# CS 579 Project: Extension Complexity of Polytopes in Combinatorial Optimization

Vasilis Livanos, Manuel Torres
Net IDs: livanos3, manuelt2

May 2, 2018

## 1 Introduction

Assuming $\mathsf{P} \neq \mathsf{NP}$, any linear program (LP) for the traveling salesperson problem ($\mathsf{TSP}$) must be of super-polynomial size, otherwise one could use the ellipsoid method or interior point method to solve a polynomial-size LP for $\mathsf{TSP}$ in polynomial time, refuting $\mathsf{P} \neq \mathsf{NP}$. It is also interesting to consider the converse of this statement: if there exists a polynomial-size LP for $\mathsf{TSP}$, then $\mathsf{P} = \mathsf{NP}$ as $\mathsf{TSP}$ is NP-complete. This motivates the following question: can we write a polynomial size LP for $\mathsf{TSP}$? The work of Fiorini *et al.* [?] attempts to resolve this question and their work is the subject of this exposition.

### 1.1 Problem Statement

The problem is simply stated: find super-polynomial lower bounds for the size of any LP for $\mathsf{TSP}$. This seemingly daunting task motivates the following definition.

**Definition 1** (Extended Formulation)**.** Let $b, d$ be column vectors and let $A$, $B$, and $C$ be real matrices with $n$, $n$, and $r$ columns, respectively. Let $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ and $Q = \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^r : Bx + Cy \leq d\}$. We say that $Q$ is an *extended formulation* (EF) of $P$ if $P = \{x : \exists y \in \mathbb{R}^r, (x, y) \in Q\}$. The size of $Q$, the EF of $P$, is the number of entries in $d$. That is, the size of the EF $Q$ is the number of inequalities defining $Q$. The *extension complexity* of $P$, denoted by $\mathrm{xc}(P)$, is the minimum size EF of $P$.

At a basic level, an EF of a polytope $P$ is a polytope $Q$ in a higher-dimensional space with a different set of constraints, but is in essence equivalent to $P$. It is equivalent in the sense that one can optimize over an EF $Q$ of $P$ to optimize over $P$. However, we gain nothing if $Q$ is more complex than $P$. Suppose, for instance, that $P$ is a polytope with $n$ variables and an exponential number of constraints. If there exists an EF $Q$ of $P$ with a polynomial number of variables and constraints in $n$, then it would be possible to optimize over $Q$ in polynomial time as a way of optimizing over $P$. It is not immediately evident that there should even be EFs that can save an exponential number of constraints at the expense of increasing the number of variables polynomially. We give an example in Section (1.2) that shows the existence of such an EF.

The notion of an EF gives a direction towards answering the question posed at the beginning of this section. In particular, if one could show that the extension complexity of the $\mathsf{TSP}$ polytope is exponential, then there would not exist a polynomial-size LP for $\mathsf{TSP}$. More formally, we are interested in the following problem.

PROBLEM STATEMENT 2. *Consider an instance to the Traveling Salesman Problem (TSP), where the input size, i.e. the number of cities is $n$. Let $TSP(n)$ be a polytope such that every point $x \in TSP(n)$ corresponds to a feasible solution to TSP, and vice-versa. Does there exist an extended formulation $Q$ of $TSP(n)$ of polynomial size, i.e. with a polynomial number of inequalities? In other words, is the extension complexity of $TSP(n)$ polynomial with respect to $n$?*

In Section (1.3), we will show interesting connections between extension complexity and communication complexity that will aid in settling this problem statement in the negative, by showing that the extension complexity of the TSP polytope is exponential with respect to the input size.

## 1.2 The Utility of Extended Formulations

In this subsection, we give an example of an extended formulation of a particular polytope that reduces the number of constraints from exponential to polynomial and only increases the number of variables by a polynomial factor. This example was given in the lecture notes by Vondrák [?].

**Example 3.** The permutahedron $P_{perm}^{(n)}$ is the convex hull of the permutation group on $[n]$. Formally, let

$$P_{perm}^{(n)} = \text{conv}(\{(\pi(1), \ldots, \pi(n)) : \pi \in S_n\}) \subset \mathbb{R}^n,$$

where conv denotes convex hull. Writing this polytope in terms of constraints, we have

$$P_{perm}^{(n)} = \left\{ x \in \mathbb{R}^n : \sum_{i=1}^n x_i = \binom{n+1}{2}; \forall S \subseteq [n], \sum_{i \in S} x_i \geq \binom{|S|+1}{2} \right\}. \tag{1}$$

To see that the sets are the same, consider a vertex $(x_1, \ldots, x_n)$ in $P_{perm}^{(n)}$. As $(x_1, \ldots, x_n)$ is a permutation of $[n]$, it follows that $\sum_{i=1}^n x_i = n(n+1)/2$, which is the first constraint in (1). Furthermore, every subset $S \subseteq \{x_1, \ldots, x_n\}$ must also sum to at least $\binom{|S|+1}{2}$, which is the other set of constraints in (1). These constraints hold for permutations of $[n]$ as $S$ will only be permutations of subsets of $[n]$. Note that $x_{i_1} + x_{i_2} + \cdots + x_{i_k} = \binom{k+1}{2}$ if and only if $x_{i_1}, \ldots, x_{i_k}$ are a permutation of $[k]$. It is then clear that any convex combination of $\{(\pi(1), \ldots, \pi(n)) : \pi \in S_n\}$ also satisfies these constraints.

