

Analysis of Sparse Cutting-planes for Sparse MILPs with Applications to Stochastic MILPs

Santanu S. Dey^{*1}, Marco Molinaro^{†2}, and Qianyi Wang^{‡1}

¹School of Industrial and Systems Engineering, Georgia Institute of Technology,
Atlanta, GA 30332, United States

²Computer Science Department, PUC-Rio, Brazil

January 5, 2016

Abstract

In this paper, we present an analysis of the strength of sparse cutting-planes for mixed integer linear programs (MILP) with sparse formulations. We examine three kinds of problems: packing problems, covering problems, and more general MILPs with the only assumption that the objective function is non-negative. Given a MILP instance of one of these three types, assume that we decide on the support of cutting-planes to be used and the strongest inequalities on these supports are added to the linear programming relaxation. Call the optimal objective function value of the linear programming relaxation together with these cuts as z^{cut} . We present bounds on the ratio of z^{cut} and the optimal objective function value of the MILP that depends only on the sparsity structure of the constraint matrix and the support of sparse cuts selected, that is, these bounds are completely data independent. These results also shed light on the strength of scenario-specific cuts for two stage stochastic MILPs.

1 Introduction

1.1 Motivation and goal

Cutting-plane technology has become one of the main pillars in the edifice that is a modern state-of-the-art mixed integer linear programming (MILP) solver. Enormous theoretical advances have been made in designing many new families of cutting-planes for general MILPs (see for example, the review papers - [19, 21]). The use of some of these cutting-planes has brought significant speedups in state-of-the-art MILP solvers [3, 17].

While significant progress has been made in developing various families of cutting-planes, lesser understanding has been obtained on the question of cutting-plane selection from a theoretical perspective. Empirically, sparsity of cutting-planes is considered an important determinant in cutting-plane selection. In a recent paper [10], we presented a geometric analysis of quality of

*santanu.dey@isye.gatech.edu

†molinaro.marco@gmail.edu

‡qwang32@gatech.edu

sparse cutting-planes as a function of the number of vertices of the integer hull, the dimension of the polytope and the level of sparsity.

In this paper, we continue to pursue the question of understanding the strength of sparse cutting-planes using completely different techniques, so that we are also able to incorporate the information that most real-life integer programming formulations have sparse constraint matrices. Moreover, the worst-case analysis we present in this paper depends on parameters that can be determined more easily than the number of vertices of the integer hull as in [10].

In the following paragraphs, we discuss the main aspects of the research direction we consider in this paper, namely: (i) The fact that solvers prefer using sparse cutting-planes, (ii) the assumption that real-life integer programs have sparse constraint matrix and (iii) why the strength of sparse cutting-planes may depend on the sparsity of the constraint matrix of the IP formulation.

What is the reason for state-of-the-art solvers to bias the selection of cutting-planes towards sparser cuts? Solving a MILP involves solving many linear programs (LP) – one at each node of the tree, and the number of nodes can easily be exponential in dimension. Because linear programming solvers can use various linear algebra routines that are able to take advantage of sparse matrices, adding dense cuts could significantly slow down the solver. In a very revealing study [22], the authors conducted the following experiment: They added a very dense valid equality constraint to other constraints in the LP relaxation at each node while solving IP instances from MIPLIB using CPLEX. This does not change the underlying polyhedron at each node, but makes the constraints dense. They observed approximately 25% increase in time to solve the instances if just 9 constraints were made artificially dense!

Is it reasonable to say that real-life integer programs have sparse constraint matrix? While this is definitely debatable (and surely “counter examples” to this statement can be provided), consider the following statistic: the average number of non-zero entries in the constraint matrix of the instances in the MIPLIB 2010 library is 1.63% and the median is 0.17% (this is excluding the non-negativity or upper bound constraints). Indeed, in our limited experience, we have never seen formulations of MILPs where the matrix is very dense, for example all the variables appearing in all the constraints. Therefore, it would be fair to say that a large number of real-life MILPs will be captured by an analysis that considers only sparse constraint matrices. We formalize later in the paper how sparsity is measured for our purposes.

Finally, why should we expect that the strength of sparse cutting-planes to be related to the sparsity of the constraint matrix of the MILP formulation? To build some intuition, consider the feasible region of the following MILP:

$$\begin{array}{rcl} A^1 x^1 & & \leq b^1 \\ & A^2 x^2 & \leq b^2 \\ x^1 \in \mathbb{Z}^{p_1} \times \mathbb{R}^{q_1}, & x^2 \in \mathbb{Z}^{p_2} \times \mathbb{R}^{q_2} & \end{array}$$

Since the constraints are completely disjoint in the x^1 and x^2 variables, the convex hull is obtained by adding valid inequalities in the support of the first $p_1 + q_1$ variables and another set of valid inequalities for the second $p_2 + q_2$ variables. Therefore, sparse cutting-planes, in the sense that their support is not on all the variables, is sufficient to obtain the convex hull. Now one would like to extend such a observation even if the constraints are not entirely decomposable, but “loosely decomposable”. Indeed this is the hypothesis that is mentioned in the classical computational paper [8]. This paper solves fairly large scale 0-1 integer programs (up to a few thousand variables) within an hour in the early 1980s, using various preprocessing

techniques and the lifted knapsack cover cutting-planes within a cut-and-branch scheme. To quote from this paper:

“All problems are characterized by sparse constraint matrix with rational data.”

“We note that the support of an inequality obtained by lifting (2.7) or (2.9) is contained in the support of the inequality (2.5) ... Therefore, the inequalities that we generate preserve the sparsity of the constraint matrix.”

Since the constraints matrices are sparse, most of the cuts that are used in this paper are sparse. Indeed, one way to view the results we obtain in this paper is to attempt a mathematical explanation for the empirical observations of quality of sparse cutting-planes obtained in [8]. Finally, we mention here in passing that the quality of Gomory mixed integer cuts were found empirically to be related to the sparsity of LP relaxation optimal tableaux in the paper [9]; however we do not explore particular families of sparse cutting-planes in this paper.

1.2 The nature of results obtained in this paper

We examine three kinds of MILPs: packing MILPs, covering MILPs, and a more general form of MILPs where the feasible region is arbitrary together with assumptions guaranteeing that the objective function value is non-negative. For each of these problems we do the following:

1. We first present a method to describe the sparsity structure of the constraint matrix.
2. Then we present a method to describe a hierarchy of cutting-planes from very sparse to completely dense. The method for describing the sparsity of the constraint matrix and that for the cuts added are closely related.
3. For a given MILP instance, we assume that once the sparsity structure of the cutting-planes (i.e. the support of the cutting-planes are decided), the strongest (or equivalently all) valid inequalities on these supports are added to the linear programming relaxation and the resulting LP is solved. Call the optimal objective function value of this LP as z^{cut} .
4. All our results are of the following kind: We present bounds on the ratio of z^{cut} and the optimal objective function value of the IP (call this z^I), where the bound depends only on the sparsity structure of the constraint matrix and the support of sparse cuts.

For example, in the packing case, since objective function is of the maximization type, we present an upper bound on $\frac{z^{cut}}{z^I}$ which, we emphasize again, depends entirely on the location of zeros in the constraint matrix and the cuts added and is independent of the actual data of the instance. We note here that the method to describe the sparsity of the matrix and cutting-planes are different for the different types of problems.

We are also able to present examples in the case of all the three types of problems, that show that the bounds we obtain are tight.

Though out this paper we will constantly refer back to the deterministic equivalent of a two-stage stochastic problem with finitely many realizations of uncertain parameters in the second stage. Such MILPs have naturally sparse formulations. Moreover, sparse cutting-planes, the so-called scenario-specific cuts (or the path inequalities), for such MILPs have been well studied.

(See details in Section 2). Therefore, any result we obtain for quality of sparse cutting-planes for sparse IPs is applicable in this setting, and this connection allows us to shed some light on the performance of scenario-specific cuts for stochastic MILPs.

We also conduct computational experiments for all these classes of MILPs to study the effectiveness of sparse cutting-planes. Our main observation is the sparse cuts usually perform much better than the worst-case bounds we obtain theoretically.

Outline the paper: We present all the definitions (of how sparsity is measured, etc.) and the main theoretical results in Section 2. Then in Section 3 we present results from an empirical study of the same questions. We make concluding remarks in Section 4. Section 5 provides proofs of all the results presented in Section 2.

2 Main results

2.1 Notation and basic definitions

Given a feasible region of a mixed integer linear program, say P , we denote the convex hull of P by P^I and denote the feasible region of the linear programming relaxation by P^{LP} .

For any natural number n , we denote the set $\{1, \dots, n\}$ by $[n]$. Given a set V , 2^V is used to represent its power set.

Definition 1 (Sparse cut on N). *Given the feasible region of a mixed integer linear program (P) with n variables, and a subset of indices $N \subseteq [n]$, we call $\alpha^T x \leq \beta$ a sparse cut on N if it is a valid inequality for P^I and the support of α is restricted to variables with index in N , that is $\{i \in [n] \mid \alpha_i \neq 0\} \subseteq N$.*

Clarification of the above definition: If $\alpha^T x \leq \beta$ is a sparse cut on N , then $\alpha_i = 0$ for all $i \in [n] \setminus N$, while α_i may also be equal to 0 for some $i \in N$.

Since we are interested in knowing how good of an approximation of P^I is obtained by the addition of all sparse cutting-planes to the linear programming relaxation, we will study the set defined next.

Definition 2 (Sparse closure on N). *Given a feasible region of a mixed integer linear program (P) with n variables and $N \subseteq [n]$, we define the sparse closure on N , denoted as $P^{(N)}$, and defined as*

$$P^{(N)} := P^{LP} \cap \bigcap_{\{(\alpha, \beta) \mid \alpha x \leq \beta \text{ is a sparse cut on } N\}} \{x \mid \alpha x \leq \beta\}.$$

2.2 Packing problems

In this section, we present our results on the quality of sparse cutting-planes for packing-type problems, that is problems of the following form:

$$\begin{aligned} \text{(P)} \quad & \max && c^T x \\ & \text{s.t.} && Ax \leq b \\ & && x_j \in \mathbb{Z}_+, \forall j \in \mathcal{L} \\ & && x_j \in \mathbb{R}_+, \forall j \in [n] \setminus \mathcal{L}, \end{aligned}$$

with $A \in \mathbb{Q}_+^{m \times n}$, $b \in \mathbb{Q}_+^m$, $c \in \mathbb{Q}_+^n$ and $\mathcal{L} \subseteq [n]$.

In order to analyze the quality of sparse cutting-planes for packing problems we will *partition the variables into blocks*. One way to think about this partition is that it allows us to understand the global effect of interactions between blocks of “similar variables”. For example, in MIPLIB instances, one can possibly rearrange the rows and columns [4, 2, 23, 1] so that one sees patterns of blocks of variables in the constraint matrices. See Figure 1(a) for an illustration of “observing patterns” in a sparse matrix. Moreover note that in what follows one can always define the blocks to be singletons, that is each block is just a single variable.

The next example illustrates an important class of problems where such partitioning of variables is natural.

Example 3 (Two-stage stochastic problem). *The deterministic equivalent of a two-stage stochastic problem with finitely many realizations of uncertain parameters in the second stage has the following form:*

$$\begin{aligned} \max \quad & c^T y + \sum_{i=1}^k (d^i)^T z^i \\ \text{s.t.} \quad & Ay \leq b \\ & A^i y + B^i z^i \leq b^i \quad \forall i \in [k], \end{aligned}$$

where y are the first stage variables and the z^i variables corresponding to each realization in the second stage. Notice there are two types of constraints:

1. Constraints involving only the first stage variables.
2. Constraints involving the first stage variables and second stage variables corresponding to one particular realization of uncertain parameters.

Note that there are no constraints in the formulation that involve variables corresponding to two different realizations of uncertain parameters.

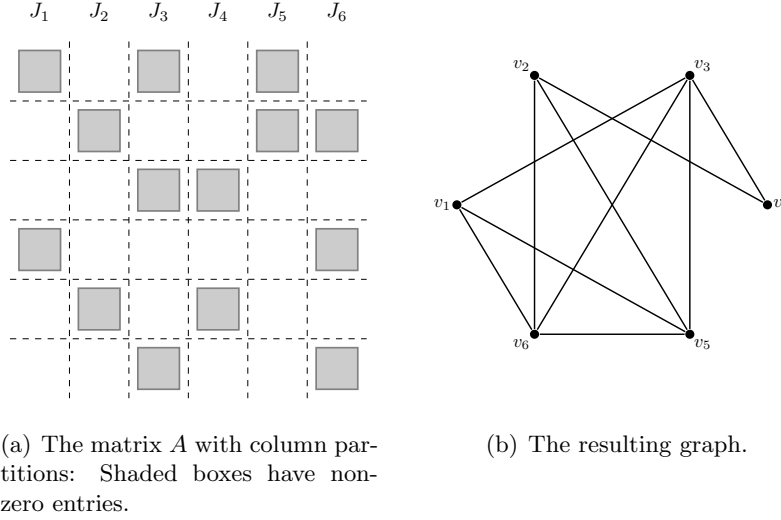
It is natural to put all the first stage variables y into one block and each of the second stage variables z^i corresponding to one realization of uncertain parameters into a separate block of variables.

To formalize the effect of the interactions between blocks of variables we define a graph that we call as the packing interaction graph. This graph will play an instrumental role in analyzing the strength of sparse cutting-planes.

Definition 4 (Packing interaction graph of A). *Consider a matrix $A \in \mathbb{Q}^{m \times n}$. Let $\mathcal{J} := \{J_1, J_2, \dots, J_q\}$ be a partition of the index set of columns of A (that is $[n]$). We define the packing interaction graph $G_{A, \mathcal{J}}^{\text{pack}} = (V, E)$ as follows:*

1. There is one node $v_j \in V$ for every part $J_j \in \mathcal{J}$.
2. For all $v_i, v_j \in V$, there is an edge $(v_i, v_j) \in E$ if and only if there is a row in A with non-zero entries in both parts J_i and J_j , namely there are $k \in [m]$, $u \in J_i$ and $w \in J_j$ such that $A_{ku} \neq 0$ and $A_{kw} \neq 0$.

Figure 1: Constructing $G_{A,\mathcal{J}}^{\text{pack}}$.



Notice that $G_{A,\mathcal{J}}^{\text{pack}}$ captures the sparsity pattern of the matrix A up to partition \mathcal{J} of columns of the matrix, i.e., this graph ignores the sparsity (or the lack of it) within each of the blocks of columns, but captures the sparsity (or the lack of it) between the blocks of the column. Finally note that if each of the blocks in \mathcal{J} were singletons, then the resulting graph is the intersection graph [12].

Figure 1 illustrate the process of constructing $G_{A,\mathcal{J}}^{\text{pack}}$. Figure 1(a) shows a matrix A , where the columns are partitioned into six variable blocks, the unshaded boxes correspond to zeros in A and the shaded boxes correspond to entries in A that are non-zero. Figure 1(b) shows $G_{A,\mathcal{J}}^{\text{pack}}$.

Example 5 (Two-stage stochastic problem: $G_{A,\mathcal{J}}^{\text{pack}}$). *Given a two-stage stochastic problem with k second stage realizations, we partition the variables in $k+1$ blocks (as discussed in Example 3). So we have a graph $G_{A,\mathcal{J}}^{\text{pack}}$ with vertex set $\{v_1, v_2, \dots, v_{k+1}\}$ and edges $(v_1, v_2), (v_1, v_3), \dots, (v_1, v_{k+1})$.*

The sparse cuts we examine will be with respect to the blocks of variables. In other words, while the sparse cuts may be dense with respect to the variables in some blocks, it can be sparse globally if its support is on very few blocks of variables. To capture this, we use a *support list* to encode which combinations of blocks cuts are allowed to be supported on; we state this in terms of subsets of nodes of the graph $G_{A,\mathcal{J}}^{\text{pack}}$.

Definition 6 (Column block-sparse closure). *Given the problem (P), let $\mathcal{J} := \{J_1, J_2, \dots, J_q\}$ be a partition of the index set of columns of A (that is $[n]$) and consider the packing interaction graph $G_{A,\mathcal{J}}^{\text{pack}} = (V, E)$.*

1. *With slight overload in notation, for a set of nodes $S \subseteq V$ we say that inequality $\alpha x \leq \beta$ is a sparse cut on S if it is a sparse cut on its corresponding variables, namely $\bigcup_{v_j \in S} J_j$. The closure of these cuts is denoted by $P^{(S)} := P^{\bigcup_{v_j \in S} J_j}$.*
2. *Given a collection \mathcal{V} of subsets of the vertices V (the support list), we use $P^{\mathcal{V},P}$ to denote*

the closure obtained by adding all sparse cuts on the sets in \mathcal{V} 's, namely

$$P^{\mathcal{V},P} := \bigcap_{S \in \mathcal{V}} P^{(S)}.$$

This definition of column block-sparse closure allows us to define various levels of sparsity of cutting-planes that can be analyzed. In particular if \mathcal{V} includes V , then we are considering completely dense cuts and indeed in that case $P^{\mathcal{V},P} = P^I$.

Let $z^I = \max\{c^T x \mid x \in P^I\}$ be the IP optimal value and $z^{\mathcal{V},P} = \max\{c^T x \mid x \in P^{\mathcal{V},P}\}$ be the optimal value obtained by employing sparse cuts on the support list \mathcal{V} . Since we are working with a maximization problem $z^{\mathcal{V},P} \geq z^I$ and our goal is to provide bounds on how much bigger $z^{\mathcal{V},P}$ can be compared to z^I . Moving forward, we will be particularly interested in two types of sparse cut support lists \mathcal{V} :

1. *Super sparse closure* ($P^{S.S.} := P^{\mathcal{V},P}$ and $z^{S.S.} := z^{\mathcal{V},P}$): We will consider the sparse cut support list $\mathcal{V} = \{\{v_1\}, \{v_2\}, \{v_3\}, \dots, \{v_{|\mathcal{V}|}\}\}$. We call this the super sparse closure, since once the partition \mathcal{J} is decided, these are the sparsest cuts to be considered.
2. *Natural sparse closure*: Let A_1, \dots, A_m be the rows of A . Let V^i be the set of nodes corresponding to block variables that have non-zero entries in A_i (that is $V^i = \{v_u \in V \mid A_{ik} \neq 0 \text{ for some } k \in J_u\}$). For the resulting sparse cut support list $\mathcal{V} = \{V^1, V^2, \dots, V^m\}$, we call the column block-sparse closure as the ‘natural’ sparse closure (and $P^{N.S.} := P^{\mathcal{V},P}$ and $z^{N.S.} := z^{\mathcal{V},P}$). The reason to consider this case is that once the partition \mathcal{J} is decided, the cuts defining $P^{N.S.}$ most closely resembles the sparsity pattern of the original constraint matrix. To see this, consider the case when $\mathcal{J} = \{\{1\}, \{2\}, \dots, \{n\}\}$, that is every block is a single variable. In this case, the sparse cut support list \mathcal{V} represents exactly the different sparsity pattern of the various rows of the IP formulation. Indeed the cuts added in [8] satisfied this sparsity pattern.

