Linear vs. Semidefinite Extended Formulations: Exponential Separation and Strong Lower Bounds

Samuel Fiorini[*1], Serge Massar[†3], Sebastian Pokutta[2], Hans Raj Tiwary[‡1], and Ronald de Wolf[§4]

[1]Department of Mathematics, Université libre de Bruxelles CP 216, Boulevard du Triomphe, 1050 Brussels, Belgium. *Email:* {sfiorini, htiwary}@ulb.ac.be
[2]Department of Mathematics, Friedrich-Alexander-Universität Erlangen-Nürnberg, Am Weichselgarten 9, 91058 Erlangen, Germany. *Email:* sebastian.pokutta@math.uni-erlangen.de
[3]Laboratoire d'Information Quantique, Université libre de Bruxelles CP 225, Boulevard du Triomphe, 1050 Brussels, Belgium. *Email:* smassar@ulb.ac.be
[4]CWI, Science Park 123, 1098 XG Amsterdam, The Netherlands. *Email:* rdewolf@cwi.nl

November 2, 2011

Abstract

We solve a 20-year old problem posed by M. Yannakakis and prove that there exists no polynomial-size linear program (LP) whose associated polytope projects to the traveling salesman polytope, even if the LP is not required to be symmetric. Moreover, we prove that this holds also for the maximum cut problem and the stable set problem. These results follow from a new connection that we make between one-way quantum communication protocols and semidefinite programming reformulations of LPs.

[*]Partially supported by the *Actions de Recherche Concertées* (ARC) fund of the French community of Belgium.
[†]Partially supported by the European Commission under the project QCS (Grant No. 255961).
[‡]Postdoctoral Researcher of the *Fonds National de la Recherche Scientifique* (F.R.S.–FNRS).
[§]Partially supported by a Vidi grant from the Netherlands Organization for Scientific Research (NWO), and by the European Commission under the project QCS (Grant No. 255961).

1 Introduction

In 1986–1987 there were attempts to prove P = NP by giving a polynomial-size LP that would solve the traveling salesman problem (TSP). Due to the large size and complicated structure of the proposed LP for TSP, it was difficult to show directly that the LP was erroneous. In a groundbreaking effort to prevent such attempts, Yannakakis [1988] (see Yannakakis [1991] for the journal version) proved that every symmetric LP for the TSP has exponential size. Here, an LP is called *symmetric* if every permutation of the cities can be extended to a permutation of the variables of the LP that preserves the LP. Because the proposed LP for TSP was symmetric, it could not possibly be correct.

In his paper, Yannakakis left as a main open problem the question of proving that the TSP admits no polynomial-size LP, *symmetric or not*. We solve this question by proving a super-polynomial lower bound on the number of inequalities in *every* LP for the TSP. Moreover, we also prove such unconditional super-polynomial lower bounds for the maximum cut and maximum stable set problems. Therefore, it is impossible to prove P = NP by giving a polynomial-size LP for any of these problems. Our approach builds on a close connection between semidefinite programming reformulations of LPs and one-way quantum communication protocols, that we introduce here.[1]

1.1 State of the Art

Solving a Problem Through an LP A combinatorial optimization problem such as the TSP comes with a natural set of binary variables. When we say that an LP solves the problem, we mean that there exists an LP over this set of variables plus extra variables that returns the correct value for all instances over the same set of natural variables, that is, for *all* choices of weights for the natural variables.

From Problems to Polytopes When encoded as 0/1-points in $\mathbb{R}^d$, the feasible solutions of a combinatorial optimization problem such as the TSP yield a polytope that is the convex hull of the resulting points. Solving an instance of the problem amounts to optimizing a linear function over this polytope. (For background on polytopes, see Appendix A.1.)

Extended Formulations Even for polynomially solvable problems, the associated polytope may have an exponential number of facets. By working in an extended space, it is often possible to decrease the number of constraints. In some cases, a polynomial increase in dimension can be traded for an exponential decrease in the number of facets. This is the idea underlying extended formulations. Formally, an *extended formulation* (EF) or *extension* of a polytope $P \subseteq \mathbb{R}^d$ is another polytope[2] $Q \subseteq \mathbb{R}^e$ such that P is the image of Q under a linear map. Optimizing a linear function f over P amounts to optimizing the linear function $f \circ \pi$ over its EF Q, where $\pi : \mathbb{R}^d \to \mathbb{R}^e$ is a linear map that projects P onto Q. We define the *size* of an EF Q as the smallest number of inequalities defining Q.[3]

The Impact of Extended Formulations EFs have for a long time pervaded discrete optimization and approximation algorithms. Indeed, Balas' disjunctive programming [Balas, 1985], the Sherali-Adams hierarchy [Sherali and Adams, 1990], the Lovász-Schrijver closures [Lovász and Schrijver,

[1] For convenience we only consider quantum states and measurements with real entries.

[2] We could allow unbounded polyhedra here, but it makes little difference because every EF of a polytope with a minimum number of facets is also a polytope.

[3] Notice in particular that the number of equalities in the description of Q has no influence on the size of Q. Another possible definition of size is the sum of the number of variables and total number of constraints (equalities or inequalities) defining the EF. Again, this makes little difference because if $P \subseteq \mathbb{R}^d$ has an EF with r inequalities, then it has an EF with $d + r$ variables, r inequalities and at most $d + r$ equalities.

1991], lift-and-project Balas et al. [1993] and configuration LPs are all based on the idea of working in an extended space. Recent surveys on EFs in the context of combinatorial optimization and integer programming are Conforti et al. [2010], Vanderbeck and Wolsey [2010], Kaibel [2011], Wolsey [2011].

Symmetry Matters Yannakakis [1991] proved a $2^{\Omega(n)}$ lower bound on the size of any *symmetric* EF of the TSP polytope $\mathrm{TSP}(n)$ (defined in Section 5). Although he remarked that he did "not think that asymmetry helps much", it was recently shown by Kaibel et al. [2010] (see also Pashkovich [2009]) that symmetry is a restriction in the sense that there exist polytopes that have polynomial-size EFs but no polynomial-size symmetric EF. This revived Yannakakis's tantalizing question about unconditional lower bounds. Those are bounds which apply to the *extension complexity* of a polytope P, defined as the minimum size of an EF of P.

0/1-polytopes with a Large Extension Complexity The strongest unconditional lower bounds so far were obtained by Rothvoß [2011]. By a counting argument inspired by Shannon's theorem [Shannon, 1949], he proved that there exist $0/1$-polytopes in $\mathbb{R}^d$ whose extension complexity is at least $2^{d/2-o(1)}$. However, Rothvoß's technique does not provide *explicit* $0/1$-polytopes with an exponential extension complexity.

The Factorization Theorem Yannakakis [1991] discovered that the extension complexity of a polytope P is determined by certain factorizations of an associated matrix, called the *slack matrix* of P, that records for each pair (F, v) where F is a facet and v is a vertex the algebraic distance of v to a hyperplane supporting F. Defining the *nonnegative rank* of a matrix M as the smallest natural number r such that M can be expressed as $M = UV$ where U and V are nonnegative matrices with r columns and r rows, respectively, it turns out the extension complexity of every polytope P is exactly the nonnegative rank of its slack matrix. This *factorization theorem* led Yannakakis to explore connections between EFs and communication complexity. Let $S = S(P)$ denote the slack matrix of the polytope P. He observed that: (i) every deterministic protocol of complexity k computing S gives rise to an EF of P of size at most 2^k, provided S is a $0/1$-matrix; (ii) the nondeterministic communication complexity of the support matrix of S is a lower bound on the extension complexity of P, or more generally, the nondeterministic communication complexity of the support matrix of every nonnegative matrix M is a lower bound on the nonnegative rank of M.