It is clear that the number of constraints in (1) is exponential. There is a straightforward way to reduce this number. Let $\pi \in S_n$. Consider the matrix $Y_\pi \in \mathbb{R}^{n \times n}$ defined such that $Y_{ij} = 1$ if $j = \pi(i)$ and 0 otherwise. In other words, $Y_{ij}$ is 1 if $i$ maps to $j$ under $\pi$. This mapping can be thought of as a perfect matching between two vertex sets $V_1 = [n]$ and $V_2 = [n]$. That is, $Y_\pi$ defines a perfect matching on $K_{n,n}$, the complete bipartite graph and therefore $\text{conv}(\{Y_\pi : \pi \in S_n\})$ is the bipartite perfect matching polytope on $K_{n,n}$. Fortunately, we know straightforward characterizations of the bipartite perfect matching polytope.

Formally, let $B_n$ be the bipartite perfect matching polytope on $K_{n,n}$. Then

$$B_n = \left\{ y \in \mathbb{R}^{n \times n} : y_{ij} \geq 0; \forall i, \sum_{j=1}^n y_{ij} = 1; \forall j, \sum_{i=1}^n y_{ij} = 1 \right\}.$$

The nontrivial constraints essentially ensure that every vertex on each side of the bipartition is matched to a vertex on the other side of the bipartition. Then we observe that each perfect matching directly corresponds to a permutation. Thus, we have the following extended formulation

$$Q = \left\{ (x, y) \in \mathbb{R}^n \times \mathbb{R}^{n \times n} : y_{ij} \geq 0, \forall i, x_i = \sum_{j=1}^n j \cdot y_{ij}; \forall i, \sum_{j=1}^n y_{ij} = 1; \forall j, \sum_{i=1}^n y_{ij} = 1 \right\}.$$

Then, we finally have that $P_{perm}^{(n)} = \{x \in \mathbb{R}^n : \exists y \in \mathbb{R}^{n \times n}, (x, y) \in Q\}$. It is clear that $Q$ has $n + n^2$ variables and a polynomial number of constraints. Thus, the extension complexity of the permutahedron is polynomial.

## 1.3 Prior and Related Work

**V:** *Mention Yannakakis's results on the Symmetric TSP polytope and also mention a couple things about approximate EFs.*  ****

## 2    Symmetric TSP Polytope

In Yannakakis's paper introducing the idea of extension complexity [**?**], he provided an important result that allowed us to reason about extension complexity in a seemingly tractable manner. We first need a couple of definitions to state Yannakakis's result.

**Definition 4** (Nonnegative Rank). Let $M \in \mathbb{R}_+^{m \times n}$. The *nonnegative rank* of $M$ is defined as $\operatorname{rank}_+(M) = \min\{r : M = TU, T \in \mathbb{R}_+^{m \times r}, U \in \mathbb{R}_+^{r \times n}\}$.

To gain intuition about the above definition, we remark that $\operatorname{rank}(M) \leq \operatorname{rank}_+(M)$, which follows from the fact the definition of $\operatorname{rank}(M)$ is the same except $T$ and $U$ are allowed to be negative. More importantly for this exposition, there is an equivalent characterization of nonnegative rank that proves useful later on. Another important fact that we will use later on is that $\operatorname{rank}_+(M)$ is equal to the minimum $r$ such that $M$ can be written as the sum of $r$ nonnegative rank-1 matrices. Formally, we have that

$$\operatorname{rank}_+(M) = \min\{r : M = \sum_{i=1}^r U_i, U_i \in \mathbb{R}_+^{m \times n}, \operatorname{rank}(U_i) = 1\}$$

Next, we introduce the so-called *slack matrix* of a polytope and then will show the connection between slack matrices and nonnegative rank.

**Definition 5** (Slack Matrix). Let $A \in \mathbb{R}^{m \times d}$, $b \in \mathbb{R}^m$, and $V = \{v_1, \ldots, v_n\} \subseteq \mathbb{R}^d$. Let $P = \{x \in \mathbb{R}^d : Ax \leq b\} = \operatorname{conv}(V)$. Then $S \in \mathbb{R}_+^{m \times n}$ defined by $S_{ij} = b_i - A_i v_j$ with $i \in [m]$, $j \in [n]$ is the *slack matrix* of $P$ with respect to $Ax \leq b$ and $V$.

Intuitively, the slack matrix of $P$ with respect to $Ax \leq b$ and $V$ encodes how much "slack" there is between every vertex in $V$ and a constraint. That is, $S_{ij}$ is essentially the slack between the $i^{th}$ constraint, $b_i$, and the $j^{th}$ vertex, $v_j$.

