Topics in Theoretical Computer Science	May 11, 2015

Lecture 12: Lower bounds on extended formulations

Lecturer: Yuri Faenza *Scribes: Jakub Tarnawski*

1 Motivation

In the course we have seen many combinatorial problems which can be solved using linear programming. The first step of this approach is to formulate the problem as a linear program, i.e., describe the convex hull of the characteristic vectors of feasible solutions via linear inequalities. Then solving the problem amounts to solving the linear program.

One example is the problem of finding a maximum-weight spanning tree in an n-vertex *complete* graph (V, E) with given weights on edges. As seen in previous lectures, the *spanning tree polytope*, i.e. the convex hull of characteristic vectors of spanning trees, is the following:

$$P_{st}(n) = \{x \in \mathbb{R}^{|E|} : x \geq 0, \quad x(E) = n-1, \quad x(E(S)) \leq |S|-1 \text{ for all } S \subseteq V, |S| \geq 2\}.$$

A bottleneck of this approach is that we have described the polytope using exponentially many inequalities. Why is this a problem? First, it is of theoretical interest to have low description complexity. Second, even though we can still optimize over the polytope in polynomial time using the ellipsoid algorithm and an efficient separation oracle, the ellipsoid algorithm is very slow and not practical. We would prefer to use algorithms which both have theoretical guarantees and are fast in practice, such as interior point methods. However, then the runtime crucially depends on the number of constraints.

This motivates the following question: can we describe this polytope using a number of inequalities which is polynomial in n?

It turns out that the answer is *yes*, and we will present such an efficient formulation. But, more interestingly, we will also show that some polytopes do not admit any such compact representation.

1.1 Spanning tree polytope

The first thing that we should probably try is to argue that most of the constraints in our current description of $P_{st}(n)$ are redundant. However, this is not the case.

Fact 1 *For any subset $S \subseteq V$ with $3 \leq |S| \leq n-3$, the polytope $P'_{st}(n)$ obtained by removing the contraint $x(E(S)) \leq |S|-1$ from the description above is strictly larger than $P_{st}(n)$.*

Proof We will exhibit a point $x \in P'_{st}(n) \setminus P_{st}(n)$. Let C_S be any cycle containing all vertices of S, and let $C_{\overline{S}}$ be any cycle containing all the other vertices. We obtain x by putting value 1 on each edge of C_S and value $\frac{n-|S|-1}{n-|S|}$ on each edge of $C_{\overline{S}}$. (Put 0 everywhere else.) Then the constraint $x(E(S)) \leq |S|-1$ is violated, as $x(E(S)) = x(C_S) = |S|$, and thus $x \notin P_{st}(n)$. However, $x \in P'_{st}(n)$:

- $x(E) = |S| + \frac{n-|S|-1}{n-|S|}(n-|S|) = n-1$,
- for any $S' \neq S$:
  - if $S \not\subseteq S'$, then $S' \cap S$ induces a forest and thus
    $$x(E(S')) = x(E(S' \cap S)) + x(E(S' \setminus S)) \leq |S' \cap S| - 1 + |S' \setminus S| \frac{n-|S|-1}{n-|S|} \leq |S' \cap S| - 1 + |S' \setminus S| = |S'| - 1,$$
  - if $S \subseteq S'$, then $S' \setminus S$ induces a forest (since $S' \neq V$) and
    $$x(E(S')) = x(E(S' \cap S)) + x(E(S' \setminus S)) \leq |S| + (|S' \setminus S| - 1) \frac{n-|S|-1}{n-|S|} \leq |S| + |S' \setminus S| - 1 = |S'| - 1.$$

■

Fact 1 implies that there are only polynomially many inequalities (with $|S| < 3$ or $n-|S| < 3$) that we could possibly remove. (It is easy to check that actually no inequality is safe to remove.) This shows that the polytope (which is $(|E|-1)$-dimensional) has exponentially many proper, $(|E|-2)$-dimensional faces, and each of them needs "its own" constraint. Therefore we cannot hope to obtain a polynomial-sized description of the spanning tree polytope – if we stay in the space $\mathbb{R}^{|E|}$.

But what if we were able to describe our polytope as a projection of another polytope in a higher-dimensional space? While this might sound cryptic as first, note that it just corresponds to introducing extra variables which help us somehow control the relationships between existing ones. This motivates the following definition.