The Clique vs. Stable Set Problem When P is the stable set polytope $\mathrm{STAB}(G)$ of a graph G (see Section 5), the slack matrix of P contains an interesting row-induced $0/1$-submatrix that is the communication matrix of the *clique vs. stable set problem* (also known as the *clique vs. independent set problem*): its rows correspond to cliques and its columns to stable sets (or independent sets) and the entry for a clique K and stable set S equals $1 - |K \cap S|$. Yannakakis [1991] gave an $O(\log^2 n)$ deterministic protocol for the clique vs. stable set problem, where n denotes the number of vertices of G. This gives a $2^{O(\log^2 n)} = n^{O(\log n)}$ size EF for $\mathrm{STAB}(G)$ whenever the whole slack matrix is $0/1$, that is, whenever G is perfect. An intriguing open problem is to determine the (deterministic or nondeterministic) communication complexity of the clique vs. stable set problem. This is a notoriously hard problem. (For recent results that explain why this problem is hard, see Kushilevitz and Weinreb [2009a,b].) The best lower bound to this day was obtained by Huang and Sudakov [2010]: they obtained a $\frac{6}{5}\log n - O(1)$ lower bound.[4] Furthermore, they state a graph-theoretical conjecture that, if true, would imply a $\Omega(\log^2 n)$ lower bound, and hence settle the communication complexity of the clique vs. stable set problem. Moreover, it would give a

[4] All logarithms in this paper are computed in base 2.

worst case $n^{\Omega(\log n)}$ lower bound on the extension complexity of stable set polytopes. However, a solution to the Huang-Sudakov conjecture seems a distant possibility.

A Tighter Connection to Communication Complexity Faenza et al. [2011] proved that the base-2 logarithm of the nonnegative rank of a matrix equals, up to a small additive constant, the minimum complexity of a randomized communication protocol (with nonnegative outputs) that computes the matrix *in expectation*. In particular, every EF of size r can be regarded as such a protocol of complexity $\log r + O(1)$ that computes a slack matrix in expectation.

1.2 Contribution

Our contribution in this paper is three-fold.

- First, we generalize the factorization theorem to *conic* EFs, that is, reformulations of an LP through a conic program. In particular, this implies a factorization theorem for *semidefinite* EFs: the *semidefinite extension complexity* of a polytope equals the *positive semidefinite rank* (shortly: *PSD rank*) of its slack matrix. To our knowledge, this generalization was also obtained by P. Parillo and R. Thomas (unpublished).

- Second, we generalize the tight connection between linear[5] EFs and classical communication complexity found by Faenza et al. [2011] to a tight connection between *semidefinite* EFs and *quantum* communication complexity. We show that any *rank-r PSD factorization* of a (nonnegative) matrix M gives rise to a one-way quantum protocol computing M in expectation that uses $\log r + O(1)$ qubits and, *vice versa*, that any one-way quantum protocol computing M in expectation that uses q qubits results in a PSD factorization of M of rank 2^q. Via the semidefinite factorization theorem, this yields a characterization of the semidefinite extension complexity of a polytope in terms of the minimum complexity of quantum protocols that compute the corresponding slack matrix in expectation.

  Then, we give a complexity $\log r + O(1)$ quantum protocol for computing a nonnegative matrix M in expectation, whenever there exists a rank-r matrix N such that M is the entry-wise square of N. This implies in particular that every d-dimensional polytope with $0/1$ slacks has a semidefinite EF of size $O(d)$, a result implicit in Gouveia et al. [2010].

  Finally, inspired by earlier work [de Wolf, 2003], we construct a $2^n \times 2^n$ matrix $S = S(n)$ that provides an exponential separation between classical and quantum protocols that compute S in expectation. On the one hand, our quantum protocol gives a rank-$O(n)$ PSD factorization of S. On the other hand, the nonnegative rank of S is $2^{\Omega(n)}$ because the nondeterministic communication complexity of the support matrix of S is $\Omega(n)$. This second part follows from an adaptation of the well-known result of Razborov [1992] on the disjointness problem, see [de Wolf, 2003].

- Third, we use the matrix $S = S(n)$ and a small-rank PSD factorization of S to prove a $2^{\Omega(n)}$ lower bound on the extension complexity of the cut polytope $\mathrm{CUT}(n)$ (see Section 5). That is, *every* linear EF of the cut polytope has an exponential number of inequalities. Via reductions, we infer from this: (i) an infinite family of graphs G such that the extension complexity of the corresponding stable set polytope $\mathrm{STAB}(G)$ is $2^{\Omega(n^{1/2})}$, where n denotes the number of vertices of G; (ii) that the extension complexity of the TSP polytope $\mathrm{TSP}(n)$ is $2^{\Omega(n^{1/4})}$. In addition to settling simultaneously the open problems of Yannakakis [1991] and Rothvoß [2011] described above, our results provide a lower bound on the extension complexity of stable set polytopes

[5] Henceforth, an EF (in the sense of the previous section) is called a *linear* EF. The use of adjectives such as "linear", "semidefinite" or "conic" will help us distinguishing the different types of EFs.

that goes well beyond what is implied by the Huang-Sudakov conjecture. Finally, we point out that although our lower bounds are strong, unconditional and apply to explicit polytopes that are well-known in combinatorial optimization, they have very accessible proofs.

1.3 Related Works

Yannakakis's paper has deeply influenced the TCS community. In addition to the works cited above, it inspired a whole series of papers on the quality of restricted *approximate* EFs such as those defined by the Sherali-Adams hierarchies and Lovász-Schrijver closures that started with Arora et al. [2002] (Arora et al. [2006] for the journal version), see Buresh-Oppenheim et al. [2006], Schoenebeck et al. [2007], Fernandez de la Vega and Mathieu [2007], Charikar et al. [2009], Georgiou et al. [2009, 2010], Benabbas and Magen [2010]. We would also like to point out, that the lower bounds established in Section 5 are essentially based on an efficient PSD factorization or, equivalently, an efficient one-way quantum communication protocol. In this sense our classical lower bounds stem from quantum considerations somewhat similar in style to Kerenidis and de Wolf [2003], Aaronson [2004], Aharonov and Regev [2004] (see Drucker and de Wolf [2011] for a survey of this line of work).

1.4 Outline

In Section 2 we state and prove the factorization theorem for arbitrary convex cones. In Section 3 we establish the equivalence of PSD factorizations of a nonnegative matrix M and one-way quantum protocols that compute M in expectation, and give an efficient quantum protocol in the case where some entry-wise square root of M has small rank. This is used in Section 4 to provide an exponential separation between quantum and classical protocols for computing a matrix in expectation, or equivalently, an exponential separation between nonnegative rank and PSD rank. In Section 5 we prove strong lower bounds on the extension complexity of the cut polytope, the stable set polytope, and the traveling salesman polytope. Concluding remarks are given in Section 6. Background on polytopes as well as some of the proofs can be found in Appendix A.

2 Conic and Semidefinite EFs

Let $Q = \{(x,y) \in \mathbb{R}^{d+k} \mid Ex + Fy = g,\ y \in C\}$ for some *closed* convex cone $C \subseteq \mathbb{R}^k$, where $E \in \mathbb{R}^{p\times d}, F \in \mathbb{R}^{p\times k}$, and $g \in \mathbb{R}^p$. Let $C^* := \{z \in \mathbb{R}^k \mid z^T y \geqslant 0, \forall y \in C\}$ denote the *dual cone* of C. We define the *projection cone of* Q as $C_Q := \{\mu \in \mathbb{R}^p \mid F^T\mu \in C^*\}$ and $\mathrm{proj}_x(Q) := \{x \in \mathbb{R}^d \mid \mu^T Ex \leqslant \mu^T g, \forall \mu \in C_Q\}$. In a first step we show that $\mathrm{proj}_x(Q)$ equals $\pi_x(Q) := \{x \in \mathbb{R}^d \mid \exists y \in \mathbb{R}^k : (x,y) \in Q\}$, the projection of Q onto the x-space. See Appendix A.2 for the proof.

