feasible fractional solution to the strengthened LP. Thus the optimal LP value is at least  $(1 - k\epsilon) \cdot n \approx n = 2k - 1$ .

On the other hand, we claim that the optimal integral solution can only choose one item and hence has value 1. For the sake of contradiction, suppose it chooses two items  $i, h \in [n]$ . Then there is some constraint (either one corresponding to  $j = i$  or  $j = h$ ) that implies that either  $x_i + \epsilon \cdot x_h \leq 1$  or  $x_h + \epsilon \cdot x_i \leq 1$ ; this constraint is violated.

Thus the integrality gap of the LP we consider is at least  $2k - 1$ , for every  $k \geq 1$ .

## 4 Better Bounds for Large Capacities

A relevant parameter in studying packing integer programs [16] is:

$$B := \min_{i \in [n], j \in [m]} \frac{c_j}{s_{ij}},$$

which is a measure of how large the capacities are relative to the sizes. In this section, we consider the  $k$ -CS-PIP problem as a function of  $B$  and obtain a better approximation ratio of  $O(k^{1/\lfloor B \rfloor})$ ; we also give a matching integrality gap.<sup>3</sup> Previously, Pritchard [14] studied  $k$ -CS-PIP when  $B > k$  and obtained a ratio of  $(1 + k/B)/(1 - k/B)$ ; in contrast, we obtain improved approximation ratios even when  $B = 2$ . In this section, it will be convenient to assume that the entries are scaled so that for every constraint  $j \in [m]$ ,  $\max_{i \in P(j)} s_{ij} = 1$ . So  $B = \min_{j \in [m]} c_j \geq 1$ .

Set  $\alpha := 8e \cdot k^{1/\lfloor B \rfloor}$ . The algorithm first solves the natural LP relaxation for  $k$ -CS-PIP to obtain fractional solution  $x$ . Then it proceeds as follows.

1. Sample each item  $i \in [n]$  independently with probability  $x_i/\alpha$ .  
Let  $\mathcal{S}$  denote the set of chosen items.
2. Define *new sizes* as follows: for every item  $i$  and constraint  $j \in N(i)$ , round up  $s_{ij}$  to  $t_{ij}$ , the next larger power of 2.
3. Set  $\mathcal{S}' \subseteq \mathcal{S}$  consists of all items  $i \in \mathcal{S}$  such that for every  $j \in N(i)$ , the items  $\{i' \in \mathcal{S} \mid t_{i'j} \geq t_{ij}\}$  have total  $t$ -size (in constraint  $j$ ) at most  $c_j$ .

<sup>3</sup>In this section, we do not attempt to optimize constants; some of the techniques from the previous section could be applied here to improve the approximation ratio.

It is clear that  $\mathcal{S}'$  is a feasible solution with probability one, since the  $s$ -sizes are at most the new  $t$ -sizes.

The approximation guarantee is proved using the following theorem.

**Theorem 4.1.** *For each  $i \in [n]$ , probability  $\Pr[i \in \mathcal{S}' \mid i \in \mathcal{S}] \geq \frac{1}{2}$ .*

*Proof.* Fix any  $i \in [n]$  and  $j \in N(i)$ . Let  $E_{ij}$  denote the event that items  $\{i' \in \mathcal{S} \mid t_{i'j} \geq t_{ij}\}$  have total  $t$ -size (in constraint  $j$ ) greater than  $c_j$ . As before,  $\Pr[i \notin \mathcal{S}' \mid i \in \mathcal{S}] \leq \sum_{j \in N(i)} \Pr[E_{ij} \mid i \in \mathcal{S}]$ .

