

Contents lists available at [SciVerse ScienceDirect](http://SciVerse.ScienceDirect.com)

Discrete Mathematics

journal homepage: www.elsevier.com/locate/disc

Combinatorial bounds on nonnegative rank and extended formulations

Samuel Fiorini^a, Volker Kaibel^b, Kanstantsin Pashkovich^b, Dirk Oliver Theis^{b,*}^a Department of Mathematics, Université libre de Bruxelles, Belgium^b Faculty of Mathematics, Otto von Guericke Universität Magdeburg, Germany

ARTICLE INFO

Article history:

Received 21 November 2011

Received in revised form 10 July 2012

Accepted 11 September 2012

Available online 11 October 2012

Keywords:

Nonnegative rank

Extended formulations

Projections of polytopes

Boolean rank

Biclique covering in bipartite graphs

ABSTRACT

An extended formulation of a polytope P is a system of linear inequalities and equations that describe some polyhedron which can be projected onto P . Extended formulations of small size (i.e., number of inequalities) are of interest, as they allow to model corresponding optimization problems as linear programs of small sizes. In this paper, we describe several aspects and new results on the main known approach to establish lower bounds on the sizes of extended formulations, which is to bound from below the number of rectangles needed to cover the support of a slack matrix of the polytope. Our main goals are to shed some light on the question how this combinatorial rectangle covering bound compares to other bounds known from the literature, and to obtain a better idea of the power as well as of the limitations of this bound. In particular, we provide geometric interpretations (and a slight sharpening) of Yannakakis' (1991) [35] result on the relation between minimal sizes of extended formulations and the nonnegative rank of slack matrices, and we describe the fooling set bound on the nonnegative rank (due to Dietzfelbinger et al. (1996) [7]) as the clique number of a certain graph. Among other results, we prove that both the cube as well as the Birkhoff polytope do not admit extended formulations with fewer inequalities than these polytopes have facets, and we show that every extended formulation of a d -dimensional neighborly polytope with $\Omega(d^2)$ vertices has size $\Omega(d^2)$.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

An *extended formulation* of a polytope¹ $P \subseteq \mathbb{R}^d$ is a system of linear inequalities and linear equations defining a polyhedron $Q \subseteq \mathbb{R}^e$ along with an affine map $\pi : \mathbb{R}^e \rightarrow \mathbb{R}^d$ satisfying $\pi(Q) = P$. Note that by translating P and/or Q , we may safely assume that π is linear. The *size* of the extended formulation is the number of inequalities in the system (the number of equations is ignored, since one can easily get rid of them). As in this setting linear optimization over P can be done by linear optimization over Q , one is interested in finding small (and simple) extended formulations. Fig. 1 illustrates an extended formulation of size 6 in \mathbb{R}^3 of a regular 8-gon. In fact, every regular 2^k -gon has an extended formulation of size $2k$ [3,21], which can, e.g., be exploited in order to approximate second-order cones.

However, several polytopes associated with combinatorial optimization problems have surprisingly small extended formulations (for recent survey articles we refer to [5,33,19]). Among the nicest examples are extended formulations of size $O(n^3)$ for the spanning tree polytopes of complete graphs with n nodes (due to Martin [27]) and of size $O(n \log n)$ for the permutahedron associated with the permutations of n elements (due to Goemans [14]). It may not be very surprising that no polynomial size extended formulations of polytopes associated with NP-hard optimization problems like the traveling

* Corresponding author.

E-mail addresses: sfiorini@ulb.ac.be (S. Fiorini), kaibel@ovgu.de (V. Kaibel), pashkovi@mail.math.uni-magdeburg.de (K. Pashkovich), dirk.theis@ovgu.de, theis@ovgu.de, theis@ovgu.de (D.O. Theis).¹ For all basic concepts and results on polyhedra and polytopes we refer to [37].

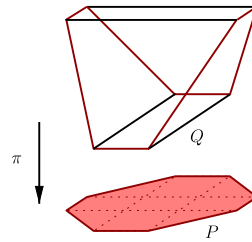


Fig. 1. A regular 8-gon as a projection of a cube.

salesman polytope are known. However, the same holds, e.g., for matching polytopes to this day. While it may well be that we simply still have missed the right techniques to construct polynomial size extended formulations for the latter polytopes, most people would probably agree that for the first ones we simply still miss the right techniques to prove lower bounds on the sizes of extended formulations of concrete polytopes. Note, however, that Rothvoß [29] recently established, by an elegant counting argument, the existence of 0/1-polytopes P in \mathbb{R}^d such that the size of every extended formulation of P is exponential in d .

There is a beautiful approach due to Yannakakis [35] for deriving lower bounds on the sizes of extended formulations that share symmetries of the polytope to be described. In fact, Yannakakis proved that neither the traveling salesman polytopes nor the matching polytopes associated with complete graphs admit polynomial size extended formulations that are invariant under permuting the nodes of the graph. These techniques can, for instance, be extended to prove that the same holds for the polytopes associated with cycles or matchings of logarithmic size in complete graphs, which do, however, have non-symmetric extended formulations of polynomial size [22], as well as to show that the permutahedra do not have symmetric extended formulations of sub-quadratic size [28].

Asking for the smallest size $\text{xc}(P)$ of an arbitrary extended formulation of a polytope P , one finds that it suffices to consider extended formulations that define *bounded* polyhedra Q , i.e., polytopes (see [10,19]). Indeed, if Q is an unbounded polyhedron that is mapped to the polytope P by the linear map π , then the recession cone of Q is contained in the kernel of π (as $\pi(Q) = P$ is bounded). Thus, the pointed polyhedron $Q' = Q \cap L$ with L being the orthogonal complement of the lineality space of Q satisfies $\pi(Q') = P$ as well. If Q' is bounded, we are done. Otherwise, choosing some vector a that satisfies $\langle a, x \rangle > 0$ for all non-zero elements of the recession cone of Q' and some $\beta \in \mathbb{R}$ such that $\beta > \langle a, v \rangle$ holds for all vertices v of Q' , we have that $Q'' := \{x \in Q' : \langle a, x \rangle = \beta\}$ is a polytope with $\pi(Q'') = P$.

Therefore, defining an *extension* of a polytope P to be a *polytope* Q along with an affine projection that maps Q to P , the extension complexity $\text{xc}(P)$ of P equals the minimal number of facets of all extensions of P (where we can even restrict our attention to full dimensional extensions). Note that since a polytope is the set of all convex combinations of its vertices, it has a simplex with the same number of vertices as an extension, showing that $\text{xc}(P)$ is at most the minimum of the numbers of vertices and facets of P .

In the same paper [35] where he established the above mentioned lower bounds on *symmetric* extended formulations for matching and traveling salesman polytopes, Yannakakis essentially also showed that the extension complexity of a polytope P equals the *nonnegative rank* of a *slack matrix* of P , where the latter is a nonnegative matrix whose rows and columns are indexed by the facets and vertices of P , respectively, storing in each row the slacks of the vertices in an inequality defining the respective facet, and the nonnegative rank of a nonnegative matrix M is the smallest number r such that M can be written as a product of two nonnegative matrices with r columns and r rows, respectively. (See [25], or [2] and the references therein.) Such a nonnegative factorization readily induces a covering of the set $\text{supp}(M)$ of non-zero positions of M by r *rectangles*, i.e., sets formed as the cartesian product of subsets of the row and of the column indices. Thus, the smallest number of rectangles that cover $\text{supp}(M)$ (the *rectangle covering number* of M) is a lower bound on the nonnegative rank of M , and hence, on the extension complexity of P if M is the slack matrix of P . Rectangle covering is the same concept as covering the edges of a bipartite graph by bicliques (e.g., [8]), and it also coincides with factorizing a Boolean matrix (e.g., [6,16]).

In fact, almost all techniques to find lower bounds on the extension complexity of polytopes are combinatorial in the sense that they yield lower bounds on the rectangle covering numbers of slack matrices rather than exploiting the true numbers in the matrices. The main goal of this paper is to present current knowledge on the concept of rectangle covering numbers in particular with respect to its relation to the extension complexity of polytopes and to add several results both on lower bounds on rectangle covering number (providing lower bounds on the extension complexity of certain polytopes) as well as on upper bounds (revealing limitations of this approach to find lower bounds on the extension complexity of polytopes). Though the paper is written from the geometric point of view, we hope that it may also be useful for people working on questions concerning the nonnegative rank from a more algebraic point of view.

The paper starts by describing (in Section 2) the relation between rectangle coverings and extensions both in the pure combinatorial setting of embeddings of face lattices as well as via the characterization of the extension complexity as the nonnegative rank of slack matrices. Here, we also give geometric explanations of Yannakakis' algebraic result relating extensions and nonnegative factorizations. We then give some results on upper bounds on rectangle covering numbers. For instance, we show that the rectangle covering number of slack matrices of polytopes with n vertices and at most k vertices in every facet is bounded by $O(k^2 \log n)$ (Proposition 3.2). In Section 4 we briefly describe the connection between the rectangle

covering number of a nonnegative matrix and deterministic as well as nondeterministic communication complexity. While the relation to deterministic communication complexity theory had already been used by Yannakakis [35], recent results of Huang and Sudakov [18] on nondeterministic communication complexity have interesting implications for questions concerning extension complexity as well. In Section 5 we then present several techniques to derive lower bounds on rectangle covering numbers, as well as some new results obtained by using these techniques. We start by proving that the logarithm of the number of faces of a polytope P is not only a lower bound on its extension complexity (as already observed by Goemans [14]) but also on the rectangle number of a slack matrix of P . Observing that the rectangle covering number of a nonnegative matrix is the chromatic number of the *rectangle graph* of the matrix, we interpret the *fooling set bound* (see Dietzfelbinger et al. [7]) as the clique number of that graph. Exploiting that bound, we show, for instance, that for both the cube as well as for the Birkhoff polytope the extension complexity equals the number of facets (Propositions 5.9 and 5.10). However, we also show that for no d -dimensional polytope the clique bound can yield a better lower bound on the extension complexity than $(d + 1)^2$. We also treat (in Section 5.4) the relation between the independence ratio of the rectangle graph and the concept of *generalized fooling sets* (due to Dietzfelbinger et al. [7]), and we prove that for d -dimensional neighborly polytopes with $\Omega(d^2)$ vertices the rectangle covering number (and thus the extension complexity) is bounded from below by $\Omega(d^2)$.

We close this introduction with a few remarks on notation. We only distinguish between row and column vectors in the context of matrix multiplications, where, as usual, vectors are meant to be understood as column vectors. For a matrix M , we denote by $M_{i,*}$ and $M_{*,j}$ the vectors in the i th row and in the j th column of M , respectively. All logarithms are meant to refer to base two.

2. Projections and derived structures

2.1. Lattice embeddings

Let us denote by $\mathcal{L}(P)$ the *face lattice* of a polytope P , i.e., the set of faces of P (including the *non-proper* faces \emptyset and P itself) partially ordered by inclusion. The following proposition describes the relation between the face lattices of two polytopes one of which is a projection of the other. The statement certainly is not new, but as we are not aware of any explicit reference for it, we include a brief proof. An *embedding* of a partially ordered set (S, \leq) into a partially ordered set (T, \sqsubseteq) is a map f such that $u \leq v$ if and only if $f(u) \sqsubseteq f(v)$. Notice that every embedding is injective.