Lemma 1. *We have* $\pi_x(Q) = \mathrm{proj}_x(Q)$. □

Let $P = \{x \in \mathbb{R}^d \mid Ax \leqslant b\} = \mathrm{conv}(V)$ be a polytope, with $A \in \mathbb{R}^{m\times d}$, $b \in \mathbb{R}^m$ and $V = \{v_1, \ldots, v_n\} \subseteq \mathbb{R}^d$. Then $Q = \{(x,y) \in \mathbb{R}^{d+k} \mid Ex + Fy = g,\ y \in C\}$ is a *conic EF* of P w.r.t. C if $P = \pi_x(Q)$. Throughout this paper we use M_i and M^j to denote, respectively, the i-th row and j-th column of M. For convenience we define $[n] := \{1, \ldots, n\}$ for $n \in \mathbb{N}$.

Definition 2. *Let P, $Ax \leq b$ and V be as above. Then $S \in \mathbb{R}_+^{m\times n}$ defined as $S_{ij} := b_i - A_i v_j$ with $i \in [m]$ and $j \in [n]$ is the* slack matrix of *P w.r.t. $Ax \leqslant b$ and V. We sometimes refer to the submatrix of the slack matrix induced by rows corresponding to facets and columns corresponding to vertices simply as* the *slack matrix of P, and denote it by $S(P)$.*

Recall that the extension complexity of polytope P is the minimum size (i.e., number of inequalities) of a linear EF of P. We denote this by $\mathrm{xc}(P)$. This is also the minimum r for which P has a conic EF w.r.t. $C = \mathbb{R}_+^r$ (see, e.g., Fiorini et al. [2011]). Recall moreover that a rank-r nonnegative

factorization of a matrix M is a factorization $M = UV$ where U and V are nonnegative matrices with r columns and r rows, respectively. The *nonnegative rank* of M, denoted by $\text{rank}_+(M)$, is the minimum rank of a nonnegative factorization of M. For linear EFs, the following factorization theorem was proved by Yannakakis. It can be stated succinctly as: $\text{xc}(P) = \text{rank}_+(S(P))$.

Theorem 3 (Yannakakis [1991])**.** *Let $P = \{x \in \mathbb{R}^d \mid Ax \leqslant b\} = \text{conv}(V)$ be a polytope. Then the slack matrix S of P w.r.t. $Ax \leqslant b$ and V has a rank-r nonnegative factorization if and only if there exists a linear EF of the form $Q = \{(x, y) \in \mathbb{R}^{d+r} \mid Ex + Fy = g,\ y \geqslant \mathbf{0}\}$.* □

We would like to remark that even though the size of a polytope is measured as the minimal number of inequalities defining it, we will not restrict the slack matrix to have rows and columns corresponding only to the facet-defining inequalities and the vertices. For the purposes of our discussion, this is not an issue since (i) existence of an EF including redundant inequalities implies the existence of a smaller EF, and (ii) the nonnegative rank of a matrix is always at least the nonnegative rank of any of its submatrices.

We now prove a factorization theorem for the slack matrix of polytopes with respect to more general closed convex cones. The proof of the theorem can be found in Appendix A.2. Yannakakis's result can be obtained as a corollary of our result by taking $C = \mathbb{R}^k_+$, and using Theorem 4 together with the fact that $(\mathbb{R}^k_+)^* = \mathbb{R}^k_+$.

Theorem 4. *Let $P = \{x \in \mathbb{R}^d \mid Ax \leqslant b\} = \text{conv}(V)$ be a polytope defined by m inequalities and n points respectively, and let S be the slack matrix of P w.r.t. $Ax \leqslant b$ and V. Also, let $C \subseteq \mathbb{R}^k$ be a closed convex cone. Then, the following are equivalent:*

(i) There exist T, U such that (the transpose of) each row of T is in C^*, each column of U is in C, and $S = TU$.

(ii) There exists a conic EF $Q = \{(x, y) \in \mathbb{R}^{d+k} \mid Ex + Fy = g,\ y \in C\}$ such that $P = \pi_x(Q)$. □

For a positive integer r, we let $\mathbb{S}^r_+$ denote the cone of $r \times r$ symmetric positive semidefinite matrices embedded in $\mathbb{R}^{r(r+1)/2}$ in such a way that, for all $y, z \in \mathbb{S}^r_+$, the scalar product $z^T y$ is the Frobenius product[6] of the corresponding matrices. A *semidefinite EF of size r* is simply a conic EF w.r.t. $C = \mathbb{S}^r_+$. The *semidefinite extension complexity* of polytope P, denoted by $\text{xc}_{SDP}(P)$, is the minimum r such that P has a semidefinite EF of size r. Observe that $(\mathbb{S}^r_+)^* = \mathbb{S}^r_+$. Hence, taking $C := \mathbb{S}^k_+$ and $k := r(r+1)/2$ in Theorem 4, we obtain the following factorization theorem for semidefinite EFs.

Corollary 5. *Let $P = \{x \in \mathbb{R}^d \mid Ax \leqslant b\} = \text{conv}(V)$ be a polytope. Then the slack matrix S of P w.r.t. $Ax \leqslant b$ and V has a factorization $S = TU$ so that $(T_i)^T, U^j \in \mathbb{S}^r_+$ if and only if there exists a semidefinite EF $Q = \{(x, y) \in \mathbb{R}^{d+r(r+1)/2} \mid Ex + Fy = g,\ y \in \mathbb{S}^r_+\}$ such that $P = \pi_x(Q)$.* □

Analogous to nonnegative factorizations and nonnegative rank, we can define PSD factorizations and PSD rank. A *rank-r PSD factorization* of an $m \times n$ matrix M is a collection of $r \times r$ symmetric positive semidefinite matrices $T_1, \ldots, T_m$ and $U^1, \ldots, U^n$ such that the Frobenius product $Tr\left[(T_i)^T U^j\right] = Tr\left[T_i U^j\right]$ equals M_{ij} for all $i \in [m], j \in [n]$. The *PSD rank* of M is the minimum r such that M has a rank-r PSD factorization. We denote this $\text{rank}_{PSD}(M)$. By Corollary 5, the semidefinite extension complexity of a polytope P is equal to PSD rank of $S(P)$: $\text{xc}_{SDP}(P) = \text{rank}_{PSD}(S(P))$. In the next section we will show that $\text{rank}_{PSD}(M)$ can be expressed in terms of the quantum communication complexity of computing M in expectation (Corollary 7).

[6]The Frobenius product is the component-wise inner product of two matrices. For matrices M and N of the same dimensions, the Frobenius product is equal to $Tr\left[M^T N\right]$. When M is symmetric this can also be written $Tr\left[MN\right]$.

3 Quantum Communication and PSD Factorizations

For a general introduction to quantum computation we refer to Nielsen and Chuang [2000], Mermin [2007], and for quantum communication complexity we refer to de Wolf [2002], Buhrman et al. [2010]. For our purposes, an r-dimensional *quantum state* ρ is an $r \times r$ PSD matrix of trace 1.[7] A *k-qubit state* is a state in dimension $r = 2^k$. If ρ has rank 1, it can be written as an outer product $|\phi\rangle\langle\phi|$ for some unit vector $|\phi\rangle$, which is sometimes called a *pure state*.

A quantum measurement (POVM) is described by a set of PSD matrices $\{E_\theta\}_{\theta\in\Theta}$, each labeled by a real number θ, and summing to the r-dimensional identity: $\sum_{\theta\in\Theta} E_\theta = I$. When measuring state ρ with this measurement, the probability of outcome θ is given by $Tr\left[E_\theta\rho\right]$. Note that if we define the PSD matrix $E := \sum_{\theta\in\Theta} \theta E_\theta$, then the *expected value* of the measurement outcome is $\sum_{\theta\in\Theta} \theta Tr\left[E_\theta\rho\right] = Tr\left[E\rho\right]$.