We show that  $\Pr[E_{ij} \mid i \in \mathcal{S}] \leq \frac{1}{2k}$ , which suffices to prove the theorem. Let  $t_{ij} = 2^{-\ell}$ , where  $\ell \in \mathbb{N}$ . Observe that all the  $t$ -sizes that are at least  $2^{-\ell}$  are actually integral multiples of  $2^{-\ell}$  (since they are all powers of two). Let  $\mathcal{I}_{ij} = \{i' \in [n] \mid t_{i'j} \geq t_{ij}\} \setminus \{i\}$ , and  $Y_{ij} := \sum_{i' \in \mathcal{I}_{ij}} t_{i'j} \cdot \mathbb{I}_{i' \in \mathcal{S}}$  where  $\mathbb{I}_{i' \in \mathcal{S}}$  are indicator random variables. The previous observation implies that  $Y_{ij}$  is always an integral multiple of  $2^{-\ell}$ . Note that

$$\begin{aligned} \Pr[E_{ij} \mid i \in \mathcal{S}] &= \Pr[Y_{ij} > c_j - 2^{-\ell} \mid i \in \mathcal{S}] \leq \Pr[Y_{ij} > \lfloor c_j \rfloor - 2^{-\ell} \mid i \in \mathcal{S}] \\ &= \Pr[Y_{ij} \geq \lfloor c_j \rfloor \mid i \in \mathcal{S}], \end{aligned}$$

where the last equality uses the fact that  $Y_{ij}$  is always a multiple of  $2^{-\ell}$ . Since each item is included into  $\mathcal{S}$  independently, we also have  $\Pr[Y_{ij} \geq \lfloor c_j \rfloor \mid i \in \mathcal{S}] = \Pr[Y_{ij} \geq \lfloor c_j \rfloor]$ . Now  $Y_{ij}$  is the sum of independent  $[0, 1]$  random variables with mean:

$$E[Y_{ij}] = \sum_{i' \in \mathcal{I}_{ij}} t_{i'j} \cdot \Pr[i' \in \mathcal{S}] \leq \sum_{i'=1}^n t_{i'j} \cdot \frac{x_{i'}}{\alpha} \leq \frac{2}{\alpha} \sum_{i'=1}^n s_{i'j} \cdot x_{i'} \leq \frac{2}{\alpha} c_j.$$

Choose  $\delta$  such that  $(\delta + 1) \cdot E[Y_{ij}] = \lfloor c_j \rfloor$ , i.e. (using  $c_j \geq 1$ ),

$$\delta + 1 = \frac{\lfloor c_j \rfloor}{E[Y_{ij}]} \geq \frac{\alpha \lfloor c_j \rfloor}{2 \cdot c_j} \geq \frac{\alpha}{4}.$$

Now using Chernoff Bound [12], we have:

$$\Pr[Y_{ij} \geq \lfloor c_j \rfloor] = \Pr[Y_{ij} \geq (1 + \delta) \cdot E[Y_{ij}]] \leq \left( \frac{e}{\delta + 1} \right)^{\lfloor c_j \rfloor} \leq \left( \frac{4e}{\alpha} \right)^{\lfloor c_j \rfloor} \leq \left( \frac{4e}{\alpha} \right)^{\lfloor B \rfloor}.$$

The last inequality uses the fact that  $c_j \geq B$ . Finally, since  $\alpha = 8e \cdot k^{1/\lfloor B \rfloor}$ , we obtain that  $\Pr[Y_{ij} \geq \lfloor c_j \rfloor] \leq \frac{1}{2k}$ . From the above, this implies  $\Pr[E_{ij} \mid i \in \mathcal{S}] \leq \frac{1}{2k}$ , which completes the proof of the theorem.  $\square$

Using this, it follows that the expected weight of solution  $\mathcal{S}'$  is at least  $\sum_{i=1}^n w_i x_i / (16e k^{1/\lfloor B \rfloor})$ . Thus we obtain:

**Theorem 4.2.** *There is an  $O(k^{1/\lfloor B \rfloor})$ -approximation algorithm for  $k$ -CS-PIP.*

## 4.1 Integrality Gap for General $B$

We show that the natural LP relaxation for  $k$ -CS-PIP has an  $\Omega(k^{1/\lfloor B \rfloor})$  integrality gap for every  $B \geq 1$ , matching the above approximation ratio up to constant factors.

For any  $B \geq 1$ , let  $t := \lfloor B \rfloor$ . We construct an instance of  $k$ -CS-PIP with  $n$  columns and  $m = \binom{n}{t+1}$  constraints. For all  $i \in [n]$ , weight  $w_i = 1$ .