2 Extended formulations

Definition 2 *Let $P \subseteq \mathbb{R}^n$ be a polytope.*[1] *Then a polytope $Q \subseteq \mathbb{R}^{n+p}$, described as*

$$Q = \left\{(x, y) \in \mathbb{R}^n \times \mathbb{R}^p : E^{=}x + F^{=}y = d^{=},\ E^{\leq}x + F^{\leq}y \leq d^{\leq}\right\}$$

is called an extended formulation for P *iff*

$$\pi_x(Q) = P, \qquad \textit{that is,} \qquad P = \{x \in \mathbb{R}^n : (\exists y \in \mathbb{R}^p)\ (x, y) \in Q\}.$$

In other words, P is the orthogonal projection of Q onto the subspace $\mathbb{R}^n \times \{(0, ..., 0)\}$.[2]

Our objective is obtaining *compact* extended formulations (or proving that they do not exist). When will we say that a formulation is small?

Definition 3 *Given P and Q as above, define the size size(Q) of the extended formulation Q to be the number of inequalities in its description (that is, the number of rows of the matrix $E^{\leq}$).*

Remark We care only about inequalities and not equalities, since we can always remove the equalities without increasing the number of inequalities.[3] Now we can also see why we distinguished inequalities and equalities in the description of the polytope Q.

Definition 4 *We define the* extension complexity *of a polytope P to be*

$$xc(P) = \min\{size(Q) : Q \textit{ is an extended formulation for } P\}.$$

The main question that we will consider is the following:

Question: Given a polytope P, is $xc(P)$ polynomial in n?

2.1 Spanning tree polytope revisited

Now we present the polynomial-sized extended formulation for the spanning tree polytope $P_{st}(n)$ (due to Kipp Martin, 1991 [Mar91]):

[1] For the spanning tree polytope, n is the number of vertices, and the dimension of the space is $|E|$. For general polytopes we will denote the dimension of the space by n.

[2] With an abuse of notation, we will denote by Q both a polytope and its linear description. In the literature Q is usually called an *extension* and a linear description of Q is called *extended formulation*.

[3] Indeed, we can take an equality, choose a variable with nonzero coefficient, compute an expression for it in terms of the other variables, and remove it completely by plugging this expression in its place in all other equalities and inequalities.

$$
\begin{aligned}
Q = \{(x,y) \in \mathbb{R}^{\binom{n}{2}} \times \mathbb{R}^{n(n-1)(n-2)} :& \\
& x_{\{v,w\}} - y_{v,w,u} - y_{w,v,u} = 0 && \text{for all distinct } u,v,w \in V, \\
& x_{\{v,w\}} + \sum_{u \in V \setminus \{v,w\}} y_{v,u,w} = 1 && \text{for all distinct } v,w \in V, \\
& x(E) = n-1, \\
& x, y \geq 0\}.
\end{aligned}
$$

Note that the pairs $\{v,w\}$ are unordered, while the triples u,v,w are ordered.

Exercise 5 *Show that Q is an extended formulation for P, that is, $P = \pi_x(Q)$.*

Proof We will prove the inclusion $\subseteq$. Let $x \in P$; we need to produce a $y \in \mathbb{R}^{n(n-1)(n-2)}$ such that $(x,y) \in Q$. WLOG we can assume that x is an extreme point of P – an indicator vector of a spanning tree T of G. Then we put

$$
y_{v,w,u} = \begin{cases} 1 & \text{if } (v,w) \in T \text{ and } u \text{ is on } w\text{'s side of the edge } (v,w) \text{ in } T, \\ 0 & \text{otherwise.} \end{cases}
$$

It is easy to verify that all constraints are satisfied.

We leave the trickier inclusion $P \supseteq \pi_x(Q)$ to the reader. ■

Clearly $\text{size}(Q) = \Theta(n^3)$ and thus $\text{xc}(P_{st}(n)) = \mathcal{O}(n^3) = \text{poly}(n)$.

Open question: is $\text{xc}(P_{st}(n)) = \Theta(n^3)$? (That is, is there a smaller extended formulation?)

3 A general approach

Even though extended formulations are not a new idea, for a long time only upper bounds (and some conditional lower bounds) were known. However, in recent years exponential lower bounds have been proved for a number of 0/1 polytopes. These developments are the product of a unified approach to extension complexity which we will present now. We also show a lower bound for one particular polytope: the correlation polytope.

3.1 The slack matrix and its nonnegative rank

The first step is to associate with P a matrix called the *slack matrix of P*, from which we will be able to extract information about the structure of the extended formulations of P.