Proposition 2.1. *If $Q \subseteq \mathbb{R}^e$ and $P \subseteq \mathbb{R}^d$ are two polytopes, and $\pi : \mathbb{R}^e \rightarrow \mathbb{R}^d$ is an affine map with $\pi(Q) = P$, then the map that assigns $h(F) := Q \cap \pi^{-1}(F)$ to each face F of P defines an embedding of $\mathcal{L}(P)$ into $\mathcal{L}(Q)$.*

Proof. For any face F of P defined by an inequality $\langle a, x \rangle \leq \beta$, we have $\langle a, \pi(y) \rangle \leq \beta$ for all $y \in Q$ with equality if and only if $\pi(y) \in F$ holds. Thus, $h(F)$ is indeed a face of Q , defined by the linear inequality $\langle a, \pi(y) \rangle \leq \beta$. Moreover, h is clearly an embedding. \square

Remark 2.2. From the definition of h , we see that $h(F \cap G) = h(F) \cap h(G)$ holds for all faces F, G . We call a lattice embedding h with this property *meet-faithful*. Clearly, not all lattice embeddings are meet-faithful.

Fig. 2 illustrates the embedding from Proposition 2.1. In the figure, P is a 4-dimensional cross-polytope, Q is a 7-dimensional simplex and π is any affine projection mapping the 8 vertices of Q to the 8 vertices of P . Denoting by e_i the i th unit vector, we have $Q = \text{conv}\{0, e_1, \dots, e_7\} \subseteq \mathbb{R}^7, P = \text{conv}\{e_1, -e_1, \dots, e_4, -e_4\} \subseteq \mathbb{R}^4$ and for instance $\pi(0) = e_1, \pi(e_1) = -e_1, \dots, \pi(e_7) = -e_4$. As the figure suggests correctly, constructing a small extended formulation for a polytope P means to hide the facets of P in the fat middle part of the face lattice of an extension with few facets.

2.2. Slack representations

Let P be a polytope in \mathbb{R}^d . Let $Ax \leq b$ be a system of linear inequalities such that

$$P = \{x \in \mathbb{R}^d : Ax \leq b\},$$

and let m denote the number of inequalities involved in $Ax \leq b$. The *slack vector* of a point $x \in \mathbb{R}^d$ is the vector $b - Ax \in \mathbb{R}^m$, and the *slack map* of P w.r.t. $Ax \leq b$ is the affine map $\sigma : \mathbb{R}^d \rightarrow \mathbb{R}^m$ that maps each vector to its slack vector. We call the polytope $\tilde{P} := \sigma(P)$, which is the image of P under the slack map σ , the *slack representation* of P w.r.t. $Ax \leq b$. Clearly the affine hull $\text{aff}(\tilde{P})$ of \tilde{P} is the image of the affine hull $\text{aff}(P)$ of P under σ . We call $\text{aff}(\tilde{P})$ the *slack space* of P w.r.t. $Ax \leq b$. Note that the polytope \tilde{P} is affinely isomorphic to the polytope P and $\tilde{P} = \text{aff}(\tilde{P}) \cap \mathbb{R}_+^m$.

A set of r nonnegative vectors $T := \{t_1, \dots, t_r\} \subseteq \mathbb{R}_+^m$ is called a *slack generating set* of P w.r.t. $Ax \leq b$ of size r if every point in \tilde{P} can be expressed as a nonnegative combination of the vectors in T . Then the following system

$$\sum_{k=1}^r \lambda_k t_k \in \text{aff}(\tilde{P}) \quad \text{and} \quad \lambda_k \geq 0 \quad \text{for all } k \in [r]$$

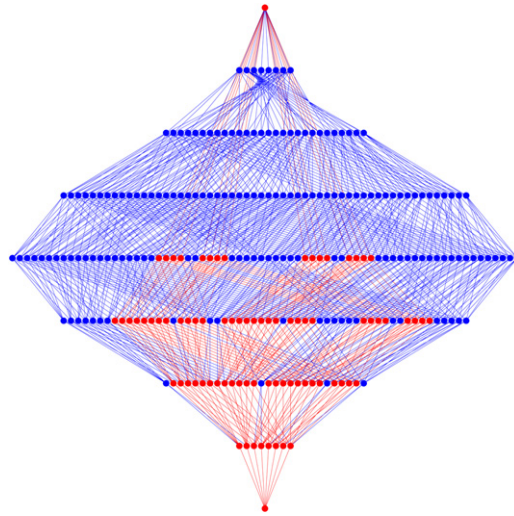


Fig. 2. Embedding of the face lattice of the 4-dimensional cross-polytope into the face lattice of the 7-dimensional simplex.

provides an extended formulation of \tilde{P} of size r via the projection map $\lambda \mapsto \sum_{k=1}^r \lambda_k t_k$. Since P and \tilde{P} are affinely isomorphic, this also yields an extension of P of size r , which is called a *slack extension* of P w.r.t. $Ax \leq b$.

Lemma 2.3. *If the inequality $\langle c, x \rangle \leq \delta$ is valid for a polytope $P := \{x \in \mathbb{R}^d \mid Ax \leq b\}$ with $\dim(P) \geq 1$, then it is a nonnegative combination of the inequalities of the system $Ax \leq b$ defining P .*

Proof. By Farkas's lemma, because $P \neq \emptyset$, there exists $\delta' \leq \delta$ such that $\langle c, x \rangle \leq \delta'$ is a nonnegative combination of the inequalities of $Ax \leq b$. Now let i be an index such that $\min\{\langle A_{i,\star}, x \rangle \mid x \in P\} < \max\{\langle A_{i,\star}, x \rangle \mid x \in P\} = b_i$. Such an index exists because $\dim(P) \geq 1$. Letting $b'_i := \min\{\langle A_{i,\star}, x \rangle \mid x \in P\}$, we see that the inequality $-\langle A_{i,\star}, x \rangle \leq -b'_i$ is valid for P . By Farkas's lemma, because of the minimality of b'_i , this last inequality is a nonnegative combination of the inequalities of $Ax \leq b$. By adding the inequality $\langle A_{i,\star}, x \rangle \leq b_i$ and then scaling by $(b_i - b'_i)^{-1}$, we infer that $0 \leq 1$ is a nonnegative combination of the inequalities of $Ax \leq b$. After scaling by $\delta - \delta'$ and adding the inequality $\langle c, x \rangle \leq \delta'$, we conclude that $\langle c, x \rangle \leq \delta$ is a nonnegative combination of the inequalities of $Ax \leq b$. \square

Lemma 2.4. *Let $P := \{x \in \mathbb{R}^d \mid Ax \leq b\}$ be a polytope with $\dim(P) \geq 1$. If the polytope P has f facets, then P has a slack generating set of size f w.r.t. $Ax \leq b$.*

Proof. Let m and r denote the total number and the number of nonredundant inequalities of $Ax \leq b$ respectively. W.l.o.g., assume that these r nonredundant inequalities come first in the system $Ax \leq b$. By Lemma 2.3, each of the $m - r$ last inequalities of $Ax \leq b$ is a nonnegative combination of the r first inequalities of $Ax \leq b$. This implies that, for each $i \in [m] \setminus [r]$, there are nonnegative coefficients $t_{k,i} \in \mathbb{R}_+$ such that

$$A_{i,\star} = \sum_{k=1}^r A_{k,\star} t_{k,i} \quad \text{and} \quad b_i = \sum_{k=1}^r b_k t_{k,i}.$$

By letting $t_{k,i} := 1$ if $i = k$ and $t_{k,i} := 0$ for $i \in [r]$, the above equations hold for all $i \in [m]$. This defines a set $T := \{t_1, \dots, t_r\}$ of nonnegative vectors $t_k \in \mathbb{R}_+^m$, $k \in [r]$.

To show that T is a slack generating set, consider for any point $x \in P$ the nonnegative combination

$$\sum_{k=1}^r (b_k - \langle A_{k,\star}, x \rangle) t_k. \tag{1}$$

For $i \in [m]$, the i th coordinate of this vector equals

$$\sum_{k=1}^r (b_k - \langle A_{k,\star}, x \rangle) t_{k,i} = \sum_{k=1}^r b_k t_{k,i} - \left\langle \sum_{k=1}^r A_{k,\star} t_{k,i}, x \right\rangle = b_i - \langle A_{i,\star}, x \rangle,$$

thus (1) equals the slack vector of x .

Moreover, in (1) the coefficients corresponding to inequalities that hold with equality for all points of P can be chosen to be zeros. Thus the corresponding vectors can be removed from T , resulting in a slack generating set of size f . \square

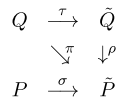


Fig. 3. The four polytopes P, \tilde{P}, Q and \tilde{Q} and the maps between them.

Theorem 2.5. *The extension complexity of a polytope P with $\dim(P) \geq 1$ is equal to the minimum size of a slack generating set of P w.r.t. any given system $Ax \leq b$ defining P .*

Proof. For an extension $Q \subseteq \mathbb{R}^e$ of P of size r we have to show that there is a slack generating set of the polytope P of size r . Since the polytopes P and \tilde{P} are affinely isomorphic, Q is an extension of the polytope \tilde{P} , i.e there exists an affine map $\pi : \mathbb{R}^e \rightarrow \mathbb{R}^m$ such that $\pi(Q) = \tilde{P}$. Let $c_i \in \mathbb{R}^e, i \in [m]$ and $g_i \in \mathbb{R}$ be such that $\pi(y)_i = g_i - \langle c_i, y \rangle$. Note that the inequality $\langle c_i, y \rangle \leq g_i$ is valid for Q , because $\pi(Q) \subseteq \tilde{P} \subseteq \mathbb{R}_+^m$.

Now, consider a system $Cy \leq g$ of n linear inequalities describing Q such that the first m inequalities of this system are $\langle c_1, y \rangle \leq g_1, \dots, \langle c_m, y \rangle \leq g_m$, and a corresponding slack map $\tau : \mathbb{R}^e \rightarrow \mathbb{R}^n$. Again, $\tilde{Q} := \tau(Q)$ is affinely isomorphic to Q . Therefore, \tilde{Q} is also an extension of \tilde{P} . Indeed, the projection $\rho : \mathbb{R}^n \rightarrow \mathbb{R}^m$ to the first m coordinates maps \tilde{Q} onto \tilde{P} , because $\pi = \rho \circ \tau$ (see Fig. 3).

By Lemma 2.4 there is a slack generating set $\{t_1, \dots, t_r\} \subseteq \mathbb{R}_+^n$ for Q of size r . We claim that stripping the last $n - m$ coordinates of these vectors results in a slack generating set $\{\rho(t_1), \dots, \rho(t_r)\}$ of the polytope P .

Indeed, consider a point $x \in P$ and its slack vector $\sigma(x) \in \tilde{P}$. Because $\tilde{P} \subseteq \pi(Q)$ there is a point $y \in Q$ such that $\pi(y) = \sigma(x)$. Because $\{t_1, \dots, t_r\}$ is a slack generating set of Q , there exist nonnegative coefficients $\lambda_1, \dots, \lambda_r \in \mathbb{R}_+$ such that $\tau(y) = \sum_{k=1}^r \lambda_k t_k$. From linearity of ρ and $\pi = \rho \circ \tau$, we have

$$\sigma(x) = \pi(y) = \rho(\tau(y)) = \rho\left(\sum_{k=1}^r \lambda_k t_k\right) = \sum_{k=1}^r \lambda_k \rho(t_k)$$

and thus P has a slack generating set $\{\rho(t_1), \dots, \rho(t_r)\}$ of size r . \square

Note that in Theorem 2.5 one may take the minimum over the slack generating sets w.r.t. any fixed system of inequalities describing P . In particular, all these minima coincide.