For every  $(t+1)$ -subset  $C \subseteq [n]$ , there is a constraint  $j(C)$  involving the variables in  $C$ : set  $s_{i,j(C)} = 1$  for all  $i \in C$ , and  $s_{i,j(C)} = 0$  for  $i \notin C$ . For each constraint  $j \in [m]$ , the capacity  $c_j = B$ . Note that the column sparsity  $k = \binom{n-1}{t} \leq (ne/t)^t$ .

Setting  $x_i = \frac{1}{2}$  for all  $i \in [n]$  is a feasible fractional solution. Indeed, each constraint is occupied to extent  $\frac{t+1}{2} \leq \frac{B+1}{2} \leq B$  (since  $B \geq 1$ ). Thus the optimal LP value is at least  $\frac{n}{2}$ .

On the other hand, the optimal integral solution has value at most  $t$ . Suppose for contradiction that the solution contains some  $t + 1$  items, indexed by  $C \subseteq [n]$ . Then consider the constraint  $j(C)$ , which is occupied to extent  $t + 1 = \lfloor B \rfloor + 1 > B$ , this contradicts the feasibility of the solution! Thus the integral optimum is  $t$ , and the integrality gap for this instance is at least  $\frac{n}{2t} \geq \frac{1}{2e} k^{1/\lfloor B \rfloor}$ .

## 5 Submodular Objective Functions

We now consider the more general case when the objective we seek to maximize is an arbitrary *monotone submodular function*  $f : 2^{[n]} \rightarrow \mathbb{R}_+$ . The problem we consider is:

$$\max \left\{ f(T) \mid \sum_{i \in T} s_{ij} \leq c_j, \forall j \in [m]; T \subseteq [n] \right\} \quad (5)$$

Again, we let  $k$  denote the column-sparseness of the underlying constraint matrix. Observe that this problem is a common generalization of maximizing submodular functions over:  $k$  partition matroids, and  $k$  knapsack constraints. Several interesting results on constrained submodular maximization have been obtained recently [2, 17, 11].

We obtain an  $O(k)$ -approximation algorithm for this Problem (5). The *extension-by-expectation* (also called the *multi-linear extension*) of a submodular function  $f$  is a continuous function  $F : [0, 1]^n \rightarrow \mathbb{R}_+$  defined as follows:

$$F(x) := \sum_{T \subseteq [n]} \prod_{i \in T} x_i \cdot \prod_{j \notin T} (1 - x_j) \cdot f(T)$$

Note that  $F(x) = f(x)$  for  $x \in \{0, 1\}^n$  and hence  $F$  is an extension of  $f$ . Even though  $F$  is a non-linear function, using the continuous greedy algorithm from Vondrák [17], we can obtain a  $(1 - \frac{1}{e})$ -approximation algorithm to the following *fractional relaxation* of (5).

$$\max \left\{ F(x) \mid \sum_{i=1}^n s_{ij} \cdot x_i \leq c_j, \forall j \in [m]; 0 \leq x_i \leq 1, \forall i \in [n] \right\} \quad (6)$$

In order to apply the algorithm from [17], one needs to solve in polynomial time the problem of maximizing a *linear* objective over the constraints  $\{\sum_{i=1}^n s_{ij} \cdot x_i \leq c_j, \forall j \in [m]; 0 \leq x_i \leq 1, \forall i \in [n]\}$ . This is indeed possible since it reduces to solving a linear program on  $n$  variables and  $m$  constraints.

Let  $x$  denote any feasible solution to Problem (6). We apply the rounding algorithm for the additive case (from the previous sections), to obtain a *feasible integral solution*  $\mathcal{S}' \subseteq [n]$  of expected value at least  $\Omega(1/k)$  times that of the fractional solution  $x$ .