Example 7 (Two-stage stochastic problem: Specific-scenario cuts, Natural sparse closure is same as relaxing “nonanticipativity” constraints in some cases). *Consider again the two-stage stochastic problem with k second stage realizations as discussed in Example 5. Consider the cuts on the support of first stages variables together with the variables corresponding to one second stage realization, the so-called specific-scenario cuts. Such cutting-planes are well-studied, see for example [13, 24]. Notice that based on the partition \mathcal{J} previously discussed, the closure of all the specific-scenario cuts is precisely equivalent to the natural sparse closure $P^{N.S.}$.*

A standard technique in stochastic integer programming is to make multiple copies of the first stage variables, which are connected through equality constraints, and relax these (“nonanticipativity”) equality constraints (via Lagrangian relaxation methods) to produce computationally strong bound [6]. It is straightforward to see that in the case where there is complete recourse, the closure of the specific-scenario cuts or equivalently the natural sparse closure, will give the same bound as this nonanticipativity dual.

To the best of our knowledge there are no known global bounds known on the quality of nonanticipativity dual. The results in this paper will be able to provide some such bounds.

In order to present our results, we require the following generalizations of standard graph-theoretic notions such a stable sets and chromatic number.

Definition 8 (Mixed stable set subordinate to \mathcal{V}). Let $G = (V, E)$ be a simple graph. Let \mathcal{V} be a collection of subsets of the vertices V . We call a collection of subsets of vertices $\mathcal{M} \subseteq 2^V$ a mixed stable set subordinate to \mathcal{V} if the following hold:

1. Every set in \mathcal{M} is contained in a set in \mathcal{V}
2. The sets in \mathcal{M} are pairwise disjoint
3. There are no edges of G with endpoints in distinct sets in \mathcal{M} .

Definition 9 (Mixed chromatic number with respect to \mathcal{V}). Consider a simple graph $G = (V, E)$ and a collection \mathcal{V} of subset of vertices.

- The mixed chromatic number $\bar{\eta}^{\mathcal{V}}(G)$ of G with respect to \mathcal{V} is the smallest number of mixed stables sets $\mathcal{M}^1, \dots, \mathcal{M}^k$ subordinate to \mathcal{V} that cover all vertices of the graph (that is, every vertex $v \in V$ belongs to a set in one of the \mathcal{M}^i 's).
- (Fractional mixed chromatic number.) Given a mixed stable set \mathcal{M} subordinate to \mathcal{V} , let $\chi_{\mathcal{M}} \in \{0, 1\}^{|V|}$ denote its incidence vector (that is, for each vertex $v \in V$, $\chi_{\mathcal{M}}(v) = 1$ if v belongs to a set in \mathcal{M} , and $\chi_{\mathcal{M}}(v) = 0$ otherwise.) Then we define the fractional mixed chromatic number

$$\begin{aligned} \eta^{\mathcal{V}}(G) = \min \sum_{\mathcal{M}} y_{\mathcal{M}} \\ \text{s.t. } \sum_{\mathcal{M}} y_{\mathcal{M}} \chi_{\mathcal{M}} \geq \mathbb{1} \\ y_{\mathcal{M}} \geq 0 \quad \forall \mathcal{M}, \end{aligned} \tag{1}$$

where the summations range over all mixed stable sets subordinate to \mathcal{V} and $\mathbb{1}$ is the vector in $\mathbb{R}^{|V|}$ of all ones.

Note that when \mathcal{V} corresponds to the super sparse closure $P^{S.S.}$, that is the elements of \mathcal{V} is the collection of singletons, the mixed stable sets subordinate to \mathcal{V} are the usual stable sets in the graph and the (resp. fractional) mixed chromatic number are the usual (resp. fractional) chromatic number.

The following simple example helps to clarify and motivate the definition of mixed stable sets: they identify sets of variables that can be set **independently** and still yield feasible solutions.

Example 10. Consider the simple packing two-stage stochastic problem:

$$\begin{aligned} \max \quad & c_1 x_1 + c_2 x_2 + c_3 x_3 \\ \text{s.t.} \quad & a_{11} x_1 + a_{12} x_2 \leq b_1 \\ & a_{21} x_1 + a_{23} x_3 \leq b_2 \\ & x \in \mathbb{Z}_+^3. \end{aligned}$$

Consider the partition $\mathcal{J} = \{\{1\}, \{2\}, \{3\}\}$ so that the graph $G_{A, \mathcal{J}}^{\text{pack}}$ equals the path $v_2 - v_1 - v_3$. Consider the support list $\mathcal{V} = \{\{1, 2\}, \{1, 3\}\}$ for the “natural sparse closure” setting. Then

the maximal mixed stable sets of subordinate to \mathcal{V} are $\mathcal{M}_1 = \{\{1, 2\}\}$, $\mathcal{M}_2 = \{\{1, 3\}\}$ and $\mathcal{M}_3 = \{\{2\}, \{3\}\}$; $\mathcal{M}_4 = \{\{1\}\}$ is a non-maximal mixed stable set.

To see that mixed stable sets identify sets of variables that can be set independently and still yield a feasible solution, for $i = 1, 2, 3$ let $x^{(i)}$ be the optimal solution to the above packing problem conditioned on $x_j = 0$ for all $j \neq i$; for example $x^{(2)} = (0, \lfloor b_1/a_{12} \rfloor, 0)$ and $x^{(3)} = (0, 0, \lfloor b_2/a_{23} \rfloor)$. Taking the mixed stable set $\mathcal{M}_3 = \{\{2\}, \{3\}\}$ we see that the combination of $x^{(2)} + x^{(3)} = (0, \lfloor b_1/a_{12} \rfloor, \lfloor b_2/a_{23} \rfloor)$ is also feasible for the problem.

Moreover, these solutions allow us to upper bound the ratio $z^{\mathcal{V},P}/z^I$, namely the quality of the column block-sparse closure. First, the integer optimum z^I is at least $\max\{c^T(x^{(2)} + x^{(3)}), c^T x^{(1)}\}$. Also, one can show that $z^{\mathcal{V},P} \leq c^T(x^{(2)} + x^{(3)}) + c^T x^{(1)}$ (this uses the fact that actually $x^{(2)} + x^{(3)}$ is the optimal solution for the problem conditioned on $x_1 = 0$, and $x^{(1)}$ the optimal solution conditioned on $x_2 = x_3 = 0$). Together this gives $z^{\mathcal{V},P}/z^I \leq 2$. Notice that the upper bound on $z^{\mathcal{V}}$ is obtained by adding up the solutions corresponding to the sets \mathcal{M}_3 and \mathcal{M}_4 , which together cover all the variables of the problem. Looking at the fractional chromatic number $\eta^{\mathcal{V}}(G_{A,\mathcal{J}}^{\text{pack}})$ allow us to provide essentially the best such bound.

Our first result gives a worst-case upper bound on $\frac{z^{\mathcal{V},P}}{z^I}$ that is, surprisingly, independent of the data A, b, c , and depends only on the packing interaction graph $G_{A,\mathcal{J}}^{\text{pack}}$ and the choice of sparse cut support list \mathcal{V} .

Theorem 11. *Consider a packing integer program as defined in (P). Let $\mathcal{J} \subseteq 2^{[n]}$ be a partition of the index set of columns of A and let $G = G_{A,\mathcal{J}}^{\text{pack}} = (V, E)$ be the packing interaction graph of A . Then for any sparse cut support list $\mathcal{V} \subseteq 2^V$ we have*

$$z^{\mathcal{V},P} \leq \eta^{\mathcal{V}}(G) \cdot z^I.$$

As discussed before, if we are considering the super sparse closure $P^{\text{S.S.}}$, $\eta^{\mathcal{V}}(G)$ is the usual fractional chromatic number. Therefore we obtain the following possibly weaker bound using Brook's theorem [5].

Corollary 12. *Consider a packing integer program as defined in (P). Let $\mathcal{J} \subseteq 2^{[n]}$ be a partition of the index set of columns of A and let $G_{A,\mathcal{J}}^{\text{pack}}$ be the packing interaction graph of A . Let Δ denote the maximum degree of G . Then we have the following bounds on the optimum value of the super sparse closure $P^{\text{S.S.}}$:*

1. *If $G_{A,\mathcal{J}}^{\text{pack}}$ is not a complete graph or an odd cycle, then*

$$z^{\text{S.S.}} \leq \Delta \cdot z^I.$$

2. *If $G_{A,\mathcal{J}}^{\text{pack}}$ is a complete graph or an odd cycle, then*

$$z^{\text{S.S.}} \leq (\Delta + 1) \cdot z^I.$$

Thus assuming the original IP is sparse and the maximum degree of $G_{A,\mathcal{J}}^{\text{pack}}$ is not very high, the above result says that we get significantly tight bounds using only super sparse cuts. In fact it is easy to show the above Corollary's bounds can be tight when $G_{A,\mathcal{J}}^{\text{pack}}$ is a 3-cycle or a star. We record this result here.

Theorem 13. For any $\epsilon > 0$:

1. There exists a packing integer program as defined in (P) and a partition $\mathcal{J} \subseteq 2^{[n]}$ of the index set of columns of A such that the graph $G_{A,\mathcal{J}}^{\text{pack}}$ is a 3-cycle and

$$z^{\text{S.S.}} \geq (3 - \epsilon)z^I.$$

2. (Strength of super sparse cuts for packing two-stage problems) There exists a packing integer program as defined in (P) and a partition $\mathcal{J} \subseteq 2^{[n]}$ of the index set of columns of A such that the graph $G_{A,\mathcal{J}}^{\text{pack}}$ is a star and

$$z^{\text{S.S.}} \geq (2 - \epsilon)z^I.$$

We mention here in passing that there are many well-known upper bounds on the fractional chromatic number with respect to other graph properties, which also highlight that for sparse graph we expect the fractional chromatic number to be small. For example, let G be a connected graph of max degree Δ and clique number $\omega(G)$. Then

1. $\eta(G) \leq \frac{\omega(G) + \Delta + 1}{2}$. ([20])
2. $\eta(G) \geq \Delta$ if and only if G is a complete graph, odd cycle, a graph with $\omega(G) = \Delta$, a square of the 8-cycle, or the strong product of 5-cycle and K_2 . Moreover if $\Delta \geq 4$ and G is not any of the graphs listed above, then $\eta(G) \leq \Delta - \frac{2}{67}$. ([15])

One question is whether we can get better bounds using the potentially denser natural sparse cuts. Equivalently, is the fractional chromatic number $\eta^{\mathcal{V}}(G_{A,\mathcal{J}}^{\text{pack}})$ much smaller when we consider the sparse cut support list \mathcal{V} corresponding to the natural sparse closure? We prove results for some special, but important, structures.

Theorem 14 (Natural sparse closure of trees). Consider a packing integer program as defined in (P). Let $\mathcal{J} \subseteq 2^{[n]}$ be a partition of the index set of columns of A and let $G_{A,\mathcal{J}}^{\text{pack}}$ be the packing interaction graph of A . Suppose $G_{A,\mathcal{J}}^{\text{pack}}$ a tree and let Δ be its maximum degree. Then:

$$z^{\text{N.S.}} \leq \left(\frac{2\Delta - 1}{\Delta} \right) z^I.$$

Compare this result for natural sparse cuts with the result for super sparse cuts on trees. While with super sparse cuts we are able to get a multiplicative bound of 2 (this is the fractional chromatic number for bipartite graphs), using natural sparse cuts the bound is always strictly less than 2.

Interestingly, this upper bound is tight even when the induced graph $G_{A,\mathcal{J}}^{\text{pack}}$ is a star, which corresponds exactly to the case of stochastic packing programs. The construction of the tight instances are based on special set systems called *affine designs*, where we exploit their particular partition and intersection properties.

Theorem 15 (Tightness of natural sparse closure of trees). For any $\epsilon > 0$, there exists a packing integer program (P) and a suitable partition \mathcal{J} of variables where $G_{A,\mathcal{J}}^{\text{pack}}$ is a star with max degree Δ such that

$$z^{\text{N.S.}} \geq \left(\frac{2\Delta - 1}{\Delta} - \epsilon \right) z^I.$$

As discussed in Example 5, for the case of two-stage stochastic problem with the right choice of \mathcal{J} the packing interaction graph is a star. So we obtain the following corollary of Theorem 14.

Corollary 16 (Strength of specific-scenario cuts for packing two-stage stochastic problems). *Consider a packing-type two-stage stochastic problem with k realization. Then*

$$z^{N.S.} \leq \left(\frac{2k-1}{k} \right) z^I,$$

where $z^{N.S.}$ is the objective function obtained after adding all specific-scenario cuts. Moreover this bound is tight.

We note that the analysis of approximation algorithm for two stage matching problem in the papers [11, 16] is related to the above result. We plan on exploring this relation in a future paper.

Finally we consider the case of natural sparse cutting-planes when $G_{A,\mathcal{J}}^{\text{pack}}$ is a cycle. Interestingly, the fractional mixed chromatic number $\eta^{\mathcal{V}}(G_{A,\mathcal{J}}^{\text{pack}})$ depends on the length of the cycle modulo 3.

Theorem 17 (Natural sparse closure of cycles). *Consider a packing integer program as defined in (P). Let $\mathcal{J} \subseteq 2^{[n]}$ be a partition of the index set of columns of A and let $G_{A,\mathcal{J}}^{\text{pack}}$ be the packing interaction graph of A . If $G_{A,\mathcal{J}}^{\text{pack}}$ is a cycle of length K , then:*

1. If $K = 3k, k \in \mathbb{Z}_{++}$, then $z^{N.S.} \leq \frac{3}{2} z^I$.
2. If $K = 3k + 1, k \in \mathbb{Z}_{++}$, then $z^{N.S.} \leq \frac{3k+1}{2k} z^I$.
3. If $K = 3k + 2, k \in \mathbb{Z}_{++}$, then $z^{N.S.} \leq \frac{3k+2}{2k+1} z^I$.

Moreover, for any $\epsilon > 0$, there exists a packing integer program with a suitable partition \mathcal{V} of variables, where $G_{A,\mathcal{J}}^{\text{pack}}$ is a cycle of length K such that

1. If $K = 3k, k \in \mathbb{Z}_{++}$, then $z^{N.S.} \geq \left(\frac{3}{2} - \epsilon \right) z^I$.
2. If $K = 3k + 1, k \in \mathbb{Z}_{++}$, then $z^{N.S.} \geq \left(\frac{3k+1}{2k} - \epsilon \right) z^I$.
3. If $K = 3k + 2, k \in \mathbb{Z}_{++}$, then $z^{N.S.} \geq \left(\frac{3k+2}{2k+1} - \epsilon \right) z^I$.

All proofs of the above results are presented in Section 5.1.

2.3 Covering problems

In this section, we present our results on the quality of sparse cutting-planes for covering-type problems, that is problems of the following form:

$$\begin{aligned} \text{(C)} \quad & \min && c^T x \\ & \text{s.t.} && Ax \geq b \\ & && x_j \in \mathbb{Z}_+, \forall j \in \mathcal{L} \\ & && x_j \in \mathbb{R}_+, \forall j \in [n] \setminus \mathcal{L} \end{aligned}$$

with $A \in \mathbb{Q}_+^{m \times n}$, $b \in \mathbb{Q}_+^m$, $c \in \mathbb{Q}_+^n$ and $\mathcal{L} \subseteq [n]$. In this case, we would like to prove lower bounds on the objective functions after adding the sparse cutting-planes.

Our first observation is a negative result: super sparse cuts as defined for the packing-type problems can be arbitrarily bad for the case of covering problems. In order to present this result, let formalize the notion of super sparse cuts in this setting. In particular, given an instance of type (C), we assume we partition the variable indices n into $\mathcal{J} = \{J_1, J_2, \dots, J_q\}$. For all $j \in [q]$ we add all possible cuts that have support on variables with index in J_j . Let $z^{S.S.}$ be the optimal objective function of the resulting LP with cuts.

Theorem 18. *For any constant $M > 0$, there exists a covering integer program (C) and partition $\mathcal{J} := \{J_1, J_2\}$ of $[n]$, such that the corresponding $z^{S.S.}$ and z^I satisfies:*

$$z^I > M \cdot z^{S.S.}.$$

Note that super sparse cuts may have support that are strict subsets of the support on the constraints of the formulation. Theorem 18 suggests that such cutting-planes in the worst case will not produce good bounds for covering problems.

It turns out that in order to analyze sparse cutting-planes for covering problems, the interesting case is when their support is at least the support of the constraints of the original formulation. Moreover, we need to work with a graph that is a “dual” of $G_{A,\mathcal{J}}^{\text{pack}}$, namely it acts on the *rows* of the problem instead of columns. For the matrix A , let A_i be the i^{th} row. We let $\text{supp}(A_i) \subseteq [n]$ be the set of variables which appear in the i^{th} constraint, that is $\text{supp}(A_i) := \{j \in [n] \mid A_{ij} \neq 0\}$.

