1 The extension complexity of the stable set polytope

Let us recall the correlation polytope and the theorem we proved in a previous lecture on its extension complexity.

Definition 1 We define the correlation polytope $\text{corr}(n)$ to be

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

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

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

This result can be used as a starting point to prove similar lower bounds on the extension complexity of many interesting polytopes. Two useful observations to achieve this are the following.

Lemma 3 Let $P$ be a polyhedron and $F$ a face of $P$. Then $xc(F) \leq xc(P)$.

Proof From polyhedral theory we know that $F = \{x \in P : Ax = b\}$ for an appropriate system $Ax = b$. Hence, for any extended formulation $Cx + Dy \leq c$ of $P$, $Cx + Dy \leq c, Ax = b$ is an extended formulation for $F$ of the same size. $\blacksquare$

Exercise 4 Assuming that in previous lemma $P$ and $F$ are polytopes, prove the statement using properties of slack matrices.

Lemma 5 Let $P \subseteq \mathbb{R}^n, Q \subseteq \mathbb{R}^{n+p}$ be polyhedra such that $Q$ is an extension of $P$. Then $xc(P) \leq xc(Q)$.

Proof Clearly any extension of $Q$ is also an extension of $P$, obtained by composing the linear maps. $\blacksquare$

Suppose now that we show the following reduction.

Lemma 6 For each $n \in \mathbb{N}$, there exists a graph $G$ with $O(n^2)$ nodes such that there exists a face $F$ of $\text{STAB}(G)$ that is an extension of $\text{corr}(n)$.

Then one can conclude using the previous lemmata that for any of those graph $G(V, E)$,

$$xc(\text{STAB}(G)) \geq xc(F) \geq xc(\text{corr}(n)) \geq 2^{\Omega(\sqrt{|V|})},$$

hence

Theorem 7 $xc(\text{STAB}(G(V, E))) \geq 2^{\Omega(\sqrt{|V|})}$.

Proof (of Lemma 6)

Let $n \in \mathbb{N}$. Consider the graph $G(V, E)$ define as follows. For each index $i \in [n]$, create nodes $ii$ and $\overline{i}$. For each pair of indices $i, j \in [n]$ with $i < j$, create nodes $ii, ij, ji, jj$. Add edges as follows: for each $i \in [n]$, add edge $(ii, \overline{i})$, for each pair of indices $i, j \in [n]$ with $i < j$, add all edges between nodes $(ii, ij, ji, jj, \overline{i}j, \overline{j}i, \overline{i}\overline{j}, \overline{j}\overline{i})$.

We will now define a linear function $\pi$ mapping a point $x \in \mathbb{R}^V$ to a point $y \in \mathbb{R}^{n \times n}$ that maps a face $F$ of $\text{STAB}(G)$ to $\text{corr}(n)$. Hence, $F$ is an extension of $\text{corr}(n)$, concluding the proof.
$F$ is the face of $\text{STAB}(G)$ given by clique inequalities

$$x_{ii} + x_{\bar{i}j} \leq 1 \text{ for all } i \in [n] \text{ and } x_{ij} + x_{ij} + x_{\bar{i}j} \leq 1 \text{ for all } i,j \in [n] \text{ with } i < j,$$

while the map $\pi$ is defined as follows:

$$y_{ij} = y_{ji} = x_{ij} \text{ for } i \leq j.$$

Consider a vertex $x$ of $F$: this is the characteristic vector of a stable set $S$ of $G$ satisfying at equality the clique inequalities above. In particular, for each $i < j$,

- either $ii$, $jj$, and $ij \in S$,
- or $ii$, $\bar{i}j$, and $ij \in S$,
- or $\bar{i}i$, $jj$, and $\bar{i}j \in S$,
- or $\bar{i}i$, $\bar{j}j$, and $\bar{i}j \in S$.

In particular, $ij \in S$ iff $ii, jj \in S$. Let $b \in \{0,1\}^n : b_i = 1$ iff $ii \in S$. Then for $y = \pi(x)$ and $i \leq j$ we have $y_{ij} = x_{ij}$ iff $ii, ij \in S$, i.e., $y = bb^T$.