We only need one more definition before we can state Yannakakis's factorization theorem.

**Definition 6** (Extension). An *extension* of $P \subseteq \mathbb{R}^n$ is a polytope $Q \in \mathbb{R}^m$ such that there is a linear map $f : \mathbb{R}^m \to \mathbb{R}^n$ such that $f(Q) = P$. We say that the *size* of $Q$ is the number of faces, or facets, it has.

The definition of extension is similar to that of an EF and this similarity is characterized in Theorem (7).

THEOREM 7 (Factorization Theorem).

**M:** *For this section, all that remains is to state and prove factorization theorem.*                    ****

## 3    Lower Bounds on Extension Complexity

Using Yannakakis's factorization theorem, we know that if the slack matrix of a given polytope $P$ has nonnegative rank at least $r$, then every EF of $P$ has size at least $r$. Said differently, if we can lower bound the nonnegative rank of the slack matrix of $P$, then this will also be a lower bound for the extension complexity of $P$. Thus, our goal now is to find an approach to lower bounding the nonnegative rank of slack matrices for particular polytopes.

### 3.1    Connections to Communication Complexity

Before we state another result of Yannakakis that aids in lower bounding the nonnegative rank of matrices, we first give some useful definitions.

**Definition 8** (Rectangle Cover). Let $M \in \{0,1\}^{2^n \times 2^n}$ with indices corresponding to bit strings. A *rectangle* is a subset of $\{0,1\}^n \times \{0,1\}^n$. We say that $R \subseteq \{0,1\}^n \times \{0,1\}^n$ is a *b-monochromatic rectangle* for $f$ if $M_{xy} = b$ for all $(x,y) \in R$. We say that a set $\mathcal{R}$ of $b$-monochromatic rectangles is a *b-rectangle cover* if $\{(x,y) \in \{0,1\}^n \times \{0,1\}^n : M_{xy} = b\} \subseteq \bigcup_{R \in \mathcal{R}} R$. We let $\chi(M)$ denote the minimum cardinality monochromatic rectangle cover of $M$.

We will let $\mathrm{suppmat}(M)$ denote the binary support matrix of the matrix $M$. That is,

$$\mathrm{suppmat}(M)_{ij} = \begin{cases} 1 & \text{if } M_{ij} \neq 0 \\ 0 & \text{if } M_{ij} = 0 \end{cases}.$$

With these definitions, we can now compare nonnegative rank and the minimum cardinality monochromatic rectangle cover.

THEOREM 9. *Let $M \in \mathbb{R}_+^{n \times n}$. Then $\mathrm{rank}_+(M) \geq \chi(\mathrm{suppmat}(M))$.*

*Proof.* Let $r = \mathrm{rank}_+(M)$. It is well-known that a matrix with nonnegative rank $r$ can be written as a sum of $r$ rank-1 nonnegative matrices. So let

$$M = \sum_{i=1}^{r} v_i u_i^T,$$

where $v_i, u_i \in \mathbb{R}^n$ for all $i \in [r]$ and $v_i u_i^T$ is nonnegative. Note that writing $M$ as a sum of $r$ rank-1 nonnegative matrices is similar to being able to write a matrix $A$ with $\mathrm{rank}(A) = k$ as a sum of $k$ rank-1 matrices.

Let $\mathrm{supp}(M) = \{(x,y) \in [n] \times [n] : M_{x,y} \neq 0\}$. Then

$$\mathrm{supp}(M) = \mathrm{supp}\left(\sum_{i=1}^{r} v_i u_i^T\right) = \bigcup_{i=1}^{r} \mathrm{supp}(v_i u_i^T),$$

where the last equality used the fact that the $v_i u_i^T$ are nonnegative matrices. Then

$$\bigcup_{i=1}^{r} \mathrm{supp}(v_i u_i^T) = \bigcup_{i=1}^{r} \mathrm{supp}(v_i) \times \mathrm{supp}(u_i^T)$$

as $(v_i u_i^T)_{kj} \neq 0$ if and only if $v_k \neq 0$ and $u_j \neq 0$ for all $k, j$. Therefore, we can construct a 1-rectangle cover for each matrix $v_i u_i^T$ by looking at the entries where $v_i u_i^T$ are nonzero. The entries where $v_i u_i^T$ are nonzero are clearly $\mathrm{supp}(v_i) \times \mathrm{supp}(u_i^T)$. As $\mathrm{supp}(v_i) \times \mathrm{supp}(u_i^T) \subset \{0,1\}^n \times \{0,1\}^n$, it follows that $\mathrm{supp}(v_i) \times \mathrm{supp}(u_i^T)$ is a valid 1-rectangle cover for $v_i u_i^T$. Thus, the set $\{\mathrm{supp}(v_i) \times \mathrm{supp}(u_i^T) : i \in [r]\}$ is a 1-rectangle cover for $\mathrm{suppmat}(M)$. This implies $\chi(\mathrm{suppmat}(M)) \leq r$. Therefore, $\mathrm{rank}_+(M) \geq \chi(\mathrm{suppmat}(M))$. $\qquad\square$

This theorem tells us that if we can show that the slack matrix of a given polytope, or a submatrix of the slack matrix, has a large monochromatic covering bound, then this would imply that any EF of the polytope is also large.