Definition 19 (Covering interaction graph of A). *Consider the matrix $A \in \mathbb{Q}^{m \times n}$. Let $\mathcal{I} = \{I_1, I_2, \dots, I_p\}$ be a partition of index set of **rows** of A (that is $[m]$). We define the covering interaction graph $G_{A,\mathcal{I}}^{\text{cover}} = (V, E)$ as follows:*

1. *There is a node $v_i \in V$ for every part $I_i \in \mathcal{I}$.*
2. *For all $v_i, v_j \in V$, there is an edge $(v_i, v_j) \in E$ if and only if there is a **column** of A with non-zero entries in both parts I_i and I_j , namely $\bigcup_{r \in I_i} \text{supp}(A_r)$ intersects $\bigcup_{r \in I_j} \text{supp}(A_r)$.*

Definition 20 (Row block-sparse closure). *Given the problem (C), let $\mathcal{I} = \{I_1, I_2, \dots, I_p\}$ be a partition of index set of rows of A (that is $[m]$) and consider the covering interaction graph $G_{A,\mathcal{I}}^{\text{cover}} = (V, E)$.*

1. *With slight overload in notation, for a set of nodes $S \subseteq V$ we say that inequality $\alpha x \leq \beta$ is a sparse cut on S if it is a sparse cut on the union of the support of the rows in S , namely $\alpha x \leq \beta$ is a sparse cut on $\bigcup_{v_i \in S} \bigcup_{r \in I_i} \text{supp}(A_r)$. The closure of these cuts is denoted by $P(S) := P(\bigcup_{v_i \in S} \bigcup_{r \in I_i} \text{supp}(A_r))$.*
2. *Given a collection \mathcal{V} of subsets of the vertices V (the row support list), we use $P^{\mathcal{V},C}$ to denote the closure obtained by adding all sparse cuts on the sets in \mathcal{V} 's, namely*

$$P^{\mathcal{V},C} := \bigcap_{S \in \mathcal{V}} P(S).$$

Moreover, we define the optimum value over the row block-sparse closure

$$z^{\mathcal{V},C} := \min \{c^T x \mid x \in P^{\mathcal{V},C}\}.$$

Example 21 (Two-stage stochastic problem: $G_{A,\mathcal{I}}^{\text{cover}}$, weak specific-scenario cuts). *Given a two-stage covering stochastic problem with k second stage realizations, we partition the rows into k blocks (each block consists of constraints between first stage variables only or first stage variables and variables corresponding to one particular realization). So we have a graph $G_{A,\mathcal{I}}^{\text{cover}}$ with $V = \{v_1, v_2, \dots, v_k\}$ which is a clique. Moreover, if we consider the closure corresponding to the row support list $\mathcal{V} = \{\{v_1\}, \{v_2\}, \{v_3\}, \dots, \{v_k\}\}$, the cuts are quite similar to specific-scenario cuts (although potentially weaker, since the supports of inequalities could possibly be strictly smaller than those allowed in the “specific-scenario cuts” in Section 2.2). Therefore we call this closure, the weak specific-scenario closure.*

We now present the main result of this section. In particular, we present a worst-case upper bound on $\frac{z^I}{z^{\mathcal{V},C}}$ that is independent of the data A , b , c , and depends only on $G_{A,\mathcal{I}}^{\text{cover}}$ and the choice of the row support list \mathcal{V} . We remind the reader that given a graph G and collection \mathcal{V} of its vertices, $\bar{\eta}^{\mathcal{V}}(G)$ is the mixed chromatic number with respect to \mathcal{V} (see Definition 8).

Theorem 22. *Consider a covering integer programming as defined in (C). Let $\mathcal{I} \subseteq 2^{[m]}$ be a partition of the index set of rows of A and let $G = G_{A,\mathcal{I}}^{\text{cover}} = (V, E)$ be the covering interaction graph of A . Then for any sparse cut support list $\mathcal{V} \subseteq 2^V$ we have*

$$z^{\mathcal{V},C} \geq \frac{1}{\bar{\eta}^{\mathcal{V}}(G)} \cdot z^I.$$

We make a few comments regarding Theorem 18:

1. While the result of Theorem 22 for covering-type IPs is very “similar” to the result of Theorem 11 for packing-type of IPs, the key ideas in the proofs are different.
2. Like the previous discussion in Section 2.2, the chromatic numbers is small for graphs with small max degree. In fact, using Brook’s Theorem [5], we can obtain a result very similar to Corollary 12 for the covering case as well.
3. The result of Theorem 18 holds even if upper bounds are present on some or all of the variables (in this case, we also need to assume that the instance is feasible).

Consider the case of two-stage covering stochastic problem with K scenario and \mathcal{I} as defined in Example 21. Since $G_{A,\mathcal{I}}^{\text{cover}}$ is a clique, its chromatic number is K . Therefore we obtain the following corollary of Theorem 22.

Corollary 23 (Strength of weak specific-scenario cuts for covering stochastic problems). *Consider a two-stage covering stochastic problem for K scenario. Let z^* be the optimal objective value obtained after adding all weak specific-scenario cuts. Then*

$$z^* \geq \frac{1}{K} z^I.$$

Next we prove that the bound presented in Corollary 23 is tight (and therefore the result of Theorem 18 is tight for $G_{A,\mathcal{I}}^{\text{cover}}$ being a clique).

Theorem 24. *Let z^* be the optimal objective value obtained after adding all weak specific-scenario cuts for a two-stage covering stochastic problem. Given any $\epsilon > 0$ with $\epsilon < K$, there exists an instance of the covering-type two-stage stochastic problem with K scenarios such that*

$$z^* \leq \frac{1}{(K - \epsilon)} \cdot z^I.$$

The proof of Theorem 24 is perhaps the most involved in this paper, as the family of instances constructed to prove the above theorem are significantly complicated.

All proofs of the above results are presented in Section 5.2.

2.4 “Packing-type” problem with arbitrary A matrix

Up until now we have considered packing and covering problems. We now present results under much milder assumptions. In particular, we consider problem (P) with arbitrary matrix $A \in \mathbb{Q}^{m \times n}$ instead of a non-negative matrix (and b is also not assumed to be non-negative). The assumptions we therefore make in this section are: c is a non-negative vector, the variables are non-negative and the objective is of the maximization-type as in (P).

We use the same definition of sparse-cutting planes as for the packing instances considered in Section 2.2. All other notation used is also the same as in Section 2.2.

As it turns out, even in this significantly more general case, it is possible to obtain tight data-independent bounds on the quality of sparse-cutting-planes. In order to present this result we introduce the notion of *corrected* average constraint density. The reason to introduce this notion is the following: the strength of cuts in this case is determined by the average density, as long as the cuts cover all the variables. Based on this, the corrected average density captures the best bound one can obtain using a given support list.

Definition 25 (Corrected average density). *Let $\mathcal{V} = \{V^1, V^2, \dots, V^t\}$ be the sparse cut support list. For any subset $\tilde{\mathcal{V}} = \{V^{u_1}, V^{u_2}, \dots, V^{u_k}\} \subseteq \mathcal{V}$ define its density as*

$$D(\tilde{\mathcal{V}}) = \frac{1}{k} \sum_{i=1}^k |V^{u_i}|.$$

We define the corrected average density of \mathcal{V} (denoted as $D_{\mathcal{V}}$) as maximum value of $D(\tilde{\mathcal{V}})$ over all $\tilde{\mathcal{V}}$'s that cover V , that is, $\bigcup_{V' \in \tilde{\mathcal{V}}} V' = V$.

Note that $D_{\mathcal{V}} \geq 1$ for any choice of sparse cut support list \mathcal{V} , and for the trivial list $\mathcal{V} = \{V(G_{A,\mathcal{J}}^{\text{pack}})\}$ that allows fully dense cuts we have $D_{\mathcal{V}} = |V(G_{A,\mathcal{J}}^{\text{pack}})|$. The following is the main result of this section.

Theorem 26. *Let (P) be defined by an arbitrary $A \in \mathbb{Q}^{m \times n}$, $b \in \mathbb{Q}^m$, $c \in \mathbb{Q}_+^n$. Let \mathcal{J} be a partition of the index set of columns of A (that is $[n]$). Let $G_{A,\mathcal{J}}^{\text{pack}} = (V, E)$ be the packing interaction graph of A and let \mathcal{V} be the sparse cut support list. If the instance is feasible, then:*

$$z^{\mathcal{V},P} \leq (|V| + 1 - D_{\mathcal{V}}) \cdot z^I.$$

Let us see some consequences of Theorem 26. Since $D_{\mathcal{V}} \geq 1$ we obtain the following result.

Corollary 27. *Given (P), with arbitrary $A \in \mathbb{Q}^{m \times n}$, $b \in \mathbb{Q}^m$, $c \in \mathbb{Q}_+^n$. Let \mathcal{J} be a partition of the index set of columns of A (that is $[n]$). If the instance is feasible, then:*

$$z^{\mathcal{V},P} \leq |V| \cdot z^I.$$

It turns out that the bound in Corollary 27 is tight when $G_{A,\mathcal{J}}^P(V, E)$ is a star.

Theorem 28 (Strength of super sparse cuts for two-stage packing-type problem with arbitrary A). *For every $\epsilon > 0$, there exists $A \in \mathbb{Q}^{m \times n}$, $b \in \mathbb{Q}^m$, $c \in \mathbb{Q}_+^n$ such that $G_{A,\mathcal{J}}^{pack}$ is a star and*

$$z^{S.S.} \geq |V| \cdot z^I - \epsilon.$$

Now let us consider the case where the sparse cut support list \mathcal{V} corresponds to the natural sparse closure, when $G_{A,\mathcal{J}}^P(V, E)$ is a star or a cycle. Clearly in both these cases we have $D_{\mathcal{V}} = 2$. Therefore we obtain the following Corollary.

Corollary 29 (Natural sparse cuts for two-stage packing-type problem with arbitrary A). *Given (P), with arbitrary $A \in \mathbb{Q}^{m \times n}$, $b \in \mathbb{Q}^m$, $c \in \mathbb{Q}_+^n$. Let \mathcal{J} be a partition of the index set of columns of A (that is $[n]$). Let $G_{A,\mathcal{J}}^{pack} = (V, E)$ be the packing interaction graph of A which is a star or a cycle. If the instance is feasible, then:*

$$z^{N.S.} \leq (|V| - 1) \cdot z^I.$$

We next show that the result of Corollary 29 is tight for $G_{A,\mathcal{J}}^{pack}$ being a star, which corresponds to two-stage packing-type problem with arbitrary A .

Theorem 30. *For every $t \in \mathbb{Z}_{++}$, there exists an instance of (P), with arbitrary $A \in \mathbb{Q}^{m \times n}$, $b \in \mathbb{Q}^m$, $c \in \mathbb{Q}_+^n$, a partition of index set of columns \mathcal{J} such that $G_{A,\mathcal{J}}^{pack}$ is a star with $K + 1$ nodes and*

$$z^{N.S.} = ((K + 1) - 1) \cdot z^I.$$

All proofs of the above results are presented in Section 5.3.

3 Computational experiments

In this section, we present our computational results on the strength of natural sparse closure of pure binary IP.

In Appendix A, we present the algorithm we implemented to estimate z^{cut} , the optimal objective function value of the natural sparse closure.

We describe the random instances we generate in section 3.1 and present the results in section 3.2. All the experiments have been carried out using CPLEX12.5.

3.1 Instance generation

We generated two kinds of problems: two-stage stochastic programming instances and random-graph based instances. We first discuss how we generated the constraint matrix for both types of instances. Then, we discuss how we generated the right-hand side based on the constraint matrix and the objective function.

3.1.1 Constraint matrix generation

To simplify the presentation consider the case of packing instances. Covering instances are generated in the same way.

First we generate the packing-type induced graph on \mathbf{nv} nodes. In case of the two-stage stochastic programming, the packing-type induced graph is a star with \mathbf{nv} nodes (i.e. $\mathbf{nv} - 1$

realizations of the second stage). For the random-graph based instances, let \mathbf{p} (parameter) be the probability that an edge exists between any pair of nodes. As a disconnected induced graph implies that the original problem is decomposable, we accept connected graph only.

Next, given the packing-type induced graph, say $G = (V, E)$ (where $\mathbf{nv} = |V|$), we construct a matrix with that can be partitioned into $|E| \times |V|$ blocks with each block of size $\mathbf{sqr} \times \mathbf{sqr}$ (where \mathbf{sqr} is a parameter). Thus the constraint matrix has $|E| \times \mathbf{sqr}$ rows and $|V| \times \mathbf{sqr}$ columns. The (i, j) th block is all zeros if edge i is not incident to node j . Else the (i, j) th block is a randomly generated dense matrix: We assign each entry the distribution of $unif\{1, \mathbf{M}\}$, where \mathbf{M} is a parameter. For packing-type with arbitrary matrix, we first generate the entry which follows $unif\{1, \mathbf{M}\}$ and then with probability 0.5, we multiply -1 . (Thus, each column block has \mathbf{sqr} variables and there are \mathbf{sqr} rows with the same support of vertices).

3.1.2 Right-hand side generation

To guarantee that the instances generated are non-trivial, we follow the following steps: Randomly select \mathbf{p}_x (parameter) from the set of $\{0.2, 0.4, 0.6, 0.8\}$. A 0-1 vector $x \in \mathbb{R}^n$ is randomly generated where for all $j \in [n]$, $x_j \sim Bernoulli(\mathbf{p}_x)$. A noise vector $\epsilon \in \mathbb{R}_+^m$ is randomly generated as: for all $i \in [m]$, $\epsilon_i \sim unif\{1, \mathbf{M}_\epsilon\}$ (\mathbf{M}_ϵ is a parameter). For a covering instance, $b = Ax - \epsilon$. Otherwise $b = Ax + \epsilon$.

3.1.3 Objective function generation

Every entry in the objective function follows the distribution of $unif\{1, \mathbf{ObjM}\}$. (\mathbf{ObjM} is a parameter.)

3.2 Computational results

3.2.1 Results for two-stage stochastic programming

We set the number of second-stage scenarios equals to 10. We set $\mathbf{sqr} = 20$, that is the number of variables for both first-stage and second-stage scenarios equals to 20. Also we set $\mathbf{M} = \mathbf{M}_\epsilon = \mathbf{ObjM} = 10$. We generated 50 instances for each of the three types of problem.

The result for packing-type problem, covering-type problem, and packing-type problem with arbitrary matrix is shown in Table 1, Table 2, and Table 3 respectively.

Table 1: Two-stage Packing SP

Avg. z^{cut}/z^{IP}	Theoretical bound on z^{cut}/z^{IP}
1.00038	1.9

Table 2: Two-stage Covering SP

Avg. z^{IP}/z^{cut}	Theoretical bound on z^{IP}/z^{cut}
1.009	10

Table 3: Two-stage Arbitrary Packing SP

Avg. z^{cut}/z^{IP}	Theoretical bound on z^{cut}/z^{IP}
1	10

3.2.2 Results for Random Graph based Instances

We set $nv = 10$, $p = 0.2$, $sqr = 20$, $M = M_\epsilon = objM = 10$. For a given random graph we generated 10 random instances, and therefore we generated 50 instances for each of the three types of problem. The result for packing-type problem, covering-type problem, and packing-type problem with arbitrary matrix is shown in Table 4, Table 5, and Table 6 respectively.

Table 4: Random Graph Based tests on Packing Problems

Graph Name	Avg. z^{cut}/z^{IP}	bound of z^{cut}/z^{IP}
Ind 1	1.0009	1.8
Ind 2	1.0028	1.75
Ind 3	1.0053	1.667
Ind 4	1.0006	1.75
Ind 5	1.003	2

Table 5: Random Graph Based tests on Covering Problems

Graph Name	Avg. z^{IP}/z^{cut}	bound of z^{IP}/z^{cut}
Ind 1	1.0045	2
Ind 2	1.0046	3
Ind 3	1.0059	3
Ind 4	1.0053	3
Ind 5	1.0052	3

Table 6: Random Graph Based tests on Arbitrary Packing Problems

Graph Name	z^{cut}/z^{IP}	bound of z^{cut}/z^{IP}
Ind 1	1	9
Ind 2	1	9
Ind 3	1	9
Ind 4	1	9
Ind 5	1	9

4 Conclusions

In this paper, we analyzed the strength of sparse cutting-planes for sparse packing, covering and more general MILP instances. The bounds obtained are completely data independent and in particular depend only on the sparsity structure of the constraint matrix and the support of sparse cuts – in this sense, these results truly provide insight into the strength of sparse cuts for

sparse MILPs. We have shown that the theoretical bounds are tight in many cases. Especially for packing, the theoretical bounds are quite strong, showing that if we have the correct sparse cutting-planes, then the bound obtained by using these cuts may be quite good.

The computational results are interesting: we observe that for all the types of problems sparse cutting planes perform significantly better than the theoretical prediction. This is perhaps not surprising since the theoretical bounds are data-free and therefore “worst-case” in nature. Hence, the empirical experiments are another justification for the main message of this paper: In many cases sparse cuts provide very good bounds for sparse IPs.

5 Proofs

5.1 Proofs for packing problems

For any vector $x \in \mathbb{R}^n$ and $N \subseteq [n]$, we use $x|_N$ to denote the projection of x on the coordinates indexed by N .

We first observe that the column sparse closure $P^{(N)}$ can be viewed essentially as the projection of P^I onto the coordinates indexed by N .

Observation 31. *Consider a mixed-integer linear set with n variables. For any $N \subseteq [n]$, let $P^I|_N$ be the projection of P^I onto the indices in N . Then $x \in P^{(N)}$ if and only if $x \in P^{LP}$ and $x|_N \in P^I|_N$.*

Observation 32. *Consider a mixed-integer set of packing type and let $\mathcal{P} \subseteq \mathbb{R}^n$ be the set of feasible solutions. Then for any set of coordinates $N \subseteq [n]$, $x \in \mathbb{R}^N$ belongs to the projection $P^I|_N$ iff the extension $\tilde{x} \in \mathbb{R}^n$ belongs to \mathcal{P}^I , where $\tilde{x}_i = x_i$ if $i \in N$ and $\tilde{x}_i = 0$ if $i \notin N$.*

5.1.1 Proof of Theorem 11

Recall we want to show that $z^{\mathcal{V},P} \leq \eta^{\mathcal{V}}(G_{A,\mathcal{J}}^{\text{pack}}) \cdot z^I$. (See Example 10 for a concrete example of how the proof works.) In this section we use P to denote the mixed-integer set corresponding to the packing problem (P).