A useful result is the following lemma of Feige [7], showing that submodular functions are also *fractionally subadditive*:

**Lemma 5.1** ([7]). *Let  $\mathcal{U}$  be a set of elements and  $\{\mathcal{A}_t \subseteq \mathcal{U}\}$  be a collection of subsets with non-negative weights  $\{\lambda_t\}$  such that  $\sum_{t: i \in \mathcal{A}_t} \lambda_t \geq 1$  for all elements  $i \in \mathcal{U}$ . Then, for any submodular function  $f$ , we have  $f(\mathcal{U}) \leq \sum_t \lambda_t f(\mathcal{A}_t)$ .*

**Lemma 5.2.** *For any  $x \in [0, 1]^n$  and  $0 \leq p \leq 1$ , let set  $\mathcal{S}$  be constructed by selecting each item  $i \in [n]$  independently with probability  $p \cdot x_i$ . Then,  $E[f(\mathcal{S})] \geq pF(x)$ .*

*Proof.* Consider the following equivalent procedure for constructing  $\mathcal{S}$ : First, construct  $\mathcal{S}_0$  by selecting each item  $i$  with probability  $x_i$ . Then construct  $\mathcal{S}$  by retaining each element in  $\mathcal{S}_0$  independently with probability  $p$ .

By definition  $E[f(\mathcal{S}_0) = F(x)]$ . For any fixed set  $T \subseteq [n]$ , let  $p_T = \Pr[\mathcal{S}_0 = T]$ . Now *conditioned* on  $\mathcal{S}_0 = T$ , set  $\mathcal{S} \subseteq \mathcal{S}_0$  is a random subset such that  $\Pr[i \in \mathcal{S} \mid \mathcal{S}_0 = T] = p$  for all  $i \in \mathcal{S}_0$ . Thus by Lemma 5.1, we have  $E[f(\mathcal{S}) \mid \mathcal{S}_0 = T] \geq p \cdot f(T)$ . Hence:

$$E[f(\mathcal{S})] = \sum_{T \subseteq [n]} \Pr[\mathcal{S}_0 = T] \cdot E[f(\mathcal{S}) \mid \mathcal{S}_0 = T] \geq \sum_{T \subseteq [n]} \Pr[\mathcal{S}_0 = T] p \cdot f(T) = p E[f(\mathcal{S}_0)] = p F(x).$$

□

Recall that our rounding algorithms first independently select each element with probability  $x_i/\alpha k$  to form a (possibly infeasible) integral set  $\mathcal{S}$ . From the previous lemma, we immediately obtain  $E[f(\mathcal{S})] \geq \frac{1}{\alpha k} F(x)$ . Next, we prune  $\mathcal{S}$  to obtain a feasible integral  $\mathcal{S}'$ ; we only need to show that, in expectation, this alteration step does not discard too much of the value of  $\mathcal{S}$ . Setting  $\alpha = 4$ , it follows from Theorem 2.5 that  $\Pr[i \in \mathcal{S}' \mid i \in \mathcal{S}] \geq 1/2$ . Thus, we would like to use the subadditivity claimed in Lemma 5.1 to show that (for instance)  $E[f(\mathcal{S}')] \geq \frac{1}{2} E[f(\mathcal{S})] \geq \frac{1}{2\alpha k} F(x)$ .

Unfortunately, the property  $\Pr[i \in \mathcal{S}' \mid i \in \mathcal{S}] \geq 1/2$  does not imply  $E[f(\mathcal{S}')] \geq \frac{1}{2c} E[f(\mathcal{S})]$  for any constant  $c$ . Consider the following example: Let set  $\mathcal{S} \subseteq [n]$  be drawn from the following distribution:

- With probability  $1/2n$ ,  $\mathcal{S} = [n]$ .
- For each  $i \in [n]$ ,  $\mathcal{S} = \{i\}$  with probability  $1/2n$ .
- With probability  $1/2 - 1/2n$ ,  $\mathcal{S} = \emptyset$ .