2.3. Non-negative factorizations

Again, let $P = \{x \in \mathbb{R}^d : Ax \leq b\}$ be a polytope, where $A \in \mathbb{R}^{m \times d}, b \in \mathbb{R}^m$. Let $V = \{v_1, \dots, v_n\} \subseteq \mathbb{R}^d$ denote any finite set such that

$$P = \text{conv}(V).$$

The slack matrix of P w.r.t. $Ax \leq b$ and V is the matrix $S = (S_{i,j}) \in \mathbb{R}_+^{m \times n}$ with

$$S_{i,j} = b_i - \langle A_{i,*}, v_j \rangle \quad \text{for all } i \in [m], j \in [n].$$

Note that the slack representation $\tilde{P} \subseteq \mathbb{R}^m$ of P (w.r.t. $Ax \leq b$) is the convex hull of the columns of S .

If the columns of a nonnegative matrix $T \in \mathbb{R}_+^{m \times r}$ form a slack generating set of P , then there is a nonnegative matrix $U \in \mathbb{R}_+^{r \times n}$ with $S = TU$. Conversely, for every factorization $S = TU$ of the slack matrix into nonnegative matrices $T \in \mathbb{R}_+^{m \times r}$ and $U \in \mathbb{R}_+^{r \times n}$, the columns of T form a slack generating set of P .

Therefore, due to Theorem 2.5, constructing an extended formulation of size r for P amounts to finding a rank- r nonnegative factorization of the slack matrix S , that is a factorization $S = TU$ into nonnegative matrices T with r columns and U with r rows. In particular, the following result follows, which is essentially due to Yannakakis [35] (he proved that extension complexity and nonnegative rank are within a factor of two of each other, when the size of an extension is defined as the sum of the number of variables and number of constraints defining the extension).

Theorem 2.6 (See Yannakakis [35]). *The extension complexity of a polytope P (which is neither empty nor a single point) is equal to the nonnegative rank of any of its slack matrices. In particular, all the slack matrices of P have the same nonnegative rank.*

Remark 2.7. It is obvious from the definition that the extension complexity is monotone on extensions: if polytope Q is an extension of polytope P , then $\text{xc}(Q) \geq \text{xc}(P)$. From Theorem 2.6, we see immediately that the extension complexity is also monotone on faces: if polytope P has F as a face, then $\text{xc}(P) \geq \text{xc}(F)$ because we can obtain a slack matrix of F from a slack matrix of P by deleting all columns that correspond to points of P which are not in F .

A nonnegative matrix T is the first factor in a nonnegative factorization of a slack matrix of P if and only if the columns of T form a slack generating set of P . In order to characterize the second factors of such nonnegative factorizations, let us consider an extension Q of P with projection π and some set $V = \{v_1, \dots, v_n\}$ with $P = \text{conv}(V)$. A *section* is a map $s : V \rightarrow Q$ such that $\pi(s(x)) = x$ holds for all $x \in V$. Clearly, every extension possesses a section map.

Recall that $(\mathbb{R}^n)^*$ denotes the dual vector space of \mathbb{R}^n , which we here regard as the set of all row vectors of size n . For an inequality $\langle c, y \rangle \leq g$ that is valid for Q , we construct a nonnegative row vector in $(\mathbb{R}^n)^*_+$ (the nonnegative orthant in $(\mathbb{R}^n)^*$), whose j th coordinate equals $g - \langle c, s(v_j) \rangle$ and call it the *section slack covector* associated with $\langle c, y \rangle \leq g$ w.r.t. set V . A set U of nonnegative row vectors in $(\mathbb{R}^n)^*_+$ is a *complete set of section slack covectors* if there is some extended formulation for P along with some section such that U is precisely the set of section slack covectors associated with the inequalities in the extended formulation. A nonnegative matrix U is the second factor in a nonnegative factorization of a slack matrix of P if and only if the rows of U form a complete set of section slack covectors (both w.r.t. to the same set V with $P = \text{conv}(V)$). From this, one in particular derives (again using Farkas's Lemma) the following characterization, where the *slack covector* associated with some valid inequality $\langle a, x \rangle \leq b$ for $P = \text{conv}(V)$ is the nonnegative row vector in $(\mathbb{R}^n)^*_+$ whose j th coordinate is $b - \langle a, v_j \rangle$.

Proposition 2.8. *A set $U \subseteq (\mathbb{R}^n)^*_+$ is a complete set of section slack covectors for $P = \text{conv}(V)$ (with $1 \leq |V| < \infty$) if and only if every slack covector associated with a valid inequality for P can be expressed as a nonnegative combination of the elements of U .*

2.4. Rectangle coverings

According to Theorem 2.6, finding lower bounds on the extension complexity of a polytope amounts to finding lower bounds on the nonnegative rank of its slack matrices. Not surprisingly, determining the nonnegative rank of a matrix is a hard problem from the algorithmic point of view. Indeed, it is NP-hard to decide whether the nonnegative rank of a matrix equals its usual rank [34] (where, of course, the first is never smaller than the second).

One way to bound the nonnegative rank $\text{rank}_+(M)$ from below is to observe that a nonnegative factorization $M = TU$, where T has r columns and U has r rows, yields a representation

$$M = \sum_{k=1}^r T_{\star,k} U_{k,\star}$$

of M as a sum of r nonnegative rank-1 matrices $T_{\star,k} U_{k,\star}$. Denoting by $\text{supp}(\cdot)$ the *support* of a vector or of a matrix (i.e., the subset of indices where the argument has a nonzero entry), we obviously have

$$\text{supp}(M) = \bigcup_{k=1}^r \text{supp}(T_{\star,k} U_{k,\star})$$

since both T and U are nonnegative.

A *rectangle* is a set of the form $I \times J$, where I and J are subsets of the row respectively column indices of M . A *rectangle covering* of M is a set of rectangles whose union equals $\text{supp}(M)$. It is important to notice that all the rectangles in any rectangle covering of M are contained in the support of M . The *rectangle covering number* of M is the smallest cardinality $\text{rc}(M)$ of any rectangle covering of M . Clearly, we have

$$\text{rc}(M) \leq \text{rank}_+(M) \tag{2}$$

for all nonnegative matrices M since

$$\text{supp}(T_{\star,k} U_{k,\star}) = \text{supp}(T_{\star,k}) \times \text{supp}(U_{k,\star})$$

holds for each $k \in [r]$. In particular, when $M = S$ is some slack matrix of the polytope P (neither empty nor a single point), then by Theorem 2.6 we have

$$\text{rc}(S) \leq \text{rank}_+(S) = \text{xc}(P).$$

Like $\text{rank}_+(S) = \text{xc}(P)$, the rectangle covering number $\text{rc}(S)$ is actually independent of the actual choice of the slack matrix S of P ; see Lemma 2.10 below.

Let us call the *support matrix* $\text{suppmat}(S)$ of S the 0/1-matrix arising from S by replacing all nonzero-entries by ones. Clearly, we have $\text{rc}(S) = \text{rc}(\text{suppmat}(S))$. The rectangles that can be part of a rectangle covering of S are called *1-rectangles* because any such rectangle induces a submatrix of $\text{suppmat}(S)$ that contains only one-entries.

Furthermore, any 0/1-matrix whose rows are indexed by some set F_1, \dots, F_m of faces of P including all facets and whose columns are indexed by some set G_1, \dots, G_n of nonempty faces of P including all vertices such that there is a one-entry at position (F_i, G_j) if and only if face F_i does not contain face G_j is called a *non-incidence matrix* for P . Associating with every inequality in $Ax \leq b$ the face of P it defines and with every point in V the smallest face of P it is contained in, one finds that the set of support matrices of slack matrices of P equals the set of non-incidence matrices of P (up to adding/removing repeated rows or columns).

Clearly, adding a row or a column to a nonnegative matrix does neither decrease the rectangle covering number nor the nonnegative rank. The following result on the rectangle covering number is easy to see.

Lemma 2.9. *The rectangle covering number remains unchanged if one adds a row or a column whose support is the union of the supports of some existing rows or columns, respectively.*

Now, we can prove that all non-incidence matrices, and thus all slack matrices, of a polytope (of dimension at least 1) have the same rectangle covering number.

Lemma 2.10. *All non-incidence matrices of a polytope (which is neither empty nor a single point) have the same rectangle covering number.*

Proof. Letting P denote the polytope in the statement of the lemma, consider a non-incidence matrix M for P . Every proper face F of P contains exactly those faces of P that are contained in all facets of P that contain F . For the non-incidence matrix M , this implies that the support of the row corresponding to F is exactly the union of the supports of the rows corresponding to the facets containing F .

Similarly, every nonempty face G of P is contained in exactly those faces of P in which all vertices of G are contained. This implies that the support of the column corresponding to G is exactly the union of the supports of the rows corresponding to the vertices contained in F .

The result then follows from Lemma 2.9 because the rectangle covering number of M equals the rectangle covering of the submatrix of M corresponding to the facets and vertices of P , and thus does not depend on M . (Notice that there might be a row in M for $F = P$, but this row is identically zero.) \square

Denoting by $\text{rc}(P)$ the rectangle covering number of any non-incidence matrix for P , we have

$$\text{rc}(P) \leq \text{xc}(P). \quad (3)$$

When studying the rectangle covering number, we may freely choose the non-incidence matrix we consider. The most natural choice is the *facet vs. vertex* non-incidence matrix, since it appears as a submatrix of every non-incidence matrix. (The rows of this 0/1-matrix are indexed by the facets, and the columns are indexed by the vertices. The entry corresponding to a facet–vertex pair is 1 if and only if the facet does not contain the vertex.) Nevertheless, we will sometimes consider non-incidence matrices with more rows.

2.5. Rectangle coverings, Boolean factorizations and lattice embeddings

As mentioned before, a rank- r nonnegative factorization of a nonnegative matrix M can be regarded as a decomposition of M into a sum of r nonnegative rank-1 matrices. Similarly, we can regard a rectangle covering of a Boolean (or 0/1-) matrix M with r rectangles as a *rank- r Boolean factorization*, that is a factorization $M = TU$ expressing M as the Boolean product of two Boolean matrices T with r columns and U with r rows. Furthermore, a rank- r Boolean factorization of a Boolean matrix M is equivalent to an embedding of the relation defined by M into the Boolean lattice $2^{[r]}$ (that is, the set of all subsets of $[r]$, partially ordered by inclusion), in the sense of the following lemma.

Lemma 2.11. *Let M be a $m \times n$ Boolean matrix. Then M admits a rank- r Boolean factorization $M = TU$ if and only if there are functions $f: [m] \rightarrow 2^{[r]}$ and $g: [n] \rightarrow 2^{[r]}$ such that $M_{ij} = 0$ if and only if $g(j) \subseteq f(i)$ for all $(i, j) \in [m] \times [n]$.*

Proof. This follows immediately by interpreting the i th row of the left factor T as the incidence vector of the *complement* of the set $f(i)$ and the j th column of the right factor U as the incidence vector of the set $g(j)$. \square

From Lemma 2.11, we can easily conclude a first lattice-combinatorial characterization of the rectangle covering number $\text{rc}(P)$ of a polytope P by taking M to be a non-incidence matrix of P .

Theorem 2.12. *Let P be a polytope with $\dim(P) \geq 1$. Then $\text{rc}(P)$ is the smallest $r \geq 1$ such that $\mathcal{L}(P)$ embeds into $2^{[r]}$.*

For a given poset, the smallest number r such that P embeds into the Boolean poset $2^{[r]}$ is known as the 2-dimension of the poset (see e.g., [17] and the references therein). We note that, independently, Gouveia et al. [15] have found a similar connection.

Proof of the Theorem 2.12. Let F_1, \dots, F_m denote the facets of P , let v_1, \dots, v_n denote the vertices of P and let M denote the facet vs. vertex non-incidence matrix for P .