Recall that any polytope can be written in two ways: as

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

for some matrix $A \in \mathbb{R}^{m \times n}$, with m being the number of constraints, or as

$$P = \text{conv}\{v_1, ..., v_d\}$$

with $v_1, ..., v_d \in \mathbb{R}^n$.

Definition 6 *Given both of these representations, the* slack matrix of P *is defined as a matrix $S \in \mathbb{R}^{m \times d}$ with*

$$S_{ij} = b_i - A_i v_j,$$

where A_i is the i-th row of A.

Since $b_i - A_i v_j$ is the slack of the i-th constraint at $v_j \in P$, we have $S_{ij} \geq 0$. Note that S is not unique for P, as it depends on the choice of $A, b, v_1, ..., v_d$.

Example 7 *Let $G = (V, E)$ be a complete graph on n vertices. A* perfect matching *is a matching covering all vertices. We can describe the* perfect matching polytope *of G as*

$$M(n) = \left\{ x \in \mathbb{R}^{|E|} : x \geq 0, \quad x(\delta(v)) = 1 \text{ for each } v \in V, \quad x(\delta(U)) \geq 1 \text{ for each odd set } U \subseteq V \right\}$$

and as

$$M(n) = conv\{\mathbb{1}_M : M \text{ is a perfect matching in } G\}.$$

Let S be the slack matrix of $M(n)$. Then the columns of S are indexed by perfect matchings M of G (and by edges e, but we ignore those), and the rows corresponding to inequalities $x(\delta(U) \geq 1$ are indexed by odd sets U. The entry of the slack matrix in column M and row corresponding to U is

$$S_{U,M} = |M \cap \delta(U)| - 1 \geq 0.$$

To say how S can be useful for us, we need another definition.

Definition 8 *Given any nonnegative matrix $S \in \mathbb{R}^{m \times d}_{\geq 0}$, we say that a pair of matrices (T, U) is a* rank-r nonnegative factorization *of S if:*

$$T \in \mathbb{R}^{m \times r}_{\geq 0}, \quad U \in \mathbb{R}^{r \times d}_{\geq 0}, \quad S = TU.$$

We define the nonnegative rank of S *as*

$$rk_+(S) = \min\{r : S \text{ has a rank-}r \text{ nonnegative factorization}\}.$$

Remark The name comes from the fact that if we skip "nonnegative" and "$_{\geq 0}$" everywhere in Definition 8, we will get an equivalent definition of the rank of a matrix.

3.2 Yannakakis' Theorem

Now we can state the fundamental result which explains our interest in slack matrices.

Theorem 9 (Yannakakis' Factorization Theorem [Yan88]) *Let P be a polytope and S its slack matrix. Suppose* $\dim(P) \geq 1$. *Then*

$$xc(P) = rk_+(P).$$

We need a few lemmas for the proof:

Exercise 10 *If S' is a minor of S, then $rk_+(S') \leq rk_+(S)$.*

Proof Given a rank-r factorization $S = TU$ and a minor $S' = S[X \times Y]$, factorize $S' = T[X \times [r]][[r] \times Y]$. ■

Exercise 11 *For a matrix $S \in \mathbb{R}^{m \times d}_{\geq 0}$ we have $rk_+(S) \leq \min(m, d)$. Moreover, $rk_+(S)$ is at most the number of nonzero rows of S.*

Proof We have the following factorizations: $S = I_{m \times m} S = S I_{d \times d}$. For the second part, suppose that the nonzero rows of S are the first r rows; then write $S = JS'$, where $J \in \mathbb{R}^{m \times r}$ is the identity matrix $I_{r \times r}$ padded with $m - r$ zero rows and $S' = S[[r] \times [m]]$ is the nonzero part of S. ■

Lemma 12 *Let $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ be a polytope with $\dim(P) \geq 1$. If the inequality $cx \leq \delta$ is valid for P,*[4] *then it can be written as a nonnegative combination of the constraints $Ax \leq b$, i.e.*

$$(\exists y \geq 0)\ yA = c,\ yb = \delta.$$

Proof idea: This statement reminds us of Farkas' lemma and LP duality. However, those results would only give us $cx \leq \delta'$ for some $\delta' \leq \delta$, instead of $cx \leq \delta$. We need the assumption that $\dim(P) \geq 1$ to fix this.