Now define  $\mathcal{S}' = \mathcal{S}$  if  $\mathcal{S} = [n]$ , and  $\mathcal{S}' = \emptyset$  otherwise. Also, define a submodular function  $f$  as follows: For each  $T \subseteq [n]$  such that  $T \neq \emptyset$ , set  $f(T) = 1$ , and set  $f(\emptyset) = 0$ . Now,  $E[f(\mathcal{S})] = 1/2 + 1/2n$ , and for each  $i \in [n]$ , we have  $\Pr[i \in \mathcal{S}' \mid i \in \mathcal{S}] = 1/2$ . On the other hand,  $E[f(\mathcal{S}')]$  is only  $1/2n \ll 1/4$ .

Why, then, does our algorithm provide a good approximation? We exploit two properties of the rounding that are *not* present in the previous example.

1. The sets  $\mathcal{S}$  constructed by our algorithm are drawn from a product distribution on the items; in contrast, the example has either  $\mathcal{S} = [n]$  or  $|\mathcal{S}| \leq 1$ .
2. Our alteration procedure has the following monotonicity property: Suppose  $T_1 \subseteq T_2 \subseteq [n]$ , and  $i \in \mathcal{S}'$  when  $\mathcal{S} = T_2$ . Then we are guaranteed that  $i \in \mathcal{S}'$  when  $\mathcal{S} = T_1$ . (That is, if  $\mathcal{S}$  contains additional items, it is more likely that  $i$  will be discarded by some constraint it participates in.) By contrast, the example above has  $\mathcal{S}' = \mathcal{S}$  if  $\mathcal{S} = [n]$ , and  $\mathcal{S}' = \emptyset$  if  $|\mathcal{S}| = 1$ .

These two properties suffice to show that  $\Pr[i \in \mathcal{S}' \mid i \in \mathcal{S}] \geq 1/2$  implies that  $E[f(\mathcal{S}')] \geq \frac{1}{2} E[f(\mathcal{S})]$ . The intuition is as follows: Because  $f$  is submodular, the marginal contribution of item  $i$  to  $\mathcal{S}$  is largest precisely when  $\mathcal{S}$  is “small”, and this is also the case when  $i$  is most likely to be retained for  $\mathcal{S}'$ . More precisely, both  $\Pr[i \in \mathcal{S}' \mid i \in \mathcal{S}]$  and the marginal contribution of  $i$  to  $f(\mathcal{S})$  are *decreasing* functions of  $\mathcal{S}$ . Thus, we prove the following generalization of Lemma 5.1, which allows us to aggregate the contributions of  $i$  over outcomes for  $\mathcal{S}$ .

**Theorem 5.3.** *Let  $\mathcal{U}$  denote the set of elements  $[n]$ , and  $x \in [0, 1]^n$ . Let  $B_\ell$ ,  $1 \leq \ell \leq 2^n$  be randomly chosen subsets of  $\mathcal{U}$ , with the probability  $p_\ell$  of choosing  $B_\ell$  given by the usual product distribution, i.e.  $p_\ell = \prod_{i \in B_\ell} x_i \prod_{i \notin B_\ell} (1 - x_i)$ . Associated with each  $B_\ell$ , we have an arbitrary distribution  $q_\ell(t)$  over subsets  $A_{\ell t}$  of  $B_\ell$ . That is, exactly one  $A_{\ell t}$  is associated with  $B_\ell$  with probability  $q_\ell(t)$ . We assume that for each  $\ell$ ,  $\sum_t q_\ell(t) = 1$ .*

Suppose that the system satisfies the following conditions.

*Covering Property:*

$$\forall i, \sum_{\ell} p_{\ell} \sum_{t: i \in A_{\ell t}} q_{\ell}(t) \geq \beta \cdot \sum_{\ell: i \in B_{\ell}} p_{\ell}. \quad (7)$$

*Monotonicity:* For any two sets  $B_\ell$  and  $B_{\ell'}$  such that  $B_\ell \subseteq B_{\ell'}$  we have

$$\forall i \in B_\ell \cap B_{\ell'}, \quad \sum_{t:i \in A_{\ell t}} q_\ell(t) \geq \sum_{t:i \in A_{\ell' t}} q_{\ell'}(t) \quad (8)$$