Suppose first that $\text{rc}(P) \leq r$, that is, M has a rank- r Boolean factorization. From maps f and g as in Lemma 2.11, we define a map h from $\mathcal{L}(P)$ to $2^{[r]}$ by letting $h(F) := \bigcap_{i:F \subseteq F_i} f(i)$. It is clear that $F \subseteq G$ implies $h(F) \subseteq h(G)$. Now assume $h(F) \subseteq h(G)$. Pick a vertex v_j of F . We have $M_{ij} = 0$ and thus $g(j) \subseteq f(i)$ for all facets F_i containing F . Hence, $g(j) \subseteq h(F)$. Because $h(F) \subseteq h(G)$, we have $g(j) \subseteq f(i)$ and thus $M_{ij} = 0$ for all facets F_i containing G . Since v_j is arbitrary, this means that every facet containing G contains all vertices of F , which implies $F \subseteq G$. Therefore, h is an embedding.

Next, if h is an embedding of $\mathcal{L}(P)$ into $2^{[r]}$, we can simply define $f(i) := h(F_i)$ and $g(j) := h(\{v_j\})$. These maps satisfy the conditions of Lemma 2.11; hence M has a rank- r Boolean factorization and $\text{rc}(P) \leq r$. The theorem follows. \square

We immediately obtain a second lattice-combinatorial characterization of the rectangle covering number of a polytope.

Corollary 2.13. *The rectangle covering number $\text{rc}(P)$ of a polytope P with $\dim(P) \geq 1$ is equal to the minimum number of facets of a polytope Q into whose face lattice $\mathcal{L}(Q)$ the face lattice $\mathcal{L}(P)$ of P can be embedded.*

Corollary 2.13 follows from Theorem 2.12 and Lemma 2.14 below.

Lemma 2.14. *If Q is a polytope with $r \geq 2$ facets then there is an embedding of $\mathcal{L}(Q)$ into $2^{[r]}$.*

Proof. We denote by G_1, \dots, G_r the facets of Q , and define a map h from $\mathcal{L}(Q)$ to $2^{[r]}$ by letting $h(G) := \{k \in [r] \mid G \not\subseteq G_k\}$ for all faces G of Q . It is easily verified that h is an embedding. \square

The embedding $\mathcal{L}(P) \rightarrow \mathcal{L}(Q)$ given by Corollary 2.13 is not always meet-faithful. It is unclear whether requiring that the embedding be meet-faithful would give a (much) better bound.

Remark 2.15. Similarly to the extension complexity, the rectangle covering number is easily shown to be monotone on extensions and on faces. If the polytope Q is an extension of the polytope P , then we may infer, e.g., from Proposition 2.1 and Corollary 2.13 that $\text{rc}(Q) \geq \text{rc}(P)$. Moreover, by reasoning on the slack matrices directly, we see that $\text{rc}(P) \geq \text{rc}(F)$ whenever polytope P has F as a face.

3. Upper bounds on the rectangle covering number

In this section, we discuss some examples of polytopes for which small rectangle coverings can be found, as well as methods for constructing such rectangle coverings. These examples illustrate cases where the rectangle covering bound shows its limitations. Cases where the rectangle covering bound can be successfully applied to obtain interesting lower bounds on the extension complexity are discussed in Section 5.

3.1. The perfect matching polytope

The perfect matching polytope $P_{\text{match}}(n)$ is defined as the convex hull of characteristic vectors of perfect matchings in the complete graph $K^n = (V_n, E_n)$. A linear description of the perfect matching polytope $P_{\text{match}}(n)$ is as follows [9]:

$$P_{\text{match}}(n) = \{x \in \mathbb{R}^{E_n} : x(\delta(v)) = 1 \text{ for all } v \in V_n, x(\delta(S)) \geq 1 \text{ for all } S \subseteq V_n, 3 \leq |S| \leq n-3, |S| \text{ odd} \\ x_e \geq 0 \text{ for all } e \in E_n\},$$

where $\delta(S)$ denotes all edges in the graph with exactly one endpoint in S , $\delta(v) := \delta(\{v\})$, and $x(F) := \sum_{e \in F} x_e$ for all edge sets $F \subseteq E_n$. Currently, no polynomial-size extension is known for the perfect matching polytope $P_{\text{match}}(n)$. Moreover, it was shown that under certain symmetry requirements no polynomial-size extension exists [35,22]. On the other side it is possible that a non-symmetric polynomial-size extension for the perfect matching polytope can be found [22]. In this context any non-trivial statement about the extension complexity of $P_{\text{match}}(n)$ is interesting.

The non-incidence matrix M (w.r.t. the above system of constraints and the characteristic vectors of perfect matchings) of the perfect matching polytope has $\Theta(n^2)$ rows corresponding to the non-negative constraints and $\Theta(2^n)$ rows corresponding to the inequalities indexed by odd subsets of vertices of K^n . A facet corresponding to some odd set S and a vertex corresponding to some matching are non-incident if and only if the matching has more than one edge in $\delta(S)$.

Lemma 3.1. *The rectangle covering number $\text{rc}(P_{\text{match}}(n))$ is $O(n^4)$.*

Proof. The non-zero entries in rows corresponding to the non-negative constraints can be trivially covered by $O(n^2)$ 1-rectangles of height one. Less obvious is the fact that the non-zero entries corresponding to the odd-subset inequalities can be covered by $O(n^4)$ 1-rectangles [35]. For this one considers the rectangles $R_{e_1, e_2} = I_{e_1, e_2} \times J_{e_1, e_2}$ indexed by unordered pair of edges e_1, e_2 , where the set I_{e_1, e_2} consists of all odd sets S such that $e_1, e_2 \in \delta(S)$ and the set J_{e_1, e_2} consists of all the matchings containing both edges e_1, e_2 . \square

Thus in the case of $P_{\text{match}}(n)$ we cannot obtain any lower bound better than $O(n^4)$ solely by reasoning on rectangle coverings.

3.2. Polytopes with few vertices on every facet

We consider a polytope P with n vertices such that every facet of P contains at most k vertices, and its facet vs. vertex non-incidence matrix M . Each row of M has at most k zeros, and at least $n-k$ ones. We now prove that the rectangle covering number of such a polytope is necessarily small.

Proposition 3.2. *If P is a polytope with n vertices and at most k vertices on each facet, then $\text{rc}(P) = O(k^2 \log n)$.*

Before giving the proof, we state the following lemma that is the main ingredient of the proof.

Lemma 3.3. Let M be the 0/1 matrix with rows indexed by $\{S \subseteq [n] : |S| \leq k\}$ and columns indexed by $\{T \subseteq [n] : |T| \leq \ell\}$, where $\ell \leq k \leq n$, such that the entry $M_{S,T}$ is non-zero if and only if the sets S, T are disjoint. Then $\text{rc}(M) = O((k + \ell)e^\ell(1 + k/\ell)^\ell \log n)$.

Proof. Consider rectangles $I_V \times J_V, V \subseteq [n]$, where $I_V := \{S \subseteq [n] : S \subseteq V, |S| \leq k\}$ and $J_V := \{T \subseteq [n] : T \cap V = \emptyset, |T| \leq \ell\}$. Obviously, every rectangle $I_V \times J_V, V \subseteq [n]$ is contained in $\text{supp}(M)$.

Pick a set $V \subseteq [n]$ by selecting the points from $[n]$ independently with probability p , for some $p \in [0, 1]$. For a fixed pair (S, T) of disjoint sets $S, T \subseteq [n]$ with $|S| \leq k$ and $|T| \leq \ell$, the probability to be covered is at least

$$q := p^k(1 - p)^\ell,$$

thus choosing the probability p equal to $\frac{k}{k+\ell}$ to maximize the value q we get

$$q = \left(\frac{k}{k + \ell}\right)^k \left(1 - \frac{k}{k + \ell}\right)^\ell \geq e^{-\ell} \left(\frac{\ell}{k + \ell}\right)^\ell.$$

Now, let us bound the natural logarithm of the expected number of entries from $\text{supp}(M)$ which are not covered if we choose independently r such rectangles. An upper bound on this quantity is:

$$\begin{aligned} \ln((n + 1)^k(n + 1)^\ell(1 - q)^r) &\leq (k + \ell) \ln(n + 1) + r \ln(1 - q) \\ &\leq (k + \ell) \ln(n + 1) - rq \\ &\leq (k + \ell) \ln(n + 1) - re^{-\ell} \left(\frac{\ell}{k + \ell}\right)^\ell. \end{aligned}$$

If the above upper bound for the logarithm of the expected number of not covered entries from $\text{supp}(M)$ is negative, we can conclude that there exists a rectangle cover for the matrix M of size r . Thus there exists a rectangle cover of size $O((k + \ell)e^\ell(1 + k/\ell)^\ell \log n)$. \square

Proof of Proposition 3.2. Let M be the facet vs. vertex non-incidence matrix of P . It suffices to show that $\text{rc}(M) = O(k^2 \log n)$. We may extend M by adding extra rows in order to obtain a matrix in which each binary vector of size n with at most k zeros appears as a row. Obviously, this operation does not decrease $\text{rc}(M)$. The result then follows from Lemma 3.3 by taking $\ell := 1$. \square

The most natural case where Proposition 3.2 applies is perhaps when P is a simplicial d -dimensional polytope with n vertices. For such a polytope there is a rectangle covering of the non-incidence matrix with $O(d^2 \log n)$ rectangles.

3.3. The edge polytope

The edge polytope $P_{\text{edge}}(G)$ of a graph G is defined as the convex hull of the incidence vectors in $\mathbb{R}^{V(G)}$ of all edges of G . Thus $P_{\text{edge}}(G)$ is a 0/1-polytope with $|E(G)|$ vertices in $\mathbb{R}^{V(G)}$. Consider a stable set S of G . Denoting by $N(S)$ the neighborhood of S in G , we see that the inequality

$$x(S) - x(N(S)) \leq 0 \tag{4}$$

is valid for $P_{\text{edge}}(G)$. It can be shown [20] that these inequalities, together with $x_v \geq 0$ for $v \in V(G)$ and $x(V(G)) = 2$, form a complete linear description of $P_{\text{edge}}(G)$.

3.3.1. A sub-quadratic size extension for all graphs

The following lemma provides an upper bound not just on the rectangle covering number of $P_{\text{edge}}(G)$, but also on the extension complexity of $P_{\text{edge}}(G)$.

Lemma 3.4. For every graph G with n vertices there exists an extension of the edge polytope $P_{\text{edge}}(G)$ of size $O(n^2 / \log n)$.

Proof. Let H denote a biclique (i.e., complete bipartite graph) with bipartition W, U . The edge polytope $P(H)$ of H has a linear description of size $O(|V(H)|)$. Namely, in addition to the equalities $x(W) = 1$ and $x(U) = 1$, the linear description consists of the nonnegativity constraints $x_v \geq 0$ for $v \in V(H)$. From any covering of the edges of G by bicliques H_1, \dots, H_t (with $H_i \subseteq G$ for all i), we can construct an extension of $P_{\text{edge}}(G)$ of the size $O(t + |V(H_1)| + \dots + |V(H_t)|)$ using disjunctive programming [1]. Since $t \leq |V(H_1)| + \dots + |V(H_t)|$ and since there is such a covering with $|V(H_1)| + \dots + |V(H_t)| = O(n^2 / \log n)$ [32] the lemma follows. \square

3.3.2. Dense graphs

For two matrices A and B of the same size, denote by $A \cdot B$ the entry-wise (or Hadamard) product: $(A \cdot B)_{i,j} = A_{i,j}B_{i,j}$. The following statements are easy to verify.