Proof

Step 1: we prove that the inequality $cx \leq \delta'$ is a nonnegative combination of the rows of $Ax \leq b$, for some $\delta' \leq \delta$.

LP duality gives us that

$$\max\{cx : Ax \leq b\} = \min\{yb : yA = c, y \geq 0\},$$

since the primal program is feasible (as $\dim(P) \geq 1$) and bounded (as $cx \leq \delta$ is valid[5]). Let δ' be the optimum primal-dual value; we have that $\delta' \leq \delta$ (again, as $cx \leq \delta$ is valid). Furthermore let $y^\star$ be an optimum dual solution. Therefore:

$$y^\star \geq 0, \quad y^\star A = c, \quad y^\star b = \delta'.$$

This is the sought combination.

Step 2: we show that $0x \leq 1$ is also a nonnegative combination of $Ax \leq b$.

Since $\dim(P) \geq 1$, at least one of the inequalities in $Ax \leq b$ is not always tight.[6] That is, for some i,

$$\min\{A_i x : x \in P\} = b'_i < b_i,$$

where A_i is the i-th row of A.[7] Therefore the inequality $A_i x \geq b'_i$ is valid, and by Step 1 it can be obtained as a nonnegative combination of $Ax \leq b$.[8] So the following inequalities can be so obtained: $-A_i x \leq -b'_i$, $A_i x \leq b_i$, and in consequence also their sum: $(A_i - A_i)x \leq b_i - b'_i$, which after scaling yields $0x \leq 1$.

Step 3: now we can take $cx \leq \delta'$ with multiplier 1, $0x \leq 1$ with multiplier $\delta - \delta' \geq 0$, and thus get $cx \leq \delta$. ■

Exercise 13 *Let*

$$P = \{x \in \mathbb{R}^n : Ax \leq b\} = conv\{v_1, ..., v_d\}$$

be a polytope with $\dim(P) \geq 1$ and S its slack matrix. Suppose $cx \leq \delta$ is a valid inequality for P. Then $rk_+(S) = rk_+(S')$, where S' is the matrix obtained by adjoining to S one extra row for the slack of the inequality $cx \leq \delta$.[9]

Proof Of course, by Exercise 10 we have $\mathrm{rk}_+(S) \leq \mathrm{rk}_+(S')$ and we need to prove the converse.

Use Lemma 12 to get a nonnegative combination

$$y \geq 0, \quad yA = c, \quad yb = \delta.$$

[4] That is, for each $x \in P$ we have $cx \leq \delta$.

[5] Also, for us a polytope is bounded by definition.

[6] Otherwise P would be an affine subspace, and since it is bounded, it would be a point ($\dim(P) = 0$).

[7] Also, $b'_i \neq -\infty$ because P is bounded.

[8] Actually, Step 1 only guarantees obtaining a stronger inequality: $Ax \geq b''_i$ for some $b''_i \geq b'_i$. However, we must also have $b''_i \leq b'_i$ since this new inequality, as a combination of $Ax \leq b$, is also valid for P (in particular, for the point where the minimum of $A_i x$ is attained). So in fact $b''_i = b'_i$.

[9] In other words, S' is the slack matrix for the polytope P written as $P = \{x \in \mathbb{R}^n : Ax \leq b, cx \leq \delta\} = \mathrm{conv}\{v_1, ..., v_d\}$.

Denote $r = \mathrm{rk}_+(S)$ and take a nonnegative rank-r factorization

$$S = TU, \quad T \in \mathbb{R}^{m\times r}_{\geq 0}, \quad U \in \mathbb{R}^{r\times d}_{\geq 0}.$$

We claim that

$$S' = T'U, \quad T' \in \mathbb{R}^{(m+1)\times r}_{\geq 0}, \quad U \in \mathbb{R}^{r\times d}_{\geq 0},$$

where T' is T with the extra row yT adjoined, is a nonnegative rank-r factorization of S'. Indeed, if we denote the extra row of S' by β ($\beta^j = \delta - cv_j$), then we have $\beta = (yT)U$, because for each column j,

$$((yT)U)^j = (y(TU))^j = (yS)^j = yS^j = \sum_{i=1}^{m} y_i S_{ij} = \sum_{i=1}^{m} y_i(b_i - A_i v_j) = yb - yAv_j = \delta - cv_j = \beta^j.$$

■

Now we can prove Yannakakis' theorem.