There is a natural identification of sets of nodes of $G_{A,\mathcal{J}}^{\text{pack}}$ with sets of indices of variables, namely if $\mathcal{J} = \{J_1, J_2, \dots, J_q\}$ is the given variable index partition and the nodes of $G_{A,\mathcal{J}}^{\text{pack}}$ are $\{v_1, v_2, \dots, v_q\}$, then the set of vertices $\{v_i\}_{i \in I}$ corresponds to the indices $\bigcup_{i \in I} J_i \subseteq [n]$. We will make use of this correspondence, and in order to make statements precise we use the function $\phi : 2^{V(G_{A,\mathcal{J}}^{\text{pack}})} \rightarrow 2^{[n]}$ to denote this correspondence; with slight abuse of notation, for a singleton set $\{v\}$ we use $\phi(v)$ instead of $\phi(\{v\})$.

Given a set of vertices $S \subseteq V(G_{A,\mathcal{J}}^{\text{pack}})$, let $x^{(S)}$ denote the optimal solution of the packing problem conditioned on all variables x_i outside S taking value 0, or more precisely, $x^{(S)} \in \text{argmax}\{c^T x \mid x \in P^I, x_i = 0 \ \forall i \notin \phi(S)\}$ (we will assume without loss of generality that $x^{(S)}$ is integral); since we are working with a packing problem, this is the roughly same as optimizing over the projection of P^I onto the variables in $\phi(S)$ (but notice $x^{(S)}$ lies in the original space).

We start by showing that, roughly speaking, the closure $P^{(S)}$ captures the original packing maximization problem as long as we ignore the coordinates outside S .

Lemma 33. *For any $x \in P^{(S)}$, we have $(c|_{\phi(S)})^T(x|_{\phi(S)}) \leq c^T x^{(S)}$.*

Proof. Given any $x \in P^{(S)}$, Observation 31 implies that $x|_{\phi(S)} \in \text{proj}_{\phi(S)}(P^I)$. Thus, there exists points $\tilde{x}^1, \dots, \tilde{x}^k \in P^I$ and $\lambda_1, \dots, \lambda_k \in [0, 1]$, such that $x|_{\phi(S)} = \sum_{i=1}^k \lambda_i \cdot (\tilde{x}^i|_{\phi(S)})$ and $\sum_{i=1}^k \lambda_i = 1$. Therefore,

$$(c|_{\phi(S)})^T(x|_{\phi(S)}) = \sum_{i=1}^k \lambda_i \cdot (c|_{\phi(S)})^T(\tilde{x}^i|_{\phi(S)}).$$

To upper bound the right-hand side, consider a point $x^i|_{\phi(S)}$. Let $\tilde{x} \in \mathbb{R}^n$ (the original space) denote the point obtained from $x^i|_{\phi(S)}$ by putting a 0 in all coordinates outside $\phi(S)$, so $\tilde{x}|_{\phi(S)} = x^i|_{\phi(S)}$ and $\tilde{x}|_{[n] \setminus \phi(S)} = \mathbf{0}$. Because P^I is of packing type, notice that \tilde{x} belongs to P^I . The optimality of $x^{(S)}$ then gives that $(c|_{\phi(S)})^T(x^i|_{\phi(S)}) = c^T \tilde{x} \leq c^T x^{(S)}$.

Employing this upper bound on the last displayed equation gives

$$(c|_{\phi(S)})^T(x|_{\phi(S)}) \leq \sum_{i=1}^k \lambda_i \cdot c^T x^{(S)} = c^T x^{(S)},$$

thus concluding the proof. \square

Now we lower bound the packing problem optimum z^I by solutions constructed via mixed stable sets.

Lemma 34. *Given any mixed stable set \mathcal{M} for $G_{A,\mathcal{J}}^{\text{pack}}$, the point $\sum_{M \in \mathcal{M}} x^{(M)}$ belongs to P^I . Thus, $z^I \geq \sum_{M \in \mathcal{M}} c^T x^{(M)}$.*

Proof. We just prove the first statement. First, notice since each $x^{(M)}$ is integral and non-negative, so is the point $\sum_{M \in \mathcal{M}} x^{(M)}$. So consider an inequality $A_i x \leq b_i$ in (P). For any two sets $M_1 \neq M_2 \in \mathcal{M}$, notice that the vector A_i either has all zeros on the indices corresponding to M_1 or on the indices corresponding to M_2 , namely either $A_i|_{\phi(M_1)} = \mathbf{0}$ or $A_i|_{\phi(M_2)} = \mathbf{0}$. Applying this to all pairs of sets in \mathcal{M} , we get that there is only one set $M^* \in \mathcal{M}$ such that $A_i|_{\phi(M^*)}$ is non-zero, which implies that

$$A_i \sum_{M \in \mathcal{M}} x^{(M)} = \sum_{M \in \mathcal{M}} A_i x^{(M)} = \sum_{M \in \mathcal{M}} (A_i|_{\phi(M)})(x^{(M)}|_{\phi(M)}) = A_i x^{(M^*)} \leq b_i,$$

where the last inequality follows from the feasibility of $x^{(M^*)}$. Thus, the point $\sum_{M \in \mathcal{M}} x^{(M)}$ satisfies all inequalities $A_i x \leq b_i$ of the system (P), concluding the proof. \square

Now, we present the proof of Theorem 11.

Proof of Theorem 11. Let $V = \{v_1, \dots, v_q\}$ denote the vertices of $G_{A,\mathcal{J}}^{\text{pack}}$, and let MSS denote the set of all mixed stable sets of $G_{A,\mathcal{J}}^{\text{pack}}$ with respect to \mathcal{V} . Let $\{y_M\}_{M \in \text{MSS}}$ be an optimal solution of linear problem (1) corresponding to the definition of mixed fractional chromatic number with respect to \mathcal{V} , and define $g = \sum_{M \in \text{MSS}} \chi_M y_M \in \mathbb{R}^q$. Based on the constraints of (1) we have that $g \geq \mathbf{1}$.

We upper bound the optimum $z^{\mathcal{V},P}$ of the column block-sparse closure. For that let

$$x^* = \text{argmax}\{c^T x \mid x \in P^{\mathcal{V},P}\},$$

be the optimal solution. Then breaking up the indices of the variables based on the nodes V and using the non-negativity of c and x^* , we have

$$\begin{aligned}
z^{\mathcal{V},P} &= c^T x^* = \sum_{j=1}^q (c|_{\phi(v_j)})^T (x^*|_{\phi(v_j)}) \leq \sum_{j=1}^q g_j \cdot (c|_{\phi(v_j)})^T (x^*|_{\phi(v_j)}) \\
&= \sum_{j=1}^q \left(\sum_{\mathcal{M} \in \text{MSS}} (\chi_{\mathcal{M}})_j \cdot y_{\mathcal{M}} \right) \cdot (c|_{\phi(v_j)})^T (x^*|_{\phi(v_j)}) \\
&= \sum_{\mathcal{M} \in \text{MSS}} y_{\mathcal{M}} \cdot \left(\sum_{j=1}^q (\chi_{\mathcal{M}})_j \cdot (c|_{\phi(v_j)})^T (x^*|_{\phi(v_j)}) \right) \\
&= \sum_{\mathcal{M} \in \text{MSS}} y_{\mathcal{M}} \cdot \left(\sum_{\{j \mid v_j \in \mathcal{M}\}} (c|_{\phi(v_j)})^T (x^*|_{\phi(v_j)}) \right) \\
&= \sum_{\mathcal{M} \in \text{MSS}} y_{\mathcal{M}} \cdot \left(\sum_{S \in \mathcal{M}} (c|_{\phi(S)})^T (x^*|_{\phi(S)}) \right).
\end{aligned}$$

To further upper bound the right-hand side consider some $\mathcal{M} \in \text{MSS}$, some $S \in \mathcal{M}$ and the term $(c|_{\phi(S)})^T (x^*|_{\phi(S)})$. First we claim that x^* belongs to the column block-sparse closure $P^{(S)}$. To see this, first recall from the definition of mixed stable set that there must be a set V_S in the support list \mathcal{V} containing S . Moreover, since $x^* \in P^{\mathcal{V},P} = \bigcap_{V' \in \mathcal{V}} P^{(V')}$, we have $x^* \in P^{(V_S)}$; finally, the monotonicity of closures implies $P^{(S)} \supseteq P^{(V_S)}$, and hence $x^* \in P^{(S)}$. Thus we can employ Lemma 33 to obtain the upper bound $(c|_{\phi(S)})^T (x^*|_{\phi(S)}) \leq c^T x^{(S)}$.

Plugging this bound on last displayed inequality and using Lemma 34 we then get

$$z^{\mathcal{V},P} \leq \sum_{\mathcal{M} \in \text{MSS}} y_{\mathcal{M}} \cdot \left(\sum_{S \in \mathcal{M}} c^T x^{(S)} \right) \leq \sum_{\mathcal{M} \in \text{MSS}} y_{\mathcal{M}} \cdot z^I = \eta^{\mathcal{V}}(G_{A,\mathcal{J}}^{\text{pack}}) \cdot z^I.$$

This concludes the proof. \square

5.1.2 Proof of Corollary 12 and Theorem 13

Brooks' Theorem [5] is the following result (recall that a proper coloring of a graph is an assignment of colors to the vertices such that no edge has the same color on both endpoints).

Theorem 35 (Brook's Theorem). *Consider a connected graph G of max degree Δ . Then G can be properly colored by Δ colors, except in two cases either when G is a complete graph or an odd cycle, in which case it can be properly colored with $\Delta + 1$ colors.*

Since the fractional chromatic number is a lower bound on the chromatic number, we obtain Corollary 12.

We now prove Theorem 13.

Proof of Theorem 13. Recall that super sparse closure corresponds to the support list

$$\mathcal{V} = \{\{v_1\}, \{v_2\}, \dots, \{v_{|V|}\}\}.$$

The proof of both parts is similar.

Part 1. We want to show an example where $z^{SS} \geq (3 - \epsilon)z^I$ for all $\epsilon > 0$ where $G_{A,\mathcal{J}}^{\text{pack}}$ is a 3-cycle. We will construct an integer program with 3 variables and $\mathcal{J} = \{\{1\}, \{2\}, \{3\}\}$. Given $\epsilon > 0$, consider the following packing integer program:

$$\begin{aligned}
\max \quad & x_1 & +x_2 & +x_3 \\
\text{s.t.} \quad & x_1 & +x_2 & & \leq 2 - \frac{2}{3}\epsilon \\
& x_1 & & +x_3 & \leq 2 - \frac{2}{3}\epsilon \\
& & x_2 & +x_3 & \leq 2 - \frac{2}{3}\epsilon \\
& x_1 \in \mathbb{Z}_+, & x_2 \in \mathbb{Z}_+, & x_3 \in \mathbb{Z}_+
\end{aligned} \tag{2}$$

Clearly $G_{A,\mathcal{J}}^{\text{pack}}$ is a 3-cycle. Note that the only valid inequalities that have support on each of the three blocks defined by \mathcal{J} is $x_i \leq 1$ for $i = 1, 2, 3$. Thus, the point $(1 - \frac{\epsilon}{3}, 1 - \frac{\epsilon}{3}, 1 - \frac{\epsilon}{3})$ belongs to the super sparse closure $P^{S.S.}$, and hence the optimum value satisfies $z^{S.S.} \geq 3 - \epsilon$. On the other hand clearly, $z^{IP} = 1$, concluding the proof.

Part 2. We want to show an example where $z^{SS} \geq (2 - \epsilon)z^I$ for all $\epsilon > 0$ where $G_{A,\mathcal{J}}^{\text{pack}}$ is a star. Take $\Delta \in \mathbb{Z}_+$. We construct a packing integer program with 2Δ variables and $\mathcal{J} = \{\{1, 2, \dots, \Delta\}, \{\Delta + 1\}, \{\Delta + 2\}, \dots, \{2\Delta\}\}$. Given $\epsilon > 0$, consider the following integer program

$$\begin{aligned}
\max \quad & \sum_{i=1}^{2\Delta} x_i \\
\text{s.t.} \quad & x_i + x_{\Delta+i} \leq 2 - \epsilon \quad \forall i \in [\Delta] \\
& x \in \mathbb{Z}_+^{2\Delta}.
\end{aligned}$$

Clearly $G_{A,\mathcal{J}}^{\text{pack}}$ is a star with Δ leaves. Letting P be the associated mixed integer set of the above integer program, note that the projection of P^I to the first block of variables $\{1, 2, \dots, \Delta\}$ equals $[0, 1]^\Delta$. Also the only valid inequalities that have support on each of the other Δ blocks $\{\Delta + i\}$ is $0 \leq x_i \leq 1$ for $i \in \{\Delta + 1, \dots, 2\Delta\}$. Thus, the point x with $x_i = 1$ for all $i \in [\Delta]$ and $x_i = 1 - \epsilon$ for all $i \in \{\Delta + 1, \dots, 2\Delta\}$ belongs to the super sparse closure $P^{S.S.}$. Thus the optimum $z^{S.S.}$ is at least $2\Delta - \Delta\epsilon$. On the other hand, clearly $z^I = \Delta$, concluding the proof. \square

5.1.3 Proof of Theorem 14

We prove the desired upper bound $z^{N.S.} \leq (\frac{2\Delta-1}{\Delta}) \cdot z^I$. Due to Theorem 11, it suffices to upper bound the fractional chromatic number $\eta^{\mathcal{V}}(G_{A,\mathcal{J}}^{\text{pack}}) \leq \frac{2\Delta-1}{\Delta}$ for \mathcal{V} set according to the natural sparse closure setting. Notice however, that in this setting every edge of $E = E(G_{A,\mathcal{J}}^{\text{pack}})$ belongs to some set in \mathcal{V} and vice-versa, and therefore $\eta^{\mathcal{V}}(G_{A,\mathcal{J}}^{\text{pack}}) = \eta^E(G_{A,\mathcal{J}}^{\text{pack}})$. Thus, it suffices to prove $\eta^E(G_{A,\mathcal{J}}^{\text{pack}}) \leq \frac{2\Delta-1}{\Delta}$.

The following is the main tool for providing an efficient mixed stable set fractional coloring.

Lemma 36. *Let $T = (V, E)$ be a tree of maximum degree Δ . Then there is a collection of $2\Delta - 1$ sets of edges $E_1, E_2, \dots, E_{2\Delta-1}$ and $2\Delta - 1$ sets of nodes $V_1, V_2, \dots, V_{2\Delta-1}$ satisfying the following:*

1. For each $i \in [2\Delta - 1]$ the collection $E_i \cup V_i$ is a mixed stable set for T subordinate to E
2. Each node of T is covered exactly Δ times by the collection of mixed stable sets $\{E_i \cup V_i\}_{i \in [2\Delta-1]}$.

Proof. If T consists of a single edge, then $\Delta = 1$ and we can simply set E_1 to be the edge of T and set $V_1 = \emptyset$ to get the desired sets. So assume that T has at least one internal node.

In order to simplify the proof we make all the degrees the same: construct the tree T' from T by adding new leaves to all internal nodes of T so that now every internal node of T' has degree exactly Δ . We will construct the desired sets $\{E'_i\}_{i \in [2\Delta-1]}$ and $\{V'_i\}_{i \in [2\Delta-1]}$ for T' via a coloring argument reminiscent of the proof of Brook's Theorem (although not the same argument).

Pick any internal node v_0 of T' and root this tree at v_0 . We label all edges and leaf nodes of T' with numbers in $[2\Delta - 1]$ according to the following BFS procedure (we use the standard meaning of "parent", "child", "depth" (where v_0 has depth 0, L is the maximum dept of any node), etc. for rooted trees):

Label each of the Δ edges incident to the root v_0 with a distinct label
for $i = 1$ to L **do**
 for every vertex v of depth i **do**
 Let S denote the set of labels assigned to all the edges incident to the parent of v and notice that $|S| = \Delta$
 if v is an internal node **then**
 Label the $\Delta - 1$ edges of v to its children with distinct labels from the set $[2\Delta - 1] \setminus S$
 else
 Assign **all** $\Delta - 1$ labels $[2\Delta - 1] \setminus S$ to v .

Then for all $j \in [2\Delta - 1]$, let E'_j (resp. V'_j) be set of edges (resp. nodes) of T' that have label j (notice that vertices have multiple labels).

It follows directly from the labeling procedure that each set $E'_i \cup V'_i$ is a mixed stable set of T' (and clearly subordinate to the edges of T'). Now to see that each node v of T' is covered exactly Δ times by the collection $\{E'_i \cup V'_i\}_{i \in [2\Delta-1]}$ we consider 2 cases: If v is an internal node, then by construction of T' it has degree exactly Δ and since $\bigcup_i E'_i = E(T')$ it is covered Δ times by the collection $\{E'_i\}_i$ and 0 times by the collection $\{V'_i\}_i$, giving the desired result. On the other hand, if v is a leaf of T' , then it is covered once by the set E'_i where i is the label of the only edge incident to v , covered by no other set E'_j , and covered by the $\Delta - 1$ sets V'_j corresponding to the labels of v . Thus, the sets $\{E'_i\}_i$ and $\{V'_i\}_i$ satisfy the desired properties with respect to the modified tree T' .

Now to get the desired sets for the original tree T , we just remove the nodes in $T' \setminus T$ from the sets $E'_i \cup V'_i$: for each $\{v, v'\}$ with $v \in V(T)$ and $v' \notin V(T)$ that belongs to $E'_i \cup V'_i$, replace it with the singleton $\{v\}$; denote the set obtained by $E_i \cup V_i$ (concretely, E_i is the set of pairs in this collection and V_i is the set of singletons in this collection). Notice that this replacement procedure does not add repeated singletons: this is because $E'_i \cup V'_i$ contains only disjoint edges (the labeling scheme above does not assign color i to two intersecting edges) and if it contains an edge (v, v') with $v' \notin V(T)$ then this implies that v is an internal node of T and hence the singleton $\{v\}$ does not belong to $E'_i \cup V'_i$.

It follows directly from this replacement operation that the sets $E_i \cup V_i$'s are mixed stable sets for T subordinate to $E(T)$ and that still each node of T is covered exactly Δ times by them. This concludes the proof. \square

The upper bound in Theorem 14 the follows from the following corollary.