Lemma 3.5. (a) For any nonnegative real matrices we have $\text{rank}_+(A \cdot B) \leq \text{rank}_+(A) \text{rank}_+(B)$.
 (b) For any real matrices we have $\text{rc}(A \cdot B) \leq \text{rc}(A) \text{rc}(B)$.

Let G be a graph with n vertices. We denote $\mathcal{I}(G)$ the set of all independent sets of G . The slack matrix of $P_{\text{edge}}(G)$ has two types of rows: those which correspond to inequalities of the form (4) for S stable in G , and those which correspond to nonnegativity inequalities $x_v \geq 0$ for $v \in V(G)$. Let M be the support of the submatrix of the slack matrix of $P_{\text{edge}}(G)$ induced by the rows corresponding to stable sets S ; let $M^{(1)}$ be the $\mathcal{I}(G) \times E(G)$ -matrix with $M_{S,e}^{(1)} = 1$ if $e \cap S = \emptyset$ and 0 otherwise; and let $M^{(2)}$ be the $\mathcal{I}(G) \times E(G)$ -matrix with $M_{S,uv}^{(2)} = 1$ if u or v has a neighbor in S . Then we have

$$M = M^{(1)} \cdot M^{(2)}. \quad (5)$$

Lemma 3.6. For the matrix $M^{(2)}$ defined above we have $\text{rc}(M^{(2)}) \leq n$.

Proof. For each $v \in V(G)$, define a 1-rectangle $I_v \times J_v$. The set I_v consists of all stable sets that contain a neighbor of v , and the set J_v consists of all edges incident to v . These n rectangles define a rectangle covering of $M^{(2)}$. \square

Now, we can prove an upper bound on $\text{rc}(P(G))$.

Proposition 3.7. Denoting by $\alpha(G)$ the stability number of G , the support of the slack matrix of the edge polytope of G admits a rectangle cover of size

$$O(\alpha(G)^3 n \log n).$$

Proof. Lemma 3.3 for $\ell := 2$ and $k := \alpha(G)$ implies $\text{rc}(M^{(1)}) = O(\alpha(G)^3 \log n)$. The result follows by combining Lemmas 3.5 and 3.6. \square

Combining Turán's Theorem [31], which states that $|E(G)| \geq \frac{|V(G)|^2}{2\alpha(G)}$, with Proposition 3.7, we obtain the final result of this section: there exists a nontrivial class of graphs for which the rectangle covering number cannot prove a strong $\Omega(|E(G)|)$ lower bound on the extension complexity.

Corollary 3.8. The support of the slack matrix of the edge polytope of G admits a rectangle cover of size $o(|E(G)|)$ if

$$\alpha(G) = o\left(\left(\frac{n}{\log n}\right)^{1/4}\right). \quad \square$$

4. Rectangle covering number and communication complexity

The aim of *communication complexity* is to quantify the amount of communication necessary to evaluate a function whose input is distributed among several players. Since its introduction by Yao in 1979 [36], it became a part of complexity theory. It was successfully applied, e.g., in the contexts of VLSI design, circuit lower bounds and lower bounds on data-structures [24].

Two players Alice and Bob are asked to evaluate a Boolean function $f : A \times B \rightarrow \{0, 1\}$ at a given input pair $(a, b) \in A \times B$, where A and B are finite sets. Alice receives the input $a \in A$ and Bob receives the input $b \in B$. Both players know the function f to be evaluated, but none of the players initially has any information about the input of the other player. They have to cooperate in order to compute $f(a, b)$. Both players can perform any kind of computation. It is only the *amount of communication* between them which is limited. In a deterministic protocol, Alice and Bob will exchange bits until one of them is able to correctly compute $f(a, b)$. The *deterministic communication complexity* of the function f is the minimum worst-case number of bits that Alice and Bob have to exchange in order to evaluate f at any given input pair (a, b) .

The function f can be encoded via its *communication matrix* $M = M(f)$ that has $M_{i,j} = f(a_i, b_j)$, where a_i is the i th element of A and b_j is the j th element of B . It is known [35] that a deterministic protocol of complexity k for f yields a decomposition of M as a sum of at most 2^k rank-one 0/1-matrices, implying $\text{rank}_+(M) \leq 2^k$. In order to apply this for constructing extensions, via Theorem 2.6, it is necessary for the slack matrix to be binary.

Let G be a graph. The *stable set polytope* $P_{\text{stab}}(G)$ is the 0/1-polytope in $\mathbb{R}^{V(G)}$ whose vertices are the characteristic vectors of stable sets of G . Because the maximum stable set problem is NP-hard, it is unlikely that a complete description of $P_{\text{stab}}(G)$ will be found for all graphs G . However, some interesting classes of valid inequalities for $P_{\text{stab}}(G)$ are known. For instance, for every clique K of G , the *clique inequality* $\sum_{v \in K} x_v \leq 1$ is valid for $P_{\text{stab}}(G)$. By collecting the slack covectors of each of these inequalities w.r.t. the vertices of $P_{\text{stab}}(G)$, we obtain a 0/1-matrix $M = M(G)$. This matrix is the communication matrix of the so-called *clique vs. stable set problem* (also known as the *clique vs. independent set problem*). Yannakakis [35] found

a $O(\log^2 n)$ complexity deterministic protocol for this communication problem. When the clique inequalities together with the non-negativity inequalities form a complete linear description of $P_{\text{stab}}(G)$, that is, when G is perfect [4], one obtains an extension of size $2^{O(\log^2 n)} = n^{O(\log n)}$ for $P_{\text{stab}}(G)$, where n denotes the number of vertices of G . For general graphs, M is not a slack matrix of $P_{\text{stab}}(G)$ but a proper submatrix. Hence $\text{rank}_+(M)$ only gives a lower bound on the extension complexity of $P_{\text{stab}}(G)$.

Consider again a Boolean function $f : A \times B \rightarrow \{0, 1\}$ and a deterministic protocol for computing f . Fix an input pair (a, b) . Given the transcript of the protocol on input pair (a, b) , each the players can verify independently of the other player if the part of the transcript that he/she is responsible for is correct. The transcript is globally correct if and only if it is correct for *both* of the players. Thus one way to persuade each player that, say, $f(a, b) = 1$ is to give them the transcript for (a, b) . If the complexity of the given deterministic protocol is k , then the players can decide whether $f(a, b) = 1$ or not based on a binary vector of size k (the transcript). This is an example of a nondeterministic protocol. We give a formal definition in the next paragraph.

In a *nondeterministic protocol*, both players are given, in addition to their own input a and b , a proof that is in the form of a binary vector π of size k (the same for both players). The players cannot communicate with each other. Each of them checks the proof π , independently of the other, and outputs one bit, denoted by $V_A(a, \pi)$ or $V_B(b, \pi)$. The result of the protocol is the *AND* of these two bits, that is, $V(a, b, \pi) := V_A(a, \pi) \wedge V_B(b, \pi)$. The aim of the protocol is to prove that $f(a, b) = 1$. In order to be correct, the protocol should be *sound* in the sense that if $f(a, b) = 0$ then $V(a, b, \pi) = 0$ for all proofs π , and *complete* in the sense that if $f(a, b) = 1$ then $V(a, b, \pi) = 1$ for some proof π . The *nondeterministic complexity* of f is the minimum proof-length k in a nondeterministic protocol for f . Letting $M = M(f)$ denote the communication matrix of f , it is well-known that the nondeterministic complexity of f is $\lceil \log(\text{rc}(M)) \rceil$; see e.g. [24].

While Yannakakis proved that the deterministic communication complexity of the clique vs. stable set problem is $O(\log^2 n)$, the exact deterministic and nondeterministic communication complexities of the clique vs. stable set problem are unknown. The best result so far is a lower bound of $\frac{6}{5} \log n - O(1)$ on the nondeterministic complexity obtained by Huang and Sudakov [18]. They also made the following graph-theoretical conjecture that, if true, would improve this bound to $\Omega(\log^2 n)$.

Recall that the *biclique partition number* of a graph G is the minimum number of bicliques (that is, complete bipartite graphs) needed to partition the edge set of G . Recall also that the *chromatic number* $\chi(G)$ of G is the minimum number of parts in a partition of the vertex set of G into stable sets.

Conjecture 4.1 (Huang and Sudakov [18]). *For each integer $k > 0$, there exists a graph G with biclique partition number k and chromatic number at least $2^{c \log^2 k}$, for some constant $c > 0$.*

This conjecture would settle the communication complexity of the clique vs. stable set problem. In terms of extensions, the result of Huang and Sudakov implies a $\Omega(n^{6/5})$ lower bound on the worst-case rectangle covering number of stable set polytopes of graphs with n vertices. Moreover, if true, the Huang–Sudakov conjecture would imply a $n^{\Omega(\log n)}$ lower bound.

5. Lower bounds on the rectangle covering number

In this section, we give lower bounds on the rectangle covering number, and apply them to prove results about the extension complexity of polytopes. Our general strategy is to focus on a specific submatrix of the slack matrix and then use simple structural properties of the support of the submatrix. Although much of the underlying geometry is lost in the process, we can still obtain interesting bounds, which are in some cases tight. Also, we compare the bounds to each other whenever possible and useful. We remark that most of the bounds discussed here are known bounds on the nondeterministic communication complexity of Boolean functions. Nevertheless, they were never studied in the context of polyhedral combinatorics.

5.1. Dimension and face counting

Lemma 5.1. *Suppose the 0/1-matrix M has h distinct rows. Then $\text{rc}(M) \geq \log h$.*

Proof. Since deleting rows does not increase the rectangle covering number, we may assume that M has exactly h rows, which are all distinct.

We proceed by contradiction, i.e., we assume that there is a collection \mathcal{R} of rectangles that covers M and contains less than $\log h$ rectangles. Each such rectangle is of the form $I \times J$, where I is a set of row indices and J is a set of column indices such that $M_{i,j} = 1$ whenever $i \in I$ and $j \in J$.

For every row-index $i \in [h]$, let \mathcal{R}_i denote the set of all rectangles $I \times J$ from \mathcal{R} such that $i \in I$. Since the number of rectangles in \mathcal{R} is less than $\log h$, there are two distinct row-indices i and i' such that \mathcal{R}_i and $\mathcal{R}_{i'}$ are equal. Because the rectangles in \mathcal{R} form a covering of the matrix M , it follows that the i th row and the i' th row of M have the same set of one-entries and are thus equal, a contradiction. \square

Obviously, the same statement holds with “rows” replaced by “columns”. It is well-known [14] that the binary logarithm of the number of faces of a polytope P is a lower bound for the number of facets in an extension of P . This number is in fact also a lower bound on the rectangle covering number of P . Furthermore, the following chain of inequalities holds.

Proposition 5.2. *For a full-dimensional polytope P in \mathbb{R}^d with f faces, we have*

$$d + 1 \leq \log f \leq \text{rc}(P).$$

Proof. The first inequality follows by induction on d . For the second inequality, consider a slack matrix for P with f rows with the i th row being the slack covector of any inequality defining the i th face of P . The support matrix M of this slack matrix has f distinct rows. The conclusion now follows from Lemma 5.1. \square

5.2. The rectangle graph

Consider a real matrix M . The *rectangle graph* of M , denoted by $G(M)$, has the pairs (i, j) such that $M_{i,j} \neq 0$ as vertices. Two vertices (i, j) and (k, ℓ) of $G(M)$ are adjacent if no rectangle contained in $\text{supp}(M)$ can be used to cover both (i, j) and (k, ℓ) . This last condition is equivalent to asking $M_{i,\ell} = 0$ or $M_{k,j} = 0$. The following result will allow us to interpret most lower bounds on $\text{rc}(M)$ in terms of graphs. We omit the proof because it is straightforward.