## 3.2 A Matrix of Exponential Nonnegative Rank

**M:** *The title of this subsection is stilted and should be fixed if possible; I want to convey that we are con-* \*\*\*\* *structing M in this subsection because that is probably the most important part about this subsection, but I don't have a succinct, precise way of stating it*
**V:** *How about "A Matrix of Exponential Nonnegative Rank"? Succinct and it conveys that we use rectangle* \*\*\*\* *cover as a lower bound on the nonnegative rank.*

As stated at the beginning of this section, it is sufficient to lower bound the nonnegative rank of the slack matrix of a given polytope to give a lower bound on the extension complexity of a polytope. Theorem (9) shows us that we can lower bound the nonnegative rank of such a slack matrix by the min-cardinality monochromatic rectangle cover of the slack matrix. However, this still leaves us in a somewhat undesirable situation: for *every* polytope we want to show has large extension complexity, we have to lower bound the corresponding slack matrix.

Fortunately, Fiorini *et al.* [?] gave a framework for reductions in this setting, which we describe formally in Section (5.1). For now, we remark that these reductions will proceed as follows. First, we will lower bound

4

the nonnegative rank of one slack matrix of a particular polytope $P$, implying a bound on the extension complexity of $P$. Then we will "reduce" other polytopes to $P$ in order show these polytopes have at least the same extension complexity as $P$.

So then we need to start by lower bounding the nonnegative rank of the slack matrix of a particular polytope. We will see in Section (4) that the following matrix $M$ is going to be a submatrix of the slack matrix whose nonnegative rank we want to lower bound.

Let $M$ be a nonnegative, real $2^n \times 2^n$ matrix indexed by bit strings and defined as follows:

$$M_{xy} = \left(1 - x^T y\right)^2$$

for all $x, y \in \{0, 1\}^n$. We also note that there is an equivalent definition of $M$:

$$M_{xy} = 1 - \langle 2\operatorname{diag}(x) - xx^T, yy^T \rangle,$$

where $\operatorname{diag}(x)$ is the $n \times n$ diagonal matrix with $x$ on its diagonal and $\langle \cdot, \cdot \rangle$ denotes the Frobenius inner product. To verify that this is an equivalent definition of $M$, we observe that

$$
\begin{aligned}
1 - \langle 2\operatorname{diag}(x) - xx^T, yy^T \rangle &= 1 - 2\langle \operatorname{diag}(x), yy^T \rangle + \langle xx^T, yy^T \rangle \\
&= 1 - 2\sum_{i=1}^{n}(\operatorname{diag}(x)yy^T)_{ii} + \sum_{i=1}^{n}(xx^T yy^T)_{ii} \\
&= 1 - 2\sum_{i=1}^{n} x_i y_i^2 + \sum_{i=1}^{n}\sum_{j=1}^{n} x_i y_j x_i y_j \\
&= 1 - 2x^T y + (x^T y)^2 = (1 - x^T y)^2
\end{aligned}
$$

De Wolf [?] showed the following theorem about the matrix $M$ constructed above.

THEOREM 10. *Every* 1-*monochromatic rectangle cover of* $\operatorname{suppmat}(M)$ *has size* $2^{\Omega(n)}$.

*Proof.* The proof critically relies on a result from [?]. The statement is as follows: there exists sets $A, B \subseteq \{0,1\}^n \times \{0,1\}^n$ and a probability distribution $\mu$ over $\{0,1\}^n \times \{0,1\}^n$ such that

1. $\mu(A) = 3/4$,
2. all $(x, y) \in A$ have $x^T y = 0$ and all $(x, y) \in B$ have $x^T y = 1$, and
3. there exists constants $\alpha, \delta > 0$ such that for all rectangles $R$, $\mu(R \cap B) \geq \alpha \cdot \mu(R \cap A) - 2^{-\delta n}$.

Now let $R_1, \ldots, R_k$ be a 1-monochromatic rectangle cover for $\operatorname{suppmat}(M)$. As $M_{xy} = 1$ for all $(x, y) \in R_i$ for any $R_i$, we have that $B \cap R_i = \emptyset$ for all $R_i$. This follows since $M_{xy} = (1 - x^T y)^2$ and $x^T y = 1$ for all $(x, y) \in B$ by (2). Using similar reasoning, we see that $\bigcup_{i=1}^{k}(A \cap R_i) = A$.