Conversely, consider point $y = bb^T$, with $b \in \{0,1\}^n$. Then the unique stable set of $G$ that contains $ii$ iff $b_i = 1$ belongs to $F$ and is in the preimage of $y$, as required.

## 2 The extension complexity of the perfect matching polytope

Recall that, to prove Theorem 2, we used the concept of rectangle cover.

**Definition 8** Given a matrix $S \in \mathbb{R}_{\geq 0}^{m \times d}$, 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 $\text{supp}(R)$ to be $X \times Y$, then in other words we want that $\text{supp}(R) \subseteq \text{supp}(S)$.

**Definition 9** 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.,

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

**Theorem 10** $r_k(S) \geq \min\{|\mathcal{R}| : \mathcal{R} \text{ is a rectangle cover of } S\}$.

We now show that rectangle covers are of no use for lower bounding the extension complexity of the perfect matching polytope. Recall that the following is a linear description of the perfect matching polytope of a graph $G(V,E)$.

$$\text{PMATCH}(G) = \{x \in \mathbb{R}^V : \begin{align*} x(\delta(v)) &= 1 \text{ for all } v \in V, \\ x(\delta(U)) &\geq 1 \text{ for } U \subseteq V, |U| \text{ odd and at least 3,} \\ x &\geq 0 \}.\$$

Note that $P(G)$ is always a face of $\text{PMATCH}(K_{|V|})$, with $K_{|V|}$ being the complete graph with $|V|$ nodes. Hence, because of Lemma 3, $xc(\text{PMATCH}(G)) \geq xc(\text{PMATCH}(K_{|V|}))$.

**Lemma 11** Let $S$ be a slack matrix of the perfect matching polytope of $K_n$. Then there exists a rectangle cover of $S$ with $O(n^4)$ rectangles.
**Theorem 12** \( xc(\text{PMATCH}(G(V,E))) = 2^O(|V|) \).

The proof is quite involved, so here we will only give a glimpse of the techniques and refer the interested reader to the original paper.

The strategy is as follows: recall that the nonnegative rank of a matrix is the minimum integer \( r \) such that the matrix can be written as the sum of \( r \) nonnegative rank-1 matrices. In particular, the support of each of those matrices is contained in the support of \( S \). Let \( \mathcal{R} \) be the family of \( 0-1 \) rank-1 matrices whose support is contained in the support of \( S \). One first shows that there exists a matrix \( S \) that is a minor of the slack matrix of \( \text{PMATCH}(K_n) \) such that \( <W,S> \) is much bigger than \( <W,R> \) for \( R \in \mathcal{R} \) (here \( <W,S> = \sum_{i,j}W_{i,j}S_{i,j} \) denotes the Frobenius product of \( W \) and \( S \)). We will then be done using the following Hyperplane separation bound.

**Lemma 13** Let \( S \in \mathbb{R}_{\geq 0}^{m \times n}, W \in \mathbb{R}^{m \times n} \). Then

\[
\text{rank}_+(S) \geq \frac{<W,S>}{\alpha ||S||_{\infty}},
\]

where \( \alpha = \max\{<W,R> : R \in \mathcal{R}\} \).

**Proof** We first show that any rank-1 matrix \( R \in [0,1]^{m \times n} \) is in the convex hull of matrices from \( \mathcal{R} \). Write \( R = uv^T \). Note that we can assume that all entries of \( u \) and \( v \) are between \( 0 \) and \( 1 \). Indeed, let \( \Delta > 1 \) be the maximum entry in, say, \( u \). Then all entries of \( v \) are between \( 0 \) and \( \frac{1}{\Delta} \), since \( R \in [0,1]^{m \times n} \).