Corollary 37. *Let H be a tree of maximum degree Δ . Then $\eta^E(H) \leq \frac{2\Delta-1}{\Delta}$.*

Proof. Consider the mixed stable sets $\mathcal{M}_i = E_i \cup V_i$ (for $i \in [2\Delta - 1]$) of H obtained from Lemma 36. Since each vertex of H is covered exactly Δ times by $\{\mathcal{M}_i\}_{i \in [2\Delta-1]}$, we have that setting $y_{\mathcal{M}_i} = \frac{1}{\Delta}$ for all i (and $y_{\mathcal{M}} = 0$ otherwise) yields a feasible solution for the mixed fractional chromatic number program (1) of value $\frac{2\Delta-1}{\Delta}$, proving the result. \square

5.1.4 Proof of Theorem 15

Fix $\epsilon > 0$; we construct an instance where $z^{N.S.} \geq \left(\frac{2\Delta-1}{\Delta} - \epsilon\right) \cdot z^I$. The construction requires the existence of the so-called *affine designs*.

Definition 38. *Given $n \in \mathbb{Z}_{++}$, we call an affine n -design a collection $\mathcal{F}_1, \dots, \mathcal{F}_n$ where each \mathcal{F}_i is a family of n -subsets of $[n^2]$ satisfying:*

1. *For any $i \in [n]$, the sets in \mathcal{F}_i partition $[n^2]$*
2. *For any $i \neq j \in [n]$ and $A \in \mathcal{F}_i$ and $B \in \mathcal{F}_j$, we have $|A \cap B| \leq 1$.*

Theorem 39 ([7], Part VII, Point 2.17). *For every prime n , an affine n -design exists.*

So consider a prime number $n \geq \Delta$ and let $\mathcal{F}_1, \dots, \mathcal{F}_n$ be an affine n -design. For a set $A \in \mathcal{F}_i$ we use $\chi_A \in \{0, 1\}^{n^2}$ to denote the indicator vector of the set A .

We will construct a packing IP in $\mathbb{R}_+^{n^2+\Delta}$ and partition the $n^2 + \Delta$ variables into $\Delta + 1$ blocks $\mathcal{J} = \{J_0, \dots, J_\Delta\}$ by setting $J_0 = \{1, \dots, n^2\}$ and $J_i = \{n^2 + i\}$, for $i = 1, \dots, \Delta$. To simplify the notation we use $x \in \mathbb{R}^{n^2}$ to represent the variables in J_0 , and $y_i, i = 1, \dots, \Delta$, to represent the variables in J_i respectively.

Let P_i be the polytope in $\mathbb{R}^{n^2+\Delta}$ given by the convex hull of the points

$$\left\{ (x, y_1, \dots, y_\Delta) \in \{0, 1\}^{n^2+\Delta} \mid \sum_j x_j \leq n \text{ and, if } y_i = 1, \text{ then } x \leq \chi_A \text{ for some } A \in \mathcal{F}_i \right\};$$

explicitly, this is the set of solutions satisfying

$$\begin{aligned} \sum_j x_j &\leq n \\ x_a + x_b + y_i &\leq 2 \quad \forall a \in A, b \in B, A \neq B \in \mathcal{F}_i \\ (x, y_1, \dots, y_\Delta) &\in [0, 1]^{n^2+\Delta}. \end{aligned} \tag{3}$$

Then the desired IP (P) is obtained by considering the integer solutions common to all these polytopes:

$$\begin{aligned} \max \sum_{j \in [n^2]} x_j + \left(\frac{n-1}{\Delta-1}\right) \sum_{j \in [\Delta]} y_j \\ (x, y_1, \dots, y_\Delta) \in \bigcap_{i \in [\Delta]} P_i \cap \mathbb{Z}^{n^2+\Delta}. \end{aligned}$$

Again we get the following interpretation for the feasible solutions for this problem: in any solution $(x, y) \in \{0, 1\}^{n^2+\Delta}$, $\sum_j x_j \leq n$ and for all $i \in [\Delta]$

$$\text{if } y_i = 1, \text{ then } x \leq \chi_A \text{ for some set } A \in \mathcal{F}_i \quad (4)$$

Let P denote the integer set corresponding to this problem. From the explicit description of the P_i 's we see that this is packing integer program whose induced graph $G_{A, \mathcal{J}}^{\text{pack}}$ is a star with maximum degree Δ .

The intuition behind this construction is the following: first, maximizing the objective function over just $P_i \cap \mathbb{Z}^{n^2+\Delta}$ (or equivalently over P_i) gives value $n + \left(\frac{n-1}{\Delta-1}\right) \cdot \Delta \approx 2n$ (by taking $x = \chi_A$ for any $A \in \mathcal{F}_i$, $y_j = 1$ for all j). Moreover, recall that the natural sparse closure w.r.t. \mathcal{J} of the full program (P) uses cuts that are only supported in (x, y_i) , for $i \in [\Delta]$; thus, roughly speaking, this closure sees each $P_i \cap \mathbb{Z}^{n^2+\Delta}$ independently, and not really capturing the fact they are being intersected. Thus, optimizing over the natural sparse closure w.r.t. \mathcal{J} still gives value $\approx 2n$. However, due to the fact the sets across the design's \mathcal{F}_i 's are almost disjoint, intersecting the regions $P_i \cap \mathbb{Z}^{n^2+\Delta}$ kills most of the solutions. A bit more precisely, the almost disjointness in the affine design and expression (4) imply that the best solution either sets many of the y_i 's to 1 and almost all x_j 's to 0, or sets all x_j 's to 1 and few y_i 's to 0; these solutions gives value $\approx n$. This gives the desired gap of ≈ 2 between the natural sparse closure and the original IP.

To make this formal, we start with the following lemma.

Lemma 40. *Setting $x = (\frac{1}{n}, \dots, \frac{1}{n})$ and $y_j = 1$ for all j gives a feasible solution to the natural sparse closure $P^{N.S.}$. Thus, $z^{N.S.} \geq n + \left(\frac{n-1}{\Delta-1}\right) \cdot \Delta$.*

Proof. Let $\bar{x} = (\frac{1}{n}, \dots, \frac{1}{n})$ and $\bar{y} = (1, \dots, 1)$ denote the desired solution.

We claim that it suffices to prove that (\bar{x}, e^i) belongs to P^I for all i , (where e^i is the i th canonical basis vector in \mathbb{R}^Δ). To see that, first notice that the natural sparse closure w.r.t. \mathcal{J} is $P^{N.S.} = \bigcap_{i \in [\Delta]} P^{(x, y_i)}$, where we use $P^{(x, y_i)}$ to denote the sparse closure of P with cuts on variables (x, y_i) (see Definition 2). Using Observations 31 and 32, it suffices to show $(\bar{x}, \bar{y}) \in P^{LP}$ and $(\bar{x}, e^i) \in P^I$. The former condition can be easily verified via equation (3), so it suffices to show $(\bar{x}, e^i) \in P^I$ for all i .

So fix $i \in [\Delta]$. Consider the collection \mathcal{F}_i and a point of the form (χ_A, e^i) for any set $A \in \mathcal{F}_i$. By definition of P_i , notice that (χ_A, e^i) belongs to $P_i \cap \mathbb{Z}^{n^2+\Delta}$. Moreover, notice that for $j \neq i$ we also have $(\chi_A, e^i) \in P_j$: this follows from the facts $\sum_j (\chi_A)_j \leq n$ and $e_j^i = 0$. Thus, we have $(\chi_A, e^i) \in P^I = \bigcap_{j \in [\Delta]} P_j \cap \mathbb{Z}^{n^2+\Delta}$. Then the average $\sum_{A \in \mathcal{F}_i} \frac{1}{n} (\chi_A, e^i)$ belongs to P^I ; since the sets in \mathcal{F}_i form a partition of $[n^2]$, $\sum_{A \in \mathcal{F}_i} \chi_A = (1, \dots, 1)$, and hence the average is $\sum_{A \in \mathcal{F}_i} \frac{1}{n} (\chi_A, e^i) = (\bar{x}, e^i) \in P^I$. This concludes the proof. \square

The next step is to understand P better.

Lemma 41. *For any solution $(x, y_1, \dots, y_\Delta) \in P$ with $\sum_{i \in [\Delta]} y_i \geq 2$ we have $\sum_{j \in [n^2]} x_j \leq 1$.*

Proof. Consider $p \neq q$ such that $y_p = y_q = 1$. By definition of P , we have that the solution $(x, y_1, \dots, y_\Delta)$ belongs to P_p and P_q . Since $y_p = y_q = 1$, this means that there are sets $A \in \mathcal{F}_p$ and $B \in \mathcal{F}_q$ such that $x \leq \chi_A$ and $x \leq \chi_B$, which further implies $x \leq \chi_{A \cap B}$. But by definition of an affine design $|A \cap B| \leq 1$, and hence $\sum_{j \in [\Delta]} x_j \leq 1$. This concludes the proof. \square

Corollary 42. We have $z^I \leq \frac{n\Delta-1}{\Delta-1}$.

Proof. Consider any feasible solution $(x, y_1, \dots, y_\Delta)$ to P^I . From the first constraint in (3) we have $\sum_j x_j \leq n$. Thus, if $\sum_i y_i \leq 1$, the solution has value at most $n + \left(\frac{n-1}{\Delta-1}\right) = \frac{n\Delta-1}{\Delta-1}$; on the other hand, using Lemma 41, if $\sum_i y_i \geq 2$ then the solution has value at most $1 + \left(\frac{n-1}{\Delta-1}\right) \Delta = \frac{n\Delta-1}{\Delta-1}$. Together these give the desired upper bound. \square

Lemma 40 and Corollary 42 give that

$$\frac{z^{N.S.}}{z^I} \geq \left(n + \left(\frac{n-1}{\Delta-1} \right) \cdot \Delta \right) \frac{\Delta-1}{n\Delta-1} = \frac{2n\Delta - n - \Delta}{n\Delta - 1} = \frac{2\Delta - 1 - \Delta/n}{\Delta - 1/n}.$$

Since $\lim_{n \rightarrow \infty} \frac{2\Delta-1-\Delta/n}{\Delta-1/n} = \frac{2\Delta-1}{\Delta}$, for a sufficiently large choice of n we get $z^{N.S.} \geq \left(\frac{2\Delta-1}{\Delta} - \epsilon\right) z^I$. This concludes the proof of Theorem 15.

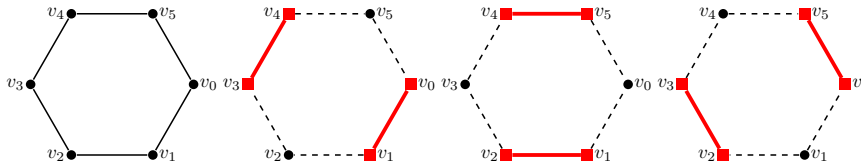
5.1.5 Proof of the first part of Theorem 17: upper bound on $z^{N.S.}$

Consider the packing interaction graph $G_{A,\mathcal{J}}^{\text{pack}}$, which is a cycle of length K . Notice that the natural sparse closure in this case corresponds to considering the support list \mathcal{V} being simply the edges of $G_{A,\mathcal{J}}^{\text{pack}}$. Thus, to prove the first part of Theorem 17 it suffices to upper bound the fractional mixed chromatic number $\eta^{E(G_{A,\mathcal{J}}^{\text{pack}})}(G_{A,\mathcal{J}}^{\text{pack}})$.

We can work more abstractly to simplify things: let $H = (V, E)$ be the cycle $v_0 - v_1 - \dots - v_{K-1} - v_0$ on K nodes, and we need to upper bound $\eta^E(H)$. To further simplify the notation, we identify v_i with $v_{i \pmod K}$ for $i \geq K$. We consider the different cases depending on $K \pmod 3$.

Case 1: $K = 3k$, $k \in \mathbb{Z}_{++}$. For $i = 0, 1, 2$, let \mathcal{M}_i denote the set of edges $\{v_j, v_{j+1}\}$ where $j = i \pmod 3$. It is clear that each \mathcal{M}_i is a mixed stable set for H subordinate to E . Moreover, since $\bigcup_{i=0}^2 \mathcal{M}_i = E$ covers each node of H exactly twice, we can find a solution for the fractional mixed chromatic number LP (1) by setting $y_{\mathcal{M}_i} = \frac{1}{2}$ for $i = 0, 1, 2$. This gives the desired bound $\eta^E(H) \leq \frac{3}{2}$.

Figure 2: Constructions of all mixed stable sets for 6-cycle



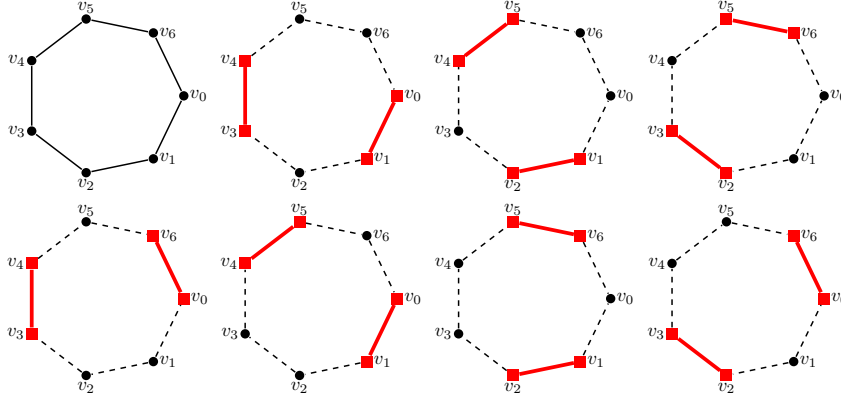
Case 2: $K = 3k + 1$, $k \in \mathbb{Z}_{++}$. We show $\eta^E(H) \leq \frac{3k+1}{2k}$. If $k = 1$, we have that H is a 4-cycle and define $\mathcal{M}_0 = \{(v_0, v_1)\}$, $\mathcal{M}_1 = \{(v_1, v_2)\}$, $\mathcal{M}_2 = \{(v_2, v_3)\}$ and $\mathcal{M}_3 = \{(v_3, v_0)\}$. Clearly these \mathcal{M}_i 's are mixed stable sets for H subordinate to E and $\bigcup_i \mathcal{M}_i$ covers each node of H exactly twice; then as in the previous case, this gives $\eta^E(H) \leq \frac{4}{2} = 2 = \frac{3k+1}{2k}$.

For $k \geq 2$, define

$$\mathcal{M}_i = \left\{ \{v_i, v_{i+1}\}, \{v_{i+3}, v_{i+4}\}, \dots, \{v_{3(k-2)+i}, v_{3(k-2)+i+1}\}, \{v_{3(k-1)+i}, v_{3(k-1)+i+1}\} \right\}$$

for $i = \{0, \dots, 3k\}$. It is straightforward to check that each \mathcal{M}_i is a mixed stable set subordinate to E and that $\bigcup_{i=0}^{3k} \mathcal{M}_i$ covers every node exactly $2k$ times. Thus again we get $\eta^E(H) \leq \frac{3k+1}{2k}$.

Figure 3: Constructions of all mixed stable sets for 7-cycle

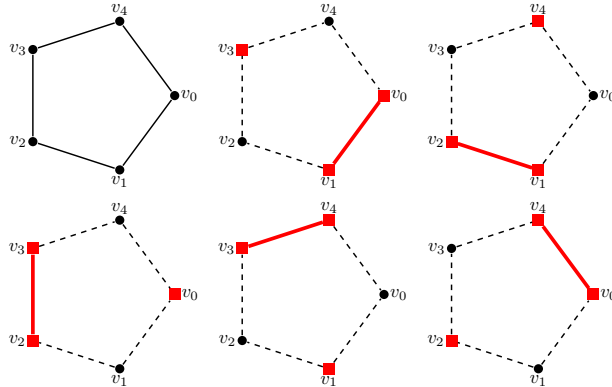


Case 3: $K = 3k + 2$, $k \in \mathbb{Z}_{++}$. Let

$$\mathcal{M}_i = \left\{ \{v_i, v_{i+1}\}, \{v_{i+3}, v_{i+4}\}, \dots, \{v_{3(k-2)+i}, v_{3(k-2)+i+1}\}, \{v_{3(k-1)+i}, v_{3(k-1)+i+1}\}, \{v_{3k+i}\} \right\}$$

for $i = \{0, \dots, 3k + 1\}$. It is straightforward to check that each \mathcal{M}_i is a mixed stable set subordinate to E and that $\bigcup_{i=0}^{3k+1} \mathcal{M}_i$ covers every node exactly $2k + 1$ times. Thus we have $\eta^E(H) \leq \frac{3k+2}{2k+1}$. This concludes the proof of the first part of the theorem.

Figure 4: Constructions of all mixed stable sets for 5-cycle



5.1.6 Proof of second part of Theorem 17: tight instances

The construction of the tight instances is similar to the one used in Theorem 15. So consider a prime number $n \geq K$ and let $\mathcal{F}_1, \dots, \mathcal{F}_n$ be an affine n -design. For a set $A \in \mathcal{F}_i$ again we use $\chi_A \in \{0, 1\}^{n^2}$ to denote the indicator vector of the set A .

We will construct a packing IP with Kn^2 variables, which are partitioned into K equally sized blocks $\mathcal{J} = \{J_0, \dots, J_{K-1}\}$, namely $J_i = \{n^2i, n^2i + 1, \dots, n^2i + n^2 - 1\}$. To simplify the notation, we use $x^i \in \mathbb{R}^{n^2}$ to represent the variables corresponding to J_i , so a solution of the IP has the form (x^0, \dots, x^{K-1}) . For $i \geq K$, we use x^i to denote $x^{i \pmod K}$.