A *one-way quantum protocol with r-dimensional messages* can be described as follows. On input i, Alice sends Bob an r-dimensional state ρ_i. On input j, Bob measures the state he receives with a POVM $\{E_\theta^j\}$ for some *nonnegative* values θ, and outputs the result. We say that such a protocol *computes a matrix S in expectation*, if the expected value of the output on respective inputs i and j, equals the matrix entry S_{ij}. We will show that such quantum protocols are essentially equivalent to PSD factorizations of S:

Theorem 6. *Let $S \in \mathbb{R}_+^{m\times n}$ be a matrix. Then the following holds:*

(i) A one-way quantum protocol with r-dimensional messages that computes S in expectation, gives a rank-r PSD factorization of S.

(ii) A rank-r PSD factorization of S gives a one-way quantum protocol with $(r+1)$-dimensional messages that computes S in expectation.

Proof. The first part is straightforward. Given a quantum protocol as above, define $E^j := \sum_{\theta\in\Theta} \theta E_\theta^j$. Clearly, on inputs i and j the expected value of the output is $Tr\left[\rho_i E^j\right] = S_{ij}$.

For the second part, suppose we are given a PSD factorization of a matrix S, so we are given PSD matrices $T_1, \ldots, T_m$ and $U^1, \ldots, U^n$ satisfying $Tr\left[T_i U^j\right] = S_{ij}$ for all i, j. In order to turn this into a quantum protocol, define $\tau = \max_i Tr\left[T_i\right]$. Let ρ_i be the $(r+1)$-dimensional quantum state obtained by adding a $(r+1)$st row and column to T_i/τ, with $1 - Tr\left[T_i\right]/\tau$ as $(r+1)$st diagonal entry, and 0s elsewhere. Note that ρ_i is indeed a PSD matrix of trace 1, so it is a well-defined quantum state. For input j, derive Bob's $(r+1)$-dimensional POVM from the PSD matrix U^j as follows. Let λ be the largest eigenvalue of U^j, and define $E_{\tau\lambda}^j$ to be U^j/λ, extended with a $(d+1)$st row and column of 0s. Let $E_0^j = I - E_{\tau\lambda}^j$. These two operators together form a well-defined POVM. The expected outcome (on inputs i, j) of the protocol induced by the states and POVMs that we just defined, is $\tau\lambda Tr\left[E_{\tau\lambda}^j \rho_i\right] = Tr\left[T_i U^j\right] = S_{ij}$, so the protocol indeed computes S in expectation. □

We obtain the following corollary which summarizes the characterization of semidefinite EFs:

Corollary 7. *For a polytope P, the following are equivalent:*

(i) P has a semidefinite EF $Q = \{(x,y) \in \mathbb{R}^{d+r(r+1)/2} \mid Ax + Fy = b, y \in \mathbb{S}_+^r\}$;

(ii) the slack matrix $S(P)$ has a rank-r PSD factorization;

(iii) there exists a one-way quantum communication protocol with $(r+1)$-dimensional messages (i.e., using $\lceil\log(r+1)\rceil$ qubits) that computes $S(P)$ in expectation (for the converse we consider r-dimensional messages). □

[7] For simplicity we will restrict to real rather than complex entries here, which doesn't significantly affect the results.

3.1 A general upper bound on quantum communication

Now, we provide a quantum protocol that efficiently computes a nonnegative matrix S in expectation, whenever there exists a low rank matrix M whose entry-wise square is S. The quantum protocol is inspired by [de Wolf, 2003, Section 3.3].

Theorem 8. *Let S be a matrix with nonnegative real entries, M be a rank-r matrix of the same dimensions such that $S_{ij} = M_{ij}^2$. Then there exists a one-way quantum protocol using $(r+1)$-dimensional pure-state messages that computes S in expectation.*

Proof. Let $M^T = U\Sigma V$ be the singular value decomposition of the transpose of M; so U and V are unitary, Σ is a matrix whose first r diagonal entries are nonzero while its other entries are 0, and $\langle j|U\Sigma V|i\rangle = M_{ij}$. Define $|\phi_i\rangle = \Sigma V|i\rangle$. Since only its first r entries can be nonzero, we will view $|\phi_i\rangle$ as an r-dimensional vector. Let $\Delta_i = \||\phi_i\rangle\|$ and $\Delta = \max_i \Delta_i$. Add one additional dimension and define the normalized $(r+1)$-dimensional pure quantum states $|\psi_i\rangle = (|\phi_i\rangle/\Delta, \sqrt{1-\Delta_i^2/\Delta^2})$. Given input i, Alice sends $|\psi_i\rangle$ to Bob. Given input j, Bob applies a 2-outcome POVM $\{E^j_{\Delta^2}, E^j_0 = I - E^j_{\Delta^2}\}$ where $E^j_{\Delta^2}$ is the projector on the pure state $U^*|j\rangle$ (which has no support in the last dimension of $|\psi_i\rangle$). If the outcome of the measurement is $E^j_{\Delta^2}$ then Bob outputs Δ^2, otherwise he outputs 0. Accordingly, the expected output of this protocol on input (i,j) is

$$\Delta^2 \Pr[\text{outcome } E^j_{\Delta^2}] = \Delta^2 \langle\psi_i|E^j_{\Delta^2}|\psi_i\rangle = \langle\phi_i|E^j_{\Delta^2}|\phi_i\rangle = |\langle j|U|\phi_i\rangle|^2 = |\langle j|U\Sigma V|i\rangle|^2 = M_{ij}^2 = S_{ij}.$$

The protocol only has two possible outputs: 0 and Δ^2, both nonnegative. Hence it computes S in expectation with an $(r+1)$-dimensional quantum message. □

Note that if S is a 0/1-matrix then $M = S$, hence any low-rank 0/1-matrix can be computed in expectation by an efficient quantum protocol. We obtain the following corollary (implicit in Theorem 4.2 of Gouveia et al. [2010]) which also implies a compact semidefinite EF for the stable set polytope of perfect graphs, reproving the previously known result by Lovász [1979, 2003].

Corollary 9. *Let P be a polytope such that $S(P)$ is a 0/1 matrix. Then* $\text{xc}_{SDP}(P) \leqslant \dim(P) + 2$. □

4 Exponential Separations: Quantum vs Classical Communication, and PSD vs Linear Factorizations

In the last section we described a close connection between PSD factorizations of a given matrix, and the dimension of quantum messages needed to compute (the entries of) that matrix in expectation. We will now look at a specific example for which we also prove an exponentially larger lower bound on the number of *classical* bits of communication needed to compute it as compared to the quantum communication complexity. Thus we obtain an exponential separation between quantum and classical communication needed to compute S in expectation. By the connection of the last section, this exponential separation in communication also provides an exponential separation between the ranks of PSD factorizations and nonnegative factorizations.

Recall that the classical nondeterministic communication complexity of a binary communication matrix is defined as $\lceil \log B \rceil$, where B is the minimum number of 1-rectangles that cover the matrix, see Kushilevitz and Nisan [1997]. This last quantity is also known as the *rectangle covering bound*. The reader should also be reminded that a classical randomized protocol in our setting computes a matrix *in expectation* and outputs only nonnegative values, see Faenza et al. [2011].

Let S be an $m \times n$ nonnegative matrix. The *support matrix* of S, denoted by $\mathrm{supp}(S)$, is the $m \times n$ binary matrix with $\mathrm{supp}(S)_{ij} = 1$ iff $S_{ij} > 0$. Then the following result is obvious (the first part is implicit in [Yannakakis, 1991] and the second part is essentially proved in Faenza et al. [2011]).

Lemma 10. *Let S denote a nonnegative matrix, and let B be the rectangle covering bound of $\mathrm{supp}(S)$. Then the following hold: (i) $\mathrm{rank}_+(S) \geqslant B$; (ii) the complexity of every classical protocol that computes S in expectation is at least the classical nondeterministic complexity $\lceil \log B \rceil$ of the communication problem that has S as communication matrix.* □

Now we turn to the main result of this section.