Proof [of Theorem 9]

Denote the description of P from which S was created by:

$$P = \{x \in \mathbb{R}^n : Ax \leq b\} = \mathrm{conv}\{v_1, ..., v_d\}.$$

Direction $\mathbf{xc}(P) \leq \mathbf{rk}_+(S)$: given a rank-$r$ nonnegative factorization of S:

$$T \in \mathbb{R}^{m\times r}_{\geq 0}, \quad U \in \mathbb{R}^{r\times d}_{\geq 0}, \quad S = TU,$$

we can obtain an extended formulation of size r as follows:[10]

$$Q = \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^r : Ax + Ty = b,\ y \geq 0\}.$$

Then:

- $\mathrm{size}(Q) = r$ (note that there are many more equalities),
- we show that $\pi_x(Q) \subseteq P$: given $(x, y) \in Q$ we have $Ty \geq 0$ since $T \geq 0$ and $y \geq 0$, so $Ax = (Ax + Ty) - Ty = b - Ty \leq b$ and thus $x \in P$,
- we show that $\pi_x(Q) \supseteq P$: given $x \in P$, we need to produce a y such that $(x, y) \in Q$. WLOG assume that x is an extreme point of P – then we must have $x = v_i$ for some i. We take $y = U^i$ to be the i-th column of U. Then $y \geq 0$ and $Ax + Ty = Av_i + TU^i = Av_i + S^i = Av_i + b - Av_i = b$, so $(x, y) \in Q$.

Direction $\mathbf{xc}(P) \geq \mathbf{rk}_+(S)$: let

$$Q = \left\{(x, y) \in \mathbb{R}^n \times \mathbb{R}^p : E^{=}x + F^{=}y = d^{=},\ E^{\leq}x + F^{\leq}y \leq d^{\leq}\right\}$$

be an extended formulation of P with $\mathrm{size}(Q) = r$ (that is, the matrix $(E^{\leq}, F^{\leq})$ has r rows). The objective is to show $\mathrm{rk}_+(S) \leq r$.

Observation: since $\pi_x(Q) = P$, all inequalities valid for P are also valid for Q. So our plan is the following: we will write down a slack matrix S' for Q, note that it has rank at most r, add the slacks of inequalities $Ax \leq b$ to it as extra rows (without increasing the rank), and then argue that S is a minor of the matrix so obtained, so S also has rank at most r.

Let then S' be the slack matrix of Q defined as follows. The first rows of S' contain the slack of inequalities $E^{\leq}x + F^{\leq}y$. The remaining rows contain the slack of constraints $E^{=}x + F^{=}y = d^{=}$, written

[10] On some level, this is the usual trick of adding nonnegative slack variables used to transform an LP of the form $Ax \leq b$ to one of the form $Ax = b,\ x \geq 0$. T having few columns means we do not have to add many slack variables.

as inequalities (two for each equality). Note that the latter are zero rows. As for the columns, we take any generating set of points of Q whose first d elements are preimages of the points $v_1, ..., v_d$: denote them by $(v_1, y_1), ..., (v_d, y_d)$. From Exercise 11, $\mathrm{rk}_+(S') \leq r$.

Now add to S' rows corresponding to slacks of all inequalities $Ax \leq b$, obtaining a matrix S''. Because these inequalities are valid for Q, by Exercise 13, $\mathrm{rk}_+(S'') = \mathrm{rk}_+(S') \leq r$.

We are done if we prove that S is a minor of S'', since then by Exercise 10 we have $\mathrm{rk}_+(S) \leq \mathrm{rk}_+(S'') \leq r$. Indeed, S is the minor corresponding to the newly-added rows and to the first d columns. ■

3.3 Rectangle covers

Definition 14 *Given a matrix $S \in \mathbb{R}^{m \times d}_{\geq 0}$, a* rectangle *$R = (X, Y)$ in S is a subset of rows X and a subset of columns Y such that all entries of the minor $S[X \times Y]$ are positive. If we define supp(R) to be $X \times Y$, then in other words we want that supp$(R) \subseteq$ supp(S).*

Definition 15 *A family $\mathcal{R}$ of rectangles of S is called a* rectangle cover *if these rectangles together cover all positive entries of S, i.e.,*

$$\bigcup_{R \in \mathcal{R}} supp(R) = supp(S).$$

We might think that a matrix which has no cover with a small number of rectangles has a complicated structure. Indeed:

Theorem 16 *$rk_+(S) \geq \min\{|\mathcal{R}| : \mathcal{R}$ is a rectangle cover of $S\}$.*

Proof Let $r = \mathrm{rk}_+(S)$ and write the nonnegative factorization:

$$S = TU = \sum_{l=1}^{r} T^l U_l$$

where T^l is the l-th column of T and U_l is the l-th row of U. Then

$$\mathrm{supp}(S) = \bigcup_{l=1}^{r} \mathrm{supp}(T^l U_l) = \bigcup_{l=1}^{r} \mathrm{supp}(T^l) \times \mathrm{supp}(U_l)$$

which yields a rectangle cover of size r. ■

Some remarks on this theorem are in order:

- One can ignore the exact values in the matrix – it only matters whether they are zero or not. This can often be helpful.
- The inequality is not tight – there are examples where $\mathrm{rk}_+(S)$ is exponential, but the minimum rectangle cover is only polynomial-sized.
- However, it will be useful for the correlation polytope, as we will see now.

4 The correlation polytope

Now we are well-equipped to prove an exponential lower bound.

Definition 17 *We define the* correlation polytope *$corr(n)$ to be*

$$corr(n) = conv\{y^b : b \in \{0,1\}^n\} \subseteq \mathbb{R}^{n\times n},$$

where $y^b \in \mathbb{R}^{n\times n}$ is the outer product

$$y^b = bb^\top.$$

Example 18

$$corr(2) = conv\left\{\begin{pmatrix}0 & 0\\ 0 & 0\end{pmatrix}, \begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}, \begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}, \begin{pmatrix}1 & 1\\ 1 & 1\end{pmatrix}\right\}.$$

Note that $\mathrm{xc}(\mathrm{corr}(n)) \le 2^n$, for $\mathrm{corr}(n)$ is a convex hull of 2^n points (see Exercise 11). The rest of the lecture is devoted to the proof of the following theorem:

Theorem 19 *$xc(corr(n)) = 2^{\Omega(n)}$.*

We begin by relating the extension complexity of $\mathrm{corr}(n)$ to the nonnegative rank of a smartly defined matrix S.[11]

Definition 20 *Define the matrix $S \in \mathbb{R}_{\ge 0}^{2^n\times 2^n}$ as follows. Let its rows and columns be indexed by vectors $a, b \in \{0,1\}^n$, and write*

$$S_{ab} = \begin{cases} 0 & \text{if } |a\cap b| = 1,\\ 1 & \text{otherwise.}\end{cases}$$

See Figure 1 for an illustration of S.

Lemma 21 *$xc(corr(n)) \ge rk_+(S)$.*

Proof By Theorem 9 and Exercise 10, it is enough to show that S is a minor of some slack matrix of $\mathrm{corr}(n)$. For the column set of the minor, we will choose all the vertices of S (the matrices y^b). For the row set, we need to come up with a family of inequalities valid for $\mathrm{corr}(n)$ parametrized by a such that the slack of the a-th inequality at y^b is 0 exactly when $|a \cap b| = 1$.

How to define the a-th inequality? Fix a, and consider first the following function of the variable $x = (x_1, ..., x_n) \in \{0,1\}^n$:

$$\pi_a : \{0,1\}^n \to \mathbb{Z}_{\ge 0}, \quad \pi_a(x) = (\langle a, x\rangle - 1)^2 \ge 0.$$

We linearize this function by expanding it into a multivariate polynomial in $x_1, ..., x_n$ and replacing all occurences of $x_i x_j$ with a variable y_{ij} and all occurences of x_i^2 (and x_i, which is equivalent since $x_i \in \{0,1\}$) with a variable y_{ii}. We thus obtain a linear functional $\rho_a : \mathbb{R}^{n\times n} \to \mathbb{R}$ with the property that $\rho_a(y^b) = \rho_a(bb^\top) = \pi_a(b)$ for any $b \in \{0,1\}^n$. Our choice of inequality is $\rho_a(y) \ge 0$. Note that it is valid for $\mathrm{corr}(n)$, since for each vertex y^b we have $\rho_a(y^b) = \pi_a(b) \ge 0$, and that for any b we have that $\rho_a(y^b) = \pi_a(b) = 0$ iff $1 = \langle a, b\rangle = |a \cap b|$. ■

Now our task is reduced to showing that $\mathrm{rk}_+(S)$ is exponentially large. We will do this by proving that S has no small *rectangle cover*.