So we have from (1) that $\mu(A) = 3/4$. Then since we argued that $A = \bigcup_{i=1}^{k}(A \cap R_i) = A$, we have $3/4 = \mu(\bigcup_{i=1}^{k}(A \cap R_i))$. By union bound, we have $3/4 = \mu(\bigcup_{i=1}^{k}(A \cap R_i)) \leq \sum_{i=1}^{k} \mu(A \cap R_i)$. Then by (3), we obtain $3/4 \leq \sum_{i=1}^{k} \mu(A \cap R_i) \leq k \cdot 2^{-\delta n}/\alpha$. Multiplying both sides by $\alpha 2^{\delta n}$, we see that $k \geq (3\alpha/4) \cdot 2^{\delta n}$, which implies the claim. $\square$

Note that Theorem (10) essentially states that the function $f : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ corresponding to $\operatorname{suppmat}(M)$ has a lower bound of $\Omega(n)$ on its communication complexity.

**M:** *should we elaborate further on this connection to communication complexity? If not, there isn't much* **\*\*\*\*** *comm. complexity in this section, so maybe we should change the name of the section to something more relevant?*

**V:** *Yep, I agree, that's why I changed the section and subsection titles to something (possibly) more mean-* **\*\*\*\*** *ingful. Take a look and tell me what you think.*

# 4   Lower Bounds for $\mathrm{CUT}(n)$ Polytope

This section describes the first of the results in [**?**], by showing that the $\mathrm{CUT}(n)$ polytope has exponential extension complexity. We briefly sketch the proof here, before proceeding more formally. The result essentially lies on a connection between the $\mathrm{CUT}(n)$ polytope and the correlation polytope, shown by [**?**]. The authors use Yannakakis's Theorem 7 to relate the extension complexity of the correlation polytope with the nonnegative rank of a matrix, and then use Yannakakis's Theorem 9 and De Wolf's Theorem 10 to lower bound it by $2^{cn}$ for some positive constant $c$, implying an exponential lower bound on the extension complexity of the $\mathrm{CUT}(n)$ polytope.

Before we proceed, we first need to provide some definitions, starting with the cut polytope $\mathrm{CUT}(n)$.

Let $K_n = (V_n, E_n)$ be the complete graph with $n$ vertices. Then, for a cut $\delta(X)$, $\chi^{\delta(X)} \in \mathbb{R}^{E_n}$ is a characteristic vector of the cut, where for a single edge $e$, $\chi_e^{\delta(X)}$ is defined as

$$\chi_e^{\delta(X)} = \begin{cases} 1 & e \in \delta(X) \\ 0 & e \notin \delta(X) \end{cases}$$

**Definition 11** (Cut Polytope)**.** The *cut polytope* $\mathrm{CUT}(n)$ is defined as the convex hull of the characteristic vectors of all possible cuts in $K_n = (V_n, E_n)$. In other words

$$\mathrm{CUT}(n) := \mathrm{conv}\left(\left\{\chi^{\delta(X)} \in \mathbb{R}^{E_n} \mid X \subseteq V_n\right\}\right)$$

Next, we define the notion of *linear isomorphism* of two polytopes.

**Definition 12** (Linearly Isomorphic Polytopes)**.** Two polytopes $P \subseteq \mathbb{R}^n$ and $Q \subseteq \mathbb{R}^m$ are called *linearly isomorphic* if there exists an invertible matrix $M \in \mathbb{R}^{n \times m}$ such that for every $x \in P$, $y = Mx \in Q$. Equivalently, for every $y \in Q$, $x = M^{-1}y \in P$.

Intuitively, if two polytopes are linearly isomorphic, you can obtain one from the other by applying an invertible linear map. It also follows easily that two linearly isomorphic polytopes have the same number of vertices and facets, and any EF of one polytope can be extended to an EF of the other polytope by using the same invertible linear map. The latter is very important for our analysis, as it implies that a bound on the extension complexity of one polytope applies to any other polytope that is linearly isomorphic to it.

Next, we provide a useful definition of the *correlation polytope*.

**Definition 13** (Correlation Polytope)**.** Let $b \in \{0,1\}^n$. Then $bb^T$ is a rank-1 binary symmetric matrix. The convex hull of all rank-1 binary symmetric matrices is called the *correlation (or Boolean quadric) polytope*, and is denoted by

$$COR(n) := \mathrm{conv}\left(\left\{bb^T \in \mathbb{R}^{n \times n} \mid b \in \{0,1\}^n\right\}\right)$$

We are now ready to show the main result of this section, that provides a lower bound on the extension complexity of the $\mathrm{CUT}(n)$ polytope.

THEOREM 14. *There exists some constant $c > 0$ such that for all $n$,*

$$\mathrm{xc}(\mathrm{CUT}(n+1)) \geq 2^{c(n)}$$

*Proof.* De Simone [**?**] showed the following Lemma, which the authors make use of and we state here without proof

LEMMA 15. *For all $n$, $\mathrm{COR}(n)$ is linearly isomorphic to $\mathrm{CUT}(n+1)$.*