Theorem 11. *For each n, there exists a nonnegative matrix $S \in \mathbb{R}^{2^n \times 2^n}$, such that any classical randomized protocol needs $\Omega(n)$ bits to compute S in expectation. Furthermore, there exists a quantum protocol that computes S in expectation using $\log n + O(1)$ qubits.*

Proof. Consider the matrix $M \in \mathbb{R}^{2^n \times 2^n}$ whose rows and columns are indexed by n-bit strings a and b, respectively, and whose entries are defined as $M_{ab} = 1 - a^T b$. Define $S \in \mathbb{R}_+^{2^n \times 2^n}$ by $S_{ab} = M_{ab}^2$. Note that M has rank $r \leqslant n + 1$ because it can be written as the sum of $n + 1$ rank-1 matrices. Hence Theorem 8 immediately implies a quantum protocol with $(n + 2)$-dimensional messages that computes S in expectation.

In order to prove an exponentially larger classical lower bound, consider the communication complexity problem whose communication matrix is the support matrix of S. This corresponds to the Boolean predicate f with $f(a, b) = 1$ iff $a^T b \neq 1$. By Lemma 10, the classical nondeterministic complexity of f is a lower bound on the complexity of a protocol that computes S in expectation. [de Wolf, 2003, Theorem 3.6] proves an $\Omega(n)$ lower bound on the classical nondeterministic communication complexity of f, hence we get the same lower bound on classical protocols that compute S in expectation. □

Together with Theorem 6 and the equivalence of randomized communication complexity (in expectation) and nonnegative rank established in Faenza et al. [2011], we immediately obtain an exponential separation between the nonnegative rank and the PSD rank.

Corollary 12. *For each n, there exists $S \in \mathbb{R}_+^{2^n \times 2^n}$, with $\mathrm{rank}_+(S) = 2^{\Omega(n)}$ and $\mathrm{rank}_{PSD}(S) = O(n)$.* □

5 Consequences: Strong Lower Bounds on Extension Complexity

Here we prove that the extension complexity of the cut polytope of the n-vertex complete graph is $2^{\Omega(n)}$, i.e., every linear EF of this polytope has an exponential number of inequalities. Then, via reductions, we prove super-polynomial lower bounds for the stable set polytope and the TSP polytope. Our starting point is the matrix $S = S(n)$ used in the previous section to obtain an exponential separation between nonnegative rank and PSD rank. We use a small-rank PSD factorization of S to embed S as a submatrix of the slack matrix of the cut polytope of K_{n+1}. The (classical) nondeterministic complexity of the support matrix of S gives a lower bound on the extension complexity of the cut polytope, implying a $2^{\Omega(n)}$-lower bound on the extension complexity of the cut polytope of K_n.

5.1 Cut Polytopes

Let $K_n = (V_n, E_n)$ denote the n-vertex complete graph. For a set X of vertices of K_n, we let $\delta(X)$ denote the set of edges of K_n with one endpoint in X and the other in its complement $\bar{X}$. This set $\delta(X)$ is known as the *cut* defined by X. For a subset F of edges of K_n, we let $\chi^F \in \mathbb{R}^{E_n}$ denote the *characteristic vector* of F, with $\chi_e^F = 1$ if $e \in F$ and $\chi_e^F = 0$ otherwise. The *cut polytope* $\mathrm{CUT}(n)$

is defined as the convex hull of the characteristic vectors of all cuts in the complete graph $K_n = (V_n, E_n)$. That is, $\mathrm{CUT}(n) := \mathrm{conv}\{\chi^{\delta(X)} \in \mathbb{R}^{E_n} \mid X \subseteq V_n\}$.

We will not deal with the cut polytopes directly, but rather with polytopes that are affinely isomorphic to them. The *correlation polytope* (or *boolean quadric polytope*) $\mathrm{COR}(n)$ is defined as the convex hull of all the rank-one binary symmetric matrices of size $n \times n$. In other words, $\mathrm{COR}(n) := \mathrm{conv}\{bb^T \in \mathbb{R}^{n\times n} \mid b \in \{0,1\}^n\}$. We will use the following known result:

Theorem 13 (De Simone [1989/90])**.** *For all n, $\mathrm{COR}(n)$ and $\mathrm{CUT}(n+1)$ are affinely isomorphic.* □

In Section 4, we defined a $2^n \times 2^n$ nonnegative matrix $S = S(n)$ with rows and columns indexed by n-bit strings such that $S_{ab} = (1 - a^T b)^2$ for all $a, b \in \{0,1\}^n$. We give an explicit PSD factorization of S. Up to normalization, this PSD factorization of S coincides with that provided by the protocol in the proof of Theorem 8. The PSD factorization is given in our next lemma, whose proof can be found in Appendix A.3.

Lemma 14. *If, for each $a, b \in \{0,1\}^n$, we let $U_a := \binom{1}{-a}\binom{1}{-a}^T$ and $V^b := \binom{1}{b}\binom{1}{b}^T$, then the matrices $\{U_a\}_{a\in\{0,1\}^n}$ and $\{V^b\}_{b\in\{0,1\}^n}$ define a PSD factorization of S.* □

Letting $\langle \cdot, \cdot \rangle$ denote the Frobenius product, we can write, for all $a, b \in \{0,1\}^n$: $S_{ab} = \langle U_a, V^b \rangle = 1 - \langle 2\,\mathrm{diag}(a) - aa^T, bb^T \rangle$, where the first equality comes from Lemma 14 and the last equality is a simple rewriting that uses the fact that b is a *binary* vector. This proves the next lemma.

Lemma 15. *For all $a \in \{0,1\}^n$, the inequality*

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

is valid for $\mathrm{COR}(n)$. Moreover, the slack of vertex $x = bb^T$ with respect to (1) is precisely S_{ab}. □

We remark that (1) is implied by the *hypermetric inequality* [Deza and Laurent, 1997] $\langle \mathrm{diag}(a) - aa^T, x \rangle \leqslant 0$. Now, we go on to prove the main result of this section.

Theorem 16. *There exists some constant $C > 0$ such that, for all n,*

$$\mathrm{xc}(\mathrm{CUT}(n+1)) = \mathrm{xc}(\mathrm{COR}(n)) \geqslant 2^{Cn}\,.$$

In particular, the extension complexity of $\mathrm{CUT}(n)$ is $2^{\Omega(n)}$.

Proof. The equality is implied by Theorem 13. Now, consider any system of linear inequalities describing $\mathrm{COR}(n)$ that starts with the 2^n inequalities (1) where $a \in \{0,1\}^n$, and a slack matrix of $\mathrm{COR}(n)$ w.r.t. this system and $\{bb^T \mid b \in \{0,1\}^n\}$. Next, delete from this slack matrix all rows except the 2^n first rows. By Lemma 15, the resulting matrix is S. By combining the fact that the nonnegative rank of any matrix is always greater or equal to the nonnegative rank of any of its submatrices with the factorization theorem (see Theorem 3 above) we find $\mathrm{xc}(\mathrm{COR}(n)) \geqslant \mathrm{rank}_+(S)$. The nonnegative rank of S is bounded from below by the rectangle covering bound for S. By [de Wolf, 2003, Theorem 3.6], the latter is $2^{\Omega(n)}$. Hence, we have $\mathrm{rank}_+(S) \geqslant 2^{Cn}$ for some constant C. The theorem follows. □

5.2 Stable set polytopes

A *stable set* S (also called an *independent set*) of a graph G is a subset $S \subseteq V$ of the vertices such that no two of them are adjacent. For a subset $S \subseteq V$, we let $\chi^S \in \mathbb{R}^n$ denote the *characteristic vector* of S, with $\chi^S_v = 1$ if $v \in S$ and $\chi^S_v = 0$ otherwise. The *stable set polytope*, denoted $\mathrm{STAB}(G)$, is the convex hull of the characteristic vectors of all stable sets in G, i.e., $\mathrm{STAB}(G) := \mathrm{conv}\{\chi^S \in \mathbb{R}^{V(G)} \mid S \text{ stable set of } G\}$. The proof of the next lemma can be found in Appendix A.3. Recall that a polytope Q is a linear EF of a polytope P if P is the image of Q under a linear projection.