First, define the integer set $Q = \{x \in \{0, 1\}^{n^2} \mid \mathbb{1}^T x \leq n\}$. Then, for $i \in \{0, \dots, K-1\}$ let P_i be the polytope in \mathbb{R}^{Kn^2} given by the convex hull of the points

$$\left\{ (x^0, \dots, x^{K-1}) \in Q^K \mid \begin{array}{l} \text{if } x^i \neq 0, \text{ then } x^{i+1} \leq \chi_A \text{ for some } A \in \mathcal{F}_i, \text{ and} \\ \text{if } x^{i+1} \neq 0, \text{ then } x^i \leq \chi_A \text{ for some } A \in \mathcal{F}_i \end{array} \right\};$$

explicitly, this is the set of solutions satisfying

$$\begin{aligned} \mathbb{1}^T x^j &\leq n && \forall j \\ x_a^i + x_b^i + x_c^i &\leq 2 && \forall a \in A, b \in B, A \neq B \in \mathcal{F}_i, \forall c \\ x_a^{i+1} + x_b^{i+1} + x_c^i &\leq 2 && \forall a \in A, b \in B, A \neq B \in \mathcal{F}_i, \forall c \\ (x^0, \dots, x^{K-1}) &\in [0, 1]^{Kn^2}. \end{aligned} \tag{5}$$

Then the desired IP (P) is obtained by considering the integer solutions common to all these polytopes:

$$\begin{aligned} \max \sum_{i=0}^{K-1} \mathbb{1}^T x^i \\ (x^0, \dots, x^{K-1}) \in \bigcap_{i=0}^{K-1} P_i \cap \mathbb{Z}^{Kn^2}. \end{aligned}$$

Again we get the following interpretation for the feasible solutions for this problem: in any solution $(x^0, \dots, x^{K-1}) \in \{0, 1\}^{Kn^2}$, $\mathbb{1}^T x^i \leq n$ for all i , and also for all i

$$\begin{aligned} \text{if } x^i \neq 0, \text{ then } x^{i+1} \leq \chi_A \text{ for some set } A \in \mathcal{F}_i, \text{ and} \\ \text{if } x^{i+1} \neq 0, \text{ then } x^i \leq \chi_A \text{ for some set } A \in \mathcal{F}_i. \end{aligned} \tag{6}$$

Let P denote the integer set corresponding to this problem. From the explicit description of the P_i 's we see that this is packing integer program whose induced graph $G_{A, \mathcal{J}}^{\text{pack}}$ is a K -cycle.

We now consider the natural sparse closure $P^{N.S.}$ and the integer hull P^I for this problem and lower bound the ratio $z^{N.S.}/z^I$. For that, given $x = (x^0, \dots, x^{K-1})$, let $\text{high}(x) = \{i \mid \mathbb{1}^T x^i \geq 2\}$, namely the set of block of variables with ‘‘high’’ value. We say that three integers are *adjacent mod K* if they are of the form $i \pmod K, i+1 \pmod K, i+2 \pmod K$.

Lemma 43. *For any solution $x \in P$, the set $\text{high}(x)$ does not contain any three adjacent mod K integers.*

Proof. By contradiction, assume that $\text{high}(x)$ contains the integers $i \pmod K, i+1 \pmod K$, and $i+2 \pmod K$. In particular, all of x^i, x^{i+1} and x^{i+2} are different from 0, and hence expression (6) implies that $x^{i+1} \leq \chi_A$ and $x^{i+1} \leq \chi_B$ for some $A \in \mathcal{F}_i \pmod K$ and $B \in \mathcal{F}_{i+1} \pmod K$; this implies that $x^{i+1} \leq \chi_{A \cap B}$. But by definition of affine design, we have $|A \cap B| \leq 1$, and hence $\mathbb{1}^T x^{i+1} \leq 1$, reaching a contradiction. \square

The following lemma can be easily checked.

Lemma 44. *Let S be a subset of $\{0, \dots, K-1\}$ that does not contain any three adjacent mod K integers. Then: (i) if $K = 3k$ or $K = 3k+1$ for $k \in \mathbb{Z}_{++}$ we have $|S| \leq 2k$, and; (ii) if $3k+2$ for $k \in \mathbb{Z}_{++}$ we have $|S| \leq 2k+1$.*

Lemma 45. *The optimal value of the integer program (P) can be upper bounded as follows: if $K = 3k$ or $K = 3k+1$ for $k \in \mathbb{Z}_{++}$, $z^I \leq (n-1) \cdot 2k + K$; if $K = 3k+2$ for $k \in \mathbb{Z}_{++}$, $z^I \leq (n-1) \cdot (2k+1) + K$.*

Proof. Let $\bar{x} = (\bar{x}^0, \dots, \bar{x}^{K-1})$ be an optimal solution to (P). Using the fact that $\mathbb{1}^T \bar{x}^i \leq n$ and the definition of $\text{high}(\bar{x})$ we get

$$\begin{aligned} z^I &= \sum_{i=0}^{K-1} \mathbb{1}^T \bar{x}^i = \sum_{i \in \text{high}(\bar{x})} \mathbb{1}^T \bar{x}^i + \sum_{i \notin \text{high}(\bar{x})} \mathbb{1}^T \bar{x}^i \\ &\leq n \cdot |\text{high}(\bar{x})| + K - |\text{high}(\bar{x})| = (n-1) \cdot |\text{high}(\bar{x})| + K. \end{aligned}$$

Upper bounding $|\text{high}(\bar{x})|$ using Lemmas 43 and 44 gives the desired result. \square

Lemma 46. *The point $\bar{x} = (\frac{1}{n}\mathbb{1}, \dots, \frac{1}{n}\mathbb{1})$ is a feasible solution to the natural sparse closure $P^{N.S.}$. Thus, $z^{N.S.} \geq Kn$.*

Proof. To simplify the notation, let $\text{zero}^i(x, x') \in \mathbb{R}^{n^2} \times \dots \times \mathbb{R}^{n^2}$ denote the vector

$$(0, \dots, 0, x, x', 0, \dots, 0)$$

where x is in the i th position and x' is in position $i+1 \pmod K$.

We claim that it suffices to prove that $\text{zero}^i(\mathbb{1}/n, \mathbb{1}/n)$ belongs to P^I for all i . To see that, first notice that the natural sparse closure w.r.t. \mathcal{J} is $P^{N.S.} = \bigcap_{i=0}^{K-1} P^{(x^i, x^{i+1})}$, where we use $P^{(x^i, x^{i+1})}$ to denote the sparse closure of P on variables (x^i, x^{i+1}) . Using Observations 31 and 32, it suffices to show $\bar{x} \in P^{LP}$ and $\text{zero}^i(\mathbb{1}/n, \mathbb{1}/n) \in P^I$. The former condition can be easily verified via equation (3), so it suffices to show $\text{zero}^i(\mathbb{1}/n, \mathbb{1}/n) \in P^I$.

So fix i . Consider the collection \mathcal{F}_i . By the definition of P_i , for each $A, B \in \mathcal{F}_i$ the point $\text{zero}^i(\chi_A, \chi_B)$ belongs to $P_i \cap \mathbb{Z}^{Kn^2}$. It also follows directly from the definition of P_j that $\text{zero}^i(\chi_A, \chi_B) \in P_j$ for all $j \neq i$. Thus, we have $\text{zero}^i(\chi_A, \chi_B) \in P^I = \bigcap_{j=0}^{K-1} P_j \cap \mathbb{Z}^{n^2+\Delta}$. Then the following average belongs to P^I :

$$\sum_{A \in \mathcal{F}_i} \frac{1}{n} \sum_{B \in \mathcal{F}_i} \frac{1}{n} \text{zero}^i(\chi_A, \chi_B) = \sum_{A \in \mathcal{F}_i} \frac{1}{n} \text{zero}^i \left(\chi_A, \sum_{B \in \mathcal{F}_i} \frac{1}{n} \chi_B \right) = \text{zero}^i \left(\sum_{A \in \mathcal{F}_i} \frac{1}{n}, \sum_{B \in \mathcal{F}_i} \frac{1}{n} \right).$$

Recalling that $\sum_{A \in \mathcal{F}_i} \chi_A = \mathbb{1}$, this average is $\text{zero}^i(\frac{1}{n}\mathbb{1}, \frac{1}{n}\mathbb{1}) \in P^I$. This concludes the proof. \square

Putting Lemmas 45 and 46 together, we get that if $K = 3k$ for $k \in \mathbb{Z}_{++}$, $\frac{z^{N.S.}}{z^I} \geq \frac{Kn}{(n-1) \cdot 2k + K} = \frac{n}{(n-1) \cdot (2/3) + 1} = \frac{1}{2/3 + 1/3n}$. Since $\lim_{n \rightarrow \infty} \frac{1}{2/3 + 1/3n} = \frac{3}{2}$, for sufficiently large n we have $z^{N.S.} \geq z^I(\frac{3}{2} - \epsilon)$, proving this part of the theorem. The other cases of $K \pmod{3}$ are similar. This concludes the proof.

5.2 Proof for covering problem

5.2.1 Proof of Theorem 18

In order to prove Theorem 18, we begin with a classical bad example for the LP relaxation of the set cover problem.

Definition 47 (Special set covering problem (SSC)). *Consider $q \in \mathbb{Z}_+$. The ground set of the set cover problem will be $\{0, 1\}^q$, and the covering sets $S(v) = \{u \in \{0, 1\}^q \setminus \{0\} \mid v^T u = 1 \pmod{2}\}$ for $v \in \{0, 1\}^q$. Then the Special Set Covering (SSC(q)) problem is defined by:*

$$\begin{aligned} (SSC(q)) \quad & \min \sum_{v \in \{0, 1\}^q} x_v \\ & \text{s.t.} \quad \sum_{v: u \in S(v)} x_v \geq 1 \quad \forall u \in \{0, 1\}^q \\ & \quad \quad x_v \in \{0, 1\} \quad \forall v \in \{0, 1\}^q. \end{aligned}$$

We refer to the left-hand matrix of SSC as A^q .

Theorem 48 ([18]). *The IP optimal value of SSC(q) is at least q , while the LP relaxation optimal value is at most 2.*

We are now ready to present the proof of Theorem 18. For that, we consider the following problem:

$$\begin{aligned} (DSC(q)) \quad & \min \sum_{v \in \{0, 1\}^q} x_v + \sum_{v \in \{0, 1\}^q} y_v \\ & \text{s.t.} \quad A^q x + A^q y \geq \mathbb{1} \\ & \quad \quad x, y \in \{0, 1\}^{2^q} \end{aligned}$$

We first argue that $DSC(q)$ preserves the gap between IP and LP from $SSC(q)$.

Lemma 49. *The IP optimal value of $DSC(q)$ is at least q , while the LP relaxation optimal value is at most 2.*

Proof. ($z^{LP} \leq 2$): By Theorem 48, there exists a feasible solution \bar{x} of the LP relaxation of $SSC(q)$ such that $\sum_v \bar{x}_v \leq 2$. Then $(\bar{x}, 0)$ is a feasible solution of the LP relaxation of $DSC(q)$, giving the desired bound.

($z^{IP} \geq q$): Assume by contradiction that (\bar{x}, \bar{y}) is a feasible solution of $DSC(q)$ with objective function less than q . Note that if $\bar{x}_v = \bar{y}_v = 1$, then we may set $\bar{y}_v = 0$ and still obtain a feasible solution with a better objective function value; similarly, if $\bar{x}_v = 0$ and $\bar{y}_v = 1$ we may set $\bar{x}_v = 1$ and $\bar{y}_v = 0$ and obtain a feasible solution with same objective value. Therefore, we may assume that $\bar{y} = 0$. In this case, \bar{x} is a feasible solution of SSC with objective function less than q , contradicting the statement of Theorem 48. \square

We consider a partition on the columns of $DSC(q)$ into two blocks: $\mathcal{J} = \{J_1, J_2\}$ where J_1 corresponding to variables x and J_2 corresponding to variables y . To complete the proof of the theorem it is sufficient to prove that the super sparse closure optimal value $z^{S.S.}$ of $DSC(q)$ is equal to optimal LP value z^{LP} of $DSC(q)$.

Lemma 50. $z^{S.S.} = z^{LP}$ for $DSC(q)$.

Proof. Let P be the integer set for $DSC(q)$. Since $P^{S.S.} = P^{(J_1)} \cap P^{(J_2)}$, Observation 31 gives that $(\bar{x}, \bar{y}) \in P^{S.S.}$ iff $(\bar{x}, \bar{y}) \in P^{LP}$ and $\bar{x} \in P^I|_{J_1}$ and $\bar{y} \in P^I|_{J_2}$. But since P is of covering-type and $SSC(q)$ is feasible, we have that $P^I|_{J_i} = [0, 1]^{2^q}$, and thus $(\bar{x}, \bar{y}) \in P^{S.S.}$ iff $(\bar{x}, \bar{y}) \in P^{LP}$. This concludes the proof. \square

5.2.2 Proof of Theorem 22

Consider a covering problem (C). As in the packing case, there is an identification of sets of nodes of $G_{A, \mathcal{I}}^{\text{cover}}$ with sets of indices of variables (the ‘‘indices in the union of their support’’), namely if $\mathcal{I} = \{I_1, I_2, \dots, I_q\}$ is the given row index partition and the nodes of $G_{A, \mathcal{I}}^{\text{cover}}$ are $\{v_1, v_2, \dots, v_q\}$, then the set of vertices $\{v_i\}_{i \in I}$ corresponds to the indices $\bigcup_{i \in I} \bigcup_{r \in I_i} \text{supp}(A_r) \subseteq [n]$. We will make use of this correspondence, and in order to make statements precise we use the function $\text{usupp} : 2^{V(G_{A, \mathcal{I}}^{\text{cover}})} \rightarrow 2^{[n]}$ to denote this correspondence; with slight abuse of notation, for a singleton set $\{v\}$ we use $\text{usupp}(v)$ instead of $\text{usupp}(\{v\})$.

Given a set of vertices $S \subseteq V(G_{A, \mathcal{I}}^{\text{cover}})$, let $x^{(S)}$ be the optimal solution of the covering problem projected to the variables relative to S , namely $x^{(S)} \in \text{argmin}\{(c|_{\text{usupp}(S)})^T y \mid y \in P^I|_{\text{usupp}(S)}\}$. Also, let $\text{zero}^S(x^{(S)}) \in \mathbb{R}^n$ denote the solution appended by zeros in the original space, namely $\text{zero}^S(x^{(S)})_i = x_i^{(S)}$ if $i \in \text{usupp}(S)$ and $\text{zero}^S(x^{(S)})_i = 0$ if $i \notin \text{usupp}(S)$.

Notice the following important property of $\text{zero}^S(x^{(S)})$ (denote $S = \{v_i\}_{i \in I}$): for any row $r \in \bigcup_{i \in I} I_i$, since the support of A_r is contained in $\text{usupp}(S)$, the constraint $A_r x \geq b_r$ is valid for $P^I|_{\text{usupp}(S)}$; therefore $A_r \text{zero}^S(x^{(S)}) = A_r|_{\text{usupp}(S)} x^{(S)} \geq b_r$. This gives the following.

Observation 51. For any subset $S = \{v_i\}_{i \in I}$ of nodes of $G_{A, \mathcal{I}}^{\text{cover}}$ and any row $r \in \bigcup_{i \in I} I_i$,

$$A_r \text{zero}^S(x^{(S)}) \geq b_r.$$

We start by showing that the solutions $x^{(M)}$, for M in a mixed stable set \mathcal{M} , can be used to provide a lower bound on the optimal value of $P^{\mathcal{V}, C}$.

Lemma 52. Let \mathcal{M} be a mixed stable set for $G_{A, \mathcal{I}}^{\text{cover}}$ subordinate to \mathcal{V} . Then

$$z^{\mathcal{V}, C} \geq \sum_{M \in \mathcal{M}} (c|_{\text{usupp}(M)})^T x^{(M)}.$$

Proof. Consider an optimal solution $x^* \in \text{argmin}\{c^T x \mid x \in P^{\mathcal{V}, C}\}$ of the row block-sparse closure. Since $P^{\mathcal{V}, C} = \bigcap_{S \in \mathcal{V}} P^{(S)}$, Observation 31 implies that $x^*|_{\text{usupp}(S)} \in P^I|_{\text{usupp}(S)}$ for all $S \in \mathcal{V}$. Moreover, since for every set M in the mixed stable set \mathcal{M} there is $S \in \mathcal{V}$ containing M , this implies that $x^*|_{\text{usupp}(M)} \in P^I|_{\text{usupp}(M)}$ for all $M \in \mathcal{M}$. Then by the optimality of $x^{(M)}$, we get $(c|_{\text{usupp}(M)})^T (x^*|_{\text{usupp}(M)}) \geq (c|_{\text{usupp}(M)})^T x^{(M)}$ for all $M \in \mathcal{M}$.

Then we can decompose the optimal solution x^* based on the variables $\text{usupp}(M)$ and use the non-negativity of c :

$$z^{\mathcal{V},C} = c^T x^* = \sum_{M \in \mathcal{M}} (c|_{\text{usupp}(M)})^T (x^*|_{\text{usupp}(M)}) + \sum_{i \notin \bigcup_{M \in \mathcal{M}} \text{usupp}(M)} c_i x_i^* \geq \sum_{M \in \mathcal{M}} (c|_{\text{usupp}(M)})^T x^{(M)},$$

where the first equality uses the fact that if $M_1, M_2 \in \mathcal{M}$, then $\text{usupp}(M_1) \cap \text{usupp}(M_2) = \emptyset$. This concludes the proof. \square

Now show how to put solutions $x^{(M)}$ together to get a feasible solution for the covering problem, thus providing an upper bound on z^I . Recall the definition of mixed chromatic number $\bar{\eta} = \bar{\eta}^{\mathcal{V}}(G_{A,\mathcal{I}}^{\text{cover}})$ and consider covering mixed stable sets $\mathcal{M}_1, \dots, \mathcal{M}_{\bar{\eta}}$ (i.e., $V(G_{A,\mathcal{I}}^{\text{cover}}) = \bigcup_i \bigcup_{M \in \mathcal{M}_i} M$).

Define $u \in \mathbb{R}^n$ as the pointwise maximum of the solutions $\{\text{zero}^M(x^{(M)})\}_{i,M \in \mathcal{M}_i}$. Since the matrix A in the problem is non-negative, Observation 51 implies that u is a feasible solution for the covering problem (C). Thus, using the non-negativity of c and of the $\text{zero}^M(x^{(M)})$'s:

$$z^I \leq c^T u \leq \sum_{i,M \in \mathcal{M}_i} c^T \text{zero}^M(x^{(M)}) = \sum_{i,M \in \mathcal{M}_i} (c|_{\text{usupp}(M)})^T x^{(M)} \leq \sum_i z^{\mathcal{V},C} = \bar{\eta} \cdot z^{\mathcal{V},C},$$

where the first inequality follows from definition of z^I and feasibility of u , the second inequality follows from non-negativity of c , and the last inequality follows from Lemma 52. This concludes the proof of Theorem 22.

5.2.3 Proof of Theorem 24

Now we prove Theorem 24 by constructing a covering instance. Since the construction is quite involved, we start with an example.