Lemma 5.3. *For every real matrix M , we have $\text{rc}(M) = \chi(G(M))$.* \square

5.3. Clique number and fooling sets

The *clique number* $\omega(G(M))$ of the rectangle graph of a real matrix M is a lower bound to its chromatic number and hence, by Lemma 5.3, a lower bound on the rectangle covering number of M . A clique in $G(M)$ corresponds to what is known as a *fooling set*: a selection of non-zeros in the matrix such that no two of them induce a rectangle contained in $\text{supp}(M)$. As is customary, we denote the clique number of a graph G by $\omega(G)$. Analogously, we denote the clique number of $G(M)$ by $\omega(M)$. We can also define the clique number $\omega(P)$ of a polytope P as $\omega(M)$, for any non-incidence matrix M of P . The next proposition shows that $\omega(P)$ is well-defined.

Proposition 5.4. *The clique number $\omega(M)$ is independent of the non-incidence matrix M chosen for the polytope P .*

Proof. Consider any non-incidence matrix M for P and a fooling set of size $q := \omega(M)$. It suffices to show that the facet–vertex non-incidence matrix of P has a fooling set of size q as well. By reordering rows and columns of M if necessary, we may assume that the fooling set is $(F_1, G_1), \dots, (F_q, G_q)$. Thus, F_1, \dots, F_q and G_1, \dots, G_q are faces of P satisfying: $F_i \not\supseteq G_i$ for all $i \in [q]$, and $F_i \supseteq G_j$ or $F_j \supseteq G_i$ for all $i, j \in [q]$ with $i \neq j$. We assume that our fooling set is chosen to maximize the number of F_i 's which are facets, plus the number of G_i 's which are vertices.

If every F_i is a facet and every G_i is a vertex, we are done because then our fooling set is contained in the facet vs. vertex submatrix of M . Otherwise, w.l.o.g., there exists an index i such that F_i is not a facet. Let i' be such that $F_{i'}$ is a facet with $F_{i'} \supseteq F_i$ but at the same time $F_{i'} \not\supseteq G_i$. If $i' > q$ then we can replace (F_i, G_i) by $(F_{i'}, G_i)$ in the fooling set, contradicting the choice of the fooling set. Otherwise, $i' \leq q$ and the fooling set contains a pair of the form $(F_{i'}, G_{i'})$. Since $F_{i'} \not\supseteq G_{i'}$, we have $F_i \not\supseteq G_{i'}$. Moreover, $F_{i'} \not\supseteq G_i$ by choice of $F_{i'}$. This contradicts the fact that $(F_i, G_i), \dots, (F_q, G_q)$ is a fooling set. \square

In this section, by constructing large fooling sets, we can give lower bounds on the extension complexity for a number of examples, including cubes (in Section 5.3.2) and the Birkhoff polytope (in Section 5.3.3). But before we do that, we will take a look at the limitations of the clique number as a lower bound for the rectangle covering number. We show that the clique number of the rectangle graph is always $O(d^2)$ for d -dimensional polytopes, and $O(d)$ for simple polytopes.

5.3.1. Limitations of the clique number as a lower bound on the rectangle covering number

Here we give some upper bounds on the sizes of fooling sets. In some cases, these bounds immediately render useless the fooling set approach to obtain a desired lower bound for the rectangle covering number or the extension complexity of a polytope. We start with an easy upper bound based on the number of zeros per row.

Lemma 5.5. *If every row of M contains at most s zeros, then $\omega(M) \leq 2s + 1$.*

Proof. If M has a fooling set of size q , then the submatrix of M induced by the rows and columns of the entries belonging to the fooling set contains at least $q(q - 1)/2$ zeros. In particular, one row of the submatrix has at least $(q - 1)/2$ zeros. By hypothesis, $(q - 1)/2 \leq s$, that is, $q \leq 2s + 1$. \square

If every vertex of a polytope is contained in at most s facets, or if every facet of a polytope contains at most s vertices, then $\omega(P) = O(s)$. In particular, for simple or simplicial polytopes, the fooling set lower bound is within a constant factor of the dimension.

We now give dimensional upper bounds on $\omega(M)$. Dietzfelbinger et al. [7] show the following.

Lemma 5.6 (Dietzfelbinger et al. [7]). *For every field \mathbb{K} and for every 0/1-matrix M , there is no fooling set larger than $\text{rank}_{\mathbb{K}}(M)^2$.*

Although the rank of the non-incidence matrix M of a polytope can be substantially larger than its dimension (this is the case already for polygons), we can nevertheless prove the following by techniques similar to those of [7].

Lemma 5.7. *For every polytope P of dimension d , we have $\omega(P) \leq (d + 1)^2$.*

Proof. Without loss of generality, assume that P is full-dimensional. Let M be a non-incidence matrix for $P = \{x \in \mathbb{R}^d : Ax \leq b\} = \text{conv}(\{v_1, \dots, v_n\})$, and $q := \omega(P)$. By reordering if necessary, we may assume that $(1, 1), \dots, (q, q)$ are the vertices of a maximum clique of the rectangle graph of M . Let $z_i := (b_i, -A_{i,\star}) \in \mathbb{R}^{d+1}$, $i = 1, \dots, q$, and $t_j := (1, v_j) \in \mathbb{R}^{d+1}$, $j = 1, \dots, q$. These vectors have the following properties:

$$\begin{aligned} \langle z_i, t_j \rangle &\geq 0 && \text{for all } i, j = 1, \dots, q; \\ \langle z_i, t_i \rangle &> 0 && \text{for all } i = 1, \dots, q; \\ \langle z_i, t_j \rangle = 0 &\text{ or } \langle z_j, t_i \rangle = 0 && \text{for all } i, j = 1, \dots, q \text{ with } i \neq j. \end{aligned} \tag{6}$$

Now consider the following $2q$ rank-one matrices: $z_i t_i^T$ for $i = 1, \dots, q$ and $t_j z_j^T$ for $j = 1, \dots, q$. Taking the usual inner product for matrices $\langle A, B \rangle = \sum_{i,j} A_{ij} B_{ij} = \text{Tr}(A^T B)$, if $i \neq j$, we have

$$\langle z_i t_i^T, t_j z_j^T \rangle = \text{Tr}(t_i z_i^T t_j z_j^T) = \text{Tr}(z_j^T t_i z_i^T t_j) = (z_j^T t_i)(z_i^T t_j) = \langle z_j, t_i \rangle \langle z_i, t_j \rangle = 0$$

but

$$\langle z_i t_i^T, t_i z_i^T \rangle = \langle z_i, t_i \rangle \langle z_i, t_i \rangle > 0.$$

This implies that the matrices $z_1 t_1^T, \dots, z_q t_q^T$ are linearly independent. Since we have q linearly independent $(d + 1) \times (d + 1)$ -matrices, we conclude that $q \leq (d + 1)^2$. \square

Remark 5.8. From a construction due to Dietzfelbinger et al. [7], we can infer the existence of a d -dimensional polytope P with $\omega(P) = \Omega\left(d^{\frac{\log 4}{\log 3}}\right)$. It is an open question which of these bounds can be improved.

We conclude this subsection with an example where the fooling set or clique number bound is particularly bad. Consider the vertex–facet non-incidence-matrix of a convex polygon with n vertices. This is a $n \times n$ 0/1-matrix M with $M_{i,j} = 0$ if and only if i is equal to j or $j + 1$ modulo n . We have $\omega(M) \leq 5$, by Lemma 5.5, whereas $\text{rc}(M) \geq \log n$ by Proposition 5.2.

5.3.2. The cube

We now apply the fooling set technique to show that d -cubes are “minimal extensions” of themselves. A *combinatorial d -cube* is a polytope whose face lattice is isomorphic to that of the unit cube $[0, 1]^d$. Compare the result of this proposition to the lower bound one can obtain as the binary logarithm of the number of faces, Proposition 5.2, which is $d \log(3) \approx 1.585 d$.

Proposition 5.9. *If P is a combinatorial d -cube then $\text{xc}(P) = 2d$.*

Proof. We obviously have $\omega(P) \leq \text{xc}(P) \leq 2d$. Below, we prove that equality holds throughout by constructing a fooling set of size $2d$, implying $\omega(P) \geq 2d$. Since $\omega(P)$ only depends on the combinatorial type of P , we may assume that $P = [0, 1]^d$. To define our fooling set, we carefully select for each facet-defining inequality of P a corresponding vertex. The vertex corresponding to the inequality $x_i \geq 0$ is the vertex v_i defined as follows:

$$(v_i)_j := \begin{cases} 1 & \text{if } 1 \leq j \leq i \\ 0 & \text{if } i < j \leq d. \end{cases}$$

For the inequality $x_i \leq 1$, take the vertex w_i defined as follows:

$$(w_i)_j := \begin{cases} 0 & \text{if } 1 \leq j \leq i \\ 1 & \text{if } i < j \leq d. \end{cases}$$

One can check that none of the facets of the cube is incident with its corresponding vertex, but that for any two facets at least one of the corresponding vertices is incident to one of them. \square

5.3.3. The Birkhoff polytope

The *n th Birkhoff polytope* is the set of doubly stochastic $n \times n$ matrices or, equivalently, the convex hull of all $n \times n$ permutation matrices. Let P denote the n th Birkhoff polytope. For $n = 1$, P is a point (and $\text{xc}(P)$ is not defined). For $n = 2$, P

is a segment and therefore $xc(P) = \omega(P) = 2$. For $n = 3$, P has 9 facets and 6 vertices, thus $xc(P) \leq 6$. As is easily verified, the facet vs. vertex non-incidence matrix for P has a fooling set of size 6. Hence, $xc(P) = \omega(P) = 6$.

Proposition 5.10. For $n \geq 4$, the extension complexity of the Birkhoff polytope is n^2 .

Proof. The Birkhoff polytope P is defined by the nonnegativity inequalities $x_{i,j} \geq 0$, $i, j \in [n]_0 = \{0, \dots, n-1\}$ (starting indices from zero will be a bit more convenient for our treatment), and the equations $\sum_i x_{i,j} = 1$, $j \in [n_0]$, $\sum_j x_{i,j} = 1$, $i \in [n_0]$. This trivially implies that $xc(P) \leq n^2$. To give a lower bound on the extension complexity, we construct a fooling set of size n^2 in the facet vs. vertex non-incidence matrix of P .

For this, for every inequality $x_{i,j} \geq 0$, we define one vertex $v_{i,j}$ such that $(v_{i,j})_{i,j} > 0$ by giving a permutation π of $[n]_0$. Namely, let $\pi(i) := j$ and $\pi(i+1) := j+1$ (all indices are to be understood modulo n). Moreover, we take $\pi(k) := i+j+1-k$ whenever $k \notin \{i, i+1\}$. This defines a permutation, and thus a vertex $v_{i,j}$ of P .