**V:** *Maybe add some intuition here if we actually find out why it holds.*                          ****

We note that, for any $a \in \{0,1\}^n$, the inequality

$$\langle 2\,\mathrm{diag}(a) - aa^T, x \rangle \leq 1 \tag{2}$$

6

is satisfied by all vertices $bb^T$ of $\mathrm{COR}(n)$, as this inequality is equivalent to $(1 - a^T b)^2 \geq 0$, as seen in Section 3.2. By convexity of $\mathrm{COR}(n)$, we get that this inequality is satisfied by all points of $\mathrm{COR}(n)$, and thus it is valid for $\mathrm{COR}(n)$. Also note that these inequalities correspond to the entries of the matrix $M$ of Section 3.2, as

$$M_{ab} = 1 - \langle 2 \operatorname{diag}(a) - aa^T, x \rangle$$

Consider now any system of linear inequalities that describes $\mathrm{COR}(n)$, and starts with the $2^n$ inequalities of (2). If we delete all but the first $2^n$ rows of any slack matrix $S$ with respect to this system of inequalities and $\{bb^T \mid b \in \{0,1\}^n\}$, the resulting $2^n \times 2^n$ matrix is exactly $M$. By Theorem 7, we get that

$$\mathrm{xc}(\mathrm{COR}(n)) = \mathrm{rank}_+(S) \tag{3}$$

Since the nonnegative rank of a matrix is at least the nonnegative rank of any of its submatrices, we have

$$\mathrm{rank}_+(S) \geq \mathrm{rank}_+(M) \tag{4}$$

By Theorem 9, $\mathrm{rank}_+(M)$ is lower bounded by the rectangle covering bound for $\mathrm{suppmat}(M)$, and finally, by Theorem 10, we have that every 1-monochromatic rectangle cover of $\mathrm{suppmat}(M)$ has size $2^{\Omega(n)}$. Therefore, there exists a constant $c > 0$, such that

$$\mathrm{rank}_+(M) \geq 2^{cn} \tag{5}$$

Combining Lemma 15 with equations (3), (4) and (5), we get

$$\mathrm{xc}(\mathrm{CUT}(n+1)) \geq 2^{cn}$$

for some constant $c > 0$. □

Theorem 14 immediately implies the following.

COROLLARY 16. *The extension complexity of* $\mathrm{CUT}(n)$ *is* $2^{\Omega(n)}$.

# 5 $\mathrm{STAB}(n)$ and $\mathrm{TSP}(n)$ Polytopes

It was made clear in the proof of Theorem (14) that the extension complexity of $\mathrm{COR}(n)$ is the same as the extension complexity of $\mathrm{CUT}(n+1)$. Thus, as we lower bounded the extension complexity of $\mathrm{CUT}(n+1)$ by $2^{\Omega(n)}$, we also have $\mathrm{xc}(\mathrm{COR}(n)) = 2^{\Omega(n)}$. As we noted in Section (3.2), we now want to use the fact that $\mathrm{xc}(\mathrm{COR}(n)) = 2^{\Omega(n)}$ and "reduce" other polytopes $Q$ to $\mathrm{COR}(n)$ to give a lower bound on $\mathrm{xc}(Q)$.

## 5.1 Framework for Reductions

We therefore want to formalize what is meant by a reduction in this context.

LEMMA 17. *Let $P$, $Q$, and $F$ be polytopes. Then*

1. *if $F$ is an extension of $P$, then $\mathrm{xc}(F) \geq \mathrm{xc}(P)$;*
2. *if $F$ is a face of $Q$, then $\mathrm{xc}(Q) \geq \mathrm{xc}(F)$.*

*Proof.* We first prove (1). Recall that $\mathrm{xc}(F)$ is the minimum size EF of $F$. Then suppose that $F'$ is the EF attaining this minimum. As $F'$ is an EF for $F$ and $F$ is an extension of $P$, it follows that $F'$ is an EF of $P$. Therefore, $\mathrm{xc}(F) \geq \mathrm{xc}(P)$.

Next, we prove (2). Let $S_Q$ be the slack matrix of $Q$ and let $Q = \mathrm{conv}(\{v_1, \ldots, v_k\})$. Further, let $S_F$ be the slack matrix of $F$. Recall that the rows of the slack matrix $S_Q$ correspond to constraints and the columns correspond to vertices of the polytope $Q$, which are $v_1, \ldots, v_k$. Since $F$ is a face of $Q$, there exists a subset $V' \subseteq \{v_1, \ldots, v_k\}$ such that the vertices of $V'$ correspond to the columns of $S_F$. Thus, by Theorem (7), $\mathrm{xc}(Q) = \mathrm{rank}_+(S_Q) \geq \mathrm{rank}_+(S_F) = \mathrm{xc}(F)$. □

Thus, Lemma (17) gives an idea of how to do the reductions we have been alluding to: show that $\mathrm{COR}(n)$ is the face of a polytope $P$ and then $\mathrm{xc}(P) \geq \mathrm{xc}(\mathrm{COR}(n)) = 2^{\Omega(n)}$.

## 5.2   Reduction from $\mathrm{STAB}(n)$

In this subsection, we will see the result of Fiorini *et al.* [?] showing that the extension complexity of the stable set polytope is exponential. This result is a great example of the reductions introduced in the previous section work. There are two results in the subsection, but most of the work is in the first result showing that there exists a graph such that the stable set polytope corresponding to this graph contains a face that is an extension of $\mathrm{COR}(n)$.