Example of the construction. We exemplify the construction for $K = 2$ and with a worse gap, and then we generalize/strengthen it (the discussion here will be somewhat informal). In this case the covering IP is the following (notice the indices of the x variables in the different constraints):

$$\min \sum_i x_i + \infty \cdot (y_1 + y_2) \tag{7}$$

$$\text{s.t.} \quad \left[\begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} x_1 + \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} x_2 \right] + \left[\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} x_3 + \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} x_4 \right] + \mathbb{1} \cdot y_1 \geq \mathbb{1} \tag{8}$$

$$\left[\begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} x_1 + \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} x_3 \right] + \left[\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} x_2 + \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} x_4 \right] + \mathbb{1} \cdot y_2 \geq \mathbb{1} \tag{9}$$

$$x \in \mathbb{Z}_+^4, \quad y \in \mathbb{Z}_+^2. \tag{10}$$

We will use the partition of rows $\mathcal{I} = \{I_1, I_2\}$, where $I_1 = \{1, 2, 3, 4\}$ (so corresponds to the first sets of covering constraints) and $I_2 = \{5, 6, 7, 8\}$.

The only (minimal) ways to satisfy the first set of constraints is to set either $x_1 = x_2 = 1$ (and all else to 0), or $x_3 = x_4 = 1$ (and all else to 0), or $y_1 = 1$ (and all else to 0); because of the cost of the y variables, actually we will always have $y = 0$ in an optimal solution. To satisfy the second set of constraints the situation is similar, but the indices on the x variables are permuted so that we need $x_1 = x_3 = 1$ or $x_2 = x_4 = 1$. So the best way to satisfy both of the constraints **simultaneously** is to set almost all x variables to 1 (actually we can just set $x_1 = x_2 = x_3 = 1$). This gives cost of 3 for the IP.

Now consider optimizing over the weak specific-scenario cuts closure $P^{\mathcal{V},C}$ (where the row support list is $\mathcal{V} = \{\{v_1\}, \{v_2\}\}$), i.e., the closure corresponding to the cuts on (x, y_1) variables and on (x, y_2) variables. Since the y_i variable can be used to satisfy the i th set of covering constraints, it is easy to see that the only undominated (x, y_1) -cuts are the ones implied only the first set of covering constraints (8), and similarly the only undominated (x, y_2) -cuts are the ones implied only by the second set of covering constraints (9). Thus, the point $x_1 = x_2 = x_3 = x_4 = \frac{1}{2}$, $y_1 = y_2 = 0$ belongs to $P^{\mathcal{V},C}$, giving $z^* = z^{\mathcal{V},C} \leq 2$.

Together, these observations give that $\frac{z^I}{z^{\mathcal{V},C}} \geq 3/2$.

General construction. We start with the special set system that is used to define the columns of the covering program.

Lemma 53. *Let $n \in \mathbb{Z}_{++}$. There is a collection $\mathcal{G}^1, \mathcal{G}^2, \dots, \mathcal{G}^n$ with the following properties:*

1. *For each $i \in [n]$, \mathcal{G}^i is a partition of $[n^n]$ and each set $G \in \mathcal{G}^i$ has size n^{n-1} .*
2. *For any selection $G^1 \in \mathcal{G}^1, G^2 \in \mathcal{G}^2, \dots, G^n \in \mathcal{G}^n$, the intersection $\bigcap_{i=1}^n G^i$ is non-empty.*

Proof. Since $|[n^n]| = |[n]^n|$, let $g : [n]^n \rightarrow [n^n]$ be any bijection between the two sets. For $j \in [n]$, define the set $G_j^i = \{g(u) \mid u \in [n]^n, u_i = j\}$. Define $\mathcal{G}^i = \{G_j^i \mid j \in [n]\}$. It is easy to check the following properties:

1. Given i , for any $j \in [n]$, $|G_j^i| = n^{n-1}$ and $\bigcup_{j=1}^n G_j^i = \{g(u) \mid u \in [n]^n\} = [n^n]$.
2. For a selection $G_{j_1}^1 \in \mathcal{G}^1, G_{j_2}^2 \in \mathcal{G}^2, \dots, G_{j_n}^n \in \mathcal{G}^n$, consider $u = (j_1, j_2, \dots, j_n)$. Then according to the definition, $g(u) \in \bigcap_{i=1}^n G_{j_i}^i$, so the intersection of these sets is non-empty.

This concludes the proof. \square

Lemma 54. *Let $n \in \mathbb{Z}_{++}$. Consider a collection $\mathcal{G}^1, \dots, \mathcal{G}^n$ satisfying the properties of Lemma 53, and consider $\bar{\mathcal{G}}^i \subseteq \mathcal{G}^i$ for $i = 1, \dots, n$. If the sets in $\bigcup_{i=1}^n \bar{\mathcal{G}}^i$ cover the whole of $[n^n]$, then there is $i \in [n]$ such that $\bar{\mathcal{G}}^i = \mathcal{G}^i$.*

Proof. By contradiction, suppose there is $G_{j_1}^1 \in \mathcal{G}^1 \setminus \bar{\mathcal{G}}^1, \dots, G_{j_n}^n \in \mathcal{G}^n \setminus \bar{\mathcal{G}}^n$. Then part 2 of Lemma 53, there exists an element $u \in \bigcup_{i=1}^n G_{j_i}^i$, and since the sets in \mathcal{G}^i partition $[n^n]$ this means that u is not covered by sets in $\bar{\mathcal{G}}^i$, for all i ; then $\bigcup_{i=1}^n \bar{\mathcal{G}}^i$ does not cover $[n^n]$, a contradiction. \square

Now pick a prime number $n \geq \max\{K, 2\}$. We will construct an instance with variables x_1, \dots, x_{n^2} and y_1, \dots, y_K , and each row-block will have n^n constraints. Let $\mathcal{G}^1, \dots, \mathcal{G}^n$ be a collection satisfying the properties from Lemma 53, and let $\mathcal{F}_1, \dots, \mathcal{F}_n$ be an affine n -design

(we remind the readers – each \mathcal{F}_i partitions $[n^2]$ with n -subsets, and for $A \in \mathcal{F}_i, B \in \mathcal{F}_j$ with $i \neq j$, $|A \cap B| \leq 1$); this affine design will be used to “permute” the indices of the x variables from one set of covering constraints to the next (see example above). We consider the explicit enumeration $\mathcal{F}_i = \{F_i^1, \dots, F_i^n\}$.

For $k \in [K]$, we define the set

$$P_k = \left\{ (x, y) \in \{0, 1\}^{n^2+K} \left| \sum_{i=1}^n \sum_{j \in F_i^k} A_j^k x_j + \mathbb{1} \cdot y_k \geq \mathbb{1} \right. \right\},$$

where the set of vectors $\{A_j^k\}_{j \in F_i^k}$ is equal to the set of vectors $\{\chi_{G_{j'}^i}\}_{j' \in [n]}$; we note that it is not important which A_j^k is assigned to which $\chi_{G_{j'}^i}$.

Then the covering integer program we consider is the following:

$$\begin{aligned} \min \quad & \sum_{j=1}^{n^2} x_j + n^n \cdot \sum_{k=1}^K y_k \\ \text{s.t.} \quad & \sum_{i=1}^n \sum_{j \in F_i^k} A_j^k x_j + \mathbb{1} \cdot y_k \geq \mathbb{1} \quad \forall k \in [K] \\ & (x, y) \in \{0, 1\}^{n^2+K}. \end{aligned}$$

(or equivalently $(x, y) \in \bigcap_{k \in [K]} P_k$). Let P denote the set of solutions for this problem.

Notice that this program has K sets of n^n covering constraints. We consider the partition of the covering constraints $\mathcal{I} = \{I_1, \dots, I_K\}$ where $I_k = \{(k-1)n^n+1, \dots, kn^n\}$ (so I_k corresponds to P_k). It is easy to see that the covering interaction graph $G_{A, \mathcal{I}}^{\text{cover}}$ is a clique.

For $\mathcal{V} = \{\{v_1\}, \dots, \{v_K\}\}$, remember that $z^* = z^{\mathcal{V}, C}$. Therefore, we want to show that $\frac{z^I}{z^{\mathcal{V}, C}} \geq K - \epsilon$ (for sufficiently large n). We start by analyzing each P_k .

Lemma 55. *A vector $(\bar{x}, 0) \in \{0, 1\}^{n^2} \times \{0, 1\}^K$ belongs to P_k if and only if there is F_i^k such that $\bar{x}_j = 1$ for all $j \in F_i^k$.*

Proof. (\Rightarrow) Let $\bar{\mathcal{G}}^i$ be the subset of the sets in \mathcal{G}^i picked by \bar{x} , namely G_t^i belongs to $\bar{\mathcal{G}}^i$ iff $\chi_{G_{j'}^i} = A_j^k$ for some j with $\bar{x}_j = 1$. Since $\bar{y} = 0$, the fact that (\bar{x}, \bar{y}) belongs to P_k implies that the sets in $\bigcup_{i=1}^n \bar{\mathcal{G}}^i$ must cover the whole of $[n^n]$. Lemma 54 then implies that there is one $\bar{\mathcal{G}}^i$ that equals \mathcal{G}^i , which translates to having $\bar{x}_j = 1$ for all $j \in F_i^k$.

(\Leftarrow) This follows from the fact that the sets in \mathcal{G}^i cover the whole of $[n^n]$. \square

We can use this to lower bound the optimal value z^I of the covering program P .

Lemma 56. $z^I \geq Kn - K^2$.

Proof. First, we claim that if $(\bar{x}, 0) \in P$, then $\sum_{j \in [n^2]} \bar{x}_j \geq Kn - K^2$. Let $S \subseteq [n^2]$ be the support of \bar{x} , so it is equivalent to show $|S| \geq nK - K^2$. Since $(\bar{x}, 0) \in P = \bigcap_{k=1}^K P_k$, using Lemma 55 we have that for every $k \in [K]$ there is $i(k)$ such that S contains $F_{i(k)}^k$, so $S \supseteq \bigcup_{k=1}^K F_{i(k)}^k$. By the inclusion-exclusion principle, we have that $|S| \geq \left| \bigcup_{k=1}^K F_{i(k)}^k \right| \geq$

$\sum_{k=1}^K |F_{i(k)}^k| - \sum_{k \neq k'} |F_{i(k)}^k \cap F_{i(k')}^{k'}|$. Using the definition of an affine n -design, get the lower bound $|S| \geq nK - K(K-1) \geq nK - K^2$.

Now consider any solution $(\bar{x}, \bar{y}) \in P$. If $\bar{y} = 0$, we have just shown that this solution has value at least $nK - K^2$; if $\bar{y} \neq 0$, this solution has value at least $n^n > nK - K^2$. This concludes the proof. \square

Finally we upper bound the optimal value of the $z^{\mathcal{V}, C}$.

Lemma 57. $z^{\mathcal{V}, C} \leq n$.

Proof. It suffices to show that the point $(\bar{x}, \bar{y}) = (\frac{1}{n}\mathbb{1}, 0) \in P^{\mathcal{V}, C}$. Recall that $P^{\mathcal{V}, C} = \bigcap_{k \in [K]} P^{\{v_k\}}$, so we show (\bar{x}, \bar{y}) belongs to all $P^{\{v_k\}}$'s. Note that (\bar{x}, \bar{y}) satisfies the linear programming relaxation; therefore, using Observation 31, to show that (\bar{x}, \bar{y}) belongs to $P^{\{v_k\}}$ it suffices to prove that $(\bar{x}, \bar{y}_k) \in P^I|_{(x, y_k)}$, where we use $P^I|_{(x, y_k)}$ to denote the projection onto the variables (x, y_k) .

Consider the following points (x^u, y) , for $u \in [n]$, constructed as:

$$\begin{aligned} y_k &= 0, \\ y_{k'} &= 1 \quad \forall k' \in [K] \setminus \{k\}, \\ x_j^u &= \begin{cases} 1 & \text{if } j \in F_u^k \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

It is straightforward to verify that $(x^u, y) \in P$ for $u \in [n]$. Thus, the average $\frac{1}{n} \sum_{u \in [n]} (x^u, y)$ belongs to P^I . It then follows that $(\bar{x}, \bar{y}_k) = (\frac{1}{n}\mathbb{1}, 0)|_{(x, y_k)}$ belongs to $P^I|_{(x, y_k)}$. This concludes the proof. \square

Putting Lemmas 56 and 57 together we get $\frac{z^I}{z^{\mathcal{S}, S}} \geq \frac{Kn - K^2}{n} = K - \frac{K^2}{n}$. For large enough n , we get $\frac{z^I}{z^{\mathcal{V}, C}} \geq K - \epsilon$. This concludes the proof of Theorem 24.

5.3 Proof for packing-type problem with arbitrary A matrix

5.3.1 Proof of Theorem 26

In this section, we use the same notation as that used in Section 5.1.1. So, let \mathcal{J} be a partition of the index set of columns of A (that is $[n]$). Let $V = \{v_1, \dots, v_q\}$ be the vertices of $G_{A, \mathcal{J}}^{\text{pack}}$ (based on Definition 4). Let $\tilde{\mathcal{V}} = \{V^{u_1}, V^{u_2}, \dots, V^{u_k}\} \subseteq \mathcal{V}$ be the subset of sparse cut support list corresponding to the definition of corrected average density $D_{\mathcal{V}}$ (see Definition 25), so $V = \bigcup_{i=1}^k V^{u_i}$ and $\frac{1}{k} \sum_{i=1}^k |V^{u_i}| = D_{\mathcal{V}}$.

Recall from Section 5.1.1 a couple of definitions: first, the function ϕ maps subsets of vertices of V to the corresponding variable indices, namely if $S = \{v_i\}_{i \in I}$ then $\phi(S) = \bigcup_{i \in I} J_i$.

For the purpose of this section, let $x^{(S)} = \operatorname{argmax} \{(c|_{\phi(S)})^T(x|_{\phi(S)}) \mid x \in P^I\}$. (Notice that this is different from the definition used in Section 5.1.1.)

Lemma 58. *For any set $\tilde{V} \in \tilde{\mathcal{V}}$ we have $z^{\mathcal{V}, P} \leq c^T x^{(\tilde{V})} + \sum_{v \in V \setminus \tilde{V}} c^T x^{(v)}$.*

Proof. Fix any $\tilde{V} \in \tilde{\mathcal{V}}$ and $S \subseteq \tilde{V}$. Let $x^* = \operatorname{argmax}\{c^T x \mid x \in P^{\mathcal{V}, P}\}$ be an optimal solution corresponding to the optimization over $P^{\mathcal{V}, P}$. Since $P^{\mathcal{V}, C} = \bigcap_{\tilde{V}' \in \tilde{\mathcal{V}}} P^{(\tilde{V}')}$, we have that $x^* \in P^{(S)} \supseteq P^{(\tilde{V}')}$. From Observation 31 we then get $x^*|_{\phi(S)} \in P^I|_{\phi(S)}$.

Thus we get $(c|_{\phi(S)})^T(x^*|_{\phi(S)}) \leq (c|_{\phi(S)})^T(x^{(S)}|_{\phi(S)}) \leq c^T x^{(S)}$, where the first inequality follows from optimality of $x^{(S)}$ and the second inequality follows from non-negativity of c and $x^{(S)}$.

In particular, since $\tilde{\mathcal{V}}$ covers V , we can apply this to the any singleton $S = \{v\}$ and get $(c|_{\phi(v)})^T(x^*|_{\phi(v)}) \leq c^T x^{(v)}$.

Applying this bound, we obtain that for any $\tilde{V} \in \mathcal{V}$

$$z^{\mathcal{V},P} = c^T x^* = (c|_{\phi(\tilde{V})})^T(x^*|_{\phi(\tilde{V})}) + \sum_{v \in V \setminus \tilde{V}} (c|_{\phi(v)})^T(x^*|_{\phi(v)}) \leq c^T x^{(\tilde{V})} + \sum_{v \in V \setminus \tilde{V}} c^T x^{(v)}.$$

This concludes the proof. \square

Now we are ready to complete the proof of the theorem. Using Lemma 58 for all sets in $\tilde{\mathcal{V}}$ and adding up these inequalities we obtain that

$$\begin{aligned} k \cdot z^{\mathcal{V},P} &\leq \sum_{i=1}^k c^T x^{(V^{u_i})} + \sum_{i=1}^k \left(\sum_{v \in V \setminus V^{u_i}} c^T x^{(v)} \right) \\ &= \sum_{i=1}^k c^T x^{(V^{u_i})} + \sum_{v \in V} \text{miss}(v) \cdot c^T x^{(v)}, \end{aligned} \quad (11)$$

where $\text{miss}(v) = |\{i \in [k] \mid v \notin V^{u_i}\}|$, that is the number of sparse-cut types in \mathcal{V} in which the variables corresponding to vertex v do not appear.

Moreover it follows from the definition of $x^{(S)}$ that $x^S \in P^I$ and therefore we have that $z^I \geq c^T x^{(S)}$ for every subset $S \subseteq V$. Thus, we obtain that

$$\begin{aligned} z^I &\geq \max \left\{ \max_{i \in [k]} \{c^T x^{(V^{u_i})}\}, \max_{v \in V} \{c^T x^{(v)}\} \right\} \\ &\geq \frac{1}{k + \sum_{v \in V} \text{miss}(v)} \left(\sum_{i=1}^k c^T x^{(V^{u_i})} + \sum_{v \in V} \text{miss}(v) \cdot c^T x^{(v)} \right) \\ &\geq \frac{k}{k + \sum_{v \in V} \text{miss}(v)} \cdot z^{\mathcal{V},P}. \end{aligned}$$

where the second inequality follows from taking a weighted average, the third inequality follows from (11). Finally, the next lemma shows that $k + \sum_v \text{miss}(v) = k + kq - kD_{\mathcal{V}}$, concluding the proof of the theorem.

Lemma 59. $kD_{\mathcal{V}} + \sum_{v \in V} \text{miss}(v) = kq$.