[11] For the remainder of the notes, n will be suppressed in the notation.

4.1 The lower bound

Theorem 19 will follow from Theorem 16 if we can show that S cannot be covered by a small number of rectangles. We will actually show a somewhat stronger statement:

Theorem 22 *Any collection $\mathcal{R}$ of rectangles in S which covers the subset S' of positive entries of S defined as follows:*

$$(a, b) \in S' \Leftrightarrow |a \cap b| = 0$$

is of size at least $2^{\Omega(n)}$.

See Figure 1 for an illustration of S'.

Figure 1: The matrix S for $n = 3$. Entries in green are those belonging to S' (see Theorem 22).

$$\begin{array}{c} \\ 000 \\ 100 \\ 010 \\ 001 \\ 110 \\ 101 \\ 011 \\ 111 \end{array} \begin{array}{c} \begin{array}{cccccccc} 000 & 100 & 010 & 001 & 110 & 101 & 011 & 111 \end{array} \\ \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix} \end{array}$$

Note that positive entries in S are those (a, b) with $|a \cap b| \neq 1$, so they have either $|a \cap b| = 0$ or $|a \cap b| \geq 2$. The set S' contains only the first ones. Recall that each rectangle of S can only contain positive entries, i.e. it contains no entry equal to 0.

To prove Theorem 22, let us first count how many entries there are to cover. The answer is simple: $|S'| = 3^n$, and this is because for each of the n elements we independently decide whether it is in a, in b, or in none of them.

So we will be done once we prove the following:

Theorem 23 *Any rectangle R in S covers at most 2^n entries of S'.*

This will imply that we need at least $3^n/2^n = (3/2)^n = 2^{\Omega(n)}$ rectangles to cover S'.

Proof [of Theorem 23]

For a rectangle R in S, we write $|R|_0 = |S' \cap R|$ (*magnitude* of R). We must prove $|R|_0 \leq 2^n$. (See Figure 2 for an example.)

The proof is by induction on n. The case $n = 1$ is easy. Suppose we have the claim for $n - 1$.

Proof idea: we will cover $R \cap S'$ using two rectangles, and show that the magnitude of each of those rectangles is at most the magnitude of a rectangle which ignores the element n. By induction hypothesis, this is at most 2^{n-1}.

Fix a rectangle $R = P \times C$. Let

$$\begin{aligned} R_1 &= P_1 \times C_1, \\ P_1 &= \{a \in P : a \ni n\} \cup \{a \in P : a \cup \{n\} \notin P\}, \\ C_1 &= \{b \in C : b \not\ni n\} \end{aligned}$$

and let

$$\overline{R}_1 = \{(a \setminus \{n\}, b) : (a, b) \in R_1\}.$$

Figure 2: An example rectangle $R = \{000, 101, 111\} \times \{000, 101, 111\}$, with $|R|_0 = 5 \leq 2^3$.

$$
\begin{array}{c}
 \\ 000 \\ 100 \\ 010 \\ 001 \\ 110 \\ 101 \\ 011 \\ 111
\end{array}
\begin{array}{c}
\begin{array}{cccccccc} 000 & 100 & 010 & 001 & 110 & 101 & 011 & 111 \end{array} \\
\begin{pmatrix}
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 0 & 1 & 1 & 0 & 0 & 1 & 0 \\
1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\
1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\
1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 \\
1 & 0 & 1 & 0 & 0 & 1 & 0 & 1 \\
1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 \\
1 & 0 & 0 & 0 & 1 & 1 & 1 & 1
\end{pmatrix}
\end{array}
$$

Note that $\overline{R}_1$ is a rectangle, for $\overline{R}_1 = \{a \setminus \{n\} : a \in P_1\} \times C_1$. Moreover, $\overline{R}_1 \subseteq \{0,1\}^{n-1} \times \{0,1\}^{n-1}$. So $|\overline{R}_1|_0 \leq 2^{n-1}$ by the induction hypothesis. Observe that $(a,b) \in S' \cap R_1$ if and only if $(a \setminus \{n\}, b) \in S' \cap \overline{R}_1$. In order to conclude $|R_1|_0 = |\overline{R}_1|_0 \leq 2^{n-1}$, it is then enough to show that for each $(a,b) \in \overline{R}_1$, exactly one of (a,b) and $(a \cup \{n\}, b)$ belongs to R_1. But this immediately follows by definition of P_1.