Before we proceed with the results, we first need some definitions. First, when we refer to a stable set of a graph $G = (V, E)$, we are referring to an independent set. For a subset $S \subseteq V$, let $\chi^S \in \mathbb{R}^V$ denote the characteristic vector of $S$. That is, $\chi_v^S = 1$ if $v \in S$ and $\chi_v^S = 0$ otherwise.

**Definition 18** (Stable Set Polytope). The *stable set polytope* $\mathrm{STAB}(G)$ is defined as the convex hull of the characteristic vectors of all stable sets in $G = (V, E)$. That is,

$$\mathrm{STAB}(G) := \mathrm{conv}\left(\left\{\chi^S \in \mathbb{R}^V | S \text{ is a stable set of } G\right\}\right).$$

We are now ready to show the main result of this section.

LEMMA 19. *For every $n$, there exists a graph $H_n$ with $O(n^2)$ vertices such that $\mathrm{STAB}(H_n)$ contains a face that is an extension of $\mathrm{COR}(n)$.*

*Proof.* Here is an outline of the proof: we first construct $H_n$, then we designate the face $F$ of $\mathrm{STAB}(H_n)$ that will be an extension of $\mathrm{COR}(n)$, and finally we give the map $\pi$ showing that $F$ is indeed an extension of $\mathrm{COR}(n)$.

First, we construct $H_n$. To do so, let $K_n$ be the complete graph on the vertices $[n]$. For an edge $ij$ in $K_n$ where $i < j$, label the edge $ij$. Now to construct $H_n$. The vertex set of $H_n$ is

$$V = \{ii, \overline{\overline{ii}} : i \in [n]\} \cup \{ij, \overline{\overline{ij}}, \underline{ij}, \overline{ij}, : i, j \in [n], i < j\}.$$

That is, we add two vertices labeled $ii$ and $\overline{\overline{ii}}$ for every vertex $i$ in $K_n$. Also, we add four vertices labeled $ij, \overline{\overline{ij}}, \underline{ij}, \overline{ij}$ for every edge label in $K_n$. It is thus easy to see that the number of vertices added is proportional to the number of vertices in $K_n$, which is $O(n^2)$. As for the edges of $H_n$, we add an edge between $ii$ and $\overline{\overline{ii}}$ for every $i$ and we make a clique out of $ij, \overline{\overline{ij}}, \underline{ij}, \overline{ij}$ for every edge label in $K_n$. Finally, we add the following edges to $H_n$ for every edge label in $K_n$:

$$\{ij, \overline{\overline{ii}}\}, \{ij, \overline{\overline{jj}}\}, \{\overline{ij}, ii\}, \{\overline{ij}, \overline{\overline{jj}}\},$$
$$\{\underline{ij}, \overline{\overline{ii}}\}, \{\underline{ij}, jj\}, \{\overline{\overline{ij}}, ii\}, \{\overline{ij}, jj\}.$$

Fiorini *et al.* [?] provide a useful figure for their construction, which we include in Figure (1). This completes the construction of $H_n$.

Next, we designate the face $F$ of $\mathrm{STAB}(H_n)$ that will be an extension of $\mathrm{COR}(n)$. Note that in the construction, every vertex $i$ in $K_n$ has two vertices associated with it: $ii$ and $\overline{\overline{ii}}$. These vertices form a clique in $H_n$. We will call these *vertex cliques*. Also, every edge label $ij$ in $K_n$ has four vertices associated with it: $ij, \overline{\overline{ij}}, \underline{ij}, \overline{ij}$. These vertices form a clique in $H_n$. We will call these *edge cliques*. Then we let $F$ be the face of $\mathrm{STAB}(H_n)$ whose vertices correspond to the stable sets containing exactly one vertex in each vertex-clique and each edge clique.

Next, we are ready to propose a linear map that will prove that $F$ is in fact an extension of $\mathrm{COR}(n)$. Let $\pi : \mathbb{R}^{V(H_n)} \to \mathbb{R}^{n \times n}$ be defined such that $\pi(x)$ maps to $y \in \mathbb{R}^{n \times n}$ where $y_{ij} = y_{ji} = x_{ij}$ for $i \le j$. We claim that $\pi(F) = \mathrm{COR}(n)$, which would imply that $F$ is an extension of $\mathrm{COR}(n)$.