Proof. We perform a simple double counting. Consider the $V/\{V^i\}_i$ incidence matrix $B \in \{0, 1\}^{k \times q}$ defined as $B_{i,v} = 1$ if $v \in V^i$ and $B_{i,v} = 0$ if $v \notin V^i$. Using the definition of $D_{\mathcal{V}}$ we have:

$$kD_{\mathcal{V}} = |\{(i, v) \in [k] \times V \mid B_{i,v} \neq 0\}|. \quad (12)$$

On the other hand, from the definition of $\text{miss}(v)$ we have that

$$\sum_{v \in V} \text{miss}(v) = \sum_{v \in V} |\{i \in [k] \mid v \notin V^i\}| = |\{(i, v) \in [k] \times V \mid B_{i,v} = 0\}|. \quad (13)$$

By (12) and (13), we have that $kD_{\mathcal{V}} + \sum_{v \in V} \text{miss}(v) = kq$. This concludes the proof. \square

5.3.2 Proof of Theorem 28

We consider the following integer program with $2K - 1$ variables:

$$\begin{aligned} \max \quad & x_K + \sum_{j=K+1}^{2K-1} x_j \\ \text{s.t.} \quad & \sum_{i=1}^K x_i = 1 \end{aligned} \tag{14}$$

$$x_i + x_j \leq 2 - \epsilon \quad \forall i \in \{1, \dots, K-1\}, \quad \forall j \in \{K+1, \dots, 2K-1\} \setminus \{K+i\} \tag{15}$$

$$x_K + x_j \leq 2 - \epsilon \quad \forall j \in \{K+1, \dots, 2K-1\} \tag{16}$$

$$x \in \{0, 1\}^{2K-1}.$$

(We assume $\epsilon < \frac{K-1}{K}$.) Let P denote the integer set relative to this problem.

We consider the partition $\mathcal{J} = \{J_1, \dots, J_K\}$ of the columns given by $J_1 = \{1, \dots, K\}$, $J_i = \{K+i-1\}$ for $i \in 2, \dots, K$. Notice that the packing interaction graph $G_{A, \mathcal{J}}^{\text{pack}}$ for this program is a star on K nodes. Writing explicitly $G_{A, \mathcal{J}}^{\text{pack}} = (V, E)$ with $V = \{v_1, \dots, v_K\}$ and

$$E = \{(v_1, v_2), (v_1, v_3), (v_1, v_4), \dots, (v_1, v_K)\}.$$

Since we are in the context of the super sparse closure, we have the support list

$$\mathcal{V} = \{\{v_1\}, \{v_2\}, \dots, \{v_K\}\}.$$

We show the bound $z^{S.S.} \geq K \cdot z^I - \epsilon$, and start by lower bounding $z^{S.S.}$.

Lemma 60. $z^{S.S.} \geq K - \epsilon$

Proof. We claim that the point \bar{x} given by $\bar{x}_j = \frac{\epsilon}{K-1}$ for all $j = \{1, \dots, K-1\}$, $\bar{x}_K = 1 - \epsilon$, $\bar{x}_j = 1$ for all $j \in \{K+1, \dots, 2K-1\}$ belongs to the natural sparse closure $P^{S.S.}$, proving it using Observation 31.

First, it is easy to check that \bar{x} belongs to the LP relaxation P^{LP} . Moreover, note that $P^I|_{J_1} = \{(x_1, \dots, x_K) \in [0, 1]^K \mid \sum_{i=1}^K x_i = 1\}$, and hence $\bar{x}|_{J_1} \in P^I|_{J_1}$. In addition, $P^I|_{J_j} = [0, 1]$ for $j \in \{2, \dots, |V|\}$, and thus $\bar{x}|_{J_j} = \bar{x}_{K+j-1} \in P^I|_{J_j}$. Since $P^{S.S.} = \bigcap_{j=1}^K P^{(J_j)}$, from Observation 31 we get that $\bar{x} \in P^{S.S.}$. \square

To complete the proof, we show that the optimal value of the IP is (at most) 1, namely exactly one of the variables $x_K, x_{K+1}, \dots, x_{2K-1}$ can take a value of 1 and the others are zero. So consider any feasible solution $\bar{x} \in \{0, 1\}^{2K-1}$. If $\bar{x}_K = 1$, then the constraints (16) imply that $x_{K+1} = x_{K+2} = \dots = x_{2K-1} = 0$. On the other hand if $\bar{x}_K = 0$, then by constraint (14) there is some $i \in [K-1]$ with $\bar{x}_i = 1$, and so constraints (15) imply $\bar{x}_j = 0$ for all $j \in \{|V|+1, \dots, 2K+1\} \setminus \{K+i\}$, and so at most \bar{x}_j can take value 0.

Since $z^I \leq 1$ and $z^{S.S.} \geq K - \epsilon$, we get the desired bound $z^{S.S.} \geq K \cdot z^I - \epsilon$, concluding the proof of Theorem 28.

5.3.3 Proof of Theorem 30

We will construct an example with $2K$ variables. To start, for $k \in [K]$ let P_k^I be the convex hull of the points

$$P_k := \{(x, y) \in \{0, 1\}^{K+K} \mid y_k = 1 \text{ if and only if either } [x_k = 1, x_i = 0 \ \forall i \neq k] \text{ or } [x_k = 0, x_i = 1 \ \forall i \neq k]\}$$

We then consider the integer program

$$\begin{aligned} \max \quad & \sum_{k=1}^K y_k \\ \text{s.t.} \quad & (x, y) \in \bigcap_{k \in [K]} P_k^I \cap \{0, 1\}^{2K}. \end{aligned}$$

Let P denote the associated integer set. The partition of variable indices we consider is $\mathcal{J} = \{J_0, J_1, J_2, \dots, J_K\}$, where J_0 corresponds to the variables x , and each J_k corresponds to variable y_k for $k \in [K]$. Notice that the packing interaction graph $G_{A, \mathcal{J}}^{\text{pack}}$ is a star on $K + 1$ nodes; explicitly, $G_{A, \mathcal{J}}^{\text{pack}} = (V, E)$ with $V = \{v_0, \dots, v_K\}$ and $E = \{\{v_0, v_1\}, \dots, \{v_0, v_K\}\}$. Recall we are in the natural sparse closure setting, so the support list \mathcal{V} in this case equals the edge set E .

We show that $z^{N.S.} \geq K \cdot z^I$. For that, we start by lower bounding $z^{N.S.}$.

Lemma 61. $z^{N.S.} \geq K$.

Proof. We show that the solution (\bar{x}, \bar{y}) given by $\bar{x} = \frac{1}{2}\mathbb{1}$ and $\bar{y} = \mathbb{1}$ belongs to $P^{N.S.}$. Following Observation 31, to show $(\bar{x}, \bar{y}) \in P^{N.S.}$ it suffices to show $(\bar{x}, \bar{y}) \in P^{LP}$ and $(\bar{x}, \bar{y}_k) \in P^I|_{(x, y_k)}$ for all $k \in [K]$. Notice that $P^{LP} = \bigcap_{k \in [K]} P_k^I$ and $P^I|_{(x, y_k)} = P_k^I|_{(x, y_k)}$ (the latter uses the fact $P_j^I|_{(x, y_k)} = [0, 1]^{K+1}$ for $j \neq k$). Thus it suffices to show $(\bar{x}, \bar{y}) \in P_k^I$ for all $k \in [K]$

For that, fix $k \in [K]$ and consider the points (x^{k1}, e^k) and (x^{k2}, e^k) , where e^i is the i th canonical basis vector and

$$\begin{aligned} x_i^{k1} &= \begin{cases} 1 & \text{if } i = k, \\ 0 & \text{otherwise} \end{cases} \\ x_i^{k2} &= \begin{cases} 0 & \text{if } i = k, \\ 1 & \text{otherwise} \end{cases} \end{aligned}$$

By definition both these points belong to P_k ; the average $\frac{1}{2}(x^{k1}, e^k) + \frac{1}{2}(x^{k2}, e^k) = (\frac{1}{2}\mathbb{1}, e^k)$ also belongs to P_k . Moreover, since the constraints defining P_k^I are independent of variable y_i for $i \neq k$, we have that $(\frac{1}{2}\mathbb{1}, \mathbb{1}) = (\bar{x}, \bar{y})$ also belongs to P_k^I . This concludes the proof. \square

Now it is easy to see from the definition of P_k that no feasible solution to the IP can set more than one y variable to 1, and hence the optimal value z^I is at most 1. Together with the previous lemma, this gives the desired bound $z^{N.S.} \geq K \cdot z^I$, thus concluding the proof of the theorem.

6 Acknowledgements

We would like to thank Carla Michini for her comments that helped improve the presentation in this paper. Santanu S. Dey and Qianyi Wang acknowledge the support from NSF CMMI Grant 1149400.

References

- [1] Acer, S., Kayaaslan, E., Aykanat, C.: A recursive bipartitioning algorithm for permuting sparse square matrices into block diagonal form with overlap. *SIAM J. Scientific Computing* **35**(1) (2013)
- [2] Bergner, M., Caprara, A., Furini, F., Lübbecke, M.E., Malaguti, E., Traversi, E.: Partial convexification of general mips by dantzig-wolfe reformulation. In: O. Günlük, G.J. Woeginger (eds.) *Integer Programming and Combinatorial Optimization - 15th International Conference, IPCO 2011, New York, NY, USA, June 15-17, 2011. Proceedings, Lecture Notes in Computer Science*, vol. 6655, pp. 39–51. Springer (2011)
- [3] Bixby, R.E., Fenelon, M., Gu, Z., Rothberg, E., R. Wunderling, R.: *Mixed-integer programming: a progress report*. chap. 18, pp. 309–326. SIAM, Philadelphia, PA (2004)
- [4] Borndörfer, R., Ferreira, C.E., Martin, A.: Decomposing matrices into blocks. *SIAM Journal on Optimization* **9**(1), 236–269 (1998)
- [5] Brooks, R.L.: On colouring the nodes of a network. *Proceedings - Cambridge Philosophical Society. Mathematical and physical sciences* **37**, 194–197 (1941)
- [6] Carøe, C.C., Schultz, R.: Dual decomposition in stochastic integer programming. *Oper. Res. Lett.* **24**(1-2), 37–45 (1999)
- [7] Colbourn, C.J., Dinitz, J.H.: *Handbook of Combinatorial Designs, Second Edition*. CRC press, Taylor and Francis Group, Boca Raton, FL (2006)
- [8] Crowder, H., Johnson, E.L., Padberg, M.W.: Solving large scale zero-one linear programming problem. *Operations Research* **31**, 803–834 (1983)
- [9] Dey, S.S., Lodi, A., Tramontani, A., Wolsey, L.A.: On the practical strength of two-row tableau cuts. *INFORMS Journal on Computing* **26**(2), 222–237 (2014)
- [10] Dey, S.S., Molinaro, M., Wang, Q.: Approximating polyhedra with sparse inequalities. *Mathematical Programming* pp. 1–24 (2015). DOI 10.1007/s10107-015-0925-y
- [11] Escoffier, B., Gourvès, L., Monnot, J., Spanjaard, O.: Two-stage stochastic matching and spanning tree problems: Polynomial instances and approximation. *European Journal of Operational Research* **205**(1), 19–30 (2010)
- [12] Fulkerson, D.R., Gross, O.: Incidence matrices and interval graphs. *Pacific Journal of Mathematics* **15**, 835–855 (1965)
- [13] Guan, Y., Ahmed, S., Nemhauser, G.L.: Cutting planes for multistage stochastic integer programs. *Operations Research* **57**(2), 287–298 (2009)

- [14] Jünger, M., Liebling, T.M., Naddef, D., Nemhauser, G.L., Pulleyblank, W.R., Reinelt, G., Rinaldi, G., Wolsey, L.A. (eds.): 50 Years of Integer Programming 1958-2008 - From the Early Years to the State-of-the-Art. Springer (2010)
- [15] King, A.D., Lu, L., Peng, X.: A fractional analogue of brooks' theorem. *SIAM J. Discrete Math.* **26**(2), 452–471 (2012)
- [16] Kong, N., Schaefer, A.J.: A factor 1/2 approximation algorithm for two-stage stochastic matching problems. *European Journal of Operational Research* **172**(3), 740–746 (2006)
- [17] Lodi, A.: Mixed integer programming computation. In: Jünger et al. [14], pp. 619–645
- [18] Lovász, L.: On the ratio of the optimal integral and fractional covers. *Discrete Mathematics* **13**, 383–390 (1975)
- [19] Marchand, H., Martin, A., Weismantel, R., Wolsey, L.A.: Cutting planes in integer and mixed integer programming. *Discrete Applied Mathematics* **123**, 397–446 (2002)
- [20] Molloy, M., Reed, B.: *Graph Colouring and the Probabilistic Method*. Springer-Verlag Berlin Heidelberg (2002)
- [21] Richard, J.P.P., Dey, S.S.: The group-theoretic approach in mixed integer programming. In: Jünger et al. [14], chap. 19, pp. 727–801
- [22] Walter, M.: Sparsity of lift-and-project cutting planes. In: S. Helber, M. Breitner, D. Rsch, C. Schn, J.M. Graf von der Schulenburg, P. Sibbertsen, M. Steinbach, S. Weber, A. Wolter (eds.) *Operations Research Proceedings 2012, Operations Research Proceedings*, pp. 9–14. Springer International Publishing (2012)
- [23] Wang, J., Ralphs, T.K.: Computational experience with hypergraph-based methods for automatic decomposition in discrete optimization. In: C.P. Gomes, M. Sellmann (eds.) *Integration of AI and OR Techniques in Constraint Programming for Combinatorial Optimization Problems, 10th International Conference, CPAIOR 2013, Yorktown Heights, NY, USA, May 18-22, 2013. Proceedings, Lecture Notes in Computer Science*, vol. 7874, pp. 394–402. Springer (2013)
- [24] Zhang, M., Küçükyavuz, S.: Finitely convergent decomposition algorithms for two-stage stochastic pure integer programs. *SIAM Journal on Optimization* **24**(4), 1933–1951

A Upper bound on z^{cut}

Assume we have the general formulation

$$\begin{aligned} \max \quad & c^T x \\ \text{s.t.} \quad & Ax \leq b \\ & x \in \mathbb{B}^n, \end{aligned}$$

where $A \in \mathbb{R}^{m \times n}$. Recall that we are interested in three type of problems: packing, covering and packing with arbitrary constraint matrix. All these categories will be written in the form of

the formulation above with different restrictions on A and c . Let $N_i = \{j \in [n] \mid A_{ij} \neq 0\}$ be the index set of non-zero entries of i^{th} row of A . Let $\mathcal{N} = \{N_1, N_2, \dots, N_t\}$, denote $P^{\mathcal{N}} = \bigcap_{i=1}^t P^{(N_i)}$ and $z^{\text{cut}} = \max_{x \in P^{\mathcal{N}}} c^T x$.

Our basic strategy is the following: we keep adding cuts on the support of some N_i and checking whether the LP solution will improve. We stop adding cuts when the objective function value does not change, thus obtaining an upper bound on z^{cut} . The formal algorithm is shown as Algorithm 1.

Algorithm 1 Estimating z^{cut}

input: $P = \{x \mid Ax \leq b\}, \mathcal{N} = \{N_1, N_2, \dots, N_t\}, z^{\text{old}} = -\infty, z^{\text{new}} = -\infty, \epsilon = 10^{-6}$

$i \leftarrow 1, \text{count} \leftarrow 0$

loop

Solve $x^* = \operatorname{argmax}_{x \in P} c^T x, z^{\text{new}} = c^T x^*$

if x^* is integral **then**

$z^{\text{cut}} = z^{\text{new}}$

Exit Loop.

else if $z^{\text{new}} - z^{\text{old}} > \epsilon$ **then**

$z^{\text{old}} \leftarrow z^{\text{new}}$

Generate a valid cut $\alpha x \leq \beta$ on the support of N_i based on Algorithm 2

$P \leftarrow P \cap \{x \mid \alpha x \leq \beta\}$

$\text{count} \leftarrow 0$

else

$z^{\text{old}} \leftarrow z^{\text{new}}$

if $\text{count} = t$ **then**

$z^{\text{cut}} = z^{\text{new}}$

Exit Loop.

else

$i \leftarrow i + 1(\text{mod } t)$

$\text{count} \leftarrow \text{count} + 1$

end loop

Once we stop adding cut on some N_i , we check whether there is a valid cut on N_{i+1} . The index count is the number of groups of supports that adding cuts will not improve the optimal objective function value. Also, as long as adding a cut produces improvement on the objective value, count will be reset as 0. The algorithm terminates when one of the following happens:

1. An integral feasible solution is found.
2. The parameter count equals to the number of supports t .

In the algorithm, we call a routine to generate the cut on N_i that is formally shown as Algorithm 2. Assume that $\alpha^T x \leq \beta$ is a valid cut on N_i for some i , then $\alpha^T \hat{x} \leq \beta$ holds for all $\hat{x} \in P^I$. However, as the formulation of P^I is implicit, we apply the technique of row generations. Let X be a subset of all integral points in P^I . At the beginning, $X = \emptyset$. And we

generate a valid cut (α^*, β^*) for X . Then we solve the following IP

$$\begin{aligned} \max \quad & \alpha^* x - \beta^* \\ \text{s.t.} \quad & x \in P^I. \end{aligned}$$

If the optimal value is less or equal to 0 then it means the cut is valid for P^I , otherwise let $X = X \cup \{x^*\}$, where x^* is the optimal solution. By re-applying this process, we will either obtain a valid cut or a certificate that no valid cut exists.

Algorithm 2 Cut generation on N_i

Input: $P = \{x | Ax \leq b\}$, $P^I = \text{conv_hull}\{x | x \in P, x \in Z^n\}$, x^*

$X \leftarrow \emptyset$, $\epsilon \leftarrow 10^{-6}$

loop

Solve $(\alpha^*, \beta^*) = \text{argmax}_{x^T \alpha \leq \beta, \forall x \in X, \|\alpha\|_1 = 1, \text{support of } \alpha = N_i} x^{*T} \alpha - \beta$

if $x^{*T} \alpha - \beta > \epsilon$ **then**

Solve $x^0 = \text{argmax}_{x \in P^I} \alpha^* x - \beta^*$

if $\alpha^* x^0 - \beta^* > \epsilon$ **then**

$X \leftarrow X \cup \{x^0\}$

else

Return (α^*, β^*)

Exit Loop

Return $(\alpha, \beta) = (\vec{0}, 0)$

Exit Loop

end loop