Now, we show that this family of inequality–vertex pairs is a fooling set. By contradiction, let us assume that for two different inequalities $x_{i,j} \geq 0$ and $x_{i',j'} \geq 0$ we have both $(v_{i,j})_{i',j'} = 1$ and $(v_{i',j'})_{i,j} = 1$. Letting π and π' denote the permutations of $[n]_0$ associated to $v_{i,j}$ and $v_{i',j'}$ respectively, this means that $\pi(i') = j'$ and $\pi'(i) = j$. Because $(i, j) \neq (i', j')$, we conclude that $i+j+1-i' = j'$ or $(i', j') = (i+1, j+1)$, and at the same time $i'+j'+1-i = j$ or $(i, j) = (i'+1, j'+1)$. Because $2 \neq 0$ and $3 \neq 0$ modulo n , these conditions lead to a contradiction. \square

The Birkhoff polytope can also be described as the convex hull of the characteristic vectors of all perfect matchings of a complete bipartite graph with n vertices on each side of the bipartition. The n th bipartite matching polytope is the convex hull of all (not necessarily perfect) matchings of a complete bipartite graph with n vertices on each side. Let P denote the n th bipartite matching polytope. For $n = 1$, P is a segment and $xc(P) = \omega(P) = 2$. For $n = 2$, P has 8 facets and 7 vertices, thus $xc(P) \leq 7$. We leave to the reader to check that $\omega(P) \geq 7$, thus $xc(P) = \omega(P) = 7$. For $n = 3$, P has 15 facets and 34 vertices, thus $xc(P) \leq 15$. It can be checked that P has a fooling set of size 15; hence $xc(P) = \omega(P) = 15$. (To the inequality $x_{i,j} \geq 0$ we associate the matching $\{(i, j), (i+1, j+1)\}$ if $i \neq j$ and $\{(i, j), (k, \ell), (\ell, k)\}$ if $i = j$ and k, ℓ denote the two other elements of $\{0, 1, 2\}$). To the inequality $\sum_j x_{i,j} \leq 1$ we associate the matching $\{(i+1, i+1), (i+2, i)\}$. To the inequality $\sum_i x_{i,j} \leq 1$ we associate the matching $\{(j+1, j+1), (j, j+2)\}$.) As before, all computations are done modulo $n = 3$. For $n \geq 4$, an argument similar to the one in the proof of Proposition 5.10 shows the following.

Proposition 5.11. For $n \geq 4$, the bipartite matching polytope has extension complexity $n^2 + 2n$.

Proof (sketch). Again, the given number is a trivial upper bound. To construct a fooling set of this size, for the nonnegativity inequalities, we take the perfect matchings constructed in the proof of the previous proposition. In addition, for the inequalities $\sum_i x_{i,j} \leq 1$, we take the vertex w_j with $(w_j)_{k+1,k} = 1$ for all $k \neq j$; for the inequalities $\sum_j x_{i,j} \leq 1$, we take the vertex u_i with $(u_i)_{k,k+1} = 1$ for all $k \neq i$. \square

5.4. Independence ratio, rectangle sizes, and generalized fooling sets

Denote by $\alpha(G)$ the maximum cardinality of an independent vertex set in G . The number $\frac{|G|}{\alpha(G)}$ is a lower bound on the chromatic number ($|G|$ stands for the number of vertices in G). Moreover, taking induced subgraphs may improve the bound: The number

$$\iota(G) := \max_{\substack{U \subseteq V(G) \\ U \neq \emptyset}} \frac{|U|}{\alpha(G[U])},$$

where $G[U]$ denotes the subgraph of G induced by the vertices in U is sometimes called the *independence ratio* of G , and is also a lower bound on the chromatic number of G .

In the context of the rectangle covering number, these bounds are known under different names. The following lemmas will make that clear. We leave the easy proofs to the reader.

Lemma 5.12. For a real matrix M , the maximum number of entries in a rectangle contained in $\text{supp}(M)$ equals $\alpha(G(M))$. \square

The concept of *generalized fooling sets* has been proposed by Dietzfelbinger et al. [7] as a lower bound on the nondeterministic communication complexity of a Boolean function. In the graph-coloring terminology, a k -fooling set is an induced subgraph H of $G(M)$ for which $\alpha(H) \leq k$ holds. If a k -fooling set on s entries of $\text{supp}(M)$ can be found, then, clearly, s/k is a lower bound on the rectangle covering number. Their *generalized fooling set lower bound* on the rectangle covering number is then just the supremum of all these fractions s/k , and it coincides with the independence ratio.

Whereas the fooling set/cliquote lower bound can be arbitrarily bad (see the example at the end of Section 5.3.1), this is not the case for the independence ratio. The following is true for general graphs (Lovász [26, Theorem 7]). It has been rediscovered by Dietzfelbinger et al. [7] for the rectangle graphs of 0/1-matrices. Lovász' argument (specialized from the more general setting of hypergraph coverings he in fact considers) proceeds by analyzing the following greedy heuristic for coloring G : pick a maximum stable set S_1 in G , then a maximum stable set S_2 in $G - S_1$, and so on. Denoting by w_k the maximum

number of nodes in G containing no stable set of size larger than k and by t_i the number of stable sets of exactly size i produced by the greedy heuristic, one finds that $\sum_{i=1}^k it_i \leq w_k$ holds for all $k = 1, \dots, \alpha = \alpha(G)$. Adding up these inequalities scaled by $1/k(k+1)$ for $k < \alpha$ and by $1/\alpha$ for $k = \alpha$ one obtains the upper bound $\sum_{i=1}^{\alpha-1} \frac{1}{i(i+1)} w_i + \frac{1}{\alpha} w_\alpha = O(\iota(G) \log(|G|))$ on the number $\sum_{i=1}^{\alpha} t_i$ of stable sets generated by the greedy procedure.

Lemma 5.13. For all graphs G , we have $\chi(G) = O(\iota(G) \log(|G|))$. \square

5.4.1. The cube revisited

We now give an alternative proof of Proposition 5.9, based on the independence number of $G(P)$ instead of the clique number.

2nd Proof of Proposition 5.9. The maximal rectangles contained in $\text{supp}(M)$ of the facet vs. vertex non-incidence matrix M are among those of the form $I \times J$, where I is a collection of facets of P that does not contain a pair of opposite facets and $J = J(I)$ is the set of those vertices of P that belong to none of the facets in I . If q denotes the cardinality of I , the size of such a rectangle is precisely $q \cdot 2^{d-q}$. Indeed, q of the coordinates of the vertices of J are determined by J , while the other $d - q$ coordinates are free, which implies $|J| = 2^{d-q}$.

As is easily verified, $q2^{d-q}$ is maximum for $q \in \{1, 2\}$. Hence, the maximum size of a rectangle in $\text{supp}(M)$ is thus 2^{d-1} . It follows that $2d$ rectangles are necessary to cover all the ones in M , because there are $2d \cdot 2^{d-1}$ ones in M . \square

5.4.2. Polytopes with few vertices on every facet

As we have seen in Lemma 5.5 above, the chances of obtaining a good lower bound based on fooling sets are poor if the matrix has few zeros in each row. A similar result is true for generalized fooling sets.

Lemma 5.14. Let M be a real matrix with at most s zeros per row and let U be a k -fooling set in M . Then $|U|/k$, namely, the lower bound on $\text{rc}(M)$ given by U , is at most $2k^2s$.

Proof. Let U be a set of vertices of $G(M)$ which is a k -fooling set, i.e. the subgraph H of $G(M)$ induced by U has $\alpha(H) \leq k$. By Turán’s theorem, we have

$$|E(H)| \geq \frac{|U|^2}{2k}.$$

Each edge of H links two entries of M in distinct rows and columns such that at least one of the two other entries in the 2×2 rectangle spanned by these entries is zero. We say that this (these) zero(s) are *responsible* for the edge. Thus for each edge of H there is at least one responsible zero (and at most two responsible zeros), located in the rectangle spanned by U . Since every row or column contains at most k elements of U , each zero in the rectangle spanned by U is responsible for at most k^2 edges of H . Hence, the number of edges in H is at most k^2 times the number of zeros in the rectangle spanned by U . Thus, since U covers at most $|U|$ rows, which together contain at most $|U|s$ zeros, we have

$$|E(H)| \leq k^2|U|s,$$

and we conclude that

$$\frac{|U|}{k} \leq 2k^2s. \quad \square$$

5.5. Fractional chromatic number

Another well-studied lower bound for the chromatic number of a graph is the *fractional* chromatic number. This number can be defined in several equivalent ways. The most convenient is probably the following: $\chi^*(G)$ is the solution of the following linear program:

$$\min \sum_S x_S \tag{7a}$$

$$\sum_{S:v \in S} x_S \geq 1 \quad \text{for every vertex } v \text{ of } G \tag{7b}$$

$$x_S \geq 0 \quad \text{for every stable set } S \text{ of } G, \tag{7c}$$

where the sums extend over all stable sets S of the graph G . In the context of nondeterministic communication complexity, this bound is well-known; see for example [23] or [24] and the references therein. In particular, it is known that $\chi(M) = O(\chi^*(G(M) \log |G(M)|))$. The following known fact from graph coloring (see e.g. [30, p. 1096]) improves this slightly.

Lemma 5.15. For all graphs G , it is true that $\chi(G) \leq (1 + \ln \alpha(G))\chi^*(G)$. \square

Clearly, if integrality conditions are imposed on the variables in the linear program (7) defining $\chi^*(G)$, then the optimal value is the chromatic number $\chi(G)$.

5.6. Neighborly d -polytopes with $\Theta(d^2)$ vertices

Suppose d is even and let $k := d/2$. Consider a neighborly polytope P in \mathbb{R}^d with vertex set $V := \text{vert}(P)$. Let $n := |V|$ denote the number of vertices of P . We assume that $n = \Theta(d^2) = \Theta(k^2)$, that is, k and d are both $\Theta(\sqrt{n})$.

Consider a non-incidence matrix that has one row for each facet, as well as one row for each $(k-1)$ -face, and one row per vertex. Let M denote the submatrix induced on the rows that correspond to the $(k-1)$ -faces. Thus the rows of M are indexed by the k -subsets of V (that is, the vertex sets of $(k-1)$ -faces of P), and the columns of M are indexed by the elements of V (that is, the vertices of P). The entry of M for k -set $F \subseteq V$ and vertex $v \in V$ is 1 if $v \notin F$ and 0 otherwise. It follows from Lemma 2.10 and the monotonicity of the rectangle covering number on submatrices that $\text{xc}(P) \geq \text{rc}(M)$.

Proposition 5.16. *Let n, k be positive integers and M denote the $\binom{n}{k} \times n$ matrix defined above. Then $\text{rc}(M) = \Omega\left(\min\left\{n, \frac{(k+1)(k+2)}{2}\right\}\right)$.*

Proof. The maximal rectangles of M are of the form $I \times J$, where J is a subset of V of size at most $n-k$ and $I = I(J)$ is the collection of all k -sets F such that $F \subseteq V - J$. We define a *cover* as a collection $\mathcal{J} = \{J_1, \dots, J_t\}$ of subsets of V , each of size at most $n-k$, such that the corresponding rectangles cover M . This last condition can be restated as follows: for each pair (F, v) with $v \notin F$ there exists an index $\ell \in [t]$ with $v \in J_\ell$ and $F \subseteq V - J_\ell$.

Consider a cover $\mathcal{J} = \{J_1, \dots, J_t\}$. For a vertex $v \in V$, consider the collection $\mathcal{J}(v)$ of sets in \mathcal{J} that contain v . There are two cases. First, it could be that $\{v\} \in \mathcal{J}(v)$. In this case, by removing v from the other sets in $\mathcal{J}(v)$, we may assume that $\mathcal{J}(v) = \{\{v\}\}$. That is, the other sets of \mathcal{J} do not contain v . Otherwise, all sets in $\mathcal{J}(v)$ contain at least one element distinct from v . In this case, it should not be possible to find a k -set F contained in $V - \{v\}$ that meets all sets in $\mathcal{J}(v)$, because otherwise \mathcal{J} would not cover the pair (F, v) . In particular, there are at least $k+1$ sets in $\mathcal{J}(v)$.