Lemma 17. *For each n, there exists a graph G_n with $O(n^2)$ vertices such that* $\text{STAB}(G_n)$ *contains a face that is a linear EF of* $\text{CUT}(n)$. □

Consider a polytope P, and a polytope Q that is a linear EF of P. Obviously $\text{xc}(P) \leqslant \text{xc}(Q)$. Next, consider a face F of P. Then, F is the projection of some face G of Q, and so G is a linear EF of F. Hence $\text{xc}(F)$ is at most the number of facets of G, which is at most the number of facets of Q. Since this holds for any linear EF Q of P, we have $\text{xc}(F) \leqslant \text{xc}(P)$. (Recall that the size of a linear EF is defined as the number of facets of the EF.) Therefore, Lemma 17 implies the following result.

Theorem 18. *For all n, one can construct a graph G'_n with n vertices such that the linear extension complexity of* $\text{STAB}(G'_n)$ *is* $2^{\Omega(n^{1/2})}$. □

5.3 TSP Polytopes

The *traveling salesman polytope of* $K_n = (V_n, E_n)$, denoted by $\text{TSP}(n)$, is defined as the convex hull of the characteristic vectors of all subsets $F \subseteq E_n$ that define a tour of K_n. That is, $\text{TSP}(n) := \text{conv}\{\chi^F \in \mathbb{R}^{E_n} \mid F \subseteq E_n \text{ is a tour of } K_n\}$. It is known that for every graph G with n vertices, $\text{STAB}(G)$ is the linear projection of a face of $\text{TSP}(n')$ with $n' = O(n^2)$, see the proof of [Yannakakis, 1991, Theorem 6]. This together with Theorem 18 gives us the following:

Theorem 19. *The linear extension complexity of* $\text{TSP}(n)$ *is* $2^{\Omega(n^{1/4})}$. □

6 Concluding Remarks

In addition to proving the first unconditional super-polynomial lower bounds on the size of EFs for the cut polytope, stable set polytope and TSP polytope, we demonstrate that the rectangle covering bound can prove strong results in the context of EFs. In particular, it can be super-polynomial in the dimension and logarithm of the number of vertices of the polytope, settling an open problem of Fiorini et al. [2011].

The exponential separation between nonnegative rank and PSD rank that we prove here (see Theorem 11) actually implies more than a super-polynomial lower bound on the extension complexity of the cut polytope. As noted in Theorem 13, the polytopes $\text{CUT}(n)$ and $\text{COR}(n-1)$ are affinely isomorphic. Let $P(n)$ denote the polyhedron isomorphic (under the same affine map) to the polyhedron defined by (1) for $a \in \{0,1\}^n$. Then (i) *every* polytope (or polyhedron) that contains $\text{CUT}(n)$ and is contained in $P(n)$ has exponential extension complexity; (ii) there exists a low complexity *spectrahedron* that contains $\text{CUT}(n)$ and is contained in $P(n)$. A spectrahedron is an intersection of the positive semidefinite cone with an affine subspace.

An important problem —also left open in Yannakakis [1991]— is whether the perfect matching polytope has a polynomial-size linear EF. Yannakakis proved that every *symmetric* EF of this polytope has exponential size, a striking result given the fact that the perfect matching problem is polynomially solvable. He conjectured that asymmetry also does not help in the case of the perfect matching polytope. Because it is based on the rectangle covering bound, our argument would not yield any super-polynomial lower bound on the extension complexity of the perfect matching polytope. Even though a polynomial-size linear EF of the matching polytope would not prove anything as surprising as P=NP, the existence of a polynomial-size EF or an unconditional super-polynomial lower bound for it remains open.

We hope that the new connections developed here will inspire more research, in particular about approximate EFs. Here are two concrete questions we leave open for future work: (i) find a *slack matrix* that has an exponential gap between nonnegative rank and PSD rank; (ii) prove that the cut polytope has no polynomial-size semidefinite EF.

Acknowledgements

We thank Giannicola Scarpa for useful discussions.

References

S. Aaronson. Lower bounds for local search by quantum arguments. In *Proc. STOC 2004*, pages 465–474, 2004.

D. Aharonov and O. Regev. Lattice problems in NP ∩ coNP. In *Proc. FOCS 2004*, pages 362–371, 2004.

S. Arora, B. Bollobás, and L. Lovász. Proving integrality gaps without knowing the linear program. In *Proc. FOCS 2002*, pages 313–322, 2002.

S. Arora, B. Bollobás, L. Lovász, and I. Tourlakis. Proving integrality gaps without knowing the linear program. *Theory Comput.*, 2:19–51, 2006.

E. Balas. Disjunctive programming and a hierarchy of relaxations for discrete optimization problems. *SIAM J. Algebraic Discrete Methods*, 6:466–486, 1985.

E. Balas, S. Ceria, and G. Cornuéjols. A lift-and-project algorithm for mixed 0-1 programs. *Math. Programming*, 58:295–324, 1993.

S. Benabbas and A. Magen. Extending SDP integrality gaps to Sherali-Adams with applications to quadratic programming and MaxCutGain. In *Proc. IPCO 2010*, volume 6080 of *Lecture Notes in Comput. Sci.*, pages 299–312. Springer, 2010.

H. Buhrman, R. Cleve, S. Massar, and R. de Wolf. Nonlocality and communication complexity. *Rev. Modern Phys.*, 82:665, 2010.

J. Buresh-Oppenheim, N. Galesi, S. Hoory, A. Magen, and T. Pitassi. Rank bounds and integrality gaps for cutting planes procedures. *Theory Comput.*, 2:65–90, 2006.

M. Charikar, K. Makarychev, and Y. Makarychev. Integrality gaps for Sherali-Adams relaxations. In *Proc. STOC 2009*, pages 283–292, 2009.

M. Conforti, G. Cornuéjols, and G. Zambelli. Extended formulations in combinatorial optimization. *4OR*, 8:1–48, 2010.

C. De Simone. The cut polytope and the Boolean quadric polytope. *Discrete Math.*, 79:71–75, 1989/90.

R. de Wolf. Quantum communication and complexity. *Theoret. Comput. Sci.*, 287:337–353, 2002.

R. de Wolf. Nondeterministic quantum query and communication complexities. *SIAM J. Comput.*, 32:681–699, 2003.

M.M. Deza and M. Laurent. *Geometry of cuts and metrics*, volume 15 of *Algorithms and Combinatorics*. Springer-Verlag, Berlin, 1997.

A. Drucker and R. de Wolf. Quantum proofs for classical theorems. *ToC Library Graduate Surveys*, 2, 2011.

Y. Faenza, S. Fiorini, R. Grappe, and H. R. Tiwary. Extended formulations, non-negative factorizations and randomized communication protocols. arXiv:1105.4127, 2011.

W. Fernandez de la Vega and C. Mathieu. Linear programming relaxation of Maxcut. In *Proc. SODA 2007*, 2007.

S. Fiorini, V. Kaibel, K. Pashkovich, and D. O. Theis. Combinatorial bounds on nonnegative rank and extended formulations. Manuscript, 2011.

K. Georgiou, A. Magen, and M. Tulsiani. Optimal Sherali-Adams gaps from pairwise independence. In *Proc. APPROX-RANDOM 2009*, pages 125–139, 2009.

K. Georgiou, A. Magen, T. Pitassi, and I. Tourlakis. Integrality gaps of $2 - o(1)$ for vertex cover SDPs in the Lovász-Schrijver hierarchy. *SIAM J. Comput.*, 39:3553–3570, 2010.

J. Gouveia, P.A. Parrilo, and R. Thomas. Theta bodies for polynomial ideals. *SIAM J. Optim.*, 20: 2097–2118, 2010.

H. Huang and B. Sudakov. A counterexample to the Alon-Saks-Seymour conjecture and related problems. arXiv:1002.4687, 2010.