We can then scale \( u \) by \( \Delta^{-1} \) and \( v \) by \( \Delta \). Let \( x, y \) be independent \( 0-1 \) random vectors, distributed so that \( P[x_i = 1] = u_i \) and \( P[y_i = 1] = v_i \). We obtain

\[
R = uv^T = \mathbb{E}[x] \mathbb{E}[y]^T = \sum_{\bar{u} \in \{0,1\}^m} P(x = \bar{u}) \bar{u} \sum_{\bar{v} \in \{0,1\}^n} P(y = \bar{v}) \bar{v} \bar{v}^T = \sum_{\bar{u} \in \{0,1\}^m, \bar{v} \in \{0,1\}^n} P(x = \bar{u}, y = \bar{v}) \bar{u} \bar{v} \bar{v}^T,
\]

hence probabilities \( P(x = \bar{u}, y = \bar{v}) \) give the nonnegative multipliers of the convex combination. Now let \( S = \sum_{i=1}^r R_i \), where the \( R_i \) are rank-1 matrices. We conclude

\[
<W,S> = \sum_{i=1}^r <W,R_i> = \sum_{i=1}^r \|S_i\|_{\infty} <W, \frac{R_i}{\|R_i\|_{\infty}} > = \sum_{i=1}^r \|R_i\|_{\infty} \alpha \leq r \|S\|_{\infty} \alpha,
\]

as required. ■

Let us remark that, unlike the rectangle covering bound, which only depends on the support of \( S \), the hyperplane separation bound depends on the specific entries of \( S \).

We now have to find a minor \( S \) of the slack matrix of \( P(K_n) \) and a matrix \( W \) for which \( <W,S> \) is much bigger than \( <W,R> \) for each \( R \in \mathcal{R} \). The columns of \( S \) will be indexed by the family \( \mathcal{M} \) of all perfect matchings of \( K_n \), while its rows will be indexed by a certain family \( \mathcal{U} \) of sets \( U \subseteq V \).
For \( \ell = 1, \ldots, n - 1 \), we let

\[
Q_\ell = \{(U, M) \in \mathcal{U} \times \mathcal{M} : |M \cap \delta(U)| = \ell\}.
\]

Note that entries in \( Q_1 \) have slack equal to 0. Hence, the support of any rectangle in \( \mathcal{R} \) does not intersect \( Q_1 \). We now define the functional \( W \) over \( \mathcal{U} \times \mathcal{M} \) as follows:

\[
W_{U,M} = \begin{cases} 
\frac{1}{|Q_3|} & \text{if } (U, M) \in Q_3 \\
-\frac{1}{(k-1)|Q_k|} & \text{if } (U, M) \in Q_k \\
0 & \text{otherwise.}
\end{cases}
\]

for an appropriate constant \( k \). \(<W,S>\) can be computed exactly

\[
< W, S > = \frac{1}{|Q_3|} |Q_3| 2 - \frac{1}{(k-1)|Q_k|} |Q_k|(k-1) = 1
\]

Upper bounding \( < W, R > \) for \( R \in \mathcal{R} \) is much more complicated. Suppose one could prove the following (which, in fact, is hard to prove).

**Theorem 14** Let \( W, S \) be defined as above. Then for each \( R \in \mathcal{R} \), one has \( < W, S > \leq 2^{-\delta n} \) for some constant \( \delta > 0 \).

Then using Lemma 13 and Theorem 14 and the properties of slack matrices, one concludes

\[
xc(P(K_n)) \geq rk^+(S) \geq \frac{< W, S >}{\|S\|_\infty} \max \{< W, R > : R \in \mathcal{R} \} \geq \frac{1}{n^{2-\delta n}} = 2^{\Theta(n)},
\]

as required.

What is the intuitive meaning of \( W \)? Note that, in the rectangle cover given by Lemma 11, entries of value \( k \) of the slack matrix with are covered \( \binom{k+1}{2} \) times. Hence, in this rectangle cover entries with big value are over-covered. This is the reason why the rectangle cover cannot be transformed into a valid nonnegative factorization of the slack matrix. \( W \) is penalizing a rectangle \( R \) for each entry of \( Q_k \) it covers. Theorem 14 implies that all rectangles that covers many entries of \( Q_3 \) will have to cover many entries of \( Q_k \) as well. So all coverings with few rectangles of the slack matrix share with the rectangle covering of Lemma 11 the property of over-covering entries with big value (indeed, \( k \) here is fixed, but any constant \( k \) big enough would work for an appropriate \( \delta \)).

## 3 Notes

Theorem 7 appeared in [2], Lemma 11 appeared in [1], while Theorem 12 appeared in [3]. In [3], Lemma 13 is attributed to Samuel Fiorini.

## References