Define analogously:

$$
\begin{aligned}
R_2 &= P_2 \times C_2, \\
P_2 &= \{a \in P : a \not\ni n\}, \\
C_2 &= \{b \in C : b \ni n\} \cup \{b \in C : b \cup \{n\} \notin C\}
\end{aligned}
$$

and let

$$\overline{R}_2 = \{(a, b \setminus \{n\}) : (a,b) \in R_2\} \subseteq \{0,1\}^{n-1} \times \{0,1\}^{n-1}.$$

Then repeating the arguments above, we deduce $|R_2|_0 \leq 2^{n-1}$.

Claim: $R \cap S' \subseteq (R_1 \cup R_2) \cap S'$.

Once we have this, we conclude that $|R|_0 \leq |R_1|_0 + |R_2|_0 \leq 2^{n-1} + 2^{n-1} = 2^n$.

So let $(a,b) \in R \cap S'$. There are four cases:

- $a \ni n$, $b \not\ni n$: then $(a,b) \in R_1$,
- $a \not\ni n$, $b \ni n$: then $(a,b) \in R_2$,
- $a \ni n$, $b \ni n$: then $n \in a \cap b$ and so $|a \cap b| \neq 0$, a contradiction with $(a,b) \in S'$,
- $a \not\ni n$, $b \not\ni n$: if $a \cup \{n\} \notin P$ or $b \cup \{n\} \notin C$, then $(a,b) \in R_1$ or $(a,b) \in R_2$, respectively. So suppose that $a \cup \{n\} \in P$ and $b \cup \{n\} \in C$, which means that $(a \cup \{n\}, b \cup \{n\}) \in P \times C = R$. On the other hand, this pair cannot be covered by R, since the corresponding entry in S is zero. Indeed, since $(a,b) \in S'$, we have $a \cap b = \emptyset$ and thus $|(a \cup \{n\}) \cap (b \cup \{n\})| = |\{n\}| = 1$, so $(a \cup \{n\}, b \cup \{n\}) \notin R$. This contradiction concludes the proof.

■

5 Discussion

An exponential lower bound on the extension complexity of the correlation polytope was first proved by Fiorini et al. in 2012 [FMP+12]. The above proof is due to Kaibel and Weltge [KW15]. It can also be made not to use Yannakakis' Theorem at all, at the cost of making it more magical.

It turns out that one can prove similar lower bounds for many polytopes, some of them using a reduction to the correlation polytope. Examples include the stable set polytope and the travelling salesman polytope.

Seeing these results, one might think that computational complexity and extension complexity are tightly related: polytopes corresponding to hard (NP-complete) problems have high extension complexity, and polytopes that can be optimized over in polynomial time (like the spanning tree polytope) have polynomial extension complexity.

However, this is not true: in 2014 Rothvoss proved that the extension complexity of the matching polytope is exponential [Rot14].

References

[FMP+12] Samuel Fiorini, Serge Massar, Sebastian Pokutta, Hans Raj Tiwary, and Ronald de Wolf. Linear vs. semidefinite extended formulations: exponential separation and strong lower bounds. In *Proceedings of the forty-fourth annual ACM symposium on Theory of computing*, pages 95–106. ACM, 2012. URL: `http://homepages.cwi.nl/~rdewolf/publ/qc/stoc130-fiorini.pdf`.

[KW15] Volker Kaibel and Stefan Weltge. A short proof that the extension complexity of the correlation polytope grows exponentially. *Discrete & Computational Geometry*, 53(2):396–401, 2015. URL: `http://bipop.inrialpes.fr/people/malick/Docs/13-kaibel.pdf`.

[Mar91] R Kipp Martin. Using separation algorithms to generate mixed integer model reformulations. *Operations Research Letters*, 10(3):119–128, 1991.

[Rot14] Thomas Rothvoß. The matching polytope has exponential extension complexity. In *Proceedings of the 46th Annual ACM Symposium on Theory of Computing*, pages 263–272. ACM, 2014. URL: `http://arxiv.org/abs/1311.2369`.

[Yan88] Mihalis Yannakakis. Expressing combinatorial optimization problems by linear programs. In *Proceedings of the twentieth annual ACM symposium on Theory of computing*, pages 223–228. ACM, 1988. URL: `http://www.researchgate.net/profile/Mihalis_Yannakakis/publication/223030131_Expressing_combinatorial_optimization_problems_by_Linear_Programs/links/53fdfe8e0cf2dca8000475c0.pdf`.