V. Kaibel. Extended formulations in combinatorial optimization. *Optima*, 85:2–7, 2011.

V. Kaibel, K. Pashkovich, and D.O. Theis. Symmetry matters for the sizes of extended formulations. In *Proc. IPCO 2010*, pages 135–148, 2010.

I. Kerenidis and R. de Wolf. Exponential lower bound for 2-query locally decodable codes via a quantum argument. In *Proc. STOC 2003*, pages 106–115, 2003.

E. Kushilevitz and N. Nisan. *Communication complexity*. Cambridge University Press, 1997.

E. Kushilevitz and E. Weinreb. The communication complexity of set-disjointness with small sets and 0-1 intersection. In *2009 50th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2009)*, pages 63–72. IEEE Computer Soc., Los Alamitos, CA, 2009a.

E. Kushilevitz and E. Weinreb. On the complexity of communication complexity. In *Proc. STOC 2009*, pages 465–474. ACM, 2009b.

C. Lemaréchal and J.B. Hiriart-Urruty. *Convex analysis and minimization algorithms I*. Springer, 1996.

L. Lovász. On the Shannon capacity of a graph. *IEEE Trans. Inform. Theory*, 25:1–7, 1979.

L. Lovász. Semidefinite programs and combinatorial optimization. In *Recent advances in algorithms and combinatorics*, volume 11 of *CMS Books Math./Ouvrages Math. SMC*, pages 137–194. Springer, New York, 2003.

L. Lovász and A. Schrijver. Cones of matrices and set-functions and 0-1 optimization. *SIAM J. Optim.*, 1:166–190, 1991.

N.D. Mermin. *Quantum computer science: an introduction*. Cambridge University Press, 2007.

M. A. Nielsen and I. L. Chuang. *Quantum computation and quantum information*. Cambridge University Press, Cambridge, 2000.

K. Pashkovich. Symmetry in extended formulations of the permutahedron. arXiv:0912.3446, 2009.

A. A. Razborov. On the distributional complexity of disjointness. *Theoret. Comput. Sci.*, 106(2): 385–390, 1992.

T. Rothvoß. Some 0/1 polytopes need exponential size extended formulations. arXiv:1105.0036, 2011.

G. Schoenebeck, L. Trevisan, and M. Tulsiani. Tight integrality gaps for Lovasz-Schrijver LP relaxations of vertex cover and max cut. In *Proc. STOC 2007*, pages 302–310. ACM, 2007.

C. E. Shannon. The synthesis of two-terminal switching circuits. *Bell System Tech. J.*, 25:59–98, 1949.

H. D. Sherali and W. P. Adams. A hierarchy of relaxations between the continuous and convex hull representations for zero-one programming problems. *SIAM J. Discrete Math.*, 3:411–430, 1990.

F. Vanderbeck and L. A. Wolsey. Reformulation and decomposition of integer programs. In M. Jünger et al., editor, *50 Years of Integer Programming 1958-2008*, pages 431–502. Springer, 2010.

L. A. Wolsey. Using extended formulations in practice. *Optima*, 85:7–9, 2011.

M. Yannakakis. Expressing combinatorial optimization problems by linear programs (extended abstract). In *Proc. STOC 1988*, pages 223–228, 1988.

M. Yannakakis. Expressing combinatorial optimization problems by linear programs. *J. Comput. System Sci.*, 43(3):441–466, 1991.

G. M. Ziegler. *Lectures on Polytopes*, volume 152 of *Graduate Texts in Mathematics*. Springer-Verlag, Berlin, 1995.

A Background and Proofs

A.1 Background on Polytopes

A *(convex) polytope* is a set $P \subseteq \mathbb{R}^d$ that is the convex hull $\text{conv}(V)$ of a finite set of points V. Equivalently, $P \subseteq \mathbb{R}^d$ is a polytope if and only if P is bounded and the intersection of a finite collection of closed halfspaces. This is equivalent to saying that P is bounded and the set of solutions of a finite system of linear inequalities (or equalities, each of which can be represented by a pair of inequalities).

Let P be a polytope in $\mathbb{R}^d$. A closed halfspace H^+ that contains P is said to be *valid* for P. In this case the hyperplane H that bounds H^+ is also said to be *valid* for P. A *face* of P is either P itself or the intersection of P with a valid hyperplane. Every face of a polytope is again a polytope. A face is called *proper* if it is neither the empty face nor the polytope itself. A *vertex* is a minimal nonempty face. A *facet* is a maximal proper face. An inequality $c^T x \leqslant \delta$ is said to be *valid* for P if it is satisfied by all points of P. The face it defines is $F := \{x \in P \mid c^T x = \delta\}$. The inequality is called *facet-defining* if F is a facet. The *dimension of a polytope* P is the dimension of the affine space $\text{aff}(P)$ containing P.

Every (finite or infinite) set V such that $P = \text{conv}(V)$ contains all the vertices of P. Conversely, letting $\text{vert}(P)$ denote the vertex set of P, we have $P = \text{conv}(\text{vert}(P))$. Suppose that P is full-dimensional, that is, $\dim(P) = d$. Then, every (finite) system $Ax \leqslant b$ such that $P = \{x \in \mathbb{R}^d \mid Ax \leqslant b\}$ contains all the facet-defining inequalities of P, up to scaling by positive numbers. Conversely, P is described by its facet-defining inequalities. In case P is not full-dimensional, these statements have to be adapted in the following way. Every (finite) system describing P contains all the facet-defining inequalities of P, up to scaling by positive numbers and adding an inequality that is satisfied with equality by *all* points of P. Conversely, a linear description of P can be obtained by picking one facet-defining inequality per facet and adding a system of equalities describing the affine hull of P.

For more background on polytopes, see the standard reference Ziegler [1995].

A.2 Proofs From Section 2

Lemma 1. *We have* $\pi_x(Q) = \text{proj}_x(Q)$.

Proof. Let $\alpha \in \pi_x(Q)$. Then there exists $y \in C$ with $E\alpha + Fy = g$. Pick any $\mu \in C_Q$. Then, $\mu^T E\alpha + \mu^T Fy = \mu^T g$ holds. Since $F^T\mu \in C^*$ and $y \in C$ we have that $(F^T\mu)^T y = \mu^T Fy \geqslant 0$. Therefore $\mu^T E\alpha \leqslant \mu^T g$ holds for all $\mu \in C_Q$. We conclude $\alpha \in \text{proj}_x(Q)$ and as α was arbitrary $\pi_x(Q) \subseteq \text{proj}_x(Q)$ follows.

Now suppose $\pi_x(Q) \neq \text{proj}_x(Q)$. Then there exists α such that $\alpha \in \text{proj}_x(Q)$ but $\alpha \notin \pi_x(Q)$. In other words there is no $y \in C$ such that $Fy = g - E\alpha$ or, equivalently, the convex cone $F(C) := \{Fy \mid y \in C\}$ does not contain the point $g - E\alpha$. Since C is a closed cone, so is $F(C)$. Therefore, by the Strong Separation Theorem there exists $\mu \in \mathbb{R}^p$ such that $\mu^T z \geqslant 0$ is valid for $F(C)$ but $\mu^T(g - E\alpha) < 0$. Then $\mu^T z = \mu^T(Fy) = (\mu^T F)y \geqslant 0$ is valid for C, i.e., $(\mu^T F)y \geqslant 0$ holds for all $y \in C$, implying $F^T\mu \in C^*$. As $\mu^T(g - E\alpha) < 0$ we have $\mu^T E\alpha > \mu^T g$. On the other hand we have $F^T\mu \in C^*$ so that $\mu \in C_Q$ implying $\mu^T E\alpha \leqslant \mu^T g$; a contradiction. Hence, $\pi_x(Q) = \text{proj}_x(Q)$ follows. □

Theorem 4. *Let* $P = \{x \in \mathbb{R}^d \mid Ax \leqslant b\} = \text{conv}(V)$ *be a polytope defined by* m *inequalities and* n *points respectively, and let* S *be the slack matrix of* P *w.r.t.* $Ax \leqslant b$ *and* V. *Also, let* $C \subseteq \mathbb{R}^k$ *be a closed convex cone. Then, the following are equivalent:*

(i) There exist T, U *such that (the transpose of) each row of* T *is in* C^*, *each column of* U *is in* C *and* $S = TU$.