Let X denote the set of the vertices v such that $\{v\} \in \mathcal{J}$, and let $s := |X|$. Pick distinct vertices v_1, v_2, \dots, v_{k+1} in $V - X$. This can be assumed to be possible because otherwise $V - X$ has at most k vertices. If $V - X$ is empty, then clearly the cover is trivial and $t \geq n$. Else $V - X$ contains an element v and, because all sets in $\mathcal{J}(v)$ are contained in $V - X$, one can easily find a k -set $F \subseteq V - \{v\}$ such that (F, v) is not covered by \mathcal{J} .

For all $\ell \in [k+1]$, it holds that the number of sets in \mathcal{J} that contain v_ℓ and none of the vertices $v_1, \dots, v_{\ell-1}$ is at least $k+2-\ell$. Otherwise, we could find $k+1-\ell$ elements in $V - \{v_1, \dots, v_\ell\}$ that, added to $\{v_1, \dots, v_{\ell-1}\}$, would define a k -set $F \subseteq V - \{v_\ell\}$ such that (F, v_ℓ) is not covered by \mathcal{J} .

Now, the number of sets in \mathcal{J} is at least

$$s + (k+1) + k + (k-1) + \dots + 1 = s + \frac{(k+1)(k+2)}{2} \geq \frac{(k+1)(k+2)}{2}.$$

The first term counts the number of singletons in \mathcal{J} , the second the number of sets containing v_1 , the third the number of sets containing v_2 but not v_1 , and so on. \square

When $n = \Theta(k^2)$, it follows from Proposition 5.16 that the minimum number of rectangles needed to cover M is $\Theta(n)$. Therefore, in this case the minimum number of facets in an extension of P is $\Theta(n)$. Note that the binary logarithm of the total number of faces is in this case $\Theta(d \log n) = \Theta(d \log d)$. In conclusion, neighborly d -polytopes with $\Theta(d^2)$ vertices give a family of polytopes such that the rectangle covering number is super-linear in both the “dimension” and the “binary logarithm of the number of faces” bounds.

6. Concluding remarks

We conclude this paper with some open problems. We begin by reiterating the first open problem in Yannakakis's paper [35]. Bounding the rectangle covering number seems to be the currently best available approach to bound the extension complexity of specific polytopes. (The new bound on the nonnegative rank proposed by Gillis and Glineur [13] does not lead to improvements for slack matrices.) Can one find other bounds? In particular, for a given polytope P , consider the smallest number of facets of a polytope Q into whose face lattice the face-lattice of P can be embedded meet-faithfully. Does this number improve substantially on the rectangle covering bound? Or is it always bounded by a polynomial in the rectangle covering bound for P ?

For some polytopes the rectangle covering number even yields optimal or near-optimal lower bounds on the extension complexity as we have demonstrated in this paper for cubes, Birkhoff polytopes, and neighborly d -polytopes with $\Theta(d^2)$ vertices. However, for other polytopes such as the perfect matching polytope, the rectangle covering bound seems rather useless. For instance, for sufficiently irregular n -gons in \mathbb{R}^2 , the extension complexity is bounded from below by $\Omega(\sqrt{n})$, while the rectangle covering number is $O(\log n)$ [12].

In all the examples we have found so far, the rectangle covering number is always polynomial in d and $\log n$, where d denotes the dimension and n the number of vertices. Can one find polytopes for which the rectangle covering number is super-polynomial in d and $\log n$?

As mentioned above, it is true that $\omega(M) \leq (\text{rank} A + 1)^2$ for every matrix A whose support is M , and Dietzfelbinger et al. [7] constructed a family of matrices M with $\text{rank}(M)^{\log 4 / \log 3} \leq \omega(M)$. Which of the two bounds can be improved?

In view of our result on the extension complexity of neighborly d -polytopes with n vertices, it is natural to ask whether the bound $\Omega(d^2)$ can be improved for $d = o(\sqrt{n})$.

Finally, in the absence of better lower bounds on the extension-complexity, it would be interesting to know the rectangle covering bound of the perfect matching polytope. As mentioned above, it is $O(n^4)$. Even a small improvement on the trivial lower bound $\Omega(n^2)$ would be interesting.

7. Final note

After submitting this paper, our third open problem was solved by Fiorini et al. [11]. They prove among other things that the rectangle covering number of the cut polytope is super-polynomial in the dimension and logarithm of the number of vertices of the polytope.

Acknowledgments

We are grateful to the referees whose comments lead to significant improvements in the presentation of the material.

References

- [1] E. Balas, Disjunctive programming and a hierarchy of relaxations for discrete optimization problems, *SIAM J. Algebr. Discrete Methods* 6 (1985) 466–486.
- [2] LeRoy B. Beasley, Thomas J. Laffey, Real rank versus nonnegative rank, *Linear Algebra Appl.* 431 (12) (2009) 2330–2335.
- [3] A. Ben-Tal, A. Nemirovski, On polyhedral approximations of the second-order cone, *Math. Oper. Res.* 26 (2) (2001) 193–205.
- [4] V. Chvátal, On certain polytopes associated with graphs, *J. Combin. Theory Ser. B* 18 (1975) 138–154.
- [5] M. Conforti, G. Cornuéjols, G. Zambelli, Extended formulations in combinatorial optimization, *4OR* 8 (1) (2010) 1–48.
- [6] D. de Caen, D.A. Gregory, N.J. Pullman, The Boolean rank of zero-one matrices, in: *Proceedings of the Third Caribbean Conference on Combinatorics and Computing* (Bridgetown, 1981), Cave Hill Campus, Barbados, 1981. Univ. West Indies, pp. 169–173.
- [7] M. Dietzfelbinger, J. Hromkovič, G. Schnitger, A comparison of two lower-bound methods for communication complexity, *Theoret. Comput. Sci.* 168 (1) (1996) 39–51. 19th International Symposium on Mathematical Foundations of Computer Science (Košice, 1994).
- [8] Faun C.C. Doherty, J. Richard Lundgren, Daluss J. Siewert, Biclique covers and partitions of bipartite graphs and related matrix ranks of $\{0, 1\}$ -matrices, in: *Proceedings of the Thirtieth Southeastern International Conference on Combinatorics, Graph Theory, and Computing* (Boca Raton, FL, 1999), volume 136, 1999, pp. 73–96.
- [9] J. Edmonds, Maximum matching and a polyhedron with 0, 1 vertices, *J. Res. Natl. Bur. Stand.* 69B (1965) 125–130.
- [10] Y. Faenza, S. Fiorini, R. Grappe, H.R. Tiwary, Extended formulations, non-negative factorizations and randomized communication protocols. [arXiv:1105.4127](https://arxiv.org/abs/1105.4127), 2011.
- [11] S. Fiorini, S. Massar, S. Pokutta, H.R. Tiwary, R. de Wolf, Linear vs. semidefinite extended formulations: exponential separation and strong lower bounds, in: *Proc. STOC 2012*, 2012 pp. 95–106.
- [12] S. Fiorini, T. Rothvoß, H.R. Tiwary, Extended formulations for polygons. [arXiv:1107.0371](https://arxiv.org/abs/1107.0371), 2011.
- [13] N. Gillis, F. Glineur, On the geometric interpretation of the nonnegative rank. [arXiv:1009.0880](https://arxiv.org/abs/1009.0880), 2010.
- [14] M.X. Goemans, Smallest compact formulation for the permutahedron. Manuscript, 2009.
- [15] João Gouveia, Pablo A. Parrilo, Rekha Thomas, Lifts of convex sets and cone factorizations, *Math. Oper. Res.* [arXiv:1111.3164](https://arxiv.org/abs/1111.3164), 2012+ (in press).
- [16] D.A. Gregory, N.J. Pullman, Semiring rank: boolean rank and nonnegative rank factorizations, *J. Comb. Inf. Syst. Sci.* 8 (3) (1983) 223–233.
- [17] M. Habib, L. Nourine, O. Raynaud, E. Thierry, Computational aspects of the 2-dimension of partially ordered sets, *Theoret. Comput. Sci.* 312 (2–3) (2004) 401–431.
- [18] H. Huang, B. Sudakov, A counterexample to the Alon–Saks–Seymour conjecture and related problems. [arXiv:1002.4687](https://arxiv.org/abs/1002.4687), 2010.
- [19] V. Kaibel, Extended formulations in combinatorial optimization, *Optima* 85 (2011) 2–7.
- [20] V. Kaibel, A. Loos, Finding descriptions of polytopes via extended formulations and liftings. Chap. 4 of: *Combinatorial Optimization - ISCO2010: Recent Progress*. Ridha Mahjoub (ed.) Wiley-ISTE, 2011 (in press). [arXiv:1109.0815](https://arxiv.org/abs/1109.0815).
- [21] V. Kaibel, K. Pashkovich, Constructing extended formulations from reflection relations, in: *Proc. IPCO 2011*, 2011, pp. 287–300.
- [22] V. Kaibel, K. Pashkovich, D.O. Theis, Symmetry matters for the sizes of extended formulations, in: *Proc. IPCO 2010*, 2010, pp. 135–148.
- [23] M. Karchmer, E. Kushilevitz, N. Nisan, Fractional covers and communication complexity, *SIAM J. Discrete Math.* 8 (1) (1995) 76–92.
- [24] E. Kushilevitz, N. Nisan, *Communication Complexity*, Cambridge University Press, Cambridge, 1997.
- [25] Cony M. Lau, Thomas L. Markham, Factorization of nonnegative matrices. II, *Linear Algebra Appl.* 20 (1) (1978) 51–56.
- [26] L. Lovász, On the ratio of optimal integral and fractional covers, *Discrete Math.* 13 (4) (1975) 383–390.
- [27] R.K. Martin, Using separation algorithms to generate mixed integer model reformulations, *Oper. Res. Lett.* 10 (3) (1991) 119–128.
- [28] K. Pashkovich, Symmetry in extended formulations of the permutahedron. [arXiv:0912.3446](https://arxiv.org/abs/0912.3446), 2009.
- [29] T. Rothvoß, Some 0/1 polytopes need exponential size extended formulations. [arXiv:1105.0036](https://arxiv.org/abs/1105.0036), 2011.
- [30] A. Schrijver, *Combinatorial Optimization. Polyhedra and Efficiency*. Vol. B, in: *Algorithms and Combinatorics*, vol. 24, Springer-Verlag, Berlin, 2003, Matroids, trees, stable sets, Chapters 39–69.
- [31] P. Turán, On an extremal problem in graph theory, *Mat. és Fizikiai Lapok* 48 (1941) 436–452.
- [32] Z. Tuza, Covering of graphs by complete bipartite subgraphs: complexity of 0–1 matrices, *Combinatorica* 4 (1984) 111–116.
- [33] F. Vanderbeck, L.A. Wolsey, Reformulation and decomposition of integer programs, in: M. Jünger, et al. (Eds.), *50 Years of Integer Programming 1958–2008*, Springer, 2010, pp. 431–502.
- [34] S.A. Vavasis, On the complexity of nonnegative matrix factorization, *SIAM J. Optim.* 20 (3) (2009) 1364–1377.
- [35] M. Yannakakis, Expressing combinatorial optimization problems by linear programs, *J. Comput. System Sci.* 43 (3) (1991) 441–466.
- [36] A.C. Yao, Some complexity questions related to distributed computing, in: *Proc. STOC 1979*, 1979, pp. 209–213.
- [37] G.M. Ziegler, *Lectures on Polytopes*, in: *Graduate Texts in Mathematics*, vol. 152, Springer-Verlag, Berlin, 1995.