We start by showing $\pi(F) \subseteq \mathrm{COR}(n)$. Let $\chi^S \in F$. Then let $b \in \{0, 1\}^n$ such that $b_i = 1$ if $ii \in S$ and $b_i = 0$ if $\overline{\overline{ii}} \in S$. As $\chi^S \in F$, we have that $S$ contains exactly one vertex from every vertex clique and exactly one vertex from every edge clique. Thus, if $ij \in S$, then $\overline{\overline{ii}}$ and $\overline{\overline{ij}}$ cannot be in $S$ since the edges $\{ij, \overline{\overline{ii}}\}$ and $\{ij, \overline{\overline{jj}}\}$ are in $H_n$. Then we have that $ii$ and $jj$ must be in $S$. Furthermore, if $ii$ and $jj$ are in $S$, then it must be the case that $ij \in S$. If this were not true, then either $\overline{\overline{ij}}, \underline{ij}$, or $\overline{ij}$ would be in $S$, but all of these
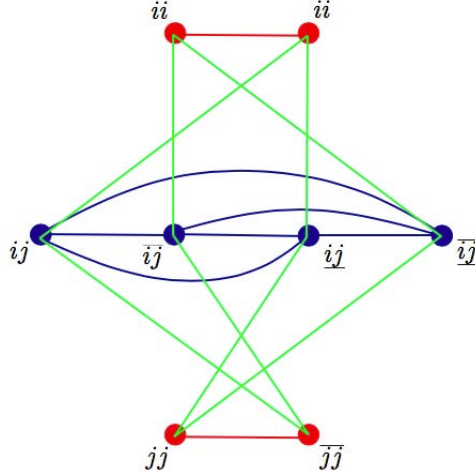
Figure 1: The set of edges and vertices added to $H_n$ for some $i, j \in [n]$ with $i < j$.

have an edge either to $ii$ or $jj$ in $H_n$, so this would contradict the fact that $S$ is a stable set. Therefore, we have that $ij \in S$ if and only if $ii$ and $jj$ are in $S$. Let $x = \chi^S$. Then $x_{ii} = 1$ if and only if $b_{ii} = 1$ by definition of $b_i$. Note that $(bb^T)_{ii} = b_{ii}$. Also, $x_{ij} = 1$ for $i < j$ if and only if $ii$ and $jj$ belong to $S$ as we argued above, which is equivalent to $b_{ii} b_{jj} = 1$, by definition of $b$. Note that $(bb^T)_{ij} = (bb^T)_{ji} = b_{ii} b_{jj}$. Thus, we have that $\pi(x) = bb^T$, which is a vertex of $\text{COR}(n)$, by Definition (13). So $\pi(F) \subseteq \text{COR}(n)$.

Finally, we show $\text{COR}(n) \subseteq \pi(F)$. Let $y = bb^T$ be a vertex of $\text{COR}(n)$. We now want to construct an $x = \chi^S \in F$ such that $\pi(x) = y$. So let $S$ be the set that contains vertex $ii$ if $b_i = 1$ and vertex $\overline{ii}$ if $b_i = 0$. By the argument in the preceding paragraph showing $ij \in S$ if and only if $ii$ and $jj$ are both in $S$, it follows that $x$ is a vertex of $F$ such that $\pi(x) = y$. Therefore, $\text{COR}(n) \subseteq \pi(F)$ and we have $\text{COR}(n) = \pi(F)$, which is what we wanted to show. $\qquad\square$

The following result is essentially a consequence of Lemma (19).

THEOREM 20. *For all $n$, there exists a graph $G_n$ with $n$ vertices such that $\text{xc}(\text{STAB}(G_n)) = 2^{\Omega(\sqrt{n})}$.*

*Proof.* Recall from the proof of Lemma (19) that the size of $H_\ell$ was $2\ell + 4\binom{\ell}{2}$, where the $2\ell$ factor came from the vertex cliques and the $4\binom{\ell}{2}$ came from the edge cliques. Thus, this construction will not work for all $n$ if we use $H_n$. Thus, we proceed as follows. Construct $G_n$ from $H_\ell$ by adding $n - (2\ell + 4\binom{\ell}{2})$ isolated vertices where $\ell$ is the largest integer such that $n - (2\ell + 4\binom{\ell}{2})$ is nonnegative. Then $\text{STAB}(H_\ell)$ is linearly isomorphic to a face of $\text{STAB}(G_n)$ (specifically, the face that contains stable sets with vertices in $H_\ell$). Thus, it easily follows that $\text{xc}(\text{STAB}(G_n)) \geq \text{xc}(\text{STAB}(H_\ell))$ as they are linearly isomorphic.

Then by Lemma (19), we know that $\text{STAB}(H_\ell)$ contains a face that is an extension of $\text{COR}(\ell)$, so by part (2) of Lemma (17), we have that $\text{xc}(\text{STAB}(H_\ell)) \geq \text{xc}(\text{COR}(\ell))$. From Theorem (14), we know that $\text{xc}(\text{COR}(\ell)) = 2^{\Omega(\ell)}$. Finally, as $\ell$ is the largest integer such that $n - (2\ell + 4\binom{\ell}{2}) = n - 2\ell^2$, it follows that $n - 2\ell^2 \geq 0$ and $n - 2(\ell+1)^2 < 0$, which implies $\ell = \Theta(\sqrt{n})$. Therefore, we have $2^{\Omega(\ell)} = 2^{\Omega(\sqrt{n})}$. Combining all of the previous statements in this paragraph, we have $\text{xc}(\text{STAB}(G_n)) = 2^{\Omega(\sqrt{n})}$. $\qquad\square$

## 5.3  Reduction from $\text{TSP}(n)$

## 6  Conclusion

9