Then, for any submodular function  $f$ ,

$$\sum_\ell p_\ell \sum_t q_\ell(t) f(A_{\ell t}) \geq \beta \cdot \sum_\ell p_\ell f(B_\ell). \quad (9)$$

*Proof.* The proof is by induction on  $n$ , the base case of  $n = 1$  is obvious. So suppose  $n \geq 2$ . For any indices  $\ell$  and  $t$ , such that  $n \in A_{\ell t}$ , by submodularity we have that  $f(A_{\ell t}) \geq f(B_\ell) - f(B_\ell - \{n\}) + f(A_{\ell t} - \{n\})$ . Thus, to show (9) it suffices to show that

$$\sum_\ell p_\ell \left( \sum_{t:n \in A_{\ell t}} q_\ell(t) (f(B_\ell) - f(B_\ell - \{n\}) + f(A_{\ell t} - \{n\})) + \sum_{t:n \notin A_{\ell t}} q_\ell(t) f(A_{\ell t}) \right) \geq \beta \cdot \sum_\ell p_\ell f(B_\ell). \quad (10)$$

By the inductive hypothesis we have that:

$$\sum_\ell \sum_t p_\ell q_\ell(t) f(A_{\ell t} - \{n\}) \geq \beta \sum_\ell p_\ell f(B_\ell - \{n\})$$

Thus to prove (10) it suffices to show that

$$\sum_\ell p_\ell \sum_{t:n \in A_{\ell t}} q_\ell(t) (f(B_\ell) - f(B_\ell - \{n\})) \geq \beta \cdot \sum_\ell p_\ell (f(B_\ell) - f(B_\ell - \{n\})). \quad (11)$$

For any set  $B_\ell$  containing  $n$ , define  $g(B_\ell) = f(B_\ell) - f(B_\ell - \{n\})$  and  $h(B_\ell) = \sum_{t:n \in A_{\ell t}} q_\ell(t)$ . Note that once we restrict to sets  $R$  that contain  $n$ , the function  $g$  is *decreasing* (due to the submodularity of  $f$ ) and non-negative. Moreover the function  $h$  is decreasing by the monotonicity condition.

Consider the probability space defined by independently choosing  $k \neq n$  with probability  $x_k$  and  $n$  with probability 1. Let  $p'_\ell = p_\ell/x_n$  denote the probability of obtain  $R = B_\ell$ . By the FKG inequality it follows that

$$\sum_\ell p'_\ell g(B_\ell) h(B_\ell) \geq \left( \sum_\ell p'_\ell g(B_\ell) \right) \left( \sum_\ell p'_\ell h(B_\ell) \right) \quad (12)$$

By the covering condition for  $k = n$ , we have that

$$\sum_\ell p_\ell \sum_{t:n \in A_{\ell t}} q_\ell(t) \geq \beta \sum_{\ell:n \in B_\ell} p_\ell$$

or equivalently

$$\sum_\ell p'_\ell h(B_\ell) \geq \beta \sum_\ell p'_\ell = \beta.$$

Combining this with (12) and using the definitions of  $g$  and  $h$  implies (11).  $\square$

*Remark:* It is easy to see that this generalizes Lemma 5.1: Let  $x_i = 1$  for each  $i \in [n]$ . The set  $B_1 = [n]$  has probability  $p_1 = 1$  and all other sets have  $p_\ell = 0$ . For this  $B_1 = [n]$ , we choose the sets  $A_{\ell t} = A_t$ , with probability distribution given by  $\{\lambda_t\}$ . The monotonicity condition is vacuous, since there is only one relevant set  $B_1$ .

**Corollary 5.4.** *Let  $S$  be a random set drawn from a product distribution on  $[n]$ . With each choice of  $S$ , associate an arbitrary subset  $S'$ . If for each  $i \in [n]$ :*

- $\Pr[i \in S' \mid i \in S] \geq \beta$ , and

- For all  $T_1 \subseteq T_2$ ,  $i \in \mathcal{S}'$  when  $\mathcal{S} = T_2$  implies that  $i \in \mathcal{S}'$  when  $\mathcal{S} = T_1$ ,

then  $E[f(\mathcal{S}')] \geq \beta E[f(\mathcal{S})]$ .