(ii) There exists a conic EF $Q = \{(x,y) \in \mathbb{R}^{d+k} \mid Ex + Fy = g,\ y \in C\}$ such that $P = \pi_x(Q)$.

Proof. We first show that a factorization induces a conic EF. Suppose there exist matrices T, U as above. We claim that Q with $E := A$, $F := T$ and $g := b$ has the desired properties. Let $v_j \in V$, then $S^j = TU^j = b - Av_j$ and so it follows that $(v_j, U^j) \in Q$ and $v_j \in \pi_x(Q)$. Now let $x \in \pi_x(Q)$. Then, there exists $y \in C$ such that $Ax + Ty = b$. Since $T_i y \geqslant 0$ for all $i \in [m]$, we have that $x \in P$. This proves the first implication.

For the converse, suppose $P = \pi_x(Q)$ with Q being a conic EF of P. By Lemma 1, $\pi_x(Q) = \{x \in \mathbb{R}^d \mid \mu^T Ex \leqslant \mu^T g, \forall \mu \in C_Q\}$, where $C_Q = \{\mu \in \mathbb{R}^p \mid F^T\mu \in C^*\}$. It suffices to prove that the submatrix of S induced by the rows corresponding to the inequalities of $Ax \leqslant b$ that define facets of P admits a factorization. Thus, we assume for the rest of the proof that all rows of S correspond to facets of P. For simplicity, we also assume that P is full-dimensional. Then, for any facet-defining inequality $A_i x \leqslant b_i$ of P there exists $\mu_i \in C_Q$ such that $A_i = \mu_i^T E$ and $b_i = \mu_i^T g$. (This follows from the fact that C_Q is closed; see also [Lemaréchal and Hiriart-Urruty, 1996, Theorem 4.3.4].) We define $T_i := \mu_i^T F$ for all i; in particular $(T_i)^T \in C^*$ as $\mu_i \in C_Q$. Now let $v_j \in V$. Since $P = \pi_x(Q)$, there exists a $y_j \in C$ such that $Ev_j + Fy_j = g$ and so $\mu_i^T Ev_j + \mu_i^T Fy_j = \mu_i^T g$. With the above we have $A_i v_j + T_i y_j = b_i$ and as $v_j \in \pi_x(Q)$ we deduce $T_i y_j \geqslant 0$. The slack of v_j w.r.t. $A_i x \leqslant b_i$ is $b_i - A_i v_j = \mu_i^T g - \mu_i^T Ev_j = \mu_i^T Fy_j = T_i y_j$. This implies the factorization $S = TU$ with $T_i = \mu_i^T F$ and $U^j = y_j$. □

A.3 Proofs From Section 5

Lemma 14. *If, for each $a, b \in \{0,1\}^n$, we let $U_a := \binom{1}{-a}\binom{1}{-a}^T$ and $V^b := \binom{1}{b}\binom{1}{b}^T$, then the matrices $\{U_a\}_{a\in\{0,1\}^n}$ and $\{V^b\}_{b\in\{0,1\}^n}$ define a PSD factorization of S.*

Proof. We have

$$\begin{aligned} Tr\left[U_a V^b\right] &= Tr\left[\binom{1}{-a}\binom{1}{-a}^T\binom{1}{b}\binom{1}{b}^T\right] \\ &= (1 - a^T b) \cdot Tr\left[\binom{1}{-a}\binom{1}{b}^T\right] \\ &= (1 - a^T b) \cdot Tr\left[\binom{1}{b}^T\binom{1}{-a}\right] \\ &= (1 - a^T b)^2 \\ &= S_{ab}. \end{aligned}$$

□

Lemma 17. *For each n, there exists a graph G_n with $O(n^2)$ vertices such that* STAB(G_n) *contains a face that is a linear EF of* CUT(n).

Proof. Because CUT(n) and COR$(n-1)$ are affinely isomorphic (see Theorem 13), it suffices to prove the result with COR$(n-1)$ instead of CUT(n). Consider the complete graph K_{n-1} with vertex set $V_{n-1} := \{1, \ldots, n-1\} = [n-1]$. For each vertex i of K_{n-1} we create two vertices labeled $ii, \overline{ii}$ in G_n and an edge between them. For each edge ij of K_{n-1}, we add to G_n four vertices labeled $ij, \overline{ij}, \underline{ij}, \underline{\overline{ij}}$ and all possible six edges between them. We further add the following eight edges to G_n:

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

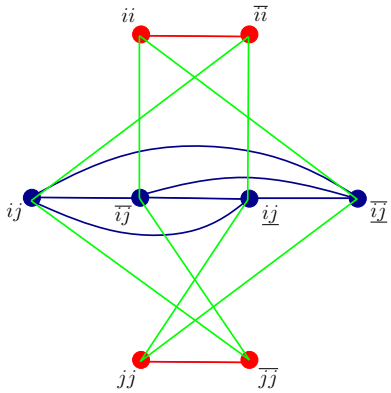

Figure 1: The edges and vertices of G_n corresponding to vertices i, j and edge ij of K_{n-1}.

See Fig. 1 for an illustration. The number of vertices in G_n is $2(n-1) + 4\binom{n-1}{2}$.

Thus the vertices and edges of K_{n-1} are represented by cliques of size 2 and 4 respectively in G_n. We will refer to these as *vertex-cliques* and *edge-cliques* respectively. Consider the face $F = F(n)$ of $\mathrm{STAB}(G_n)$ whose vertices correspond to the stable sets containing exactly one vertex in each vertex-clique and each edge-clique. (The vertices in this face correspond exactly to stable sets of G_n with maximum cardinality.)

Consider the linear map $\pi : \mathbb{R}^{V(G_n)} \to \mathbb{R}^{(n-1)\times(n-1)}$ mapping a point $x \in \mathbb{R}^{V(G_n)}$ to the point $y \in \mathbb{R}^{(n-1)\times(n-1)}$ such that $y_{ij} = y_{ji} = x_{ij}$ for $i \leqslant j$. In this equation, the subscripts in y_{ij} and y_{ji} refer to an ordered pair of elements in $[n-1]$, while the subscript in x_{ij} refers to a vertex of G_n that corresponds either to a vertex of K_{n-1} (if $i = j$) or to an edge of K_{n-1} (if $i \neq j$).

We claim that the image of F under π is $\mathrm{COR}(n-1)$, hence F is a linear EF of $\mathrm{COR}(n-1)$ (and thus of $\mathrm{CUT}(n)$). Indeed, pick an arbitrary stable set S of G_n such that $x := \underline{\chi}^S$ is on face F. Then define $b \in \{0,1\}^{n-1}$ by letting $b_i := 1$ if $ii \in S$ and $b_i := 0$ otherwise (i.e., $\overline{ii} \in S$). Notice that for the edge ij of K_{n-1} we have $ij \in S$ if and only if both vertices ii and jj belong to S. Hence, $\pi(x) = y = bb^T$ is a vertex of $\mathrm{COR}(n-1)$. This proves $\pi(F) \subseteq \mathrm{COR}(n-1)$. Now pick a vertex $y := bb^T$ of $\mathrm{COR}(n-1)$ and consider the unique maximum stable set S that contains vertex ii if $b_i = 1$ and vertex $\overline{ii}$ if $b_i = 0$. Then $x := \chi^S$ is a vertex of F with $\pi(x) = y$. Hence, $\pi(F) \supseteq \mathrm{COR}(n-1)$. Thus $\pi(F) = \mathrm{COR}(n-1)$. This concludes the proof. □