*Proof.* Immediate from Theorem 5.3; we simply associate the single set  $A_{\ell_1} = \mathcal{S}'$  with each choice  $B_\ell$  of  $\mathcal{S}$ . The two conditions on the construction of  $\mathcal{S}'$  imply the Covering and Monotonicity properties respectively.  $\square$

The algorithm in Section 2 selects item  $i$  for  $\mathcal{S}$  with probability  $x_i/4k$  and guarantees the conditions of Corollary 5.4 with  $\beta = 1/2$ . The algorithm in Section 3 uses probability  $x_i/k$  and guarantees  $\beta \geq 1/(e + o(1))$ . We combine Lemma 5.2 and Corollary 5.4 with the fact that  $x$  is an  $\frac{e}{e-1}$ -approximate solution to the continuous relaxation (6) to prove the main result of this section.

**Theorem 5.5.** *There are randomized algorithms for maximizing any monotone submodular function over  $k$ -column sparse packing constraints achieving approximation ratios  $\frac{8e}{e-1}k$  and  $\frac{e^2}{e-1}k + o(k)$ .*

**Acknowledgements:** NK thanks Chandra Chekuri and Alina Ene for detailed discussions on  $k$ -CS-PIP, including the proof of Lemma 2.3. We also thank Deeparnab Chakarabarty and David Pritchard for discussions and sharing a copy of [3]. We thank Jan Vondrak and Chandra Chekuri for pointing out an error in the original proof of Theorem 5.5, which prompted us to prove Theorem 5.3.

## References

- [1] Piotr Berman, A  $d/2$  approximation for maximum weight independent set in  $d$ -claw free graphs, *Nordic Journal of Computing*, 7(3), 178-184, 2000.
- [2] G. Calinescu, C. Chekuri, M. Pál and J. Vondrák. Maximizing a monotone submodular function under a matroid constraint, In *IPCO*, 2007.
- [3] D. Chakrabarty and D. Pritchard, Personal Communication, 2009.
- [4] C. Chekuri, A. Ene and N. Korula. Unsplittable Flow in Paths and Trees and Column-Restricted Packing Integer Programs. In *APPROX*, 2009.
- [5] C. Chekuri, A. Ene and N. Korula, Personal Communication, 2009.
- [6] C. Chekuri, M. Mydlarz and B. Shepherd, Multicommodity Demand Flow in a Tree and Packing Integer Programs, *ACM Transactions on Algorithms*, 3(3), 2007.
- [7] U. Feige, On maximizing welfare when utility functions are subadditive, In *STOC*, 2006, 41-50.
- [8] E. Hazan, S. Safra and O. Schwartz, On the complexity of approximating  $k$ -set packing, *Computational Complexity*, 15(1), 20-39, 2003.
- [9] A. J. Hurkens, 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. Discrete Math.*, 2(1), 68-72 1989.
- [10] S. Khot, Improved Inapproximability Results for MaxClique, Chromatic Number and Approximate Graph Coloring. In *FOCS*, 2001, 600–609.
- [11] A. Kulik, H. Shachnai and T. Tamir. Maximizing submodular functions subject to multiple linear constraints. In *SODA*, 2009.
- [12] R. Motwani and P. Raghavan. *Randomized Algorithms*, Cambridge University Press, 1995.
- [13] G. L. Nemhauser, L. A. Wolsey and M. L. Fisher. An analysis of approximations for maximizing submodular set functions II. *Mathematical Programming Study* 8 (1978), 73–87.
- [14] D. Pritchard, Approximability of Sparse Integer Programs, In *ESA*, 2009.

- [15] B. Shepherd and A. Vetta, The demand matching problem, *Mathematics of Operations Research*, 32, 563-578, 2007.
- [16] A. Srinivasan, Improved Approximation Guarantees for Packing and Covering Integer Programs, *SIAM J. Comput.*, 29(2), 1999, 648-670.
- [17] J. Vondrak, Optimal approximation for the submodular welfare problem in the value oracle model. In *STOC*, 2008, 67-74.